Hybrid Massive MIMO Architecture with SWIPT in NOMA Systems

by

Ahlam JAWARNEH

THESIS PRESENTED TO ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
IN PARTIAL FULFILLMENT FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
Ph.D.

MONTREAL, DECEMBER 13,2022

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

BOARD OF EXAMINERS

THIS THESIS HAS BEEN EVALUATED

BY THE FOLLOWING BOARD OF EXAMINERS

Mr. Michel Kadoch, Thesis supervisor
Department of Electrical Engineering, École de technologie supérieure

Mrs. Nadjia Kara, Chair, Board of Examiners
Département de génie logiciel et des TI, École de technologie supérieure

Mr. Abdelouahed Gherbi, Member of the Jury
Department of Software and IT Engineering, École de technologie supérieure

Mrs. Halima Elbiaze, External Independent Examiner
Computer Science departement, Université du Québec à Montréal

THIS THESIS WAS PRESENTED AND DEFENDED

IN THE PRESENCE OF A BOARD OF EXAMINERS AND THE PUBLIC

ON NOVEMBER 23,2022

AT ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor Professor Michel Kadoch for his unlimited support, valuable advice, and recommendations during my Ph.D. study.

I would also like to convey my deepest appreciation to my lovely daughters Hala and Jana for Their unconditional love, support, and great patience throughout whole difficulties we faced. I would also like to thank my mother, sisters, brothers, and friends for everything they have done for me.

I would like to thank Dr. Zaid Albataineh from Yarmouk University for his support and help during our Ph.D. journey. He gave me his valuable time and extensive experience which constituted a great addition to the research work, as his guidance and advice were the beacons that I used in my entire research work.

Last, but by no means least, I would thank Baker for believing in me and for his support.

Architecture MIMO massive hybride avec SWIPT dans les systèmes NOMA

Ahlam JAWARNEH

RÉSUMÉ

La récupération d'énergie radiofréquence (EH) est apparue comme une possibilité importante pour l'augmentation de l'efficacité énergétique des réseaux existants et futurs. L'EH a également émergé comme une option durable, permettant la réutilisation d'une partie de l'énergie transmise à travers un cycle énergétique. En outre, l'énorme besoin de débits de données élevés et l'augmentation du trafic dans les réseaux sans fil ont suscité des recherches sur le spectre des ondes millimétriques (ondes mm) pour les futurs systèmes de communication (5G). Les réseaux d'antennes massives sont utilisés pour cette haute fréquence. Cependant, les coûts de déploiement élevés et la complexité du traitement du signal rendent les systèmes de formation de faisceau entièrement numériques (homogènes) coûteux et inefficaces. De plus, l'utilisation d'un système de formation de faisceau purement analogique réduit les performances du système en raison de ses limites matérielles. Pour résoudre les problèmes susmentionnés, la formation de faisceau hybride a été proposée. Trouver une conception appropriée reste une difficulté importante pour le chercheur. Dans cette thèse, nous formulons et présentons une nouvelle méthode basée sur l'algorithme LRE-CG (Long Recurrence Enlarged Conjugate Gradient) pour réaliser itérativement la méthode de détection MMSE, tout en évitant les complexités d'inversion de la matrice dans les systèmes MIMO massifs. Nous démontrons que ses performances sont supérieures à celles des méthodes actuelles et compétitives. Nous adaptons ensuite son application aux systèmes MIMO massifs et aux systèmes MIMO massifs à ondes millimétriques avec le système NOMA. Les résultats de la simulation ont permis de vérifier que la méthode présentée surpasse les méthodes conventionnelles trouvées dans la littérature, y compris l'approche basée sur l'approximation des séries de Neumann et l'itération de Gauss-Siedel et les méthodes itératives de Gauss-Siedel (GS) pour les systèmes MIMO massifs. L'algorithme proposé parvient à atteindre une performance quasi optimale à celle de l'algorithme MMSE typique avec un niveau minimal d'itérations. Deuxièmement, nous abordons le problème de la consommation d'énergie associée aux éléments de signaux mixtes tels que les éléments analogiques-numériques dans les systèmes MIMO massifs à ondes millimétriques (mmWave). Nous utilisons la méthode d'accès multiple non orthogonal (NOMA) dans les systèmes MIMO massifs à ondes millimétriques (mmWave) pour améliorer l'efficacité spectrale. La technologie simultanée de transmission d'informations et d'énergie sans fil (SWIPT) sera utilisée dans les systèmes MIMO massifs à ondes millimétriques. L'utilisation de SWIPT contribue à prolonger la durée de vie de la batterie des utilisateurs mobiles (MU) et à améliorer l'efficacité énergétique (EE) du système, en particulier dans le cas des systèmes MIMO où l'interférence d'inter-utilisateur peut être réutilisée pour la récupération de l'énergie (EH). Cependant, nous avons initialement conçu un algorithme de regroupement(clustering) des utilisateurs basé sur l'algorithme de regroupement(clustering) par propagation d'affinité, qui rassemble préférentiellement les équipements d'utilisateur (UE) en fonction de leur canal de corrélation et de leur distance. Subséquemment, nous concevons un encodeur RF analogique basé sur la sélection du groupe d'utilisateurs pour tous les faisceaux, suivi d'une conception d'un encodeur numérique en bande de base de faible dimension pour

atténuer davantage les interférences entre faisceaux, et maximiser le débit total réalisable pour le système considéré. Par la suite, nous transformons le problème d'optimisation original en deux sous problèmes : le premier concerne d'allocation de puissance conjointe et le deuxième est sur la répartition de puissance. Le problème d'optimisation non convexe considéré est difficile à résoudre, en raison de la présence de variables jointes et d'interférences entre utilisateurs. Pour faire face à ce problème, une approche disjointe est adoptée, dans laquelle l'allocation et la répartition de la puissance sont séparées, et les sous-problèmes correspondants sont résolus à l'aide de la méthode de dualité lagrangienne. Les résultats de la simulation confirment l'efficacité de la méthode proposée et démontrent qu'elle est quasi-optimale et qu'elle bénéficie d'une meilleure efficacité spectrale et énergétique par rapport aux conceptions abordées dans l'état de l'art, et à la méthode conventionnelle à savoir le Système MIMO-NOMA à ondes millimétriques compatible avec SWIPT.

Mots-clés: Récupération d'énergie, Onde Millimétrique, Formation de Faisceau Hybride, Massifs MIMO, Gradient Conjugué Elargi à Longue Récurrence, Accès Multiple Non Orthogonal, La Technologie de Transmission Simultanée D'informations et D'énergie Sans Fil.

Hybrid Massive MIMO Architecture with SWIPT in NOMA Systems

Ahlam JAWARNEH

ABSTRACT

It is becoming apparent that radio frequency energy harvesting (EH) presents a vital potential for improving the energy efficiency of existing and future networks. EH has also emerged as a viable option, allowing for the re-use of a portion of the transmitted energy through an energy cycle. Moreover, the enormous need for large data speeds and increased traffic in wireless networks has prompted research into the millimeter wave (mm-Wave) spectrum for future (5G) communication systems. Massive antenna arrays are used for this high frequency. However, high implementation costs and signal processing complexity make completely digital (homogeneous) beamforming systems expensive and inefficient, while employing purely analog beamforming reduces system performance owing to hardware limits. To address the aforementioned issues, hybrid beamforming has been proposed. Finding an appropriate design remains a significant difficulty for the researcher. In this thesis, we first formulate and present a new method based on the long recurrence enlarged conjugate gradient (LRE-CG) algorithm for iteratively realizing the MMSE detection method while avoiding the matrix inversion complexities in massive MIMO systems. We demonstrate its superior performance to present competing methods. Then we tailor its application to massive MIMO systems and mmWave massive MIMO systems with NOMA System. It has been verified by the simulation results that the presented method outperforms the conventional methods in the literature including the Neumann series approximation-based approach and Gauss-Siedel iterative (GS) methods for massive MIMO systems. The proposed algorithm succeeds to attain the near-optimal performance of a typical MMSE algorithm with a minimal level of iterations. Secondly, we deal with the issue of power consumption in millimeter-wave (mmWave) huge MIMO systems caused by mixed-signal components like analog-to-digital converters. In order to increase the efficiency of the spectrum, we use non-orthogonal multiple access (NOMA) in mmWave large MIMO systems. Massive multi-input multi-output (MIMO) systems operating at millimeter waves will make advantage of the simultaneous wireless transmission of information and power (SWIPT). In the NOMA scenario, where inter-user interference may be utilized for energy harvesting, SWIPT is particularly helpful in extending the battery life of mobile users (MUs) and improving the system energy efficiency (EE) (EH). To begin with, we developed a user grouping algorithm based on affinity propagation clustering, which groups user equipment (UE) based on channel correlation and distance. Following that, we designed the analog RF precoder for all beams using the selected user grouping, then we designed a low-dimensional digital baseband precoder to achieve a maximum sum-rate as well as minimize inter-beam interference for the system. Afterward, we formed a joint power allocating and power splitting optimization problem. Given the existence of linked variables and inter-user interference, the studied non-convex optimization issue is challenging to handle. In order to address this issue, a decoupled approach is used, in which the power allocation and power splitting are treated as independent issues, and the Lagrangian duality technique is used to solve these sub-problems. The simulation findings validate the efficacy of the proposed technique and show that it is near-optimal and has superior spectrum and energy efficiency compared to the previous designs and the existing SWIPT-enabled mmWave MIMO-NOMA system.

Keywords: Energy Harvesting, millimeter wave, Hybrid Beamforming, Massive MIMO, Long Recurrence Enlarged Conjugate Gradient, Simultaneous Wireless Information and Power Transmission technology, Non-Orthogonal Multiple Access.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF ALGORITHMS

LIST OF ABBREVIATIONS

5G	Fifth generation technology standard
3GPP	Third Generation Partnership Project
ADC	Analog to digital converter
AoA	Angle of Arrival
AoD	Angle of Departure
AWGN	Additive White Gaussian Noise
BF	Beamforming
BEP	Bit Error Probability.
BER	Bit error rate
BS	Base station
BW	Band Width
CB	Conjugate Beamforming
CG	Conjugate Gradient
CSI	Channel State Information
D2D	Device to Device.
DAC	Digital to Analog Converter
DoA	Direction of Arrival
DoF	Degree of Freedom
DSP	Digital Signal Processing

EE	Energy Efficiency
EH	Energy Harvesting
EGC	Equal Gain Combining
FPGA	Field Programmable Gate Array
GS	Gauss-Seidel
HB	Hybrid beamforming
HP	Hybrid Precoding
ID	Information Decoding
IID	Independent Identically Distributed
IM	Index Modulation
IoT	Internet of Things
ISI	Inter Symbol Interference
IUI	Inter User Interference
LNA	Low Noise Amplifier
LoS	Line of Sight
LRE-CG	Long Recurrence Enlarged Conjugate Gradient
LSV	Left Singular Vector
LTE	Long Term Evolution
LTE-A	Long Term Evolution Advanced
MIMO	Multiple-input and multiple-output

MISO	Multiple Input Single Output
ML	Maximum Liklihood
MMSE	Minimum Mean Square Error
mMTC	Massive Machine-Type Communications
mmWave	Millimeter-Wave
MS	Mobile station
MUs	Mobile users
MU-MIMO	Multi-user
MU-MIMO	Multi-user MIMO
NOMA	Non-orthogonal Multiple Access
NSE	Neumann Serious Expansion
OFDM	Orthogonal Frequency Division Multiplexing
OMP	Orthogonal Matching Pursuit
PA	Power Amplifier
PAPR	Peak to Average Power Ratio
PDF	Probability Density Function
PS	Power Splitting
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QRD	QR Decomposition

QSM	Quadrature Spatial Modulation
RF	Radio frequency
RMS	Root Mean Square
RSV	Right Singular Vector
SE	Spectral Efficiency
SIMO	Single Input Multiple Output
SINR	Signal to Interference and Noise Ratio
SISO	Single Input Single Output
SLNR	Signal to Leakage and Noise Ratio
SWIPT	Simultaneous Wireless Information and Power Transmission Technology
SM	Spatial Modulation
SNR	Signal-to-noise Ratio
SU	Single User
SVD	Singular Value Decomposition
TDD	Time Division Duplex
3D	Three dimensional
UE	User Equipment
ULA	Uniform linear array
UPA	Uniform planar array
WLAN	Wireless Local Area Network

WPT	Wireless Power Transfer
WSMSE	Weighed sum mean square error
ZF	Zero forcing

LIST OF SYMBOLS AND UNITS OF MEASUREMENTS

μ	Channel gain
δ	Noise and interference variance
E	Equivalent channel matrix
G	Gram matrix
H	Channel matrix
$\hbar_k$	signal to interference plus noise ratio for the k-th user
K	Number of users
N	Number of antennas
n	additive white Gaussian noise vector
s	Transmitted signal vector
W	MMSE linear detection matrix
y	Received signal vector
σ	Standard Deviation
s	Transmitted signal vector
$\zeta_{K,b}$	of the k-th user

INTRODUCTION

Meeting the massive data traffic created by the fast proliferation of mobile devices and countless innovative apps is a key problem for mobile networks. According to Cisco (Cisco, 2022), worldwide mobile data traffic increased by 63% in 2016 to 7.2 exabytes per month at the end of 2016, up from 4.4 exabytes per month at the end of 2015. Between 2016 and 2022, the volume will increase seven-fold, reaching 49.0 exabytes per month 2022. Furthermore, it is estimated that there will be 11.6 billion mobile devices, which is three times the worldwide predicted population in 2022. The evolution of wireless communication systems always starts with novel top-level criteria, which include 5G mobile technologies and subsequent generations. They now have new operational modes: massive machine-type communications (mMTC), ultra-reliable low-latency communications URLLC, and enhanced mobile broadband (eMBB) (3rd Generation Partnership Project, 2014). Each of these modes has a set of underlying properties Figure 0.1 that enable the creation of innovative applications. Smart cities, Industry automation, and virtual reality and further applications as demonstrated on 0.1 are only a few of the widely utilized application cases described in the research pillar (Gupta & Jha, 2015; Dai *et al.*, 2015; Osseiran *et al.*, 2014). High dependability, energy efficiency (EE), and spectrum efficiency (SE) become essential factors for them.

The fourth generation (4G) wireless communication, Advanced (LTE-A), is now in use. LTE-A improves capacity and coverage by implementing technology such as Carrier Aggregation (CA), Coordinated Multipoint (CoMP), Advanced Multiple Input Multiple Output (MIMO), Relays, Heterogeneous Networks (HetNets), and improved Inter-Cell Interference Cancellation (eICIC) (3rd Generation Partnership Project, 2014). LTE-A intends to achieve peak rates of one GBps in the downlink and 500 Mbps in the uplink, mobility of up to 350 km/h, and latency of fewer than fifty milliseconds. These performance requirements, however, will not be able to respond to the massive development of future mobile data networks, as demand for mobile services and applications is rapidly overloading wireless spectrum capacity. The implementation

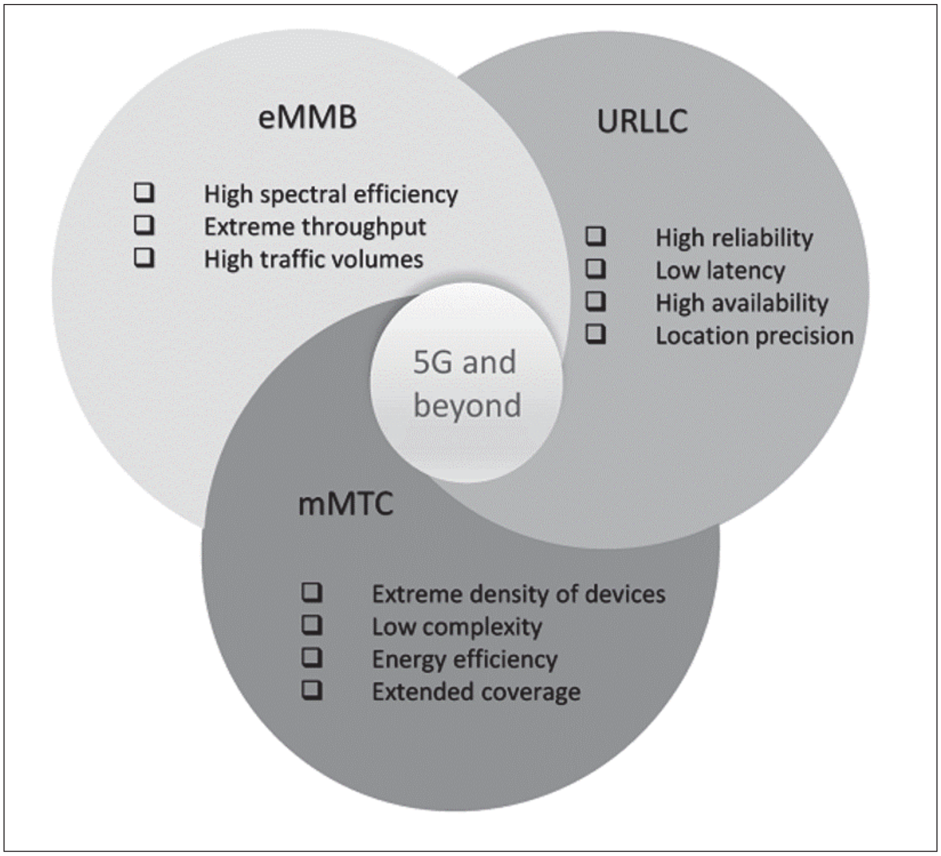


Figure 0.1 Operation modes of 5G wireless networks
Taken from Pérez et al. (2021)

of the fifth generation (5G) wireless communications technology has commenced. Many technologies, such as large MIMO, mmWave communication (Gupta & Jha, 2015; Dai *et al.*, 2015), and non-orthogonal multiple access (NOMA), have been suggested. Device-to-Device (D2D) communication (Doppler, Rinne, Wijting, Ribeiro & Hugl, 2009) is another approach for increasing the possible rate in wireless communications, in which devices may exchange data directly without the requirement for a base station. 5G is planned. It supports many applications subject to relevant technological requirements (Osseiran *et al.*, 2014). 5G promises to handle data-intensive applications such as streaming video, online gaming, and virtual reality. When compared to existing 4G networks, the data rate is 1000 times greater (1 to 10 Gb per second). Furthermore, future wireless networks will be diverse, including technologies such as the Internet of Things (IoT) and machine-to-machine (M2M) connections. Furthermore, the Next Generation Network focuses on creating a more environmentally friendly communication infrastructure. Due to sustainable growth, environmental, and economic considerations, increased transmit powers will not result in a significant rise in data rates. Instead, the thousand-time data-rate increase will have to be accomplished with less energy use. As a result, improved spectrum and power efficiency are high-priority targets for future network architectures. As can be seen, future wireless networks must not only be resilient, adaptable, and energy efficient, but they must also deliver high-quality, low-cost services to customers. Nevertheless, (Figure 0.1) also shows EE as a future network feature, with particular significance for enabling a sustainable Internet of Things (IoT). In the future years, billions of devices, including sensors, actuators, radio-frequency identifiers (RFIDs), mobile phones, and so on, will be linked to the Internet, receiving and delivering data on a regular basis. It should be noted that doing maintenance chores on a large number of terminals would be difficult or impracticable. Because most of these gadgets use batteries as their primary power source, battery replacement is one of the most frequent maintenance procedures. However, in certain circumstances, such as sensors in an industrial setting, sensors implanted in the human body (body area networks), or others embedded in building structures, a battery replacement would be difficult and dangerous. Then, the continual replacement of energy sources becomes a high-cost activity that takes a long time and poses an environmental risk (Singh, Ponnuru & Madhow, 2009). As a result, new technologies that contribute to avoiding these challenges are necessary. Several attempts have been made in the previous several years in terms of EE, and new low-power-consuming networking solutions have been offered. Lora/LoraWAN, SigFox (Björnson, Hoydis & Sanguinetti, 2017), narrowband IoT (NB-IoT) (Lu, Li, Swindlehurst, Ashikhmin & Zhang, 2014), and Zigbee (Sohrabi & Yu, 2015) have all played important roles as IoT industry pioneers. However, EE cannot be seen as a need merely on the end device. Instead, it is a notion that must be implemented throughout the network, including the BSs. Energy harvesting (EH) solutions and hardware advances, in combination with energy-efficient communication approaches, seem to be the key to achieving EE within the 'green communications' paradigm.

0.1 Motivation

Many communication systems, including the fourth generation (4G) cellular system, IEEE 802.11n wireless local area network system (Cisco, 2022; 3rd Generation Partnership Project, 2014), Long-Term Evolution Advanced (LTE-A) (Cisco, 2022; 3rd Generation Partnership Project, 2014), and many more, have demonstrated the benefits of using multiple inputs and multiple outputs (MIMO). It has received widespread praise from communication specialists as a promising core technology that has the potential to be used in a variety of wireless communication systems in the near future (Gupta & Jha, 2015). Large-scale MIMO differs from the more common small-scale MIMO technology. In LTE-A, regular MIMO is typically equipped with eight antennas; however, large-scale MIMO is provided with a huge number of antennas, which might be as many as 128 or even more. This technology, according to a newly proposed method, would allow these antennas in the base station to simultaneously service many user equipment devices (Dai *et al.*, 2015). There are theoretical pieces of evidence that large-scale MIMO systems are capable of achieving high energy efficiency while still achieving orders of magnitude increases in the spectrum (Doppler *et al.*, 2009). The complications caused by the linear detector with a perfect inversion matrix increased in tandem with the increase in the number of users in large-scale MIMO systems, making them prohibitively expensive. There have been a number of studies undertaken that have focused on the Neumann serious expansion (NSE) for approximation purposes in order to overcome the precise matrix inversion (Gu, Liu, Mo & Chi, 2004; Albataineh, 2018; Albataineh & Salem, 2017; Wu, Dick, Cavallaro & Studer, 2016; Ng, Lo & Schober, 2013). However, it has been shown that when the NSE number is greater than 2, the amount of complexity increases significantly once again. To further examine the problem associated with the previously mentioned issue, we propose in this thesis that the matrix inversion-less signal detector technique with a low degree of complexity attached to it might be employed for a large-scale MIMO system in an effort to investigate the problem. The proposed technique is based on the long recurrence expanded conjugate gradient (LRE-CG) method (Wang *et al.*, 2019), which makes it suitable for large-scale MIMO systems due to its low computational complexity. Instead of focusing on identifying new research areas, we believe that establishing an orthonormal basis for Krylov subspace with a big dimension is far more important at this time. In addition to being utilized to update the solution, the full basis is also employed to prevent the occurrence of excessively intricate matrix inversions. The method's convergence rate is also projected to be increased to a more acceptable level as a result of this improvement. The convergence of the suggested signal detection method is also demonstrated in this work, hence ensuring its practicability and viability in the real world. This proposed approach, which has been validated with the help of simulation results, has the ability to efficiently address the matrix inversion issue inside the iterative procedure up to the point where the required accuracy direction is attained. To that end, improved spectrum efficiency, along with improvements in energy efficiency, are among the essential performance indicators (KPIs) for 5G, which are projected to result in an approximately 100-fold increase in spectral efficiency compared to present 4G wireless communications. Toward this end, SWIPT, presented for the first time in (Wu *et al.*, 2016), has gained wide acceptance in the last few years (Gu *et al.*, 2004; Albataineh, 2018; Albataineh & Salem, 2017). SWIPT proposes that the same received RF signals may include both information and energy, and that this may be accomplished using power-splitting receivers in practice. SWIPT is a tool used to increase the battery life of wireless communication devices by harvesting energy from RF signals. This can advance networks such as the Internet of Things, especially in IoT with many wireless devices. Careful consideration of the trade-off between information rate and harvested energy level is necessary when SWIPT is employed in multi-user systems because inter-user interference might negatively impact the ID while supporting the EH. Indeed, initiatives have been put out to address this issue. In addition, in (Prasad, 2004), the transmit power was reduced under the signal-to-interference-plus-noise ratio (SINR) and Quality of service (QoS) requirements for multi-user MIMO systems to minimize interference and noise. A further aspect of interest is the combined transceiver and power-splitting SWIPT down-link design, which also uses the mean squared error (MSE) criteria (Gesbert, Shafi, Shiu, Smith & Naguib, 2003). The combined transceiver and power splitting design was explored to enhance the energy efficiency in multi-cell multi-user down-link SWIPT systems. Even though SWIPT is capable of providing efficient wireless communications, it has only been tested on single-user systems, where future challenges to the joint transceiver and power splitting optimization will emerge.

0.2 Problem Statement

Because of the growing need for high throughput and data rate, 5G cellular networks will require a new spectrum (Uwaechia, Mahyuddin, Ain, Latiff & Za'bah, 2019). Such is the utilization of mm-wave frequencies, which only have a considerable amount of bandwidth at these high frequencies. In recent years, there has been a lot of interest in using 28 GHz (for cellular access and backhaul) as well as 60 GHz (for wireless LAN) (Huang, Lin, Zhang & Zhao, 2019). Because of the substantial free-space pathloss at these frequencies, adaptive antenna arrays with high gain and, therefore, a large number of antenna components will be required. At the base station (BS), hundreds of antenna elements will be used, resulting in a large MIMO operating regime. The signal for each antenna element may be beamformed in the baseband and then upconverted via an RF chain (and similarly at the reception) to get the best performance; in this instance, an information-theoretically optimal system utilizing digital beamforming (Zhu, Wang, Dai & Wang, 2016) might be used. However, using hundreds of RF chains is too costly for mm-wave systems. Furthermore, the high power consumption of the mixer and analog-digital converters (ADC) limits mm-wave system deployment (Choi, Lee & Evans, 2019). As can be observed, any realistic solution will therefore need low-complexity implementations. To solve the hardware limitation of establishing an RF chain with mm-wave bandwidth, analog-only beamformers were first suggested to relocate all needed signal processing to the analog domain, which is typically accomplished by phase shifters. TX and RX are both provided with a single RF chain. Meanwhile, codebook-based beam training is used to discover the optimal TX/RX beam pair without knowledge of CSI. Despite the fact that this kind of system saves hardware complexity, it is confined to serving one data stream on the same frequency-and-time resource block. As a consequence, more flexible, low-complexity designs are required to support the coupling of massive MIMO and mm-wave communications. Aside from the hardware limitation, the performance benefits of massive MIMO systems have been established on the premise of flawless full CSI at the BS, which can be obtained by up-link training (Gao, Dai, Hu, Wang & Wang, 2014) in time division duplexing (TDD) systems but is difficult to accomplish in mm-wave systems. Because most of these gadgets use batteries as their primary power source, battery replacement is one of the most frequent maintenance procedures. However, in certain circumstances, such as sensors in an industrial setting, sensors implanted in the human body (body area networks), or others embedded in building structures, a battery replacement would be difficult and dangerous. Then, the continual replacement of energy sources becomes a high-cost activity that takes a long time and poses an environmental risk (Dai *et al.*, 2014). As a result, new technologies that contribute to avoiding these challenges are necessary. Several attempts have been made in the previous several years in terms of EE, and new low-power-consuming networking solutions have been offered. Lora/LoraWAN, SigFox, narrowband IoT (NB-IoT), and Zigbee have all played important roles as IoT industry pioneers. However, EE cannot be seen as a need merely on the end device. Instead, it is a notion that must be implemented throughout the network, including the BSs. Energy harvesting (EH) solutions and hardware advances, in combination with energy-efficient communication approaches, seem to be the key to achieving EE within the 'green communications' paradigm. In this thesis, we are interested in a new system that can exist by combining the spectrum-efficient mmWave massive MIMO-NOMA systems with energy-efficient SWIPT. This work presents a new way to solve the joint power allocation, power splitting, and joint precoding problem in SWIPT-enabled mmWave MIMO-NOMA systems by incorporating user groupings.

0.3 Research Objectives

This thesis consists of two parts, each with a different focus. Part one of this thesis examines the use of Massive MIMO in uplink as receiver complexity increases. The resulting work should solve the following main issues:

- How to solve the inversion matrix problem numerically at the detection part rather than using the exact MMSE method which is impractical on massive MIMO.
- Search for a iterative numerical solution with the least number of iterations to reduce complexity in the process and to reduce processing time.

We shall address the energy and spectral efficiency issues in massive MIMO downlink in the second section of thesis, thus we will primarily concentrate on these goals:

- To develop a new design for downlink massive mimo systems that trades off spectral and energy efficiency.
- To present new solutions for hybrid beamforming that approaches the performance of a fully digital beamformer in terms of MIMO multiplexing gains.
- Employing the new hybrid approach on other access system rather using OFDMA.
- developing a more adaptable hybrid beamforming system using the strategies that take energy harvesting into account.

0.4 Methodology Overview

In the first stage of our methodology, We employ the long recurrence enlarged conjugate gradient (LRE-CG) method in massive MIMO uplink systems to reduce the data detection complexity and enhance performance. We avoided the complicated matrix invasion by exploiting the LRE-CG method which offers a detection signal with a low-complexity near-optimal signal. The efficiency of the proposed method has been improved by a projection technique that requires a search direction in each sub-domain rather than forming all search directions conjugate to each other. We use a diagonal-approximate initial solution to the presented method. The LRE-CG method has never been applied for signal detection. Therefore, this work provides the first attempt to employ this approach in large-scale MIMO uplink systems. we analyze and identify our problem through a very well study literature review throughout the whole work. This helped us to define the specification and requirements to design the system and the Mathematical model development. A mathematical model to implement SWIPT in Hybrid Precoding with mmWave massive MIMO-NOMA systems to achieve the energy and spectrum efficient wireless communications that were designed. For testing and evaluating the validation of our proposed system, we first use MATLAB to build the model and the channel and rebuild another traditional scheme, then simulate the results and compare these previously proposed models with our proposed system with respect to energy and spectral efficiency. In such a system, joint optimization of the transceiver for information decoding and power splitting for energy harvesting, which includes hybrid precoding design, power allocation, user grouping, and power splitting criteria considerations were implemented.

0.5 Thesis Contribution

The main Contribution of this thesis is to formulate and present a new method based on the long recurrence enlarged conjugate gradient (LRE-CG) algorithm for iteratively realizing the MMSE detection method while avoiding the matrix inversion complexities in massive MIMO systems. We illustrate its superior performance by presenting a comparison with other previous methods. Secondly, we address the problem of energy consumption associated with mixed-signal components such as analog-to-digital components in millimeter-wave (mmWave) massive

MIMO systems. We employ non-orthogonal multiple access (NOMA) in millimeter-wave (mmWave) massive MIMO systems to enhance the spectrum efficiency. The simultaneous wireless information and power transmission technology (SWIPT) will be used in mmWave massive Multiple-Input multiple-Output MIMO systems. The utilization of SWIPT contributes to prolonging the battery life of mobile users (MUs) and enhances the system energy efficiency (EE), especially in the NOMA scenario where the inter-user interference can be reused for energy harvesting (EH). However, we initially designed a user grouping algorithm based on the affinity propagation clustering algorithm, which preferentially groups the user equipment (UE) based on distance and channel correlation. Following that, the analog RF precoder is designed for all beams using the selected user grouping, then a low-dimensional digital baseband precoder is designed to achieve a maximum sum rate as well as minimize inter-beam interference for the system. Afterward, a joint power allocating and power splitting optimization problem is formed. . The considered non-convex optimization problem is arduous to tackle, resulting from the presence of coupled variables and inter-user interference. To cope with this problem, a decoupled approach is adopted, in which the power allocation and power splitting are separated, and the corresponding sub-problems are solved using the Lagrangian duality method. Our contributions can be summarized as follows:

1. The long recurrence enlarged conjugate gradient (LRE-CG) approach is employed and studied in this thesis as a way to iteratively realize the MMMS algorithm while avoiding the complications of matrix inversion. In addition, a diagonal-approximate starting solution to the LRE-CG approach was used to speed up the conversion rate and reduce the complications required. It has been discovered that the LRE-CG-based approach has the ability to significantly reduce computational complexity. By comparing simulation results, it's clear that this new methodology surpasses well-established ways like the Neumann series approximation-based method and the Gauss-Siedel iterative method.

2. We explore hybrid analog/digital precoding and power splitting optimization to create SWIPT-enabled mmWave mMIMO-NOMA systems with hybrid analog-digital recording. To focus on the clustering process, we first propose a new affinity propagation clustering method for user grouping to help with the initial cluster formation process. The parameters for this algorithm include the channel correlation and channel distance values. In this case, we consider the hybrid analog-digital precoder, power allocation, and power slitting factor optimization problem as a sum-rate maximization problem. We seek to maximize the overall power and minimum rate values under the set power and rate restrictions for each UE.
3. We have now set out to build a hybrid mmWave MIMO-NOMA precoding matrix to overcome this challenge. In the first step, the analog precoder is intended to ensure that all beams acquire the maximum equivalent channel gain, depending on the user groupings. Finally, we construct the digital precoding vector for each UE, which prioritizes those users with the most substantial equivalent channel gain per beam to minimize inter-user interference. To simplify our total power and minimum rate restrictions at each UE, we frame the issue as a combined optimization of power allocation and power-splitting factors. The added requirement is that both variables are limited.
4. To optimize the attainable data rate of the system given the restrictions of transmit power and EH need, the combined power allocation and splitting control issue is mathematically modeled. Because of the interrelationship between the linked variables, non-convex and complicated issues emerge.
5. In contrast to (Uwaechia *et al.*, 2019; Alkhateeb, Leus & Heath, 2015), we propose decoupling the joint power allocation and transmit power. Before attempting to optimize the PS ratio assignment with fixed power allocation, we address the subproblem of optimizing the PS ratio assignment with varying power allocation. The Lagrangian duality approach helps solve both sub-problems. Convergence is established when this technique is performed several times.

0.6 Thesis Outline

The remainder of the thesis is arranged as follows: in Chapter 1, we offer some background information regarding wireless communication, as well as the literature-based energy harvested in further depth. In chapter 2, The long recurrence enlarged conjugate gradient (LRE-CG) approach is proposed in this study as a way to iteratively realize the MMMS algorithm while avoiding the complications of matrix inversion. In addition, a diagonal-approximate starting solution to the LRE-CG approach was used to speed up the conversion rate and reduce the complications required. It has been discovered that the LRE-CG-based approach has the ability to significantly reduce computational complexity. By comparing simulation results, it's clear that this new methodology surpasses well-established ways like the Neumann series approximation-based method and the Gauss-Siedel iterative method. With a small number of iterations, the suggested approach achieves the near-optimal performance of a standard MMSE algorithm. In chapter 3, We employ non-orthogonal multiple access (NOMA) in millimeter-wave (mmWave) massive MIMO systems to enhance the spectrum efficiency. The simultaneous wireless information and power transmission technology (SWIPT) will be used in mmWave massive Multiple-Input multiple-Output MIMO systems. The utilization of SWIPT contributes to prolonging the battery life of mobile users (MUs) and enhances the system energy efficiency (EE), especially in the NOMA scenario where the inter-user interference can be reused for energy harvesting (EH). However, we initially designed a user grouping algorithm based on the affinity propagation clustering algorithm, which preferentially groups the user equipment (UE) according to the channel correlation and distance. Following that, an analog RF precoder is designed for all beams using the selected user grouping, followed by a low-dimensional digital baseband precoder to maximize sum rate and minimize inter-beam interference. After that, a joint power allocation and power splitting optimization problem is solved. The considered non-convex optimization problem is arduous to tackle, resulting from the presence of coupled variables and inter-user interference. To cope with this problem, a decoupled approach is adopted, in which the power allocation and power splitting are separated, and the corresponding sub-problems are solved using the Lagrangian duality method. Simulation results confirm the effectiveness of the proposed method and demonstrate that the proposed method is near-optimal and enjoys higher spectrum and energy efficiency compared with state-of-the-art designs and the conventional SWIPT-enabled mmWave MIMO-NOMA system. Finally, we conclude our work and discuss possible research directions.

CHAPTER 1

RELATED WORK AND BACKGROUND INFORMATION

1.1 Introduction

In this chapter, we first explain some background information about the fundamentals of wireless communication with the main feature of the next generations(5G) and we introduce the literature review on the concepts of different beamforming architectures available. Finally, we present a brief explanation of massive MIMO NOMA systems.

1.1.1 mmWave Communications Based on Massive-MIMO

A mmWave frequency range cannot accommodate major wireless technologies used below 6 GHz due to the poor propagation characteristics in the high-frequency spectrum. To model mmWave transmission, the following features need to be considered

- Sparsity: Almost all objects along the propagation path, such as walls, vehicles, etc, contain non-translucent surfaces larger than mmWave communications signals' wavelength, causing mmWave signals to be more sensitive to blockages than microwave signals. Using steerable antennas to identify scattering and reflecting objects or to avoid signal-absorbing objects can also facilitate the transmission of the signal between the transmitter and receiver. An illustration of attenuation caused by atmospheric factors for mmWave is provided in Figure (1.1). Furthermore, only a few scatterers and reflectors will contribute to propagation since many objects will absorb the mmWave signals. As a result, channel impulse responses will usually be sparse at these frequencies.
- Pathloss of Transmission Power: All wireless communication systems attenuate transmitted signals based on their distance. A path loss increases directly in proportion to the inverse of the square of the wavelength as in the equation

$$P_R = \frac{P_T G_T G_R \lambda^2}{(4\pi d)^2} \tag{1.1}$$

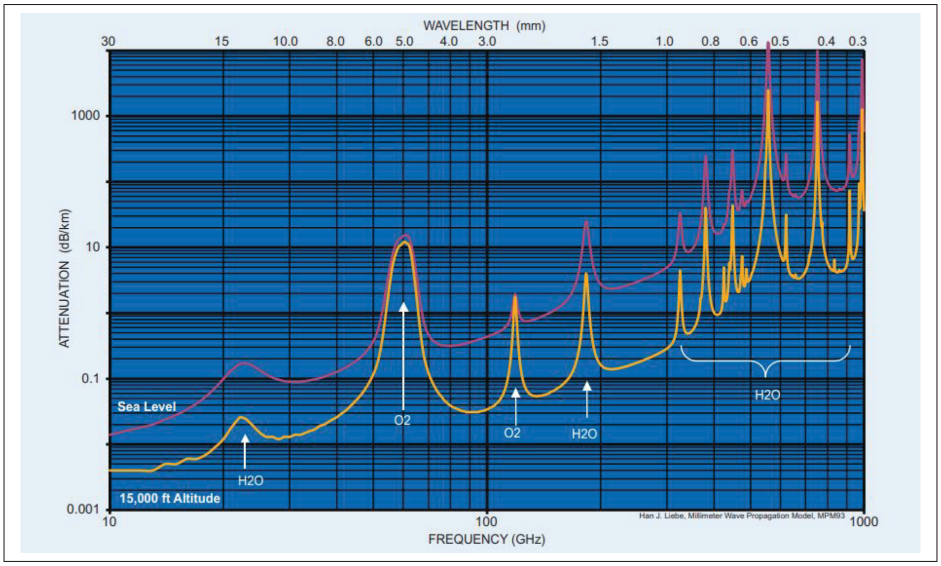


Figure 1.1 Atmospheric absorption of millimeter waves
Taken from Liebe (1989).

P_T and P_R are defined as the transmitter and receiver power, respectively. G_T and G_R are the transmitters and receiver gains. where λ is the wavelength and d represents the distance between the transmitter and receiver. Because of mmWave signals' shorter wavelengths, they experience a higher pathloss than microwave signals at fixed antenna gains. Other conditions are equal, therefore, the received power is much lower than in lower frequency bands. This high path loss can, however, be compensated for with directional transmissions using high gain antennas (Heath, Gonzalez-Prelcic, Rangan, Roh & Sayeed, 2016). An antenna's directivity gain (either as a transmitter or receiver) can be expressed as a function of its effective area A and wavelength, thus G / A. Through the use of antenna arrays, the aperture size, as well as the antenna gain, can be effectively increased. (Rappaport, Heath Jr, Daniels & Murdock, 2015) While mmWave bands present several obstacles, they provide higher data rates (high bandwidths on the order of 20 GHz). Also, mmWaves can be utilized to narrow coverage areas, resulting in denser communication network links that boost capacity by utilizing spatial reuse. Massive-MIMO systems can be implemented using mmWave communications since a larger number of smaller antenna elements can be deployed, resulting in additional performance gains. Communication between multiple users relies heavily on interference suppression, which is primarily accomplished by beamforming at the transmitter and receiver. The MIMO technique is also useful for multiplexing and spatial diversity besides beamforming. Wireless channels can be enhanced by the exploitation of spatial diversity, along with other sources such as time and frequency, to defeat fading. As a result of MIMO, it is possible to transfer parallel data streams without consuming more bandwidth or power, thus increasing the number of spatial dimensions for communication. To fully take advantage of the benefits of operating at mmWave frequencies, several challenges must be addressed, as outlined below.

- Massive-MIMO is an energy-consuming feature for mmWave communications. Higher-resolution directional beamforming is made possible by a higher number of antennas but at the expense of increased complexity and power consumption. As more antennas are added, there is a scaling up of the power consumed by RF chains since each antenna must have its own RF chain complete with mixers, data converters, and power amplifiers. A hybrid A/D processing technique will be essential for mmWave massive-MIMO systems, more details will be provided next section.
- Coherence time in the mmWave is affected by a large doppler spread. Mobile users have a very limited range, which becomes a major problem. Therefore, in mmWave cellular networks, channel estimation, modulation, and coding must take into account the rapidly chaining wireless channel.
- It has always been difficult to achieve theoretical results in RF design limits of wireless communication systems. High carrier frequencies with large bandwidths are particularly problematic when designing RF integrated circuits for mmWave systems, resulting in nonlinear distortions, phase noise, and IQ imbalances.

1.1.2 Technologies Using Massive-MIMO mmWave Communications

Even though mmWave has been used for radar, military, and wireless backhaul technology, indoor communications were the first commercial application. WLANs and wireless personal area networks (WPNs) use mmWave technology to operate in the 60 GHz unlicensed spectrum for the first industrial applications of mmWave technology. mmWave wireless signals are used

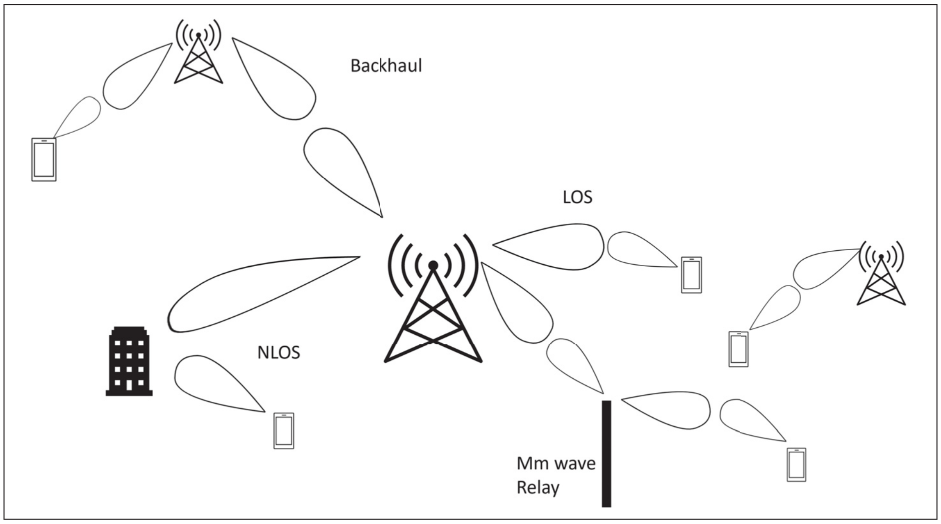


Figure 1.2 mmWave cellular network
Adapted from (Rappaport *et al.*, 2015).

for wireless connections with high bandwidth, replacing wired connections, and WirelessHD can handle high-definition multi-media uncompressed data. The operation of WLANs at 60 GHz can also be enabled by IEEE 802.11ad, a common amendment to IEEE 802.11. There are numerous potential applications of mmWave that are discussed below as a result of these early and successful applications.

- Both research and industry have focused on peer-to-peer networks for many years. In recent years, P2P networks have become increasingly important. Particularly, vehicle-to-vehicle communication, autonomous vehicles, and vehicle-to-infrastructure communication have attracted considerable interest. The benefits of mmWave are numerous, and its use is currently prevalent in automotive radar applications, so we can anticipate that it will have an important role to play in vehicular networks in the near future.
- beyond-Fifth-Generation (B5G) Wireless Communication Systems: mmWave transmission is known to be capable of gigabit data transmission rate in point-to-point backhaul links and, indoor communications but existing systems require high hardware costs to maintain robust long-range connectivity. The mmWave technology is now able to be effectively utilized for outdoor communication networks as well as indoor communication networks due to massive-MIMO and cost-effective mmWave technologies. Cellular networks using mmWave

technology, however, are likely to be built differently than those using microwave technology, as illustrated in Fig. 2.2.

- Personal networks: Low-power, small-scale wireless devices can establish high-speed connections via advanced wireless technology. Accessing smart devices via wireless networks is a direct application of mmWave. This includes mobile phones, watches, tablets, tracking devices, virtual-reality tools, and glasses. These networks are interested in mmWave due to its space limitation, along with its high data rate and low communication latency, which are also important requirements in this case.

1.2 Massive-MIMO mmWave Communications

There will be a thorough review of the literature at the start of each chapter In this chapter. However, this section focuses on mmWave communications, massive-MIMO, and mmWave communications. Massive-MIMO communications frameworks are presented, followed by mmWave communications frameworks.

1.2.1 mmWave Communications

It was in indoor communications that mmWave was initially applied to enable multigigabit wireless links; commercialized systems even implemented this technology. In view of the densely crowded microwave spectrum not being able to meet the demands of 5G, mmWave bands between 30 and 100 GHz provide a large bandwidth of over 20 GHz for the next generation of cellular wireless systems (Rappaport *et al.*, 2015). The first step to mmWave communications is to understand propagation effects and channel modeling in depth. There have already been studies of mmWave propagation for indoor communications. Nonetheless, models of the mmWave MIMO channel have been updated in recent years and many articles have been presented in this area (Rappaport *et al.*, 2015; Zhang, Xin & Liang, 2009). mmWave communications have a major drawback for outdoor use, which is the significant Pathloss, is caused by a combination of penetration loss, atmospheric absorption, rain attenuation, and free-space pathloss. Pathloss can be effectively defeated by beamforming by using a larger number of antennas at both transmitter and receiver. A massive MIMO system is necessary to enable outdoor mmWave communications. Additionally, due to form factor limitations, massive-MIMO can only be implemented in mmWave bands. For future 5G wireless networks, integrating these techniques will be a fortunate opportunity because of their compatibility.

1.2.2 Massive-MIMO

Massive-MIMO has been developed conceptually and theoretically based on the advanced development of MIMO technologies in cellular systems and wireless local areas. But massive-MIMO did not become a realistic possibility until more than a decade had passed (Rusek *et al.*, 2012; Larsson, Edfors, Tufvesson & Marzetta, 2014). Numerous research studies have been conducted on massive-MIMO systems, but many questions remain (Björnson *et al.*, 2017). Despite massive-MIMO being established as a crucial technology for 5G, various factors, challenges, and possibilities are still being investigated. In this thesis, hybrid architectures are used to implement massive-MIMO systems.

1.2.3 Hybrid Analogue-Digital Signal System

A MIMO system can use precoding if it knows the radio channel information. Because of mmWave transmission's severe pathloss effects, precoding is particularly necessary for massive-MIMO. It is inefficient to implement massive-MIMO using conventional FD because such a system requires each antenna to be driven with its own RF chain, thereby increasing energy consumption and system costcitepel2014spatially. Massive-MIMO systems were simplified and made more energy-efficient with hybrid analog/digital systems. The number of required RF chains can be reduced by cascading an analog (RF) precoder following the baseband digital precoder. In (Zhang, Molisch & Kung, 2005), the first attempt was made to implement a single symbol stream of FDP with hybrid A/D precoding with two RF chains. In (Sohrabi & Yu, 2015) the number of symbols per transmission is required to be the same as the number of RF chains in multi-stream data hybrid designs. A constant modulus constraint attached to phase-shifters is required in the entries of the RF precoder matrix resulting in an extremely complex structure in the HBF, For this reason, the resulting precoder optimization is non-convex. The use of heuristic iterative algorithms or reconstruction algorithms has been directly used in many publications to design HBF citepel2014spatially. A further discussion of practical power constraints is provided in (Garcia-Rodriguez, Venkateswaran, Rulikowski & Masouros, 2016).

1.2.4 MIMO-OFDM Hybrid Beamforming

Several wireless systems, including 4G and WiFi, utilize orthogonal frequency division multiplexing (OFDM). In next-generation wireless systems, OFDM will remain the dominant multi-carrier (MC) technique because it is a simple and flexible system. However, massive-MIMO systems have posed even greater difficulties than narrow-band systems for implementing OFDM. In massive MIMO-OFDM, hybrid architecture reduces the number of RF chains but complicates the design of RF signal processing networks. A hybrid scheme must exhibit an equal response to all subcarriers, so designing it is more difficult than creating a narrow-band system. It has been proposed to design wide-band hybrid precoders (Alkhateeb *et al.*, 2015; Chi, Peng, Huang, Tsai & Ma, 2008); however, these studies are still in their infancy and there are many questions left to be addressed (Sohrabi & Yu, 2015; Yu, Shen, Zhang & Letaief, 2016; Zhu, Zhang & Yang, 2017).

1.2.5 mmWave Massive-MIMO Channel Estimation

Earlier, we discussed the development of modeling hybrid beamforming for narrow- and wide-band communication. The underlying channel parameters must be accurately estimated before any of these techniques can be applied. As a consequence, when massive-MIMO implementation uses the hybrid architecture, the conventional FD channel estimation approaches are unable to be effectively applied; therefore, it is necessary to use the HBF-based channel estimation approach. (Alkhateeb *et al.*, 2015) presented one of the first algorithms utilizing channel estimation for the narrow-band scenarios; this method was then applied to the wideband transmission. Numerous studies have been conducted on the channel parameters estimation for massive mmWave MIMO-OFDM systems (Lin *et al.*, 2017; Zhang *et al.*, 2009; Gao *et al.*, 2014).

It is evident from all of these studies that the sparsity of mmWave channels is used in HSP-based channel estimators; despite this, alternative approaches could still be examined and investigated.

1.3 Multiple-Access Techniques

The communication system needs to be capable of serving multiple users at the same time. Thus, by effectively assigning time/code/frequency resource allocations, more than one user is served in a single resource unit. The assignment of a dedicated resource block to each user is commonly referred to as orthogonal multiple access (OMA). OMA techniques commonly used in wireless communication include the following.

- Frequency-division multiple access (FDMA): Each user is given a distinct sub-band in FDMA, which divides the system's overall bandwidth into several of orthogonal channels (Goldsmith, 2005). To prevent any interference, these channels typically have guard bands between them.
- Time-division multiple access (TDMA): With the TDMA, numerous users time-share access to the same frequency channel (Goldsmith, 2005). During their designated time slot, the users communicate sequentially.
- Code-division multiple access (CDMA): CDMA modulates the signals of each user using different unique codes, which spread the signals into a larger bandwidth than necessary to handle the data rate, and all transmitted signals share the same bandwidth (Goldsmith, 2005). The receiver decodes the signals to split up the different users.
- Orthogonal frequency-division multiple access (OFDMA): The OFDMA is a multi-user adaptation of orthogonal frequency division multiplexing (OFDM). In both OFDM and OFDMA, the system bandwidth is divided into multiple frequency allocations, so they will not interfere with each other despite the absence of the guard bands between them. In parallel, these subcarriers can be modulated separately. In this way, multiple access is achieved by allocating subcarriers based and channel quality to multiple users (Morelli, Kuo & Pun, 2007).

Throughout the development of wireless technologies, beginning with the first generation and continuing through the fourth, OMA techniques have many principal advantages such as Improving network capacity, reducing the overall cost, and reducing the consumption of valuable resources for communication. Nevertheless, these conventional OMA technologies have reached their limits in terms of performance enhancements, and cannot fulfill the requirements of future wireless networks. As opposed to orthogonal multiple-access (OMA), non-orthogonal multiple-access (NOMA) enables multiple users to access one frequency, time, and code, as well as improving spectral efficiency, massive connectivity, and latency. Therefore, in our work, we focus on the performance of massive MIMO using NOMA.

1.3.1 Non-Orthogonal Multiple Access (NOMA)

The NOMA technique utilizes the transmission power to create an efficient multiplexing method as seen in Figure(1.3) (Vaezi, Schober, Ding & Poor, 2019). Multiple access technologies are now extended beyond the conventional time/code/frequency domain.

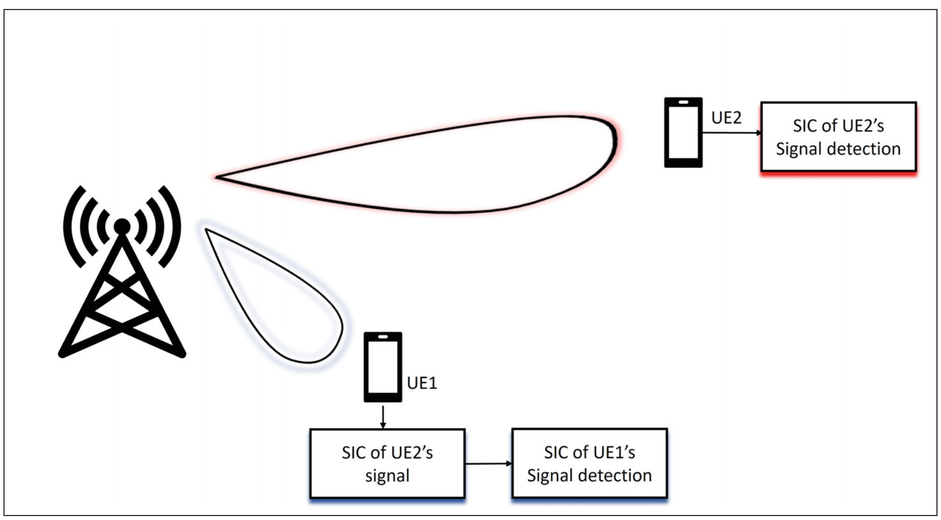


Figure 1.3 NOMA power domain multiplexing for two users
Adapted from Vaezi *et al.* (2019)

So, NOMA is an approach that can integrate with the massive MIMO system. SC and SIC are the two main techniques behind NOMA (Vanka *et al.*, 2012). It is not entirely new that these technologies have been investigated, both theoretically and practically, for several decades (Vanka *et al.*, 2012). As an example, SIC has been incorporated into many systems, including IEEE 802.15.4, multi-user MIMO (Gelal *et al.*, 2012), V-BLAST (Wolniansky, Foschini, Golden & Valenzuela, 1998), and CDMA (Patel & Holtzman, 1994). Thus, NOMA can be seamlessly integrated into the future of wireless networks. It is possible to transmit information from multiple users at the same time using SC. Specifically, SC allows encoding signals with a weak channel at a lower rate and superimposing them on signals with a stronger channel. In this situation, the transmitter must be able to order users based on channel conditions. As part of a traditional OMA scheme, stronger signals are usually allocated higher transmit power. In contrast, NOMA assigns more power to users with a weak channel to ensure that they can decode their intended messages by treating other users' messages as noise (Vaezi *et al.*, 2019; Senel, Cheng, Björnson & Larsson, 2019). The users with strong channel conditions apply the SIC technique to subtract the interference from the users with poor channel conditions. At the receiver, the SIC technique is used to achieve interference cancellation. A stronger channel user decodes the message intended for a weaker channel user first. In the process of generating the received signal, the stronger user subtracts the interference from the weaker user's signal. After that, the stronger user decodes its message without being affected by the weaker user's interference. Also, the receiver end can be designed with low complexity hardware (Andrews, 2005).

1.4 Simultaneous Wireless Information and Power Transfer (SWIPT)

Systems that are spectrum-efficient often have high energy consumption and are therefore inefficient in terms of energy use. In wireless research, energy-efficient design has been examined for a variety of scenarios, which is typically defined as the amount of information capable of being reliably transmitted per unit of energy consumed (Mahapatra, Nijsure, Kaddoum, Hassan & Yuen, 2015). There has begun a shift from spectrum-efficiency optimization to energy-efficiency optimization in communication, where the aim isn't just to maximize data rates, but also to maximize the amount of information that can be reliably transmitted per unit of energy consumed (Li *et al.*, 2011). Several advanced algorithms have been suggested to achieve high energy efficiency in future networks since energy efficiency has become as a primary concern. Although we are still in the early stages of research, there are many open and challenging issues must be addressed. Further improving energy efficiency is possible with advanced techniques such as SWIPT and wireless powered communications. Therefore, new approaches to energy-efficient applications are needed to accommodate different operating conditions. Wireless transceivers and sensors, such as wireless medical sensors inside the human body, and unsupervised low-power devices in the internet of things are increasingly becoming power-limited and reduced in size due to the rapid advancement of semiconductor technology. A new approach is needed, where wireless devices can harvest energy by gathering radio frequency signals from the environment rather than using an electric grid and conventional energy sources. This will reduce expenses and the trouble of replacing or recharging these device batteries in these situations. With SWIPT (Zhang & Ho, 2013), wireless information and power transfer can be combined within a variety of wireless communication systems. The operation of energy harvesting on a radio frequency signal will destroy the information content of the signal, so receivers cannot simultaneously perform energy harvesting (EH) and information decoding operations. For SWIPT schemes to be implemented properly, the receiver setup must be carefully designed. For practical implementation of SWIPT, several specific receiver structures have been proposed, such as time-switching receivers Perera, Jayakody, Sharma, Chatzinotas & Li (2017), separated receivers (Zhang & Ho, 2013), antenna selecting receivers (Shi, Liu, Xu & Zhang, 2014), and power-splitting receivers (Ding *et al.*, 2015). Moreover, the minimum required power at an information receiver is incredibly higher than the power at an energy harvesting receiver. To satisfy this different requirement, modern signal processing methods and massive multiple antenna with beamforming filters (Perera *et al.*, 2017) are emerging and gaining popularity in SWIPT enabled systems. The efficiency of wireless information and power transfer can be improved by well-designed beamforming systems. Maintaining energy efficiency can be achieved by integrating SWIPT with wireless communication systems. The development of SWIPT systems will face a variety of limitations and challenges in the future (Shi *et al.*, 2014). It is important to conduct performance analysis and test the results in various energy harvesters, especially nonlinear models of imperfect energy harvesters (Clerckx *et al.*, 2018). Many existing works have applied the linear EH model due to its ease of solving optimization problems. Wireless power conversion via EH circuits, however, occurs in practice in a nonlinear manner (Lu, Xiong, Fan, Zhong & Letaief, 2018). Consequently, the linear EH model cannot capture the power-dependent EH efficiency, resulting in a mismatch with practical systems and potentially leading to invalid optimizations. To adapt the linear EH model for practical systems, many schemes and solutions will have to be reevaluated.

CHAPTER 2

DATA DETECTION METHOD FOR UPLINK MASSIVE MIMO SYSTEMS BASED ON THE LONG RECURRENCE ENLARGED CONJUGATE GRADIENT

2.1 Introduction

Many communication systems, including the fourth generation (4G) cellular system, IEEE 802.11n wireless local area network system, Long-Term Evolution Advanced (LTE-A), and many more, have demonstrated the benefits of using multiple inputs and multiple outputs (MIMO) (Larsson *et al.*, 2014). It has received widespread praise from communication specialists as a promising core technology that has the potential to be used in a variety of wireless communication systems in the near future (Andrews *et al.*, 2014). Large-scale MIMO differs from the more common small-scale MIMO technology. In LTE-A, regular MIMO is typically equipped with eight antennas; however, large-scale MIMO is provided with a huge number of antennas, which might be as many as 128 or even more. This technology, according to a newly proposed method, would allow these antennas in the base station to simultaneously service many user equipment devices (Rusek *et al.*, 2012). There is theoretical evidence that large-scale MIMO systems are capable of achieving high energy efficiency while still achieving orders of magnitude increases in the spectrum (Lyu & Ling, 2018). In the course of evaluating the practical advantages of large-scale MIMO, various difficulties have been observed. For example, increasing the performance of the practical signal detection algorithm in the uplink to accommodate multiuser interferences. Growth in the number of transmit antennas has been shown to cause a fast increase in the complexity of ideal maximum likelihood (ML) detectors (Alwakeel & Mehana, 2019). As a result, it becomes impracticable for large-scale MIMO systems, and their relevance diminishes as a result of this. In order to achieve near-optimal performance while reducing the degree of complexities, nonlinear detection algorithms, such as fixed-complexity sphere decoding (Albataineh, 2019) and tabular search (Albataineh, 2021) are proposed. However, this low degree of complexity continues to be a concern when the MIMO system is vast in size or when the modulation order is high (Chen, 2017b) (for example, when there are 128 antennas at the base station and 64-quadrature amplitude modulation) (QAM). Every UE in the coverage area is serviced by every AP in the communication range in the traditional cell-free massive MIMO topology (Hanchate & Nema, 2018; Yin, Wu, Cavallaro & Studer, 2014; Gao *et al.*, 2014). In (Yin, Wu, Studer, Cavallaro & Dick, 2013), a typology is proposed in respect of uplink receiver coordination across APs with CPU, ranging from entirely dispersed (Level 2) to totally centralized (Level 4) implementations, with the highest level of cooperation being the most cooperative. Although scalability issues for channel estimation, data decoding/precoding, and fronthaul signaling have been highlighted in recent work (Dai *et al.*, 2014), it is imperative that these issues be overcome in order to enable large-scale deployments of cell-free networks on a global scale. It has encouraged researchers to propose user-centric ways to selectively service a subset of APs in wide coverage areas, due to the fact that the majority of APs in a broad coverage area have insignificant channel gains to one or more specific UEs (Grigori, Moufawad & Nataf, 2016). In order to deal with the complexities while maintaining high performance, a low-complexity linear detection algorithm such as zero-forcing and minimum mean square error (MMSE) could potentially be used for up-linking the multiuser large-scale MIMO systems (Alabed, 2018). MMSE is a linear detection algorithm with near-optimal performance for up-linking the multiuser large-scale MIMO systems. This approach, on the other hand, employs a matrix inversion that is both difficult and unfavorable in nature. For translating matrix inversion into matrix-vector multiplication series (Hu, Wang, Gaol & Ning, 2014), the Neumann series approximation approach was recently introduced. Although this algorithm has the potential to reduce complexity, the reduction is not very substantial. The complications caused by the linear detector with a perfect inversion matrix increased in tandem with the increase in the number of users in large-scale MIMO systems, making them prohibitively expensive. There have been a number of studies undertaken that have focused on the Neumann serious expansion (NSE) for approximation purposes in order to overcome the precise matrix inversion (Gu *et al.*, 2004; Alnabelsi, Salameh & Albataineh, 2020; Albataineh, 2018; Albataineh & Salem, 2017). However, it has been shown that when the NSE number is greater than 2, the amount of complexity increases significantly once again. There have been other iterative linear algorithms suggested recently to achieve a better balance between performance and complexity, including the conjugate (CG) method (Prasad, 2004; Gesbert *et al.*, 2003). The Gauss-Seidel (GS) algorithm (Belschner, Rakocevic & Habermann, 2019), and the successive over-relaxation (SOR) algorithm (Chen, 2019) among others. In order to get better MIMO detection with less complexity, these techniques are believed to be beneficial. In (Raviteja, Hong & Viterbo, 2017), they propose to include the dynamic cooperative grouping methodology from the connectivity MIMO research into cell-free massive MIMO. There may be an overlap between the AP groupings that service various UEs, and the groups are chosen based on the demands of the users. The authors of (Jiang, Li, Gong & Su, 2018) take the position that the dynamic cooperative grouping may be used with both centralized (Level 4) and completely dispersed (Level 2) uplink implementations in the same network. However, with DCC, the Level 3 implementation (based on the taxonomy in (Lee, 2017) has not been addressed because it is not required. When the CPU reaches Level 3, it adds a second layer of decoding, known as largescale fading decoding (LSFD), in order to reduce interference. When compared to Level 2 in the original cell-free massive MIMO, this distributed processing technique has been demonstrated to significantly enhance SE. However, this method has not been investigated in user-centric networks. The best SE performance among the levels is achieved by using Level 4, but this requires the computation of centralized receive combiners at the CPU, which has significantly higher dimensions when contrasted to Level 3 and Level 2 local beamforming and thus increases the complexity of the algorithm of the level. To further examine the problem associated with the previously mentioned issue, we suggest in this work that the matrix inversion-less signal detector technique with a low degree of complexity attached to it might be employed for a large-scale MIMO system in an effort to investigate the problem. The suggested technique is based on the long recurrence expanded conjugate gradient (LRE-CG) method (Wang *et al.*, 2019), which makes it suitable for large-scale MIMO systems due to its low computational complexity. Instead of focusing on identifying new research areas, we believe that establishing an orthonormal basis for the Krylov subspace with a big dimension is far more important at this time. In addition to being utilized to update the solution, the full basis is also employed to prevent the occurrence of excessively intricate matrix inversions. The method's convergence rate is also projected to be increased to a more acceptable level as a result of this improvement. The convergence of the suggested signal detection method is also demonstrated in this work, hence ensuring its practicability and viability in the real world. This paper's approach, which has been validated with the help of simulation results, has the ability to efficiently address the matrix inversion issue inside the iterative procedure up to the point where the required accuracy direction is attained. According to a survey of current literature and research effort relevant to this subject matter, our work represents the first and only attempt to employ the LRE-CG approach for the process of signal detection in an uplink large-scale MIMO system that has been made. This chapter has been divided into sections to help readers to have a comprehensive and clear grasp of the problem that has been recognized and the solution that has been suggested in the study. Section 2.2 of this chapter offers a brief overview of the system modeling methodology. It has been attempted in Section 2.3 to define the suggested low complexity signal detection method, as well as the process of its convergence and a study of the complexity associated with it. Section 2.4 presents the findings of the bit error (BER) stimulation of the performance of our suggested system performance. Section 2.5 concludes with a synopsis of the complete piece of work.

2.2 System Model

For the system model, first, we will consider an uplink large-scale MIMO system where N antennas are employed at the base system and K selected single antenna UE devices are simultaneously served for communicating. The $N >> K$ assumption is made in this case, e.g., $N = 128$ and $K = 16$ (Prasad, 2004). In the parallel transmitted bit stream, K different users' signals are encoded separately at first. In order to map it to the constellation system, the channel encoder encode the data first. In order to conduct the mapping, values are extracted from the energy-normalized modulation constellation Q. s in this model represents the $K \times 1$ transmitted signal vector which includes the transmissions from all the K users and $H \in C^{N \times K}$ is used to denote the flat Rayleigh fading channel matrix with zero mean and unit variance in which all the entries are considered to be independent as well as identically disturbed. The signal vector y in the $N \times 1$ receiver can be expressed as:

$$y = Hs + n \tag{2.1}$$

In this equation, n is a $N \times 1$ additive white Gaussian noise vector whose entries follow$CN\left(0,\ \sigma^2\right)$. Multiuser signal detection work has been performed at the base station BS in order to get the estimated transmitted signal vector s from the noisy signals vector y received. It is important to note here that the channel matrix H is usually obtainable through time domain and frequency domain training pilots (Jiang *et al.*, 2018; Lee, 2017). Now, the estimated transmitted signal vector $\hat{s}$ that is obtained by the MMSE linear detection method can be expressed as

$$\hat{s} = \left(H^H H - \sigma^2 I_K\right)^{-1} H^H y = W^{-1} y_{MF} \tag{2.2}$$

Here the $y_{MF} = H^H y$ is the matched-filter output of y, and the MMSE filtering matrix W is denoted by

$$W = G - \sigma^2 I_K \tag{2.3}$$

where $G = H^H H$ represents the Gram matrix. Using the estimated results for soft-input channel decoding, the LLRs (log-likelihood ratios) of the transmitted signal vector can be derived. The assumption at this point is that the equivalent channel matrix is $E = W^{-1}G$ and $U = W^{-1}H^H\left(W^{-1}H^H\right)^H = W^{-1}GW^{-1}$. Therefore, with (2.1) and (2.2) combined, the MMSE estimate $\hat{s}$ is as follows

$$\hat{s} = Es + W^{-1}H^H n \tag{2.4}$$

According to the following equation, we may predict that the sent symbol for the k-th user will be as follows:

$$\widehat{s_k} = \mu_k s_k + \delta_k \tag{2.5}$$

where s_k denotes the symbol employed to represent the k-th element of the vector of the transmitting signal s. $\mu_k = E_{kk}$ denotes the channel gain, and $\delta_k^2 = \sum_{m \neq k}^{K} |E_{mk}|^2 + U_{kk}\sigma^2$ represents the noise plus interference (NPI) variance; U_{kk} denotes the one component of matrix U in the k-th row and k-th column and E_{mk} denote the one component of matrix E in the m-th row and k-th column. However, the LLR $\mathcal{L}_{k,b}$ of the k-th user can be expressed as follows

(Albataineh, 2018)

$$\mathcal{L}_{k,b} = \hbar_k \left(\min_{\tau \in S_b^0} \left| \frac{\widehat{s_k}}{\mu_k} - \tau \right|^2 - \min_{\bar{\tau} \in S_b^1} \left| \frac{\widehat{s_k}}{\mu_k} - \bar{\tau} \right|^2 \right) \tag{2.6}$$

where $\hbar_k = \frac{\mu_k^2}{\delta_k^2}$ denotes the signal to interference plus noise ratio for the k-th user, and S_b^0, S_b^1 represents the group that consists of the signs of the constellation Q. From the above, it is clearly demonstrated that the MMSE linear detection algorithm is almost optimal for the process of uplinking the multiuser large-scale MIMO systems. However, it has also been verified that it is not possible to avoid the sophisticated matrix inversion W^{-1} included in the MMSE algorithm. To calculate the final LLRs for soft-input channel decoding, MMSE estimates, channel gain, and noise plus interference (NPI) variance are essentially needed where they can be computed by using the matrix inversion. The complexity computing of matrix inversion is $O(K^3)$ which is considered high because K is typically very large in the uplink large-scale MIMO system (Albataineh, 2018). Here, it has been proposed to use a low-complexity signal detection technique in which the iterative LRE-CG algorithm is been employed to estimate the MMSE without the need for matrix inversion. Adding a diagonal approximate initial solution (Albataineh & Salem, 2017) to the LRE-CG method, we have used it to enhance the convergence and reduce the degree of complexity. Alongside, we also propose for estimating the channel gain and NPI variance for LLR computation by employing an approximated method that is not required to compute the exact matrix inversion. To sum up, the overall analysis of the proposed LRE-GC algorithm has been presented to demonstrate that there are certain advantages of this algorithm over other typical and conventional sophisticated methods found in the literature.

2.3 The proposed Method

For an uplink large-scale MIMO system, the channel matrix H is an asymptotically orthogonal column full-rank matrix according to the suggested technique. It guarantees the Hermitian positive definiteness of the MMSE filtering matrix W. The LRE-CG technique (Dai *et al.*, 2014) may be used to iteratively solve (2.2) in the absence of matrix inversion because of its particular characteristic. N-dimensional linear equation $Ax = b$ has been solved using the

LRE-CG technique, while, A represents the N-dimensional Hermitian positive definite matrix, x denotes the N-dimensional solution vector, and b represents the N-dimensional measurement vector. With the LRE-CG approach, which differs from the usual method in that it does not use a computer at all to solve the equation of $A^{-1}b = x$ repeatedly, the complexity of solving $Ax = b$ is kept to an absolute minimum. W is a Hermitian positive definite matrix, hence we may decompose it as a Hermitian positive definite matrix.

$$W = D + L + L^H \tag{2.7}$$

Matrix W's diagonal and lower triangular halves are referred to as D and L, respectively. LRE-CG technique is used to estimate the transmitted signal vector s once this step has been completed. Krylov projection technique of the LRE-CG is used to solve a linear system of equations (Dai *et al.*, 2014). Without the matrix inversion, the LRE-CG approach may be able to solve the issue of (2.3) by addressing the following optimization problem.

$$\hat{s} = arg \min_{\hat{s} \in C^U} \|H^H b - A\hat{s}\| \tag{2.8}$$

where $A = H^H H + N_0 I_U \in C^{U \times U}$ denotes a positive definite matrix, which represents the regularized uplink Gram matrix. The method in (Dai *et al.*, 2014) may be used to iteratively compute the solution, utilizing the LRE-CG technique with minimal computational cost. As an alternative, LRE-CG may be used to determine the transmitted signal vector s at the i-th iteration as follows.

$$\hat{s}_i = \hat{s}_{i-1} + P_i \alpha_i \tag{2.9}$$

where P_i represents the $U \times t$ matrix that consists of the t sub-domain search-directions, and α_i represents the vector of size t. Our LRE-CG-based technique for soft output data identification is summarized in Algorithm 1. Our LRE-CG approach is based on the algorithm in (Dai *et al.*, 2014). Even with an infinite number of repeats, the suggested algorithm reduces the complexity of the MMSE technique from $O(K3)$ to $O(K2)$ in algorithm 2.1.

Algorithm 2.1 LRE-CG for soft-output MMSE detection

Input:	**Number of Flops**
$\boldsymbol{A}$, the $\boldsymbol{n \times n}$ symmetric positive definite matrix	
$\boldsymbol{b}$, the $\boldsymbol{n \times 1}$ observed vector	
$\boldsymbol{x_0}$, the initial guess	
$\boldsymbol{\epsilon}$, the stopping tolerance	
$\boldsymbol{t}$, Number of the sub-domains(search direction)	
$\boldsymbol{Itr_{max}}$, the maximum allowed iterations	
Output:	
$\boldsymbol{x_{itr}}$, the approximate solution	
$\boldsymbol{r = b - Ax_0, Itr = 1}$	$\boldsymbol{2nnz - 1}$
$\boldsymbol{W = \mathcal{T}(r_0), Q = W}$	$\boldsymbol{2nnz + n(t-1)}$
A-orthonormalize $\boldsymbol{P_1}$	
while $\boldsymbol{(Itr < itr_{max})}$ **do**	
$\boldsymbol{G = G^{-1}(Q^t r)}$	$\boldsymbol{(2nnz - 1)t + (2n-1)t^2}$
$\boldsymbol{\alpha = Q^t AQ}$	$\boldsymbol{(2n-1)n}$
$\boldsymbol{x_{itr} = x_{itr-1} + Q\alpha}$	$\boldsymbol{2nt}$
$\boldsymbol{r = r - AQ\alpha}$	$\boldsymbol{2nt}$
$\boldsymbol{W = AW}$	$\boldsymbol{(2nnz - n)t}$
A-orthonormalize $\boldsymbol{W}$ using modified Gram Shmidt	$\boldsymbol{nnzt^2 + nt^2}$
$\boldsymbol{Q = QW}$	$\boldsymbol{(2nnz - n)t}$
$\boldsymbol{Itr = Itr + 1}$	$\boldsymbol{1}$
end while	

2.3.1 Computational Complexity

According to Algorithm 2.1, the computational complexity is assessed as follows. The calculated number of multiplications for each step in the proposed approach is used to calculate the final result. At each iteration, the total number of multiplications is given by:

$$O = 4Ut^2 + 8Ut + 2U \tag{2.10}$$

where t denotes the number of search directions. Since the presented method aims to decrease the computational cost, Itr_{max} should be made considerably less than U so that the presented algorithm's computational complexity is less than $O(U^3)$. Also, we investigate computational complexity as it pertains to various approaches found in the literature for comparison in the next section.

2.4 Results and Discussion

Monte Carlo simulations have been carried out in a coded 20-MHz MIMO-OFDM uplink system with 2048 subcarriers in order to evaluate the performance of the proposed technique in terms of error-rate performance. 1200 of these are used for data transmissions, such as in LTE Advanced (LTE-A) (Prasad, 2004) and other networks. The 64-QAM modulation scheme is used in conjunction with Gray mapping and a rate-3/4 turbo code. The channel matrices were also created in order to get the spatial and frequency correlation, for which we utilized the WINNER-Phase-2 model (Gesbert *et al.*, 2003) with 7.8 cm antenna spacing, similar to the models used in (Yin *et al.*, 2014; Alnabelsi *et al.*, 2020). It has been decided to use a log-MAP turbo encoder for the purpose of decoding the channel. A bit error rate is also supplied, which is calculated by coding over one OFDM signal with 1200 data subcarriers and calculating the bit error rate. In this regard, we concentrate on a number of massive MIMO detection systems. The experiments were conducted due to the MATLAB program on an Intel Core i7 CPU with a 2.4-GHz processor and 4G MB RAM, as well as a MATLAB environment. Figure 2.1 shows a comparison of the bit error rate (BER) for the presented method in the study, as well as for other precise and approximate data-detection algorithms utilized for huge MU-MIMO systems with various antenna configurations, as shown in the paper. We have specifically acquired the BER findings for the Neumann series detection (Albataineh, 2019), the CG-based detection (Hanchate & Nema, 2018), and the Gauss-Seidel (GS)-based detection (Yin *et al.*, 2013) techniques. In addition to this, we have supplied a reference equalization that is an exact linear MMSE equalizer as well. Three rounds of BER versus SNR of the described techniques are shown in figure 2.1 with simulation results for each iteration. It is set up with the following parameters: N=128 antennas, U=8 users, and SNR values ranging from -10 to 20 decibels (decibels per kilometer). Figure 2.1 illustrates that the suggested technique, which is based on the LRE-CG method, is capable of approaching the performance of the MMSE algorithm while consistently delivering the lowest BER when compared to other algorithms described in the literature. Also included is a comparison of the average CPU timings for the various techniques, which is presented in Table 2.1. As demonstrated in Table 2.1, the suggested approach is comparable to the other algorithms, and it even outperforms the other methods when it comes to BER. Furthermore, it is undeniable that the CG technique and the Richardson method are less difficult algorithms than the other algorithms available. However, with the introduction of increasingly powerful computer systems, such as Graphics Processing Units (GPUs), the accuracy of performance measurements has gained importance.

Table 2.1 Average computational times for each method (in sec) for $U \times N = 8 \times 32$case

	MMSE	**CG**	**SOR**	**Neur**	**Richard**	**GS**	**LRE-CG**
$\boldsymbol{Itr = 3}$	1.03 e-04	8.15 e-05	1.31 e-04	1.26 e-04	5.41e-05	1.31 e-04	5.53e-05

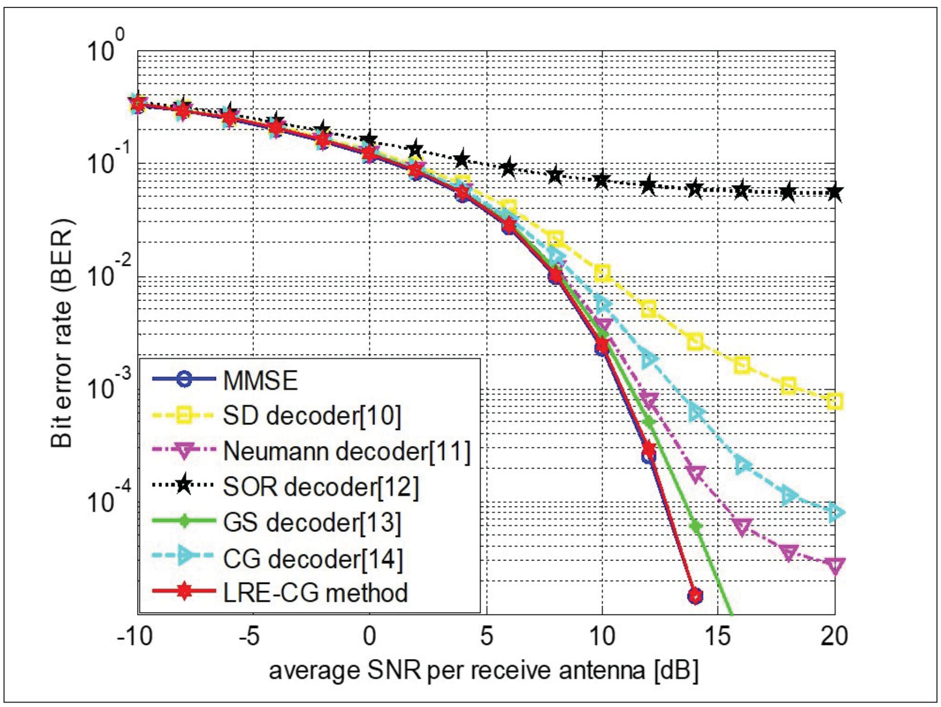


Figure 2.1 The BER compared the proposed estimated technique and alternative ways to calculate LLRs for N=128 antennas with 8 users and different SNRs

Figures 2.2 and 2.3 show the performance of the suggested algorithm, which is based on the LRE-CG technique and the other methods discussed above, when N=128 and U=16, respectively. As seen in Figures 2.2 and 2.3, the suggested method comes close to the performance of the MMSE algorithm while outperforming other algorithms that have been reported in the literature. Table 2.2 also includes a comparison of typical computation times, which illustrates the difference between the two approaches. As demonstrated in Table 2.2, the suggested technique is comparable to the other methods, and it even outperforms them in terms of BER performance.

Table 2.2 Average computational times for each method (in sec) for $U \times N = 16 \times 128$ case

	MMSE	**CG**	**SOR**	**Neur**	**Richard**	**GS**	**LRE-CG**
$\boldsymbol{Itr = 3}$	1.42e-04	9.32 e-05	2.35e-04	3.25e-04	1.05e-05	2.27 e-04	1.11e-04
$\boldsymbol{Itr = 4}$	1.44 e-04	1.23 e-04	2.84e-04	4.12e-04	1.27e-04	2.83e-04	1.26e-04

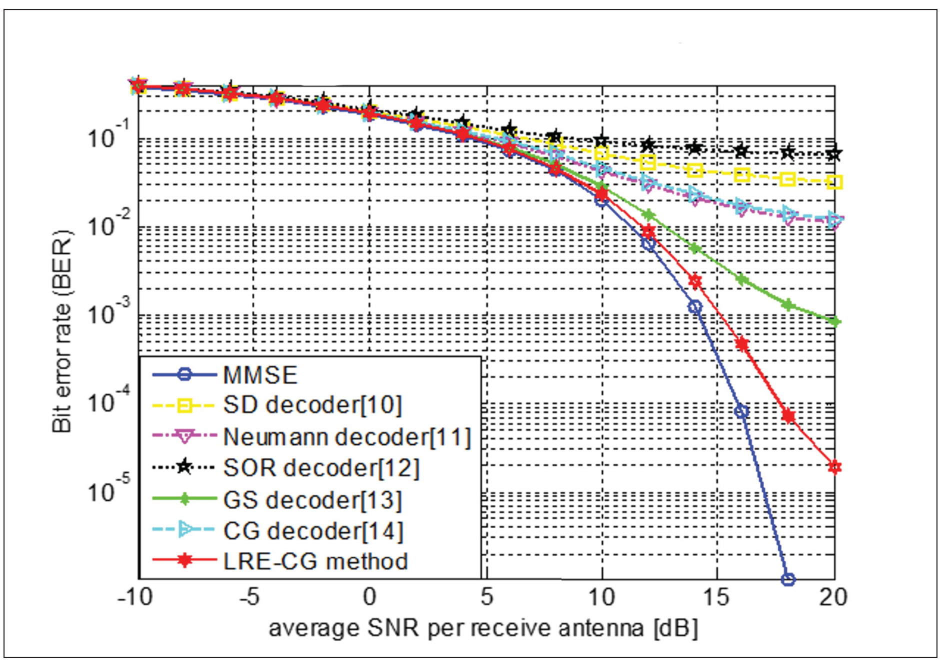


Figure 2.2 The evaluation of proposed approximation technique's BER to alternative approaches for calculating LLRs for a system with 128 antennas and 16 users, using $Itr_{max} = 3$

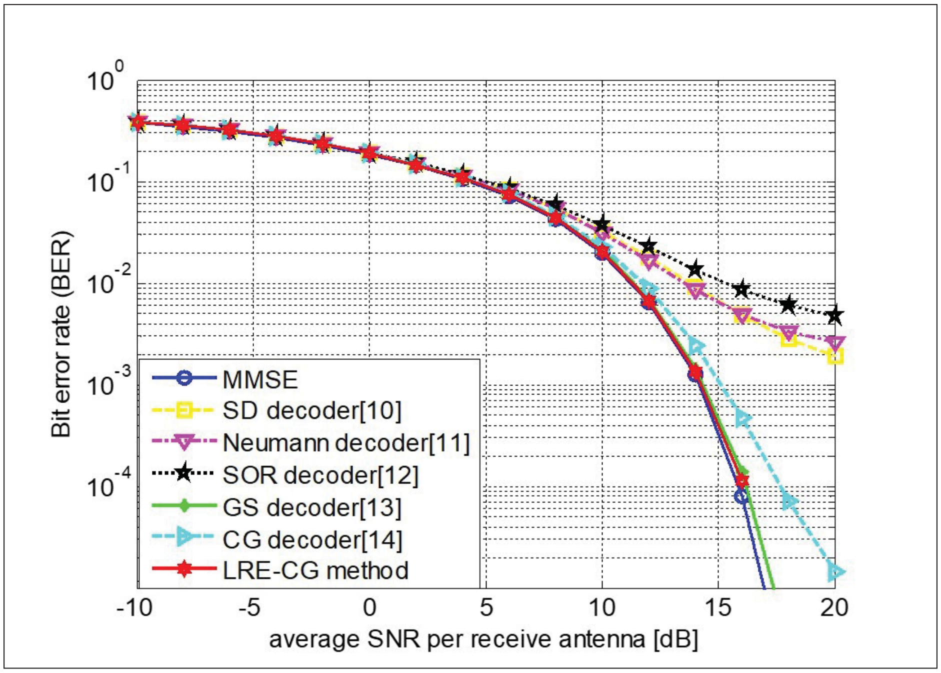
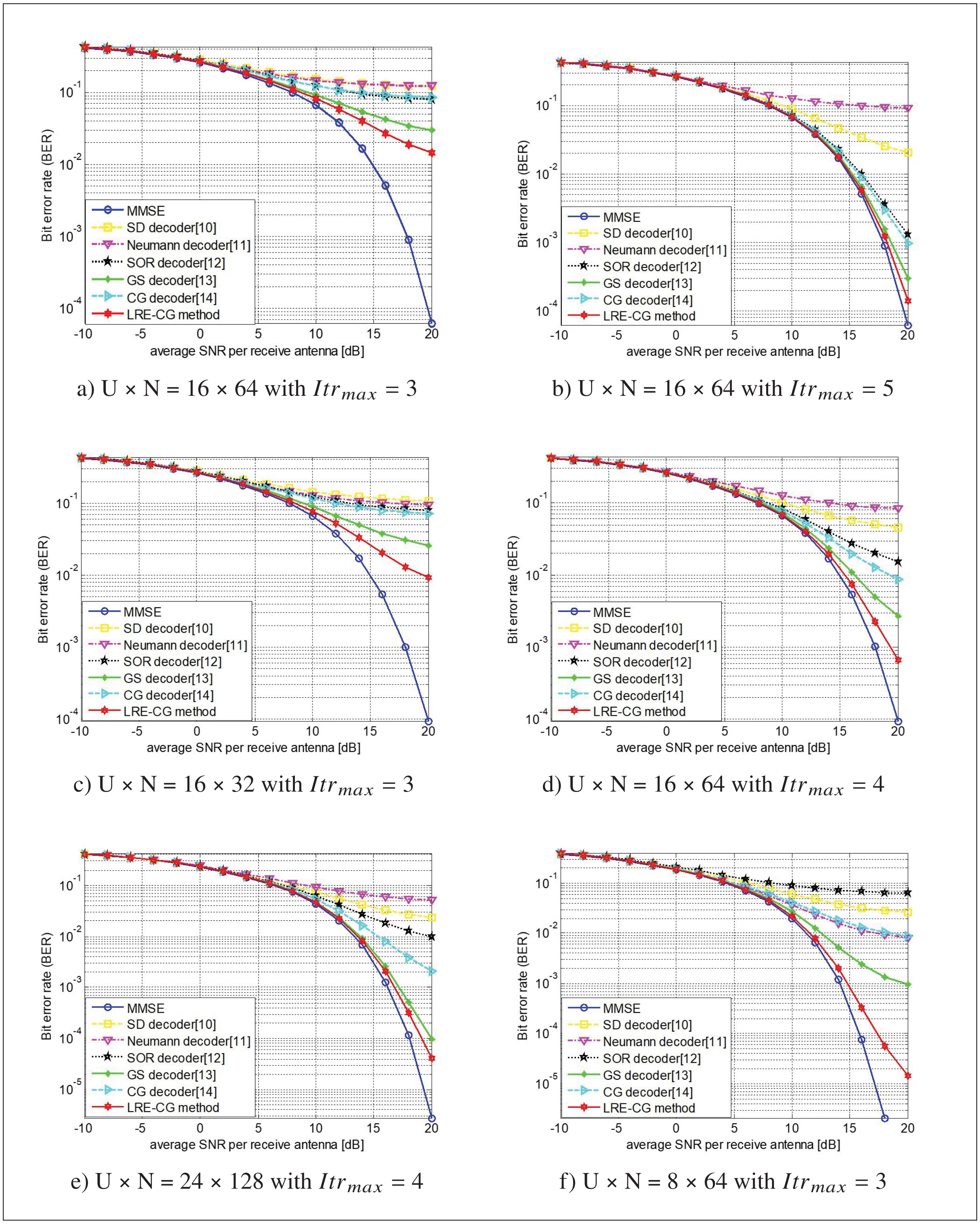


Figure 2.3 The evaluation of proposed approximation technique's BER to alternative approaches for calculating LLRs for a system with 128 antennas and 16 users, using $Itr_{max} = 4$

After that, in Figure 2.4, we compare the BER performance of the proposed algorithm with the SOR technique, the GS-based approach, the standard algorithm based on Neumann, and other algorithms in the literature using a variety of situations. It has been found that the suggested method operates admirably with a variety of antenna and user configurations. It is also demonstrated that when the number of iterations of the MMSE algorithm increases, the BER performance of the method approaches that of all traditional techniques in terms of BER. However, when a comparable number of iterations is used, the suggested technique is found to be superior when compared to the other approaches in terms of performance. As shown in Figure 2.4, we also offer simulation results that are based on a comparison between the number of antennas at the base stations and the BER performance of the proposed method when a certain number of users is taken into consideration. It can be observed that as the value of N grows, the performance of the MMSE technique improves in a corresponding fashion. The precise performance of the algorithm may be attained by the suggested technique, regardless of the number of antennas used, when the number of iterations is kept to a bare minimum, such as three iterations. On the contrary, the performance of the GS-based and Neumann-based algorithms improves when the number of iterations is increased, while there is still a performance loss due to the lack of negligibility in the algorithms. According to the results of this comparison, the other standard algorithms in the literature are less superior to the suggested method. The Neumann series approach also performs well in the scenario ($N \times U = 128 \times 8$) which reinforces the impression in (Gao *et al.*, 2014) that this method requires a high user-to-BS ratio ($p = N/U$), which is supported by the results of this study.

2.5 Conclusion

In conclusion, the proposed detection with approximation LLR calculation has high resilience against changes in channel correlation and loading factor, which is summarized in this paper. In our numerical findings, it has been demonstrated that, for relatively high ratios between the base station and user antennas, the proposed detection strategy rapidly corresponds to the performance of an accurate detection technique. So the proposed methodology is capable of producing performances that are comparable to those of an accurate inversion method while needing (in many cases) less computing complexity. Further to this point, the approximate Neumann series inversion and other schemes suggested in the literature are outperformed by the proposed scheme in terms of both efficiency and complication, and our system is less complicated. The proposed detector is efficient and can be used in a variety of antenna configurations in large MIMO systems with a variety of antenna types.



a) U × N = 16 × 64 with $Itr_{max} = 3$

b) U × N = 16 × 64 with $Itr_{max} = 5$

c) U × N = 16 × 32 with $Itr_{max} = 3$

d) U × N = 16 × 64 with $Itr_{max} = 4$

e) U × N = 24 × 128 with $Itr_{max} = 4$

f) U × N = 8 × 64 with $Itr_{max} = 3$

Figure 2.4 BER performance comparison in the massive MIMO uplink

CHAPTER 3

DECOUPLING ENERGY EFFICIENT APPROACH FOR HYBRID PRECODING-BASED MMWAVE MASSIVE MIMO-NOMA WITH SWIPT

3.1 Introduction

With 5G wireless communication networks, it is increasingly important to provide services with much higher quality, including improving the system capacity within the constrained service power and spectrum resources (Uwaechia *et al.*, 2019). Massive MIMO, utilizing millimeter waves (mmWave), is an emerging technology for 5G/6G wireless communications because it offers higher bandwidth and better spectrum efficiency (Mumtaz, Rodriguez & Dai, 2016). Throughput and spectral efficiency are improved by orders of magnitude when the mmWave bandwidth is increased (Hemadeh, Satyanarayana, El-Hajjar & Hanzo, 2017). This makes 5G wireless communication an appealing technology for the future. Theoretically speaking, the capacity for multiuser MIMO (massive MIMO) to enhance spectral efficiency by order of magnitude has been proven to be a more significant multiuser gain (Heath *et al.*, 2016). However, the use of non-orthogonal multiple access (NOMA) in millimeter-wave large MIMO systems has recently been investigated to improve spectrum efficiency (Dai, Wang, Peng & Chen, 2018; Uwaechia & Mahyuddin, 2020; Zhu *et al.*, 2019a; Alkhateeb *et al.*, 2015). Through the combination of multiple power levels on the same frequency resource block, NOMA can enhance spectral efficiency across the entire system. This has led to the emergence of NOMA as a contender for 5G wireless communication technologies (Dai *et al.*, 2018). Overall, mmWave with higher frequencies is better suited for antenna arrays with a massive MIMO system due to the small physical size of the huge antenna array. In addition, a large antenna array can use precoding to avoid free space path loss of mmWave signals, thereby achieving significant array gain for connections with a quality Signal-to-noise ratio (SNR) (Uwaechia & Mahyuddin, 2020).

Massive MIMO systems employ a significant number of antennas, each of which has a single RF chain, resulting in higher costs and more energy usage. A solution to this problem has been offered in the form of hybrid precoding (HP), which helps to effectively decrease the number of RF chains needed in mmWave massive MIMO without causing a visible drop in performance (Zhu *et al.*, 2019a; Alkhateeb *et al.*, 2015). HP focuses on developing completely digital precoders, which are composed of several analog and RF chains, to boost antenna gain and, as a result, reception quality (Choi *et al.*, 2019). It is often common to see HP networks with both fully connected and sub-connected topologies (Cheng, Yue, Yu, Liang & Li, 2019). Sub-connected architectures are predicted to provide greater energy efficiency.

Although there are various ways to improve the system's energy efficiency, improving the longevity of numerous limited-power mobile devices and enhancing the energy efficiency of the system are also critical considerations for 5G networks, particularly in the application scenarios of the internet of things (IoT) and Massive Machine-Type Communications (mMTC). A revolutionary technology termed SWIPT was introduced in (Zhu *et al.*, 2016) as a result of the advancement and development of wireless power transfer (WPT). Although SWIPT has certain advantages, the significant disparity in signal strength between the information decoder and rectifier circuit causes this technology to be underutilized (Huang *et al.*, 2019). Two effective receiving methods, time switching (TS) and power splitting (PS), were developed in (Gao, Dai, Han, Chih-Lin & Heath, 2016) to solve this problem. These schemes used time switching (TS) and power splitting (PS), with information decoding (ID) and EH, performed in a separate time and power domains, respectively. As a result, SWIPT enables an improved system EE, a viable green communication option for future wireless networks. Therefore, it has been noticed by both researchers and industrial people (Zhu *et al.*, 2019b; Zhang, Dong, Jin & Yuan, 2017).

Precoding is done fully in the digital domain to eliminate interference between distinct data streams in the standard cellular frequency spectrum (e.g., 2–3 GHz). Because of the higher energy demands, each antenna requires a specialized RF chain (including a digital-to-analog converter, up converter, etc.) with a total energy usage of approximately 250 mW per RF chain (Zhu *et al.*, 2016). A significant number of RF chains will be required for a mmWave massive MIMO system with 64 antennas because of the usual digital precoding method. A hybrid analog-digital precoding solution was developed to address this problem. Instead of using traditional digital precoding, RF chains are used to obtain these results, and an analog precoder is implemented using a large number of analog phase shifters (PSs) (Zhang *et al.*, 2017). There is no performance difference between digital and hybrid precoding because hybrid precoding uses fewer RF chains while delivering equivalent energy efficiency.

Two distinct classifications may be used for the current hybrid precoding strategies. the preliminary works (Boyd, Boyd & Vandenberghe, 2004) that described the use of sparse precoding to hybrid precoding is called "precoding with sparse precoding." (Tang *et al.*, 2020) presented an efficient method called orthogonal matching pursuit (OMP) to attain nearly optimum performance. In the second hybrid precoding method, which involves iterative searching among predefined codebooks (Zhang *et al.*, 2009), the best hybrid precoding matrix was found iteratively by sequentially passing through the codebooks. Each RF chain is linked to all base-station (BS) antennas through PSs. Under the assumption that there are a huge number of BS antennas (e.g., 256, as studied in (Ng *et al.*, 2013), the fully connected design will require thousands of PSs, which might introduce three new limitations: 1) in order to generate more energy, the larger phased array radar needs to absorb more energy for excitation; 2) in order to compensate for the insertion loss of PS, the larger phased array radar requires more energy; 3) because of the higher computational complexity, the larger phased array radar consumes more energy. While the hybrid precoding method with the sub-connected design uses fewer PSs, it requires all RF chains to be linked to each BS antenna. Because the sub-connected architecture is projected to be more energy efficient and simpler to implement for mmWave MIMO systems, it follows that the sub-connected architecture is expected to be more energy efficient and easier to implement for mmWave MIMO systems. The initial challenge of hybrid precoding with a fully connected architecture is difficult because of the new limitations imposed by the sub-connected architecture (Pal, Srinivas & Chaitanya, 2018).

The NOMA technique was previously used for beamspace MIMO for the first time in (M. Elmagzoub, 2020), which may be considered a straightforward realization of HP, and power allocation was adjusted to maximize the sum rate that could be achieved. Furthermore, in (Zhu *et al.*, 2019a), the HP architecture employed NOMA overall, and digital precoding was implemented using digital block diagonalization (BD) precoding. In addition, more complex digital precoding was suggested in (Zhao *et al.*, 2018), known as minimization maximization (MM)-based precoding. Therefore, the power allocation for mmWave large MIMO-NOMA systems was adjusted to improve their energy efficiency, and an iterative technique was suggested to optimize the power allocation.

Improved spectrum efficiency, along with improvements in energy efficiency, is among the most key performance indicators (KPIs) for 5G, which are projected to result in an approximately 100-fold increase in spectral efficiency compared to the present 4G wireless communications. Toward this end, SWIPT, presented for the first time in (Raviteja *et al.*, 2017), has gained wide acceptance in the last few years. SWIPT proposes that the same received RF signals may include both information and energy, and that this may be accomplished using power-splitting receivers in practice. SWIPT is a tool used to increase the battery life of wireless communication devices by harvesting energy from RF signals. This can advance networks such as the Internet of Things, especially in IoT with many wireless devices. Careful consideration of the trade-off between information rate and harvested energy level is necessary when SWIPT is employed in multiuser systems because inter-user interferences might negatively impact the ID while supporting the EH (Aljumaily & Li, 2019). Indeed, initiatives have been put out to address this issue. In addition, in (Zhu *et al.*, 2019b), the transmit power was reduced under the signal-to-interference-plus-noise ratio (SINR) and Quality of service (QoS) requirements for multiuser MIMO systems to minimize interference and noise.

A further aspect of interest is the combined transceiver and power-splitting SWIPT downlink design, which also uses the mean squared error (MSE) criteria (Zhao *et al.*, 2019). The combined transceiver and power splitting design was explored to enhance the energy efficiency in multicell multiuser downlink SWIPT systems. Even though SWIPT is capable of providing efficient wireless communications, it has only been tested on single-user systems, where future challenges to the joint transceiver and power splitting optimization will emerge.

In this thesis, we are interested in a new system that can exist by combining the spectrum-efficient mmWave massive MIMO-NOMA systems with energy-efficient SWIPT. This work presents a new way to solve the joint power allocation, power splitting, and joint precoding problem in SWIPT-enabled mmWave MIMO-NOMA systems by incorporating user groupings.

as we will explain in the chapter We explore hybrid analog/digital precoding and power splitting optimization to create SWIPT-enabled mmWave mMIMO-NOMA systems with hybrid analog-digital recording. To focus on the clustering process, In section 3.2 propose a new affinity propagation clustering method for user grouping to help with the initial cluster formation process. The parameters for this algorithm include the channel correlation and channel distance values. In this case, we consider the hybrid analog-digital precoder, power allocation, and power slitting factor optimization problem as a sum-rate maximization problem. We seek to maximize the overall power and minimum rate values under the set power and rate restrictions for each UE. We have now set out to build a hybrid mmWave MIMO-NOMA precoding matrix to overcome this challenge. In the first step, the analog precoder is intended to ensure that all beams acquire the maximum equivalent channel gain, depending on the user groupings. Finally, we construct the digital precoding vector for each UE, which prioritizes those users with the most substantial equivalent channel gain per beam to minimize inter-user interference. To simplify our total power and minimum rate restrictions at each UE, we frame the issue as a combined optimization of power allocation and power-splitting factors. The added requirement is that both variables are limited. To optimize the attainable data rate of the system given the restrictions of transmit power and EH need, the combined power allocation and splitting control issue is mathematically modeled. Because of the interrelationship between the linked variables, non-convex and complicated issues emerge. . In contrast to (Uwaechia *et al.*, 2019; Alkhateeb *et al.*, 2015), we propose decoupling the joint power allocation and transmit power. Before attempting to optimize the PS ratio assignment with fixed power allocation, we address the subproblem of optimizing the PS ratio assignment with varying power allocation. The Lagrangian duality approach helps solve both sub-problems. Convergence is established when this technique is performed several times. HP-based mmWave massive MIMO-NOMA systems with SWIPT were simulated to evaluate energy efficiency and spectrum efficiency. The results showed an enhancement in energy efficiency and spectrum. The proposed method for mmWave massive

MIMO-NOMA systems with SWIPT can outperform those of mmWave massive MIMO-OMA systems with SWIPT by achieving greater spectrum and energy efficiency.

The remainder of this chapter is structured as follows. Specifically, Section 3.2 describes the system model of the SWIPT-enabled mmWave mMIMO-NOMA system with hybrid analog-digital precoding as well as the sum-rate issue formulation. Section 3.3 describes the design of the user-grouping algorithm. The hybrid analog-digital precoder design is presented in Section 3.4. In Section 3.5, the formulation of the problem itself and an iterative optimization technique to further simplify the solution of the non-convex issue, are presented. Section 3.6 presents the results of the simulations for attainable rates and energy efficiency. Section 3.7 concludes the paper with a summary of the findings.

3.2 System Model

Take into account a single-cell downlink mmWave massive MIMO-NOMA system. The base station (BS) consists of N_{RF} RF chains and N_t transmitted antennas to serve K, single antenna users. In this study, we assume that the user equipment is supplied with a power-splitting receiver for SWIPT.

Each antenna is connected to a dedicated RF chain in a fully digital MIMO system, as shown in 3.1. Moreover, the required number of RF chains is equal to the number of antennas, which causes high power consumption and expensive hardware costs. The fully hybrid precoding architecture is shown in 3.2. It is evident that the required number of RF chains in the hybrid precoding architecture is less than the number of antennas. Each of the N_{RF} RF chains in the fully hybrid precoding is linked to all N antennas owing to phase shifters. However, the required phase shifters are equal to NN_{RF} and each RF chain can employ the full array gain. In the subconnected hybrid precoding architecture 3.3, the required phase shifters are equal to N because each RF chain is linked to a subset of N base-station antennas.

In (Dai *et al.*, 2018), it has been shown that the number of RF chains is larger than or equal to the number of beams, and each beam can only tolerate one user in hybrid precoding based

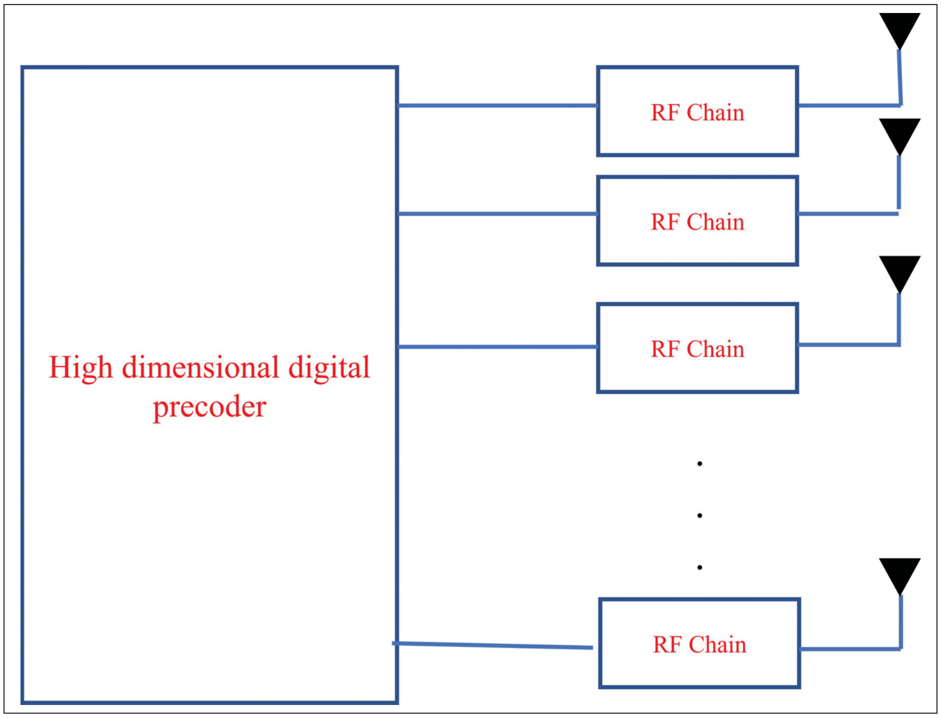


Figure 3.1 mmWave massive MIMO-NOMA with fully digital precoding architecture

on mmWave massive MIMO systems. However, we assume that the number of beams, G, equals the number of RF chains, N_{RF} to obtain the full multiplexing gain. Moreover, NOMA technology can be employed to make each beam tolerate more than one user. Consider $S_g \forall g = 1, 2, \ldots, G$ represents the set of users supported by the gth beam with $|S_g| \geq 1$, and we have $S_i \cap S_j = \emptyset \; \forall i \neq j$, thus $\sum_{g=1}^{G} |S_g| = K$. Then, the received signal at the mth user in the gth beam is given by:

$$y_{g,m} = h_{g,m}^H A \sum_{i=1}^{G} \sum_{j=1}^{|S_i|} d_i \sqrt{p_{i,j}} s_{i,j} + v_{g,m} \tag{3.1}$$

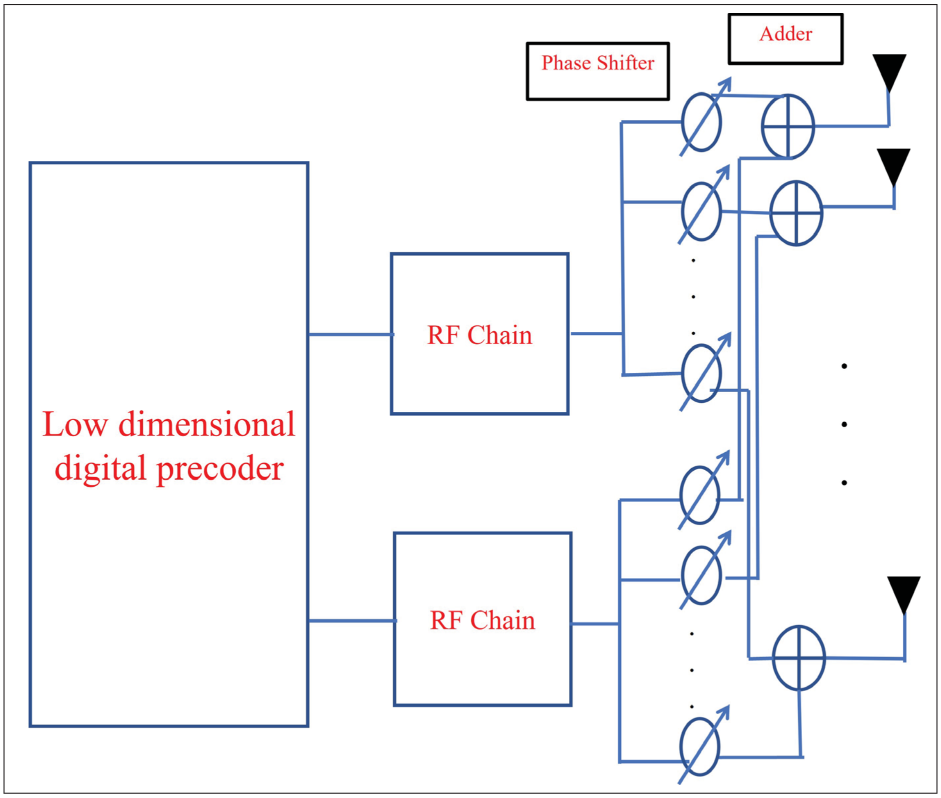


Figure 3.2 mmWave massive MIMO-NOMA with fully connected HP architecture

$$y_{g,m} = h_{g,m}^{H} A d_g \sqrt{p_{g,m}} s_{g,m} + h_{g,m}^{H} A d_g \left(\sum_{j=1}^{m-1} \sqrt{p_{g,j}} s_{g,j} + \sum_{j=m+1}^{|S_g|} \sqrt{p_{g,j}} s_{g,j} \right) + h_{g,m}^{H} A \sum_{i \neq g} \sum_{j=1}^{|S_i|} d_i \sqrt{p_{i,j}} s_{i,j} + v_{g,m} \quad (3.2)$$

In equation 3.2, the first, second, third, and last terms represent the desired signal, intra-beam interference, inter-beam interference, and noise, respectively. Where $s_{g,m}$ denotes the transmitted signal with $E\left\{\left|S_{g,m}\right|^2\right\} = 1$, $p_{g,m}$ represents the transmitted power of the mth user in the gth beam, $v_{g,m} \in CN\left(0, \sigma_v^2\right)$ is the complex noise, $d_g \in C^{N_{RF} \times 1}$represents the digital precoding vector of the gth beam, and $A \in C^{N \times N_{RF}}$ denotes the analog precoding matrix, where

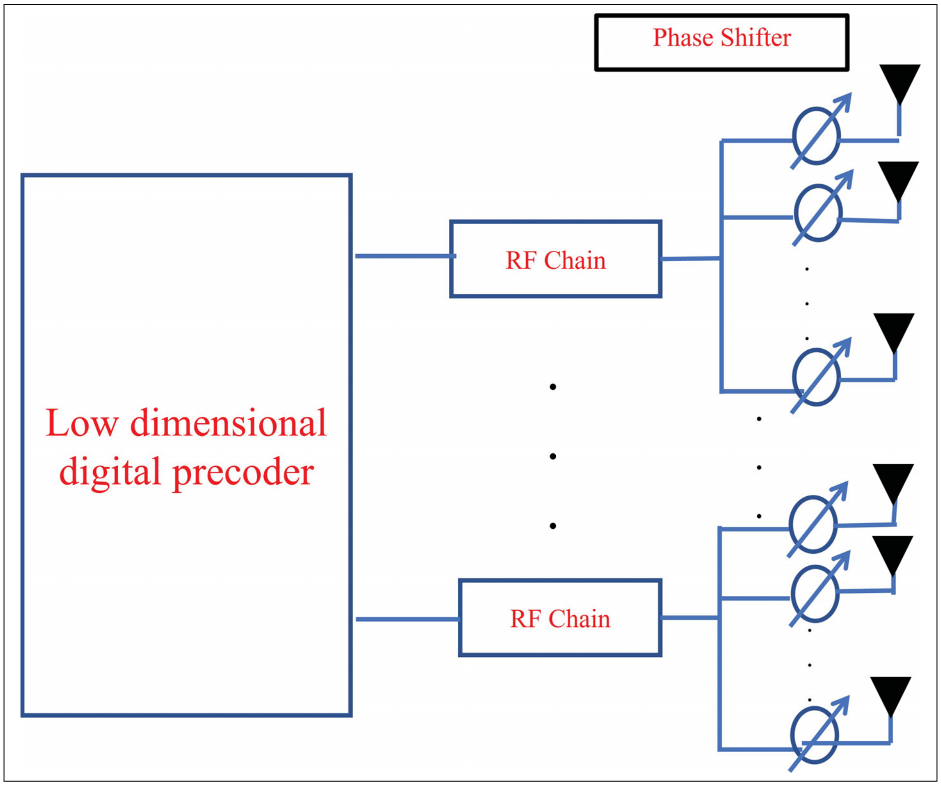


Figure 3.3 mmWave massive MIMO-NOMA with Sub-connected HP architecture

$||Adg||_2 = 1 \forall g = 1, 2, \cdots, G$. For the fully hybrid precoding architecture, the analog precoding matrix $A^{(full)}$ is given by:

$$A^{(full)} = \left[\bar{a}_1^{(full)},\ \bar{a}_2^{(full)},\ \ldots,\ \bar{a}_{N_{RF}}^{(full)}\right] \tag{3.3}$$

where $\bar{a}_n^{(full)} \in \mathbb{C}^{N\times 1}\ \forall\ n = 1,\ 2,\ \ldots,\ N_{RF}$ is the steering vector with the same amplitude of $\frac{1}{\sqrt{N}}$ and different phases (Heath *et al.*, 2016). For the sub-hybrid precoding architecture, the analog precoding matrix $A^{(sub)}$ is given by:

$$A^{(sub)} = \begin{bmatrix} \bar{a}_1^{(sub)} & 0 & \ldots & 0 \\ 0 & \bar{a}_2^{(sub)} & \ldots & 0 \\ 0 & 0 & \ldots & \bar{a}_{RF}^{(sub)} \end{bmatrix} \tag{3.4}$$

With no loss of generality, let us assume that $M = \frac{N}{N_{RF}}$ is an integer, and each RF chain is linked with M antennas in the sub-hybrid precoding architecture. $\bar{a}_n^{(sub)} \in \mathbb{C}^{M\times 1}\ \forall\ n = 1,\ 2,\ \ldots,\ N_{RF}$ is the steering vector with the same amplitude of $\frac{1}{\sqrt{M}}$ (Dai *et al.*, 2018; Uwaechia & Mahyuddin, 2020).

Let us consider the mmWave MIMO channel model (Heath *et al.*, 2016; Dai *et al.*, 2018; Uwaechia & Mahyuddin, 2020), where the $N \times 1$ channel vector $h_{g,m}$ of the mth user in the gth beam is given by

$$h_{g,m} = \sqrt{\frac{N}{L_{g,m}}} \sum_{l=1}^{L_{g,m}} \alpha_{g,m}^{(l)} a\left(\vartheta_{g,m}^{(l)}, \theta_{g,m}^{(l)}\right) \tag{3.5}$$

where $L_{g,m}$ represents the number of paths of the mth user in the gth beam. $\alpha_{g,m}^{(l)}, \vartheta_{g,m}^{(l)} and \theta_{g,m}^{(l)}$ denote the complex gain, the azimuth angle of departure (AoD) and the elevation angle of departure of the lth path respectively. $a\left(\vartheta_{g,m}^{(l)}, \theta_{g,m}^{(l)}\right) is the N \times 1$ steering vector.

For a uniform linear array (ULA) with N_1 elements in the horizon and N_2 elements in the vertical direction, the array steering vector $a\left(\vartheta_{g,m}^{(l)}, \theta_{g,m}^{(l)}\right)$ is given by

$$a\left(\vartheta_{g,m}^{(l)}, \theta_{g,m}^{(l)}\right) = a_{az}\left(\vartheta_{g,m}^{(l)}\right) \otimes a_{el}\left(\theta_{g,m}^{(l)}\right) \tag{3.6}$$

Where

$$a_{az}\left(\vartheta_{g,m}^{(l)}\right) = \frac{1}{\sqrt{N_1}} \left[e^{j2\pi i\left(\frac{d_1}{\lambda}\right) sin\left(\vartheta_{g,m}^{(l)}\right)} \right]_{i\in J(N_1)} \tag{3.7}$$

And

$$a_{el}\left(\theta_{g,m}^{(l)}\right) = \frac{1}{\sqrt{N_2}} \left[e^{j2\pi i\left(\frac{d_2}{\lambda}\right) sin\left(\vartheta_{g,m}^{(l)}\right)} \right]_{i\in J(N_2)} \tag{3.8}$$

where $J(n) = \{0,\ 1,\ \ldots,\ n-1\}, \lambda$ is the signal wavelength, d_1 and d_2 are the horizontal and antenna spacings, respectively. We usually assume that $d_1 = d_2 = \frac{\lambda}{2}$ for mmWave communication systems (Dai *et al.*, 2018).

Power splitting receivers allow one to split the received signal into two parts. While some of the signals are used for information decoding (ID), others can be used for energy harvesting (EH) (Zhang *et al.*, 2017).

The signal for energy harvesting is expressed as:

$$y_{g,m}^{EH} = \sqrt{1 - \beta_{g,m}} y_{g,m} \tag{3.9}$$

where $\beta_{g,m} \in [0, 1]$ is the power factor for the mth user in the gth beam, and the harvested energy is given by:

$$P_{g,m}^{EH} = \eta \left(1 - \beta_{g,m}\right) \left(\sum_{i=1}^{G} \sum_{j=1}^{|S_j|} \|\bar{h}_{g,m}^{H} d_i\|^2 P_{i,j} + \sigma_v^2 \right) \tag{3.10}$$

where $\bar{h}_{g,m}^{H} = h_{g,m}^{H} A$ represents the equivalent of the channel vector and $\eta \in [0, 1]$ denotes the energy conversion efficiency. However, the signal for information decoding is given by

$$y_{g,m}^{ID} = \sqrt{\beta_{g,m}} y_{g,m} + u_{g,m} \tag{3.11}$$

where $u_{g,m} \in \mathbb{CN} \left(0, \sigma_u^2\right)$ represents the noise of the power splitter.

Based on the NOMA at each beam, SIC at the receiver was performed as well as intra-beam superposition coding at the transmitter. With no loss of generality, let us assume that $\|\bar{h}_{g,2}^{H} d_i\|^2 \geq \ldots \geq \|\bar{h}_{g,|s_i|}^{H} d_i\|^2 \forall\ g = 1, 2, \ldots, G$. Then, the mth user in the gth beam can be diminished the interference from the jth user (for all $j > m$) in the gth beam using the SIC method (Zhu *et al.*, 2016). The signal for information decoding at the mth user in the gth beam is as follows:

$$y_{g,m}^{ID} = \sqrt{\beta_{g,m}}\left(\bar{h}_{g,m}^{H} d_g \sqrt{p_{g,m}} s_{g,m} + \quad \bar{h}_{g,m}^{H} d_g \sum_{j=1}^{m-1} \sqrt{p_{g,j}} s_{g,j} + \bar{h}_{g,m}^{H} \sum_{i \neq g} \sum_{j=1}^{|S_i|} d_i \sqrt{p_{i,j}} s_{i,j} + \quad v_{g,m}\right) + u_{g,m} \tag{3.12}$$

Then, the SINR at the mth user in the gth beam is expressed as:

$$\gamma_{g,m} = \frac{\|\bar{h}_{g,m}^{H} d_g\|_2^2 p_{g,m}}{\varepsilon_{g,m}} \tag{3.13}$$

Where

$$\varepsilon_{g,m} = \|\bar{h}_{g,m}^{H} d_g\|_2^2 \sum_{m-1}^{j=1} p_{g,j} + \sum_{i \neq g} \|\bar{h}_{g,m}^{H} d_g\|_2^2 \sum_{|S_j|}^{j=1} p_i, j + \sigma v2 + \sigma u^2 \beta_{g,m} \tag{3.14}$$

Accordingly, the achievable rate is given by:

$$R_{g,m} = \log_2\left(1 + \gamma_{g,m}\right) \tag{3.15}$$

Lastly, the achievable sum rate is given by:

$$R_{sum} = \sum_{g=1}^{G} \sum_{m=1}^{|S_g|} R_{g,m} \tag{3.16}$$

Nevertheless, the achievable sum rate in (3.16) can be enhanced by designing user grouping, the analog RF precoder matrix A^{RF} digital baseband precoders d_g for the g-th UE, power allocation, and power splitting factors.

3.3 User Grouping

As the number of users (K) is larger than that of the RF chain N_{RF}, that is, $K > N_{RF}$, we need to schedule the user into G groups, that is, $G = N_{RF}$. To this end, we propose an intuitive algorithm for user grouping. Owing to the spatial directivity of the SWIPT-based mmWave Massive MIMO NOMA System, we use the affinity propagation clustering algorithm to implement the user grouping (Huang *et al.*, 2019; Gao *et al.*, 2016) For mmWave large MIMO systems, a clustering technique based on user multidimensional attributes is described in order to increase system performance by considering the similarity of users' characteristics. Our solution, which uses mmWave Massive MIMO NOMA technology, calculates the relevance between users based on their characteristics to cluster them efficiently and precisely. We consider two types of features: h and p, which represent the user channel vector and the distance between users respectively and we define the feature vector $V = (h, p)$. Furthermore, prior to clustering, it is necessary to normalize the multidimensional aspects of the user characteristics. Here, the linear normalizing approach is modified to regulate the outcomes in the range of $[0, 1]$ in order to achieve better control. Our distance measures the similarity between user channels. We utilize the Euclidean distance to measure the similarity between users' relative locations, which is a vector because the transmission channel is also a vector. The statement of the relevance between users i and j, on the other hand, is defined as (Huang *et al.*, 2019).

$$S_1 = H_{ij} = \operatorname{arcsch} \frac{h_i^H hjh^H jh_j}{\|h_i\|\|h_j\|} \tag{3.17}$$

$$S_2 = P_{ij} = d_{ij} \tag{3.18}$$

$$S_{i,j} = -\sqrt{w_1 S_1^2 + w_2 S_2^2} \tag{3.19}$$

where $\sum_{i=1}^{2} w_i = 1 \forall\, w_i \in [0, 1]$ is the weight factor associated with the characteristics that meet the criteria. The higher the similarity between two users, the closer the distance between them. Thus, we utilize the negative distance to make it positively linked.

The affinity propagation (AP) clustering algorithm (Huang *et al.*, 2019) is a semi-supervised clustering algorithm that does not require the user to specify the initial cluster center or the number of clusters in advance. It has good clustering stability and a low error rate and is widely used. We utilize the idea of information transmission of the AP method (Huang *et al.*, 2019) based on multidimensional similarity for grouping users, as described in detail below.

1. By calculating and assigning the median of the similarity matrix for each user K in the similarity matrix $[S]$, the reference degree of user K may be determined and assigned to the vector $s(i,\ k)$.
2. Create a 0 in the responsibility $r(i,k)$ and availability $a(i,k)$ matrix to represent the initial state. Calculate the right number of iterations,Itr, as well as the damping factor (λ);
3. Use the following procedure to repeatedly compute the responsibility and availability for each user k with respect to user i in Itr times:

$$r_{t+1}(i,k) = s\,(i,k) - \max_{k \neq k\prime} \{a_t\,(i,k\prime) + s\,(i,k\prime)\}$$

$$a_{t+1}\,(i,k) = min\left\{0, r_t\,(k,k) + \sum_{i\prime \notin \{i,k\}} max\,\{0, r_t\,(i\prime,k)\}\right\},\ i \neq k$$

$$a_{t+1}\,(i,k) = \sum_{i\prime \neq k} max\,\{0, r_t\,(i\prime,k)\}$$

4. Calculate the responsibility $r_t(i,k)\forall i = 1,\ldots,n$ and availability $a_t\,(i,k)\,\forall i = 1,\ldots,n$. In order to update information in the AP method, one can incorporate the attenuation coefficient (γ), which is a real number between 0 and 1, with a typical value of between 0.5 and 0.9:
5.

$$\hat{r}_{t+1}(i,k) = (1-\gamma)r_{t+1}(i,k) + \gamma r_t(i,k)\hat{a}_{t+1}\,(i,k) = (1-\gamma)a_{t+1}\,(i,k) + \gamma a_t\,(i,k)$$

6. Update the responsibility $r_{t+1}(i,k)\forall i = 1,\ldots,n$ and availability $a_{t+1}\,(i,k)\,\forall i = 1,\ldots,n$.
7. Calculate $e(k,k) = r(k,k) + a(k+k)\forall k = 1, 2, ..., K$, and if $e(k,k) > 0, k$ is the center of the cluster. After that, the cluster center set of users is established. Each user is allocated to the appropriate cluster based on the concept of the minimal distance between the two clusters.

Algorithm 1 provides the pseudocode for the improved AP scheme, which is a mathematical representation of the code.

Algorithm 3.1 Presented User Grouping Method

Input:
Number of UEs: $K > N^{RF}$
Number of RF chains: N^{RF}
Number of beams: G
Channel Matrix: $\boldsymbol{H} = [\boldsymbol{h_1}, \boldsymbol{h_2}, \ldots, \boldsymbol{h_K}]$
Number of BS antennas: N
Initialization: $M = 0^G$
Set predefined threshold: $0 \leq \gamma \leq 1$
Output:
Optimized User Grouping: $\mho = \hat{g}_1, \hat{g}_2, \ldots, \hat{g}_G$

$\mathcal{K} = \{1,\ 2,\ \ldots,\ K\}$
Initilize $\Omega_m^{(1)} = k_m \in \mathcal{K} \forall\ m = 1, 2, \ldots, G$
$\Psi = \|h_1\|_2, \|h_2\|_2, ..., \|h_k\|_2$
$\bar{\boldsymbol{H}} = \left[\frac{\boldsymbol{h_1}}{\|\boldsymbol{h_1}\|_2}, \frac{\boldsymbol{h_2}}{\|\boldsymbol{h_2}\|_2}, \ldots, \frac{\boldsymbol{h_K}}{\|\boldsymbol{h_K}\|_2}\right]$
Calculate S

$$\boldsymbol{S_{i,j}} = -\sqrt{w_1 S_1^2 + w_2 S_2^2}$$

$t = 1$
while $\left\{\Omega_m^{(t)} \neq \Omega_m^{(t-1)}\right\}$ **do**
  Initialize $\boldsymbol{\hat{\mathscr{g}}_m = \Omega_m^{(t)}}$
  for $k \in \frac{\mathcal{K}}{\left\{\Omega_m^{(t)}\right\}}$ **do**
    $\boldsymbol{g = arg \max_{1 \leq g \leq M} S_{i,j}}$
    $\boldsymbol{\mho = \hat{\mathscr{g}} \cup k}$
  end for
  $\boldsymbol{t = t + 1}$
  Update $\boldsymbol{\Omega_m^{(t)} for}\ m = 1, 2, \ldots, G$
end while
Return $\mho = \hat{\mathscr{g}}_1, \hat{\mathscr{g}}_2,\ \ldots, \hat{\mathscr{g}}_G$

3.4 Hybrid Precoding Design

To maximize (3.11) for each UE, we should reduce the inter-beam interference while simultaneously increasing the effective channel gain. Zero forcing (ZF) is a technique that may be used in conventional multiuser MIMO (MU-MIMO) systems (Dai *et al.*, 2018; Alkhateeb *et al.*, 2015; Pal *et al.*, 2018). We propose to use phase-only array response adjustment to link the N_{RF} RF chain outputs with the N_{BS} BS antennas, using low-cost phase shifters, in order to decrease hardware restrictions while still realizing the full potential of mmWave huge MIMO-NOMA systems. Unfortunately, because of the elementwise constant-magnitude limitation on the analog precoder, that is, $\left|\left[F^{RF}\right]_{i,j}\right| = \frac{1}{\sqrt{N_{BS}}}$, $\forall i, j$, they cannot be used directly in the hybrid analog-digital precoding method (Dai *et al.*, 2018; Alkhateeb *et al.*, 2015; Pal *et al.*, 2018). Because of the constant-magnitude restriction, the subsets of feasible areas are not convex; thus, the solution is non-convex. Consequently, we are considering creating the analog RF precoder and the digital baseband precoder in distinct phases of the development process. Based on (Dai *et al.*, 2018; Choi *et al.*, 2019), we present an efficient analog RF precoding algorithm to design F^{RF} and a low-dimensional digital baseband precoding algorithm to design F^{BB} for downlink multiuser mmWave massive MIMO-NOMA systems. As a first step, we designed the analog RF precoding matrix.

3.4.1 Analog RF precoding method

Our goal with the analog RF precoder is for the phases of $H = [h_1,\ h_2,\ \ldots,\ h_K]$ to be aligned so that the high array gain delivered by the massive MIMO system can be harvested effectively. Using Algorithm 2, we can quickly review the analog RF precoder architecture. For simplicity, it is preferable to focus on the main element of the proposed algorithm rather than providing a redundant demonstration. Initially, we start the analog precoder as an all-zero matrix to ensure that it operates correctly. It is necessary to extract the phases of the conjugate transpose of the aggregate downlink mmWave massive MIMO-NOMA channel from the BS to numerous users in Step 4 to complete the computation. Phase alignment of channel components is performed in Step 10 to build the analog RF precoder in order to harvest a significant array gain. Subsequently, once the effective baseband channel has been coupled with the ideal analog RF precoder acquired, the digital baseband precoder design is carried out to minimize interference and maximize the sum rate that can be accomplished.

3.4.2 Digital bseband pecoding mehod

The digital baseband precoding matrix is designed such that only the UEs in each beam with strong channels are selected to eliminate inter-user interference. To avoid inter-beam interference, the structure of digital precoding is transformed into a typical massive MIMO-NOMA precoding issue. As shown in (Dai *et al.*, 2018; Alkhateeb *et al.*, 2015), the low-complexity zero-forcing (ZF) precoding technique is used for digital precoding without sacrificing generality. Specifically, we present an algorithmic solution based on the concepts of (Dai *et al.*, 2018; Alkhateeb *et al.*, 2015) after designing the analog RF precoder (F^{RF}). The pseudocode for the digital baseband precoder is given in Algorithm 3. We first set the number of UEs (K), number of RF chains N^{RF}, number of BS antennas (N), channel matrix H, the optimized analog RF precoder ($\hat{F}^{RF}$) from Algorithm 2, and the optimized user grouping from Algorithm 1. The precoding algorithm then employs a zero-force precoding algorithm to reduce inter-user interference. As a result, the digital baseband precoder can be represented as

$$\hat{F}^{BB} = H^{H}\left(HH^{H}\right)^{-1} \tag{3.20}$$

Then, we normalize the digital precoder as follows.

$$\hat{F}^{BB} = \left[\frac{\hat{f}_1^{BB}}{f_1^{BB*}}, \frac{\hat{f}_2^{BB}}{f_2^{BB*}}, \dots, \frac{\hat{f}_{N^{RF}}^{BB}}{f_{N^{RF}}^{BB*}},\right] \tag{3.21}$$

where $f_n^{BB*} = repmat(\|F^{RF} f_n^{BB*}\|_2, N_{RF}, 1)\, \forall\, n = 1, \dots, N^{RF}$ and $\|F^{RF} f_n^{BB*}\|_2 = \sqrt{\sum |F^{RF} f_n^{BB*}|^2}$.

Algorithm 3.2 Presented Analog RF Precoding Method for mmWave Massive MIMO-NOMA Systems with SWIPT

Input:
Number of UEs: $K > N^{RF}$
Number of RF chains: N^{RF}
Channel Matrix: $\boldsymbol{H} = [\boldsymbol{h_1}, \boldsymbol{h_2}, \ldots, \boldsymbol{h_K}]$
Optimized User Grouping: $\{\hat{\mathscr{g}}_1, \hat{\mathscr{g}}_1, \ldots, \hat{\mathscr{g}}_G\}$
Number of BS antennas: N
Initialization: $F^{RF} = 0^{N \times N^{RF}}$
Number of quantization bits: B
Output:
Optimal analog RF precoding: $\boldsymbol{F^{RF}}$

Set the phase: $\Lambda = \left\{ \frac{2\pi n}{2^B},\ n = 0,\ 1,\ \ldots.,\ 2^{B-1} \right\}$
for : $g = 1$ *to* G **do**
  Recall the optimized user grouping: $\{\hat{\mathscr{g}}_1, \hat{\mathscr{g}}_1, \ldots, \hat{\mathscr{g}}_G\}$
  Set the aggregate downlink channel: $\bar{\boldsymbol{H}} = [\boldsymbol{H}]_{:, \hat{\mathscr{g}}_d}$
  Extract phase of the $\bar{\boldsymbol{H}} : \mathfrak{G} = \angle\, \bar{\boldsymbol{H}}$
  Initialize angle: $\boldsymbol{\vartheta} = \boldsymbol{0}^N$
  for $m = 1$ *to* $|S_g|$ **do**
    $[\sim, \boldsymbol{k}] = \boldsymbol{min}\, |[\mathfrak{G}]_m - \Lambda|$
    $\boldsymbol{\vartheta}(\boldsymbol{m}) = [\Lambda]_k$
  end for
  Compute Optimal analog RF precoding: $\boldsymbol{F^{RF}}$

$$\boldsymbol{F^{RF}}(:, g) = exp(j\ \vartheta)$$

end for

3.5 Joint Optimization of Power Allocation and Power Splitting

In this section, we have investigated the combined power allocation and power splitting optimization to achieve the highest possible data rate in mmWave Massive MIMO-NOMA systems with SWIPT. Because of the presence of both inter-group and intra-group interferences in MIMO-NOMA systems with SWIPT, the current optimization algorithms for solving the joint optimization problem of power allocation and power splitting in MIMO systems with SWIPT are not suitable to be directly applied in MIMO-NOMA systems with SWIPT, where there are several groups and many users in each group. As a result, obtaining optimal solutions is quite difficult. To address this complex challenge, an iterative optimization technique is created in this section, which allows for the generation of suboptimal solutions while fulfilling the intended EH restrictions and the transmit power constraint requirements. Furthermore, the following formulation may be used to precisely express the issue of combined power allocation and power-splitting optimization:

$$\max_{\{p_{g,m}\},\{\beta_m\}} R_{sum}\left(p_{g,m}, \beta_m\right) \tag{3.22}$$

$$s.t. \qquad \sum_{g=1}^{G} \sum_{m=1}^{|S_g|} P_{g,m} \ \leq P_T \tag{3.23}$$

$$0 \leq \beta_m \leq 1 \qquad \forall\, m \tag{3.24}$$

$$p_{g,m} \geq 0 \quad \forall\, g, m \tag{3.25}$$

$$P_{g,m}^{EH} \geq p_{g,m}^{req} \tag{3.26}$$

Constraint 3.22 indicates that the transmitted power constraint, that is, $\sum_{g=1}^{G} \sum_{m=1}^{|S_g|} P_{g,m}$, cannot exceed the threshold of P_T being the maximum total transmission power of the BS. Constraint 3.24 limits the power splitting factor β_m for the mth user to be in the range of $[0, 1]$. Constraint 3.25 indicates the non-negativity of the power allocated to the mth user in the gth beam. Constraint 3.26 shows that each mth user in the gth beam is required to harvest at least $p_{g,m}^{req} W$ Power is the minimum amount of energy harvested for each mth user in the gth beam.

As a consequence of the objective function and the coupling of the multiple variables, the optimization issue of the attainable data rate described in equations (3.22-3.26) is neither convex

Algorithm 3.3 Presented Digital Baseband Precoding Method for mmWave Massive MIMO-NOMA Systems with SWIPT

Input:
Number of UEs: $K > N^{RF}$
Number of RF chains: N^{RF}
Channel Matrix: $\boldsymbol{H} = [\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_K]$
Optimized User Grouping: $\{\hat{\mathcal{G}}_1, \hat{\mathcal{G}}_1, \ \ldots, \ \hat{\mathcal{G}}_G\}$
Number of BS antennas: N
Optimal analog RF precoding: F^{RF}
Number of quantization bits: B
Output:
Optimal analog RF precoding: $\boldsymbol{F}^{\boldsymbol{RF}}$

Set the phase: $\Lambda = \left\{\frac{2\pi n}{2^B}, \ n = 0, \ 1, \ \ldots., 2^{B-1}\right\}$
$\bar{\boldsymbol{H}} = \boldsymbol{H}^H \boldsymbol{F}^{RF}$
$\tilde{\boldsymbol{H}} = \left[\bar{\boldsymbol{H}}\right]_{:,\hat{\mathcal{G}}_1}$
$\hat{\boldsymbol{F}}^{BB} = \tilde{\boldsymbol{H}}^H \left(\tilde{\boldsymbol{H}}\tilde{\boldsymbol{H}}^H\right)^{-1}$
$\hat{\boldsymbol{F}}^{BB} = \left[\frac{\hat{\boldsymbol{F}}_1^{BB}}{\|\hat{\boldsymbol{F}}_1^{BB*}\|_2}, \frac{\hat{\boldsymbol{F}}_2^{BB}}{\|\hat{\boldsymbol{F}}_2^{BB*}\|_2}, \ldots, \frac{\hat{\boldsymbol{F}}_{N^{RF}}^{BB}}{\|\hat{\boldsymbol{F}}_K^{BB*}\|_{N^{RF}}}\right]$
  where
  $\boldsymbol{f}^{BB*} = \boldsymbol{repmat}\left(\|\boldsymbol{f}^{RF}\boldsymbol{f}_n^{BB*}\|_2, \boldsymbol{N}^{RF}, \boldsymbol{1}\right) \forall \boldsymbol{n} = \boldsymbol{1}, \ldots, \boldsymbol{N}^{RF}$
Initialize baseband precoding: $\boldsymbol{F}^{BB} = \boldsymbol{0}^{N^{RF} \times K}$
$\left[\boldsymbol{F}^{BB}\right]_{:,\hat{\mathcal{G}}_1} = \hat{\boldsymbol{F}}^{BB}$
for : $g = 1$ *to* G **do**
  $\boldsymbol{\Lambda} = \boldsymbol{nonzeros}\left([\boldsymbol{\Lambda}]_{\hat{\mathcal{G}}_1}\right)^T$
  for $m = 2$ *to* $|\Lambda|$ **do**
    $\boldsymbol{F}^{BB}(:, \boldsymbol{\Lambda}_n) = \left[\boldsymbol{F}^{BB}\right]_{:,\hat{\mathcal{G}}_g}$
  end for
end for

nor linear owing to the objective function. Furthermore, the optimization problem mentioned above is a well-known NP-hard problem, and as a result, the solution is complex and cannot be easily achieved. There is a possibility that an exhaustive search approach will provide a solution to this problem. The computational complexity of the exhaustive search technique, on the other hand, increases substantially as the number of users increases. As a result, this technique is far from feasible, particularly in the context of IoT, where there is a desire for massive MIMO systems. We will create an iterative strategy to tackle this problem based on the Lagrangian duality methodology in this section, which will be as follows: It is feasible for any optimization issue containing many variables to deal with the sub-problem over a subset of variables while treating the remainder as constants and then dealing with the sub-problem over the remaining variables. This is supported by the literature (Boyd *et al.*, 2004; Tang *et al.*, 2020). This separation of $p_{g,m}$ and β_m allows us to create a realistic and effective solution for the studied optimization issue in equations (3.22-3.26). First, we examine the scenario in which all the components of the power allocation, $p_{g,m}\ \forall\ g, m$, are constants. Here, we focus on optimizing the power splitting factors $\beta_m\ \forall\ m$ under the fixed power allocation $p_{g,m}\ \forall\ g, m$. Therefore, the optimization subproblem can be rewritten as follows:

$$\max_{\{p_{g,m}\},\beta_{g,m}} \quad R_{sum}\left(\beta_m\right) \tag{3.27}$$

$$s.t. \qquad 0 \leq \beta_m \leq 1 \qquad \forall\ m \tag{3.28}$$

$$P^{EH}_{g,m} \geq p^{req}_{g,m} \tag{3.29}$$

According to equation (3.10) and constraint 3.29, $\beta_m\ \ \forall\ m$ is required to satisfy the following condition:

$$\beta_m \quad \leq 1 - \frac{p^{req}_{g,m}}{\eta\|\Sigma^{G}_{i=1}\Sigma^{|S_i|}_{j=1}\|\bar{h}^{H}_{g,m}d_i\|^2_2 p_{i,j} + \sigma^2_v} \cong \beta^{UB}_m \tag{3.30}$$

Considering equation (3.28) and (3.30) together, the supposed optimization problem is infeasible unless the $\beta^{UB}_m > 0\ \ \forall\ m$.

***Proposition*1 :** Assume that the process of power splitting in the receiver is almost idealized, and the noise power for all users in the gth beam is equal, that is, $|\sigma_u|^2 \rightarrow 0$. The considered optimization problem in equations (3.27-3.29) is convex with respect to the power splitting factors $\beta_{,m}\ \ \forall\ m$.

***Proof* :** First, we ensure that the viable power splitting factor area is not empty and convex to guarantee the convexity of the optimization issue in equations (3.27-3.29). Because of the limitation of $\beta_m^{UB} > 0 \ \ \forall\, m$, the feasible area of the power splitting factor is not empty, and its convexity can be determined using Equations equations (3.29) and (3.30) , respectively. After that, we conclude that the objective function in equation (3.27) is concave on the power splitting factors $\beta_m \quad \forall\, m$. Let us recall the equation (3.15) equations ,(3.31) and (3.32) can be expressed as shown

$$R_{g,m} = \log_2\left(1 + \frac{\|\bar{h}_{g,m}^H d_g\|^2 p_{g,m}}{\beta_m \left(\|\bar{h}_{g,m}^H d_g\|^2 |\Sigma_{j=1}^{m-1} p_{g,j} + \Sigma_{i\neq g}\|\bar{h}_{g,m}^H d_i\|_2^2 \Sigma_{j=1}^{|S_i|} p_{i,j} + \sigma_v^2\right) + \frac{\sigma_u^2}{\beta_m}}\right) \tag{3.31}$$

$$R_{g,m} = \log_2\left(1 + \frac{\|\bar{h}_{g,m}^H d_g\|^2 \beta_m p_{g,m}}{\beta_m \left(\|\bar{h}_{g,m}^H d_g\|^2 |\Sigma_{j=1}^{m-1} p_{g,j} + \Sigma_{i\neq g}\|\bar{h}_{g,m}^H d_i\|_2^2 \Sigma_{j=1}^{|S_i|} p_{i,j} + \sigma_v^2\right) + \sigma_u^2}\right) \tag{3.32}$$

Let us assume that

$$p_{g,m} = \|\bar{h}_{g,m}^H d_i\|_2^2 p_{g,m} \tag{3.33}$$

$$\beta_{g,m} = \|\bar{h}_{g,m}^H d_g\|^2 |\Sigma_{j=1}^{m-1} p_{g,j} + \Sigma_{i\neq g}\|\bar{h}_{g,m}^H d_i\|_2^2 \Sigma_{j=1}^{|S_i|} p_{i,j} + \sigma_v^2 \tag{3.34}$$

Given

$$R_g = \Sigma_{m=1}^{|S_g|} \log_2\left(1 + \frac{A_{g,m}\beta_{g,m}}{A_{g,m}\beta_{g,m} + \sigma_u^2}\right) \tag{3.35}$$

Which represents the achievable data rate on the gth beam. Thus, $R_s um(\beta_m)$ in equaion (3.27) is given as

$$R_{sum}(\beta_m) = \Sigma_{g=1}^{G} R_g \tag{3.36}$$

$$R_{sum}(\beta_m) = \Sigma_{g=1}^{G} \Sigma_{m=1}^{|S_g|} \log_2\left(1 + \frac{A_{g,m}\beta_{g,m}}{A_{g,m}\beta_{g,m} + \sigma_u^2}\right) \tag{3.37}$$

Then the first derivative of R_g with respect to β_m is given by

$$\frac{\partial R_g}{\partial \beta_m} = \frac{1}{\ln 2} \cdot \frac{A_{g,m}\sigma_u^2}{(A_{g,m}\beta_m + B_{g,m}\beta_m + \sigma_u^2)(B_{g,m}\beta_m + \sigma_u^2)} \tag{3.38}$$

Moreover, the second derivative of R_g with respect to β_m is given by

$$\frac{\partial^2 R_g}{\partial \beta_m^2} = \frac{1}{\ln 2} \cdot \frac{A_{g,m}\sigma_u^2(2(A_{g,m} + B_{g,m})B_{g,m}\beta_m + 2B_{g,m}\sigma_u^2 + A_{g,m}\sigma_u^2)}{(A_{g,m}\beta_m + B_{g,m}\beta_m + \sigma_u^2)(B_{g,m}\beta_m + \sigma_u^2)} \tag{3.39}$$

And

$$\frac{\partial^2 R_g}{\partial \beta_n \partial \beta_m} = 0 \qquad \forall n \neq m \tag{3.40}$$

According to the equation above, the corresponding Hessian matrix $\boldsymbol{H}$ is given by

$$H = \begin{pmatrix} H_{\mathbf{1}} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & H_{|S_g|} \end{pmatrix} \tag{3.41}$$

where $\boldsymbol{H}_m = \frac{\partial^2 R_g}{\partial {\beta_m}^2} \leq 0 \ \forall \ m \in \left[1, \left|S_g\right|\right]$. Correspondingly, the Hessian matrix is negative or equal to zero for all values of $\beta_m \ \forall \ m \in \left[1, \left|S_g\right|\right]$, then the R_g is concave with respect to β_m. Therefore, the objective function in equation (3.27) is concave on the power splitting factors $\beta_m \quad \forall m$ because it represents the finite summation of concave functions. To that end, one can obtain the near-optimal solution for the optimization problem in equations(3.27)-(3.27) by using the Lagrangian duality-based method (Boyd *et al.*, 2004). The corresponding Lagrangian function is formulated as in the following equation

$$\Upsilon\left(\beta,\lambda,\ \mu,\upsilon\right) =$$
$$\Sigma_{g=1}^{G}\Sigma_{m=1}^{|S_g|}\log_2\left(1+\frac{\beta_m\|\bar{h}_{g,m}^{H}d_i\|_2^2 p_{g,m}}{\beta_m\left(\|\bar{h}_{g,m}^{H}d_g\|^2|\Sigma_{j=1}^{m-1}p_{g,j}+\Sigma_{i\neq g}\|\bar{h}_{g,m}^{H}d_i\|_2^2\Sigma_{j=1}^{|S_i|}p_{i,j}+\sigma_v^2\right)+\sigma_u^2}\right)+$$
$$\Sigma_{m=1}^{|S_g|}\lambda_m\beta_m+\Sigma_{m=1}^{|S_g|}\mu_m\left(1-\beta_m\right)+\Sigma_{m=1}^{|S_g|}v_m\left(\mu_m\left(1-\beta_m\right)\left(|\Sigma_{i=1}^{G}\Sigma_{j=1}^{|S_j|}\|\bar{h}_{g,m}^{H}d_i\|^2 p_{i,j}+\sigma_v^2\right)-p_{g,m}^{req}\right) \tag{3.42}$$

where $\boldsymbol{\lambda} = \left[\lambda_1,\ \lambda_2,\ ...,\ \lambda_{|S_g|}\right]^T$ and $\boldsymbol{\mu} = \left[\mu_1,\ \mu_2,\ ...,\ \mu_{|S_g|}\right]^T$ are non-negative Lagrange multipliers, which correspond to constraint (3.28). $\boldsymbol{\upsilon} = \left[\nu_1,\ \nu_2,\ ...,\ \nu_{|S_g|}\right]^T$ is a non-negative Lagrange multiplier corresponding to constraint (3.29). Accordingly, one can express the Lagrange dual objective function as follows

$$\Gamma\left(\lambda,\mu,\upsilon\right) = \max_{\beta}\Upsilon\left(\beta,\lambda,\mu,\upsilon\right) \tag{3.43}$$

Then, one can model the Lagrange dual optimization problem as follows

$$\min_{\lambda,\mu,\upsilon}\Gamma\left(\lambda,\mu,\upsilon\right) \tag{3.44}$$

$$s.t. \qquad \lambda\succcurlyeq 0,\ \mu\ \succcurlyeq 0,\upsilon\ \succcurlyeq 0 \tag{3.45}$$

To solve the Lagrange dual issue mentioned earlier, we first optimize the PS factor $\boldsymbol{\beta}$ using the provided dual variables $(\boldsymbol{\lambda},\ \boldsymbol{\mu},\boldsymbol{\upsilon})$ using the gradient ascent technique, and then update the dual variables $(\boldsymbol{\lambda},\ \boldsymbol{\mu},\boldsymbol{\upsilon})$ with the optimized $\boldsymbol{\beta}$ using a well-known sub-gradient methodology (Zhang *et al.*, 2009) to obtain the optimal.

1. We find the gradient direction of the Lagrange objective function in (Raviteja *et al.*, 2017) regarding to power splitting factor $\boldsymbol{\beta_m}\ \forall\, \boldsymbol{m}$ to optimize the $\boldsymbol{\beta_m}$ with given variables $(\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\upsilon})$ as follows

$$\nabla_{\beta_m}\Upsilon = \sum_{g=1}^{G}\frac{1}{\ln 2}\cdot\frac{A_{g,m}\sigma_u^2}{\left(A_{g,m}\beta_m + B_{g,m}\beta_m + \sigma_u^2\right)\left(B_{g,m}\beta_m + \sigma_u^2\right)}$$
$$+\lambda_m - \mu_m + v_m\left(\eta\left(\sum_{i=1}^{G}\sum_{j=1}^{|S_j|}\|\bar{h}_{g,m}^H d_i\|^2 p_{i,j} + \sigma_v^2\right) - p_{g,m}^{req}\right) \tag{3.46}$$

where $\boldsymbol{A_{g,m}}$ and $\boldsymbol{B_{g,m}}$ are defined in (3.33) and (3.34), respectively. Particularly, $\boldsymbol{\beta_m}$ can be updated using the following formula

$$\beta_m\left(Itr\ +1\right) = \beta_m\left(Itr\right)\ + \varepsilon\left(Itr\right)\nabla_{\beta_m(Itr)}\Upsilon \tag{3.47}$$

where $\boldsymbol{\beta_m\ (Itr\ )}$ and $\boldsymbol{\beta_m\ (Itr\ +1)}$ represent the $\boldsymbol{\beta_m}$ in the Itr-th and $(Itr\ +1)$-th iterations, respectively. $\boldsymbol{\varepsilon\ (Itr)}$ defines the updated step size for the $\boldsymbol{\beta_m}$ in the Itr-th iteration and satisfies the following condition:

$$\varepsilon\left(Itr\right) = \arg\max_{\varepsilon}\Upsilon\left(\beta\left(Itr\ +1\right), \lambda, \mu, \upsilon\right)|_{\beta_m(Itr\ +1)=\beta_m(Itr)\ +\varepsilon(Itr)\nabla_{\beta_m(Itr)}\Upsilon} \tag{3.48}$$

Process in (46) is repeated until $\left|\boldsymbol{\nabla_{\beta_m(Itr)}\Upsilon}\right| \leq \boldsymbol{\epsilon_1}\ \forall\ \boldsymbol{m}$, and the optimal power splitting factor is denoted as $\boldsymbol{\beta^*}$. Therefore, the Lagrange dual-objective function in equation (3.43) is given by

$$\Gamma\left(\lambda, \mu, \upsilon\right) = \Upsilon\left(\beta^*, \lambda, \mu, \upsilon\right) \tag{3.49}$$

2. We update and determine the optimal Lagrange multipliers $\boldsymbol{(\lambda, \mu, \upsilon)}$ by solving the Lagrange dual optimization problem in (3.50)–(3.51) as follows

$$\min_{\lambda,\mu,\upsilon}\Gamma\left(\lambda, \mu, \upsilon\right) \tag{3.50}$$

$$s.t. \qquad \lambda \succcurlyeq 0,\ \mu\ \succcurlyeq 0, \upsilon\ \succcurlyeq 0 \tag{3.51}$$

To state it bluntly, the dual issue is convex on the set of Lagrange multipliers $\boldsymbol{(\lambda, \mu, \upsilon)}$). As a result, to maximize the dual variables, a one-dimensional search strategy can be used. Nonetheless, the objective function (3.44) is not always differentiable; therefore, this gradient-based method is not always possible in all situations. The dual variables $(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\upsilon})$ are determined using the widely used sub-gradient approach (as shown below), with the sub-gradient directions being applied as follows:

$$\nabla_{\lambda_m}\Gamma = \beta_m^* \tag{3.52}$$

$$\nabla_{\mu_m}\Gamma = 1 - \beta_m^* \tag{3.53}$$

$$\nabla_{\mu_m}\Gamma = \eta\left(1 - \beta_m^*\right)\left(\sum_{i=1}^{G}\sum_{j=1}^{|S_j|}\|\bar{h}_{g,m}^H d_i\|^2 p_{i,j} + \sigma_v^2\right) - p_{g,m}^{req} \tag{3.54}$$

To that end, the value of $\boldsymbol{\lambda_m}$ decreases if the $\boldsymbol{\nabla_{\lambda_m}\Gamma > 0}$, the value of $\boldsymbol{\mu_m}$ decreases if the $\boldsymbol{\nabla_{\mu_m}\Gamma > 0}$, and the value of $\boldsymbol{\nu_m}$ decreases if the $\boldsymbol{\nabla_{\nu_m}\Gamma > 0}$. Based on this remark, we employ the binary search method (Zhang *et al.*, 2009) with an error tolerance $\boldsymbol{\epsilon_2}$ to identify the best Lagrange multipliers $(\boldsymbol{\lambda^*}, \boldsymbol{\mu^*}, \boldsymbol{\upsilon^*})$ for the particular scenario. Thus, the algorithms developed in steps 1 and 2 operate alternately until the duality gap no longer changes, that is,

$$|R_{sum}\left(\beta^*\right) - \Gamma\left(\lambda^*,\ \mu^*, \upsilon^*\right)| = Const \tag{3.55}$$

Where $\boldsymbol{Const}$ represents a non-negative constant value. Second, we optimize the power allocation with a fixed power splitting factor in the optimization problem equations (3.22)–(3.26). We aim to find the power allocation $\boldsymbol{p_{g,m}}$ $\forall$ $\boldsymbol{g, m}$ under the optimized power splitting factor $\boldsymbol{\beta^*}$. However, we can rewrite the optimization problem in equations (3.22)–(3.26) as follows:

$$\max_{p_{g,m}} R_{sum}\left(p_m\right) \tag{3.56}$$

$$s.t. \quad \Sigma_{g=1}^{G}\Sigma_{m=1}^{|S_g|}P_{g,m} \leq P_T \tag{3.57}$$

$$p_{g,m} \geq 0 \quad \forall\, g, m \tag{3.58}$$

$$P_{g,m}^{EH} \geq p_{g,m}^{req} \tag{3.59}$$

***Proposition*2 :** Assume that the process of power splitting in the receiver is almost idealized, and the noise power for all users in the gth beam is equal, that is, $|\sigma_u|^2 \rightarrow 0$. In (56)–(59), the convexity of the sub-optimization issue is determined by whether or not the feasible domain is empty.

***Proof* :** It should be noted that the feasible domain of the sub-problems (56)–(59) is assumed to be non-empty and its convexity can be easily deduced from the constraints in (57)–(59). Next, we will examine the concavity of the objective function (56) in relation to the power allocation $p_{g,m}\ \forall\ g, m$. Based on the assumption above, the objective function can be written as

$$R_{sum}\left(p_{g,m}\right) = \sum_{g=1}^{G} R_g \tag{3.60}$$

$$R_{sum}\left(p_{g,m}\right) = \sum_{g=1}^{G}\sum_{m=1}^{|S_g|} \log_2\left(1 + \frac{\|\bar{h}_{g,m}^H d_g\|_2^2 p_{g,m}}{\left(\|\bar{h}_{g,m}^H d_g\|^2 |\Sigma_{j=1}^{m-1} p_{g,j} + \Sigma_{i\neq g}\|\bar{h}_{g,m}^H d_i\|_2^2 \Sigma_{j=1}^{|S_i|} p_{i,j} + \sigma_v^2\right)}\right) \tag{3.61}$$

Where

$$R_g = \sum_{m=1}^{|S_g|} \log_2\left(1 + \frac{\|\bar{h}_{g,m}^H d_g\|_2^2 p_{g,m}}{\left(\|\bar{h}_{g,m}^H d_g\|_2^2 |\Sigma_{j=1}^{m-1} p_{g,j} + \Sigma_{i\neq g}\|\bar{h}_{g,m}^H d_i\|_2^2 \Sigma_{j=1}^{|S_i|} p_{i,j} + \sigma_v^2\right)}\right) \tag{3.62}$$

Let us define the relationship between the m-th user and its decoding order as $m = \psi(m)$. Because the process of power splitting in the receiver is almost idealized and the noise power for all users in the gth beam is equal, the objective function can be rewritten as follows:

$$R_g = \sum_{m=1}^{|S_g|} \log_2\left(1 + \frac{\|\bar{h}_{g,\psi(m)}^H d_g\|_2^2 p_{g,\psi(m)}}{\left(\|\bar{h}_{g,\psi(m)}^H d_g\|_2^2 |\Sigma_{j=m+1}^{|S_g|} p_{g,\psi(j)} + \sigma_v^2\right)}\right) \tag{3.63}$$

$$R_g = \sum_{m=1}^{|S_g|} \log_2\left(\frac{\|\bar{h}_{g,\psi(m)}^H d_g\|_2^2 \Theta_{g,m} + \sigma_v^2}{\|\bar{h}_{g,\psi(m)}^H d_g\|_2^2 \Theta_{g,m+1} + \sigma_v^2}\right) \tag{3.64}$$

$$R_g = \sum_{m=1}^{|S_g|} \log_2\left(\|\bar{h}^H_{g,\psi(m)} d_g\|_2^2 \Theta_{g,m} + \sigma_v^2\right) - \log_2\left(\|\bar{h}^H_{g,\psi(m)} d_g\|_2^2 \Theta_{g,m+1} + \sigma_v^2\right) \tag{3.65}$$

where $\Theta_{g,m} = \sum_{j=m}^{|S_g|} p_{g,\psi(j)}$ and $\Theta_{g,m+1} = \sum_{j=m+1}^{|S_g|} p_{g,\psi(j)}$. Now, one can find the first derivative of R_g with respect to $p_{g,\psi(m)}$ as follows:

$$\frac{\partial R_g}{\partial p_{g,\psi(m)}} = \frac{1}{ln2} \cdot \left(\frac{\|\bar{h}^H_{g,\psi(1)} d_g\|_2^2}{\|\bar{h}^H_{g,\psi(1)} d_g\|_2^2 \Theta_{g,1} + \sigma_v^2}\right) \quad \forall m = 1 \tag{3.66}$$

And

$$\begin{aligned}\frac{\partial R_g}{\partial p_{g,\psi(m)}} =& \frac{1}{\ln 2} \cdot \frac{\|\bar{h}^H_{g,\psi(1)} d_g\|_2^2}{\|\bar{h}^H_{g,\psi(1)} d_g\|_2^2 \Theta_{g,1} + \sigma_v^2} \\ &+ \sum_{l=2}^{m}\left(\frac{\|\bar{h}^H_{g,\psi(l)} d_g\|_2^2}{\|\bar{h}^H_{g,\psi(l)} d_g\|_2^2 \Theta_{g,l} + \sigma_v^2} - \frac{\|\bar{h}^H_{g,\psi(l-1)} d_g\|_2^2}{\|\bar{h}^H_{g,\psi(l-1)} d_g\|_2^2 \Theta_{g,l} + \sigma_v^2}\right) \\ &\forall \qquad 2 \le m \le |S_g|\end{aligned} \tag{3.67}$$

Moreover, the second derivative of R_g with respect to $p_{g,\psi(m)}$ is given by

$$\begin{aligned}\frac{\partial^2 R_g}{\partial p_{g,\psi(m)} \partial p_{g,\psi(n)}} =& -\frac{1}{\ln 2} \cdot \frac{\|\bar{h}^H_{g,\psi(1)} d_g\|_2^4}{\left(\|\bar{h}^H_{g,\psi(1)} d_g\|_2^2 \Theta_{g,1} + \sigma_v^2\right)^2} \\ &- \frac{1}{\ln 2} \cdot \sum_{l=2}^{m}\left(\frac{\|\bar{h}^H_{g,\psi(1)} d_g\|_2^4}{\left(\|\bar{h}^H_{g,\psi(1)} d_g\|_2^2 \Theta_{g,1} + \sigma_v^2\right)^2} - \frac{\|\bar{h}^H_{g,\psi(1)} d_g\|_2^4}{\left(\|\bar{h}^H_{g,\psi(1)} d_g\|_2^2 \Theta_{g,1} + \sigma_v^2\right)^2}\right) \\ &\forall \qquad m\end{aligned} \tag{3.68}$$

According to (3.68), it can easily be inferred that the Hessian matrix of R_g with respect to $p_{g,m}$ $\forall$ g, m is negative or equal to zero. Consequently, the R_g is concave with respect to $p_{g,m}$. Therefore, because the sum of a finite number of concave functions stays concave, the objective function in (3.56) is concave on the power allocations $p_{g,m}$ $\forall$ g, m. Additionally, this study uses the Lagrangian duality-based method to obtain the near-optimal power allocation (Boyd *et al.*, 2004). The corresponding Lagrangian function for the sub-problem in (3.56)–(3.59) is formulated as follows:

$$\begin{aligned}\bar{\Upsilon}\left(p,\alpha,\eta,\kappa\right) = &\sum_{g=1}^{G}\sum_{m=1}^{|S_g|}\log_2\left(1+\frac{\|\bar{h}^{H}_{g,\psi(m)}d_g\|_2^2 p_{g,\psi(m)}}{\left(\|\bar{h}^{H}_{g,\psi(m)}d_g\|_2^2|\Sigma_{j=m+1}^{|S_g|}p_{g,\psi(j)}+\sigma_v^2\right)}\right)\\ &+\sum_{g=1}^{G}\sum_{m=1}^{|S_g|}\alpha_{g,m}p_{g,\psi(m)}+\sum_{m=1}^{|S_g|}\eta_m\left(P_T-\sum_{g=1}^{G}p_{g,\psi(m)}\right)\\ &+\sum_{m=1}^{|S_g|}\kappa_m\left(\eta(1-\beta_m^*)\left(\sum_{i=1}^{G}\sum_{j=1}^{|S_j|}\|\bar{h}^{H}_{g,\psi(j)}d_{\psi(j)}\|_2^2 p_{i,\psi(j)}+\sigma_v^2\right)-p^{req}_{g,\psi(m)}\right)\end{aligned} \tag{3.69}$$

where $\alpha = \left[\alpha_1, \alpha_2, \cdots, \alpha_{|S_g|}\right]^T$, $\eta = \left[\eta_1,\ \eta_2, \cdots,\ \eta_{|S_g|}\right]^T$ and $\kappa = \left[\kappa_1,\ \kappa_2, \cdots,\ \kappa_{|S_g|}\right]^T$ are non-negative Lagrange multipliers that correspond to the constraints in (3.57), (3.58), and (3.59), respectively. Notably, $\alpha_n = \left[\alpha_{n,1}, \alpha_{n,2}, \cdots, \alpha_{n,|S_g|}\right]^T$ is a non-negative Lagrange multiplier corresponding to constraint (3.57). Accordingly, one can express the Lagrange dual objective function as follows

$$\Gamma\left(\alpha,\eta,\kappa\right) = \max_{p}\Upsilon\left(p,\alpha,\eta,\kappa\right) \tag{3.70}$$

Then, one can model the Lagrange dual optimization problem as follows

$$\min_{\alpha,\eta,\kappa}\Gamma\left(\alpha,\eta,\kappa\right) \tag{3.71}$$

$$s.t. \qquad \alpha \succcurlyeq 0,\ \eta \succcurlyeq 0, \kappa \succcurlyeq 0 \tag{3.72}$$

The proposed algorithm to solve the corresponding optimization problems consists of the following two steps: First, we employed the gradient ascent method to determine the optimal power allocation p^*. The gradient direction of the Lagrangian function with respect to the power allocation is given as

$$
\begin{aligned}
\nabla_{p_{g,\psi(m)}}\bar{\Upsilon} = \frac{1}{\ln 2}\cdot&\left(\frac{\|\bar{h}^H_{g,\psi(1)}d_g\|_2^2}{\left(\|\bar{h}^H_{g,\psi(1)}d_g\|_2^2|\Theta_{g,1}+\sigma_v^2\right)}\right)\\
&+\sum_{l=2}^{m}\left(\frac{\|\bar{h}^H_{g,\psi(l)}d_g\|_2^2}{\left(\|\bar{h}^H_{g,\psi(l)}d_g\|_2^2|\Theta_{g,l}+\sigma_v^2\right)}-\frac{\|\bar{h}^H_{g,\psi(l-1)}d_g\|_2^2}{\left(\|\bar{h}^H_{g,\psi(l-1)}d_g\|_2^2|\Theta_{g,l}+\sigma_v^2\right)}\right)\\
&+\alpha_{g,m}-\eta_m\left(\sum_{j=1}^{|S_g|}\kappa_j\eta(1-\beta_j^*)\|\bar{h}^H_{g,\psi(j)}d_g\|_2^2\right)
\end{aligned}
\tag{3.73}
$$

In particular, the power allocation for each user on the g-th beam $(1 \le g \le G)$ can be sequentially updated using the following expressions:

$$
p_{g,\psi(m)}\left(Itr\ +1\right) = p_{g,\psi(m)}\left(Itr\right)\ +\bar{\varepsilon}\left(Itr\right)\nabla_{p_{g,\psi(m)}(Itr)\bar{\Upsilon}} \tag{3.74}
$$

where $p_{g,\psi(m)}\left(Itr\ \right)$ and $p_{g,\psi(m)}\left(Itr\ +1\right)$ represents the $p_{g,\psi(m)}$ in the Itr-th and $(Itr\ +1)$-th iterations, respectively. $\bar{\varepsilon}\left(Itr\right)$ defines the updated step size for the $p_{g,\psi(m)}$ in the Itr-th iteration and it satisfies the condition $\left|\nabla_{p_{g,\psi(m)}}\Upsilon\right| \le \epsilon_3\ \forall\ 1 \le$. The updated process in (3.73) and (3.74) for the power allocation on the g-th beam is repeated until $\left|\nabla_{p_{g,\psi(m)}}\Upsilon\right| \le \epsilon_3\ \forall\ 1 \le m \le \left|S_g\right|$And the optimal power allocation is denoted as p^*. Therefore, the Lagrange dual-objective function in (3.70) is given by

$$
\Gamma\left(\lambda,\mu,\upsilon\right) = \Upsilon\left(p^*,\alpha,\eta,\kappa\right) \tag{3.75}
$$

Next, we can update and determine the optimal Lagrange multipliers(α,η,κ) by solving the Lagrange dual optimization problem in (3.71)–(3.72) as follows:

$$
\min_{\lambda,\mu,\upsilon}\Gamma\left(\alpha,\eta,\kappa\right) \tag{3.76}
$$

$$
s.t. \qquad \alpha \succcurlyeq 0,\ \eta\ \succcurlyeq 0, \kappa\ \succcurlyeq 0 \tag{3.77}
$$

We utilize the commonly used sub-gradient technique to find the dual variables $(\alpha,\ \eta, \kappa)$, for which the sub-gradient directions are applied in the following ways:

$$\nabla_{\alpha_m}\Gamma = p_{g,\psi(m)} \tag{3.78}$$

$$\nabla_{\eta_m}\Gamma = P_T - \sum_{m=1}^{G} p_{g,\psi(m)} \tag{3.79}$$

$$\nabla_{\mu_m}\Gamma = \eta\left(1-\beta_m^*\right)\left(\sum_{i=1}^{G}\sum_{j=1}^{|S_j|}\|\bar{h}_{g,\psi(j)}^{H} d_{\psi(j)}\|^2 p_{i,\psi(j)} + \sigma_v^2\right) - p_{g,\psi(m)}^{req} \tag{3.80}$$

In this chapter, we apply the binary search technique with error tolerance ϵ_4 to find the optimal solution of the Lagrange multipliers in their many forms $(\alpha^*,\ \eta^*, \kappa^*)$. As a result, the proposed algorithm runs alternatively until the duality gap no longer changes, that is,

$$\left|R_{sum}\left(p^*\right) - \bar{\Gamma}\left(\alpha^*,\ \eta^*, \kappa^*\right)\right| = Const \tag{3.81}$$

Where Const represents a non-negative constant value. To that end, we have developed a solution to the sub-problems to optimize the power allocation and power splitting factor. Nevertheless, the algorithm developed for the joint optimization problem in (3.22)–(3.26) is presented in Algorithm (3.4). The computational complexity of the developed method is given as

$$O\left(G\left|S_g\right|^2 \log\left(\frac{1}{\epsilon_1{}^2}\right)\log\left(\frac{1}{\epsilon_2{}^2}\right)\log\left(\frac{1}{\epsilon_3{}^2}\right)\log\left(\frac{1}{\epsilon_4{}^2}\right)\right) \tag{3.82}$$

3.6 Simulation Results

Spectral efficiency is defined as the sum rate attained when operating within a given spectrum (3.17). In contrast, energy efficiency refers to the ratio between the sum rate obtained and the

Algorithm 3.4 Proposed Method for mmWave Massive MIMO-NOMA systems with SWIPT

Input:
Channel vectors: $\boldsymbol{h}_{g,m} \quad \forall\, g, m$
Digital precoding vectors: $\boldsymbol{d_g} \quad \forall\, g$
Noise variance: σ_v^2
Maximum iteration times: Itr_{max}
Output:
Optimal power allocation: $\boldsymbol{p}^* = p^*_{g,m} \quad \forall\, g, m$
Optimal power splitting factors: $\boldsymbol{\beta}^* = \beta_m{}^* \quad \forall m$

Initialize $\boldsymbol{p}$ and stop criteria $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$
repeat
    repeat
        Step 1: Optimize the power splitting factors $\beta_m \quad \forall m$ under fixed power allocation p
        repeat
            Initialize dual variables: $(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\upsilon})$
            Solve the problem in (3.27) to obtain the optimal power splitting factors $\boldsymbol{\beta}^*$ according to (3.46)-(3.48).
        until $\left|\nabla_{\beta_m(Itr)}\, \Upsilon\right| \leq \epsilon_1 \;\forall\, m$;
        Determine the optimal Lagrange dual multipliers $(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\upsilon})$ according to (3.52)-(3.53).
    until $|R_{sum}\,(\boldsymbol{\beta}^*) - \Gamma\,(\lambda^*,\ \mu^*, \upsilon^*)| = Const$;
    repeat
        Step 2: Optimize the power allocation with fixed the power splitting factors $\boldsymbol{\beta}^*$
        repeat
            Initialize the power splitting factors $\boldsymbol{\beta}^*$ to Solve the problem in (3.56) to obtain the optimal power allocation p^* according to (3.70)-(3.74).
        until $\left|\nabla_{p_{g,\psi(m)}}\, \Upsilon\right| \leq \epsilon_3 \;\forall\, 1 \leq m \leq \left|S_g\right|$;
        Determine the optimal Lagrange dual multipliers $(\boldsymbol{\alpha^*}, \boldsymbol{\eta^*}, \boldsymbol{\kappa^*})$ according to (3.78)-(3.80).
    until $\left|R_{sum}\,(p^*) - \bar{\Gamma}\,(\alpha^*,\ \eta^*, \kappa^*)\right| = Const$;
until $R_{sum}\,(p^*) = R_{sum}\,(\beta^*)$;

total power consumed (Huang *et al.*, 2019) i.e .

$$EE = \frac{Achievable\ sum\ rate}{Total\ power\ consumption} \tag{3.83}$$

$$EE = \frac{R_{Sum}}{P_t + N_{RF}P_{RF} + N_{phaseshift}P_{phaseshift} + P_{BB}} \tag{3.84}$$

where $P_t = \sum_{g=1}^{G} \sum_{m=1}^{|S_g|} p_{g,m}$ is the total transmitted power, P_{RF} is the power consumed by each RF chain, P_{BB} represents the baseband power consumption, and $P_{phaseshift}$ is the power consumption of each phase shift. In particular, $P_{RF} = 300\ mW$, $P_{phaseshift} = 40\ mW\ \forall\ B = 4$ bit phase shifter, and $P_{BB} = 200\ mW$ are adopted as the typical values. In addition, $N_{phaseshift}$ is the number of phase shifters and is equal to NN_{RF} for hybrid precoding. Moreover, all presented results are averaged over 100 random channel realizations. To demonstrate the performance of the proposed technique, we present the simulation results to illustrate both the spectrum efficiency and energy efficiency of the hybrid precoding architecture. The following parameters are provided for the simulation: the system's bandwidth is defined as $1Hz$, corresponding to a rate as high as possible (3.15). The BS and UE are equipped with uniform linear antennas (ULAs) with half-wavelength spacing. The BS is equipped with $N = 64$ antennas and $N_{RF} = 4$ RF chains, and can serve up to $K \geq N_{RF}$ UEs simultaneously. All K UEs are clustered into $G = N, RF = 4$ beams, with each beam consisting of more than one user simultaneously. According to equation (3.5), a channel(3.5) vector for the mth user in the gth beam is created by considering one line-of-sight (LoS) components as well as two non-line-of-sight (NLoS) components, that is, the number of routes that the mth user takes in the gth beam $(L_{g,m} = 3)$.The complex gain of the LoS path is $\alpha_{g,m}^{(1)}/simCN\,(0, 1)$ and the complex gains of the NLoS paths are $\alpha_{g,m}^{(l)}/simCN\,(0,\ 0.1)\,\forall 2 \leq l \leq L_{g,m}$. The azimuth angle of departure (AoD) is $\vartheta_{g,m}^{(l)}$ and elevation angle of departure $\theta_{g,m}^{(l)}$ of the lth path is assumed to follow the uniform distribution $\mathcal{U}\,(-\pi, \pi)\ \forall 1 \quad \leq l \leq L_{g,m}$. The bit resolution $B = 4$ is used to quantize the phase shifters. The SNR is defined as the ratio of signal to noise $\left(p_t/\sigma^2\right)$, where the maximum transmitted power $p_t = 30mW$, the minimal achievable rate for each user, is $R_{g,m}^{min}/10$, where $R_{g,m}^{min}$ is the lowest possible rate among all users when completely digital ZF precoding is used, and the lowest amount of energy collected by each user is $p_{g,m}^{min} = 0.1\ mW$.

In the simulations, we consider the proposed method with the following four methods of mmWave massive MIMO systems with SWIPT for comparison: (1) "SWIPT-Fully digital ZF Precoding," (2) "SWIPT-Hybrid Precoding NOMA proposed in (Dai *et al.*, 2018)," (3) "SWIPT-Hybrid Precoding NOMA proposed in (Zhao *et al.*, 2019)," and (4) "SWIPT-Hybrid Precoding OMA,"

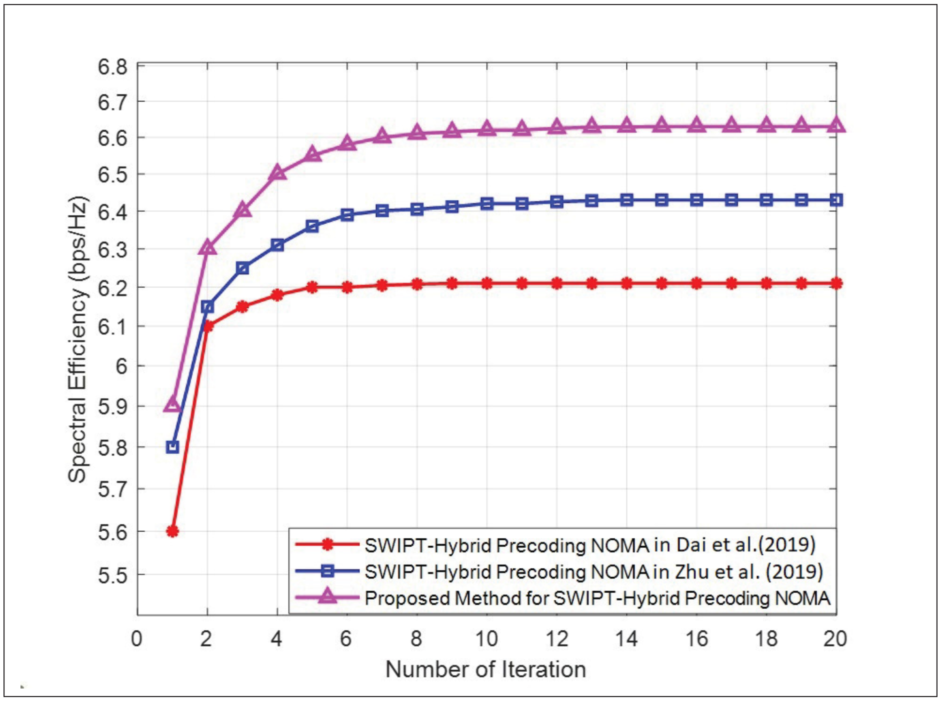


Figure 3.4 Spectrum efficiency of HP system versus the number of iterations for the joint power allocation and power splitting optimization.

where OMA is implemented for UEs in each beam. An Intel Core $i5 - 2400S@1.6GHz$ (4 cores) and $8GB$ of RAM were used to run the simulations. Figure (3.4) shows the spectrum efficiency as a function of the number of iterations, where the number of users is fixed at K = 6, and the SNR is set to 0 dB. The curves depicted in Figure (3.4) illustrate the convergence of the proposed method described in Section IV, which addresses the problem of joint power allocation and power splitting for systems with fixed K users. From Figure (3.4), the spectrum efficiency appears to have stabilized after the proposed method in Section IV has been iterated 13 times, which demonstrates the convergence of the proposed method. However, our proposed techniques require approximately 13 iterations for the combined power allocation and power splitting optimization to converge, whereas the SWIPT-Hybrid Precoding NOMA described in (Dai *et al.*, 2018) requires approximately nine iterations for convergence. The SWIPT-Hybrid Precoding NOMA described in (Dai *et al.*, 2018) converges to a greater spectrum efficiency than our proposed method. According to the SWIPT-Hybrid Precoding NOMA described in (Zhao *et al.*, 2019), the joint power allocation and power splitting optimization require approximately 12 iterations to converge. Therefore, to guarantee that each scheme can remain stable during the simulations, the number of iterations for the power allocation and power slitting optimization is set to 14.

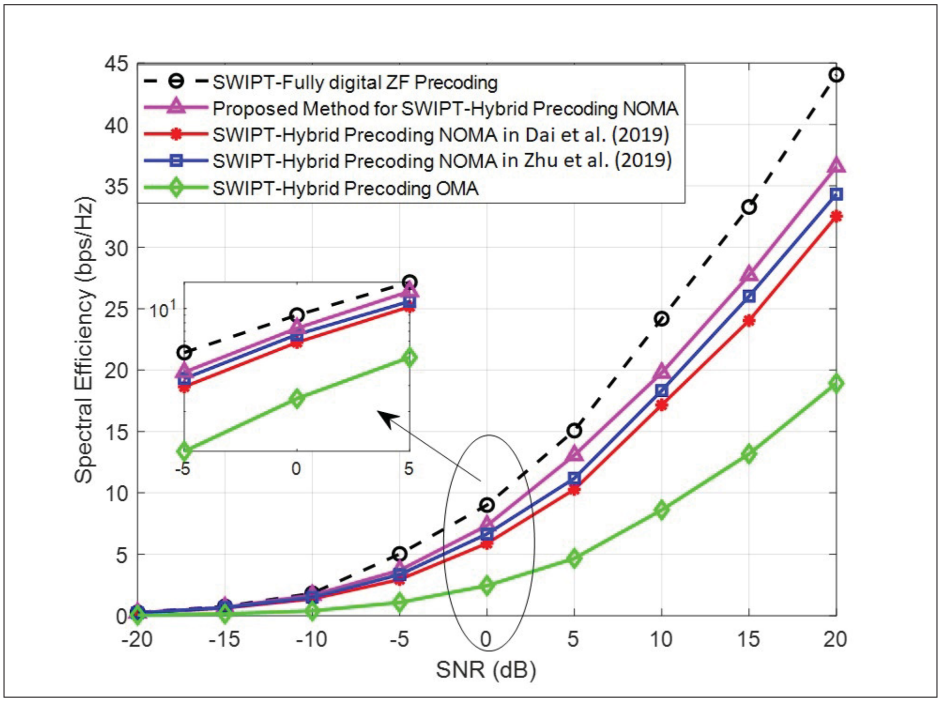


Figure 3.5 Spectrum efficiency against SNR.

The spectrum efficiency of the system is illustrated in Figure (3.5) We consider the spectrum efficiency, SNR, and the number of users to determine which of the four signal-processing methods offers the best tradeoff between performance and cost. Because of NOMA's greater spectrum efficiency, we can say that the mmWave massive MIMO-NOMA systems with SWIPT that we proposed can provide better spectrum efficiency than that of mmWave massive MIMO-OMA systems with SWIPT. As can be seen in Figure (3.5), the spectrum efficiency increases as the SNR increases for all the methods being examined. SWIPT-Full-digital ZF Precoding performs better in increasing the overall spectral efficiency compared to all the precoding schemes, but it requires more processing than other methods.

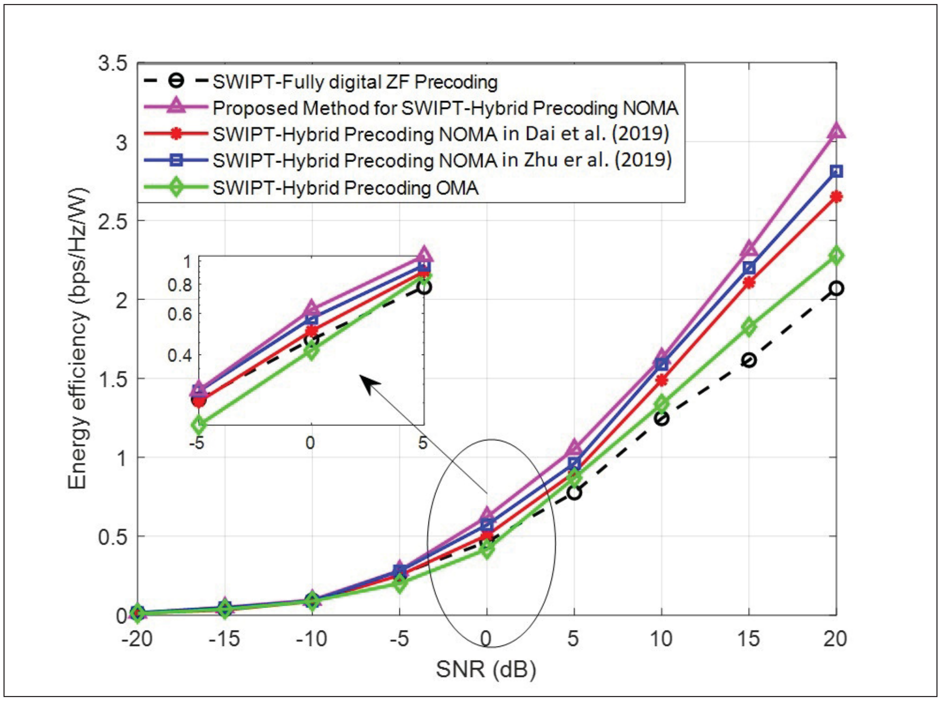


Figure 3.6   Energy efficiency against SNR.

Figure (3.6) depicts the SNR-adjusted energy efficiency, which can accommodate up to six users. According to our findings in Figure (3.6), the proposed mmWave massive MIMO-NOMA systems with SWIPT achieved greater energy efficiency than both mmWave massive MIMO-OMA systems with SWIPT and completely digital MIMO systems with SWIPT. With RF chains, as in fully digital MIMO systems, each RF chain needs 300 mW of power. Contrary to this statement, with SWIPT-Hybrid Precoding NOMA systems, the number of RF chains is significantly lower than the number of antennas. Therefore, compared to completely digital MIMO systems, RF chains generate much less energy. Furthermore, the SWIPT-enabled mmWave massive MIMO-NOMA system with hybrid precoding is shown to perform better than current systems in moderate to high SNR regimes because of the usage of NOMA.

Figure (3.7) depicts a comparison of the spectrum efficiency vs. the number of UEs for all five schemes under discussion, with the SNR fixed at 10 dB for all five schemes considered. As shown in Figure (3.7), the efficiency of the spectrum increases for all curves as the number of UEs increases. In this case, several UEs can share the same time-frequency resource block

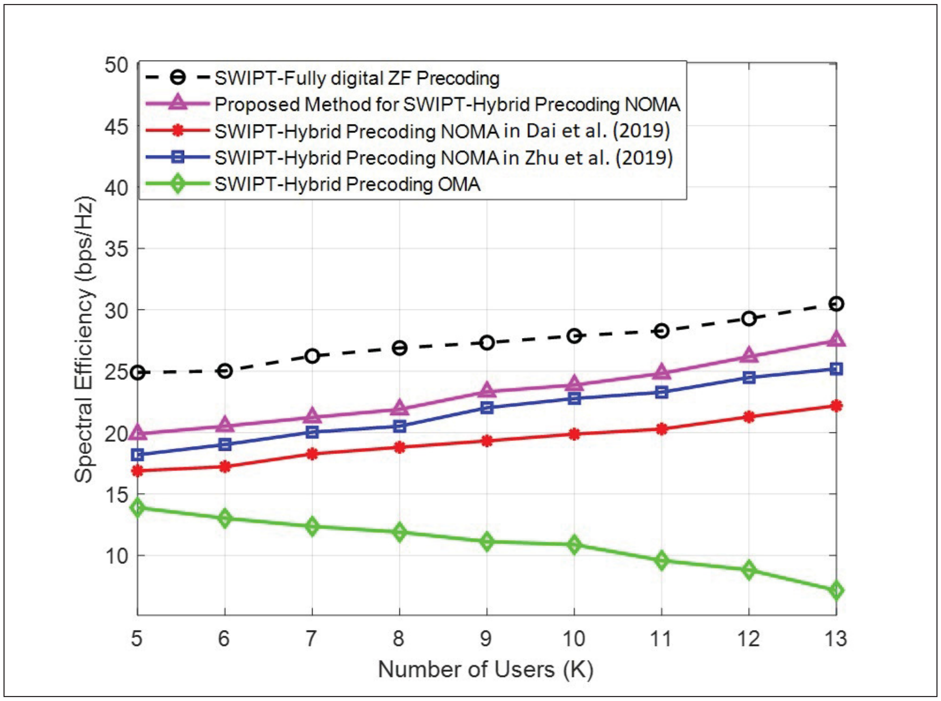


Figure 3.7 Spectrum efficiency of hybrid precoding system versus the number of users for the joint power allocation and power splitting optimization.

by utilizing intra-beam superposition coding at the base station and SIC at the receiver, which allows for greater efficiency. The proposed SWIPT-enabled mmWave huge MIMO-NOMA systems with hybrid precoding, on the other hand, outperform the other methods and achieve performance that is comparable to the SWIPT-Full-digital ZF Precoding. As a result, using the suggested user grouping, analog RF precoder and digital baseband precoder design methods are helpful for interbeam interference cancellation while also enhancing the overall system performance and efficiency. The energy efficiency versus the number of users is shown in Figure (3.8). The SNR was adjusted to 10 dB. For illustration, Figure (3.8) depicts several curves with various degrees of curvature. The energy efficiency of the SWIPT-Hybrid Precoding OMA system decreases with an increasing number of UEs. Another important observation is that the energy efficiency of the SWIPT-Full-digital ZF Precoding scheme increases with the number of UEs. Moreover, we have also noticed that the SWIPT-enabled mmWave mMIMO-NOMA

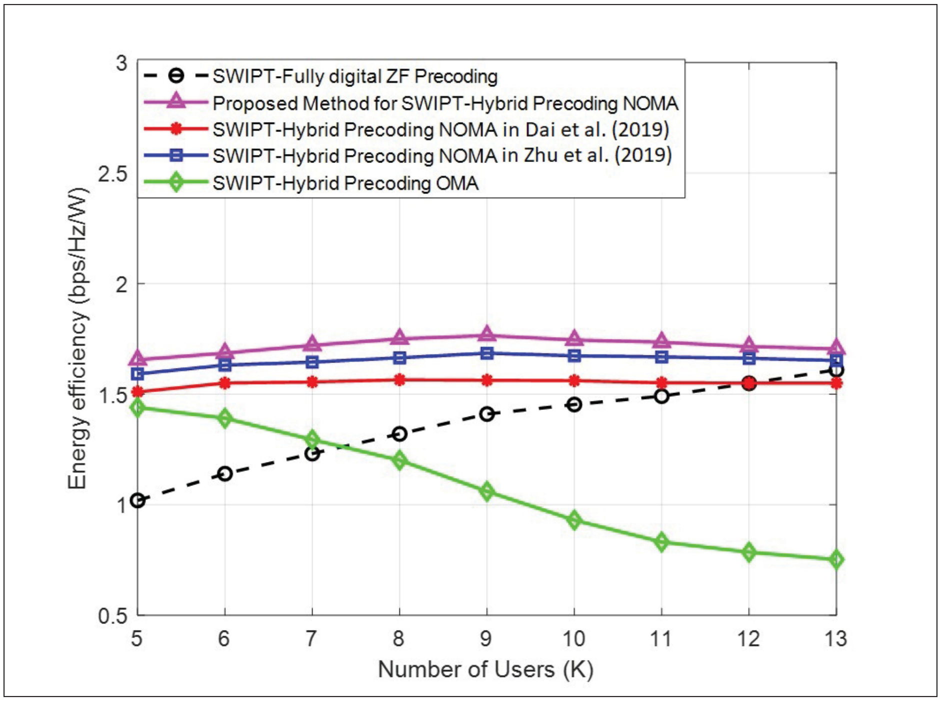


Figure 3.8 Energy efficiency of hybrid precoding system versus the number of users for the joint power allocation and power splitting optimization.

system with SWIPT MMIMO-NOMA capability shows superior energy efficiency at a low and medium number of users. It increases efficiency as we go up with the number of users.

3.7 Conclusion

In this chapter , hybrid precoding for SWIPT-enabled mmWave mMIMO-NOMA systems to enhance the attainable sum rate and total energy efficiency. The optimization of user grouping is given first, followed by the creation of hybrid analog-digital precoders. Then, given the maximum transmit power budget restrictions and minimal EH need, we examined the feasible data rate maximization problem for SWIPT-enabled mmWave mMIMO-NOMA systems with PS receivers. Because of the coupling of many variables and the presence of inter-user interference, the maximization issue was non-convex, making it difficult to obtain the best solution directly. We used a decoupled strategy to solve this problem, in which the linked variables, such as power allocation and PS ratio assignment, were separated. The Lagrangian duality-based technique was then used to solve the associated subproblems. The proposed technique with hybrid precoding considerably increased the spectrum efficiency and energy efficiency of the studied system compared to existing state-of-the-art systems, demonstrating its efficacy. Furthermore, mmWave MIMO-NOMA continues to outperform mmWave MIMO-OMA.

CONCLUSION AND RECOMMENDATIONS

Within the scope of this thesis, we tackled the problem of computational complexity in Massive MIMO uplink. Due to the approximation LLR calculation employed in our detection approach, it is very robust to variations in channel correlation and loading factor. Our numerical findings demonstrate that the performance of the suggested detection method soon converges to that of an accurate detection technique for relatively large base station-to-user antenna ratios. Thus, the proposed approach may provide results that are competitive with those obtained using an accurate inversion method, while often requiring less computer complexity. In addition, the efficacy and simplicity of the proposed method surpass those of the approximation Neumann series inversion and other recommended systems in the literature. Furthermore, our system is simpler. The proposed detector works well and may be used with large MIMO systems using different antenna setups. Next, we looked at the topic of 5G communication systems' energy efficiency by developing, modeling, and analyzing new energy and spectral efficiency trade-offs in massive MIMO downlink. Critical to the design is the exploration of the compatibility of massive MIMO, NOMA, hybrid beamforming, and SWIPT. For mmWave mMIMO-NOMA systems with SWIPT support, hybrid precoding improves both the possible sum rate and the total energy efficiency. When user clustering has been optimized, hybrid analog-digital precoder construction is described. We next examined the viable data rate optimisation issue for SWIPT-enabled mmWave mMIMO-NOMA systems with PS receivers, taking into account the maximum transmit power budget limits and low EH demand. Due to the existence of inter-user interference and the coupling of several variables, the optimization problem was non-convex and hence difficult to solve directly. To address this issue, we used a decoupled strategy, which included untangling interdependent factors such as power distribution and PS ratio assignment. Subproblems were then solved using the Lagrangian duality method. Spectrum efficiency and energy efficiency were both greatly enhanced by the suggested approach with hybrid precoding as compared to state-of-the-art systems. Also, mmWave MIMO-NOMA continues to outperform mmWave

MIMO-OMA. In future research, we may apply the proposed model to circumstances with a larger number of users and larger antenna arrays. We may also include measurements utilizing software-defined radios for more realistic results. Furthermore, one of the most challenging difficulties with any proposed algorithm is determining whether or not it can work online. Real-time implementation is essential for assessing the efficacy of a particular approach. As a consequence, we will be expanding our new algorithms to enable for online deployment. The use of a mix of block-based and real-time techniques, such as a block LMS type structure ((Mumtaz *et al.*, 2016; Zhu *et al.*, 2019a; Aljumaily & Li, 2019)) is a new strategy for implementing the algorithm. Furthermore, the design and analysis of extreme Massive MIMO models for 6G wireless networks will be the key focus of our future research. The analysis and conclusions of our dissertation imply that the design and analysis of massive MIMO systems employing SWIPT through the development of energy- and spectrum-efficient hybrid NOMA schemes is an exciting field for future study in 6G.

APPENDIX I

LIST OF PUBLICATIONS

We have presented the main content of this thesis in three publishing works which contain two categories as follows

1. Published Journals

- "Data detection method for uplink massive MIMO systems based on the long recurrence enlarged conjugate gradient." which has bee published in International Journal of Electrical and Computer Engineering 12.4 (2022): 3911-3921 (Jawarn, Albataineh & Kadoch, 2022)
- "Decoupling energy efficient approach for hybrid precoding-based mmWave massive MIMO-NOMA with SWIPT." has bee published in IEEE Access 10 (2022): 28868-28884. (Jawarneh, Kadoch & Albataineh, 2022)

2. Published Confrences

- "Iterative signal detection based on LRE-CG method for uplink massive MIMO systems." has been published in International Wireless Communications and Mobile Computing (IWCMC). IEEE, 2021.(Jawarneh & Kadoch, 2021)

BIBLIOGRAPHY

3rd Generation Partnership Project. (2014). *Requirements for further advancements for) (Release Evolved Universal Terrestrial Radio Access (E-UTRA) (LTE-Advanced 12).*

Alabed, S. (2018). Performance analysis of differential beamforming in decentralized networks. *Int. J. Electr. Comput. Eng.*, 8(3), 1692–1700.

Albataineh, Z. (2018). Robust blind channel estimation algorithm for linear STBC systems using fourth order cumulant matrices. *Telecommunication Systems*, 68(3), 573–582.

Albataineh, Z. (2019). Iterative signal detection based on MSD-CG method for uplink massive MIMO systems. *2019 16th International Multi-Conference on Systems, Signals & Devices (SSD)*, pp. 539–544.

Albataineh, Z. (2021). Low-complexity near-optimal iterative signal detection based on MSD-CG method for uplink massive MIMO systems. *Wireless Personal Communications*, 116(3), 2549–2563.

Albataineh, Z. & Salem, F. M. (2017). Adaptive blind CDMA receivers based on ICA filtered structures. *Circuits, Systems, and Signal Processing*, 36(8), 3320–3348.

Aljumaily, M. S. & Li, H. (2019). Machine learning aided hybrid beamforming in massive-MIMO millimeter wave systems. *2019 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, pp. 1–6.

Alkhateeb, A., Leus, G. & Heath, R. W. (2015). Limited feedback hybrid precoding for multi-user millimeter wave systems. *IEEE transactions on wireless communications*, 14(11), 6481–6494.

Almeida, V. A., Doneda, D. & Monteiro, M. (2015). Governance challenges for the Internet of Things. *IEEE Internet Computing*, 19(4), 56–59.

Alnabelsi, S. H., Salameh, H. A. B. & Albataineh, Z. M. (2020). Dynamic resource allocation for opportunistic software-defined IoT networks: stochastic optimization framework. *International Journal of Electrical and Computer Engineering*, 10(4), 3854.

Alwakeel, A. S. & Mehana, A. H. (2019). Multi-cell MMSE data detection for massive MIMO: new simplified bounds. *IET Communications*, 13(15), 2386–2394.

Andrews, J. G. (2005). Interference cancellation for cellular systems: a contemporary overview. *IEEE Wireless Communications*, 12(2), 19–29.

Andrews, J. G., Buzzi, S., Choi, W., Hanly, S. V., Lozano, A., Soong, A. C. & Zhang, J. C. (2014). What will 5G be? *IEEE Journal on selected areas in communications*, 32(6), 1065–1082.

Belschner, J., Rakocevic, V. & Habermann, J. (2019). Complexity of coordinated beamforming and scheduling for OFDMA based heterogeneous networks. *Wireless Networks*, 25(5), 2233–2248.

Björnson, E., Hoydis, J. & Sanguinetti, L. (2017). Massive MIMO has unlimited capacity. *IEEE Transactions on Wireless Communications*, 17(1), 574–590.

Boyd, S., Boyd, S. P. & Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.

Chen, J.-C. (2017a). Hybrid beamforming with discrete phase shifters for millimeter-wave massive MIMO systems. *IEEE Transactions on Vehicular Technology*, 66(8), 7604–7608.

Chen, J.-C. (2017b). Manifold optimization approach for data detection in massive multiuser MIMO systems. *IEEE Transactions on Vehicular Technology*, 67(4), 3652–3657.

Chen, J.-C. (2019). Computationally Efficient Data Detection Algorithm for Massive MU-MIMO Systems Using PSK Modulations. *IEEE Communications Letters*, 23(6), 983–986.

Cheng, L., Yue, G., Yu, D., Liang, Y. & Li, S. (2019). Millimeter wave time-varying channel estimation via exploiting block-sparse and low-rank structures. *IEEE Access*, 7, 123355–123366.

Chi, C.-Y., Peng, C.-H., Huang, K.-C., Tsai, T.-H. & Ma, W.-K. (2008). A block-by-block blind post-FFT multistage beamforming algorithm for multiuser OFDM systems based on subcarrier averaging. *IEEE transactions on wireless communications*, 7(8), 3238–3251.

Choi, J., Lee, G. & Evans, B. L. (2019). User scheduling for millimeter wave hybrid beamforming systems with low-resolution ADCs. *IEEE Transactions on Wireless Communications*, 18(4), 2401–2414.

Cisco, R. (2022, Jan). Cisco annual internet Report - Cisco Annual Internet Report (2018–2023) White Paper. Cisco. Retrieved from: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html.

Clerckx, B., Zhang, R., Schober, R., Ng, D. W. K., Kim, D. I. & Poor, H. V. (2018). Fundamentals of wireless information and power transfer: From RF energy harvester models to signal and system designs. *IEEE Journal on Selected Areas in Communications*, 37(1), 4–33.

Dai, L., Gao, X., Su, X., Han, S., Chih-Lin, I. & Wang, Z. (2014). Low-complexity soft-output signal detection based on Gauss–Seidel method for uplink multiuser large-scale MIMO systems. *IEEE Transactions on Vehicular Technology*, 64(10), 4839–4845.

Dai, L., Wang, B., Yuan, Y., Han, S., Chih-Lin, I. & Wang, Z. (2015). Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends. *IEEE Communications Magazine*, 53(9), 74–81.

Dai, L., Wang, B., Peng, M. & Chen, S. (2018). Hybrid precoding-based millimeter-wave massive MIMO-NOMA with simultaneous wireless information and power transfer. *IEEE Journal on Selected Areas in Communications*, 37(1), 131–141.

Ding, Z., Zhong, C., Ng, D. W. K., Peng, M., Suraweera, H. A., Schober, R. & Poor, H. V. (2015). Application of smart antenna technologies in simultaneous wireless information and power transfer. *IEEE Communications Magazine*, 53(4), 86–93.

Doppler, K., Rinne, M., Wijting, C., Ribeiro, C. B. & Hugl, K. (2009). Device-to-device communication as an underlay to LTE-advanced networks. *IEEE communications magazine*, 47(12), 42–49.

El Ayach, O., Rajagopal, S., Abu-Surra, S., Pi, Z. & Heath, R. W. (2014). Spatially sparse precoding in millimeter wave MIMO systems. *IEEE transactions on wireless communications*, 13(3), 1499–1513.

Gao, X., Dai, L., Hu, Y., Wang, Z. & Wang, Z. (2014). Matrix inversion-less signal detection using SOR method for uplink large-scale MIMO systems. *2014 IEEE Global Communications Conference*, pp. 3291–3295.

Gao, X., Dai, L., Han, S., Chih-Lin, I. & Heath, R. W. (2016). Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays. *IEEE Journal on Selected Areas in Communications*, 34(4), 998–1009.

Garcia-Rodriguez, A., Venkateswaran, V., Rulikowski, P. & Masouros, C. (2016). Hybrid analog–digital precoding revisited under realistic RF modeling. *IEEE Wireless Communications Letters*, 5(5), 528–531.

Gelal, E., Ning, J., Pelechrinis, K., Kim, T.-S., Broustis, I., Krishnamurthy, S. V. & Rao, B. D. (2012). Topology control for effective interference cancellation in multiuser MIMO networks. *IEEE/ACM Transactions on Networking*, 21(2), 455–468.

Gesbert, D., Shafi, M., Shiu, D.-s., Smith, P. J. & Naguib, A. (2003). From theory to practice: An overview of MIMO space-time coded wireless systems. *IEEE Journal on selected areas in Communications*, 21(3), 281–302.

Goldsmith, A. (2005). *Wireless communications*. Cambridge university press.

Grigori, L., Moufawad, S. & Nataf, F. (2016). Enlarged Krylov subspace conjugate gradient methods for reducing communication. *SIAM Journal on Matrix Analysis and Applications*, 37(2), 744–773.

Gu, T., Liu, X., Mo, Z. & Chi, X. (2004). Multiple search direction conjugate gradient method I: methods and their propositions. *International Journal of Computer Mathematics*, 81(9), 1133–1143.

Gupta, A. & Jha, R. K. (2015). A survey of 5G network: Architecture and emerging technologies. *IEEE access*, 3, 1206–1232.

Hanchate, S. M. & Nema, S. (2018). Energy Efficient and Low Complexity Signal Detection with Generalized Spatial Modulation for Massive MIMO System. *2018 3rd International Conference for Convergence in Technology (I2CT)*, pp. 1–8.

Heath, R. W., Gonzalez-Prelcic, N., Rangan, S., Roh, W. & Sayeed, A. M. (2016). An overview of signal processing techniques for millimeter wave MIMO systems. *IEEE journal of selected topics in signal processing*, 10(3), 436–453.

Hemadeh, I. A., Satyanarayana, K., El-Hajjar, M. & Hanzo, L. (2017). Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget. *IEEE Communications Surveys & Tutorials*, 20(2), 870–913.

Hu, Y., Wang, Z., Gaol, X. & Ning, J. (2014). Low-complexity signal detection using CG method for uplink large-scale MIMO systems. *2014 IEEE international conference on communication systems*, pp. 477–481.

Huang, L., Lin, H., Zhang, H. & Zhao, Y. (2019). A Multi-Dimensional Features-based Clustering Algorithm for Massive MIMO System. *2019 14th International Conference on Computer Science & Education (ICCSE)*, pp. 423–427.

Humar, I., Ge, X., Xiang, L., Jo, M., Chen, M. & Zhang, J. (2011). Rethinking energy efficiency models of cellular networks with embodied energy. *IEEE network*, 25(2), 40–49.

Jaber, M., Imran, M. A., Tafazolli, R. & Tukmanov, A. (2016). 5G backhaul challenges and emerging research directions: A survey. *IEEE access*, 4, 1743–1766.

Jawarn, A., Albataineh, Z. & Kadoch, M. (2022). Data detection method for uplink massive MIMO systems based on the long recurrence enlarged conjugate gradient. *International Journal of Electrical & Computer Engineering (2088-8708)*, 12(4), 2222–2236.

Jawarneh, A. & Kadoch, M. (2021). Iterative signal detection based on LRE-CG method for uplink massive MIMO systems. *2021 International Wireless Communications and Mobile Computing (IWCMC)*, pp. 1798–1801.

Jawarneh, A., Kadoch, M. & Albataineh, Z. (2022). Decoupling energy efficient approach for hybrid precoding-based mmWave massive MIMO-NOMA with SWIPT. *IEEE Access*, 10, 28868–28884.

Jiang, F., Li, C., Gong, Z. & Su, R. (2018). Stair matrix and its applications to massive MIMO uplink data detection. *IEEE Transactions on Communications*, 66(6), 2437–2455.

Larsson, E. G., Edfors, O., Tufvesson, F. & Marzetta, T. L. (2014). Massive MIMO for next generation wireless systems. *IEEE communications magazine*, 52(2), 186–195.

Lee, Y. (2017). Decision-aided Jacobi iteration for signal detection in massive MIMO systems. *Electronics Letters*, 53(23), 1552–1554.

Lee, Y.-Y., Wang, C.-H. & Huang, Y.-H. (2014). A hybrid RF/baseband precoding processor based on parallel-index-selection matrix-inversion-bypass simultaneous orthogonal matching pursuit for millimeter wave MIMO systems. *IEEE Transactions on Signal Processing*, 63(2), 305–317.

Li, G. Y., Xu, Z., Xiong, C., Yang, C., Zhang, S., Chen, Y. & Xu, S. (2011). Energy-efficient wireless communications: tutorial, survey, and open issues. *IEEE Wireless communications*, 18(6), 28–35.

Liebe, H. J. (1989). MPM—An atmospheric millimeter-wave propagation model. *International Journal of Infrared and millimeter waves*, 10(6), 631–650.

Lin, X., Wu, S., Kuang, L., Ni, Z., Meng, X. & Jiang, C. (2017). Estimation of sparse massive MIMO-OFDM channels with approximately common support. *IEEE Communications Letters*, 21(5), 1179–1182.

Lu, L., Li, G. Y., Swindlehurst, A. L., Ashikhmin, A. & Zhang, R. (2014). An overview of massive MIMO: Benefits and challenges. *IEEE journal of selected topics in signal processing*, 8(5), 742–758.

Lu, Y., Xiong, K., Fan, P., Zhong, Z. & Letaief, K. B. (2018). Coordinated beamforming with artificial noise for secure SWIPT under non-linear EH model: Centralized and distributed designs. *IEEE Journal on Selected Areas in Communications*, 36(7), 1544–1563.

Lyu, S. & Ling, C. (2018). Hybrid vector perturbation precoding: The blessing of approximate message passing. *IEEE Transactions on Signal Processing*, 67(1), 178–193.

M. Elmagzoub, H. (2020). On the MMSE-based multiuser millimeter wave MIMO hybrid precoding design. *International Journal of Communication Systems*, 33(11), e4409.

Mahapatra, R., Nijsure, Y., Kaddoum, G., Hassan, N. U. & Yuen, C. (2015). Energy efficiency tradeoff mechanism towards wireless green communication: A survey. *IEEE Communications Surveys & Tutorials*, 18(1), 686–705.

Masoudi, M., Khafagy, M. G., Conte, A., El-Amine, A., Françoise, B., Nadjahi, C., Salem, F. E., Labidi, W., Süral, A., Gati, A. et al. (2019). Green mobile networks for 5G and beyond. *IEEE Access*, 7, 107270–107299.

Morelli, M., Kuo, C.-C. J. & Pun, M.-O. (2007). Synchronization techniques for orthogonal frequency division multiple access (OFDMA): A tutorial review. *Proceedings of the IEEE*, 95(7), 1394–1427.

Mumtaz, S., Rodriguez, J. & Dai, L. (2016). *MmWave massive MIMO: a paradigm for 5G*. Academic Press.

Ng, D. W. K., Lo, E. S. & Schober, R. (2013). Wireless information and power transfer: Energy efficiency optimization in OFDMA systems. *IEEE Transactions on Wireless Communications*, 12(12), 6352–6370.

Osseiran, A., Boccardi, F., Braun, V., Kusume, K., Marsch, P., Maternia, M., Queseth, O., Schellmann, M., Schotten, H., Taoka, H. et al. (2014). Scenarios for 5G mobile and wireless communications: the vision of the METIS project. *IEEE communications magazine*, 52(5), 26–35.

Pal, R., Srinivas, K. & Chaitanya, A. K. (2018). A beam selection algorithm for millimeter-wave multi-user MIMO systems. *IEEE Communications Letters*, 22(4), 852–855.

Patel, P. & Holtzman, J. (1994). Analysis of a simple successive interference cancellation scheme in a DS/CDMA system. *IEEE journal on selected areas in communications*, 12(5), 796–807.

Patrik Cerwall,Stephen Carson, P. J. (2021). *Ericsson mobility report*.

Perera, T. D. P., Jayakody, D. N. K., Sharma, S. K., Chatzinotas, S. & Li, J. (2017). Simultaneous wireless information and power transfer (SWIPT): Recent advances and future challenges. *IEEE Communications Surveys & Tutorials*, 20(1), 264–302.

Prasad, R. (2004). *OFDM for wireless communications systems*. Artech House.

Pyla, S., Raju, K. P. & Subrahmanyam, N. B. (2018). Performance analysis of adaptive filter channel estimated MIMO OFDM communication system. *International Journal of Electrical & Computer Engineering*, 8(5), 2088–8708.

Rappaport, T. S., Heath Jr, R. W., Daniels, R. C. & Murdock, J. N. (2015). *Millimeter wave wireless communications*. Pearson Education.

Raviteja, P., Hong, Y. & Viterbo, E. (2017). Analog beamforming with low resolution phase shifters. *IEEE Wireless Communications Letters*, 6(4), 502–505.

Roh, W., Seol, J.-Y., Park, J., Lee, B., Lee, J., Kim, Y., Cho, J., Cheun, K. & Aryanfar, F. (2014). Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results. *IEEE communications magazine*, 52(2), 106–113.

Rusek, F., Persson, D., Lau, B. K., Larsson, E. G., Marzetta, T. L., Edfors, O. & Tufvesson, F. (2012). Scaling up MIMO: Opportunities and challenges with very large arrays. *IEEE signal processing magazine*, 30(1), 40–60.

Senel, K., Cheng, H. V., Björnson, E. & Larsson, E. G. (2019). What role can NOMA play in massive MIMO? *IEEE Journal of Selected Topics in Signal Processing*, 13(3), 597–611.

Shi, Q., Liu, L., Xu, W. & Zhang, R. (2014). Joint transmit beamforming and receive power splitting for MISO SWIPT systems. *IEEE Transactions on Wireless Communications*, 13(6), 3269–3280.

Singh, J., Ponnuru, S. & Madhow, U. (2009). Multi-gigabit communication: The ADC bottleneck. *2009 IEEE International Conference on Ultra-Wideband*, pp. 22–27.

Sohrabi, F. & Yu, W. (2015). Hybrid beamforming with finite-resolution phase shifters for large-scale MIMO systems. *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 136–140.

Tang, J., Luo, J., Ou, J., Zhang, X., Zhao, N., So, D. K. C. & Wong, K.-K. (2020). Decoupling or learning: Joint power splitting and allocation in MC-NOMA with SWIPT. *IEEE Transactions on Communications*, 68(9), 5834–5848.

Uwaechia, A. N. & Mahyuddin, N. M. (2020). Spectrum and energy efficiency optimization for hybrid precoding-based SWIPT-enabled mmwave mMIMO-NOMA systems. *IEEE Access*, 8, 139994–140007.

Uwaechia, A. N., Mahyuddin, N. M., Ain, M. F., Latiff, N. M. A. & Za'bah, N. F. (2019). On the spectral-efficiency of low-complexity and resolution hybrid precoding and combining transceivers for mmWave MIMO systems. *IEEE Access*, 7, 109259–109277.

Vaezi, M., Schober, R., Ding, Z. & Poor, H. V. (2019). Non-orthogonal multiple access: Common myths and critical questions. *IEEE Wireless Communications*, 26(5), 174–180.

Vanka, S., Srinivasa, S., Gong, Z., Vizi, P., Stamatiou, K. & Haenggi, M. (2012). Superposition coding strategies: Design and experimental evaluation. *IEEE Transactions on Wireless Communications*, 11(7), 2628–2639.

Wang, W., Cheng, N., Teh, K. C., Lin, X., Zhuang, W. & Shen, X. (2019). On countermeasures of pilot spoofing attack in massive MIMO systems: A double channel training based approach. *IEEE Transactions on Vehicular Technology*, 68(7), 6697–6708.

Wolniansky, P. W., Foschini, G. J., Golden, G. D. & Valenzuela, R. A. (1998). V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel. *1998 URSI international symposium on signals, systems, and electronics. Conference proceedings (Cat. No. 98EX167)*, pp. 295–300.

Wu, M., Dick, C., Cavallaro, J. R. & Studer, C. (2016). High-throughput data detection for massive MU-MIMO-OFDM using coordinate descent. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 63(12), 2357–2367.

Xiao, Z., Xia, P. & Xia, X.-G. (2016). Codebook design for millimeter-wave channel estimation with hybrid precoding structure. *IEEE Transactions on Wireless Communications*, 16(1), 141–153.

Yin, B., Wu, M., Studer, C., Cavallaro, J. R. & Dick, C. (2013). Implementation trade-offs for linear detection in large-scale MIMO systems. *2013 IEEE international conference on acoustics, speech and signal processing*, pp. 2679–2683.

Yin, B., Wu, M., Cavallaro, J. R. & Studer, C. (2014). Conjugate gradient-based soft-output detection and precoding in massive MIMO systems. *2014 IEEE Global Communications Conference*, pp. 3696–3701.

Yu, X., Shen, J.-C., Zhang, J. & Letaief, K. B. (2016). Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems. *IEEE Journal of Selected Topics in Signal Processing*, 10(3), 485–500.

Zhang, H., Dong, A., Jin, S. & Yuan, D. (2017). Joint transceiver and power splitting optimization for multiuser MIMO SWIPT under MSE QoS constraints. *IEEE Transactions on Vehicular Technology*, 66(8), 7123–7135.

Zhang, L., Xin, Y. & Liang, Y.-C. (2009). Weighted sum rate optimization for cognitive radio MIMO broadcast channels. *IEEE Transactions on Wireless Communications*, 8(6), 2950–2959.

Zhang, R. & Ho, C. K. (2013). MIMO broadcasting for simultaneous wireless information and power transfer. *IEEE Transactions on Wireless Communications*, 12(5), 1989–2001.

Zhang, X., Molisch, A. F. & Kung, S.-Y. (2005). Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection. *IEEE Transactions on Signal Processing*, 53(11), 4091–4103.

Zhang, Y., Wang, D., Wang, J. & You, X. (2016). Channel estimation for massive MIMO-OFDM systems by tracking the joint angle-delay subspace. *ieee access*, 4, 10166–10179.

Zhao, N., Wang, W., Wang, J., Chen, Y., Lin, Y., Ding, Z. & Beaulieu, N. C. (2018). Joint beamforming and jamming optimization for secure transmission in MISO-NOMA networks. *IEEE Transactions on Communications*, 67(3), 2294–2305.

Zhao, X., Zhang, Y., Geng, S., Du, F., Zhou, Z. & Yang, L. (2019). Hybrid precoding for an adaptive interference decoding SWIPT system with full-duplex IoT devices. *IEEE internet of things journal*, 7(2), 1164–1177.

Zhu, L., Zhang, J., Xiao, Z., Cao, X., Wu, D. O. & Xia, X.-G. (2019a). Joint Tx-Rx beamforming and power allocation for 5G millimeter-wave non-orthogonal multiple access networks. *IEEE Transactions on Communications*, 67(7), 5114–5125.

Zhu, L., Zhang, J., Xiao, Z., Cao, X., Wu, D. O. & Xia, X.-G. (2019b). Millimeter-wave NOMA with user grouping, power allocation and hybrid beamforming. *IEEE Transactions on Wireless Communications*, 18(11), 5065–5079.

Zhu, X., Wang, Z., Dai, L. & Wang, Q. (2016). Adaptive hybrid precoding for multiuser massive MIMO. *IEEE Communications Letters*, 20(4), 776–779.

Zhu, Y., Zhang, Q. & Yang, T. (2017). Low-complexity hybrid precoding with dynamic beam assignment in mmWave OFDM systems. *IEEE Transactions on Vehicular Technology*, 67(4), 3685–3689.