Hybrid Generalized Non-Orthogonal Multiple Access for the 5G Wireless Networks

Masters Thesis

Samson Manyani Zitha

A thesis submitted in fulfilment of the requirement for the degree of

MASTERS OF SCIENCE IN ENGINEERING
(COMPUTER ENGINEERING)

School of Electrical, Electronic & Computer Engineering
Durban
South Africa

Thesis submitted December, 2018

Hybrid Generalized Non-Orthogonal Multiple Access for the 5G Wireless Networks

Samson Manyani Zitha

Supervisor: Dr. Tom Mmbasu Walingo

A thesis submitted in fulfilment of the requirement for the
degree of

MASTERS OF SCIENCE IN ENGINEERING
(COMPUTER ENGINEERING)

School of Electrical, Electronic & Computer Engineering
University of KwaZulu-Natal
South Africa

EXAMINER'S COPY

Thesis submitted December, 2018

As the candidate’s supervisor, I have approved this thesis for submission.

Signed..Date...

Name: Dr. Tom Mmbasu Walingo

Declaration 1 - Plagiarism

I, **Samson Manyani Zitha**, declare that;

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
    (a) Their words have been re-written but the general information attributed to them has been referenced.
    (b) Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the references sections.

Signed..Date...

Declaration 2 - Publication

The following papers emanating from this work have either been published or are under review:

1. **Samson Manyani Zitha, Tom Walingo** and **James Okello**, “Performance Analysis of Generalized-NOMA Techniques in 5G Wireless Networks ”, Proceedings of SATNAC 2018, Harmanus, South Africa.
2. **Samson Manyani Zitha** and **Tom Walingo**, “Energy-Efficient Hybrid G-NOMA Resource Allocation scheme for 5G Heterogeneous Networks”, IEEE Communication Letters, 2018 (Under Review)
3. **Samson Manyani Zitha** and **Tom Walingo**, “Hybrid G-NOMA Resource Allocation scheme for 5G Small-Cell Networks”, South African Institute of Electrical Engineers (SAIEE), 2018 (Under Review).

Dedication

Above all, all thanks goes to the almigthy God. I dedicate this research thesis to my family, my future wife Sharlotte Thulsile Mangane and my daughter Thandolwethu Awande Zitha.

Acknowledgements

I wish to extend my sincere gratitude and appreciation to God and all the individuals who gave their full support and assistance in many ways during this work. I am greatly indebted to my supervisor, Dr. Tom Walingo for his wise and expect guidance, constructive corrections, comments, and encouragement that contributed to the completion of the study. I would like to thank Prof James Okello for his support and opportunity to begin this study at first.

To my family, without their love, care and support, nothing would have been possible, no words can express the appreciation, love and thanks I have for you, you will forever resides in my heart. I would like to thank the University of KwaZulu Natal and the CRART center for all the financial support you provided me with, may you continue to support and empower many more to come. To all my friends, colleagues and roomates "ngiyabonga". To all the reviewers and Mrs Shiellah Makhubele, thank you for all the hard work you have invested into making this work a success. Finally, to my future wife Sharlotte Thulsile Mangane and my daughter Thandolwethu Awande Zitha, thank you for believing in me and I dedicate this work to you.

Contents

List of Figures

List of Figures

List of Tables

List of Tables

List of Acronyms

1G	**F**irst **G**eneration
2G	**S**econd **G**eneration
3G	**T**hird **G**eneration
4G	**F**our **G**eneration
5G	**F**ifth **G**eneration
AMPS	**A**dvanced **M**obile **P**hone **S**ystem
APP	**A** **P**osteriori **P**robability
AWGN	**A**dditive **W**hite **G**aussian **N**oise
BER	**B**it **E**rror **R**ate
BS	**B**ase **S**tation
CBC	**C**hip **B**y **C**hip
CDMA	**C**ode **D**ivision **M**ultiple **A**ccess
CD-NOMA	**C**ode **D**omain **N**on-**O**rthogonal **M**ultiple **A**ccess
CSI	**C**hannel **S**tate **I**nformation
C-RAN	**C**loud **R**adio **A**ccess **N**etworks
DEC	**D**ecoding
EARTH	**E**nergy-**A**ware **R**adio and Network **T**echnology
EE	**E**nergy-**E**fficiency
ENC	**E**ncoding

ESE	**E**lementary **S**ignal **E**stimator
FBS	**F**emto-cell **B**ase **S**tation
FDE	**F**requency **D**omain **E**quilization
FDMA	**F**requency **D**ivision **M**ultiple **A**ccess
FUEs	**F**emto-cell **U**ser **E**quipments
FRA	**F**uture **R**adio **A**ccess
Gbps	**G**iga**B**ytes per **S**econd
G-NOMA	**G**eneralized **N**on-**O**rthogonal **M**ultiple **A**ccess
GSM	**G**lobal **S**ystem for **M**obile **C**ommunication
GPRS	**G**eneral **P**acket **R**atio **S**ervice
GPS	**G**lobal **P**osition **S**ystem
HetNets	**H**eterogenoues **N**etworks
HG-NOMA	**H**ybrid **G**eneralized **N**on-**O**rthogonal **M**ultiple **A**ccess
HG-NOMA-SIC	**H**ybrid **G**eneralized **N**on-**O**rthogonal **M**ultiple **A**ccess **S**uccessive **I**nterference **C**ancellation
IDMA	**I**nterleave **D**ivision **M**ultiple **A**ccess
IMEI	**I**nternational **M**obile **E**quipment **I**dentity
IoT	**I**nternet of **T**hings
IP	**I**nternet **P**rotocol
ITU	**I**nternational **T**elecomunication **U**nion
ITU-R	**I**nternational **T**elecomunication **U**nion-**Radio** comunication
	LDS **L**ow **D**ensity **S**preading **LDSMA** **L**ow **D**ensity **S**preading **M**ultiple **A**ccess **LTE** **L**ong **T**erm **E**volution **A**dvance **LSD-CDMA** **L**ow **D**ensity **S**preading-**C**ode **D**ivision **M**ultiple **A**ccess
M2M	**M**achine-to-**M**achine
MAC	**M**ultiple **A**ccess

MAC	**M**eduim **A**ccess **C**ontrol
MBS	**M**acro-cell **B**ase **S**tation
MC-CDMA	**M**ulti-**C**arrier **C**ode **D**ivision **M**ultiple **A**ccess
MIMO	**M**ultiple **I**put **M**ultiple **O**utput
MIMO-NOMA	**M**ultiple **I**put **M**ultiple **O**utput-**N**on-**O**rthogonal **M**ultiple **A**ccess
MMSE	**M**inimum **M**ean **S**quare **E**rror
mMTC	**M**assive **M**achine-**T**ype **C**ommunication
mmWave	**M**illimeter **W**ave
MUSA	**M**ulti-**U**ser **S**hared **A**ccess
MUEs	**M**acro-cell **U**ser **E**quipments
MUD	**M**ulti-**U**ser **D**etection
MUDD	**M**ulti-**U**ser **D**etection and **D**ecoding
MPA	**M**essage **P**assing **A**lgorithm
MTC	**M**achine-**T**ype **C**ommunication
NMA	**N**ovel **M**ultiple **A**ccess
NOMA	**N**on-**O**rthogonal **M**ultiple **A**ccess
NLM	**N**ew **L**ogistic **M**ap **I**nterleaver
OMA	**O**rthogonal **M**ultiple **A**ccess
OFDMA	**O**rthogonal **F**requency **D**ivision **M**ultiple **A**ccess
PDMA	**P**attern **D**ivision **M**ultiple **A**ccess
PD-NOMA	**P**ower **D**omain **N**on-**O**rthogonal **M**ultiple **A**ccess
PHY	**Ph**ysical
QAM	**Q**uadrature **A**mplitude **M**odulation
QoS	**Q**uality of **S**ervices
QPSK	**Q**uadrature **P**hase **S**hift **K**eying
RE	**R**esource **E**lements

SAMA	**S**uccessive **I**nterference **C**ancellation **A**menable **M**ultiple **A**ccess
SDR	**S**oftware **D**efined **R**adio
SoDeMA	**S**oftware **D**efined **M**ultiple **A**ccess
SOMSA	**S**ub-**O**ptimal **M**atching **S**cheme of **S**ubchannel **A**ssignment
SIC	**S**uccessive **I**nterference **C**ancellation
SINR	**S**ignal **I**nterference to **N**oise **R**atio
SCMA	**S**parse **C**ode **M**ultiple **A**ccess
uMTC	**U**ltrareliable **M**achine-**T**ype **C**ommunication
UMTS	**U**niversal **M**obile **T**eleommunication **S**ystem

Abstract

The deployment of 5G networks will lead to an increase in capacity, spectral efficiency, low latency and massive connectivity for wireless networks. They will still face the challenges of resource and power optimization, increasing spectrum efficiency and energy optimization, among others. Furthermore, the standardized technologies to mitigate against the challenges need to be developed and are a challenge themselves. In the current predecessor LTE-A networks, orthogonal frequency multiple access (OFDMA) scheme is used as the baseline multiple access scheme. It allows users to be served orthogonally in either time or frequency to alleviate narrowband interference and impulse noise. Further spectrum limitations of orthogonal multiple access (OMA) schemes have resulted in the development of non-orthogonal multiple access (NOMA) schemes to enable 5G networks to achieve high spectral efficiency and high data rates. NOMA schemes unorthogonally co-multiplex different users on the same resource elements (RE) (i.e. time-frequency domain, OFDMA subcarrier, or spreading code) via power domain (PD) or code domain (CD) at the transmitter and successfully separating them at the receiver by applying multi-user detection (MUD) algorithms. The current developed NOMA schemes, refered to as generalized-NOMA (G-NOMA) technologies includes; Interleaver Division Multiple Access (IDMA, Sparse code multiple access (SCMA), Low-density spreading multiple access (LDSMA), Multi-user shared access (MUSA) scheme and the Pattern Division Multiple Access (PDMA). These protocols are currently still under refinement, their performance and applicability has not been thoroughly investigated. The first part of this work undertakes a thorough investigation and analysis of the performance of the existing G-NOMA schemes and their applicability.

Generally, G-NOMA schemes perceives overloading by non-orthogonal spectrum resource allocation, which enables massive connectivity of users and devices, and offers improved system spectral efficiency. Like any other technologies, the G-NOMA schemes need to be improved to further harvest their benefits on 5G networks leading to the requirement of Hybrid G-NOMA (G-NOMA) schemes. The second part of this work develops a HG-NOMA scheme to alleviate the

5G challenges of resource allocation, inter and cross-tier interference management and energy efficiency. This work develops and investigates the performance of an Energy Efficient HG-NOMA resource allocation scheme for a two-tier heterogeneous network that alleviates the cross-tier interference and improves the system throughput via spectrum resource optimization. By considering the combinatorial problem of resource pattern assignment and power allocation, the HG-NOMA scheme will enable a new transmission policy that allows more than two macro-user equipment's (MUEs) and femto-user equipment's (FUEs) to be co-multiplexed on the same time-frequency RE increasing the spectral efficiency. The performance of the developed model is shown to be superior to the PD-NOMA and OFDMA schemes.

Preface

"We cannot seek achievement for ourselves and forget about progress and prosperity of our community. Our ambitions must be broad enough to include the aspirations and needs of others, for their sakes and for our own"

— Cesar Chavez

Samson Manyani Zitha

University of KwaZulu-Natal, March 5, 2019

Part I

Introduction

Introduction and Background

Wireless communication access technologies have evolved through different paths targeting improved performance, high efficiency, and full guaranteed QoS in a mobile communication environment. The first generation (1G) network fulfilled the fundamental basics of mobile voice, while the second generation (2G) networks introduced the concept of network capacity and extended coverage. The third generation (3G) system, offered high speed internet access for multimedia applications, presenting a step change from the 2G networks. The fourth generation (4G) networks is a pure data wireless connection network, providing access to a wide range of end-to-end internet protocol (IP) communication services, which includes advanced services such as mobility, flexibility, reliability, and affordability. The fifth generation (5G) networks is highly anticipated to offer low-latency, high throughput and massive connectivity of billions of devices around the world. For its realization the 5G faces the challenges of proliferation of data traffic, high energy consumption, increasing spectral resource efficiency, and power optimization. Therefore, this work focuses on multiple access schemes and energy-efficiency spectrum resource optimization algorithms to alleviate the challenges of 5G heterogeneous networks (HetNets). This chapter highlights the evolution of mobile communications networks, 5G networks background, challenges, research motivation, and the main contributions of this research.

1 Evolution of Mobile Communication Networks

The evolution of mobile networks from the 1G system until today and the forecasts for the next decade clearly indicates a growth of both the network itself (in terms of installed equipment) and carried traffic (in terms of transmitted bits) [1]. Today major network operators still count on their 2G global system for mobile (GSM) communication network, for revenues generated from voice traffic. They offer broadband and more advanced services based on the 3G universal mobile telecommunications system (UMTS) network. Recently, they began to deploy the 4G long-term evolution (LTE) technology, promising faster broadband connections to their customers. At the same time, the telecommunications' industry started to think about new scenarios and new features for what the next generation of mobile networks for the year 2020 and beyond will be like, i.e., 5G network technology [1], [2]. A telecommunication network constitutes a major sector of information communication and technology (ICT) and has undergone enormous growth. Capacity issues and the delivery of complex real-time services are some drivers of the evolution of telecommunication [3].

During the 1980s, the 1G mobile telecommunication systems were deployed for commercial use. They became the first analog cellular system, or advanced mobile phone system (AMPS), to be deployed in North America. With the increasing demand for capacity and high quality of communications, the development of digital cellular technology thereafter became significant in transforming human lives. The 2G mobile networks were commercially launched in Finland by Radiolinja in 1991 to improve the 1G networks. The 2G networks used the GSM standard [4], to enhance spectral-efficiency and offered improved mobile data services than the AMPS. The introduction of the general packet radio service (GPRS) became the first major step to revolutionize the GSM networks towards the development of 3G UMTS network. In the early 1980s, the 3G networks were developed by the international telecommunication union (ITU). Compared to the 2G networks, the 3G offered higher data rates and greater security. Furthermore, using bandwidth and location information of 3G devices, the global positioning system (GPS), location-based services, mobile Internet access, video calls, and mobile TV were developed into applications. A new generation of cellular standards has appeared approximately every ten years since the 1G were introduced. In March 2008, the standard requirements for the 4G networks were specified by the international telecommunications union-radio communications sector (ITU-R) [5]. A significant step forward from 3G to 4G was that, the 4G supports all-internet protocol (IP) based communication, such as IP telephony, instead of traditional circuit-switched telephony service. The spread spectrum radio technology used in the 3G was abandoned in 4G networks. The main technologies in 4G systems, were orthogonal frequency division multiple access (OFDMA) multi-carrier transmission and frequency domain equalization (FDE) schemes. From the beginning of mobile digital cellular communications, orthogonal multiple access (OMA) has been the only means to separate users. In the 1G system the frequency division multiple access (FDMA) was used to partition the available spectrum into non-overlapping frequency sub-bands, with each accommodating one digital data stream. On the other hand, in 2G networks, a round-robin digital data stream fashion termed time division multiple access (TDMA), was used to partition users into time slots. Moreover, the improvements from GSM led to the introduction of the 3G, which applied the code domain multiple Access (CDMA) scheme. The basic idea behind CDMA is utilizing a pseudo-random digital spread spectrum technique, to enable the full use of the available spectrum. Based on the orthogonal frequency division multiplexing (OFDM) waveform, a multi-carrier orthogonal frequency division multiple access (OFDMA) was implemented for the 4G. This enabled tight constraint of orthogonal frequency-domain packing, with subcarrier spacing inverse to the symbol duration [6]. In light of this, the potential benefits of the OMA schemes includes simplified transceiver design and intra-cell or co-channel interference avoidance. However, the constraints are too obvious. i.e., the number of users that can be served simultaneously, is limited by the pool of radio resource restrictions. Furthermore, user scheduling requires dedicated feedback channels, at the expense of signaling overhead to ensure orthogonality. Thus, these problems needs to addressed further in the upcoming generations.

The trend analysis from analog phone calls across all IP services, including voice, video, and messaging, has been motivated by the need to meet the demands of new mobile network technologies. Needless to say that, mobile networks are currently facing a new challenge of birthing a hyper-connected society through the 5G services. The 5G technology is expected to roll out by 2020 to meet the overwhelming QoS of data rates and spectral-efficiency. Hence, various technologies have been proposed in recent years, and these include: massive multiple-input multiple output (MIMO) [7], millimeter wave communications (mmWave) [8], LTE in unlicensed spectrum (LTE-U) [9], and cloud radio access networks [10]. Subsequently, from the future radio access technology viewpoint, a step change in data-speed, spectral efficiency, massive connectivity of users access, and a significant reduction in end-to-end latency are major requirements for 5G networks. Since the evolution of the mobile-Internet and the Internet of Things (IoT) exponentially accelerating the demand for higher data-rate applications, the 5G network-level data-rates are expected be 10-20 GigaBytes per second (Gbps) (which is, 10-20 times the peak data-rate in 4G networks), and user-experienced data-rates expected to be 1 Gbps (100 times the user-experienced data rate in 4G). They will have a latency (end-to-end round-trip delay) of 1 millisecond (one-fifth of the latency in 4G) [11]. A summary of the mobile networks standards, is given in Table 1. This work focuses on the development of access standards for the 5G.

2 Fifth Generation (5G) Networks

2.1 LTE or LTE-A Networks

The 4G long-term evolution (LTE) mobile communication network was commercially launched on December 2009 by Telia Sonera from Norway and Sweden [13]. The LTE or system architecture

Table 1: Evolution of wireless networks [12]

Generations	Access Technology		Data Rate	Frequency Band	Bandwidth	Forward Error Correction	Switching	Application
1G	Advanced Mobile Phone Service (AMPS) Frequency Division Multiple Access (FDMA)		2.4 kbps	800 MHz	30 KHz	NA	Circuit	Voice
2G	Global System for mobile communications (GSM)(Time Division Multiple Access (TDMA))		10 kbps	850/900/1800/1900 MHz	200 KHz	NA	Circuit	Voice Data
	Code Division Multiple Access (CDMA)		10 kbps		1.25 MHz			
2.5G	General Packet Radio Service (GPRS)		50 kbps		200 KHz		Circuit/ Packet	
	Enhanced Data Rate for GSM Evolution (EDGE)		200 kbps		200 KHz			
3G	Wideband Code Division Multiple Access (WCDMA)/ Universal Mobile Telecommunications Systems (UMTS)		384 kbps	800/850/900 /1800/1900/ 2100 MHz	5 MHz	Turbo Codes	Circuit/ Packet	Voice + Data+ Video Calling
	Code Division Multiple Access (CDMA) 2000		384 kbps		1.25 MHz		Circuit/ Packet	
3.5G	High Speed Uplink/ Downlink Packet Access (HSUPA/HSDPA)		5-30 Mbps		5 MHz		Packet	
	Evolution-Data Optimized (EVDO)		5-30 Mbps		1.25 MHz		Packet	
3.75G	Long Term Evolution (LTE) (Orthogonal / Single Carrier Frequency Division Multiple Access) (OFDMA / SC-FDMA)		100-200 Mbps	1.8 GHz, 2.6 GHz	1.4 MHz to 20 MHz	Concatenated codes	Packet	Online gaming + High Definition Television
	Worldwide Interoperability for Microwaves Access (WIMAX) (Scalable Orthogonal Frequency Division Multiple Access (SOFDMA)	Fixed WiMAX	100-200 Mbps	3.5 GHz and 5.8 GHz initially	3.5 MHz and 7 MHz in 3.5 GHz band; 10 MHz in 5.8 GHz band			
4G	Long Term Evolution (LTE) (Orthogonal / Single Carrier Frequency Division Multiple Access) (OFDMA / SC-FDMA)		DL 3Gbps UL 1.5Gbps	1.8GHz 2.6GHz	1.4MHz to 20 MHz	Turbo Codes	Packet	Online aming + High Definition Television
	Worldwide Interoperability for Microwaves Access (WIMAX)(Scalable Orthogonal Frequency Division Multiple Access (SOFDMA)	Mobile WiMAX	100-200 Mbps	2.3GHz, 2.5GHz, and 3.5GHz initially	3.5MHz, 7MHz, 5MHz, 10MHz and 8.75MHZ initially			
5G	Beam Division Multiple Access (BDMA) and Non- and quasi-orthogonal or Filter Bank multi Carrier (FBMC) multiple access		10-50 Gbps (expected)	1.8, 2.6 GHz and expected 30-300 GHz	60 GHz	Low Density Parity Check codes (LDPC)	Packet	Ultra High Definition video + Virtual Reality application

evolution (LTE/SAE) was specified by the third generation partnership project (3GPP) as an improved standard of UMTS [14]. More specifically, the LTE networks were designed as packet-based with fewer network elements attribute, such as improving system capacity and coverage, offering low end-to-end latency compared with the previous networks, while providing QoS in terms of HetNets deployment, flexible bandwidth operation and seamless integration with other existing wireless communication systems [15]. Particularly, the current mobile system's worldwide provides broadband wireless service through LTE and WiMAX telecommunication technologies. Both LTE and WiMAX use frequency division duplex (FDD) and Time division duplex (TDD) to realize multiplexing. To keep up with the proliferation of data traffic, the currently deployed LTE architecture infrastructures should be able to handle the limitations of extension to higher data-rates, which is difficult to achieve using OMA techniques. This is because of excessive interference between users, spectrum scarcity, and the difficulties of providing a wide range of multi-rate requirements [16]. Furthermore, the challenges of deploying small-cell networks related to LTE-Advanced networks standardization add more complications [17]. The security aspects of the LTE and LTE-A networks where investigated in [15]. The authors [18] investigated the issues of standardization, protocols, and convergence of 4G technologies. Even though LTE/LTE-A is continuously updated through new releases, and with LTE Advanced Pro Release 13 being the latest one, the development of the 5G wireless networks has commenced.

2.2 5G Networks

5G networks have been necessitated by the numerous challenges facing 4G networks, which includes the need for higher data-rate, improved capacity, lower cost devices, low end-to-end latency, and massive user access connectivity [19]. The 5G network is expected to be standardized in 2018 and deployed in 2020 [20]. Furthermore, the limitations of OFDMA has led to the development of the non-orthogonal multiple access (NOMA) technology, which has recently been put forward for consideration by the 3GPP LTE-A. NOMA introduces a new multiplexing domain by considering the combinatorial classic time/frequency/code domains. By doing so, significant bandwidth efficiency enhancement can be achieved over the traditional OMA techniques [21]. This subsection provides an overview of the 5G architecture technology and its features i.e, macro and small cells networks and multiple access technologies.

2.2.1 5G Network Architecture

Through persistent effort and determination, networks operators are implementing a digital transformation to create a better digital world. With the rapid increase in wireless data traffic, the challenges of increasing network capacity and improving energy-efficiency, cost, and spectrum utilization, as well as providing better scalability for handling the increasing number of the connected devices arise [12]. To meet these demands of the user and network operators, and to overcome the challenges put forward for the 5G system, a radical change in the strategy of designing the 5G wireless cellular architecture is more than required. Regarding recent progress for the 5G networks, the requirements are to be decided by the ITU-R. On the other hand, the working party (WP) 5D group is currently preparing evaluation criteria and submissions of proposals and analysis of candidate 5G technologies. Hence, by late 2019 the first certified 5G standards requirements, are expected to be concluded. Furthermore, the ITU-T has recently completed a study based on the system-wide view of 5G architectures and proofs-of-concept into networking towards innovations to the development of the 5G systems through a focus group on IMT-2020-ITU. However, the usage trend patterns of the 5G (IMT 2020) are not limited to mobile broadband. In fact, the IMT-2020 is envisioned to support a diverse heterogeneity of usage scenarios/use cases into three categories [22].

1. Enhanced mobile broadband (eMBB): eMBB is one of three primary 5G New Radio (NR) use cases defined by the 3GPP, in order to provide provision for the requirements of 5G eMBB, significant improvements in bandwidth, wide-area coverage, spectral-efficiency and signaling efficiency are to be achieved compared to 4G [23]. Enabled by the key versatile radio technologies, which includes, offloading, hyper-densification, increased bandwidth usage in the mmWave spectrum and massive MIMO technologies will result in higher performance of eMBB.

2. Ultra-reliable and low latency communications (URLLC): URLLC is a service category to support stringent requirements for reliability, latency-sensitive, availability, remote control, autonomous driving, disaster relief, and tactile Internet etc [24]. Even though the latency of 4G LTE networks has been significantly enhanced from its 3G predecessor networks, the end-to-end latency still remains in the range of 30–100 ms. Hence, the 5G networks are expected to limit the end-to-end round-trip delay up to 1 millisecond which is one-fifth of the latency in 4G [11].

3. Massive machine type communications (mMTC): mMTC are characterized by fully automatic data generation, exchange, processing, and actuation among intelligent machines, without or

with a low intervention of humans. However, the deployment mMTC will consist of a large number of devices with a relatively low (or relatively high) volume of non-delay-sensitive data. Hence, low-cost devices and very long power-constrained devices are required from now and beyond. Furthermore, diverse mMTC services will exhibit different traffic patterns, thus, making the problem of spectrum resource allocation very challenging.

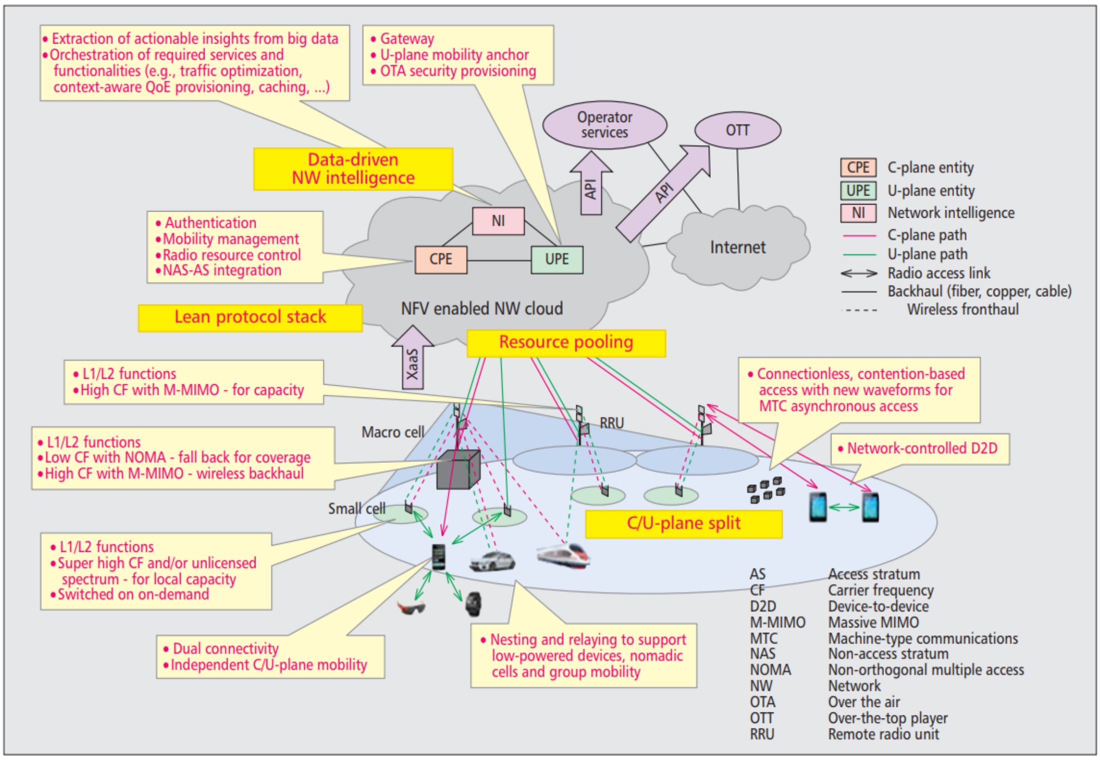


Fig. 1: 5G Network Architecture [25]

The authors in [25] presented a two-layer 5G network architecture consisting of a radio access network and a network cloud, which integrated many enabling technologies such as small-cells, massive MIMO (mMIMO), software-defined networks (SDN), and network functions virtualization (NVF) to facilitate optimal use of network resources for quality-of-experience (QoE) provisioning and planning. This architecture is shown in Fig 1. In their work, an initial proof of concept was presented in order to demonstrate the technical feasibility of the proposed architecture. The crucial issues that needed to be addressed and resolved to realize a complete 5G architecture vision were thoroughly discussed. In addition, [26] proposed an all-SDN network architecture with hierarchical network control capabilities as a simplified and unified approach to mobility and routing management for 5G networks. Beyond this, the novel architecture supports connectivity management as a service, which may be offered to provide QoS differentiation with a range of options to protect flows against subscriber mobility at different price levels without the utilization of tunneling protocols. Performance results showed that the proposed architecture is a viable solution for mobile Device-to-device (D2D) communications with end-to-end latency below 5 ms. The efficiency issues of the current 3GPP centralized approach to mobility management using core anchoring were discussed in [27]. They presented and discussed multiple dimensions of distributing the mobility functions, and suggested how the mobile terminal can utilize them (terminal based solution) in orchestration for efficient communication over a 5G flat mobile network. [28] argued that there is a need to integrate functionality in D2D and infrastructure-to-device (I2D) modes, and proposed a multi-modal proportional fairness algorithm to achieve this goal. Furthermore, they evaluated the impact of D2D in a two-tier HetNets scenario, where macrocell and microcell network coverage are combined. Simulation results showed that, although I2D retains a clear edge for general-purpose downloading, D2D is a viable solution for localized transfers as well as viral content, but apparently cannot entirely be replace the deployment of small-cells. It is useful to complement them and mitigate the effects of a reduced infrastructure deployment together with enabling new services based on proximity in a scalable way [29].

2.2.2 Macro and Small-Cells Networks

The densification of cellular networks, based on a large scale deployment of small-cells presents a promising solution that meets the requirements of the 5G wireless systems, in terms of network capacity and throughput. Considering the 5G IoT scenario, the existing licensed spectrum assigned to, by networks carrier's, act as the solution to the limited spectrum shared by frequency reuse and spatial diversity, through the deployed small-cells, [30]. This densification via small-cells with different sizes in addition to macro base stations constitute HetNets with high level of performance and service quality [31]. The advantages of this approach over macrocell enhancements includes the fact that the cost of deployment of small cells in HetNets is much lower than that of macrocells. Unlike a macrocell, where a significant portion of the recurring cost comes from fiber to each cell site, power usage, and real estate, there is no big operating cost in user deployed small-cell networks. Again, HetNets are energy-efficient as they can be utilized intelligently and opportunistically. Depending on the traffic demand, small-cells can be in a dormant state, to alleviates energy consumption and interference management. Lastly, hyper-dense HetNets can realize the always best-connected principle by seamless handover and smart offloading. Proximity-based over-air congestion control and fast inter-cell load balancing in HetNets improves the overall spatial reuse, as many small-cells are deployed indoors, offloading is not only for indoor user traffic, but also the outdoor traffic to indoor small cells providing a huge gain [32, 33]. The characteristics of small-cells in terms of size, number of users, power and range is shown in Table 2.

Table 2: Comparison of small-cells [33]

Licensed Small-Cells				
	Femto	Pico	Micro/Metro	Macro
Indoor/Outdoor	Indoor	Indoor or Outdoor	Outdoor	Outdoor
Number of users	4 to 16	32 to 100	200	200 to 1000+
Maximum output power	20 to 100mW	250mW	2 to 10W	40 to 100W
Maximum cell radius	10 to 50m	200m	2 km	10 to 40 km
Bandwidth	10 MHz	20 MHz	20,40 MHz	60 to 75 MHz
Technology	3G/4G/Wi-Fi	3G/4G/Wi-Fi	3G/4G/Wi-Fi	3G/4G
MIMO	2×2	2×2	4×4	4×4
Backhaul	DSL,coble,fiber	Microwave, mm	Fiber, microwave	Fiber, microwave

2.2.2.1 Macrocell-Networks - From the deployment of the first conventional cellular network architecture, macrocell or microsite have been used to provide radio-access coverage (in kilometers) served by a high macrocell base station (MBS) power cell site incorporated with antennas, towers, or mast. Particularly, macrocells provide larger coverage than the entire existing small-cells category i.e., microcell, picocells, femtocells. Macrocells exhibits the characteristic of ground-based masts antennas mounted, rooftops and other existing structures, with a reasonable height that provides a clear view over the surrounding buildings and terrain. In addition, MBS's are associated with a high power output of typically 10 of watts. Hence, the performance of macrocell can be improved by increasing the efficiency of the transceiver.

2.2.2.2 Microcell-Networks - The micro-cellular concept and its implications for cellular telecommunications in urban areas have been widely understood for their capability to support low-cost devices and improve the coverage areas, of the 2G GSM networks [34]. As depicted in Table 2, they are good for outdoor deployment, with the capability to accommodate up to 200 end-users over a cell radius of 2 km [33]. Incorporated within macrocell coverage network, the benefits of higher spectral-efficiency and higher data-rate are achieved [34].

2.2.2.3 Picocell-Networks - Picocell's networks classified as regular eNBs, long existed since the 1990s [35], and are uniquely identified with the characteristics of lower transmit power than the traditional macrocells. They are, typically, perfect when shorter transmission distance coupled with low power are considered, to enhance both the achievable capacity and signal-to-interference-noise-ratio (SINR) within the cell [36]. Moreover, they are equipped with full-duplexed-directional antennas, i.e., not partitioned. Picocells are deployed indoors or outdoors in order to provide cellular coverage within residential and wireless coverage of urban areas. They have the capability to transmit maximum power between 250 mW and approximately 2 W for outdoor deployments, while utilizing 100 mW or less for indoor end-users deployments [33].

2.2.2.4 Femtocell-Networks - Femtocells are categorized as the smallest small-cells of the digital communication networks; hence, they include the characteristics of low power, short range, small size, and low cost. They are specifically deployed by consumers, to act as home base stations (BS) overlaid in the macrocell networks to provide enhanced coverage for indoor users and improve the capacity of the whole system [37], [33]. The connection of femtocell is established through the broadband communications links such as fiber optic or digital subscriber line (DSL) links [38]. From the end-user side and network operator side, the technology has brought many benefits. Although, end-users are placed as the prime entities to benefit from deploying femtocells, as they enjoy better throughput, signal quality, and affordable power saving. Network operators also yield great benefits i.e., improved network capacity and spectral-efficiency [39]. Femtocells enables reduced transmit power while maintaining good indoor coverage. Penetration losses insulate the femtocell from the surrounding femtocell transmissions. To summarize [38], the capacity benefits of femtocells are attributed to:

- Reduced distance between the femtocell and the user, which leads to higher received signal strength.
- Low transmit power, and interference mitigation from neighboring MUEs and FUEs, due to outdoor propagation and penetration losses.
- Femtocells accommodate around one to 16 users. They devote a larger portion of their resources (transmit power and bandwidth) to each subscriber. A macrocell, on the other hand, has a larger coverage area (10 - 40km radius) and a larger number of users, which makes it difficult to provide QoS to all users.

Femtocells are classified according to the type of access control they have towards the system modeled [30], and include;

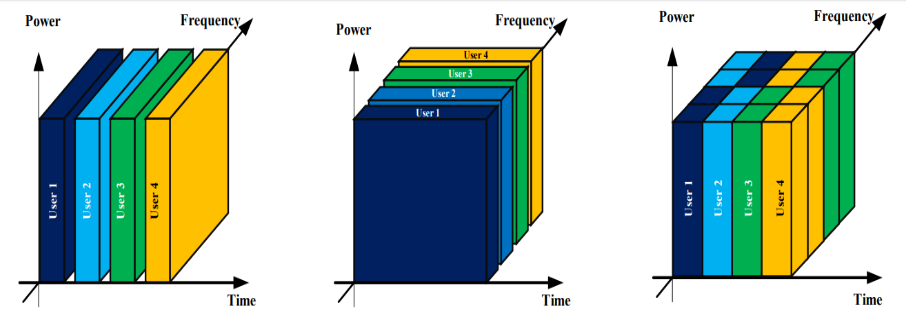
- A closed subscriber group (CSG) is a type of access which is limited by the number of users connected to the femtocell coverage network area. Considering a femtocell configured in a CSG mode, femtocells resource access will only be permitted to those users within the femtocell access control list. This type of femtocell access control strategy is usually applicable in residential deployment scenarios.

- On the other hand, a femtocell configured in an open access group (OSG), allows access to any user in the femtocell. However, in public places such as airports and shopping malls, open access mode of femtocells can also be used where any user can access the femtocell and benefit from its services. This access mode is usually used to improve indoor coverage.
- In hybrid access mode, non-registered and registered end-users can access the femtocell, but preference would be given to those users subscribed to the femtocell. Thus, limiting the backhaul connection and femtocell load access. Hybrid access mode is preferable for small business or enterprise deployment scenarios.

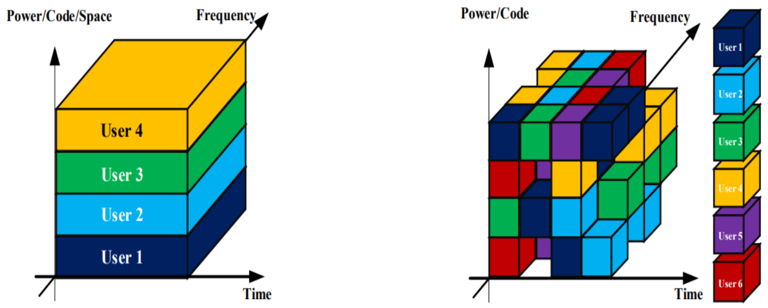
The co-existence of femtocells networks comes with potential challenges including technical, business, and regulatory issues that need to be addressed further. One of the major challenges is interference management i.e., between neighboring femtocell equipment, and between femtocell and macrocell equipment. In general, two types of interferences that occur in a two-tier HetNets architecture includes cross-tier and co-tier interference, which is well discussed in the sections to follow. Severe interference may lead to "Deadzones," i.e., areas where the QoS degrades significantly [40]. Hence, deadzones are created due to the asymmetric level of transmission power within the network and the distance between the macrocell user and MBS. Thus, it is essential to adopt effective and robust interference management schemes to mitigate co-tier interference and reduce the cross-tier interference considerably in order to enhance the throughput and spectral-efficiency of the overall network. Other challenges in femtocell deployments include handoff and mobility management, timing and synchronization, auto-configuration, and security [40, 41]. An effective and efficient mobility management and handover schemes (macrocell-to-femtocell, femtocell-to-macrocell, and femtocell-to-femtocell) are necessary for mass deployment of femtocells in UMTS and LTE networks. The scheme should have low complexity and signaling cost, deal with different access modes and perform proper resource management beforehand for efficient handover. Timing and synchronization are other major challenges for femtocells since synchronization over IP backhaul is difficult, and inconsistent delays may occur due to varying traffic congestion. Since the femtocells are required to operate on a "plug-and-play" basis, it is important that femtocells can organize and configure autonomously and access the radio network intelligently so that they only cause minimal impact on the existing macrocell network. Since femtocells could be vulnerable to malicious attacks (e.g., masquerading, eavesdropping, man-in-the-middle attack etc.), enhanced authentication and key agreement mechanisms are required to secure femtocell networks [40, 42, 43]. This work focuses on mitigating the intra and cross-tier challenges of the femtocells, to maximum the overall system energy-efficiency.

2.3 5G Multiple Access Schemes

5G wireless networks face various challenges in order to support large-scale heterogeneous traffic and users, therefore new modulation and multiple access (MA) schemes are being developed to meet the changing demands [44]. An overview of the most currently developed modulation and multiple access schemes for 5G networks is comprehensively detailed in this subsection. MA schemes were always regarded as the landmark of each mobile communication generation from 1G to 4G i.e., FDMA for 1G, TDMA for 2G, CDMA for 3G, OFDMA and single-carrier FDMA (SC-FDMA) for 4G. These are categorized as OMA schemes as shown in Fig 2, in which users are separated via orthogonal resources using frequency, time or code domain resources, making it possible to build a system with a low complexity linear receiver [45], [46]. With the proliferation of data over wireless channels not seen in the previous generations, NOMA schemes present a promising technology towards the emerging 5G networks. Hence, NOMA is is the focus of this work, and will be explored in details.



a) TDMA b) FDMA c) OFDMA



d) CDMA/SDMA e) Possible NoMA Solution

Fig. 2: Wireless communication Multiple Access schemes [6]

2.3.1 Non-Orthogonal Multiple Access (NOMA)

From the evolution of mobile communication networks, radio access technologies have been the underlying physical link for cellular networks, which are implemented in the radio access networks (RANs). Basically, RAN operates using multiple access channels to provide transmission between mobile-terminal connections to the core networks. Particularly, multiple access techniques development is one of the foremost aspects of improving system spectral-efficiency and capacity [47]. Multiple access techniques can broadly be categorized into two different approaches, namely, OMA and NOMA. In the conventional OMA, different users can only be served orthogonally in either time or frequency to alleviate inter-user interference [48]. The number of available simultaneously accessing users is strictly proportional to the number of orthogonal resources, which is limited. NOMA was first introduced in [49], as a promising multiple access scheme for future radio access (FRA) towards the 2020s and beyond. It is an efficient way to share radio resources and can root back to the network information theory [50]. Massive connectivity of users access is realized in NOMA by multiplexing multiple users over the same time-frequency channel within the same cell, while offering a number of advantages i.e, low transmission latency, improved spectral efficiency (SE), relaxed channel feedback, and higher cell-edge throughput. Multiuser detection (MUD) algorithms, such as successive interference cancellation (SIC) are employed to detect the desired signals at the receiver sides [11]. Generally, NOMA is categorized into cooperative and non-cooperative schemes where non-cooperative schemes include power-domain and code-domain NOMA schemes [51].

– **NOMA Transmission** - Thanks to Cover in 1972 [52], who proposed the concept of superposition Coding (SC), which is the underlying building blocks of multiple coding schemes motivated for achieving the capacity of the scalar Gaussian BC [53]. More particularly, Bergman's article paper published in 1973 [54] and Gallaghers' [55] showed theoretically that using the SC, capacities approaching both the Gaussian broadcast channel and general broadcast channel can be achieved. The objective of superposition coding is to communicate two message simultaneously by encoding them into a single signal in two layers, as shown in Fig 3 [56]. The design technique for SC using a finite library of finite-block length point-to-point codes was proposed in [57] for finite constellations. Hence, SC was used as the optimal scheme with SIC at the receivers [58]. In addition, [59] and [60], further proved that based on information theory SC incorporated with non-orthogonal multiplexing at the transmitter and SIC at the receiver, the capacity region of the downlink broadcast channels not only outperforms classic orthogonal multiplexing but is also optimal. Inspired by the fundamentals of information theory, researchers were motivated to investigate SC with diverse channels, such as Raleigh channels, interference channels, and multiple access channels. Despite the fact that, the aforementioned contributions have played a vital role, with the use of SC from a theoretical perspective, further research improvements have been made to evolve the concept of SC technique from theory to practice [57, 61].

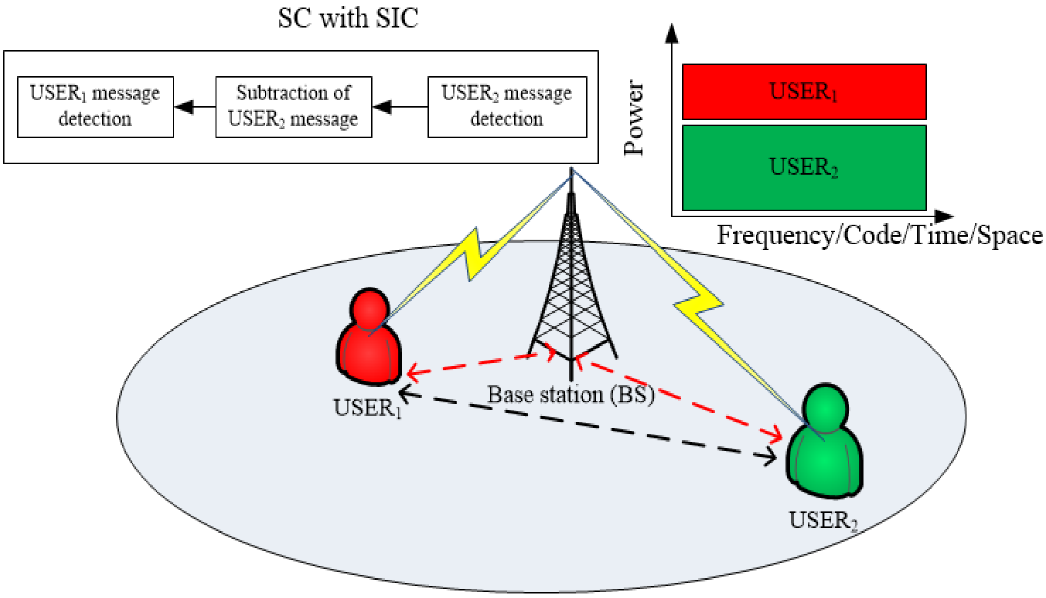


Fig. 3: Uplink superposition Coding with SIC.

- **NOMA Receiver** - To decouple the superimposed signals at the receiver, SIC technique was proposed in [52] as an efficient interference management scheme to improve network capacity, as shown in Fig 3. Particularly, the fundamental idea behind SIC is successively decoding user signals according to their signal strength. The procedure for performing SIC procedure is as follows; After one user with the strongest signal power is decoded, it is then subtracted from the coupled signal before the succeeding user's signal is decoded [11]. NOMA implements SIC MUD as the baseline receiver for robust multiple access schemes, considering the evolution of device processing capabilities, SIC in the receiver has a physical layer capability that allows a receiver to decode signals that are detected simultaneously. Furthermore, the challenges of the conventional SIC were discussed in [62], and a new secure SIC scheme was proposed using MAC address and IMEI, which are dedicated to the smartphone as private keys. Considering an uplink transmission scenario, where a single-cell BS receives superposed coded information of two user equipment's UEs, equipped with single antennas at both transmitter and receiver. Assuming, an overall system transmission bandwidth B of 1Hz, and users defined as UE$_1$ and UE$_2$ which transmits the signals $\{x_1\}$ and $\{x_2\}$ multiplexed to $\{x_i\}_{i=1}^{2}$, with $E[x_i]^2$, and the transmission power denoted as $\{P_i\}_{i=1}^{2}$. The sum of P_i is limited to P^{max} denoted as the transmit maximum power QoS. Then, the superposed received signal at BS is denoted as

$$y = h_1\sqrt{P_1}x_1 + h_2\sqrt{P_2}x_2 + z, \tag{1}$$

where h_1 and h_2 are the complex channel gain from the corresponding UE$_i$ to the BS. Without the loss generality, we assume that $|h_1| > |h_2|$, and both are perfectly known, z denotes the AWGN characteristics , with the power density given by σ_z^2. The superimposed signals of the UE$_i$ are successfully decoded at the receiver, with the aid of SIC. The optimal order for decoding the UE$_i$ is to decode the UE$_1$ with the higher channel gains followed by the UE$_2$ which has less channel gains. Additionally, the channel gains are normalized by the spectral noise power density and inter-cell interference characteristics, as follows $|h_i|^2/N_{0,i}$. Based on this order, assuming an error-free detection, the throughput of each UE$_i$, denoted as R_i, is given by

$$R_1 = B\log_2\left(1 + \frac{P_1|h_1|^2}{\sigma_z^2}\right), \tag{2}$$

$$R_2 = B\log_2\left(1 + \frac{P_2|h_2|^2}{P_1|h_2|^2 + \sigma_z^2}\right). \tag{3}$$

From (3) and (4), it can be concluded that the power-allocation for each UE greatly impacts the UEs throughput performance and thus the modulation and coding scheme (MCS) used for data transmission of each UE$_1$. Considering users that are distributed in a single cell cellular network tier, with SIC at the receivers, The boundary of the capacity region is given by

$$R_k = B\log_2\left(1 + \frac{P_k|h_k|^2}{\sum\limits_{j\neq k}^{K} P_j|h_k|^2 + \sigma_z^2}\right), \tag{4}$$

for all possible splits $P = \sum\limits_{k=1}^{K} P_k$ of the total power at the base station. The optimal points are achieved by superposition coding at the transmitter and SIC at each of the receivers. The cancellation order at every receiver is always to decode the weaker users before decoding its own data. Hence, the complexity increases with the increase in the number of the users.

The basic idea of NOMA is encouraged by spectrum sharing, through intra-cell multi-user multiplexing scheme that utilizes code-domain or power domain. Thus, the key distinguishing attribute of NOMA from OMA is supporting a high number of connected users than the number of orthogonal resource elements utilized, with the aid of non-orthogonal resource allocation. This may be perceived by the sophisticated interference management cancellation schemes at the cost of increased receiver complexity [59]. A comprehensive comparison of the G-NOMA technologies was done in [59] and [63], where their key features, principles, advantages, and disadvantages as a major solution to improve spectral-efficiency, system capacity and receiver complexity were provided. Moreover, challenges and opportunities for future researchers in this field, regarding NOMA trend design were also emphasized to provide insight on the potential future works. Finally, the concept of software-defined multiple access (SoDeMA) was proposed, to allow various MA schemes including both traditional OMA and current NOMA, to enables adaptive configuration to support diverse services and applications for the future 5G networks. Regarding the issues of resource allocation, investigations were also conducted for NOMA apart from performance analysis. In [64], a NOMA resource allocation for uplink was proposed considering a single cell. The authors in [65] proposed a solution to optimize the combinatorial problem of the receiver and radio resource allocation of uplink NOMA. The proposed scheme was iterative in nature, multiuser detection and decoding (MUDD) to improve the performance of the multiuser by utilizing the information derived from the channel decoder was considered. In addition, a novel subcarrier and power allocation algorithm was proposed that maximizes the users' weighted sum-rate in [66]. In [67], a resource allocation scheme was developed for a downlink multi-user system to facilitate dealing with a large number of users in a typical NOMA system. Driven by the exponential increase of the energy cost, a resource allocation scheme that jointly optimizes the power and bandwidth, in order to maximize the system energy-efficiency was proposed in [68]. The scheme achieved better energy -efficiency performance compared to the conventional OMA. The authors in [69] proposed a sub-optimal subchannel assignment and power allocation to maximize the energy-efficiency for the downlink NOMA network, which showed superiority in terms of sum rate and energy-efficiency when compared to the conventional OMA scheme. However, NOMA techniques have been refined into G-NOMA schemes, each dealing with different aspects.

2.3.2 Generalized NOMA technologies

The emerging 5G networks are expected to significantly increase the network data rates, throughput, massive connectivity of user access and devices, improved spectrum efficiency, low latency, and high EE, when compared with the currently deployed LTE or LTE-A networks. To meet such demands for the 5G networks, innovative technologies on RAN and radio air-interface are more than desired in physical layer designs. Recently the development of G-NOMA schemes has attracted increasing research interests from both academic and industrial fields. Developed G-NOMA schemes include, SCMA, MUSA, PDMA, and resource spread multiple access (RSMA) proposed by Huawei, ZTE, DT mobile, Qualcomm, respectively. In this subsection, we provide a comprehensive overview of the G-NOMA technologies for the 5G networks.

2.3.2.1 Cooperative NOMA Cooperative communications in wireless networks have gained a great deal of attention due to its ability to offer spatial diversity to mitigate fading while resolving the difficulties of mounting multiple antennas on small communications terminals [70]. The basic idea underlying the concept of cooperative communications is deploying relay nodes to communicate forwarding information from the source to the desired destinations. Therefore, cooperative communications integrated with NOMA yields further improvements in system capacity and transmission reliability. The authors in [71] investigated a cooperative NOMA scheme which exploited prior information available in NOMA systems. The scheme relied on users with better channel conditions to decode the messages for the others, to improve reception reliability for users with poor channel conditions. Hence, cooperative communications for users with better channel conditions than others can be employed by using short-range communications techniques, such as ultra-wideband and Bluetooth [11].

2.3.2.2 Cognitive Radio Inspired NOMA To further enhance the issues of spectrum scarcity in wireless communications, cognitive ratio (CR) networks, which provides the capability to enable unlicensed users to utilize licensed spectrum resource in an opportunistic manner without causing interference to the licensed users have been proposed as a solution. CR networks are envisioned as a key enabling technology of dynamic spectrum access techniques to provide high bandwidth to mobile users via heterogeneous wireless architectures [72, 73]. More particularly, the concept of spectrum sensing technique which is the enabler for CR has been well investigated in [74, 75], together with the trade-off between sensing overhead and CR throughput. CR network is regarded as a special case for the NOMA technology perceived in location domain while the classic NOMA can be realized in power, code, or multiple domains [76]. For example, considering an unlicensed user that shares the same licensed spectrum, with a licensed user limited by the distance constraint, the interference generated by the unlicensed user will be generally small while achieving improved data-rates. The authors in [77], proposed a special case for CR networks, termed D2D communications which gained popular research interest lately, regarded as a simple CR network for a two user case. Furthermore, location domain enables CR to be exploited in spatial domain [78] and the frequency-spatial domain [79]. The convergence to cognitive-radio-inspired NOMA (CRNOMA) was proposed in [80], which was totally different from the conventional scheme with fixed power allocation. The closed-form expressions of the outage probability have been derived in [81] for NOMA-based underlay CR networks. In addition, power allocation schemes for CR-NOMA and NOMA-based CR networks have been developed in [82], to maximize the system sum-rate and EE.

2.3.2.3 Power-Domain NOMA Power-domain NOMA is regarded as the fundamental MA scheme for the 5G networks [83]. More specifically, it has been regarded as multiuser superposition transmission (MUST) for the downlink version of NOMA, which was proposed for the 3GPP-LTE-A networks in [49]. In literature, it has been shown that NOMA can improve system capacity and user experiences. Hence, recent work [84], investigating a downlink multiuser superposition transmission for LTE networks has been accepted by 3GPP LTE Release 14, which aimed to identify the necessary techniques to enable LTE to support the downlink intra-cell multiuser superposition transmission. In [11], the basic principles of various power-domain NOMA related techniques, including power allocation in NOMA, multiple antenna based NOMA, and cooperative NOMA were investigated. Power-domain NOMA realizes multiplexing by allocating different power to multiple users within the same time/frequency/code spectrum resource elements according to their channel conditions. It should be noted that excellent performance can be achieved when a cell-center user and cell-edge user are scheduled with moderate computational complexity when the SIC is applied in the receiver. However, power-domain multi-user superposition poses additional challenges due to more interfering neighboring users. Hence, properly transmit power control for NOMA users within the same cell becomes critical.

2.3.2.4 Code-domain NOMA

- **Sparse Code Multiple Access** SCMA was proposed in [85], as non-orthogonal multiplexing scheme, which applies the concept of sparse codebook similar to the low-density spreading (LDS) signature matrix, as shown in Fig 4. Spectrum resource elements are utilized to realize the massive connectivity of users through spreading codes [86]. SCMA utilizes the property of multi-dimensional constellations to avoid incurring collisions, while reducing the sophisticated receiver complexity, in order to enhance the massive connectivity of the IoT and spectral-efficiency. Following the attribution of SCMA to the multi-dimension property, one resource block can be projected into its subspace from the constellation referred as a codebook. Due to the sparsity of the codewords matrix and the minimum distance of the multi-dimensional constellation lattice, the message passing algorithm (MPA) detection is employed [85, 87].

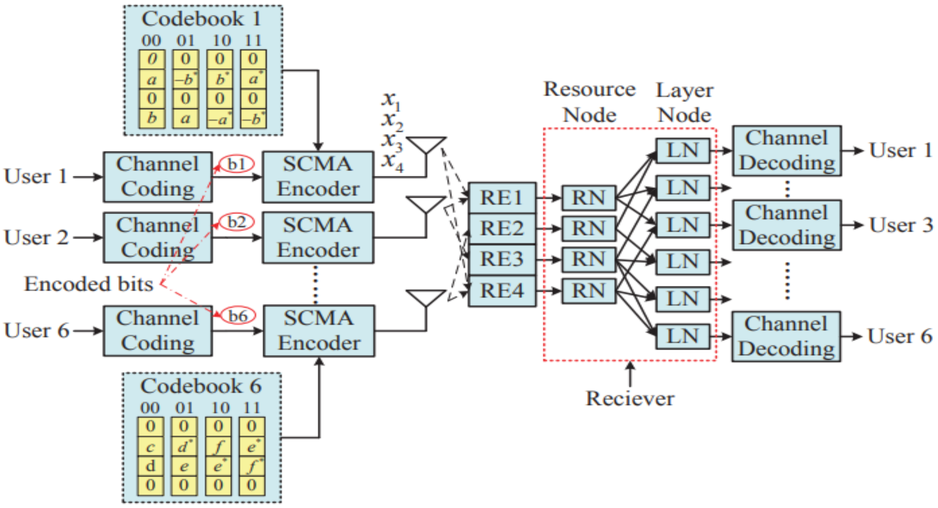


Fig. 4: SCMA transceiver model [88].

Regardless of the sparse signature sequences, SCMA is associated with the challenges of high detection and decoding complexity. Hence, the complexity is even higher when a large number of users and constellation are employed. In addition, because the codebook size is limited, collision is a serious problem in SCMA, which hinders its overloading capacity [87]. Therefore, based on the design principles of lattice constellations, the authors in [89] proposed a systematic approach to design SCMA codebooks which followed the code design for point-to-point communication over fast fading channel. Furthermore, codebook design based on multi-dimensional SCMA (MD-SCMA) considering interleaving and constellation rotation method was proposed for the downlink in [90], which was constructed by the subset of lattice Z^2. Lastly, an improved method based on the star-QAM signaling constellations was proposed in [91] for designing the SCMA codebooks. With regard to receiver complexity reduction, the authors in [92] and [93] proposed a low complexity decoding techniques taking advantage of the SCMA codebook structure. In addition, the emphasis on iteration reduction, convergence speedup, computation simplification, and implementation of SCMA decoder based on deterministic message passing algorithm, where addressed in [94]. The author in [95] proposed an SCMA-based uplink grant-free multiple-access system based on the receiver design utilizing the concept of blind detection to enable massive connectivity of IoT. Moreover, the authors [96] proposed a new approach for multiple access in the 5G networks called power domain sparse code multiple access. In addition, a novel resource allocation problem for PSMA-based HetNets with the aim of maximizing the sum rate with certain constraints was proposed. In [97] a unified framework to analyze the EE of SCMA scheme and a low complexity decoding algorithm was proposed for prototyping.

- **Multi-User shared access -** MUSA is one of the possible non-orthogonal transmissions and grant-free multiple access schemes for the 5G communication systems, which operates both in power-domain and code-domain [87]. Data of each user is spread with a family of short length complex spreading codes that are designed to support massive connectivity while minimizing signaling overhead and power consumption between users that share the same resource block, as depicted in Fig. 5. [51, 87, 98].

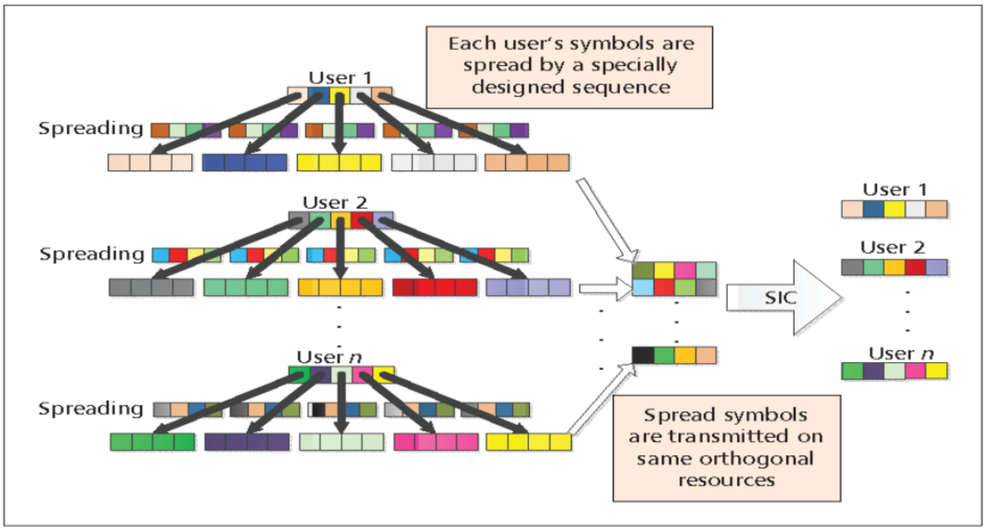


Fig. 5: MUSA transceiver model [45].

With characteristics of grant-free access and non-orthogonal transmission, each transmitting user can autonomously choose its spreading code, therefore eliminating the need for resource coordination by BS [98]. [63]. Blind detection is applied in MUSA, in which the minimum mean square error successive interference cancellation (MMSE-SIC) is employed to mitigate interference between users. The supremacy of MUSA is reflected by high overloading, robust blind detection and the true sense of grant-free transmission. Moreover, the possibility of collision due to same spreading sequences is significantly small since a large number of spreading codes can be accommodated. Additionally, user detection is realized without the knowledge of the spreading codes. However, one of the main challenges in MUSA is performance loss due to error propagation caused by the implementation of SIC [87]. Thus, the design of the complex spreading sequence is important to MUSA since it determines the interference between different users and system performance. Moreover, the impact on the complexity of SIC implementation also needs to be considered. The authors [99], studied the design of the non-binary spreading sequence to achieve low cross-correlation with very short length.

2.3.2.5 Interleaved-Based NOMA The IDMA-framework was proposed by L.Ping at el in [100, 101] as a NOMA scheme, to accommodate massive connectivity of user. The basic idea behind IDMA is employing user-specific interleavers combined with low-rate channel coding in separating users as shown in Fig 6. It is a wideband scheme that allows the use of multiuser detectors with moderate complexity. IDMA inherits desirable feature characteristics with CDMA, especially mitigation against fading to coordinate user interference from other cells [102]. These interleavers can be selected randomly (or deterministically for practical convenience) and orthogonality is not essential in IDMA. The main advantages of IDMA are its high user overloading and high spectral efficiency. Hence, [103] outlines the attractive features of IDMA as follows:

- IDMA facilitates a low-cost iterative technique for multi-user detection (MUD).
- IDMA with proper power control achieves multi-user sum-rate close to the near-capacity.
- IDMA with decentralized power control can offer significantly higher throughput than conventional ALOHA in random-access environments.
- IDMA together with data-aided channel estimation (DACE) can fully take advantage of massive multiple-input multiple-output (MIMO) systems.

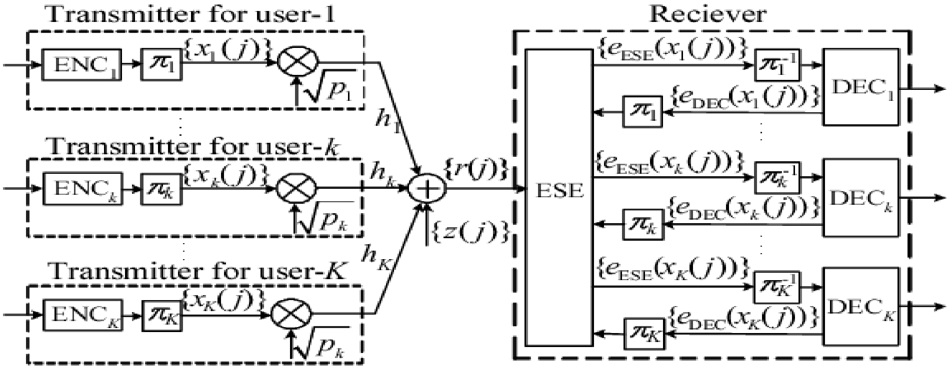


Fig. 6: IDMA transceiver model [45].

With all the benefits brought by IDMA, large decoding complexity and decoding latency, remains the main challenges of IDMA, especially when a large number of users is supported. In addition, interference among users is inevitable in IDMA, but can be suppressed by a low-cost iterative multi-user detection (MUD) procedure also called chip-by-chip detection [104]. Again, the additional pilots or long preamble needed to estimate the user's channel also remains a challenge [87]. The challenges of transmission and detection principles of IDMA where investigated in [105]. In [106], a power allocation based on linear programming methods problem for practically coded interleave-division multiple-access (IDMA) systems with iterative multi-user detection (MUD) over multiple access channels (MACs) was investigated. Furthermore, an optimized power allocation technique based on an interior-point method (IPM) for practically coded CDMA or interleave-division multiple-access (IDMA) systems over multiple access channels was proposed in [107], and very low-cost iterative detection algorithm derived for the IDMA scheme based on a chip-by-chip detection principle was proposed in [108]. The authors in [102] extended IDMA to MIMO multiuser systems employing spatial multiplexing. An iterative receiver for MIMO-IDMA that incorporated an efficient soft multiuser detector whose complexity is linear in the number of users was further proposed.

2.3.2.6 NOMA Multiplexing in Multiple Domains PDMA, a novel-NOMA scheme based on code patterns was proposed in [109]. Non-orthogonal patterns are used to realize multiplexing of different users mapped on the same spectrum RE as shown in Fig. 7. PDMA can significantly improve spectral-efficiency and massive connectivity of user's access. The patterns are carefully designed to cope with multiple domains i.e., code, power, and space domain, to gain SIC-amenable characteristics. In light of this, a low-complexity SIC-based MPA MUD algorithm with reliable performance is employed at the receiver side. PDMA sparse pattern matrix is similar to that of SCMA, with the difference in the number of spectrum resource elements occupied by each user [110], [111].

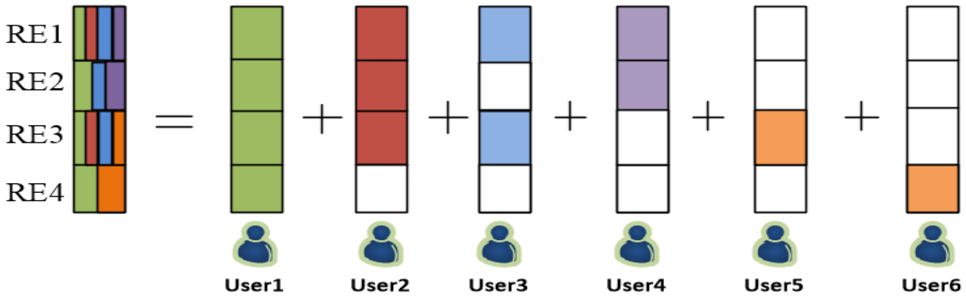


Fig. 7: Resource mapping of PDMA pattern for six users on four Resource [45].

The appropriate disparity in diversity and power control where investigated in [111], with aim of effectively exploiting a low-complexity SIC-based detector to realize the near-perfect cancellation of multi-user interference. Hence, link-level simulations illustrated that PDMA is capable of accommodating a 300 percent overload, while it still enjoyed the transmission reliability close to conventional OMA schemes [111]. However, the main challenge of PDMA is its high complexity for higher modulation order when joint detection is applied. Additionally, as a result of a limited number of patterns, high probability of collision occurs when grant-free access is assumed. Channel estimation is a big problem due to the requirements of additional pilots or preamble needed [87]. The authors in [110] exploited the advantages of PDMA in a downlink wireless network which was fully loaded. A power-domain PDMA using SIC to improve system throughput and enhance the spectrum efficiency was further proposed. Furthermore, the investigation performance of the outage probability, as well as the achievable sum data rate for uplink PDMA system, was carried out in [112]. In particular, the closed-form expressions of outage probability for three users sharing on two resource blocks with interference cancellation receiver was investigated. The analysis transmission latency of grant based PDMA and grant free was investigated in [113]. Hence, a detailed definition of GF-PDMA resource and analysis of the scalability to confront pattern collision for GF-PDMA was proposed.

2.3.2.7 Hybrid Non-Orthogonal Multiplexing Domains Recently, related work has emerged to investigate resource allocation problems for hybrid NOMA systems to optimize the system data rates and spectral-efficiency. To further facilitate sum-rate optimization, a hybrid multiple access, which combines the properties of NOMA and OFDMA was proposed in [67]. Simulation results confirmed the superiority of the proposed hybrid multiple access scheme over conventional NOMA. The authors in [114], proposed a hybrid MC-NOMA resource allocation model which incorporates NOMA and OMA to fully exploit the potential advantages of the two kinds of multiple access schemes. In the proposed algorithm, the resource allocation included the selection of multiple access modes, user clustering, subcarrier assignment and power allocation. By performance analysis, it was found that the hybrid MC-NOMA significantly outperformed the conventional MC-NOMA and OMA in terms of user fairness and system efficiency. Furthermore, a hybrid transmission scheme which consists of the NOMA and TDMA was proposed in [115]. However, the corresponding energy minimization problem was too complex to deal with, compared to the traditional NOMA schemes. Additional, the above contributions, only focused on sum-rate optimization. Therefore, this motivated us to propose a new transmission scheme called HG-NOMA and the corresponding energy-efficiency resource optimization problem. Hence, the proposed HG-NOMA scheme combined the PD-NOMA and PDMA schemes into one unified framework for a two tier HetNets.

2.3.3 Comparison of NOMA with OFDMA

We discuss the basic NOMA with SIC and analyze its performance gain compared to OFDMA. We consider, a popular NOMA scheme that uses power domain to achieve multiplexing.

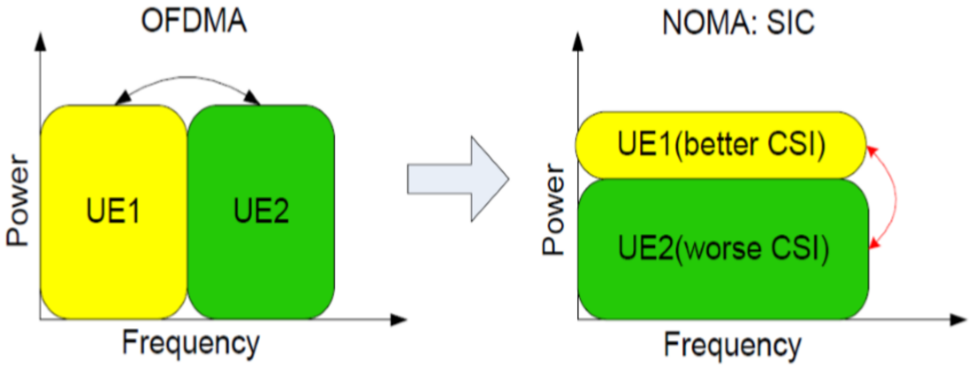


Fig. 8: Comparison of NOMA and OFDMA. [116]

Considering an uplink NOMA, with users UE$_1$ and UE$_2$, and transmits the modulated symbols x_1 and x_2, with the transmission power P_1 and P_2, as shown in Fig. 8. The received signal at the BS is given as

$$y = h_1\sqrt{P_1}x_1 + h_2\sqrt{P_2}x_2 + z, \tag{5}$$

where h_1 and h_2 are the channel gain from UE$_1$ and UE$_2$ to the corresponding BS, z denotes the AWGN noise characteristics, with the power spectral density given by σ_z^2. In NOMA, x_1 and x_2 are transmitted using the same time-frequency channel and interfere with each other. In the uplink, the SIC is implemented at the BS receiver. With the SIC, the receiver decodes x_1 and x_2 in two stages. In the first stage, the receiver decodes x_2, treating x_1 as Gaussian interference. Once the receiver correctly decodes x_2, it can subtract the x_1 component from aggregate received signal y and then decode x_1. Since only the receiver Gaussian noise is remaining in the system, the maximum throughput is achieved for UE$_i$ the same as its single-user bound. Thus, the throughput of user i, R_i, is represented as

$$R_1^{NOMA} = B\log_2\left(1+\frac{P_1|h_1|^2}{\sigma_z^2}\right), \quad R_2^{NOMA} = B\log_2\left(1+\frac{P_2|h_2|^2}{P_1|h_2|^2+\sigma_z^2}\right). \tag{6}$$

If the decoding order is opposite, R_i becomes

$$R_1^{NOMA} = B\log_2\left(1+\frac{P_1|h_1|^2}{P_2|h_1|^2+\sigma_z^2}\right), \quad R_2^{NOMA} = B\log_2\left(1+\frac{P_2|h_2|^2}{\sigma_z^2}\right). \tag{7}$$

Note that in both (6) and (7), regardless of the decoding order in the SIC, the total throughput, $R_1 + R_2$, is the same, and is given by

$$R_{NOMA} = B\log_2\left(1 + \frac{P_1|h_1|^2}{\sigma_z^2}\right) + B\log_2\left(1 + \frac{P_2|h_2|^2}{P_1|h_2|^2 + \sigma_z^2}\right). \tag{8}$$

This is indeed the maximum total achievable throughput in this uplink multiple access channel. On the other hand, for uplink OFDMA, the total bandwidth B is equally occupied by UE$_1$ and UE$_2$ (assume that each user has half bandwidth). Therefore, the data-rates of UE$_1$ and UE$_2$ in OFDMA, can be respectively represented as

$$R_1^{OFDMA} = \frac{1}{2}B\log_2\left(1 + \frac{P_1|h_1|^2}{\sigma_z^2}\right), \quad R_2^{OFDMA} = \frac{1}{2}B\log_2\left(1 + \frac{P_2|h_2|^2}{\sigma_z^2}\right). \tag{9}$$

Thus, the sum rate can be written as

$$R_{OFDMA} = \frac{1}{2}B\log_2\left(1 + \frac{P_1|h_1|^2}{\sigma_z^2}\right) + \frac{1}{2}B\log_2\left(1 + \frac{P_2|h_2|^2}{\sigma_z^2}\right). \tag{10}$$

Case study: Consider a two UE case with a cell-interior UE$_1$ and a cell-edge UE$_2$, where $|h_1|^2 < |h_2|^2$. We set to $P_1|h_1|^2/\sigma_z^2 = 20$dB and $P_2|h_2|^2/\sigma_z^2 = 0$dB, respectively, in OFDMA, and assume that the total bandwidth B is occupied by the two users (assume that each user has half bandwidth i.e $\alpha = 0.5$) assuming proportional fairness (PF) criteria. On the other hand, for the NOMA systems, the bandwidth of the subchannel is B and the power proportional factors for UE$_1$ and UE$_2$ are β_1 and β_1, respectively. Using the parameter values in [47], and the SIC implemented at the receiver for the NOMA system, the UE rates are calculated as follows; For OFDMA system $R_1 = 3.33$ bps and $R_2 = 0.50$ bps, respectively. On the other hand, in NOMA, when the power allocation is conducted as $P_1 = 1/5P$ and $P_2 = 4/5P$, the user rates are calculated according to (7) as $R_1 = 4.39$ bps and $R_2 = 0.74$ bps, respectively. The corresponding gains of NOMA from OFDMA are 32% and 48% for UE$_1$ and UE$_2$, respectively. Therefore, $R_{OFDMA} = 3.33 + 0.50 = 3.83$ bps, and $R_{NOMA} = 4.39 + 0.74 = 5.13$ bps. Based on the above numerical analysis, it can be concluded that the sum data-rate of the NOMA system will be improved when the channel gains increase, This is because OMA must allocate most of the bandwidth exclusively to the user under good channel conditions to achieve the maximum total throughput. It is due to this that we advocate for NOMA techniques and make them a focus of this work.

3 Challenges of the 5G Networks and Mitigation Approaches

This section details some challenges examined throughout this work. Mitigation approaches for the considered challenges are also provided. While the emerging 5G networks come with great benefits in terms of enhanced throughput, data-rates, low-latency, and high spectral-efficiency, a number of challenges also arises, which includes: the development of advanced modulation and multiple access schemes, interference management, network architecture design, spectrum-resource management, and energy-efficiency. This is because of the demand for higher data rates to accommodate the heterogeneous traffic generated by the billion devices around the world. However, deploying small-cells sufficiently improves spectral-efficiency through higher frequency reuse spectrum resources, and the transmitted power is significantly reduced such that the power lost through propagation will be lower [117]. However, they also come with their own challenges. This research work focuses on the mitigation of the following challenges; multiple access schemes, interference management due to HetNets, spectrum resource scarcity and EE optimization.

3.1 Modulation and Multiple Access for 5G Networks

The emerging 5G wireless networks to support large-scale heterogeneous traffic and users presents various challenges that needs to be addressed urgently. Therefore, new modulation and multiple access (MA) schemes are being developed to meet these challenging demands brought by this traffic proliferation. Hence, it becomes more crucial to keep track of the most currently developed promising modulation and MA schemes approach for 5G networks [44]. Orthogonal frequency division multiplexing (OFDM) is the physical layer waveform technology adopted in recent wireless communication networks, which has many key features such as its robustness, ease of implementation and high spectrum-efficiency. Needless to say that, it is limited by a number of disadvantages that conflict's with 5G technology requirements in achieving MBB and IoT for many reasons. These include high peak to average power ratio (PAPR), signaling overhead in time and high energy consumption caused by the synchronization for maintaining orthogonality. Furthermore, it has high out-of-band emission (OBE), to the overhead created using cyclic prefix which needs space within the data streams [118]. This work investigates the performance analysis of the G-NOMA schemes as a potential solution to the OMA technologies, to address the issues of massive connectivity and improved system capacity.

3.2 Heterogeneous Networks

Network operators are already feeling the strain. The actual problem they are facing is not coverage, but capacity, which is now nearly universal. There are just too many mobile users and devices demanding too much data. As a solution, HetNets based on a large-scale deployment of small-cells

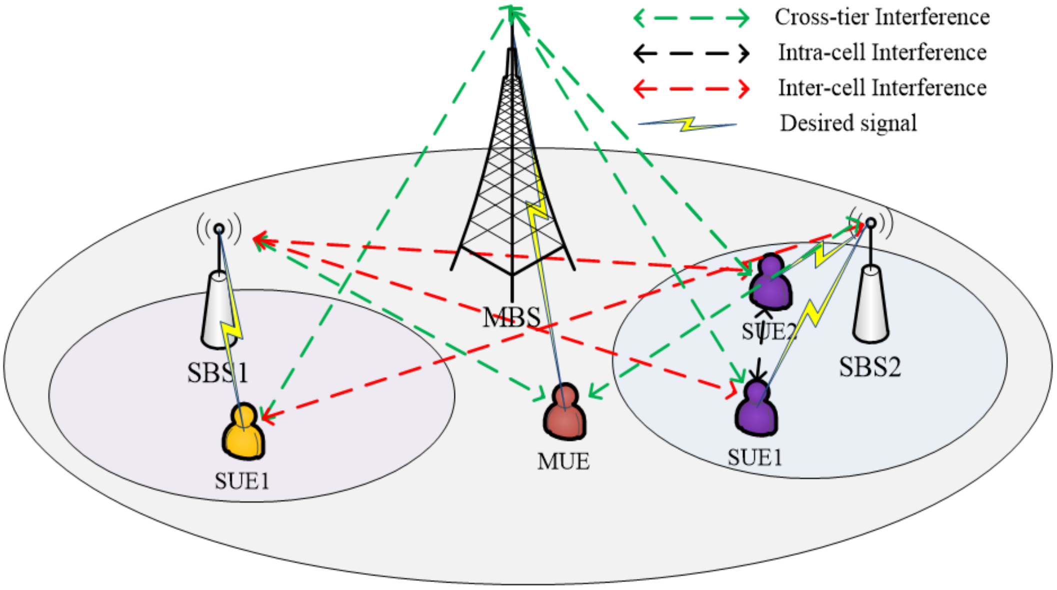


Fig. 9: Types of interference in HetNets environment [33].

with low power nodes overlaid in macrocells have been recently proposed to cope with spectrum scarcity for the 5G networks [119]. In particular, small-cells networks are motivated by the principle of rethinking network design to offer unprecedented network capacity and improved spectrum-efficiency, while minimizing energy consumption by adopting energy-efficient small-cell architectures [120]. Small-cells are typically used in enterprise and residential deployments targeting the indoor coverage, and the higher powered picocells deployed to extended outdoor coverage to fill the macrocell coverage holes [117] [121]. Nearly all benefits brought by 5G HetNets deployment affect the nature of interference management, especially with the deployment of unplanned small-cells, where network operators have limited control over the location of the small-cells. Traditional methods of interference management including frequency reuse or BS coordination no longer translates to HetNets [121]. Nevertheless, the concurrent operation of small-cells and macrocells results in irregular shaped cell sizes, causing interference challenges, which then requires advanced power control and resource scheduling methods to mitigate. Fig. 9, clearly shows an example of the type of interference in HetNets environment.

In HetNets, co-tier and cross-tier interference occur either on the uplink or downlink when the aggressor (sources of interference) and victim use the same spectrum resource elements. Table 3 provides a summary to better understand the interference between aggressor and victims.

1. The cross-tier interference occurs between network elements that belong to different tiers or network-layers. The aggressors are network elements that produce the interference, while the victims are the network elements that suffer the interference originating from the aggressor. For

example, considering the uplink-domain, the direction of interference is either from the MUE to the SBS or SUEs to the MBS as shown in Fig. 9. Similarly, in the downlink, the direction of interference is either from the MBS to the SUEs or SBS to the MUE [122].

2. The co-tier interference occurs among network elements of the same network-layer or tier which may include inter and intra small-cell interference nor inter-macrocell interference. In the case of a two-tier femtocells networks, the co-tier interference occurs between neighboring femtocells. For example, the aggressor i.e, SUEs in this case, sources co-uplink interference to neighboring SBS(s), which are then considered to be victims. On the other hand, an SBS is as a source of co-downlink interference affecting adjacent SUE(s) as depicted in Fig. 9 [122].

Table 3: Interference scenarios in two-tier HetNets [33]

Agressors	**Victims**	**Interference type**	**Transmission mode**
MBS	small-cell UE (SUE)	Cross-tier interference	Downlink
macrocell UE (MUE)	SBS	Cross-tier interference	Uplink
SBS	MUE	Cross-tier interference	Downlink
SUE	MBS	Cross-tier interference	Uplink
SBS	SUE	Co-tier interference	Downlink
SUE	SBS	Co-tier interference	Uplink
SUE	SUE	Intra-cell interference	Uplink/Dowlink

The theoretical aspects for understanding HetNets , and the practical constraints, and the potential challenges that limit operators to maximize these networks capacity were discussed in [123]. Based on range-expansion mobile association scheme, a downlink radio resource-allocation framework for OFDMA in HetNets and intracellular CoMP was investigated in [124]. In [125], resource sharing algorithms to mitigate against interference was achieved through the application of heterogeneous cloud radio access networks (H-CRAN). Furthermore, the authors in [126] proposed a decentralized resource allocation scheme to alleviate interference in HetNets. Additionally, a HetNets dynamic resource allocation IR-HARQ schemes under the presence of interference was investigated in [127]. However, most of these works on resource allocation do not feature a combination of NOMA technologies for multiple users multiplexing on HetNets. This is done in this work.

3.3 Spectrum Resource Scarcity

The overwhelmingly critical issue right now is that, we are running out of broadband spectrum. Therefore spectrum resource sharing schemes are essential to achieve the ambitious goal set by the 5G to increase the network capacity of current predecessor 4G networks by 1000 folds [128, 129]. The need to manage spectrum is motivated by the aspects and attributes of multi-dimensional resources which do not only include time, frequency, space, power, multi-user diversity but also includes computation complexity, and energy resources. Due to the ubiquitous channel state information (CSI) and QoS/QoE, the so-called 5G IoT will results in tremendous data traffic, which will cause significant signaling overhead challenges. Fortunately, the dense deployment of indoor and outdoor small-cells provides a promising solution to provision spectrum resource utilization and to alleviate interference management [130]. However, spectrum resource optimization in ultra-dense networks involves resource scheduling, interference coordination, throughput maximization [129, 131–133] i.e,. capacity [134, 135] , sum-rate [136, 137] and spectral-efficiency [138]. This work focuses on EE spectrum resource allocation for a two-tier HetNets technology, which is considered essential to alleviate interference, spectrum-efficiency and sum-rate capacity.

3.4 Energy Efficiency in 5G Networks

The evolution of mobile technologies from 4G to 5G, corresponds to the requirements of EE optimization, which is growing at a fast rate [139]. At present, ICT systems are responsible for about 5% of the world's CO_2 emissions [140], [141], since they currently deploy wireless communication systems, which are powered by traditional carbon-based energy sources. In addition, it is anticipated that 75% of the ICT sector will be wireless by 2020 [139], thus implying that wireless communications will become the critical sector to manage as far as reducing energy consumption [142]. Considering the current networks which are designed to maximize capacity by scaling up the transmit powers; such an approach is not robust, given the exponential growth of data traffic. However, using more energy to maximize network capacity results in unacceptable operation costs. Hence, a paradigm shift from throughput to energy-efficiency optimization has long begun. More specifically, a comprehensive survey of recent advances in wireless networks EE was conducted in [143]. A cooperative HetNets for 5G, which supports centralized energy saving, cooperative radio and baseband processing was studied in [144]. As EE constitutes to the key performance indicators for the 5G networks, potential solutions needs to be developed to address EE under various dynamic situations like varying traffic load, mobility, and interference conditions. Most of this approaches useful for maximizing EE in wireless networks includes; resource allocation, network planning, and deployment, hardware solutions, energy transfer and harvesting [139]. Although, the deployment of small-cells has solved the distance and transmit power challenges, interference and energy consumption are practical problems that degrades network performance. As a result, existing multiplexing schemes to minimize energy have been investigated literature. Network architectures' based on energy-efficient classifications were addressed in [120, 145, 146]. Furthermore, the authors in [147] examined the trade-off between bandwidth and EE for wireless networks. It should be noted that, EE in wireless networks EE can also be extended over different protocol layers, which consumes energy by the application of power amplifiers, mixers, processors, registers, filters etc. Therefore, energy-efficient performance metrics are primary key when optimizing EE, since they are directly related to optimization decisions across all the protocol layers [116, 148]. This thesis considers EE resource allocation, in which the system EE is defined using bit-per-joule performance metric to ensure a reliable and secure transmission, and is given as follows [142],

$$EE = \frac{Tf(\gamma)}{T(\mu p + P_c)} = \frac{f(\gamma)}{\mu p + P_c} \quad \big[\text{bits/Joules}\big]. \tag{11}$$

As a common feature, all of the numerator's performance metrics are measured in [bits/s] subject to signal-to-noise ratio (SNR) (or SINR) of the transmission link, denoted by γ. Thus, we express the system benefits by a function of $f(\gamma)$, where f is defined according to the particular performance metric to be optimized. From (11), μp and P_c denote the transmit and circuit power consumption. The circuit power consumption is simply an additional power device consumption, which includes signal processing and active circuit blocks such as: an analog-to-digital converter, digital-to-analog converter, synthesizer, and mixer during the transmission [149]. From (11) it can be observed that, EE is measured in [bits/Joule]. Thus, presenting effectiveness with which each Joule of energy is used for information transmitted.

4 Research Motivation

The 5G wireless communication network, is highly anticipated by its ability to achieve increased network capacity, higher spectral-efficiency, low-latency and massive connectivity of user's and devices, etc. Additionally, the aspect of EE in wireless networks had received attention, due to the environmental as well as economical motives for network operators. Since minimizing energy consumption to maximize network capacity incurs unacceptable energy costs and contributes toward environmental issues. Among the proposed solutions, NOMA and HetNets have been envisioned as the key technologies to offer improved network capacity and spectral-efficiency. Considering HetNets, EE is maximized by overlaying low powered small-cells in macrocell infrastructures. As for NOMA, EE maximization is achieved via power domain multiplexing. Moreover, NOMA strikes a good balance between system throughput and user fairness compared to the traditional orthogonal user scheduling. Femtocells are considered as the small-cells throughout in this work, to mitigate interference, while maximizing EE. However, the benefits of NOMA schemes have not been well investigated, and needs further development to better hybrid schemes, to meet the 5G challenges. This thesis, provides a systematic approach of combining hybrid NOMA (H-NOMA) schemes in HetNets. Particularly, such deployment brings new technical challenges such as mutual co-tier and cross-tier interference from the neighboring cells and network users that this work needs to solve. Furthermore, the application of H-NOMA in HetNets makes spectrum resource allocation more difficult and challenging. To this end, interference management and spectrum resource allocation to maximize the achievable sum-rates and EE, are the challenges that this work addresses.

5 Research Objectives

The following is an enumerated summary of the main objectives of this thesis:

1. To conduct a thorough critical literature review of the existing NOMA, referred as G-NOMA algorithm in this work, and evaluate their performance.
2. To develop HG-NOMA resource allocation algorithm featuring Additive White Gaussian Noise (AWGN) channel, Rayleigh fading channel, e.t.c. To conduct a performance analysis of the developed model in HetNets environment.
3. To modify and conduct an EE performance analysis of the developed HG-NOMA resource allocation scheme.

6 Research Methodology

In the contribution to knowledge, this work considered the following standards analytical tools, apart from simulations in Matlab and C based programming language; convex optimization and Lagrange dual method. The two techniques were considered when formulating the energy-efficiency optimization problem. Convex optimization have been long used in LTE-Advanced networks, and today, it serves as a new indispensable computational tool, which increases the ability to solve problems such as linear programming to a much richer and larger class of problems. This is discussed in this section.

6.1 Convex Optimization Approach

Resource management plays an important role to improve EE in wireless communication networks [150]. The main resource management attributes for wireless communications is frequency, time and power optimization. However, there are other different mechanisms for resource management in wireless networks, which includes: congestion control, routing and spectrum resource allocation [151]. In this thesis, we mainly focus on resource allocation (user scheduling) and power control to maximize the system EE in NOMA-HetNets. Resource allocation means that the scheduler needs to assign multiple users to a number of spectrum resources. Needless to say that, resource allocation schemes perform differently, under various channel conditions. On the other hand, power resources means that, the scheduler needs to allocate different powers to the users. Generally, these problems are analytically modelled as maximization or minimization optimization functions subject to some constraints. One of the most common and effective mathematical tool to solve problems in wireless networks is convex optimization method. In literature, convex optimization is defined as the combination of one of the three different disciplines, convex analysis, optimization, and numerical computation. Recently this tool is of importance in wireless communication and networking, enabling viable solutions to very large, practical problems such as, reliability and performance efficiency [152]. In this thesis, the convex optimization was considered when designing and formulating the EE optimization problem. To elaborate in details, let us consider a standard form of a convex problem given as

$$\min_{x \in Z \subseteq \mathrm{R}^n} f(x) \tag{12}$$

where x describes a vector known as the optimization variable, $f\colon \mathrm{R}^n \rightarrow \mathrm{R}$ denotes a convex function which needs to be minimized, and Z is a convex set describing the set of feasible solutions. From a computational perspective, convex optimization problems are interesting in the sense that, any local optimal solution will always be guaranteed to be globally optimal. One of the effective concepts of convex optimization considered in this thesis is the Lagrange Duality, which is associated with the Karush-Kuhn-Tucker (KKT) conditions and Lagrange multipliers attributes. These two methods provides important and efficient optimality conditions for convex optimization problems.

6.2 Lagrange Dual Method Optimization

In general, the Lagrange duality theory is known to be the study of optimal solutions to convex optimization problems. More specifically, the optimal solution to a dual problem is a vector of Karush-Kuhn-Tucker (KKT) multipliers, which are used in the KKT conditions to ensure that a solution is optimal. In this work, Lagrange criterion was used to optimize the number of participating FUEs allocated per spectrum resource, by optimizing the transmitting power of FUEs in each small-cell of the HetNets. To be specific, an EE optimization problem was formulated and solved using the Lagrange dual method, considering time and fractional transformation [153]. Based on user scheduling scheme [154], an optimal resource and power allocation scheme was derived using the Lagrange criterion method and achieved high QoS, compared to other optimization algorithms [155, 156]. For example, the Lagrange criterion problem can be formulated as

$$\min_{x \in Z \subseteq \mathrm{R}^n} f_i(x) \tag{13}$$

s.t

$$\begin{aligned} g_i(x) &\leq 0 \quad i = 1, 2, ..., m, \\ h_i(x) &= 0 \quad i = 1, 2, ..., p, \end{aligned} \tag{14}$$

where $f_i(x)$ is the objective or cost function, $x \in Z$ is the optimization variable. $g_i(x)$ and $h_i(x)$ are the inequality and equality constraint functions, and are affine. Eq(13) and (14) describes the problem of finding an x that minimizes $f_i(x)$ among all x values that satisfies the conditions $f_i(x) = 0, i = 1, 2, ..., m, h_i(x) = 0, i = 1, 2, ..., p$ and $x \in Z$. Given a convex constrained minimization problem of the form (13), the Lagrangian function is defined as $L\colon \mathrm{R}^n \times \mathrm{R}^m \times \mathrm{R}^p$ and is given by

$$L\big(x, \alpha, \beta\big) = f_0(x) + \sum_{i=1}^{m} \alpha_i g_i(x) + \sum_{i=1}^{p} \beta_i h_i(x) \tag{15}$$

where the first argument of the Lagrangian is a vector $x \in \mathrm{R}^n$ referred to as the primal variable, whose dimension matches that of the optimization variable in the original optimization problem, $\alpha \in \mathrm{R}^m$ and $\beta \in \mathrm{R}^p$ are variables for α_i and β_i for each of the m convex inequality constraints and p affine equality constraints in the original optimization problem. These elements of α and β are collectively known as the dual variables of the Lagrangian or Lagrange multipliers [155, 156]. The optimal value of the optimization problem is denoted as p^*, and is equal to the minimum possible value of the objective function in the feasible region, and is given by

$$p^* = \min\left\{f_0(x)\colon g_i(x) \leq 0, i = 1, 2, ..., m, h_i(x) = 0, i = 1, 2, ..., p\right\} \tag{16}$$

to take on the values $+\infty$ and $-\infty$ when the problem is either infeasible (the feasible region is empty) or unbounded below (there exists feasible points such that $f(x) \rightarrow -\infty$), respectively. It is said that x^* is an optimal point if $f(x^*) = p^*$. Note that there can be more than one optimal point, even when the optimal value is finite [156]. These methods are commonly used in solving resource allocation problems and are used in this thesis.

7 Research Main Contributions

The work done in this thesis has resulted into several publications. Two of them are outlined as below.

7.1 Paper A: Performance Analysis of Generalized-NOMA Techniques in 5G Wireless Networks

Abstract: Studies the generalized non-orthogonal multiple access (G-NOMA) for the fifth generation (5G) wireless networks under perfect channel state information (CSI). This work investigated the performance analysis of the G-NOMA schemes, namely Interleave Division Multiple Access (IDMA), Multi-User Shared Access (MUSA), pattern division multiple access (PDMA), Sparse Code Multiple Access (SCMA) that address the issues of massive connectivity and improved system capacity. The performance of the G-NOMA schemes is investigated in terms of the bit error rate (BER) and achievable sum rate, and all simulations are conducted in a typically Rayleigh flat fading wireless channel. Simulation results shows that IDMA outperforms the other G-NOMA schemes due to the near-optimal design of user-specific interleavers.

7.2 Paper B: Hybrid G-NOMA Resource Allocation scheme for 5G Small-Cell Networks

Abstract: Non-orthogonal multiple access (NOMA) schemes improve the spectral-efficiency and data-rates for fifth generation (5G) heterogeneous networks (HetNets). Advancement of these schemes will further maximize the performance of the networks. This work develops a hybrid generalized-NOMA (G-NOMA) scheme for a two-tier HetNet. The hybrid G-NOMA (HG-NOMA) scheme combines different resource pattern assignment and power allocation for the different users multiplexed on the same spectrum resource element (SRE) of the network. The resource assignment and power allocation problem is formulated as an energy-efficiency (EE) maximization problem with the aim of maximizing user's connectivity, EE and sum-rate capacity of the small-cells. For reception, a low-complexity hybrid G-NOMA successive interference cancellation (SIC) receiver that combines power-levels and diversity pattern gain to realize multi-user detection is proposed. Performance results show the superiority of the HG-NOMA scheme as compared to the traditional orthogonal multiple access schemes for small-cells in terms of sum-rate capacity, EE and complexity, hence demonstrating their suitability.

References

[1] D. Rapone, D. Sabella, and M. Fodrini, "Energy efficiency solutions for the mobile network evolution towards 5g: an operator perspective," in *2015 Sustainable Internet and ICT for Sustainability (SustainIT)*. IEEE, 2015, pp. 1–9.

[2] G. P. Fettweis, K.-C. Chen, and R. Tafazoli, "Green radio: Energy efficiency in wireless networks," *Journal of Communications and Networks*, vol. 12, no. 2, pp. 99–102, 2010.

[3] G. Koutitas and P. Demestichas, "A review of energy efficiency in telecommunication networks," *Telfor journal*, vol. 2, no. 1, pp. 2–7, 2010.

[4] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[5] I. R. Sector, "Requirements related to technical performance for imt-advanced radio interface (s)," *Report ITU*, 2008.

[6] Y. Chen, A. Bayesteh, Y. Wu, B. Ren, S. Kang, S. Sun, Q. Xiong, C. Qian, B. Yu, Z. Ding *et al.*, "Toward the standardization of non-orthogonal multiple access for next generation wireless networks," *IEEE Communications Magazine*, vol. 56, no. 3, pp. 19–27, 2018.

[7] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for fdd massive mimo," *IEEE Transactions on Signal Processing*, vol. 63, no. 23, pp. 6169–6183, 2015.

[8] X. Gao, L. Dai, S. Han, I. Chih-Lin, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmwave mimo systems with large antenna arrays," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 998–1009, 2016.

[9] H. Zhang, X. Chu, W. Guo, and S. Wang, "Coexistence of wi-fi and heterogeneous small cell networks sharing unlicensed spectrum," *IEEE Communications Magazine*, vol. 53, no. 3, pp. 158–164, 2015.

[10] H. Zhang, C. Jiang, J. Cheng, and V. C. Leung, "Cooperative interference mitigation and handover management for heterogeneous cloud small cell networks," *IEEE Wireless Communications*, vol. 22, no. 3, pp. 92–99, 2015.

[11] S. R. Islam, N. Avazov, O. A. Dobre, and K.-S. Kwak, "Power-domain non-orthogonal multiple access (noma) in 5g systems: Potentials and challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 721–742, 2017.

[12] A. Gupta and R. K. Jha, "A survey of 5g network: Architecture and emerging technologies," *IEEE access*, vol. 3, pp. 1206–1232, 2015.

[13] E. Ezhilarasan and M. Dinakaran, "A review on mobile technologies: 3g, 4g and 5g," in *Recent Trends and Challenges in Computational Models (ICRTCCM), 2017 Second International Conference on*. IEEE, 2017, pp. 369–373.

[14] S. Sesia, M. Baker, and I. Toufik, *LTE-the UMTS long term evolution: from theory to practice*. John Wiley & Sons, 2011.

[15] J. Cao, M. Ma, H. Li, Y. Zhang, and Z. Luo, "A survey on security aspects for lte and lte-a networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 283–302, 2014.

[16] C. CISCO, "Visual networking index: Global mobile data traffic forecast update 2014-2019," Technical Report, CISCO February 2015, Tech. Rep., 2011.

[17] Z. Bharucha, E. Calvanese, J. Chen, X. Chu, A. Feki, A. De Domenico, A. Galindo-Serrano, W. Guo, R. Kwan, J. Liu *et al.*, "Small cell deployments: Recent advances and research challenges," *arXiv preprint arXiv:1211.0575*, 2012.

[18] M. Rahman and F. Mir, "Fourth generation (4g) mobile networks-features, technologies & issues," in *3G and Beyond, 2005 6th IEE International Conference on*. IET, 2005, pp. 1–5.

[19] N. T. Le, M. A. Hossain, A. Islam, D.-y. Kim, Y.-J. Choi, and Y. M. Jang, "Survey of promising technologies for 5g networks," *Mobile information systems*, 2016.

[20] M. Lauridsen, L. C. Gimenez, I. Rodriguez, T. B. Sorensen, and P. Mogensen, "From lte to 5g for connected mobility," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 156–162, 2017.

[21] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal multiple access for 5g and beyond," *Proceedings of the IEEE*, vol. 105, no. 12, pp. 2347–2381, 2017.

[22] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5g: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201–1221, 2017.

[23] H. Gamage, N. Rajatheva, and M. Latva-aho, "Channel coding for enhanced mobile broadband communication in 5g systems," in *Networks and Communications (EuCNC), 2017 European Conference on*. IEEE, 2017, pp. 1–6.

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

[25] P. K. Agyapong, M. Iwamura, D. Staehle, W. Kiess, and A. Benjebbour, "Design considerations for a 5g network architecture," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 65–75, 2014.

[26] V. Yazıcı, U. C. Kozat, and M. O. Sunay, "A new control plane for 5g network architecture with a case study on unified handoff, mobility, and routing management," *IEEE communications magazine*, vol. 52, no. 11, pp. 76–85, 2014.

[27] A. Yegin, J. Park, K. Kweon, and J. Lee, "Terminal-centric distribution and orchestration of ip mobility for 5g networks," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 86–92, 2014.

[28] F. Malandrino, C. Casetti, and C.-F. Chiasserini, "Toward d2d-enhanced heterogeneous networks," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 94–100, 2014.

[29] D. Soldani, K. Pentikousis, R. Tafazolli, and D. Franceschini, “5g networks: End-to-end architecture and infrastructure [guest editorial],” *IEEE Communications Magazine*, vol. 52, no. 11, pp. 62–64, 2014.

[30] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, “Femtocells: Past, present, and future,” *IEEE Journal on Selected Areas in communications*, vol. 30, no. 3, pp. 497–508, 2012.

[31] I. Allal, B. Mongazon-Cazavet, K. Al Agha, S.-M. Senouci, and Y. Gourhant, “A green small cells deployment in 5g—switch on/off via iot networks & energy efficient mesh backhauling,” in *IFIP Networking Conference (IFIP Networking) and Workshops, 2017*. IEEE, 2017, pp. 1–2.

[32] I. Hwang, B. Song, and S. S. Soliman, “A holistic view on hyper-dense heterogeneous and small cell networks,” *IEEE Communications Magazine*, vol. 51, no. 6, pp. 20–27, 2013.

[33] L. Frenzel, “Understanding the small-cell and hetnet movement,” *Electronic design*, vol. 18, 2013.

[34] R. Coombs and R. Steele, “Introducing microcells into macrocellular networks: A case study,” *IEEE Transactions on Communications*, vol. 47, no. 4, pp. 568–576, 1999.

[35] R. Iyer, J. Parker, and P. Sood, “Intelligent networking for digital cellular systems and the wireless world,” in *Global Telecommunications Conference, 1990, and Exhibition.’Communications: Connecting the Future’, GLOBECOM’90., IEEE*. IEEE, 1990, pp. 475–479.

[36] S. Ramanath, E. Altman, V. Kumar, and M. Debbah, “Optimizing cell size in pico-cell networks,” in *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, 2009. WiOPT 2009. 7th International Symposium on*. IEEE, 2009, pp. 1–9.

[37] S. Jafarpour-Alamdari, V. Solouk, H. Kalbkhani, and M. G. Shayesteh, “Interference analysis of femto-macrocell networks in 6-sector macrocell layout,” in *Telecommunications (IST), 2012 Sixth International Symposium on*. IEEE, 2012, pp. 277–282.

[38] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, “Femtocell networks: a survey,” *IEEE Communications magazine*, vol. 46, no. 9, 2008.

[39] M. Z. Chowdhury, Y. M. Jang, and Z. J. Haas, “Cost-effective frequency planning for capacity enhancement of femtocellular networks,” *Wireless Personal Communications*, vol. 60, no. 1, pp. 83–104, 2011.

[40] N. Saquib, E. Hossain, L. B. Le, and D. I. Kim, “Interference management in ofdma femtocell networks: Issues and approaches,” *IEEE Wireless Communications*, vol. 19, no. 3, 2012.

[41] T. Zahir, K. Arshad, A. Nakata, and K. Moessner, “Interference management in femtocells,” *IEEE communications surveys & tutorials*, vol. 15, no. 1, pp. 293–311, 2013.

[42] Z. Qin, Y. Liu, Z. Ding, Y. Gao, and M. Elkashlan, “Physical layer security for 5g non-orthogonal multiple access in large-scale networks,” in *Communications (ICC), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–6.

[43] L. Mohjazi, M. Al-Qutayri, H. Barada, K. Poon, and R. Shubair, “Deployment challenges of femtocells in future indoor wireless networks,” in *GCC Conference and Exhibition (GCC), 2011 IEEE*. IEEE, 2011, pp. 405–408.

[44] Y. Cai, Z. Qin, F. Cui, G. Y. Li, and J. A. McCann, “Modulation and multiple access for 5g networks,” *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 629–646, 2018.

[45] Y. Wang, B. Ren, S. Sun, S. Kang, and X. Yue, “Analysis of non-orthogonal multiple access for 5g,” *China Communications*, vol. 13, no. Supplement2, pp. 52–66, 2016.

[46] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. Bhargava, “A survey on non-orthogonal multiple access for 5g networks: Research challenges and future trends,” *arXiv preprint arXiv:1706.05347*, 2017.

[47] K. Higuchi and A. Benjebbour, “Non-orthogonal multiple access (noma) with successive interference cancellation for future radio access,” *IEICE Transactions on Communications*, vol. 98, no. 3, pp. 403–414, 2015.

[48] B. Wang, K. Wang, Z. Lu, T. Xie, and J. Quan, “Comparison study of non-orthogonal multiple access schemes for 5g,” in *Broadband Multimedia Systems and Broadcasting (BMSB), 2015 IEEE International Symposium on*. IEEE, 2015, pp. 1–5.

[49] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, “Non-orthogonal multiple access (noma) for cellular future radio access,” in *Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th*. IEEE, 2013, pp. 1–5.

[50] A. Morgado, K. M. S. Huq, S. Mumtaz, and J. Rodriguez, “A survey of 5g technologies: regulatory, standardization and industrial perspectives,” *Digital Communications and Networks*, vol. 4, no. 2, pp. 87–97, 2018.

[51] M. Basharat, W. Ejaz, M. Naeem, A. M. Khattak, and A. Anpalagan, “A survey and taxonomy on nonorthogonal multiple-access schemes for 5g networks,” *Transactions on Emerging Telecommunications Technologies*, vol. 29, no. 1, p. e3202, 2018.

[52] T. Cover, “Broadcast channels,” *IEEE Transactions on Information Theory*, vol. 18, no. 1, pp. 2–14, 1972.

[53] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[54] P. Bergmans, “Random coding theorem for broadcast channels with degraded components,” *IEEE Transactions on Information Theory*, vol. 19, no. 2, pp. 197–207, 1973.

[55] R. G. Gallager, “Capacity and coding for degraded broadcast channels,” *Problemy Peredachi Informatsii*, vol. 10, no. 3, pp. 3–14, 1974.

[56] L. Wang, E. Şaşoğlu, B. Bandemer, and Y.-H. Kim, “A comparison of superposition coding schemes,” in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*. IEEE, 2013, pp. 2970–2974.

[57] S. Vanka, S. Srinivasa, Z. Gong, P. Vizi, K. Stamatiou, and M. Haenggi, "Superposition coding strategies: Design and experimental evaluation," *IEEE Transactions on Wireless Communications*, vol. 11, no. 7, pp. 2628–2639, 2012.

[58] A. Zafar, M. Shaqfeh, M.-S. Alouini, and H. Alnuweiri, "On multiple users scheduling using superposition coding over rayleigh fading channels," *IEEE Communications Letters*, vol. 17, no. 4, pp. 733–736, 2013.

[59] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A survey of non-orthogonal multiple access for 5g," *IEEE Communications Surveys & Tutorials*, 2018.

[60] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[61] R. Zhang and L. Hanzo, "A unified treatment of superposition coding aided communications: Theory and practice," *IEEE communications surveys & tutorials*, vol. 13, no. 3, pp. 503–520, 2011.

[62] G. B. Satrya and S. Y. Shin, "Security enhancement to successive interference cancellation algorithm for non-orthogonal multiple access (noma)," in *Personal, Indoor, and Mobile Radio Communications (PIMRC), 2017 IEEE 28th Annual International Symposium on*. IEEE, 2017, pp. 1–5.

[63] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang, "Non-orthogonal multiple access for 5g: solutions, challenges, opportunities, and future research trends," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74–81, 2015.

[64] R. Ruby, S. Zhong, H. Yang, and K. Wu, "Enhanced uplink resource allocation in non-orthogonal multiple access systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1432–1444, 2018.

[65] M. Al-Imari, P. Xiao, and M. A. Imran, "Receiver and resource allocation optimization for uplink noma in 5g wireless networks," in *Wireless Communication Systems (ISWCS), 2015 International Symposium on*. IEEE, 2015, pp. 151–155.

[66] R. Singh, H. Zhu, and J. Wang, "Performance of non-orthogonal multiple access (noma) in a c-ran system," in *Vehicular Technology Conference (VTC-Fall), 2017 IEEE 86th*. IEEE, 2017, pp. 1–5.

[67] Z. Q. Al-Abbasi and D. K. So, "Resource allocation in non-orthogonal and hybrid multiple access system with proportional rate constraint," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6309–6320, 2017.

[68] J. Wang, H. Xu, L. Fan, B. Zhu, and A. Zhou, "Energy-efficient joint power and bandwidth allocation for noma systems," *IEEE Communications Letters*, vol. 22, no. 4, pp. 780–783, 2018.

[69] F. Fang, H. Zhang, J. Cheng, and V. C. Leung, "Energy-efficient resource allocation for downlink non-orthogonal multiple access network," *IEEE Transactions on Communications*, vol. 64, no. 9, pp. 3722–3732, 2016.

[70] A. S. Ibrahim, A. K. Sadek, W. Su, and K. R. Liu, “Cooperative communications with relay-selection: when to cooperate and whom to cooperate with?” *IEEE Transactions on Wireless Communications*, vol. 7, no. 7, 2008.

[71] Z. Ding, M. Peng, and H. V. Poor, “Cooperative non-orthogonal multiple access in 5g systems,” *IEEE Communications Letters*, vol. 19, no. 8, pp. 1462–1465, 2015.

[72] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, “A survey on spectrum management in cognitive radio networks,” *IEEE Communications magazine*, vol. 46, no. 4, 2008.

[73] K. B. Letaief and W. Zhang, “Cooperative communications for cognitive radio networks,” *Proceedings of the IEEE*, vol. 97, no. 5, pp. 878–893, 2009.

[74] F. Salahdine, “Spectrum sensing techniques for cognitive radio networks,” *arXiv preprint arXiv:1710.02668*, 2017.

[75] D. Rawat, G. Yan, and C. Bajracharya, “Signal processing techniques for spectrum sensing in cognitive radio networks,” *International Journal of Ultra Wideband Communications and Systems*, vol. 10, no. x/x, pp. 1–10, 2010.

[76] L. Lv, J. Chen, Q. Ni, Z. Ding, and H. Jiang, “Cognitive non-orthogonal multiple access with cooperative relaying: A new wireless frontier for 5g spectrum sharing,” *IEEE Communications Magazine*, vol. 56, no. 4, pp. 188–195, 2018.

[77] S. Dinh-Van, V.-D. Nguyen, and O.-S. Shin, “Interference-aware transmission for d2d communications in a cellular network,” *Wireless Personal Communications*, vol. 98, no. 1, pp. 1489–1503, 2018.

[78] W. Guibène and D. Slock, “Cooperative spectrum sensing and localization in cognitive radio systems using compressed sensing,” *Journal of Sensors*, vol. 2013, 2013.

[79] X. Chen, H.-H. Chen, and W. Meng, “Cooperative communications for cognitive radio networks—from theory to applications,” *IEEE Communications surveys & tutorials*, vol. 16, no. 3, pp. 1180–1192, 2014.

[80] M. F. Kader and S. Y. Shin, “Cooperative hybrid spectrum sharing: A noma-based approach,” *Wireless Personal Communications*, vol. 96, no. 1, pp. 1–21, 2017.

[81] Y. Liu, Z. Ding, M. Elkashlan, and J. Yuan, “Nonorthogonal multiple access in large-scale underlay cognitive radio networks,” *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 10 152–10 157, 2016.

[82] Z. Yang, J. A. Hussein, P. Xu, Z. Ding, and Y. Wu, “Power allocation study for non-orthogonal multiple access networks with multicast-unicast transmission,” *IEEE Transactions on Wireless Communications*, 2018.

[83] K. Higuchi and Y. Kishiyama, “Non-orthogonal access with random beamforming and intra-beam sic for cellular mimo downlink,” in *Vehicular Technology Conference (VTC Fall), 2013 IEEE 78th*. IEEE, 2013, pp. 1–5.

[84] H. Lee, S. Kim, and J.-H. Lim, “Multiuser superposition transmission (must) for lte-a systems,” in *Communications (ICC), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–6.

[85] H. Nikopour and H. Baligh, “Sparse code multiple access,” in *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*. IEEE, 2013, pp. 332–336.

[86] M. R. Bhatia, “Multiple access techniques for 5g networks.”

[87] C. YAN, Z. YUAN, W. LI, and Y. YUAN, “Non-orthogonal multiple access schemes for 5g,” *ZTE Communications*, vol. 14, no. 4, pp. 11–16, 2016.

[88] M. Kulhandjian and C. D’Amours, “Design of permutation-based sparse code multiple access system,” in *Personal, Indoor, and Mobile Radio Communications (PIMRC), 2017 IEEE 28th Annual International Symposium on*. IEEE, 2017, pp. 1–6.

[89] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh, “Scma codebook design,” in *Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th*. IEEE, 2014, pp. 1–5.

[90] D. Cai, P. Fan, X. Lei, Y. Liu, and D. Chen, “Multi-dimensional scma codebook design based on constellation rotation and interleaving,” in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*. IEEE, 2016, pp. 1–5.

[91] L. Yu, X. Lei, P. Fan, and D. Chen, “An optimized design of scma codebook based on star-qam signaling constellations,” in *Wireless Communications & Signal Processing (WCSP), 2015 International Conference on*. IEEE, 2015, pp. 1–5.

[92] A. Bayesteh, H. Nikopour, M. Taherzadeh, H. Baligh, and J. Ma, “Low complexity techniques for scma detection,” in *Globecom Workshops (GC Wkshps), 2015 IEEE*. IEEE, 2015, pp. 1–6.

[93] Y. Du, B. Dong, Z. Chen, J. Fang, P. Gao, and Z. Liu, “Low-complexity detector in sparse code multiple access systems,” *IEEE Communications Letters*, vol. 20, no. 9, pp. 1812–1815, 2016.

[94] C. Zhang, C. Yang, W. Xu, S. Zhang, Z. Zhang, and X. You, “Efficient sparse code multiple access decoder based on deterministic message passing algorithm,” *arXiv preprint arXiv:1804.00180*, 2018.

[95] A. Bayesteh, E. Yi, H. Nikopour, and H. Baligh, “Blind detection of scma for uplink grant-free multiple-access,” in *Wireless Communications Systems (ISWCS), 2014 11th International Symposium on*. IEEE, 2014, pp. 853–857.

[96] M. Moltafet, N. Mokari, M. R. Javan, H. Saeedi, and H. Pishro-Nik, “A new multiple access technique for 5g: Power domain sparse code multiple access (psma),” *IEEE Access*, vol. 6, pp. 747–759, 2018.

[97] S. Zhang, X. Xu, L. Lu, Y. Wu, G. He, and Y. Chen, “Sparse code multiple access: An energy efficient uplink approach for 5g wireless systems,” in *Global Communications Conference (GLOBECOM), 2014 IEEE*. IEEE, 2014, pp. 4782–4787.

[98] Z. Yuan, G. Yu, W. Li, Y. Yuan, X. Wang, and J. Xu, “Multi-user shared access for internet of things,” in *Vehicular Technology Conference (VTC Spring), 2016 IEEE 83rd*. IEEE, 2016, pp. 1–5.

[99] Y. Yuan, Z. Yuan, G. Yu, C.-h. Hwang, P.-k. Liao, A. Li, and K. Takeda, "Non-orthogonal transmission technology in lte evolution," *IEEE Communications Magazine*, vol. 54, no. 7, pp. 68–74, 2016.

[100] L. Ping, L. Liu, K. Wu, and W. Leung, "A unified approach to multiuser detection and space-time coding with low complexity and nearly optimal performance," in *Proceedings of the Annual Allerton Conference on Communication Control and Computing*, vol. 40, no. 1. The University; 1998, 2002, pp. 170–179.

[101] L. Liu, W. Leung, and L. Ping, "Simple iterative chip-by-chip multiuser detection for cdma systems," in *IEEE Vehicular Technology Conference*, vol. 3. IEEE; 1999, 2003, pp. 2157–2161.

[102] C. Novak, F. Hlawatsch, and G. Matz, "Mimo-idma: Uplink multiuser mimo communications using interleave-division multiple access and low-complexity iterative receivers," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 3. IEEE, 2007, pp. III–225.

[103] Y. Hu and L. Ping, "Interleave division multiple access (idma)," in *Multiple Access Techniques for 5G Wireless Networks and Beyond.* Springer, 2019, pp. 417–449.

[104] L. Ping, P. Wang, and X. Wang, "Recent progress in interleave-division multiple-access (idma)," in *Military Communications Conference, 2007. MILCOM 2007. IEEE.* IEEE, 2007, pp. 1–7.

[105] L. Ping, L. Liu, K. Wu, W. Leung *et al.*, "Interleave-division multiple-access (idma) communications," in *Proc. 3rd International Symposium on Turbo Codes and Related Topics*. Citeseer, 2003, p. 173180.

[106] P. Wang, L. Ping, and L. Liu, "Power allocation for multiple access systems with practical coding and iterative multi-user detection," in *Communications, 2006. ICC'06. IEEE International Conference on*, vol. 11. IEEE, 2006, pp. 4971–4976.

[107] ——, "Optimized power allocation for multiple access systems with practical coding and iterative multi-user detection," in *Proc. 2006 Int. Symp. Turbo Codes Related Topics Connection Int. ITG-Conf. Source Channel Coding*, 2006.

[108] L. Ping, L. Liu, and W. Leung, "A simple approach to near-optimal multiuser detection: interleave-division multiple-access," in *Wireless Communications and Networking, 2003. WCNC 2003. 2003 IEEE*, vol. 1. IEEE, 2003, pp. 391–396.

[109] S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, and K. Niu, "Pattern division multiple access—a novel nonorthogonal multiple access for fifth-generation radio networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3185–3196, 2017.

[110] J. Zeng, B. Li, X. Su, L. Rong, and R. Xing, "Pattern division multiple access (pdma) for cellular future radio access," in *Wireless Communications & Signal Processing (WCSP), 2015 International Conference on*. IEEE, 2015, pp. 1–5.

[111] X. Dai, Z. Zhang, B. Bai, S. Chen, and S. Sun, "Pattern division multiple access: A new multiple access technology for 5g," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 54–60, 2018.

[112] J. Zeng, D. Kong, X. Su, L. Rong, and X. Xu, “On the performance of pattern division multiple access in 5g systems,” in *Wireless Communications & Signal Processing (WCSP), 2016 8th International Conference on*. IEEE, 2016, pp. 1–5.

[113] W. Tang, S. Kang, B. Ren, and X. Yue, “Uplink grant-free pattern division multiple access (gf-pdma) for 5g radio access,” *China Communications*, vol. 15, no. 4, pp. 153–163, 2018.

[114] Z. Song, Q. Ni, and X. Sun, “Spectrum and energy efficient resource allocation with qos requirements for hybrid mc-noma 5g systems,” *IEEE Access*, vol. 6, pp. 37 055–37 069, 2018.

[115] Y. Xu, C. Shen, T.-H. Chang, S.-C. Lin, Y. Zhao, and G. Zhu, “On energy-efficient noma designs for heterogeneous low-latency downlink transmissions,” *arXiv preprint arXiv:1806.09144*, 2018.

[116] F. Fang, H. Zhang, J. Cheng, and V. C. Leung, “Energy-efficient resource scheduling for noma systems with imperfect channel state information,” in *Communications (ICC), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1–5.

[117] W. H. Chin, Z. Fan, and R. Haines, “Emerging technologies and research challenges for 5g wireless networks,” *IEEE Wireless Communications*, vol. 21, no. 2, pp. 106–112, 2014.

[118] M. B. Yassein, S. Aljawarneh, and A. Al-Sadi, “Challenges and features of iot communications in 5g networks,” in *Electrical and Computing Technologies and Applications (ICECTA), 2017 International Conference on*. IEEE, 2017, pp. 1–5.

[119] H. Zhang, C. Jiang, R. Q. Hu, and Y. Qian, “Self-organization in disaster-resilient heterogeneous small cell networks,” *IEEE Network*, vol. 30, no. 2, pp. 116–121, 2016.

[120] H. Zhang, F. Fang, J. Cheng, K. Long, W. Wang, and V. C. Leung, “Energy-efficient resource allocation in noma heterogeneous networks,” *IEEE Wireless Communications*, vol. 25, no. 2, pp. 48–53, 2018.

[121] N. Al-Falahy and O. Y. Alani, “Technologies for 5g networks: challenges and opportunities,” *IT Professional*, vol. 19, no. 1, pp. 12–20, 2017.

[122] C. Jarray, A. Bouabid, and B. Chibani, “Enabling and challenges for 5g technologies,” in *Information Technology and Computer Applications Congress (WCITCA), 2015 World Congress on*. IEEE, 2015, pp. 1–9.

[123] A. Ghosh, N. Mangalvedhe, R. Ratasuk, B. Mondal, M. Cudak, E. Visotsky, T. A. Thomas, J. G. Andrews, P. Xia, H. S. Jo *et al.*, “Heterogeneous cellular networks: From theory to practice,” *IEEE communications magazine*, vol. 50, no. 6, 2012.

[124] Q. Li, R. Q. Hu, Y. Qian, and G. Wu, “Intracell cooperation and resource allocation in a heterogeneous network with relays,” *IEEE Transactions on Vehicular Technology*, vol. 62, no. 4, pp. 1770–1784, 2013.

[125] H. Dahrouj, A. Douik, O. Dhifallah, T. Y. Al-Naffouri, and M.-S. Alouini, “Resource allocation in heterogeneous cloud radio access networks: advances and challenges,” *IEEE Wireless Communications*, vol. 22, no. 3, pp. 66–73, 2015.

[126] O. Oguejiofor, L. Zhang, and M. Alhabo, "Decentralized resource allocation for heterogeneous cellular networks," in *Wireless Communication Systems (ISWCS), 2017 International Symposium on*. IEEE, 2017, pp. 158–163.

[127] T. Villa, R. Merz, and R. Knopp, "Dynamic resource allocation in heterogeneous networks." in *GLOBECOM*, 2013, pp. 1915–1920.

[128] H. Peng, Y. Xiao, Y.-N. Ruyue, and Y. Yifei, "Ultra dense network: challenges, enabling technologies and new trends," *China Communications*, vol. 13, no. 2, pp. 30–40, 2016.

[129] Y. Teng, M. Liu, F. R. Yu, V. C. Leung, M. Song, and Y. Zhang, "Resource allocation for ultra-dense networks: A survey, some research issues and challenges," *IEEE Communications Surveys & Tutorials*, 2018.

[130] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-dense networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2522–2545, 2016.

[131] Q. Li, G. Wu, and R. Q. Hu, "Analytical study on network spectrum efficiency of ultra dense networks," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*. IEEE, 2013, pp. 2764–2768.

[132] M. Liu, Y. Teng, and M. Song, "Tradeoff analysis between spectrum efficiency and energy efficiency in heterogeneous networks (hetnets) using bias factor," *Journal of Communications*, vol. 10, no. 10, pp. 784–789, 2015.

[133] C. C. Coskun and E. Ayanoglu, "Energy-spectral efficiency tradeoff for heterogeneous networks with qos constraints," in *2017 IEEE International Conference on Communications (ICC)*. IEEE, 2017, pp. 1–7.

[134] Z. Xiao, J. Yu, T. Li, Z. Xiang, D. Wang, and W. Chen, "Resource allocation via hierarchical clustering in dense small cell networks: a correlated equilibrium approach," in *Personal, Indoor, and Mobile Radio Communications (PIMRC), 2016 IEEE 27th Annual International Symposium on*. IEEE, 2016, pp. 1–5.

[135] S. A. Kazmi, N. H. Tran, T. M. Ho, T. Z. Oo, T. LeAnh, S. Moon, and C. S. Hong, "Resource management in dense heterogeneous networks," in *Network Operations and Management Symposium (APNOMS), 2015 17th Asia-Pacific*. IEEE, 2015, pp. 440–443.

[136] Y. Luo and C. Hua, "Resource allocation and user-centric clustering in ultra-dense networks with wireless backhaul," in *Wireless Communications & Signal Processing (WCSP), 2016 8th International Conference on*. IEEE, 2016, pp. 1–5.

[137] C. Hua, Y. Luo, and H. Liu, "Wireless backhaul resource allocation and user-centric clustering in ultra-dense wireless networks," *IET Communications*, vol. 10, no. 15, pp. 1858–1864, 2016.

[138] E. Pateromichelakis, M. Shariat, A. Quddus, M. Dianati, and R. Tafazolli, "Dynamic clustering framework for multi-cell scheduling in dense small cell networks," *IEEE Communications Letters*, vol. 17, no. 9, pp. 1802–1805, 2013.

[139] A. Abrol and R. K. Jha, “Power optimization in 5g networks: A step towards green communication,” *IEEE Access*, vol. 4, pp. 1355–1374, 2016.

[140] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok, “The global footprint of mobile communications: The ecological and economic perspective,” *IEEE Communications Magazine*, vol. 49, no. 8, 2011.

[141] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. A. Imran, D. Sabella, M. J. Gonzalez, O. Blume *et al.*, “How much energy is needed to run a wireless network?” *IEEE Wireless Communications*, vol. 18, no. 5, 2011.

[142] S. Buzzi, I. Chih-Lin, T. E. Klein, H. V. Poor, C. Yang, and A. Zappone, “A survey of energy-efficient techniques for 5g networks and challenges ahead,” *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 697–709, 2016.

[143] X. Cao, L. Liu, Y. Cheng, and X. S. Shen, “Towards energy-efficient wireless networking in the big data era: A survey,” *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 303–332, 2018.

[144] R. Q. Hu and Y. Qian, “An energy efficient and spectrum efficient wireless heterogeneous network framework for 5g systems,” *IEEE Communications Magazine*, vol. 52, no. 5, pp. 94–101, 2014.

[145] V. Mor and H. Kumar, “Energy efficient wireless mobile networks: A review,” in *Optimization, Reliabilty, and Information Technology (ICROIT), 2014 International Conference on*. IEEE, 2014, pp. 281–285.

[146] M. De Sanctis, E. Cianca, and V. Joshi, “Energy efficient wireless networks towards green communications,” *Wireless Personal Communications*, vol. 59, no. 3, pp. 537–552, 2011.

[147] C. Bae and W. E. Stark, “Energy and bandwidth efficiency in wireless networks,” in *Communications, Circuits and Systems Proceedings, 2006 International Conference on*, vol. 2. IEEE, 2006, pp. 1297–1302.

[148] D. Feng, C. Jiang, G. Lim, L. J. Cimini, G. Feng, and G. Y. Li, “A survey of energy-efficient wireless communications,” *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 167–178, 2013.

[149] F. Fang, “Energy efficient resource allocation for non-orthogonal multiple access (noma) systems,” Ph.D. dissertation, University of British Columbia, 2017.

[150] J. Reed, M. Vassiliou, and S. Shah, “The role of new technologies in solving the spectrum shortage [point of view],” *Proceedings of the IEEE*, vol. 104, no. 6, pp. 1163–1168, 2016.

[151] A. J. Goldsmith and S. B. Wicker, “Design challenges for energy-constrained ad hoc wireless networks,” *IEEE wireless communications*, vol. 9, no. 4, pp. 8–27, 2002.

[152] E. Tsakonas, “Convex optimization for assignment and generalized linear regression problems,” Ph.D. dissertation, KTH Royal Institute of Technology, 2014.

[153] W. Dinkelbach, “On nonlinear fractional programming,” *Management science*, vol. 13, no. 7, pp. 492–498, 1967.

[154] F. Fang, H. Zhang, J. Cheng, S. Roy, and V. C. Leung, “Joint user scheduling and power allocation optimization for energy efficient noma systems with imperfect csi,” *IEEE J. Sel. Areas Commun*, vol. 35, no. 12, pp. 2874–2885, 2017.

[155] C. B. Do, “Convex optimization overview (cnt’d),” 2009.

[156] Z. Kolter, “Convex optimization overview,” 2008.

[157] R. M. Alaez, E. Chirivella-Perez, J. M. A. Calero, and Q. Wang, “New topology management scheme in lte and 5g networks,” in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*. IEEE, 2018, pp. 1–5.

[158] J.-M. Kelif, J.-M. Gorce, and A. Gati, “Performance and energy in green superposition coding wireless networks: An analytical model,” in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.

[159] Y. Liu, Z. Qin, M. Elkashlan, Y. Gao, and A. Nallanathan, “Non-orthogonal multiple access in massive mimo aided heterogeneous networks,” in *Global Communications Conference (GLOBECOM), 2016 IEEE*. IEEE, 2016, pp. 1–6.

[160] Y. Liu, Z. Qin, M. Elkashlan, A. Nallanathan, and J. A. McCann, “Non-orthogonal multiple access in large-scale heterogeneous networks,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 12, pp. 2667–2680, 2017.

Part II

Papers

Paper A

Performance Analysis of Generalized-NOMA Techniques in 5G Wireless Networks

Samson Manyani Zitha, Tom Walingo and James Okello

Proceedings of SATNAC 2018, Harmanus, South Africa.

The layout has been revised.

Abstract

The generalized non-orthogonal multiple access (G-NOMA) enables fifth generation (5G) wireless networks to meet their demands of massive connectivity of users and devices, low latency, and high spectral efficiency. This work investigates the performance analysis of the G-NOMA schemes, namely Interleave Division Multiple Access (IDMA), Multi-User Shared Access (MUSA), pattern division multiple access (PDMA), Sparse Code Multiple Access (SCMA) that address the issues of massive connectivity and improved system capacity. The performance of the G-NOMA schemes is investigated in terms of the bit error rate (BER) and achievable sum rate. Perfect channel state information (CSI) is assumed in this work, and all simulations are conducted in a typical Rayleigh flat fading wireless channel. Simulation results show that IDMA outperforms the other G-NOMA schemes due to the near-optimal design of user-specific interleavers.

1 Introduction

The proliferation of data traffic over wireless channels not seen in previous generations has necessitated the need for newer multiple access technologies for the fifth generation (5G) communication networks [1]. Moreover, due to the rapid development of mobile internet and internet of things (IoT), 5G networks have diverse quality of services (QoS) demands such as, the need to support massive connectivity of users and devices, provide higher spectral efficiency and provide low latency among others [2] - [3]. Generally, multiple access technologies are classified into two categories i.e. orthogonal multiple access (OMA) and non-orthogonal multiple access (NOMA). In the conventional OMA, such as frequency division multiple access (FDMA), time division multiple access (TDMA), code division multiple access (CDMA) and orthogonal frequency division multiple access (OFDMA), different users can only be served orthogonally in either time or frequency to alleviate inter-user interference [2]. Hence the number of available simultaneously accessing users is strictly proportional to the number of orthogonal resources, thus limited [4]. The 5G demands are rather challenging and difficult to satisfy using OMA schemes. On the other hand, NOMA technologies allow multiple users to share the same time and frequency on the same spatial layer via code domain or power domain multiplexing. Hence, it has been considered an auspicious candidate to address the QoS demands of the 5G system [1] - [3]. A performance investigation of this schemes is necessary to determine their application benefits in 5G networks.

In this paper, we focus on G-NOMA schemes that address the issues of massive connectivity and improved system capacity. Recently, several NOMA schemes have attracted a lot of attention and have been studied in both industry and academia [3]. NOMA technologies, however, permit controllable interferences by non-orthogonal resource allocation at the cost of a slight increase in receiver complexity [2], [5]. NOMA schemes are generally divided into two categories; power domain and code domain multiplexing [3] - [6]. In power domain NOMA, multiplexing occurs by allocating different power levels, to different users in the same orthogonal resource block according to their channel conditions. While code domain multiplexing occurs by assigning sparse codes designed to support multiple users that share the same time-frequency spectrum domain [7], [8]. Non-orthogonal multiple access schemes include, low-density spreading multiple access (LDSMA) [9], sparse code multiple access (SCMA) [10], multi-user shared access (MUSA) [11], pattern division multiple access (PDMA) [12], and interleave division multiple access (IDMA) [13]. These are the ones investigated in this work.

LDSMA is a special form of the conventional code division multiple access (CDMA), in which data symbols are spread by low density spreading sequences. A chip-level iterative multiuser decoding based on the message passing algorithm (MPA) receiver is used to effectively exploit the low-density spreading (LDS) structure [9] - [11]. In SCMA, users bits streams are directly mapped to sparse codewords selected to user specific codebooks, which enables massive connectivity in the 5G networks. Hence, users data is transmitted within the same radio resources in time-frequency domain. The sparsity of the codewords makes it possible to use the near-optimal multiuser decoding practically feasible [1] - [2], [10]. MUSA is a non-orthogonal transmission and grant-free multiple access schemes. Data of each user is spread with a family of short length complex spreading codes that are designed to support massive connectivity while minimizing signaling overhead and power consumption between users that share the same resource block. Blind detection is applied in MUSA, in which the minimum mean square error successive interference cancellation (MMSE-SIC) is used to cancel interference between users [11], [14]. PDMA improves spectral efficiency by the realization of using sparse non-orthogonal patterns to map recourses to different users. Lastly, the IDMA explores the possibilities that allow multiple users to share time and frequency by employing user specific interleavers to separate users. Not only does IDMA inherit many advantages of CDMA, but it also allows a simple chip-by-chip (CBC) iterative sub-optimal multi-user detection (MUD) strategy [13] - [15]. In fact, G-NOMA schemes realize overloading by non-orthogonal resource allocation, which enables massive connectivity and improves system spectral efficiency. While

G-NOMA techniques have the capability of increasing the channel sum rate capacity, there is no clear quantification of the systems performance analysis nor their comparison. A comparison of the G-NOMA techniques have been carried out in [2], [14], [16]. However, most of it is a theoretical comparison, if not, no comprehensive results are developed for the schemes. Hence, this work investigates a comprehensive system performance comparison of the G-NOMA techniques incorporating different user overloading.

The rest of the paper is organized as follows, In Section II, G-NOMA schemes features, principles and system models, are discussed. In Section III simulation results and performance comparison analysis is done. Finally, Section IV, concludes the paper.

Notation: The bold upper and lower-case letters denote matrices and column vectors, respectively. $I_{N\times N}$ is an N by N identity matrix. The following superscripts $(\cdot)^H$, $(\cdot)^{-1}$ and $(\cdot)^T$ represents Hermitian, inverse and transpose operator, respectively. $|\cdot|$ and $||\cdot||$ denote the absolute value of a scalar and the Frobenius norm. $\odot$ Indicates the element-wise dot product of two matrices.

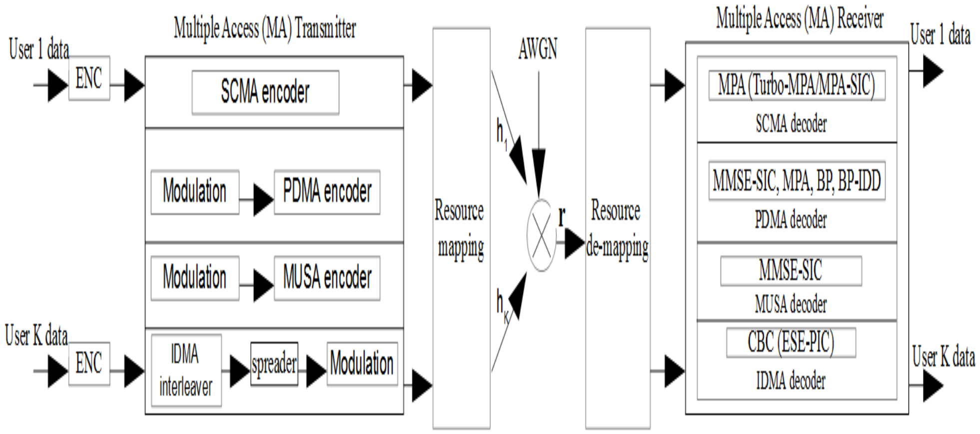


Fig. A.1: General framework of G-NOMA schemes.

2 G-NOMA TECHNIQUES

2.1 Sparse Code Multiple Access (SCMA)

In SCMA, incoming encoded bits of each user are directly mapped to a multi-dimensional sparse codeword of an SCMA codebook. The SCMA codebook is the joint optimization of multi-dimensional modulation and low density spreading. SCMA requires joint-decoding, because a user message at the different REs is jointly encoded as shown in Fig. A.1, where the MPA receiver is used to ensure low complexity detection [8].

System-model: In the SCMA transmitter system, we assume the SCMA encoder contains K separate layers (or users) and each layer has a codebook χ_k, where each codebook contains J dimensional complex sparse codewords N, i.e. $\boldsymbol{\chi_k} = \big(x_k, x_{k1}, ..., x_{KN}\big)$. We consider the case $J \ll K$, namely, overloading and shaping gain is achieved by SCMA and the demand for massive connectivity in 5G is accommodated. Let $\big\{\boldsymbol{x_{n,k}}\big\}_{n=1}^{N}$ denote the vector $\big(x_k, x_{k1}, ..., x_{KN}\big)^T$ then the n^{th} codeword can be written as $\boldsymbol{x_{kn}} = \big\{x_{jkn}\big\}_{j=1}^{J}$. In each layer, $\boldsymbol{d_k}$ bits of each K layer are coded by a low rate channel encoder $\big(ENC_k\big)$, and directly mapped to one sparse codeword on the k^{th} different codebooks. The N codewords are therefore multiplexed over J shared orthogonal resources, such as OFDMA tones or MIMO spatial layer, which are therefore affected by the channel gain vector $\boldsymbol{h_k} = \big(h_{1k}, h_{2k}, ..., h_{JK}\big)$. Let $diag\big(\boldsymbol{h_k}\big)$ denote the diagonal matrix where its m^{th} diagonal element is the m^{th} element of $\boldsymbol{h_k}$. Hence the received signal $\boldsymbol{r}$ is therefore expressed as

$$\boldsymbol{r} = \sum_{k=1}^{K} \boldsymbol{h_k x_k} + \boldsymbol{z}, \tag{A.1}$$

where $\boldsymbol{x_k} = \big(x_{1k}, x_{1k}, ..., x_{JK}\big)^T$ is the k^{th} codeword and $z \sim \mathcal{N}(0, \sigma^2)$ is additive white Gaussian noise (AWGN). The overloading factor is the ratio between the number of k^{th} users and j^{th} orthogonal resources given by $\gamma = K/J$. According to the method of decoding and interference cancellation, the signal to noise-plus interference ratio $SINR$ at the BS can be expressed as

$$SINR_{n,k} = \frac{p_{n,k} g_{n,k} \mid h_{n,k} \mid^2}{\sum_{j \neq k}^{K} p_{n,j} g_{n,j} \mid h_{n,j} \mid^2 + \sigma^2}, \tag{A.2}$$

where $g_{n,k}$ is a variable that shows the connection between the users and the REs i.e., $g_{n,k} = \big\{0, 1\big\}$, where "1」 means the user's data are mapped to the corresponding resource, while the element "0」 means the opposite. Therefore, the achievable sum rate for each user is defined as

$$R_k = \sum_{n=1}^{N} \log_2 \Big(1 + SINR_{n,k}\Big), \tag{A.3}$$

where $p_{n,k} = 1$ is normalized for equal power allocation.

SCMA-receiver: With SCMA making full utilization of the sparsity of codewords joint-decoding is realized in the receiver. Particularly, MUD based on the message passing algorithm (MPA) is realized for separating user's data with acceptable low-complexity detection. The MPA is typically a joint-decoding receiver algorithm that can be explained by the bipartite factor graph [2], which includes the variable and factor nodes. Particularly, in SCMA the variable nodes consist of the transmitted codewords for k^{th} user, while the factor nodes can be seen as the received signals over j^{th} subcarriers. **Algorithm 1** shows a detailed procedure of the MPA Algorithm [5] .

Algorithm 1 MPA detection algorithm for SCMA

```
**Variable Definition**
$V^t_{k\to j}(x_k):$ the message sent from $k^{th}$ user node to $j^{th}$ resource node during $i^{th}$ iteration
$U^t_{j\to k}(x_k):$ the message sent from $j^{th}$ resource node to $k^{th}$ user node during $i^{th}$ iteration
Initializing the marginal probability of each codeword $V^0_{k\to j}(x_{jm}):\ \leftarrow \frac{1}{M}$, for $j \in [J]$ and $k \in [K], m \in [M]$
for $i = 1, 2, ..., N \quad iterations$ do
    update the messages from the resource nodes
    for $j \in [J], k \in [K], m \in [M]:\ (j,k)\exists$ factor graph do
        $U^t_{j\to k}(x_k) \leftarrow \sum_{(x_p)\theta_k\setminus j} \frac{1}{\pi N_0} exp\Big[-\frac{1}{N_0} \parallel r_j - \quad h_{j,k}x_{j,k,m} - \sum_{p\in\theta_j\setminus k} h_{k,p}x_{j,p} \parallel^2\Big] \prod_{p\in\theta_j\setminus k} V^t_{p\to j}(x_p)$
    end for
    update the messages from the resource nodes
    for $j \in [J], k \in [K], m \in [M]:\ (j,k)\exists$ factor graph do
        $V^t_{j\to k}(x_k) \leftarrow p(x_j) \prod_{s\in\theta_k\setminus j} U^{t-1}_{s\to k}(x_{k,m})$
    end for
    Normalize the probabilities of $V^t_{j\to k}(x_k)$ and $U^t_{j\to k}(x_k)$ to        keep them numerically stable
    for $j, k$ with $c_{j,k} = 1$ and $x_j$ in $\chi^j$ do
        $V^t_{k\to k}(x_{j,m}) \leftarrow V^t_{k\to k}(x_{j,m}) / \sum_{m=1}^{M} V^t_{k\to k}(x_{j,m})$
    end for
end for
    Make a decision after some $i^{th}$ iteration
for $k = 1, 2, ..., K$ and $m = 1, 2, ..., M$ do
    $V_k(x_{k,m}) \leftarrow p(x_j) \prod_{j\in\theta_k\setminus j} U^{N\ iterations}_{j\to k}(x_{k,m})$
end for
Finally, the one in $\{x_{k,m}\}^M_{m=1}$ which maximizes $V_k(\cdot)$ is regarded as the transmitted codeword in the $k^{th}$ layer
```

2.2 Multi-User Shared Access (MUSA)

MUSA is a novel multiple access scheme realized in code domain multiplexing. The key principle of MUSA is that, non-orthogonal complex spreading codes with short length are used by multiple users for grant-free data transmissions on the same radio resources. Blind detection based on the SIC receiver is used to decouple the received signals. Different from conventional CDMA, MUSA uses a family of low cross-correlation complex spreading sequences in the uplink to enable overloading and reduce the interference among users [17], while the classical MC-CDMA uses long-binary spreading sequences [3].

System-model: Fig. A.1 illustrates the transceiver structure of the G-NOMA schemes. Generally, for MUSA, K simultaneous users transmit bits of information which are encoded by a low rate encoder (FEC) and modulated by an M-QAM modulator resulting in $\{\boldsymbol{x_k}\}_{k=1}^{K}$. Then each k^{th} user, randomly picks a complex spreading code s_k of short length, and spreads the modulated symbol x_k, which is therefore transmitted over the same n^{th} of the N orthogonal REs, such as OFDMA tones or MIMO spatial layer [11]. Similar to SCMA, overloading is realized by MUSA i.e. $N < K$. Therefore the received signal on the n^{th} RE can be represented as

$$\boldsymbol{r_n} = \sum_{k=1}^{K} \boldsymbol{h_{n,k} s_{n,k} x_k} + \boldsymbol{z_n}, \tag{A.4}$$

where $\boldsymbol{x_k} = \big(x_1, x_2, ..., x_K\big)^T$ is the modulated symbol and $h_{n,k}$ is the channel gain from the k^{th} user to the base station (BS) at the n^{th} RE, $s_{n,k}$ is the n^{th} component of the spreading sequence s_k of k^{th} user. Finally, $z_n \sim \mathcal{N}(0, \sigma^2)$ is additive white Gaussian noise (AWGN) sample.

MUSA-Receiver: MUD at the receiver is implemented by using the minimum mean square error (MMSE) for decoding the superposed user's signals. **Algorithm 2** [2] provides a detailed MMSE-SIC decoding scheme for the system model of eq(A.4). Since the decoding order doesn't influence the sum rate capacity, we assume that the users are decoded in the order of decreasing $SINR$ of their indices without loss of generality. Then the interference that the k^{th} user experiences on the n^{th} resource is denoted as

$$I_{n,k} = \sum_{i \neq k}^{K} b_{i,k} s_{i,k} h_{i,k}, \tag{A.5}$$

where $b_{i,k} = 1$, if the n^{th} resource is allocated to the k^{th} user and $b_{i,k} = 0$ otherwise. Then the sum rate of the k^{th} user on the n^{th} RE is given by eq(A.3).

Algorithm 2 MMSE-SIC based ordering for MUSA

```
Input Variables
$\boldsymbol{r}$ :  the received signal for all orthogonal subcarriers
$\boldsymbol{H}$ :  channel response matrix
$K$ :  number of users

for $k = 1, 2, ..., K$ do
    Estimate the MMSE transformation weight matrix
    $\boldsymbol{W_{MMSE}} = \underset{\boldsymbol{W_{MMSE}}min}{E} \left[ \parallel \boldsymbol{x} - \boldsymbol{W_r} \parallel^2 \right]$
                $= \left(\boldsymbol{H^H H} + \sigma^2 I\right)^{-1} \boldsymbol{H^H}$
    Where $I$ is an identity matrix based on the $k^{th}$ users
    for $i = 1, 2, ..., K - \left(k - 1\right)$ do
        Calculate the post $SINR$ for each $k^{th}$ user
        $SINR_i = \frac{|\boldsymbol{W_{i,MMSE} H_i}|^2}{\sum_{l \neq i} |\boldsymbol{W_{i,MMSE} H_i}| + \sigma^2 \|\boldsymbol{W_{i,MMSE}}\|^2}$
    end for
    Perform post-ordering of $SINR$ in decreasing order
    Estimate $\hat{\boldsymbol{x}}$ of the $k^{th}$user with highest $SINR$
                $\hat{\boldsymbol{x_k}} = \boldsymbol{W_{k,MMSE} * r}$
                $\hat{\boldsymbol{r_k}} = \boldsymbol{r} - \boldsymbol{h_k} \hat{\boldsymbol{x_k}}$
    Remove $\boldsymbol{h(j)}$ from the channel matrix $\boldsymbol{H}$
end for
```

2.3 Pattern Divison Multiple Access (PDMA)

Pattern division multiple access (PDMA), is a novel G-NOMA scheme based on the successive interference cancellation amenable multiple access (SAMA) technique. In order to address the challenges of massive connectivity and higher spectral efficiency for the 5G networks. Different from SCMA with no strict rule of using low-density spreading codes [8], PDMA uses PDMA pattern to define a sparse mapping from data to a group of resources. The pattern is introduced to differentiate signals of users sharing the same resources [6]. This work investigates PDMA in code domain to make full use of the various channel resources. PDMA can be detected by either a SIC-based receiver or MPA receiver. The PDMA is originally designed based on the SIC receiver [14]. In this paper we adopt the MPA receiver, for low complexity detection.

System-model: Fig. A.1, as for the PDMA system, non-orthogonal patterns are realized to maximize the diversity and to minimize the interference of multiple users. K users and their data are mapped onto N REs by using distinguished PDMA patterns g_k.The k^{th} users PDMA patterns on N REs construct a PDMA pattern matrix $\boldsymbol{G}_{\boldsymbol{PDMA}}^{\left[\boldsymbol{N},\boldsymbol{F}\right]} = \left[g_1, g_2, \ldots, g_K\right]$ with the dimensions of $N \times K$. For PDMA system, the received signal at the receiver from the transmitter is obtained by spreading the user's modulation symbol x_k according to the PDMA pattern g_k [12]. In the uplink, the received signal on the n^{th} RE can be expressed as

$$\boldsymbol{r_n} = \sum_{k=1}^{K} \boldsymbol{H_{PDMA}}\big(\boldsymbol{n}, \boldsymbol{k}\big)\sqrt{p_{n,k}}\boldsymbol{x_k} + \boldsymbol{z_n}, \tag{A.6}$$

where

$$\begin{aligned} \boldsymbol{H_{PDMA}} &= \boldsymbol{H_{Ch}} \odot \boldsymbol{G}_{\boldsymbol{PDMA}}^{\left[\boldsymbol{N},\boldsymbol{K}\right]} \\ &= \left[h_1, h_2, \ldots, h_K\right] \odot \left[g_1, g_2, \ldots, g_K\right], \\ &= \begin{bmatrix} h_{11}g_{11} & h_{12}g_{12} & \cdots & h_{1K}g_{1K} \\ h_{21}g_{21} & h_{22}g_{22} & \cdots & h_{2K}g_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N1}g_{N1} & h_{N2}g_{N2} & \cdots & h_{NK}g_{NK} \end{bmatrix}, \end{aligned} \tag{A.7}$$

indicates the PDMA equivalent channel response matrix from users to BS with dimension $N \times K$, $\boldsymbol{H_{Ch}} = \left[h_1, h_2, \ldots, h_K\right]$ is the channel gain of the k^{th} users, x_k is the transmit symbol from the k^{th} user to BS, $p_{n,k}$ denotes the allocated transmit power of k^{th} user at n^{th} REs. We assume that the total transmit power P allocated in each RE for each user is equal. In this work, the transmit power is normalized $p_{n,k} = 1$ for equal power allocation. Then eq(A.6) becomes:

$$\boldsymbol{r_n} = \sum_{k=1}^{K} \boldsymbol{H_{PDMA}(n,k)x_k} + \boldsymbol{z_n}, \tag{A.8}$$

where $z_n \sim \mathcal{N}(0,\sigma^2)$ denotes the corresponding complex AWGN at the n^{th} RE. The signal for the j^{th} user, where $j \neq k$, will be treated as interference as the BS detects the signal of k^{th} user. Hence, the achievable sum rate of the k^{th} user on the n^{th} RE is as in eq(A.3). Whereby the SINR given by

$$SINR_{n,k} = \frac{p_{n,k} g_{n,k} \mid h_{n,k} \mid^2}{\sum\limits_{j \neq k}^{K} p_{n,j} g_{n,j} \mid h_{n,j} \mid^2 + \sigma^2}, \tag{A.9}$$

denotes the k^{th} user's signal-to-Interference-plus-Noise Ratio (SINR) at the n^{th} RB in the receiver.

PDMA-receiver: At the receiver, the MPA-receiver (refer to **Algorithm 1**) scheme is implemented to facilitate interference cancellation. However, for non-orthogonal transmission employing a SIC receiver, the error propagation problem is introduced. The MPA is used to maintain a low complexity detection. However, the MPA receiver suffers from the detection complexity problem. Therefore, improvements of the MPA can be used to reduce complexity, such as using the projected constellation methodology to construct the codebooks [3], [18]. In addition, the Gaussian approximation of interference (GAI), which models the interference-plus-noise as Gaussian distributed, and such approximation tends to be more accurate as the amount of connectivity becomes larger in 5G [3].

2.4 Interleave Division Multiple Access (IDMA)

The idea behind IDMA is non-orthogonal transmission and the application of using, user-specific interleavers as the only means for user separation. IDMA allows a very simple chip-by-chip (CBC) iterative sub-optimal MUD strategy in the receiver. Regarding IDMA, some researchers have developed different interleaver's designs, including Random Interleaver, New Logistic Map Interleavers (NLM Interleavers). Helical Interleaver, with the aim of reducing the bandwidth, memory resources, and increased spectral efficiency [13] - [19].

System-model: From Fig. A.1, the IDMA system utilizes the bit level interleaver MA encoder. Generally, the information bits of K users is sequentially processed by a low rate channel encoder (FEC), interleaved by using a unique user interleaver [15] and spread by a binary sequences, which is different from MUSA low-correlation complex spreading codes. The spreaded bit sequence is then modulated using M-QAM modulator to yield $\{\boldsymbol{x_k}\}_{k=1}^{K}$. The received signal on each j^{th} frame at the base station is given by

$$\boldsymbol{r_n} = \sum_{k=1}^{K} \boldsymbol{h_k x_k}(\boldsymbol{j}) + \boldsymbol{z}(\boldsymbol{j}), \tag{A.10}$$

where x_k, h_k and $z\big(j\big)$ is the transmitted signal of k^{th} user, complex channel coefficient between the k^{th} users and the base station, and the samples of the additive white Gaussian noise (AWGN) with $\sigma^2 = N_0/2$ respectively. We assume that the channel coefcients h_k are known a priori at the receiver.

IDMA-receiver: We adopt an iterative sub-optimal receiver as Chip-by-chip (CBC) MUD scheme, as illustrated in the transceiver model of Fig A.1, which consists of an elementary signal estimator (ESE) and a posteriori probability (APP) decoders (DECs) that iteratively exchange extrinsic information of the k^{th} users. **Algorithm 3** [13], provides a detailed CBC decoding schemes for the IDMA. Denoted by $SINR_k^{\big(l\big)}$ the average $SNIR$ for the outputs of the ESE after the l^{th} iteration. Let $f_k SINR_k^{\big(l\big)}$ be the average variance of the outputs of DEC driven by an input sequence with $SINR_k^{\big(l\big)}$. Then the $SINR$ is given by

$$SINR_k^{\big(l+1\big)} = \frac{p_k \mid \boldsymbol{h_k} \mid^2}{\sum_{j\neq k}^{K} p_i \mid \boldsymbol{h_i} f_k SINR_k^{\big(l\big)} \mid^2 + \sigma^2}, \tag{A.11}$$

where l is the maximum number of iterations. Hence the achievable sum rate for each user is defined as in eq(A.3),where $p_k = 1$ is normalized for equal power allocation.

3 Simulation Results and Discussions

The uplink BER rate and achievable sum rate performance are investigated and presented in this section. The codebook design in SCMA is according to [10] and the PDMA non-orthogonal patterns design is based on the principle in [12]. Spreading sequences in MUSA are generated by a pseudo-random sequence whose real and imaginary part is a 3-ary set which includes $\big\{-1, 0, 1\big\}$, so each element of the complex spreading codes is from the set $\big\{-1+i, -1, -1-i, i, 0, -i, 1+i, 1, 1-i\big\}$ before normalization. The IDMA is employed with random interleaving as in [13]. The simulation parameters are presented in Table A.1. In addition, the OFDMA simulation parameters are according to [7] for performance bench-marking.

Algorithm 3 Chip-by-Chip Detection for IDMA

```
Input Variables
$\boldsymbol{r}$ :  the received signal for all orthogonal subcarriers
$\boldsymbol{H}$ :  channel response matrix
$K$ :  number of users
for $n = 1, 2, ..., N \quad iterations$ do
    Perform each task for each user
    for $k = 1, 2, ..., K$ do
        Calculate the expected value of each user
        $E(\boldsymbol{x_k}(\boldsymbol{j})) = tanh(e_{DEC}(\boldsymbol{x_k}(j))/2)$
        Calculate the variance of each users data
        $Var(\boldsymbol{x_k}(j)) = 1 - (E(\boldsymbol{x_k}(j)))^2$
        Sum of the mean of each user
        $E(\boldsymbol{r}(\boldsymbol{j})) = \sum_{k=1}^{K} \boldsymbol{h_k} E(\boldsymbol{x_k}(j))$
        Sum of the variance of each user
        $Var\Big(\boldsymbol{r}(\boldsymbol{j})\Big) = \sum_{k=1}^{K} \mid \boldsymbol{h_k} \mid^2 Var(x_k(j)) + \sigma^2$
        Compute the Extrinsic LLR from ESE of each user
        $e_{ESE}(\boldsymbol{x_k}(\boldsymbol{j})) = 2\boldsymbol{h_k} * \frac{\boldsymbol{r}(j) - (E(\boldsymbol{r}(\boldsymbol{j})) - \boldsymbol{h_k} E(\boldsymbol{x_k}(\boldsymbol{j})))}{Var(\boldsymbol{r}(\boldsymbol{j})) - |h_k|^2 Var(\boldsymbol{x_k}(\boldsymbol{j}))}$
    end for
end for
```

Table A.1: Simulation parameters

Parameters	**Assumptions**
System Bandwidth	10MHz
Carrier frequency	2GH
Channel Model	Rayleigh Flat fading
Channel estimation	Perfect
Overloading Factors	150 and 200
Decoding Iterations	10 for both MPA and ESE
Modulation coding rate	QPSK; Convolutional coding, Random interleaver for IDMA
Receiver model	MPA receiver for SCMA and PDMA, MMSE-SIC for MUSA, linear MMSE receiver for OFDMA, ESE Chip-by-Chip receiver for IDMA

Figs. A.2 - A.3 illustrates the BER performance vs the SNR of the G-NOMA schemes with diverse overloading factors. Without loss of generality, we assume the allocated power for all REs is kept the same for all G-NOMA and OMA schemes. The results in Fig. A.2, show that G-NOMA technologies outperform the OMA when users are overloaded with 150 percent. This is due to the diversity gain from the sparsity of codewords, PDMA pattern, bitstream interleaving, and complex spreading. MUSA on the other hand, shows a different tendency when 150 percent overloading case is exhibited against OFDMA as in Fig. A.2. This is because MUSA adopted a linear MMSE-SIC receiver to mitigate inter-user interference. It reduces the power of the desired signal and causes error propagation problem which suppresses the interference. Hence diversity gain loss occurs. IDMA on the other hand, shows excellent performance when compared to other G-NOMA schemes even in an uncoded environment. This can be attributed to by the type of interleaver implemented to differentiate the users. Moreover, it can be observed that SCMA with 150 percent user overloading has a higher gain when compared to PDMA and MUSA with 150 percent user overloading.

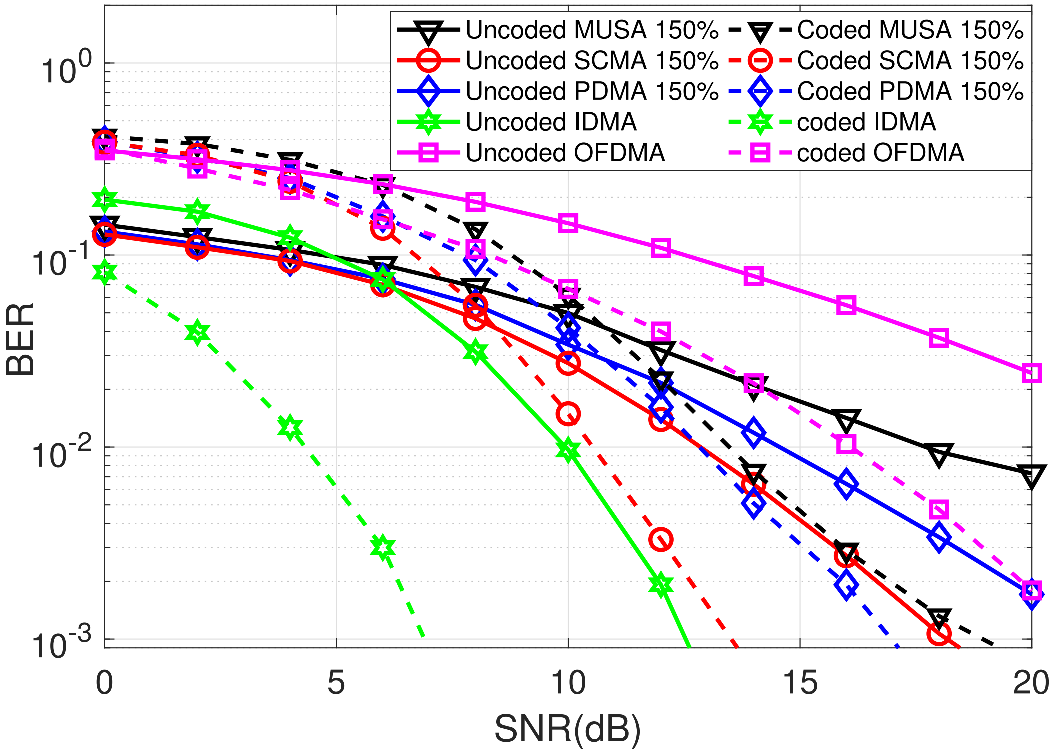


Fig. A.2: BER performance of G-NOMA schemes with 150 percent overloading

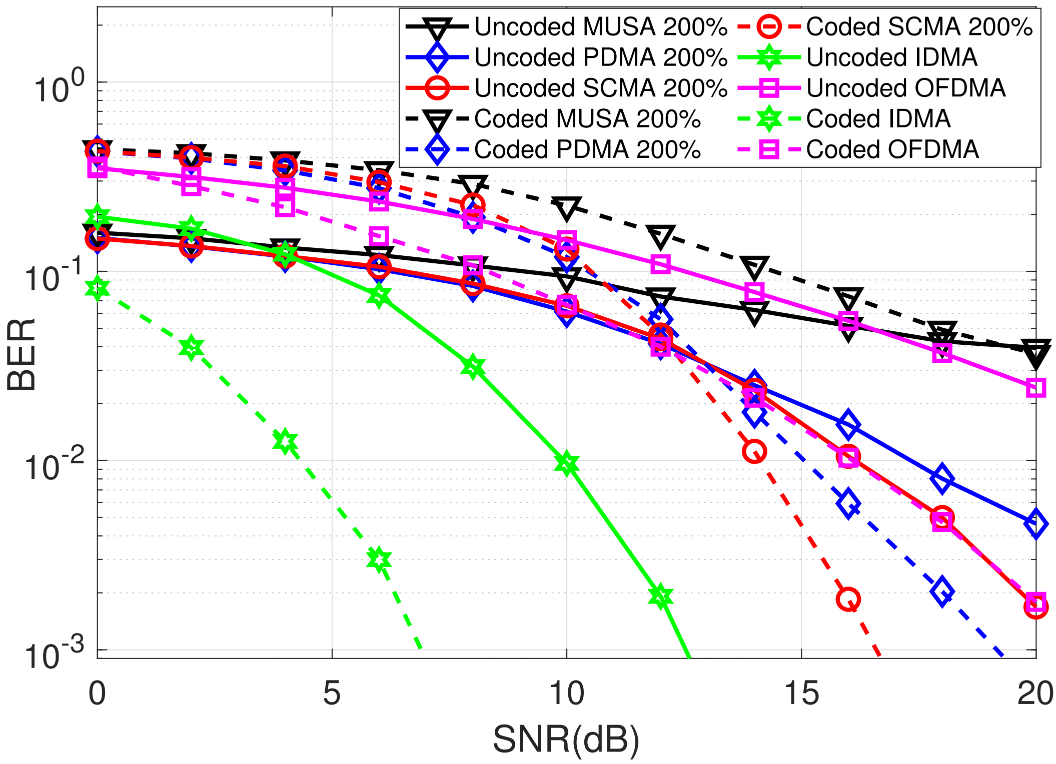


Fig. A.3: BER performance of G-NOMA schemes with 200 percent overloading

Fig. A.3. Exhibits the BER performance of the G-NOMA techniques with 200 percent overloading case. The BER curves of MUSA with 200 percent user overloading has a unique tendency. Hence it is saturated. This is because the spreading factor is much smaller than the number of the overloaded users, therefore the effective channel matrix for the MMSE weight filter does not have rank enough to differentiate the larger number of users' signals. This means that it is more difficult to separate each users' signal as the overloading factor increases. Moreover, the diversity gain of MUSA is relatively low compared to the gain in other G-NOMA schemes. However it should be noted that not all the G-NOMA schemes outperform the OFDMA. Even so, MUSA can overcome this problem with the aid of moderate user overloading factors.

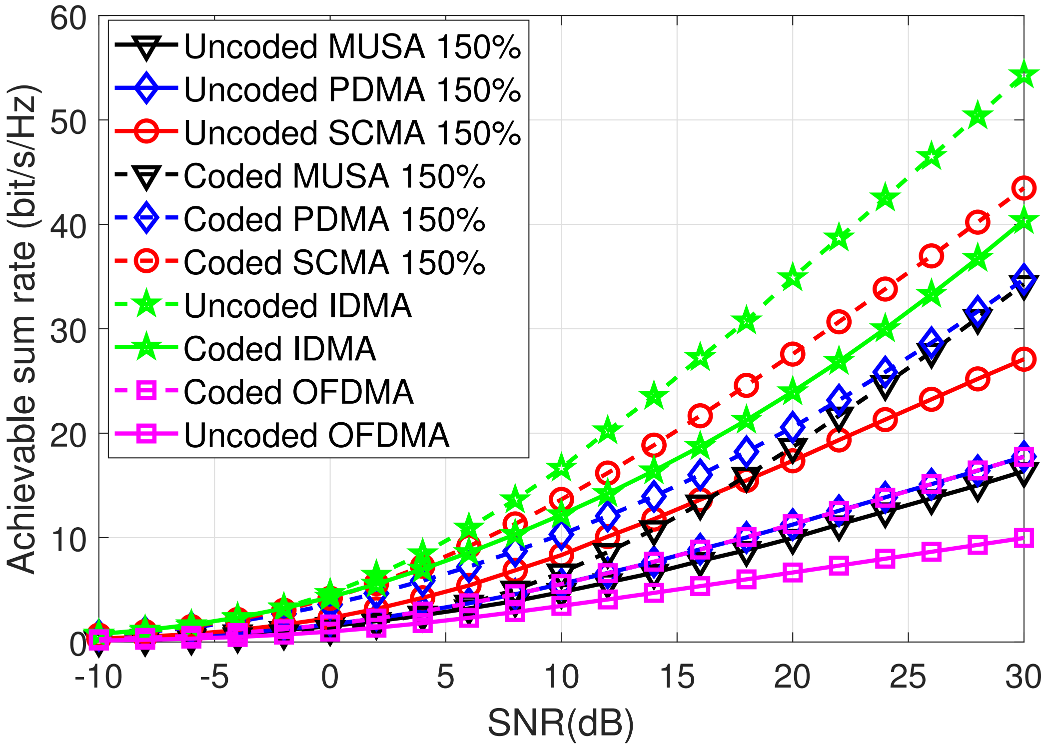


Fig. A.4: Achievable sum rate with 150 percent overloading

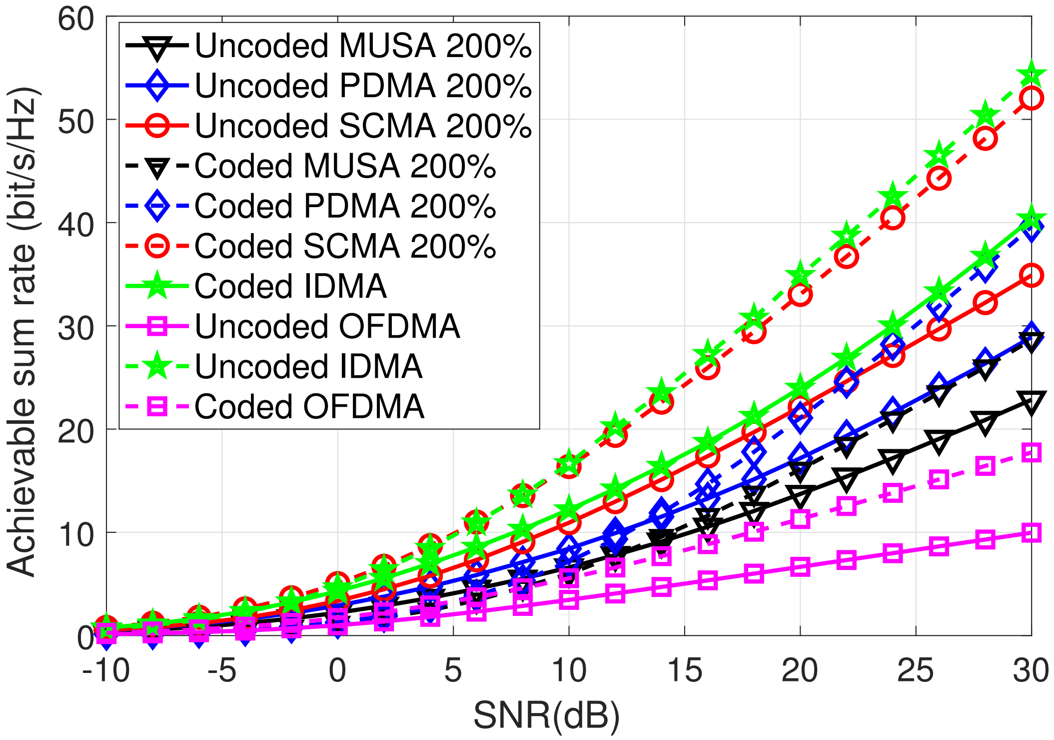


Fig. A.5: Achievable sum rate with 200 percent overloading

Figs. A.4 - A.5 shows the achievable sum rate performance of the G-NOMA techniques with 150 and

200 percent overloading case. From Fig. A.4 we observe that the G-NOMA schemes, performs better than the OMA scheme. However, for uncoded PDMA and MUSA the results are almost the same, when compared to the OFDMA. On the other hand, from Fig. A.5, when the overloading factor is 200 percent, all the G-NOMA schemes outperform the OFDMA techniques. The increased overloading also increase the achievable sum rate performance.

4 Conclusion

In this paper, we have compared the performance of four typical G-NOMA schemes, in terms of BER and achievable sum rate. It was found that IDMA has excellent performance and SCMA has the best performance due to the near-optimal design of sparse codewords together with the near-optimal MPA receiver, while PDMA outperforms MUSA. This comparison study reveals that to obtain better system performance for G-NOMA techniques, advance IDMA interleavers and the design of sparse codebooks in SCMA, low-correlation spreading sequences in MUSA, and non-orthogonal patterns in PDMA should be optimized. Additionally, robust receiver with low complexity is also expected for NOMA.

References

[1] K. Au, L. Zhang, H. Nikopour, E. Yi, A. Bayesteh, U. Vilaipornsawai, J. Ma, and P. Zhu, "Uplink contention based scma for 5g radio access," in *Globecom Workshops (GC Wkshps), 2014*. IEEE, 2014, pp. 900–905.

[2] B. Wang, K. Wang, Z. Lu, T. Xie, and J. Quan, "Comparison study of non-orthogonal multiple access schemes for 5g," in *Broadband Multimedia Systems and Broadcasting (BMSB), 2015 IEEE International Symposium on*. IEEE, 2015, pp. 1–5.

[3] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang, "Non-orthogonal multiple access for 5g: solutions, challenges, opportunities, and future research trends," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74–81, 2015.

[4] Z. Gao, L. Dai, S. Han, I. Chih-Lin, Z. Wang, and L. Hanzo, "Compressive sensing techniques for next-generation wireless communications," *IEEE Wireless Communications*, 2018.

[5] C. Zhang, Y. Luo, and Y. Chen, "A low complexity detection algorithm for scma," *arXiv preprint arXiv:1611.08853*, 2016.

[6] B. Ren, Y. Wang, X. Dai, K. Niu, and W. Tang, "Pattern matrix design of pdma for 5g ul applications," *China Communications*, vol. 13, no. Supplement2, pp. 159–173, 2017.

[7] A. Idris, A. N. Farhana, H. Adiba, and M. Kassim, "Ber and papr analysis of mimo ofdma and scfdma system using different diversity techniques," in *Control System, Computing and Engineering (ICCSCE), 2017 7th IEEE International Conference on*. IEEE, 2017, pp. 293–298.

[8] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. Bhargava, "A survey on non-orthogonal multiple access for 5g networks: Research challenges and future trends," *arXiv preprint arXiv:1706.05347*, 2017.

[9] A.-I. Mohammed, M. A. Imran, and R. Tafazolli, "Low density spreading for next generation multicarrier cellular systems," in *Future Communication Networks (ICFCN), 2012 International Conference on*. IEEE, 2012, pp. 52–57.

[10] H. Nikopour and H. Baligh, "Sparse code multiple access," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*. IEEE, 2013, pp. 332–336.

[11] Z. Yuan, G. Yu, W. Li, Y. Yuan, X. Wang, and J. Xu, "Multi-user shared access for internet of things," in *Vehicular Technology Conference (VTC Spring), 2016 IEEE 83rd*. IEEE, 2016, pp. 1–5.

[12] S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, and K. Niu, "Pattern division multiple access—a novel nonorthogonal multiple access for fifth-generation radio networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3185–3196, 2017.

[13] L. Ping, L. Liu, K. Wu, and W. K. Leung, "Interleave division multiple-access," *IEEE Transactions on Wireless Communications*, vol. 5, no. 4, pp. 938–947, 2006.

[14] C. YAN, Z. YUAN, W. LI, and Y. YUAN, “Non-orthogonal multiple access schemes for 5g,” *ZTE Communications*, vol. 14, no. 4, pp. 11–16, 2016.

[15] S. Bajpai and D. Srivastava, “Performance analysis of idma scheme using different coding techniques with receiver diversity using random interleaver,” *IJSRET), ISSN*, pp. 2278–0882, 2014.

[16] Y. Wang, B. Ren, S. Sun, S. Kang, and X. Yue, “Analysis of non-orthogonal multiple access for 5g,” *China Communications*, vol. 13, no. Supplement2, pp. 52–66, 2016.

[17] R1-166404, “Reciever details and link performance for musa,” *3GPP TSG RAN WG1*, pp. 22–26, August 2016.

[18] G. Song and X. Wang, “Comparison of interference cancellation schemes for non-orthogonal multiple access system,” in *Vehicular Technology Conference (VTC Spring), 2016 IEEE 83rd.* IEEE, 2016, pp. 1–5.

[19] P. Wang, Q. Chen, L. Ping, and L. Liu, “Power allocation for practically coded idma systems over broadcast channels,” in *Information Theory Workshop, 2006. ITW’06 Chengdu. IEEE.* IEEE, 2006, pp. 673–677.

Paper B

Hybrid G-NOMA Spectrum Resource Allocation Scheme for 5G Small-Cell Networks

Samson Manyani Zitha and Tom Walingo

The paper is under review at the

South African Institute of Electrical Engineers (SAIEE) Journals

The layout has been revised.

Abstract

Non-orthogonal multiple access (NOMA) schemes improve the spectral-efficiency and data-rates of fifth generation (5G) heterogeneous networks (HetNets). Advancement of these schemes will further maximize the performance of the networks. This work develops a hybrid generalized-NOMA (G-NOMA) scheme for a two-tier HetNet. The hybrid G-NOMA (HG-NOMA) scheme combines different resource pattern assignment and power allocation for the different users multiplexed on the same spectrum resource element (SRE) of the network. The resource assignment and power allocation problem is formulated as an energy-efficiency (EE) maximization problem with the aim of maximizing user's connectivity, EE and sum-rate capacity of the small-cells. For reception, a low-complexity hybrid G-NOMA successive interference cancellation (SIC) receiver that combines power-levels and diversity pattern gain to realize multi-user detection is proposed. Performance results show the superiority of the HG-NOMA scheme as compared to the traditional orthogonal multiple access schemes for small-cells in terms of sum-rate capacity, EE and complexity, hence demonstrating their suitability.

1 Introduction

The proliferation of modern wireless communication technologies and their applications has accelerated the demand for high data-rates and spectral-efficiency. As a result, energy-efficiency (EE), spectrum scarcity and hence spectrum resource utilization, have emerged as major challenges for the fifth generation (5G) networks. To mitigate these challenges, new standardized communication technologies need to be developed and well investigated. In the current predecessor LTE-A networks, orthogonal frequency multiple access (OFDMA) scheme is used for spectrum access. It allows the user's to be served orthogonally in time-frequency spectrum to alleviate narrowband interference and impulsive noise. Even though OFDMA increased the spectrum capacity of previous networks, more is required for future networks. Driven by these spectrum limitations of orthogonal multiple access (OMA) schemes, non-orthogonal multiple access (NOMA) technologies, envisioned for 5G networks, have been proposed for their potential ability to yield high spectrum-efficiency and enhanced massive connectivity of user access [1]. The significance of NOMA is co-multiplexing users on the same spectrum resource elements (SREs) (i.e. time slot, OFDMA subcarrier, or spreading code) via power domain (PD) or code domain (CD) at the transmitter and successfully separating them at the receiver by multi-user detection (MUD) schemes. NOMA schemes permit controllable interference by non-orthogonal resource allocation albeit increase in receiver complexity [2].

The generalized-NOMA (G-NOMA) schemes proposed to alleviate the 5G multiple access demands include: Interleave division multiple access (IDMA) [3], which explores multiplexing by employing user specific interleavers; PD-NOMA where multiplexing occurs by allocating different power levels to different users on the same orthogonal SRE according to their channel conditions [4]; Sparse code multiple access (SCMA) [5] and low-density spreading multiple access (LDSMA) [6] based on the idea of user information being spread over multiple subcarriers using code words; Multi-user shared access (MUSA) [7], an improved version of the code division multiple access (CDMA), which uses low-correlation spreading codes to overcome overloading at the transmitter and employing successive interference cancellation (SIC) to realize MUD at the receiver; Finally, the pattern division multiple access (PDMA) [8], which uses non-orthogonal patterns to maximize diversity and the overlapping of multiple users. Generally, G-NOMA schemes achieve overloading by non-orthogonal spectrum resource allocation, which enables massive connectivity and improves system spectral-efficiency. The development and performance improvement of the G-NOMA techniques together with their challenges over the traditional OMA techniques is exhaustively done in [9], [10], and need not be emphasized.

HetNets based on a large-scale deployment of small-cells with low power nodes overlaid in macrocells present a promising solution to cope with spectrum scarcity for the 5G networks [11]. In particular, small-cell networks are motivated by the principle of rethinking network design to offer unprecedented network capacity and improved spectrum-efficiency, while minimizing energy consumption by adopting energy-efficient small-cell architectures [12]. This work primarily focuses on small-cells to maximize the system sum-rate and EE. Due to their small sizes, they present the challenges of spectrum resource allocation and interference management (i.e., cross-tier and intra-tier interference). To alleviate the cross-tier interference, resource optimization involving hybrid G-NOMA (HG-NOMA) scheme is proposed for the small-cells. Resource optimization for OFDMA to improve spectral-efficiency and data-rates in the point-to-point transmission scenario has been extensively done for single and HetNets. Resource allocation strategies for OFDMA downlink have been investigated in [13], [14] for single networks. In [15], [16], resource allocation schemes for heterogeneous macrocell/femtocell networks have been developed. Uplink resource allocation strategies and EE in OFDMA have been done in [17] [18]. The same cannot be said on the development and investigation of G-NOMA schemes.

1.1 Related Works

Several aspects of the G-NOMA power & resource allocation schemes have been investigated. The performance improvement of NOMA by optimizing the receiver design and the radio resource allocation was investigated in [19], at the cost of restrained computational complexity and limited number of users mapped per SRE. An advanced user pairing and power allocation based resource allocation for a downlink NOMA system is studied in [20]. In [21] and [22], the proportional fairness scheduling was investigated for downlink NOMA. However, little is known about the computational complexity and tractability of the NOMA resource allocation algorithm. A genetic algorithm based NOMA-OFDM downlink radio resource allocation scheme was proposed in [23], to reach a target solution which balances the trade-off between system throughput and user fairness. In addition, the coverage and throughput performances of HetNets with the non-coordinated NOMA and coordinated joint transmission NOMA schemes were proposed in [24]. The works in [25] and [26] extend resource allocation to the convergence of HetNets with NOMA in multi-tier networks, showing greater improvement in system capacity and throughput. An iterative distributed power control was studied in [27] to maximize the sum-rate of NOMA HetNet under the constraints of total transmit power and users QoS requirement. This showed greater improvement in spectral efficiency and lower outage performance when compared to OMA HetNets. An elegant uplink subchannel and power allocation scheme for NOMA system was proposed in [28]. The authors in [29] and [30] proposed iterative resource allocation algorithms for a typical NOMA system, but did not consider EE. Most of these resource allocation works do not feature a combination of EE resource allocation strategies for multiple users on the uplink of a NOMA HetNet.

Recently, the potential of EE in 5G HetNets has been explored, and various energy-efficient resource optimization problems studied for NOMA. For a downlink NOMA system, energy-efficient subchannel power allocation schemes were proposed in [12, 31]. Nevertheless, the number of users was limited to two in the subchannel and power allocation algorithm, greatly restraining the application of NOMA. In [32], NOMA was implemented in ultra-dense HetNet to optimize the system EE and fairness. In [33], the comparison analysis of PD-NOMA and SCMA was investigated for downlink HetNets, where SCMA achieved better sum-rate capacity performance at the expense of system complexity. The trade-off between data-rate performance and energy consumption in NOMA was examined in [12] as the problem of EE user scheduling and power optimization in downlink NOMA HetNets for perfect and imperfect CSI, respectively. Considering, dual connectivity, coordinated multi-point transmission and PD-NOMA, the downlink energy efficiency of heterogeneous cloud radio access network (H-CRAN) was maximized in [34]. The authors in [35] investigated an energy cooperation resource allocation for a two-tier HetNets with NOMA, where the BS's are powered by both renewable energy sources and the conventional grid. A combination of EE resource allocation NOMA features mentioned earlier was not considered in these NOMA schemes.

Few attempts to develop hybrid schemes have been undertaken. In [1], a downlink hybrid HetNets framework was proposed, where NOMA is adopted in small cells and MIMO employed in macrocells. The trade-off for achieving spectral-efficiency and EE with users' minimum rate requirements in hybrid multi-carrier non-orthogonal multiple access systems incorporating both NOMA and OMA modes into one unified framework was investigated in [36]. A system-level evaluation of the impact of NOMA in the downlink capacity dimension of a HetNet with hybrid multiple access where NOMA and orthogonal multiple access coexist was investigated in [26]. Different from the conventional OMA HetNets design , G-NOMA enhanced HetNets design posses additional challenges of co-channel and cross-tier interference, EE problems, and results in high-complexity receivers [13]. These hybrid schemes do not combine all the NOMA features. They do not address the problem of how small-cell improves the system EE and sum-rate, by considering joint resource assignment and power allocation in a macrocell environment. Furthermore, they do not employ hybrid MUD receivers that combine power resource and patterns in the reception. Accordingly, while using the traditional power-domain NOMA in HetNets, the appropriate allocation of different powers to different users to improve the EE and sum-rate performances of the overall system is still not very clear, in these works

1.2 Contribution and Methodology

This work proposes an EE resource allocation framework for NOMA enhanced small-cells to maximize the systems EE. The primary contributions and methodology of this paper is summarized as follows:

- We propose a hybrid G-NOMA (HG-NOMA) resource allocation scheme that integrates PD-NOMA and PDMA on the uplink of 5G small-cells network. Power resources are chosen as the fundamental multiplexing domain between the MUEs and FUEs, and code domain as the key multiplexing domain in the sparse pattern mapping of the FUEs.
- Based on the proposed model, we formulate an EE maximization problem by considering joint spectrum and power allocation, with the aim of maximizing massive connectivity of user access,

system EE and sum-rate capacity enhancement of the small cells. In addition, we propose a low-complexity HG-SIC receiver that combines both the power-levels $\&$ diversity (patterns) gain to realize MUD.

- To solve the non-convex EE problem of each FBS, we invoke the langrange dual problem to iteratively update the resource assignment and power allocation vector by solving the approximate convex problem. The proposed algorithm is convergent and the solution satisfies the Karush-Kuhn-Tucker (KKT) conditions.
- We demonstrate that the proposed EE HG-NOMA scheme outperforms the traditional OMA based small-cells in terms of sum-rate capacity and EE. This demonstrates the practicality of the proposed scheme.
- Finally, we show that, the computational complexity of the proposed EE HG-NOMA increases with the number of FBS, and users in each FBS.

1.3 Organization and Notations

The rest of the paper is organized as follows, Section II presents the system model and problem formulation. The proposed energy-efficient resource allocation schemes is detailed in Section III. In Section IV, the simulation results and performance evaluation are discussed. Finally, Section V concludes the paper.

Notation: The bold upper and lower-case letters denote matrices and column vectors, respectively. $I_{N\ \times N}$ Is an N by N identity matrix. The following superscripts $\big(\cdot\big)^{H}$, $\big(\cdot\big)^{-1}$ and $\big(\cdot\big)^{T}$ represents Hermitian, inverse and transpose operator. $|\cdot|$ and $||\cdot||$ denote the absolute value of a scalar and the Frobenius norm. $\odot$ indicates the element-wise dot product of two matrices.

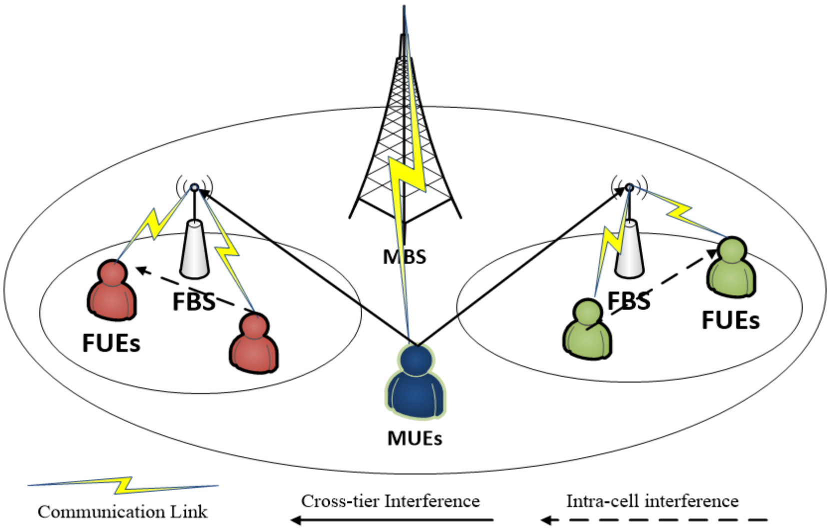


Fig. B.1: Illustration of an uplink HG-NOMA HetNets works with one MBS and FBS.

2 System Model and Problem Formulation

We consider a two-tier conventional HetNet model, with a centralized macro-cell base station (MBS) in the macro-cell, as shown in Fig. B.1. We assume that M MUEs are uniformly distributed within the macrocell overlaid with F femtocells characterized by low-power constraint which limits their coverage. In each of the F femtocells, there exists a centralized femtocell base station (FBS) and K FUEs that are uniformly distributed within the femtocell. In this work, the FUEs and MUEs are co-multiplexed over the same time-frequency spectrum resource elements (SREs) to alleviate massive connectivity and improve the overall EE under the assumption of perfect channel state information (CSI). Block fading is adopted in this work, where the fading environment of each SRE is deduced to be the same within an SRE, but differs independently across different SREs. We specifically focus on the small-cells, by considering the intra-tier interference and cross-tier interference from the MUEs. Particularly, the MUEs, FUEs, and FBS are equipped with a single antenna.

2.1 Transmitter power and pattern allocation

The network has a total bandwidth B, divided into N SREs, with each SRE occupying a bandwidth $B_{sc} = B/N$. The transmitter assigns $P_{i,k,n}^{FUE}$ and $P_{m,n}^{MUE}$ power levels to the k^{th} FUE in the i^{th} FBS on the n^{th} SRE and transmit power to the m^{th} MUE on the n^{th} SRE, respectively. Defining $h_{i,k,n}^{FUE}$ and $h_{i,m,n}^{MF}$, as the channel gain from the k^{th} FUE to i^{th} FBS on the n^{th} SRE and the channel gain on the n^{th} SRE link of the m^{th} MUE to the i^{th} FBS. Let $G_{FUE}^{[I,K,N]} = [\alpha_{i,k,n}]_{F\times K\times N}$ be the SRE

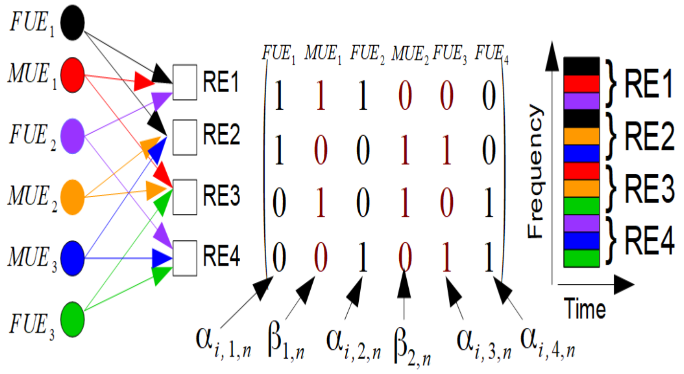


Fig. B.2: Channel model of the uplink Hybrid G-NOMA enabled HetNets.

HG-NOMA transmitter pattern matrix for the small cells, where $\alpha_{i,k,n} = 1$ denotes that n^{th} SRE is assigned to the k^{th} FUE communicating to the i^{th} FBS. Similarly, denoting $G_{MUE}^{\left[M,N\right]} = \left[\beta_{m,n}\right]_{M \times N}$ as SRE HG-NOMA pattern transmitter matrix for the MUEs, where $\beta_{m,n} = 1$ denotes that n^{th} SRE is assigned to the m^{th} MUE. The SRE pattern allocation matrix is shown in Fig. B.2. Consequently, the received signal by the i^{th} FBS for k^{th} FUE on the n^{th} SRE is given by

$$y_{i,k,n} = \underbrace{G_{FUE}^{\left[I,K,N\right]}\sqrt{P_{i,k,n}^{FUE}}x_k}_{\text{FUE desired signal}} + \underbrace{\sum_{j \neq k}^{K} G_{FUE}^{\left[I,J,N\right]}\sqrt{P_{i,j,n}^{FUE}}x_j}_{\text{intra-cell interference}} + \underbrace{\sum_{m=1}^{M} G_{MUE}^{\left[M,N\right]}\sqrt{P_{n,m}^{MUE}}x_m}_{\text{Cross-tier interference}} + z_{i,k,n} \tag{B.1}$$

where x_k, x_j and x_k are the transmitted symbols, and $z_{i,k,n}$ is the AWGN noise. Hence, the data-rate of the k^{th} FUE to the i^{th} FBS on the n^{th} SRE is given by

$$\Gamma_{i,k,n}^{FBS} = \frac{P_{i,k,n}^{FUE} h_{i,k,n}^{FUE} G_{FUE}^{\left[I,K,N\right]}}{\sum\limits_{j=1,j\neq k}^{K} P_{i,j,n}^{FUE} h_{i,j,n}^{FUE} G_{FUE}^{\left[I,J,N\right]} + I_n^M + \sigma_z^2}. \tag{B.2}$$

where $I_n^M = \sum_{m=1}^{M} G_{MUE}^{\left[M,N\right]} P_{m,n}^{MUE} h_{i,m,n}^{MF}$ is the cross-tier interference caused by the MUEs, and σ_z^2 is the received background noise. Without the loss of generality, $P_{i,k,n}^{FUE} \leq P_{m,n}^{MUE}$ is for the PD-NOMA separation.

2.2 Receiver Model

A MUD algorithm based on SIC is proposed, as shown in Algorithm 1. The HG-SIC receiver jointly considers the FUEs specific signal power and diversity order. Using $H_p(i,k,n) = \left[h_{i,k,n}^{FUE}\right]_{F \times K \times N} \odot G_{FUE}^{\left[I,N,K\right]}$, the equivalent channel matrix, the FUEs are detected based on the characteristics of high signal power and low diversity order. Hence, the signal for the j^{th} FUE and m^{th} MUE, where $j \neq k$ is

Algorithm 4 H-MMSE-SIC receiver scheme

```
Input Variables: $y$ (received signal), $H_p$, $\sigma_z^2$
Output Variable: $\hat{x}$ is the estimated symbol
while $i \leq F$ do
    for $k = 1, 2, ..., K$ do
    Estimate the MMSE transformation weight matrix
    $W_{MMSE}(i,k) = \min_{W_{MMSE}(i,k)} E\big[ \parallel x_{i,k} - W_y(i,k) \parallel^2 \big]$
                $= \big(H_p(i,k)^{H_p(i,k)} H_p(i,k) + \sigma_z^2 I\big)^{-1} H_p(i,k)^{H_p(i,k)}$
    where $I$ is an identity matrix
        for $l = 1, 2, ..., K - \big(k - 1\big)$ do
            Compute $SINR_{i,l}$ for $k^{th}$ FUE on the $i^{th}$ FBS
            $SINR_{i,l} = \frac{|W_{MMSE}(i,l) H_p(i,l)|^2}{\sum_{l \neq k}^{K} |W_{MMSE}(i,l) H_p(i,l)| + \sigma_z^2 \| W_{MMSE}(i,l) \|^2}$
            Compute the diversity order $N_{div}(i,l)$ for each $l^{th}$ FUE on
            the $i^{th}$ FBS
            $N_{div}(i,l)\big(i\big) = \arg\min A\big(G_{FUE}^{\big[I,K,:\big]}\big)$
        end for
        Perform post-ordering of $SINR_{i,l}$ and lowest $N_{div}(i,l)$
        $\delta_{i,k}^{max} = \delta\big(max(SINR_{i,l}), min(N_{div}(i,l))\big)$
        Estimate $\hat{x}_{i,k}$ of the $k^{th}$ FUE on the $i^{th}$ FBS, which satisfies $\delta_k^{max}$
                $\hat{x}_{i,k} = W_{MMSE}(i,k) * y_{i,k}$
                $\hat{y}_{i,k} = y_{i,k} - H_p(i,k)\hat{x}_{i,k}$
         Remove the $k^{th}$ FUE and perform signal reconstruction
    end for
end while
```

treated as interference. Consequently, the sum-rate, $R_{i,k,n}^{FUE}$, at the i^{th} FBS for the k^{th} FUE on the n^{th} SRE is given by

$$R_{i,k,n}^{FUE} = \log_2 \left(1 + \Gamma_{i,k,n}^{FBS}\right). \tag{B.3}$$

Based on (B.3), the achievable sum-rate capacity at the FBS, C^F, of the considered network can be formulated as

$$C^F = \sum_{i=1}^{F} \sum_{k=1}^{K} \sum_{n=1}^{N} R_{i,k,n}^{FUE}. \tag{B.4}$$

2.3 Energy-Efficiency Problem Formulation

The spectrum resource assignment is computed by the FBS for each FUEs under the following constraints:

- Maximum power constraint: The total allocated power of the k^{th} users on the n^{th} SREs of the i^{th} FBS cannot exceed P_{max}

  $$\sum_{n=1}^{N} \alpha_{i,k,n} P_{i,k,n}^{FUE} \leq P_{max}, \tag{B.5}$$

  where P_{max} is the maximum transmit power of each FBS.

- Rate requirement QoS: The rate QoS is guaranteed for each k^{th} FUE to ensure its maximum performance and should exceed R_k^{min}. Thus, the constraint is given as

  $$\sum_{i=1}^{F} \sum_{k=1}^{K} \sum_{n=1}^{N} R_{i,k,n}^{FUE} \geq R_k^{min}. \tag{B.6}$$

- Cross-tier interference management constraint: The interference temperature constraint is considered to limit the cross-tier interference experienced by a MBS from femtocells to macrocell. The interference temperature constraint is given by

  $$\sum_{i=1}^{F} \sum_{k=1}^{K} \alpha_{i,k,n} P_{i,k,n}^{FUE} h_{i,k,n}^{MF} \leq I_n^{th} \quad , \forall n, \tag{B.7}$$

  where I_n^{th} is the maximum tolerable cross-tier interference temperature on the n^{th} SRE in the macrocell.

- FUE spectrum resource constraint: The SRE cannot be assigned to no more than L FUEs at a time in a femtocell. Therefore, the SRE assignment constraint is given by

  $$\sum_{i=1}^{F} \alpha_{i,k,n} \leq L. \tag{B.8}$$

Our purpose is to maximize the overall EE capacity by considering the joint combinatorial decision policies of power allocation $\mathrm{P} = \{P_{i,k,n}^{FUE}|\forall i,k,n\}$ and SRE assignment $\mathrm{X} = \{\alpha_{i,k,n}|\forall i,k,n\}$. The system EE is defined as a ratio of the system sum-rate to the total power utilization expressed as

$$EE\big(\mathrm{P,X}\big) = \frac{C^F\big(\mathrm{P,X}\big)}{P_T^F\big(\mathrm{P,X}\big) + P_c^F}, \tag{B.9}$$

where total power consumption at the FBS is modelled as $P_T^F\big(\mathrm{P,X}\big) = \sum_{i=1}^{F}\sum_{k=1}^{K}\sum_{n=1}^{N}\zeta\alpha_{i,k,n}P_{i,k,n}^{FUE}$, and $\zeta \in \big(0,1\big)$ is the energy conversion inefficiency, P_c^F is the circuit power consumption. Thus, the corresponding optimization problem is expressed as

$$\max_{\{\mathrm{P,X}\}} EE\big(\mathrm{P,X}\big) \tag{B.10}$$

s.t

$$C1\colon \sum_{n=1}^{N}\alpha_{i,k,n}P_{i,k,n}^{FUE} \leq P_{max} \quad ,\forall i,k, \tag{B.11a}$$

$$C2\colon P_{i,k,n}^{FUE} \geq 0 \quad ,\forall i,k,n \tag{B.11b}$$

$$C3\colon \sum_{i=1}^{F}\sum_{k=1}^{K}\sum_{n=1}^{N}R_{i,k,n}^{FUE} \geq R_k^{min}, \tag{B.11c}$$

$$C4\colon \sum_{i=1}^{F}\sum_{k=1}^{K}\alpha_{i,k,n}P_{i,k,n}^{FUE}h_{i,k,n}^{MF} \leq I_n^{th} \quad ,\forall n, \tag{B.11d}$$

$$C5\colon \sum_{i=1}^{F}\alpha_{i,k,n} \leq L \quad ,\forall i,k,n, \tag{B.11e}$$

$$C6\colon \alpha_{i,k,n} \in \big\{0,1\big\} \quad ,\forall i,k,n \tag{B.11f}$$

C2 shows that the allocated power is non-negative. C6 denotes the allocation of the number of users allocated per SRE.

3 Proposed Energy-Efficient resource allocation schemes

The objective function in (B.10) is a mixed integer and non-convex optimization problem. Mathematical transformation and convex optimization techniques are carried out to transform the original problem into a concave one. The langrange dual method is then applied to solve the resultant transformed problem. The non-linearity and non-convex characteristics of (B.10) is dual decomposed as outer and inner loop problems respectively [37]. This effectively transforms it into a concave optimization problem. The transformation is described below.

3.1 Outer Loop Problem

To deal with the non-convexity caused by the consideration of intricate interference distributions, and to make the optimization problem easier to solve, the binary indicator $\alpha_{i,k,n}$ is relaxed to a continuous variable i.e. $\hat{\alpha}_{i,k,n} \in \big[0,1\big]$, where $\hat{\alpha}_{i,k,n}$ is deemed as a time-sharing QoS for the n^{th} SRE. For notational brevity, the actual power allocated to k^{th} FUE in the i^{th} FBS on the n^{th} SRE is denoted as $\hat{p}^{FUE}_{i,k,n} = \hat{\alpha}_{i,k,n} p^{FUE}_{i,k,n}$. Without the loss of generality, constraints (C1) and (C2) can be rewritten as

$$\hat{C}1\colon \sum_{n=1}^{N} \hat{p}^{FUE}_{i,k,n} \leq P_{max}, \quad \hat{C}2\colon \hat{p}^{FUE}_{i,k,n} \geq 0. \tag{B.12}$$

An approximation of transmission rate $R^{FBS}_{i,k,n}$ (B.3) is used as [13] and is given by

$$\hat{R}^{FUE}_{i,k,n} = B_{sc} \log_2 \Big(1 + \hat{\Gamma}^{FBS}_{i,k,n}\Big), \tag{B.13}$$

where

$$\hat{\Gamma}^{FBS}_{i,k,n} = \frac{\hat{P}^{FUE}_{i,k,n} h^{FUE}_{i,k,n}}{\sum\limits_{j=1, j\neq k}^{K} \hat{P}^{FUE}_{i,j,n} h^{FUE}_{i,j,n} + \hat{\alpha}_{i,k,n} I^{M}_{n} + \hat{\alpha}_{i,k,n} \sigma^2_z}. \tag{B.14}$$

3.2 Inner Loop Problem

To deal with the non-linearity which comes from the fractional objective function defined in (B.10), the Dinkelbach method introduced in [38] is used. Introducing U as a parameter to scale the total power consumption for the femtocells, we find the solution for a given value of U to the problem, expressed as $\alpha_{i,k,n}$ and $P^{FUE}_{i,k,n}$. Hence, we define

$$\begin{aligned} f\big(U\big) = \max_{\{\mathrm{P,X}\}} & \sum_{i=1}^{F} \sum_{k=1}^{K} \sum_{n=1}^{N} R^{FUE}_{i,k,n} - \\ & U\big(\sum_{i=1}^{F} \sum_{k=1}^{K} \sum_{n=1}^{N} \zeta \alpha_{i,k,n} P^{FUE}_{i,k,n} + P^{F}_{c}\big). \end{aligned} \tag{B.15}$$

It is observed that $f\big(U\big)$ is negative when U approaches infinity, while $f\big(U\big)$ is positive when U approaches minus infinity. Therefore, $f\big(U\big)$ is convex with respect to U. Define $\mathbf{X}^* = \big[\hat{\alpha}^*_{i,k,n}\big]$ and $\mathbf{P}^* = \big[\hat{P}^{FUE,*}_{i,k,n}\big]$ as the optimal SRE assignment and power allocation policy for problem (B.15) [37]. The maximum EE U^* can be achieved if and only if

$$\begin{aligned} f\big(U^*\big) &= \max_{\{\mathrm{P,X}\}} C^F(\mathbf{P}^*, \mathbf{X}^*) - U^*(P^F_T(\mathbf{P}^*, \mathbf{X}^*) + P^F_C) \\ &= C^F(\mathbf{P}^*, \mathbf{X}^*) - U^*\big(P^F_T(\mathbf{P}^*, \mathbf{X}^*)\big) \\ &= 0, \end{aligned} \tag{B.16}$$

where the maximum EE U^* for the femtocells is given by

$$U^* = \frac{C^F\big(\{\alpha^{\star}_{i,k,n}\}, \{P^{FUE,\star}_{i,k,n}\}\big)}{P^F_T\big(\{\alpha^{\star}_{i,k,n}\}, \{P^{FUE,\star}_{i,k,n}\}\big) + P^F_c}. \tag{B.17}$$

The original objective function in (B.10) can be transformed into an equivalent subtractive form as

$$\max_{\{\mathrm{P,X}\}} \sum_{i=1}^{F}\sum_{k=1}^{K}\sum_{n=1}^{N} \alpha_{i,k,n}\hat{R}^{FUE}_{i,k,n} - U\big(\sum_{i=1}^{F}\sum_{k=1}^{K}\sum_{n=1}^{N} \zeta\hat{P}^{F}_{i,k,n} + P^F_c\big) \tag{B.18}$$

s.t

$$\hat{C}1\colon \sum_{n=1}^{N} \hat{P}^{FUE}_{i,k,n} \leq P_{max} \quad ,\forall i,k, \tag{B.19a}$$

$$\hat{C}2\colon \hat{P}^{FUE}_{i,k,n} \geq 0 \quad ,\forall i,k,n \tag{B.19b}$$

$$\hat{C}3\colon \sum_{i=1}^{F}\sum_{k=1}^{K}\sum_{n=1}^{N} \alpha_{i,k,n}\hat{R}^{FUE}_{i,k,n} \geq R^{min}_k, \tag{B.19c}$$

$$\hat{C}4\colon \sum_{i=1}^{F}\sum_{k=1}^{K} \hat{P}^{FUE}_{i,k,n} h^{MF}_{i,k,n} \leq I^{th}_n \quad ,\forall n, \tag{B.19d}$$

$$\hat{C}5\colon \sum_{i=1}^{F} \hat{\alpha}_{i,k,n} \leq L \quad ,\forall k,n \tag{B.19e}$$

$$\hat{C}6\colon \hat{\alpha}_{i,k,n} \in \big[0,1\big] \quad ,\forall i,k,n. \tag{B.19f}$$

For a given value U, the Hessian matrix of the objective function in (B.18) is concave, subject to the convex inequality constraints of (B.19a) - (B.19f). Thus, with the feasible set of objective function being convex, the Hessian matrix can be proved to be negative semi-definite with respect to $\hat{p}^{FUE}_{i,k,n}$ and $\hat{\alpha}_{i,k,n}$ [39]. Moreover, a distinctive optimal solution to the transformed optimization problem in (B.18) can be procured in polynomial time [15, 39].

3.3 Lagrange dual method

In this subsection, the transformed optimization problem in (B.18) is jointly concave with respect to $\{\mathrm{P, X}\}$ and is solved by adopting the Lagrangian dual method [15, 39]. The Lagrangian function used to solve (B.18) can be given by

$$\begin{aligned}
&L\big(\{\mathrm{P}\},\{\mathrm{X}\},U,\Omega\big)\\
&=\sum_{i=1}^{F}\sum_{k=1}^{K}\sum_{n=1}^{N}\alpha_{i,k,n}\hat{R}_{i,k,n}^{FUE}-U\bigg(\sum_{i=1}^{F}\sum_{k=1}^{K}\sum_{n=1}^{N}\hat{P}_{i,k,n}^{FUE}+P_c^F\bigg)\\
&+\sum_{i=1}^{F}\sum_{k=1}^{K}\lambda_{i,k}\bigg(P_{max}-\sum_{n=1}^{N}\hat{P}_{i,k,n}^{FUE}\bigg)\\
&+\sum_{n=1}^{N}\mu_n\bigg(I_n^{th}-\sum_{i=1}^{F}\sum_{k=1}^{K}\beta_{i,m,n}P_{m,n}^{MUE}h_{i,k,n}^{MF}\bigg)\\
&+\sum_{i=1}^{F}\sum_{n=1}^{N}\eta_{i,n}\bigg(L-\sum_{k=1}^{K}\hat{\alpha}_{i,k,n}\bigg)\\
&+\sum_{i=1}^{F}\sum_{k=1}^{K}\nu_{i,k}\bigg(\sum_{n=1}^{N}\alpha_{i,k,n}\hat{R}_{i,k,n}^{FUE}-R_k^{min}\bigg),
\end{aligned}\tag{B.20}$$

where $\Omega=\big(\lambda\geq 0,\mu\geq 0,\eta\geq 0,\nu\geq 0\big)$ are the Lagrange multipliers for transformed constraints $\hat{C}1,\hat{C}2,\hat{C}3,\hat{C}4$, and $\hat{C}5$ in (B.18), respectively. The boundary constraints $\hat{C}2$ and $\hat{C}6$ in (B.18) will be satisfied in the Karush-Kuhn-Tucker (KKT) conditions [38]. Thus, the Lagrangian dual function is defined as

$$g\big(U,\lambda,\mu,\eta,\nu\big)=\max_{\{\mathrm{P,X}\}}L\big(\{\mathrm{P}\},\{\mathrm{X}\},U,\Omega\big).\tag{B.21}$$

Given U, the dual problem is expressed as

$$\begin{aligned}
&\min_{(\lambda,\mu,\eta,\nu)}\ g\big(U,\lambda,\mu,\eta,\nu\big)\\
&\text{s.t}\quad \lambda\geq\ 0,\mu\geq\ 0,\eta\geq\ 0,\nu\geq\ 0.
\end{aligned}\tag{B.22}$$

We decompose the Lagrangian dual function of (B.20) into a master problem and $K\times N$ subproblems. The dual problem can be solved iteratively with each FBS solving the corresponding local subproblem in each iteration using local information [15, 39]. Accordingly, the Lagrangian function in (B.20) is rewritten as

$$\begin{aligned}
&L\big(\{\mathrm{P}\},\{\mathrm{X}\},U,\Omega\big)\\
&=\sum_{i=1}^{F}\sum_{n=1}^{N}L_{i,n}\big(\{\mathrm{P}\},\{\mathrm{X}\},\Omega\big)+\sum_{i=1}^{F}\sum_{k=1}^{K}\lambda_{i,k}P_{max}-UP_c^{FBS}\\
&+\sum_{n=1}^{N}\mu_nI_n^{th}+\sum_{i=1}^{F}\sum_{n=1}^{N}\eta_{i,n}\big(L\big)+\sum_{i=1}^{F}\sum_{k=1}^{K}\nu_{i,k}R_k^{min},
\end{aligned}\tag{B.23}$$

where

$$\begin{aligned} &L_{i,n}\big(\{\mathrm{P}\},\{\mathrm{X}\},U,\Omega\big)\\ &=\sum_{k=1}^{K}\alpha_{i,k,n}\hat{R}_{i,k,n}^{FUE}+\sum_{k=1}^{K}U\zeta\hat{P}_{i,k,n}^{FUE}-\lambda_{i,k}\hat{p}_{i,k,n}^{FUE}-\sum_{k=1}^{K}\eta_{i,n}\alpha_{i,k,n}\\ &-\mu_n\sum_{k=1}^{K}\beta_{i,m,n}P_{m,n}^{MUE}h_{i,k,n}^{MF}+\sum_{k=1}^{K}\nu_{i,k}\alpha_{i,k,n}\hat{R}_{i,k,n}^{FUE}. \end{aligned} \tag{B.24}$$

Using standard optimization techniques and the KKT conditions, we obtain the optimal transmit power solution given $\big(\Omega=\lambda\geq\ 0,\mu\geq\ 0,\eta\geq\ 0,\nu\geq\ 0\big)$ as (B.25),

$$P_{i,k,n}^{FUE}=\frac{\hat{p}_{i,k,n}^{FUE,*}}{\alpha_{i,k,n}}=\left[\frac{B_{sc}\big(1+\nu_{i,k}\big)}{\sum\limits_{j=1}^{k-1}B_{sc}\big(1+\nu_{i,j}\big)\big(\Gamma_{i,j,n}^{FBS}\big)+\ln\big(2\big)\big(U\zeta+\lambda_{i,k}+\mu_n h_{i,k,n}^{MF}\big)}\right]^{+}. \tag{B.25}$$

The optimal solution typically follows water filling algorithm, where $\big[x\big]^{+}=\max\big[0,1\big]$. From (B.24), the partial derivative of the Lagrangian is expressed as

$$\frac{\partial L_{i,n}(...)}{\partial\alpha_{i,k,n}}=\Delta_{i,k,n}-\eta_{i,n}\begin{cases}<0, & \alpha_{i,k,n}=0.\\ =1, & 0\leq\alpha_{i,k,n}\leq 1.\\ >0, & \alpha_{i,k,n}=1,\end{cases} \tag{B.26}$$

where

$$\begin{aligned} &\Delta_{i,k,n}\\ &=\big(1+\nu_{k,f}\big)B_{sc}log_2\left(1+\frac{P_{i,k,n}^{FUE}|h_{i,k,n}^{F}|^2}{\sum\limits_{j=k+1}^{K}P_{i,j,n}^{FUE}|h_{i,k,n}^{F}|^2+I_n^M+\sigma_z^2}\right)\\ &-\big(1+\nu_{k,f}\big)\frac{B_{sc}}{N}\left(1+\frac{P_{i,k,n}^{FUE}|h_{i,k,n}^{F}|^2}{\sum\limits_{j=k+1}^{K}P_{i,j,n}^{FUE}|h_{i,k,n}^{F}|^2+I_n^M+\sigma_z^2}\right)\\ &-\big(t+\lambda_{i,k}\big)p_{i,k,n}^{FUE}-\mu_n P_{i,k,n}^{FUE}h_{i,k,n}^{MF}. \end{aligned} \tag{B.27}$$

Moreover, by applying the concept of marginal benefits, the policy assignment of the n^{th} SRE to the optimal user is given as

$$\alpha_{i,k,n}^{*}=1\mid k^{*}=\max_{k}\Delta_{i,k,n}. \tag{B.28}$$

Finally, based on the sub-gradient method, the dual variables are updated according to the following expressions:

$$\lambda_{i,k}\big(\ell+1\big)=\Bigg[\lambda_{i,k}\big(\ell\big)-\phi_1\big(\ell\big)\Bigg(P_{max}-\hat{P}^{FUE}_{i,k,n}\Bigg)\Bigg]^+, \tag{B.29a}$$

$$\nu_{i,k}\big(\ell+1\big)=\Bigg[\nu_{i,k}\big(\ell\big)-\phi_2\big(\ell\big)\Bigg(\hat{R}^{FUE}_{i,k,n}-R^{min}_{k}\Bigg)\Bigg]^+, \tag{B.29b}$$

$$\mu_{n}\big(\ell+1\big)=\Bigg[\mu_{n}\big(\ell\big)-\phi_3\big(\ell\big)\Bigg(I^{th}_{n}-\hat{P}^{FUE}_{i,k,n}h_{i,k,n}\Bigg)\Bigg]^+, \tag{B.29c}$$

where ℓ denotes the number of iteration. ϕ_1, ϕ_2 and ϕ_3 are the corresponding small step sizes. When the step sizes are sufficiently small, the Lagrangian multipliers converges to equilibrium points. The complete EE spectrum resource allocation for the HetNets is evaluated as in Algorithm 2.

Algorithm 5 EE Spectrum Resource Allocation Scheme

```
Initialize the maximum number of iterations $I_{max}$ and the maximum tolerance $\varepsilon$, the EE $U$, the iteration index $i=0$, and $p^{FUE}_{i,k,n}$ with equal power allocation across all SRE
while | $\frac{C^F(\alpha(i-1)),P^F(i-1))}{P^F_T(\alpha(i-1)),P^F(i-1))+P^F_c}$ |≥ $\varepsilon$ or $i \leq I_{max}$ do
    Initialize the maximum number of iterations $L_{max}$ and set inner loop iteration $l=0$, and initialize Lagrange multipliers $(\lambda,\mu,\nu)$.
    repeat
        for $i=1$ do
            for $k=1$ do
                for $n=1$ do
                    Given the EE $U$ update $P^{FUE}_{i,k,n}$ according to (B.25)
                    Calculate $\Delta_{i,k,n}$ according to (B.27)
                    FBS update $\alpha^*_{i,k,n}$ according to (B.28)
                    FBS update $\lambda$ and $\nu$ according to (B.29a)-(B.29b)
                end for
            end for
        end for
        MBS updates $\mu$ according to (B.29c), and broadcasts the updated     value to all FBSs via backhaul.
    until  Converges or $l = L_{max}$
    Set $i=i+1$ and $U(i)=\frac{C^F(\alpha(i-1)),P^F(i-1))}{P^F_T(\alpha(i-1)),P^F(i-1))+P^F_c}$
end while
```

3.4 Complexity analysis

Following the work of [15], the computational complexity of the proposed EE spectrum resource allocation algorithm, Algorithm 2, can be derived. In Algorithm 2, during each iteration, the following steps are executed to yield a low computational complexity order.

- In line 8-10 : the calculation of (B.25), (B.27), and (B.28) for every FUE on each SRE in every femtocell requires FKN operations, independently.
- In line 11 : $O(KF)$ calculations required for updating the dual variables λ and ν according to (B.29a)-(B.29b), respectively. Assuming, that the sub-gradient method is used to updated the dual variables, Algorithm 2 will require ϑ_f iterations to converge.
- In line 15 : the parameter μ is computed by the MBS for each SRE and its complexity follows the order $O(N)$ per iteration, therefore, ϑ_f is a polynomial function of (K^2F^2N).

Therefore, the total computational complexity is given by $O((KFN)^2\vartheta_f))$. However, compared with the exhaustive search for SRE allocation, which has a worst-case complexity of $O(KF^N)$ [15]. Moreover, ϑ_f can be made small enough if the initial values of λ, ν and μ are well chosen, together with suitable values of iteration step sizes.

Table B.1: Simulation parameters

Parameters	**Symbol**	**Values**
Carrier frequency		2GHz
Path loss exponent		2.8
SRE Bandwidth	B_{sc}	10 MHz
Number of Femtocells	F	50
Number of FUES per femtocell	K	6
Number of SREs	N	24
Assigned FUEs per SRE on each FBS	L	4
FUE static power	P_c^F	21 dBm
Processing noise	σ_z^2	-125 dBm
Minimum transmission rate	R_k^{min}	5 Mbps/Hz
Maximum transmission power	P_{max}	20 dBm
Interference threshold	I_n^{th}	$10^{-5.5}$ W
Power amplifier inefficiency	ζ	0.9

4 Performance Evaluation

The performance of the system is evaluated based on Algorithm 2, and in terms of EE, sum-rate and complexity. We consider a random distribution of the FUEs in a small-cells coverage area of 50 meters radius. Subsequently, the small-cells are randomly overlaid within a macrocell of 500 meters radius. The minimum distance between the MBS and small-cells is 40 meters. Without loss of generality, we assume that the channel model accounts for small-scale Rayleigh fading, large-scale path loss, and shadowing (log-normally distribution). The summary of the simulation parameters is given in Table 1. The algorithms for OFDMA and PD-NOMA are similar to the ones in [15] and [40] respectively, albeit the introduction of EE resource allocation proposed in this work.

Figure B.3 shows the EE versus the number of iterations performance of the proposed scheme against OFDMA and PD-NOMA. We set the number of FBS to 30. From the results, we observe that all schemes converge within a small number of iterations. However, the proposed scheme converges to an optimal solution with higher EE compared to PD-NOMA and OFDMA. Thus, the proposed HG-NOMA scheme is suitable for practical implementation and improving performance.

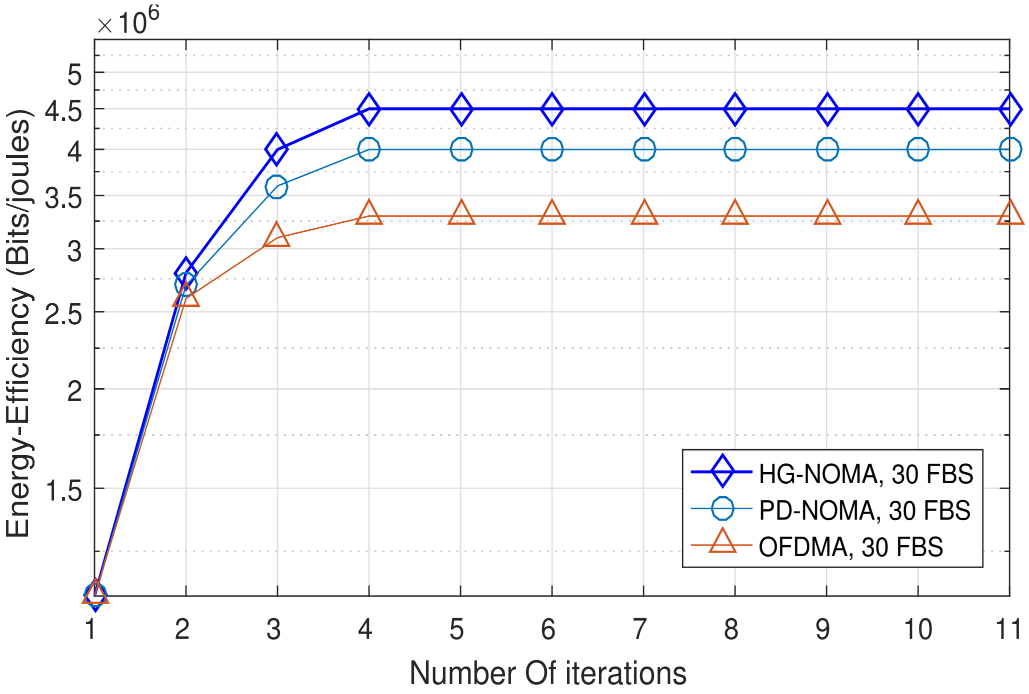


Fig. B.3: Energy-Efficiency vs number of Iterations

Figure B.4 shows the overall system EE versus the number of FBS performance. The results show that the system EE increases with the increase in the number of FBS. Furthermore, we observe that the HG-NOMA outperforms both the PD-NOMA and OFDMA, due to the efficient use of the spectrum resources and the assumption that the total number of FUEs allowed for spectrum access is always larger than the number of SRE to be scheduled.

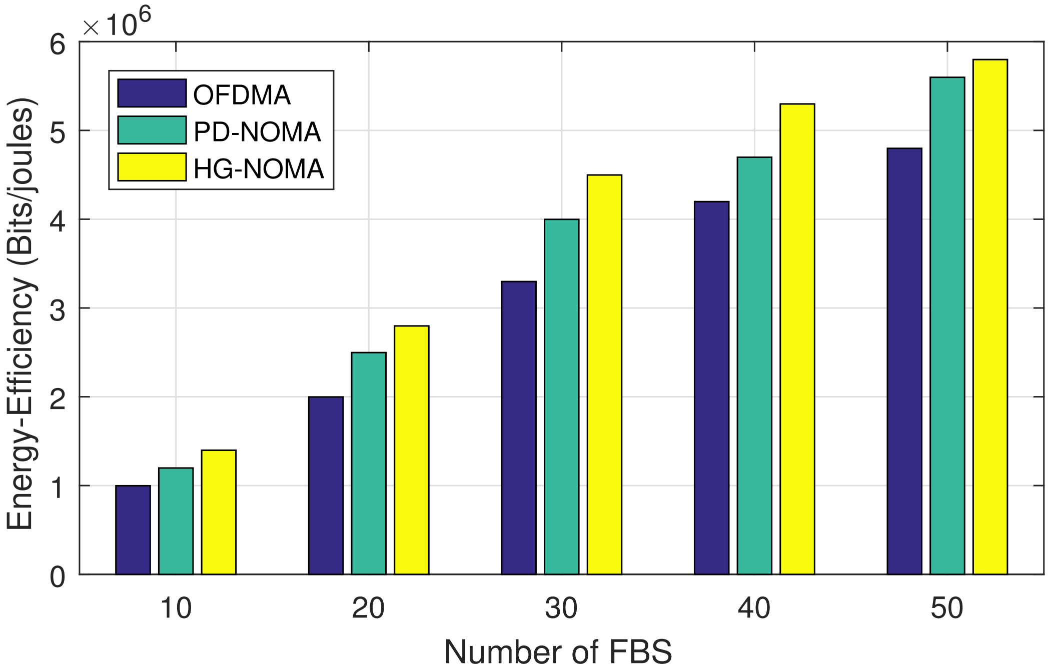


Fig. B.4: Energy-Efficiency vs number of FBS

Figure B.5 shows the system achievable sum-rate capacity of the small-cells versus SNR performance. The results indicate that the sum-rate increases monotonically with the SNR. It is also observed that HG-NOMA achieves a higher sum-rate compared to the PD-NOMA due to the introduced patterns between the potential multiplexed users. Besides, we also notice that the OFDMA system achieves a lower sum rate compared to the both NOMA techniques since OMA leads to underutilized spectrum resource.

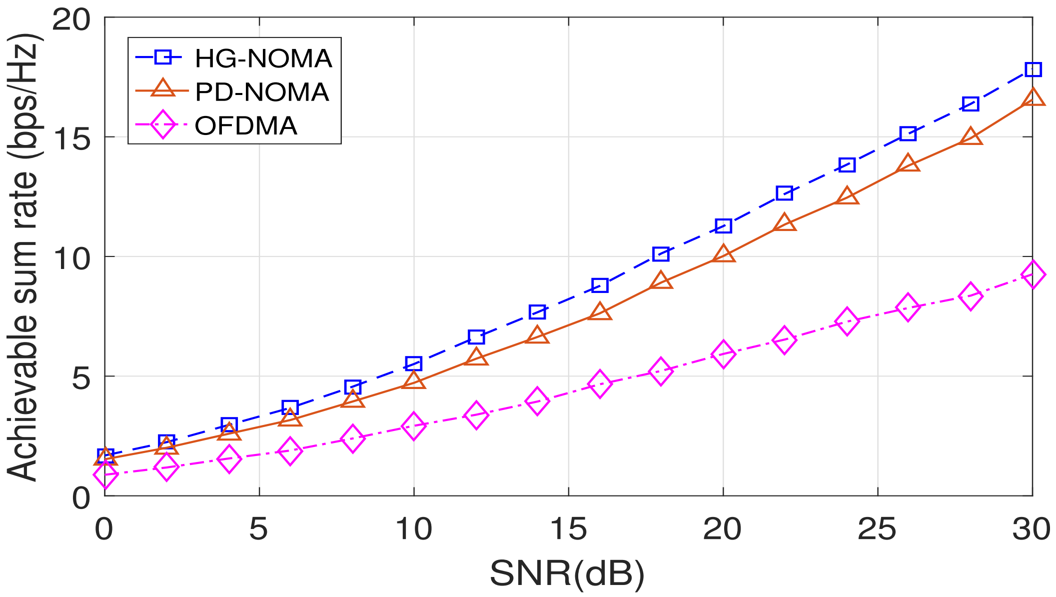


Fig. B.5: Average sum-rate capacity vs SNR (dB)

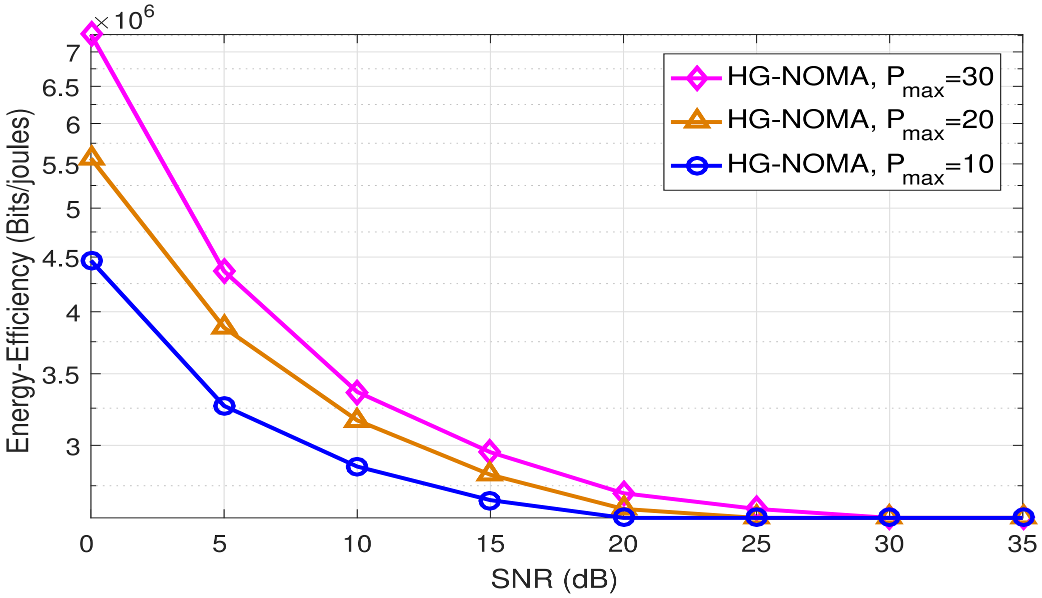


Fig. B.6: Energy-Efficiency vs SNR (dB)

Figure B.6 depicts the performance of the proposed system EE with respect to the SNR. It can be observed that the EE decreases with increasing SNR. This is because the maximum transmit power consumption increases faster than the overall system EE. Furthermore, the EE of the proposed HG-NOMA scheme is approximate to the comprehensive search solution, especially at higher transmit power.

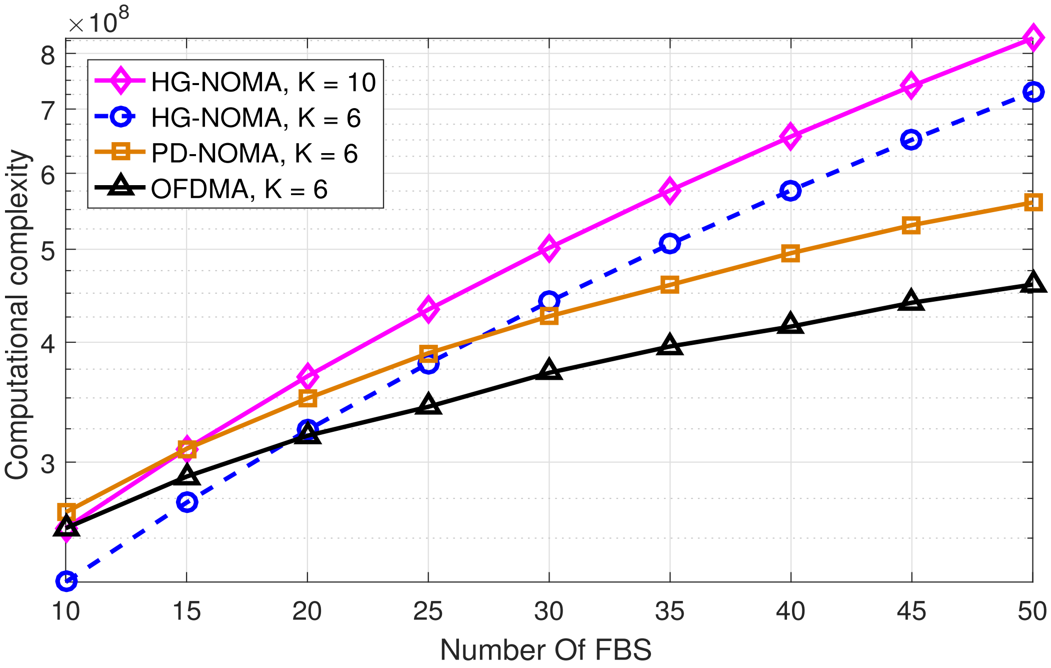


Fig. B.7: Computational complexity analysis vs Number of FBS

Figure B.7 shows the computational complexity versus the number of FBS for the proposed algorithm. Needless to say that the computational complexity increases with an increase in number of any of FBS, FUEs, and SRE. Even though the HG-NOMA provides benefits of EE and improved sum-rate, as expected, it comes at the cost of added complexity. This is because of the introduced diversity patterns, and HG-NOMA overloading.

5 Conclusion

In this paper, a HG-NOMA resource allocation scheme that integrates PD-NOMA and PDMA have been designed for interference management of 5G small-cells networks. Power resources were chosen as the fundamental multiplexing domain between the MUEs and FUEs, and code resources as the key multiplexing domain in the sparse pattern mapping. We formulated an EE maximization problem by considering joint spectrum allocation and power control, to maximize massive connectivity of user access, EE and sum-rate capacity of small cells, which in turn maximizes the overall system performance. Additionally, we have proposed a low-complexity HG-SIC receiver that combines both the power-levels and diversity patterns gain to realize MUD. The simulation results demonstrate superior performance of the proposed scheme in terms of sum-rate capacity and EE. This confirms the practical implementation of proposed scheme for the 5G small-cells network system. Future research trends could consider the application of simultaneous wireless and power transfer for HG-NOMA and extend the work to Multiple-input Multiple-Output (MIMO) systems.

References

[1] Y. Liu, Z. Qin, M. Elkashlan, A. Nallanathan, and J. A. McCann, “Non-orthogonal multiple access in large-scale heterogeneous networks,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 12, pp. 2667–2680, 2017.

[2] Y. Cai, Z. Qin, F. Cui, G. Y. Li, and J. A. McCann, “Modulation and multiple access for 5g networks,” *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 629–646, 2018.

[3] L. Ping, L. Liu, K. Wu, and W. K. Leung, “Interleave division multiple-access,” *IEEE Transactions on Wireless Communications*, vol. 5, no. 4, pp. 938–947, 2006.

[4] S. R. Islam, N. Avazov, O. A. Dobre, and K.-S. Kwak, “Power-domain non-orthogonal multiple access (noma) in 5g systems: Potentials and challenges,” *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 721–742, 2017.

[5] H. Nikopour and H. Baligh, “Sparse code multiple access,” in *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*. IEEE, 2013, pp. 332–336.

[6] A.-I. Mohammed, M. A. Imran, and R. Tafazolli, “Low density spreading for next generation multicarrier cellular systems,” in *Future Communication Networks (ICFCN), 2012 International Conference on*. IEEE, 2012, pp. 52–57.

[7] Z. Yuan, G. Yu, W. Li, Y. Yuan, X. Wang, and J. Xu, “Multi-user shared access for internet of things,” in *Vehicular Technology Conference (VTC Spring), 2016 IEEE 83rd*. IEEE, 2016, pp. 1–5.

[8] S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, and K. Niu, “Pattern division multiple access—a novel nonorthogonal multiple access for fifth-generation radio networks,” *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3185–3196, 2017.

[9] B. Wang, K. Wang, Z. Lu, T. Xie, and J. Quan, “Comparison study of non-orthogonal multiple access schemes for 5g,” in *Broadband Multimedia Systems and Broadcasting (BMSB), 2015 IEEE International Symposium on*. IEEE, 2015, pp. 1–5.

[10] W. Shin, M. Vaezi, B. Lee, D. J. Love, J. Lee, and H. V. Poor, “Non-orthogonal multiple access in multi-cell networks: Theory, performance, and practical challenges,” *IEEE Communications Magazine*, vol. 55, no. 10, pp. 176–183, 2017.

[11] H. Zhang, C. Jiang, R. Q. Hu, and Y. Qian, “Self-organization in disaster-resilient heterogeneous small cell networks,” *IEEE Network*, vol. 30, no. 2, pp. 116–121, 2016.

[12] H. Zhang, B. Wang, C. Jiang, K. Long, A. Nallanathan, V. C. Leung, and H. V. Poor, “Energy efficient dynamic resource optimization in noma system,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 5671–5683, 2018.

[13] M. S. Al-Janabi, C. C. Tsimenidis, B. S. Sharif, and S. Y. Le Goff, “Optimized resource allocation strategy for a single-cell downlink amc-ofdma systems,” in *Electronics, Circuits, and Systems (ICECS), 2013 IEEE 20th International Conference on*. IEEE, 2013, pp. 209–212.

[14] S. Ismail, C. K. Ng, and N. K. Noordin, “Fairness resource allocation for downlink ofdma systems,” in *Communications (MICC), 2009 IEEE 9th Malaysia International Conference on*. IEEE, 2009, pp. 575–579.

[15] H. Zhang, C. Jiang, N. C. Beaulieu, X. Chu, X. Wen, and M. Tao, “Resource allocation in spectrum-sharing ofdma femtocells with heterogeneous services.” *IEEE Trans. Communications*, vol. 62, no. 7, pp. 2366–2377, 2014.

[16] T. Lee, H. Kim, J. Park, and J. Shin, “An efficient resource allocation in ofdma femtocells networks,” in *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*. IEEE, 2010, pp. 1–5.

[17] A. Hatoum, R. Langar, N. Aitsaadi, R. Boutaba, and G. Pujolle, “Cluster-based resource management in ofdma femtocell networks with qos guarantees,” *IEEE Transactions on Vehicular Technology*, vol. 63, no. 5, pp. 2378–2391, 2014.

[18] A. Abdelnasser and E. Hossain, “Joint subchannel and power allocation in two-tier ofdma hetnets with clustered femtocells,” in *Communications (ICC), 2013 IEEE International Conference on*. IEEE, 2013, pp. 6002–6007.

[19] M. Al-Imari, P. Xiao, and M. A. Imran, “Receiver and resource allocation optimization for uplink noma in 5g wireless networks,” in *Wireless Communication Systems (ISWCS), 2015 International Symposium on*. IEEE, 2015, pp. 151–155.

[20] S. Islam, M. Zeng, O. A. Dobre, and K.-S. Kwak, “Resource allocation for downlink noma systems: Key techniques and open issues,” *IEEE Wirel. Commun*, vol. 25, no. 2, pp. 40–47, 2018.

[21] J. Choi, “Power allocation for max-sum rate and max-min rate proportional fairness in noma.” *IEEE Communications Letters*, vol. 20, no. 10, pp. 2055–2058, 2016.

[22] L. Lei, D. Yuan, C. K. Ho, and S. Sun, “Power and channel allocation for non-orthogonal multiple access in 5g systems: Tractability and computation,” *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 8580–8594, 2016.

[23] Ö. F. Gemici, F. Kara, I. Hokelek, G. K. Kurt, and H. A. Çırpan, “Resource allocation for noma downlink systems: Genetic algorithm approach,” in *Telecommunications and Signal Processing (TSP), 2017 40th International Conference on*. IEEE, 2017, pp. 114–118.

[24] C.-H. Liu and D.-C. Liang, “Heterogeneous networks with power-domain noma: Coverage, throughput, and power allocation analysis,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 5, pp. 3524–3539, 2018.

[25] J. Zeng, B. Li, Y. Mao, X. Su, and Y. Kuang, “Research of heterogeneous networks convergence with noma,” in *Communications in China (ICCC), 2015 IEEE/CIC International Conference on*. IEEE, 2015, pp. 1–5.

[26] A. S. Marcano and H. L. Christiansen, “Impact of noma on network capacity dimensioning for 5g hetnets,” *Ieee Access*, 2018.

[27] D. Ni, L. Hao, Q. T. Tran, and X. Qian, “Power allocation for downlink noma heterogeneous networks,” *IEEE Access*, 2018.

[28] R. Ruby, S. Zhong, H. Yang, and K. Wu, “Enhanced uplink resource allocation in non-orthogonal multiple access systems,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1432–1444, 2018.

[29] G. Lv, X. Li, R. Shang, P. Xue, and Y. Jin, “Dynamic resource allocation for uplink non-orthogonal multiple access systems,” *IET Communications*, vol. 12, no. 6, pp. 649–655, 2018.

[30] J. Zhao, Y. Liu, K. K. Chai, A. Nallanathan, Y. Chen, and Z. Han, “Spectrum allocation and power control for non-orthogonal multiple access in hetnets,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 5825–5837, 2017.

[31] F. Fang, H. Zhang, J. Cheng, and V. C. Leung, “Energy-efficient resource allocation for downlink non-orthogonal multiple access network,” *IEEE Transactions on Communications*, vol. 64, no. 9, pp. 3722–3732, 2016.

[32] L. Xiang and H. Chen, “Energy-efficient and fair power allocation approach for noma in ultra-dense heterogeneous networks,” in *2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*. IEEE, 2017, pp. 89–94.

[33] M. Moltafet, N. M. Yamchi, M. R. Javan, and P. Azmi, “Comparison study between pd-noma and scma,” *IEEE Transactions on Vehicular Technology*, vol. 67, no. 2, pp. 1830–1834, 2018.

[34] M. Moltafet, R. Joda, N. Mokari, M. R. Sabagh, and M. Zorzi, “Joint access and fronthaul radio resource allocation in pd-noma based 5g networks enabling dual connectivity and comp,” *IEEE Transactions on Communications*, 2018.

[35] B. Xu, Y. Chen, J. R. Carrión, and T. Zhang, “Resource allocation in energy-cooperation enabled two-tier noma hetnets toward green 5g,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 12, pp. 2758–2770, 2017.

[36] Z. Song, Q. Ni, and X. Sun, “Spectrum and energy efficient resource allocation with qos requirements for hybrid mc-noma 5g systems,” *IEEE Access*, vol. 6, pp. 37 055–37 069, 2018.

[37] S.-F. Hung, P.-R. Li, K.-T. Feng, and Y.-T. Lin, “Joint wireless charging and hybrid power based resource allocation for lte-a wireless network,” in *Wireless Communications and Networking Conference (WCNC), 2017 IEEE*. IEEE, 2017, pp. 1–6.

[38] W. Dinkelbach, “On nonlinear fractional programming,” *Management science*, vol. 13, no. 7, pp. 492–498, 1967.

[39] Y. Xu, G. Li, J. Tang, and L. Luan, “Robust resource allocation for uplink sum rate maximization in multi-cell heterogeneous networks,” in *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, 2018, pp. 1–6.

[40] M. Moltafet, P. Azmi, N. Mokari, M. R. Javan, and A. Mokdad, “Optimal and fair energy efficient resource allocation for energy harvesting-enabled-pd-noma-based hetnets,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 2054–2067, 2018.

Part III

Conclusion

Conclusion

We conclude this thesis by summarizing the contributions and present potential future research trends that are related to our accomplished work.

The introduction detailed the evolution of mobile communication networks and 5G networks background, The research motivation, methodology, and contributions were also clearly stated in the introduction. Despite the potential benefits brought by the 5G networks to revolutionize cellular networks. The proliferation of data traffic over wireless channels have resulted to the challenges of massive connectivity of both users, and devices, higher spectral-efficiency, low-latency, and improved data-rates. Which prompts the need for advance modulation & multiple access for the 5G networks. Among the challenges brought by the development of the 5G, interference management and EE remain the major challenges, as macrocells are overlaid by the low power small-cells, to maximize the system capacity and EE. Thus, the investigation for efficient solutions to encourage improved network performance are the main objective this of research. The summary of the research contributions is comprehensively outlined in the section below, which sets the direction for the future works.

1 Summary of research contribution

In paper A, we studied the performance analysis of G-NOMA technologies for the 5G wireless networks under perfect CSI. The investigated G-NOMA schemes, included: IDMA, MUSA, PDMA, and SCMA that addressed the issues of massive connectivity and improved system-capacity. The investigation was carried in terms of the BER and achievable sum-rate capacity, and the Rayleigh flat fading wireless channel was considered. simulation results showed that IDMA outperformed the other G-NOMA schemes due to the near-optimal design of user-specific interleavers.

In paper B, an energy-efficient HG-NOMA resource allocation scheme was proposed to address the issues of interference management, improved system-capacity and energy-efficiency for a two-tier HetNets. The HG-NOMA resource allocation scheme integrated PD-NOMA and PDMA, through resource patterns allocation and power control to enable MUEs and FUEs to be unorthogonaly multiplexed on the same time-frequency SRE. EE maximization problem was formulated, considering joint spectrum allocation and power control. Furthermore, a low-complexity H-MMSE-SIC receiver was proposed that combines the power-resources & diversity pattern gain to realize MUD is proposed. Simulation results showed that, the proposed scheme has superior performance over the conventional OMA based small-cells in terms of sum-rate capacity and EE, which demonstrated the practicality of the proposed framework.

2 Possible Future Work

This section provides some insight on possible future trends that could extend to this work and other contributions related to EE and interference management for the future of mobile networks. While the models presented in this work are not extensively complex, further refinements and upgrades are still expected to provide better results, conclusions, and recommendations. Following the lines of this thesis, a major upgrade could be that of considering, simalutanous wireless information and power transfer. In light of the conclusion, with respect to the papers of this thesis on the future research items includes:

1. G-NOMA with SWIPT: While NOMA integrated with HetNets technology have the capability of increasing channel capacity and maximize energy-efficiency in ultra-dense networks, there is no evidence yet that indicates the absence of alternative dimensions within the realm of power transfer and wireless charging resource allocation capabilities, to maximize the overall systems data-rates, spectral-efficiency and EE in 5G NOMA-HetNets.
2. G-NOMA with imperfect CSI: Considering EE optimization challenge, future work could consider the effect of imperfect CSI in the SWIPT optimization. Imperfect CSI in EE optimization is a practical problem that degrades the performance of the network, and should be considered. Additionally, the channel model could also be investigated.