Resource Allocation in Non-Orthogonal Multiple Access Technologies for 5G Networks and Beyond

Simon Kariuki Chege

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

Doctor of Philosophy in Engineering
(ELECTRONIC ENGINEERING)

School of Engineering
University of KwaZulu-Natal,
Durban, South Africa

July 22, 2022

Resource Allocation in Non-Orthogonal Multiple Access Technologies for 5G Networks and Beyond

Simon Kariuki Chege

Supervisor: Prof. Tom Walingo

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

Doctor of Philosophy in Engineering
(ELECTRONIC ENGINEERING)

School of Engineering
University of KwaZulu-Natal
South Africa

July 22, 2022

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

Signed.....Date.....

Declaration 1 - Plagiarism

I, **Simon Kariuki Chege**, 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.....22/07/2022

Declaration 2 - Publication

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

1. S. Chege and T. Walingo, "Energy efficient resource allocation for uplink hybrid power domain sparse code nonorthogonal multiple access heterogeneous networks with statistical channel estimation," *Trans Emerging Tel Tech.* 2021; 32:e4185. https://doi.org/10.1002/ett.4185

2. S. Chege and T. Walingo, "Multiplexing Capacity of Hybrid PD-SCMA Heterogeneous Networks," in IEEE Transactions on Vehicular Technology, vol. 71, no. 6, pp. 6424-6438, June 2022, doi: 10.1109/TVT.2022.3162304.

3. S. Chege and T. Walingo, "MIMO based hybrid PD-SCMA Uplink Transceiver System," in *IEEE Communication Letters* (*Under Review*)

The following book chapter emanating from this work has been published:

1. S. Chege and T. Walingo, "Dual Parameter Ranking Based Resource Allocation for PD-SCMA Cognitive Radio Networks," In: Woungang I, Dhurandher SK., eds. *4th International Conference on Wireless, Intelligent and Distributed Environment for Communication (WIDECOM 2021)* Springer International Publishing; 2022; Cham: 123–138.

The following conference paper emanating from this work has been published:

1. S. Chege and T. Walingo, "Resource Allocation for Uplink SCMA NOMA in Heterogeneous Networks," *2019 IEEE AFRICON*, 2019, pp. 1-6, doi: 10.1109/AFRICON46755.2019.9133742.

Signed:.....

Dedication

To my late grandmother Njeri Chege whose sacrifices, unwavering support and prayers have seen me this far.

For sacrifices of all the unsung heroes and heroines of my adventures.

Acknowledgements

"*In the fullness of time, He makes all things beautiful!*"

First, I am eternally grateful to the Most High, for His love, strength, courage, wisdom, inspiration and patience to complete this work. Indeed, He is a wonderful God and I will always attest to His greatness.

My heartfelt appreciation goes to my supervisor, Professor Tom Walingo, for being a friend and a guide who never got weary in holding my hand during my PhD studies. His invaluable expertise and dedication led to the successful completion of my PhD studies. The friendship we did build over time is the friendship we shall keep.

I would also like to thank the University of Kwazulu-Natal, particularly the CRART center for sponsoring my conferences and other related costs during the research period. Special thanks to the academic, technical and administrative staff of the Discipline of Electrical, Electronic and Computer Engineering for their tremendous support whenever called upon. Indeed, it was a pleasure working with you. Many thanks to the AGNES Intra-Africa Mobility Grant for the award of a 2 months research mobility stay at Makerere University, Uganda and The ACADEMY Intra-Africa Project for the 10 months PhD credit-seeking mobility award at The University of Tlemcen, Algeria.

Special recognition to Moi University, Kenya for granting me the chance, and her contribution towards the successful completion of this degree.

My profound gratitude to my cheer leaders, number one being my wife Wanjiku Chege. Thank you for loving me, filling the gaps in my absence, being patient with me, and taking care of our beautiful children, Wambui Chege and Mburu Chege. To my parents, brothers and sisters, I am deeply thankful for your enigmatic love and support that kept me going during very difficult periods. To my friends, "The Braai family", thank you for your immense words of encouragement, the occasional weekend braai, and the friendship we sow. You made Durban home away from home.

To everyone who supported me in any way, and wished me well, ***"Ni ngaatho cia ngoro gitina"***.

Abstract

The increasing demand of mobile and device connectivity poses challenging requirements for 5G wireless communications, such as high energy- and spectral-efficiency and low latency. This necessitates a shift from orthogonal multiple access (OMA) to Non-Orthogonal Multiple Access (NOMA) techniques, namely, power-domain NOMA (PD-NOMA) and code-domain NOMA (CD-NOMA). The basic idea behind NOMA schemes is to co-multiplex different users on the same resource elements (time slot, OFDMA sub-carrier, or spreading code) via power domain (PD) or code domain (CD) at the transmitter while permitting controllable interference, and their successful multi-user detection (MUD) at the receiver albeit, increased computational complexity.

In this work, an analysis on the performance of the existing NOMA schemes is carried out. Furthermore, we investigate the feasibility of a proposed uplink hybrid-NOMA scheme namely power domain sparse code multiple access (PD-SCMA) that integrates PD-NOMA and CD-NOMA based sparse code multiple access (SCMA) on heterogeneous networks (HetNets). Such hybrid schemes come with resource allocation (RA) challenges namely; codebook allocation, user pairing and power allocation. Therefore, hybrid RA schemes namely: Successive Codebook Ordering Assignment (SCOA) for codebook assignment (CA), opportunistic macro cell user equipment (MUE)- small cell user equipment (SUE) pairing (OMSP) for user pairing (UP), and a QoS-aware power allocation (QAPA) for power allocation (PA) are developed for an energy efficient (EE) system. The performance of the RA schemes is analyzed alongside an analytical RA optimization algorithm. Through numerical results, the proposed schemes show significant improvements in the EE of the small cells in comparison with the prevalent schemes. Additionally, there is significant sum rate performance improvement over the conventional SCMA and PD-NOMA.

Secondly, we investigate the multiplexing capacity of the hybrid PD-SCMA scheme in HetNets. Particularly, we investigate and derive closed-form solutions for codebook capacity, MUE multiplexing and power capacity bounds. The system's performance results into low outage when the system's point of operation is within the multiplexing bounds. To alleviate the RA challenges of such

a system at the transmitter, dual parameter ranking (DPR) and alternate search method (ASM) based RA schemes are proposed. The results show significant capacity gain with DPR-RA in comparison with conventional RA schemes.

Lastly, we investigate the feasibility of integrating the hybrid PD-SCMA with multiple-input multiple-output (MIMO) technique namely, M-PD-SCMA. The attention to M-PD-SCMA resides in the need of lower number of antennas while preserving the system capacity thanks to the overload in PD-SCMA. To enhance spectral efficiency and error performance we propose spatial multiplexing at the transmitter and a low complex joint MUD scheme based on successive interference cancellation (SIC) and expectation propagation algorithm (EPA) at the receiver are proposed. Numerical results exhibit performance benchmark with PD-SCMA schemes and the proposed receiver achieves guaranteed bit error rate (BER) performance with a bounded increase in the number of transmit and receive antennas. Thus, the feasibility of an M-PD-SCMA system is validated.

Contents

List of Figures

List of Figures

List of Tables

List of Tables

List of Acronyms

1G	First Generation Communication Systems
2G	Second Generation Communication Systems
3G	Third Generation Communication Systems
4G	Fourth Generation Communication Systems
5G	Fifth Generation Communication Systems
3GGP	3rd Generation Partnership Project
AMPS	Advanced Mobile Phone System
ASM	Alternate Search Method
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
CDF	Cumulative Density Function
CDMA	Code Division Multiple Access
CIR	Channel Impulse Response
CoMP	Coordinated Multi-point
C-RANs	Cloud Radio Access Networks
CS	Compressive Sensing
CSI	Channel State Information
D2D	Device-to Device
DPR	Dual Parameter Ranking

E2E	End to End
EDGE	Enhanced Data GSM Evolution
EE	Energy Efficiency
eMBB	Enhanced Mobile Broadband
eNodeB	E-UTRAN NodeB
EPS	Evolved Packet System
EVDO	Evolution data only
FDMA	Frequency Division Multiple Access
Gbps	Giga bits per second
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communication
HD	High Definition
HetNets	Heterogeneous Networks
HSDPA	High Speed Downlink Packet access
HSPA	High-Speed Packet Access
HSUPA	High-Speed Uplink Packet Access
ICI	Inter-Cell Interference
IDMA	Interleave Division Multiple Access
IMT	International Mobile Communications
IoT	Internet of Things
LDS	Low Density Spreading
LTE-A	Long Term Evolution Advanced
M2M	Machine to Machine
MA	Multiple Access
MAC	Media Access Control
Mbps	Mega bits per second

MBS	Macro Cell Base Station
METIS	Mobile and wireless communications Enablers for the Twenty -Twenty Information Society
MIP	Mixed Integer Problem
mMTC	massive Machine Type Communication
M-RAT	Multiple Radio Access Technology
MPA	Message Passing Algorithm
MUE	Macro Cell User Equipment
MUD	Multi-User Detection
NGWNs	Next Generation Wireless Networks
NFV	Network Function Virtualization
NOMA	Non-Orthogonal Multiple Access
NMT	Nordic Mobile Telephone
OFDMA	Orthogonal Frequency Multiple Access
OMA	Orthogonal Multiple Access
PDN	Packet Data Network
PD-NOMA	Power Domain Non-Orthogonal Multiple Access
QoE	Quality of Experience
QoS	Quality of Service
RA	Resource Allocation
RAN	Radio Access Network
RE	Resource Element
RF	Radio Frequency
RTT	Radio Transmission Technology
SBS	Small cell Base Station
SC-FDMA	Single Carrier Frequency Division Multiple Access

SCMA	Sparse Code Multiple Access
SIC	Successive Interference Cancellation
SINR	Signal to Interference Noise Ratio
SNR	Signal to Noise Ratio
SMS	Short Messaging Services
SUE	Small cell User Equipment
TACS	Total Access Communication System
TCP	Transmission Control Protocol
TTI	Transmission Time Interval
UE	User Equipment
UMTS	Universal Mobile Telecommunications Service
URLLC	Ultra reliable and low latency communications
V2I	Vehicle-to-Infrastructure
V2V	Vehicle-to-Vehicle
VLC	Visible Light Communication
VoLTE	Voice over LTE

Preface

The research discussed in this thesis is carried out in the College of Agriculture, Engineering and Science of the University of Kwa-Zulu Natal, Durban, from February 2019 until April 2022 by Simon Kariuki Chege under the supervision of Professor Tom Walingo.

As the candidate's supervisor, I, Tom Walingo, agree to the submission of this dissertation.

Signed:..... Date:.....

I, Simon Kariuki Chege, hereby declare that all the material incorporated in this thesis are my own original work, except where acknowledgement is made by name or in the form of a reference. The work contained in herein has not been submitted in any form for any degree or diploma to any other institution.

Signed:..... Date:22/07/2022.....

University of KwaZulu-Natal, July 22, 2022

Part I

Introduction

1 Introduction and Background

Wireless communication networks are arguably among the rapidly growing communications technology in the last decade. Communication networks continues to re-define different aspects of work and life in the society around the world. The burgeoning of smart devices creates a platform for a wide array of end-user applications that have increased meteorically in the recent past. Amongst the extensive range of mobile enabled services is the internet, which has overwhelmingly been embraced as a decisive technology of the information age. Internet provides ubiquitous capacity of multi-modal, interactive communication in chosen time and transcending space. Today's smart devices are equipped with applications that enable users perform seamless, advanced and reliable real-time data acquisition and processing in medical, business, security, entertainment, education, research, science and engineering related fields. Such applications require particular quality of service (QoS) requirements from service providers such as; high traffic and data rate, low power consumption, ultra-low latency real time services, high mobility and reliability support, radio spectrum sharing and massively connected devices. Ultimately, mobile communication is a prerequisite rather than luxury for sustaining a reasonable quality of modern life.

The surging QoS demands for wireless communication services has necessitated a re-design in access to network resources by both the service providers and the end user. Compared to 4G LTE, the QoS demands namely; spectral efficiency has increased by factors of 5 to 15, connectivity density target expected to be ten times higher, i.e., at least $10^6/km^2$, radio latency $\leq$ 1ms and low-cost efficiency of more than 100 times are required for the support of diverse compelling services [1]. In order to meet these meticulous requirements, multi-dimensional solutions are being conceived. These include implementation of enhanced multi-radio access protocols, deployment of multiple-input multiple-output (MIMO), utilization of millimeter wave frequency bands and reinforcement of infrastructural and architectural environment (including Heterogeneous networks (HetNets) where heterogeneity of cells with different transmit power, coverage range and cost of deployment) to increase capacity.

The radio spectrum resource is finite and scarce. Therefore, resource management is pivotal in servicing user and network demands in the next generation wireless networks (NGWNs). Motivated by such observations, this work focuses on the analysis of existing NOMA technologies with an objective of developing a hybrid NOMA that blends different technologies for enhanced spectral efficiency. In particular, we investigate the feasibility of integrating power–domain NOMA and code–domain NOMA in an uplink hierarchical HetNet system. Parametric hybrid resource allocation

(RA) schemes for such a system model are developed. An alternative dual-decomposition based RA techniques are utilized to analytically assess the system's performance. At the receiver, hybrid joint multi user detectors (MUD) are proposed. Due to the potential of these technologies towards achieving optimal spectrum sharing, this work further investigates the multiplexing capacity of the hybrid NOMA. In order to improve the benefits of system throughput, capacity and diversity, this work develops and investigates the application of multiple-input multiple-output (MIMO) based hybrid NOMA transceiver system on an uplink HetNet, aimed at achieving a balance on the number of antennas and capacity/spectral efficiency.

This chapter outlines the evolution and advancements of wireless networks, conducts literature review on fifth generation (5G) networks including the architecture, challenges, technical requirements and potential challenges. Furthermore, NOMA technologies are presented. Existing NOMA RA schemes are discussed with the aim of confronting the hybrid NOMA challenges. Having outlined the 5G challenges and the potential 5G facilitators, we formulate the research problem followed by a vivid explanation of the research objectives and the methodology.

2 Evolution of wireless networks

The first wireless transmission in history was achieved in the year 1895 when an Italian inventor, Marconi used radio waves wirelessly to transmit Morse code signals a distance of 3.2KMs. Since then, engineers have been seeking to efficiently utilize radio frequency waves. In the mid-19th century, wired telephony was already popular. Due to its restricted mobility, scientists were already designing devices that did not require wired connection but rather transmit voice using radio waves. The first prototype of mobile hand held devices conducted in the mid-1970s. This was deemed as the turning point in wireless communications that led to an evolution of technologies and standards [2].

The first generation (1G) mobile network was introduced in the early 1980s. Owing to the ever-increasing demand for more connections worldwide, there was need for rapid advancements in the mobile communication standards. The 1G transmitted only voice signals at a frequency range of 800 MHz and 900 MHz using analog switching technology at limited bandwidth of 10 MHz. However, 1G experienced demerits of poor voice quality due to excessive interference, user support limitation and cell coverage, degraded battery life and insecurity issues [3].

In late 1980s, the second generation (2G) of mobile communications was presented by the Global System for Mobile communication (GSM) technology that would later become a global wireless standard. GSM standard capability of supporting 14.4 to 64 Kbps maximum data rate was adequate

for short messaging services (SMS) and email services besides better quality voice transmission. The standard employed digital switching at a bandwidth of 30 to 200 KHz. In mid 1990s, Qualcomm launched Code Division Multiple Access (CDMA) system which exhibited enhanced spectral efficiency, number of users and data rate. A further improvement led to what was called 2.5G and 2.75G that saw the establishment of General Packet Radio Service (GPRS), CDMA2000 and Enhanced Data GSM Evolution (EDGE). The maximum data rate was enhanced to 171 Kbps, 384 Kbps and 473.6 Kbps for GPRS, CDMA2000 and EDGE respectively [4], [5].

The introduction of Universal Mobile Terrestrial / Telecommunication Systems (UMTS) with a data rate of 384kbps led to the third generation (3G) of mobile communications. 3G was apt in multimedia chat, email, video calling, games, social media, location tracking, maps and healthcare. For enhanced QoS, two technology improvements namely; High Speed Downlink Packet access (HSDPA) and High-Speed Uplink Packet Access (HSUPA) that saw data rate improved to 2 Mbps were introduced resulting to 3.5G. The 3.75G is a refinement of the 3G system employing the High-Speed Packet Access plus (HSPA+) technology. Nevertheless, 3G roll out suffered expensive spectrum licenses, high bandwidth requirements, compatibility issues with previous technologies, high costs of infrastructure, devices and implementation [6].

Developed by IEEE, the fourth generation (4G) of mobile communications offered advanced data rates, handled more advanced multimedia services while deploying the 3rd Generation Partnership Project (3GPP) defined standards namely; long term evolution (LTE) and LTE advanced (LTE-A). 4G uses complex modulation schemes and carrier conglomeration for simultaneous uplink and downlink voice and data transmission over internet protocol (IP) packets. Some key features exhibited by 4G include data rate of up to 1Gbps, enhanced security and mobility, reduced latency, high definition (HD) gaming and streaming and enhanced voice over LTE network VoLTE. However, 4G still exhibits challenges of infrastructural implementation costs, expensive spectrum, costly end user mobile devices and time-consuming upgrades [7].

The fifth generation (5G) of mobile communications, currently at testing stage exhibits potential in delivering ultra-fast internet, multimedia experience for customers, ultra-low latency significant for mission critical applications, higher security and mobility. 5G employs heterogeneous cells, non-orthogonal multiple access and beam forming technologies as well as cloud-based infrastructure to improve capacity and spectrum efficiency [8–10]. Table 1 highlights the comparison of 1G to 5G technologies.

Table 1: Comparison of 1G to 5G technologies

Generation	Data Rate	Frequency band	Technology	Switching	Key features
1G (1970-1980s)	14.4 Kbps	800 MHz	AMPS,NMT ,TACS	Circuit	Voice only services
2G (1990-2000)	9.6/14.4Kbps	850/900/ 1800/1900	TDMA, CDMA	Circuit	Voice and Data services
2.5G to 2.75G (2001-2004)	171.2 Kbps 20-40 Kbps	MHz 850/900/ 1800/1900 MHz	GPRS	Circuit	Voice, Data and web mobile internet, low speed streaming and email services
3G (2001-2004)	3.1 Mbps 500-700 Mbps	850/900/ 1800/1900/ 2100MHz	CDMA2000 UMTS and EDGE (1xRTT ,EVDO)	Circuit/ Packet	Voice, Data, Multimedia, smart phone applications, faster web browsing,video calling and TV streaming.
3.5G (2006-2010)	14.4 Mbps 1-3 Mbps	850/900/ 1800/1900/ 2100MHz	HSPA	Packet	All 3G services with enhanced speed and mobility support.
4G (2010 - onwards)	100-300 Mbps 3-5 Mbps 100 Mbps (Wi-Fi)	1.8/2.6GHz 2.3/2.5/ 3.5 GHz	WiMAX, LTE and Wi-Fi	Packet	High speed high , quality VoIP HD multi-media streaming, 3D gaming,HD video conferencing and worldwide roaming.
5G (2020 - onwards)	1-10 Gbps	1.8/2.6GHz Expected 30-300 GHz	LTE-A, OMA and NOMA	Packet	Super-fast mobile internet, low latency network for mission critical applications, IoT, security and surveillance, HD multi-media, autonomous driving, smart health-care applications.

3 Fifth generation (5G) networks

In this work, the focus is on the capacity and spectral efficiency improvement of 5G networks by considering architectural alternatives and multi-radio access techniques enhancements. The following sections present the 5G application scenarios, challenges and potential facilitators to address the challenges and the design fundamentals.

3.1 5G Application scenarios

There is considerable pressure with telecommunication giants to define the key 5G requirements namely; improving cellular network architecture in order to address challenges of increased capacity, improved data rate, minimized latency, expendable connectivity and robust QoS. According to Mobile and wireless communications enablers for the Twenty -Twenty Information Society (METIS) deliverables [11], [12] and International Mobile Communications framework IMT-2020 [13], 5G targets to address three broad generic services [14] namely;

1. *Enhanced Mobile Broadband (eMBB)*: eMBB confronts the growing traffic volume and data rate requirements by the accelerated number of new applicants such as virtual and augmented reality. The reliability and robustness of the bandwidth, wide-area coverage, hotspots, spectral

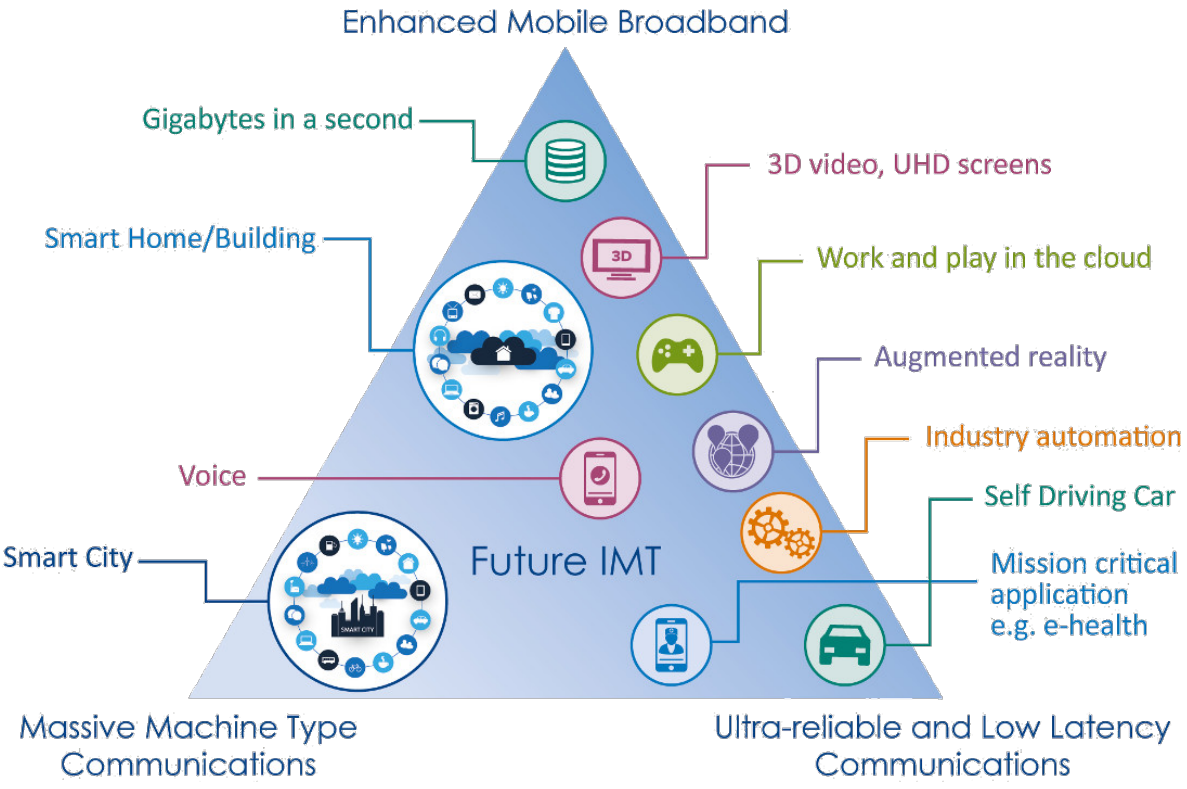

Fig. 1: 5G application scenarios [15].

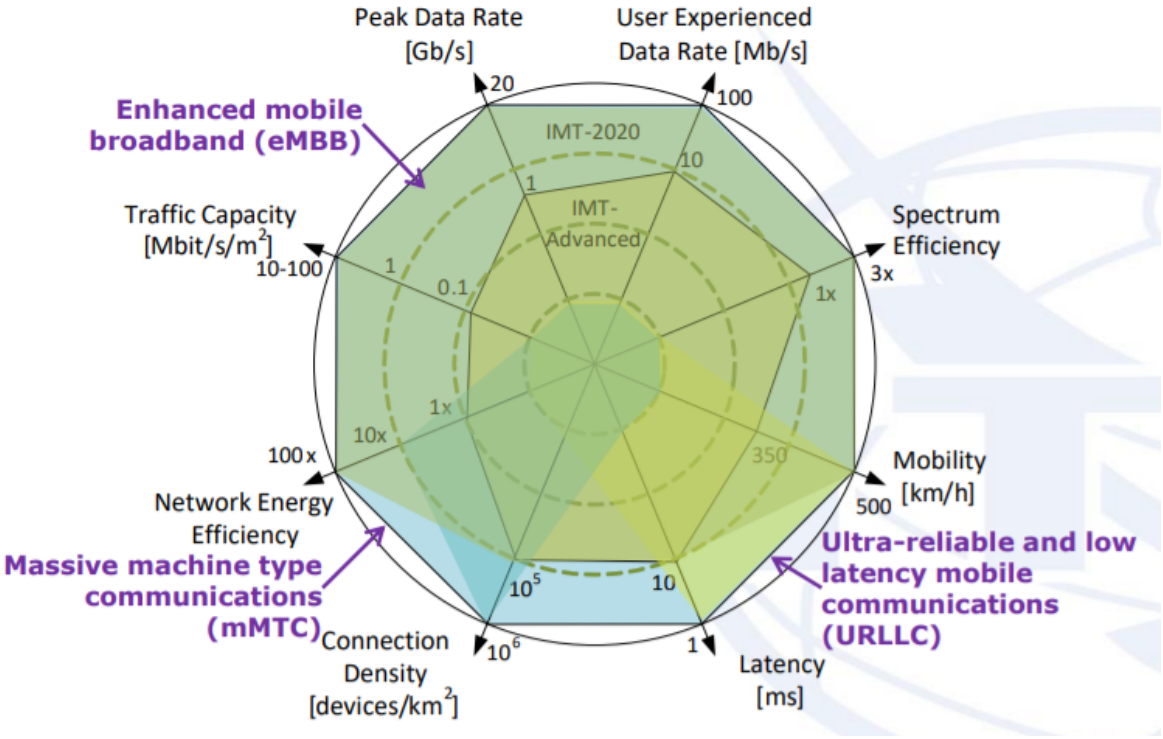

Fig. 2: 5G requirements (IMT 2020) [15].

and signal efficiency will be requisite by new 5G radio systems [16].

2. *massive Machine Type Communication (mMTC)* : mMTC involves the massive deployments of devices with non-delay sensitive data such as sensors and actuators, remote controlled and readable utility meters. 5G must guarantee connectivity solutions for tens of billions of devices with minimal human interaction. Since the devices experience diverse data traffic needs, divergent RA solutions are vital for an efficient system [17].

3. *Ultra reliable and low latency communications (URLLC)*: URLLC concerns the provision of service levels with very high probability [18]. Applications where low delay and uncompromising latency requirements are critical factors such as remote driving, smart grids, industrial control, and haptic communication enabling remote work in, e.g., hazardous environments or remote surgery can be supported with URLLC service. Fig. 1 and Fig. 2 illustrate the application scenarios of the 5G services and 5G requirements according to IMT 2020 [15]. respectively.

3.2 5G Challenges

Research on 5G standardization and deployment targeting to address the ubiquitous network demands is active [19]. In this undertaking, researchers and operators are awake to the challenges,

not adequately addressed by already deployed 4G networks, that 5G mobile networks must address. Challenges namely; higher capacity, higher data rate, interference management, lower end-to-end (E2E) latency, massive device connectivity, reduced capital and operations cost, and consistent quality of experience (QoE) provisioning are reviewed in [20], [21] and emphasized in this work.

3.2.1 Capacity and data rate

The current data traffic demands are as a result of exponential increase compared to demands in 2010. 5G and beyond mobile networks are expected to support a 1000-fold increase in the number of mobile devices, IoTs, connected cars and homes, moving robots and sensors [8]. According to IMT-2020 [21], traffic capacity is expected to be $10-100$ Mbit/s/m^2. Around 70% of this traffic capacity arise in indoor usage such as in homesteads, offices, malls, train stations, and other public building places [22]. Even as data traffic increases, it is observed that signalling traffic is growing 50% faster than data traffic [23]. The spectral efficiency has sharply increased 10 times while peak data rate of 10 Gb/s and 1 Gb/s for low mobility and high mobility respectively is experienced. These scenarios call for enhanced capacity not only in the radio access network (RAN) but also in the backbone, back-haul, and front-haul capacities. Consensually among operators, the provision of more spectrum through deployment of millimeter wave band, enhanced spectrum efficiency, network densification, and offloading will be necessary to address these challenges in the RAN [24].

3.2.2 Interference management

Due to the unusual demand for high bandwidth and high spectrum sharing requirement, inter-cell interference in ultra-dense HetNets becomes inevitable. Consequently, inter-cell interference (ICI) management plays an increasingly important role in mobile cellular networks [25]. In dense hierarchical networks (i.e., with several tiers), cross-tier and co-tier interference is inevitable. Cross-tier interference prevails between users sharing spectrum resources in different layers whereas co-tier interference occurs amongst users in the same network layer [26]. Co-tier interference can be observed either as inter- or intra-cell interference. Due to operators' inadequate control in planning the small cell's location occasioned by spontaneous deployment of small cell, inter-cell interference presents a major challenge. In addition, the disparities in traffic loads occasioned by varying transmit powers in different BSs leads to more challenging interference management and RA problems in HetNets compared to conventional single-tier systems. Further, users in different tiers experiencing different traffic restrictions (e.g., public, private, hybrid, etc.) thus exhibit diverse interference levels [27]. Furthermore, carrier aggregation and provision of cooperation among BSs and peer-to-peer communication (P2P) complicate the interference dynamics. Interference presents

constraints in achieving 5G capacity. Potential ICI management schemes including advanced air-interface techniques (E.g. NOMA) and advanced coordinated communications where coordination is performed at both the network and device sides need to be designed [28].

3.2.3 End-to-end latency

In the recent decade, there has been a surge in emerging applications with ultra-low end-to-end (E2E) latency. E2E latency is critical for real time applications such as remote-controlled medical robotics and industrial applications that require swift feedback control cycles for proper functioning. Critical safety-aware applications employing vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications require available and reliable quick request-response and feedback control cycles. Applications such as augmented and virtual reality require very fast feedback cycles to relieve cyber sickness. 5G networks therefore, must be able to reliably support below 1 ms E2E latency [29], [30]. To realize the 5G ultra-low latency requirement, operators are employing new air interface techniques with shorter transmission time interval (TTI). Enhanced higher-layer protocols such as case and network-aware admission/congestion control algorithms can be used to replace the conventional TCP slow start. Other techniques include use of D2D and HetNets to bring communication end-points closer [31]. Furthermore, E2E latency can be reduced by adding more intelligence at the network edges through pre-fetching and caching techniques, use of service-dependent location of control plane (C-plane) protocols and automated orchestration [32].

3.2.4 Connectivity

The number of interconnected devices has meteorically increased in the recent years by between $10\tilde{}100$ fold. Research by IMT-2020 outlines that 5G networks are expected to achieve a connectivity density of 1 Million devices per square kilometre (10^6/km^2) [29]. The devices differ in resource requirements; devices with meagre resources and requiring only intermittent connectivity for their operations such as sensors and those that demand an always-on connectivity for monitoring and/or tracking for proper functioning such as traffic management services, SCADA and security support devices. Such range of devices exhibit diverse connection and service requirement challenges for scalable and efficient operation. A number of solution-oriented techniques can be deployed to support hyper-connectivity including advances in air interface design, multi-radio resource access for spectrum sharing, signalling optimization, digital signal processing techniques, intelligent clustering and relaying techniques [33].

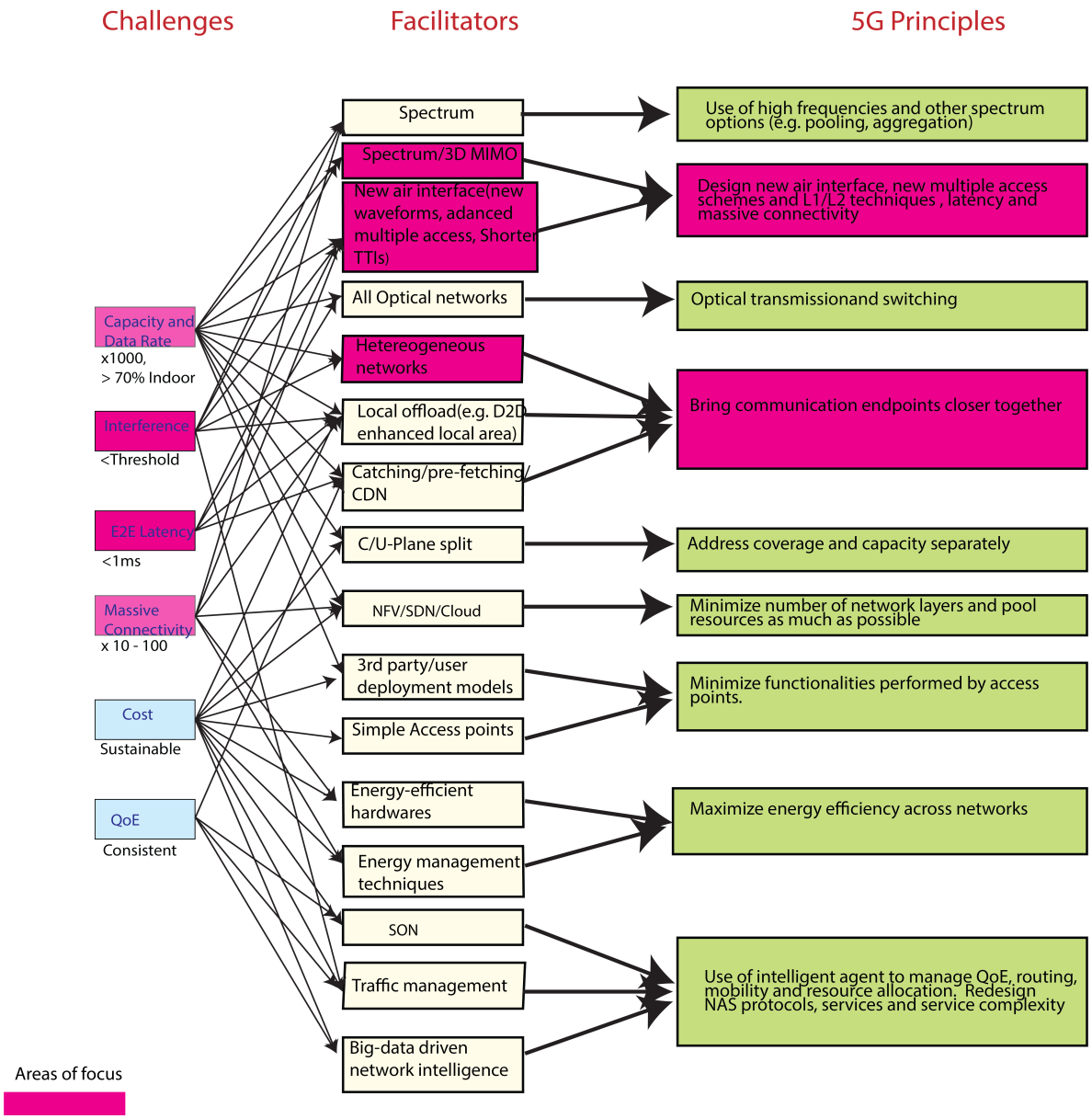

Fig. 3: 5G challenges, facilitators and design principles [29]

3.2.5 Cost

In the current age, connectivity is an indispensable pillar for accelerating digitalisation, driving innovation and overall socio-economic development. As a result, it is vital to consider the investment cost of infrastructure including roll-out costs, maintenance, managerial, and operational costs to make connectivity a globally available, accessible, and tenable utility. These costs are huge and operators are trying to minimize passing the costs to the consumers. As a requisite, 5G networks from RAN, core, backbone routers, and backhaul need to devise ways that provide consumer demands at sustainable service provision. Some 5G enablers such as deployment of ultra-dense HetNets bring about huge equipment, maintenance and operational costs. One of the ways to cub cost is to reduce BSs functionalities through use of one-layer functionality and keeping higher layer functionalities to a network cloud. Another challenge is the network energy consumption, with RAN consuming approximately 80% of the energy requirements. Smart energy management solutions that put BS to sleep when not in use, energy efficient hardware designs, low-power backhaul, network function virtualization (NFV) and software defined networking (SDN) can all contribute to reducing the cost of operating a 5G network [29], [34].

3.2.6 Quality of Experience

Quality of experience (QoE) characterises the user's perception on the application and service reception. Since different users have unique demands, QoE is uniquely application specific and also user-specific and thus exclusive. For instance, the quality of the encoded and delivered video affect its QoE. Insufficient QoE leads to user discontent whereas excessive QoE strains the available user and network resources. Depending on the application and user needs, network parameter such as bandwidth, delay, latency and context parameters of device, user, and environment are used to describe QoE. Most 5G enablers can significantly improve QoE. Besides, traffic optimization techniques, cache and computing resources installation at the network edge could be employed to meet increasing QoE expectations. Fig.3 illustrates the 5G challenges, facilitators and design principles.

3.3 5G Technical Requirements

The minimum technical performance requirements for 5G are defined in [21]. They can be summarised as follows;

1. *Peak data rate*: which is the maximum received data bits (in bit/s) assuming error-free conditions assignable to a single mobile station. The minimum downlink and uplink peak data

rate are 20 Gbit/s and 10 Gbit/s respectively.

2. *Peak spectral efficiency*: This is the maximum data rate under ideal conditions normalised by channel bandwidth (in bit/s/Hz). This requirement is purposed for the evaluation of eMBB usage scenario. The minimum downlink and uplink peak spectral efficiencies are 30 bit/s/Hz and 15 bit/s/Hz respectively.

3. *User experienced data rate*:Based on the user throughput defined as the number of correctly received bits, user experienced data rate is the 5% point of user throughput cumulative distribution function. IMT-2020 outlines the minimum downlink and uplink user experienced data rate for eMMB usage as 100 Mbit/s and 50 Mbit/s respectively.

4. 5^{th} *percentile user spectral efficiency*: This is the 5% point of the CDF of the normalized user throughput. The normalized user throughput is defined as the number of correctly received bits, over a certain period of time, divided by the channel bandwidth and is measured in bit/s/Hz. The minimum requirement in downlink and uplink is 0.3 bit/s/Hz and 0.21 bit/s/Hz respectively.

5. *Average spectral efficiency*: This is the aggregate throughput of all users divided by the channel bandwidth of a specific band divided by the number of TRxPs and is measured in bit/s/Hz/TRxP. For dense eMBB, the minimum requirement in downlink and uplink is 7.8 bit/s/Hz/TRxP and 5.4 bit/s/Hz/TRxP respectively.

6. *Area traffic capacity*: This is the total traffic throughput served per geographic area (in Mbit/s/m^2). The target value for Area traffic capacity in downlink is 10 Mbit/s/m^2 in the Indoor Hotspot – eMBB test environment.

7. *Latency*: Two types of latency are defined for eMBB and URLLC usage scenarios. User plane latency is total time the radio network takes to send a packet from source to destination (in ms), the minimum given as 4 ms for eMBB and 1 ms for URLLC for both downlink and uplink. On the other hand, control plane latency refers to the transition time from idle state to the start of continuous data transfer (active state) advised to be lower than 20 ms.

8. *Connection density*: Defined for mMTC usage scenario, connection density is the total number of devices fulfilling a specific quality of service (QoS) per unit area (per km^2). The minimum requirement for connection density is 1 million devices per km^2.

9. *Energy efficiency*: Network energy efficiency is the capability of a radio interface technology (RIT) to minimize the radio access network energy consumption in relation to the traffic capacity provided. Device energy efficiency is the capability of the RIT to minimize the power consumed

by the device modem in relation to the traffic characteristics.

10. *Reliability*: Defined for uRLLC, reliability basically is the success probability of transmitting a layer at least two thirds within a required maximum time. Proponents encourage reliability is $1 - 10^{-5}$ success probability.

11. *Mobility*: This is the maximum mobile station speed at which a defined QoS can be achieved (in km/h) defined for different classes of mobility namely; 0 km/h for stationary, $0 - 10$ km/h for pedestrian, $10 - 120$ km/h for vehicular and $120 - 500$ km/h for high speed vehicular. It is envisaged that the minimum traffic channel link data rate normalized by bandwidth should be 1.5, 1.12, 0.8 and 0.45 bit/s/Hz for 10, 30, 120 and 500 km/h mobility respectively.

12. *Mobility interruption time*: It is the shortest time duration supported by the system during which a user terminal cannot exchange user plane packets with any base station during transitions. This includes time required for network procedure execution, control signalling protocol time and other message exchange time. The minimum requirement for mobility interruption time is 0 ms for both eMBB and URLLC usage scenarios.

13. *Bandwidth*: Defined as the maximum aggregated system bandwidth supported by single or multiple radio frequency (RF) carriers. The requirement for bandwidth is at least 100 MHz.

3.4 Potential 5G facilitating technologies

Recent research points out to six key facilitators that will have an immense impact on 5G progression. They include; dense small cell deployment, massive MIMO (m-MIMO), device to device (D2D), machine to machine (M2M), multiple radio access technologies (M-RAT) and millimetre-wave communications. Additionally, advanced waveforms, futuristic coordinated multipoint (CoMP), carrier aggregation, structured coding techniques, use of network virtualization and deployment of cloud radio access networks (C-RANs) will boast 5G networks realization [19]. The 5G multitier network architecture presented in Fig. 4, [20] includes macro-cell, small-cells, Wi-Fi, massive multiple-input, multiple-output (M-MIMO) with beamforming, internet of things (IoT), cognitive radio networks (CRN), visible light communication (VLC), device-to-device (D2D), machine-to machine (M2M) communications and NFV enabled network cloud. To alleviate some of the highlighted challenges, this work focuses on establishing efficient multiple access technologies. In particular, hybrid non-orthogonal multiple access (NOMA) technologies are proposed due to the potential to transmit multiple users compared to the available orthogonal resource elements.

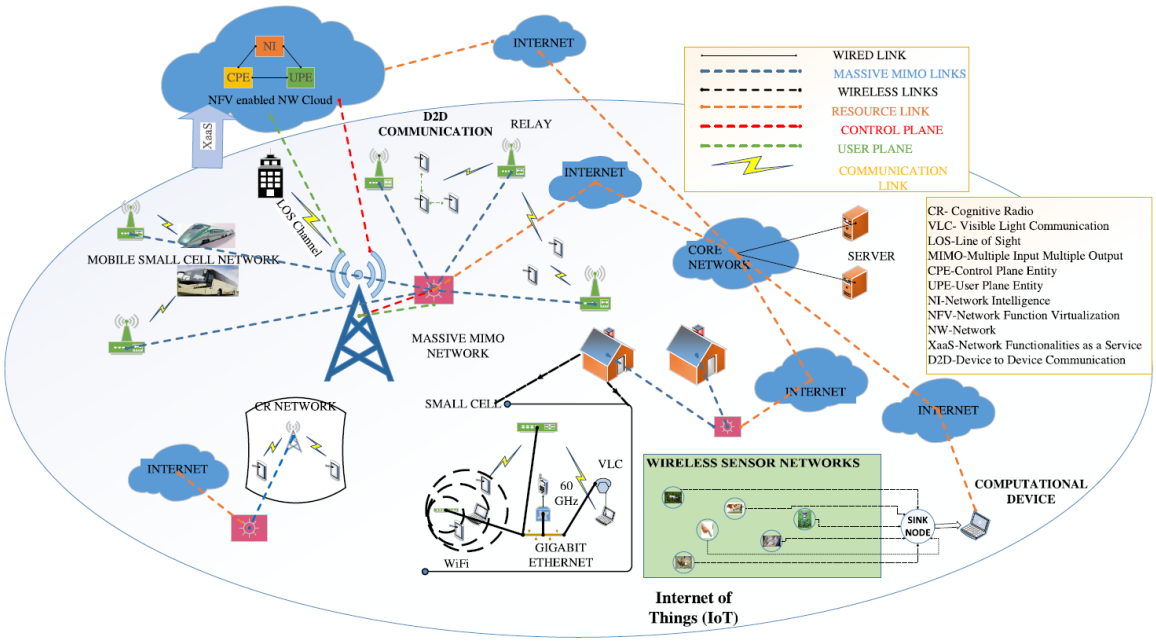

Fig. 4: 5G network architecture [20]

3.4.1 Dense Heterogeneous Networks

Initially launched in 4G networks, dense HetNets involve deployment and concurrent operation of sophisticated underlaid small cells in a macro-cell. The small cells are limited radius networks complete with small base stations (SBS) and access points (APs) for spectrum re-use. Hence, they help achieve increased spectral efficiency, network capacity and significantly reduce the power consumption of mobile device due to its communication with nearby SBS. Small cells can be employed in indoor and outdoor environments. However, HetNets are faced with design challenges and require innovation in hardware miniaturization and cost reduction in the SBS design. Furthermore, interference management presents a challenge due to the uncoordinated nature of HetNet deployments [35]. Another potential challenge is the significant rise in the handoff rate resulting to call drops since mobile stations (MS) must move around many hotspots. To avert this challenge, control/user plane splitting can be deployed where MS are enabled to issue access requests to BS and signalling techniques employed to mitigate the handover failure rate [29]. This work focus on a design model of a hybrid NOMA technique on multi-tier HetNet consisting of small cell user equipment's (SUEs) and the macro user equipment's (MUEs). In particular, we consider the small cell spectrum resource re-use with the macro cell users, without compromising the small cell user's QoS requirements. To achieve the specifications, proper design of resource allocation at the transmitter and optimal multi-user detection schemes are prudent.

3.4.2 Massive MIMO and beamforming.

The deployment of massive MIMO (*m*-MIMO) involves increasing the degree of freedom in wireless channels by use on multiple transceiver antennas. Beamforming, on the other hand, involves concentrating a large gain limited beam-width power in a specific direction. Hence, a significant performance improvement in terms of reliability, cell throughput and better cell edge performance can be obtained [36]. M-MIMO suffers pilot contamination from neighbouring cells as number of antennas grow. Other challenges include channel estimation accuracy which necessitate sophisticated algorithms and the large-scale m-MIMO architecture requirement is a point of concern [37].

3.4.3 Device to Device Communications.

Device to Device Communications (D2D) is a potential 5G technology that enables new peer-to-peer and location-based applications and services like public safety networks [38], [39] thereby improving spectral resource utilization, reduce energy consumption and substantially reduce latency [40], thus suitable for delay-sensitive application scenarios. D2D communications necessitate effective radio resource management strategies to properly coordinate mutual interference between cellular and D2D users in the reuse mode [41].

3.4.4 Machine to Machine Communications

Machine to Machine (M2M) communications involve applications that transmit only minimal amounts data and only intermittently. The data exchanged in an M2M network is generally sparse since it can be from sensors, computer devices. This kind of data creates major signalling overheads on the mobile network [42]. Such applications include smart meters, smart traffic systems, vending machines, security alerts and tracking. To develop an application with M2M communication there are many issues, inter-operability, coverage, energy efficiency, portability, and scalability [9]. For interference management and to avoid collisions, efficient MAC protocols are deployed.

3.4.5 Millimetre-wave frequency band.

Previous generations of mobile network overwhelmingly operated on the microwave band, due to its convenient propagating characteristics, making it too scarce [20]. The ubiquitous demands for higher capacity, connectivity and better QoS necessitate additional spectrum. Availability of substantial quantity of spectrum can be made accessible if the millimeter-wave band is utilized to fulfil all the 5G requirements. It has been suggested that 5G uses mm-wave bands from $20 - 90$ GHz due to the availability of unused bandwidth. This band comes with challenges of propagation, atmospheric

absorption and hardware constraints in comparison with microwave band which can be mitigated by using beamforming and a larger antenna array [43].

3.4.6 Multi-radio access technologies.

In the recent past, the radio access technologies have undergone tremendous evolvement. In particular, for 1G, 2G, 3G, and 4G network systems, frequency division multiple access (FDMA), time division multiple access (TDMA), code division multiple access (CDMA), and orthogonal frequency division multiple access (OFDMA) have respectively been employed as the corresponding key multiple access technologies [44]. These technologies are designed to orthogonally allocate wireless resources to multiple users (OMA). Faced with increasing number of supported users by the limited resources and channel impairments that distort the orthogonality, it remains a challenge for OMA to meet the stringent spectral efficiency and massive connectivity 5G requirements. Consequently, NOMA concept has been proposed. Basically, NOMA supports non-orthogonal RA among the users, i.e., more users are supported than the number of accessible orthogonal time-, frequency-, code-domain resources or their combinations, at the expense of increased receiver computational complexity [1]. Several NOMA solutions are discussed in literature and broadly categorized in three broad categories namely; power-domain NOMA, code-domain NOMA and interleave division multiple access (IDMA). NOMA being the focus of this work, the basic principles and theoretical analysis of the NOMA techniques is discussed in the subsequent sections. The work further examines the possibility of blending existing NOMA techniques for a possible hybrid NOMA with the purpose of enhancing connectivity, spectral and energy efficiency.

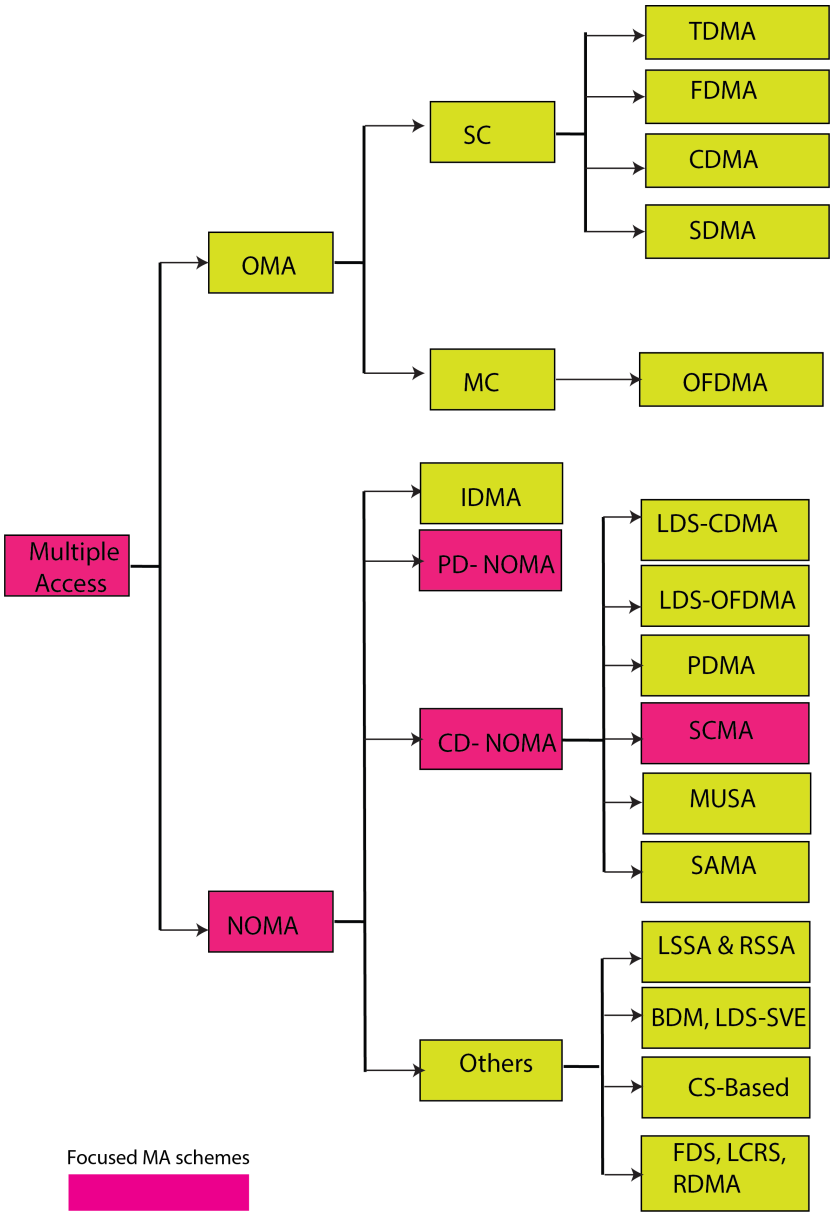

Fig. 5: OMA and NOMA categories of multiple access.

4 Non-orthogonal multiple access (NOMA) technologies

In this section, we detail NOMA basic principles as contrasted to OMA. The OMA and NOMA categories are presented in Fig. 5. In OMA techniques namely; FDMA, TDMA, CDMA and OFDMA respectively used for 1G, 2G, 3G, and 4G, multiple users are assigned to orthogonal radio resources in the time-, frequency-, code-domain or to their combinations. In FDMA, users transmit unique, user-specific signals over unique frequency resources. At the receiver, each users' data in their corresponding frequency bands are readily detected. Likewise, in TDMA, exclusive time slots are allocated to individual users, hence little effort to distinguish the different users' signals at the

receivers in the time domain is required. CDMA technique involves mapping of different users to orthogonal spreading sequences like Walsh-Hadamard codes and then transmit while sharing the same time-frequency resources. A decorrelation MUD is then utilized at the receiver. In OFDMA, radio resources are split in the time-frequency grid and can be observed as a coalescing of FDMA and TDMA [45]. Though the orthogonality in OMA resource utilization results in minimal interference and hence low complex linear detection is sufficient, the maximal number of admissible users is strictly limited by the number of orthogonal resources accessible in conventional OMA schemes, which becomes a tight bound when massive connectivity is required for 5G. Besides, OMA places a hard limit on the maximum achievable sum-rate of multi-user wireless systems [46].

NOMA schemes promise to circumvent the highlighted OMA limitations. Through non-orthogonal allocation of resources and controlled interference, NOMA can achieve overloading by multiplexing more users than the number of orthogonal resources. At the receiver, sophisticated MUDs with compromising polynomial or exponential computational complexity orders as compared to OMA receivers are employed. Generally, NOMA access schemes are classified into three broad categories namely; inter-leaver division multiple access (IDMA), power domain NOMA (PD-NOMA) and code domain NOMA (CD-NOMA). NOMA is the focus areas of this work and the categories are discussed exhaustively in the sections below.

4.1 Power - Domain NOMA

The concept of PD-NOMA was first presented in [47]. Theoretically, PD-NOMA scheme draws its roots in multi-user information theory. They include scalar and vector multiple access, superposition coding, dirty paper coding, iterative water-filling, joint decoding including SIC and other transceiver related schemes [48]. Through superposition coding [49], channel coding and modulation at the transmitter, multiple users are directly superimposed on each other and simultaneously share the time-frequency resources. PD-NOMA is achieved by allocating distinct power levels to different

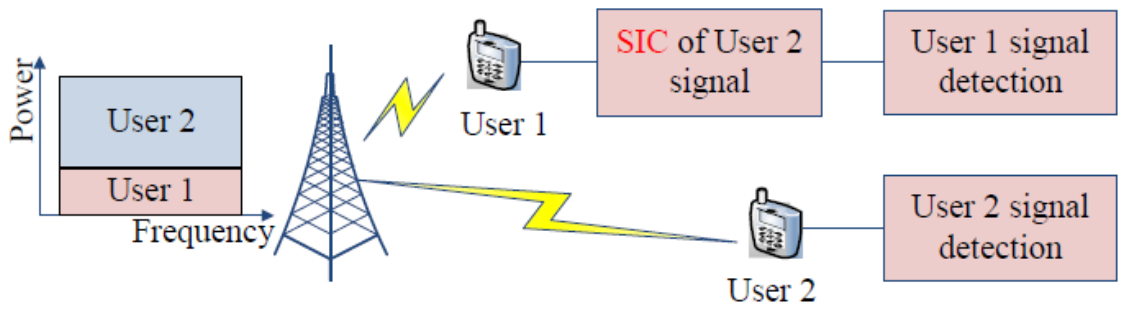

Fig. 6: Downlink PD-NOMA transceiver model [1]

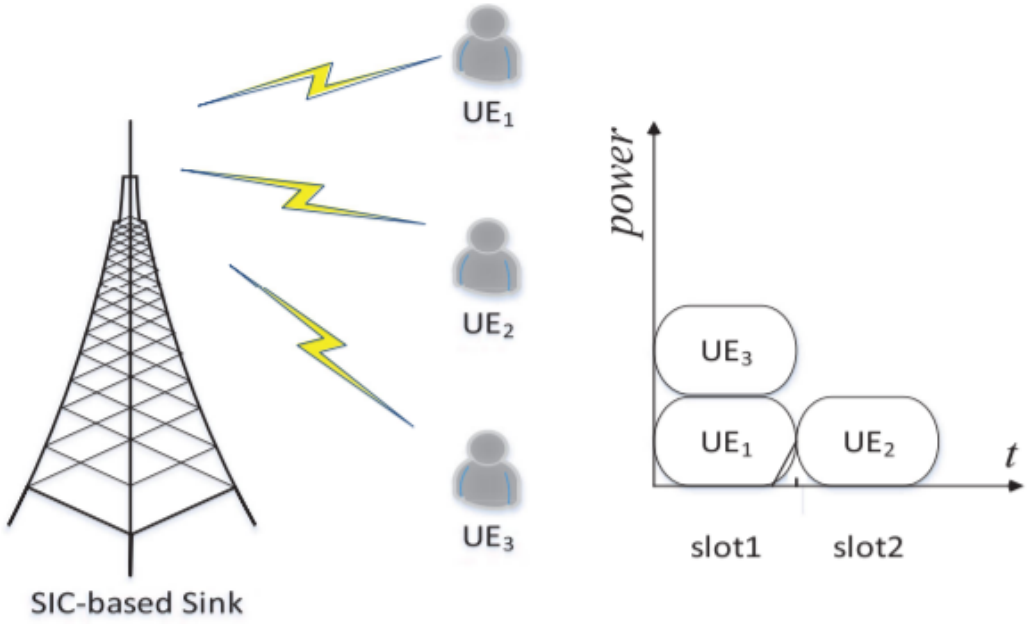

Fig. 7: Uplink PD-NOMA transceiver model [1]

users. At the receiver, user detection is achieved using MUD algorithms such as SIC.

A PD-NOMA applied in the downlink and uplink scenarios for two single antenna users is shown in Fig. 6 and Fig. 7 respectively. The total power is allocated to all the users subject to defined power-scaling coefficients, then at the BS the users are superimposed. At the receiver, the optimal SIC detection order relies on detecting the strongest to the weakest user. This implies that, any user can detect its message with little interference-contamination exacted by the other users with smaller normalized channel gain in the order. Therefore, strongest users cancel interference from weak users. In the uplink, the BS computes the signal power loss as the signal propagates through the wireless channel from user to BS and uses the parameter to evaluate the decoding order. We assume that gains of all users remain constant during a frame, which is realistic for slow fading channels [50]. In an uplink system with K users, the BS decodes $j-th(j<i)$ user signal first before it detects the $i-th$ user's signal then removes $(i-1)$ users' signals from the observed signal $\mathbf{y}$. The rest $(K-i)$ signals are treated as interference.

The performance of PD-NOMA depends majorly on the optimality of SIC. Since SIC performance is conditioned on the number of distinct power levels and QoS requirements, several works have focused on classic PD-NOMA power allocation schemes [51–56]. The authors show that power allocation for the superposition process at the transmitter and SIC process at the receiver should be conducted diligently since the power distinctiveness between the weak and strong users can be immense with

increasing number of users or modulation orders. The research on integration of PD-NOMA with other existing techniques such as MIMO, cognitive radio networks (CRNs), cooperative communication and HetNets is actively on-going. Proposed for a NOMA system in [57] is a power domain cyclic spread multiple access (PDCSMA). Though PD-NOMA offers merits of spectral and capacity enhancement, its performance is prone to additional receiver computational complexity compared to OMA. There is also the possibility of errors rippling and overspreading to consecutive users through the decoding process on account a user makes an error [58].

4.2 Code - Domain NOMA

In this section, we discuss CDMA inspired code domain NOMA (CD-NOMA) schemes which achieves user multiplexing by employing unique user-specific spreading sequences. The multiple users then share the same time-frequency resources. Unlike in CDMA, the CD-NOMA spreading sequences are strictly sparse and non-orthogonal with low cross-correlation.

4.2.1 Low-density spreading CDMA (LDS-CDMA) and Low-density spreading OFDM (LDS-OFDM)

The basic principles of LDS-CDMA, observed as the very initial NOMA technique, are discussed in [59]. LDS-CDMA draws its inspiration from the conventional CDMA and low-density parity check (LDPC) coding [60], with the aim of reducing user interference and receiver computational complexity. The block model of LDS-CDMA is shown in Fig. 8. In CDMA structure transmitting over memoryless symbol-synchronous channel, every chip of the received signal carries the contribution from all the users in the system. That is to say, each user observes the contribution from all the other users at every received chip. Contrary, the LDS-CDMA structure switches off a large portion of the spreading chips. This implies that each user only spreads its message over a few number of chips. Since the number of superimposed signals carried at each chip is less than the number of users, the imposed interference significantly reduces. By careful design of the spreading sequences, multi-user interference can be mitigated [1]. In [61], the authors propose a structured approach to design LDS codes for LDS-CDMA scheme. The approach involves mapping of the signature constellation elements to the spreading matrix that hosts the spreading sequences. Authors in [62] proposed a near-optimal MUD using belief propagation (BP) with low-computational complexity. A dynamic factor graph-based MUD that provides a good trade-off between bit error rate (BER) performance and computational computational complexity is proposed in [63].

The LDS-OFDM on the other hand, can be seen as a blend of OFDM and LDS-CDMA techniques.

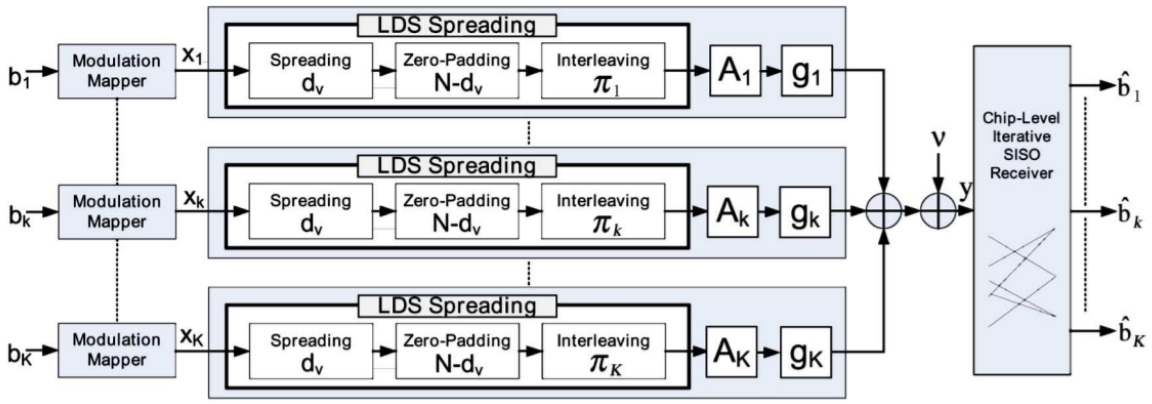

Fig. 8: LDS-CDMA system [59]

Here, every user's symbol is spread across a discreetly selected number of sub-carriers and overlaid on top of each other in the frequency domain. In the conventional OFDMA, one user symbol is mapped to a sub-carrier and dissimilar users transmitted on disparate orthogonal subcarriers, eliminating the interference between each other. This means that the number of sub-carriers limit the number of user symbols transmitted. Contrary, in LDS-OFDM, user symbols are firstly convoluted with the LDS sequences with length corresponding to the number of subcarriers. The resultant chips are transmitted on disparate subcarriers. Here, only a fraction of the sub-carriers transmits the spread symbols and therefore each subcarrier transmits only the chips [64]. To improve receiver computational complexity, authors in [65] introduced an upper bound on the number of users per subcarriers. Authors in [66] investigated link-level and system-level performance while [67] proposed a joint subcarrier and power allocation method to improve the achievable LDS-OFDM performance.

4.2.2 Sparse Code Multiple Access (SCMA)

Sparse code multiple access (SCMA) is a novel NOMA scheme that draws its concepts from LDS-CDMA and OFDMA [68]. Contrary to LDS-CDMA encoder that consists of a quadrature amplitude modulation (QAM) mapper and a spreader to expand a QAM symbol to a complex symbol sequence using a CDMA signature, SCMA structure implements a joint design of multi-dimensional modulation and low-density spreading [69]. In SCMA, binary information is directly encoded to multi-dimensional complex codewords from a pre-defined set of codebooks. Multiple codebooks are generated to achieve multiple access through exclusive and sparse allocation of available resource elements (REs). With conventional SCMA, each codebook is assigned to one user [70]. The application of multiple users in a code book is our major pioneering and the focus of this research. Due to sparsity, iterative MUD algorithms such as message passing algorithm (MPA) [59] and recently proposed expectation propagation algorithm [71] can be used to detect the multiplexed

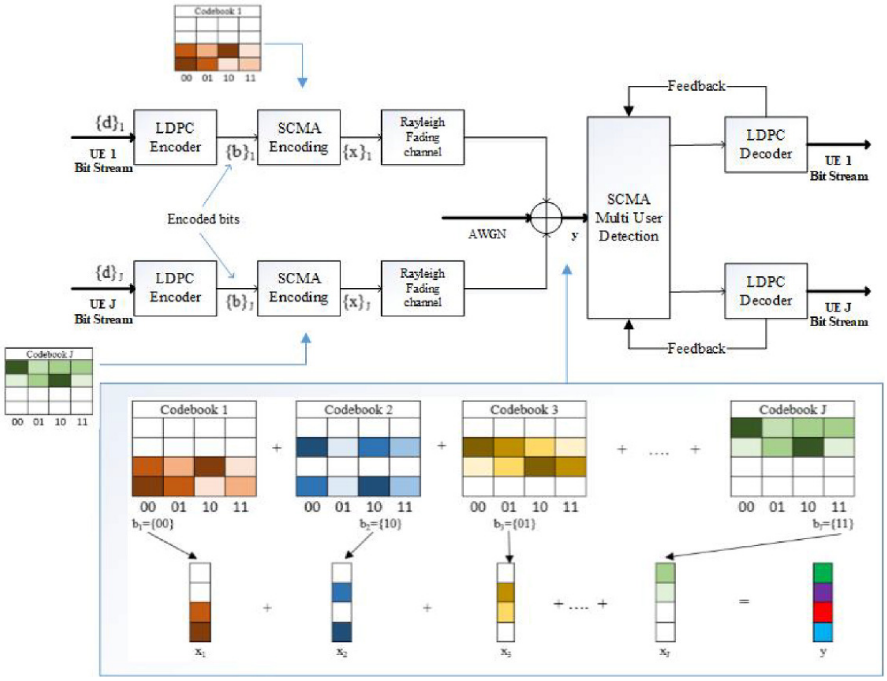

Fig. 9: SCMA transceiver system [72]

codewords with acceptable computational complexity. An SCMA transceiver system is illustrated in Fig. 9.

SCMA is characterized by advantages of user overloading, sparse codewords for moderate receiver computational complexity and shaping gain contributing to multiple dimensions for multiplexing. However, SCMA exhibits challenges namely; the multi-dimensional lattice constellation for optimal codebook design, RA i.e., REs, codebooks and power assignment and increased decoding computational complexity order for high number of users and constellations. Several works in literature present codebook designs such as constellation rotation and interleaving based techniques [71–76], golden angle modulation [77] and extended mother codebook [78] amongst other others. On SCMA RA, heuristic based RA [79], biological based RA [80], joint subcarrier and power allocation [81] and parametric based RA [82] among other RA schemes are proposed for spectral and capacity enhancement. Recently, research on integrating SCMA with other multiple access technologies has gained traction. Authors in [83] and [84] investigate the performance of MIMO based SCMA as the number of users and antennas varies. To further enhance the spectral and energy efficiency, the work in [85] integrates SCMA with PD-NOMA in a homogeneous network i.e., encoding same type of users in both the code-domain and the power-domain. This work develops a hybrid NOMA system that encodes low-power small cell users in SCMA codebooks in the code-domain and superimposes higher-power macro-cell users in the SCMA codebooks using power-domain multiplexing in a HetNet system [86]. The outcome of such a technique is the

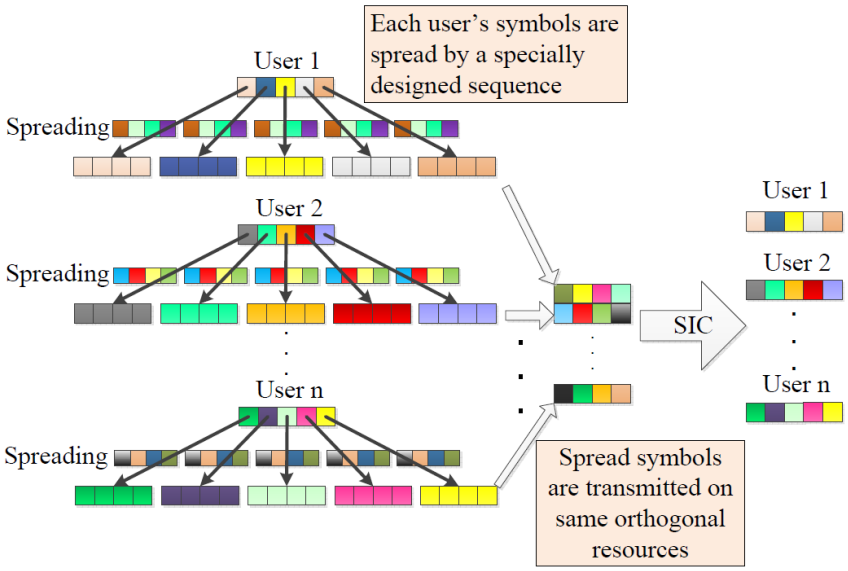

Fig. 10: MUSA system [1]

increased number of multiple users sharing a resource block albeit the increased receiver computational complexity.

4.2.3 Multi-user shared access (MUSA)

MUSA is a grant-free NOMA technique that is considered as an enhanced CDMA design where the transmitted user symbols are multiplied by the same low cross correlation spreading sequence [87]. The resulting sequences are then transmitted over the same OFDM resources as shown in Fig. 10. Before superposition in the downlink, the users are clustered where in each cluster the users' symbols are weighted using different power-scaling coefficients. In order to spread the superimposed symbols, orthogonal sequences with the groups' length are used as the spreading sequences. This implies that users in the same cluster utilize the same spreading sequence which are orthogonal across the different clusters. Inter-group interferences are minimized and SIC can be employed by making use of the associated power distinctiveness. MUSA exhibits enhanced downlink capacity, an explicit benefit from the SINR difference and SIC linearity. Additionally, MUSA can guarantee user fairness without compromising the capacity [1], [88]. Authors in [89] propose the 5-ary codes for an uplink MUSA. The analysis is done while varying the SIC detection order. In [90], the authors propose a PA algorithm that guarantees sufficient power difference between signals to mitigate error propagation in a mMTC application scenario. Authors in [91] employ MUSA to support IoT overloading requirements. In [92], real Fourier-related transform spreading methods are proposed in order to minimize the peak-to-average power ratio (PAPR) while increasing the stability in a fading channel environment. A deep neural network (DNN)-based MUD that observes the correlation between the received signal and the

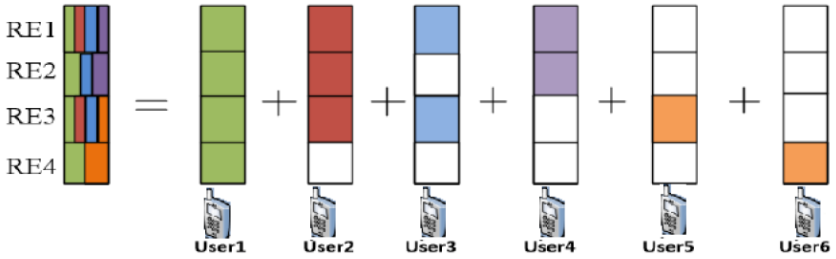

Fig. 11: PDMA transceiver system [94]

user-specific spreading sequences for MUSA is proposed in [93].

4.2.4 Pattern Division Multiple Access (PDMA)

PDMA is another promising NOMA technique implemented in multiple domains to realize distinct transmission diversity order [94]. Non-orthogonal patterns which are designed through maximization of the diversity and minimization of the multiple users' correlation are employed in PDMA at the transmitter. Multiplexing is then realized in the power-, code- or spatial-domains, or their combinations. Such design of PDMA patterns not only distinguish the user symbols utilizing similar resources but also enhance system performance with affordable detection computational complexity. A PDMA pattern resource mapping with six users multiplexed on four REs is illustrated in Fig. 11. Each PDMA pattern is assigned to a single user. From Fig. 11, user_1's message is mapped to all four REs in the group, and user_2's message is mapped to the REs 1, 2 and 3, etc. The transmission diversity order for the six users is 4, 3, 2, 2, 1, and 1, respectively [95]. Recent studies on PDMA include outage performance and achievable sum data rate performance investigation for 5G systems [96] A discrete Fourier transform spread generalized multi-carrier based PDMA is proposed in [97]. A joint PDMA transceiver proposed in [98] is designed for pattern mapping that utilizes power and beam allocation to enhance the achievable sum rate and the access connectivity at the transmitter respectively. A spatial filter to control the inter-beam interference and SIC are employed at the receiver.

4.3 Interleave Division Multiple Access (IDMA)

IDMA is an interleave based NOMA technique that characterize different users based on their different bit-level inter-leavers [99]. The IDMA transceiver system is illustrated in Fig. 12. The symbols are first multiplied with the spreading sequences then IDMA interleaves the chips. Accordingly, IDMA can be regarded as an effectively chip-interleaved CDMA. Fundamentally, the IDMA idea involves utilizing user-specific inter-leavers together with low-rate channel coding as

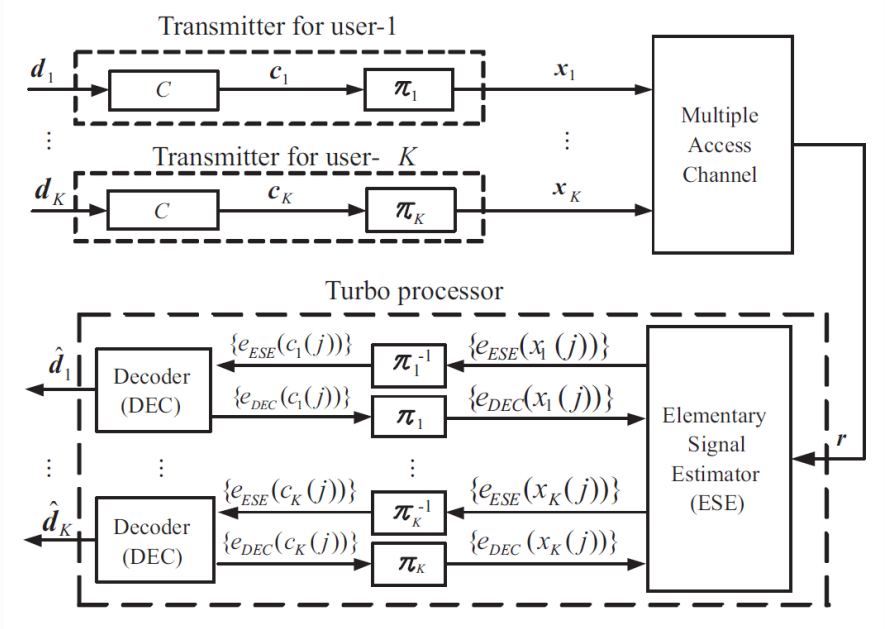

Fig. 12: IDMA transceiver system [100]

illustrated in Fig.12, [100]. In comparison to CDMA, IDMA can achieve 1 dB SNR gain in overloaded systems and increased diversity gain [101]. Authors in [102] outline IDMA merits including allowing use of iterative low-cost MUD methods and realization of near-capacity multi-user sum-rate especially for power controlled IDMA. In addition, decentralized power control IDMA achieves higher throughput in comparison to the conventional ALOHA in random access scenarios. Lastly, data aided channel estimation (DACE) coupled IDMA can exploit massive multiple input multiple output (MIMO) systems.

4.4 Other NOMA Schemes

Besides the prominent schemes, alternative novel NOMA schemes are proposed and investigated in literature for spectral efficiency and connectivity enhancement. Successive interference cancellation aided multiple access (SAMA) technique [103] is analogous to MUSA except that in SAMA, the non-zero elements of its spreading sequence are equal to one for each user. The spreading sequence matrix is designed such that the number of groups with different 1's in the spreading sequence is maximized while the number of the overlapping spreading sequences with the same number of 1's is minimized. The maximal number of users supported by N orthogonal sub-carriers can be given by $2^N - 1$. Another CD-NOMA scheme is the spatial division multiple access (SDMA) whereby, instead of using unique, user-specific spreading sequences, SDMA utilizes unique, user specific channel impulse response (CIR) to preserve the orthogonality which may be destroyed by the convolution of user symbols and CIR over dispersive channels.

Another category of alternative NOMA is the signature-based NOMA that constitutes the low code rate and signature based shared access (LSSA) [104] and resource spread multiple access (RSMA) [105]. LSSA multiplexes user symbols using specific signature patterns consisting of a complex/binary sequence, reference signal (RS) and a short length vector permutation pattern. Similarly, RSMA assigns unique signatures to distinguish the different users and spreads those user symbols over all the accessible time-frequency resources. Both signature-based NOMA potentially allows synchronous access and grant free transmissions.

Other schemes include Bit division multiplexing (BDM) NOMA which depends on structured modulation while the time-frequency resources are split-up at the symbol level [106], compressive sensing (CS)-based NOMA that smartly exploits either the user activity or the data sparsity [107], low density spreading-signature vector extension (LDS-SVE) established by Fujitsu [108], frequency domain spreading (FDS) [109], low code rate spreading (LCRS) [109] schemes initiated by Intel and Repetition division multiple access (RDMA) developed by MTK [110].

This work focuses mainly on two NOMA schemes namely; PD-NOMA and SCMA. PD-NOMA derives its preference on the ability to superimpose different assigned users based on the power imbalance between the user signals through superposition coding. Through extensive research, PD-NOMA exhibits merits of higher achievable user rates and greater spectrum efficiency compared to OFDMA. Besides, PD-NOMA boasts of established successive interference cancellation schemes at the receiver. Although the optimal power allocation (PA) technique to guarantee optimal power distinctiveness amongst users is intractable, researchers have proposed near-optimal PA schemes (such as water filling technique) to strike a balance between serving the desired QoS and interference management. On the other hand, the code-domain based SCMA directly encodes the user symbols into codebooks derived from low-density multi-dimensional constellations. SCMA enjoys overloading, sparsity and shaping gain thereby contributing to multiple dimensions for multiplexing of the users. Research has shown that SCMA achieves better link level performance, spectral and energy efficiency outperforming other CD-NOMA schemes albeit compromised MUD computational complexity order. Research outlines that optimal resource (i.e., codebooks, resource elements and power) allocation enhances the performance of SCMA systems. Besides, SCMA enjoys the low computational complexity reception techniques such as message passing algorithm (MPA), max-log-MPA, log-MPA and expectation propagation algorithm (EPA) amongst others, due to the sparsity of SCMA codewords. With the above in mind, the feasibility and development of a hybrid NOMA technique that encodes user signals in code-domain using SCMA and superimposes user signals in power-domain applying PD-NOMA necessitates an investigation.

5 Research Problem, Motivation and Objectives

5.1 Research Problem and Motivation

In the past decade, network operators have witnessed an explosive increase in the number of connected mobile and IoT devices that exhibit wide-ranging wireless traffic data requirements. The prevailing challenges of resource allocation and scheduling, interference management, massive connectivity, traffic densification, high capacity and data rate, end to end latency sensitivity, cost, quality of experience and transceiver modelling pose considerable limitations in the development of modern sophisticated generation of networks. One of the potential facilitators for 5G realization and alleviation of the challenges is the design of advanced multi-radio access technologies. Research in the recent past recommends a paradigm shift from orthogonal multiple users to resources access (OMA) to non-orthogonal access (NOMA). NOMA supports a higher number of users than the number of orthogonal resource slots with the aid of non-orthogonal resource allocation. This is realized through sophisticated inter-user interference cancellation, superposition coding in power-domain NOMA and user mapping to user-specific low density spreading sparse sequences in code-domain NOMA. Through research, NOMA outperforms OMA in many aspects including massive connections of devices, data throughput and spectral efficiency. However, these benefits come at the expense of increased computational complexity at the receiver.

Motivated by this perspective and further realize the benefits of NOMA, this work explores the possibility of extending the NOMA techniques to hybrid NOMA in order to alleviate the established performance challenges. In particular, the feasibility and potential of a hybrid NOMA that integrates PD-NOMA and SCMA in a heterogeneous multi-tier network is investigated. Such a hybrid NOMA system will come with its own challenges of resource (i.e., resource elements, codebooks and power) allocation, user clustering and multi-user detection computational complexity. Furthermore, this work seeks to answer the questions; what is the multiplexing capacity of such a hybrid NOMA system? Can we develop efficiently optimized resource allocation algorithms for such hybrid-NOMA for dynamic 5G and beyond networks? Are joint low complex MUD schemes appropriate for such technology feasible? And lastly, to enhance spectral efficiency and throughput, can we integrate the hybrid NOMA technique with multiple transceiver systems?

5.2 Research Objectives

The following identified objectives are the focus of this thesis:

1. To provide detailed literature review on NOMA and identify appropriate resource allocation

algorithm for NOMA schemes.

2. To develop and investigate sophisticated Hybrid-NOMA schemes build from efficiently integrating PD-NOMA with CD-NOMA schemes feasible for 5G networks.
3. To develop efficiently optimized resource allocation algorithms for such Hybrid-NOMA for dynamic 5G and beyond networks.
4. To investigate the multiplexing capacity bounds for the proposed Hybrid-NOMA technologies.
5. To integrate hybrid NOMA with existing MIMO technologies and examine the performance of such a system.

6 Thesis Overview and Contributions

This research work contributes to the wealth of knowledge in wireless communication networks by developing an uplink hybrid-NOMA technique in heterogeneous networks environment and proposing the associated hybrid resource allocation and multi-user detection schemes. In particular, we demonstrate the feasibility of a hybrid power-domain sparse code non-orthogonal multiple access (PD-SCMA) that integrates both power-domain non-orthogonal multiple access (PD-NOMA) and code-domain based sparse code multiple access (SCMA) in an uplink hierarchical HetNet system. Moreover, the associated resource allocation schemes, multi-user detection schemes and the multiplexing capacity bounds are summarised in sub-sections 6.1, 6.2 and 6.3 respectively. The research contributions are outlined in sub-section 6.4. The work covered in this thesis is detailed in paper A, paper B and paper C as presented in Part II, Part III and Part IV respectively. The conclusion of the work and suggestions for possible future research directions are highlighted in Part V.

6.1 Resource Allocation Schemes

The hybrid-NOMA technology come with their RA challenges: macro- and small-cell users (MUE and SUE) efficient power allocation (PA), SUEs optimal codebook assignment (CA), MUEs vs SUEs pairing interference management among others. In order to achieve an energy efficient (EE) PD-SCMA system, desirable and reliable resource (i.e., codebooks and power) assignment and user pairing (UP) and clustering schemes are necessary to retain the numerous interferences below the set temperature threshold and therefore maintain an acceptable QoS at a decent receiver computational complexity. Firstly, a PD-SCMA system where at most one MUE is paired with SUE in a codebook is considered and hybrid parametric based RA schemes are proposed. For such a scenario, a successive codebook ordering assignment (SCOA) scheme featuring channel quality metric

codebook ordering for codebook assignment to SUEs is proposed for the CA. For UP, a scheme featuring the differences in channel qualities and pairing interference metric for MUE-SUE pairing namely opportunistic MUE-SUE pairing (OMSP) is proposed. For PA, a QoS aware power allocation (QAPA) scheme that incorporates BS to user distance and QoS awareness metric is proposed. To achieve a stable RA combining the CA, UP and PA iteratively, a simulation-based joint energy efficiency (EE) resource allocation (JEERA) algorithm is then proposed. The EE performance of the simulation-based JEERA algorithm is validated by comparing with the proposed analytical RA based on modified dual decomposition (DDEEARA) optimization algorithm which employs a parametric transformation of the Dinkelbach method and Lagrange dual decomposition.

Secondly, we then extend the investigation to consider a PD-SCMA system where multiple MUEs are superimposed onto SUE assigned to the same codebook towards achieving near-optimal spectrum sharing for future heterogeneous networks. To alleviate the RA challenges of such a system at the transmitter, we decompose the RA problem into CA, UP and PA and propose dual parameter ranking (DPR) schemes based RA schemes namely; dual parameter ranking for CA (DPR-CA), dual parameter ranking for UP (DPR-UP) and dual parameter ranking for PA (DPR-PA). DPR-RA schemes utilizes ranking metrics proposed for the dual players that is, users and resources (code and power), to rank and contest for each other and perfectly match each other while employing a stable matching algorithm. We then propose an alternate search method (ASM) to iteratively admit the individual DPR-RA policies at the transmitter. In both scenarios, the proposed RA schemes result in enhanced sum achieved rate and system capacity. Besides, the RA schemes provide significantly improved codebook and pairing interference management, thereby enhancing the detection experience.

In the first two scenarios, we considered users and BSs with single antennas. This work extends the investigation to the application of multiple transceiver antennas. In particular, we employ spatial multiplexing based MIMO technique and integrate with the proposed hybrid PD-SCMA with the aim of striking a balance on the number of antennas, users and network capacity. Performance analysis show that employing MIMO on PD-SCMA exhibit reduced bit-error rates, enhanced reliability and spectrum efficiency, therefore validating the system.

6.2 Multi-User Detection schemes

The PD-SCMA system exhibits exponential multi-user detection (MUD) computational complexity occasioned by the non-linear detection techniques. The computational complexity monotonically rises as the number of assigned codebooks, paired MUEs, antennas and distinct power level increase.

This work proposes two hybrid low computational complexity MUD schemes. Firstly, a joint SIC and logarithmic message passing algorithm (Log-MPA) named SIC-Log-MPA receiver is proposed. The combined performance effect of the SIC-Log-MPA significantly minimizes the computational complexity order compared to the SIC-MPA. Log-MPA is capable of saving more than 40% of the required multiplications while completely eliminating the exponential operations confronted by the conventional MPA during the iterations. Secondly, we propose and employ a different MUD receiver based on modified expectation propagation algorithm (EPA) and SIC (J-EPA-SIC). Differently from MPA, modified EPA only pursues the means and variances of the transmitted messages during the iterative detection on the factor graph. The resultant effect is a reduced linearly varying computational complexity order unlike the exponential order exhibited with MPA. Both SIC-Log-MPA and J-EPA-SIC can be implemented with ease on condition that the point of operation is withing the multiplexing bound.

6.3 Multiplexing Capacity Bounds

The PD-SCMA system has various multiplexing dimensions namely; codebook multiplexing capacity, pairing multiplexing capacity and power multiplexing. By utilizing measures of sparsity and their attributes, we investigate and derive codebook capacity bounds, that gives the number of admissible codebooks in a PD-SCMA system for assignment to SUEs. In pairing capacity, this work explores the MUE multiplexing bounds in each codebook based on a proposed pairing metric. Basically, we investigate the upper bound on the number of MUEs that can be multiplexed with a SUE in an assigned codebook. Lastly, we investigate the maximal number of MUEs based on maximum number of discrete and distinct power levels to enable optimal detection at the receiver subject to outage performance and QoS constraints. It can be observed from analysis that the codebook capacity is bounded by the positional, overloading and the derived sparsity bounds. The pairing capacity is bounded by the proposed parameter ranking metric while the power capacity is bounded by the number of distinct power levels to guarantee SIC and minimum QoS.

6.4 Research Contributions

The research has resulted into the following papers;

Paper A:

S. Chege and T. Walingo, "Energy efficient resource allocation for uplink hybrid power domain sparse code non-orthogonal multiple access heterogeneous networks with statistical channel estimation," *Transactions on Emerging Tel Tech*. 2021; 32:e4185. https://doi.org/10.1002/ett.4185

The summary of this paper is as follows:

Modern 5G heterogeneous networks (HetNets) require hybrid multiple access technology for optimal performance. The feasibility of a hybrid power domain sparse code non orthogonal multiple access (PD-SCMA) that integrates both power domain non orthogonal multiple access (PD-NOMA) and sparse code multiple access (SCMA) for an uplink hierarchical HetNet system is demonstrated. Hybrid schemes namely: Successive Codebook Ordering Assignment (SCOA) for codebook assignment (CA), opportunistic MUE-SUE pairing (OMSP) for user pairing (UP), and a QoS-aware power allocation (QAPA) for power allocation (PA) are developed. The SCOA algorithm is based on channel quality ordering metric, OMSP algorithm is based on channel quality diversity and pairing interference metric while the QAPA algorithm features a QoS awareness metric. A joint energy efficiency (EE) resource allocation (JEERA) algorithm that iteratively performs CA, UP, and PA for small cell user equipment (SUE) and the macro user equipment (MUE) to limit interference, improve spectral and energy efficiency is presented. The problem is formulated as a mixed integer non-convex system EE resource allocation optimization for the small cells under QoS constraints of minimum sum-rate, interference temperature, maximum power, and SCMA structure for a hybrid low computational complexity joint SIC-Log-MPA receiver. A modified near-optimal dual decomposition analytical methodology featuring Dinkelbach fractional transformations is utilized to assess the system's performance on an imperfect wireless channel. Through numerical results, the proposed schemes are shown to improve the EE of the small cells in comparison with the prevalent schemes.

Paper B:

S. Chege and T. Walingo, "Multiplexing Capacity of Hybrid PD-SCMA Heterogeneous Networks," in IEEE Transactions on Vehicular Technology, vol. 71, no. 6, pp. 6424-6438, June 2022, doi: 10.1109/TVT.2022.3162304.

The summary of this paper is as follows:

Hybrid multiple access schemes are considered potential technologies towards achieving optimal spectrum sharing for future heterogeneous networks. Multiple users are multiplexed on a single resource unit in the code domain for sparse code multiple access (SCMA) or power domain for power domain non-orthogonal multiple access (PD-NOMA) and in both domains for the hybrid power domain sparse code non-orthogonal multiple access (PD-SCMA). This allows for effective spectrum usage but comes at a cost of increased detection computational complexity resulting in loss of performance in terms of outages. It is therefore imperative to determine the user multiplexing capacity for effective performance. This work investigates codebook capacity bounds in small cells,

pairing and power capacity bounds for the number of small cell user equipment's (SUEs) and macro cell user equipment's (MUEs) that can be multiplexed on a codebook for the developed PD-SCMA technology. Closed-form solutions for codebook, pairing and power multiplexing capacity bounds are derived. The performance of the system results into low outage when the system's point of operation is within the multiplexing bounds. To alleviate the resource allocation (RA) challenges of such a system at the transmitter, dual parameter ranking (DPR) and alternate search method (ASM) based RA schemes are proposed. The results show significant capacity gain with DPR-RA in comparison with conventional schemes.

Paper C:

S. Chege and T. Walingo, "MIMO based hybrid PD-SCMA Uplink Transceiver System," in *IEEE Communication Letters*, (*Under review*)

The summary of this paper is as follows:

The application of multiple-input multiple-output (MIMO) on power domain sparse code multiple access (PD-SCMA) system would enhance their performance by increasing the multiplexing and diversity gains at the cost of increased detection computational complexity as more users and antennas are deployed. This work develops and investigates the performance of spatial multiplexing MIMO based hybrid PD-SCMA system (M-PD-SCMA) transceiver on an uplink heterogeneous network with the aim of achieving a balance on the number of antennas and capacity/spectral efficiency. Numerical results exhibit performance benchmark with PD-SCMA schemes and the proposed receiver achieves guaranteed bit error rate (BER) performance with a bounded increase in the number of transmit and receive antennas. Thus, the feasibility of an M-PD-SCMA system is validated.

References

[1] 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*, vol. 20, no. 3, pp. 2294–2323, 2018.

[2] J. V. Lopa, "Evolution of mobile generation technology: 1g to 5g and review of upcoming wireless technology 5g," *International Journal of Modern Trends in Engineering and Research*, vol. 2, 2015.

[3] P. Kaveh and K. Prashant, *Principles of wireless networks: A unified approach*. Prentice Hall PTR, 2011.

[4] P. Nisha, S. Shantanu, and K. S. Awadhesh, "A survey on 5g: The next generation of mobile communication," *Physical Communication*, vol. 18, 2016.

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

[6] M. Mohammad and K. Sumit, "Evolution of mobile wireless technology from 0 g to 5 g," in *International journal of Computer Science and Information Technologies*, vol. 6, no. 3, 2015, p. 2545–2551.

[7] R. B. Mudit and B. Vardhan, "Generations of mobile wireless technology: A survey," *International Journal of Computer Applications*, vol. 5, no. 4, pp. 26–32, 2010.

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

[9] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5g wireless networks: A comprehensive survey," *IEEE Communications Surveys Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[10] M. António, M. S. H. Kazi, M. Shahid, and R. Jonathan, "A survey of 5g technologies: regulatory, standardization and industrial perspectives," *Digital Communications and Networks*, vol. 4, no. 2, pp. 187–97, 2018.

[11] The METIS Project, [Online]. Available: https://metis2020.com/.

[12] *Deliverable D6.3 : Intermediate system evaluation results*, ICT-317669 METIS, 2014, [Online]. Available: https://www.metis2020.com/documents/deliverables/.

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

[14] H. Droste, G. Zimmermann, M. Stamatelatos, N. Lindqvist, O. Bulakci, J. Eichinger, V. Venkatasubramanian, U. Dotsch, and H. Tullberg, "The metis 5g architecture: A summary of metis work on 5g architectures," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, 2015, pp. 1–5.

[15] K. G. Liolis, S. R. A, and et al, "Use cases and scenarios of 5g integrated satellite-terrestrial networks for enhanced mobile broadband: The sat5g approach," *Int J Satell Commun Network*, vol. 37, no. 2, p. 91–112, 2019.

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

[17] J. Ivan, F. Ivan, and P. Marko, "Massive machine-type communications: An overview and perspectives towards 5g," in *The 3rd International Virtual Research Conference In Technical Disciplines, RCITD*, vol. 3, 2015.

[18] 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.

[19] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5g be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 2014.

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

[21] *Minimum Requirements Related to Technical Performance for IMT-2020 Radio Interface(s), document ITU-R M. [IMT-2020.TECH PERF REQ]*, ITU, 2016.

[22] Qualcomm, "The 1000x mobile data challenge," White Paper, Nov 2018.

[23] NSN, "Signaling is growing 50% faster than data traffic," White Paper, 2012.

[24] Y. Kishiyama, A. Benjebbour, T. Nakamura, and H. Ishii, "Future steps of lte-a: evolution toward integration of local area and wide area systems," *IEEE Wireless Communications*, vol. 20, no. 1, pp. 12–18, 2013.

[25] Y. Zhou, L. Liu, H. Du, L. Tian, X. Wang, and J. Shi, "An overview on intercell interference management in mobile cellular networks: From 2g to 5g," in *2014 IEEE International Conference on Communication Systems*, 2014, pp. 217–221.

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

[27] E. Hossain and M. Hasan, "5g cellular: key enabling technologies and research challenges," *IEEE Instrumentation Measurement Magazine*, vol. 18, no. 3, pp. 11–21, 2015.

[28] W. Nam, D. Bai, J. Lee, and I. Kang, "Advanced interference management for 5g cellular networks," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 52–60, 2014.

[29] 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.

[30] P. F. Gerhard, "A 5g wireless communications vision," Microwave Journal, December 2012.

[31] D. Rico and P. Merino, "A survey of end-to-end solutions for reliable low-latency communications in 5g networks," *IEEE Access*, vol. 8, pp. 192 808–192 834, 2020.

[32] D. Xu, T. Li, Y. Li, X. Su, S. Tarkoma, T. Jiang, J. Crowcroft, and P. Hui, "Edge intelligence: Empowering intelligence to the edge of network," *Proceedings of the IEEE*, vol. 109, no. 11, pp. 1778–1837, 2021.

[33] M. Agiwal, H. Kwon, S. Park, and H. Jin, "A survey on 4g-5g dual connectivity: Road to 5g implementation," *IEEE Access*, vol. 9, pp. 16 193–16 210, 2021.

[34] F. Grijpink, M. Alexandre, S. Halldor, and V. Nemanja, "The road to 5g: The inevitable growth of infrastructure cost," TECH, MEDIA AND TELECOM PRACTICE, McKinsey Telecommunications, 2018. [Online]. Available: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/

[35] J. F. Valenzuela-Valdés, Palomares, J. C. González-Macías, A. Valenzuela-Valdés, P. Padilla, and F. Luna-Valero, "On the ultra-dense small cell deployment for 5g networks," in *2018 IEEE 5G World Forum (5GWF)*, 2018, pp. 369–372.

[36] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid beamforming for massive mimo: A survey," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 134–141, 2017.

[37] I. Chih-Lin, C. Rowell, S. Han, Z. Xu, G. Li, and Z. Pan, "Toward green and soft: a 5g perspective," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 66–73, 2014.

[38] G. Fodor, E. Dahlman, G. Mildh, S. Parkvall, N. Reider, G. Miklós, and Z. Turányi, "Design aspects of network assisted device-to-device communications," *IEEE Communications Magazine*, vol. 50, no. 3, pp. 170–177, 2012.

[39] T. Doumi, M. F. Dolan, S. Tatesh, A. Casati, G. Tsirtsis, K. Anchan, and D. Flore, "Lte for public safety networks," *IEEE Communications Magazine*, vol. 51, no. 2, pp. 106–112, 2013.

[40] S. A. R. Naqvi, H. Pervaiz, S. A. Hassan, L. Musavian, Q. Ni, M. A. Imran, X. Ge, and R. Tafazolli, "Energy-aware radio resource management in d2d-enabled multi-tier hetnets," *IEEE Access*, vol. 6, pp. 16 610–16 622, 2018.

[41] L. Liang, G. Y. Li, and W. Xu, "Resource allocation for d2d-enabled vehicular communications," *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 3186–3197, 2017.

[42] S. Počuča and D. Giljević, "Machine to machine (m2m) communication impacts on mobile network capacity and behaviour," in *2012 Proceedings of the 35th International Convention MIPRO*, 2012, pp. 607–611.

[43] S. G. Larew, T. A. Thomas, M. Cudak, and A. Ghosh, "Air interface design and ray tracing study for 5g millimeter wave communications," in *2013 IEEE globecom workshops (gc wkshps)*. IEEE, 2013, pp. 117–122.

[44] A. W. Scott and R. Frobenius, *Multiple Access Techniques: FDMA, TDMA, AND CDMA*, 2008, pp. 413–429.

[45] 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, 2010.

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

[47] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (noma) for cellular future radio access," in *2013 IEEE 77th vehicular technology conference (VTC Spring)*. IEEE, 2013, pp. 1–5.

[48] O. Maraqa, A. S. Rajasekaran, S. Al-Ahmadi, H. Yanikomeroglu, and S. M. Sait, "A survey of rate-optimal power domain noma with enabling technologies of future wireless networks," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2192–2235', year=.

[49] T. M. Cover, "Broadcast channels," *IEEE Transactions on information theory*, vol. 18, no. 1, pp. 2–14, 1972.

[50] C. Xu, M. Wu, Y. Xu, and Y. Fang, "Uplink low-power scheduling for delay-bounded industrial wireless networks based on imperfect power-domain noma," *IEEE Systems Journal*, vol. 14, no. 2, pp. 2443–2454, 2019.

[51] C.-L. Wang, J.-Y. Chen, and Y.-J. Chen, "Power allocation for a downlink non-orthogonal multiple access system," *IEEE wireless communications letters*, vol. 5, no. 5, pp. 532–535, 2016.

[52] M. S. Ali, H. Tabassum, and E. Hossain, "Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (noma) systems," *IEEE access*, vol. 4, pp. 6325–6343, 2016.

[53] 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.

[54] M.-R. Hojeij, C. A. Nour, J. Farah, and C. Douillard, "Waterfilling-based proportional fairness scheduler for downlink non-orthogonal multiple access," *IEEE Wireless Communications Letters*, vol. 6, no. 2, pp. 230–233, 2017.

[55] X. Wang, R. Chen, Y. Xu, and Q. Meng, "Low-complexity power allocation in noma systems with imperfect sic for maximizing weighted sum-rate," *IEEE Access*, vol. 7, pp. 94 238–94 253, 2019.

[56] Y. Iraqi and A. Al-Dweik, "Power allocation for reliable sic detection of rectangular qam-based noma systems," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 8355–8360, 2021.

[57] R. B. Shalini and S. L. Stuwart, "Power domain cyclic spread multiple access: An interference-resistant mixed noma strategy," *International Journal of communication systems*, vol. 32, no. 13, p. e4025, 2019.

[58] S. Islam, M. Zeng, and O. A. Dobre, "Noma in 5g systems: Exciting possibilities for enhancing spectral efficiency," *IEEE 5G Tech Focus*, vol. 1, no. 2, pp. 1–6, 2017.

[59] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous cdma systems over awgn channel," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1616–1626, 2008.

[60] R. Gallager, "Low-density parity-check codes," *IRE Transactions on information theory*, vol. 8, no. 1, pp. 21–28, 1962.

[61] J. Van De Beek and B. M. Popovic, "Multiple access with low-density signatures," in *GLOBECOM 2009-2009 IEEE Global Telecommunications Conference*, 2009, pp. 1–6.

[62] D. Guo and C.-C. Wang, "Multiuser detection of sparsely spread cdma," *IEEE journal on selected areas in communications*, vol. 26, no. 3, pp. 421–431, 2008.

[63] Y. Du, B. Dong, P. Gao, Z. Chen, J. Fang, and S. Wang, "Low-complexity lds-cdma detection based on dynamic factor graph," in *2016 IEEE Globecom Workshops (GC Wkshps)*, 2016, pp. 1–6.

[64] R. Hoshyar, R. Razavi, and M. Al-Imari, "Lds-ofdm an efficient multiple access technique," in *2010 IEEE 71st Vehicular Technology Conference*, 2010, pp. 1–5.

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

[66] A.-I. Mohammed, M. A. Imran, R. Tafazolli, and D. Chen, "Performance evaluation of low density spreading multiple access," in *2012 8th international wireless communications and mobile computing conference (IWCMC)*, 2012, pp. 383–388.

[67] M. Al-Imari, M. A. Imran, R. Tafazolli, and D. Chen, "Subcarrier and power allocation for lds-ofdm system," in *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, 2011, pp. 1–5.

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

[69] Y. Wu, C. Wang, Y. Chen, and A. Bayesteh, "Sparse code multiple access for 5g radio transmission," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, 2017, pp. 1–6.

[70] 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 *2014 IEEE Global Communications Conference*, 2014, pp. 4782–4787.

[71] X. Meng, Y. Wu, Y. Chen, and M. Cheng, "Low complexity receiver for uplink scma system via expectation propagation," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, 2017, pp. 1–5.

[72] B. Ghani, F. Launay, J. P. Cances, C. Perrine, and Y. Pousset, "Iterative decoding for scma systems using log-mpa with feedback ldpc decoding," in *International Symposium on Ubiquitous Networking*, 2019, pp. 18–31.

[73] M. Alam and Q. Zhang, "Performance study of scma codebook design," in *2017 IEEE wireless communications and networking conference (WCNC)*, 2017, pp. 1–5.

[74] 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)*, 2016, pp. 1–5.

[75] S. Liu, J. Wang, J. Bao, and C. Liu, "Optimized scma codebook design by qam constellation segmentation with maximized med," *IEEE Access*, vol. 6, pp. 63 232–63 242, 2018.

[76] X. Zhang, G. Han, D. Zhang, D. Zhang, and L. Yang, "An efficient scma codebook design based on lattice theory for information-centric iot," *IEEE Access*, vol. 7, pp. 133 865–133 875, 2019.

[77] Z. Mheich, L. Wen, P. Xiao, and A. Maaref, "Design of scma codebooks based on golden angle modulation," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1501–1509, 2018.

[78] Z. Hou, Z. Xiang, P. Ren, and B. Cao, "Scma codebook design based on divided extended mother codebook," *IEEE Access*, vol. 9, pp. 71 563–71 576, 2021.

[79] N. M. Balasubramanya, S. Payami, and M. Sellathurai, "Uplink resource allocation for shared lte and scma iot systems," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, 2018, pp. 1–5.

[80] T. Sefako and T. Walingo, "Application of biological resource allocation techniques to scma noma networks," in *2019 IEEE AFRICON*, 2019, pp. 1–7.

[81] W. Zhu, L. Qiu, and Z. Chen, "Joint subcarrier assignment and power allocation in downlink scma systems," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, 2017, pp. 1–5.

[82] A. Jehan, M. Zeeshan, and T. Ashraf, "Parametric analysis of scma performance with different codebooks and orthogonal resources," in *2021 International Conference on Robotics and Automation in Industry (ICRAI)*, 2021, pp. 1–5.

[83] Y. Du, B. Dong, Z. Chen, P. Gao, and J. Fang, "Joint sparse graph-detector design for downlink mimo-scma systems," *IEEE Wireless Communications Letters*, vol. 6, no. 1, pp. 14–17, 2016.

[84] S. Tang, L. Hao, and Z. Ma, "Low complexity joint mpa detection for downlink mimo-scma," in *2016 IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1–4.

[85] 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, 2017.

[86] S. Chege and T. Walingo, "Energy efficient resource allocation for uplink hybrid power domain sparse code nonorthogonal multiple access heterogeneous networks with statistical channel estimation," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 1, p. e4185, 2021.

[87] Z. Yuan, C. Yan, Y. Yuan, and W. Li, "Blind multiple user detection for grant-free musa without reference signal," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, 2017, pp. 1–5.

[88] Y. Tao, L. Liu, S. Liu, and Z. Zhang, "A survey: Several technologies of non-orthogonal transmission for 5g," *China communications*, vol. 12, no. 10, pp. 1–15, 2015.

[89] E. M. Eid, M. M. Fouda, A. S. T. Eldien, and M. M. Tantawy, "Performance analysis of musa with different spreading codes using ordered sic methods," in *2017 12th international conference on computer engineering and systems (ICCES)*, 2017, pp. 101–106.

[90] W. B. Ameur, P. Mary, M. Dumay, J.-F. Hélard, and J. Schwoerer, "Power allocation for ber minimization in an uplink musa scenario," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, 2020, pp. 1–5.

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

[92] Y. Ma, Z. Yuan, Y. Hu, and W. Li, "A real fourier-related transform spreading ofdm multi-user shared access system," in *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, 2019, pp. 1–5.

[93] T. Sivalingam, S. Ali, N. H. Mahmood, N. Rajatheva, and M. Latva-Aho, "Deep neural network-based blind multiple user detection for grant-free multi-user shared access," in *2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2021, pp. 1–7.

[94] 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.

[95] S. Chen, X. Dai, and S. Sun, "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.

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

[97] X. Bian, J. Tang, H. Wang, M. Li, and R. Song, "An uplink transmission scheme for pattern division multiple access based on dft spread generalized multi-carrier modulation," *IEEE Access*, vol. 6, 2018.

[98] Y. Jiang, P. Li, Z. Ding, F.-C. Zheng, M. Ma, and X. You, "Joint transmitter and receiver design for pattern division multiple access," *IEEE Transactions on Mobile Computing*, vol. 18, no. 4, pp. 885–895, 2019.

[99] H. Bizaki, Ed., *Non-Orthogonal Multiple Access (NOMA) for 5G Networks*, ser. Towards 5G Wireless Networks - A Physical Layer Perspective. London, United Kingdom: IntechOpen, 2016.

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

[101] K. Kusume, G. Bauch, and W. Utschick, "Idma vs. cdma: Analysis and comparison of two multiple access schemes," *IEEE Transactions on Wireless Communications*, vol. 11, no. 1, pp. 78–87, 2012.

[102] L. Ping, L. Liu, K. Wu, and W. Leung, "Interleave-division multiple-access (idma) communications," in *Proc. 3rd International Symposium on Turbo Codes and Related Topics*, 2003, pp. 173–180.

[103] X. Dai, S. Chen, S. Sun, S. Kang, Y. Wang, Z. Shen, and J. Xu, "Successive interference cancelation amenable multiple access (sama) for future wireless communications," in *2014 IEEE International Conference on Communication Systems*, 2014, pp. 222–226.

[104] 3GPP, "Low code rate and signature based multiple access scheme for new radio," TSG RAN1 #85, Nanjing, China,, 23rd-27th, May. 2016.

[105] .3GPP, "Discussion on multiple access for new radio interface," TSG RAN1 WG #84, Busan,Korea, 11th-15th, April. 2016.

[106] J. Huang, K. Peng, C. Pan, F. Yang, and H. Jin, "Scalable video broadcasting using bit division multiplexing," *IEEE Transactions on Broadcasting*, vol. 60, no. 4, pp. 701–706, 2014.

[107] B. Wang, L. Dai, Y. Yuan, and Z. Wang, "Compressive sensing based multi-user detection for uplink grant-free non-orthogonal multiple access," in *2015 IEEE 82nd Vehicular Technology Conference (VTC2015-Fall)*, 2015, pp. 1–5.

[108] 3GPP, "Initial lls results for ul non-orthogonal multiple access," TSG RAN1 WG1 #85, Nanjing, China, 23rd-27th, May 2016.

[109] .3GPP, "Multiple access schemes for new radio interface," TSG RAN1 WG 84, Busan,Korea, 11th-15th, April 2016.

[110] 3GPP, "New uplink non-orthogonal multiple access schemes for nr," TSG RAN1 WG1 86, Gothenburg, Sweden, 22nd-26th, August 2016.

Part II

Paper A

Paper A

Energy Efficient Resource Allocation for Uplink Hybrid Power Domain Sparse Code Non Orthogonal Multiple Access Heterogeneous Networks with Statistical Channel Estimation

Simon Chege and Tom Walingo

Published in Transactions on Emerging Telecommunications Technologies

https://doi.org/10.1002/ett.4185

Abstract

Modern 5G heterogeneous networks (HetNets) require hybrid multiple access technology for optimal performance. The feasibility of a hybrid power domain sparse code non orthogonal multiple access (PD-SCMA) that integrates both power domain non orthogonal multiple access (PD-NOMA) and sparse code multiple access (SCMA) for an uplink hierarchical HetNet system is demonstrated. Hybrid schemes namely: Successive Codebook Ordering Assignment (SCOA) for codebook assignment (CA), opportunistic MUE-SUE pairing (OMSP) for user pairing (UP), and a QoS-aware power allocation (QAPA) for power allocation (PA) are developed. The SCOA algorithm is based on channel quality ordering metric, OMSP algorithm is based on channel quality diversity and pairing interference metric while the QAPA algorithm features a QoS awareness metric. A joint energy efficiency (EE) resource allocation (JEERA) algorithm that iteratively performs CA, UP, and PA for small cell user equipment (SUE) and the macro user equipment (MUE) to limit interference, improve spectral and energy efficiency is presented. The problem is formulated as a mixed integer non-convex system EE resource allocation optimization for the small cells under QoS constraints of minimum sum-rate, interference temperature, maximum power, and SCMA structure for a hybrid low complexity joint SIC-Log-MPA receiver. A modified near-optimal dual decomposition analytical methodology featuring Dinkelbach fractional transformations is utilized to assess the system's performance on an imperfect wireless channel. Through numerical results, the proposed schemes are shown to improve the EE of the small cells in comparison with the prevalent schemes.

1 Introduction

The demand for increased capacity in 5G networks has fueled a paradigm shift from orthogonal multiple access (OMA) to Non-Orthogonal Multiple Access (NOMA) techniques. Unlike OMA techniques that allow only one user to use a resource unit (RU), orthogonal frequency division multiple access (OFDMA) subcarrier, time division multiple access (TDMA) timeslot or code division multiple access (CDMA)'s code, NOMA allows multiplexing of several users on the same RU resulting in high spectral efficiency, user fairness and massive connectivity [1]. NOMA schemes permit controlled interference by non-orthogonal resource allocation. The interference and RU limitation constraints have necessitated the development of new combined optimal radio resource allocation (RA), user pairing (UP) and power allocation (PA) schemes for these networks.

Generally, NOMA access schemes are classified into Interleaver Division Multiple Access (IDMA), power and code domain multiplexing [2]. IDMA utilizes user specific interleavers for multiplexing [3] while Power domain NOMA (PD-NOMA) [4] superimposes multiple users by allocating distinct power levels to different users at the transmitter. In code domain NOMA (CD-NOMA) [5], multiple users share the same time-frequency resources by mapping incoming bits to unique user-specific sparse, non-orthogonal low cross-correlation spreading sequences. In addition, different categories of CD-NOMA have been applied on wireless networks: Sparse code multiple access (SCMA) [6], [7] and low-density spreading multiple access (LDSMA) [8] are based on the idea of user information being spread over multiple sub-carriers using multi-dimensional complex codewords. Multi-user shared access (MUSA) [9], an enhanced CDMA-style scheme, employs spreading sequences characterized by low-correlation to support overloading at the transmitter and facilitate a near optimal interference cancellation at the receiver side. Finally, Pattern Division Multiple Access (PDMA) [10], [11], which uses non-orthogonal patterns designed through maximization diversity and minimization of multiple users' correlation. These techniques employ multi-user detection (MUD) techniques for separation of transmitted user symbols at the receiver. Consequently, NOMA allows considerable controlled interference due to the non-orthogonal resource sharing at the expense of magnified complexity at the receiver.

Due to their heterogeneous nature and performance requirements, future networks will require to blend multiple access schemes resulting in hybrid NOMA schemes. At the core of PD-NOMA is the successive interference cancellation (SIC) technique that makes it possible to assign at least one user to a single RU by cancelling the interference resulting from the non-exclusive employment of the RUs. On the contrary, SCMA scheme employs codebooks which are generated from a

multidimensional constellation resulting in constellation shaping gain [5]. The SCMA decoder employs an iterative message passing algorithm (MPA) to effectively cancel interference and improve the decoding performance at the expense of increased computational complexity. In particular, a combination of PD-NOMA and SCMA is feasible on a heterogeneous multitier network consisting of small cell user equipment (SUE) and the macro user equipment (MUE). This work considers multiplexing MUEs and SUEs using PD-NOMA into the codebook while allocating different power levels and hence apply SIC for their separation at the receiver. On the contrary, different codes are exclusively assigned to the SUE's in the small cells hence employing MPA for their separation. The receiver features a low-complexity iterative hybrid MUD scheme based on SIC and log-domain MPA (Log-MPA). This results into a hybrid power domain sparse code non-orthogonal multiple access (PD-SCMA) scheme that combines the benefits of the two different access schemes to increase the networks capacity.

The hybrid PD-SCMA multi-tier heterogeneous network (HetNet) schemes come with their own challenges: MUE vs SUE power allocation (PA), SUEs vs SUEs codebook assignment (CA), MUEs vs SUEs pairing interference among others. For the system to be effective, proper resource (i.e., codebooks and power) allocation and user pairing (UP) schemes are required to keep the multiple interferences at the required levels to maintain acceptable quality of service (QoS) at decent receiver complexity. Only then we can realize the benefits of applying hybrid PD-SCMA scheme on a HetNet. Furthermore, the robustness and effectiveness of the developed system should be investigated on a realistic imperfect channel with proper imperfect channel models due to numerous co-channels, different degrees of fading and transmission delays. Analytical evaluation of the complex system featuring multitier HetNets with both SUEs and MUEs employing multiple QoS schemes (CA, PA, UP) is not trivial and cannot be easily evaluated by the exhaustive search methods. This work develops an iterative joint energy efficiency (EE) resource allocation algorithm that optimizes the performance of such a system. By decomposing the original problem into codebook assignment, user pairing and power allocation sub-problems, we propose tractable and practical EE resource allocation schemes.

1.1 Related Work

NOMA schemes have received significant attention in the last decade as potential multiple access schemes for 5G technologies. In [12], a PD-NOMA scheme that clusters users based on their channel quality differences is proposed. The scheme performs optimal PA that maximizes the sum-throughput of all users. In [13], a PA scheme based on the sum of normalized rates for two users

with a possible extension to more users is proposed. In [14], various fair ordering methods have been investigated for tradeoff among the energy efficiency, fairness, harvested energy and system sum rate in PD-NOMA based HetNets. Resource allocation in downlink HetNet NOMA for two scenarios namely, exclusive subcarrier allocation case and intercell interference case for sum-rate maximization are investigated in [15]. The proposed scheme decouples the optimization problem into RU allocation and PA subproblems, which are solved by using the mesh adaptive direct search algorithm and different successive convex approximation methods, respectively. In [16], different SCMA configurations are employed to explore performance aspects for edge Internet of Things (IOT) systems. A scenario where the network operator shares the uplink spectrum between the original long-term evolution (LTE) users and new users using SCMA is investigated in [17]. A comparative study between SCMA and PD-NOMA schemes has been presented in [18] indicating that the former scheme achieved better performance than the latter at the expense of increased computational complexity. In summary, a technical analysis has been reported in [19] and [20] on the feasibility, development, performance improvement and challenges of generalized NOMA (G-NOMA) techniques over the traditional OMA. On the contrary, the same has not been done for an uplink hybrid HetNet with optimal resource allocation.

One of the challenges arising for SCMA is codebook resource assignment, becoming more adverse in HetNets [21]. Dynamic CA methods proposed in [22] utilize the available channel estimation information (CSI) to achieve user fairness and performance in a single cell network system. By decoupling CA and PA, cost efficient EE optimization schemes are proposed for a single cell downlink network in [23]. Specifically, a novel CA scheme employing equal power distribution is developed. Furthermore, the authors exploit the quasi-concavity of the resultant PA problem to devise an optimal derivative-bisection algorithm. In [24], a three-step iterative joint CA and PA algorithm is proposed for the uplink sum-rate optimization problem. A specific codebook design method which can cancel the effect of the dependency between the non-zero entries of codewords is presented. Consequently, the objective function becomes a concave function in which convex optimization is employed to obtain optimal PA. In [25], an iterative algorithm that jointly performs CA and PA decoupled sub-problems in a SCMA-HetNet system is proposed. Since PA is a non-convex sub-problem, a dual approach considering successive convex approximation (SCA) is employed, while CA is in integer linear programing (ILP) form and therefore solved using ILP softwares. Energy saving oriented codebook resource management schemes for both uplink and downlink SCMA networks are proposed in [26]. Specifically, for the uplink networks, a dual coordinate search approach is adopted to determine the mapping matrix and CA for minimal detection complexity. For the downlink networks, codebook design and assignment and PA are jointly solved for minimal total

power consumption. A Lagrangian dual decomposition technique is employed to propose a fast but low complex iterative algorithm. In a downlink single-cell based PSMA system, joint resource allocation and SIC ordering based on matching game and submodularity approach for sum-rate maximization is proposed [27]. Herein, several SUEs are assigned a CB resource at the transmitter while SIC ordering and detection employed at the receiver. Particularly, users' CA is modelled, first, as a many-to-many matching game where players' preference lists are determined based on the average channel gains. To solve the matching game, a Gale-Shapley algorithm is employed. Secondly, the CA is modelled as a submodular optimization problem and an iterative difference of submodular function (DSF) algorithm is proposed. The PA sub-problem employs a SCA-based difference of convex functions (DC) method. This work forms part of our comparison algorithms.

Research on SCMA based Device-to-Device (D2D) underlay communications systems is ongoing. In [28], an interference aware hypergraph CA algorithm that realizes full exploitation of the available CBs is developed. Here, a codebook can be occupied by one cellular user and more than one D2D links. [29] investigates SCMA sum rate maximization by assigning codewords of the codebook to the cellular user based on the lower bound of the achievable user rate. An opportunistic scheduling approach where cellular users are allocated CBs based on the available CSI is developed in [30]. A D2D enabled cellular hybrid network based on SCMA and efficient resource management schemes are proposed in [31]. Particularly in a downlink system, CBs are assigned cellular users (CUs) by initially assuming equal PA, interference constraint, minimum data rate constraint. The D2D admission for CB re-use and pairing is based on a proposed conflict graph featuring rate aware codebook selection for the D2D system (RACBS-D2D). Lastly, PA is performed using the geometric water filling (GWF) approach.This work is also useful in our comparisons. In the above reviewed contributions in this subject, none is applied in an uplink PD-SCMA HetNet model. In constrast, we develop a root mean square (RMS) channel quality metric codebook assignment that takes into account all the RUs appended in a codebook resulting in fast assignment which consequently improves the sum rate performance at the SBS.

The development of hybrid NOMA systems is gaining traction for HetNets. A hybrid scheme that combines topological interference management (TIM) and NOMA in a two-stage decoding process is introduced in [32]. TIM is applied to manage the inter-group interference while NOMA is employed for intra-group interference through SIC. Multi-carrier NOMA (Mc-NOMA), a variation of hybrid NOMA where the users in a network are divided into multiple groups and user with particular groups served in the same RU to realize overloading is developed in [33]. In [34], a hybrid HetNet scheme is proposed where NOMA is applied in the small cells and massive multiple-input multiple-output

(MIMO) employed in macro cells. An uplink hybrid nonorthogonal multiple access ($\boldsymbol{h}-$NOMA) scheme utilizing power domain multiplexing for OFDM-based systems is proposed in [35]. $\boldsymbol{h}-$NOMA employs precoding techniques to address the high peak to average power ratio (PAPR) and minimum mean square error (MMSE) receiver for low interference, low-complexity MUD and consistent fairness performance. A new power domain sparse code multiple access (PSMA) is proposed in [36]. In PSMA, multiple SUEs can reuse a codebook more than once in each cell where for each user PD-NOMA is employed to transmit non-orthogonally. Due to unrealistic network performance bounds occasioned by common analytical techniques, [37] proposes alternative biological resource allocation schemes that include: the ant colony optimization, particle swarm optimization and a hybrid adaptive particle swarm optimization algorithm for a hybrid NOMA system. Clearly, a combination of power and code domain NOMA is more feasible on a multi-tier multi-user network than interleaver and any of the two domains due to interleaver's user power imbalance scenario limitations [4]. In contrast to the previous works, this paper develops a hybrid PD-SCMA HetNet through exclusive CA to SUE in the small cells and codebook re-use for MUEs in the macro-cell subject to an interference threshold.

User pairing (UP) is important for optimal resource allocation in NOMA HetNet systems to avoid pairing interference from MUE-SUE pair transmitting on the same codebook simultaneously. In [38], a game-theoretical scheme using EE resource allocation and interference pricing for interference limited HetNet is proposed. A modified swap matching algorithm based on stable matching theory is proposed in [39] to allocate the spectrum resource. In [40], UP in uplink NOMA where the BS divides the set of users into disjunct pairs and assigns the available resources to these pairs is developed. UP schemes based on Channel Quality Indicator (CQI) for sum capacity maximization over uplink NOMA system are proposed for both perfect and imperfect SIC in [41]. A joint UP and PA algorithm for uplink NOMA aimed at enhancing the proportional fairness of the users is proposed in [42]. In [43], a joint UP and access point assignment problem evaluated for near far users and far users in D2D communication enabled dense HetNet is developed. The impacts of UP on the performance of two NOMA systems, NOMA with fixed PA and cognitive radio inspired NOMA is characterized in [44]. Recent research has focused on addressing UP in the current LTE-A schemes and homogeneous PD-NOMA scenarios. This has not been implemented in a hybrid HetNet with SUEs and MUEs. An attempt to pair the MUEs to SUE assigned codebooks via a channel quality difference and minimal pairing interference metric is proposed in this work. This is observed to improve SIC decoding at the SBS due to minimal pairing interference contamination and adverse channel quality difference amongst paired users.

In summary, this work demonstrates the feasibility of an uplink hybrid PD-SCMA HetNet with SUEs and MUEs on a stochastic channel model in maximizing small cells' EE while maintaining QoS at the MBS, featuring the following:

- An uplink hybrid PD-SCMA HetNet model that integrates PD-NOMA and SCMA. Multiplexing between the MUEs and SUEs is done in the power domain with PD-NOMA while SUEs multiplexing is done in the code domain with SCMA. The developed receiver features a novel low-complexity MUD scheme based on joint SIC and Log-MPA. Unlike the receiver in [36], the complexity order of the proposed MUD reduces significantly for the simplified sub-optimal version of MPA.
- Hybrid CA, UP and PA schemes for SUEs and MUEs. The hybrid successive codebook ordering assignment (SCOA) for CA scheme features the channel quality metric ordering of codebooks for allocation to MUEs and SUEs. The hybrid opportunistic MUE-SUE pairing (OMSP) for UP features the differences in channel qualities and pairing interference metric for MUE-SUE pairing while the hybrid QoS aware power allocation (QAPA) for PA incorporates distance and QoS awareness metric. A practical iterative joint EE resource allocation (JEERA) algorithm that combines CA, UP and PA is then presented. These hybrid schemes greatly limit network interference, improve spectral and energy efficiency.
- A modified dual decomposition-based EE analytical resource allocation (DDEEARA) optimization algorithm employing a parametric transformation from the Dinkelbach method and Lagrange dual decomposition for system EE performance evaluation is proposed. Even though this approach has been used extensively in literature [36], [45], [46], its application to multiple parameters (CA, UP, PA) is not trivial since it has been developed and results generated.
- Comparative results between the simulation-based JEERA algorithm and the analytical-based DDEEARA algorithm,the proposed CA, UP and PA schemes versus resource allocation schemes of [27] and [31] and lastly, PD-SCMA versus SCMA and PD-NOMA. Results show significant energy efficiency performance improvement of the proposed PD-SCMA system over the conventional SCMA and PD-NOMA.

1.2 Organization

The remainder of the paper is organized as follows: the system and receiver models besides problem formulation are developed in Section 2. In Section 3, details of the proposed joint EE resource

Table A.1: Notation and Variables

Notation	Meaning	Notation	Meaning
C	Codebooks	$y_{f,k,c}^{SUE}$	Received signal at SBS
J	Codebook size	η_{EE}	Energy efficiency
K	SUEs in a SBS	P_T	Total transmit power
M	MUEs in the MBS	P_c	Circuit power
B	System bandwidth	R_{min}^k	Minimum SUE rate requirement
N	Resource units (RUs)	$\mathcal{I}_{k\to m}^c$	Pairing Interference (SUEs to MBS)
$\mathcal{M}_c$	Set of MUEs paired on codebook c	$\mathcal{I}_{m\to k}^c$	Pairing Interference (MUEs to SBS)
B_{ru}	RU bandwidth	σ_ε^2	Estimation error variance
F	SBSs/Small cells	d_u	No. of users in a RU
$\star_{f,k,c}^{SUE}$	Relation of the k^{th} SUE on codebook c of the f^{th} SBS	L	No. of users in a codebook
$\star_{m,c}^{MUE}$	Relation of m^{th} MUE on codebook c	d_v	No. of RUs in a codecook
$q_{f,k,c}^{SUE}, q_{m,c}^{MUE}$	Codebook assignment policies	$\sigma_{f,k,c}^2$	Noise signal variance
$A_{k,m}^c$	User pairing policy	$z_{f,k,c}$	AWGN signal
$P_{f,k,c}^{SUE}, P_{m,c}^{MUE}$	Power allocation policies	$r_{f,k,c}^{SUE}, r_{m,c}^{MUE}$	Achievable user sum rate
$h_{f,k,c}^{SUE}, h_{m,c}^{MUE}$	Channel fading gains	ρ	Resource mapping matrix
$\varepsilon_{f,k,c}^{SUE}, \varepsilon_{m,c}^{MUE}$	Channel estimation errors	R	Achivable total system sum rate
σ_ε^2	Channel error variance	G	Equivalent channel matrix
$x_{f,k,c}^{SUE}$, $x_{m,c}^{MUE}$	Transmitted signals	$\gamma_{f,k,c}^{SUE}, \gamma_{m,c}^{MUE}$	Signal to Interference and Noise Ratio

allocation (JEERA) algorithm and the individual schemes are provided. Section 4 outlines the dual decomposition-based EE analytical resource allocation (DDEEARA) optimization algorithm. In Section 5, the computational complexity of the proposed schemes is discussed with simulation results and performance evaluation discussed in Section 6. Finally, the concluding remarks are presented in Section 7.

1.3 Notation

The following notations are adopted in this work. The superscripts $(\cdot)^H$, $(\cdot)^{-1}$ and $(\cdot)^T$ represents Hermitian, inverse and transpose operators respectively. $|\cdot|$ and $\|\cdot\|$ denotes the absolute value of a scalar and the Euclidean norm. $\odot$ indicates the element-wise dot product of two matrices while $\ominus$

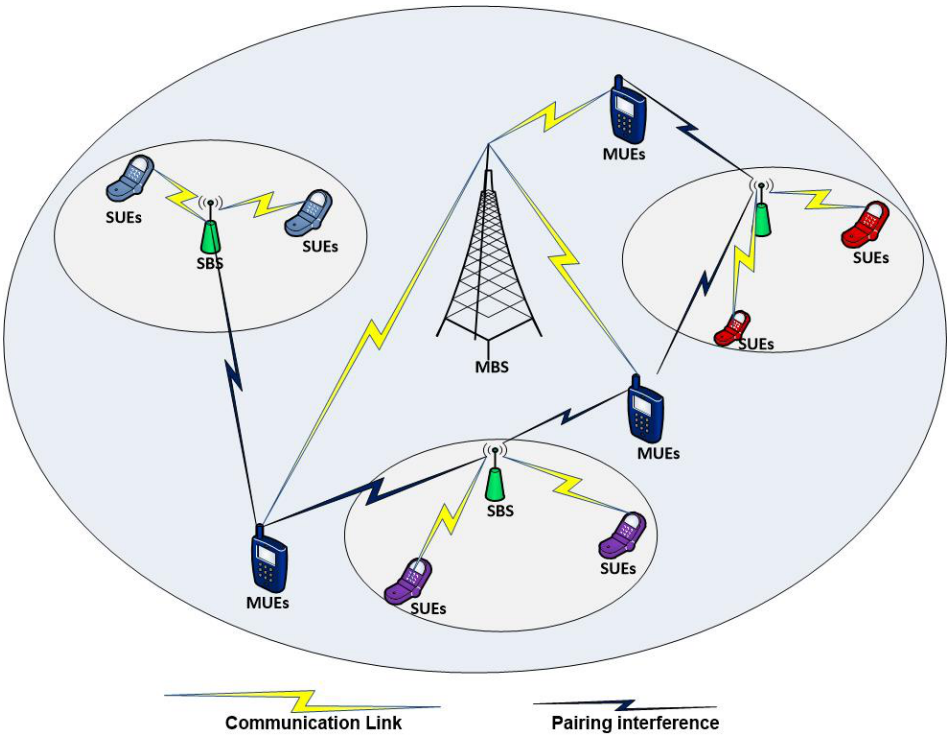

Fig. A.1: Uplink hybrid PD-SCMA HetNet model

evaluates the difference of every value in a $N \times 1$ vector from every value in a $1 \times N$ vector to form a $N \times N$ matrix of the differences. Without loss of generality, the notation $\bullet_{f,k,c}^{SUE}$ will be used to denote the relation between k^{th} SUE on codebook c of the f^{th} SBS, and $\bullet_{m,c}^{MUE}$, the relation between m^{th} MUE on codebook c. The summary list of all notations and variables is given in Table A.1.

2 SYSTEM MODEL AND PROBLEM FORMULATION

2.1 System Model

The proposed network model is an uplink hybrid PD-SCMA based two-tier wireless HetNet with a centralized macro base station (MBS) serving M randomly distributed MUEs as shown in Fig. A.1. The macro cell is overlaid by F small cells, each characterized by a centralized low power small cell base station (SBS) serving K uniformly distributed SUEs. MUEs and SUEs are co-multiplexed on the same time-frequency RU for enhanced spectral efficiency in an imperfect channel estimation environment. The MUEs, SUEs, and SBS are assumed to be equipped with a single antenna. The total network bandwidth B is equally shared by N RUs with each RU of bandwidth $B_{ru} = B/N$. The following policies are defined:

- The PA policy $\mathbf{P} = \left\{P_{F,K,C}^{SUE}, P_{M,C}^{MUE}\right\}$ is such that the transmitter allocates $P_{F,K,C}^{SUE} = \left[P_{f,k,c}^{SUE}\right]_{F\times K\times C}$ power to the k^{th} SUE on codebook c of the f^{th} small cell and

$P^{MUE}_{M,C} = \left[P^{MUE}_{m,c}\right]_{M\times C}$ to the m^{th} MUE utilizing the same codebook c. Note that $P^{MUE}_{m,c} >> P^{SUE}_{f,k,c}$.

- The CA policy $\mathbf{Q} = \left\{Q^{MUE}_{M,C}, Q^{SUE}_{F,K,C}\right\}$ where $Q^{MUE}_{M,C} = \left[q^{MUE}_{m,c}\right]_{M\times C}$ and $Q^{SUE}_{F,K,C} = \left[q^{SUE}_{f,k,c}\right]_{F\times K\times C}$ denote the macro cell and small cells transmitter codebook assignment matrix respectively. Also, $\left[q^{MUE}_{m,c}\right] = 1$ implies the m^{th} MUE is assigned codebook c, otherwise $\left[q^{MUE}_{m,c}\right] = 0$. Similarly, $\left[q^{SUE}_{f,k,c}\right] = 1$ means that codebook c is assigned to the k^{th} SUE in the f^{th} small cell.

- The UP policy $\mathbf{A} = \left[A^{c}_{k,m}\right]_{K\times M}$, where $A^{c}_{k,m} = 1$ denotes that the k^{th} SUE is paired with the m^{th} MUE on codebook c, while $A^{c}_{k,m} = 0$, denotes otherwise. The policy matrix $\mathbf{A}$ pairs $Q^{MUE}_{M,C}$ with $Q^{SUE}_{F,K,C}$.

Let $\mathbf{H} = \left\{H^{MUE}_{M,C}, H^{SUE}_{F,K,C}\right\}$, where $H^{MUE}_{M,C} = \left[h^{MUE}_{m,c}\right]_{M\times C}$ and $H^{SUE}_{F,K,C} = \left[h^{SUE}_{f,k,c}\right]_{F\times K\times C}$ denote channel-fading gains of the m^{th} MUE and k^{th} SUE in the f^{th} small cell on codebook c. The statistical channel coefficients are modelled as:

$$h^{MUE}_{m,c} = \hat{h}^{MUE}_{m,c} + \varepsilon^{MUE}_{m,c}, \tag{A.1}$$

$$h^{SUE}_{f,k,c} = \hat{h}^{SUE}_{f,k,c} + \varepsilon^{SUE}_{f,k,c}, \tag{A.2}$$

where $\hat{h}^{MUE}_{m,c} \sim \mathcal{CN}\left(0, \hat{\sigma}^2_{m,c}\right)$ and $\hat{h}^{SUE}_{f,k,c} \sim \mathcal{CN}\left(0, \hat{\sigma}^2_{f,k,c}\right)$ are the estimated channel gains with variances $\hat{\sigma}^2_{m,c}$ and $\hat{\sigma}^2_{f,k,c}$ respectively. $\varepsilon^{MUE}_{m,c} \sim \mathcal{CN}\left(0, \sigma^2_{\varepsilon}\right)$ and $\varepsilon^{SUE}_{f,k,c} \sim \mathcal{CN}\left(0, \sigma^2_{\varepsilon}\right)$ denote the channel estimation errors with variance σ^2_{ε} . It is assumed that estimated channel gains and errors are uncorrelated stationary and ergodic random processes.

Considering that each codebook is exclusively assigned to a single SUE at each SBS but re-used by MUEs in the macro-cell, the received signal $y^{SUE}_{f,k,c}$ is given by:

$$y^{SUE}_{f,k,c} = \underbrace{q^{SUE}_{f,k,c} h^{SUE}_{f,k,c} \sqrt{P^{SUE}_{f,k,c}} x^{SUE}_{f,k,c}}_{Desired Signal} + \underbrace{\sum_{m\in\mathcal{M}_c} A^{c}_{k,m} h^{MUE}_{m,c} \sqrt{P^{MUE}_{m,c}} x^{MUE}_{m,c}}_{Pairing Interference from MUEs} + z_{f,k,c}, \tag{A.3}$$

where $x^{SUE}_{f,k,c}$ and $x^{MUE}_{m,c}$ are the transmitted signals of the SUE and MUE on codebook c, respectively, $\mathcal{M}_c$ denotes the set of MUEs paired with SUE k on codebook c and $z_{f,k,c} \sim \mathcal{CN}\left(0, \sigma^2_{f,k,c}\right)$ is the additive white Gaussian noise power (AWGN). Note that users on different codebooks do not interfere with each other and we assume negligible inter-cell interference

due to same codebook re-used in different small-cells. The achievable throughput at the MBS, $r_{m,c}^{MUE}$ and at each SBS, $r_{f,k,c}^{SUE}$ are respectively given by:

$$r_{m,c}^{MUE} = B_{ru} \log_2 \left(1 + \gamma_{m,c}^{MUE}\right), \tag{A.4}$$

$$r_{f,k,c}^{SUE} = B_{ru} \log_2 \left(1 + \gamma_{f,k,c}^{SUE}\right), \tag{A.5}$$

where $\gamma_{m,c}^{MUE}$ and $\gamma_{f,k,c}^{SUE}$ are the signal to interference and noise ratio (SINR) of the m^{th} MUE and k^{th} SUE in the f^{th} small cell on codebook c respectively, given by (A.6) and (A.7).

$$\gamma_{m,c}^{MUE} = \frac{q_{m,c}^{MUE} P_{m,c}^{MUE} \left|h_{m,c}^{MUE}\right|^2}{\underbrace{q_{f,k,c}^{SUE} P_{f,k,c}^{SUE} \left|h_{f,k,c}^{SUE}\right|^2}_{PairingInterference, \mathcal{I}_{k\to m}^{c}} + \sigma_{m,c}^2}, \tag{A.6}$$

$$\gamma_{f,k,c}^{SUE} = \frac{q_{f,k,c}^{SUE} P_{f,k,c}^{SUE} \left|h_{f,k,c}^{SUE}\right|^2}{\underbrace{\sum_{m\in\mathcal{M}_c} A_{k,m}^{c} P_{m,c}^{MUE} \left|h_{m,c}^{MUE}\right|^2}_{PairingInterference, \mathcal{I}_{m\to k}^{c}} + \sigma_{f,k,c}^2}. \tag{A.7}$$

Note that there is no intra-codebook interference and at most one MUE is paired to an SUE on a codebook.As such, there is no intra-codebook MUE to MUE interference. The system's achievable sum rate of the SBS is given by:

$$\mathbf{R}\big(\mathbf{Q}, \mathbf{A}, \mathbf{P}\big) = \sum_{f=1}^{F} \sum_{k=1}^{K} \sum_{c=1}^{C} r_{f,k,c}^{SUE}. \tag{A.8}$$

2.2 Receiver Model

This work proposes a receiver featuring low-complexity MUD scheme based on a joint SIC-Log-MPA receiver. Firstly, the iterative Log-MPA algorithm detects the transmitted signal pair on each codebook. Thereafter, SIC sorting is applied to separate the MUE and SUE on the same codebook. Given the received signal $\mathbf{y}$, codebook size J, the joint SIC-Log-MPA receiver is a three-step decoding process that computes the estimate of the desired signal $\hat{\mathbf{x}}$ and the bit-wise log-like ratio (LLR). This process is shown in Algorithm 1 and described as follows:

Step I: Symbol iterative MPA decoding process: With the variable nodes (VN) and function nodes (FNs) representing K active users and N RUs, respectively, the algorithm performs iterative computations of log-domain extrinsic information (LDEI) sent from FN n to VN $t \in \Psi_t$, $S_{n\to t}^{\tau}(c_t)$,

and from VN k to FN $n \in \Psi_n$, $D^{\tau}_{t\to n}(c_t)$ for the transmitted codeword c_t of MUE-SUE pair t at the τ^{th} iteration as:

$$S^{\tau}_{n\to t}(c_t) = \max_{c_i|i\in\Psi_t\backslash t}\left\{\frac{-1}{\sigma^2_{f,k,n}}\left\|y^{SUE}_{f,k,n} - \sum_{t,c_{t,n}\neq 0} h^{SUE}_{f,t,n}c_{t,n}\right\|^2 + \sum_{i\in\Psi_t\backslash t} D^{\tau-1}_{i\to n}(c_i)\right\}, \tag{A.9}$$

$$D^{\tau}_{t\to n}(c_t) = \log\big(\frac{1}{J}\big) + \sum_{i\in\Psi_n\backslash n} S^{\tau-1}_{i\to t}(c_i), \tag{A.10}$$

where Ψ_t denotes the set of MUE-SUE pairs transmitting on the n^{th} RU and Ψ_n the set of RUs in a codebook such that $c_{t,n}$, the element of c_t on the n^{th} RU, is not equal to 0. The log-domain a posteriori probability of codeword c_t is computed after L iterations and given by:

$$\mathcal{Z}_{f,c_t} = \log\big(\frac{1}{J}\big) + \sum_{n\in\Psi_n\backslash n} S^{\tau_{max}}_{n\to t}(c_t). \tag{A.11}$$

Note that the low-complexity log-MPA applies a Jacobian logarithm, $\log\big(e^a + e^b\big) \approx \max\big(a, b\big)$, in the simplification and elimination of the normalization operation which in effect eliminates the exponential operations. It also simplifies the dynamic range of euclidean distance between the observed signal and the estimated transmit symbols. This significantly reduces the number of computationally time consuming multiplications encountered in the original MPA [47], [48].

Step II: SIC decoding [5], [49]: The SBS detects the SUE associated signal in a codeword by regarding the MUE signal as interference using linear detection employing a MMSE detector. Using the equivalent channel matrix $G\big(f,k,c\big) = \big[h^{SUE}_{f,k,c}\big]_{FxKxC} \odot Q^{SUE}_{F,K,C}$, the MMSE transformation weight matrix estimate is given as

$$\mathcal{W}_{MMSE}\big(f,k,c_t\big) = \Big[\big(G\big(f,k,c_t\big)\big)^H G\big(f,k,c_t\big) + \sigma^2_{f,k,c}\boldsymbol{I}\Big]^{-1}\big(G\big(f,k,c_t\big)\big)^H. \tag{A.12}$$

where I is an identity matrix. The estimate of the k^{th} SUE to the f^{th} SBS on the codebook c, $\hat{x}_{f,k,c_t}$ is derived as

$$\hat{x}_{f,k,c_t} = \mathcal{W}_{MMSE} * \mathcal{Z}_{f,c_t}. \tag{A.13}$$

In the case of more multiplexed users in one codebook, the resultant output $\hat{y}_{f,c_t}$, after decoding $\hat{x}_{f,k,c_t}$, is obtained as:

Algorithm 1 Joint-SIC-Log-MPA Receiver

Input variables Q,A, P, G, $\sigma^2_{f,k,c}$ and **y** the received signal
Initialization
Initialize the vector $D^0_{t\to n}(c_t) = -\log_2(J), \forall c_t$
Output variables $\hat{\mathbf{x}}, b_i$
for $f = 1 : F$ **do**
  for $c_t = 1, 2, \cdots, C$ **do**
    Step I: log MPA iterative process
    for $\tau = 1$ to τ_{max} **do** do
      Compute $S^{\tau}_{n\to t}(c_t)$, eqn. (A.9)
      Compute $D^{\tau}_{t\to n}(c_t)$, eqn. (A.10)
      Compute $\mathcal{Z}_{f,c_t}$, eqn. (A.11)
    end for
  end for
  Step II: SIC decoding
  for $c_t = 1, 2, \cdots, C$ **do**
    Determine $\mathcal{W}_{MMSE}\big(f, k, c_t\big)$, eqn. (A.12)
    Compute $\hat{x}_{f,k,c_t}$, eqn. (A.13)
  end for
  Step III: User Bits Reconstruction
  for $i \in 1, 2, \cdots, \log_2(J)$ **do**
    Compute $LLR(b_i)$, eqn (A.15)
    if $LLR(b_i) \leq 0$, **then** $\hat{b}_i = 1$
    else $\hat{b}_i = 0$
    end if
  end for
end for

$$\hat{y}_{f,c_t} = \mathcal{Z}_{f,c_t} - G\big(f,k,c_t\big) \odot \hat{x}_{f,k,c_t}. \tag{A.14}$$

Step III: LLR computation: The bit-wise LLR, $LLR(b_i)$, given by (A.15), estimates each user bit $\hat{b}_i$ by comparing the computed LLR to zero.

$$LLR(b_i) = \max_{c_t \in \mathcal{C}|b_i=0} \big(\hat{x}_{f,k,c_t}\big) - \max_{c_t \in \mathcal{C}|b_i=1} \big(\hat{x}_{f,k,c_t}\big), \tag{A.15}$$

2.3 Problem Formulation

The objective of this work is to maximize the overall sum rate with the unit power cost under the various QoS requirements in the small cells. The system energy efficiency is formulated as a ratio of the system sum rate to the total power consumption as [45]:

$$\eta_{EE}\big(\mathbf{Q},\mathbf{A},\mathbf{P}\big) = \frac{\mathbf{R}\big(\mathbf{Q},\mathbf{A},\mathbf{P}\big)}{P_T\big(\mathbf{Q},\mathbf{A},\mathbf{P}\big) + P_c} \tag{A.16}$$

where P_c, $P_T = \sum_{f=1}^{F}\sum_{k=1}^{K}\sum_{c=1}^{C} \xi q_{f,k,c}^{SUE} P_{f,k,c}^{SUE}$ and $\xi \in (0,1)$ are the circuit power consumption, total transmission power and energy conversion inefficiency respectively. Therefore, the corresponding energy efficiency optimization problem can be expressed as:

$$\begin{aligned}
\max_{\mathbf{Q},\mathbf{A},\mathbf{P}} \quad & \eta_{EE}\big(\mathbf{Q},\mathbf{A},\mathbf{P}\big) \\
\text{s.t.} \quad & \mathbf{C1}: \sum_{c=1}^{C} r_{f,k,c}^{SUE} \geq R_{min}^{k},\ \forall f,k \\
& \mathbf{C2}: \sum_{k=1}^{K} q_{f,k,c}^{SUE} \leq d_u,\ \forall f,c \\
& \mathbf{C3}: \sum_{m \in \mathcal{M}_c} A_{k,m}^{c} P_{m,c}^{MUE} \big|h_{m,c}^{MUE}\big|^2 \leq \mathcal{I}_{m\to k}^{c,th},\ \forall k \\
& \mathbf{C4}: q_{f,k,c}^{SUE} \in (0,1),\ \forall f,k,c \\
& \mathbf{C5}: P_{f,k,c}^{SUE} \geq 0,\ \forall f,k,c \\
& \mathbf{C6}: \sum_{c=1}^{C} q_{f,k,c}^{SUE} P_{f,k,c}^{SUE} \leq P_{max}^{SUE},\ \forall f,k \\
& \mathbf{C7}: A_{k,m}^{c} P_{m,c}^{MUE} \big|h_{m,c}^{MUE}\big|^2 \geq q_{f,k,c}^{SUE} P_{f,k,c}^{SUE} \big|h_{f,k,c}^{SUE}\big|^2,\ \forall f,k,c
\end{aligned} \tag{A.17}$$

In this problem, constraint **C1** sets the HetNet QoS rate requirement to ensure guaranteed performance of the k^{th} SUEs in the f^{th} small cell. Constraint **C2** limits the number of users multiplexed on one RU link. Constraint **C3** sets the tolerable pairing interference on each codebook on which a MUE

is co-multiplexed with a SUE. Constraint **C4** guarantees that a codebook is allocated to at most one user in each small cell. Constraints **C5** and **C6** ensure that the SUE transmit power is non-negative and is within the maximum SUE transmit power. Lastly, constraint **C7** ensures successful interference cancellation at the SBS.

However, solving (A.17) is difficult due to its mixed combinatorial feature. In addition, combined computational cost encountered in CA, UP and PA exhaustive search for optimal EE is prohibitive. Thus, sub-optimal but practical algorithms are preferred in practice. For this reason, we decouple CA, UP and PA to devise tractable algorithms. An iterative algorithm, JEERA algorithm, that jointly optimizes the system EE based on the proposed individual schemes is presented in Section 3 together with the proposed resource allocation schemes. An analytical evaluation scheme for the EE optimization, DDEEARA algorithm is later presented in Section 4.

3 Proposed Energy Efficient Resource Allocation Schemes

In this section, JEERA algorithm, **Algorithm 2** and the proposed CA, UP and PA schemes are discussed. In Algorithm 2, the system EE improves at each iteration and converges at the end of the procedure. The CA and UP are optimized using the proposed Algorithm 3 and Algorithm 4 respectively. Based on the derived CA and UP results, the PA scheme can be updated using Algorithm 5. These proposed individual EE schemes and their matrices **Q**, **A** and **P** are presented in the sub-sections that follow.

Algorithm 2 JEERA Algorithm

1: **Initialization**; Initialize the matrix $\mathbf{Q}(0)$, $\mathbf{A}(0)$ and $\mathbf{P}(0) = \left[\frac{P_{max}^{SBS}}{K}, P_M^{MBS} d_m^{-\varpi}\right]$ at $t = 0$ (iteration number), maximum iterations t_{max} and maximum tolerance ω.
2: **while** $\left(\frac{|\eta_{EE}(t+1) - \eta_{EE}(t)|}{|\eta_{EE}(t+1)|}\right) \geq \omega$ or $t \leq t_{max}$ **do**
3: Using $\mathbf{P}(t)$, SBS and MBS perform codebook assignment to obtain CA policy $\mathbf{Q}(t)$ using SCOA Algorithm, Algorithm 3.
4: SBS performs user pairing to obtain the UP policy $\mathbf{A}(t)$ using OMSP Algorithm, Algorithm 4.
5: Based on $\mathbf{Q}(t)$ and $\mathbf{A}(t)$, update the PA $\mathbf{P}(t)$ using QAPA Algorithm, Algorithm 5.
6: Set $t = t + 1$ and compute $\eta_{EE}(t+1)$.
7: **end while**

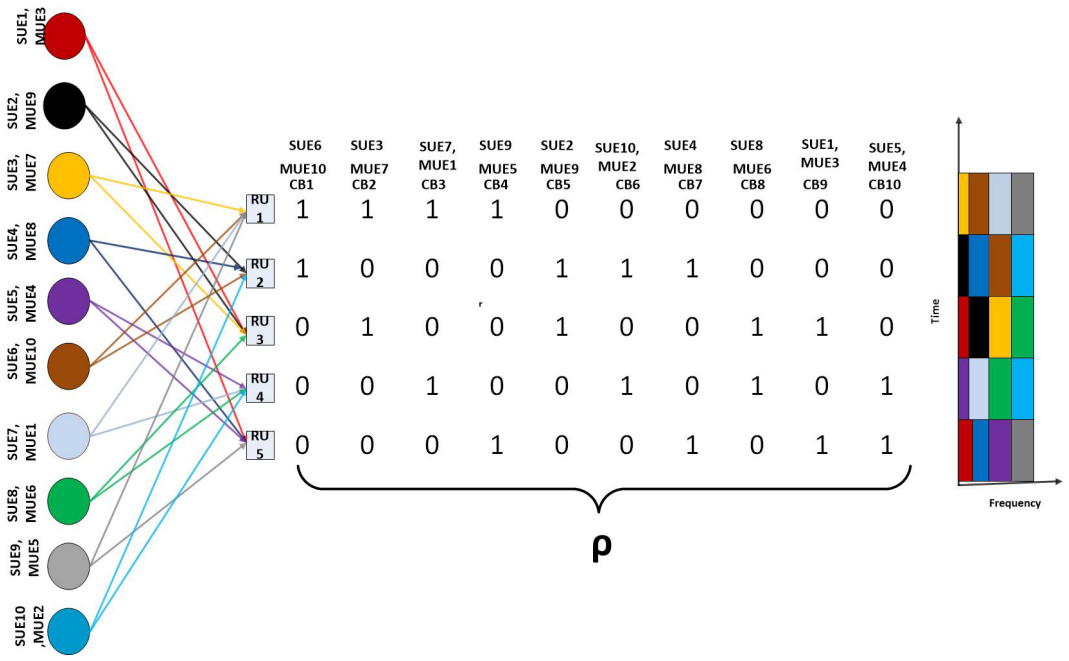

Fig. A.2: An uplink hybrid PD-SCMA resource mapping matrix ρ with $N = 5$, $d_v = 2$ and $C = 10$, with user codebook assignment $\{\mathbf{Q}\}$ and pairing policy $\{\mathbf{A}\}$ appended.

3.1 Successive Codebook Ordering Assignment (SCOA) Algorithm $\{\mathbf{Q}\}$

This consists of the RU codebook mapping and the user codebook mapping. First, the codebook RU assignment is determined by an assignment matrix $\{\rho\}$. The SCMA scheme maps a stream of $\log_2(J)$ binary bits to an N-dimensional complex codewords selected from a predefined codebook of size J [6]. The $N-$ dimensional codewords are sparse vectors with $d_v, (d_v < N)$ non-zero entries corresponding to d_v specific RUs for a user. There are $C(N, d_v) = \dfrac{N!}{(N-d_v)!d_v!}$ codebooks for a system with d_v RUs, and the set of codebooks is denoted by $\mathcal{C} = \{1, 2, \cdots, C\}$. These vectors form a codebook resource mapping matrix $\rho = \left[\rho_{c,n}\right]_{C\times N}$, where $\rho_{c,n} = 1$ represents the mapping between RU n and codebook c, $\rho_{c,n} = 0$, otherwise. The matrix elements are normally fixed for SCMA transmitter resource assignment system [18], [23]. As an example of a hybrid PD-SCMA with $N = 5$, $d_u = 4$ and $d_v = 2$, the matrix ρ is illustrated in Fig. A.2 with RU's as rows and codebooks as columns. The CA and UP are appended to the matrix for further clarity. From the figure, $\rho_{3,9} = 1$ and $\rho_{5,9} = 1$ implies that RU3 and RU5 are mapped to CB9. The appended information indicates that SUE1 and MUE3 are paired on CB9 such that the SUE CA matrix $Q^{SUE}_{F,K,C}$ element $q^{SUE}_{f,1,9} = 1$ implies SUE1 is assigned CB9 while for $Q^{MUE}_{M,C}$ the element $q^{MUE}_{3,9} = 1$ implies MUE3 is assigned CB9.

Second, the user codebook assignment matrix scheme $\{\mathbf{Q}\}$ utilizes the Successive Codebook Ordering Assignment (SCOA), **Algorithm 3**. Optimal SUE and initial MUE CA are achieved by SUE and MUE ordering based on the channel quality metrics $\varphi^{SUE}_{f,k,c}$ and $\varphi^{MUE}_{m,c}$ as given in (A.18) and (A.19)

respectively.

$$\varphi_{f,k,c}^{SUE} = \sqrt{\frac{1}{d_v} \sum_{i \in \Psi_n} \frac{P_{max}^{SBS}}{K} \frac{\left| g_{f,k,i}^{SUE} \right|^2}{\sigma_{f,k,i}^2}} \tag{A.18}$$

$$\varphi_{m,c}^{MUE} = \sqrt{\frac{1}{d_v} \sum_{j \in \Psi_n} P_M^{MBS} d_m^{-\varpi} \frac{\left| g_{m,j}^{MUE} \right|^2}{\sigma_{m,j}^2}} \tag{A.19}$$

where P_{max}^{SBS} denotes maximum SBS transmit power, P_M^{MBS} is the proportion of MBS power to serve M MUEs transmission, $g_{f,k,i}^{SUE} \sim \mathcal{CN}(0,1)$ and $g_{m,n}^{MUE}$ denote the Rayleigh fading coefficients for the n^{th} RU link of the k^{th} SUE and m^{th} MUE , d_m represents the distance of the m^{th} MUE to the MBS and ϖ is the path-loss exponent. It is clear that the metric is maximized for a user with high gain on most of the RUs and close to the BS then minimized for a poor gain user at the cell edge. For CA policy $\mathbf{Q}$, the elements of $\mathbf{Q}_{M,C}^{MUE} = \left[q_{m,c}^{MUE}\right]_{M \times C}$ and $\mathbf{Q}_{F,K,C}^{SUE} = \left[q_{f,k,c}^{SUE}\right]_{F \times K \times C}$ are determined as follows: Select the indices of the d_v largest gain coefficient RU links associated with $\max \varphi_{f,k,c}^{SUE}$ while each RU link is assigned to not more than d_u SUEs. At each index position, replace the index value with 1 to form the codeword vector $\left[\rho_{c,n}\right]_{C \times N}^{SUE}$. This codeword is assigned to the first SUE close to the SBS i.e. $q_{f,1,c}^{SUE} = 1$. The same is done for MUEs while considering $\max \varphi_{m,c}^{MUE}$. The procedure is repeated for the next user until all users are assigned to the corresponding codebooks.

3.2 Opportunistic MUE-SUE pairing (OMSP) Algorithm $\{\mathrm{A}\}$

The MUE and SUE are paired when the MUEs and SUEs are assigned the same codebook such that $q_{m,c}^{MUE} := q_{f,k,c}^{SUE}$. This implies that, $\mathbf{A} = \left[A_{k,m}^c = 1\right]_{K \times M}$. Based on Figure A.2, the user pairing matrix element $A_{1,3}^9 = 1$ implies that SUE1 pairs with MUE3 and transmits on RUs associated with codebook 9. The SUE and MUE channel quality metrics $\varphi_{f,k,c}^{SUE}$ and $\varphi_{m,c}^{MUE}$ evaluated in (A.18) and (A.19) form the vectors $\mathbf{U}_K^{SUE} = \left[\varphi_{f,1,c}^{SUE}, \cdots, \varphi_{f,K,c}^{SUE}\right]$ and $\mathbf{V}_M^{MUE} = \left[\varphi_{1,c}^{MUE}, \cdots, \varphi_{M,c}^{MUE}\right]$ respectively. The hybrid opportunistic MUE-SUE pairing (OMSP), **Algorithm 4**, is based on maximizing the channel quality metric difference (CQD), $\theta_{K \times M} = \left| \mathbf{U}_K^{SUE} \ominus \mathbf{V}_M^{MUE} \right|$ whose elements are $\theta_{k,m}$, and minimizing the pairing interference, $\mathbf{I}_{M \times K}^c$. The pairing metric, $\chi_{m,k}^c$, is given by:

$$\chi_{m,k}^c = \delta\Big(\max\left(\theta_{K \times M}\right), \min\left(\mathbf{I}_{min}^c\right) \Big) \tag{A.20}$$

In this work, an assumption that a codebook can be utilized by a maximum of one SUE in the SBS and paired to a maximum of one MUE in the MBS is made for simplicity in determining $\mathbf{Q}$ and $\mathbf{A}$ for

Algorithm 3 SCOA Algorithm

Input variables; d_u, K,M, C, $P_{F,K,C}^{SUE}$ and $P_{M,C}^{MUE}$
Initialization; Initialize the vectors $Q_{M,C}^{MUE} = [\mathbf{0}]$, $Q_{F,K,C}^{SUE} = [\mathbf{0}]$
Output variable: $\mathbf{Q^{opt}}$
for $f = 1 : F$ **do**
  for $k = 1 : K$ **do**
    Determine $\varphi_{f,k,c}^{SUE}$, eqn. (A.18)
    for $m = 1 : M$ **do**
      Determine $\varphi_{m,c}^{MUE}$, eqn. (A.19)
      $i = 1, j = 1$
      while number of users assigned on each RU $\leq d_u$ **do**
        for $l = 1 : d_v$ **do**
          Determine $t \leftarrow$ index of l^{th} largest gain coefficient d_v RU with the $\max \varphi_{f,k,c}^{SUE}$
          Determine $\left[\rho_{i,t}\right]_{C\times N}^{SUE}$ by replacing the index value of position l with 1
          Assign $q_{f,k,i}^{SUE} = 1$
        end for
        $i = i + 1$
        for $l = 1 : d_v$ **do**
          Determine $t \leftarrow$ index of l^{th} largest gain coefficient d_v RU with the $\max \varphi_{m,c}^{MUE}$
          Determine $\left[\rho_{j,t}\right]_{C\times N}^{MUE}$ by replacing the index value of position l with 1
          Assign $q_{m,j}^{MUE} = 1$
        end for
        $j = j + 1$
      end while
    end for
  end for
end for
$Q_{F,K,C}^{SUE} = \left[q_{f,k,c}^{SUE}\right]_{F\times K\times C}$, $Q_{M,C}^{MUE} = \left[q_{m,c}^{MUE}\right]_{M\times C}$

Algorithm 4 OMSP Algorithm

Input variables; $Q_{F,K,C}^{SUE}$, $Q_{M,C}^{MUE}$
Initialization; Initialize the matrix $\mathbf{I}_{M\times K}^{c} = \left[I_{m\to k}^{c}\right]_{M\times K} = [\mathbf{0}]$, $\theta_{K\times M} = [\mathbf{0}]$
Initialize the vector $\mathbf{U}_{K}^{SUE} = [\mathbf{0}]$, $\mathbf{V}_{M}^{MUE} = [\mathbf{0}]$
Output variable; $\mathbf{A^{opt}}$
for $k = 1 : K$ **do**
    Evaluate the matrix $\mathbf{U}_{K}^{SUE}$
    for $m = 1 : M$ **do**
        Evaluate the matrix $\mathbf{V}_{M}^{MUE}$
    end for
end for
Evaluate the CQD, $\theta_{K\times M}$
for $k = 1 : K$ **do**
    for $m = 1 : M$ **do**
        Compute interference temperature coefficients, $I_{m\to k}^{c}$
        Determine matrix $\mathbf{I}_{M\times K}^{c}$, from elements $I_{m\to k}^{c}$
    end for
end for
Determine $\mathbf{I}_{min}^{c}$, Kuhn-Munkres on $\mathbf{I}_{M\times K}^{c}$
Evaluate $\chi_{m,k}^{c}$, eqn. (A.20)
Perform post-pairing
Output $\mathbf{A^{opt}} = \left[A_{k,m}^{c}\right]_{K\times M}$

PD-SCMA.

3.3 QoS Aware Power Allocation (QAPA) Algorithm $\{\mathbf{P}\}$

Once the CA (**Q**) and UP (**A**) schemes are obtained, the power allocation scheme matrix $\mathbf{P} = \left\{P_{F,K,C}^{SUE}, P_{M,C}^{MUE}\right\}$ applies the QoS satisfaction-based metrics at the MBS and at each SBS [50]. The SUE transmission power is constrained such that the MUE can satisfy their minimum QoS. Let R_{min}^{m} denote the minimum achievable rate for the m^{th} MUE on codebook c. Using (A.4) and (A.6), the MUE QoS requirement is given as:

$$B_{ru}\log_2\left(1+\frac{q_{m,c}^{MUE}P_{m,c}^{MUE}\left|h_{m,c}^{MUE}\right|^2}{q_{f,k,c}^{SUE}P_{f,k,c}^{SUE}\left|h_{f,k,c}^{SUE}\right|^2+\sigma_{m,c}^2}\right)\geq R_{min}^{m}. \tag{A.21}$$

Algorithm 5 QAPA Algorithm

Input; R^k_{min}, R^m_{min}, $\mathbf{H}$, $\sigma^2_{m,c}$, $\sigma^2_{f,k,c}$, $\mathcal{I}^{c,th}_{k\to m}$
Initialization; Initialize $P^{SUE}_{f,k,c} = P^{SBS}_{f,k,c}/K$, $P^{MUE,init}_{m,c} = P^{MBS}_M d_m^{-\varpi}$
Output variable; $\mathbf{P^{opt}}$
Step I: SUE Power Allocation
for $f = 1 : F$ **do**
  for $k = 1 : K$ **do**
    Based on (A.21), evaluate $P^{SUE}_{f,k,c}$, eqn. (A.22)
    Compute $\bar{P}^{SUE}_{f,k,c}$, eqn. (A.23)
    Compute $P^{SUE,min}_{f,k,c}$ for minimum SUE QoS requirement, eqn. (A.25)
    Compute $\omega^{SUE}_{f,k,c}$, eqn. (A.26)
  end for
end for
Step II: MUE Power Allocation
for $m = 1 : M$ **do**
  Determine MUE rate requirement, eqn. (A.21)
  Compute $P^{MUE,min}_{m,c}$, eqn. (A.27)
  Compute $\omega^{MUE}_{m,c}$, eqn. (A.28)
end for

From (A.21), the SUE transmit power $P^{SUE}_{f,k,c}$ is constrained by:

$$P^{SUE}_{f,k,c} \leq \frac{1}{q^{SUE}_{f,k,c}\left|h^{SUE}_{f,k,c}\right|^2}\left(\frac{q^{MUE}_{m,c}P^{MUE}_{m,c}\left|h^{MUE}_{m,c}\right|^2}{2^{\frac{R^m_{min}}{B_{ru}}}-1} - \sigma^2_{m,c}\right) \tag{A.22}$$

$P^{SUE}_{f,k,c}$ is also limited due to the predetermined interference threshold $\mathcal{I}^{c,th}_{k\to m} = q^{SUE}_{f,k,c}\bar{P}^{SUE}_{f,k,c}\left|h^{MUE}_{m,c}\right|^2$. Hence, by this provision, the SUE transmit power under interference awareness, $\bar{P}^{SUE}_{f,k,c}$ is computed as:

$$\bar{P}^{SUE}_{f,k,c} \leq \frac{\mathcal{I}^{c,th}_{k\to m}}{q^{SUE}_{f,k,c}\left|h^{SUE}_{f,k,c}\right|^2} \tag{A.23}$$

Similarly, using (A.5) and (A.7), the SUE transmit power constraint for its minimum QoS is computed as in (A.24). Therefore, the minimum SUE transmit power for minimum QoS, $P^{SUE,min}_{f,k,c}$ is computed by (A.25).

$$P^{SUE}_{f,k,c} \geq \frac{2^{\frac{R^m_{min}}{B_{ru}}}-1}{q^{SUE}_{f,k,c}\left|h^{SUE}_{f,k,c}\right|^2}\left(q^{MUE}_{m,c}P^{MUE}_{m,c}\left|h^{MUE}_{m,c}\right|^2 + \sigma^2_{f,k,c}\right) \tag{A.24}$$

$$P_{f,k,c}^{SUE,min} = \frac{2^{\frac{R_{min}^{k}}{B_{ru}}} - 1}{q_{f,k,c}^{SUE}\left|h_{f,k,c}^{SUE}\right|^2}\left(q_{m,c}^{MUE}P_{m,c}^{MUE}\left|h_{m,c}^{MUE}\right|^2 + \sigma_{f,k,c}^2\right) \tag{A.25}$$

Therefore, the final constrained SUE transmit power metric is given by:

$$\omega_{f,k,c}^{SUE} = \begin{cases} \min\left(\bar{P}_{f,k,c}^{SUE}, \max\left(P_{f,k,c}^{SUE}, P_{f,k,c}^{SUE,min}\right), P_{f,k,c}^{max}\right), & \text{if } \Delta \geq P_{f,k,c}^{SUE,min} \\ Infeasible, & \text{otherwise} \end{cases} \tag{A.26}$$

where $\Delta = \min\left(P_{f,k,c}^{SUE}, \bar{P}_{f,k,c}^{SUE}\right)$. The power allocation matrix $P_{f,k,c}^{SUE}$ is increased and decreased in the bounds of (A.26). The minimum power requirement for each MUE to meet its own QoS is then given by:

$$P_{m,c}^{MUE,min} = \frac{2^{\frac{R_{min}^{m}}{B_{ru}}} - 1}{q_{m,c}^{MUE}\left|h_{m,c}^{MUE}\right|^2}\left(q_{f,k,c}^{SUE}P_{f,k,c}^{SUE}\left|h_{f,k,c}^{SUE}\right|^2 + \sigma_{m,c}^2\right) \tag{A.27}$$

As such, the power allocation metric, $\omega_{m,c}^{MUE}$, for the MUEs will be defined as

$$\omega_{m,c}^{MUE} = \min\left(P_{m,c}^{MUE,init}, P_{m,c}^{MUE,min}\right) \tag{A.28}$$

A QoS aware power allocation (QAPA) strategy for SUEs and MUEs is illustrated in **Algorithm 5**.

4 Dual Decomposition-based EE Analytical Resource Allocation (DDEEARA) Evaluation Technique

The problem in (A.17) is a mixed integer non-linear programming (MINLP) and cannot be solved by classical mathematical methods. An alternative near-optimal analytical EE resource allocation optimization, DDEEARA is discussed in this section. To solve the problem, we employ mathematical transformations and convex optimization techniques to transform (A.17) into a concave problem. Thereafter, the Lagrange dual method used in [45] and [51] is applied to solve the resultant transformed problem. However, from the Lagrangian, we obtain independent optimal points for the CA, UP and PA based on the associated dual variables.

4.1 Problem Transformations

Since (A.17) is a non-convex fractional programming problem, we first perform integer relaxation of CA indicators $q_{f,k,c}^{SUE}$ and $q_{m,c}^{MUE}$ as well as the pairing indicator $A_{k,m}^{c}$ to obtain continuous variables $\hat{q}_{f,k,c}^{SUE} \in [0,1]$, $\hat{q}_{m,c}^{MUE} \in [0,1]$ and $\hat{A}_{k,m}^{c} \in [0,1]$ respectively. The relaxed continuous indicators can

intuitively be interpreted as time-sharing QoS for codebooks. Following the integer relaxation, the CA, UP and PA policies become $\hat{\mathbf{Q}} = \left[\hat{Q}^{SUE}_{F,K,C}, \hat{Q}^{MUE}_{M,C}\right]$, $\hat{\mathbf{A}} = \left[\hat{A}^{c}_{k,m}\right]_{KxM}$ and $\hat{\mathbf{P}} = \left[\hat{P}^{SUE}_{F,K,C}, \hat{P}^{MUE}_{M,C}\right]$ respectively. The approximation of the sum-rate $r^{SUE}_{f,k,c}$ as used in [[51]] is given as $\hat{r}^{SUE}_{f,k,c} = B_{ru}\log_2\left(1+\hat{\gamma}_{f,k,c}\right)$, with $\hat{\gamma}_{f,k,c}$ given by:

$$\hat{\gamma}^{SUE}_{f,k,c} = \frac{\hat{q}^{SUE}_{f,k,c}\hat{P}^{SUE}_{f,k,c}\left|h^{SUE}_{f,k,c}\right|^2}{\underbrace{\sum_{m\in\mathcal{M}_c}\hat{A}^{c}_{k,m}\hat{P}^{MUE}_{m,c}\left|h^{MUE}_{m,c}\right|^2}_{PairingInterference,\mathcal{I}^{c}_{m\to k}} + \hat{q}^{SUE}_{f,k,c}\sigma^2_{f,k,c}} \tag{A.29}$$

Secondly, to deal with the non-linearity of the fractional objective function, a parametric transformation from the Dinkelbach method [52] is used to transform the fractional program to a non-fractional one. The maximum EE of the system can be defined as:

$$b^* = \max_{\mathbf{Q},\mathbf{A},\mathbf{P}} \frac{\mathbf{R}\left(\hat{\mathbf{Q}},\hat{\mathbf{A}},\hat{\mathbf{P}}\right)}{P_T\left(\hat{\mathbf{Q}},\hat{\mathbf{A}},\hat{\mathbf{P}}\right)+P_c} = \frac{\mathbf{R}\left(\hat{\mathbf{Q}}^*,\hat{\mathbf{A}}^*,\hat{\mathbf{P}}^*\right)}{P^*_T\left(\hat{\mathbf{Q}}^*,\hat{\mathbf{A}}^*,\hat{\mathbf{P}}^*\right)+P_c} \tag{A.30}$$

which is a non-linear concave-convex fractional program. This can be transformed into an equivalent parameterized non-fractional formform given by [46]:

$$\max_{\mathbf{Q},\mathbf{A},\mathbf{P}} \left\{\mathbf{R}\left(\hat{\mathbf{Q}},\hat{\mathbf{A}},\hat{\mathbf{P}}\right) - b\left(P_T\left(\hat{\mathbf{Q}},\hat{\mathbf{A}},\hat{\mathbf{P}}\right)+P_c\right)\right\} \tag{A.31}$$

with the conditions $\mathbf{C1}$ - $\mathbf{C7}$ of (A.17) modified to include the relaxed indicators. b is a parameter introduced to scale the weight of $P_T\left(\hat{\mathbf{Q}},\hat{\mathbf{A}},\hat{\mathbf{P}}\right)+P_c$. For a given value of b, the corresponding resource allocation policy is denoted as $\hat{\mathbf{Q}}$, $\hat{\mathbf{A}}$, $\hat{\mathbf{P}}$ and the optimal solution to (A.30) defined as $\mathbf{Q}^*$, $\mathbf{A}^*$, $\mathbf{P}^*$. Note that the solution in (A.30) is equivalent to finding the maximum energy efficiency b^*. It can be proven that the optimal solution of the subtractive form of (A.31) is reached when $\mathbf{R}\left(\hat{\mathbf{Q}},\hat{\mathbf{A}},\hat{\mathbf{P}}\right) - b\left(P_T\left(\hat{\mathbf{Q}},\hat{\mathbf{A}},\hat{\mathbf{P}}\right)+P_c\right)$ approaches zero [46].

4.2 Lagrange Dual Decomposition

The transformed optimization problem in (A.31) is jointly concave with respect to $\{\mathbf{Q},\mathbf{A},\mathbf{P}\}$. It is solved by adopting the Lagrange dual decomposition method. The difference between the primal and dual solution is zero when strong duality holds [45]. In this case, we solve the primal problem of (A.31) by solving its associated dual problem. The Lagrangian function can be written as:

$$\mathcal{L}\big(\mathbf{Q},\mathbf{A},\mathbf{P},b,\aleph\big)=\sum_{f=1}^{F}\sum_{k=1}^{K}\sum_{c=1}^{C}\hat{r}_{f,k,c}^{SUE}-b\Bigg(\sum_{f=1}^{F}\sum_{k=1}^{K}\sum_{c=1}^{C}\xi\hat{P}_{f,k,c}^{SUE}+P_c\Bigg)+\sum_{f=1}^{F}\sum_{k=1}^{K}\lambda_{f,k}\Bigg(\sum_{c=1}^{C}\hat{r}_{f,k,c}^{SUE}-R_{min}^{k}\Bigg)$$
$$+\sum_{f=1}^{F}\sum_{c=1}^{C}\alpha_{f,c}\Bigg(d_u-\sum_{k=1}^{K}\hat{q}_{f,k,c}^{SUE}\Bigg)+\sum_{k=1}^{K}\beta_k\Bigg(\mathcal{I}_{m\rightarrow k}^{c}-\sum_{m\in\mathcal{M}_c}\hat{A}_{m,c}^{c}\hat{P}_{m,c}^{MUE}\big|h_{m,c}^{MUE}\big|^2\Bigg)$$
$$+\sum_{f=1}^{F}\sum_{k=1}^{K}\nu_{f,k}\Bigg(P_{max}^{SUE}-\sum_{c=1}^{C}\hat{P}_{f,k,c}^{SUE}\Bigg)+\sum_{f=1}^{F}\sum_{k=1}^{K}\sum_{c=1}^{C}\varrho_{f,k,c}\Bigg(\hat{A}_{m,c}^{c}P_{m,c}^{MUE}\big|h_{m,c}^{MUE}\big|^2$$
$$-\hat{q}_{f,k,c}^{SUE}P_{f,k,c}^{SUE}\big|h_{f,k,c}^{SUE}\big|^2\Bigg). \tag{A.32}$$

where $\aleph=\big(\lambda\geq\mathbf{0},\alpha\geq\mathbf{0},\beta\geq\mathbf{0},\nu\geq\mathbf{0},\varrho\geq\mathbf{0}\big)$ are the Lagrange multipliers corresponding to the modified constraints of $\mathbf{C1}$, $\mathbf{C2}$, $\mathbf{C3}$, $\mathbf{C6}$ and $\mathbf{C7}$. The modified boundary constraints of $\mathbf{C4}$ and $\mathbf{C5}$ are absorbed in the Karush-Kuhn-Tucker (KKT) conditions [53]. The dual function is defined as:

$$g\big(b,\lambda,\alpha,\beta,\nu,\varrho\big)=\max_{\mathbf{Q},\mathbf{A},\mathbf{P}}\mathcal{L}\big(\mathbf{Q},\mathbf{A},\mathbf{P},b,\aleph\big) \tag{A.33}$$

Given b, the dual problem of (A.17) can be expressed as:

$$\min_{\lambda,\alpha,\beta,\nu,\varrho}g\big(b,\lambda,\alpha,\beta,\nu,\varrho\big) \tag{A.34}$$

$$s.t.,\lambda\geq\mathbf{0},\alpha\geq\mathbf{0},\beta\geq\mathbf{0},\nu\geq\mathbf{0},\varrho\geq\mathbf{0} \tag{A.35}$$

Since each f^{th} SBS solves for optimal resource allocation for the k associated SUEs using local information, (A.31) can be decomposed into f sub-problems as:

$$\mathcal{L}\big(\mathbf{Q},\mathbf{A},\mathbf{P},b,\aleph\big)=\mathcal{L}_f\big(\mathbf{Q},\mathbf{A},\mathbf{P},b,\aleph\big)-bP_c-\sum_{f=1}^{F}\sum_{k=1}^{K}\lambda_{f,k}R_{min}^{k}+\sum_{f=1}^{F}\sum_{c=1}^{C}\alpha_{f,c}d_u+\sum_{k=1}^{K}\beta_k\mathcal{I}_{m\rightarrow k}^{c}$$
$$+\sum_{f=1}^{F}\sum_{k=1}^{K}\nu_{f,k}P_{max}^{SUE} \tag{A.36}$$

where,

$$\begin{aligned}\mathcal{L}_f\left(\mathbf{Q},\mathbf{A},\mathbf{P},b,\aleph\right) =& \sum_{k=1}^{K}\sum_{c=1}^{C}(1+\lambda_{f,k})\hat{r}_{f,k,c}^{SUE} - \sum_{k=1}^{K}\sum_{c=1}^{C}(b\xi+\nu_{f,k})\hat{P}_{f,k,c}^{SUE} - \sum_{k=1}^{K}\sum_{c=1}^{C}\alpha_{f,c}\hat{P}_{f,k,c}^{SUE} \\ &- \sum_{k=1}^{K}\sum_{m\in\mathcal{M}_c}\hat{A}_{m,c}^{c}\hat{P}_{m,c}^{MUE}\left|h_{m,c}^{MUE}\right|^2 + \sum_{k=1}^{K}\sum_{c=1}^{C}\varrho_{f,k,c}\hat{A}_{m,c}^{c}P_{m,c}^{MUE}\left|h_{m,c}^{MUE}\right|^2 \\ &- \sum_{k=1}^{K}\sum_{c=1}^{C}\varrho_{f,k,c}\hat{q}_{f,k,c}^{SUE}P_{f,k,c}^{SUE}\left|h_{f,k,c}^{SUE}\right|^2.\end{aligned} \tag{A.37}$$

Optimal resource allocation solutions are obtained by applying KKT conditions in combination with optimization methods to find the stationary point of (A.37) with respect to $\mathbf{Q}$, $\mathbf{A}$, and $\mathbf{P}$ respectively, with λ, α, β, ν and ϱ being fixed. Therefore, the optimal power allocation solution $P_{f,k,c}^{SUE}$ can be described mathematically as:

$$P_{f,k,c}^{SUE} = \left[\frac{B_{ru}(1+\lambda_{f,k})}{\ln 2\left(b\xi+\nu_{f,k}+\varrho_{f,k,c}\left|h_{f,k,c}^{SUE}\right|^2\right)} - \frac{1}{\Delta_{f,k,c}^{SUE}}\right]^{+} \tag{A.38}$$

where $[x]^{+} = \max(0,x)$ and $\Delta_{f,k,c}^{SUE} = \frac{q_{f,k,c}^{SUE}\left|h_{f,k,c}^{SUE}\right|^2}{A_{m,c}^{c}P_{m,c}^{MUE}\left|h_{m,c}^{MUE}\right|^2+\sigma_{f,k,c}^2}$. Using the concept of marginal benefits [45], the policy assignment of codebook c to the optimal user can be defined by:

$$q_{f,k,c}^{SUE} = 1|k^* = \arg\max_{k}\Gamma_{f,k,c} \tag{A.39}$$

where

$$\Gamma_{f,k,c} = B_{ru}(1+\lambda_{f,k})\frac{1}{ln2}\frac{\frac{P_{f,k,c}^{SUE}\left|h_{f,k,c}^{SUE}\right|^2}{\mathcal{I}_{m\to k}^{c}+\sigma_{f,k,c}^2}}{\left(1+\frac{P_{f,k,c}^{SUE}\left|h_{f,k,c}^{SUE}\right|^2}{\mathcal{I}_{m\to k}^{c}+\sigma_{f,k,c}^2}\right)} - \alpha_{f,c} - (b\xi+\nu_{f,k})P_{f,k,c}^{SUE} - \varrho_{f,k,c}P_{f,k,c}^{SUE}\left|h_{f,k,c}^{SUE}\right|^2. \tag{A.40}$$

The optimal pairing policy, $A_{k,m}^{c}$ can be defined as:

$$A_{k,m}^{c} = 1|m^* = \arg\max_{m}\Xi_{f,k,c} \tag{A.41}$$

with $\Xi_{f,k,c}$ given by:

Algorithm 6 DDEEARA Algorithm

Initialization:
Initialize maximum number of the iterations I_{max}, maximum tolerance ω, set $\tau = 0$
Initialize $P_{f,k,c}^{SUE} = P_{f,k,c}^{SBS}/K$, $P_{m,c}^{MUE,init} = P_M^{MBS} d_m^{-\varpi}$
Initialize the energy efficiency, b
while $\left|\mathbf{R}(\mathbf{Q},\mathbf{A},\mathbf{P})^{\tau} - b^{*}\left(P_T(\mathbf{Q},\mathbf{A},\mathbf{P})^{\tau} + P_c\right)\right| \geq \omega$ or $\tau \leq I_{max}$ **do**
  Initialize maximum number of inner loop iterations L_{max} and set $l = 0$
  for $c = 1 : C$ **do**
    for $m = 1 : M$ **do**
      for $k = 1 : K$ **do**
        Initialize dual variables λ, α, β, ν and ϱ
        Using b^{τ}, update $P_{f,k,c}^{SUE}$ eqn. (A.38)
        Compute $q_{f,k,c}^{SUE}$, eqn. (A.39)
        Compute $A_{k,m}^{c}$, eqn. (A.41)
        Compute the sum-rate, eqn. (A.8)
        while dual variables have not converged or $l \leq L_{max}$ **do**
          SBS updates $\lambda_{f,k}^{l+1}$, $\alpha_{f,c}^{l+1}$, $\nu_{f,k}^{l+1}$, $\varrho_{f,k,c}^{l+1}$ using (A.43), (A.44), (A.46), (A.47)
          MBS updates β_k^{l+1} according to (A.45)
        end while
      end for
    end for
  end for
  Set $\tau = \tau + 1$ and let $b^{\tau} = \dfrac{\mathbf{R}(\mathbf{Q},\mathbf{A},\mathbf{P})^{\tau-1}}{\left(P_T^{*}(\mathbf{Q},\mathbf{A},\mathbf{P})^{\tau-1} + P_c\right)}$
end while

$$\begin{aligned}\Xi_{f,k,c} =& B_{ru}(1+\lambda_{f,k})\log_2\left(1+\frac{q_{f,k,c}^{SUE}P_{f,k,c}^{SUE}\left|h_{f,k,c}^{SUE}\right|^2}{P_{m,c}^{MUE}\left|h_{m,c}^{MUE}\right|^2+\sigma_{f,k,c}^2}\right) - \beta_k P_{m,c}^{MUE}\left|h_{m,c}^{MUE}\right|^2 \\ & -\varrho_{f,k,c}P_{m,c}^{MUE}\left|h_{m,c}^{MUE}\right|^2.\end{aligned} \tag{A.42}$$

After obtaining the optimal CA, UP and PA, we invoke the sub-gradient method to iteratively update the dual multipliers given by:

$$\lambda_{f,k}^{l+1} = \left[\lambda_{f,k}^{l} - \zeta_1^{l}\left(\sum_{c=1}^{C} r_{f,k,c}^{SUE} - R_{min}^{k}\right)\right]^{+} \tag{A.43}$$

$$\alpha_{f,c}^{l+1} = \left[\alpha_{f,c}^{l} - \zeta_2^l\left(d_u - \sum_{k=1}^{K} q_{f,k,c}^{SUE}\right)\right]^+ \tag{A.44}$$

$$\beta_{k}^{l+1} = \left[\beta_{k}^{l} - \zeta_3^l\left(\mathcal{I}_{m\to k}^{c,th} - \sum_{m\in\mathcal{M}_c} A_{k,m}^{c} P_{m,c}^{MUE}\left|h_{m,c}^{MUE}\right|^2\right)\right]^+ \tag{A.45}$$

$$\nu_{f,k}^{l+1} = \left[\nu_{f,k}^{l} - \zeta_4^l\left(P_{max}^{SUE} - \sum_{c=1}^{C} q_{f,k,c}^{SUE} P_{f,k,c}^{SUE}\right)\right]^+ \tag{A.46}$$

$$\varrho_{f,k,c}^{l+1} = \left[\varrho_{f,k,c}^{l} - \zeta_5^l\left(A_{k,m}^{c} P_{m,c}^{MUE}\left|h_{m,c}^{MUE}\right|^2 - q_{f,k,c}^{SUE} P_{f,k,c}^{SUE}\left|h_{f,k,c}^{SUE}\right|^2\right)\right]^+ \tag{A.47}$$

with ζ_1, ζ_2,ζ_3, ζ_4 and ζ_5 being positive infinitesimal step sizes at iteration l. The iterative DDEEARA can then be described by **Algorithm 6**.

5 CONVERGENCE AND COMPLEXITY ANALYSIS OF THE PROPOSED SCHEMES

In this section, we outline analytical discussion of convergence and the computational complexity of the simulation-based and analytical-based EE resource allocation schemes for the PD-SCMA system.

5.1 CONVERGENCE ANALYSIS

5.1.1 JEERA algorithm

The convergence of the JEERA algorithm is given in the following Lemma.

Lemma 5.1. *For a given EE value η_{EE} and minimum deviation ω , the JEERA algorithm converges within a limited finite number of iterations.*

Proof. The proof of Lemma 5.1 follows similar proof of joint resource allocation algorithm (NVRA Algorithm 4) in [54]. The maximum power of the SBS and the number of codebooks are finite. Consequently, the EE of the system (A.17) is upper bounded. For a given $\mathbf{Q}$ and $\mathbf{P}$, OMSP algorithm computes the optimal $\mathbf{A^{opt}}$. At each iteration t, OMSP computes the CQD for maximum CQD, $\max\theta_{K\times M}$ and pairing interference for $\mathbf{I}_{min}^{c}$ to optimize the pairing metric $\chi_{m,k}^{c}$. Due to its increasing monotonicity, it can be observed that OMSP converges to $\mathbf{A^{opt}}$. Mathematically, it can be written as

$$\mathbf{A^0} \rightarrow \mathbf{A^1} \rightarrow \mathbf{A^2} \rightarrow \cdots \rightarrow \mathbf{A^{opt-1}} \rightarrow \cdots \rightarrow \mathbf{A^{opt}} \tag{A.48}$$

Since EE is upper bounded and increased in each iteration, OMSP algorithm must converge to $\mathbf{A^{opt}}$ within a limited finite number of iterations. To further enhance the EE, QAPA algorithm is updated in step 3 with $\mathbf{Q}$ and $\mathbf{A}$ fixed. Consequently, the EE is increased from $\eta_{EE}(t)$ to $\eta_{EE}(t+1)$. Then a new cycle for SUE codebook assignment and MUE pairing is started. Overall, the $\eta_{EE}(t)$ is monotonically increasing and upper bounded, which guarantees that the total number of iterations is limited for a given deviation of ω. This is further illustrated in the numerical results. □

5.1.2 DDEEARA algorithm

The convergence of the DDEEARA algorithm is given in the following Lemma.

Lemma 5.2. *Algorithm 6 will converge to the optimum of dual problem A.34 given that A.35 holds, in a finite number of iterations.*

Proof. Let $F(\eta_{EE}) = \mathbf{R}\big(\hat{\mathbf{Q}}, \hat{\mathbf{A}}, \hat{\mathbf{P}}\big) - b\big(P_T\big(\hat{\mathbf{Q}}, \hat{\mathbf{A}}, \hat{\mathbf{P}}\big) + P_c\big)$ denote the Dinkelbach transformation expression of the EE problem. For the outer layer, since $F(\eta_{EE})$ is strictly monotonic decreasing, then an optimal solution η^*_{EE}, is guaranteed when

$$0 = F(\eta^*_{EE}) \geq F(\eta_{EE}) = 0 \tag{A.49}$$

The EE $F(\eta_{EE})$ improves after each outer layer iteration τ, until the convergence (eqn. A.49 is satisfied). The convergence and optimality proof of the Dinkelbach based equivalence scheme can be found in [52]. For the inner layer problem, the subgradient method with the constant step size is employed to tackle the dual problem minimization. Selection of the step size is largely inspired by its practical significance to achieve primal solutions and improve the convergence rate analysis. At each inner layer iteration l, the Lagrange multipliers $\lambda, \alpha, \beta, \nu, \varrho$ are cautiously selected to enhance the square summation of step sizes in order to ensure guaranteed convergence. By introducing Theorem 4 in [55], if step size $\zeta_i^l, (i = 1, 2, 3, 4)$ satisfies following two conditions

$$\lim_{l\to\infty} \zeta_i^l = 0, \qquad \sum_{l=0}^{\infty} \zeta_i^l = \infty \tag{A.50}$$

then, starting from any initial value of $\mathbf{Q}$, $\mathbf{A}$ and $\mathbf{P}$, the dual problem of A.17 can converge to the optimal solution of A.34. It can be shown that for a constant step size, the solution strictly converges

to the optimal solution of A.34. It follows that, if the inner and outer layer converges, the proof of Lemma 5.2 is complete.

□

5.2 COMPLEXITY ANALYSIS

5.2.1 Receiver Complexity

The complexity of PD-NOMA and SCMA have been derived in [18]. In this analysis, G, J, d_u, d_v, L and τ are used to denote the equivalent channel matrix, codebook set size, number users superimposed in one RU, number of RUs assigned to each user, number of users simultaneously multiplexed in one codebook and the number of iterations respectively. For PD-NOMA with $S - QAM$ constellation [56], computing the estimated received symbol requires $S(\log_2 S + 1)$ multiplications and $2(S - 1)$ additions for each RU. By considering the MMSE detector at the receiver, the solution involves the computation of MMSE transformation weight matrix estimate given by $(G^H G + \sigma^2 I)^{-1} G^H$. The computational complexity order of λ^{-1} and $\lambda^H \lambda$ matrices (with size $b \times b$) is $\mathcal{O}(b^3)$. For SCMA, the original MPA and log-MPA detector complexity is investigated in [57]. Extending this knowledge to PD-SCMA, the Log-MPA algorithm increases the number of addition operations when compared to the original MPA. However, it can save more than 40% of the multiplications and completely eliminates the exponential operations. Assuming that the number of users paired on a codebook is L, 2 in our case, the SBS should apply log MPA once to detect and decode the C codebook layers. Then for each codebook layer, SIC is applied $L - 1$ times to correctly decode the transmitted data. The number of additions, multiplications, exponential and comparison operations, and the resultant overall approximate computational complexity order are summarized in Table A.2.

5.2.2 Codebook Assignment Complexity

The codebook resource assignment at the SBS is an assignment between C codebooks and K SUEs. The SBS exhaustively searches the best assignment. Since each codebook can only be assigned to at most one SUE, there are $K \times C$ possible assignments with the complexity of the optimal exhaustive search being $\mathcal{O}(2^{KC})$. The SCOA algorithm is considered as an ordering problem with respect to the channel quality with complexity $\mathcal{O}(K^2)$. Hence the total complexity of the SCOA algorithm is given as $\mathcal{O}(CK^2)$.

Table A.2: Receiver Complexity of PD-NOMA, SCMA and PD-SCMA

Algorithms	**PD-NOMA**	**SCMA**	**PD-SCMA**
ADD	$2\tau d_u d_v(S-1)$	$J^{d_u}Nd_u\tau(2d_u+1)+JNd_u\tau(d_v-2)+2JNd_u$	$2L(S-1)+(3J^{d_u}Nd_u\tau+4JNd_u)$
MUL	$2\tau d_u d_v S(\log_2 S+1)$	$J^{d_u}Nd_u\tau(d_u+1)+4JNd_u$	$3L(\log_2 S+1)+(J^{d_u}Nd_u\tau+2JNd_u)$
EXP	0	$J^{d_u}Nd_u\tau$	0
CMP	0	0	$L(J^{d_u}Nd_u\tau+JNd_u\tau(d_v-2))$
Overall complexity order	$\mathcal{O}\big((L^3+(2L^2d_v)(L-1))+(\log_2 S)\big)$	$\mathcal{O}\big((\tau_{max})\big\lvert J\big\rvert^{d_v}\big)$	$\mathcal{O}\big((\tau_{max})\log_2\big\lvert J\big\rvert^{d_v}+Lb^3\log_2 S\big)$

5.2.3 User Pairing Complexity

In the OMSP algorithm, pairing between SUE k and MUE m on a codebook c and is based on the maximum channel quality difference $(\theta_{K\times M})$ and the minimum interference temperature $\mathcal{I}^c_{min}$ obtained through the Kuhn Munkres algorithm which has a complexity of $\mathcal{O}(MK^2)$. Since each MUE calculates its $(\theta_{k,m})$ and $\mathcal{I}^c_{min}$ corresponding to each SUE allocated on codebook c, the post-ordering has a complexity of $\mathcal{O}(MK)$. Therefore, the combined pairing complexity can be approximated to $\mathcal{O}(MK^2)$.

5.2.4 Power Allocation Complexity

In the QAPA algorithm, a SUE considers the rate of the pairing MUE at each cycle. The SUE then considers its own power limitation due to interference threshold. Finally, the SUE considers its own minimum QoS requirements before deciding on its own power limits as given in (A.26). The complexity is of order $\mathcal{O}(MK)$. The MUE power allocation is based on its own minimum QoS satisfaction and requires $\mathcal{O}(M)$. As such, the QAPA algorithm complexity can be defined to be of order $\mathcal{O}(MK)$.

5.2.5 DDEEARA Complexity

Following the work of [51], the computational complexity of the proposed DDEEARA, **Algorithm 6**, can be derived. In **Algorithm 6**, a worst-case complexity of searching (A.38) needs FKC operations in each iteration, and the calculation of (A.40) for every SUE on each codebook in every small cell

Table A.3: Simulation parameters

Parameters	**Symbol**	**Values**
Carrier frequency		$2GHz$
MBS coverage radius		$500m$
SBS coverage radius		$50m$
Path loss exponent	ϖ	3
Error variance	σ_{ε}^2	0.01
RU Bandwidth	B_{ru}	$10MHz$
Number of SBS	F	60
Number of SUEs per Small cell	K	20
Number of Codebooks	C	20
Number of RUs	N	6
Number of RUs per codebook	d_v	3
Number of users in a codebook	L	2
Signal to Noise Ratio	SNR	$32dB$
SUE circuit power	p_c	$21dBm$
Noise variance	$\sigma_{f,k,n}^2$	$-125dBm$
Minimum SUE transmission rate	R_{min}^k	$5Mbps/Hz$
Maximum transmission power	$P_{f,k,c}^{SUE}$	$25dBm$
Interference threshold	$\mathcal{I}_{m\to k}^c$	$10^{-5.5}W$
Interference threshold	$\mathcal{I}_{k\to m}^c$	$10^{-5.5}W$
Power amplifier inefficiency	ξ	0.9

entails FKC operations. While using the sub-gradient method, the algorithm requires π iterations to converge, $\mathcal{O}(KF)$ calculations are required for updating each dual variable $\lambda_{f,k}$, $\alpha_{f,c}$ and $\nu_{f,k}$. $\mathcal{O}(K)$ and $\mathcal{O}(KFC)$ calculations are required to update β_k and $\varrho_(f,k,c)$ respectively. Therefore, π is a polynomial function of $F^2K^2C^2$, and the complexity order of algorithm is $\mathcal{O}(F^2K^2C^2\pi)$. The choice of the dual variables greatly affects the number of iterations and therefore must be carefully chosen.

6 SIMULATION RESULTS AND DISCUSSION

The performance of the proposed resource allocation algorithms in terms of EE, sum-rate and complexity is presented herein. In the simulation model, spectrum-sharing small cells are uniformly distributed in a cellular macro-cell with randomly distributed SUEs in a circular SBS. All the parameters are as provided in Table A.3 unless otherwise specified. The performance of the simulation-based JEERA algorithm is compared to the analytical DDEEARA algorithm. The algorithms for PD-NOMA and SCMA are similar to the ones in [14] and [18] respectively, albeit the EE resource allocation proposed in this work. Furthermore, we incorporate resource allocation schemes of [27] and [31] in the comparative analysis. Particularly, the CA and UP algorithms in PD-SCMA are compared with the modified matching game and submodularity approach-based CA while PA is compared with the SCA-based DC method of PA in [27]. In addition, the proposed EE schemes are compared with the resource management schemes in D2D enabled cellular hybrid network of [31], albeit for a downlink system. Here, SCOA algorithm is compared with the proposed CA based on equal power allocation, interference constraint and minimum data rate constraint, OMSP algorithm with RACBS-D2D and lastly, QAPA with GWF algorithm.

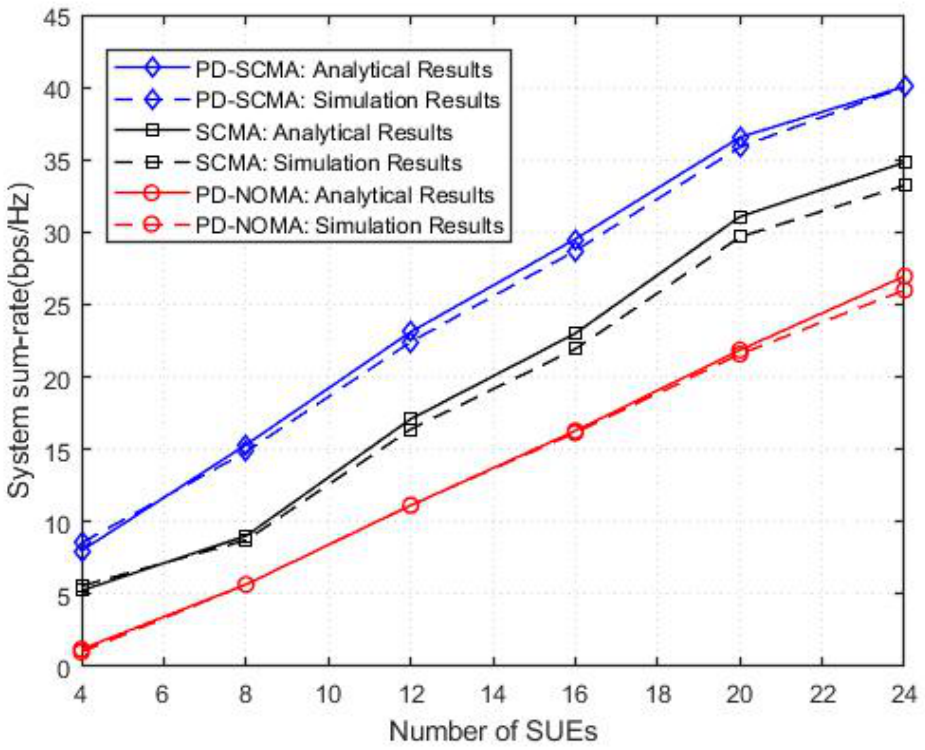

Fig. A.3: Sum-rate vs number of SUEs for different schemes

Figs.A.3 and A.4 present the results of performance analysis. From Fig.A.3, it is observed that the

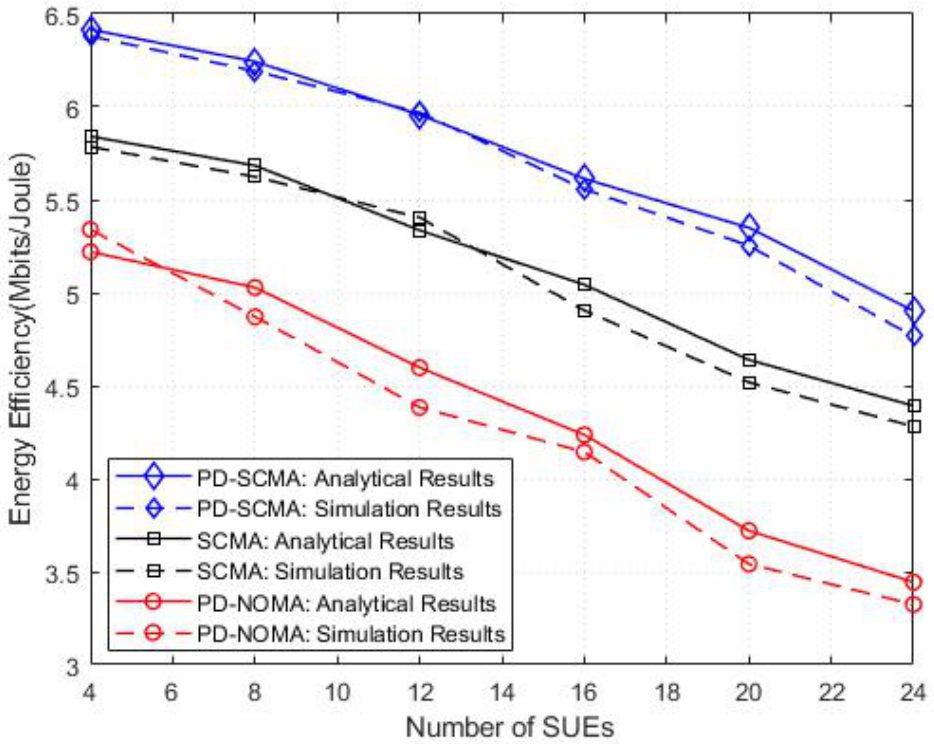

Fig. A.4: Energy efficiency vs number of SUEs for different schemes

system sum-rate increases with the number of SUEs in all the schemes. PD-SCMA performs better than SCMA, then PD-NOMA in that order. This can be attributed to the efficient spectral RUs utilization from non-exclusive allocation of codebooks performed at the transmitter therefore eliminating the codebook interference in PD-SCMA. Additionally, the deployment of the combined CA, UP and PA schemes further enhances PD-SCMA sum-rate performance compared to SCMA with similar CA and PD-NOMA with similar PA schemes. In PD-SCMA, increased number of SUEs and therefore codebooks translates to increased number of RUs. Consequently, more RUs can be assigned to each user, therefore, the sum rate increases. Furthermore, the performance of the simulated and analytical schemes is comparatively close. In Fig.A.4, it can be observed that as the number of users increases the EE of the systems decreases across all the schemes. In comparison to the sum-rate performance, the simulated and analytical results are comparable for the protocols. Although the EE is higher in the beginning, it starts to deteriorate as more users transmit within the same SBS coverage. Additional users beyond the saturation limit compromise the performance of the system. Note that the sum-rate is closely linked to the EE protocols behavior and henceforth, we use either one of them in subsequent discussions.

Fig.A.5 depicts the system sum-rate performance for PD-SCMA with the proposed resource allocation schemes employed against signal to noise ratio (SNR). The results show that the sum-rate increases

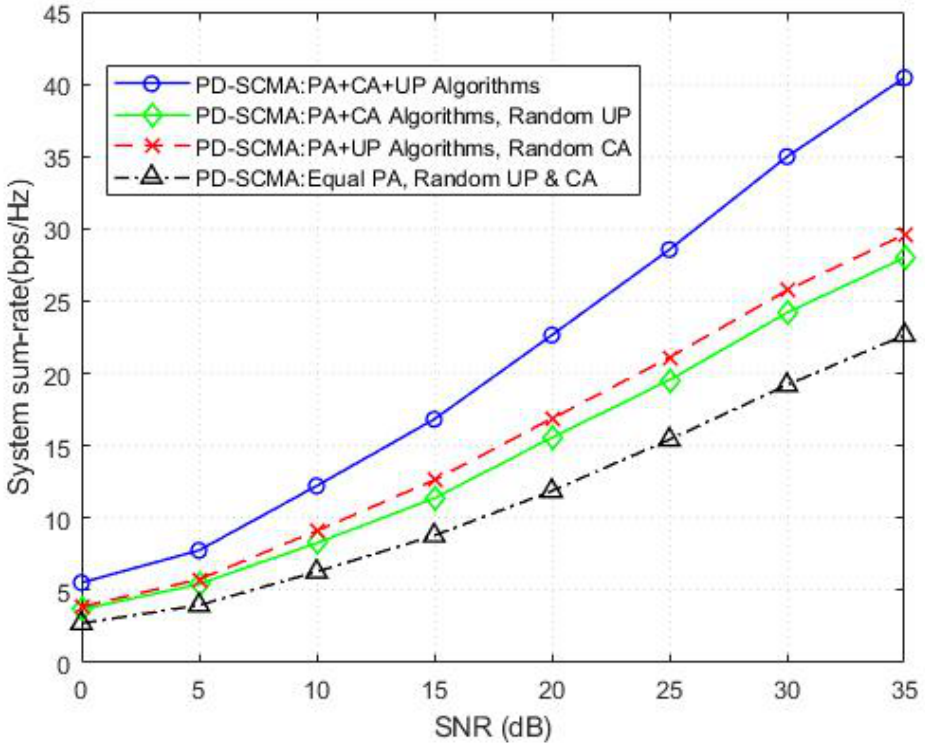

Fig. A.5: Sum-rate vs SNR for PD-SCMA resource allocation schemes

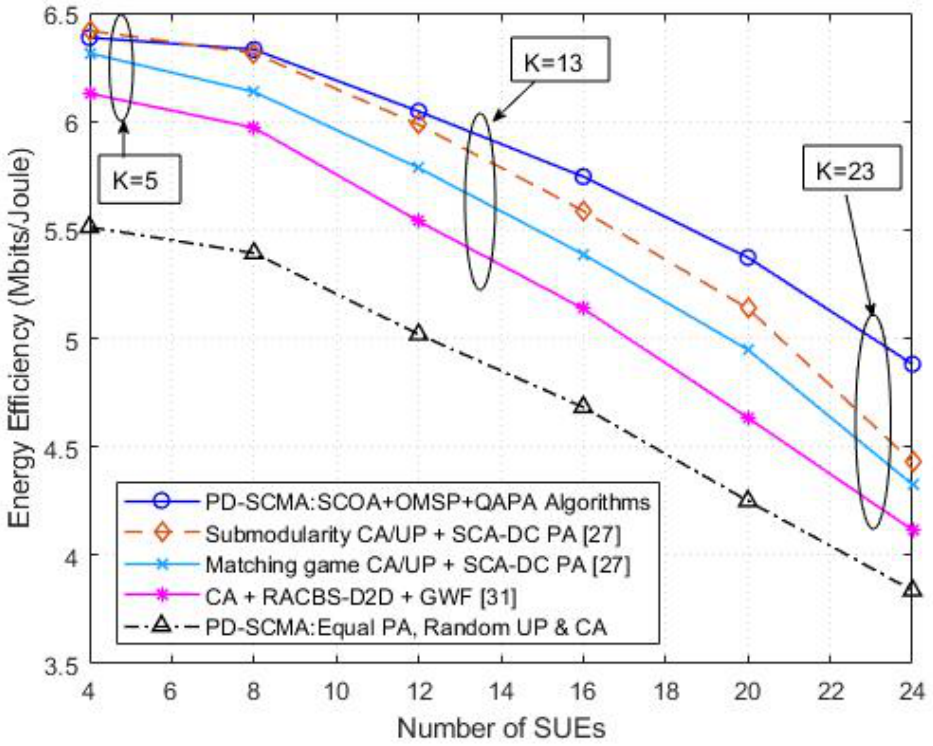

Fig. A.6: EE resource allocation schemes vs number of SUEs

monotonically with the SNR. In fact, it can be observed that PD-SCMA with CA, UP and PA schemes achieve a higher sum-rate compared to the performance when any of the other schemes is used. This confirms the significant performance improvement of applying the CA and UP schemes, albeit higher performance with UP scheme. This confirms the importance of optimal pairing of users in a codebook which gives a considerate performance improvement as opposed to the applied CA scheme.

Fig.A.6 illustrates the EE performance of the proposed resource allocation schemes versus the number of SUEs in a SBS. It can be observed that the system EE significantly decreases with increase in number of SUEs across all the resource allocation scheme considerations. As expected, the system exhibits significant performance improvement when the proposed EE resource allocation algorithms are applied and compared to SUE equal PA with random CA and UP. The proposed EE schemes outperforms matching and submodularity resource allocation-based schemes of [27], specifically for an increasing number of SUEs due to the interference suppression during pairing employed by the OMSP algorithm. A performance difference in the range of 8% for few users to 20% for higher number of users, in comparison with the efficient resource management schemes of [31], can be observed. Despite efficient utilization of spectral resources, power dissipation in the system increases faster than the system EE. Furthermore, PD-SCMA with the proposed EE resource allocation schemes

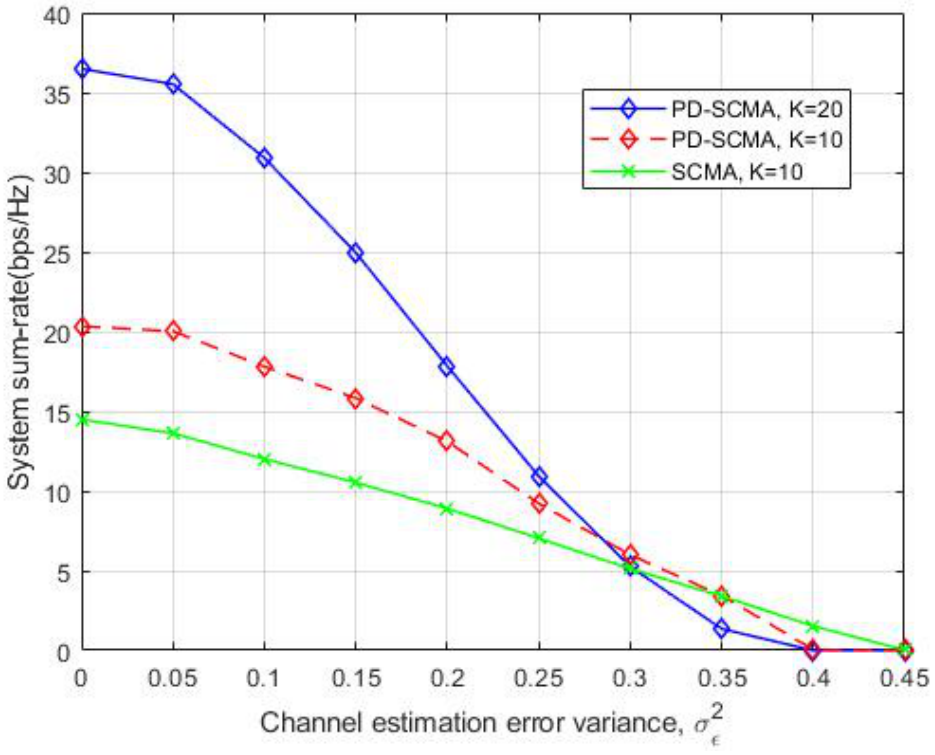

Fig. A.7: Sum-rate capacity vs channel estimation error

achieve higher performance compared to uniform PA and random CA and UP.

Fig.A.7 shows the system sum-rate capacity performance versus channel estimation error variance, σ_{ε}^2. There is a sharp decrease in the system sum-rate capacity. Furthermore, PD-SCMA with 20 users experience much degradation than with 10 users. SCMA with 10 users give a much slower degradation than PD-SCMA. As the number of users increase, the system experiences higher CSI imperfection from the additional noise terms. For enhanced performance, careful trade-off between CSI error variance and the number of users is required.

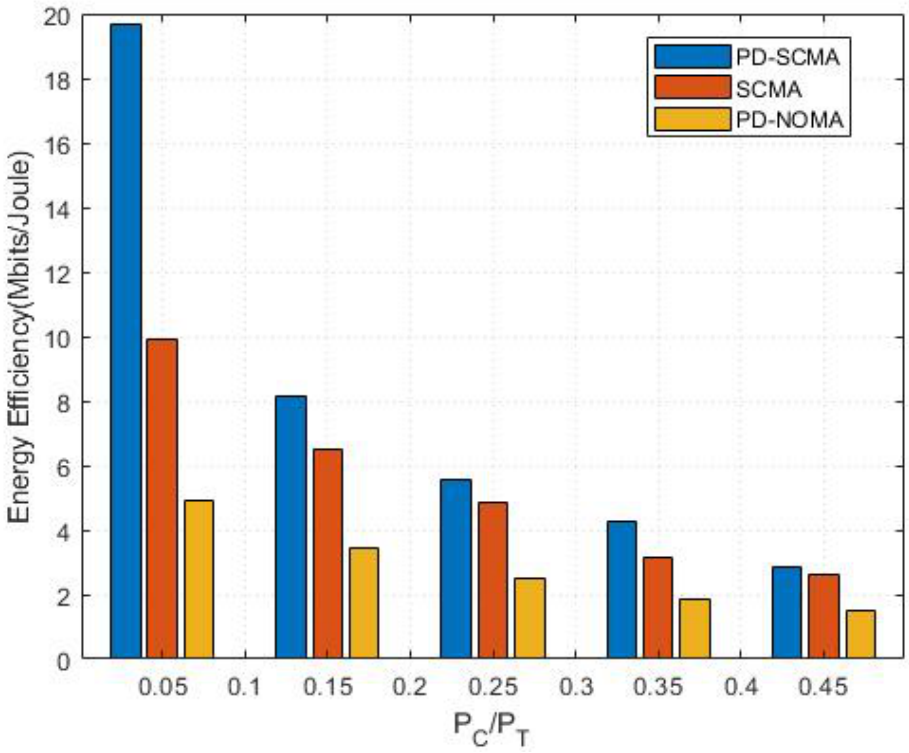

Fig. A.8: Energy efficiency vs circuit to total transmit power ratio

The EE performance versus circuit power to SUE transmit power ratio P_c/P_T is presented in Fig.A.8. With maximum SUE transmit power, $P_{max}^{SUE} = 25dBm$, the system performs less energy-efficiently when the circuit power ratio increases for all the schemes. Following the relation in (A.16) of energy efficiency, its value will become smaller when P_c increases. However, the PD-SCMA system equipped with the proposed resource allocation algorithms still outperforms SCMA and PD-NOMA schemes.

In Fig.A.9, the EE performance of the proposed PD-SCMA simulation and analytical schemes at different number of iterations in comparison to SCMA and PD-NOMA is shown. The results indicate that all schemes converge within 600 iterations, reiterating the convergence analysis of Section 5.1. However, PD-SCMA schemes converge to an optimal solution with higher EE compared to SCMA

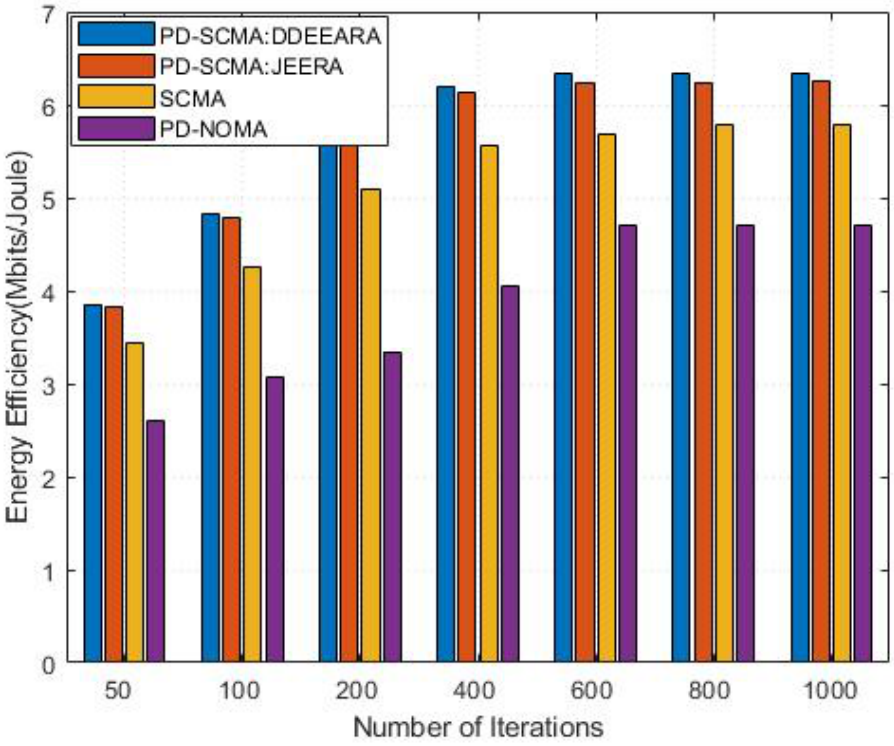

Fig. A.9: Energy efficiency vs Iterations

and PD-NOMA. As such, PD-SCMA is feasible for practical implementation.

Lastly, the computational complexity versus the number of SUEs for the proposed algorithm is shown in Fig.A.10. The total PD-SCMA complexity analyzed is the combinational complexity of the proposed schemes. It is observed that the computational complexity increases with an increase in number of SUEs. Although PD-SCMA provides improved EE and sum-rate performances, the implementation suffers increased complexity cost than SCMA and PD-NOMA. This is largely attributed to the number of mathematical computations required at the MUD receiver, MUE-SUE pairing and PD-SCMA overloading. However, PD-SCMA exhibits lower complexity than PSMA of [36] which implements the original MPA receiver

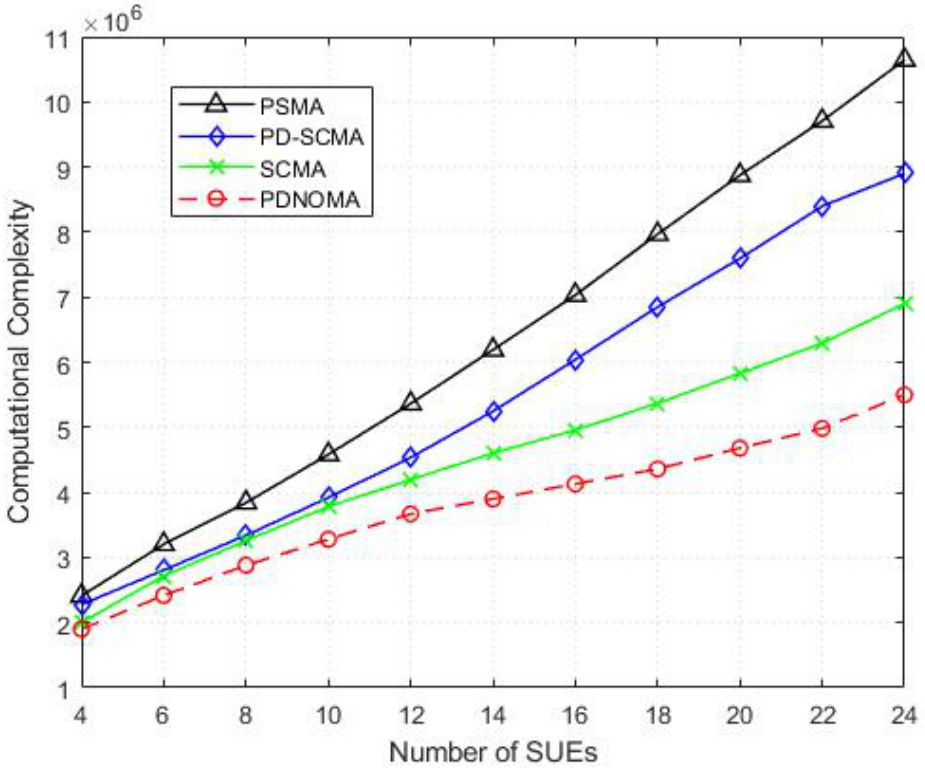

Fig. A.10: Computational complexity vs number of SUEs

7 Conclusion

In this paper, a hybrid PD-SCMA HetNet model that combines power and code domain to multiplex MUEs and SUEs on 5G uplink networks was developed. The model employs SCOA for UP, OMSP for CA and QAPA for PA algorithms. The designed receiver utilizes a low-complexity MUD scheme based on joint SIC and Log- MPA. Simulation and analytical-based EE resource allocation performance for the small cells under QoS constraints of minimum sum-rate, interference temperature, system maximum power is evaluated. Based on the results, the feasibility of hybrid PD-SCMA as a multiplexing technique for MUEs and SUEs for future networks is confirmed. In addition, the performance of PD-SCMA is better compared to that of SCMA and PD-NOMA with or without channel estimation error albeit with increased complexity. The developed schemes namely: SCOA for UP, OMSP for CA and QAPA for PA greatly improve the performance of PD-SCMA. Our future work will be to investigate the multiplexing of more SUEs and MUEs on a codebook and the subsequent development of the MUD schemes for such a system. Furthermore, the developed protocols should be investigated from the computational model's perspective. Besides, other performance parameters like computational times will need to be explored in detail.

References

[1] Y. Liu, Z. Qin, M. Elkashlan, A. Nallanathan, and J. 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] 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.

[3] L. Ping, L. Liu, K. Wu, and W. Leung, "Interleave Division Multiple-Access," *IEEE Transactions on Wireless Communications*, vol. 5, no. 4, pp. 938–947, 2006.

[4] S. Islam, N. Avazov, O. Dobre, and K. 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, 2016.

[5] 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*, vol. 20, no. 3, pp. 2294–2323, 2018.

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

[7] A. Farhadi Zavleh and H. Bakhshi, "Resource Allocation in Sparse Code Multiple Access-based Systems for Cloud-radio Access network in 5g networks," *Transactions on Emerging Telecommunications Technologies*, p. e4153, 2020.

[8] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel Low-Density Signature for Synchronous CDMA Systems over AWGN Channel," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1616–1626, 2008.

[9] Z. Yuan, G. Yu, W. Li, Y. Yuan, X. Wang, and J. Xu, "Multi-User Shared Access for Internet of Things," ser. 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring), 2016, pp. 1–5.

[10] 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, 2016.

[11] A. Hooshiary, P. Azmi, and N. Mokari, "Energy efficient transmission in dynamic pdma-based systems with rf energy harvesting," *Transactions on Emerging Telecommunications Technologies*, vol. 31, no. 4, p. e3923, 2020.

[12] M. S. Ali, H. Tabassum, and E. Hossain, "Dynamic User Clustering and Power Allocation for Uplink and Downlink Non-orthogonal Multiple Access (NOMA) Systems," *IEEE Access*, vol. 4, pp. 6325–6343, 2016.

[13] 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.

[14] 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.

[15] M. Moltafet, P. Azmi, M. R. Javan, N. Mokari, and A. Mokdad, "Optimal radio resource allocation to achieve a low ber in pd-noma–based heterogeneous cellular networks," *Transactions on Emerging Telecommunications Technologies*, vol. 30, no. 5, p. e3572, 2019.

[16] A. Alnoman, S. Erkucuk, and A. Anpalagan, "Sparse Code Multiple Access-based Edge Computing for IoT Systems," *IEEE Internet of Things Journal*, 2019.

[17] N. M. Balasubramanya, S. Payami, and M. Sellathurai, "Uplink Resource Allocation for Shared LTE and SCMA IoT Systems," ser. 2018 IEEE 87th Vehicular Technology Conference (VTC Spring), 2018, pp. 1–5.

[18] 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, 2017.

[19] H. Zhang, C. Jiang, N. C. Beaulieu, X. Chu, X. Wen, and M. Tao, "Resource Allocation in Spectrum-Sharing OFDMA Femtocells with Feterogeneous Services," *IEEE Transactions on Communications*, vol. 62, no. 7, pp. 2366–2377, 2014.

[20] T. Lee, H. Kim, J. Park, and J. Shin, "An Efficient Resource Allocation in OFDMA Femtocells Networks," ser. 2010 IEEE 72nd Vehicular Technology Conference (VTC-Fall), 2010, pp. 1–5.

[21] S. Chege and T. Walingo, "Resource Allocation for Uplink SCMA NOMA in Heterogeneous Networks," ser. 2019 IEEE AFRICON, 2019, pp. 1–6.

[22] M. Dabiri and H. Saeedi, "Dynamic SCMA Codebook Aassignment Methods: A Comparative Study," *IEEE Communications Letters*, vol. 22, no. 2, pp. 364–367, 2017.

[23] Y. Li, M. Sheng, Z. Sun, Y. Sun, L. Liu, D. Zhai, and J. Li, "Cost-Efficient Codebook Assignment and Power Allocation for Energy Efficiency Maximization in SCMA Networks," ser. 2016 IEEE 84th Vehicular Technology Conference (VTC-Fall), 2016, pp. 1–5.

[24] L. Ziyang, C. Wen, W. Fan, W. Feng, X. Xiuqiang, and C. Yan, "Joint Codebook Assignment and Power Allocation for SCMA based on Capacity with Gaussian Input," ser. 2016 IEEE/CIC International Conference on Communications in China (ICCC), 2016, pp. 1–6.

[25] M. Moltafet, P. Azmi, A. K. Kelayeh, and M. Foruzesh, "Resource Allocation in SCMA-based System," *The Modares Journal of Electrical Engineering*, vol. 15, no. 1, pp. 9–14, 2015.

[26] D. Zhai, "Adaptive Codebook Design and Assignment for Energy Saving in SCMA Networks," *IEEE Access*, vol. 5, pp. 23 550–23 562, 2017.

[27] A. Zakeri, M. Moltafet, and N. Mokari, "Joint Radio Resource Allocation and SIC Ordering in NOMA-Based Networks using Submodularity and Matching Theory," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9761–9773, 2019.

[28] Y. Dai, M. Sheng, K. Zhao, L. Liu, J. Liu, and J. Li, "Interference-Aware Resource Allocation for D2D Underlaid Cellular Network using SCMA: A Hypergraph Approach," ser. 2016 IEEE Wireless Communications and Networking Conference, 2016, pp. 1–6.

[29] H. M. Kim, H. V. Nguyen, G.-M. Kang, Y. Shin, and O.-S. Shin, "Device-to-Device Communications Underlaying an Uplink SCMA System," *IEEE Access*, vol. 7, pp. 21 756–21 768, 2019.

[30] K. Zhao, Y. Shi, Y. Dai, L. Liu, J. Liu, M. Sheng, and J. Li, "Resource Allocation in Device-to-Device Communication Underlaid Cellular Network Using SCMA: An Opportunistic Approach," ser. 2015 IEEE International Conference on Communications in China (ICCC), 2015, pp. 1–6.

[31] A. Sultana, I. Woungang, A. Anpalagan, L. Zhao, and F. Lilatul, "Efficient Resource Allocation in SCMA-Enabled Device-to-Device Communication for 5g Networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5343–5354, 2020.

[32] V. Kalokidou, O. Johnson, and R. Piechocki, "A Hybrid TIM-NOMA Scheme for the SISO Broadcast Channel," ser. 2015 IEEE International Conference on Communication Workshop (ICCW), 2015, pp. 387–392.

[33] 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, 2016.

[34] 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.

[35] V. K. Trivedi, K. Ramadan, P. Kumar, M. I. Dessouky, and F. E. Abd El-Samie, "Trigonometric Transforms and Precoding Strategies for OFDM-based Uplink Hybrid multi-carrier Nonorthogonal Multiple Access," *Transactions on Emerging Telecommunications Technologies*, vol. 30, no. 12, p. e3694, 2019.

[36] 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, 2017.

[37] T. Sefako and T. Walingo, "Biological Resource Allocation Algorithms for Heterogeneous Uplink PD-SCMA NOMA Networks," *IEEE Access*, vol. 8, pp. 194 950–194 963, 2020.

[38] S. Bu, F. R. Yu, and H. Yanikomeroglu, "Interference-Aware Energy-Efficient Resource Allocation for OFDMA-based Heterogeneous Networks with Incomplete Channel State Information," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 3, pp. 1036–1050, 2014.

[39] G. Liu, H. Zhao, and D. Li, "Resource Allocation in Heterogeneous Networks : A Modified Many-to-One Swap Matching," ser. 2017 IEEE 17th International Conference on Communication Technology (ICCT), 2017, pp. 508–512.

[40] M. A. Sedaghat and R. R. Müller, "On User Pairing in Uplink NOMA," *IEEE Transactions on Wireless Communications*, vol. 17, no. 5, pp. 3474–3486, 2018.

[41] F. W. Murti and S. Y. Shin, "User Pairing Schemes based on Channel Quality Indicator for Uplink Non-orthogonal Multiple Access," ser. 2017 9th International Conference on Ubiquitous and Future Networks (ICUFN), 2017, pp. 225–230.

[42] L. Chen, L. Ma, and Y. Xu, "Proportional Fairness based User Pairing and Power Allocation Algorithm for Non-Orthogonal Multiple Access System," *IEEE Access*, vol. 7, pp. 19 602–19 615, 2019.

[43] J. Liu, G. Wu, S. Xiao, X. Zhou, G. Y. Li, S. Guo, and S. Li, "Joint Power Allocation and User Scheduling for Device-to-Device-Enabled Heterogeneous Networks With Non-Orthogonal Multiple Access," *IEEE Access*, vol. 7, pp. 62 657–62 671, 2019.

[44] Z. Ding, P. Fan, and H. V. Poor, "Impact of User Pairing on 5G Non Orthogonal Multiple Access Downlink Transmissions," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6010–6023, 2015.

[45] 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 Transactions on Communications*, vol. 62, no. 7, pp. 2366–2377, 2014.

[46] F. Fang, H. Zhang, J. Cheng, S. Roy, and V. C. M. Leung, "Joint User Scheduling and Power Allocation Optimization for Energy Efficient NOMA Systems With Imperfect CSI," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 12, pp. 2874–2885, December 2017.

[47] 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," ser. 2014 IEEE Global Communications Conference, 2014, pp. 4782–4787.

[48] W. B. Ameur, P. Mary, M. Dumay, J. Hélard, and J. Schwoerer, "Performance Study of MPA, Log-MPA and MAX-Log-MPA for an Uplink SCMA Scenario," ser. 2019 26th International Conference on Telecommunications (ICT), 2019, pp. 411–417.

[49] B. Wang, K. Wang, Z. Lu, T. Xie, and J. Quan, "Comparison Study of Non-Orthogonal Multiple Access Schemes for 5G," ser. 2015 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, 2015, pp. 1–5.

[50] A. Syed, N. Raza, P. Haris, A. H. Syed, L. Musavian, N. Qiang, A. I. Muhammad, G. Xiaohu, and T. Rahim, "Energy-Aware Radio Resource Management in D2D-Enabled Multi-Tier Hetnets," *IEEE Access*, vol. 6, no. 7, pp. 16 610–16 622, 2018.

[51] S. Hung, P. Li, K. Feng, and Y. Lin, "Joint Wireless Charging and Hybrid Power based Resource Allocation for LTE-A Wireless Network," ser. 2017 IEEE Wireless Communications and Networking Conference (WCNC), 2017, pp. 1–6.

[52] W. Dinkelbach, "On Nonlinear Fractional Programming," *Management Science*, vol. 13, no. 7, p. 492–498, 1967.

[53] M. S. Al-Janabi, C. C. Tsimenidis, B. S. Sharif, and S. Y. L. Goff, "Optimized Resource Allocation Strategy for a Single-cell downlink AMC-OFDMA Systems," ser. 2013 IEEE 20th International Conference on Electronics, Circuits, and Systems (ICECS), 2013, pp. 209–212.

[54] H. Zheng, H. Li, S. Hou, and Z. Song, "Joint Resource Allocation with Weighted Max-Min Fairness for NOMA-Enabled V2X Communications," *IEEE Access*, vol. 6, pp. 65 449–65 462, 2018.

[55] A. M. Rush and M. Collins, "A Tutorial on Dual Decomposition and Lagrangian Relaxation for Inference in Natural Language Processing," *Journal of Artificial Intelligence Research*, vol. 45, pp. 305–362, 2012.

[56] G. Yue, N. Prasad, and S. Rangarajan, "Iterative MMSE-SIC Receiver with Low-Complexity Soft Symbol and Residual Interference Estimations," ser. 2013 Asilomar Conference on Signals, Systems and Computers, Nov 2013, pp. 2113–2117.

[57] J. Liu, G. Wu, S. Li, and O. Tirkkonen, "On Fixed - Point Implementation of Log - MPA for SCMA Signals," *IEEE Wireless Communications Letters*, vol. 5, no. 3, pp. 324–327, June 2016.

Part III

Paper B

Paper B

Multiplexing Capacity of Hybrid PD-SCMA Heterogeneous Networks

Simon Chege and Tom Walingo

Published in IEEE Transactions on Vehicular Technology
doi: 10.1109/TVT.2022.3162304

Abstract

Hybrid multiple access schemes are considered potential technologies towards achieving optimal spectrum sharing for future heterogeneous networks. Multiple users are multiplexed on a single resource unit in the code domain for sparse code multiple access (SCMA) or power domain for power domain non-orthogonal multiple access (PD-NOMA) and in both domains for the hybrid power domain sparse code non-orthogonal multiple access (PD-SCMA). This allows for effective spectrum usage but comes at a cost of increased detection complexity resulting in loss of performance in terms of outages. It is therefore imperative to determine the user multiplexing capacity for effective performance. This work investigates codebook capacity bounds in small cells, pairing and power capacity bounds for the number of small cell user equipment's (SUEs) and macro cell user equipment's (MUEs) that can be multiplexed on a codebook for the developed PD-SCMA technology. Closed-form solutions for codebook, pairing and power multiplexing capacity bounds are derived. The performance of the system results into low outage when the system's point of operation is within the multiplexing bounds. To alleviate the resource allocation (RA) challenges of such a system at the transmitter, dual parameter ranking (DPR) and alternate search method (ASM) based RA schemes are proposed. The results show significant capacity gain with DPR-RA in comparison with conventional schemes.

1 Introduction

The demand for increased capacity for 5G networks has fueled a paradigm shift from orthogonal multiple access (OMA) to Non-Orthogonal Multiple Access (NOMA) techniques, namely, power-domain NOMA (PD-NOMA) and code-domain NOMA (CD-NOMA). Recent progress on PD-NOMA performance is reviewed in [1] while a variant scheme referred to as NOMA-2000 is investigated in [2]. Sparse code multiple access (SCMA) is a promising CD-NOMA technique offering overloading, sparse codewords and shaping gain contributing to multiple dimensions for multiplexing [3], [4]. Due to their heterogeneous nature and performance requirements, future networks will blend multiple access schemes resulting in hybrid NOMA schemes. Downlink transceiver schemes combining PD-NOMA and CD-NOMA with SCMA are proposed in [5], [6] and [7]. In particular, a combination of PD-NOMA and SCMA on a heterogeneous multi-tier network (HetNet) consisting of small cell user equipment's (SUEs) and the macro user equipment's (MUEs) yields a novel hybrid power domain sparse code non-orthogonal multiple access (PD-SCMA) scheme. Recent works in [5], [7] [8], and [9] show that PD-SCMA exhibits spectral and energy efficiency performance superiority over PD-NOMA and SCMA.

In hybrid PD-SCMA, MUEs and SUEs are co-multiplexed using PD-NOMA into a codebook, while allocating different power levels and hence applying successive interference cancellation (SIC) for their detection at the receiver. Furthermore, different codes are exclusively assigned to the MUE-SUE clusters hence employing a modified log-domain message passing algorithm (log-MPA) for interference cancellation. The codebook and pairing interference, resource unit (RU) limitation and receiver design complexity constraints necessitate an investigation on the multiplexing capacity performance and bounds of the number of MUEs and SUEs that can be multiplexed in one codebook for specified outage and Quality of Service (QoS) guarantees.

Multiplexing capacity has been studied for various NOMA schemes. For PD-NOMA schemes, closed-form capacity and outage expressions can be derived due to the linear successive decoding algorithms employed [10]. On the contrary, the explicit outage and capacity expressions of uplink SCMA related systems are challenging due to the iterative multi-user detection (MUD) nature and the interactions among the multi-dimensional signals from different users, unlike with the successive signal detection in conventional PD-NOMA. Based on the capacity region of the multiple access channel (MAC), the capacity region theory of an uplink SCMA has been derived in [11]. In [12], the authors investigate admissible user bounds in uplink SCMA subject to user-rate constraints. Lower bounds of user capacity considering maximal ratio combining receiver (MRC) and SIC-MRC

receivers are presented. The per-tier outage probability and capacity considering log-normal fading channels for small cells and Rayleigh fading channels for the macro-cell in an SCMA system is derived in [13].

In the power domain sparse code multiple access (PSMA) of [5], SUEs multiplexed in a codebook in each small cell are detected by employing original MPA and SIC. In [6], the authors formulate a modified optimization problem by introducing a SIC ordering optimization variable besides codebook and power allocation (PA). The result is an improved SIC ordering algorithm based on network, channel gain, and available resource conditions. In [7], an investigation of user pairing (UP), joint impact of the non-uniform power allocation and multi-dimensional code sparsity on the wireless systems is conducted. In [8], a HetNet scheme that pairs one SUE with at most one MUE and transmits over a codebook was considered. To enhance codebook re-use beyond the work presented in [8], this work investigates the bounds of pairable MUEs in a codebook subject to QoS constraints and link requirements for massive connectivity. The capacity of these systems have not been effectively evaluated. Furthermore, PD-SCMA features cross-tier codebook resource re-use by the SUEs and MUEs resulting in user pairing interference. This leads to the following research questions: 1) How many codebooks can we simultaneously assign to SUEs in a small cell?, 2) How many MUEs can be paired with a SUE on a codebook subject to an acceptable outage and QoS at the receiver? and 3) How many power levels for guaranteed optimal SIC at each codebook? These questions have not been tackled and as such, they are the focus of this work.

The multiplexing of MUEs and SUEs on a PD-SCMA system at the transmitter necessitates efficient codebook assignment, power allocation and MUE vs SUE pairing policies. For codebook assignment (CA), optimal codebook generation and capacity enhancement is required. Typical codebook generation operations include complex conjugation, phase rotation and dimensional permutation [12]. Optimizing these operators makes the colliding users at the RU distinct, easing the decoding process thus enhancing the multiplexing capacity. Increasing the dimensionality for enhanced codebook capacity has received much attention in recent times. In [14], an overloading technique that transmit extra users with one RU is proposed albeit increased complexity of order one. In [15], design of high dimensional codebooks involves increasing the rows and columns of the resource mapping matrix for a constant overloading factor and complexity per user. In [16], the constraint in SCMA that a fixed number of nonzero elements per codeword is relaxed. Instead, a user could use a varying number of RUs to allow for a flexible system design. These techniques of increasing codebook capacity are limited by the overloading factor and per-user complexity. By utilizing measures of sparsity [17], [18] and their attributes, we propose codebook capacity bounds, i.e., the

number of admissible codebooks in a PD-SCMA system for assignment to SUEs.

In each codebook assigned to a SUE, we propose to superimpose multiple MUEs in PD-SCMA scheme through a diversified power allocation technique. SIC capacity is bounded by complexity, number of distinct power levels, QoS requirements and outage performance. In particular, a PA technique that maximizes the number of power levels to enable multiple user SIC decoding ensures adequate outage performance at the receiver. In [19], PA constraint requirements in downlink and uplink systems for effective SIC are outlined. In an uplink cluster, users experience distinct channel conditions. At the base station (BS), SIC performance relies on the distinctiveness of the received signals. A PA based on power back-off scheme to ensure diverse arrived power in uplink NOMA is proposed in [20]. Feasibility conditions in an energy efficiency maximization problem for effective SIC are investigated in [21]. Authors in [22] redefine user fairness index that measures the difference between the rates that can be achieved by the users and the fair rates suggested by the power distribution among them, a predominant parameter in interference management and detection process. A dynamic resource allocation (RA) that neither fixes the number of multiplexed users on each RU nor restricts the rate assignments on sub-carriers of each user is proposed in [23]. A general PA scheme that strictly guarantees performance gain in downlink and uplink NOMA systems is proposed in [24]. In particular, the exact expressions for the outage probability and the average rate achieved by the proposed scheme are derived based on the PA coefficients. However, the PA schemes do not address a HetNet scenario of MUE-SUE codebook re-use. We propose HetNet QoS aware user-power ranking metrics for PA to SUEs and MUEs in the network.

To minimize the SUE-MUE cluster pairing interference in a codebook, a robust UP policy is imperative. Authors in [25] characterize the impact of UP on the performance of two NOMA systems, NOMA with fixed PA and cognitive radio inspired NOMA. In the proposed hybrid-domain NOMA [26], pairing is achieved by clustering the users in small path loss (strong) and large path loss (weak) groups. Spectral efficiency is maximized through relaxation of RA non-convexity problem constraints by using successive convex approximation (SCA) and re-weighted minimization approaches. A pairing metric based on channel quality difference and minimal pairing interference is proposed in [8] for an SUE-MUE pair in a codebook. In [9], users are clustered by employing alternative biological RA schemes. To further extend the research in [8], this work proposes a UP policy that multiplexes at least one MUE to a SUE in a codebook based on ranking parameters, then cluster users based on a matching game.

At the transmitter, we propose the following dual parameter ranking resource allocation (DPR-RA) techniques: dual parameter ranking for CA (DPR-CA), dual parameter ranking for UP (DPR-UP) and

dual parameter ranking for PA (DPR-PA). DPR-RA utilizes proposed ranking metrics for the dual players i.e., users and resources (code and power), to rank and optimally match each other while employing a matching algorithm. The individual DPR-RA policies are admitted iteratively at the transmitter while exploiting the alternate search method (ASM) [6]. At the receiver, a dynamic ordered low complexity hybrid SIC-log-MPA (HSLM) MUD featuring PD- and CD-NOMA that detects multiple multiplexed users in RU is developed. First, the iterative Log-MPA algorithm detects the transmitted symbols on the codebook. Thereafter, SIC sorting is applied to decode the MUEs and SUE symbols on each codebook. Given a dynamic decoding order based on the user received signal power, the users are decoded linearly employing the minimum mean square error (MMSE) detector. Unlike the receiver in [5], [6] and [7], the complexity order of the proposed MUD reduces significantly for the log-MPA [27]. Since the number of pairable MUEs in a codebook is constrained, outage analysis in SIC capacity estimation is not trivial. This work utilizes the outage performance general expressions proposed in [28] to analyze the codebook outage performance for each decoding order based on the observation of each MUE's received power strength as the paired number MUEs increases.

In summary, this work investigates user multiplexing capacity for a PD-SCMA system. Specifically, we explore the bounds of assigned codebooks in a small base station (SBS) using measures of sparsity. In addition, we investigate the MUE multiplexing bounds in each codebook based on the pairing metric employed and lastly, the maximal number of MUEs based on maximum number of distinct power levels to enable optimal detection at the receiver subject to outage performance and QoS constraints. To improve performance, this work proposes DPR-RA and ASM schemes at the transmitter. An improved low complexity hybrid MUD is employed at the receiver. The performance of the system is investigated considering capacity, outage probability, complexity and convergence that are derived. From the analysis, the codebook capacity is bounded by the positional, overloading and the derived sparsity bounds.. The pairing capacity is bounded by the proposed parameter ranking metric while the power capacity is bounded by the number of distinct power levels to guarantee SIC and minimum QoS. Moreover, we validate the analytical results with numerical based Monte Carlo simulations. The capacity bounds significantly improve with the proposed DPR-RA policy in comparison with prevalent schemes. Lastly, it is shown that the performance of the system results in low outage when the system's point of operation is within the multiplexing bounds.

1.1 Organization

The rest of the paper is organized as follows: the system model is developed in Section II. Section III details the proposed DPR-RA schemes. Section IV outlines the hybrid receiver. The multiplexing capacity bounds are discussed in section V. Outage probability is analyzed in section VI while the computational complexity and convergence of the proposed schemes is presented in section VII. In Section VIII the results and discussions are presented. Finally, Section IX concludes the paper.

1.2 Notation

The notation $*_{f,k,c}^{SUE}$and $*_{m,c}^{MUE}$, both in upper and lower case, represents the relation of $*$ to the k^{th} SUE on codebook c of the f^{th} SBS and the relation m^{th} MUE on codebook c, respectively. $\boldsymbol{I}$ is an identity matrix. The superscript $(\cdot)^H$ and $(\cdot)^{-1}$ represents Hermitian and inverse operator respectively while $|\cdot|$ denotes the absolute value of a scalar.

2 System Model

The system model for an uplink hybrid PD-SCMA based two-tier HetNet model of Fig. B.1 is considered. It comprises of a centralized macro base station (MBS) serving a set of $\mathcal{M} = \{1, \cdots M\}$ randomly distributed macro cell users (MUEs) and underlaid with a set of $\mathcal{F} = \{1, \cdots F\}$ small cells, each characterized by a centralized low power small cell base station (SBS) serving a set of $\mathcal{K} = \{1, \cdots K\}$ uniformly distributed SUEs. MUEs and SUEs are co-multiplexed on the set of available codebooks (CB) from set $\mathcal{C} = \{1, \cdots C\}$ designed from complex mapping of the time-frequency $\mathcal{N} = \{1, \cdots N\}$ RUs. The $N-$dimensional codewords of a codebook are sparse vectors with d_v ($d_v < N$) nonzero entries corresponding to d_v specific RUs for a user.

The SUE transmit signal vector is denoted as $\mathbf{x}_k = \left\{x_{f,k,c}^{SUE}\right\}_{k=1}^{K}$ while the MUE transmit signal vector is $\mathbf{x}_m = \left\{x_{m,c}^{MUE}\right\}_{m=1}^{M}$. Consequently, we define the following policies:

- The SUE codebook assignment policy $\mathbf{Q} = \left[q_{f,k,c}^{SUE}\right]_{F\times K\times C}$ is a matrix representation of the small cells transmitter resource assignment. $q_{f,k,c}^{SUE} = 1$ implies that the k^{th} SUE is assigned codebook c in the f^{th} small cell and $q_{f,k,c}^{SUE} = 0$, otherwise.
- The user pairing policy $\mathbf{A} = \left[A_{k,m}^{c}\right]_{K\times M}$ is a matrix for pairing of the m^{th} MUE to the k^{th} SUE on codebook c. $A_{k,m}^{c} = 1$ when SUE k is paired with the MUE $m \in \mathcal{M}_c$ on codebook c and $A_{k,m}^{c} = 0$ denotes the contrary. Here $\mathcal{M}_c = \{1, \cdots m, \cdots, J\}$ is the set of paired MUEs on codebook c used by SUE k with cardinality $|\mathcal{M}_c| = J$ and $\mathcal{M}_c \subset \mathcal{M}$.

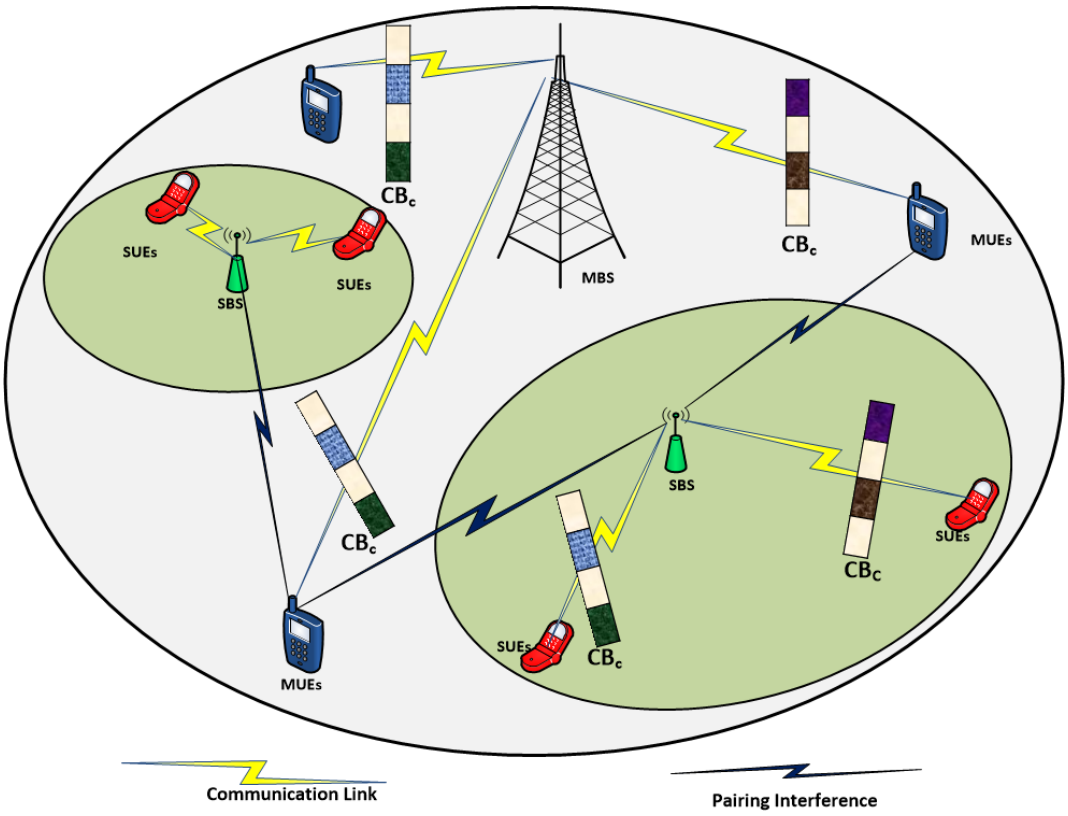

Fig. B.1: Uplink hybrid PD-SCMA HetNet model

- The power allocation policy at the transmitter $\mathbf{P} = \left\{P^{SUE}_{F,K,C}, P^{MUE}_{M,C}\right\}$. Here, $P^{SUE}_{F,K,C} = \left[P^{SUE}_{f,k,c}\right]_{F\times K\times C}$ where $P^{SUE}_{f,k,c}$ is the allocated power to the k^{th} SUE on codebook c of the f^{th} small cell. $P^{MUE}_{M,C} = \left[P^{MUE}_{m,c}\right]_{M\times C}$ where $P^{MUE}_{m,c}$ is the distinct power level allocated to the m^{th} MUE utilizing codebook c.

Let $\mathbf{H} = \left\{H^{SUE}_{F,K,C}, H^{MUE}_{M,C}\right\}$, where $H^{MUE}_{M,C} = \left[h^{MUE}_{m,c}\right]_{M\times C}$ and $H^{SUE}_{F,K,C} = \left[h^{SUE}_{f,k,c}\right]_{F\times K}$. Here, $h^{MUE}_{m,c} = \sqrt{\vartheta_{m,c}}g^{MUE}_{m,c}$ and $h^{SUE}_{f,k,c} = \sqrt{\vartheta_{f,k,c}}g^{SUE}_{f,k,c}$ denotes channel coefficients of MUE and SUE averaged over the d_v RUs in each codebook with $\vartheta_{m,c}$ and $\vartheta_{f,k,c}$ denoting the channel power (i.e. large-scale fading parameter) and $g^{MUE}_{m,c}$ and $g^{SUE}_{f,k,c}$, the small-scale fading assumed to be independent and identically distributed (i.i.d.) complex Gaussian $\mathcal{CN}(0,1)$. Each codebook is mapped to specific RUs. The RUs to codebooks mapping matrix $\rho = \left[\rho_{n,c}\right]_{N\times C}$ represents an assignment such that $\rho_{n,c} = 1$ if codebook c utilizes RU n, $\rho_{n,c} = 0$ otherwise. Each codebook $c \in \mathcal{C}$ contains D sparse codeword vectors with d_v $(d_v \leq N)$ non-zero entries corresponding to d_v specific RUs allocated. The multiplexing model for the SUEs and MUEs is illustrated in Fig. B.2 The resultant ρ matrix is given

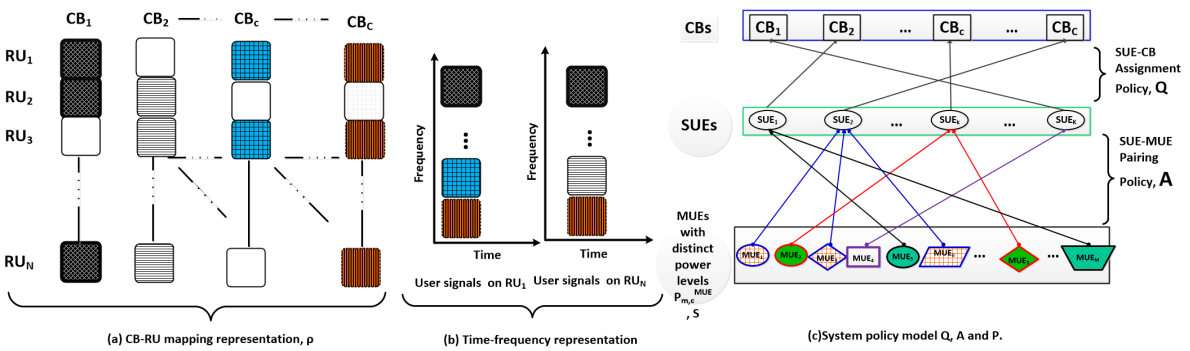

Fig. B.2: Multiplexing model for RU mapping, SUE codebook assignment and MUE pairing.

as

$$\rho = \begin{bmatrix} 1 & 0 & \cdots & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 0 & \cdots & 1 \end{bmatrix} \tag{B.1}$$

As an example, in Fig.B.2 (a), $\rho_{1,1} = 1$, $\rho_{1,c} = 1$ and $\rho_{1,C} = 1$ implies that RU_1 is mapped CB_1, CB_c and CB_C respectively, where $|\mathcal{C}|$ is the maximum number of assigned codebooks. (b) illustrates the time-frequency representation of codebooks utilizing n^{th} RU and lastly, (c) illustrates the transmitter codebook resource matrix $\mathbf{Q}$ given by (B.2) such that SUE_1 is allocated to CB_2, SUE_K to CB_1 and so on.

$$\mathbf{Q} = \begin{bmatrix} 0 & 1 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \tag{B.2}$$

and the user pairing matrix $\mathbf{A} = \left[A^c_{k,m}\right]_{K\times M}, \forall m \in \mathcal{M}_c, \mathcal{M}_c \subset \mathcal{M}$, given by

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & \cdots & 0 & \cdots & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 1 & 0 & 0 & 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \tag{B.3}$$

From Fig. B.2 (c) and matrix $\mathbf{A}$ in (B.3), we can deduce that SUE_2 is paired with MUE_1, MUE_3

and MUE_6 on codebook CB_C.

For the PA policy, MUEs m in a codebook c are allocated distinct power levels such that the codebook PA vector P_c is given as

$$\mathbf{P_c} = \left[P_{1,c}^{MUE} P_{2,c}^{MUE} \cdots P_{m,c}^{MUE} \cdots P_{J,c}^{MUE} P_{f,k,c}^{SUE} \right]^T \tag{B.4}$$

where $P_{m,c}^{MUE} = S, S \in \mathcal{S}$ are the distinct power levels in a set of power levels $\mathcal{S}$. In Fig. B.1 (c), the distinct power levels (shapes and shades) associated with codebook C assigned to SUE_2, $P_C = \left[P_{1,C}^{MUE} P_{3,C}^{MUE} P_{6,C}^{MUE} P_{f,2,C}^{SUE} \right]^T$. The PA policy for the PD-SCMA system is a vector given by

$$\mathbf{P} = \left[\mathbf{P_1 P_2} \cdots \mathbf{P_c} \cdots \mathbf{P_C} \right] \tag{B.5}$$

The bounds on the codebook assignment policy and the number of admissible codebooks $(\mathbf{Q}, C)$, the optimal user pairing scheme and the admissible limit on the pairable number MUEs $(\mathbf{A}, J)$ and lastly, maximal distinct power levels $(\mathbf{P}, S)$ to guarantee optimal number of MUEs S successively decoded, determine the multiplexing capacity and are the subject of this investigation.

Algorithm 7 Transmitter ASM-RA Algorithm

Initialization: Initialize the matrix $\mathbf{Q}(0)$, $\mathbf{A}(0)$ and $\mathbf{P}(0)$ at iteration $t = 0$ and maximum iterations t_{max}

1: **while** $t \leq t_{max}$ or Convergence is false **do**
2: **for** SBS $f = 1 : F$ **do**
3: Determine DPR-PA $P_{F,K,C}^{SUE}(t)$ eqn. (B.11), (B.12) and Algorithm 9
4: Determine DPR-CA $\mathbf{Q}(t)$ eqn. (B.6), (B.7) and Algorithm 9
5: **end for**
6: Determine DPR-PA $P_{M,C}^{MUE}(t)$ eqn. (B.13), (B.14) and Algorithm 9
7: Determine DPR-UP $\mathbf{A}(t)$ eqn. (B.8), (B.9) and Algorithm 9
8: Set $t = t + 1$ and update $Q(t)$, $A(t)$, $P(t)$.
9: **end while**
10: **RETURN** $\mathbf{Q}, \mathbf{A}, \mathbf{P}$

3 Dual Parameter Ranking Resource Allocation

In this section, the proposed hybrid PD-SCMA resource allocation scheme is presented. Based on the ASM [6], the transmitter jointly and iteratively allocates individual resources in PA, CA and UP as in Algorithm 7. The CA, UP and PA resource allocation policies are formulated as dual parameter

ranking (DPR) problems and solved using a one-to-many matching game algorithm proposed for stable matching [29], [30], [31] as DPR-CA, DPR-UP and DPR-PA respectively. The DPR observes players X and Y of sets $\mathcal{X}$ and $\mathcal{Y}$ respectively, ranks and matches them accordingly based on the preference metrics. Denote by $\Theta_X(Y) = \psi^{RA}_{Y\to X}$ the ranking of player Y by player X. Similarly, let $\Theta_Y(X) = \psi^{RA}_{X\to Y}$ denote the ranking of player X by player Y. The preference metrics $\Theta_X(Y)$ and $\Theta_Y(X)$ are defined for each DPR-RA policy in the subsequent sub-sections.

The detailed individual DPR-RA matching game-based algorithm for the RA policies is presented in Algorithm 9 of Appendix 10.1. The convergence of the DPR-RA algorithm can be verified by observing the preference formulation of players. Preference relations of players $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$ are fixed for a given PA. Hence, given fixed preference relations of $\mathcal{X}$ and $\mathcal{Y}$, the DPR-RA algorithm is known as the deferred acceptance algorithm in the two-sided matching which converges to a stable matching denoted by μ^*_{RA} [30] i.e., μ^*_{CA}, μ^*_{UP} and μ^*_{PA} for CA, UP and PA respectively.

3.1 Codebook Assignment,$\{\rho, \mathbf{Q}\}$, DPR-CA

The RUs to codebooks mapping matrix ρ is usually predetermined and fixed [4], [15]. The determination of the number of codebooks C is detailed in the multiplexing section, Section V.B. In the DPR-CA problem, codebook set $\mathcal{C}$ represents set of players $\mathcal{X}$, while set of SUE's $\mathcal{K}$ represents set of players $\mathcal{Y}$. A codebook $c \in \mathcal{C}$ ranks the SUEs employing the achievable rate $R^{SUE}_{f,k,c}$ at the receiver, determined in Section 4, while SUE $k \in \mathcal{K}$ ranks the codebooks employing a channel quality-based ranking metric (C.4). The ranking metrics are given by

$$\psi^{CA}_{c\to k} = R^{SUE}_{f,k,c} \tag{B.6}$$

$$\psi^{CA}_{k\to c} = \sqrt{\frac{1}{d_v}\sum_{i\in N_c}\frac{P^{SBS}_{max}}{K}\frac{\left|g^{SUE}_{f,k,i}\right|^2}{\sigma^2_{f,k,i}}} \tag{B.7}$$

where P^{SBS}_{max} denotes maximum SBS transmit power and N_c is the set of d_v RUs utilized by codebook c. Consequently, for the DPR-CA $\Theta_X(Y) = \psi^{CA}_{k\to c}$, $\Theta_Y(X) = \psi^{CA}_{c\to k}$ and $J_{CA} = 1$. It can be deduced that the metric $\psi^{CA}_{c\to k}$ is maximized for a SUE with high gain on most of the RUs and close to the SBS and minimized for a poor gain SUE at the cell edge. Similarly, a codebook prefers a SUE with high achievable rate. The output of the algorithm is a stable codebook-SUE matching μ^*_{CA} for the CA assignment policy $\mathbf{Q} = \left[q^{SUE}_{f,k,c}\right]_{F\times K\times C}$.

3.2 User Pairing,$\{\mathbf{A}\}$, DPR-UP

The SUE-MUEs pairing is defined by the pairing policy matrix $\mathbf{A}$ such that $\mathbf{A} = \left[A^c_{m,c} = 1\right]_{K\times M}$. Similar to DPR-CA, the DPR-UP problem is modelled as a one-to-many matching game with set $\mathcal{X}$

and set $\mathcal{Y}$ representing SUE set $\mathcal{K}$ and MUE set $\mathcal{M}$ respectively. The ranking metrics are given as follows

$$\psi_{m\to k}^{UP} = R_{f,k,c}^{SUE} \tag{B.8}$$

$$\psi_{k\to m}^{UP} = \frac{P_{m,c}^{MUE}\left|h_{m,c}^{MUE}\right|^2}{\Lambda_{m,c}} - \beta\delta_k P_{m,c}^{MUE}\left|h_{m,c}^{MUE}\right|^2 \tag{B.9}$$

where the first term in (B.9) represents MUE m channel gain with $\Lambda_{m,c} = 2^{R_{m,c}^{min}} - 1$, the signal to interference plus noise ratio (SINR) threshold corresponding to the minimum MUE rate requirement, $R_{m,c}^{min}$. The second term measures the relative loss that MUE m advances to the SUE k over the codebook c, defined as the interference cost imposed by SBS to MUE m. β denotes a fixed coefficient with unit interference of MBS due to the SUE on codebook c. The parameter δ_k, given by $\delta_k = \max\left(0, (\sum_{i=1}^{m} A_{k,i}^{c} P_{i,c}^{MUE}\left|h_{i,c}^{MUE}\right|^2 - \mathcal{I}_{m\to k}^{c,th})/\mathcal{I}_{m\to k}^{c,th}\right)$ measures the degree of violation of the pairing interference temperature threshold at the SBS, $\mathcal{I}_{m\to k}^{c,th}$ given by

$$\sum_{m=1}^{J} A_{k,m}^{c} P_{m,c}^{MUE}\left|h_{m,c}^{MUE}\right|^2 \leq \mathcal{I}_{m\to k}^{c,th} \tag{B.10}$$

The MUEs rank SUEs based on SUEs achievable rate-based metric $R_{f,k,c}^{SUE}$. Consequently, for the DPR-UP $\Theta_X(Y) = \psi_{k\to m}^{UP}$ and $\Theta_Y(X) = \psi_{m\to k}^{UP}$ while $J_{UP} \geq 1$ represents the MUE pairing quota, the bounds of which will be subsequently determined in section 5.3. The matching game results to a stable SUE-MUE matching μ_{UP}^{*} for the user pairing policy $\mathbf{A} = [A_{k,m}^{MUE}]_{K\times M}$.

3.3 Power Allocation, $\{\mathbf{P}\}$, DPR-PA

The DPR-PA provides the PA policy P that allocates power to SUEs and MUEs. In DPR-PA, the set of discrete power levels S represent the set of players X while the user sets i.e., SUEs, $\mathcal{K}$ and MUEs, $\mathcal{M}$ individually represent the set of players Y. The amount of power in the power levels S decreases as the number of power levels increase. The ranking metrics are given as follows;

$$\psi_{s\to k}^{PA} = 2^{R_{f,k,c}^{SUE}} - 1 \tag{B.11}$$

$$\psi_{k\to s}^{PA} = \begin{cases} \min\left(\bar{P}_{f,k,c}^{SUE}, \max\left(P_{f,k,c}^{SUE}, P_{f,k,c}^{SUE,min}\right), \quad P_{f,k,c}^{max}\right), \\ \text{if } \Delta \geq P_{f,k,c}^{SUE,min} \\ Infeasible, \qquad\qquad\qquad\qquad \text{otherwise.} \end{cases} \tag{B.12}$$

$$\psi_{s\to m}^{PA} = 2^{R_{m,c}^{MUE}} - 1 \tag{B.13}$$

$$\psi_{m\to s}^{PA} = \min\left(P_{m,c}^{MUE}, P_{max}^{MUE}\right) \tag{B.14}$$

where in (B.11) and (B.13), $\psi_{s\to i}^{PA}$ gives the achievable SINR-based metric employed by the power levels to rank users $i \in \{k, m\}$. $R_{f,k,c}^{SUE}$ and $R_{m,c}^{MUE}$ denotes the achievable rate of SUEs and MUEs on codebook c respectively.

The users rank the power levels based on a $\psi_{i\to s}^{PA}$,$i \in \{k, m\}$ determined by the individual sets of MUEs and SUEs. In (B.12), $\psi_{k\to s}^{PA}$ represents the QoS aware power allocation metric [8] employed by the SUEs to rank the power levels $P_{f,k,c}^{SUE}$. In (B.14), $\psi_{m\to s}^{PA}$ represents the minimum SINR based metric employed by the MUEs to rank the power levels $P_{m,c}^{MUE}$, derived as follows; Suppose that each MUE m is required to achieve its minimum SINR $\Lambda_{m,c}$. It is desirable to provide this quality with minimum powers. Assuming perfect SIC, we propose a PA solution that provides for individual powers $\bar{P}_{m,c}^{MUE}$ [32]. Based on the instantaneous SINR derived in Section 4, the MUE optimal PA problem reduces to

$$\Lambda_{m,c} = \frac{\bar{P}_{m,c}^{MUE}\left|h_{m,c}^{MUE}\right|^2}{\sum\limits_{i=m+1}^{S} \bar{P}_{i,c}^{MUE}\left|h_{i,c}^{MUE}\right|^2 + \phi_c} \tag{B.15}$$

where $\phi_c = q_{f,k,c}^{SUE} P_{f,k,c}^{SUE}\left|h_{f,k,c}^{SUE}\right|^2 + \sigma_{m,c}^2$ is the combined SUE interference plus noise for codebook c. After several algebraic computations, the compact form solution of (B.15) can be given as

$$\bar{P}_{m,c}^{MUE} = \frac{\phi_c}{\left|h_{m,c}^{MUE}\right|^2} \times \frac{\kappa_{m,c}\Lambda_{m,c}}{1+\Lambda_{m,c}} \times \frac{1}{\left(1 - \sum\limits_{i=1}^{S} \frac{\kappa_{i,c}\Lambda_{i,c}}{1+\Lambda_{i,c}}\right)} \tag{B.16}$$

where

$$\kappa_{m,c} = \begin{cases} 1, & \text{m=1} \\ \Pi_{j=1}^{k-1} a_{j,c}, & \text{m >1} \end{cases} \tag{B.17}$$

with $a_{j,c} = \dfrac{2}{1+\Lambda_{j,c}}$. Given that P_{max}^{MUE} is the maximum MUE power requirement, the MUE to power level ranking metric $\psi_{m\to p}^{PA}$ can be given as in (B.14). Consequently, for the DPR-PA $\Theta_X(Y) = \psi_{p\to i}^{PA}$ and $\Theta_Y(X) = \psi_{i\to p}^{PA}, i \in \{m, k\}$ and $J_{PA} = 1$. The algorithm outputs a stable power levels to users (i.e., SUEs, MUEs) matching μ_{PA}^* for the PA assignment policy $\mathbf{P} = \left\{P_{F,K,C}^{SUE}, P_{M,C}^{MUE}\right\}$.

4 PD-SCMA Receiver

In this section, the receiver MUD is presented. The received vector over an SCMA block, $\mathbf{y_{ru}} = \left[y_1^{ru} y_2^{ru} \cdots y_N^{ru}\right]^T$ is a $N \times 1$ vector, with y_n^{ru} as the received symbol on RU n given by

$$\begin{aligned} y_n^{ru} = &\sum_{k=1}^{K} \rho_{n,c} h_{f,k,n}^{SUE} \sqrt{P_{f,k,n}^{SUE}} x_{f,k,n}^{SUE} \\ &+ \sum_{k=1}^{K} \sum_{m \in \mathcal{M}_c} A_{k,m}^{c} h_{m,n}^{MUE} \sqrt{P_{m,n}^{MUE}} x_{m,n}^{MUE} + z_{f,k,n}. \end{aligned} \tag{B.18}$$

where $z_{f,k,n} \sim \mathcal{CN}(0, \sigma_{f,k,n}^2)$ is the additive white Gaussian noise (AWGN). The iterative Log-MPA decoding process uses the information of $\mathbf{y_{ru}}$ to compute the received observation vector over codebooks, $\mathbf{y_{cb}} = \left[y_1^{cb}, \cdots, y_C^{cb}\right]^T$ with observation y_c^{cb} as the approximate received symbol for the J MUEs and SUE of codebook c given by

$$\begin{aligned} y_c^{cb} = &q_{f,k,c}^{SUE} h_{f,k,c}^{SUE} \sqrt{P_{f,k,c}^{SUE}} x_{f,k,c}^{SUE} + \\ &\sum_{m \in \mathcal{M}_c} A_{k,m}^{c} h_{m,c}^{MUE} \sqrt{P_{m,c}^{MUE}} x_{m,c}^{MUE} + z_{f,k,c}. \end{aligned} \tag{B.19}$$

Note that users on different codebooks do not interfere with each other and we assume negligible inter-cell interference resulting from codebook re-use in different small-cells. Without loss of generality, let $w_{MUE_{m,c}} = P_{m,c}^{MUE} \left|h_{m,c}^{MUE}\right|^2$ and $\lambda_{MUE_{m,c}} = \dfrac{1}{E(w_{MUE_{m,c}})}$ denote the instantaneous received signal power and its mean value respectively. Prior to decoding, the receiver determines the instant decoding order π based on the instantaneous received user signal power [28]. Subsequently, users are decoded in the sequence of $\left[MUE_1, MUE_2, \cdots, MUE_J, SUE_k\right]$ with the instantaneous signal power relation $\left[w_{MUE_{1,c}}, w_{MUE_{2,c}}, \cdots, w_{MUE_{J,c}}, w_{SUE_{k,c}}\right]$. The highest ranked user experiences interference from all users while the lowest channel gain user effectively enjoys interference-free transmission.

Algorithm 8 HSLM Receiver

Input: Q,A, P, G, $\sigma^2_{f,k,c}$ and $\mathbf{y}$ the received signal

Output: $\hat{x}$

for SBS $f = 1 : F$ **do**
  Perform iterative log-MPA process on $\mathbf{y_{ru}}$ to compute $\mathbf{y_{cb}}$ for each codebook for $\tau_m ax$ iterations.
  for each codebook $c = 1 : C$ **do**
    for each user $i \in \{k, m\}$ in c **do**
      Compute $\aleph(i)$, (B.26)
      Perform post-ordering of $\aleph(i)$ in decreasing order
      $\bar{\mathbf{y}}_{\mathbf{cb}} = \mathbf{y_{cb}} \times \aleph(i)$
      Apply SIC on resulting ordered signal
      $\bar{\mathbf{y}}_{\mathbf{cb}} = \bar{\mathbf{y}}_{\mathbf{cb}} - A^c_{k,m} h^{MUE}_{m,c} \sqrt{P^{MUE}_{m,c}} x^{MUE}_{m,c}$
      $\aleph(i) = \aleph(i-1)$
    end for
    Perform user bits reconstruction by computing LLRs.
  end for
end for

The instantaneous SINR of the m^{th} MUE multiplexed at codebook c, $\gamma^{MUE}_{m,c}$ is given by

$$\gamma^{MUE}_{m,c} = \frac{A^c_{m,c} w_{MUE_{m,c}}}{\sum_{j=m+1}^{J} A^c_{j,c} w_{MUE_{j,c}} + \phi_c}, \tag{B.20}$$

while that of the J^{th} MUE is given by

$$\gamma^{MUE}_{J,c} = \frac{A^c_{J,c} w_{MUE_{J,c}}}{\phi_c}, \tag{B.21}$$

The SINR of the k^{th} SUE $\gamma^{SUE}_{f,k,c}$, after successful SIC of all MUEs $m \in \mathcal{M}_c$ in each codebook is given by

$$\gamma^{SUE}_{f,k,c} = \frac{q^{SUE}_{f,k,c} P^{SUE}_{f,k,c} |h^{SUE}_{f,k,c}|^2}{\sigma^2_{f,k,c}}. \tag{B.22}$$

The achievable data rate, $R^{SUE}_{f,k,c}$ is given as

$$R^{SUE}_{f,k,c} = \log_2 \left(1 + \gamma^{SUE}_{f,k,c}\right), \tag{B.23}$$

while that of the MUEs is similarly derived and given as

$$R^{MUE}_{m,c} = \log_2 \left(1 + \gamma^{MUE}_{m,c}\right), \tag{B.24}$$

A hybrid PD-SCMA receiver is proposed in Algorithm 8. By applying a Jacobian logarithm approximation [23], the log-MPA computes the symbol estimate transmitted over each codebook y_c^{cb}. This estimate is passed to the next step of SIC decoding. Define $\mathbf{G} = \mathbf{P} \odot \mathbf{H}$ as the equivalent received signal power matrix for the multiplexed users. The MMSE transformation weight matrix estimate is given as [33]

$$\mathcal{W}_{MMSE} = \left[\mathbf{G}^H\mathbf{G} + \sigma^2\boldsymbol{I}\right]^{-1}\mathbf{G}^H. \tag{B.25}$$

Using the MMSE matrix, the decoding order metric $\aleph(i)$ for each user on codebook c can be computed by

$$\aleph(i) = \frac{|\mathcal{W}_{MMSE}(i)\mathbf{G}(i)|^2}{\sum_{i\neq j}^{M+1}|\mathcal{W}_{MMSE}(i)\mathbf{G}(i)| + \sigma_i^2\|\mathcal{W}_{MMSE}(i)\mathbf{G}(i)\|^2} \tag{B.26}$$

Lastly, after SIC decoding, the symbol estimates for each user over a codebook can be obtained by log-like ratio (LLR) computation [8].

5 Multiplexing Capacity Bounds

In this section, we first analyze the PD-SCMA capacity region. Secondly, we investigate codebook, pairing and power multiplexing capacities of the hybrid PD-SCMA technology.

5.1 Capacity Region of PD-SCMA

Similar to SCMA capacity region C_{SCMA} derived based on multiple access channels (MAC) [11], the capacity region for the uplink PD-SCMA for both SUEs and MUEs, C_{PDSCMA} can be derived as (B.27). The data rate $\mathbf{R}$ is a $(J+1)\times C\times S$ matrix for J MUEs with S distinct power levels and a SUE in a codebook for C codebooks.

$$\begin{aligned} C_{PDSCMA}(\mathbf{P},\mathbf{H}) = \Bigg\{ \mathbf{R} : & \left(\sum_{c\in\mathcal{C}} R_{f,k,c}^{SUE} + \sum_{c\in\mathcal{C}}\sum_{m\in\mathcal{M}_c} R_{m,c}^{MUE}\right) \\ & \leq \log_2\left(1 + \frac{1}{\sigma_{f,k,c}^2}\sum_{c\in\mathcal{C}} q_{f,k,c}^{SUE}\rho_{n,c}\left|h_{f,k,c}^{SUE}\right|^2 P_{f,k,c}^{SUE}\right) + \\ & \log_2\left(1 + \frac{1}{\sigma_{m,c}^2}\sum_{c\in\mathcal{C}}\sum_{m\in\mathcal{M}_c} A_{m,c}^{c}\rho_{n,c}\left|h_{m,c}^{MUE}\right|^2 P_{m,c}^{MUE}\right)\Bigg\}, \\ & \mathcal{C}\subset\{1,\cdots,C\}, \mathcal{M}_c\subset\{1,\cdots,J\}, P_{m,c}^{MUE}\in S, S\subset\{1,\cdots,S\}. \end{aligned} \tag{B.27}$$

The capacity region $C_{PDSCMA}(\mathbf{P},\mathbf{H})$ represents the upper bound or maximum achievable rates of all the users and is independent of any practical codebook design. Besides, it can be observed that the sparsity of ρ matrix, number of pairable MUEs $|\mathcal{M}_c| = J$ and power levels S influence the capacity region of (B.27). For a sparser ρ, i.e., a smaller number of $\rho_{n,c} = 1$, the upper bound of

the right hand (B.27) is smaller and therefore, the achievable rate is smaller. But a sparser ρ means less codebook interference and, therefore, the complexity of the MUD will be smaller. Increasing the number of paired MUEs J increases capacity subject to the minimum QoS requirements and lastly, distinct power levels in $\mathbf{P}$ determine the upper bounds of J given by S for optimal SIC capacity. It is therefore imperative to re-evaluate the individual resource bounds for the system capacity.

5.2 Codebook Multiplexing Capacity, C

The number of codebooks available in a SCMA dimension is a combinational problem expressed by the following function

$$C = f(N, d_v, \lambda) \tag{B.28}$$

For the capacity, the following theorem is derived.

Proposition 1. *Given the different considerations, the number of admissible codebooks can be given by the minimum of positional bound* $(C1)$*, overloading bound* $(C2)$*, sparsity bound* $(C3)$

$$C = \min\{C1, C2, C3\}, \left\{ \underbrace{\begin{pmatrix} N \\ d_v \end{pmatrix}}_{C1}, \underbrace{(\lambda \cdot N)}_{C2}, \underbrace{\begin{pmatrix} N \\ \frac{N}{2} \end{pmatrix}}_{C3} \right\} \tag{B.29}$$

Proof. The proof of the theorem can be supported by the following Lemmas. □

Lemma 5.1. *Positional bound:* $C1 = \begin{pmatrix} N \\ d_v \end{pmatrix}$

Proof. $C1$ represents a unique solution to C resulting from inserting $N - d_v$ all-zero row vectors within the rows of identity I_N [11]. Alternatively, the generation of C codebooks is equivalent to selecting d_v positions out of N RUs, hence $C \leq \begin{pmatrix} N \\ d_v \end{pmatrix}$. □

Lemma 5.2. *Overloading bound :* $C2 = (\lambda \cdot N)$

Proof. $C2$ gives an important bound on the total number of codebooks which is a function of the overloading factor λ defined as $\lambda = \frac{C}{N}$. Though desired for OMA, a large λ results in increased number of colliding codebooks per RU, limited number of differentiable constellations generated from the mother constellation hence difficulties in decoding the mixed signal for the MPA. Therefore, from the overloading factor, $C2 \leq (\lambda \cdot N)$. □

Lemma 5.3. *Sparsity bound :* $C3 = \begin{pmatrix} N \\ \frac{N}{2} \end{pmatrix}$.

Proof. The proof of Lemma 5.3 is presented in Appendix 10.2. □

5.3 MUE Multiplexing Capacity, J

In uplink PD-SCMA, a SUE is paired with J MUEs to transmit over the same codebook c. The bounds of J is subject to the pairing interference and is conditioned on the ranking parameter $\psi^{UP}_{k\to m}$.

Proposition 2. *Given the pairing metric* $\psi^{UP}_{k\to m}$ *and assuming* $\delta_k \neq 0$*, the number of MUEs* J *that can additively be paired with a SUE in codebook c is given by*

$$J = \left\{ \sum_{i=1}^{m} i : \psi^{UP}_{k\to m} \geq \frac{w_{MUE_{m,c}}}{\Lambda_{m,c}}, m \in \mathcal{M}_c \right\} \tag{B.30}$$

The upper bound of J *can be given when* $\psi^{UP}_{k\to m} = \dfrac{w_{MUE_{i,c}}}{\Lambda_{m,c}}$.

Proof. Given the pairing interference temperature threshold $\mathcal{I}^{c,th}_{m\to k}$ in (B.10) and assuming that $\delta_k \neq 0$, then the pairing metric (B.9) can be written as

$$\psi^{UP}_{k\to m} = w_{MUE_{m,c}} \left[\frac{1}{\Lambda_{m,c}} - \beta \left(\frac{\sum_{i=1}^{m} w_{MUE_{i,c}} - \mathcal{I}^{c,th}_{m\to k}}{\mathcal{I}^{c,th}_{m\to k}} \right) \right] \tag{B.31}$$

For codebook c, the MUEs are additively multiplexed with SUE k in the order of the ranking parameter $\left\{\psi^{UP}_{k\to 1}\right\}, \left\{\psi^{UP}_{k\to 2}\right\}, \cdots, \left\{\psi^{UP}_{k\to m}\right\}, \cdots, \left\{\psi^{UP}_{k\to J}\right\}$. With each MUE additively paired, the pairing interference $\sum_{i=1}^{m} w_{MUE_{i,c}}$ monotonically increases and the term in the inner bracket, consequently decreases. To protect the codebook QoS, it is required that the pairing metric guarantees (B.30), the upper bound of J. □

5.4 Power Multiplexing Capacity, S

The maximum MUE pairing capacity in a codebook S is a function of the power policy. The capacity is conditioned on the number of distinct power levels that can guarantee SIC decoding. Without loss of generality, it can be deduced that $P^{MUE}_{m,c}\left|h^{MUE}_{m,c}\right|^2 \geq P^{SUE}_{f,k,c}\left|h^{SUE}_{f,k,c}\right|^2$. Accordingly, of interest is the paired MUEs power level distinctions to guarantee optimal SIC of MUEs, subject to maximum power allocation by MBS. If Γ_c denotes the minimum power level difference required to distinguish between the MUE to be decoded and the remaining non-decoded MUEs, then the feasibility of effective SIC is governed by a PA policy given by [19]:

$$\frac{w_{MUE_{m,c}}}{\phi_c} - \sum_{i=m+1}^{S} \frac{w_{MUE_{i,c}}}{\phi_c} \geq \Gamma_c, m = 1, 2, \cdots, (S-1), \tag{B.32}$$

Given the PA (B.15), (B.16) and (B.32), the power multiplexing capacity can be characterized using the following proposition.

Proposition 3. *Considering a linear and dynamic SIC receiver cancelling $L = 1$ MUEs successively, the power multiplexing capacity S is bounded such that*

$$S < 1 + \frac{1}{a_c \Lambda_c}, \tag{B.33}$$

where Λ_c denotes the MUEs uniform minimum SINR on codebook c and $a_c = 2/(1 + \Lambda_c)$.

Proof. Given that the MUEs' desired $\Lambda_{m,c}$ can be accommodated within the feasible dynamic power constraints, (B.32), we can then determine the maximum number of MUEs, S in a codebook. For any $\Lambda_{m,c} < \infty$, there exists a positive and limited power solution, feasible only if the conditions (B.34) derived from (B.15), (B.16) and (B.32) are satisfied.

$$\sum_{i=1}^{S} \frac{\kappa_{i,c}\Lambda_{i,c}}{1 + \Lambda_{i,c}} < 1 \quad \text{and} \quad \sum_{i=m+1}^{S} w_{MUE_{i,c}} \geq \frac{\phi_c(\Gamma_c - \Lambda_{m,c})}{\Lambda_{m,c} - 1} \tag{B.34}$$

With SIC linearity, we can assume that when L MUEs are successively canceled, there will be one set of users experiencing canceled interference and another employing conventional single-user detection [32].

$$\underbrace{1, 2, \cdots, L}_{\text{Cancelled signals}}, \underbrace{L+1, L+2, \cdots, S}_{\text{Uncancelled signals}} \tag{B.35}$$

Considering a case with MUEs in a codebook having uniform minimum SINR Λ_c and that $a_c = 2/(1 + \Lambda_c)$, the maximum number of MUEs S must fulfill the feasibility condition

$$S < 1 + L + a_c^{-L}\left(1 + \frac{1}{\Lambda_c} - \frac{a_c^{L+1} - 1}{a_c - 1}\right), L < S \tag{B.36}$$

When only one MUE $L = 1$ is successively cancelled, then using (B.36) we can obtain the result of Proposition 3. □

6 Outage Probability Analysis

In this section, we investigate pairing outage probability based on the time-varying recieved power strength for each MUE. Analogous to conventional uplink NOMA [34], the sum rate remains consistent irrespective of the decoding order for each codebook. The maximum number of paired MUEs S is a function of the pairing metric determined based on pairing interference, maximum number of power levels and SIC receiver complexity. For each codebook, each pairing user independently transmits its signal at either maximum transmit power or controlled transmit power. At the receiver, the decoding order is based on the instantaneous user received signal power.

With deteriorating channel conditions and SIC constraints, transmission reliability is not guaranteed i.e., if the achievable rate is less than target rate, then outage occurs. Let $E^c_{m|\pi}$ denote the event that achievable rate $R^{MUE}_{m,c}$ of MUE m is larger than the target rate $\bar{R}^{MUE}_{m,c}$ under a given decoding order π. Using (B.20), (B.21) and (B.24), the probability of $E^c_{m\|\pi}$ can be formulated as [28]

$$P(E^c_{m|\pi}) = P\big(R^{MUE}_{m,c} \geq \bar{R}^{MUE}_{m,c}|\pi\big) \tag{B.37}$$

After obtaining $P(E^c_{m|\pi})$, the outage probability of MUE m under the given decoding order is expressed as

$$P^{out}_{m|\pi} = 1 - \prod_{l=1}^{m} P(E^c_{l|\pi}) \tag{B.38}$$

Now, equation (B.37) can be expressed as a joint probability $P\big(R^{MUE}_{m,c} \geq \bar{R}^{MUE}_{m,c}, \pi\big)$ that events $R^{MUE}_{m,c} \geq \bar{R}^{MUE}_{m,c}$ and π occur simultaneously given by

$$P(E^c_{m|\pi}) = \frac{P\big(R^{MUE}_{m,c} \geq \bar{R}^{MUE}_{m,c}, \pi\big)}{P(\pi)} \tag{B.39}$$

The closed-form expression of (B.39) can be derived by first determining the probability of the decoding order π (denominator) in the set of all possible decoding order given as

$$P(\pi) = \frac{\prod\limits_{i=2}^{S} \lambda_{MUE_{i,c}}}{\prod\limits_{i=2}^{S} \left(\sum\limits_{k=1}^{i} \lambda_{MUE_{i,c}} \right)} \tag{B.40}$$

From (B.20) and (B.24), and for simplicity, let $\bar{w}^S_{m+1} = \sum\limits_{i=m+1}^{S} w_{MUE_{i,c}}$. Then, for $m \neq S$ the numerator in (B.39) can be given as

$$P\big(R^{MUE}_{m,c} \geq \bar{R}^{MUE}_{m,c}, \pi\big) = P\big(w_{MUE_{m,c}} \geq \beta_m \bar{w}^S_{m+1}, \pi\big) \tag{B.41}$$

where $\beta_m = 2^{\bar{R}^{MUE}_{m,c}} - 1$. The derivation of joint probability in (B.41) is highly determined by β_m, and therefore we can derive it for different cases of β_m. Without losing generality, we assume that $\beta_m \geq 1$ and only express the joint probability $P\big(w_{MUE_{m,c}} > \beta_m \bar{w}^S_{m+1}, \pi\big)$ for $m \neq S$ given as

$$P\big(w_{MUE_{m,c}} > \beta_m \bar{w}^S_{m+1}, \pi\big) = \frac{\prod\limits_{i=2}^{S} \lambda_{MUE_{i,c}}}{\prod\limits_{i=2}^{m} \left(\sum\limits_{k=1}^{i} \lambda_{MUE_{i,c}} \right) \prod\limits_{i=m+1}^{S} \left(\sum\limits_{k=m+1}^{i} \lambda_{MUE_{i,c}} + \nu \sum\limits_{j=1}^{m} \lambda_{MUE_{j,c}} \right)} \tag{B.42}$$

where $\nu = (i-m)\beta_m$. Besides, for the last decoded MUE S, the conditional probability $P(E^c_{S|\pi})$ is given by

$$P(E^c_{S|\pi}) = P\Bigg(w_{MUE_{S,c}} > \beta_S \phi_c, \pi\Bigg) = \frac{\prod\limits_{i=2}^{S} \lambda_{MUE_{i,c}} \exp\Bigg(-\sum\limits_{i=1}^{S} \lambda_{MUE_{i,c}} \beta_S \phi_c\Bigg)}{\prod\limits_{i=2}^{S}\Bigg(\sum\limits_{k=1}^{i} \lambda_{MUE_{k,c}}\Bigg)} \tag{B.43}$$

Combining (B.40) and (B.42), we can evaluate (B.39) given as (B.44).

$$P(E^c_{m|\pi}) = \frac{1}{\prod\limits_{i=m+1}^{S}\Bigg(\sum\limits_{k=m+1}^{i} \lambda_{MUE_{i,c}} + \nu \sum\limits_{j=1}^{m} \lambda_{MUE_{j,c}}\Bigg)} \tag{B.44}$$

Finally, using (B.38), the outage probability expression under the condition of the rate threshold β_m and the maximum number of pairing MUEs, S is obtained.

7 Complexity and Convergence Analysis

In this section, complexity and convergence of the system is analyzed.

7.1 Complexity Analysis

The complexity of PD-NOMA and SCMA has been derived in [4]. In the uplink PD-SCMA model, we assume that each codebook is assigned to $S+1$ users, (i.e., one SUE and S MUEs) simultaneously. At the SBS, HSLM receiver applies the iterative log-MPA once to detect and decode the C codebook layers. Then for each codebook layer, SIC is applied S times to correctly decode the transmitted data. By considering the MMSE detector at the receiver, the solution involves the computation of MMSE transformation weight matrix estimate given by $\left[\mathbf{G}^H\mathbf{G} + \sigma^2\boldsymbol{I}\right]^{-1}\mathbf{G}^H$. The computational complexity order of λ^{-1} and $\lambda^H\lambda$ (with size $b \times b$) is $\mathcal{O}(b^3)$. Given that each user is assigned at most one codebook, the complexity order of the uplink PD-SCMA can approximately be given by $\mathcal{O}\Bigg(\big(\tau_{max}\big)\log\big(|D|_{d_v}\big) + Sb^3\Bigg)$.

The overall complexity order of solution for the ASM method is basically a linear combination of the complexity of each DPR-RA solution. The DPR-CA, DPR-UP and DPR-PA are modelled as one-to-many matching games. The computational complexity of a matching algorithm is given as a linear function of the size of its preference lists [6]. The complexity order of Algorithm 9 can be given by $\mathcal{O}\big(C \times K\big)$ for DPR-CA and $\mathcal{O}\big(K \times S\big)$ for DPR-UP. The complexity order of the SUEs DPR-PA can be approximated to $\mathcal{O}\big(K^2\big)$ and $\mathcal{O}\big(M^2\big)$ for MUEs DPR-PA.

7.2 Convergence of ASM

With the ASM method, each iteration enhances the overall solution to convergence through improvement of each individual resource ($\mathbf{Q}$, $\mathbf{A}$ and $\mathbf{P}$).

Proposition 4. *Having a feasible initialization, the transmitter resource allocation algorithm converges to a local maximum.*

Proof. Similar to the proof [6], let $\mathcal{O}(\mathbf{Q}, \mathbf{A}, \mathbf{P})$ denote the combined solution. For each two consecutive iterations, we have the following relations:

$$\begin{aligned}
\mathcal{O}\big(\mathbf{P}[t], \mathbf{Q}[t], \mathbf{A}[t]\big) &= \max_{\mathbf{P}} \mathcal{O}\big(\mathbf{P}, \mathbf{Q}[t], \mathbf{A}[t]\big) \\
&\geq \mathcal{O}\big(\mathbf{P}[t-1], \mathbf{Q}[t], \mathbf{A}[t]\big) \\
&= \max_{\mathbf{Q}} \mathcal{O}\big(\mathbf{P}[t-1], \mathbf{Q}, \mathbf{A}[t]\big) \\
&\geq \mathcal{O}\big(\mathbf{P}[t-1], \mathbf{Q}[t-1], \mathbf{A}[t]\big) \\
&= \max_{\mathbf{A}} \mathcal{O}\big(\mathbf{P}[t-1], \mathbf{Q}[t-1], \mathbf{A}\big) \\
&\geq \mathcal{O}\big(\mathbf{P}[t-1], \mathbf{Q}[t-1], \mathbf{A}[t-1]\big)
\end{aligned} \tag{B.45}$$

This means that the transmitter resource allocation ASM based algorithm gives non-decreasing sum rate as the iterations continue, hence is ensured to converge to a locally optimal solution. □

8 Results and Discussion

The analytical evaluation of the uplink hybrid PD-SCMA system is presented in this section. First, we compare the performance of the proposed transmitter DPR-RA algorithm with the following; regular CA, random UP and equal PA (Equal & regular RA), biological based RA algorithms in PD-SCMA [9], sub-modularity-based CA+UP with successive convex approximation based difference of convex functions (SCA-DC) PA [6] and lastly, efficient resource management schemes of [35]. Secondly, we demonstrate the multiplexing capacity bound results of Propositions 1, 2 and 3 based on the RA schemes. Besides, we present the system capacity performance with respect to the number of small cell codebooks and multiplexed MUEs. Thirdly, outage analysis for any given decoding order considering varying power levels and SNR is presented. To clearly understand the analytical outage performance results, we provide the theoretical evaluations and the Monte Carlo simulations. We assume that the path loss of each user can be offset by the path loss compensation part [21]. To cancel the pairing interference successively, we ensure the minimum power level difference in each codebook Γ_c, (B.32). With no loss of generality, we characterize the received signal to noise ratio (SNR) $\frac{P}{\sigma^2_{m,c}}$, with P set

Table B.1: Simulation Parameters

Parameters	Values
Center carrier frequency	$2\,GHz$
MBS coverage radius	$500\,m$
SBS coverage radius	$50\,m$
Maximum transmission power	$23\,dBm$
Noise variance, σ_x^2	$-174\,dBm$
Distance path loss	$PL(dB) = 128.1 + 37.6\log_{10} D$
Fast fading channel model	$g_{f,k,c}^{SUE}, g_{m,c}^{MUE} \sim \mathcal{CN}(0,1)$
RU Bandwidth	$200\,kHz$
Minimum transmission rate	$5\,Mbps/Hz$
Interference threshold	$10^{-5.5}\,W$

as the received power for the user with the worst large-scale fading. The detailed system parameters and assumptions are presented in Table B.1.

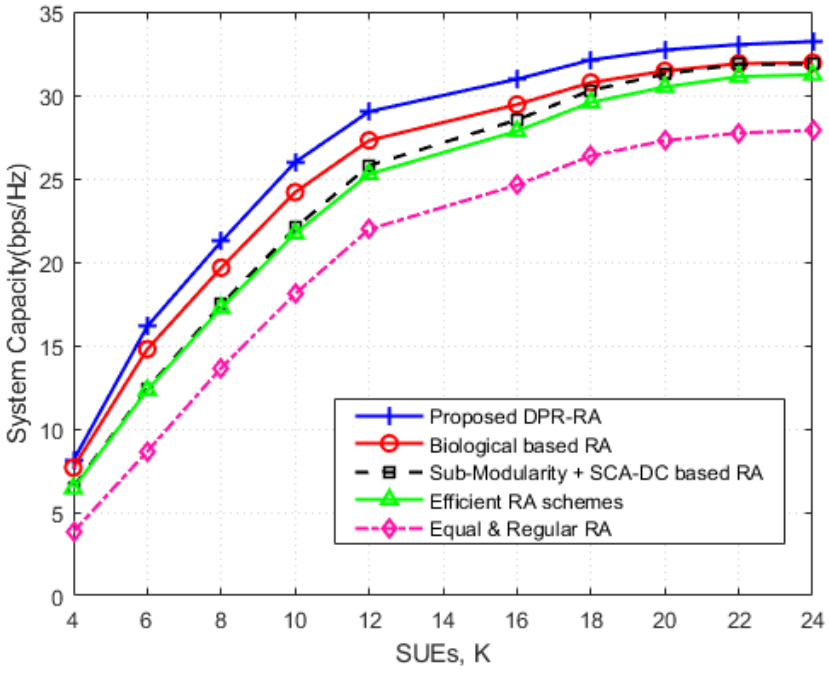

Fig. B.3: System Capacity vs number of SUEs.

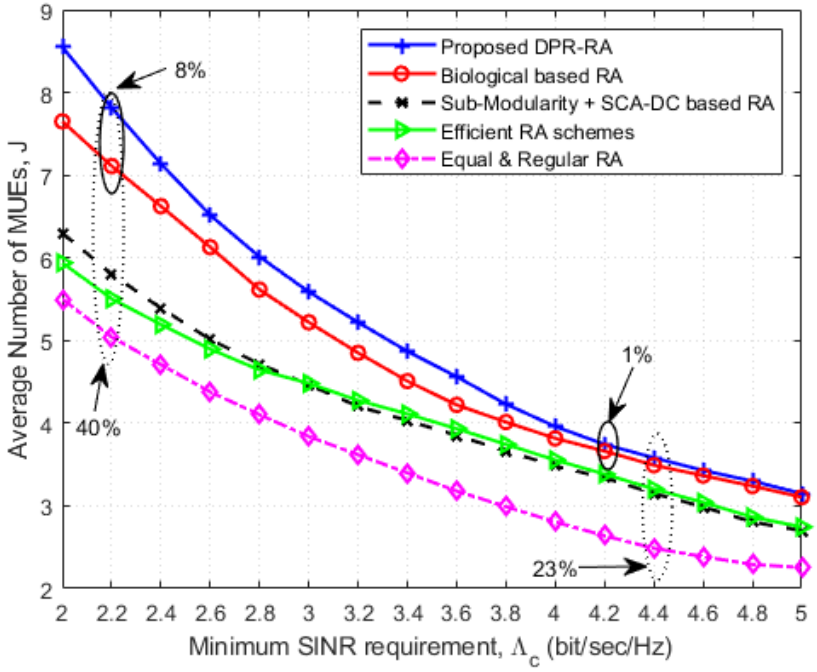

Fig. B.4: Average number of multiplexed MUEs, J vs. minimum MUE SINR Λ_c.

Fig. B.3 illustrates the overall system capacity performance versus the number of SUEs K. In all the schemes, the system capacity increases with increasing number of SUEs. It can be observed that PD-SCMA employing the DPR-RA outperforms other RA schemes due to non-exclusive allocation of codebooks performed at the transmitter eliminating codebook interference. Additionally, the RA schemes results in better capacity performance than the equal power and regular CB assignment, attributed to the efficient spectral RU utilization associated with the RA schemes [5].

Fig. B.4 shows the average number of paired MUEs J against minimal SINR requirements Λ_c for various RA schemes. The proposed DPR-RA algorithm outperforms the considered RA comparison schemes at high and low minimal SINR requirements. Compared with [9], the proposed DPR-RA scheme exhibit 8% and 1% higher the number of pairing users for high and low values of Λ_c respectively. When compared with the worst case scenario RA, i.e., equal power, regular CA and uniform UP, the improvement is much higher with 40% and 25% difference achieved for high and low Λ_c respectively.

The codebook bounds C discussed in Lemma 5.1, 5.2 and 5.3 are presented in Table B.2 in terms of N, C, d_v and λ. In Lemma 5.1, d_v is kept constant irregardless of the number of spectral RUs, N. In Lemma 5.2, the number of codebooks C is a function of the overloading factor λ. By varying λ for a fixed N, the number of codebooks and users vary at the expense of codebook interference

Table B.2: Codebook Capacity Bounds

N =	4	6	8	10	12	14
Lemma 1 : $C = \binom{N}{d_v}$	6	15	28	45	66	91
Lemma 2 : $C = \lambda \cdot N$	12	18	24	30	36	42
Lemma 3: $C = \binom{N}{\frac{N}{2}}$	6	20	70	252	924	3432

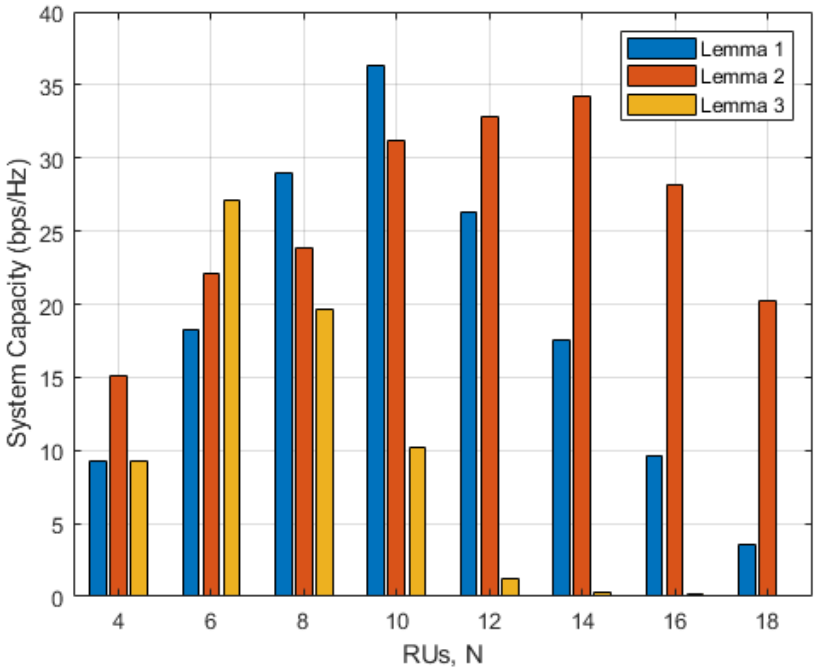

Fig. B.5: System Capacity versus Resource Units.

and difficulties in signal detection. Lemma 5.3 proposes that the number of RUs per codebook, d_v is variable and a factor of N. C is then bounded such that the upper bound of $d_v = N/2$. It is worthy noting that fixing d_v in Lemma 5.1 limits C even as additional RUs N are utilized. To optimize the RUs and the C, Lemma 5.3 is proposed. However, it is observed that as d_v approaches the upper bound $d_v = N/2$, C becomes very large which is impractical. Therefore, the choice of variable d_v is subject to acceptable receiver design complexity and is bounded.

The effect of the Lemmas 5.1-5.3 on the capacity is illustrated in Fig. B.5. As it can be observed, maximal capacity is obtained when the ratio of codebooks to RUs is optimized. Though capacity increases with RUs in Lemma 5.2, maintaining a constant overloading factor results in wasteful use

of RUs for the same capacity. Varying d_v towards the upper bound $d_v = N/2$ in Lemma 5.3 achieves same optimal capacity with less RUs albeit higher overloading and hence increased receiver complexity. As a result, beyond the system codebook bound, capacity deteriorates drastically. It can be observed that Lemma 5.1 gives optimal capacity performance at $N = 10$ and $C = 45$, Lemma 5.2 at $N = 14$ and $C = 42$ and Lemma 5.3 at $N = 6$ and $C = 20$, beyond which this operating points, the system capacity deteriorates. It should be noted that the minimum in Table B.2 is Lemma 5.2 but it has not reached its maximum capacity.

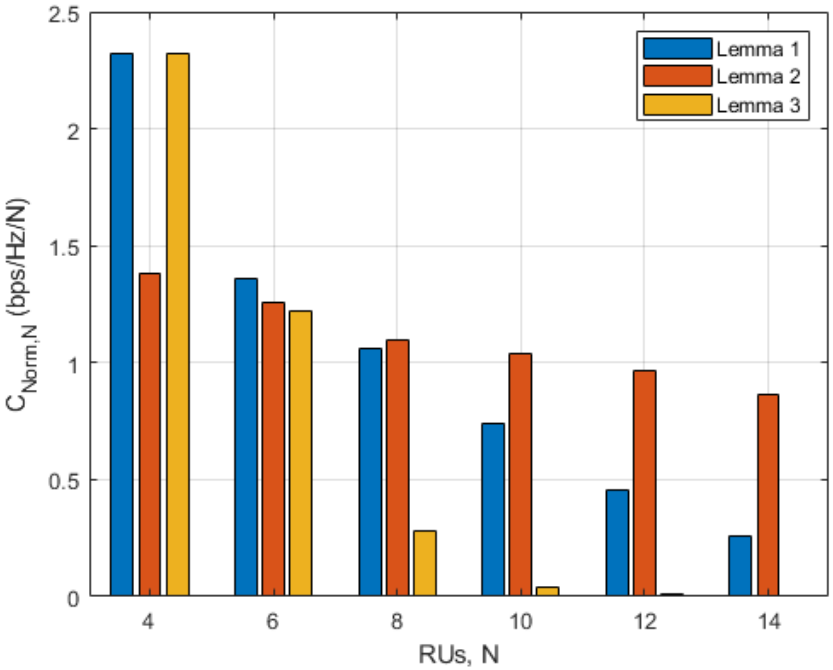

Fig. B.6: System Capacity versus Resource Units.

Denote by $C_{Norm,N}$, the normalized system capacity for each value of N given by $C_{Norm,N} = Capacity/N$. Fig. B.6 illustrates $C_{Norm,N}$ against the number of RUs N with respect to Lemmas 5.1, 5.2 and 5.3. For each Lemma, it can be observed that the normalized capacity is maximized at specific operating points. Generally, $C_{Norm,N}$ diminishes with increasing number of RUs N for the three Lemmas. However, it can be observed that Lemma 5.3 exhibits a drastic drop as N increases, followed by Lemma 5.1 while Lemma 5.2 demonstrate a steady reduction. Beyond $N = 8$, the $C_{Norm,N}$ falls below one, indicating a dwindling capacity performance. Note that the developed $C_{Norm,N}$ performance results validates codebook capacity bounds as given in proposition 1 and further validates the bounds in Table B.2.

The effect of varying d_v in Lemma 5.1 towards the upper bound proposed in Lemma 5.3 on capacity is illustrated in Fig. B.7. Since the number of codebooks C is a combinatorial function, for any N

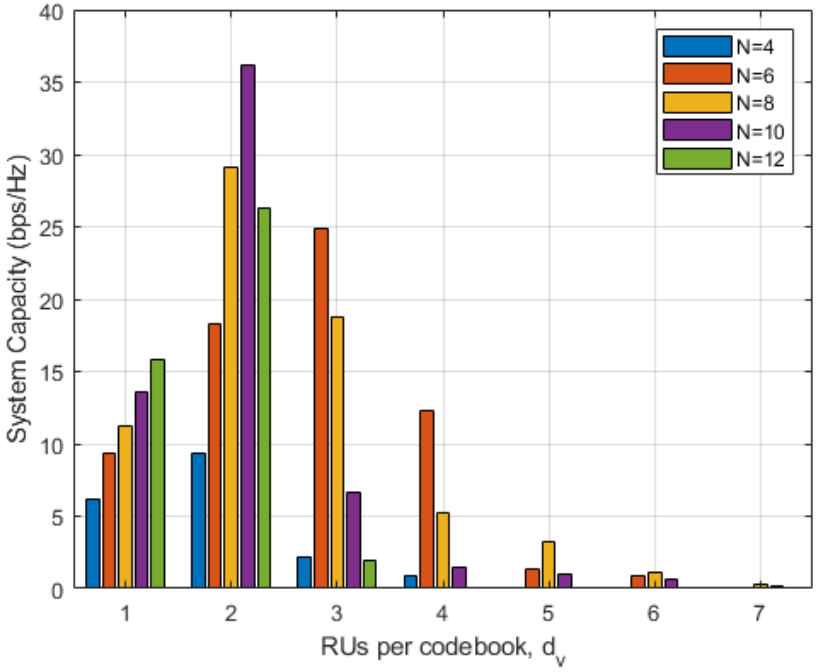

Fig. B.7: System Capacity versus d_v.

chosen, C increases drastically. As a result, for low values of d_v, capacity improves gradually to the optimal value beyond which the system capacity degenerates for each N chosen. As d_v approaches the upper bound $d_v = N/2$, the resulting value of C exceeds the optimal value rapidly especially for

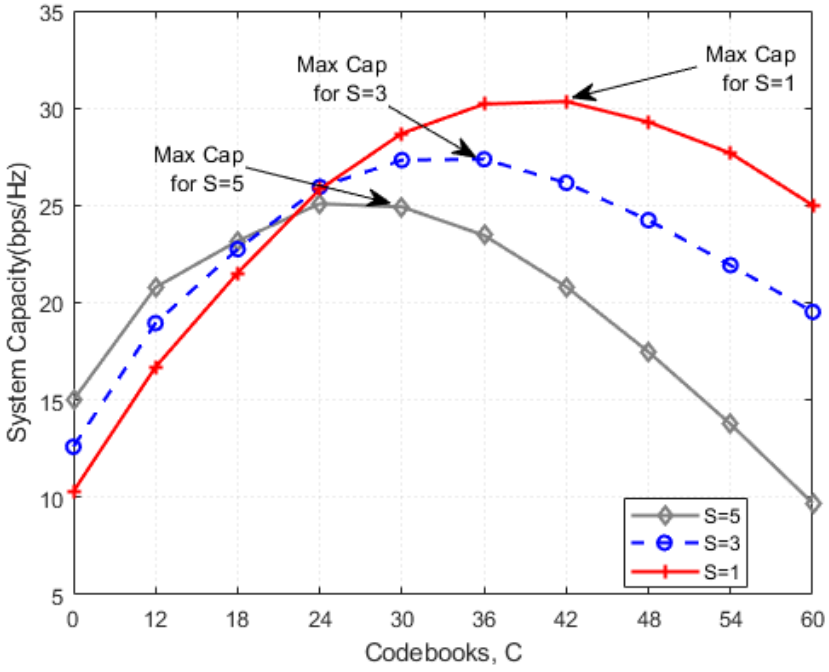

Fig. B.8: System Capacity versus Codebooks at various power level, S.

higher values of N and strains the system performance. Beyond the codebook bound for any N and d_v, the system yields due to overwhelming interference and complexity in the receiver. Besides, there is improved performance while using $N = 8$ for extended values of d_vs.

Fig. B.8 illustrates the system capacity against codebooks C for different power levels values S. It can be observed that different power levels achieve the maximum capacity at different values of codebooks beyond which the capacity performance deteriorates. At low value of S, capacity increases gradually and achieves the maximum value while utilizing more codebooks than at higher S values. This can be attributed to availability of resources at small cell to multiplex more MUEs, unlike with higher S where the system experiences additional receiver complexity. From the figure, maximum capacity is achieved at 24, 36 and 42 for power levels $S = 5$, $S = 3$ and $S = 1$ respectively.

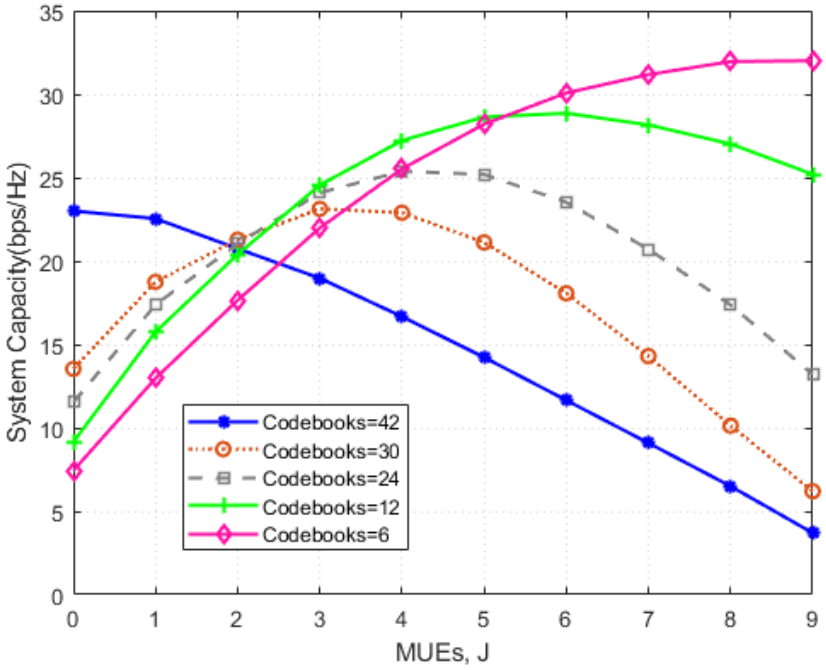

Fig. B.9: System Capacity vs MUEs, J for different number of codebooks

The system capacity performance versus paired MUEs J for disparate number of codebooks C is shown in Fig. B.9 as presented in proposition 2. As J increases, capacity improves to a maximum value, beyond which, multiplexing additional MUEs deteriorates the performance due to aggravated pairing interference in the codebook. With few codebooks assigned, the system can multiplex a larger J with minimal performance deterioration unlike for higher number of codebooks. As an example, with 6 codebooks, the system can multiplex 7 MUEs with significant performance while at 42 codebooks, the system can barely multiplex one MUE, $J < 1$.

Denote by $\zeta(J, C)$, the ratio of the number of pairable MUEs in set $\mathcal{M}$ and and the number of

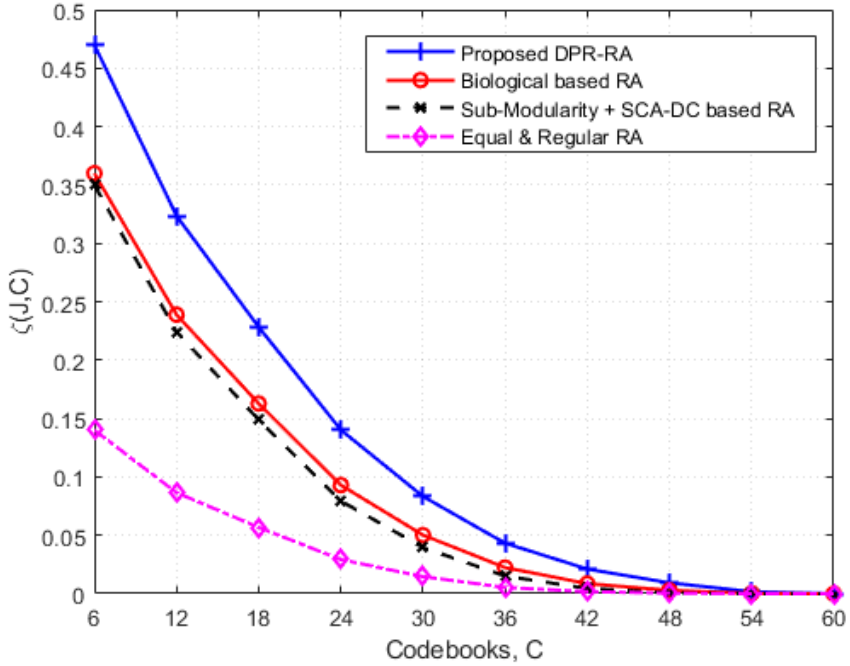

Fig. B.10: The ratio $\zeta(J, C)$ versus the number of Codebooks.

codebooks in set $\mathcal{C}$ given as $\zeta(J, C) = \frac{J}{C}$. Since there is a one-to-one assignment between codebooks and SUEs of set $\mathcal{K}$, then $\zeta(J, C)$ illustrates how many MUEs in the macro-cell can be paired with SUEs in a codebook and achieve desirable QoS. Fig. B.10 depicts the $\zeta(J, C)$ versus the number of codebooks. It can be observed that $\zeta(J, C)$ decreases with increase in the number of codebooks. Since the small cells have a higher priority, physical resources are first allocated to them. Due to the receiver complexity and pairing interference constraints, SUEs are prioritized for higher QoS. As the number of codebooks increase, the resource requirement for the SUEs increases, leading to the resources allocated to the MUEs decreasing in the case of limited physical resources.

Fig. B.11 represents the outage performance against power levels S for different codebook values. Additive deployment of power levels subsequently increases the maximal number of MUEs that can be multiplexed in a codebook. Power being a constrained resource, increasing the power levels gradually violate the minimum SINR based PA (B.16) and the SIC constraint (B.32). There is substantial outage with higher number of codebooks assigned compared to utilizing fewer codebooks. As the number of codebooks increases, it becomes increasingly complex to decode both at MPA and SIC, therefore resulting to outage. Based on a specified outage operating point and Λ_c, the maximum number of power levels S which subsequently determine the optimal number of MUEs can be determined. For example, designing the system to operate at an outage of 0.7, simulation results give the bounds of S for 6, 12 and 24 codebooks is 3, 2 and 1 respectively, which validates simulation bound results of

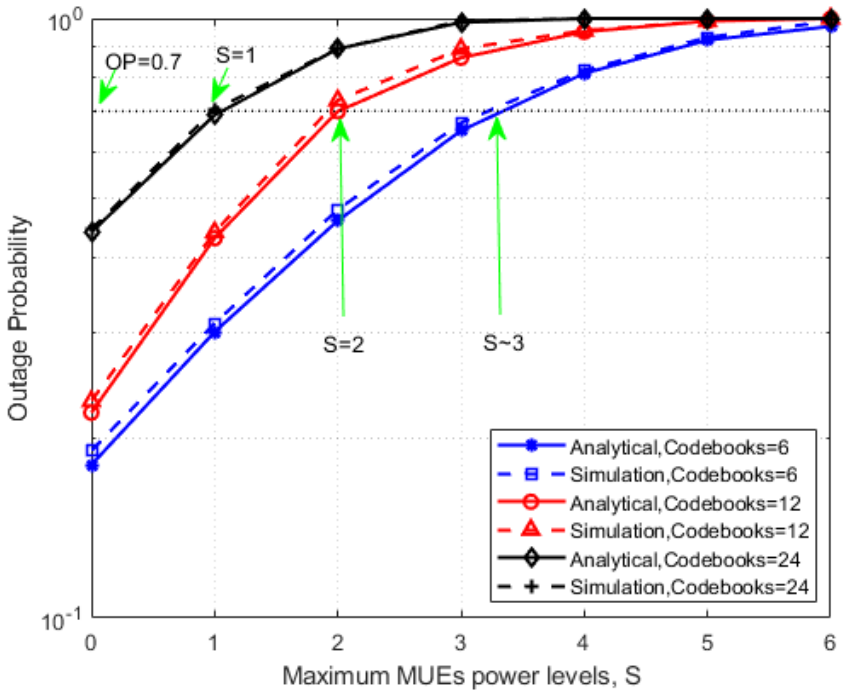

Fig. B.11: Outage versus the maximum MUEs' power levels S.

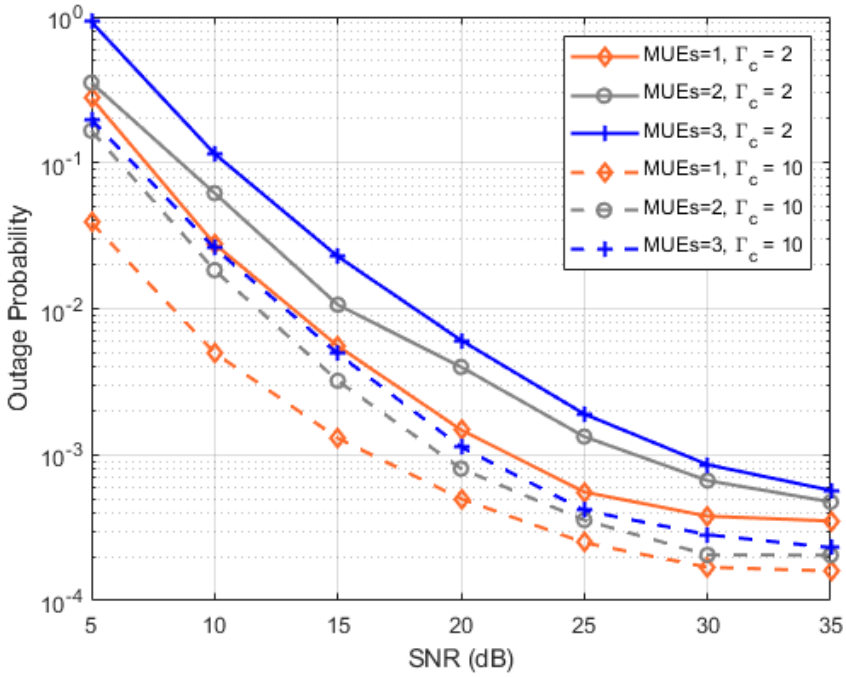

Fig. B.12: Outage versus SNR for different values of Γ_c.

proposition 3.

In Fig. B.12, MUEs outage probability vis-à-vis SNR for the considered dynamic ordered SIC under different Γ_c is investigated. In the considered scenarios, MUEs target data rate is set equal at

0.7bps/Hz and number of codebooks assigned $C = 6$. It is observed that a larger Γ_c improves outage performance since the distinction in the power domain provides SIC advantage in detecting the superimposed MUEs in a codebook. With the MUEs having equal target data rate, the larger Γ_c weakens the pairing interference resulting in the great gaps of MUEs.

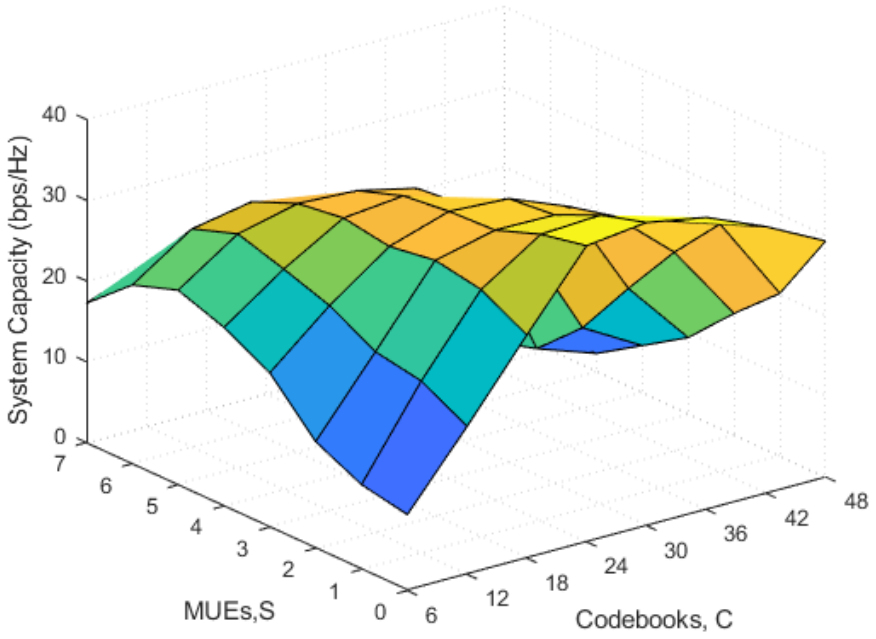

Fig. B.13: System Capacity vs MUEs vs number of codebooks.

In Fig. B.13, the capacity is presented with varying codebooks and number of MUEs. It can be observed that as both codebook and maximum number of MUEs S increases, the capacity increases to a maximum value, beyond which the performance deteriorates. Assigning more codebooks increases the MPA complexity order while multiplexing many MUEs results in violation of SIC constraints therefore deteriorating receiver capacity. From the figure, different combinations of S and number of codebooks can be identified to design the desirable system.

9 Conclusion

In this work, the multiplexing capacity bounds for the uplink hybrid PD-SCMA is investigated apropos the number of codebooks C, number of MUEs that can be paired with a SUE in codebook J and the maximal MUEs S that can be multiplexed in a codebook subject to the optimal number of power levels. To alleviate the RA challenges of uplink PD-SCMA, DPR-RA schemes are proposed and their performance investigated. It can be observed that DPR-RA based schemes greatly improve the performance of PD-SCMA. By using CA, UP and PA ranking parameters, measures of sparsity,

SIC constraints and users' QoS requirements, we obtain the codebook capacity bounds, MUE multiplexing bounds and power multiplexing bounds. In addition, the numerical system capacity and outage probability results validate our analytical multiplexing bounds. With a specified outage, the maximum number of MUEs to be multiplexed with SUE in a codebook for PD-SCMA system with a given number of codebooks is determined. We open a discussion on the multiplexing capacity of the hybrid PD-SCMA scheme for a feasible NOMA technology.

10 Appendices

10.1 A. DPR-RA Algorithm

We formulate the DPR-RA μ_{RA} matching game as follows [30]; Define a one-to-many RA matching game by a tuple $\big(\mathcal{X}, \mathcal{Y}, \succ_{\mathcal{X},RA}, \succ_{\mathcal{Y},RA} \quad \big)$. The notation $\succ_{\mathcal{X},RA}= \quad \{\succ_{X,RA}\}_{X\in\mathcal{X}}$ and $\succ_{\mathcal{Y},RA}= \{\succ_{Y,RA}\}_{Y\in\mathcal{Y}}$ represent the set of preference relations of player X and Y respectively. Let the metrics $\Theta_X(Y)$ and $\Theta_Y(X)$ denote the preference metric of player X in ranking Y and that of player Y in ranking X respectively.

Definition 1. *Given two disjoint finite sets of players X and Y, a matching μ_{RA} is defined as a function $\mu_{RA} : \mathcal{Y} \mapsto \mathcal{X}$, such that:*

- $X = \mu_{RA}(Y) \leftrightarrow Y = \mu_{RA}(X)$*;*
- $|\mu_{RA}(X)| \leq J_{RA}$ *and* $|\mu_{RA}(Y)| \leq 1$

Here, 1) indicates that, if player Y is matched to player X i.e., $X = \mu_{RA}(Y)$, then player X is also matched player Y i.e., $Y = \mu_{RA}(X)$ as well. 2) outlines the matching constraints such that each player X can only be assigned to at most J_{RA} players $\big(\mu_{RA}(X) \leq J_{RA}\big)$ and the condition $\big(\mu_{RA}(Y) \leq 1\big)$ guarantees that at most one X player can be matched to the player Y under the matching μ_{RA}. In the matching game, a preference $X_1 \succ_{Y,RA} X_2$ is such that player Y prefers X_1 to X_2 if $\Theta_1(X) > \Theta_1(X), X_1, X_2 \in \mathcal{X}$. Similarly, $Y_1 \succ_{X,RA} Y_2$ shows that X prefers Y_1 to Y_2 under the condition that $\Theta_1(Y) > \Theta_1(Y), Y_1, Y_2 \in \mathcal{Y}$.

Definition 2. *A pair $(X, Y) \neq \mu_{RA}, X \in \mathcal{X}, Y \in \mathcal{Y}$ is said to be a blocking pair for the matching μ_{RA} if it is not blocked by an individual player X and Y, and there exists another matching $\acute{\mu}_{RA} \in \mu_{RA}(X, Y)$ such that player X and player Y can achieve a higher utility metric. This mathematically implies that $\acute{\mu}_{RA} \succ_X \mu_{RA}$ and $\acute{\mu}_{RA} \succ_Y \mu_{RA}$. A matching μ_{RA} is said to be stable if it is not blocked by an individual player X and Y or any pair.*

Definition 3. *A bid value $b^{RA}_{Y\to X}(t) = 1$ when player Y prefers to associate with player X, otherwise $b^{RA}_{Y\to X}(t) = 0$.*

10.2 B. Proof of Lemma 5.3

The sparseness of codewords can be completely characterized by ρ. Let $\varrho_n = \{c : 1 \leq c \leq C, \rho_{n,c} = 1\}$ denote the index of users that contribute to the n^{th} RU, then $|\varrho_n|$is equal to the row weight d_v of the k^{th} row of this matrix. From signal processing, the *Gini index* (GI),

Algorithm 9 DPR-RA algorithm

Initialization: Sets $\mathcal{X}$, $\mathcal{Y}$ and channel gain matrix $\mathbf{H}$

Set of matched Y players,$\mathsf{Y}_X = \emptyset$ and

Maximum number of matched players in set Y_X, J_{RA}

Set of requested Y players $\mathsf{Y}_X^{req} = \emptyset$

Set of rejected Y players $\mathsf{Y}_X^{rej} = \emptyset$

Initialize bid request $b_{Y\to X}^{RA}(t)$

Output: a stable matching μ_{RA}^*

Evaluate stable matching μ_{RA}^*.

Each player $Y \in \mathcal{Y}$ constructs $\succ_{Y,RA}$ using the metric $\Theta_Y(X)$.

while $\sum_{\forall X,Y} b_{Y\to X}^{UP}(t) \neq 0$ **do**

  for each un-associated Y: **do**

    Find $X = \arg\max_{X\in\succ_{Y,RA}} \Theta_Y(X)$

    Send a request $b_{Y\to X}^{RA}(t) = 1$ to player X

    for each player X **do**

      Find $X = \arg\max_{X\in\succ_{Y,RA}} \Theta_Y(X)$

      Update $\mathsf{Y}_X^{req} \leftarrow \left\{Y : b_{Y\to X}^{RA}(t) = 1, Y \in \mathcal{Y}\right\}$

      Construct $\succ_{Y,RA}$ using the metric $\Theta_X(Y)$

    end for

    if $\mathsf{Y}_X \leq J_{RA}$ **then**

      $\mathcal{Y}_X \leftarrow \mathsf{Y}_X$

    else

      Repeat

      Accept $Y = \arg\max_{Y\in\succ_{X,RA}} \sum_{Y\in\mathcal{Y}_X} \Theta_X(Y)$

      Update $\mathcal{Y}_X \leftarrow \mathcal{Y}_X \cup Y$

      Until $\left|\mathcal{Y}_X\right| = J_{RA}$

    end if

    Update $\mathsf{Y}_X^{rej} \leftarrow \left\{\mathsf{Y}_X^{rej} \setminus \mathsf{Y}_X\right\}$

    Remove player $X \in\succ_{Y,RA}, \forall Y \in \mathcal{Y}_X^{rej}$

  end for

end while

$\Xi(\vec{r})$ best describes the attributes of signal sparsity [17], [18]. We model the problem i.e., number of RUs in a codebook, d_v and using GI, we develop the codebook bounds.

Definition 4. *Given a vector $\vec{r} = [r_1 r_2 \cdots r_N]$ with its elements re-ordered and represented by $r_{[N]}$ for $n = 1, 2, \cdots N$ where $|r_{[1]}| \leq |r_{[2]}| \leq \cdots \leq |r_{[N]}|$, the Gini Index $\Xi(\vec{r})$ is given by*

$$\Xi(\vec{r}) = 1 - 2\sum_{n=1}^{N} \frac{r_{[n]}}{\|\vec{r}\|_1} \left(\frac{N - n + 0.5}{N} \right) \tag{B.46}$$

where $\|\vec{r}\|_1$ is the l_1 norm of $\vec{r}$ given by $\|\vec{r}\|_1 = \sum_n r_n$.

Contrary to other measures of sparsity, the GI, $\Xi(\vec{r})$ is normalized and assumes values between 0, for least sparse vector with most non-zero elements, and 1 for most sparse vector with least non-zero elements. Moreover, the value of $\Xi(\vec{r})$ is independent of the size of the vector, and hence can be used to determine sparsity of different size vectors.

Proposition 5. *Given an $N-$length "binary" vector, the necessary and sufficient sparsity GI value can be given as $\Xi(\vec{r}) \geq \frac{1}{2}$. The lower bound GI value can be given as $\Xi(\vec{r}) = \frac{1}{2}$.*

Proof. The proof is by induction: Consider three vectors $\vec{r}_1 = [1011]^T$, $\vec{r}_2 = [1001]^T$ and $\vec{r}_3 = [1000]^T$ of length $N = 4$. By using the l_0 norm, $\vec{r}_3$ is the sparsest followed by $\vec{r}_2$ and $\vec{r}_1$ respectively. In our context, codebook associated with $\vec{r}_1$ utilizes more d_v RUs than codebook associated with $\vec{r}_2$. The GI values $\Xi(\vec{r}_1) = 0.25$, $\Xi(\vec{r}_2) = 0.5$ and $\Xi(\vec{r}_) = 0.75$. It can be observed that increasing d_v RUs increases $\sum_{n=1}^{N} \frac{r_{[n]}}{\|\vec{r}\|_1} \left(\frac{N - n + 0.5}{N} \right)$. We observe that, $d_v \leq \frac{1}{2} \cdot N$ results in $\Xi(\vec{r} \geq 0.5$ sufficient sparsity for a vector of any length N. Thus, a lower bound GI value $\Xi(\vec{r} = 0.5$ results in an upper bound $d_v = \frac{N}{2}$ □

References

[1] S. M. 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.

[2] A. A. Khansa, Y. Yin, G. Gui, and H. Sari, "Power-domain noma or noma-2000?" in *2019 25th Asia-Pacific Conference on Communications (APCC)*, 2019, pp. 336–341.

[3] 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*, vol. 20, no. 3, pp. 2294–2323, 2018.

[4] 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.

[5] 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, 2018.

[6] A. Zakeri, M. Moltafet, and N. Mokari, "Joint radio resource allocation and sic ordering in noma-based networks using submodularity and matching theory," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9761–9773, 2019.

[7] S. Sharma, K. Deka, V. Bhatia, and A. Gupta, "Joint power-domain and scma-based noma system for downlink in 5g and beyond," *IEEE Communications Letters*, vol. 23, no. 6, pp. 971–974, 2019.

[8] S. Chege and T. Walingo, "Energy efficient resource allocation for uplink hybrid power domain sparse code nonorthogonal multiple access heterogeneous networks with statistical channel estimation," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 1, p. e4185, 2021.

[9] T. Sefako and T. Walingo, "Biological resource allocation algorithms for heterogeneous uplink pd-scma noma networks," *IEEE Access*, vol. 8, pp. 194 950–194 963, 2020.

[10] X. Yue, Z. Qin, Y. Liu, X. Dai, and Y. Chen, "Outage performance of a unified non-orthogonal multiple access framework," in *2018 IEEE International Conference on Communications (ICC)*, 2018, pp. 1–6.

[11] J. Chen, Z. Wang, W. Xiang, and S. Chen, "Outage probability region and optimal power allocation for uplink scma systems," *IEEE Transactions on Communications*, vol. 66, no. 10, pp. 4965–4980, 2018.

[12] G. Doaa, M.Ahmed, and E. Khaled, "User capacity for uplink scma system," *Phys. Commun.*, vol. 39, no. C, 2020.

[13] S. Samarakoon, N. Rajatheva, M. Bennis, and M. Latva-aho, "Outage probability and capacity for two-tier femtocell networks by approximating ratio of rayleigh and log normal random variables," in *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, 2013, pp. 1–5.

[14] M. Alam and Q. Zhang, "Performance study of scma codebook design," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, San Francisco, CA, USA, 2017, pp. 1–5.

[15] L. Li, Z. Ma, P. Z. Fan, and L. Hanzo, "High-dimensional codebook design for the scma down link," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 10 118–10 122, 2018.

[16] M. Zhao, S. Zhou, W. Zhou, and J. Zhu, "An improved uplink sparse coded multiple access," *IEEE Communications Letters*, vol. 21, no. 1, pp. 176–179, 2017.

[17] N. Hurley and S. Rickard, "Comparing measures of sparsity," *IEEE Transactions on Information Theory*, vol. 55, no. 10, pp. 4723–4741, 2009.

[18] D. Zonoobi, A. A. Kassim, and Y. V. Venkatesh, "Gini index as sparsity measure for signal reconstruction from compressive samples," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 927–932, 2011.

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

[20] P. Parida and S. S. Das, "Power allocation in ofdm based noma systems: A dc programming approach," in *2014 IEEE Globecom Workshops (GC Wkshps)*, Austin, TX, USA, 2014, pp. 1026–1031.

[21] N. Zhang, J. Wang, G. Kang, and Y. Liu, "Uplink nonorthogonal multiple access in 5g systems," *IEEE Communications Letters*, vol. 20, no. 3, pp. 458–461, 2016.

[22] G. Gui, H. Sari, and E. Biglieri, "A new definition of fairness for non-orthogonal multiple access," *IEEE Communications Letters*, vol. 23, no. 7, pp. 1267–1271, 2019.

[23] M. Zeng, A. Yadav, O. A. Dobre, and H. V. Poor, "Energy-efficient power allocation for uplink noma," in *2018 IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, United Arab Emirates, 2018, pp. 1–6.

[24] X. Li, C. Li, and Y. Jin, "Dynamic resource allocation for transmit power minimization in ofdm-based noma systems," *IEEE Communications Letters*, vol. 20, no. 12, pp. 2558–2561, 2016.

[25] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5g nonorthogonal multiple-access downlink transmissions," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6010–6023, 2016.

[26] C. Quan, A. Yadav, B. Geng, P. K. Varshney, and H. V. Poor, "A novel spectrally-efficient uplink hybrid-domain noma system," *IEEE Communications Letters*, vol. 24, no. 11, pp. 2609–2613, 2020.

[27] W. B. Ameur, P. Mary, M. Dumay, J. F. Hélard, and J. Schwoerer, "Performance study of mpa, log-mpa and max-log-mpa for an uplink scma scenario," in *2019 26th International Conference on Telecommunications (ICT)*, Hanoi, Vietnam,, 2019, pp. 411–416.

[28] J. Wang, B. Xia, K. Xiao, Y. Gao, and S. Ma, "Outage performance analysis for wireless non-orthogonal multiple access systems," *IEEE Access*, vol. 6, pp. 3611–3618, 2018.

[29] Y. Gu, W. Saad, M. Bennis, M. Debbah, and Z. Han, "Matching theory for future wireless networks: fundamentals and applications," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 52–59, 2015.

[30] T. LeAnh, N. H. Tran, W. Saad, L. B. Le, D. Niyato, T. M. Ho, and C. S. Hong, "Matching theory for distributed user association and resource allocation in cognitive femtocell networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 9, pp. 8413–8428, 2017.

[31] S. Bayat, Y. Li, L. Song, and Z. Han, "Matching theory : Applications in wireless communications," *IEEE Signal Processing Magazine*, vol. 33, no. 6, pp. 103–122, 2016.

[32] F. Berggrenand and S. Slimane, "Linear successive interference cancellation in ds-cdma systems," *Wireless Communications and Mobile Computing*, vol. 3, no. 7, pp. 847–859, 2003.

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

[34] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir, "A general power allocation scheme to guarantee quality of service in downlink and uplink noma systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7244–7257, 2016.

[35] A. Sultana, I. Woungang, A. Anpalagan, L. Zhao, and L. Ferdouse, "Efficient resource allocation in scma-enabled device-to-device communication for 5g networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5343–5354, 2020.

Part IV

Paper C

Paper C

MIMO based hybrid PD-SCMA Uplink Transceiver System

Simon Chege and Tom Walingo

Under Review

Submitted to IEEE Communication Letters

Abstract

The application of multiple-input multiple-output (MIMO) on power domain sparse code multiple access (PD-SCMA) system would enhance their performance by increasing the multiplexing and diversity gains at the cost of increased detection complexity as more users and antennas are deployed. This work develops and investigates the performance of spatial multiplexing MIMO based hybrid PD-SCMA system (M-PD-SCMA) transceiver on an uplink heterogeneous network with the aim of achieving a balance on the number of antennas and capacity/spectral efficiency. Numerical results exhibit performance benchmark with PD-SCMA schemes and the proposed receiver achieves guaranteed bit error rate (BER) performance with an increase in the number of transmit and receive antennas. Thus, the feasibility of an M-PD-SCMA system is validated.

1 Introduction

The merits and demerits of the evolution of non-orthogonal multiple access (NOMA) schemes from power domain NOMA (PD-NOMA), sparse code multiple access (SCMA), to the hybrid schemes like power domain sparse code non-orthogonal multiple access (PD-SCMA) and their applications in current communication networks have been well investigated [1–3]. This evolution is fueled by the creation of more capacity and spectral efficiency.

Multiple-input multiple-output (MIMO) technologies have greatly improved throughput in communication systems. Moreover, the integration of MIMO for improving the capacity of NOMA systems is being embraced [4]. In [5], authors investigate MIMO performance with multiple clustered users and analytically proves the superiority of MIMO-NOMA sum channel and ergodic capacity over MIMO-OMA. MIMO-NOMA finds applications in small packet transmissions in Internet of Things (IoT) users where users have diversified Quality of Service (QoS) requirements [6]. In [7], diversity and multiplexing based MIMO schemes for uplink SCMA system are investigated. It is demonstrated that improved spectral efficiency, capacity performance with reduced antennas [8] and BER performance [9] can be achieved for a MIMO-SCMA system. Authors in [10], [11] investigate the combined downlink detection of MIMO based SCMA for a near-optimal BER performance and notable reduced complexity.

Though the design and application of MIMO based NOMA technologies is still in its early stages, the modelling in hybrid NOMA schemes has so far not been undertaken to the best of our knowledge. This could lead to improving on the benefits of the hybrid system in throughput, capacity and diversity. However, this is fraught with the challenges of multiplexing at the transmitter and surging complexity at the receiver as the number of users and antennas grows. Inspired by [10] and [11], the main contribution of this work is the development and integration of MIMO schemes on a hybrid PD-SCMA (M-PD-SCMA) uplink system. We alleviate the integration challenges by employing spatial multiplexing (SM) based MIMO scheme at the transmitter where each transmit antenna at each layer transmits an independent SCMA user codeword. At the receiver, we develop a joint multi-user detection (MUD) based on modified expectation propagation algorithm (EPA) and successive interference cancellation (SIC) (J-EPA-SIC). From the system analysis, it can be observed that BER performance is dependent on the number of transceiver antennas. The system capacity references PD-SCMA capacity when system's point of operation is within the multiplexing bounds investigated in [3] and the receiver exhibits reduced complexity. The feasibility of an M-PD-SCMA system is thus validated. We denote by x, $\mathbf{x}$, $\mathbf{X}$ and $\mathcal{X}$ a scalar, vector, matrix and set respectively.

2 Proposed MIMO Based PD-SCMA

A PD-SCMA transmitter employing the SM based MIMO on a two-tier heterogeneous network (HetNet) model is considered. The HetNet model comprises of a centralized macro base station (MBS) serving a set of $\mathcal{U}, (|\mathcal{U}| = U)$ randomly distributed macro cell users (MUEs) and underlaid set of $\mathcal{F}, (|\mathcal{F}| = F)$ small cells, each characterized by a centralized low power small cell base station (SBS) serving a set of $\mathcal{J}, (|\mathcal{J}| = J)$ uniformly distributed SUEs [3]. Each user and BS has N_t and N_r transmit and receive antennas respectively.

The M-PD-SCMA transmitter consists of three stages; Firstly, the resource allocation (RA) stage, where the resource elements (REs), codebooks and power are assigned to user symbols and individual users respectively, and user pairing and clustering is done. Secondly, layered power domain (PD) multiplexing stage where codeword selection and multiplexing of the selected codewords from clustered users in PD is done. Lastly, antenna assignment stage where summed codewords from each layer are allocated to an antenna. Similarly, to recover the approximate transmitted user symbol $\hat{\mathbf{s}}$, the MUD operates the received signal at n_r in two stages; Firstly, modified EPA iterative process where iterative detection of the codewords is executed followed by user symbols reconstruction realized by computing the posterior log likelihood ratios (LLRs). Secondly, SIC process where the signals of users with weaker channel conditions are decoded and subtracted from the received codeword.

2.1 MIMO-based PD-SCMA Transmitter

The block diagram of the uplink M-PD-SCMA transmitter for n_t-th antenna is shown in Fig. C.1. Users are paired to form L clusters, where each cluster is assigned a unique codebook utilizing distinct REs. Similar to the conventional SCMA, a PD-SCMA transmitter operates L layers (of set $\mathcal{L}$), on which L independent symbol streams are transmitted. A layer is constructed by drawing select codewords from each user in the cluster matched to the layer i.e., $J = 1$ SUE and V MUEs from the set $\mathcal{V}_{CB}, (|\mathcal{V}_{CB}| = V, \mathcal{V}_{CB} \in \mathcal{U})$. This implies that each layer constitutes of $M = (V + 1)$ users' codewords and SUE to layer is a one-to-one matching, $L = J$. Prior to transmission, the M-PD-SCMA performs the following steps;

1. *Resource allocation*: Followed by V-BLAST encoding, forward error correction (FEC) and interleaving, every $log_2(M)$-bit user symbols are mapped, according to SCMA encoding, to a length$-K$ sparse vector resulting into complex codewords $\mathbf{s}^{SUE_{j,n_t}}$ and $\mathbf{s}^{MUE_{j,n_t}}$ respectively given by

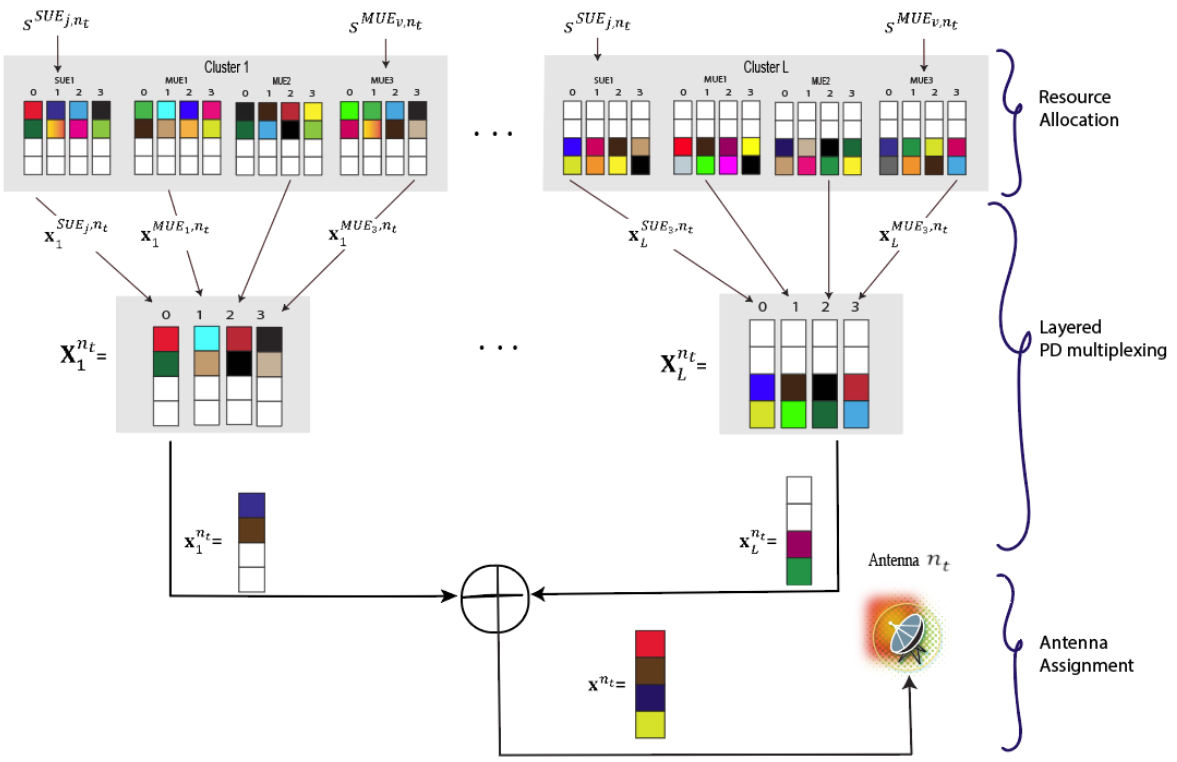

Fig. C.1: System model of the proposed uplink spatial multiplexing-based M-PD-SCMA system with $S = 4$ codewords.

$$\begin{aligned} \mathbf{s}^{SUE_{j,n_t}} &= \left[s^{SUE_{j,1,n_t}}, \cdots, s^{SUE_{j,K,n_t}}\right]^T, \\ \mathbf{s}^{MUE_{j,n_t}} &= \left[s^{MUE_{v,1,n_t}}, \cdots, s^{MUE_{v,K,n_t}}\right]^T \end{aligned} \tag{C.1}$$

These vectors belong to a finite set of $\mathcal{M}, (|\mathcal{M}| = M)$ codewords of a codebook CB. Here, each CB comprises of $M = 4$ codewords, $\mathcal{M} = \{0, 1, 2, 3\}$. The entries $\mathbf{s}^{SUE_{j,n_t}}$ and $\mathbf{s}^{MUE_{v,n_t}}$ denote respectively the j^{th} SUE and v^{th} MUE mapped to the k^{th} RE on a CB for the n_t-th antenna. As an example from Fig. C.1., the user information symbols utilize RE_1 and RE_2 of codebook CB_1. Subsequently, the transmitter performs codebook and power allocation utilizing RA schemes proposed in [3]. After RA, MUEs are paired with a SUE on a codebook using a user pairing scheme [3] to form the L clusters.

2. *Layered PD multiplexing*: For the transmitting antenna n_t, a codeword from each pairing user in the cluster is selected. The selected codewords are consolidated resulting into a layer $\mathbf{X}_l^{n_t}$ given as

$$\mathbf{X}_l^{n_t} = \left[\mathbf{x}_l^{SUE_{j,n_t}} \mathbf{x}_l^{MUE_{1,n_t}} \cdots \mathbf{x}_l^{MUE_{V,n_t}}, \right] \in \mathrm{C}^{K \times M} \tag{C.2}$$

The entries $\mathbf{x}_l^{SUE_{j,n_t}} = \sqrt{P_l^{SUE_j}} \cdot \mathbf{s}^{SUE_{j,n_t}}$ and $\mathbf{x}_l^{MUE_{v,n_t}} = \sqrt{P_l^{MUE_v}} \cdot \mathbf{s}^{MUE_{v,n_t}}$ and $P_l^{SUE_j}$ and $P_l^{MUE_v}$ are the normalized SUE and MUE power levels respectively. The codewords in

Algorithm 10 M-PD-SCMA Transmitter algorithm

Initialization Initialize the sets: $\mathcal{F}, \mathcal{U}, \mathcal{V}, \mathcal{J}$ and $\mathcal{L}$.
Stage I: Resource allocation
  Sparse encoding of incoming user symbols, eqn. (10).
for MUEs $v = 1 : V$ **do**
  MUE power allocation, $P_l^{MUE_v}$.
end for
for SBS $f = 1 : F$ **do**
  for SUE $j = 1 : J$ and $l = 1 : L$ **do**
    SUE power allocation, $P_l^{SUE_j}$.
    Codebook assignment to SUEs.
  end for
  SUE-MUEs pairing and clustering.
Stage II: Layered PD multiplexing
  for SUE $n_t = 1 : N_t$ and $m = 1 : M$ **do**
    Select codeword m from each user in cluster l and integrate them to obtain $\mathbf{X}_l^{n_t}$, eqn. (C.2).
    Perform PD multiplexing of message vectors in $\mathbf{X}_l^{n_t}$ to obtain $\mathbf{x}_l^{n_t}$.
Stage III: Antenna assignment
    Sum the codewords $\mathbf{x}_l^{n_t}$ from all the L layers to obtain $\mathbf{x}^{n_t}$.
    Perform layer - antenna assignment.
    Transmit $\mathbf{x}^{n_t}$ through antenna n_t.
  end for
end for

$\mathbf{X}_l^{n_t}$ are then multiplexed in power domain by diversifying the allocated power levels of the users in the clusters resulting to the layer message vector $\mathbf{x}_l^{n_t} \in \mathcal{C}^{K\times 1}$.

3. *Antenna assignment*: The vectors $\mathbf{x}_l^{n_t}$ from all the L layers are then summed together to obtain the transmit vector $\mathbf{x}^{n_t} \in \mathcal{C}^{K\times 1}$. The transmit message vector $\mathbf{x}^{n_t}$ is assigned to the n_t-th antenna and transmitted over the K subcarriers. Note that in the subsequent antenna n_{t+1}, the transmitter selects and transmits different codewords from the users in a cluster. The transmitter algorithm is presented in Algorithm 10.

Under the constraint that no two layers should be assigned all the same REs for an affordable complexity order, the system loading is given as $\lambda = M \times \begin{pmatrix} L \\ d_v \end{pmatrix}$. The received signal vector after

at the $n_r - th$ receiving, $\mathbf{y}^{n_r}$ reads,

$$\mathbf{y}^{n_r} = \sum_{n_t=1}^{N_t} \sum_{l=1}^{L} diag(\mathbf{h}_l^{n_t,n_r}) \mathbf{x}_l^{n_t} + \mathbf{z}^{n_r} \tag{C.3}$$

where $\mathbf{h}_l^{n_t,n_r} = \left[h_{l,SUE_j}^{n_t,n_r} h_{l,MUE_1}^{n_t,n_r} \cdots h_{l,MUE_V}^{n_t,n_r}\right]$. Through the $n_r - th$ receive antenna, $h_{l,SUE_j}^{n_t,n_r}$ and $h_{l,MUE_v}^{n_t,n_r}$ denote SUE and MUE channel coefficients averaged over the d_v in each layer l respectively. In the uplink, since the PD-SCMA codewords from different clusters are not multiplied by the same fading channel, a modified EPA MUD based should be considered. Consequently, we employ a normalized channel coefficient $g_l^{n_t,n_r}$ of the multiplexed layer signal given by

$$g_l^{n_t,n_r} = \sqrt{\left(|h_{l,SUE_j}^{n_t,n_r}|^2 + \sum_{v=1,v\in\mathcal{V}_c}^{V} |h_{l,MUE_v}^{n_t,n_r}|^2\right)} \tag{C.4}$$

2.2 MIMO-based PD-SCMA Receiver

A low complex modified joint EPA-SIC receiver is proposed for the uplink system. Unlike in MPA [10], modified EPA only pursues the means and variances of the transmitted messages during the iterative detection on the factor graph. The M-PD-SCMA MUD performs two steps repetitively at antenna n_r for all the layers;

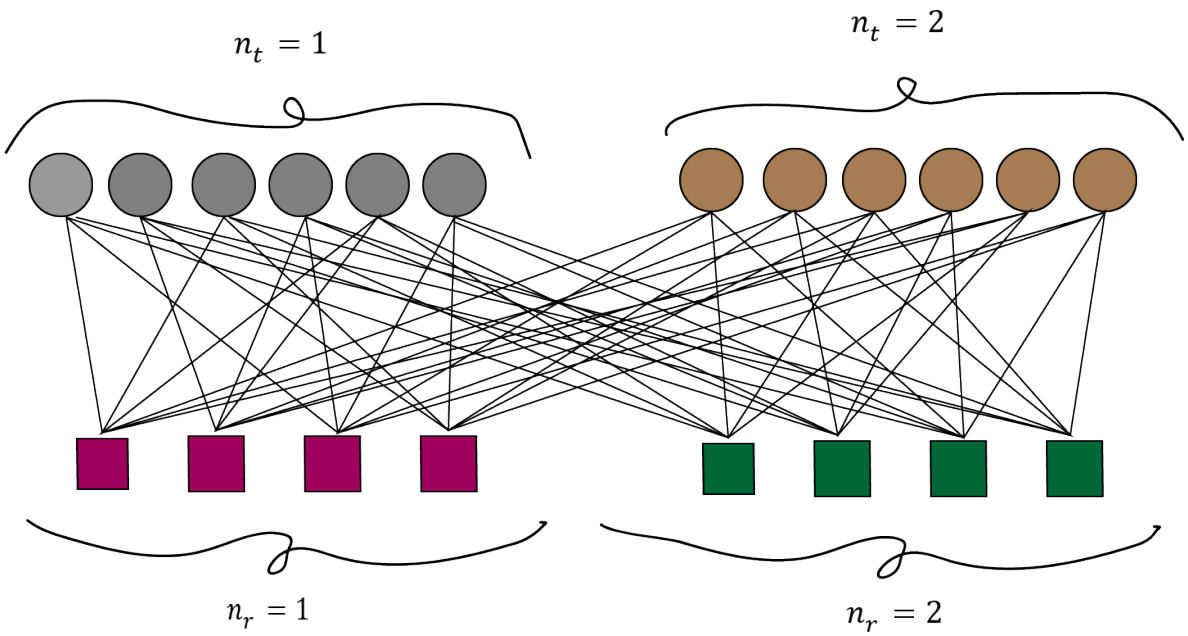

Fig. C.2: Factor graph representation for a SM based M-PD-SCMA with $L = 6$ and $K = 4$ for $n_t = 2$ and $n_r = 2$. codewords.

1. *Modified EPA process*: A factor graph representation for $n_t = n_r = 2$ is illustrated in Fig. C.2. The circles and squares represent the layer node (variable node, (VNs)) and receive node (resource node, (RNs)) respectively. Here, the factor graph is comprised of LN_t variable nodes v_l and KN_r resource nodes r_k, for which each variable node is a PD multiplexed symbol of M users. Denote by $I^{(\tau)}_{r_k \longrightarrow v_l}(\mathbf{x})$(respectively $I^{(\tau)}_{v_l \longrightarrow r_k}(\mathbf{x})$) the message corresponding to codeword $\mathbf{x}$ transmitted by (to) VN v_l to (by) r_k at the τ^{th} iteration. Based on [12], the modified EPA iterative process consists of the following steps;

   (a) Compute the posterior belief approximation $q^{n_t(\tau)}\big(\mathbf{x}_l^{n_t}|\mathbf{y}\big)$ for all $\mathbf{x}_l^{n_t} \in \mathcal{C}_l^{n_t}$ for each variable node $v_{l\in\mathcal{L}}$,

$$q^{n_t(\tau)}\big(\mathbf{x}_l^{n_t}|\mathbf{y}\big) \propto I_{\Delta\to k}(\mathbf{x}_l^{n_t}) \prod_{k\in\mathcal{I}_v(l)} I^{(\tau-1)}_{r_k\longrightarrow v_l}\big(\mathbf{x}_{kl}^{n_t}\big) \tag{C.5}$$

   where $\propto$ denotes equality up to scale and $\mathcal{I}_v(l)$ denotes the set of resource node indices connected to variable node v_l. We assume a uniform a-priori probability $I_{\Delta\to k}(\mathbf{x}_l^{n_t}) = \frac{1}{M}$.

   (b) Compute the posterior mean $\mu_{kl}^{(\tau)}$ and variance $\sigma_{kl}^{(\tau)}$ for each variable node $v_{l\in\mathcal{L}}$ and resource node $r_{k\in\mathcal{I}_v(l)}$ as follows;

$$\begin{aligned}
\mu_{kl}^{n_t(\tau)} &= \sum_{\mathbf{x}_l^{n_t}\in\mathcal{C}_l^{n_t}} q^{n_t(\tau)}\big(\mathbf{x}_l^{n_t}|\mathbf{y}\big)\cdot\mathbf{x}_{kl}^{n_t} \\
\sigma_{kl}^{n_t(\tau)} &= \sum_{\mathbf{x}_l^{n_t}\in\mathcal{C}_l^{n_t}} q^{n_t(\tau)}\big(\mathbf{x}_l^{n_t}|\mathbf{y}\big)\cdot\big|\mathbf{x}_{kl}^{n_t}-\mu_{kl}^{n_t(\tau)}\big|^2
\end{aligned} \tag{C.6}$$

   (c) Evaluate the means $\mu^{n_t(\tau)}_{v_l\longrightarrow r_k}$ and the variances $\sigma^{n_t(\tau)}_{v_l\longrightarrow r_k}$ of the messages $I^{n_t(\tau)}_{v_l\longrightarrow r_k} \propto \mathcal{CN}\Big(\mu^{n_t(\tau)}_{v_l\longrightarrow r_k}, \sigma^{n_t(\tau)}_{v_l\longrightarrow r_k}\Big)$;

$$\begin{aligned}
\sigma^{n_t(\tau)}_{v_l\longrightarrow r_k} &= \left(\frac{1}{\sigma_{kl}^{(\tau)}} - \frac{1}{\sigma^{n_t(\tau-1)}_{r_k\longrightarrow v_l}}\right)^{-1} \\
\mu^{n_t(\tau)}_{v_l\longrightarrow r_k} &= \left(\frac{\mu_{kl}^{n_t(\tau)}}{\sigma_{kl}^{n_t(\tau)}} - \frac{\mu^{n_t(\tau-1)}_{r_k\longrightarrow v_l}}{\sigma^{n_t(\tau-1)}_{r_k\longrightarrow v_l}}\right)^{-1}
\end{aligned} \tag{C.7}$$

   (d) Determine the means $\mu^{n_t(\tau)}_{r_k\longrightarrow v_l}$ and the variances $\sigma^{n_t(\tau)}_{r_k\longrightarrow v_l}$ of the messages $I^{n_t(\tau)}_{r_k\longrightarrow v_l} \propto \mathcal{CN}\Big(\mu^{n_t(\tau)}_{r_k\longrightarrow v_l}, \sigma^{n_t(\tau)}_{r_k\longrightarrow v_l}\Big)$;

$$\begin{aligned}
\mu^{n_t(\tau)}_{r_k\longrightarrow v_l} &= \frac{1}{g_l^{n_t,n_r}}\left(y_k^{n_r} - \sum_{i\in\mathcal{I}_r(k), i\neq l} g_l^{n_t,n_r}\cdot\mu^{n_t(\tau)}_{v_i\longrightarrow r_k}\right) \\
\sigma^{n_t(\tau)}_{r_k\longrightarrow v_l} &= \frac{1}{\big|g_l^{n_t,n_r}\big|^2}\left(N_0 - \sum_{i\in\mathcal{I}_r(k), i\neq l} \big|g_l^{n_t,n_r}\big|^2\cdot\sigma^{n_t(\tau)}_{v_i\longrightarrow r_k}\right)
\end{aligned} \tag{C.8}$$

   where $\mathcal{I}_r(k)$ denotes the set of variable node indices connected to resource node r_k.

After τ_{max} iterations on (C.6), (C.7) and (C.8) we can obtain $q^{n_t(\tau_{max})}\left(\mathbf{x}_l^{n_t}|\mathbf{y}\right)$ and the posterior LLRs $\Lambda_{kl}^{n_t}$ can then be computed using (5) in a similar way to [12]. We initialize the iterations with $\mu_{r_k \longrightarrow v_l}^{n_t(0)} = 0$ and $\sigma_{r_k \longrightarrow v_l}^{n_t(0)} = \infty$ where ∞ is taken as a large positive constant.

2. *SIC process*: Having successfully recovered a layer, SIC is employed to detect the M PD multiplexed users. Prior to decoding, the receiver computes the decoding order metric $\mathcal{N}_l$ proposed in [3] for layer l. Then, the instant decoding order π is determined based on the instantaneous received user signal power. Subsequently, users are decoded in the sequence $\left[MUE_1, MUE_2, \cdots, MUE_V, SUE_j\right]$. The highest ranked user experiences interference from all users while the lowest channel gain user effectively enjoys interference-free transmission.

The complexity of the EPA can be approximated to $\mathcal{O}\left(N_t K N_r M d_f\right)$, observed to linearly scale the both M and the degree of superposition d_f on a given RE which is lower than the message passing algorithm (MPA) counterpart exhibiting $\mathcal{O}\left(K N^r M^{N_t d_f}\right)$. The complexity of SIC is primarily in the computation of the decoding order metric for each user multiplexed in the layer, and is given as $\mathcal{O}(b^3)$ for a MMSE transformation weight matrix of $b \times b$. Consequently, the overall J-EPA-SIC receiver complexity can approximately be given by $\mathcal{O}\left(N_t K N_r M d_f + M b^3\right)$.

3 Results and Discussion

In this section, the bit error rate (BER), capacity and complexity performance of the uplink M-PD-SCMA system are presented. Denote by ρ the maximum number of bits per user transmitted by two antennas during two transmission channel slots. We consider a SM based M-PD-SCMA that transmits 4 codewords which is equivalent to $\rho = 8$ bits/user/2 transmit antennas/2 channel use periods, when a codebook of size $M = 4$ codewords is employed. Other simulation properties are as employed in [3].

Similar to numerical analysis done in [10], Fig. C.3 represents the BER performance as the number of antennas grows for both MPA- and EPA- based M-PD-SCMA schemes. For both receiver schemes, the BER performance improves as the number of receive antennas increases. The EPA based receiver closely approximates the near-optimal MPA based receiver even for higher number of antennas. Therefore, by employing more antennas, the EPA detector can achieve a near optimal performance with low complexity.

Fig. C.4 depicts the system capacity versus the number of SUEs/layers in comparison with other NOMA schemes. It can be observed that the system capacity for all schemes increase sharply for low number of SUEs up to approximately 12 SUEs (12 layers), beyond which the capacity growth is

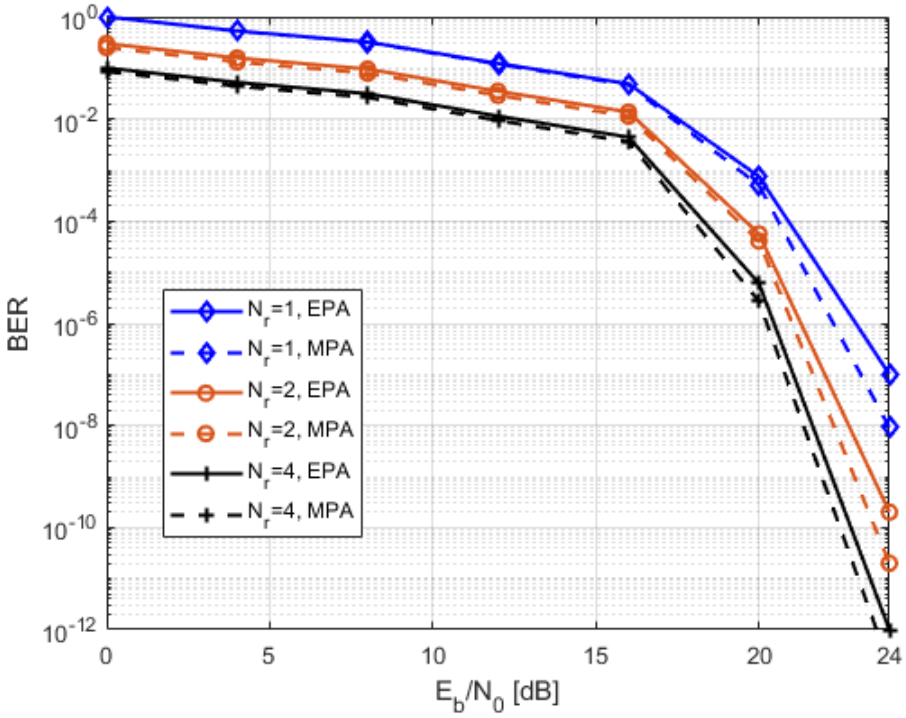

Fig. C.3: BER performance versus N_r with $\rho = 4$.

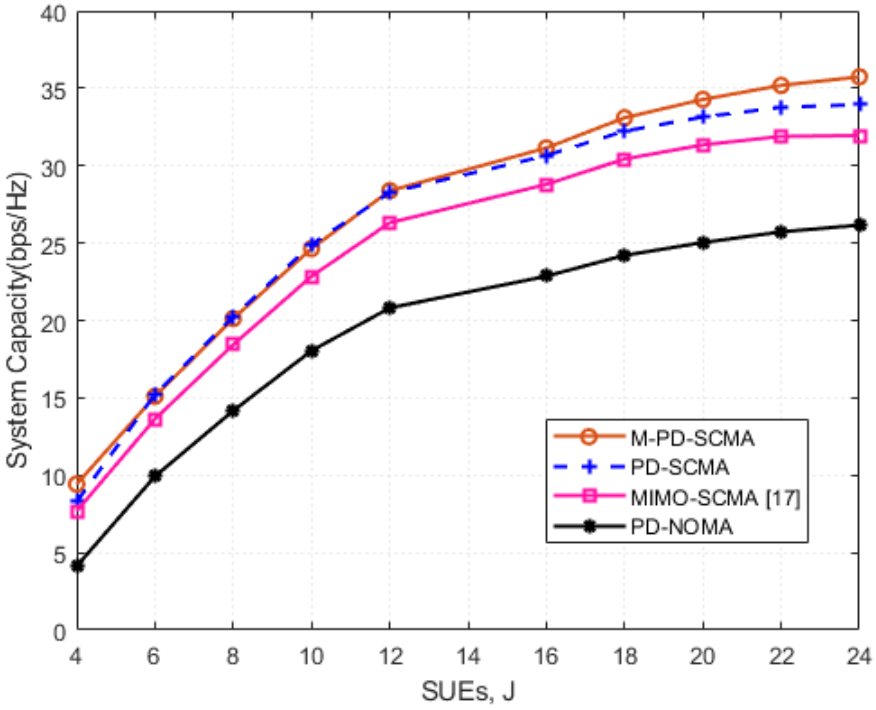

Fig. C.4: Capacity vs number of SUEs.

gradual. Increasing the number of layers results in aggravated interference that degrades the performance. This implies that the optimal capacity can be obtained when the number of users is

within the multiplexing bounds. It can be observed that M-PD-SCMA capacity benchmarks PD-SCMA and evidently outperforms MIMO-SCMA and the PD-NOMA scheme.

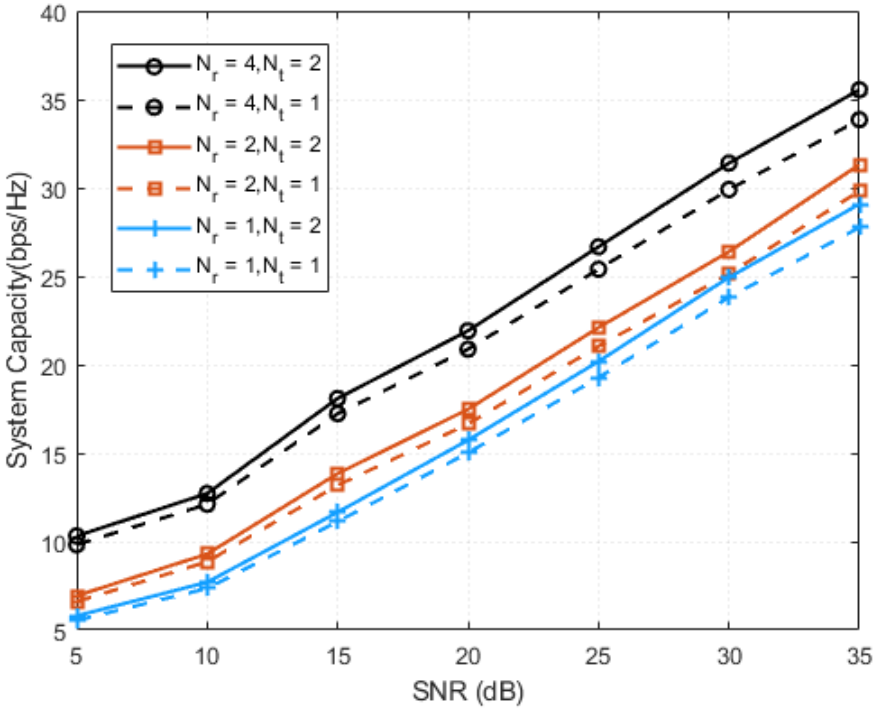

Fig. C.5: Capacity versus SNR.

Fig. C.5 illustrates the system capacity versus signal to noise ratio (SNR) for different number of transmit and receive antennas. Using 12 layers and $\rho = 4$, capacity increases monotonically with the SNR for different values of N_r. In fact, it can be observed that employing higher number of antennas achieves a higher capacity due to enhanced spatial multiplexing order. In varying the N_t, the capacity closely follows the capacity for different values of N_r hence satisfactorily justifying the use of lowered number of transmit antennas for the same achieved system capacity.

Lastly, the computational complexity vs the number of SUEs or rather the number of layers employed for the proposed MUD algorithm is shown in Fig. C.6. In this case we consider a fixed modulation order $M = 4$ and compare the EPA with MPA based MUDs complexity orders. It can be observed that the receiver MUD becomes more complex for both MPA and EPA as the value of L increases. This can be attributed to the increased number of VN indices connected to a single RN r_k. Predictably from the results, EPA based MUD exhibits significantly lower complexity order than MPA. Since only the means and variances of the messages are followed iteratively, EPA results in a linearly scaling complexity order unlike the exponential order resulting with MPA.

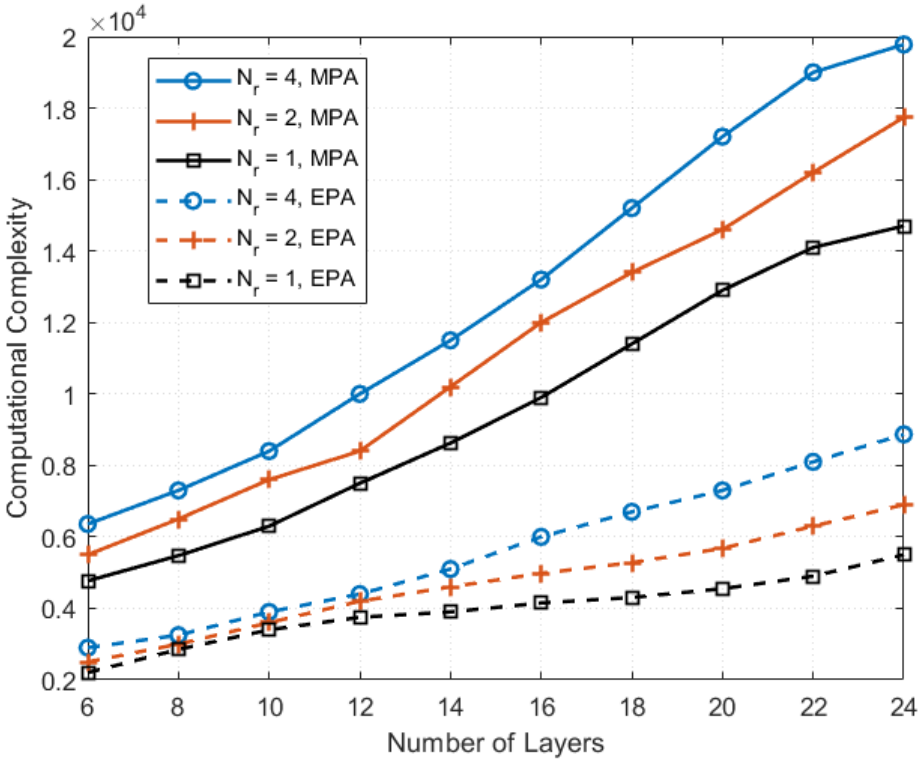

Fig. C.6: Receiver complexity versus number of layers

4 Conclusion

We have investigated the performance of spatial multiplexing uplink MIMO based hybrid power domain sparse code multiple access (M-PD-SCMA) system. To enhance performance at the receiver, a joint MUD based on EPA and SIC is proposed. The system's performance is analyzed based on the parameters of capacity, BER and complexity. Results show that M-PD-SCMA capacity benchmarks the PD-SCMA capacity. Employing EPA based MUD significantly reduces the complexity order compared to MPA hence a good candidate MUD even as the number of antennas grow.

References

[1] 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.

[2] S. Chege and T. Walingo, "Energy efficient resource allocation for uplink hybrid power domain sparse code nonorthogonal multiple access heterogeneous networks with statistical channel estimation," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 1, p. e4185, 2021.

[3] S. Chege and T. Walingo., "Multiplexing capacity of hybrid PD-SCMA heterogeneous networks," *IEEE Transactions on Vehicular Technology*, 2022, [Online]. Available: http://doi.org/10.1109/TVT.2022.3162304.

[4] Y. Huang, C. Zhang, J. Wang, Y. Jing, L. Yang, and X. You, "Signal processing for mimo-noma: Present and future challenges," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 32–38, 2018.

[5] M. Zeng, A. Yadav, O. A. Dobre, G. I. Tsiropoulos, and H. V. Poor, "Capacity comparison between mimo-noma and mimo-oma with multiple users in a cluster," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2413–2424, 2017.

[6] Z. Ding, L. Dai, and H. V. Poor, "Mimo-noma design for small packet transmission in the internet of things," *IEEE access*, vol. 4, pp. 1393–1405, 2016.

[7] Z. Mheich, I. A. Hemadeh, Z. Liu, and P. Xiao, "Low-complexity expectation propagation detection for uplink mimo-scma systems," in *2020 International Conference on UK-China Emerging Technologies (UCET)*, 2020, pp. 1–4.

[8] S. Han, C. Guo, W. Meng, C. Li, Y. Cui, and W. Tang, "The uplink and downlink design of mimo-scma system," in *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2016, pp. 56–60.

[9] Z. Pan, J. Luo, J. Lei, L. Wen, and C. Tang, "Uplink spatial modulation scma system," *IEEE Communications Letters*, vol. 23, no. 1, pp. 184–187, 2018.

[10] Y. Du, B. Dong, Z. Chen, P. Gao, and J. Fang, "Joint sparse graph-detector design for downlink mimo-scma systems," *IEEE Wireless Communications Letters*, vol. 6, no. 1, pp. 14–17, 2016.

[11] S. Tang, L. Hao, and Z. Ma, "Low complexity joint mpa detection for downlink mimo-scma," in *2016 IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1–4.

[12] X. Meng, Y. Wu, Y. Chen, and M. Cheng, "Low complexity receiver for uplink scma system via expectation propagation," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, 2017, pp. 1–5.

Part V

Conclusion

1 Conclusion

To conclude the thesis, this section provides a summary of the research contributions realized in this work and provides future research directions.

In the introduction, we present a detailed description of the evolution wireless communication networks (including 4G and 5G networks). The architecture, challenges, technical requirements and potential facilitating technologies of 5G are well elaborated. This work focuses on NOMA and therefore, an exhaustive analysis on the performance of the existing NOMA schemes is carried out with the focus to develop a potential hybrid NOMA technique. We then formulate the research problem, draw the research objectives, highlight the research overview and lastly present a summary of the main contributions of the research. Despite their merits of enhanced data rate and capacity, 5G networks exhibit challenges of resource allocation, interference management, spectral and energy efficiency occasioned by HetNet deployments. One of the potential solutions for 5G realization is the design of appropriate resource allocation and MUD schemes crucial in optimizing the performance of these networks. This work develops a multi-tier HetNet architecture for 5G NOMA networks on the uplink. The network modelling, feasibility, proposed schemes and their performance analysis is expounded in the research outputs as presented in the paper summaries that follow.

In paper A, a hybrid PD-SCMA HetNet model that combines power and code domain to multiplex MUEs and SUEs on 5G uplink networks was developed. The model employs SCOA for UP, OMSP for CA and QAPA for PA algorithms. The designed receiver utilizes a low-complexity MUD scheme based on joint SIC and Log- MPA. Simulation and analytical-based EE resource allocation performance for the small cells under QoS constraints of minimum sum-rate, interference temperature, system maximum power is evaluated. Based on the results, the feasibility of hybrid PD-SCMA as a multiplexing technique for MUEs and SUEs for future networks is validated. In addition, PD-SCMA outperforms SCMA and PD-NOMA with or without channel estimation error albeit with increased complexity. The developed RA schemes greatly improve the performance of PD-SCMA.

In paper B, the multiplexing capacity bounds for the uplink hybrid PD-SCMA is investigated with reference to the number of admissible codebooks, number of multiplexed MUEs that can be paired with a SUE in codebook and the maximal MUEs S that can be multiplexed in a codebook subject to the optimal number of power levels. Furthermore, to alleviate the RA challenges in the uplink PD-SCMA, we proposed and investigated the performance of DPR-RA schemes. It can be observed that DPR-RA based schemes greatly improve the performance of PD-SCMA. By employing the

proposed CA, UP and PA ranking parameters namely the measures of sparsity, SIC constraints and users' QoS requirements, the codebook capacity bounds, MUE multiplexing bounds and power multiplexing bounds are derived. In addition, the numerical system capacity and outage probability results validate our analytical multiplexing bounds. With a specified outage, the maximum number of MUEs to be multiplexed with SUE in a codebook for PD-SCMA system with a given number of codebooks is determined. The work opens a discussion on the multiplexing capacity of the hybrid PDSCMA scheme for a feasible NOMA technology.

In paper C, the feasibility and performance of spatial multiplexing uplink MIMO based hybrid PD-SCMA (M-PD-SCMA) system is investigated. To enhance performance at the receiver, a joint MUD based on EPA and SIC is proposed. The system's performance is analysed based on the parameters of capacity, BER and complexity. Results show that M-PD-SCMA capacity benchmarks the PD-SCMA capacity. Employing EPA based MUD significantly reduces the complexity order compared to MPA hence a good candidate MUD even as the number of antennas grows.

2 Future Research

The developed hybrid NOMA and the proposed RA and MUD schemes has potential for further improvement. Since PD-SCMA utilizes uniquely designed SCMA codebooks at the code-domain, the multiplexing performance can be enhanced through prudent design of multi-user multi-dimensional codebooks. The developed RA protocols should be investigated from the computational model's perspective. Besides, the performance parameters like computational times can be further explored in detail for implementation. The application of hybrid NOMA and the proposed RA schemes can be extended to other network environments such as cognitive radio, visible light communications, device to device, machine to machine and vehicle to vehicle and infrastructure amongst others.