UNIVERSITY OF SOUTHAMPTON

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING

SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE

Channel Estimation, User Activity Identification and Signal Detection in Grant-Free Multiple Access Systems

by

Jiatian Zhang

B. Eng., MSc.

A doctoral thesis report submitted in partial fulfilment of
the requirements for the award of Doctor of Philosophy
at the University of Southampton

Sep 2021

SUPERVISOR:

Professor Lie-Liang Yang

BEng MEng PhD, FIEEE, FIET

Next Generation of Wireless (NGW) Group, School of Electronics and Computer Science

and

Professor Robert G. Maunder

PhD, CEng, FIET, SMIEEE, SFHEA

Next Generation of Wireless (NGW) Group, School of Electronics and Computer Science

University of Southampton

Southampton SO17 1BJ

United Kingdom

Dedicated to my parents, Professor Zhefu Zhang, Mrs Fengjie Wang and to my fiancé Mr Sholto Holdsworth and his family.

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING
SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

Channel Estimation, User Activity Identification and Signal Detection in Grant-Free Multiple Access Systems

by Jiatian Zhang

The 5th and beyond wireless systems are expected to be device-centric. It has been widely recognized that in the device-centric systems, the traditional grant-based multiple access (GMA) methods are low efficiency due to the hand-shaking procedure resulted resource consumption and latency. Hence, the grant-free multiple access (GFMA) techniques have been proposed and studied, in order to solve the problems experienced by the GMA. This thesis focuses on the GFMA for massive machine-type communications (mMTC), with emphasis on the physical layer techniques, including channel estimation, user equipment (UE) activity identification (UAI) and information detection. To these objectives, we first provide a literature review in terms of the research background and various methods for achieving GFMA.

Then, the performance of the dynamic direct sequence code-division multiple-access (DyDS-CDMA) and multicarrier CDMA (DyMC-CDMA) with minimum mean-square error (MMSE) and MMSE-assisted successive interference cancellation (MMSE-SIC) detection schemes are studied. In our studies, we assume that each base station (BS) or access point (AP) with limited degrees-of-freedom is capable to support a massive number of potential UEs, while each UE becomes active with a small probability. Hence, the active UEs and the number of them are highly dynamic, making the number of active UEs possibly higher than the degrees-of-freedom of the system. Assuming ideal UAI and ideal channel estimation of active UEs, we study the potential performance achievable by the DyDS-CDMA and DyMC-CDMA systems employing the MMSE-assisted successive interference cancellation detection (MMSE-SICD), to demonstrate the feasibility of MMSE-SICD for operation in massive GFMA (mGFMA) systems, where the number of active UEs may be much higher than the systems' degree-of-freedom. Our studies reveal that the DyDS-CDMA and DyMC-CDMA systems aided by the MMSE-SICD are highly efficient for operation in mGFMA environments. Near single-user performance is achievable, even when the average number of active UEs per time-slot reaches two times of the system's degrees-of-freedom, and hence, the number of active UEs of a time-slot may be

much higher than two times of the system's degrees-of-freedom.

Following the above preparation work, then, we investigate the channel estimation and propose UAI algorithms for mGFMA systems. Specifically, channel estimation is studied from several aspects by assuming different levels of knowledge to the AP, and based on which five UAI approaches are proposed. We study the performance of channel estimation, the statistics of estimated channels, and the performance of UAI algorithms. Our studies show that the proposed approaches are capable of circumventing some of the shortcomings of the existing techniques designed based on compressive sensing (CS) and message passing algorithms (MPAs). They are robust for operation in the mGFMA systems where the active UEs and the number of them are highly dynamic.

Then, we investigate a multicarrier mGFMA system with an assumption that a big number of highly dynamic UEs are monitored by an AP with multiple receive antennas (MRA), which is referred to as the MRA/MC-mGFMA system. The channel estimation, UAI and information detection are separately or jointly addressed. To be more specific, firstly, the channels of both active and inactive UEs are estimated in the principle of MMSE. Then, based on the estimated channels, a low-complexity threshold-based UAI (TB-UAI) is proposed to detect the activities of UEs. Finally, information of active UEs is detected in the principle of MMSE-SIC. Furthermore, a joint algorithm, referred to as SIC-MMSE-JCUD, is proposed for joint channel estimation, UAI and data detection in the principle of MMSE-SIC. Additionally, considering that for a practically limited bandwidth, no set of the well-designed signature sequences can support a big number of UEs in a mGFMA system, we propose a class of sequences by combining the Gold-sequences with the Zadoff-Chu (ZC) sequences. Our studies show that deploying multiple receive antennas at AP is beneficial to channel estimation, UAI and information detection. Aided by the multiple receive antennas of AP, a low-complexity TB-UAI algorithm is highly efficient for UAI. Furthermore, our proposed class of signature sequences allows to attain much better performance than the random sequences.

Then, we extend our studies to the massive distributed grant-free multiple-access (MDR-GFMA) systems, to investigate the mGFMA with the cell-free scenario. In our studies, we assume a MDR-GFMA system where remote radio heads (RRHs) or APs, or simply distributed antennas (DAs) are randomly distributed in a given area based on the point Poisson (PP) distribution, while UEs are uniformly distributed. We assume that signals transmitted by UEs experience both the large-scale fading of propagation path-loss and shadowing as well as the small-scale Rayleigh fading. Signals received by different RRHs/APs are forwarded to a so-called signal processing central unit (SPCU), where channel estimation, UAI and data detection are carried out. In terms of signal processing at SPCU, channel

estimation is achieved in the principle of MMSE. Following channel estimation, an orthogonal matching pursuit (OMP) relied algorithm is implemented to attain initial UAI, which is enhanced with the aid of the pilot detection of each initially identified active UE. Finally, the data sent by active UEs is detected using either MMSE detection or the MMSE-SIC detection. Our studies show that the proposed algorithms are effective, and achieve expected performance in the MDR-GFMA systems with various dynamics, including active UEs and the number of them, locations of DAs and the number of them serving different UEs, geographically resulted large-scale fading of propagation path-loss and shadowing.

Declaration of Authorship

I, **Jiatian Zhang**, declare that the thesis entitled **Channel Estimation, User Identification and Signal Detection in Grant-Free Multiple Access Systems** and the work presented in it are my own and has been generated by me as the result of my own original research. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University;
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- Where I have consulted the published work of others, this is always clearly attributed;
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- Parts of this work have been published.

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

Acknowledgements

I would like to express my heartfelt gratitude to my supervisor, Professor Lie-liang Yang, who initiated me to the art of research. What I respect is not only his fundamental constructive knowledge, also he has been a faithful friend, a dedicated research collaborator, my father-like spiritual mentor, and a very supportive supervisor. I also would like to express my heartfelt gratitude to my secondary supervisor Professor Robert Maunder for his kind and generous support and help whenever I was in need. Many thanks to Professor Soon Xin Ng and Dr Mohammed El-Hajjar for their kind guidance and assistance ever since my PhD study. I am also very grateful to Professor Lajos Hanzo and Professor Sheng Chen for their helpful advice.

I would like to thank all my colleagues in the group of the Next Generation of Wireless in the University of Southampton. It was a great pleasure to work in cooperation with Dr. Chao Xu, Mr. Yifeng Xiong in a variety of interesting research topics. Many thanks to Dr. Peng Pan from Hangzhou Dianzi University and Dr. Hongkui Shi from Nanjing University of Posts and Telecommunications for their constructive comments and suggestions. I also would like to especially thank Mr. Yifeng Xiong for his kind and helpful advice. I truly cherish my time researching in Southampton, where I will always look back with very nostalgic memories. These make me feel reminiscent of Jack Kerouac's book "The Dharma Bums": "O ever youthful, O ever weeping".

I would like to express my appreciation to the Holdsworth family, my fiancé's parents Mr. Jonathan Holdsworth and Mrs. Annabel Holdsworth. They provided me a very safe and very quiet research room in the non-expectable and unpredictable lockdown time. I appreciate it so much in terms of they have looked after me like their own daughter with the situation of the last two years pandemic. Also Mr. Sholto Holdsworth himself is a researcher who uses his love, motivations and inspirations raising me up all the time.

The world has changed a lot by the impact of COVID-19. Looking back in a way to which has led me here, I think that I do not have any regrets over any time I wasted. Nor will I feel ashamed of not accomplishing anything. Four years ago I decided to resign my job in the central Tokyo and to pursue a PhD in the University of Southampton, my future was vaguely in front of me at that moment. Personally, there was a huge gap from the industry back to academia. However, I have never forgotten why I started, and even my mission is sometimes struggled. I want to thank myself for having the courage to do my PhD, to conquer all of the stress, to pursued myself offering such a luxury life to accomplish my PhD, to achieve a tranquil inside accompanying the cutting edge technology's

research.

I would like to express my heartfelt gratitude to my mum Ms Fengjie Wang and my dad Professor Zhefu Zhang for their unconditional love and support. They have endured hardship in their lives and the golden age of their lives had been destroyed by political unsettlement, so that I could have opportunities they themselves never had. I am their eyes, when I get towards to search the horizontal of the world. When I achieve anything meaningful, I will always be thankful for them a million times. They have both already retired, but the good spirit inherited from them make me feel strong, to stand on mountains, to walk on stormy seas. Especially, my father, who is currently suffering from the his physical condition, I try to let him see my success at least accomplishing my PhD more than ever.

I have a lot of genuine friends, although some of them are even have 8 or 9 hours jet-lag with me, they care and encourage me for all the time. I would like to thank my brother Sphinx, my life long friends Xiaojing Wang, Zhongya Zhu, Lanyu Gao, Yue Zhang, Hui Zhang and Yinzuo Zhang, who helped me go through the hardest times of my life. Xiaojing Wang and I have known eachother for 21 years, as a university's lecture, she is very wise and modest. To me, she has been a lighthouse for all the time in my life. Also, I would never forget the time living in 122 Broadlands Road and 11 Hartley Avenue, both of them are in Southampton, I spent a half my research time in those two houses, when troubles come and my heart burdened be, Miss Lanyu Gao would play piano to therapy me, Mr. Yue Zhang, Miss Lingyu Wang and Dr. Fei He used their unique humour sense to provide a lot of help and assured me very much, I am really appreciative. I also would like to thank the Dog and Duck pub in Outwood, some of the key equations were inspired from the corner table in that pub, Waine, Donna and the locals treat me like family.

List of Publications

Journals:

1. **J. Zhang**, P. Pan, L. -L. Yang and R. G. Maunder, "Channel Estimation and User Activity Identification in Massive Grant-Free Multiple-Access", IEEE Open Journal of Vehicular Technology, vol.1, pp.296-316, August 2020, (Invited Paper).

2. P. Pan, **J. Zhang** and L. -L. Yang, "Massive Distributed Antenna Systems: Channel Estimation and Signal Detection", IEEE Access, vol.8, pp.186055-186070, October 2020.

3. **J. Zhang**, P. Pan and L. -L. Yang, "Joint Channel Estimation, User Activity Identification and Information Detection in Multi-Antenna mGFMA Systems", ready to submit as an IEEE journal paper.

4. **J. Zhang**, P. Pan and L. -L. Yang, "Channel Estimation, User Activity Identification and Data Detection in Distributed Antenna mGFMA Systems", ready to submit as an IEEE journal paper.

5. H. K. Shi, **J. Zhang** and L. -L. Yang, "A Low-Complexity CNN aided Performance Improving Scheme for Grant-Free LDS-OFDM Network", ready to submit as an IEEE journal paper.

Conferences:

1. **J. Zhang**, P. Pan and L. -L. Yang, "Dynamic DS-CDMA Aided by Successive Interference Cancellation for Massive Grant-Free Multiple-Access", 2018 10th International Conference on Wireless Communications and Signal Processing (WCSP), pp.1-7, Hangzhou, China, 18-20 October 2018.

Strenuis Ardua Cedunt (The Heights Yield to Endeavour)

– University of Southampton Motto

Contents

List of Acronyms and Symbols

List of Acronyms

$3GPP$: The 3rd generation partnership Project

$5G$: The 5th generation mobile networks

$6G$: The 6th generation mobile networks

APs: Access points

$BOMA$: Building block sparse-constellation based orthogonal multiple access

BSs: Base-stations

CoF: Compute-and-forward

$C-RNTI$: Cell- radio network temporary identity

CS: Compressive Sensing

$CTSMA$: coded tandem spreading multiple access

DAs: Distributed antennas

DL: Downlink

$DMRS$: Demodulation of reference signal

$DyDS-CDMA$: dynamic direct sequence code-division multiple-access

EAE: Eigen-analysis enhanced

$eMBB$: Enhanced Mobile Broadband

$FDMA$: Frequency Division Multiple Access

FDS: Frequency diverse subaperturing

$GFMA$: Grant-free multiple-access

$GOCA$: Group orthogonal coded access

gNB: Next generation NodeB

HTC: Human-type communications

$IDMA$: Interleave division multiple access

$IGMA$: Interleave-grid multiple access

$IRSA$: Irregular repetition slotted ALOHA

IoT: Internet of Things

ISI: Inter-signal interference

LBT: Listen before talk

$LCRS$: Low code rate spreading

LDS: Low-density signature

$LDS-SVE$: Low density spreading signature vector extension

$LPMA$: Lattice partition multiple access

$LSSA$: low code rate and signature based shared access

LTE: Long-Term Evolution

$LTE-A$: Long-Term Evolution-Advanced

MA: Multiple access

$MC-CDMA$: Multi carrier-sequence code division multiple access

$MC-mGFMA$: Multicarrier mGFMA

$MDR-GFMA$: massive distributed grant-free multiple-access

$MIMO$: Multi-Input Multi-Output

ML: Machine Learning

$mMTC$: Massive Machine-Type Communications

$MMSE$: Minimum mean square error

MPA: Message passing algorithm

MRA: Multiple receive antennas

MUD: Multi-user detection

$MUSA$: Multi-user shared access

$NCMA$: Nonlinear modified constant modulus algorithm

$NOCA$: Notification-oriented computer architecture

$NOMA$: Non-orthogonal Frequency Division Multiplexing

NR: New Radio

$OFDM$: Orthogonal Frequency Division Multiplexing

OMA: Orthogonal multiple access

OMP: orthogonal matching pursuit

$PDMA$: Polarization-division multiple access

$PD-NOMA$: Power-domain non orthogonal multiple access

POs: PUSCH Occasions

PP: Poisson point

$PRACH$: Physical Random Access Channel

$PRBs$: Physical Resource Blocks

$PUSCH$: Physical Uplink Shared Channel

RA: Random access

$RACH$: Random access channel

RAN: Radio access network

RAR: Random-access response

RB: Resource block

$RDMA$: Repetition division multiple access

$RRHs$: Remote radio heads

$RSMA$: Resource spread multiple access

PRU: PUSCH resource unit

ROs: PRACH Occasions

$SAMA$: Spread ALOHA multiple access

$SCMA$: Sparse code multiple access

$SDMA$: Spatial division multiple access

SIC: Successive interference cancellation

$SICD$: Successive interference cancellation detection

$SPCU$: signal processing central unit

TB: Threshold-based

UAI: User activity identification

UEs: User Equipments

UL: Uplink

$URLLC$: Ultra-Reliable Low-Latency Communications

$WG1$: Working group 1 (3GPP)

ZC: Zadoff-Chu

Special Symbols

$(\cdot)^H$: The Hermitian transpose of a matrix.

$(\cdot)^T$: The transpose of a matrix.

$(\cdot)^*$: The conjugate of a complex symbol/vector/matrix.

$\|\cdot\|^2$: The Euclidean norm of a vector/matrix.

$\otimes$: The Kronecker product

$\sum$: The summation of all elements.

$\prod$: The product of all elements.

$\int$: The integration function.

Chapter 1

Introduction

The Internet of Things (IoT) has been hitting its stride embarking on the 5th Generation (5G) wireless networks, 5G beyond and the 6th Generation (6G) wireless networks [1], which are expected to support numerous devices and operated in three main scenarios [2]: Ultra-Reliable Low-Latency Communications (URLLC), enhanced Mobile Broadband (eMBB) and massive Machine-Type Communications (mMTC). Therefore, one of the targets to the future wireless network has been frequently described as device-centric, full end-to-end latency reduction, network energy efficiency, connection density and reliability, etc. In this thesis, our focus is on the mMTC.

1.1 Research Background and Motivation

In the context of above mentioned three distinctive scenarios of the network implementations, more recent attention has focused on the provision of the latency-sensitive applications that demand wide coverage and high capacity. It is a widely recognised view that most of the machines/devices will continuously work in the New Radio (NR) [3] scenarios, where sensors are committed to carry on data collections and report the real-time wireless channel states by sending short-burst messages. In the NR sensing scenarios, sensors are defined [4] as the mechanical or electrical devices sensitive to light, temperature, radiation levels, etc., that transmits signals to measuring or controlling instruments. Within the context of the IoT applications in wireless systems: a sensor is usually defined as the devices which is generally small and mechanical but is sensitive to a measurable physical parameter and provides a signal level result directly related to the measured amount of the physical parameter [5]. In the wide variety use of sensors, the parameter being measured represents the physical phenomenon, which is converted to a measurable signal. Therefore, the main characteristics of the sensor User Equipments (UEs) in IoT or mMTC networks are that

of providing fully/partially communication in a sporadic transmission manner mainly over uplink, with a very small transmit-data size per device but under the requirement of an extremely high energy efficiency [6]. These UEs usually have very low tolerance of latency. An example of these UEs is the one studied in [7], in which the latency of critically augmented reality surgeries and robots was targeted to be less than $750\mu s$ and $2ms$, respectively. In parallel to the work on the mMTC, Ericsson predicted that the number of IoT UEs may reach 31.4 billion by 2023, accounting for 60% in mMTC [8].

The conventional Long-Term Evolution (LTE) and Long-Term Evolution-Advanced (LTE-A) systems, mainly support eMBB that predominately focus on the downlink (DL) human-type communications (HTC) in a large data size manner. On the contrary, driven by the unaccountable IoT applications, future wireless networks are expected to support numerous devices and to provide massive connectivity [9]. As a result, many wireless systems will be device-centric, constituted by densely distributed devices and machines, in contrast sharply to the human-centric 1G - 4G wireless systems. As most devices/machines carry out tasks such as sensing, data collection, etc., they often need to report via wireless channels by sending short-burst messages. Explicitly, in the device-centric wireless networks, the legacy human-centric systems, such as LTE/LTE-A in 4G systems, may not be efficient for supporting the sporadic short-burst applications required by IoT devices (and machines). Instead, novel communications protocols need to be designed to allow a huge number of ultra-densely distributed devices to communicate with their base-stations (BSs) or access points (APs) reliably and efficiently [10, 11]. To address this problem, low latency uplink traffic is expected to dominate in the future sensors/device-centric systems [12]. Therefore, this thesis focuses on the delivery of the uplink machine-type UE traffic.

To support a wide range of connectivity requirements in the NR networks, a lot of uplink access approaches have been proposed in both industry and academia. Explicitly, the multiple access schemes for the UL of mMTC are summarized in Tables 1.1-1.2, along with the key functions including the main contributions receiver types. More specifically, the multiple access (MA) schemes for mMTC are mainly designed based on the different MA signatures, which can be categorized into power domain, codebook, spreading sequences, scrambling and interleaving, etc. From the list, we can see that some of the references have considered the mMTC aided IoT scenarios where supporting massive connectivity via loading multiple UEs in one RB is considered. However, most of these perspectives are still based on the view of HTC, where APs and UEs need to hold a priori knowledge about each other and strict system synchronization is also required in terms of the centralized scheduling networks. Therefore, an extra consumption of the overhead signalling may increase the challenge, as the results of the strict synchronizations and the

requirements for that each device requests an assigned data transmission slot through a contention-based random access (RA) procedure, which is as defined in the grant-based scheme. Nevertheless, the concept of MA in mMTC is motivated by the expectation that the applications are "always on" [3] or "arrive and go" so as to achieve a substantially reduced latency. However, in reality, the time performance loss imposed by the conventional grant-based access schemes again indicates that even the presence of improved grant-based access procedures may be struggling to cater the mMTC traffic. In order to improve the access performance in the mMTC aided NR systems, 3GPP RAN working group 1 (WG1) has already proposed to employ the multiple random access channel (RACH) preamble formats to flexibly support the shorter or longer preamble lengths for possibly avoiding the outage resulted from the congestion in the system, as detailed in [13, 14]

Although a series of significant improvements have already been available in the grant-based schemes, there is only one RACH timeline supporting all the cases in the HTC grant-based procedures. Keeping in mind that mMTC is required to flexibly serve variety communications scenarios under different latency constraints, the one timeline scheme is obvious not sufficient. Therefore, the design of novel RACH schemes is urgent and critical. To address this, a "2-step" RACH scheme based on grant-free concept was proposed in 3GPP, as shown in Fig. 1.1 (a), With the "4-step" scheme, the UE transmits a contention-based Physical Random Access Channel (PRACH) preamble, also known as Msg1. After detecting preamble bits, the next generation NodeB (gNB) responds with a random-access response (RAR), also known as Msg2. Here the RAR includes the detected preamble ID, a time-advance command, a cell- Radio Network Temporary Identity (C-RNTI) and an uplink grant for scheduling a Physical Uplink Shared Channel (PUSCH) transmission from the UE known as Msg3. Then, UE sends Msg3 in respondence to the RAR, which includes an ID for contention resolution. Upon receiving Msg3, the gNB finally sends the "contention resolution" message of Msg4, which contents the contention resolution ID. Once UE receives Msg4 and finds its contention-resolution ID, it sends an acknowledgement to a PUCCH to complete the "4-step" random access procedure. By contrast, in a "2-step" scheme, as shown in Fig. 1.1 (b), MsgA consists of a PRACH preamble and a PUSCH transmission, they are defined as MsgA PRACH and MsgA PUSCH respectively. The MsgA PRACH preambles are different from the "4-step" RACH preambles. However, they can be transmitted by the same PRACH Occasions (ROs) as the preambles in the "4-step" RACH. Furthermore, they can be sent in separate ROs. The PUSCH transmissions are organized in the PUSCH Occasions (POs), which span multiple symbols and Physical Resource Blocks (PRBs), with optional guard intervals and guard bands between consecutive POs. Each PO consists of multiple demodulation of reference signal (DMRS) sequences, and each pair of DMRS sequence is known as a PUSCH resource unit (PRU).

Table 1.1: A Summary of Proposed Multiple Access Schemes [6] (Part I).

Access Schemes	Categories	Contributions	Receiver Types
SCMA [15–19]	codebook based	Codebooks work via multi-dimensional constellation mapping, to provide more diversity on constellation shaping gain; bit-to-constellation mapping and spreading are combined to make sure that each user has their unique codebook.	MPA
PDMA [20–23]	codebook based	Users can be multiplexed in a variety domains, i.e., code, power, space and their combinations via employing sparsity of spreading sequences	MPA/SIC
PD-NOMA [24–27]	power based	UEs transmit their data at different power levels in the same resource block (RB), and the APs receive the different power via MUD SIC receivers.	SIC
LDS-CDMA [28–31]	spreading sequences based	UEs can share a RB across unique user specific spreading sequences by exploiting LDS to constrain the interference on each chip.	MPA
LDS-OFDM [32–35]	spreading sequences based	The scheme can be treated as a combination of the traditional OFDM and the LDS-CDMA, mapping the multiplied symbols with LDS sequences over different OFDM subcarriers.	MPA
LDS-SVE [36]	spreading sequences based	Inspired by the traditional LDS, a larger signature vector is built for exploiting diversity gain.	MPA
SAMA [37]	spreading sequence based	The spreading sequences in SAMA appear high sparsity to support the UEs data spreading in a variety number of resources.	MPA
MUSA [38]	spreading sequences based	Data of each user is spread with a family of complex spreading sequences on the real parts and imaginary parts with short length codes. Moreover, each user can choose its spreading code autonomously.	SIC
NCMA [39–41]	spreading sequences based	Grassmannian line packaging problem aided spreading codes obtaining.	PIC
NOCA [42]	spreading sequences based	Spreading codes are defined by LTE	SIC

Table 1.2: A Summary of Proposed Multiple Access Schemes [6] (Part II).

Access Schemes	Categories	Contributions	Receiver Types
FDS [43]	spreading sequences based	Modulation symbols spread directly with multiple orthogonal or quasi-orthogonal codes.	SIC
LCRS [44]	spreading sequence based	Modulated symbols directly spread with multiple orthogonal codes.	SIC
GOCA [45]	spreading sequences based	Non-orthogonal sequences provide group separation, where GOCA has a two-stage spreading structure. Nevertheless, devices in a group are supported by a set of orthogonal sequences.	SIC
IRSA [46]	power based	A packet generated at each AP is duplicated to the number defined by the system and is transmitted in different time slots selected by the AP.	SIC
LSSA [47]	scrambling based	Low rate FEC is used for symbol-level permutation patterns.	SIC
RSMA [48–50]	scrambling based	Long spreading sequences in a good correlation manner combining with very low FEC rate, to support different interleavers.	SIC
IDMA [51]	interleaving based	Symbol-level interleaving for specific users, low rate FEC may also be involved.	ESE
IGMA [52]	interleaving based	symbol-level interleavers and bit-level grid mapping are used for UEs separation.	ESE
RDMA [45]	interleaving based	Bit-level cyclic-shift repetition patterns are used for designing device-specific signatures.	SIC
SDMA [53–55]	others	Due to a variety of channel impulse responses, multiplex UEs can be responded by user specific CIR. Therefore, a large number of UEs can work under a comparably accurate estimated channel situation.	PIC
LPMA [56–58]	others	LPMA supports flexibly multiplexing due to the multiple degrees of freedom. Moreover, multi-level lattice superposition codes can be allocated at different levels to multiplex UEs.	SIC
BOMA [59]	others	Carry out multiplexing by attaching data information from good channel UEs to those relatively bad channel UEs. Furthermore, the bad channel UEs can apply coarse constellation with large minimum distance.	LCR

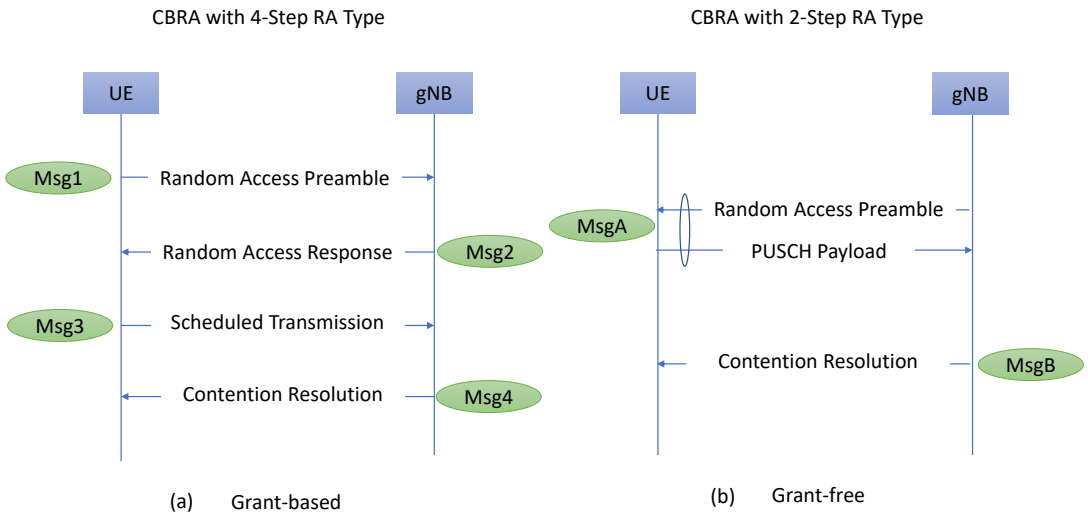


Figure 1.1: An example of 4-Step RA Type and 2-Step RA Type based on the contention-based random access (CBRA) for grant-based and grant-free MA [60].

Therefore, the "2-step" RACH supports both "one-to-one" and "multiple-to-one" mappings between the preambles and PRUs.

According to Fig. 1.1, we can infer that the grant-based "4-step" random-access procedure requires two round-trip cycles between UE and gNB, which not only increases the latency but also consumes additional overhead for control-signalling. By contrast the grant-free "2-step" RACH is able to reduce the latency and control-signaling overhead by improving a single round trip cycle between UE and gNB. This is achieved by combining the preamble (Msg1) and the scheduled PUSCH transmission (Msg3) into a single message (MsgA) at the UE side. Then, the random-access response (Msg2) and the contention resolution message (Msg4) are combined into a single MsgB at the gNB side. Furthermore, for operation in the unlicensed spectrum in the future wireless systems, the number of Listen Before Talk (LBT) attempts can be reduced by reducing the number of messages transmitted between the UE and the gNB.

Currently intensive researches on grant-based access methods are still ongoing. However, in device-centric mMTC networks, all devices are expected to be operated with high energy-efficiency, which for some applications is critical due to the practical constraint on battery changing, battery re-charging, etc. In order to meet the above-mentioned challenges, grant-free multiple-access (GFMA) has been proposed to support sporadic short-burst applications [61]. Specifically in [62], the authors compared the performance of the transmissions of short-burst data blocks in both traditional LTE network and in new 5G grant-free networks. The studies show that data transmission in grant-free systems should be in the manner of arrive-and-go, and that using the traditional communication schemes

to transmit sporadic short-burst traffic is inefficient. Hense, it has been widely recognized that the GFMA [60, 63] is a more promising solution, which can support massive applications in small package data services in the sporadically uplink dominant scenarios.

1.2 Overview of Non-orthogonal Multiple-Access

Based on the above discussion, it is important to design new wireless scheme for GFMA to resolve the limitations of the existing MA procedures. With regard to this context, non-orthogonal multiple access (NOMA) has been proposed as a competitive MA scheme for the future IoT networks. The definition of NOMA is in contrast to the concept of orthogonal multiple access (OMA). In OMA, gNBs allocate an appropriate resource to one UE. Hence, it is clear that the OMA methods are hard to meet the massive connectivities in a URLLC system, due to the limited number of OMA signatures. By contrast, NOMA can support multiple UEs in one resource block (RB) synchronously [64–69]. As some examples, in [70], Tararefa *et al*, considered the frequency division multiple access (FDMA) scheme, supported by the joint user association and resource allocation for IoT uplink. Sutton *et al* [71] investigated the performance of a time division multiple access (TDMA)-assisted IoT uplink. In [72], Hu *et al* introduced the cognitive radio principles to code Division Multiple Access (CDMA), in order to address the problem of resource shortage in IoT systems. Furthermore, in [73], the NOMA technique was introduced for supporting IoT cloud services.

Although NOMA can offer efficient multiple access over one RB, the data initiated by UEs in the conventional networks can not be transmitted to the BSs without firstly sending a connection request. To minimize the latency in a multiple access environment, various NOMA-based techniques have been proposed by researchers, so that a large number of UEs can efficiently share the same resources units. However, the signalling overhead demanding a lot of resource consumption which is still disruptive, as the uplink access scheduling is based on the response of the available channel resources allocated by BS/access point (AP) in LTE/LTE-A. This is regarded as an obvious ceiling in the uplink access of IoT networks. Hence, the compatible and simple access procedure should be considered by reducing the signalling overhead as much as possible.

Furthermore, if GFMA and OMA are combined, collisions will occur due to the special resource allocating mechanism. The authors in [74] analysed how the collisions are generated in OMA, as shown in Fig. 1.2. When two UEs simultaneously initiate uplink access in the OMA-GFMA systems, it is highly risky to meet each other in the same grant-free time slot. In this case, they have to trigger a retransmission mechanism in the

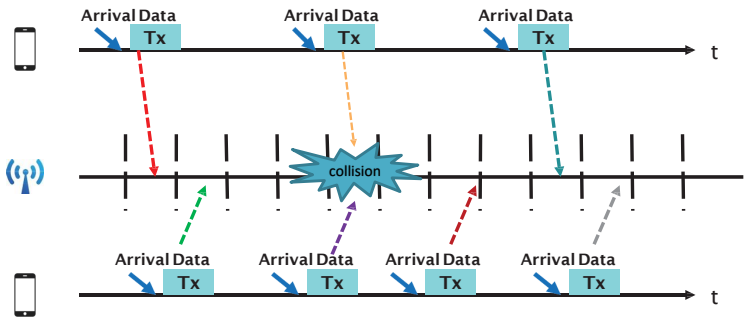


Figure 1.2: [74] Illustration of collision in a combined OMA-GFMA system.

next random back-off time window. Therefore, researchers have be motivated to propose novel uplink access methods without relying on grant in NOMA, i.e. GFMA. In [75–77], NOMA has been proved to be a robust MA strategy for the 5G and future wireless systems. In literature, various NOMA designs have been proposed mainly based on power allocation [78–82], spreading schemes [16, 83–87], scrambling sequences [88] and interleaving algorithms [89] [90], as summarized in Tables 1.1-1.2. The essence of NOMA is employing multiusers detection (MUD) for applying the specific signature codes to transmit data over the same RBs. In [91–93], the performance of NOMA was compared with that of OMA. The authors of [94] illustrated that NOMA techniques allow many devices to share limited resources. Meanwhile, the authors in [95] proposed a coded tandem spreading multiple access (CTSMA) scheme for enhancing the multiple-access capability of NOMA systems. Furthermore, in [96], a so-called comprehensive NOMA scheme was proposed for MTC which allows synchronization, channel estimation, user identification and data detection to be performed in one single shot. In [97, 98], the authors investigated the multiple receive antenna assisted multi-cell system with cooperation, as a further extension of a cloud-radio structure. In [99], a novel preamble design was proposed for attaining a higher success rate in the MIMO relied NOMA systems. Additionally, in [100] [101], the authors explored a novel interleave-division multiple-access (IDMA) transceiver design, which makes use of the asynchronous characters of random signals in the detection and channel estimation in the GFMA-NOMA systems.

However, what stands out in the GFMA-based UL NOMA systems is that even the dynamic grant scheduling is not required for transmissions. This means that a UE can transmit its data when it needs. In this manner, the end-to-end latency may be significantly reduced along with the reduction of signalling overhead. Moreover, multiple access should be allowed to share the same time-frequency resources through the principles of NOMA, which makes NOMA well suite for the UL GFMA systems.

1.3 Overview of Grant-Free Multiple-Access

Due to the drawbacks of grant-based MA and OMA schemes as discussed above, GFMA has been considered for the uplink access in mMTC communications. In this section, we provide an overview for GFMA in terms of its research issues.

Consider numerous UEs staying in a sleep status in a BS/AP serving area with some of them occasionally waking up to transmit information without permission, i.e, in the GFMA way, and only under the comparably simple protocol, such as, of time-slot based transmission. Most signal processing will be done by the BS or AP. In order to achieve GFMA, some a-priori information is needed. These include UE's IDs or the signature codes assigned to UEs. Additionally, in GFMA, pilot symbols are sent along with data. The signature codes and pilot symbols are used by BS(AP) to achieve UAI, channel estimation and finally, data detection of active UEs.

Therefore, to achieve the signal detection in GFMA systems, the signatures are supposed to be pre-configured. In principle, all methods listed in Table 1.1 and Table 1.2 are able to be satisfied to implement GFMA. However, as the power-based NOMA is heavily depended on the various receiving power levels in different UEs, it may improve a huge challenge in terms of maintaining sufficient power differences from multiple UEs at the BS side. Therefore, in general, the power-based methods need an operational near-far solution in reality. However, this potential near-far problem may bring hurdles to the design of successive interference cancellation (SIC) receivers for the UL GFMA systems [102, 103]. Consequently, the power based NOMA schemes are in general not suitable for operation in mGFMA systems.

Below we provide a brief overview of the different GFMA schemes which are clarified in the ways as shown in Table 1.3

1.3.1 Compressive Sensing Based GFMA

Compressive Sensing (CS) exploits the sparsity of a signal and then recovers it from far fewer samples than the required by the Nyquist criteria. In most IoT networks, the number of active users is usually much smaller than the number of potential users in the whole system, which means the IoT network appears sparsity [104]. Therefore, active UE detection and data recovery of active UEs always take the priority, where the BS/AP needs to identify the active UEs out of the potential UEs and then detects the transmitted data by active UEs. To meet these requirements, the CS assisted approaches have gained massive research interest for the receiver design in GFMA IoT networks. So that the inherent spar-

Table 1.3: Categorization of GFMA schemes.

GFMA Schemes	Main Characteristics
Comprehensive sensing (CS) based GFMA	1. Sparse code multiple access (SCMA). 2. CS allows to support more UEs than systems degrees of freedom. 3. Blind MUD methods to detect active UEs is possible. 4. User-specific signature patterns provide distinctive features among different UEs.
Message passing algorithm (MPA) based GFMA	1. Each user has its own codebook and adaptivity to enable dimensional mapping. 2. MAP receivers can support more users to appear more diversities.
Machine learning (ML) based GFMA	1. Training sequences is essential. 2. To exploit ML scheme enabling an achievable optimum decision.
Compute-and-forward (CoF) based GFMA	1. UEs encode their data into inner codes and outer codes. 2. Inner codes are in charge of error correlation while outer codes provide support for MUD.
MA signature design based GFMA	1. UEs use flexible signature patterns as preambles, pilots, etc, in the spreading/interleaving way. 2. MUD and SIC for detection. 3. One MA signature is defined over one OFDM RB.

sity can be exploited by the CS based MUD to enhance the system performance. As an example in [105], the authors investigated the CS-based MUD in GFMA systems.

In the CS-assisted GFMA, blind MUD can be implemented to jointly perform user activity detection and data detection [106]. This is achieved by exploiting the sparsity existing in the GFMA systems. Hence, the CS based NOMA schemes allow UEs to transmit their data at any time slot in the grant-free way.

Specifically, there are two types of vector measuring models considered in literature for the CS-based GFMA. The first model is a single-measurement vector based CS (SMV-CS) [107], which represents the received signals of the one burst of transmission data as a vector that is the product of the UE's data vector and the sensing matrix continuing by, such as spreading, channels, etc. However, in the SMV-CS, the size of the sensing matrix may become extreme when the number of UEs is high. In this case, the systems performance degrades due to the increasing number of UEs and the long processing time. To improve the performance, the multiple-measurement vector CS (MMV-CS) [108] has been proposed to reduce the size of the sensing matrix aided and reduce the complexity

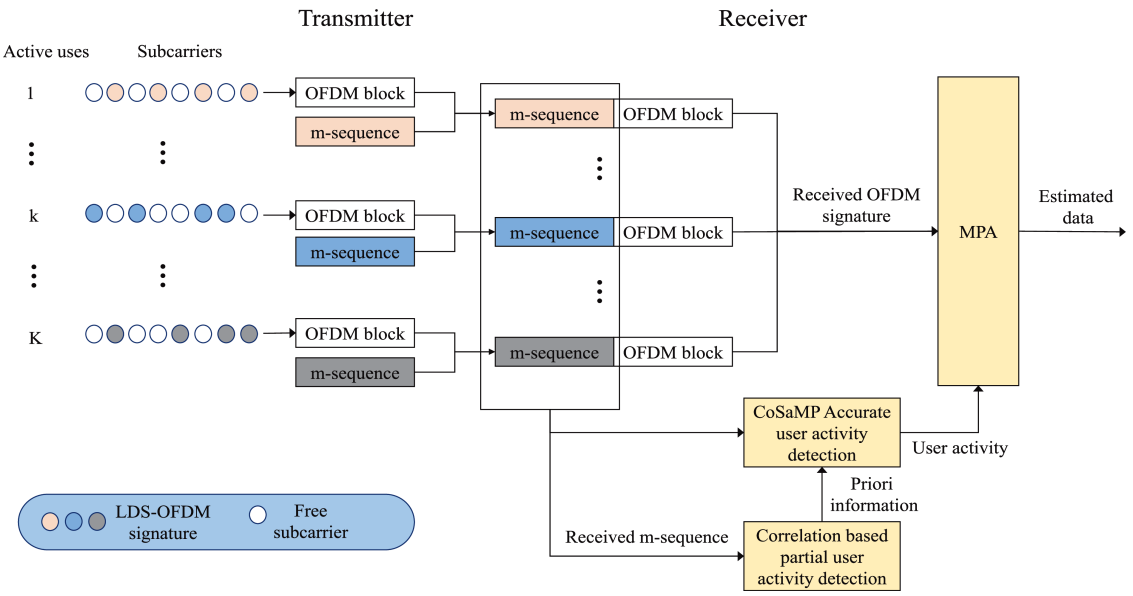


Figure 1.3: A typical CS-based GFMA system with MPA-assisted detection in [6] and [115]

of receiving processing. It is shown in literature that MMV-CS is an effective candidate method for implementations in URRLC scenarios [109–111].

In the CS-assisted GFMA, there are also other methods and schemes, such as the user activity detection algorithms proposed in [6], the three different algorithms proposed in [61], which are the active pilots detection and channel estimation, focal underdetermined system solver (FOCUSS) and expectation maximization (EM). Moreover, a joint sparsity-inspired sphere decoding (SI-SD) combined blind detection algorithm was proposed for CS GFMA in [112]. With this algorithm, each user's status and transmission data can be jointly detected by introducing one additional all-zero codeword. Additionally, an enhanced detection based group orthogonal matching pursuit (DGOMP) MUD algorithm was proposed in [113] for UE activity and data detection in the massive GFMA systems with SCMA assisted transmissions. Furthermore, the authors in [114] studied synchronization, channel estimation, user detection and data decoding in GFMA systems.

1.3.2 Message Passing Algorithm (MPA) Based GFMA

The CS based GFMA method is capable of achieving good performance in sparse IoT networks. In literature, various methods have been proposed to enhance the performance of the CS-based GFMA systems. MPA-based algorithm can be regarded as the most promising one, for example, having the structure as shown in Fig. 1.3 from Fig.3. of [115].

To be more specific, in [116], the authors enhanced the CS-based approach by propos-

ing a greedy algorithm relied on the MAP criterion, which performs UE identification and data detection jointly by exploiting the a-posteriori probabilities of each other. In [115], the authors described the MPA relied CS-based GFMA in detail, where the two optimization steps MUD designed for the LDS-OFDM GFMA systems. Specifically, during the first stage, the correlation-based activity is detected, which generates an approximated support fed into the second stage, when the CoSaMP algorithm is carried out. In [117], the authors considered the variations of the sparsity level in the received UE signals and proposed a switching mechanism to adaptively choose between a CS-based MUD and a traditional MUD. Similarly, considering that the sparsity of UEs appears variety in the time domain, a dynamic low complexity CS-based MUD was proposed in [118, 119]. This method was proposed based on the principle that although users can do random access (RA) at any time they may generally transmit data in adjacent time slots, which leads to the temporal correlation of the active user sets [120]. By exploiting this temporal correlation, the output of the estimated active user set in one particular time-slot may be used as the initial set of active UEs in the adjacent time-slot.

In the MPA-assisted GFMA, current near-optimal MUD based on the MPA algorithms has been widely studied to achieve the near optimal MPA performance. However, in these detectors, the receiver assumes that the BS/AP knows ideally the UE's activities, which is in general impractical. In reality, GFMA systems are highly dynamic, making the MUD yet highly challenging. To mitigate this problem, CS and MPA were jointly used to design a CS-MPA detector to achieve joint user activity and information detection in the LDS based GFMA systems [117], where a so-called CoSaMP was proposed by using a CS-based sparse signal recovery algorithms for the user activity detection by identifying the positions of non-zero elements [117]. It is shown that this algorithm has low complexity and good robustness against noise. Furthermore, the CoSaMP can make full use of the sparsity of the LDS structures, allowing the receiver to employ a low complexity MPA-based processing. Additionally, the authors of [117] proposed another low-complexity MPA-based algorithm for the GFMA systems to perform MUD.

1.3.3 Multiple-Access Signature (MAS) Based GFMA

In GFMA systems, BS may not have the complete information about the UEs and various other transmission parameters may also be unknown or only partially-known. In this case, to enable GFMA, a contention-based multiple-access (MA) [121] resource is defined, which comprises of a physical resource (a time-frequency block) and a MA signature. The MA signatures may include at least one of the following, referring to Tables 1.1-1.2: codebooks/codewords, sequences, interleavers, mapping patterns, demodulation reference

signals, power-dimensions, spatial-dimensions, preambles, etc. For the MA signature selection, one option is that the UEs performs random selection and the other option is that UEs' signature is pre-configured/pre-determined [122–124]. It can be shown that based on these MA signatures, various GFMA systems can be designed.

1.3.4 Compute-and-forward (CoF) Based GFMA

The GFMA schemes considered so far are mainly based on the MA signatures by employing user-specific sequences, such as, spreading, scrambling, interleaving sequences, or any combination of these, to enable MUD at the receiver. Another approach towards GFMA is proposed in [125], based on the principle of CoF [126]. This scheme relies on the codes with a linear structure, such as the nested lattice codes. This is because the linearity of the codebooks ensures that any combinations of codewords are themselves codewords. Therefore, the destination is free to determine which linear equation to recover [125]. Based on this principle, in the CoF-based GFMA systems, the UEs encode their messages into two different channel codes, where one is used for error correction and the other one is used for user detection. Correspondingly, at the receiver side, the BS/AP first decodes the sum of the received codewords to identify the active UEs, and then decodes the data of active UEs. If the BS/AP cannot recover the sum of codewords correctly, all the transmitted data of active UEs are lost, which is similar to the signature-based GFMA schemes discussed earlier.

1.4 Contributions and Thesis Outline

During the study for my PhD, I have authored and been a co-author for 5 journal papers as well 1 conference paper. This thesis has been written based on the 3 journal papers that I published and are under review as the primary author.

The main body of this thesis is focused on the investigation of channel estimation, UE activity identification and information detection in GFMA systems. The organization of this report is outlined in Fig.1.4.

The main contributions of this thesis can be summarized as follows:

Chapter 2: Signal Detection in Dynamic DS-CDMA and MC-CDMA Systems

In **Chapter 2**, we study the dynamic direct sequence code-division multiple-access (DyDS-CDMA) and multicarrier CDMA (DyMC-CDMA) systems, where the dynamic property is typical in the mGFMA systems to be considered in the following chapters. Spe-

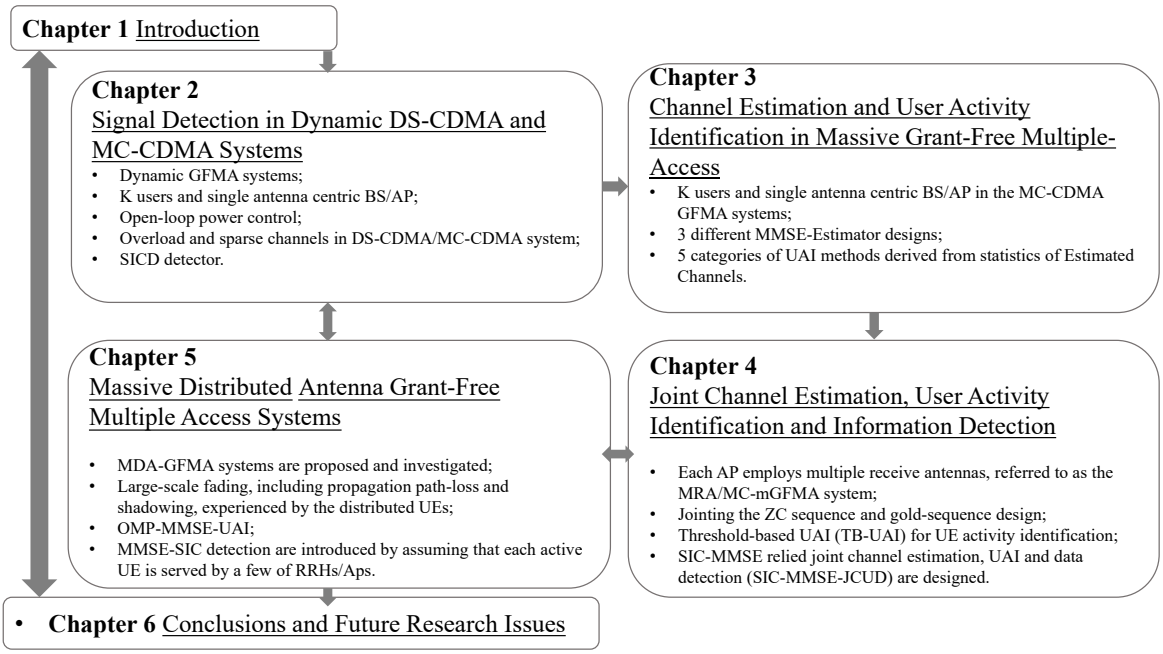


Figure 1.4: Outline of this treatise.

cially, this chapter assumes that each BS/AP with limited degrees of freedom is expected to control a massive number of UEs. We assume that UEs become active to transmit data on the basis of time-slots and during a time-slot, each of UEs activates independently and follows a common activation probability. While for the DyDS-CDMA systems we assume flat Rayleigh fading, we assume that the DyMC-CDMA systems experience frequency-selective Rayleigh fading. A low-complexity detector based on minimum mean-square error and successive interference cancellation, referred to as the MMSE-SICD, is proposed for signal detection in DyDS-CDMA and DyMC-CDMA systems. Assuming ideal active UE identification and ideal channel estimation of active UEs, in this chapter, we focus on the potential performance achievable by the DyDS-CDMA and DyMC-CDMA systems employing the MMSE-SICD, to demonstrate the feasibility of MMSE-SICD for operation in mGFMA systems.

- DyDS-CDMA and Dy-MC-CDMA systems are investigated, where a big number of potential UEs is supported and during each time-slot, each UE becomes active to transmit with a small probability. Hence, the active UEs are also highly dynamic.
- A low-complexity MMSE-SICD is investigated for the signal detection in DyDS-CDMA and DyMC-CDMA systems, in order to achieve the detection that is robust to the dynamics of these system, while without making much trade-off of the system performance.

- The bit error rate (BER) performance of DyDS-CDMA and DyMC-CDMA systems is investigated with assumption that DyDS-CDMA signals experience flat fading, while DyMC-CDMA signals experience frequency-selective fading.

- Furthermore, the BER performance of DyMC-CDMA systems is investigated and compared, when both dense spreading and sparse spreading are respectively applied. Note that, the sparse spreading is specifically introduced to DyMC-CDMA systems, because DyMC-CDMA systems may experience the peak-to-average power ratio (PAPR) problem [127].

Our studies demonstrate that the MMSE-SICD is highly efficient for operation in the considered dynamic systems. Furthermore, the DyMC-CDMA scheme with MMSE-SICD has low-complexity and is robust to the system's dynamics and is feasible for achieving the frequency-selective diversity. Hence, the DyMC-CDMA scheme with MMSE-SICD constitutes a promising scheme for operation in mGFMA systems.

Chapter 3: Channel Estimation and User Activity Identification in Massive Grant-Free Multiple-Access

In **Chapter 3**, we study the channel estimation and UAI in the mGFMA scenario, where the number of potential UEs may be much bigger than the number of resource units and each UE has a small activation probability. Hence, the active UEs and the number of them are highly dynamic over time, and the number of active UEs at a time may be higher than the number of resource units. Specifically, we first consider the channel estimation of both active and inactive UEs in the principle of minimum mean-square error (MMSE), when AP is assumed to have the ideal knowledge, no knowledge or the partial knowledge of active UEs. In our channel estimation, AP will make the best use of the information provided by the transmission of pilot symbols and data (payload) symbols.

The main contributions of this chapter can be summarized as follows:

- Channel estimation in dynamic multicarrier mGFMA (MC-mGFMA) scenarios is investigated, where the channels of both active and inactive UEs are estimated in MMSE principle, when AP is assumed to employ ideal knowledge, no knowledge or partial knowledge about the active UEs. It is shown that the proposed channel estimation approaches are robust to the dynamics in mGFMA systems. Furthermore, it is shown that the added knowledge about active UEs results in the enhanced reliability of channel estimation. These properties can be exploited for the design of UAI algorithms.

- The statistical properties of the estimated channels of both active and inactive UEs

are analyzed, showing that the estimated channels of active UEs and that of inactive UEs have distinct features, which can be exploited for distinguishing active UEs from inactive UEs.

- Hence, based on the studies of channel estimation and of the characteristics of estimated channels, five UAI algorithms are proposed for MC-mGFMA and their performance is studied. These UAI algorithms are listed as follows:

  a. Threshold-based UAI (TB-UAI);

  b. Eigen-analysis enhanced UAI (EAE-UAI);

  c. Successive interference cancellation (SIC) assisted UAIa (SIC-UAIa);

  d. SIC assisted UAIb (SIC-UAIb);

  e. Auto-correlation matrix evolving UAI (AME-UAI).

- Furthermore, These UAI algorithms have their distinctive characteristics and present different performance-complexity trade-off. They however have one common feature that all of them are suitable for operation in the mGFMA systems having highly dynamic UEs in terms of the active UEs and the number of them.

Our studies show that any added information about the active UEs can be exploited to enhance the reliability of channel estimation. This is reflected in practical mGFMA systems that the activity patterns of UEs may be correlated or one active UE may have several packets to transmit in continuous time-slots. Then, this kind of information can be exploited to enhance the reliability of channel estimation in mGFMA systems.

Chapter 4: Joint Channel Estimation, User Activity Identification and Information Detection

In **Chapter 4**, we investigate a multicarrier massive grant-free multiple-access (MC-mGFMA) system by assuming that a big number of highly dynamic user equipments (UEs) are monitored by an access point (AP) with multiple receive antennas (MRA), which is referred to as the MRA/MC-mGFMA system. The channel estimation, UE activity identification (UAI) and information detection are addressed in the MRA/MC-mGFMA system. First, the channels of both active and inactive UEs are detected in the principle of minimum mean-square error (MMSE). Then, based on the estimated channels, a low-complexity threshold-based UAI (TB-UAI) is proposed to detect the activities of UEs. Finally, information of active UEs is detected in the principle of the successive interference cancellation (SIC) assisted MMSE (SIC-MMSE). Furthermore, a joint algorithm, referred to as SIC-MMSE-JCUD, is proposed to carry out channel estimation, UAI and information detection

jointly in the principle of SIC-MMSE. Additionally, considering that no set of the well-designed signature sequences can support the big number of potential UEs in mGFMA systems, we propose a class of sequences designed by combining the Gold-sequences with the Zadoff-Chu (ZC) sequences. The performance of different schemes is studied and compared based on Monte-Carlo simulations. Our studies show that deploying multiple receiver antennas at AP is beneficial to the channel estimation, UAI and information detection. Aided by the multiple receive antennas of AP, a low-complexity TB-UAI algorithm is highly efficient for UE activity identification. Furthermore, our proposed class of signature sequences allows to attain much better performance than the random sequences.

The main contributions of this chapter can be summarized as follows:

- We study a multicarrier mGFMA system where each AP employs multiple receive antennas, which is referred to as the MRA/MC-mGFMA system for convenience of description.
- We extend our studies in [128] to the MRA/MC-mGFMA system, and investigate the impact of multi-antenna AP on the design and performance of channel estimation, UAI and information detection.
- Jointing the ZC sequence and gold-sequence, we design a class of signature sequences so as to support the possibly huge number of UEs in a mGFMA systems. We also investigate the effect of the proposed class of sequences on the performance of channel estimation, UAI and information detection.
- We propose the channel estimation, UAI, signal detection and joint channel estimation, UAI and information detection algorithms for the proposed MRA/MC-mGFMA system. Specifically, our channel estimation is in the principle of minimum mean-square error (MMSE) under the assumption that AP has full knowledge, partial knowledge or no knowledge about the active UEs. Then, assuming that AP has no knowledge about active UEs, we study the threshold-based UAI (TB-UAI) for UE activity identification. Finally, based on the channel estimation and/or TB-UAI, the successive interference cancellation (SIC) assisted MMSE (SIC-MMSE) and the SIC-MMSE relied joint channel estimation, UAI and data detection (SIC-MMSE-JCUD) are designed.
- Based on numerical simulations, we study and compare the performance of the channel estimation, UAI and information detection algorithms, when various aspects are considered.

Our studies show that when multiple receive antennas at APs are available, the low-complexity threshold-based UAI (TB-UAI) is capable of achieving promising performance, which significantly improved with the increase of the number of receive antennas. Furthermore, as the number of receive antennas of AP increases, the threshold setting in TB-UAI becomes relatively easier. This implies that the performance of TB-UAI becomes less sensitive to the threshold applied, and the near-optimum threshold can be relatively easily achieved, as the number of receive antennas of AP increases. In terms of the signature sequences, the studies show that the proposed class of signature sequences outperforms the random sequences. Furthermore, for signal detection, both SIC-MMSE and SIC-MMSE-JCUD are highly effective for operation in the MRA/MC-mGFMA systems, where both active UEs and the number of them are highly dynamic.

Chapter 5: Channel Estimation, User Activity Identification and Data Detection in Massive Distributed Antenna Grant-Free Multiple Access Systems

In **Chapter 5**, we consider the massive distributed grant-free multiple-access (MDR-GFMA), where RRHs/APs are distributed in the communications area. Each RRH/AP only serves the UEs around it, while each UE is only connected with a few of RRHs/APs around it. Furthermore, the signals received by an RRH/AP from different UEs have different power. All the above-mentioned settings/assumptions impose new challenges to the design of the algorithms for the UAI and data detection in MDR-GFMA systems. Hence, in this chapter, we are motivated to study the low-complexity while still efficient methods for the channel estimation, UAI and data detection in the MDR-GFMA systems. The main contributions and novelties can be summarized as follows:

- A MDA-GFMA system is proposed and investigated, where the RRHs/APs (or simply distributed antennas (DAs)) are distributed based on the Poisson point (PP) distribution and UEs are randomly and uniformly distributed in the coverage area. In the considered MDA-GFMA system, RRHs/APs are assumed to be connected to a SPCU, where signal processing is carried out. UEs become active randomly with a small activation probability. Each of the active UEs is only connected to a few of the DAs close to the UE. Hence, no power control is needed as there is no path-loss, the operation is UE centric and the system is cell-free.

- In terms of signal processing, first, the MMSE-assisted channel estimation is implemented to estimate the channels of UEs, regardless of them being active or inactive, with assumption that the SPCU has different levels of knowledge about the large-scale fading, including propagation path-loss and shadowing, experienced by the UEs. Furthermore, the characteristics of the estimated channels for active and inactive UEs are analyzed and demonstrated.

- Second, based on the MMSE-relied channel estimation, an orthogonal matching pursuit assisted UAI, referred to as the OMP-MMSE-UAI, algorithm is proposed and investigated, when each of UEs is assumed to be served by the RRHs/APs within a given area centred at the UE. The OMP-MMSE-UAI provides the initial UAI, which is further enhanced via the detection of the pilot symbols expected for different UEs.

- Third, after the channel estimation and UAI, the data transmitted by active UEs are detected. Both MMSE detection and the successive interference cancellation assisted MMSE (MMSE-SIC) detection are introduced, again, by assuming that each active UE is served by a few of RRHs/APs around the UE.

- The performance of channel estimation, UAI and data detection in MDR-GFMA systems is investigated with the aid of Monte-Carlo simulations. The results demonstrate that the proposed algorithms are effective. The MMSE-based channel estimation is low-complexity and is capable of achieving promising performance, which significantly improves with the increase of the density of the distributed RRHs/APs. The OMP-MMSE-UAI algorithm can effectively limit the miss of active UEs, while avoiding the false-alarms of inactive UEs. Furthermore, the MMSE-SIC is not only efficient for operation, but for performance in the MDR-GFMA systems.

Our studies show that the MMSE-assisted channel estimation is efficient for operation in the MDR-GFMA systems. Explicitly, the estimated channels of active UEs and inactive UEs have significant difference in terms of the MSE performance. This means that the estimated channels for a UE may be exploited to identify whether the UE is active. Secondly, in our OMP-MMSE-UAI algorithm, the embedded OMP algorithm is capable of providing an initial UAI with a small miss probability, while the followed UAI enhancement aided by the pilot detection can efficiently identify the false-alarmed UEs. Furthermore, concerning the two detection schemes considered, the MMSE-SIC detector is highly efficient for data detection in MDR-GFMA systems, which experience a range of dynamics such as the active UEs and the number of them, the geography resulted randomness, and the number of DAs serving a UE.

Finally, in **Chapter 6**, the conclusions derived from the research are summarized. Furthermore, some future works related to the research are discussed.

Chapter 2

Signal Detection in Dynamic DS-CDMA and MC-CDMA Systems

In this chapter, we study the dynamic direct sequence code-division multiple-access (DyDS-CDMA) and multicarrier CDMA (DyMC-CDMA) systems, where the dynamic property is typical in the massive grant-free multiple-access (mGFMA) systems to be considered in the following chapters. Specially, this chapter assumes that each base-station (BS) or access point (AP) with limited degrees of freedom is expected to control a massive number of devices, called user equipments (UEs). We assume that UEs become active to transmit data on the basis of time-slots and during a time-slot, each of UEs activates independently and following a common activation probability. While for the DyDS-CDMA systems we assume flat Rayleigh fading, we assume that the DyMC-CDMA systems experience frequency-selective Rayleigh fading. A low-complexity detector based on minimum mean-square error and successive interference cancellation, referred to as the MMSE-SICD, is proposed for signal detection in DyDS-CDMA and DyMC-CDMA systems. Assuming ideal active UE identification and ideal channel estimation of active UEs, in this chapter, we focus on analysing/measuring performance achievable by the DyDS-CDMA and DyMC-CDMA systems employing the MMSE-SICD, to demonstrate the feasibility of MMSE-SICD for operation in mGFMA systems. Our studies reveal that the DyDS-CDMA and DyMC-CDMA systems aided by the MMSE-SICD are high-efficiency for operation in mGFMA environments. Near single-user performance is achievable, when the average number of active UEs per time-slot reaches two times of the system's degrees-of-freedom, and hence, the number of active UEs of a time-slot may be much higher than two times of the system's degrees-of-freedom.

2.1 Introduction

Driven by the unaccountable IoT applications, the 5th generation (5G) wireless networks are expected to support numerous devices and provide massive connectivity [9]. As one of the results, many 5G wireless systems will be device-centric, constituted by densely distributed devices and machines, in contrast to the human-centric 1-4G wireless systems. As most of devices/machines carry out the sensing, data collection, etc., tasks, and they often need to report via wireless channels by sending short-burst messages, we may expect that in the device-centric 5G wireless systems, uplink traffics will become dominate [12].

In the device-centric wireless networks, explicitly, the legacy human-centric systems, such as LTE/LTE-A in 4G systems, will not be efficient for supporting the sporadic short-burst applications required by IoT devices (and machines). Instead, novel communications protocols need to be designed to allow a huge number of ultra-densely distributed devices to communicate with their base-stations (or access points (APs)) reliably and efficiently [10, 11]. Furthermore, in these device-centric wireless networks, all devices are expected to be operated with high energy-efficiency, which for some applications is critical due to the practical constraint on battery changing, battery re-charging, etc. In order to meet the above-mentioned challenges, grant-free multiple-access (GFMA) has been proposed to support sporadic short-burst applications [61]. Specifically in [62], the authors have compared the performance of transmission of short-burst data blocks in both traditional LTE network and in new 5G grant-free networks. The studies show that data transmission in grant-free systems should be in the manner of arrive-and-go, and that using the traditional communication schemes to transmit sporadic short-burst traffic is inefficient.

For uplink multiple-access in 5G IoT networks, in literature, there are a range of optional techniques proposed. As some examples, in [70], Tararefa *et al*. have considered the FDMA scheme supported by the joint user association and resource allocation for IoT uplink. Sutton *et al*. [71] investigated the performance of a TDMA-assisted IoT uplink. In [72], Hu *et al*. have introduced the cognitive radio principles to CDMA, in order to address the problem of resource shortage in IoT systems. Furthermore, in [73], the non-orthogonal multiple-access (NOMA) technique has been introduced for supporting IoT cloud services. Specifically, in the domain of GFMA, the authors of [94] have proposed the sparse code multiple-access (SCMA), which is one of NOMA techniques, for supporting many devices to share limited resources. In [129], the authors have investigated joint active user identification and information detection in GFMA systems. By contrast, the authors in [95] have proposed a coded tandem spreading multiple access (CTSMA) scheme for enhancing the multiple-access capability of GFMA systems. Furthermore, in [96], a so-called comprehensive grant-free random access scheme has been proposed for machine-

type communications, which allows synchronization, channel estimation, user identification and data detection are all performed in one single shot.

In the context of the signal detection in GFMA systems, various techniques have been proposed based on the factor graph dependent sparse code multiple access (SCMA) [62] and compressive sensing [118]. While both the techniques employ a range of advantages, as argued in the above-mentioned references, there are also critical limitations for them to apply in the GFMA systems, in particular, of supporting ultra-dense devices requiring massive connectivity, i.e., the massive GFMA (mGFMA). While the achievable performance of SCMA systems is depended closely on the well-structured factor graphs, the recovery performance of compressive sensing relying systems is limited by the restricted isometric property (RIP) condition [130–132], which constrains that the number of active devices should be significantly lower than the degrees of freedom (or the number of resource units). However, in the mGFMA systems, a typical characteristic is that the number of active devices is highly dynamic and time-varying, making the design of high-efficiency radio receiver for APs highly challenging. For example, consider an AP supporting 1000 potential devices each having an independent and uniform activation probability of 0.05. Assume that data is transmitted based on time-slots, and devices become active and start transmission at the start of time-slots. Then, we can find that the average number of active devices per time-slot is $1000 \times 0.05 = 50$. The probabilities that the number of active devices per time-slot is more than 50 and more than 100 are 0.4625 and 0.0207, respectively, both of which may significantly affect system's reliability, if they are not coped with carefully in design. Explicitly, a mGFMA system would not be efficient, which is designed based on the average number of active devices of 50, which results in low-reliability in case there is a big number of active devices. It is also inefficient if it is designed based on 100 or more active devices for attaining the required reliability, which however results in low-efficiency resource utility, due to the fact that the number of active devices is often significantly lower than 100. Instead, a mGFMA system should be designed to maximize the resource utility, while enabling high-reliability and robust operation in the highly dynamic IoT environments.

In the mobile communication system, when different users are ready to utilize the same transmission resource, their signals must be generated following a specific multiple access scheme, such that the signals of the different users can be separated at the receiver. The above discussed detector design based on DS-CDMA as a multiple access technique. Furthermore, multi-carrier systems offer a good bandwidth efficiency with immunity to channel dispersion. In OFDM, the orthogonal property can be held as long as the subcarrier frequency separation maintains the integer multiplex over one inverse of OFDM symbol duration. The combination of OFDM and DS-CDMA is MC-CDMA. In MC-

CDMA technique, the original data stream are first multiplied with the spreading sequence and then modulated on the different carriers. Moreover, in GFMA system, the transmitting data symbol durations become obviously smaller than the delay due to the multi-path channel. MS-CDMA is a robust solution to deal with the inter symbol interference (ISI) and to deal with the the capacity receded by MAI. The research in has been proved [133], the probability of bit error in MC-CDMA system with 16 users is the same as the single user case in the DS-CDMA, where it is assumed that the time is perfectly synchronized at the receiver in both systems. Thus, the transmitter of each user derives a clock signal and a carrier oscillator signal from a network synchronization reference signal [134]. Due to these properties, it makes MC-CDMA system enable to be a candidate to consider in GFMA implementation. It motivate us to discuss channel estimation, UAI and information detection relied on MC-CDMA system from the next chapter.

Based on the above consideration, in this chapter, we are motivated to study a multiple-access protocol for mGFMA and to design a corresponding low-complexity detection algorithm. We do not impose the limit on the number of active devices per time-slot, but aim to provide robust detection for any active devices no matter how many of them there are. Our detection algorithm is also not dependent on factor graphs, as well-structured factor graphs in mGFMA are often unavailable. Alternatively, we return to the convention principles of successive interference cancellation (SIC), and design a high-efficiency SIC scheme for supporting dynamic DS-CDMA (DyDS-CDMA) to achieve mGFMA. In this chapter, we assume ideal active device identification and also ideal channel estimation of active devices, but focus our attention on signal detection in mGFMA environments, in order to demonstrate the potentials of DyDS-CDMA. For this purpose, a detector based minimum mean-square error and SIC, referred to as the MMSE-SICD, is developed. Finally, the performance of the DyDS-CDMA systems employing the MMSE-SICD is investigated and compared from different perspectives based on Monte-Carlo simulations. Our studies and performance results show that the DyDS-CDMA aided by the proposed SIC is high-efficiency for operating in mGFMA environments, which allows to achieve near single-user performance, even when the average number of active devices per time-slot reaches two times of the system's degrees-of-freedom.

The remainder of this chapter is outlined as follows. In Section 2.2, we describe the system model for both DyDS-CDMA and DyMC-CDMA systems. Section 2.3 addresses the signal detection by proposing the MMSE-SICD algorithm. Section 2.4 demonstrates the performance results of both the DyDS-CDMA and DyMC-CDMA systems with various considerations. Finally, we derive the conclusions from research in Section 2.5.

2.2 Description of Dynamic DS-CDMA and MC-CDMA Systems

We consider a single cell with a base-station (BS), or access point (AP), locating at the center and K potential mobile UEs randomly distributed in the cell. The communications between these UEs and BS is based on DS-CDMA or MC-CDMA [72, 127, 135]. We assume that the UEs transmit information to BS on the basis of time-slots. Given a time-slot, we assume that each UE has a probability of P_a to become active to transmit a frame of data. Furthermore, we assume that the events of activation are independent in terms of time-slots and UEs. Based on the above assumptions, it can be shown that, for a given time-slot, the number of active UEs, K_A, obeys the binomial distribution, given by

$$P(K_A) = \binom{K}{K_A} P_a^{K_A}(1-P_a)^{K-K_A},\ K_A = 0,1,\ldots,K \tag{2.1}$$

Let us express the spreading factor of both DyDS-CDMA and DyMC-CDMA as N. Since the activation probability P_a of a UE is usually very small, about $0.05 \sim 0.1$ [136], N is usually designed to be significantly smaller than K. Therefore, from (2.1) we are implied that, given a considered time-slot, the number of active UEs, i.e. K_A, can be either lower or higher than N, resulting in the system either under-load or over-load. In literature, such as in [137], CS has typically been considered for active UEs identification and their signal detection, when GFMA is considered. However, the successful recovery under CS has to be constrained by the restricted isometry property (RIP) condition [130–132]. This condition makes the CS approaches in general only valid, when the number of active UEs is significantly lower than the number of resource units. In other words, CS approaches are not suitable for the actually overload scenarios, where the number of active UEs is higher than the number of resource units. Specifically for the DyDS-CDMA or DyMC-CDMA systems considered in this chapter, CS approaches are only suitable for the cases, when K_A is significantly lower than N.

More specifically, when given a spreading factor N of the DyDS-CDMA or DyMC-CDMA, we can readily find that the probability of that the systems are overload is

$$P(K_A > N) = \sum_{k=N+1}^{K} \binom{K}{k} P_a^k(1-P_a)^{K-k} \tag{2.2}$$

For example, let us assume $P_a = 0.05$ and $N = 32$. Then, it can be shown that $P(> N) = 0.9963$ if $K = 1000$, and $P(> N) \approx 1$ if $K = 1500$. Therefore, in the DyDS-CDMA or DyMC-CDMA systems having ultra-dense UEs, making the number of UEs per cell above a thousand, the systems will be surely always overloaded.

In this chapter, for the sake of simplicity but focusing our attention on principles, we assume that the signals sent from UEs to BS experience flat Rayleigh fading in DyDS-CDMA systems[1], while experience frequency-selective fading in DyMC-CDMA systems. Let the spreading code assigned to the kth UE be expressed as $\boldsymbol{c}_k = [c_{k0}, c_{k1}, \ldots, c_{k(N-1)}]^T$ ($k = 1, 2, \ldots, K$), which may be used as the identity of UE k during the stage of active UEs identification, as shown in the following chapters, and as the spreading sequence for the kth UE's data transmission. Let us consider a time-slot, and define a set of corresponding active UEs as $\mathcal{I}_A$, which contains the indices of the active UEs. We assume low-rate communication UEs, all of which use binary-phase shift-keying (BPSK) modulation in order to save energy. Then, when the DyDS-CDMA systems are considered, the received discrete signals at BS corresponding to the active UEs can be expressed as

$$\boldsymbol{y} = \sum_{k \in \mathcal{I}_A} \sqrt{P_k} h_k \boldsymbol{c}_k b_k + \boldsymbol{n} \tag{2.3}$$

where $\boldsymbol{y}$ is the N-length observation vector, b_k denotes a binary bit sent by the kth UE, h_k represents the kth channel's fading gain, which obeys the complex Gaussian distribution with zero mean and unit variance, P_k denotes the power received from UE k, while $\boldsymbol{n}$ is a N-length Gaussian noise vector distributed with zero mean and a covariance matrix of $\sigma^2 \boldsymbol{I}_N$ with $\boldsymbol{I}_N$ being a $(N \times N)$ identity matrix.

For convenience of following description, we represent (2.3) in a compact form of

$$\begin{aligned} \boldsymbol{y} =& \boldsymbol{CGPb} + \boldsymbol{n} \\ =& \boldsymbol{Hb} + \boldsymbol{n} \end{aligned} \tag{2.4}$$

where by definition, $\boldsymbol{C} = [\boldsymbol{c}_1, \boldsymbol{c}_2, \cdots, \boldsymbol{c}_{|\mathcal{I}_A|}]$, where $|\mathcal{I}_A|$ denotes the number of elements in $\mathcal{I}_A$, $\boldsymbol{G} = \mathrm{diag}\{h_1, h_2, \cdots, h_{|\mathcal{I}_A|}\}$, $\boldsymbol{P} = \mathrm{diag}\{\sqrt{P_1}, \sqrt{P_2}, \cdots, \sqrt{P_{|\mathcal{I}_A|}}\}$, $\boldsymbol{b} = [b_1, b_2, \cdots, b_{|\mathcal{I}_A|}]^T$, and finally, $\boldsymbol{H} = \boldsymbol{CGP}$.

In the context of the DyMC-CDMA, we assume that it is operated in the OFDM principle [127] with sufficient length of cyclic prefixing (CP). The frequency-selective fading channel is assumed to have L number of resolvable paths in the time-domain, whose time-domain impulse response can be expressed as $\boldsymbol{h}_{k,t} = [h_{k1,t}, h_{k2,t}, \ldots, h_{kL,t}]^T$, $k = 1, 2, \ldots, K$, where $h_{kl,t}$ is independent complex Gaussian distributed with zero mean and variance of $1/L$. Correspondingly, the frequency-domain channel gains experienced by the N subcarriers can be obtained as [127]

$$\boldsymbol{h}_k = \boldsymbol{F}_N \boldsymbol{\Psi}_{N \times L} \boldsymbol{h}_{k,t} \tag{2.5}$$

[1]Note that, in frequency-selective fading channels, when the block-based detection in DyDS-CDMA systems is considered, the detection principle is the same as that considered in this chapter.

where $\mathbf{\Psi}_{N\times L}$ is a $N \times L$ mapping matrix obtained from the first L columns of the identity matrix $\boldsymbol{I}_N$, and $\boldsymbol{F}_N$ is a N-point fast Fourier transform (FFT) matrix. Based on the above assumptions and assuming BPSK modulation, it can be shown that an observation equation corresponding to (2.3) can be written as

$$\boldsymbol{y} = \sum_{k\in\mathcal{I}_A} \sqrt{P_k}\boldsymbol{H}_k\boldsymbol{c}_k b_k + \boldsymbol{n} \tag{2.6}$$

where $\boldsymbol{H}_k = \text{diag}(\boldsymbol{h}_k)$ is a $N \times N$ diagonal matrix.

Corresponding to (2.4), (2.6) can also be represented in the forms of

$$\boldsymbol{y} = \sum_{k\in\mathcal{I}_A} \tilde{\boldsymbol{h}}_k b_k + \boldsymbol{n} \tag{2.7}$$

$$= \boldsymbol{H}\boldsymbol{b} + \boldsymbol{n} \tag{2.8}$$

where $\tilde{\boldsymbol{h}}_k = \sqrt{P_k}\boldsymbol{H}_k\boldsymbol{c}_k$, $\boldsymbol{H} = \left[\tilde{\boldsymbol{h}}_1, \tilde{\boldsymbol{h}}_2, \ldots, \tilde{\boldsymbol{h}}_{|\mathcal{I}_A|}\right]$, and $\boldsymbol{b}$ is the same as that previously defined for DyDS-CDMA.

For the spreading codes assigned to mobile UEs, for DyDS-CDMA systems, we simply assume that they are binary random codes having the elements of $\pm 1/\sqrt{N}$. By contrast for DyMC-CDMA systems, we investigate two classes of random codes. The first class is the same as that assumed for DyDS-CDMA systems. By contrast, considering that the multicarrier signaling in the DyMC-CDMA results in the PAPR problem [127], the second class of spreading codes are the sparse spreading codes, which contain a big fraction of zeros in addition to the binary elements of $+1$ and -1. Specifically, we still assume the random spreading codes, but an element of which has a small probability to be $+1$ or -1, while has the rest probability to be 0. For example, each element has a probability of $1/10$ to be $+1$, of $1/10$ to be -1, and of $8/10$ to be zero, making the average sparsity of the N-length sequence to be $N \times 2/10 = 0.2N$. The average sparsity of spreading sequences is expressed as $\bar{d}_x$. Additionally, the sparse spreading sequences are all normalized to have unity energy of $\|\boldsymbol{c}_k\| = 1$. Explicitly, with the employment of sparse spreading codes, the PAPR problem of DyMC-CDMA systems may be effectively mitigated. However, the sparsity of spreading codes may make a trade-off with the achievable performance of DyMC-CDMA systems. Hence, in this chapter, we investigate the effect of the spreading codes' sparsity on the achievable performance of DyMC-CDMA systems, when communicating over frequency-selective fading channels. Our later performance study will show that when given the frequency-selectivity (such as L resolvable paths of frequency-selective fading channels), the near-full frequency-diversity gain can only be achievable, when the sparsity of spreading sequences is sufficiently high, usually slightly higher than L.

Both the DyDS-CDMA and DyMC-CDMA considered in this chapter are studied in GFMA environments [138]. In these systems, UEs send BSs information without handshaking with the BSs. In this case, closed-loop power-control [139] is impossible. However, in such DyDS-CDMA or DyMC-CDMA systems, BSs may send pilot signals periodically, such as, for mobile UEs to carry out required synchronization, and estimate their distances from BSs through measuring the pilot signals' strength. Therefore, mobile UEs are able to execute open-loop power-control [139] based on the strength of received pilot signals. With the above consideration, we assume that open-loop power-control is implemented. Correspondingly, we assume that the power received by BS from a mobile UE obeys a truncated Gaussian distribution represented as

$$f_{P_k}(p) = \mu \exp\left(-\frac{(p-P_0)^2}{2\sigma_P^2}\right), \quad \max\{0, P_0 - P_E\} \leq p \leq P_0 + P_E \tag{2.9}$$

where P_E denotes the maximum power-control error with respect to P_0 of the received power in ideal power-control, and μ is a constant to ensure that the integration of $f_{P_k}(p)$ over the range of $\max\{0, P_0 - P_E\} \leq p \leq P_0 + P_E$ is one.

Let us below consider the signal detection of the active UEs.

2.3 Signal Detection of Active UEs

In our considered dynamic CDMA systems with the motivation for GFMA, the number of active UEs accessing a BS during a time-slot is highly dynamic. In the references addressing GFMA, CS algorithms have been typically introduced for signal detection [140]. However, the recovery capability of CS-relying detectors is constrained by the RIP condition [130–132]. This constraint implies that the system should be significantly under-load, meaning that $K_A << N$, in order to achieve the satisfactory performance of signal recovery. However, in practice when there is a big number of potential UEs of K, such as over 1000 per BS, there is usually a very high probability that the number of active UEs exceeds the recovery capability of the CS-relying detectors. Furthermore, in the ultra-dense cases, as previous examples show, it is nearly the certain event that the system becomes overload, i.e., $K_A > N$. In these cases, the CS-relying detectors are incompetent. Therefore, in practice, we should design the detection algorithms, which are efficient for operation in the dynamic user environment where the constraint on the number of active UEs is hardly satisfied. With the above motivation, in this chapter, we return to the conventional SIC-assisted multiuser detection, to introduce a SIC based detector operated with the minimum mean-square error principles [141], which is referred to as the MMSE-SICD. We extend it

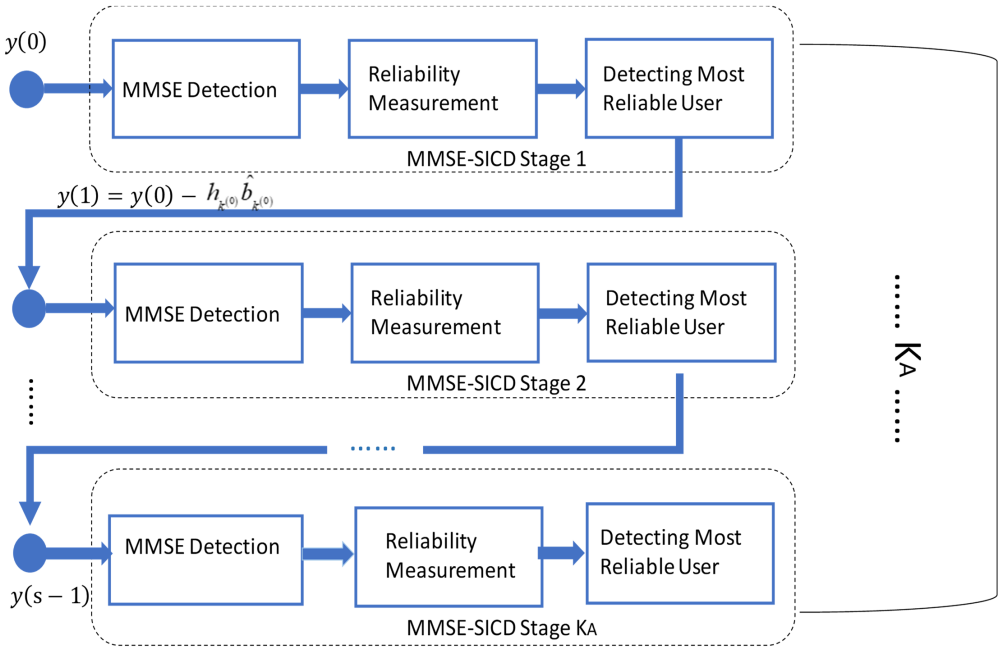


Figure 2.1: Demonstration of the MMSE-SICD for DyDS-CDMA and DyMC-CDMA systems.

for operation in the dynamic communication environment, where the number active UEs is a random variable obeying the binomial distribution of (2.1). Correspondingly, we investigate its efficiency and demonstrate the effect of UEs' dynamic behavior on the achievable performance.

The MMSE-SICD considered is an iterative multiuser detector. During an iteration, MMSE-based multiuser interference suppression (MMSE-MUIS), reliability measurement to identify a most reliable UE, and interference cancellation are respectively operated, the principles of which are detailed in the following subsections.

2.3.1 MMSE-Based Multiuser Interference Suppression

Below the MMSE-SICD is analyzed in the context of the DyDS-CDMA with the observation equation of (2.4). The analysis of the MMSE-SICD for the DYMC-CDMA with the observation equation of (2.8) is the same.

Consider the observation equation of (2.4) and assume that BS employs the knowledge of $\boldsymbol{H}$, i.e., of $\boldsymbol{C}$, $\boldsymbol{G}$ and $\boldsymbol{P}$. Then, the decision variables for detecting the data of active UEs

in the sense of MMSE can be expressed as [127]

$$\begin{aligned}\boldsymbol{z} =& \boldsymbol{H}^H \left(\boldsymbol{H}^H\boldsymbol{H} + \sigma^2\boldsymbol{I}_{|\mathcal{I}_A|}\right)^{-1}\boldsymbol{y} \\ =& \boldsymbol{H}^H\boldsymbol{R}_y^{-1}\boldsymbol{y}\end{aligned} \tag{2.10}$$

where we defined $\boldsymbol{R}_y = \boldsymbol{H}^H\boldsymbol{H} + \sigma^2\boldsymbol{I}_{|\mathcal{I}_A|}$ for convenience of following description. If we apply $\boldsymbol{H} = [\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_{|\mathcal{I}_A|}]$, where $\boldsymbol{h}_k$ is for UE k, the decision variables for the active UEs can be respectively written as

$$z_k = \boldsymbol{h}_k^H\boldsymbol{R}_y^{-1}\boldsymbol{y},\ k = 1, 2, \ldots, |\mathcal{I}_A| \tag{2.11}$$

It can be understood that, if the $|\mathcal{I}_A|$ UEs' information is detected based on (2.11), the information of different UEs will be detected with different reliabilities, contributed by channel fading, Gaussian noise and the different power received from different UEs. Therefore, it is desirable to enhance the detection performance with the aid of SIC by exploiting the reliability information revealed from (2.11).

2.3.2 Reliability Measurement

In [142], a near-optimum reliability measurement scheme was proposed, which allows a CDMA (or SDMA) system of N degrees of freedom to achieve near-optimum BER performance over Rayleigh fading channels, even when the number of users simultaneously detected is as high as $2N$, which corresponds to a system loading factor of two. Based on the reliability measurement scheme proposed in [142], the reliability of detection of the kth active UE can be estimated as

$$\begin{aligned}L_k &= (1 + \tilde{\gamma}_k)|z_k| \\ &= \left(\frac{1}{1 - \boldsymbol{h}_k^H\boldsymbol{R}_y^{-1}\boldsymbol{h}_k}\right)|z_k|,\ k = 1, 2, \ldots, K_{|\mathcal{I}_A|}\end{aligned} \tag{2.12}$$

where $\tilde{\gamma}_k = \boldsymbol{h}_k^H\boldsymbol{R}_y^{-1}\boldsymbol{h}_k / (1 - \boldsymbol{h}_k^H\boldsymbol{R}_y^{-1}\boldsymbol{h}_k)$ is the signal-to-interference-plus-noise ratio (SINR), if b_k is detected based on z_k of (2.11).

Once the reliabilities of the active UEs to be detected are available, the MMSE-SICD can be carried out, which is described as follows.

2.3.3 MMSE-SICD Algorithm

When there are $|\mathcal{I}_A|$ UEs to detect, the MMSE-SICD needs to execute $|\mathcal{I}_A|$ iterations, with each iteration detecting one of the $|\mathcal{I}_A|$ active UEs. As mentioned previously, during

an iteration, the MMSE-MUIS, reliability measurement to identify the most reliable UE, detection of the most reliable UE, and interference cancellation of the most reliable UE from the other not detected UEs are respectively executed. In summary, the MMSE-SICD algorithm can be stated as follows.

Algorithm 2.3.3 MMSE-SICD Algorithm.

Initialization: $\boldsymbol{y}^{(0)} = \boldsymbol{y}, \boldsymbol{R}_y^{(0)} = \boldsymbol{R}_y, \boldsymbol{H}^{(0)} = \boldsymbol{H}, \boldsymbol{W}^{(0)} = \boldsymbol{R}_y^{-1}\boldsymbol{H}$.

Detection:
for $s = 1, 2, \ldots, |\mathcal{I}_A|$,

1. *Computing the decision variables for the UEs not detected*:
   $\boldsymbol{z}^{(s)} = \Re\{(\boldsymbol{W}^{(s-1)})^{\boldsymbol{H}}\boldsymbol{y}^{(s-1)}\}$, where $\boldsymbol{W}^{(s-1)} = (\boldsymbol{R}_y^{(s-1)})^{-1}\tilde{\boldsymbol{H}}^{(s-1)}$, $\Re\{x\}$ returns the real-part of x.

2. *Determining the most reliable UE among the UEs not detected*: Among the $|\mathcal{I}_A| - s + 1$ UEs having not been detected, compute their reliabilities according to (2.12), and then find the most reliable UE as
   $k^{(s)} = \text{argmax}_{k_i'}\{L_{k_1'}, L_{k_2'}, \ldots, L_{k'_{|\mathcal{I}_A|-s+1}}\}$.

3. *Detecting the most reliable UE*: $\hat{b}_{k^{(s)}} = \text{sgn}(z_{k^{(s)}})$, where $z_{k^{(s)}}$ is the k_sth entry of $\boldsymbol{z}^{(s)}$.

4. *Interference cancellation*:
   $\boldsymbol{y}^{(s)} = \boldsymbol{y}^{(s-1)} - \boldsymbol{h}_{k^{(s)}}\hat{b}_{k^{(s)}}$;

5. *Update*: $\boldsymbol{W}^{(s)} \leftarrow \boldsymbol{W}^{(s-1)}$.

In the above algorithm, the weight matrix $\boldsymbol{W}^{(s-1)}$ needs to be updated to $\boldsymbol{W}^{(s)}$, which has the formula of $\boldsymbol{W}^{(s)} = (\boldsymbol{R}_y^{(s)})^{-1}\tilde{\boldsymbol{H}}^{(s)}$. After the interference cancellation at the sth iteration, we have

$$\tilde{\boldsymbol{H}}^{(s)} = \tilde{\boldsymbol{H}}^{(s-1)}\boldsymbol{P}^{(s)}, \tag{2.13}$$

$$\boldsymbol{R}_y^{(s)} = \boldsymbol{R}_y^{(s-1)} - \boldsymbol{h}_{k^{(s)}}(\boldsymbol{h}_{k^{(s)}})^H \tag{2.14}$$

where $\boldsymbol{P}^{(s)}$ is a permutation matrix obtained from $\boldsymbol{I}_{|\mathcal{I}_A|}$ by removing from it the s columns corresponding to the s active UEs having detected. Then, by using (2.13), we can express

$\boldsymbol{W}^{(s)}$ as

$$\begin{aligned}\boldsymbol{W}^{(s)} &= \left(\boldsymbol{R}_y^{(s)}\right)^{-1} \tilde{\boldsymbol{H}}^{(s)} \\ &= \left[\boldsymbol{R}_y^{(s-1)} - \boldsymbol{h}_{k^{(s)}} (\boldsymbol{h}_{k^{(s)}})^H\right]^{-1} \tilde{\boldsymbol{H}}^{(s-1)} \boldsymbol{P}^{(s)} \end{aligned} \tag{2.15}$$

Upon applying the matrix inverse lemma, and expressing $\boldsymbol{w}_{k^{(s)}}^{(s-1)} = \left(\boldsymbol{R}_y^{(s-1)}\right)^{-1} \boldsymbol{h}_{k^{(s)}}$, we obtain

$$\boldsymbol{W}^{(s)} = \left[\boldsymbol{W}^{(s-1)} + \frac{\boldsymbol{w}_{k^{(s)}}^{(s-1)} (\boldsymbol{h}_{k^{(s)}})^H \boldsymbol{W}^{(s-1)}}{\boldsymbol{Q}^{(s)}(k^{(s)}, k^{(s)})}\right] \boldsymbol{P}^{(s)} \tag{2.16}$$

where $\boldsymbol{Q}^{(s)}(k^{(s)}, k^{(s)})$ is the $(k^{(s)}, k^{(s)})$ entry of $\boldsymbol{Q}^{(s)} = \boldsymbol{I}_{|\mathcal{I}_A|-s+1} - \left(\tilde{\boldsymbol{H}}^{(s-1)}\right)^H \boldsymbol{W}^{(s-1)}$;

The complexity of the MMSE-SICD is mainly depended on the computation of $\boldsymbol{R}_y^{-1}$, which is however only required to be computed once for $\boldsymbol{W}^{(0)}$ during the detection of the $|\mathcal{I}_A|$ active UEs, as shown in the above-stated algorithm. The detection delay is dependent on the number of active UEs, which is however dynamic, when the number of active UEs is dynamic.

2.4 Performance Results and Analysis

Below we provide the simulation results to show the error performance of the DyDS-CDMA and DyMC-CDMA systems with the MMSE-SICD assisted detection. In our simulations, we assume that the maximum power-control error is 0.1 times the transmit power obtained in ideal power-control. Since we normalize the transmit power under ideal power-control to unity, we hence have the normalized $P_E = 0.1$. Correspondingly, in all the figures, the average SNR represents the SNR averaged with respect to fading channels as well as power-control error.

2.4.1 Performance of DyDS-CDMA Systems

In Fig. 2.2, we depict the BER performance of the DyDS-CDMA systems employing the MMSE-SICD, when each cell supports $K = 200$, 1000 or 1500 potential UEs, and assuming that each UE has an activation probability of $P_a = 0.05$. Hence, the average number of active UEs is $K_A = 10$, 50 and 75. Furthermore, given the spreading factor of $N = 32$ considered, we can find from (2.2) that the probability that the system is overload is 0.9963 for $K = 1000$, or nearly 1 for 1500. The results in Fig. 2.2 show that, for $K = 200$ or 1000, the achieved BER performance is nearly the same as that of the single-user system, i.e., $K = 1$, communicating over Rayleigh fading channels. It is worth noting that, given

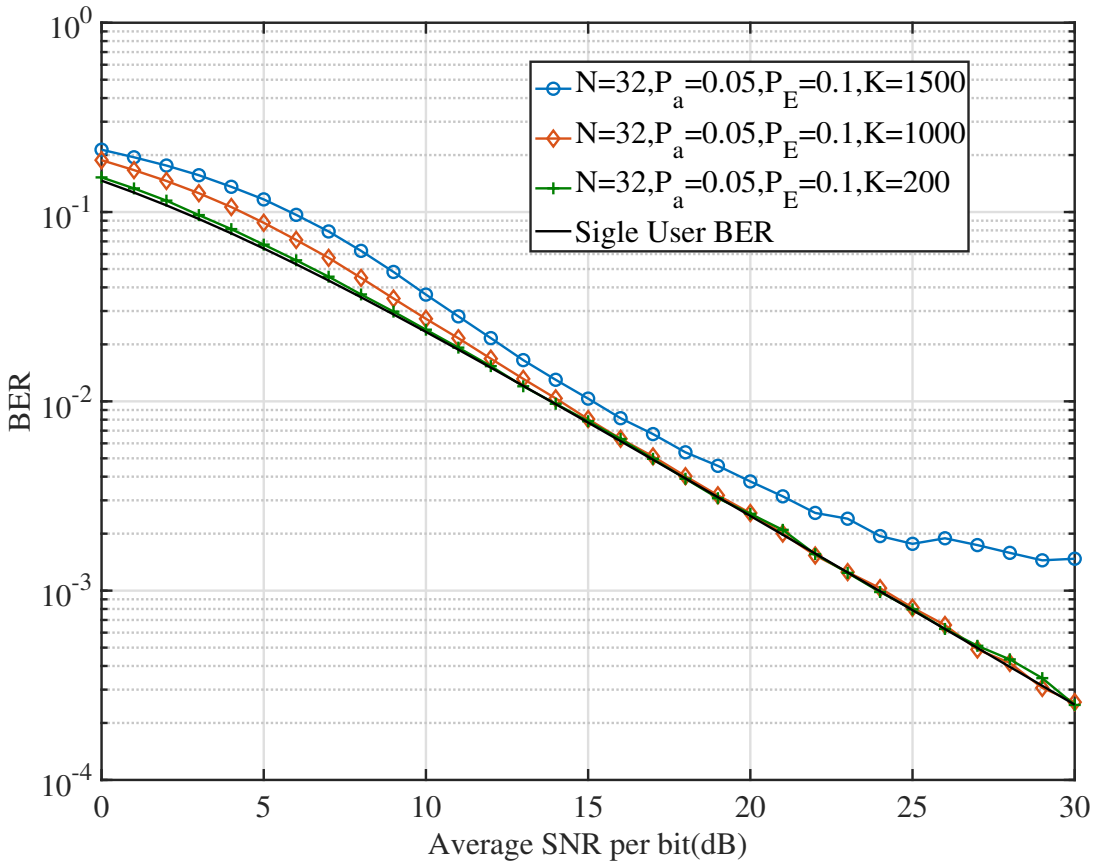


Figure 2.2: BER versus average SNR per bit performance of DyDS-CDMA $(N = 32)$ systems using MMSE-SIC detection, when communicating over Rayleigh fading channels.

$K = 1000$, we can calculate that the probability that there are $64(= 2N)$ or more active UEs is 0.0207. Therefore, the DyDS-CDMA system is capable of achieving near single-user BER performance, even when it is heavily overloaded, with a system loading factor reaching two or higher. By contrast, when there are $K = 1500$ potential UEs supported, the BER is still very close to the single-user BER performance, when the SNR is below 20 dB, or the (uncoded) BER is about 0.003, which is practically desirable in most applications. This is because, when error-control codes are employed, the decoded BER can be readily improved to 10^{-6} or lower, provided that the uncoded BER is between 0.001 to 0.01. We should note that when $K = 1500$, the probabilities of that the number of active UEs is $64(= 2N)$ or more, and $96(= 3N)$ or more are 0.8952 and 0.0069, respectively. These probabilities imply that the DyDS-CDMA system employing the MMSE-SICD has the capability to be efficiently operated in the communication environments of extremely overload.

Fig. 2.3 demonstrates the BER performance of the DyDS-CDMA systems supporting $K = 200$ potential UEs per cell, when UEs have different activation probabilities, as shown

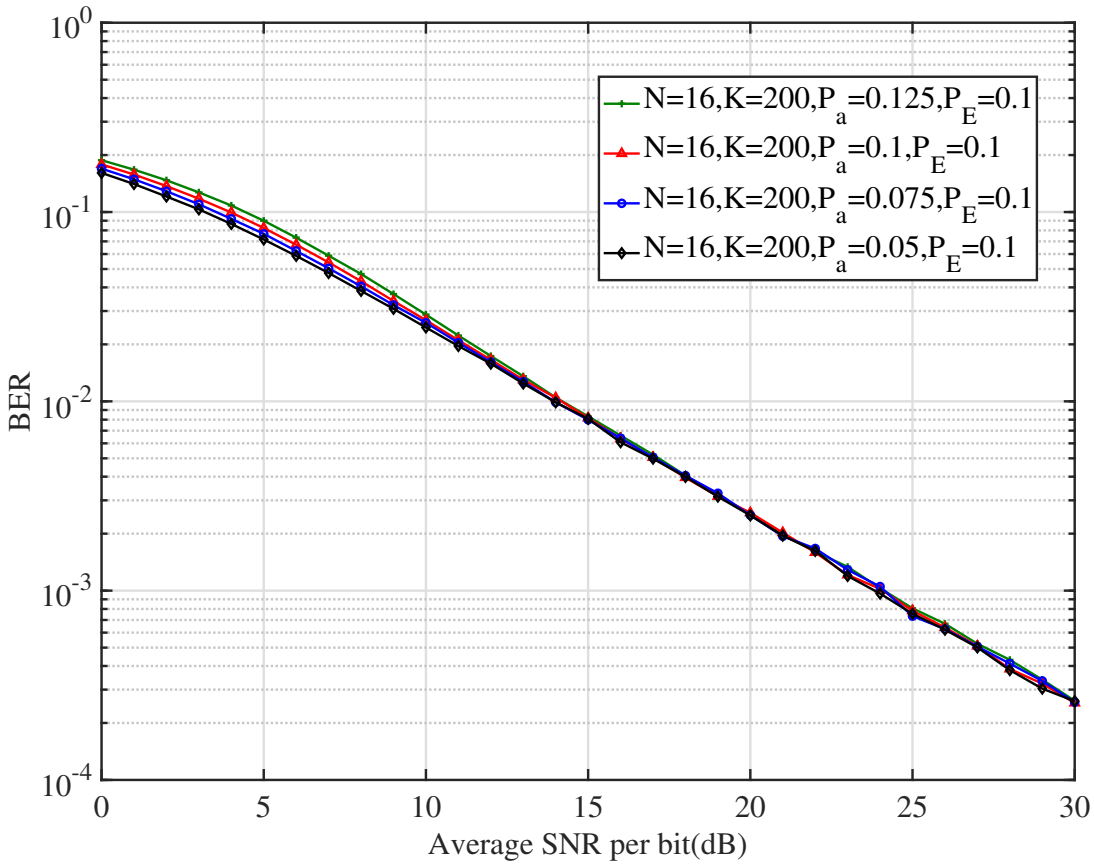


Figure 2.3: BER versus average SNR per bit performance of the DyDS-CDMA systems, when UEs have different probability of activation.

in the figure. Correspondingly, we can calculate that the average numbers of active UEs are 10, 15, 20 and 25. Explicitly, for all these cases, the BER performance that is close to the single-user BER performance can be achieved.

Fig. 2.4 compares the BER versus average SNR per bit performance of the DyDS-CDMA systems employing the MMSE multiuser detection (MMSE-MUD) and the MMSE-SICD proposed in this chapter. While the MMSE-SICD is capable of achieving near single-user BER performance, the conventional MMSE-MUD is unable to cope with the dynamics of the UEs, resulting in very poor BER performance over all the SNR range considered.

Finally, in Fig. 2.5, we illustrate the effect of UEs' activities on the BER performance of the DyDS-CDMA systems employing the MMSE-SICD or the conventional MMSE-MUD, when the number of potential UEs per cell is $K = 200$ and the average SNR is 15 or 20 dB. From the results we observe that the BER of the DyDS-CDMA systems increases steadily, as the activation probability of UEs increases from $P_a = 0.01$ to $P_a = 0.2$, which correspond to an average number of active UEs of $K_A = 2$ to $K_A = 40$. From literature [127] we know that, when the number of users is lower than N, the conventional

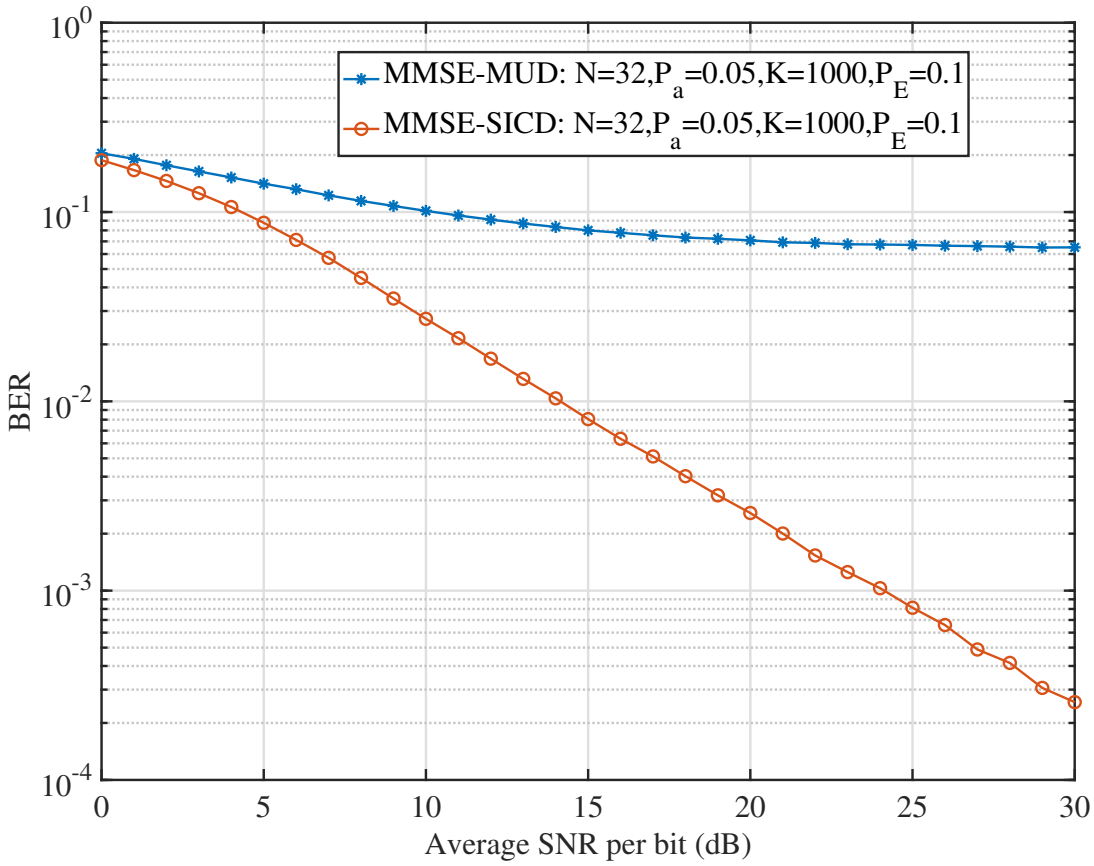


Figure 2.4: Comparison of BER versus average SNR per bit performance of the DyDS-CDMA systems employing the MMSE-MUD and MMSE-SICD.

MMSE-MUD operated in granted scenarios is usually not very sensitive to the increase of the number of users supported. The above observation means that the conventional MMSE-MUD is inefficient for operation in mGFMA systems, where the number of UEs is highly dynamic. By contrast, when the MMSE-SICD is employed, the BER performance of the DyDS-CDMA systems is very robust to the increase of the activation probability of UEs. As shown in the figure, the BER performance maintains near unchanged, until $P_a = 0.16$. After that point, the BER performance also slight degrades, as the UEs' activation probability increases. Over all the ranges of P_a, the MMSE-SICD significantly outperforms the MMSE-MUD.

2.4.2 Performance of DyMC-CDMA Systems

Below let us demonstrate the performance of DyMC-CDMA systems communicating over frequency-selective fading channels. Note that, for convenience of comparison with the DyDS-CDMA systems over flat-fading channels, the system parameters in Fig. 2.6 - Fig. 2.8 are set the same.

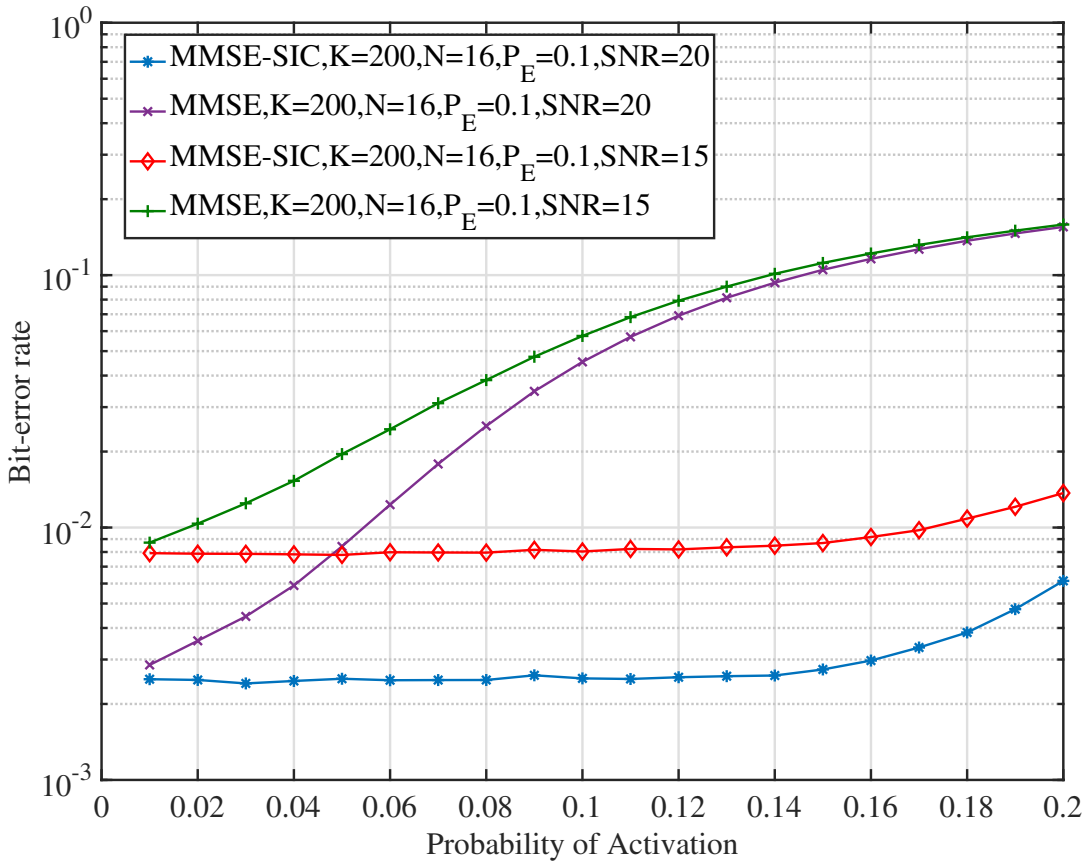


Figure 2.5: Effect of UEs' activities on the BER versus average SNR per bit performance of the DyDS-CDMA systems employing MMSE-SICD or the conventional MMSE-MUD.

First, corresponding to Fig. 2.2, in Fig. 2.6, we show the BER performance of the DyDS-CDMA systems employing the MMSE-SICD, when communicating over the frequency-selective fading channels with a time-domain resolvable paths of $L = 8$. The other parameters were the same as that for Fig. 2.2. In comparison with Fig. 2.2, first, owing to the diversity gain provided by the frequency-selective fading channels, the DyMC-CDMA systems with $K = 1$ and $K = 200$ are capable of achieving the near-Gaussian channel BER performance. Furthermore, the BER performance of the system with $K = 200$ is near the same as the single-user BER performance. Second, when $K = 1000$, the BER performance is promising and much better than the corresponding BER performance in Fig. 2.2, when SNR is lower than about 15 dB. However, when SNR is higher than 15 dB, the BER performance presents error-floor, due to the un-compressible multiuser interference. Additionally, when the number of potential UEs is $K = 1500$, the performance becomes worse than that in Fig. 2.2. This is because in DyMC-CDMA systems, the subcarriers experience frequency-selective fading, which modifies the binary spreading sequences in frequency-domain, making it difficult to suppress the multiuser interference by the MMSE operation.

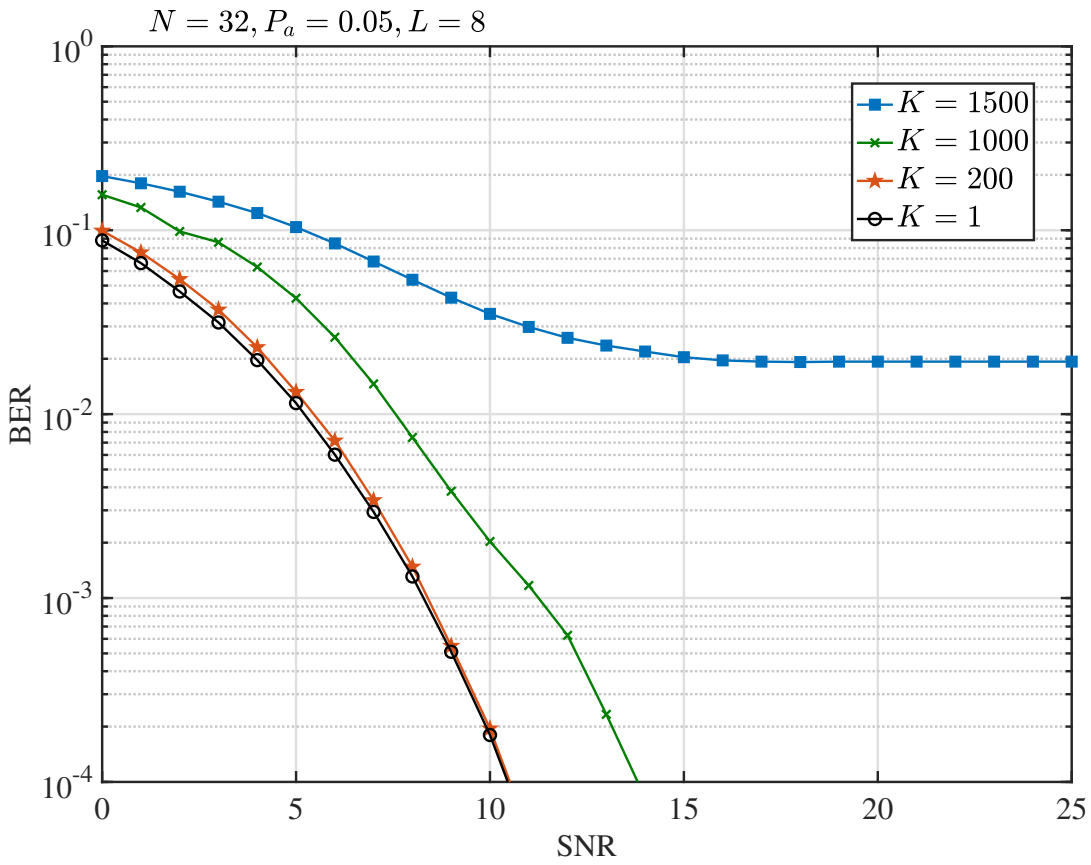


Figure 2.6: BER versus average SNR per bit performance of DyMC-CDMA $(N = 32)$ systems with MMSE-SICD, when dense spreading sequences are employed.

In correspondence to Fig. 2.4, in Fig. 2.7, the BER performance of the DyMC-CDMA systems over the frequency-selective fading with $L = 8$ is compared, when MMSE-MUD and MMSE-SICD are employed, respectively. Again, significant performance gain can be attained by the MMSE-SICD over the MMSE-MUD. Therefore, MMSE-SICD is also robust for operation in the DyMC-CDMA systems, where the number of potential UEs and the number of active UEs are time-variant. Furthermore, when communicating over frequency-selective fading channels, the MMSE-SICD is efficienct to achieve the frequency diversity gain.

Fig. 2.8 is in correspondence to Fig. 2.5 to investigate the effect of the activation probability on the achievable BER performance of the DyMC-CDMA systems, when communicating over the frequency-selective fading channels with $L = 8$. First, when comparing the BER performance in Fig. 2.5 with the corresponding one in Fig. 2.8, we can observe that for a given activation probability, the BER performance of the DyMC-CDMA is better (much better at relatively low activation probability) that that of DyDS-CDMA, owing to the fact that DyMC-CDMA can exploit the frequency diversity. However, when the activation probability is too high, making the multiuser interference dominate the achievable

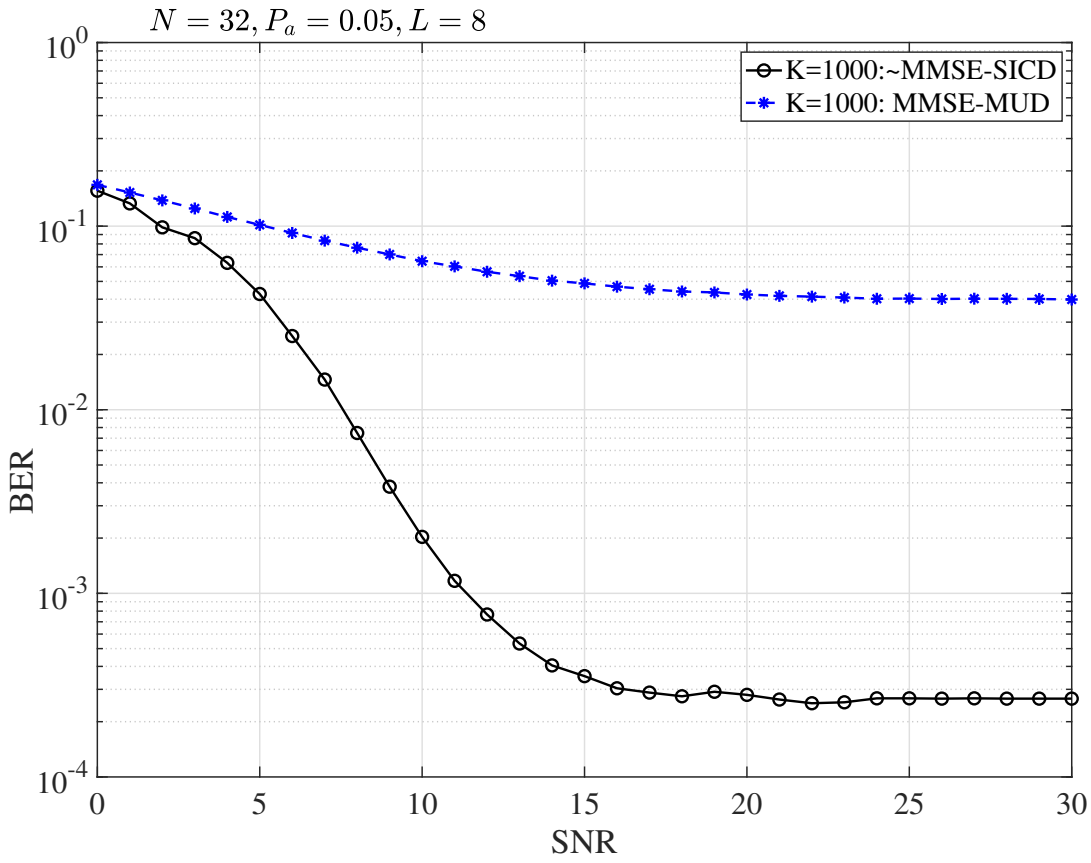


Figure 2.7: Comparison of BER versus average SNR per bit performance of the DyMC-CDMA systems employing the MMSE-MUD and MMSE-SICD, when dense spreading sequences are employed.

BER performance, both schemes attain a similar BER performance at a given activation probability. Second, while the BER performance of the DyDS-CDMA with MMSE-SID is robust to the varying activation probability, the BER performance of the DyMC-CDMA continuously increases, as the activation probability increases. Again, as shown in Fig. 2.8, the MMSE-SICD significantly outperforms the MMSE-MUD.

Fig. 2.9 shows the effect of the frequency-selectivity of fading channels on the BER performance of DyMC-CDMA systems, when the MMSE-SICD is employed. As shown in Fig. 2.9, we set the number of time-domain resolvable paths to be $L = 1, 2, 4, 8, 16$, reflecting that the channel is from flat fading ($L = 1$) to very frequency selective ($L = 16$). Again, the results demonstrate that the MMSE-SICD is effective for attaining the frequency diversity gain. When the channel becomes more frequency selective, the BER performance continuously improves, when L increases from $L = 1$ to $L = 16$. Specifically, when $L = 1$, the BER performance is the same as that of the single-user BPSK scheme communicating over flat Rayleigh fading channels, as shown in Fig. 2.2. In contrast, when $L = 16$, the BER performance is near the BER of the BPSK over Gaussian

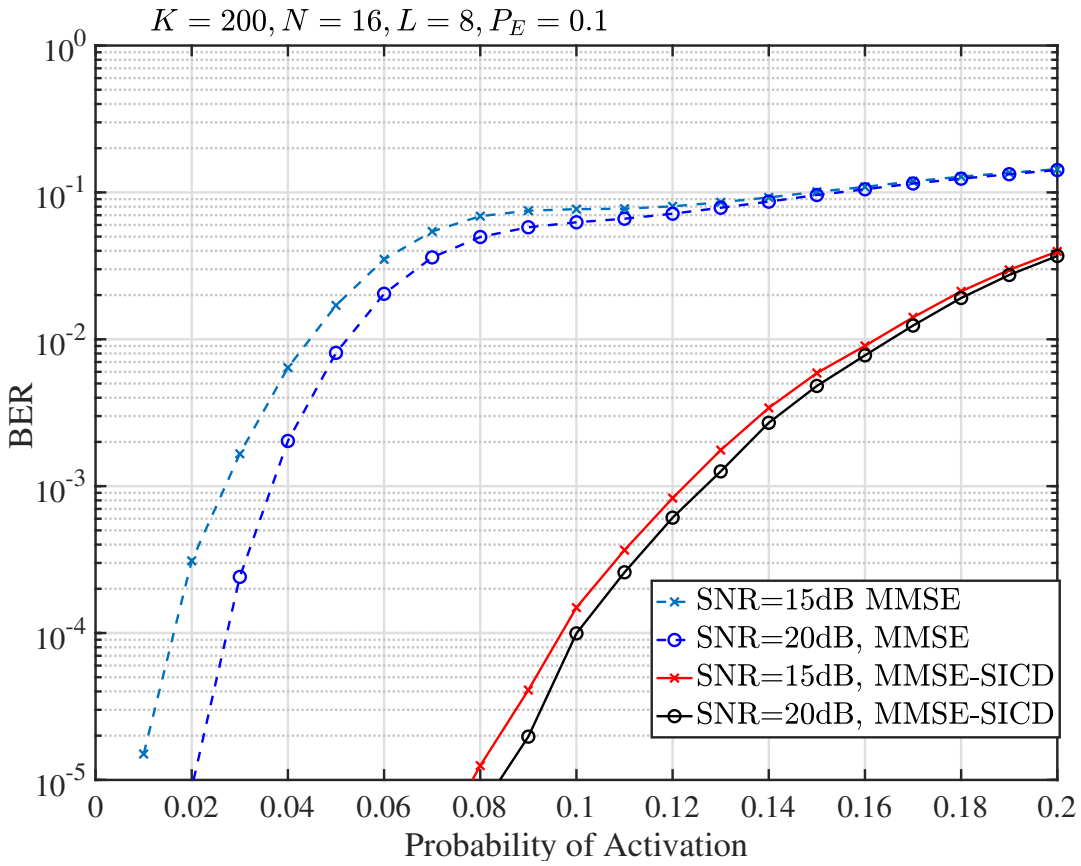


Figure 2.8: Effect of UEs' activities on the BER versus average SNR per bit performance of the DyMC-CDMA systems employing MMSE-SICD or the conventional MMSE-MUD, when dense spreading sequences are employed.

channels [143].

All the above results are obtained by assuming that the spreading sequences in DyDS-CDMA and DyMC-CDMA systems are binary random sequences with elements $+1$ and -1, which are hence dense sequences. In the following two figures, we specifically investigate the effect of the sparsity of spreading sequences on the achievable BER performance of the DyMC-CDMA systems over frequency-selective fading channels. For this purpose, we assume that the spreading sequences are also random sequences but with elements of $+1$, -1 and 0. When an element is 0, it means that the UE does not send information on the corresponding subcarrier. For the sparsity, it is reflected by the average weight, expressed as $\bar{d}_x$, of the spreading sequences. Note that, to avoid that a spreading sequence in simulation has all 0 elements, the minimum weight of a spreading sequence is set to 1.

As shown in Fig. 2.10, when the average sparsity $\bar{d}_x$ varies from $\bar{d}_x = 3$ to $\bar{d}_x = 18$, the achieved BER performance improves at a given SNR. In more detail, when the average sparsity is as low as $\bar{d}_x = 3$, we can observe that significant diversity gain is lost in comparison with the available diversity gain of $L = 8$, as seen in Fig. 2.9. The

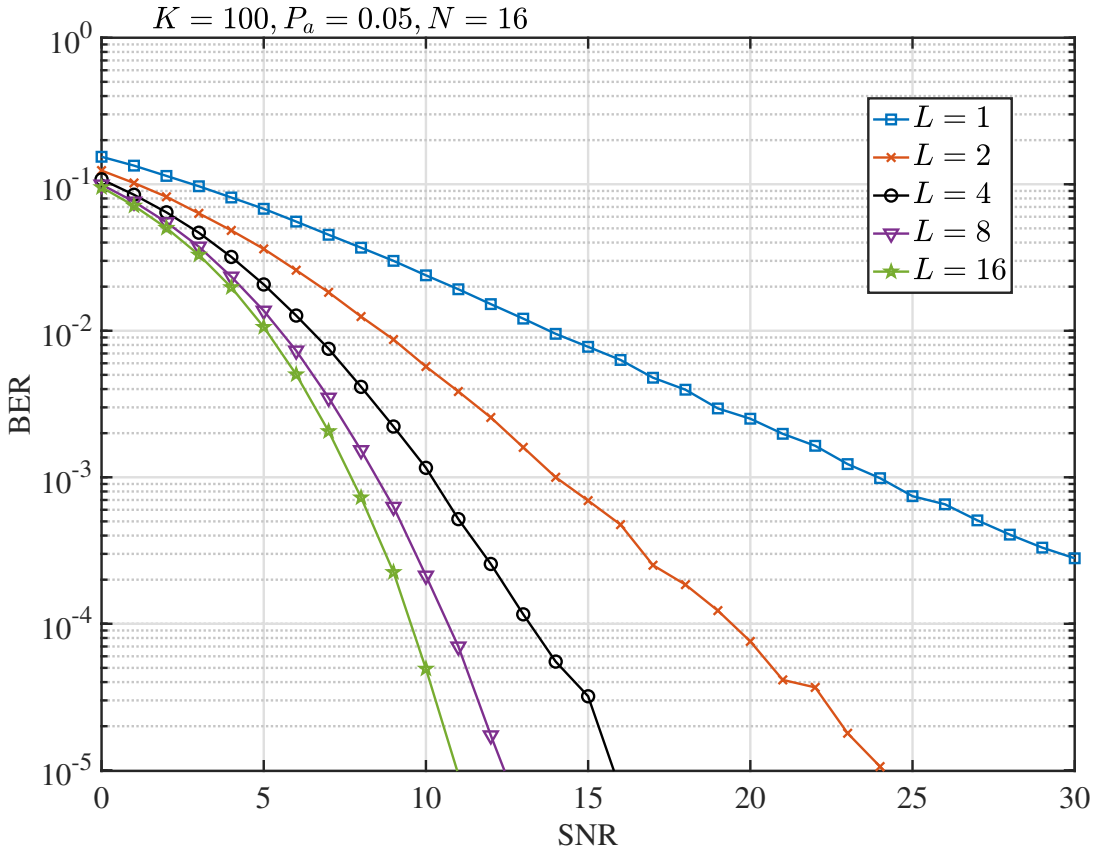


Figure 2.9: BER versus average SNR per bit performance of DyMC-CDMA systems with MMSE-SICD, when dense spreading sequences are employed.

reason behind is that when $\bar{d}_x = 3$, the average number of subcarriers activated by a UE is also 3 in average. Hence, the diversity order achievable is only about 3, instead of $L = 8$. Furthermore, at $\bar{d}_x = 3$, the probability that two or more UEs share the same subcarriers increases, which results in that the multiuser interference is unable to be effectively mitigated by the MMSE-SICD. Consequently, as seen in Fig. 2.10, the curve of $\bar{d}_x = 3$ appears error-floor in relatively high SNR region. In contrast, when the average sparsity is relatively high, such as, $\bar{d}_x > 9$, each UE sends its information on more than $L = 8$ subcarriers. Correspondingly, nearly all available diversity gain can be achieved, which can be seen by comparing the results in Fig. 2.10 with the curve of $L = 8$ in Fig. 2.9. From the comparison, we observe that the BER performance of the DyMC-CDMA systems with dense spreading can be achieved by applying the sparse spreading sequences with the average sparsity about half of the spreading sequences' length. Specifically for the example considered, $\bar{d}_x \geq 15$ allows to achieve nearly the same BER performance of the DyMC-CDMA systems with dense spreading.

Finally, by setting $\bar{d}_x = 15$ and 18, respectively, in Fig. 2.11, we investigate the BER performance of the DyMC-CDMA systems, when the number of potential UEs are $K =$

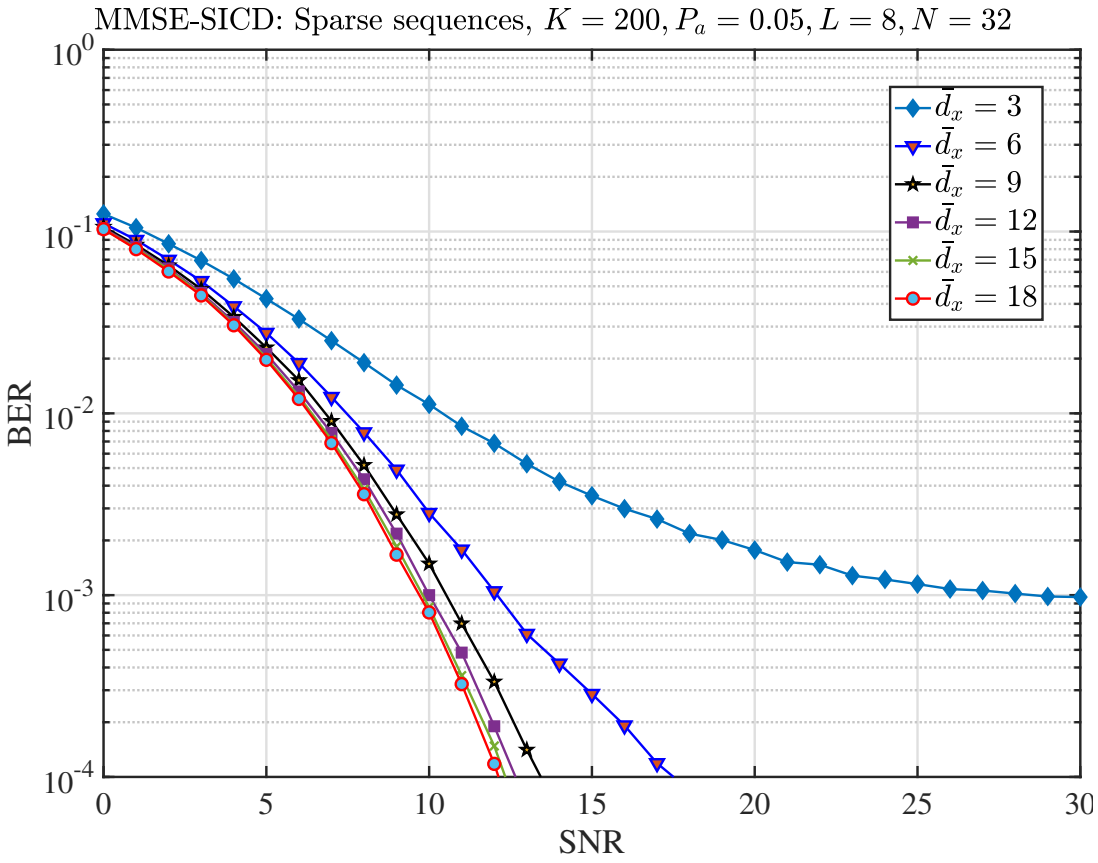


Figure 2.10: BER versus average SNR per bit performance of DyMC-CDMA systems with MMSE-SICD, when sparse spreading sequences are employed.

200, 1000 and 2000, respectively, in order to demonstrate whether the MMSE-SICD is efficient for operation in the DyMC-CDMA systems with sparse spreading. By comparing the results in Fig. 2.11 and that in Fig. 2.6, we can readily realize that the MMSE-SICD is similarly effective for operation in the DyMC-CDMA systems with sparse spreading. As seen in the two figures, the BER performance achieved in both the dense and sparse spreading cases is similar for $K = 200$ or $K = 1000$.

2.5 Chapter Conclusions

In this chapter, we have presented a DyDS-CDMA and a DyMC-CDMA schemes for supporting ultra-densely deployed wireless IoT devices, which result in that each BS may need to support a big number of potential UEs, while each of the devices activates to transmit during a time-slot with a small probability. Hence, for a given time-slot, the active UEs and the number of active UEs simultaneously transmit information is dynamic. In this chapter, we assume the ideal identification of active UEs and also assume the ideal channel estimation of the active UEs, with the to demonstrate the potentially achievable performance

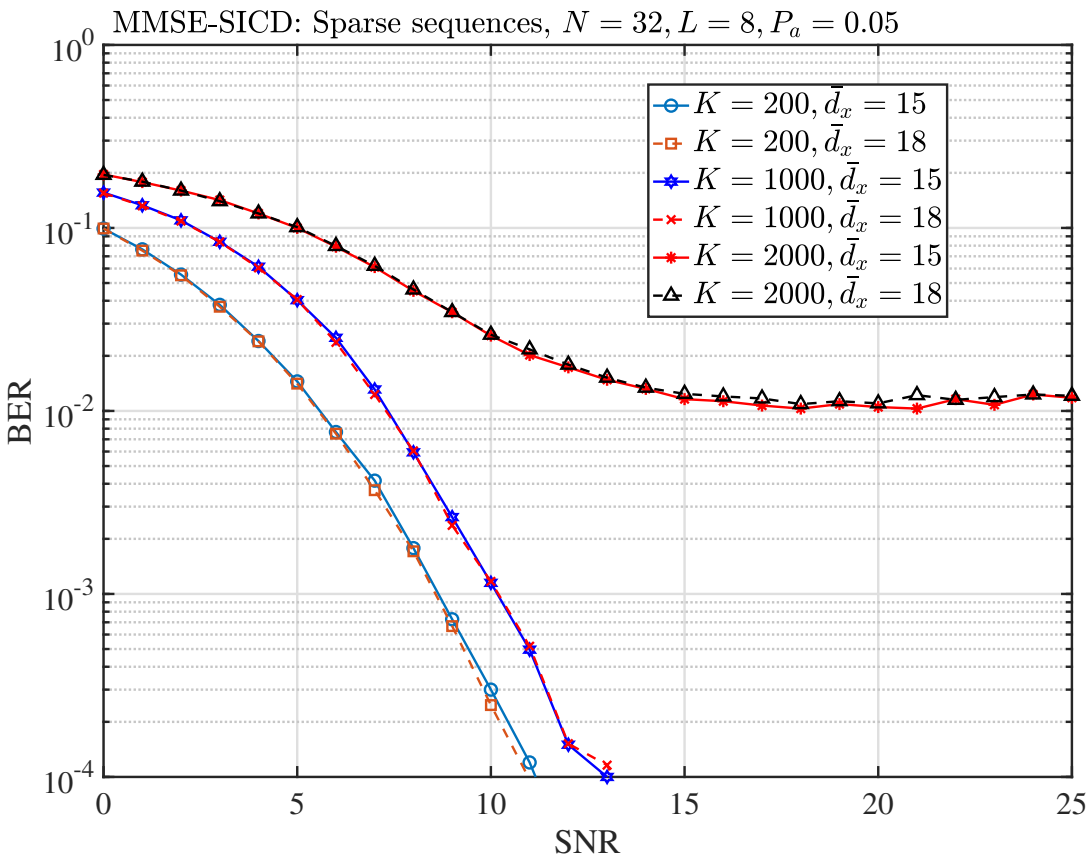


Figure 2.11: BER versus average SNR per bit performance of DyMC-CDMA systems with MMSE-SICD, when sparse spreading sequences are employed.

of the low-complexity MMSE-SICD operated in the DyDS-CDMA and DyMC-CDMA systems. Therefore, based on the Monte-Carlo simulations, we investigated the BER performance of the DyDS-CDMA and DyMC-CDMA systems employing the MMSE-SICD, when different settings are considered. Furthermore, the MMSE-SICD is compared with the MMSE-MUD from different perspectives. Our studies demonstrate that the MMSE-SICD is a high-efficiency and robust detection scheme for operation in the DyDS-CDMA and DyMC-CDMA systems operated in mGFMA environments. It is capable of achieving near single-user BER performance, even when the average number of active UEs per time-slot reaches two times of N of the systems' degrees-of-freedom. With the aid of the MMSE-SICD, the DyMC-CDMA system employing dense spreading is feasible for achieving the frequency diversity, when communicating over frequency-selective fading channels. In order to reduce the PAPR problem, we have also studied the BER performance of the DyMC-CDMA systems employing the MMSE-SICD. It is shown that the near-full gain of frequency diversity may only be achievable, when the sparsity of spreading sequences is sufficiently high. Specifically, a typical sparsity for achieving the near-full gain of frequency diversity is the one that is slightly larger than the number of time-domain resolvable paths of the frequency-selective fading channels.

Owing to the above-mentioned merits of DyMC-CDMA and MMSE-SICD, in the following chapters, the DyMC-CDMA will be suggested for the implementation of mGFMA, and the MMSE-SICD will be employed for the signal detection in mGFMA systems.

Chapter 3

Channel Estimation and User Activity Identification in Massive Grant-Free Multiple-Access

GFMA allows to significantly reduce the overhead in comparison with granted multiple-access. However, information detection in GFMA is challenging, as it has to be executed along with the activity detection of UEs and channel estimation. In this chapter, we study the channel estimation and propose the active UE identification (AUI) algorithms for the mGFMA systems. For the above mentioned purposes, the channel estimation is studied from several aspects by assuming different levels of knowledge to the access point, and based on which five AUI approaches are proposed. We study the performance of channel estimation, the statistics of estimated channels, and the performance of AUI algorithms. Our studies show that the proposed approaches are capable of circumventing some of the shortcomings of the existing techniques designed based on compressive sensing and message passing algorithms. They are robust for use in the mGFMA systems where the active UEs and the number of them are highly dynamic.

3.1 Introduction

So far, in wireless communications systems, a wireless terminal connecting its AP or base-station (BS) is typically achieved via GMA, where a handshaking process between wireless terminal and AP is operated to fulfill the functions of multiple-access. GMA has shown its advantages in the conventional wireless communication systems, from the 1G to 4G, where traffic is mainly activated by human being. However, starting from the 5G, wireless systems are rendered to be device (machine) centric, with the traffic initialized by numer-

ous devices (machines). As the result, the characteristics of the traffic in 5G and beyond wireless systems will be very different from that in the 1G to 4G wireless systems. In these future wireless systems, there may be various types of sporadic short-burst traffic generated, for example, by sensing nodes and IoT devices of many kinds, nodes for sending control commands, etc. [12]. Explicitly, transmission of these kinds of sporadic short-burst traffic using the traditional communications schemes, such as that in 4G LTE/LTE-A systems, is often inefficient, due to their reliance on a huge amount of overhead for signaling and control, yielding an extremely high overhead-to-data ratio as well as introducing extra communication latency.

In order to improve the efficiency of the transmission of sporadic traffic, save the overhead for corresponding network maintenance and to support massive connectivity, in recent years, various GFMA schemes have been proposed and investigated [6, 62, 144]. In GFMA, a UE, which may be a device, machine, etc., becomes active randomly to transmit data to its AP simply under the control of an assigned random access protocol, such as slotted ALOHA, without requesting for a grant from its AP [74, 144]. GFMA has the advantage of greatly saving the resources that are otherwise consumed for the extra overhead. However, it also imposes new challenges for AP to recover the data sent by active UEs. This becomes even more significant when GFMA is employed to support massive connectivity, forming the mGFMA. In mGFMA, each UE usually only has a small probability to become active at a time(-slot). Hence, active UEs as well as the number of them are highly dynamic, both of which are usually not *a-priori* knowledge to AP. Therefore, in order to recover the data sent by active UEs, an AP has to accomplish first the functions of active UE identification and the channel estimation for active UEs, or implement them in joint ways [6, 61, 118, 129, 137, 145–147].

The investigation of UAI in random-access scenarios was already started along with the conventional CDMA systems [148–151]. Recently, it has drawn a lot of research attention with the mGFMA in IoT networks and in particular, in mMTC, where UAI is often jointly implemented with channel estimation and even data detection [61, 129, 145, 146]. As above-mentioned, in mGFMA systems, the number of potential UEs supported by an AP is usually very big, up to thousands, tens of thousands or even more, while each UE only has a very small probability that is normally less than 10% to become active. Consequently, when without the knowledge of active UEs, signal detection in mGFMA is naturally a CS relied signal recovery problem. Owing to this, in recent years, various CS-based joint UAI, channel estimation and/or MUD algorithms have been developed and investigated in the context of mGFMA, when various sparsity structures are considered [61, 116, 118, 137, 140, 146, 152–161]. While CS-based methods have some outstanding merits as claimed in references, they are not appearing for operation in the mGFMA

systems where the number of active UEs is large [162]. This is because the recovery performance of CS-based methods is limited by the restricted isometric property (RIP) condition [163]. This condition implies that the number of active UEs (sparsity) should be significantly lower than the number of resource units (subcarriers, antennas, spreading factor, time-slots, etc.) of an AP, as seen, e.g., in [145, 146, 153, 155, 156, 158], and hence resulting in a low efficiency of resource usage. Furthermore, in mGFMA systems, the number of active UEs is highly dynamic, making the systems often *real-time* overload[1] and substantially violating the RIP condition. For example, for an AP using $N = 32$ resource units to support $K = 1000$ potential UEs of each having an independent activation probability of 5%, the overloading probability is $P(K_A > N) = 0.9963$. Explicitly, the CS-based methods are not suitable for operation in this mGFMA system, and other efficient approaches are required for the UAI, channel estimation and data detection.

In mGFMA, the number of potential UEs is usually significantly higher than the number of orthogonal resource units. Hence, mGFMA is generally implemented in the principles of non-orthogonal multiple-access (NOMA). In the NOMA family, sparse-code multiple-access (SCMA) is an attractive scheme, which allows to support the multiple-access of significant overloading but still achieve near optimum performance by exploiting the message-passing algorithm (MPA)-assisted detection [164]. Hence, SCMA has been introduced to GFMA, e.g., in [18,62,147,165,166]. Specifically in [94], the joint UE activity detection and data detection were considered, when assuming that all UEs (both active and inactive) transmit the headers for activity detection, while the channels of all UEs are known to BS. Considering an uplink GFMA, [147] investigated the MPA-assisted receiver design for joint channel estimation and signal detection. In [165], a faster-than-Nyquist (FTN) singling SCMA was proposed to support uplink GFMA, where only very small systems (6 UEs supported by 4 resource units) were investigated. In [166], the authors considered the MPA-assisted information detection in a rate-less SCMA-based GFMA systems, when assuming a constant number of active UEs, and that receiver has ideal knowledge about the active users and their channels. While one UE employs only one codebook in the above-mentioned references, the authors in [18] proposed to map a UE's consecutive data symbols to the codewords chosen from different codebooks, which achieves code diversity. In general, SCMA achieves the near optimum performance depending on the MPA-assisted detection operated on *good* factor graphs. However, when we consider the mGFMA system in which an AP supports a huge number of potential UEs, while it has no

[1]Here, *real-time* overload means that the number of active UEs is more than the number of resource units. By contrast, the overload claimed in some references considering CS-assisted GFMA means that the number of potential UEs is higher than the number of resource units. In the rest of this chapter, we mean overload the real-time overload, unless otherwise notified.

knowledge of active UEs, the number of them and of their channels, maintaining good factor graphs in such highly dynamic communication scenarios is extremely challenging. In addition to the SCMA-based GFMA, a semi-grant-free power-domain NOMA based transmission scheme was proposed in [167], where one UE is operated as a primary UE under a conventional grant-based protocol, while the other secondary UEs opportunistically access the same channel in grant-free manner.

In addition to the above-mentioned, GFMA and mGFMA have also been investigated by invoking other techniques. In [116], the authors enhanced the CS-based approach by proposing a greedy algorithm relied on the maximum a-posteriori probability (MAP) criterion, which performs UE identification and data detection jointly by exploiting the a-posteriori probabilities of each other. An expectation propagation based joint UE activity detection and channel estimation scheme was introduced in [168] for achieving GFMA in mMTC networks. In [169], the authors investigated the preamble design for UE detection and channel estimation, so as to attain a high success rate in grant-free random access in massive multiple-input multiple-output (MIMO) systems. The authors of [170] compared two types of preamble structures on the performance of grant-free random access massive MIMO. Similarly, GFMA in massive MIMO systems was considered in [171], where a fixed number of UEs access one of the available channels and orthogonal preambles are transmitted for UE activity identification and channel estimation. Furthermore, the GFMA massive MIMO scenario was considered in [172], where the number of active UEs are first estimated, followed by the joint detection of active UE and their messages with the aid of a so-called ensemble independent component analysis decoding algorithm, without relying on explicit channel estimation. Reed-muller sequences were introduced in [173] to aid UE identification and channel estimation in the GFMA systems supporting massive connectivity. In [174], UE activity and signal detection were studied with mGFMA by designing preambles and exploiting the interleave-division multiple access (IDMA), when user signals are randomly and asynchronously transmitted. More recently, machine-learning (ML) approaches have been introduced to mGFMA [155, 161, 175–177]. To be more specific, the authors of [155] investigated the joint UE activity detection and channel estimation by formulating them as a block sparse signal recovery problem, which is solved by a block sparse Bayesian learning (BSBL) algorithm. In [175], the asynchronous sparse Bayesian learning algorithm was adopted for channel estimation while the support vector machine method was applied for UE activity detection, when the number of active users is assumed to be known to BS. To alleviate the convergence problem of the BSBL algorithm [155], a deep neural network-aided message passing-based block sparse Bayesian learning algorithm was proposed in [176] to achieve the joint UE activity detection and channel estimation in mGFMA scenarios. The joint optimization of finite-alphabet spreading sequences

and multi-user detection were addressed in [177], where deep learning principle was introduced for the design of both encoder and decoder. Furthermore, in [161], deep learning was used to predict the activities of UEs, followed by a CS algorithm for data detection and further enhancing the UE activity identification. While nearly all the above-mentioned works considered the grant-free access techniques at physical layer, the authors of [178] investigated the GFMA from the MAC layer perspective in the IoT systems with sporadic traffic.

In this chapter, we investigate the channel estimation and UAI in the mGFMA scenario, where the number of potential UEs may be much bigger than the number of resource units and each UE has a small activation probability. Hence, the active UEs and the number of them are highly dynamic over time, and the number of active UEs at a time may be higher than the number of resource units. Specifically, we first consider the channel estimation of both active and inactive UEs in the principle of minimum mean-square error (MMSE), when AP is assumed to have the ideal knowledge, no knowledge or the partial knowledge of active UEs. In our channel estimation, AP will make the best use of the information provided by the transmission of pilot symbols and data (payload) symbols. Our studies show that any added knowledge about the active UEs can be exploited to enhance the reliability of channel estimation. This is reflected in practical mGFMA systems that the activity patterns of UEs may be correlated or one active UE may have several packets to transmit in continuous time-slots. Then, this kind of information can be exploited to enhance the reliability of channel estimation in mGFMA systems.

Following the channel estimation, we then study the statistical properties of the estimated channels of both active and inactive UEs. We explicitly show that the estimated channels of active UEs usually have relatively good mean-square error (MSE) performance, which improves for example, with the increase of signal-to-noise ratio (SNR), of the number of pilot symbols, etc. By contrast, as there are no transmissions from inactive UEs, the estimated channels for them always have a high MSE resulted from noise and the interference from active UEs, which is also insensitive to the investment of resources, such as, transmit power, pilot symbols, etc. We also study the distribution of estimated channel's power, showing that the estimated channel's power of active UEs is usually much higher than that of inactive UEs, when SNR is in the range of practical applications.

Finally, by exploiting the channel estimation and the characteristics of estimated channels, we propose a range of UAI algorithms, in which some of them make use of the estimated channel's knowledge, while the others carry out joint channel estimation and UAI. In detail, the first one is the simplest threshold-based UAI (TB-UAI) algorithm, which however makes a sharp trade-off between false-alarm and miss identification, and is also

Table 3.1: Comparison of this work with the related works in literature.

Contributions	System model	Channel model	Random No. of UEs	Dynamic UEs	Priori knowledge	Real overloading	CE	UAI method	Joint UAI&CE
This work	SP-NOMA	Correlated FS	✓	✓		✓	✓	CE	✓
[179]	SP-NOMA	Uncorrelated FS		✓	UEs' channels			CS	
[180]	SP-NOMA	Uncorrelated FS	✓	✓*	UEs' channels			CS	
[137]	mMIMO	Flat	✓	✓			✓	CS	✓
[181]	SP-NOMA	Correlated FS		✓	UEs' channels			CS	
[182]	SP-NOMA	Flat		✓*			✓	CS	UAI&CE&DD
[183]	SP-NOMA	Correlated FS	✓	✓	UEs' channels	✓		MAP	
[184]	mMIMO	flat	✓	✓			✓	CS	✓
[185]	mMIMO	flat	✓	✓	No. of UEs			CS/learning	UAI&CE&DD
[186]	SP-NOMA	Uncorrelated FS		✓	UEs' channels			CS	
[187]	SP-NOMA	Uncorrelated FS	✓	✓			✓	CS/learning	✓
[94]	SCMA	Gaussian**	✓***	✓***	UEs' channels	✓		MP	
[188]	SCMA	Correlated FS	✓	✓	No. of UEs	✓	✓	ML/MP	

Notes: ✓: yes; Empty: no; SP: spreading; FS: frequency-selective fading; CE: channel estimation; mMIMO: massive or multiuser MIMO; DD: data detection; MAP: maximum a posteriori probability; MP: message-passing; AMP: approximate MP; *: temporally correlated user activity; ML: maximum likelihood; **: Gaussian channel for performance; ***: All UEs transmit headers for UAI;

hard to set an appropriate threshold. The second one is referred to as the eigen-analysis enhanced UAI (EAE-UAI) algorithm, which exploits the eigen-analysis to improve the performance of UAI. In order to further enhance the performance of UAI, we then propose two successive interference cancellation (SIC) assisted UAI algorithms, namely, the SIC-UAIa and SIC-UAIb, which improve the performance of channel estimation and UAI with the aid of the SIC strategies as proposed. Finally, as above-mentioned, any added knowledge about active UEs can be exploited to enhance the performance of channel estimation. Correspondingly, we propose an auto-correlation matrix evolving UAI (AME-UAI) algorithm, which makes use of the earlier identified UEs to improve the channel estimation and UAI of the UEs processed later. In the above proposed five algorithms, the last three carry out joint channel estimation and UAI. Furthermore, in comparison with the TB-UAI, the other four algorithms make no or little trade-off between false-alarm and miss, in addition to the performance enhancement.

The novelty of our work is compared with the related works in Table 3.1 and the main contributions can be summarized as follows:

- Channel estimation in dynamic multicarrier mGFMA (MC-mGFMA) scenarios is investigated, where the channels of both active and inactive UEs are estimated in MMSE principle, when AP is assumed to employ ideal knowledge, no knowledge or partial knowledge about the active UEs. It is shown that the proposed channel estimation approaches are robust to the dynamics existing in mGFMA systems. Furthermore, we show that the added knowledge about active UEs results in the enhanced reliability of channel estimation.

- The statistical properties of the estimated channels of both active and inactive UEs are analyzed, showing that the estimated channels of active UEs and that of inactive

UEs have distinct features, which can be exploited for distinguishing active UEs from inactive UEs.

- Hence, based on the studies of channel estimation and of the characteristics of estimated channels, five UAI algorithms are proposed for MC-mGFMA and their performance is studied. These UAI algorithms are listed as follows:

  1. Threshold-based UAI (TB-UAI);
  2. Eigen-analysis enhanced UAI (EAE-UAI);
  3. Successive interference cancellation (SIC) assisted UAIa (SIC-UAIa);
  4. Successive interference cancellation (SIC) assisted UAIb (SIC-UAIb);
  5. Auto-correlation matrix evolving UAI (AME-UAI).

  These UAI algorithms have their distinctive characteristics and present different performance-complexity trade-off. They however have one common feature that all of them are suitable for operation in the mGFMA systems having highly dynamic UEs in terms of the active UEs and the number of them.

The remainder of this chapter is organized as follows. In Section 3.2, we describe the mGFMA system model. Section 3.3 focuses on channel estimation, while Section 3.4 demonstrates the performance of channel estimation. Furthermore, in Section 3.5, we study the statistical properties of the channels estimated by the approaches in Section 3.3. The UAI algorithms and their complexity are addressed in Section 3.6, and the performance of UAI is demonstrated in Section 3.7. Finally, the conclusions from our research are presented in Section 3.8.

3.2 Description of mGFMA System

We consider a single-cell multicarrier mGFMA (MC-mGFMA) system in the principles of the frequency-domain spread orthogonal frequency-division multiple-access (OFDMA) [127]. The MC-mGFMA system has N subcarriers and one access point (AP) (or base-station (BS)) located at the center of the cell. Since most mGFMA systems are motivated for low-rate IoT services, we assume only the binary phase shift keying (BPSK) modulation, although it can be readily extended to other quadrature amplitude modulations (QAM) modulation. In this MC-mGFMA system, the signature code assigned to the kth UE equipment (UE) is expressed as $\boldsymbol{c}_k = [c_k(0), c_k(1), \ldots, c_k(N-1)]^T$, where $c_k(n) \in \{+1/\sqrt{N}, -1/\sqrt{N}\}$, making $\|\boldsymbol{c}_k\|^2 = 1$. The signature $\boldsymbol{c}_k$ of UE k serves for both

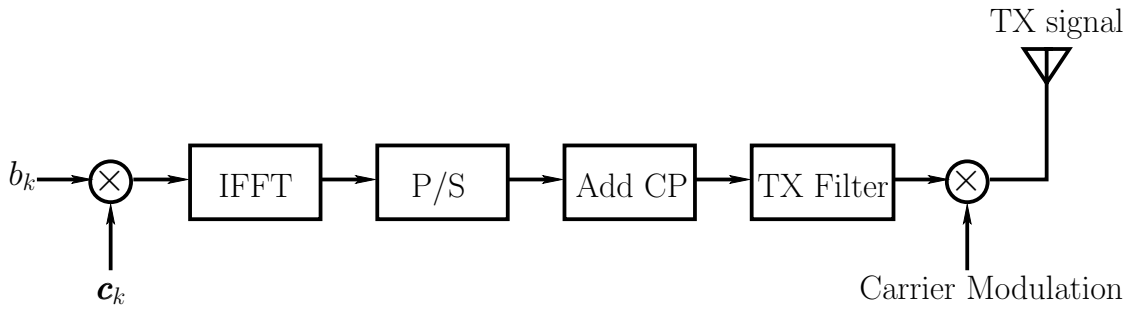


Figure 3.1: Transmitter schematic diagram for GFMA systems.

the spreading of transmitted signal, and the identity (ID) of UE k for AP to identify its activity during the UE identification stage. If UE k is identified to be active, $\boldsymbol{c}_k$ is also used by AP to demodulate the signals received from UE k during the data detection stage. In our considered MC-mGFMA system, IFFT/FFT techniques are introduced for subcarrier modulation/demodulation, and a sufficiently long cyclic prefix (CP) is added to avoid the interference between adjacent OFDM blocks. We assume that the MC-mGFMA system supports K potential UEs, which are mobile IoT devices randomly distributed in the cell. In massive connectivity application scenarios, we have $K >> N$. We assume that the system implements a synchronization sub-system, which enables all UEs to synchronize with the AP's clock. A UE has a small probability of P_a to become active, and an active UE starts transmission at the beginning of the next time-slot. One frame is transmitted per time-slot by each active UE. We express K_A the number of active UEs during a time-slot, which is a random variable having its average satisfying $\bar{K}_A << K$, as the result of a small P_a.

3.2.1 Transmitter Modeling

The transmitter schematic diagram for the considered MC-mGFMA system is shown in Fig. 3.1, which follows the typical framework of OFDM transmitter [127]. To transmit a binary bit $b_k \in \{+1, -1\}$, it is first spread by $\boldsymbol{c}_k$, the products of which are transformed to the time-domain using IFFT. Then, after the parallel-to-serial (P/S) conversion, adding CP, transmitter (TX) filtering, and carrier modulation, the resultant signal is transmitted from the transmit antenna of UE k.

In our study, we assume block-based fading, i.e., the fading of a UE maintains constant over a frame duration, but is independent frame-by-frame. Let the frame length be expressed as $N_F = N_I + N_P + N_L$, where N_I denotes the number of information bits per frame, and N_P is the number of pilot bits per frame, inserted for channel estimation and UE identification. Furthermore, at the beginning of each frame, N_L bits are transmitted

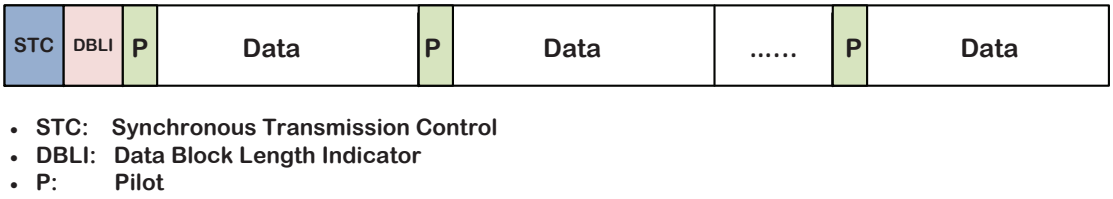


Figure 3.2: Frame structure for operation of proposed MC-mGFMA.

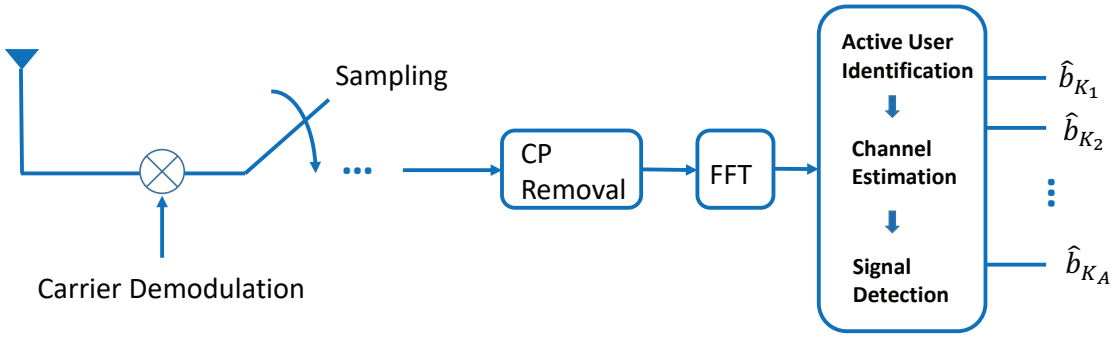


Figure 3.3: Receiver schematic diagram of MC-mGFMA system.

to represent the data block length indicator (DBLI), which tells AP the number of frames left to transmit. In our studies, we typically set $N_L = 2$. Hence, once a UE becomes active, the maximum number of frames to transmit in one session is 4, which is informed to AP by setting the DBLI bits in the first frame to '11'. This value reduces following the sequence '10', '01', until '00' in the 4th frame. Then, no matter whether this UE finishes transmission, it stops transmission. If it has more data to transmit, it activates again and starts a new transmission session. The benefit of employing the DBLI is that the AP does not need to identify every active UE, but those newly activating ones, which enhances the performance of active UE identification, channel estimation and hence of signal detection. For illustration, the frame arrangement in our MC-mGFMA is depicted in Fig. 3.2. As seen in Fig. 3.2, there is also a 'STC' domain, which is used to make the transmissions by different UEs be approximately synchronized at AP, as done in LTE/LTE-A systems [189]. To implement this, each UE measures the transmission delay between it and its AP based on the pilot signals sent from the AP. Then, the UE can adjust its starting transmission point in the 'STC' region, so that the transmissions from different UEs are approximately synchronized when they arrive at the AP.

3.2.2 Representation of Received Signal

The receiver block diagram for the MC-mGFMA system is shown in Fig. 3.3, which follows the receiving principles of OFDM until after the FFT operation. Hence, after the FFT

operation, the received signals from N subcarriers can be represented as

$$\boldsymbol{y}_m = \sum_{k=1}^{K} \boldsymbol{C}_k \boldsymbol{\mathcal{F}} \boldsymbol{\Psi}_k \boldsymbol{h}_k b_k^{(m)} I_k + \boldsymbol{n}_m, \tag{3.1}$$

$$= \sum_{k'=1}^{K_A} \boldsymbol{C}_{k'} \boldsymbol{\mathcal{F}} \boldsymbol{\Psi}_{k'} \boldsymbol{h}_{k'} b_{k'}^{(m)} + \boldsymbol{n}_m,\ m = 1, 2, \ldots, N_F \tag{3.2}$$

where $b_k^{(m)}$ is the mth bit in a frame sent from UE k, $\boldsymbol{y}_m$ and $\boldsymbol{n}_m$ are N-length observation and noise vectors, respectively, and $\boldsymbol{n}_m$ follows the complex Gaussian distribution with zero mean and covariance matrix $\sigma^2 \boldsymbol{I}_N$, expressed as $\mathcal{CN}(0, \sigma^2 \boldsymbol{I}_N)$, where $\sigma^2 = 1/SNR$ with SNR being the signal-to-noise ratio (SNR). In (3.1), $\boldsymbol{C}_k = \text{diag}\{c_k(0), c_k(1), \cdots, c_k(N-1)\}$, $\boldsymbol{h}_k = [h_{k,0}, \cdots, h_{k,L-1}]^T$ is the time-domain channel impulse response (CIR) of UE k, where $h_{k,l}$ obeys the distribution $\mathcal{CN}(0, 1/L)$. Correspondingly, the frequency-domain CIRs, i.e., the fading gains of N subcarriers are given by $\boldsymbol{\mathcal{F}} \boldsymbol{\Psi}_k \boldsymbol{h}_k$ [127], where $\boldsymbol{\mathcal{F}}$ is the $(N \times N)$ FFT matrix, and $\boldsymbol{\Psi}_k$ is the mapping matrix constituted by the first L columns of the identity matrix $\boldsymbol{I}_N$. Furthermore, in (3.1), I_k is the activity indicator, $I_k = 1$ or $I_k = 0$ indicates that UE k is active or inactive. For simplicity, let in (3.1) $\boldsymbol{A}_k = \boldsymbol{C}_k \boldsymbol{\mathcal{F}} \boldsymbol{\Psi}_k$, which is a $(N \times L)$ matrix. Then, we have

$$\boldsymbol{y}_m = \sum_{k=1}^{K} I_k \boldsymbol{A}_k \boldsymbol{h}_k b_k^{(m)} + \boldsymbol{n}_m, \tag{3.3}$$

$$= \sum_{k'=1}^{K_A} \boldsymbol{A}_{k'} \boldsymbol{h}_{k'} b_{k'}^{(m)} + \boldsymbol{n}_m, m = 1, 2, \ldots, N_F \tag{3.4}$$

Specifically for the pilots $p = 1, 2, \cdots, N_P$, the received signals can be expressed as

$$\boldsymbol{y}_p = \sum_{k=1}^{K} b_k^{(p)} I_k \boldsymbol{A}_k \boldsymbol{h}_k + \boldsymbol{n}_p$$

$$= \sum_{k'=1}^{K_A} b_{k'}^{(p)} \boldsymbol{A}_{k'} \boldsymbol{h}_{k'} + \boldsymbol{n}_p,\ p = 1, 2, \cdots, N_P \tag{3.5}$$

Let us collect the observations from the N_P pilot symbols to a vector $\boldsymbol{y} = [\boldsymbol{y}_1^T, \boldsymbol{y}_2^T, \cdots, \boldsymbol{y}_{N_P}^T]^T$. Then, it can be expressed as

$$\boldsymbol{y} = \sum_{k=1}^{K} I_k (\boldsymbol{p}_k \otimes \boldsymbol{A}_k) \boldsymbol{h}_k + \boldsymbol{n} \tag{3.6}$$

$$= \sum_{k'=1}^{K_A} (\boldsymbol{p}_{k'} \otimes \boldsymbol{A}_{k'}) \boldsymbol{h}_{k'} + \boldsymbol{n} \tag{3.7}$$

where $\boldsymbol{p}_k = [b_k^{(1)}, b_k^{(2)}, \cdots, b_k^{(N_P)}]^T$ contains the N_P pilot bits of UE k, $\otimes$ is the Kronecker product, while $\boldsymbol{n}$ is a NN_P-length vector following the distribution of $\mathcal{CN}(0, \sigma^2 \boldsymbol{I}_{NN_P})$.

3.3 Channel Estimation

In this section, we consider the channel estimation in the considered MC-mGFMA system, all based on the minimum mean-square error (MMSE) [190]. First, we derive the channel estimator by assuming that AP knows ideally the active UEs. Then, the channel estimator is obtained by assuming that AP has no knowledge at all about the active UEs. Explicitly, both the above cases are extreme cases. In practice, the situation may be that AP knows some active UEs, e.g., from the DBLI bits sent by UEs, as we considered in Section 3.2.1, but does not know the newly activated UEs. Therefore, we also consider the channel estimation in this more practical scenario. Note that we choose MMSE instead of zero-forcing (ZF) (or least square) estimator, as ZF estimator requires to invert a large sized matrix that is dependent on the number of potential UEs, and it is well-known that MMSE estimator achieves better estimation performance than ZF estimator.

3.3.1 Channel Estimation with Active UEs Known to AP - Estimator-K

When AP knows all the active UEs, the channels of these UEs can be estimated based on (3.7). Hence, when MMSE-relied estimation is employed, the estimated channel of the ith active UE can be expressed as

$$\hat{\boldsymbol{h}}_i = \boldsymbol{W}_i^H \boldsymbol{y},\ i = 1, 2, \cdots, K_A \tag{3.8}$$

where according to the principles of MMSE, the weight matrix $\boldsymbol{W}_i$ can be expressed as [190]

$$\boldsymbol{W}_i = \boldsymbol{R}_y^{-1} \boldsymbol{R}_{y,i} \tag{3.9}$$

In (3.9), $\boldsymbol{R}_y$ is the autocorrelation matrix of $\boldsymbol{y}$, which can be derived from (3.7) as

$$\begin{aligned}
\boldsymbol{R}_y &= E\left[\boldsymbol{y}\boldsymbol{y}^H\right] \\
&= E\left[\left(\sum_{k'=1}^{K_A} (\boldsymbol{p}_{k'} \otimes \boldsymbol{A}_{k'})\boldsymbol{h}_{k'} + \boldsymbol{n}\right)\left(\sum_{l'=1}^{K_A} (\boldsymbol{p}_{l'} \otimes \boldsymbol{A}_{l'})\boldsymbol{h}_{l'} + \boldsymbol{n}\right)^H\right] \\
&= \sum_{k'=1}^{K_A} (\boldsymbol{p}_{k'} \otimes \boldsymbol{A}_{k'}) E\left[\boldsymbol{h}_{k'}\boldsymbol{h}_{k'}^H\right] (\boldsymbol{p}_{k'} \otimes \boldsymbol{A}_{k'})^H + \sigma^2 \boldsymbol{I}_{N_P N}
\end{aligned} \tag{3.10}$$

With the aid of the relationships of $(\boldsymbol{A} \otimes \boldsymbol{B})^H = \boldsymbol{A}^H \otimes \boldsymbol{B}^H$ and $(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = \boldsymbol{AC} \otimes \boldsymbol{BD}$ [191], we can derive that if long-term average cross many frames is imagined,

we have $E\left[\boldsymbol{h}_{k'}\boldsymbol{h}_{k'}^H\right] = \boldsymbol{I}_L/L$, and (3.10) can be simplified to

$$\boldsymbol{R}_y = \sum_{k'=1}^{K_A}\left(\boldsymbol{p}_{k'}\boldsymbol{p}_{k'}^H \otimes \frac{\boldsymbol{A}_{k'}\boldsymbol{A}_{k'}^H}{L}\right) + \sigma^2\boldsymbol{I}_{N_PN} \tag{3.11}$$

By contrast, if only one frame is considered, $\boldsymbol{h}_{k'}$ is a constant vector. In this case, $\boldsymbol{R}_y$ has the form of

$$\boldsymbol{R}_y = \sum_{k'=1}^{K_A}\left(\boldsymbol{p}_{k'}\boldsymbol{p}_{k'}^H \otimes \boldsymbol{A}_{k'}\boldsymbol{h}_{k'}\boldsymbol{h}_{k'}^H\boldsymbol{A}_{k'}^H\right) + \sigma^2\boldsymbol{I}_{N_PN} \tag{3.12}$$

However, in MC-mGFMA systems, this $\boldsymbol{R}_y$ is inaccessible to the AP, as it is unable to construct it by exploiting the available knowledge about the UEs, while it is also incapable of estimating it using the received signals of one frame. By contrast, the AP is able to construct the $\boldsymbol{R}_y$ in (3.11), as it knows the pilot symbols and the spreading codes of the active UEs.

In (3.9), $\boldsymbol{R}_{y,i}$ is the cross-correlation matrix between $\boldsymbol{y}$ and $\boldsymbol{h}_i$, which can be shown to be

$$\boldsymbol{R}_{y,i} = \frac{1}{L}(\boldsymbol{p}_i \otimes \boldsymbol{A}_i) \tag{3.13}$$

Consequently, upon substituting (3.11) and (3.13) into (3.9), the weight matrix for estimating the ith UE's channel is given by

$$\begin{aligned}\boldsymbol{W}_i &= \left[\sum_{k'=1}^{K_A}\left(\boldsymbol{p}_{k'}\boldsymbol{p}_{k'}^H \otimes \boldsymbol{A}_{k'}\boldsymbol{A}_{k'}^H\right) + \sigma^2\boldsymbol{I}_{N_PN}\right]^{-1}(\boldsymbol{p}_i \otimes \boldsymbol{A}_i)/L \\ &= \left[\sum_{k'=1}^{K_A}\left(\boldsymbol{p}_{k'}\boldsymbol{p}_{k'}^H \otimes \boldsymbol{A}_{k'}\boldsymbol{A}_{k'}^H\right) + L\sigma^2\boldsymbol{I}_{N_PN}\right]^{-1}(\boldsymbol{p}_i \otimes \boldsymbol{A}_i)\end{aligned} \tag{3.14}$$

Similarly, if (3.12) is considered, we have

$$\begin{aligned}\boldsymbol{W}_i = &\left[\sum_{k'=1}^{K_A}\left(\boldsymbol{p}_{k'}\boldsymbol{p}_{k'}^H \otimes \boldsymbol{A}_{k'}\boldsymbol{h}_{k'}\boldsymbol{h}_{k'}^H\boldsymbol{A}_{k'}^H\right) + \sigma^2\boldsymbol{I}_{N_PN}\right]^{-1} \\ &\times (\boldsymbol{p}_i \otimes \boldsymbol{A}_i)/L\end{aligned} \tag{3.15}$$

Note that although (3.15) is not implementable, it is insightful for deriving the other estimators, as shown in our forthcoming discourses. Furthermore, from (3.11) we can deduce that the $(N \times N)$ diagonal sub-matrices of $\boldsymbol{R}_y$ are given by

$$\begin{aligned}\boldsymbol{R}_y(n,n) =& \boldsymbol{R}_a \\ =& \sum_{k'=1}^{K_A}\frac{\boldsymbol{A}_{k'}\boldsymbol{A}_{k'}^H}{L} + \sigma^2\boldsymbol{I}_N, n = 1,2,\ldots,N_P\end{aligned} \tag{3.16}$$

By contrast, from (3.12) we can know that these diagonal sub-matrices are

$$\begin{aligned} \boldsymbol{R}_a &= \sum_{k'=1}^{K_A} \boldsymbol{A}_{k'} \boldsymbol{h}_{k'} \boldsymbol{h}_{k'}^H \boldsymbol{A}_{k'}^H + \sigma^2 \boldsymbol{I}_N, \\ & n = 1, 2, \ldots, N_P \end{aligned} \tag{3.17}$$

These results will be useful later.

The MSE of the channel estimation for an active UE can be derived as

$$\begin{aligned} M_{SE}(A) =& E\left[\|\boldsymbol{h}_i - \boldsymbol{W}_i^H \boldsymbol{y}\|^2\right] \\ =& \mathrm{Tr}\left(\boldsymbol{I}_L / L - \boldsymbol{R}_{y_{h_i}}^H \boldsymbol{W}_i\right) \\ =& 1 - \mathrm{Tr}\left(\boldsymbol{R}_{y,i}^H \boldsymbol{W}_i\right) \end{aligned} \tag{3.18}$$

If the AP mistakenly estimates the channel of a UE that is actually not active, the AP can still form $\boldsymbol{W}_i$ in the form of (3.14). Correspondingly, the MSE of the channel estimation for an inactive UE is

$$\begin{aligned} M_{SE}(\bar{A}) =& E\left[\| - \boldsymbol{W}_i^H \boldsymbol{y}\|^2\right] \\ =& \mathrm{Tr}\left(\boldsymbol{W}_i^H \boldsymbol{R}_y \boldsymbol{W}_i\right) \end{aligned} \tag{3.19}$$

Later in Section 3.6, we will investigate the statistics of estimated channels conditioned on that a UE is active or inactive. This statistical information will then be exploited for active UE identification.

Finally, we note that the MMSE estimator is a biased estimator [190]. In order to obtain an unbiased estimation of $\boldsymbol{h}_i$, after the MMSE estimation, we can let

$$\hat{\boldsymbol{h}}_i = \boldsymbol{Q}_i \boldsymbol{W}_i^H \boldsymbol{y},\ i = 1, 2, \cdots, K_A \tag{3.20}$$

where $\boldsymbol{Q}_i$ is applied to achieve the unbiased estimation, making the estimated channel satisfy $E[\hat{\boldsymbol{h}}_i] = \boldsymbol{h}_i$. Therefore, when substituting (3.7) into the above equation and completing the expectation, we can obtain

$$\boldsymbol{Q}_i = \left(\boldsymbol{W}_i^H (\boldsymbol{p}_i \otimes \boldsymbol{A}_i)\right)^{-1} \tag{3.21}$$

3.3.2 Channel Estimation with Active UEs Unknown to AP - Estimator-uK

In contrast to the extreme case that AP knows ideally the active UEs, another extreme case in MC-mGFMA is that AP only knows there are K potential UEs, but does not know the

active UEs and even the number of active UEs K_A. In this case, first, AP has to treat the data sent from different UEs as independent identically distributed (iid) random variables, and use $E[\boldsymbol{p}_{k'}\boldsymbol{p}_{k'}^H]$ to replace $\boldsymbol{p}_{k'}\boldsymbol{p}_{k'}^H$ in (3.12), where

$$E[\boldsymbol{p}_{k'}\boldsymbol{p}_{k'}^H] = \boldsymbol{I}_{N_P} \tag{3.22}$$

When applying this result to (3.12), we obtain

$$\begin{aligned}\boldsymbol{R}_y &= \sum_{k'=1}^{K_A}\left(\boldsymbol{I}_{N_P}\otimes\boldsymbol{A}_{k'}\boldsymbol{h}_{k'}\boldsymbol{h}_{k'}^H\boldsymbol{A}_{k'}^H\right)+\sigma^2\boldsymbol{I}_{N_PN}\\ &= \left(\boldsymbol{I}_{N_P}\otimes\left[\sum_{k'=1}^{K_A}\boldsymbol{A}_{k'}\boldsymbol{h}_{k'}\boldsymbol{h}_{k'}^H\boldsymbol{A}_{k'}^H+\sigma^2\boldsymbol{I}_N\right]\right)\\ &= \left(\boldsymbol{I}_{N_P}\otimes\boldsymbol{R}_a\right)\end{aligned} \tag{3.23}$$

where $\boldsymbol{R}_a$ from (3.17) is substituted. Note that in (3.23), K_A is still unknown to AP, which will be addressed later in this subsection.

When AP attempts to estimate the channel of UE i, it can use the related information (spreading code and its pilot symbols) to construct the cross-correlation matrix $\boldsymbol{R}_{y,i}$ of (3.13). Consequently, AP can form the weight matrix $\boldsymbol{W}_i$ as

$$\begin{aligned}\boldsymbol{W}_i &= \boldsymbol{R}_y^{-1}\boldsymbol{R}_{y_{h_i}}\\ &= \left(\boldsymbol{I}_{N_P}\otimes\boldsymbol{R}_a\right)^{-1}\left(\boldsymbol{p}_i\otimes\boldsymbol{A}_i\right)/L\\ &= \left(\boldsymbol{I}_{N_P}\otimes\boldsymbol{R}_a^{-1}\right)\left(\boldsymbol{p}_i\otimes\boldsymbol{A}_i\right)/L\\ &= \left(\boldsymbol{p}_i\otimes\boldsymbol{R}_a^{-1}\boldsymbol{A}_i/L\right)\end{aligned} \tag{3.24}$$

Let us express $\boldsymbol{W}_i$ as

$$\boldsymbol{W}_i = \left[\boldsymbol{W}_{i,1}^T, \boldsymbol{W}_{i,2}^T, \cdots, \boldsymbol{W}_{i,N_p}^T\right]^T \tag{3.25}$$

Then, it can be readily shown that

$$\boldsymbol{W}_{i,p} = b_i^{(p)}\boldsymbol{R}_a^{-1}\boldsymbol{A}_i/L,\ p = 1,2,\ldots,N_P \tag{3.26}$$

Furthermore, when substituting (3.25) into (3.8), we obtain

$$\hat{\boldsymbol{h}}_i' = \boldsymbol{W}_i^H\boldsymbol{y} = \sum_{p=1}^{N_P}\boldsymbol{W}_{i,p}^H\boldsymbol{y}_p = \sum_{p=1}^{N_P}\boldsymbol{h}_i^{(p)} \tag{3.27}$$

where

$$\hat{\boldsymbol{h}}_i^{(p)} = \boldsymbol{W}_{i,p}^H\boldsymbol{y}_p,\ p = 1,2,\ldots,N_P \tag{3.28}$$

which can be shown is the MMSE-based estimation of $\boldsymbol{h}_i$ using only the signals related to the pth pilot symbol.

Therefore, due to the application of (3.22), the estimation of (3.27) does not directly provide an estimate for $\boldsymbol{h}_i$ in MMSE sense. Instead, each component of $\hat{\boldsymbol{h}}_i^{(p)}$, as seen in (3.28), is the MMSE estimation of $\boldsymbol{h}_i$ based on the observation $\boldsymbol{y}_p$ corresponding to the pth pilot symbol. In other words, (3.27) is the sum of the N_P estimates to $\boldsymbol{h}_i$ based on the N_P pilot symbols. Hence, in order to obtain an unbiased estimate to $\boldsymbol{h}_i$, we can form the estimation as

$$\hat{\boldsymbol{h}}_i = \frac{1}{N_P}\sum_{p=1}^{N_P} \boldsymbol{Q}_i^{(p)} \boldsymbol{W}_{i,p}^H \boldsymbol{y}_p = \frac{1}{N_P}\sum_{p=1}^{N_P} \boldsymbol{Q}_i^{(p)} \hat{\boldsymbol{h}}_i^{(p)} \tag{3.29}$$

where, explicitly, the factor $1/N_P$ takes account of the average over N_P pilots, while $\boldsymbol{Q}_i^{(p)}$ is for obtaining the unbiased estimation of $\boldsymbol{h}_i$ based on the pth pilot symbol, which can be found to be

$$\begin{aligned} \boldsymbol{Q}_i^{(p)} &= \left(\boldsymbol{W}_{i,p}^H b_i^{(p)} \boldsymbol{A}_i\right)^{-1} \\ &= \left(\boldsymbol{A}_i^H \boldsymbol{R}_a^{-1} \boldsymbol{A}_i / L\right)^{-1}, \; i = 1, 2, \ldots, N_P \end{aligned} \tag{3.30}$$

From (3.27) and (3.29) we can conceive that when the AP does not know the active UEs, it has to first carry out the symbol-based channel estimation, and then average the estimates obtained from the N_P pilot symbols to give the final estimation. Furthermore, while AP knows $b_i^{(p)}$ from the pilot sequences of UEs and it can also construct $\boldsymbol{A}_i$ from the UEs' spreading sequences, AP has to know $\boldsymbol{R}_a$, in order to compute $\boldsymbol{W}_{i,p}$ of (3.26) to fulfill the channel estimation based on (3.29). However, AP does not have the knowledge about the active UEs and even the number of them, $\boldsymbol{R}_a$ is unable to be constructed using the knowledge available to AP, but has to be obtained from alternative approaches. In this chapter, we propose to estimate $\boldsymbol{R}_a$ from the received signals as

$$\hat{\boldsymbol{R}}_a \approx \frac{1}{N_F}\sum_{n=1}^{N_F} \boldsymbol{y}_n \boldsymbol{y}_n^H \tag{3.31}$$

The above equation implies that $\boldsymbol{R}_a$ can be estimated by making use of all the signals received within a frame-duration, which is beneficial to lessening the effect of transmitted data. When applying (3.4) into the above equation, we obtain an approximation of

$$\hat{\boldsymbol{R}}_a \approx \sum_{k'=1}^{K_A} \boldsymbol{A}_{k'} \boldsymbol{h}_{k'} \boldsymbol{h}_{k'}^H \boldsymbol{A}_{k'}^H + \sigma^2 \boldsymbol{I}_N \tag{3.32}$$

the righthand side of which is exactly the $\boldsymbol{R}_a$ of (3.17).

Hence and in summary, for AP to estimate the ith UE's channel when it has no knowledge about the active UEs, the AP first uses all the received signals over one frame to

$$\boldsymbol{R}_y = \begin{pmatrix} \boldsymbol{R}_a & \sum\limits_{k\in\bar{\mathcal{K}}_A} b_k^{(1)} b_k^{(2)} \dfrac{\boldsymbol{A}_k \boldsymbol{A}_k^H}{L} & \cdots & \sum\limits_{k\in\bar{\mathcal{K}}_A} b_k^{(1)} b_k^{(N_P)} \dfrac{\boldsymbol{A}_k \boldsymbol{A}_k^H}{L} \\ \sum\limits_{k\in\bar{\mathcal{K}}_A} b_k^{(2)} b_k^{(1)} \dfrac{\boldsymbol{A}_k \boldsymbol{A}_k^H}{L} & \boldsymbol{R}_a & \cdots & \sum\limits_{k\in\bar{\mathcal{K}}_A} b_k^{(2)} b_k^{(N_P)} \dfrac{\boldsymbol{A}_k \boldsymbol{A}_k^H}{L} \\ \vdots & \vdots & \ddots & \vdots \\ \sum\limits_{k\in\bar{\mathcal{K}}_A} b_k^{(N_P)} b_k^{(1)} \dfrac{\boldsymbol{A}_k \boldsymbol{A}_k^H}{L} & \sum\limits_{k\in\bar{\mathcal{K}}_A} b_k^{(N_P)} b_k^{(2)} \dfrac{\boldsymbol{A}_k \boldsymbol{A}_k^H}{L} & \cdots & \boldsymbol{R}_a \end{pmatrix} \tag{3.33}$$

estimate $\boldsymbol{R}_a$ based on (3.31). Then, it constructs $\boldsymbol{A}_i$ using the spreading sequence assigned to UE i. Furthermore, with the above and the pilots of UE i, AP forms the weight matrices based on (3.26) and $\boldsymbol{Q}_i^{(p)}$'s based on (3.30). Finally, the channel of UE i is estimated based on (3.29).

3.3.3 Channel Estimation with Active UEs Partially Known to AP - Estimator-pK

In the previous two subsections, we have considered two extreme cases. Specially in Subsection 3.3.1, it is assumed that AP has ideal knowledge about active UEs, which should yield the upper-bound estimation performance. By contract, in 3.3.2, AP is assumed to have no knowledge about active UEs, which results in the worst estimation performance. In practical mGFMA systems, the operational condition may be more like that AP knows in part the active UEs, as the result that some UEs have more than one frame to send, as indicated to AP by the DBLI bits. However, AP does not know the newly activated UEs, whose activities need to be identified.

Therefore, if AP knows some UEs as *a-priori*, this information can be exploited to improve the performance of channel estimation in MC-mGFMA systems. Specifically, we can construct $\boldsymbol{R}_y$ in the form of (3.11) or (3.12) as follows.

First, AP estimates for $\boldsymbol{R}_a$ using (3.31). Then, AP constructs $\boldsymbol{A}_k$ for all the active UEs known to AP. With the aid of this information and the pilot symbols, AP is capable of constructing a $\boldsymbol{R}_y$ as shown on the top of the next page.

In (3.33), $\bar{\mathcal{K}}_A$ is the set of active UEs known to AP.

The cross-correlation matrices for all UEs are available to AP, which are the same as (3.13). Consequently, the AP can construct the weight matrices $\boldsymbol{W}_i$ as (3.9), and finally,

estimate the channels of UEs in the same form as (3.20).

3.3.4 Discussion

The above-considered three channel estimation schemes are all in the principle of MMSE. Hence, all of them have relatively low-complexity for practical implementation. However, the Estimator-K and Estimator-pK need to invert a relatively large size matrix (autocorrelation matrix), they have higher complexity than the Estimator-uK. Nevertheless, during one time-slot, the autocorrelation matrix is only required to be inverted once, and then can be used for the channel estimation of all UEs[2]. Therefore, its contribution to the overall complexity should be insignificant.

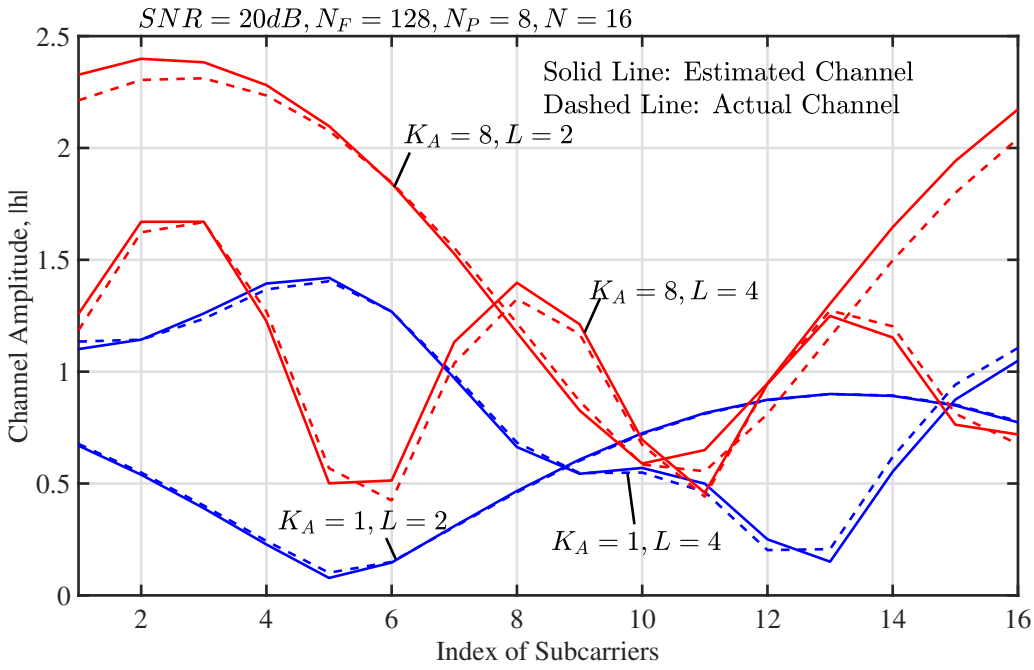
In terms of the delay introduced by channel estimation, if K UEs' channels are estimated successively, the delay is proportional to K. However, from the principles of these estimators, we can readily know that all the three estimators can be implemented in parallel, i.e., all the K UEs' channels can be simultaneously estimated, once the inverse of the autocorrelation matrix is computed. In this way, the delay of channel estimation is independent of the number of UEs involved.

3.4 Performance of Channel Estimation

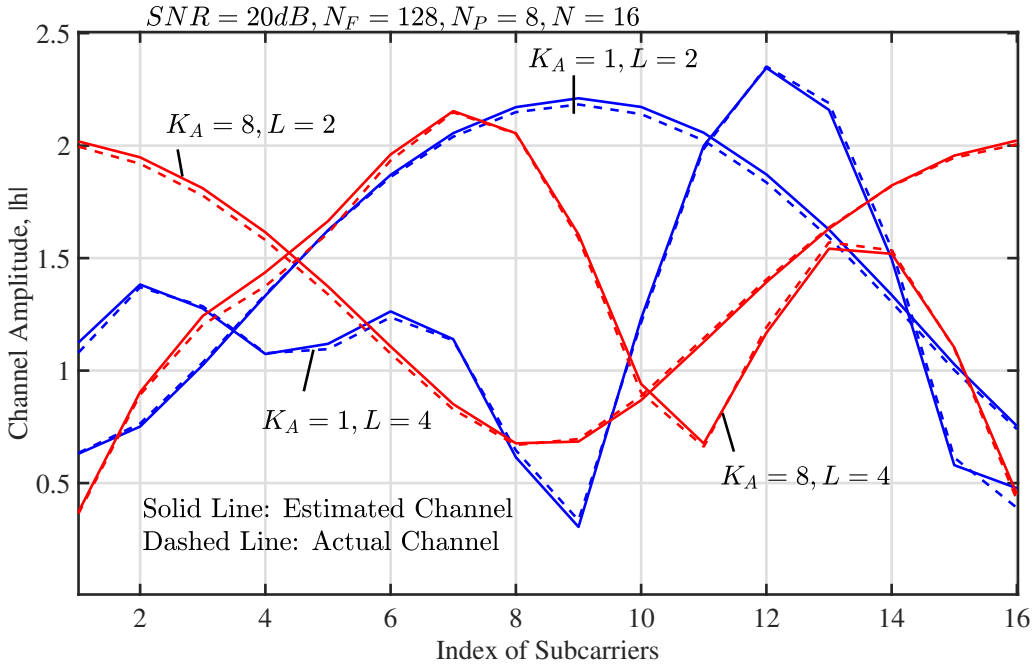
In this section, we provide some results to demonstrate the performance of the respective channel estimation schemes considered in the above section. The observations from them will be exploited for UAI in Section 3.6.

In Figs.3.4a and 3.4b, we show the snapshots of the MMSE-based channel estimation considered in Section 3.3. In our demonstrations, we assume a MC-mGFMA system with $N = 16$ subcarriers and supporting $K_A = K = 1$ or 8 active UEs known to AP. The frame length is set to $N_F = 128$, in which $N_p = 8$ pilots are used for channel estimation. We assume the frequency-selective fading channels with $L = 2$ or 4. Furthermore for Fig.3.4a, we assume that the channel is estimated using the symbol-based estimator of (3.29), while for Fig.3.4b, the channel is estimated by the block-based estimator of (3.20). From the results we can have the following observations. First, when the channel becomes more frequency selective resulted from L being changed from $L = 2$ to $L = 4$, the subcarrier channels become less correlated, or the channel is more variant in the frequency domain. Second, when the number of UEs is increased from $K_A = 1$ to $K_A = 8$, the symbol-based estimator shows slight performance degradation as the result of multiuser interference. By

[2]Also, this autocorrelation matrix and its inverse can be used for the data detection of active UEs



(a) Symbol-based channel estimation of (3.29).



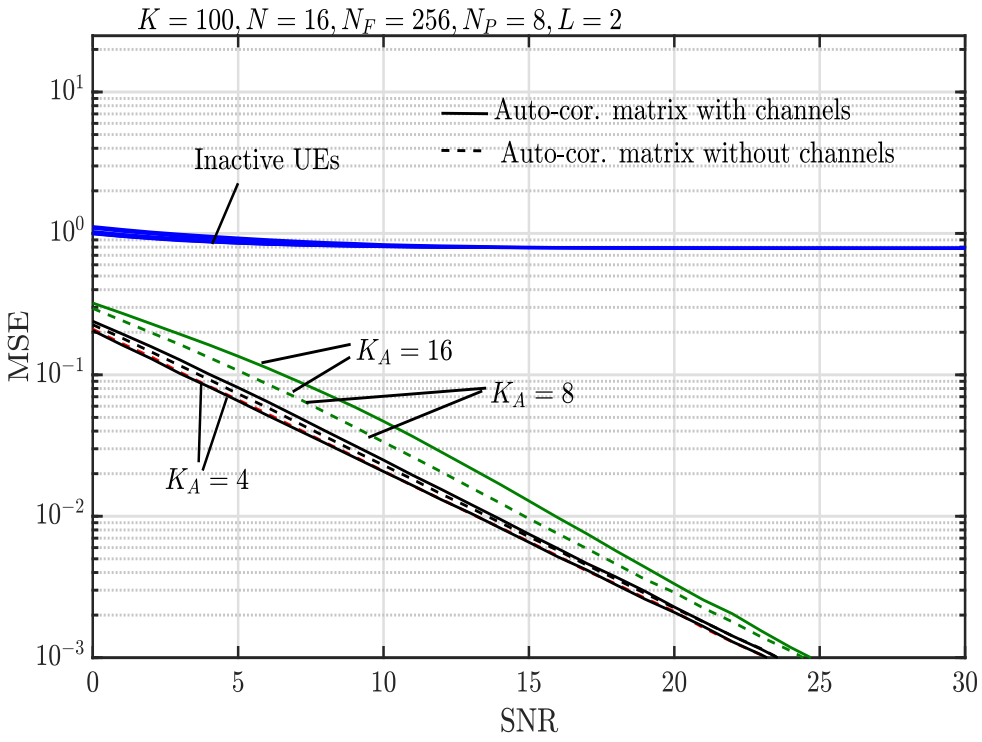
(b) Block-based channel estimation of (3.20).

Figure 3.4: Snapshots of estimated channels.

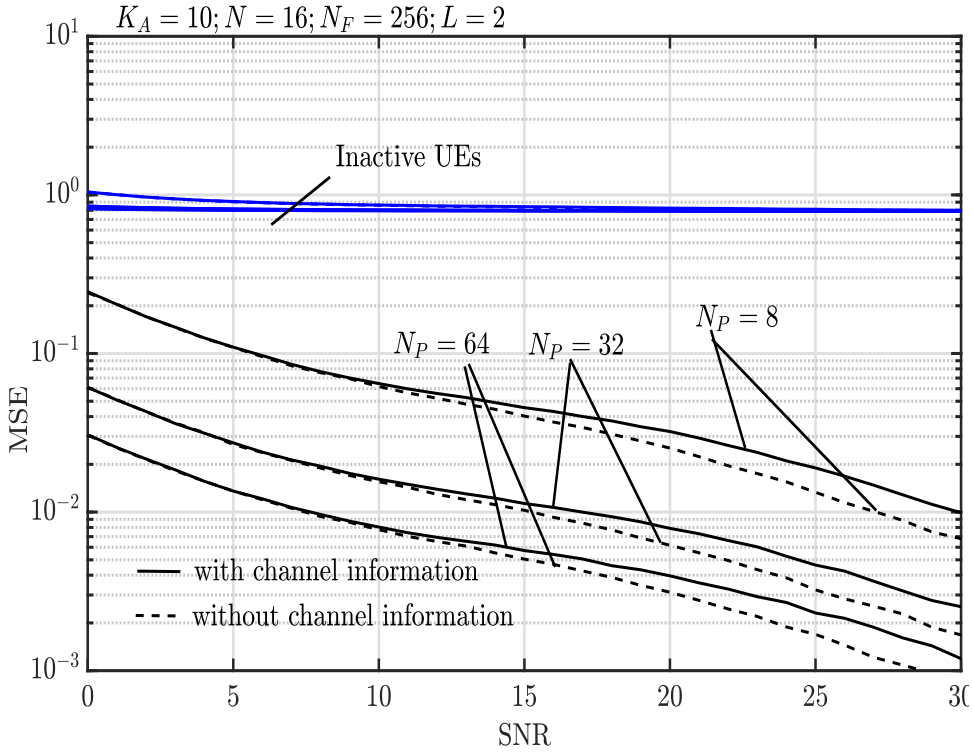
contrast, the block-based estimator demonstrates that in both cases, the estimated channels are close to the actual ones. The reason behind is that when $N = 16$ and $N_p = 8$, the symbol-based estimator has only $N = 16$ degrees of freedom, while the block-based estimator has $NN_p = 128$ degrees of freedom available for channel estimation. Third, by comparing Fig.3.4a with Fig.3.4b, we can explicitly see that the block-based estimator achieves more accurate estimation than the symbol-based estimator. Hence, once AP has some knowledge about active UEs, it can be exploited for improving the performance of channel estimation. This implies that among the three channel estimators, Estimator-K should outperform both Estimator-pK and Estimator-uK, while Estimator-pK outperforms Estimator-uK, as shown below.

In Section 3.3.1, we discussed that when the auto-correlation matrix $\boldsymbol{R}_a$ is estimated using (3.32), and $\boldsymbol{R}_y$ is constructed based on the estimated $\boldsymbol{R}_a$, the channels would be invoked in these autocorrelation matrices. However and ideally, it is desirable that these auto-correlation matrices are free from the channels to be estimated, as seen in (3.11). In Fig 3.5a and Fig. 3.5b, the impact of the invoked channels on the estimation performance is demonstrated. From both figures we can find that the auto-correlation matrix including the channels to be estimated only yields slight performance loss, in comparison to the auto-correlation matrix free from channels. Since the performance loss is insignificant, the estimated $\boldsymbol{R}_a$ is valid for application. Figs. 3.5a and 3.5b show that the MSE of the estimated channels for inactive UEs is usually very high and furthermore, does not change much with the increase of SNR. There will be more discussion about it in association with following figures.

In Fig.3.6, we investigate the MSE performance of the channel estimation for both active and inactive UEs, when AP either has ideal knowledge about active UEs (Fig.3.6a) or has no knowledge about active UEs (Fig. 3.6b) and hence, the channel estimation follows the principles in Section 3.3.2. The parameters used in our studies are detailed with the figures. From the results of Fig.3.6 we can observe that for inactive UEs, their channels are always estimated with very high MSE. This is because in this case, the receiver at AP can only estimate interference and noise, due to the fact that there are no signals transmitted by inactive UEs. By contrast, if the channel estimation is for an active UE, the MSE presents to reduce as SNR increases. Specifically, when the Estimator-K is considered, as shown in Fig.3.6a, the MSE performance is very promising due to the simultaneous use of all the $N_P = 32$ pilots for channel estimation. As SNR increases, even for the case of $K_A = 48$, the MSE performance converges to that of a single active UE. The reason behind is that there are in total $N_P \times N = 32 \times 32 = 1024$ degrees of freedom used for channel estimation, which is significantly larger than $K_A \cdot L = 48 \times 4 = 192$ of the number of variables estimated. Furthermore, as shown in Fig.3.6a, in the low SNR region,

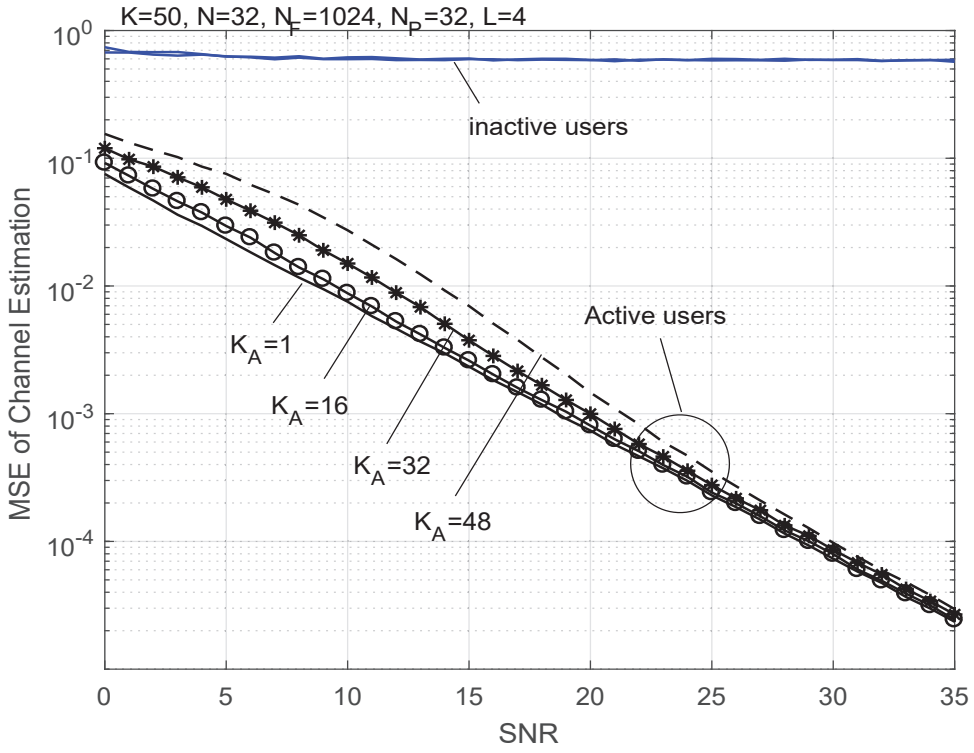


(a) Estimator-K.

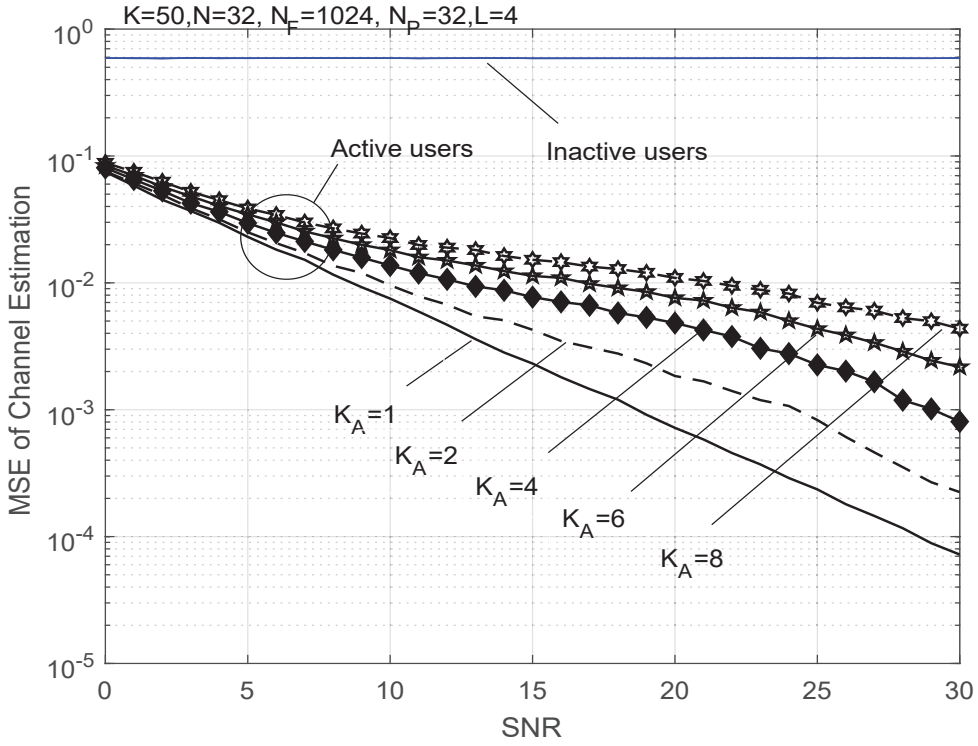


(b) Estimator-uK.

Figure 3.5: MSE performance of channel Estimator-K and Estimator-uK, when channel is invoked or not invoked in auto-correlation matrix.



(a) Estimator-K.



(b) Estimator-uK.

Figure 3.6: MSE performance of channel Estimator-K and Estimator-uK for both active and inactive UEs.

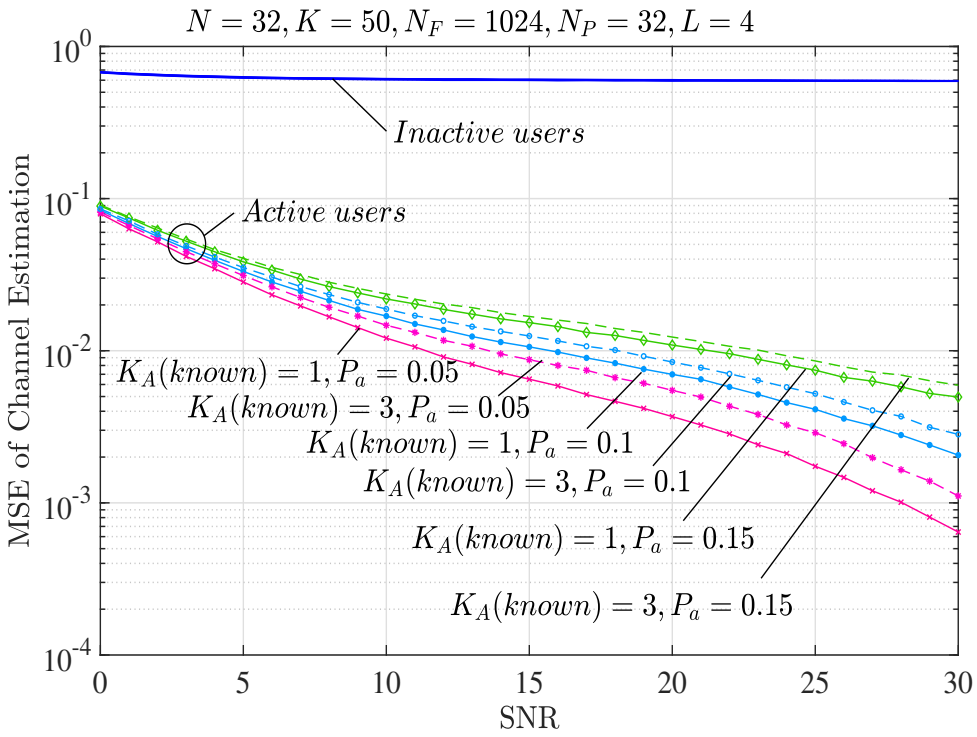


Figure 3.7: MSE performance of channel Estimator-pK for both active and inactive UEs.

the MSE performance slightly degrades, as the number of active UEs increases.

By contrast, as shown in Fig.3.6b for Estimator-uK, as only the symbol-based channel estimator can be implemented, multiuser interference imposes a big effect on the MSE performance and the MSE performance significantly degrades, when the number of active UEs increases. The reason behind this observation is that the channel from each UE has $L = 4$ variables to be estimated. However, corresponding to one symbol, there are only $N = 32$ observations (degrees-of-freedom). Hence, when there is one more active UE, there are 4 extra variables to be estimated, which is a relatively big number with respect to $N = 32$ and hence, significant performance degradation is observed, as shown in Fig.3.6b. Furthermore, this can explain why the MSE performance corresponding to $K_A = 8$ is very poor. In this case, there are in total $K_A \cdot L = 32$ variables to be estimated based on $N = 32$ observations per symbol.

When comparing Fig.3.6b with Fig.3.6a, a very important observation can be derived, i.e., the MSE performance shown in Fig.3.6a is much better than that in Fig.3.6b in the case of $K_A > 1$. Furthermore, as shown in Fig. 3.6a, the MSE performance of the considered cases converges, if SNR is sufficiently high. These observations imply that whenever there is some information about the activity of UEs available, this information may be exploited for improving the performance of channel estimation. This is the fundamental for the

Estimator-pK.

Therefore, Fig. 3.7 shows the MSE performance of channel estimation, when AP has partial knowledge of the active UEs. In the figure, $K_A(known) = 1$ or 3 means that the AP knows 1 or 3 active UEs, while the other UEs become active independently with the activation probability shown in the figure. As seen in the figure, the MSE performance degrades, as the activation probability per UE increases, resulting in an increase of active UEs in average. For a given activation probability, when the number of active UEs is changed from $K_A + 1$ to $K_A + 3$, the MSE performance only slightly degrades, as AP can make use of the Estimator-pK. Again, for the inactive UEs, the MSE of channel estimation is very high.

3.5 Statistics of Estimated Channels

In the above section, we have considered the channel estimation in MC-mGFMA systems. In order to achieve reliable signal detection of UEs, it is essential for the AP receiver to know which UEs are active. Before we detail the UE activity identification (UAI) schemes, let us first have a look of the statistical properties of the estimated channels corresponding to the active and inactive UEs.

Following the above section, the time-domain CIR of the ith UE is $\boldsymbol{h}_i$, which is assumed to have L taps. After the MMSE-assisted channel estimation in Section 3.3, the estimate to $\boldsymbol{h}_i$ can be written as

$$\hat{\boldsymbol{h}}_i = \begin{cases} \boldsymbol{h}_i + \boldsymbol{n}_i, & \text{if UE } i \text{ is active} \\ \boldsymbol{n}'_i, & \text{if UE } i \text{ is inactive} \end{cases} \tag{3.34}$$

where $\boldsymbol{n}_i$ is the channel estimation error of an active UE, which can be Gaussian approximated with a PDF of $\mathcal{CN}\left(\boldsymbol{0}, \sigma_1^2 \boldsymbol{I}_L\right)$, where σ_1^2 is the variance of channel estimation error. By contrast, $\boldsymbol{n}'_i$ is the estimate to an inactive UE, which can also be approximated to have the Gaussian distribution with a PDF of $\mathcal{CN}\left(\boldsymbol{0}, \sigma_0^2 \boldsymbol{I}_L\right)$, where σ_0^2 may be different from σ_1^2.

Let us now consider the statistics of $|\hat{\boldsymbol{h}}_i|^2$ on the condition that UE i is either active or inactive. First, when UE i is active, the elements of $\hat{\boldsymbol{h}}_i$ are iid complex Gaussian random variables, each of which has zero mean and a variance of $1/L + \sigma_1^2$, when Rayleigh fading channels are assumed, and when all component paths have the same power of $1/L$. Therefore, $|\hat{\boldsymbol{h}}_i|^2$ obeys the centre χ^2-distributions with $2L$ degrees of freedom, and the PDF

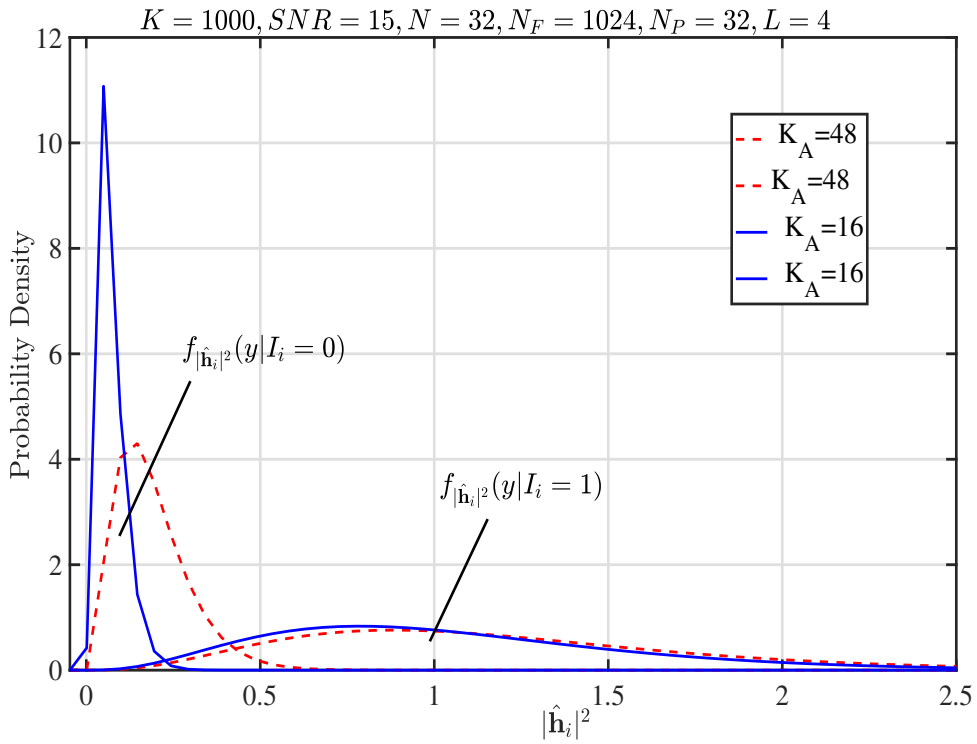


Figure 3.8: PDFs of $|\hat{\boldsymbol{h}}_i|^2$ on condition that UE i is active $(I_i = 1)$ or inactive $(I_i = 0)$, when channels are estimated based on the assumption that AP has no knowledge about the active UEs.

is [192]

$$f_{|\hat{\boldsymbol{h}}_i|^2}(\boldsymbol{y}|I_i = 1, K_A) = \frac{1}{(L-1)!(1/L + \sigma_1^2(K_A))^L}\boldsymbol{y}^{L-1} \times \exp\left(-\frac{\boldsymbol{y}}{1/L + \sigma_1^2(K_A)}\right), \; y \geq 0 \tag{3.35}$$

where we explicitly show that $\sigma_1^2(K_A)$ is related to the number of active UEs. By contrast, when the ith UE is inactive, it can be shown that $|\hat{\boldsymbol{h}}_i|^2$ follows the centre χ^2-distribution of

$$f_{|\hat{\boldsymbol{h}}_i|^2}(\boldsymbol{y}|I_i = 0, K_A) = \frac{1}{(L-1)!\sigma_0^{2L}(K_A)}\boldsymbol{y}^{L-1} \times \exp\left(-\frac{\boldsymbol{y}}{\sigma_0^2(K_A)}\right), y \geq 0 \tag{3.36}$$

Both (3.35) and (3.36) are central χ^2-distributions. For given K_A, we in general have $1/L + \sigma_1^2(K_A) > \sigma_0^2(K_A)$. Hence, from the properties of χ^2-distribution we can know that the $|\hat{\boldsymbol{h}}_i|^2$ in (3.35) corresponding to active UE is usually larger than that in (3.36) corresponding to inactive UE, as shown below.

Below are some exemplified PDFs obtained from simulations, which are shown in Figs. 3.8 - 3.12, when different settings and assumptions are considered. Note that, the

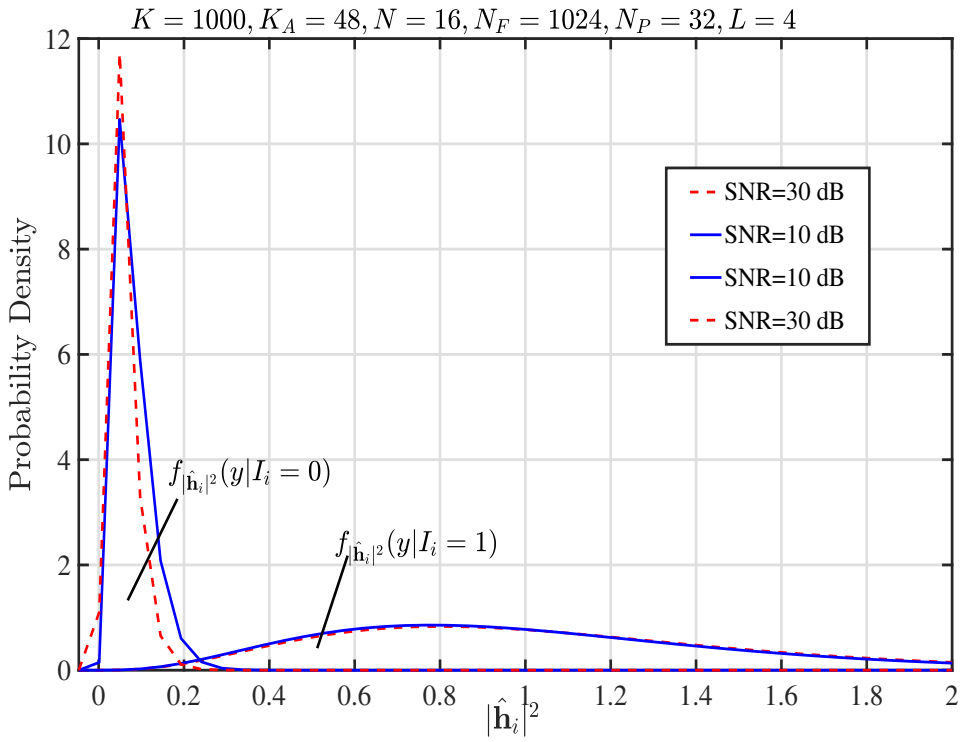


Figure 3.9: PDFs of $|\hat{\boldsymbol{h}}_i|^2$ on condition that UE i is active $(I_i = 1)$ or inactive $(I_i = 0)$, when channels are estimated based on the assumption that AP has no knowledge about the active UEs.

PDFs of $|\hat{\boldsymbol{h}}_i|^2$ on condition that UE i is active or inactive can be directly computed from (3.35) or (3.36), if we know σ_1^2 or σ_0^2. However, these variances need to be obtained from simulations. Therefore, we directly generate the PDFs from simulations. Specifically, in Figs. 3.8 - 3.10, the PDFs are obtained under the assumption that AP has no knowledge about the active UEs, corresponding to the channel estimation in Section 3.3.2. By contrast, in Figs. 3.11 and 3.12, the AP is assumed to know all the active UEs, which can hence jointly exploit the N_P pilot symbols used for channel estimation, as shown in Section 3.3.1. From these PDFs, we can explicitly conceive that if a UE is active, its estimated channel's magnitude, i.e., $|\hat{\boldsymbol{h}}_i|^2$, usually has a relatively high value. By contrast, if a UE is inactive, the value of $|\hat{\boldsymbol{h}}_i|^2$ is usually low and distributed mainly in the region close to zero. This distribution information of $|\hat{\boldsymbol{h}}_i|^2$ on condition of UE i being active or inactive can be exploited for the design of the UAI schemes with low-complexity, as shown in the next section.

3.6 Active UE Identification

In this section, we propose a range of UAI schemes designed based on the signal and

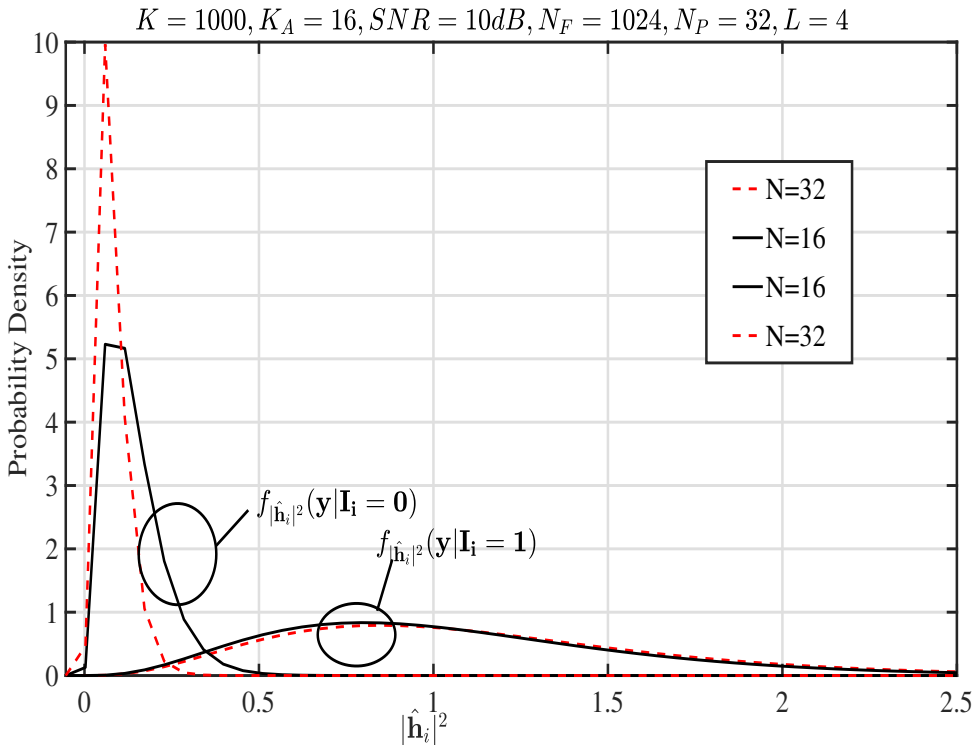


Figure 3.10: PDFs of $|\hat{\boldsymbol{h}}_i|^2$ on condition that UE i is active $(I_i = 1)$ or inactive $(I_i = 0)$, when channels are estimated based on the assumption that AP has no knowledge about the active UEs.

channel characteristics, and with the aid of the channel estimation approaches considered in Section 3.3. Furthermore, the UAI schemes are analyzed by considering some practical issues, and their performance is investigated and compared in Section 3.7. Let us first consider the threshold-based UAI (TB-UAI).

3.6.1 Threshold-Based UAI (TB-UAI)

From the statistical analysis in Section 3.5 we can explicitly see that $|\hat{\boldsymbol{h}}_i|^2$ usually has a small value, if UE i is inactive. By contrast, if UE i is active, $|\hat{\boldsymbol{h}}_i|^2$ usually takes a relatively big value. These observations straightforwardly imply the threshold-based UAI approach.

Let T_h be a threshold set for the TB-UAI. Then, when given the estimated CIR of $|\hat{\boldsymbol{h}}_i|^2$ for UE i, the TB-UAI identifies the ith UE's state according to

$$\hat{I}_i = \begin{cases} 1, & \text{if } |\hat{\boldsymbol{h}}_i|^2 \geq T_h, \\ 0, & \text{else} \end{cases} \tag{3.37}$$

where $\hat{I}_i = 1$ or 0 represents UE i is active or inactive.

Note that in the case that AP has no knowledge about the active UEs, the channels used in (3.37) are estimated using the Estimator-uK. If AP can obtain partial information about

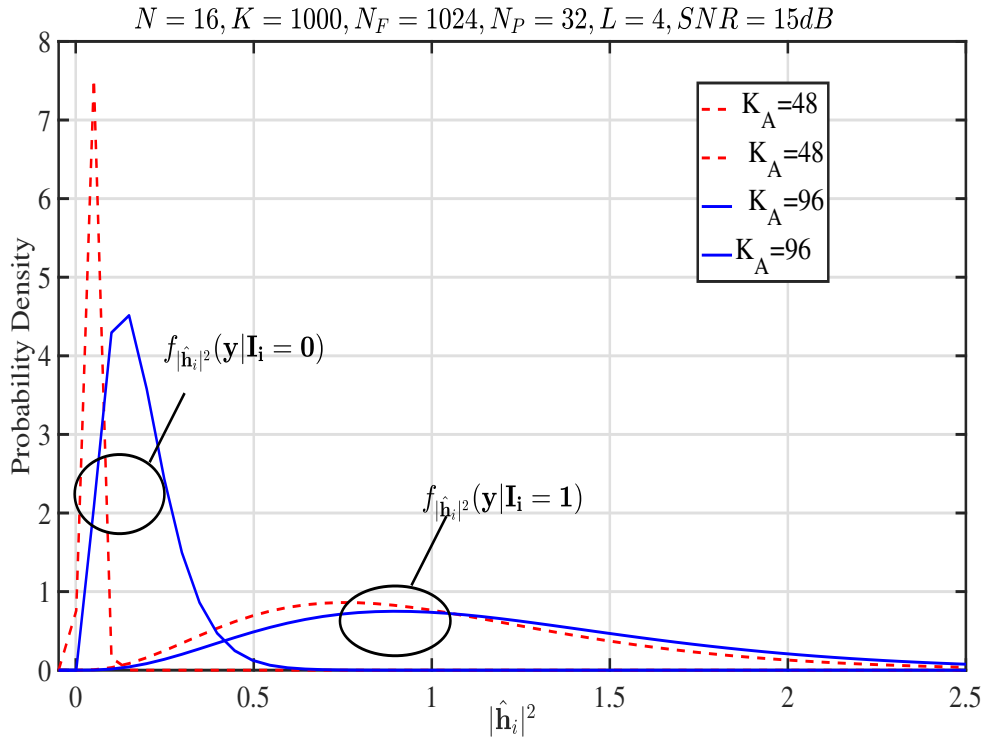


Figure 3.11: PDFs of $|\hat{\boldsymbol{h}}_i|^2$ on condition that UE i is active $(I_i = 1)$ or inactive $(I_i = 0)$, when channels are estimated based on the assumption that AP knows the active UEs.

the active UEs, the channels in (3.37) can be estimated by the Estimator-pK.

With the aid of the statistics of $|\hat{\boldsymbol{h}}_i|^2$ given in Section 3.5, we can derive the miss and false-alarm probabilities of the TB-UAI as follows. Firstly, the miss probability conditioned on K_A active UEs is

$$P_M(K_A) = \int_0^{T_h} (\boldsymbol{y}|I_i = 1, K_A)dy \tag{3.38}$$

Upon substituting (3.35) into (3.38) and completing the integration, we obtain

$$\begin{aligned} P_M(K_A) =& 1 - \exp\left(-\frac{LT_h}{1 + L\sigma_1^2(K_A)}\right) \\ &\times \sum_{k=0}^{L-1} \frac{1}{k!} \left(\frac{LT_h}{1 + L\sigma_1^2(K_A)}\right)^k \end{aligned} \tag{3.39}$$

Secondly, the false-alarm probability conditioned on K_A active UEs is

$$P_F(K_A) = \int_{T_h}^{\infty} f_{|\hat{\boldsymbol{h}}_i|^2}(\boldsymbol{y}|I_i = 0, K_A)dy \tag{3.40}$$

When substituting (3.36) into (3.40), we obtain

$$P_F(K_A) = \exp\left(-\frac{T_h}{\sigma_0^2(K_A)}\right) \sum_{k=0}^{L-1} \frac{1}{k!} \left(\frac{T_h}{\sigma_0^2(K_A)}\right)^k \tag{3.41}$$

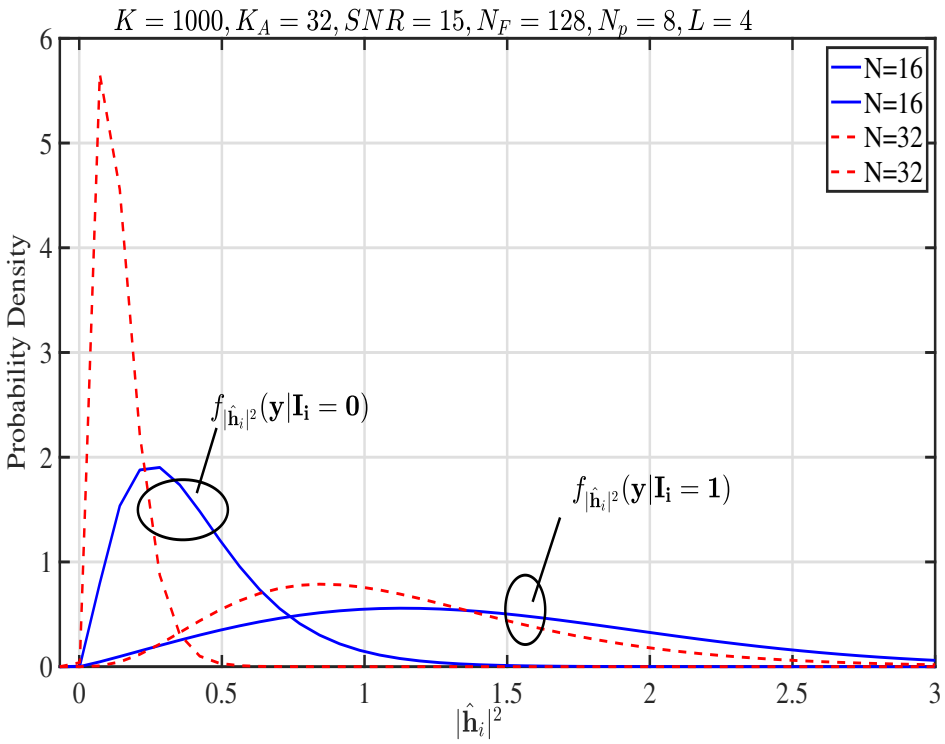


Figure 3.12: PDFs of $|\hat{\boldsymbol{h}}_i|^2$ on condition that UE i is active $(I_i = 1)$ or inactive $(I_i = 0)$, when channels are estimated based on the assumption that AP knows the active UEs.

Furthermore, the probability of erroneous identification conditioned on K_A active UEs is

$$P_E(K_A) = P_a P_M(K_A) + (1 - P_a) P_F(K_A) \tag{3.42}$$

where P_a is the activation probability of a UE.

Assume that UEs become active independently and all have the same activation probability P_a. Then, the number of active UEs K_A at a given time-slot obeys the binomial distribution with the probability mass function (PMF) of

$$P(K_A) = \binom{K}{K_A} P_a^{K_A} (1 - P_a)^{K - K_A}, \quad K_A = 0, 1, \cdots, K \tag{3.43}$$

With the aid of this distribution, the average miss probability is then given by

$$P_M = \sum_{K_A=1}^{K} P(K_A) P_M(K_A) \tag{3.44}$$

and the average false-alarm probability is

$$P_F = \sum_{K_A=0}^{K-1} P(K_A) P_F(K_A) \tag{3.45}$$

Finally, the average probability of erroneous identification is

$$P_E = P_a P_M + (1 - P_a) P_F \tag{3.46}$$

From (3.39) and (3.41) we can know that when T_h increases, the false-alarm probability of (3.41) reduces, while the miss probability of (3.39) increases. Therefore, the miss and false-alarm probabilities make trade-off against each other. In the TB-UAI algorithm, the threshold T_h may be chosen according to different requirements. First, it may be selected to minimize the erroneous probability P_E. Second, according to the practical requirement, it is sometimes desirable to have a small false-alarm probability guaranteed. In this case, the threshold can be chosen as $T_h = \arg\{\max\{P_F(T'_h)\} \leq \bar{P}_F\}$, where $\bar{P}_F$ is the maximum false-alarm probability allowed. Furthermore, in some applications, the miss of an active UE may be critical, but some false-alarms can be tolerated. In this case, the threshold can be selected to protect a small miss probability. Hence, when given a small miss probability of $\bar{P}_M$, the threshold is chosen as $T_h = \arg\{\max\{P_M(T'_h)\} \leq \bar{P}_M\}$.

However, in any of the above cases, determining an appropriate threshold for attaining near-optimum performance is highly challenging, as the distribution of $|\hat{\boldsymbol{h}}_i|^2$ is dependent on K_A of the number of active UEs, or on their activation probability P_a, if the average miss and false-alarm probabilities are considered, as shown in Figs. 3.8 - 3.12, and the formulas above in Section 3.6.1. In order to circumvent these challenges, below we propose the alternative UAI schemes, which are not dependent on threshold setting.

3.6.2 Eigen-Analysis Enhanced UAI (EAE-UAI)

It is well-known that with the aid of eigen-analysis, an AP may acquire the information about the number of active transmissions [148]. This approach can be introduced to enhance the UAI in our MC-mGFMA system, if the activation probability per UE is small or/and the number of potential UEs is small, resulting in that the number of active UEs has a small probability to exceed N. In detail, the EAE-UAI is operated as follows.

First, the AP estimates the autocorrelation matrix based on (3.31), which is a $(N \times N)$ symmetric matrix. Since this autocorrelation matrix is estimated by exploiting both pilot symbols and data symbols, i.e., exploiting all the N_F symbols in one frame, it should be sufficiently accurate for deriving the observations via the eigen-analysis on it. Then, according to the matrix theory [193], $\hat{\boldsymbol{R}}_a$ can be decomposed into a signal subspace having the dimensions of K_A if $K_A \leq N$, or of N if $K_A \geq N$. Let us execute the eigenvalue

decomposition (EVD) on $\hat{\boldsymbol{R}}_a$, we obtain

$$\hat{\boldsymbol{R}}_a = \boldsymbol{U}_s \boldsymbol{\Lambda}_s \boldsymbol{U}_s^H + \boldsymbol{U}_n \boldsymbol{\Lambda}_n \boldsymbol{U}_n^H \tag{3.47}$$

where $\boldsymbol{U}_s$ gives the signal subspace and has K_A columns, if $K_A \leq N$, while $\boldsymbol{U}_n$ is the noise (or null) subspace having $(N - K_A)$ columns, also if $K_A < N$. Hence, if $K_A < N$, both $\boldsymbol{\Lambda}_s$ and $\boldsymbol{\Lambda}_n$ in (3.47) are diagonal matrices with real eigenvalues being the diagonal elements. Let us assume that the N eigenvalues of $\hat{\boldsymbol{R}}_a$ are expressed as $\lambda_1, \lambda_2, \cdots, \lambda_N$. Then, according to the principles of EVD, ideally, we have

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{K_A} > \lambda_{K_A+1} = \lambda_{K_A+2} = \cdots = \lambda_N \tag{3.48}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_{K_A}$ are the eigenvalues corresponding to the signal subspace or the diagonal elements in $\boldsymbol{\Lambda}_s$. By contrast, $\lambda_{K_A+1}, \cdots, \lambda_N$ are the eigenvalues corresponding to the noise subspace, which form the diagonal elements of $\boldsymbol{\Lambda}_n$. Furthermore, according to the principles of EVD, these eigenvalues can be expressed as

$$\lambda_i = \begin{cases} P_{S_i} + \sigma^2, & \text{if } i \leq K_A \\ \sigma^2, & \text{if } i > K_A \end{cases} \tag{3.49}$$

where P_{S_i} is the signal power of the ith component, and σ^2 is the noise power.

Hence, with the aid of the eigenvalues obtained from the EVD of $\hat{\boldsymbol{R}}_a$, the number of active UEs K_A can be identified as the number of eigenvalues whose values are larger than the noise power.

In practice, the border between the signal and noise subspaces might not be very clear. In this case, AP may set a threshold based on the noise variance. Any eigenvalues larger than the threshold are identified as corresponding to the active UEs.

Once AP has the knowledge of K_A of the number of active UEs, the EAE-UAI can identify the K_A active UEs as the K_A UEs, whose channel magnitudes $|\hat{\boldsymbol{h}}_i|^2$ are the largest among the K UEs in the system. In summary, the EAE-UAI algorithm is stated Algorithm 3.6.2.

Algorithm 3.6.2 Eigen-analysis enhanced active UE identification (EAE-UAI) algorithm.

Input: $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_{N_F}$, $\{\boldsymbol{c}_k\}$, pilot symbols.

1. *Estimate the autocorrelation matrix using* (3.31).
2. *Derive K_A using the EVD in Section 3.6.2.*

3. *Estimate the channels of K UEs following Section 3.3 (using either Estimator-uK or Estimator-pK, depending on the UE activity knowledge available to AP).*

4. *Order the channels according to their power as* $|\hat{\boldsymbol{h}}_{1'}|^2 \geq |\hat{\boldsymbol{h}}_{2'}|^2 \geq \cdots \geq |\hat{\boldsymbol{h}}_{K'}|^2$, and identify the number of active UEs K_A.

5. *UE Identification*:

   if $K_A < N$, the UEs corresponding to the K_A largest terms of $|\hat{\boldsymbol{h}}_{K'}|^2$ are identified as the active UEs;

   if $K_A \geq N$, then the threshold-based UAI is used to determine the active UEs.

Note that at the last step concerning $K_A \geq N$, the AP may simply select a few of UEs more than N as the active UEs, under the constraint that the outage probability of

$$P_{out} = \sum_{k=K_A+1}^{K} \binom{K}{k} P_a^k (1-P_a)^{K-k} \tag{3.50}$$

is below a pre-set value, such as, 10^{-5}. In this way, the probability that active UEs are missed is small. However, there are possibly some false-alarm UEs, which should be readily removed after their data detection and with the aid of the pilots and/or error detection/correction. This is because signal detection of a false-alarm UE should have a bit error rate (BER) of about 0.5, which can be easily identified by the embedded error control coding [194]. Alternatively, the AP receiver can compare the detected pilot symbols of a UE with its actual pilots, and rendering the UE a false-alarm, if the number of erroneously detected pilot symbols is higher than a given value.

Nevertheless, whenever there is a false-alarm UE, the AP receiver has to contribute time and complexity for its signal detection. In order to further improve the UAI in our MC-mGFMA, below we propose the successive interference cancellation assisted UAI (SIC-UAI).

3.6.3 Successive Interference Cancellation Assisted UAI (SIC-UAI)

One of the motivations for introducing SIC to the UAI process is to solve the problem that in the EVE-UAI, AP is unable to know the number of active UEs, when $K_A \geq N$, and hence not sure when the identification process should end. Below we propose two SIC-UAI approaches for achieving the UAI objective. With the first approach, EVD is employed to determine whether the UAI process is completed. The second approach does

not make use of EVD, instead, it implements pilot detection to check when the UAI process can be finished.

The fundamental principle behind the two approaches is the involvement of the SIC operation. In detail, whenever a UE is identified to be active and its channel is estimated with relatively high reliability, its contribution to the received signals corresponding to pilots can then be removed with the aid of the information available to the AP. In this way, the following UEs' identification would become more reliable.

Specifically, let the ith UE be the identified active UE, and its estimated channel be expressed as $\hat{\boldsymbol{h}}_i$. Since AP knows the pilot symbols sent by UE i and also the spreading codes of UE i, based on (3.5), AP can carry out the SIC as

$$\boldsymbol{y}'_p = \boldsymbol{y}_p - b_i^{(p)} \boldsymbol{A}_i \hat{\boldsymbol{h}}_i, \ p = 1, 2, \cdots, N_P \tag{3.51}$$

After the SIC of UE i, the autocorrelation matrix can be updated for the identification of other active UEs.

Initially, let AP use the pilot signals to estimate the autocorrelation matrix as

$$\hat{\boldsymbol{R}}_a = \frac{1}{N_P} \sum_{p=1}^{N_P} \boldsymbol{y}_p \boldsymbol{y}_p^H \tag{3.52}$$

Note that, in contrast to (3.31) that can use all the received signals of a frame to estimate $\boldsymbol{R}_a$, here only pilot symbols are used, due to the employment of SIC. Therefore, we can be conceived that this kind of algorithms require that there are sufficient number of pilots, so that $\boldsymbol{R}_a$ can be estimated with sufficient accuracy. As our simulation results show, having 16 pilots is already capable of providing reasonable estimation.

Note furthermore that if there are active UEs known to AP, their contributions to observations can be cancelled by, first, estimating their channels based on the principles of Estimator-pK, and then, carrying out the interference cancellation based on (3.51). In the following two SIC-UAI schemes, we assume that all active UEs are unknown to AP, and therefore, only the Estimator-uK can be implemented.

With the above preparations, the SIC-UAI algorithms corresponding to the two approaches can now be stated as Algorithm 3.6.3.a and Algorithm 3.6.3.b, respectively, as follows. In these algorithms, I is the maximum number of iterations allowed, which can be set by assuming a small outage probability, such as 10^{-6}, base on (3.50).

Algorithm 3.6.3.a Successive Interference Cancellation Assisted UE Identification - Approach-a (SIC-UAIa).

Initialization: $\boldsymbol{y}_1^{(0)} = \boldsymbol{y}_1, \boldsymbol{y}_2^{(0)} = \boldsymbol{y}_2, \cdots \boldsymbol{y}_{N_P}^{(0)} = \boldsymbol{y}_{N_P}$;
$\{\boldsymbol{A}_k\}$, $\{b_{k,n}\}$, I.

Compute autocorrelation matrix $\hat{\boldsymbol{R}}_a^{(0)}$ *using* (3.52).

For $i = 1, 2, \ldots, I$:

1. *EVD on* $\hat{R}_a^{(i-1)}$ *to obtain* N *eigenvalues, which are ordered in descending order as* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$.
2. *If it is identified that* $K_A < N$, *active UEs are identified by following the EAE-UAI algorithm. Then, UE identification finishes.*
3. *If* $K_A > N$ *active UEs are implied from the EVD, executing the following operations:*
   - *Estimate the channels of the* $(K - i + 1)$ *unidentified UEs;*
   - *Find the UE with the maximum of* $|\hat{\boldsymbol{h}}_i|^2$, *expressed as* $k' = \arg\max_k \{|\hat{\boldsymbol{h}}_{1'}|^2, |\hat{\boldsymbol{h}}_{2'}|^2, \cdots, |\hat{\boldsymbol{h}}_{(K-i+1)'}|^2\}$;
   - *Interference cancellation of UE* k' *by following* (3.51);
   - *Update:* $\hat{\boldsymbol{R}}_a^{(i)} = \hat{\boldsymbol{R}}_a^{(i-1)} - \boldsymbol{A}_{k'} \hat{\boldsymbol{h}}_{k'} \hat{\boldsymbol{h}}_{k'}^H \boldsymbol{A}_{k'}^H$.

Outputs: Active UEs and their estimated channels.

Note that the SIC-UAIa algorithm may be extended by allowing to cancel several UE signals simultaneously at Step 3). This will accelerate the identification process, but at the cost of some performance loss.

The SIC-UAIa algorithm makes use of the properties of EVD to terminate the identification process. However, if the number of pilots is insufficient, making the estimation of $\boldsymbol{R}_a$ noisy, EVD might not be able to provide a clear boundary between signals and noise. This would degrade the UAI performance. Hence, we propose the SIC-UAI approach-b (SIC-UAIb) as Algorithm 3.6.3.b, which does not invoke EVD. Instead, whenever a UE is identified to be active, after its channel estimation, AP tries to detect the pilot symbols sent by the UE. If the number of erroneously detected pilot symbols is below a threshold value, such as $1/10$ of the pilot symbols, the activity of the UE is confirmed. Otherwise, if the error ratio is high, such as near 0.5, the UE is then rendered inactive.

Algorithm 3.6.3.b Successive Interference Cancellation Assisted UE Identification - Approach-b (SIC-UAIb).

Initialization: $\boldsymbol{y}_1^{(0)} = \boldsymbol{y}_1, \boldsymbol{y}_2^{(0)} = \boldsymbol{y}_2, \cdots \boldsymbol{y}_{N_P}^{(0)} = \boldsymbol{y}_{N_P}$;
$\{\boldsymbol{A}_k\}$, $\{b_{k,n}\}$, I.

Compute autocorrelation matrix $\hat{\boldsymbol{R}}_a^{(0)}$ *using* (3.52)*, and* $\left(\hat{\boldsymbol{R}}_a^{(0)}\right)^{-1}$.

For $i = 1, 2, \ldots, I$:

1. *Estimate the channels of the* $(K - i + 1)$ *unidentified UEs, and find the most reliable active UE as*
   $k' = \arg\max_k \{|\hat{\boldsymbol{h}}_{1'}|^2, |\hat{\boldsymbol{h}}_{2'}|^2, \cdots, |\hat{\boldsymbol{h}}_{(K-i+1)'}|^2\}$.

2. *Use the estimated channel* $\hat{\boldsymbol{h}}_{k'}$ *and* $\hat{\boldsymbol{R}}_a^{(i-1)}$ *to form* $\boldsymbol{w}_{k'}^{(i-1)}$ *for detecting the* N_P *pilot symbols of UE* k':

   - **if** *the number of erroneous pilot symbols is higher than a preset threshold* N_C*, the UAI process finishes;*
   - **if** *the number of erroneous pilot symbols is lower than or equal to* N_C*, UE* k' *is identified as active, and execute the following operations:*

     a) Interference cancellation*:*
     $\boldsymbol{y}_n^{(i)} = \boldsymbol{y}_n^{(i-1)} - b_{k',n}\boldsymbol{A}_{k'}\hat{\boldsymbol{h}}_{k'}$, for $n = 1, 2, \cdots, N_P$;
     b) *Update:* $\hat{\boldsymbol{R}}_a^{(i)} = \hat{\boldsymbol{R}}_a^{(i-1)} - \boldsymbol{A}_{k'}\hat{\boldsymbol{h}}_{k'}\hat{\boldsymbol{h}}_{k'}^H\boldsymbol{A}_{k'}^H$*, and*
     $\left(\hat{\boldsymbol{R}}_a^{(i)}\right)^{-1} = \left(\hat{\boldsymbol{R}}_a^{(i-1)}\right)^{-1} + \dfrac{\boldsymbol{w}_{k'}^{(i-1)}\left(\boldsymbol{w}_{k'}^{(i-1)}\right)^H}{1 - \hat{\boldsymbol{h}}_{k'}^H\boldsymbol{A}_{k'}^H\boldsymbol{w}_{k'}^{(i-1)}}$
     c) *Return 1).*

Outputs: Active UEs and their estimated channels.

Note that, as shown in Algorithm 3.6.3.b, in Step 2), the SIC-UAIb algorithm detects the most reliably identified UE. In MMSE detection principle, during the ith iteration, the decision variables corresponding to the N_P pilots can be formed as

$$z_{k'}(n) = \left(\boldsymbol{w}_{k'}^{(i-1)}\right)^H \boldsymbol{y}_n^{(i-1)},\ n = 1, 2, \ldots, N_P \tag{3.53}$$

where $\boldsymbol{w}_{k'}^{(i-1)} = \left(\boldsymbol{R}_a^{(i-1)}\right)^{-1}\boldsymbol{A}_{k'}\hat{\boldsymbol{h}}_{k'}$. Then, the decision is made as $\hat{b}_{k',n} = 1$, if $\Re\{z_{k'}(n)\} > 0$, and $\hat{b}_{k',n} = -1$, otherwise.

Note furthermore that at Step 2) when the most reliable UE k' is cancelled from the received signals, both the autocorrelation matrix and its inverse are updated. Here, the inverse is obtained by applying the *matrix inverse lemma* on $\left(\hat{\boldsymbol{R}}_a^{(i-1)} - \boldsymbol{A}_{k'}\hat{\boldsymbol{h}}_{k'}\hat{\boldsymbol{h}}_{k'}^H\boldsymbol{A}_{k'}^H\right)^{-1}$, yielding

$$\begin{aligned}\left(\hat{\boldsymbol{R}}_a^{(i)}\right)^{-1} &= \left(\hat{\boldsymbol{R}}_a^{(i-1)}\right)^{-1} \\ &+ \frac{\left(\boldsymbol{R}_a^{(i-1)}\right)^{-1}\boldsymbol{A}_{k'}\hat{\boldsymbol{h}}_{k'}\left[\left(\boldsymbol{R}_a^{(i-1)}\right)^{-1}\boldsymbol{A}_{k'}\hat{\boldsymbol{h}}_{k'}\right]^H}{1-\hat{\boldsymbol{h}}_{k'}^H\boldsymbol{A}_{k'}^H\left(\boldsymbol{R}_a^{(i-1)}\right)^{-1}\boldsymbol{A}_{k'}\hat{\boldsymbol{h}}_{k'}}\end{aligned} \tag{3.54}$$

Substituting $\boldsymbol{w}_{k'}^{(i-1)} = \left(\boldsymbol{R}_a^{(i-1)}\right)^{-1}\boldsymbol{A}_{k'}\hat{\boldsymbol{h}}_{k'}$ into the above equation, we obtain the updating formula in the algorithm. In this way, the algorithm only needs to compute the inverse of autocorrelation matrix once, which allows to significantly reduce complexity.

Algorithm 3.6.3.b assumes that each iteration identifies only one UE, which results in an UAI delay proportional to the actual number of active UEs. In order to shorten the UAI process, several *most reliable* UEs can be simultaneously identified and processed during one iteration. In this way, the UAI delay can be reduced, but at the cost of some performance loss. Furthermore, in order to mitigate the performance loss as above-mentioned, the number of UEs identified in one iteration can be set in an adaptive way. At first, it may be set to a relatively large value, when the *most reliable* UEs can be easily identified. Then, the number is gradually reduced to improve the reliability of identification of the relatively weak UEs.

3.6.4 Auto-correlation Matrix Evolving UAI (AME-UAI)

As the channel estimation performance results in Section 3.4 suggest, whenever AP employs some knowledge about active UEs, this knowledge can be exploited by the Estimator-pK to improve channel estimation. This property can also be exploited to enhance the performance of UAI, yielding the AME-UAI algorithm, which is described as Algorithm 3.6.4.

Algorithm 3.6.4 Auto-correlation matrix evolving UAI (AME-UAI).

Initialization: $\boldsymbol{y}_1^{(0)} = \boldsymbol{y}_1, \boldsymbol{y}_2^{(0)} = \boldsymbol{y}_2, \cdots \boldsymbol{y}_{N_P}^{(0)} = \boldsymbol{y}_{N_P}$;
$\{\boldsymbol{A}_k\}$, $\{b_{k,n}\}$, I.

1) Compute autocorrelation matrix $\hat{\boldsymbol{R}}_a$ using (3.31)*; 2) Construct $\hat{\boldsymbol{R}}_y^{(0)}$ of* (3.33) *with $\hat{\boldsymbol{R}}_a$ being the diagonal elements and zero for all other elements; 3) Compute* $\left(\hat{\boldsymbol{R}}_y^{(0)}\right)^{-1}$.

For $i = 1, 2, \ldots, I$:

1. *Estimate the channels of the* $(K - i + 1)$ *unidentified UEs based on Estimator-pK in Section 3.3.3, and find the most reliable active UE as* $k' = \arg\max_k \{|\hat{\boldsymbol{h}}_{1'}|^2, |\hat{\boldsymbol{h}}_{2'}|^2, \cdots, |\hat{\boldsymbol{h}}_{(K-i+1)'}|^2\}$.

2. *Use the estimated channel* $\hat{\boldsymbol{h}}_{k'}$ *and* $\hat{\boldsymbol{R}}_a$ *to form* $\boldsymbol{w}_{k'}^{(i-1)}$ *for detecting the* N_P *pilot symbols of UE* k':

   - **if** *the number of erroneous pilot symbols is higher than a preset threshold* N_C*, the UAI process finishes;*
   - **if** *the number of erroneous pilot symbols is lower than or equal to* N_C*, UE* k' *is identified as active. Then, execute the following operations:*
     a) *Updating* $\hat{\boldsymbol{R}}_y^{(i-1)}$ *to* $\hat{\boldsymbol{R}}_y^{(i)}$ *following* (3.33) *by adding the non-diagonal component contributed by UE* k'*;*
     b) *Updating* $\left(\hat{\boldsymbol{R}}_y^{(i-1)}\right)^{-1}$ *to* $\left(\hat{\boldsymbol{R}}_y^{(i)}\right)^{-1}$*;*
     c) *Return 1).*

   Outputs: Active UEs and their estimated channels.

As shown in Algorithm 3.6.4, it first estimates the auto-correlation matrix $\hat{\boldsymbol{R}}_a$. Since there is no interference cancellation operation, this estimation can exploit all the received signals in one frame, including both the pilot signals and data carrying signals, i.e., $\hat{\boldsymbol{R}}_a$ can be estimated based on (3.31). Based on $\hat{\boldsymbol{R}}_a$, the initial auto-correlation matrix $\hat{\boldsymbol{R}}_y^{(0)}$ considering all the N_P pilots can be constructed, with $\hat{\boldsymbol{R}}_a$ being the diagonal block matrices, while all the other elements being zero. Then, the Algorithm forwards to the UAI stage, which iteratively identifies the active UEs and updates the auto-correlation matrix $\hat{\boldsymbol{R}}_y$. In detail, given $\hat{\boldsymbol{R}}_y^{(i-1)}$ obtained at the $(i-1)$-th iteration, the $(K - i + 1)$ unidentified UEs' channels are estimated by following the principles of Estimator-pK. The corresponding weight matrix for the k-th unidentified UE is expressed as

$$\boldsymbol{W}_k = \left(\hat{\boldsymbol{R}}_y^{(i-1)}\right)^{-1} (\boldsymbol{p}_k \otimes \boldsymbol{A}_k)/L \tag{3.55}$$

Then, the UE has the highest reliability being an active UE is identified. Furthermore, following the identification and with the aid of the estimated channels, the pilot symbols of the identified UE are detected and compared with the pilots symbols of the UE. If the number of errors is high, such as, close to 0.5, the UAI process can be rendered completed.

Otherwise, if the number of errors is relatively low, and in favor of a positive identification, the algorithm updates $\hat{\boldsymbol{R}}_y^{(i-1)}$ to $\hat{\boldsymbol{R}}_y^{(i)}$, and prepares to forward to the next iteration.

3.6.5 Complexity Analysis and Discussion

Above five UAI algorithms have been proposed to carry out UAI (TB-UAI and EAE-UAI) or jointly execute UAI and channel estimation (SIC-UAIa, SIC-UAIb and AME-UAI). In this subsection, we analyze their complexity. In our analysis, we assume that estimating the channel of one UE needs the computations expressed as C_e, so as to focus our attention on the complexity of UAI schemes. Note that in practice, the channel estimation of all UEs can be conducted in parallel, if latency is a critical consideration.

First, for the TB-UAI scheme, estimating K UEs channels requires KC_e computations. Then, computing $|\hat{\boldsymbol{h}}_i|^2$ for K UEs requires $2KL$ real multiplications and $K(L-1)$ real additions. Finally, there are K comparisons required to compare with the threshold. Hence in total, the number of operations required by the TB-UAI scheme is $KC_e + 3KL$. Furthermore, we can be implied from Section 3.3 that $C_e >> 3L$. Hence, the complexity of the TB-UAI scheme is determined by KC_e.

Second, for the EAE-UAI algorithm, the autocorrelation matrix has been obtained during the channel estimation stage. Based on the QR algorithm, the EVD algorithm has a complexity of $\mathcal{O}(N^3)$ [195]. Sorting the N eigenvalues using the quick-sort algorithm [196] needs in average $2N \ln N$ comparisons. Again, estimating K UEs channels requires KC_e computations, and computing $|\hat{\boldsymbol{h}}_i|^2$ of K UEs requires $2KL$ real multiplications and $K(L-1)$ real additions. Sorting the K number of $|\hat{\boldsymbol{h}}_i|^2$ needs in average $2K \ln K$ comparison. Hence in total, the complexity of the EAE-UAI algorithm is proportional to $KC_e + N^3 + 2N \ln N + 3KL + 2K \ln K$.

Third, for the SIC-UAIa algorithm, let us assume that in average $\bar{I}$ SIC iterations are implemented, and $K >> I$, due to the fact that the number of potential UEs is usually significantly larger than the number of active UEs. Then, after ignoring those insignificant operations, it can be shown that during these iterations, the complexity is proportional to $\bar{I}\left(KC_e + N^3 + 2N^2 + NL\right)$. At the final iteration without SIC, the EAE-UAI algorithm is operated, which has the complexity proportional to $KC_e + N^3 + 2N \ln N + 3KL + 2K \ln K$. Hence, the total number of operations is proportional to $(\bar{I}+1)KC_e + (\bar{I}+1)N^3 + 2N \ln N + 3KL + 2K \ln K + 2\bar{I}N^2 + \bar{I}NL$.

For the SIC-UAIb algorithm, again, we assume that the average number of iterations is $\bar{I}$, and $K >> I$. First, estimating the channels and finding the most reliable one need KC_e and K operations, respectively. Second, computing the weight vector $\boldsymbol{w}_{k'}^{(i-1)}$ needs

Table 3.2: Complexity of UAI schemes in Chapter 3.

UAI scheme	Complexity	Characteristic	Advantages	Disadvantages
TB-UAI	KC_e	Designed based on the distinct properties of the estimated channels for active and inactive UEs.	Identify directly.	
EAE-UAI	$KC_e + N^3 + 2N\ln N + 3KL + 2K\ln K$	Eigen-analysis is introduced to find the number of active UEs during the UAI process.	The eigenvalues obtained from the EVD of $\hat{\boldsymbol{R}}_a$, the number of active UEs K_A can be identified as the number of eigenvalues whose values are larger than the noise power.	AP is unable to know the number of active UEs, when $K_A \geq N$.
SIC-UAIa	$(\bar{I}+1)KC_e + (\bar{I}+1)N^3 + 2N\ln N$	Whenever a UE is identified to be active and its channel is estimated with relatively high reliability, its contribution to the received signals corresponding to pilots can then be removed with the aid of the information available to the AP.	It makes use of the properties of EVD.	If the number of pilots is insufficient, making the estimation of $\boldsymbol{R}_a$ noisy.
SIC-UAIb	$\bar{I}\left(KC_e + 2N(2N+L+N_P)\right)$	Whenever a UE is identified to be active and its channel is estimated with relatively high reliability, its contribution to the received signals corresponding to pilots can then be removed with the aid of the information available to the AP.	The most reliable UEs can be easily identified.	Each iteration identifies only one UE.
AME-UAI	$\bar{I}\left(KC_e + N^3N_P^3\right)$	When AP employs some knowledge about active UEs, this knowledge can be exploited by the Estimator-pK to improve channel estimation.	Since there is no interference cancellation operation, this estimation can exploit all the received signals in one frame, including both the pilot signals and data carrying signals.	

$2N(N+L)$ multiplications and additions. Detecting N_P pilots needs about $2NN_P$ multiplications. Decision making, finding the number of errors and interference cancellation are all insignificant, and their contribution to complexity can be ignored. Finally, since $\boldsymbol{A}_{k'}\hat{\boldsymbol{h}}_{k'}$ has been computed when computing $\boldsymbol{w}_{k'}^{(i-1)}$, it can be shown that updating the autocorrelation matrix needs about $2N^2$ operations. Hence, the complexity of the SIC-UAIb algorithm is proportional to $\bar{I}\left(KC_e + 2N(2N+L+N_P)\right)$.

Finally for the AME-UAI algorithm, the complexity is without any doubt dominated by the channel estimation and the inverse of $\hat{\boldsymbol{R}}_y^{(i)}$, which has the dimensions of $(NN_P \times NN_P)$. Hence, when assuming that the average number of iterations is $\bar{I}$, the computation amount of KC_e for channel estimation per iteration, and the computation requirement of $\mu N^3N_P^3$ for the inverse of $\hat{\boldsymbol{R}}_y^{(i)}$, the total computations of AME-UAI is proportional to $\bar{I}\left(KC_e + N^3N_P^3\right)$.

In summary, the complexity of the UAI schemes are listed in Table 3.2. Obviously, TB-UAI algorithm has the lowest complexity. The complexity of SIC-UAIa algorithm is higher than that of the EAE-UAI algorithm, as SIC-UAIa algorithm carries out SIC and several EVD operations may be executed during one identification process. When comparing SIC-UAIa with SIC-UAIb algorithms, SIC-UAIa algorithm needs to execute EVD, the number of iterations required by SIC-UAIa algorithm is usually significantly less than that required by SIC-UAIb algorithm. Finally, the AME-UAI has the highest complexity, as it needs a similar number of iterations as the SIC-UAIb algorithm. Furthermore, SIC-UAIb only

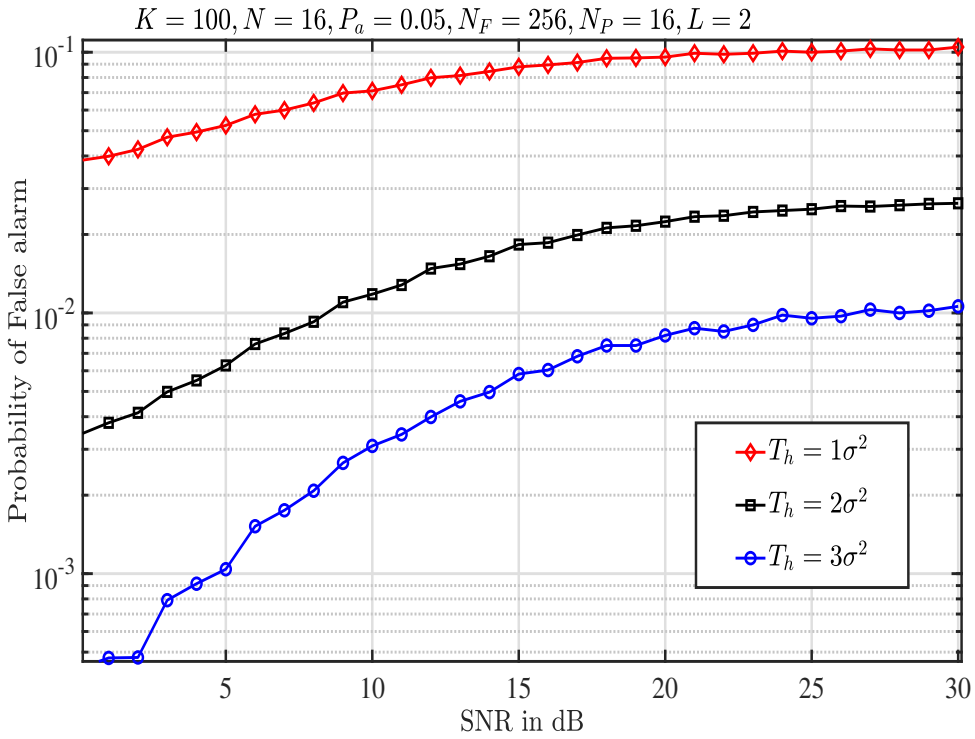
needs to update the matrix $\hat{R}_a^{(i-1)}$ to $\hat{R}_a^{(i)}$, which is size $(N \times N)$ and can also be achieved using the existing low-complexity algorithm. By contrast, in the AME-UAI algorithm, we did not find an existing algorithm to efficiently update $\hat{R}_a^{(i-1)}$ to $\hat{R}_a^{(i)}$, which has the dimensions of $(NN_P \times NN_P)$.

3.7 Performance of Active UE Identification

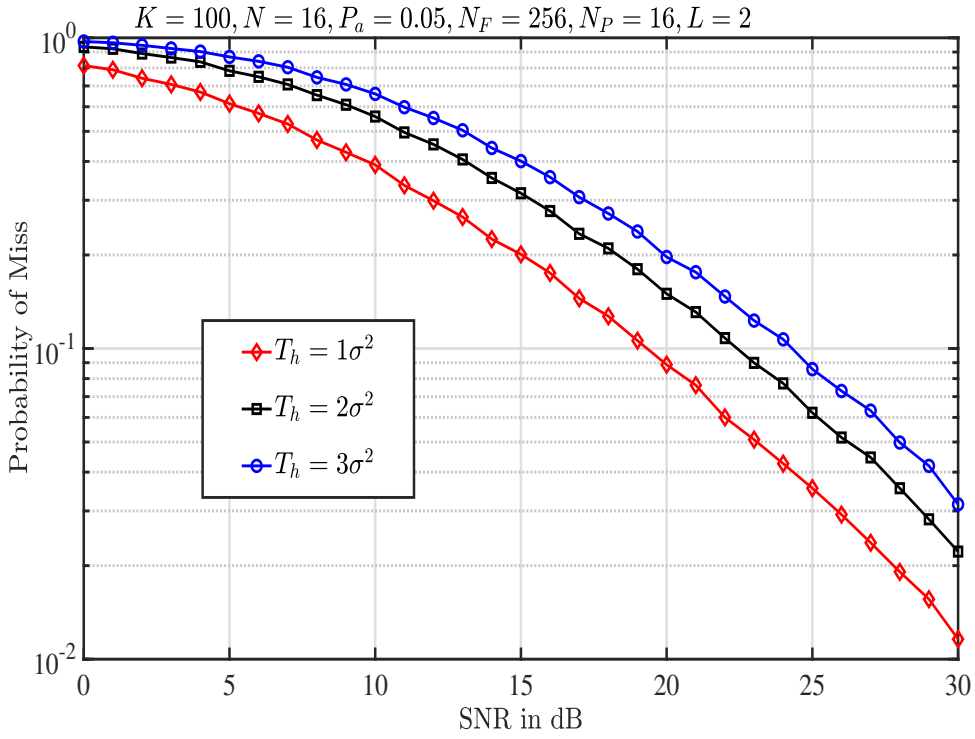
In this section, the performance of the different UAI algorithms considered in the last section is investigated. We emphasize both the miss and false-alarm probabilities.

In Figs. 3.13a and 3.13b, we demonstrate the probabilities of false-alarm and miss of the TB-UAI algorithm, when a MC-mGFMA system with $N = 16$ subcarriers and supports $K = 100$ potential UEs of each having the activation probability of $P_a = 0.05$. Furthermore, as shown in the figures, the frame length is $N_F = 256$ bits, among which $N_P = 16$ pilots are used, and the frequency-selective fading channel has $L = 2$ paths. Additionally, the thresholds are set to $T_h = \sigma^2$, $2\sigma^2$ or $3\sigma^2$, where $\sigma^2 = 1/SNR$. The results show that the miss probability increases and simultaneously, the false-alarm probability reduces, as the threshold increases. This implies that there is a trade-off between the miss and false-alarm probabilities, as shown in Fig. 3.14. Hence, when given a false-alarm (or miss) probability, a suitable threshold may be found to minimize the miss (or false-alarm) probability. However, near optimum threshold is hard to derive, as it is sensitive to the SNR and in particular, to the number of active UEs, which is dynamic in mGFMA systems. The results in Figs. 3.13a and 3.13b also show that for a given $T_h = \sigma^2$, $2\sigma^2$ or $3\sigma^2$, the false-alarm probability increases and the miss probability decreases, as SNR increases. This can be explained with the aid of the statistics of estimated channels, for example, as shown in Fig. 3.9. As shown in this figure, when SNR is changed from 10 dB to 30 dB, the distributions do not appear significant differences. However, when SNR increases, the threshold $T_h = \sigma^2 = 1/SNR$ significantly reduces. Specifically, the threshold is reduced from 0.1 to 0.001, when SNR is changed from 10 dB to 30 dB. Due to these effects, the false-alarm probability increases and the miss probability decreases, when SNR increases.

In Fig. 3.15, we show the probabilities of false-alarm and miss generated by the EAE-UAI algorithm, with the parameters as detailed with the figure. From the results we observe that both the false-alarm and miss probabilities decrease, when SNR increases. This is the result that the EVD in EAE-UAI algorithm allows more confident distinction between signals and noise, when SNR becomes higher. Furthermore, as seen in Fig. 3.15, both the false-alarm and miss probabilities reduce, when N_P increases. This is because when N_P increases, a more accurate estimation to the autocorrelation matrix can be attained, which



(a) Probability of false-alarm



(b) Probability of miss

Figure 3.13: Probability of false-alarm and miss, when active UEs are identified using the TB-UAI algorithm.

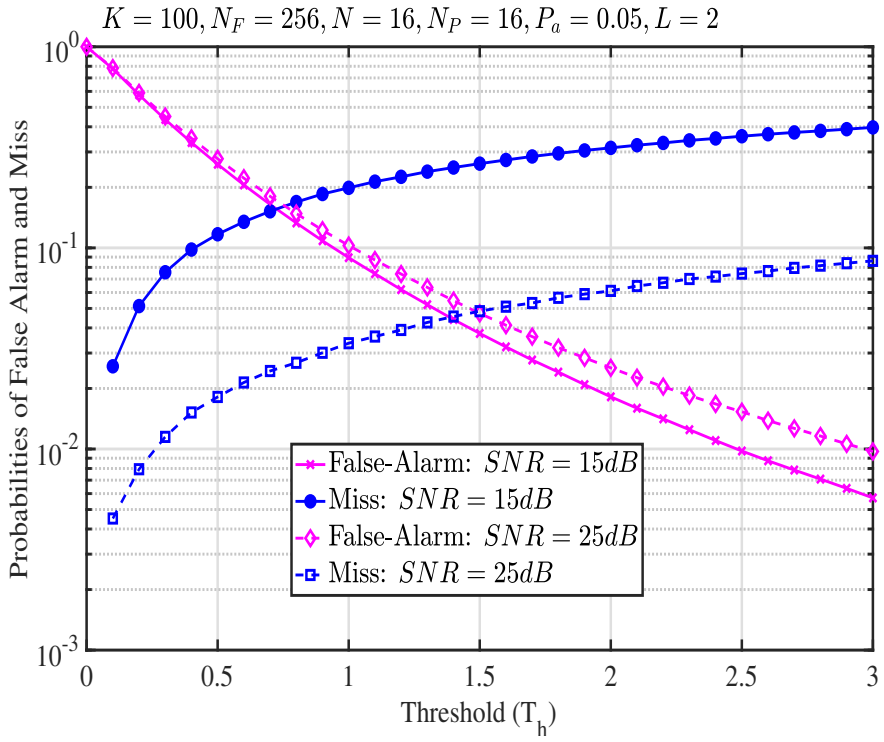
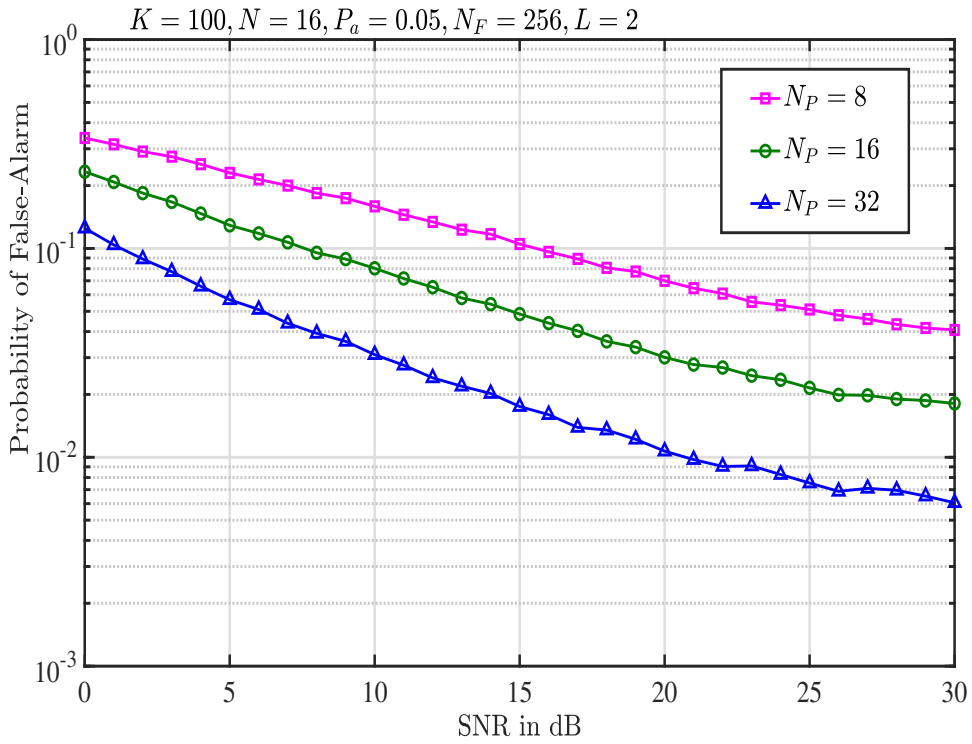


Figure 3.14: Trade-off between the probabilities of false-alarm and miss resulted by the TB-UAI algorithm.

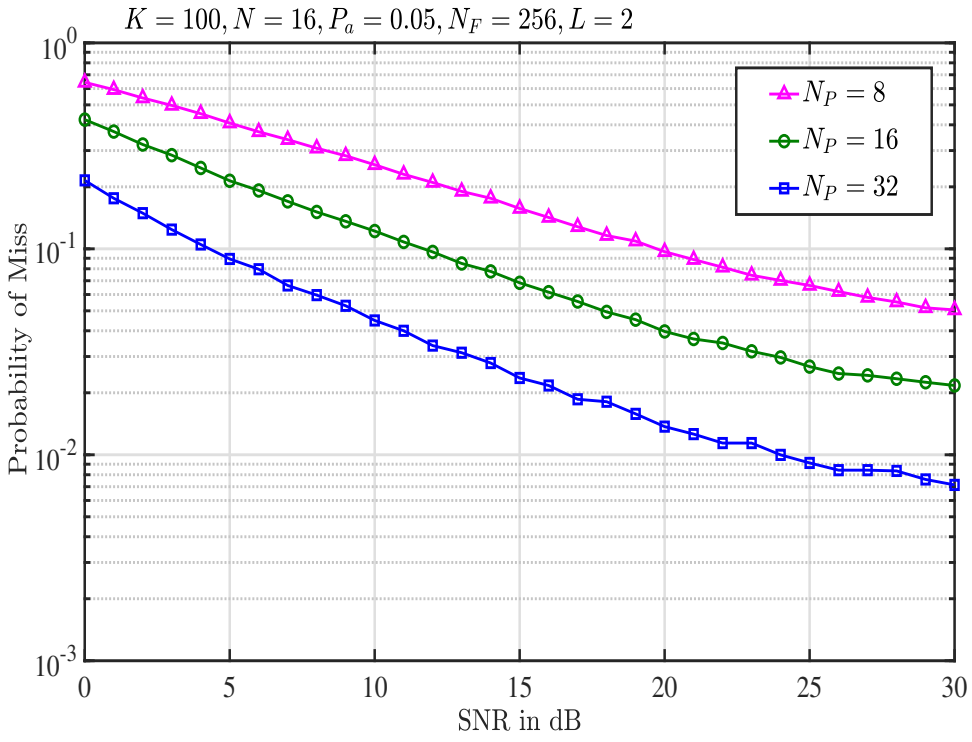
in turn enhances the distinction between signals and noise.

In Fig.3.16, we depict the probabilities of false-alarm and miss yielded by the SIC-UAIa algorithm, when it is operated with the MC-mGFMA systems with their parameters as shown in the figures. For the SIC-UAIa algorithm, three operational cases are considered, which are: (a) the number of active UEs is known to AP; (b) the number of active UEs is unknown to AP, or (c) the active UEs are partially known to AP. Hence in the case of (a), the AP only needs to identify which are the active UEs. By contrast, in the case of (b), the AP requires to identify both the number of active UEs and who they are. Finally, in the context of the case of (c), we assume that any newly activated UE at a time has the probabilities of P_1, P_2, P_3 and P_4 to transmit 1, 2, 3, or 4 frames. We assume that these probabilities obey the exponential distribution with $P_i = Ce^{-\alpha i}$, $i = 1, 2, 3, 4$, where $C = 0.641$ and $\alpha = 0.4196$. Hence, except the first frame, the AP knows a part of active UEs during all the following frames, which can be exploited to enhance the UAI performance, as analyzed previously in Section 3.6.

The results of Fig.3.16 validate the prediction for the performance of the SIC-UAIa algorithm operated in the case of (c). The partial information about the active UEs can be exploited to significantly enhance the performance of UAI. In general, the performance of case (a) is better than that of case (c), and the performance of case (b) is the worst.

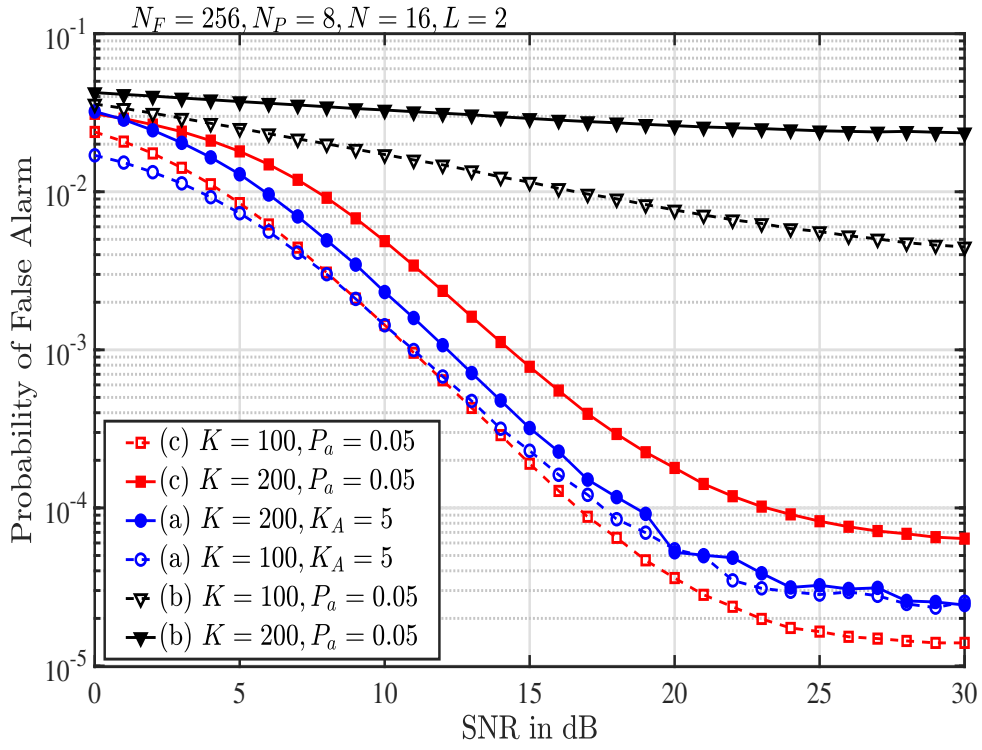


(a) Probability of false-alarm

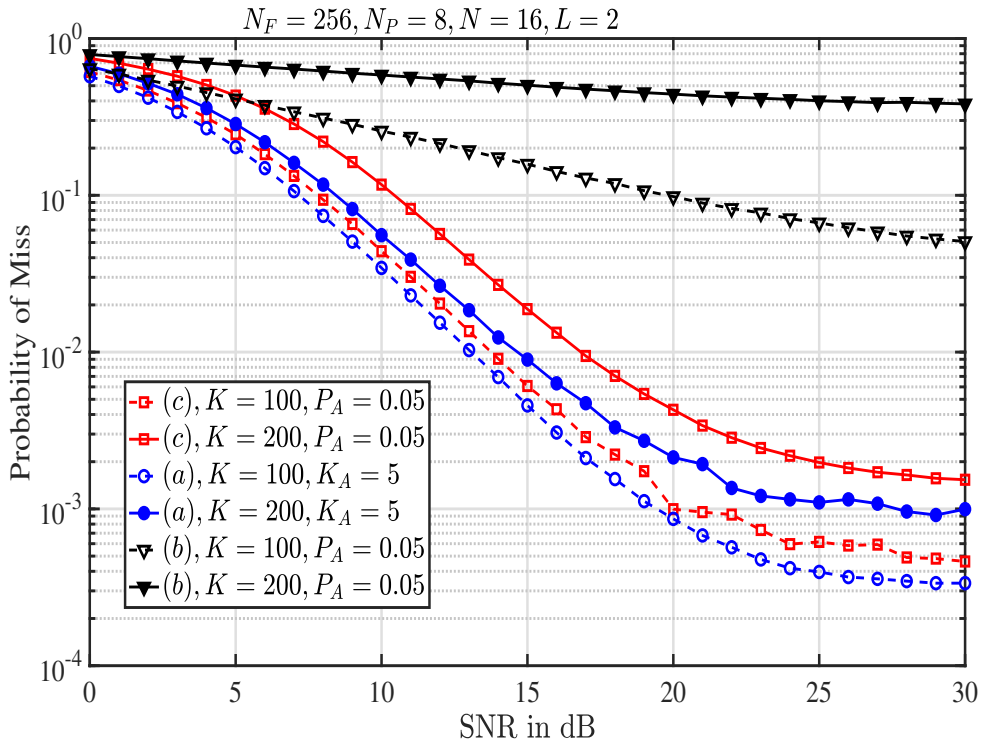


(b) Probability of miss

Figure 3.15: Probability of false-alarm and miss, when active UEs are identified using the EAE-UAI algorithm.



(a) Probability of false-alarm



(b) Probability of miss

Figure 3.16: Probabilities of false alarm and miss, when active UEs are identified using the SIC-UAIa algorithm in the case: (a) the number of active UEs is known to AP, (b) the number of active UEs is not known to AP, or (c) active UEs are partially known to AP.

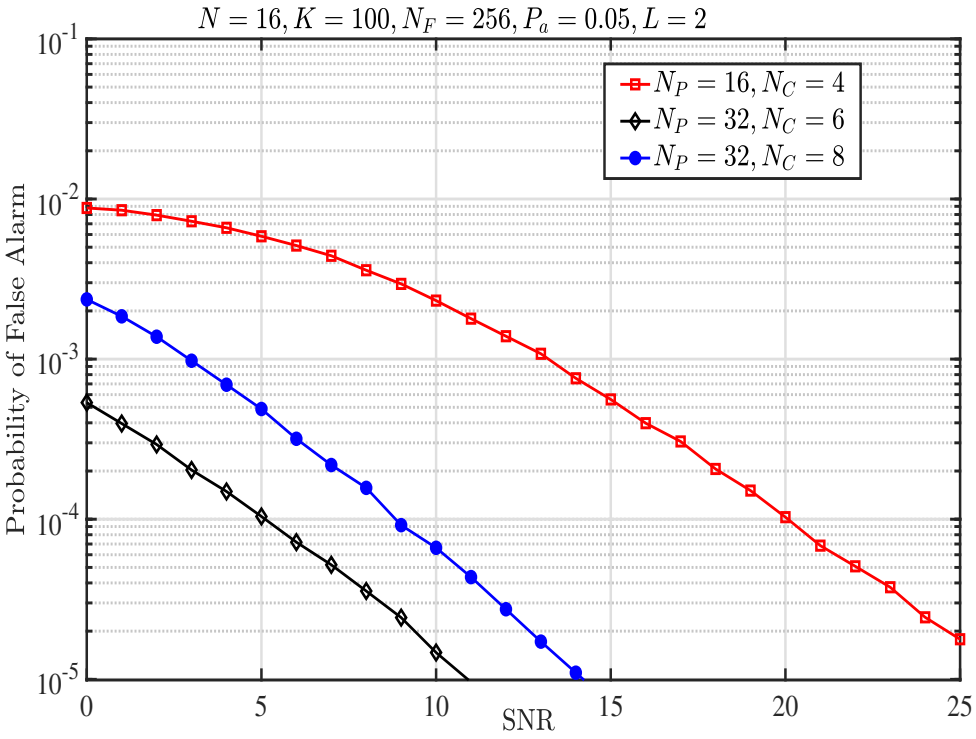
However, as seen in Fig. 3.16a for the false-alarm probability, the performance in the cases of (a) and (c) for $K = 100$ appears crossing. This is the result that the average number of active UEs in the case of (c) is low, and it also has the partial information for the active UEs (known who they are). By contrast, in the case of (a), AP only knows the number of UEs, but not know who they are. Finally, as seen in Fig.3.16, all the performance curves present error floor. This is because of channel estimation errors as well as the interference cancellation resulted from incorrect UAI.

In Figs.3.17a and 3.17b, we demonstrate the probabilities of false alarm and miss, respectively, generated by the SIC-UAIb algorithm, when AP has no knowledge about both the active UEs and the number of active UEs. It is shown that when given the ratio of N_C/N_P, transmitting more pilot symbols improve the performance of UAI. By contrast, given the number of pilot symbols to $N_P = 32$, if N_C is reduced from 8 to 6, the miss probability only slightly increases, while the false-alarm probability reduces significantly. Hence, the performance trade-off (between false-alarm and miss) problem of the SIC-UAIb algorithm is much less severer than that of the TB-UAI algorithm.

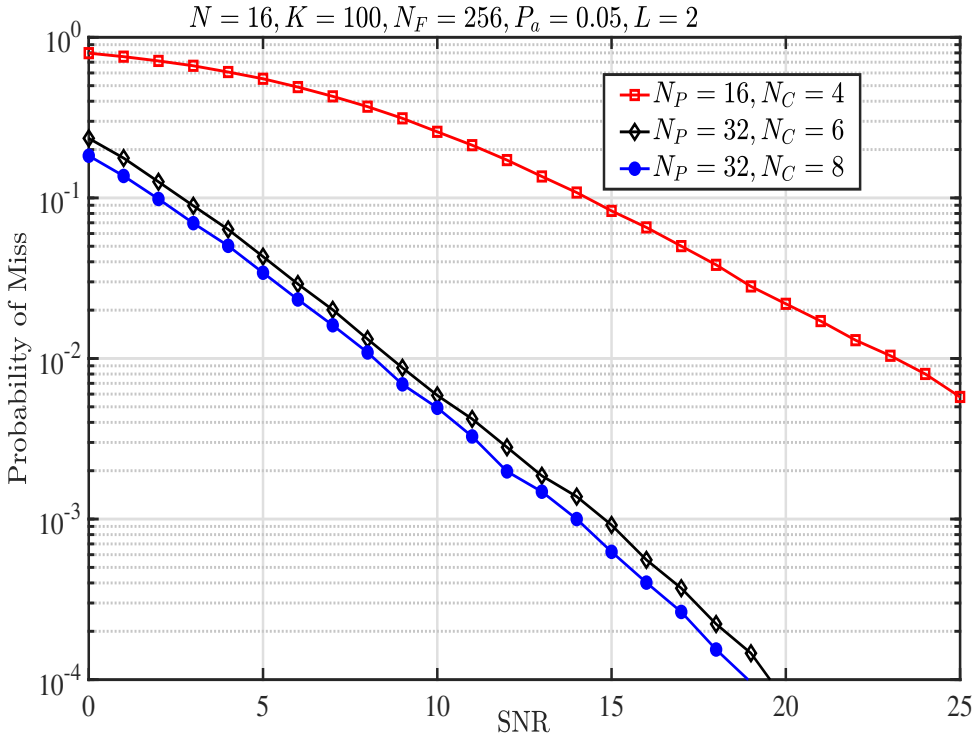
Finally, Figs.3.18a and 3.18b depict the false-alarm and miss probabilities of the MC-mGFMA systems employing the AME-UAI algorithm, when the parameters are as shown with the figure. Here we demonstrate the impact of the value of N_c on the false-alarm and miss probabilities. Explicitly, as N_c increases, the false-alarm probability increases, while the miss probability decreases. From the results we can be implied that $N_c = N_P/4$ is a reasonable value for practical operation. As shown in the figure, when $N_c = 4 = N_P/4$, the false-alarm probability is about 10^{-3} at the SNR of 20 dB. Simultaneously, at this SNR, the miss probability is below 10^{-3}. Again, the false-alarm probability is more sensitive to the value of N_C than the miss probability. However, the performance trade-off problem is severer than that of the SIC-UAIb algorithm. In contrast to the TB-UAI algorithm, the desired value of N_C in the AME-UAI algorithm can be relatively easy to find, which can be obtained such as based on numerical simulations.

3.8 Chapter Conclusions

We have investigated the joint channel estimation and UAI in MC-mGFMA systems. First, the channel estimation is addressed by assuming that AP has the full knowledge, no knowledge or the partial knowledge about active UEs. The studies allow us to reveal the fact that any added knowledge about active UEs can be exploited for enhancing the channel estimation and furthermore, for the design of novel UAI algorithms. Then, the statistics of the estimated channels of a UE is studied on the condition that the UE is active or inactive.

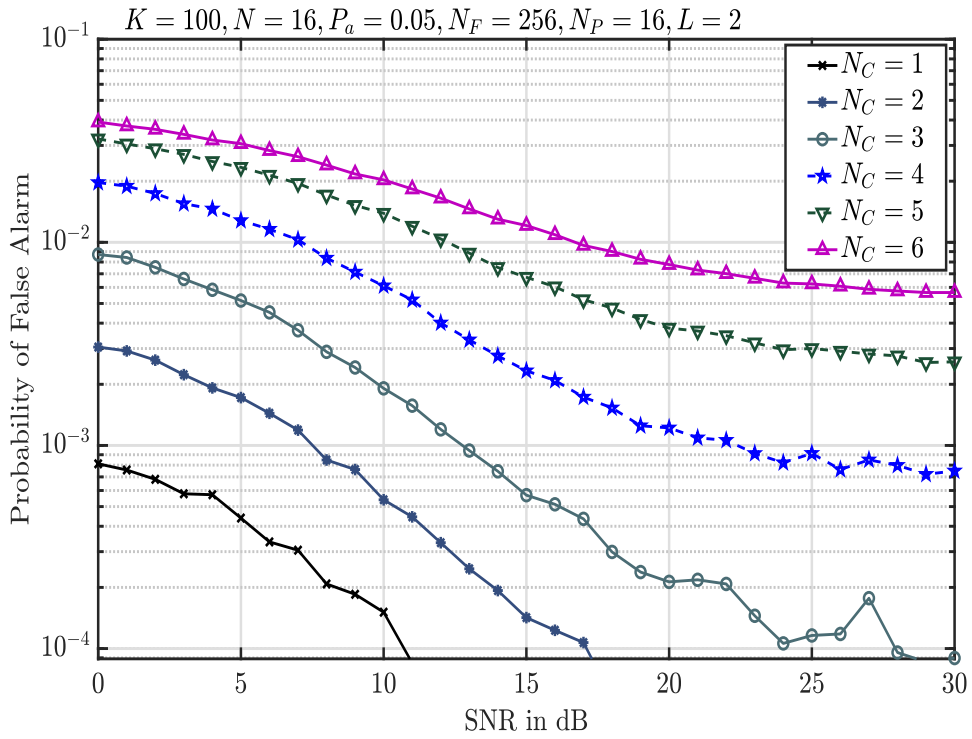


(a) Probability of false-alarm

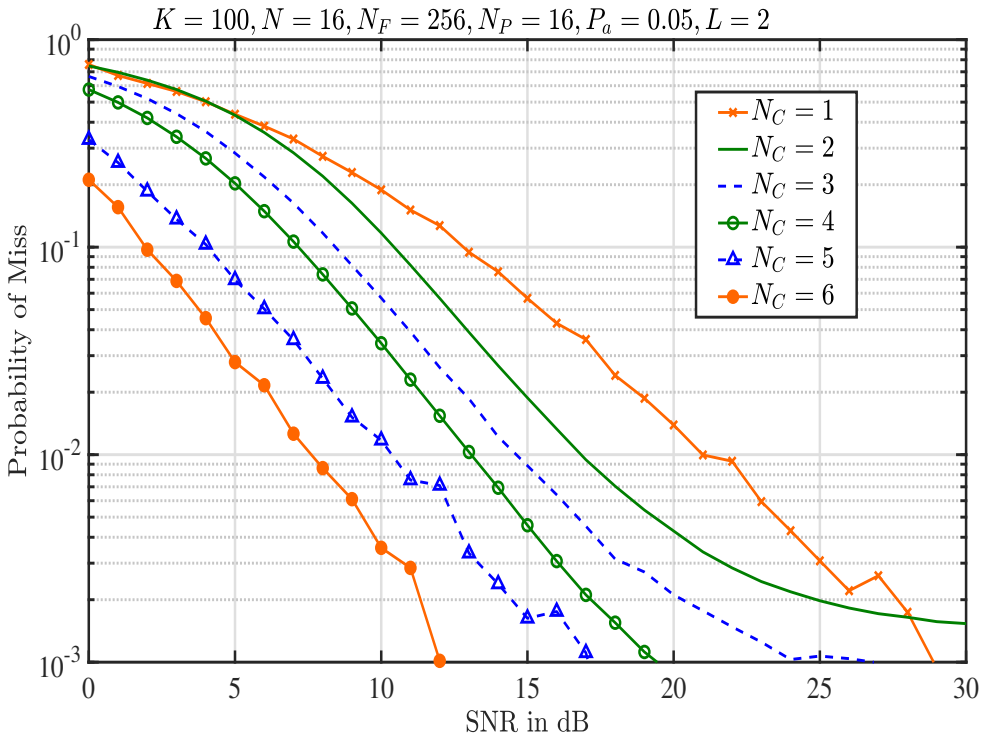


(b) Probability of miss

Figure 3.17: Probabilities of false alarm and miss, when UEs are identified using the SIC-UAIb algorithm.



(a) Probability of false-alarm



(b) Probability of miss

Figure 3.18: Probabilities of false alarm and miss, when UEs are identified using the AME-UAI algorithm.

It is shown that the estimated channels of active UEs and inactive UEs present distinctive features, explaining that the estimated channels of UEs can be utilized for UAI. Finally, based on the studies of channel estimation and their statistics, five UAI algorithms have been proposed and their performance has been studied. Our studies show that while TB-UAI has the lowest complexity, it yields a sharp trade-off between false-alarm and miss probabilities, making it hard to be implemented in the practical high-dynamic mGFMA environments. By contrast, the other four UAI algorithms are capable of providing significant performance enhancement in comparison with TB-UAI. Furthermore, the EAE-UAI and SIC-UAIa algorithms impose no trade-off between false-alarm and miss probabilities and the SIC-UAIb algorithm only makes little trade-off between false-alarm and miss probabilities. Although the AME-UAI algorithm brings in some trade-off between false-alarm and miss probabilities, the related parameter N_c can be relatively easily set in mGFMA systems.

In comparison with existing approaches, the UAI algorithms proposed in this chapter do not experience the RIP constraint of CS and do not depend on the factor graphs required by the MPA-based approaches. Instead, the signatures considered are random sequences and the number of active UEs can be highly dynamic, making the actual number of active UEs sometimes be significantly higher than the number of resource units of the system. Therefore, it can be expected that our proposed approaches are robust for operation in mGFMA systems.

We will address the joint channel estimation, UAI and the information detection of active UEs, in order to further enhance the performance of mGFMA systems.

Chapter 4

Joint Channel Estimation, User Activity Identification and Information Detection in Multi-Antenna mGFMA Systems

In this chapter, we investigate a multicarrier massive grant-free multiple-access (MC-mGFMA) system by assuming that a big number of highly dynamic user equipments (UEs) are monitored by an access point (AP) with multiple receive antennas (MRA), which is referred to as the MRA/MC-mGFMA system. The channel estimation, UE activity identification (UAI) and information detection are addressed. First, the channels of both active and inactive UEs are detected in the principle of minimum mean-square error (MMSE). Then, based on the estimated channels, a low-complexity threshold-based UAI (TB-UAI) is proposed to detect the activities of UEs. Finally, information of active UEs is detected in the principle of the successive interference cancellation (SIC) assisted MMSE (SIC-MMSE). Furthermore, a joint algorithm, referred to as SIC-MMSE-JCUD, is proposed to carry out channel estimation, UAI and information detection jointly in the principle of SIC-MMSE. Additionally, considering that no set of the well-designed signature sequences is enough to support the big number of potential UEs in mGFMA system, we propose a class of sequences designed by combining the Gold-sequences with the Zadoff-Chu (ZC) sequences. The performance of different schemes is studied and compared based on Monte-Carlo simulations. Our studies show that deploying multiple receive antennas at AP is beneficial to the channel estimation, UAI and information detection. Aided by the multiple receive antennas of AP, a low-complexity TB-UAI algorithm is highly efficient for UE activity identification. Furthermore, our proposed class of signature sequences allows to attain much better performance than the random sequences.

4.1 Introduction

One of the important applications of mGFMA is for uplink access in URLLC. Numerous researches have been carried out on this topic, especially, when considering the nowadays practical wireless systems, where an AP (or a BS is usually deployed with multiple antennas for signal receiving [144, 169]. Conventionally, when an uplink access request is initialized without AP's grant, the ALOHA style access scheme is usually implemented [197] [198] [199]. This kind of access is suitable for short-burst transmission and has the advantages of shortening the access latency, reducing overhead and resource consumption, etc. This principle can also applied in the mGFMA IoT systems, where massive UEs are monitored by the APs having multiple antennas. However, in a mGFMA system [200], there are potentially a huge number of UEs that are ultra-densely deployed and transmit their data to APs in the principles of non-orthogonal multiple-access (NOMA) supported by unique spreading sequences [63]. In mGFMA systems, each of the potential UEs has only a small probability to become active for data transmission during a time-slot. Hence, mGFMA systems are highly dynamic in terms of the active UEs as well as the number of them during a time. On the AP side, no *a-priori* knowledge about the number of active UEs and who are the active UEs is available. Hence, an AP has to firstly identify the active UEs and then detect their data, or do these in a joint way.

Recently, there have been a lot of research on mGFMA. Most of the researches based on MC-CDMA considered the user activity identification (UAI) and data detection. Specifically, in terms of the URLLC-demanded IoT networks [201–203], UAI and data detection were carried out jointly with channel estimation. As above-mentioned, in mGFMA systems, each UE has only a very small probability to become active for data transmission within a time-slot. Hence, UAI and data detection in mGFMA systems is a typical compressive sensing (CS) problem, for which a lot of researches have been carried out, see for example [140, 179, 204]. However, information recovery in the CS-represented problems experience the limitation of restricted isometry property (RIP) [205]. As the result of the RIP, the number of active UEs should be significantly lower than the number of resources units (determined by the number of antennas, time-slots, spreading factor, etc.), in order to achieve the recovery performance of practically meaningful. Unfortunately, in mGFMA systems, the number of active UEs at a time is highly dynamic, making it hard to meet the constraint of RIP. Otherwise, the number of potential UEs supportable should be small, which results in the low-efficiency of resource usage, and is not suitable for the mGFMA scenarios where UEs are supposed to be a big number and ultra-high density.

In the context of the signal detection in mGFMA systems, the massage passing algorithm (MPA) and approximate massage passing (AMP) have received intensive atten-

tion owing to their near-optimum performance. For example, in [147, 206], the authors introduced the AMP for signal detection in the sparse-code multiple access (SCMA) assisted GFMA systems, showing that improved bit-error-rate (BER) performance is achievable. It is well-known that mGFMA systems are in general operated in the principle of NOMA. Therefore, SCMA [207] constitutes one of the promising candidates that have received widely attention. Specifically, in [208], a SCMA system was suggested, where the MPA was operated with the dynamic factor graph for signal detection so as to reduce the computational complexity. In [209], a GFMA system was designed based on the rateless SCMA, where data of active UEs are detected using MPA under the assumption that the receiver has the knowledge of active UEs as well as their channel information. Furthermore, in [165], a faster-than-Nyquist (FTN) singling SCMA was proposed to support uplink GFMA, for which a new algorithm was developed for joint channel estimation, user activity detection and data detection. It can be shown that all the MPA-assisted detection methods considered in the above-mentioned references are capable of achieving near optimum performance, provided that good factor graphs are formed for signal detection [210, 211]. However, in a real mGFMA environment where exist a huge number of potential UEs and the active UEs as well as their number are highly dynamic, maintaining good factor graphs for signal detection is very challenging.

There were also many other methods proposed for supporting mGFMA systems. As some examples, In [97, 98], the authors investigated the multi-antenna MIMO assisted multi-cell systems, which work cooperatively to form an extended cloud-radio structure so as to help the edge nodes (EN) to access the network in GFMA principle. In [99], a novel preamble was designed so as to attain the higher success rate of 5G-based NGEO satellite communications in the MIMO-relied mGFMA systems. In [100, 101], a novel interleave-division multiple-access (IDMA) transceiver was proposed to support mGFMA communication, where the asynchronous characteristics of random signals were exploited for signal detection and channel estimation. Recently, with the advance of machine learning (ML) and intensive research related, a range of ML-based approaches have also been proposed for signal detection and user identification in mGFMA systems. Specifically, in [187], the authors introduced a novel block sparse Bayesian learning (BSBL) method for supporting massive machine type communications, where a vector learning algorithm was proposed for user identification and signal detection. Meanwhile, in [212,213], the authors considered the recovery of an active signal in a block by employing BSBL, which is operated in an adaptive way for UAI and signal detection. In [214, 215], the deep-learning combined with MPA was studied to detect signals in mGFMA networks. Additionally, in [216], the authors proposed the successive interference cancellation (SIC) method in assisting the signal detection in the mGFMA systems implemented by the so-called tandem

spreading multiple access (TSMA) technique.

Against the background, in [128], we have investigated the channel estimation and UAI in the multicarrier (MC) mGFMA systems, where a range of UAI or joint channel estimation and UAI algorithms are proposed. Our studies demonstrated that the proposed approaches are capable of circumventing some of the shortcomings of the techniques developed based on CS and MPA. They are suitable for operation in the mGFMA systems where both active UEs and the number of them are highly dynamic. However, in [128], first, we assumed that an AP (or BS) employs only single antenna, which is impractical as in future wireless systems APs are typically multi-antenna terminals. Second, in [128], we assumed the random spreading codes, due to the problem that in mGFMA systems the number of potential UEs is usually much large than the spreading factors, meaning that no set of well-designed codes, such as, m-sequences, Gold-sequences, Zadoff-Chu (ZC) sequence, etc., are enough for supporting the potential UEs. Furthermore, in [128], we only considered the channel estimation and UAI, but without considering the information detection. Therefore, in this chapter, we extend our studies in [128] to the more general scenario by addressing the above mentioned issues. To be more specific, the novelties and contributions can be summarized as follows:

- We introduce a multicarrier mGFMA system where each AP employs multiple receive antennas, which is referred to as the MRA/MC-mGFMA system for convenience of description.

- We extend our studies in [128] to the MRA/MC-mGFMA system, and investigate the impact of multi-antenna AP on the design and performance of channel estimation, UAI and information detection.

- Jointing the ZC sequence and gold-sequence, we design a class of signature sequences so as to support the possibly huge number of UEs in a mGFMA systems. We also investigate the effect of the proposed class of sequences on the performance of channel estimation, UAI and information detection.

- We propose the channel estimation, UAI, signal detection and joint channel estimation, UAI and information detection algorithms for the proposed MRA/MC-mGFMA system. Specifically, our channel estimation is in the principle of minimum mean-square error (MMSE) under the assumption that AP has full knowledge, partial knowledge or no knowledge about the active UEs. Then, assuming that AP has no knowledge about active UEs, we introduce the threshold-based UAI (TB-UAI) for UE activity identification. Finally, based on the channel estimation and/or TB-UAI, the successive interference cancellation (SIC) assisted MMSE (SIC-MMSE)

and the SIC-MMSE relied joint channel estimation, UAI and data detection (SIC-MMSE-JCUD) are designed.

- Based on numerical simulations, we study and compare the performance of the channel estimation, UAI and information detection algorithms, when various aspects are considered.

Our studies show that when multiple receive antennas at APs are available, the low-complexity TB-UAI is capable of achieving promising performance, which significantly improves with the increase of the number of receive antennas. Furthermore, as the number of receive antennas of AP increases, the threshold setting in TB-UAI becomes relatively easier. This implies that the performance of TB-UAI becomes less sensitive to the threshold applied, and the near-optimum threshold can be relatively easily attached, as the number of receive antennas of AP increases. In terms of the signature sequences, the studies show that the proposed class of signature sequences outperforms the random sequences. Furthermore, for signal detection, both SIC-MMSE and SIC-MMSE-JCUD are highly effective for operation with the MRA/MC-mGFMA systems, where both active UEs and the number of them are highly dynamic.

The remainder of this chapter is outlined as follows. In Section 4.2, we describe the MRA/MC-mGFMA system model, define the main system parameters and provide the representation of signaling. Section 4.3 deals with the MMSE-assisted channel estimation and performance of channel estimation. UAI and performance of UAI are addressed in Section 4.4. Section 4.5 considers the SIC-MMSE detection after the principle of MMSE detection, while Section 4.6 states the joint channel, UAI and data detection algorithm. Section 4.7 dedicates to the performance results and discussion. Finally, the conclusions from research are summarized in Section 4.8.

4.2 Description of Massive Grant-Free Multiple-Access Systems with Multiple Receive Antennas

We consider a single-cell multicarrier (MC) mGFMA system equipped with multiple receive antennas (MRA), referred to as the MRA/MC-mGFMA for convenience of description. For each user equipment (UE), the signaling is in the principles of the frequency-domain spread orthogonal frequency-division multiple-access (OFDMA) [127]. The MRA/MC-mGFMA system has one cell with one access point (AP) (or base-station (BS)) located at the center of the cell, and there are N subcarriers for frequency-domain spreading. Since most mGFMA systems are motivated for supporting low-rate IoT services, we as-

sume for simplicity the binary phase shift keying (BPSK) modulation, although it can be readily extended to other quadrature amplitude modulations (QAM) schemes. In the MRA/MC-mGFMA system, the signature code assigned to the kth UE is expressed as $\boldsymbol{c}_k = [c_k(0), c_k(1), \ldots, c_k(N-1)]^T$, where $c_k(n) \in \{+1/\sqrt{N}, -1/\sqrt{N}\}$, hence we have $\|\boldsymbol{c}_k\|^2 = 1$. The signature $\boldsymbol{c}_k$ of UE k serves for both the spreading of transmitted signal, and the identity (ID) of UE k for AP to identify its activity during the UE identification stage. If UE k is identified to be active, $\boldsymbol{c}_k$ is further used by AP to demodulate the signals received from UE k during the data detection stage. In our MRA/MC-mGFMA system, IFFT/FFT techniques are introduced for subcarrier modulation/demodulation. A sufficiently long cyclic prefix (CP) is introduced to avoid the inter-block interference between adjacent OFDM blocks. We assume that the MRA/MC-mGFMA system supports K potential UEs, which are mobile IoT devices randomly distributed in the cell. As the massive connectivity applications are considered, we have $K \gg N$. We assume that the system embeds a synchronization sub-system, which enables all UEs to synchronize with the AP's clock. Furthermore, we assume that a UE has a small probability P_a to become active during a time-slot for transmitting a frame. An active UE starts transmission at the beginning of a time-slot and one frame is transmitted per time-slot by each active UE. We express K_A the number of active UEs during a time-slot, which is a random variable having the average satisfying $\bar{K}_A \ll K$, as the result of a small activation probability of $P_a \ll 1$.

4.2.1 Transmitter Modeling

The transmitter in the MRA/MC-mGFMA system follows the typical framework of the frequency domain spread OFDMA. To transmit a binary bit $b_k \in \{+1, -1\}$ of UE k, it is first spread by $\boldsymbol{c}_k$, the products of which are transformed to the time-domain using IFFT. Then, after the parallel-to-serial (P/S) conversion, adding CP, transmitter (TX) filtering, and carrier modulation, the resultant signal is transmitted from the transmit antenna of UE k.

In our study, we assume block fading, i.e., the fading of a UE maintains constant over a frame duration, but is independent frame-by-frame. The frame has the structure as shown in Fig. 4.1. The frame length can be expressed as $N_F = N_I + N_P + N_L$, where N_I denotes the number of information bits per frame, N_P is the number of pilot bits per frame, inserted for channel estimation and UE identification. Furthermore, as shown in Fig. 4.1, at the beginning of each frame, N_L bits, referred to as the data block length indicator (DBLI), are transmitted to indicate the length in frames of a data block transmitted by a UE. The DBLI domain is used so that AP does not need to repeatedly identify the activities of the UEs having several frames to send. Note that in our studies, we typically set $N_L = 2$.

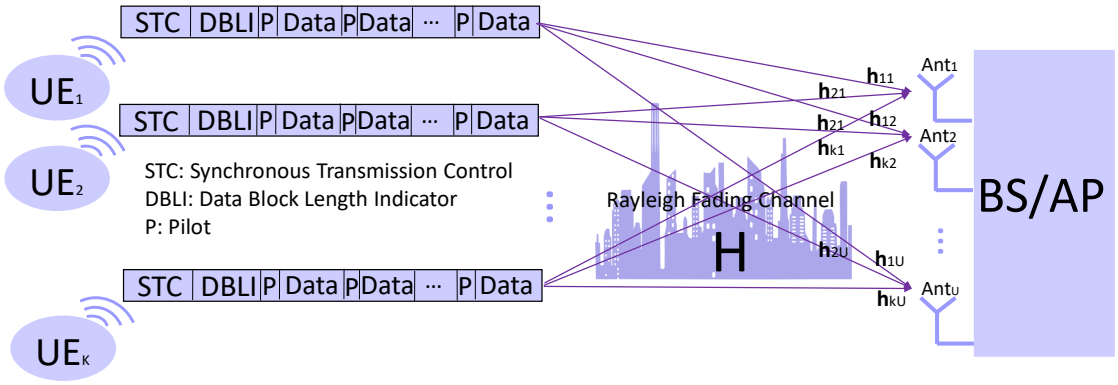


Figure 4.1: Illustration of frame structure and the system structure of MRA/MC-mGFMA systems.

This explains that once a UE becomes active, it can transmit upto 4 frames, but AP only needs to identify its activity once. This is implemented by setting the DBLI bits in the first frame to '11', which are reduced following the sequence of '10', '01', until '00' in the 4th frame to tell AP the end of current session. However, if the UE has more than 4 frames to transmit, it keeps transmitting in the next time-slot but it is treated as a new transmission session, and AP will re-identify the activity of this UE. As above mentioned, the benefit of employing a DBLI domain is that AP does not have to identify every active UE, but only those newly activated ones. Hence, the performance of UE's activity identification (UAI), channel estimation and data detection can be enhanced. Additionally, as shown in Fig. 4.1, there is a 'STC' domain in a time-slot. This is used to make the transmissions from different UEs be synchronized at AP.

4.2.2 Representation of Received Signal by Multiple Antennas

The receiver follows the receiving principles of OFDM. After removing CP, serial-to-parallel conversion and FFT operation, the received signals by the uth receive antenna can be expressed as

$$\begin{aligned}\boldsymbol{y}_m^{(u)} &= \sum_{k=1}^{K} \boldsymbol{C}_k \boldsymbol{\mathcal{F}} \boldsymbol{\Psi}_k \boldsymbol{h}_k^{(u)} b_{k,m} I_k + \boldsymbol{n}_m^{(u)}, \\ &= \sum_{k'=1}^{K_A} \boldsymbol{C}_{k'} \boldsymbol{\mathcal{F}} \boldsymbol{\Psi}_{k'} \boldsymbol{h}_{k'}(u) b_{k',m} + \boldsymbol{n}_m^{(u)}, \\ & m = 1, 2, \ldots, N_F; \\ & u = 1, 2, \ldots, U \end{aligned} \tag{4.1}$$

where U is the number of receive antennas, $I_k = 1$ or $I_k = 0$ indicates whether UE k is active or inactive, $b_{k,m} \in \{+1, -1\}$ is the mth bit in a frame sent from UE k, $\boldsymbol{y}_m^{(u)}$

and $\boldsymbol{n}_m^{(u)}$ are N-length observation and noise vectors, respectively, and $\boldsymbol{n}_m^{(u)}$ follows the complex Gaussian distribution with zero mean and covariance matrix $\sigma^2\boldsymbol{I}_N$, expressed as $\mathcal{CN}(0,\sigma^2\boldsymbol{I}_N)$. In (4.1), $\boldsymbol{C}_k = \text{diag}\{c_k^{(0)}, c_k^{(1)}, \cdots, c_k^{(N-1)}\}$, $\boldsymbol{h}_k^{(u)} = [h_{k,0}^{(u)}, \cdots, h_{k,L-1}^{(u)}]^T$ is the time-domain channel impulse response (CIR) from UE k to the uth receive antenna at AP, where $h_{k,l}^{(u)}$ obeys the distribution of $\mathcal{CN}(0,1/L)$. Correspondingly, the frequency-domain CIRs, i.e., the fading gains of N subcarriers are given by $\boldsymbol{\mathcal{F}}\boldsymbol{\Psi}_k\boldsymbol{h}_k^{(u)}$ [127], where $\boldsymbol{\mathcal{F}}$ is the $(N \times N)$ FFT matrix, and $\boldsymbol{\Psi}_k$ is the mapping matrix constituted by the first L columns of the identity matrix $\boldsymbol{I}_N$. Note that in (4.1), the second equality only considers the active UEs. Additionally, in (4.1), N_F is the number of bits sent in a frame. For simplicity, let in (4.1) $\boldsymbol{A}_k = \boldsymbol{C}_k\boldsymbol{\mathcal{F}}\boldsymbol{\Psi}_k$, which is a $(N \times L)$ matrix. Then, we have

$$\begin{aligned}\boldsymbol{y}_m^{(u)} &= \sum_{k=1}^{K} I_k\boldsymbol{A}_k\boldsymbol{h}_k^{(u)}b_{k,m} + \boldsymbol{n}_m^{(u)},\\ &= \sum_{k'=1}^{K_A} \boldsymbol{A}_{k'}\boldsymbol{h}_{k'}^{(u)}b_{k',m} + \boldsymbol{n}_m^{(u)},\\ &\quad m = 1,2,\ldots,N_F;\\ &\quad u = 1,2,\ldots,U\end{aligned} \tag{4.2}$$

Let $\boldsymbol{y}_m = [(\boldsymbol{y}_m^{(1)})^T, (\boldsymbol{y}_m^{(2)})^T, \cdots, (\boldsymbol{y}_m^{(U)})^T]^T$. Then, it can be shown that

$$\begin{aligned}\boldsymbol{y}_m &= \sum_{k=1}^{K} I_k(\boldsymbol{I}_U \otimes \boldsymbol{A}_k)\boldsymbol{h}_k b_{k,m} + \boldsymbol{n}_m\\ &= \sum_{k=1}^{K} I_k\boldsymbol{B}_k\boldsymbol{h}_k b_{k,m} + \boldsymbol{n}_m\\ &= \sum_{k'=1}^{K_A} \boldsymbol{B}_{k'}\boldsymbol{h}_{k'}b_{k',m} + \boldsymbol{n}_m, m = 1,2,\ldots,N_F\end{aligned} \tag{4.3}$$

In the above equation, we defined $\boldsymbol{B}_k = \boldsymbol{I}_U \otimes \boldsymbol{A}_k$, which is a $(NU \times UL)$ matrix, $\boldsymbol{h}_k = [(\boldsymbol{h}_k^{(1)})^T, (\boldsymbol{h}_k^{(2)})^T, \cdots, (\boldsymbol{h}_k^{(U)})^T]^T$ is a UL-length vector, containing the channels from UE k to the U receive antennas of AP, which is complex Gaussian distributed with zero mean and a covariance matrix $\boldsymbol{I}_{UL}/L$. $\boldsymbol{n}_m = [(\boldsymbol{n}_m^{(1)})^T, (\boldsymbol{n}_m^{(2)})^T, \cdots, (\boldsymbol{n}_m^{(U)})^T]^T$ is a UN-length Gaussian noise vector, following the distribution of $\mathcal{CN}(0,\sigma^2\boldsymbol{I}_{UN})$ with $\sigma^2 = 1/SNR$.

If we only consider the pilot symbols for the purpose of channel estimation, we can write (4.3) as

$$\begin{aligned}\boldsymbol{y}_p &= \sum_{k=1}^{K} I_k\boldsymbol{B}_k\boldsymbol{h}_k b_{k,p} + \boldsymbol{n}_p\\ &= \sum_{k'=1}^{K_A} \boldsymbol{B}_{k'}\boldsymbol{h}_{k'}b_{k',p} + \boldsymbol{n}_p,\ p = 1,2,\ldots,N_P\end{aligned} \tag{4.4}$$

Let $\boldsymbol{y} = [(\boldsymbol{y}_1)^T, (\boldsymbol{y}_2)^T, \cdots, (\boldsymbol{y}_{N_p})^T]^T$, Then we obtain

$$\begin{aligned}\boldsymbol{y} &= \sum_{k=1}^{K} I_k(\boldsymbol{p}_k \otimes \boldsymbol{B}_k)\boldsymbol{h}_k + \boldsymbol{n} \\ &= \sum_{k'=1}^{K_A} (\boldsymbol{p}_{k'} \otimes \boldsymbol{B}_{k'})\boldsymbol{h}_{k'} + \boldsymbol{n}\end{aligned} \tag{4.5}$$

where $\boldsymbol{p}_k = [b_{k,1}, b_{k,2}, \cdots, b_{k,N_P}]^T$, $\boldsymbol{n} = [(\boldsymbol{n}_1)^T, (\boldsymbol{n}_2)^T, \cdots, (\boldsymbol{n}_{N_p})^T]^T$ obeys the distribution of $\mathcal{CN}(0, \sigma^2 \boldsymbol{I}_{N_P UN})$ and $\boldsymbol{y}$ is a $(N_P UN \times 1)$ matrix.

Alternatively, we can represent the observations in the form as below. Given $\boldsymbol{y}_m^{(u)}$ as shown in (4.2), let $\boldsymbol{y}^{(u)} = [(\boldsymbol{y}_1^{(u)})^T, (\boldsymbol{y}_2^{(u)})^T, \cdots, (\boldsymbol{y}_{N_P}^{(u)})^T]^T$, $u = 1, 2, \ldots, U$, which is a $N_P N$-length vector. Then, it can be shown that

$$\begin{aligned}\boldsymbol{y}^{(u)} &= \sum_{k=1}^{K} I_k(\boldsymbol{p}_k \otimes \boldsymbol{A}_k)\boldsymbol{h}_k^{(u)} + \boldsymbol{n}^{(u)} \\ &= \sum_{k'=1}^{K_A} (\boldsymbol{p}_{k'} \otimes \boldsymbol{A}_{k'})\boldsymbol{h}_{k'}^{(u)} + \boldsymbol{n}^{(u)},\ u = 1, 2, \ldots, U\end{aligned} \tag{4.6}$$

where $\boldsymbol{n}_u$ is a $N_P N$-length Gaussian noise vector, following the distribution of $\mathcal{CN}(0, \sigma^2 \boldsymbol{I}_{N_P N})$. Furthermore, let $\boldsymbol{y} = [(\boldsymbol{y}^{(1)})^T, (\boldsymbol{y}^{(2)})^T, \cdots, (\boldsymbol{y}^{(U)})^T]^T$, $\boldsymbol{n} = [(\boldsymbol{n}^{(1)})^T, (\boldsymbol{n}^{(2)})^T, \cdots, (\boldsymbol{n}^{(U)})^T]^T$, which follows the Gaussian distribution of $\mathcal{CN}(0, \sigma^2 \boldsymbol{I}_{UN_P N})$, $\boldsymbol{h}_k = [(\boldsymbol{h}_k^{(1)})^T, (\boldsymbol{h}_k^{(2)})^T, \cdots, (\boldsymbol{h}_k^{(U)})^T]^T$, which is a UL-length vector obeying the distribution of $\mathcal{CN}(0, \boldsymbol{I}_{UL}/L)$. Then, we have the representation of

$$\begin{aligned}\boldsymbol{y} &= \sum_{k=1}^{K} I_k(\boldsymbol{I}_U \otimes (\boldsymbol{p}_k \otimes \boldsymbol{A}_k))\boldsymbol{h}_k + \boldsymbol{n} \\ &= \sum_{k'=1}^{K_A} (\boldsymbol{I}_U \otimes (\boldsymbol{p}_{k'} \otimes \boldsymbol{A}_{k'}))\boldsymbol{h}_{k'} + \boldsymbol{n}\end{aligned} \tag{4.7}$$

Above we have represented the received signals at AP in different forms, which will be used later for different purposes.

4.2.3 Open-loop Power Control

Since random access is usually implemented in mGFMA systems, closed-loop power-control [139] is impossible to implement. However, in mGFMA systems, the pilot signals sent by APs periodically for time synchronization may be used by UEs to estimate their distances from APs via measuring the pilot signal's strength. In this case, UEs are able to implement open-loop power-control [139] based on the strength of received pilot

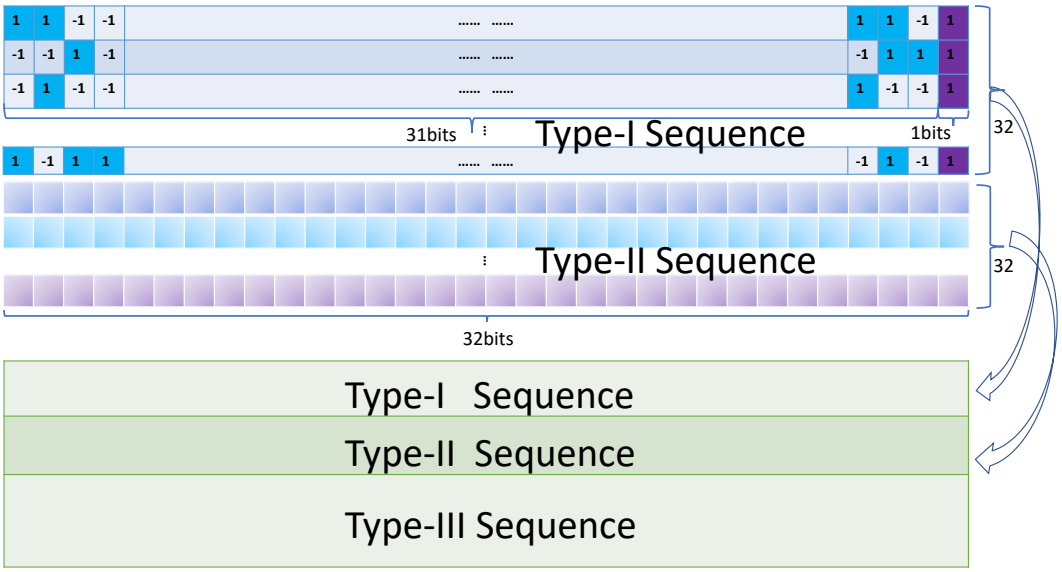


Figure 4.2: Signature sequences generated from Gold sequences and Zadoff-Chu sequences.

signals. With the above consideration, we assume that open-loop power-control can be implemented in MRA/MC-mGFMA systems. Correspondingly, we assume that the power received from UE k by one receive antenna of BS obeys a truncated Gaussian distribution represented as

$$f_{G_k}(g) = \mu \exp\left(-\frac{(g-G_0)^2}{2\sigma_G^2}\right), \quad \max\{0, G_0 - G_E\} \leq g \leq G_0 + G_E \tag{4.8}$$

where G_k denotes the power received from UE k, G_E denotes the maximum power-control error, and G_0 is the received power in ideal power-control, while μ is a constant to ensure that the integration of $f_{G_k}(g)$ over the range of $max\{0, G_0 - G_E\} \leq g \leq G_0 + G_E$ is 1.

Upon involving power control, (4.7) can now be modified to a formula of

$$\begin{aligned} \boldsymbol{y} &= \sum_{k=1}^{K} I_k \sqrt{G_k} (\boldsymbol{I}_U \otimes (\boldsymbol{p}_k \otimes \boldsymbol{A}_k)) \boldsymbol{h}_k + \boldsymbol{n} \\ &= \sum_{k'=1}^{K_A} \sqrt{G_{k'}} (\boldsymbol{I}_U \otimes (\boldsymbol{p}_{k'} \otimes \boldsymbol{A}_{k'})) \boldsymbol{h}_{k'} + \boldsymbol{n} \end{aligned} \tag{4.9}$$

Furthermore, since (4.7) is a normalized form, we have $G_0 = 1$.

4.2.4 Signature Sequences

In mGFMA, the signature sequences required can be significantly more than N of the number of subcarriers. In our previous study [128], random spreading sequences are as-

sumed. However, in practical systems, pure random sequences are rarely used. Therefore, in the MRA/MC-mGFMA system considered in this chapter, we propose to generate enough signature sequences from the products of Gold sequences [217] and Zadoff-Chu sequences [218]. Fig. 4.2 shows an example to constructed the new class of sequences.

As shown in Fig. 4.2, we construct three types of sequences, namely, the Type-I, Type-II and Type-III sequences with each type consisting of the sequences of length 32. Specifically, Type-I sequences are constructed from the 31-length Gold sequences by randomly appending a binary symbol, forming the 32-length sequences. By contrast, Type-II sequences are obtained by removing some heading and trailing symbols from 37-length Zadoff-Chu sequences. After the above processing, both Type-I and Type-II sequences are 32-length, which are convenient for product operations. Based on the Type-I and Type-II sequences, the Type-III class of sequences are constructed by the bitwise multiplication of Type-I sequences and Type-II sequences. In total, we can form $32 \times 32 + 64 = 1086$ sequences, where 64 sequences are contributed by the 32 Type-I and 32 Type-II sequences.

4.3 Channel Estimation

In this section, we investigate the channel estimation in MRA/MC-mGFMA systems, all based on the minimum mean-square error (MMSE) principles. Let us first consider the channel estimation when AP is assumed to know the active UEs, from which we gain the insight for designing the other channel estimators.

4.3.1 Channel Estimation with Active UEs Known to AP - Estimator-K

When AP knows which are active UEs, the channels of these active UEs can be estimated as follows. With the observations as shown in (4.7) prepared, the channels from UE i to the U receive antennas of AP can be expressed as

$$\hat{\boldsymbol{h}}_i = [(\hat{\boldsymbol{h}}_i^{(1)})^T, (\hat{\boldsymbol{h}}_i^{(2)})^T, \cdots, (\hat{\boldsymbol{h}}_i^{(U)})^T]^T = \boldsymbol{W}_i^H \boldsymbol{y} \tag{4.10}$$

where the weight matrix $\boldsymbol{W}_i$ obtained via MMSE optimization can be expressed as

$$\boldsymbol{W}_i = \boldsymbol{R}_y^{-1} \boldsymbol{R}_{y,h_i} \tag{4.11}$$

In (4.11), $\boldsymbol{R}_y$ is the autocorrelation matrix of $\boldsymbol{y}$ in (4.7), given by

$$\begin{aligned}\boldsymbol{R}_y &= E\left[\boldsymbol{y}\boldsymbol{y}^H\right] \\ &= \sum_{k'=1}^{K_A} \frac{1}{L}\left(\boldsymbol{I}_U \otimes (\boldsymbol{p}_{k'}\boldsymbol{p}_{k'}^H \otimes \boldsymbol{A}_{k'}\boldsymbol{A}_{k'}^H)\right) + \sigma^2 \boldsymbol{I}_{UN_PN},\end{aligned} \tag{4.12}$$

$\boldsymbol{R}_{y,h_i}$ is the cross-correlation matrix between $\boldsymbol{y}$ of (4.7) and $\boldsymbol{h}_i$, which can be derived as

$$\begin{aligned}\boldsymbol{R}_{y,h_i} &= E\left[\boldsymbol{y}\boldsymbol{h}_i^H\right] \\ &= \frac{(\boldsymbol{I}_U \otimes (\boldsymbol{p}_i \otimes \boldsymbol{A}_i))}{L}\end{aligned} \tag{4.13}$$

Explicitly, $\boldsymbol{R}_y$ of (4.12) can be expressed as

$$\begin{aligned}\boldsymbol{R}_y &= \left(\boldsymbol{I}_U \otimes \sum_{k'=1}^{K_A} \frac{1}{L}(\boldsymbol{p}_{k'}\boldsymbol{p}_{k'}^H \otimes \boldsymbol{A}_{k'}\boldsymbol{A}_{k'}^H)\right) + \sigma^2 \boldsymbol{I}_{UN_PN} \\ &= \left(\boldsymbol{I}_U \otimes \left(\sum_{k'=1}^{K_A} \frac{1}{L}(\boldsymbol{p}_{k'}\boldsymbol{p}_{k'}^H \otimes \boldsymbol{A}_{k'}\boldsymbol{A}_{k'}^H) + \sigma^2 \boldsymbol{I}_{N_PN}\right)\right)\end{aligned} \tag{4.14}$$

Substituting (4.14) and (4.13) into (4.11), we obtain

$$\begin{aligned}\boldsymbol{W}_i &= \left(\boldsymbol{I}_U \otimes \left(\sum_{k'=1}^{K_A} \frac{1}{L}(\boldsymbol{p}_{k'}\boldsymbol{p}_{k'}^H \otimes \boldsymbol{A}_{k'}\boldsymbol{A}_{k'}^H) + \sigma^2 \boldsymbol{I}_{N_PN}\right)\right)^{-1} \\ &\quad \times \left(\frac{(\boldsymbol{I}_U \otimes (\boldsymbol{p}_i \otimes \boldsymbol{A}_i))}{L}\right) \\ &= \left(\boldsymbol{I}_U \otimes \boldsymbol{R}_{y^{(u)}}^{-1} \boldsymbol{R}_{y^{(u)},h_i^{(u)}}\right) \\ &= \left(\boldsymbol{I}_U \otimes \boldsymbol{W}_i^{(u)}\right)\end{aligned} \tag{4.15}$$

where $\boldsymbol{W}_i^{(u)} = \boldsymbol{R}_{y^{(u)}}^{-1} \boldsymbol{R}_{y^{(u)}h_i^{(u)}}$ is the weight matrix to process the received signals from the uth receive antenna, which can be expressed as

$$\begin{aligned}\boldsymbol{W}_i^{(u)} =& \boldsymbol{R}_{y^{(u)}}^{-1} \boldsymbol{R}_{y^{(u)}h_i} \\ &= \left(\sum_{k'=1}^{K_A} (\boldsymbol{p}_{k'}\boldsymbol{p}_{k'}^H \otimes \boldsymbol{A}_{k'}\boldsymbol{A}_{k'}^H) + L\sigma^2 \boldsymbol{I}_{N_PN}\right)^{-1} \\ &\quad \times (\boldsymbol{p}_i \otimes \boldsymbol{A}_i), \\ &\quad i = 1, 2, \cdots, K,\ u = 1, 2, \cdots, U\end{aligned} \tag{4.16}$$

From (4.16) we can observe that $\boldsymbol{W}_i^{(u)}$ is in fact independent of u, i.e, of the index of receive antennas. In other words, the same weight matrix can be used for processing the received signals of all receive antennas. This has double-fold benefit. First, it is straightforward that this can reduce the complexity for computing the weight matrices in terms of

different receive antennas. Second, this allows to estimate a more accurate weight matrix by applying the received (pilot) signals by all receive antennas, instead of that by one.

Upon substituting (4.15) and (4.6) into (4.10), we can obtain

$$\begin{aligned}\hat{\boldsymbol{h}}_i &= \boldsymbol{W}_i^H \boldsymbol{y} \\ &= \left(\boldsymbol{I}_U \otimes \boldsymbol{W}_i^{(u)}\right)^H [(\boldsymbol{y}^{(1)})^T, (\boldsymbol{y}^{(2)})^T, \cdots, (\boldsymbol{y}^{(U)})^T]^T \end{aligned} \tag{4.17}$$

from which we can decompose to obtain

$$\hat{\boldsymbol{h}}_i^{(u)} = \left(\boldsymbol{W}_i^{(u)}\right)^H \boldsymbol{y}^{(u)},\ u = 1, 2, \cdots, U \tag{4.18}$$

The above analysis show that the channels of individual receive antennas can be estimated separately in parallel without performance loss. However, as previously discussed, $\boldsymbol{W}_i^{(u)}$ is common to all receive antennas, which can be more accurately estimated by exploiting the received signals by all the antennas of AP. Hence, the performance of channel estimation can be enhanced, when AP has more receive antennas.

Since in this subsection we assume that AP has the knowledge about the active UEs, AP is capable of constructing $\boldsymbol{W}_i^{(u)}$ of (4.16), provided that it knows the noise variance. Alternatively, if the number of received antennas at AP is big, $\boldsymbol{W}_i^{(u)}$ can also be obtained by first estimating $\boldsymbol{R}_{y^{(u)}}$ as

$$\boldsymbol{R}_{y^{(u)}} = \frac{1}{U} \sum_{u=1}^{U} \boldsymbol{y}^{(u)} (\boldsymbol{y}^{(u)})^H \tag{4.19}$$

Explicitly, when U becomes larger, a more accurate $\boldsymbol{R}_{y^{(u)}}$ can be obtained. The other term $\boldsymbol{p}_i \otimes \boldsymbol{A}_i$ for $\boldsymbol{W}_i^{(u)}$ as seen in (4.16) can be directly constructed by AP.

Following the principles of MMSE, it can be shown that the total MSE of the MMSE estimator for an active UE is given by

$$\begin{aligned} M_{SE}(A) &= \frac{1}{U} E\left[\|\boldsymbol{h}_i - \boldsymbol{W}_i^H \boldsymbol{y}\|^2\right] \\ &= \frac{1}{U} \mathrm{Tr}\left(\boldsymbol{I}_{UL}/L - \boldsymbol{R}_{y_{h_i}}^H \boldsymbol{W}_i\right) \\ &= 1 - \frac{1}{U} \mathrm{Tr}\left(\boldsymbol{R}_{y,i}^H \boldsymbol{W}_i\right) \end{aligned} \tag{4.20}$$

per receive antenna. By contrast, the MSE of the channel estimation for an inactive UE normalized by the number of receive antennas is

$$\begin{aligned} M_{SE}(\bar{A}) &= \frac{1}{U} E\left[\| - \boldsymbol{W}_i^H \boldsymbol{y}\|^2\right] \\ &= \frac{1}{U} \mathrm{Tr}\left(\boldsymbol{W}_i^H \boldsymbol{R}_y \boldsymbol{W}_i\right) \end{aligned} \tag{4.21}$$

Above the channel estimation is carried out by assuming that AP has the knowledge of active UEs. However, in mGFMA communications, this is hard to achieve as UE's access using random access. Therefore, in the next subsection, we consider the channel estimation when assuming that AP has no knowledge of the active UEs.

4.3.2 Channel Estimation with Active UEs Unknown to AP - Estimator-uK

When AP has no knowledge about the active UEs, it has to estimate the weight matrices so as to estimate the channels of active/inactive UEs. It can be shown that the channels of a UE can be either estimated based on (4.2) in terms of individual receive antennas or based on (4.3) to jointly estimate all the channels of U receive antennas in one formula. However, we should note that when the signals received by different receive antennas are uncorrelated, both estimation methods are equivalent. By contrast, if receive antennas have correlation, the joint estimation should outperform the individual estimation. However, in our studies, antennas are assumed to be uncorrelated.

Let us first consider the channel estimation at the individual antenna level. To proceed, the MMSE-based channel estimator for UE i can be derived from (4.2), which gives

$$\hat{\boldsymbol{h}}_i^{(u)} = \frac{1}{N_p}\sum_{p=1}^{N_p} \boldsymbol{C}_{i,p}^{(u)} \left(\boldsymbol{W}_{i,p}^{(u)}\right)^H \boldsymbol{y}_p^{(u)},\ u = 1,2,\ldots,U \\ i = 1,2,\ldots,K \tag{4.22}$$

where the index p emphasizes 'pilot', $\boldsymbol{C}_{i,p}^{(u)}$ is for achieving unbiased estimation, and $\boldsymbol{W}_{i,p}^{(u)}$ can be expressed as

$$\boldsymbol{W}_{i,p}^{(u)} = \boldsymbol{R}_{\boldsymbol{y}_p^{(u)}}^{-1} \boldsymbol{R}_{\boldsymbol{y}_p^{(u)}\boldsymbol{h}_i^{(u)}} \tag{4.23}$$

In (4.23), $\boldsymbol{R}_{\boldsymbol{y}_p^{(u)}}$ is the autocorrelation matrix of $\boldsymbol{y}_p^{(u)}$. Since AP does not know which are active UEs, it has to estimate $\boldsymbol{R}_{\boldsymbol{y}_p^{(u)}}$. To achieve this, let us first derive the theoretical expression of $\boldsymbol{R}_{\boldsymbol{y}_p^{(u)}}$, which can be expressed as

$$\begin{aligned} \boldsymbol{R}_{\boldsymbol{y}_p^{(u)}} =& E\left[\boldsymbol{y}_p^{(u)} \left(\boldsymbol{y}_p^{(u)}\right)^H\right] \\ =& \frac{1}{L}\sum_{k'=1}^{K_A} \boldsymbol{A}_{k'} \left(\boldsymbol{A}_{k'}\right)^H + \sigma^2 \boldsymbol{I}_N \end{aligned} \tag{4.24}$$

Equation (4.24) shows that $\boldsymbol{R}_{\boldsymbol{y}_p^{(u)}}$ is independent of the symbol index and also the antenna index. This means that we can use all the signals received by different antennas of AP for

estimation of $\boldsymbol{R}_{\boldsymbol{y}_p^{(u)}}$, i.e., we can estimate it as

$$\begin{aligned}\hat{\boldsymbol{R}}_{\boldsymbol{y}_p^{(u)}} = \boldsymbol{R}_a =& \frac{1}{N_p}\sum_{p=1}^{N_p}\boldsymbol{y}_p^{(u)}\left(\boldsymbol{y}_p^{(u)}\right)^H \\ =& \frac{1}{N_F}\sum_{m=1}^{N_F}\boldsymbol{y}_m^{(u)}\left(\boldsymbol{y}_m^{(u)}\right)^H \\ =& \frac{1}{UN_F}\sum_{u=1}^{U}\sum_{m=1}^{N_F}\boldsymbol{y}_m^{(u)}\left(\boldsymbol{y}_m^{(u)}\right)^H \end{aligned} \tag{4.25}$$

where the first, second and third equalities mean that $\boldsymbol{R}_{\boldsymbol{y}_p^{(u)}}$ is estimated using only the pilot symbols received by a specific antenna, all the symbols received by a specific antenna and all the symbols received by all the receive antennas of AP, respectively. Certainly, exploiting more received signals gives more accurate estimation.

In (4.23), $\boldsymbol{R}_{\boldsymbol{y}_p^{(u)}\boldsymbol{h}_i^{(u)}}$ in theory is

$$\begin{aligned}\boldsymbol{R}_{\boldsymbol{y}_p^{(u)}\boldsymbol{h}_i^{(u)}} =& E\left[\boldsymbol{y}_p^{(u)}\left(\boldsymbol{h}_i^{(u)}\right)^H\right] \\ =& \boldsymbol{A}_i b_{i,p}/L \end{aligned} \tag{4.26}$$

implying that AP can construct it, as AP knows the pilots and signatures of a UE to be estimated.

Finally, it can be shown that $\boldsymbol{C}_{i,p}^{(u)}$ in (4.22) is given by $\boldsymbol{C}_{i,p}^{(u)} = L\left(\boldsymbol{A}_i^H\hat{\boldsymbol{R}}_a^{-1}\boldsymbol{A}_i\right)^{-1}$.

Upon substituting the above derived terms into (4.23) and then into (4.22), we can obtain the estimate of

$$\begin{aligned}\hat{\boldsymbol{h}}_i^{(u)} = \left(\boldsymbol{A}_i^H\hat{\boldsymbol{R}}_a^{-1}\boldsymbol{A}_i\right)^{-1}\boldsymbol{A}_i^H\hat{\boldsymbol{R}}_a^{-1}\times\frac{1}{N_p}\sum_{p=1}^{N_p}b_{i,p}\boldsymbol{y}_p^{(u)}, \\ u = 1,2,\ldots,U;\ i = 1,2,\ldots,K \end{aligned} \tag{4.27}$$

From Eq. 4.27 we can know that for $\hat{\boldsymbol{h}}_i^{(u)}$, we first average the observations corresponding to the pilots by invoking the associated pilot bits. Then, the averaged outputs are processed in MMSE principle to give the estimates.

Additionally, it can be shown that the MSE of the estimation for an active UE is

$$MSE(A) = 1 - \frac{1}{L^2}\mathrm{Tr}\left(\boldsymbol{A}_i^H\left[\hat{\boldsymbol{R}}_{\boldsymbol{y}_p^{(u)}}\right]^{-1}\boldsymbol{A}_i\right) \tag{4.28}$$

per receive antenna. The MSE of the estimation for an inactive UE is

$$MSE(\bar{A}) = \frac{1}{L^2}\mathrm{Tr}\left(\boldsymbol{A}_i^H\left[\hat{\boldsymbol{R}}_{\boldsymbol{y}_p^{(u)}}\right]^{-1}\boldsymbol{A}_i\right) \tag{4.29}$$

per receive antenna.

Similarly, when the channels of all U receive antennas are jointly estimated, based on (4.3), we have the expression for estimation as

$$\hat{\boldsymbol{h}}_i = \frac{1}{N_p} \sum_{p=1}^{N_p} \boldsymbol{C}_{i,p} \boldsymbol{W}_{i,p}^H \boldsymbol{y}_p,\ i = 1, 2, \ldots, K \tag{4.30}$$

where $\boldsymbol{y}_p$ is given by (4.3) for pilots, and $\boldsymbol{W}_{i,p}$ is

$$\boldsymbol{W}_{i,p} = \boldsymbol{R}_{\boldsymbol{y}_p}^{-1} \boldsymbol{R}_{\boldsymbol{y}_p \boldsymbol{h}_i} \tag{4.31}$$

In (4.31), $\boldsymbol{R}_{\boldsymbol{y}_p}$ can be estimated as

$$\begin{aligned} \hat{\boldsymbol{R}}_{\boldsymbol{y}_p} = \hat{\boldsymbol{R}}_b =& \frac{1}{N_p} \sum_{p=1}^{N_p} \boldsymbol{y}_p \boldsymbol{y}_p^H \\ =& \frac{1}{N_F} \sum_{m=1}^{N_F} \boldsymbol{y}_m \boldsymbol{y}_m^H \end{aligned} \tag{4.32}$$

which is independent of the p index. $\boldsymbol{R}_{\boldsymbol{y}_p \boldsymbol{h}_i}$ is given by

$$\boldsymbol{R}_{\boldsymbol{y}_p \boldsymbol{h}_i} = \boldsymbol{B}_i b_{i,p} / L \tag{4.33}$$

Furthermore, it can be shown that $\boldsymbol{C}_{i,p} = L \left(\boldsymbol{B}_i^H \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{B}_i \right)^{-1}$. Therefore, we have

$$\boldsymbol{W}_{i,p} = \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{B}_i b_{i,p} / L \tag{4.34}$$

Substituting it into (4.30), we obtain

$$\begin{aligned} \hat{\boldsymbol{h}}_i =& \left(\boldsymbol{B}_i^H \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{B}_i \right)^{-1} \boldsymbol{B}_i^H \hat{\boldsymbol{R}}_b^{-1} \times \frac{1}{N_p} \sum_{p=1}^{N_p} b_{i,p} \boldsymbol{y}_p, \\ & i = 1, 2, \ldots, K \end{aligned} \tag{4.35}$$

4.3.3 Channel Estimation with Active UEs Partially Known to AP - Estimator-pK

In the realistic MRA/MC-mGFMA systems, AP may know a part of the active UEs in the system, as some active UEs may have several frames to transmit, as indicated by the DBLI in Fig.4.1. In this case, channels may be estimated in the joint way as in section 4.3.1, but with a partially estimated auto-correlative matrix. Specifically, when only one receive antenna is considered, we can construct $\boldsymbol{R}_{y^{(u)}}$ as (4.36) with $\hat{\boldsymbol{R}}_{\boldsymbol{y}_p^{(u)}}$ obtained by (4.25), as shown on the top of the next page, where $\bar{\mathcal{K}}_A$ is the set of active UEs known to AP.

$$\boldsymbol{R}_y^{(u)} = \begin{pmatrix} \hat{\boldsymbol{R}}_{\boldsymbol{y}_p^{(u)}} & \sum\limits_{k\in\bar{\mathcal{K}}_A} b_k^{(1)} b_k^{(2)} \dfrac{\boldsymbol{A}_k \boldsymbol{A}_k^H}{L} & \cdots & \sum\limits_{k\in\bar{\mathcal{K}}_A} b_k^{(1)} b_k^{(N_P)} \dfrac{\boldsymbol{A}_k \boldsymbol{A}_k^H}{L} \\ \sum\limits_{k\in\bar{\mathcal{K}}_A} b_k^{(2)} b_k^{(1)} \dfrac{\boldsymbol{A}_k \boldsymbol{A}_k^H}{L} & \hat{\boldsymbol{R}}_{\boldsymbol{y}_p^{(u)}} & \cdots & \sum\limits_{k\in\bar{\mathcal{K}}_A} b_k^{(2)} b_k^{(N_P)} \dfrac{\boldsymbol{A}_k \boldsymbol{A}_k^H}{L} \\ \vdots & \vdots & \ddots & \vdots \\ \sum\limits_{k\in\bar{\mathcal{K}}_A} b_k^{(N_P)} b_k^{(1)} \dfrac{\boldsymbol{A}_k \boldsymbol{A}_k^H}{L} & \sum\limits_{k\in\bar{\mathcal{K}}_A} b_k^{(N_P)} b_k^{(2)} \dfrac{\boldsymbol{A}_k \boldsymbol{A}_k^H}{L} & \cdots & \hat{\boldsymbol{R}}_{\boldsymbol{y}_p^{(u)}} \end{pmatrix} \tag{4.36}$$

The cross-correlation matrix is given $(\boldsymbol{p}_i \otimes \boldsymbol{A}_i)/L$, which the AP can construct. Consequently, the AP can construct the weight matrices $\boldsymbol{W}_i^{(u)}$ based on the first equation in (4.16), and finally, estimate the channels of UEs based on (4.18).

It should be noted that with the aid of the multiple receive antennas employed by AP, $\boldsymbol{R}_{y^{(u)}}$ can be directly estimated in the form of (4.19), provided that the number of receive antennas is sufficient. When comparing (4.19) and (4.25), we can know that $\hat{\boldsymbol{R}}_{\boldsymbol{y}_p^{(u)}}$ can be estimated with higher accuracy than $\boldsymbol{R}_{y^{(u)}}$. Therefore, to improve the accuracy of $\boldsymbol{R}_{y^{(u)}}$, we can first estimate it using (4.19) and then, replace the diagonal matrices corresponding to $\boldsymbol{R}_{\boldsymbol{y}_p^{(u)}}$ by the estimated $\hat{\boldsymbol{R}}_{\boldsymbol{y}_p^{(u)}}$.

4.3.4 Performance of Channel Estimation

Figs. 4.3 - 4.5 demonstrate the MSE performance of the channel estimation for both active and inactive UEs, when AP has respectively ideal knowledge, no knowledge and partial knowledge about the active UEs. For convenience, the parameter values used in the studies are associated with the figures. Note specifically that in the case of considering the partial knowledge about active UEs, as shown in Fig.4.1, we assume that each UE may have 1, 2, 3 or 4 frames of the same probability to transmit, once become active. If a UE has more than one frame to transmit, AP only needs to identify it once when it transmits the first frame. For the remaining frames, the activity of the UE is known to the AP.

From the results of Figs. 4.3 - 4.5, we can clearly observe that in general, if AP tries to estimate the channels of an inactive UE, regardless of the number of receive antennas and the knowledge about UEs to AP, the MSE is always relatively high. By contrast, if the channel estimation is for an active UE, the MSE of all the three cases is relatively small, which decreases with the increase of SNR. At a given SNR, when the number of receive antennas increases, the MSE difference between that of the estimation for inactive

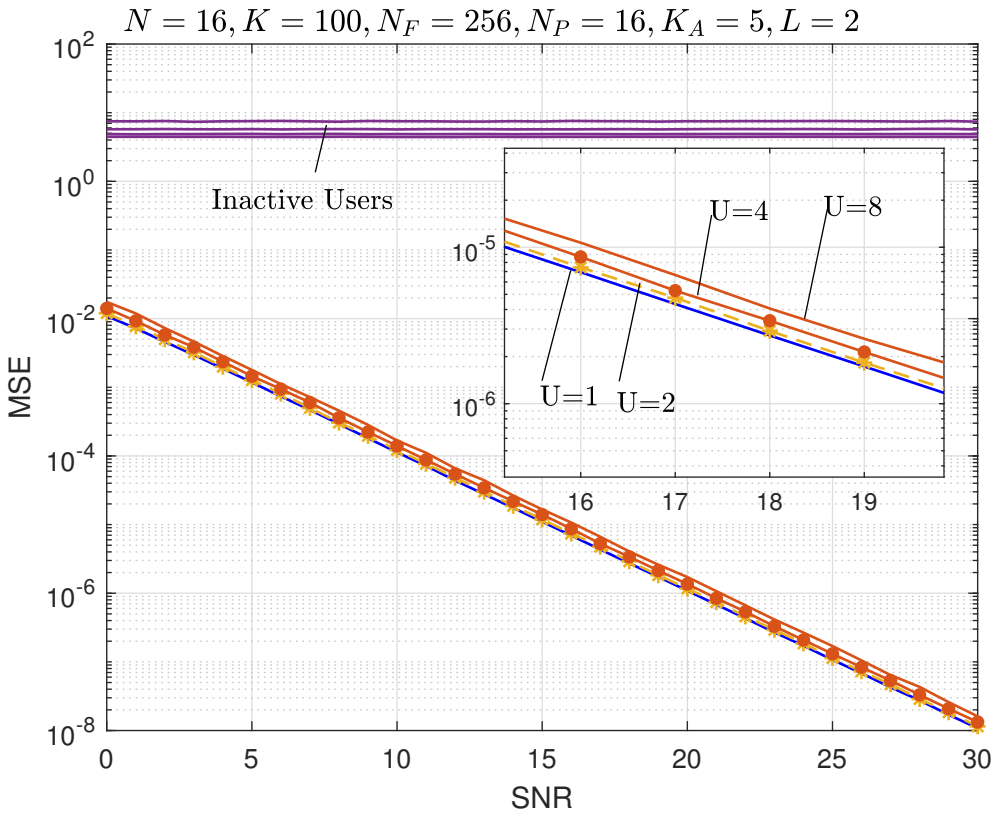


Figure 4.3: MSE performance of the channel estimation for both active UEs and inactive UEs, when AP has ideal knowledge about the active and inactive UEs.

UE and that of the estimation for active UE increases, which is more declared, when SNR increases. This observation implies that employing more receive antennas is beneficial to UAI. However, at a given SNR, the MSE slightly increases with the increase of U. This is because the MSE is linearly proportional to U. If the MSE is normalized by dividing U, the normalized MSE should decreases with the increase of U.

4.4 User Identification Based on the Estimated Multiple Antenna Channels

In the above section, we have investigated the channel estimation in MRA/MC-mGFMA systems. To achieve the reliable signal detection of UEs, it is essential for AP to know which UEs are active. Before we detail the UE activity identification (UAI) scheme, let us first investigate the statistical properties of the estimated channels corresponding to the active and inactive UEs.

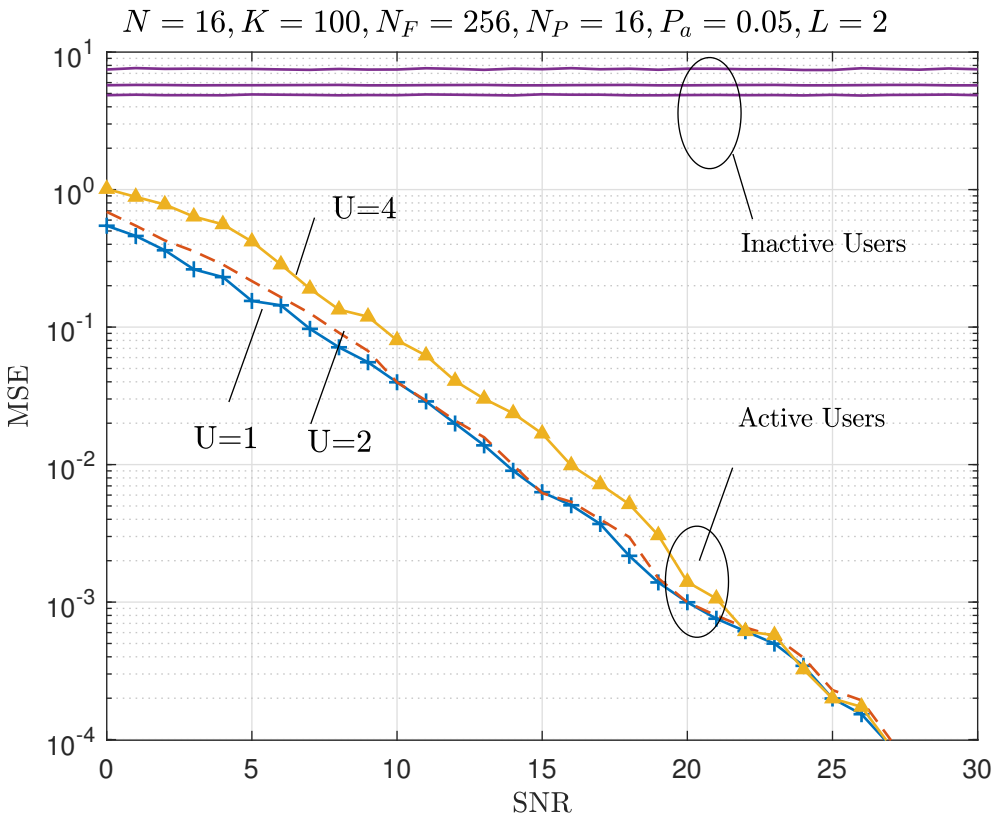


Figure 4.4: MSE performance of the channel estimation for both active UEs and inactive UEs, when AP has no knowledge about the active and inactive UEs.

4.4.1 Statistical Characteristics of Estimated Channels

Following the above analysis we know that the time-domain CIR of the ith UE is $\boldsymbol{h}_i$, which has in total UL taps. After the MMSE-assisted channel estimation in Section 4.3, the estimate to $\boldsymbol{h}_i^{(u)}$ can be written as

$$\hat{\boldsymbol{h}}_i = \begin{cases} \boldsymbol{h}_i + \boldsymbol{n}_i, & \text{if UE } i \text{ is active} \\ \boldsymbol{n}_i', & \text{if UE } i \text{ is inactive} \end{cases} \tag{4.37}$$

where $\boldsymbol{n}_i$ is the channel estimation error of an active UE, which can be Gaussian approximated with a PDF of $\mathcal{CN}\left(\boldsymbol{0}, \sigma_1^2 \boldsymbol{I}_{UL}\right)$, with σ_1^2 being the variance of channel estimation error of active UEs. By contrast, $\boldsymbol{n}_i'$ is the estimate to an inactive UE, which can also be approximated by the Gaussian distribution with the PDF of $\mathcal{CN}\left(\boldsymbol{0}, \sigma_0^2 \boldsymbol{I}_{UL}\right)$, where σ_0^2 is in general different from σ_1^2.

For comparison, let us consider the statistics of $|\hat{\boldsymbol{h}}_i|^2/U$ on the condition that UE i is active or inactive. First, when UE i is active, the elements of $\hat{\boldsymbol{h}}_i$ are iid complex Gaussian random variables, each of which has zero mean and a variance of $(1/L + \sigma_1^2)$, when Rayleigh fading channels are assumed, and when all component paths have the same power of $1/L$. Therefore, $|\hat{\boldsymbol{h}}_i|^2$ obeys the centre χ^2-distributions with $2UL$ degrees of freedom [192], and

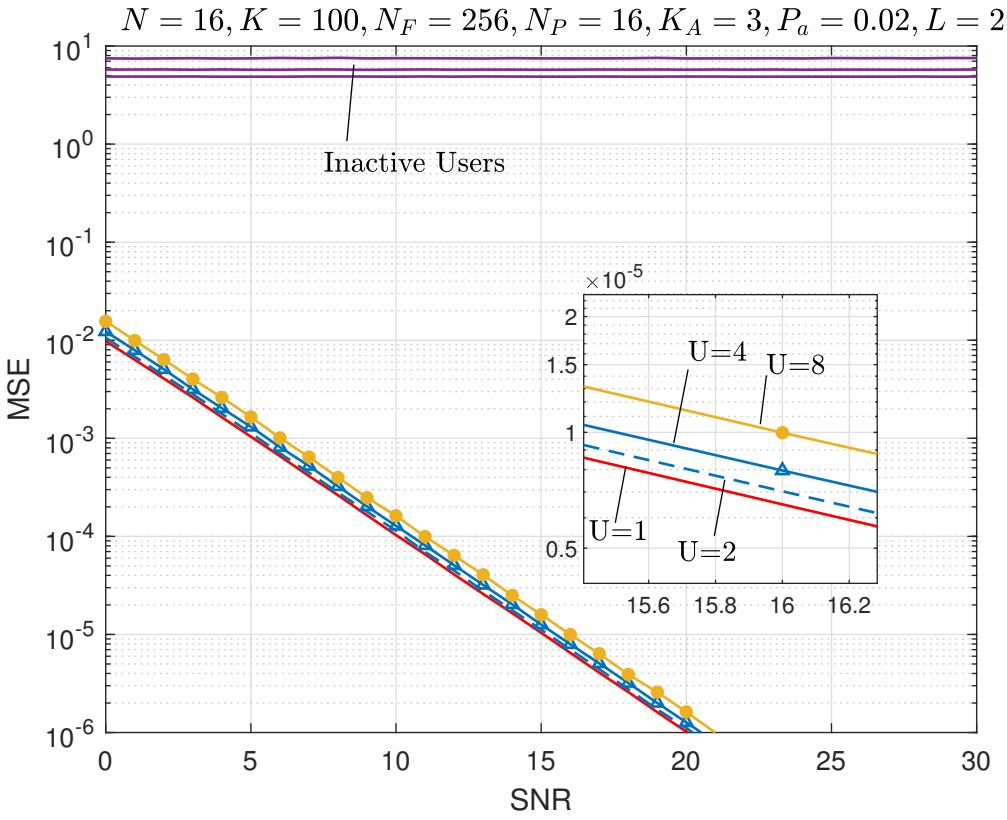


Figure 4.5: MSE performance of the channel estimation for both active UEs and inactive UEs, when AP has partial knowledge about the active and inactive UEs.

the PDF of $|\hat{\boldsymbol{h}}_i|^2/U$ can be found to be

$$f_{|\hat{\boldsymbol{h}}_i(U)|^2/U}(y|I_i = 1, K_A) = \frac{U^{UL}}{(UL-1)!(1/L+\sigma_1^2(K_A))^{UL}} \times y^{UL-1} \exp\left(-\frac{Uy}{1/L+\sigma_1^2(K_A)}\right), \; y \geq 0 \tag{4.38}$$

where we explicitly show that $\sigma_1^2(K_A)$ is related to the number of active UEs. By contrast, when the ith UE is inactive, it can be shown that $|\hat{\boldsymbol{h}}_i|^2/U$ follows the centre χ^2-distribution of

$$f_{|\hat{\boldsymbol{h}}_i|^2/U}(y|I_i = 0, K_A) = \frac{U^{UL}}{(UL-1)!(\sigma_0^2(K_A))^{UL}} y^{UL-1} \times \exp\left(-\frac{Uy}{\sigma_0^2(K_A)}\right), y \geq 0 \tag{4.39}$$

However, directly evaluating the PDFs of (4.38) and (4.39) is highly challenging as $\sigma_1^2(K_A)$ and $\sigma_0^2(K_A)$ are related to the MMSE estimation and also to the number of access UEs. Therefore, we below demonstrate the PDFs, as shown for example in Figs. 4.6 - 4.8, when different parameters in terms of L and U are set.

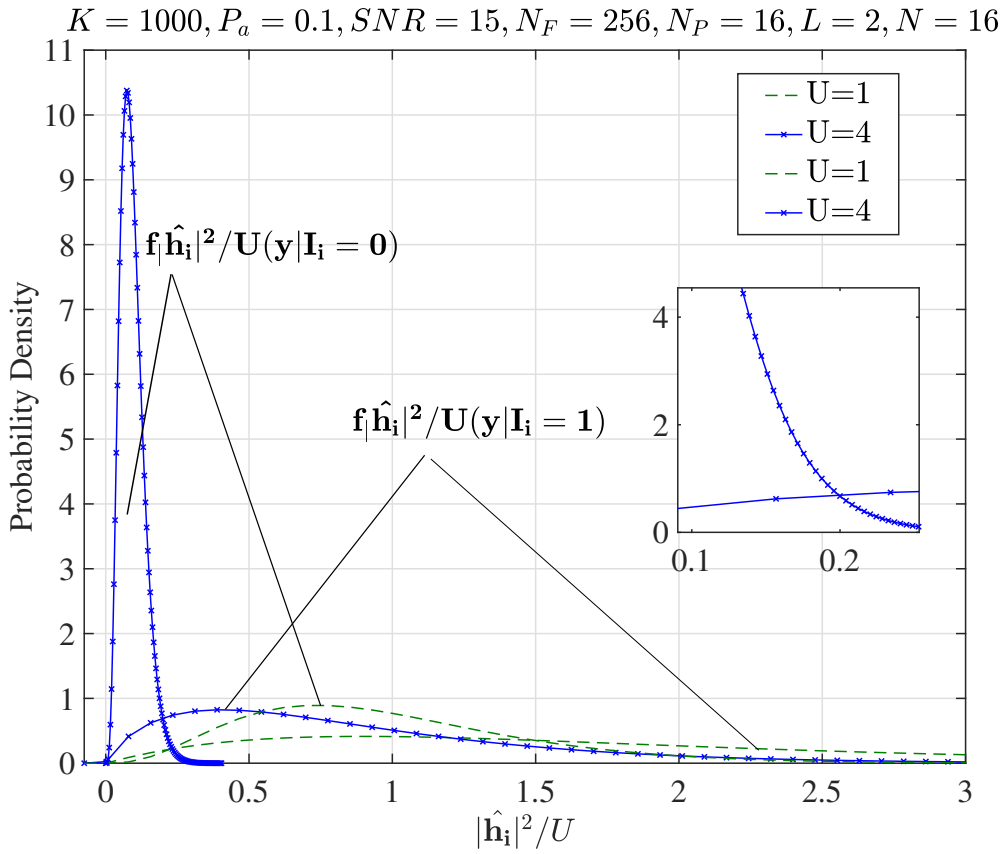


Figure 4.6: PDFs of $|\hat{\boldsymbol{h}}_i|^2/U$ on condition that UE i is active $(I_i = 1)$ or inactive $(I_i = 0)$, when the channels of receive antennas experience iid Rayleigh fading.

From the results of Figs. 4.6 - 4.8 and also implied by the properties of χ^2-distributions given by (4.38) and (4.39), we can know that first, when a UE is active, the value of $|\hat{\boldsymbol{h}}_i^{(u)}|^2/U$ is usually distributed in the range having relatively high value. By contrast, for an inactive UE, the value of $|\hat{\boldsymbol{h}}_i^{(u)}|^2/U$ is usually small. Second, when U becomes larger, this difference between the PDFs of active and inactive UEs becomes more significant. This can be explicitly conceived by Fig. 4.6. This property implies that UEs' activities can be identified based on the values of $|\hat{\boldsymbol{h}}_i^{(u)}|^2/U$, in particular, when the number of receive antennas is high. This is beneficial to practice implementation, as the threshold-based UAI (TB-UAI) usually has low-complexity for implementation, if the threshold can be relatively easily set to a near-optimal value. Additionally, from Figs. 4.6 - 4.8, we can observe that when the value of L increases, the difference between the PDFs of $|\hat{\boldsymbol{h}}_i^{(u)}|^2/U$ for active and inactive UEs is also slightly enhanced, although the enhancement is not as significant as the case when increasing U. Therefore, for a MRA/MC-mGFMA system experiencing highly frequency-selective fading and also employing a relatively big number of receive antennas at APs, the low-complexity threshold-based UAI (TB-UAI) scheme can be highly efficient, which is investigated below in the next subsection.

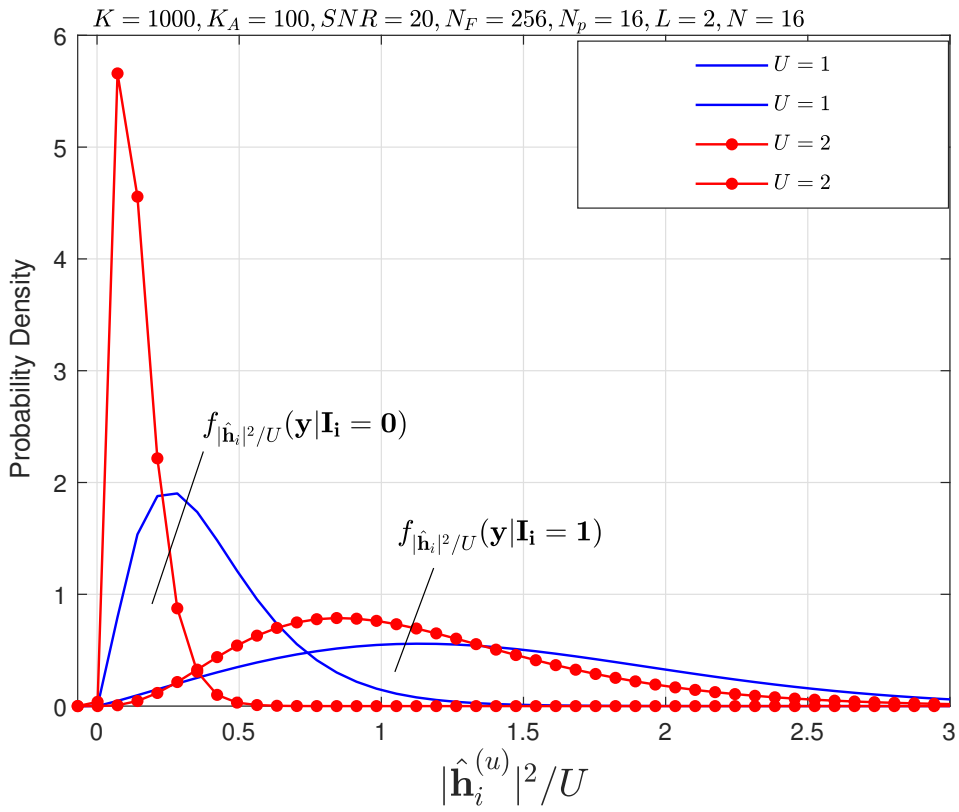


Figure 4.7: PDFs of $|\hat{\boldsymbol{h}}_i|^2/U$ on condition that UE i is active $(I_i = 1)$ or inactive $(I_i = 0)$, when the channels of receive antennas experience iid Rayleigh fading.

4.4.2 Threshold-Based UE Activity Identification (TB-UAI)

Let T_h be the corresponding threshold set in the TB-AUI scheme. Then, when given the estimated CIR of $|\hat{\boldsymbol{h}}_i^{(u)}|^2$ for UE i, the TB-AUI identifies the ith UE's state according to

$$\hat{I}_i = \begin{cases} 1, & \text{if } |\hat{\boldsymbol{h}}_i|^2/U \geq T_h, \\ 0, & \text{else} \end{cases} \tag{4.40}$$

where $\hat{I}_i^{(u)} = 1$ or 0 represents UE i is identified to be active or inactive.

With the aid of the statistics of $|\hat{\boldsymbol{h}}_i|^2/U$ given above, we can derive the miss and false-alarm probabilities of the TB-AUI scheme as follows. Firstly, the miss probability conditioned on K_A active UEs is

$$P_M(K_A) = \int_0^{T_h} f_{|\hat{\boldsymbol{h}}_i|^2/U}(\boldsymbol{y}|I_i = 1, K_A)dy \tag{4.41}$$

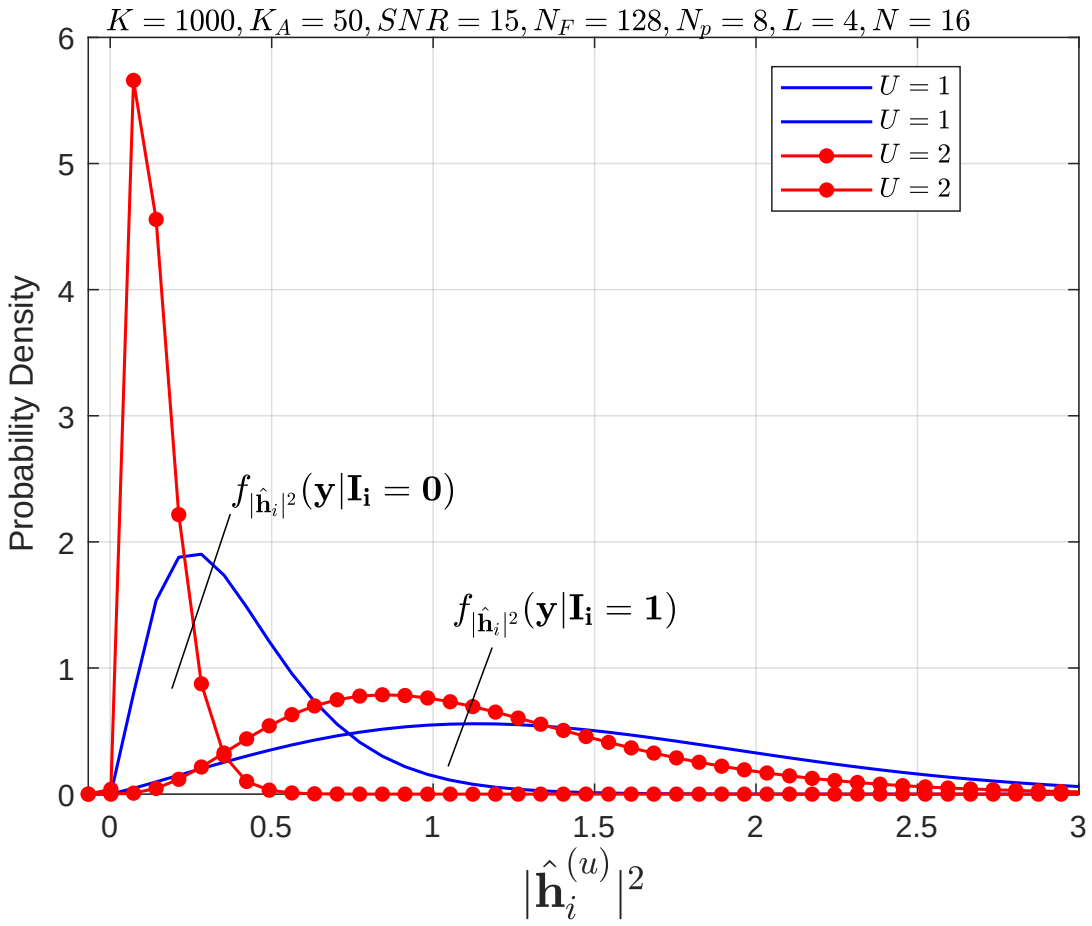


Figure 4.8: PDFs of $|\hat{\boldsymbol{h}}_i|^2/U$ on condition that UE i is active $(I_i = 1)$ or inactive $(I_i = 0)$, when the channels of receive antennas experience iid Rayleigh fading.

Upon substituting (4.38) into (4.41) and completing the integration, we can obtain

$$P_M(K_A) = 1 - \exp\left(-\frac{ULT_h}{1 + L\sigma_1^2(K_A)}\right) \times \sum_{k=0}^{UL-1} \frac{1}{k!}\left(\frac{ULT_h}{1 + L\sigma_1^2(K_A)}\right)^k \tag{4.42}$$

Secondly, the false-alarm probability when there are K_A active UEs is given by

$$P_F(K_A) = \int_{T_h}^{\infty} f_{|\hat{\boldsymbol{h}}_i^{(u)}|^2}(\boldsymbol{y}|I_i = 0, K_A)dy \tag{4.43}$$

After substituting (4.39) into (3.40), we obtain

$$P_F(K_A) = \exp\left(-\frac{UT_h}{\sigma_0^2(K_A)}\right) \sum_{k=0}^{UL-1} \frac{1}{k!}\left(\frac{UT_h}{\sigma_0^2(K_A)}\right)^k \tag{4.44}$$

It can be shown that the miss probability of (4.42) and the false-alarm probability of (4.44) can be approximated with the aid of the Taylor Theorem for the remainders of

exponential functions, when the number of receive antennas is sufficiently large [219]. Let us express

$$P(x) = \exp(-x) \sum_{k=0}^{UL-1} \frac{x^k}{k!} \tag{4.45}$$

Then, we can modify it to

$$\begin{aligned} P(x) =& \exp(-x) \left(\exp(x) - \sum_{k=UL}^{\infty} \frac{x^k}{k!} \right) \\ =& 1 - \exp(-x) \sum_{k=UL}^{\infty} \frac{x^k}{k!} \end{aligned} \tag{4.46}$$

Now, applying the Neyman-Pearson approximation [219], we obtain

$$\begin{aligned} P(x) &\approx 1 - \gamma \left(\exp(-x) \cdot \frac{\exp(x)}{(UL)!} x^{UL} \right) \\ &= 1 - \frac{\gamma x^{UL}}{(UL)!} \end{aligned} \tag{4.47}$$

where $\gamma = \exp^{(a-1)x}$ is a function of x.

Applying Lagrange Theorem of the remainders of exponential functions

$$P(x) = 1 - \exp(-x) \frac{\exp^{\theta}}{(UL)!} x^{UL}, \textit{where } \theta \in (0, x) \tag{4.48}$$

let $\theta = ax$, with $0 < a < 1$, we have

$$\begin{aligned} P(x) =& 1 - \exp(-x) \exp^{(ax)} \frac{x^{UL}}{(UL)!} \\ =& 1 - \exp^{(a-1)x} \frac{x^{UL}}{(UL)!} \end{aligned} \tag{4.49}$$

then, define $\gamma = \exp^{(a-1)x}$.

With the approximated result of (4.47), the miss and false-alarm probabilities can now be expressed as

$$P_M(K_A) = \frac{\gamma_M}{(UL)!} \left(\frac{ULT_h}{1 + L\sigma_1^2(K_A)} \right)^{UL} \tag{4.50}$$

and

$$P_F(K_A) = 1 - \frac{\gamma_F}{(UL)!} \left(\frac{UT_h}{\sigma_0^2(K_A)} \right)^{UL} \tag{4.51}$$

respectively.

Having the miss and false-alarm probabilities prepared, the probability of erroneous UAI conditioned on K_A active UEs can be obtained as

$$P_E(K_A) = P_a P_M(K_A) + (1 - P_a) P_F(K_A) \tag{4.52}$$

where P_a is the small activation probability of a UE. Furthermore, if we assume that UEs activate independently and have the same activation probability P_a, the number of active UEs K_A within a given time-slot obeys the binomial distribution with the probability mass function (PMF) of

$$P(K_A) = \binom{K}{K_A} P_a^{K_A} (1 - P_a)^{K - K_A}, K_A = 0, 1, \cdots, K \tag{4.53}$$

Therefore, the average miss probability is

$$P_M = \sum_{K_A=1}^{K} P(K_A) P_M(K_A) \tag{4.54}$$

The average false-alarm probability is

$$P_F = \sum_{K_A=0}^{K-1} P(K_A) P_F(K_A) \tag{4.55}$$

Finally, the average probability of erroneous UAI is

$$P_E = P_a P_M + (1 - P_a) P_F \tag{4.56}$$

Note that, in the TB-AUI algorithm, the threshold T_h may be chosen according to different requirements. First, it may be selected to minimize the erroneous probability P_E. Second, it can be selected to protect a small false-alarm probability. In this case, the threshold can be chosen as $T_h = \arg\{\max\{P_F(T_h')\} \leq \bar{P}_F\}$, where $\bar{P}_F$ is the maximum false-alarm probability allowed. Furthermore, to guarantee a small miss probability, the threshold can be selected as $T_h = \arg\{\max\{P_M(T_h')\} \leq \bar{P}_M\}$.

Note furthermore that in comparison with the single-antenna AP, the multiple antennas of AP will benefit the threshold selection in TB-AUI. This is because with the aid of the multiple channels corresponding to multiple receive antennas estimated by AP, the estimated channels for an active UE should appear significant difference from that for an inactive UE.

4.4.3 Performance of TB-UAI

Below we provide a range of results to demonstrate the performance of the TB-UAI scheme analyzed in Section 4.4.2.

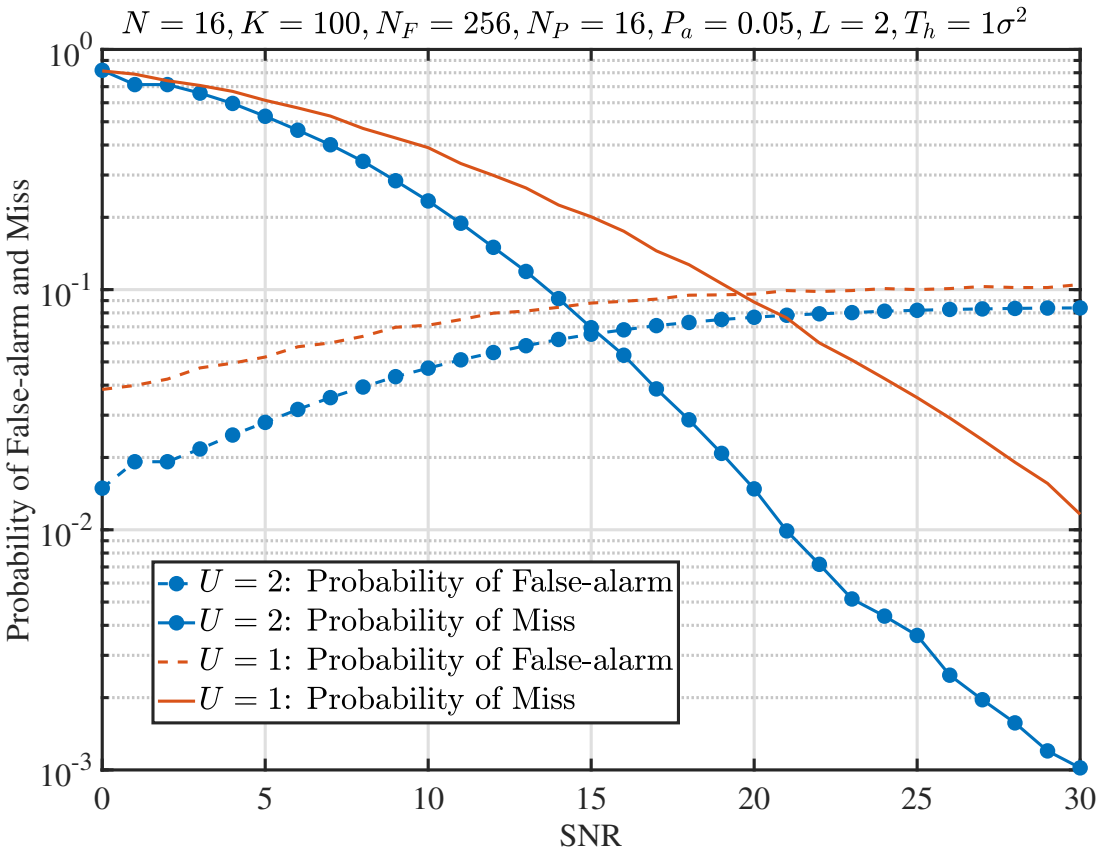
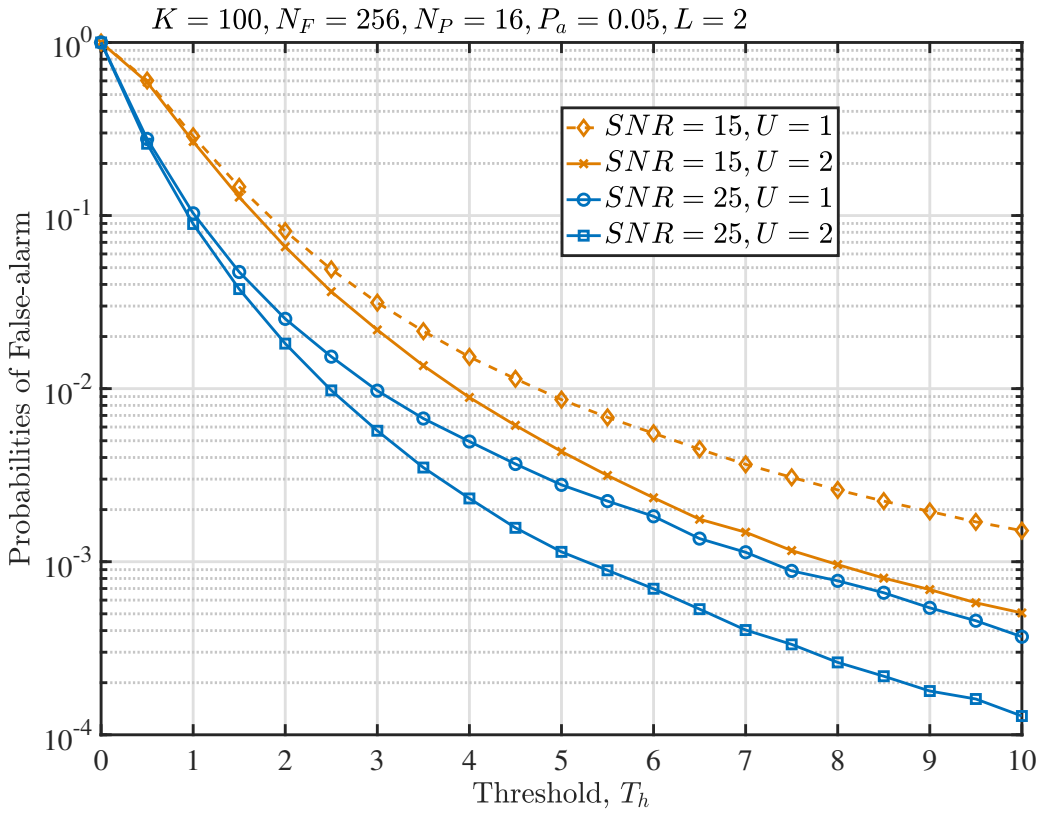


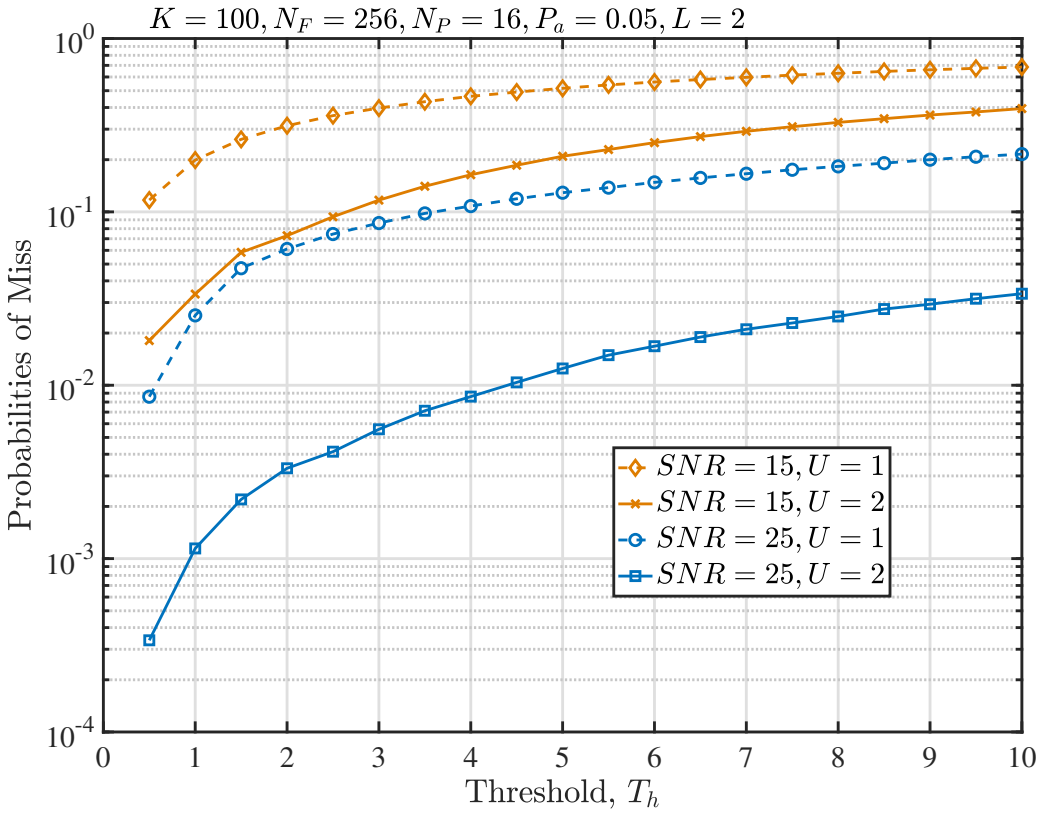
Figure 4.9: Probabilities of false alarm and miss, when UEs are identified using the TB-UAI scheme.

First in Fig. 4.9, we demonstrate the effect of SNR on the performance of TB-UAI, when the parameters as shown on the top of the figure are employed. When the threshold is set to $T_h = \sigma^2 = 1/SNR$, we observe that the miss probability decreases significantly, as the SNR increases. This is because the channels can be more reliably estimated, when SNR is higher. Furthermore, as SNR increases, the threshold $T_h = \sigma^2 = 1/SNR$ decreases. Both of the above-mentioned effects result in the decrease of the miss probability. By contrast, the false-alarm probability slightly increases with the increase SNR, which is mainly because of the reduction of T_h resulted from the increase of SNR. As shown in Fig. 4.9, when the number of receive antennas is increased from $U = 1$ to $U = 2$, the miss probability yields a significant decrease at a given SNR and at the same time, the false-alarm also reduces. Hence, we can be implied that using multiple receive antennas at AP is capable of significantly improve the performance of UAI, even when the low-complexity TB-UAI scheme is employed.

In Fig. 4.10, we demonstrate the false-alarm and miss probabilities with respect to threshold for the MRA/MC-mGFMA systems, when channels are estimated based on $N_P = 16$ pilots in the frames of length $N_F = 256$. The other parameters are detailed on



(a) Prob. of false-alarm



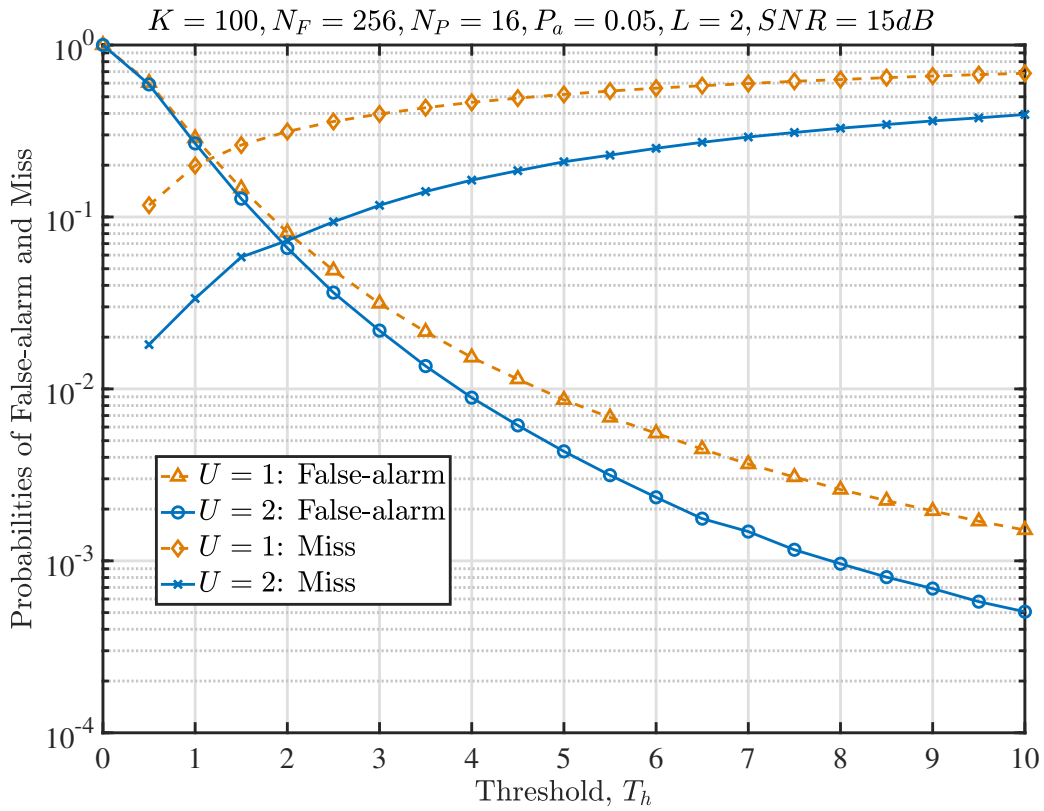
(b) Prob. of miss

Figure 4.10: Probabilities of false-alarm and miss, when UEs are identified using the channel-estimation based TB-UAI.

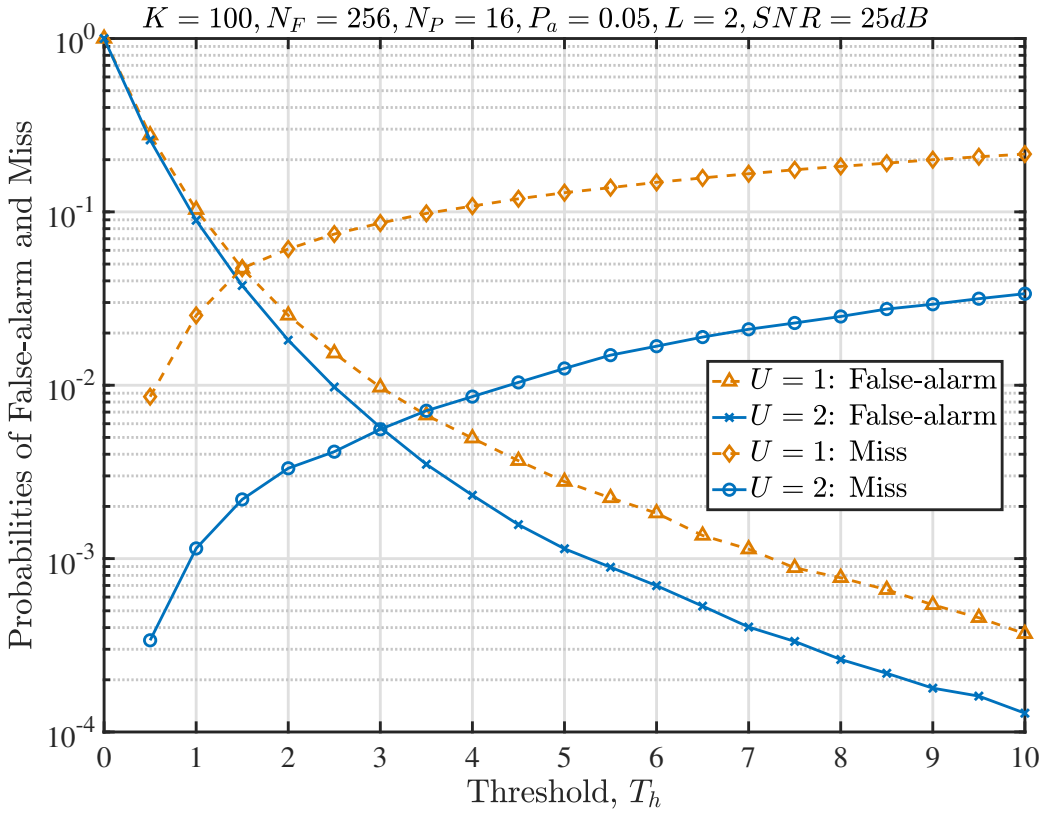
the tops of the figures. Note that, the threshold shown in the figures is normalized by the noise variance of σ^2. From the results we observe that when T_h increases, the false-alarm probability decreases while the miss probability increases. Hence, there exists trade-off between false-alarm probability and miss probability, which was shown in Fig. 4.9 and will also be demonstrated in the figures below. As shown in Fig. 4.10, both the false-alarm and miss probabilities reduced, when the SNR is increased from $SNR = 15$ dB to $SNR = 25$ dB. Furthermore, when the number of receive antennas at AP is increased from $U = 1$ to $U = 2$, while the false-alarm probability has a mild improvement, the miss probability presents significant improvement.

Fig. 4.11 shows the trade-off between the false-alarm probability and miss probability, which are drawn in terms of the threshold that was normalized by the noise power σ^2. Explicitly, the false-alarm probability decreases while the miss probability increases, when the threshold increases. Therefore, if we motivate small miss probability, the false-alarm probability would be relatively high. On the other side, if a small false-alarm probability is the design objective, the system should be able to tolerate a relatively big miss probability. However, if AP is deployed with a relatively big number of antennas, both the false-alarm probability and miss probability can be small values. From this observation we are implied that when the AP has very low number of receive antennas, such as $U = 1$ in the last chapter, the TB-UAI scheme cannot work well, as in this case, the UAI performance is very sensitive to the threshold. By contrast, when AP has multiple receive antennas, as implied by the results in Fig. 4.11, we may easily achieve the required small false-alarm and miss probabilities by setting the threshold near the point where the false-alarm probability curve and miss probability curve cross.

Finally, in Fig. 4.12, we compare the performance achieved by our proposed signature code and that attained by the random signature codes. Explicitly, for any given threshold, both the miss probability and the false-alarm probability achieved by our proposed signature coders are smaller than the corresponding ones by the random signature codes. Hence, we can expect that the MRA/MC-mGFMA systems employing the proposed signature codes are capable of achieving better UAI performance than the MRA/MC-mGFMA systems employing random signature codes. Furthermore, when data detection is considered, the MRA/MC-mGFMA systems employing the proposed signature codes should also achieve higher detection reliability than the MRA/MC-mGFMA systems employing random signature codes, as to be studied in Section 4.7.



(a) Trade-off between Prob. of false-alarm and miss



(b) Trade-off between Prob. of false-alarm and miss

Figure 4.11: Trade-off between the probabilities of false-alarm and miss, when UEs are identified using the channel estimation based TB-UAI.

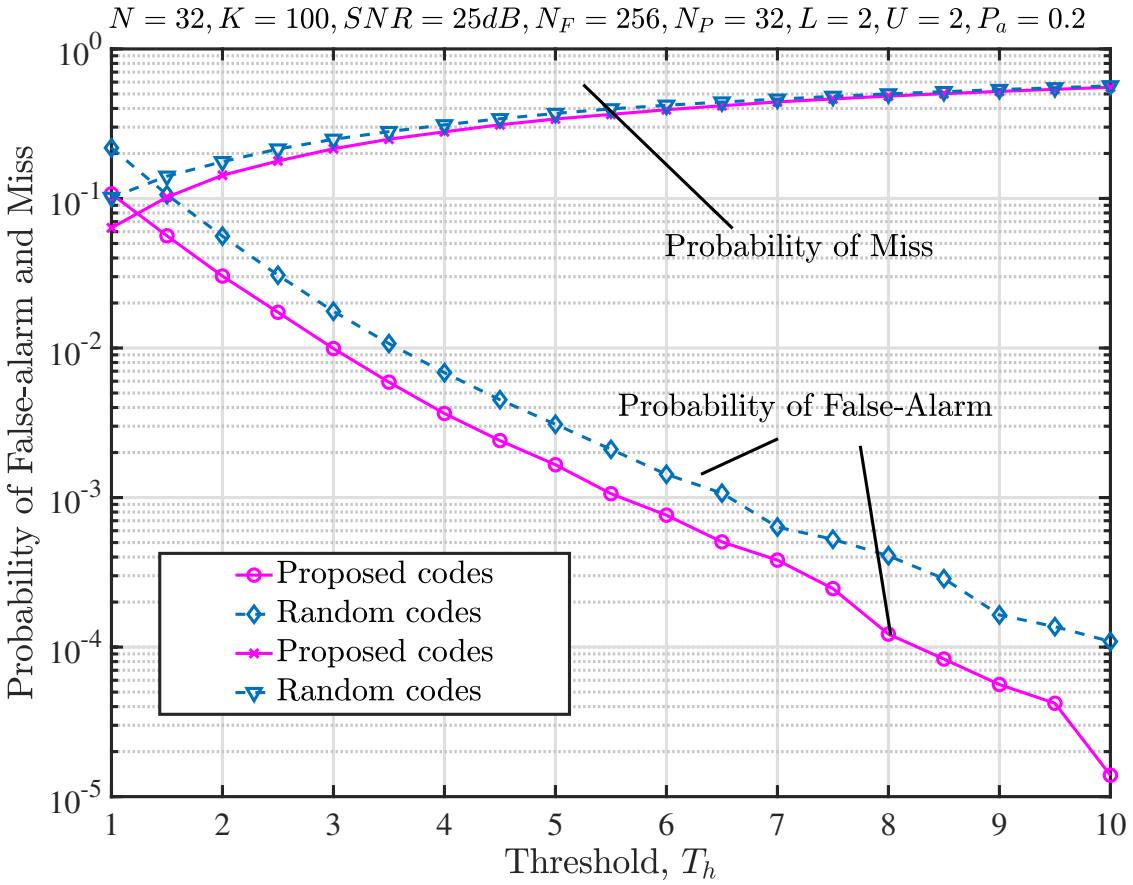


Figure 4.12: Trade-off between the probabilities of false-alarm and miss, when UEs are identified using the channel estimation based TB-UAI.

4.5 Successive Interference Cancellation Assisted Data Detection

After channel estimation and UAI, data detection can be carried out. To provide low-complexity detection while still achieve near optimum detection performance, we propose the successive interference cancellation (SIC)-assisted detection, which requires MMSE processing, reliability measurement and successive interference cancellation, as detailed below.

4.5.1 Minimum Meam-Square Error Based Detection

To detect UE i in the principle of MMSE, the decision variable is formed as

$$z_{i,m} = \boldsymbol{w}_i^H \boldsymbol{y}_m, m = 1, 2, \ldots, N_F; \\ i = 1, 2, \ldots, K_A \tag{4.57}$$

where the data symbols in one frame are all detected, regardless of the pilot symbols or data symbols. Assuming this has two folds of consideration. First, it is for convenience of notification. Second, the detection of pilots may have further application, such as, for further identifying a falsely identified UE by the TB-UAI scheme. This issue will be further studied in Section 4.6.

In (4.57), $\boldsymbol{y}_m$ is given by (4.3) for the mth symbol, the weight vector $\boldsymbol{w}_i$ is expressed as

$$\boldsymbol{w}_i = \boldsymbol{R}_{y_m}^{-1} \boldsymbol{r}_{y_m,i},\ i = 1, 2, \ldots, K \tag{4.58}$$

where the autocorrelation matrix is

$$\boldsymbol{R}_{y_m} = E\left[\boldsymbol{y}_m \boldsymbol{y}_m^H\right] \tag{4.59}$$

which has the expression of

$$\boldsymbol{R}_{y_m} = \sum_{k'=1}^{K_A} \boldsymbol{B}_{k'} \hat{\boldsymbol{h}}_{k'} \hat{\boldsymbol{h}}_{k'}^H \boldsymbol{B}_{k'}^H + \sigma^2 \boldsymbol{I}_{UN} \tag{4.60}$$

where $\hat{\boldsymbol{h}}_{k'}$ are estimated channels. Alternatively and especially when active UEs cannot be accurately identified, $\boldsymbol{R}_{y_m}$ can be estimated as

$$\hat{\boldsymbol{R}}_{y_m} = \frac{1}{N_F} \sum_{m=1}^{N_F} \boldsymbol{y}_m \boldsymbol{y}_m^H \tag{4.61}$$

In (4.58), $\boldsymbol{r}_{y_m,i}$ is the cross-correlation vector between $\boldsymbol{y}_m$ and $b_{i,m}$, which is given by

$$\boldsymbol{r}_{y_m,i} = E\left[\boldsymbol{y}_m b_{i,m}\right] = \boldsymbol{B}_i \boldsymbol{h}_i \tag{4.62}$$

Since the channels of UE i have been estimated and $\boldsymbol{B}_i$ is known to AP, AP can construct $\boldsymbol{r}_{y_m,i}$ as

$$\boldsymbol{r}_{y_m,i} = \boldsymbol{B}_i \hat{\boldsymbol{h}}_i,\ i = 1, 2, \ldots, K \tag{4.63}$$

Finally, if decisions are needed, based on $z_{i,m}$, the binary data symbols are detected as

$$\begin{aligned} \hat{b}_{i,m} = &\begin{cases} +1, & \text{if } z_{i,m} > 0, \\ -1, & \text{else} \end{cases} \\ & m = 1, 2, \ldots, N_F; \\ & i = 1, 2, \ldots, K_A \end{aligned} \tag{4.64}$$

4.5.2 Reliability Measurement

Owing to its relatively low complexity for implementation and capable of achieving near-optimum bit error rate (BER) performance, the SIC-assisted detection has drawn a lot of research attention in multiuser communications scenarios [127]. In [127], the author investigated a near-optimum reliability measurement method, which is high efficiency for operation in both space-division multiple-access (SDMA) and CDMA systems. In [127, 142], the simulation results showed that near-optimum performance is achievable, when communicating over Rayleigh fading channels. Furthermore, the research results in [142, 220] showed that even when the loading factor, which is the number of supported users normalized by the number of resource units, such as, spreading factor, is as high as 2, the system can still work efficiently and achieve near-optimum performance. Due to its successive implementation structure, the SIC-assisted detection scheme proposed in [142] is well suitable for operation in the dynamic mGFMA systems, including our MRA/MC-mGFMA system.

To implement the SIC-assisted detection, the detection reliabilities of bits are required to be measured, so that UEs are detected from the most reliable one to the least reliable one. In this chapter, we introduce the method proposed in [142], which for our MRA/MC-mGFMA system measures the reliability according to

$$\begin{aligned} L_{i,m} =& (1+\tilde{\gamma}_i) z_{i,m} \\ =& \left(\frac{1}{1-\hat{\boldsymbol{h}}_i^H \hat{\boldsymbol{R}}_{y_m}^{-1} \hat{\boldsymbol{h}}_i} \right) z_{i,m}, \\ & m = 1,2,\ldots,N_F; i = 1,2,\ldots,K_A \end{aligned} \tag{4.65}$$

where $\tilde{\gamma}_i = \hat{\boldsymbol{h}}_i^H \hat{\boldsymbol{R}}_{y_m}^{-1} \hat{\boldsymbol{h}}_i / (1 - \hat{\boldsymbol{h}}_i^H \hat{\boldsymbol{R}}_{y_m}^{-1} \hat{\boldsymbol{h}}_i)$ is the measured signal-to-interference-plus-noise ratio (SINR) of UE i, $\hat{\boldsymbol{h}}_i$ is the estimated channel of UE i, given in Section 4.3.2 or Section 4.3.3, $\hat{\boldsymbol{R}}_{y_m}$ and $z_{i,m}$ are given by (4.61) and (4.57), respectively. Additionally, it is assumed that there are K_A active UEs identified. According to [142], when the value of $L_{i,m}$ is larger, the detection of the corresponding bit of a UE is more reliable.

4.5.3 SIC-Assisted MMSE Detection in Dynamic MRA/MC-mGFMA Systems

Assume that there are $K_A = |\mathcal{K}_A|$ UEs identified active. Then, during each symbol duration, the SIC-assisted MMSE (SIC-MMSE) detector operates K_A iterations to detect the K_A active UEs, each iteration detects one bit of a UE. Note that, in order to reduce the

detection latency, the bits of one frame can be detected in parallel without affecting the detection performance. This is because once the channels are estimated and the active UEs are identified, the detections of the different bits in one frame are independent. Hence, our description below ignores the index m for brevity. Additionally, we note that to further reduce the detection latency, in one iteration, all the bits of K_A UEs with their reliabilities higher than a pre-set threshold may be detected.

Algorithm 4.5.3 SIC-Assisted MMSE (SIC-MMSE) Detection Algorithm.

Initialization:

$\boldsymbol{y}^{(0)} = \boldsymbol{y}_m$, $\{\boldsymbol{B}_k\}$, K_A, $\boldsymbol{R}_y^{(0)} = \hat{\boldsymbol{R}}_{y_m}$, $\{\hat{\boldsymbol{h}}_k\}$, $k \in \mathcal{K}_A$.

Preparation:

$\tilde{\boldsymbol{H}}^{(0)} = \left[\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \ldots, \hat{\boldsymbol{h}}_{K_A}\right]$, $\boldsymbol{W}^{(0)} = \left(\boldsymbol{R}_y^{(0)}\right)^{-1} \tilde{\boldsymbol{H}}^{(0)}$, $\boldsymbol{Q}^{(0)} = \boldsymbol{I}_{K_A} - (\tilde{\boldsymbol{H}}^{(0)})^H \boldsymbol{W}^{(0)}$;

Detection:

for $s = 1, 2, \ldots, K_A$:

1. *Forming decision variables*:
   $\boldsymbol{z}^{(s)} = \Re\left\{\left(\boldsymbol{W}^{(s-1)}\right)^H \boldsymbol{y}^{(s-1)}\right\}$;
2. *Determine the most reliable UE*: For the UEs $k'_1, k'_2, \ldots, k'_{K_A-s+1}$ not detected, compute their reliabilities according to (5.43), and find the most reliable UE as $k^{(s)} = \arg\max_{k'_i}\{L_{k'_1}, L_{k'_2}, \ldots, L_{k'_{K_A-s+1}}\}$;
3. *Detect the most reliable UE*: $\hat{b}_{k^{(s)}} = \text{sgn}\left(\boldsymbol{z}_{k^{(s)}}^{(s)}\right)$, where $\boldsymbol{z}_{k^{(s)}}^{(s)}$ is the $k^{(s)}$th entry of $\boldsymbol{z}^{(s)}$;
4. *Interference cancellation*:
   $\boldsymbol{y}^{(s)} = \boldsymbol{y}^{(s-1)} - \boldsymbol{B}_{k^{(s)}} \boldsymbol{h}_{k^{(s)}} \hat{b}^{(s)}$;
5. *Update*:

$$\begin{aligned} \boldsymbol{W}^{(s)} &= \left[\boldsymbol{W}^{(s-1)} + \frac{\boldsymbol{w}_{k^{(s)}}^{(s-1)} \left(\boldsymbol{h}_{k^{(s)}}\right)^H \boldsymbol{W}^{(s-1)}}{\boldsymbol{Q}^{(s-1)}(k^{(s)}, k^{(s)})}\right] \boldsymbol{P}^{(s)}, \\ \tilde{\boldsymbol{H}}^{(s)} &= \tilde{\boldsymbol{H}}^{(s-1)} \boldsymbol{P}^{(s)}, \\ \boldsymbol{Q}^{(s)} &= \boldsymbol{I} - (\tilde{\boldsymbol{H}}^{(s)})^H \boldsymbol{W}^{(s)}. \end{aligned} \tag{4.66}$$

end for

Outputs: Detected data bits of K_A active UEs.

The detection is summarized in Algorithm 4.5.3. In the algorithm, $\boldsymbol{Q}^{(s-1)}(k^{(s)},k^{(s)})$ is the $(k^{(s)},k^{(s)})$th entry of $\boldsymbol{Q}^{(s-1)}$, while $\boldsymbol{w}_{k^{(s)}}^{(s-1)}$ is the $k^{(s)}$th column of $\boldsymbol{W}^{(s-1)}$. Furthermore, $\tilde{\boldsymbol{H}}^{(s)}$ is obtained from $\tilde{\boldsymbol{H}}^{(s-1)}$ by deleting the column of $\boldsymbol{h}^{(s)}$ corresponding to the UE just detected at the sth detection iteration, and $\boldsymbol{P}^{(s)}$ is a permutation matrix obtained from $\boldsymbol{I}_{K_A}$ after deleting the s columns corresponding to the UEs already detected.

From the description of Algorithm 4.5.3, we can know that the SIC-MMSE detector operates K_A iterations with each iteration detecting one most reliable UE. Then the interference of the detected UE on the not yet detected UEs is removed, so as to enhance the detection performance of the following UEs. Furthermore, at the end of a detection iteration, the weight matrix as well as the channel matrix for the remaining UEs are updated.

As discussed in [142], The SIC-MMSE detector can be efficiently implemented using the existing algorithms proposed in the context of the V-BLAST detection [221], yielding the detection complexity of the SIC-MMSE being on the order of $\mathcal{O}\left(c_1K_AN+c_2N^2\right)$ per UE, where c_1 and c_2 are certain constants. The detection delay is also linearly depended on K_A and the maximum detection delay is K_A symbol (bit) durations, when N_F symbols per frame are detected in parallel.

4.6 Joint Channel Estimation, UE Activity Identification and Data Detection

Algorithm 4.6 SIC-MMSE assisted joint channel estimation, UAI and data detection.

Initialization:

$\boldsymbol{y}_m{}^{(0)}=\boldsymbol{y}_m, m=1,2,\ldots,N_F$,
$\{\boldsymbol{A}_k\}$ or $\{\boldsymbol{B}_k\}$, K_M, pilot bits.

Preparation:

Estimate $\hat{\boldsymbol{R}}_b$ based on (4.32), and set $\hat{\boldsymbol{R}}_b^{(0)}=\hat{\boldsymbol{R}}_b$.

For $i=1,2,\ldots,K_M$:

1. *Estimate the channels of the* $(K-i+1)$ *unidentified UEs based on* $\boldsymbol{R}_b^{(i-1)}$ *and* (4.35). *Let the estimated channels be expressed as* $\hat{\boldsymbol{h}}_{1'},\hat{\boldsymbol{h}}_{2'},\cdots,\hat{\boldsymbol{h}}_{(K-i+1)'}$.

2. *Identify the most reliable active UE as*
   $k' = \arg\max_k \{|\hat{\boldsymbol{h}}_{1'}|^2, |\hat{\boldsymbol{h}}_{2'}|^2, \cdots, |\hat{\boldsymbol{h}}_{(K-i+1)'}|^2\}$.

3. *Compute* $\boldsymbol{w}_{k'} = \boldsymbol{R}_b^{(i-1)} \hat{\boldsymbol{h}}_{k'}$, *and update* $\boldsymbol{R}_b^{(i)} = \boldsymbol{R}_b^{(i-1)} - \boldsymbol{B}_{k'} \hat{\boldsymbol{h}}_{k'} \hat{\boldsymbol{h}}_{k'}^H \boldsymbol{B}_{k'}^H$

4. *Detect the pilots and data bits of the most reliable UE as*
   $\hat{b}_{k',m} = \text{sgn}\left(z_{k',m} = \Re\{\boldsymbol{w}_{k'}^H \boldsymbol{y}_m^{(i-1)}\}\right), m = 1, \ldots, N_F$.

   - **if** *the number of erroneous pilot symbols is higher than a preset threshold* N_C, *the UAI process finishes;*
   - **if** *the number of erroneous pilot symbols is lower than or equal to* N_C, *UE* k' *is identified as active, and execute the following operations:*

     a) Interference cancellation*:*
     $\boldsymbol{y}_m^{(i)} = \boldsymbol{y}_m^{(i-1)} - \boldsymbol{B}_{k'} \hat{\boldsymbol{h}}_{k'} \hat{b}_{k',m}$, for $m = 1, 2, \cdots, N_F$;

     b) *Return 1).*

   Outputs: Active UEs and their information bits.

In the above Section 4.5.3, data detection is carried out after both channel estimation and UAI are completed. However, the channel estimation addressed in Section 4.3 experiences interference among the active UEs. If the interference from stronger UEs can be cancelled before the channel estimation of relatively weaker UEs, the overall performance of MRA/MC-mGFMA systems should be improved. In this section, we propose a joint channel estimation, UAI and data detection algorithm that is designed based on the principles of MMSE and SIC, which is referred to as the *SIC-MMSE-JCUD* for convenience of description, to improve the reliability of MRA/MC-mGFMA systems. The algorithm is stated as Algorithm 4.6, which is explained as follows.

First, according to the number K of potential UEs and the activation probability P_a, a number K_M is determined so that the probability of more than K_M active UEs does not exceed a small probability, such as, 10^{-5}. The algorithm terminates either before or when the number of iterations reach this number.

As shown in Algorithm 4.6, during the ith iteration, meaning that there are $(i-1)$ UEs identified and detected, the algorithm first estimates the channels of the $(K-i+1)$ UEs, as shown at Step 1). Then, at Step 2), the most reliable UE is identified as the UE whose channel has the largest magnitude, by following the principle in Section 4.4.1. For this most reliable UE, as shown at Step 3), its weight vector for the detections of pilots

Table 4.1: Complexity of UAI schemes in Chapter 4.

UAI scheme	Complexity	Characteristic	Advantages	Disadvantages
TB-UAI	KC_e	Designed based on the distinct properties of the estimated channels for active and inactive UEs.	Identify directly.	
JCUD	$\mathcal{O}\left(c_1 K_A N + c_2 N^2\right)$	According to the number K of potential UEs and the activation probability P_a, a number K_M is determined so that the probability of more than K_M active UEs does not exceed a small probability, such as, 10^{-5}. The algorithm terminates either before or when the number of iterations reach this number.	High reliability.	Only one UE can be identify in each of the iteration.

and data is computed, and furthermore, the auto-correlation matrix is updated to $\boldsymbol{R}_b^{(i)}$ for use in the next iteration. Then, at Step 4), the pilots and data bits of the most reliable UE are first detected. Furthermore, the number of erroneous pilot symbols is compared with a pre-set number N_C. If the number of erroneous pilot symbols exceeds N_C, meaning that the identified UE is most probably an inactive UE or its channel is too poor to be detected, the algorithm terminates and outputs the active UEs as well as their data bits. Otherwise, if the number of erroneous pilot symbols is smaller than N_C, meaning that the identified UE is active, $\boldsymbol{y}_m^{(i-1)}$ is then updated to $\boldsymbol{y}_m^{(i)}$ for $m = 1, 2, \ldots, N_F$ by removing the contribution of the detected UE. Then, the algorithm returns to Step 1) to start the next iteration unless K_M iterations are operated.

Note that, Algorithm 4.6 assumes that each iteration handles only one UE, which results in that the processing delay is proportional to the product of the number of active UEs and the frame length of N_F. In order to shorten the processing delay, several *most reliable* UEs may be simultaneously processed during one iteration. Without any doubt, the operation in this way results in the trade-off between the processing delay and the performance. An improved approach is to use the adaptive approach, with which the UEs having their channel magnitudes higher than a pre-set threshold are simultaneously processed. Otherwise, when all the channel magnitudes are lower than a pre-set threshold, only the UE having the highest channel magnitude is processed.

4.7 Performance of MRA/MC-mGFMA Systems

In this section, we show the performance of the dynamic MRA/MC-mGFMA systems employing the SIC-MMSE detection provided in Section 4.5 or the SIC-MMSE-JCUD considered in Section 4.6. Let us first consider the dynamic MRA/MC-mGFMA systems employing the SIC-MMSE detection.

First in Fig. 4.13, we study the BER performance of the MRA/MC-mGFMA systems

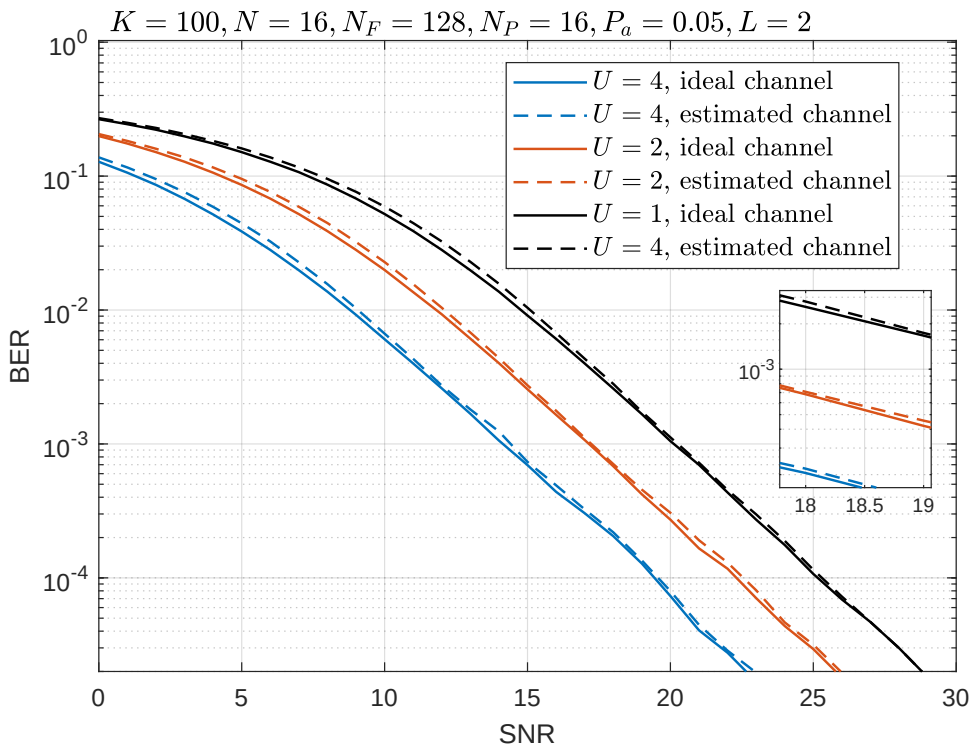


Figure 4.13: BER performance of MRA/MC-mGFMA systems employing SIC-MMSE detection.

employing the SIC-MMSE detection, when assuming that AP employs either ideal channel knowledge or the channels are estimated based on the pilots sent by UEs. As shown on the top of the figure, we assume that the number of potential UEs is $K = 100$ and each UE has an activation probability of $P_a = 0.05$. For channel estimation, 16 pilots are transmitted over a frame of length 128. Additionally, the spreading factor is set to $N = 16$ and the number of multipaths is $L = 2$. From the results of Fig. 4.13 we can clearly observe that the BER performance of MRA/MC-mGFMA systems improves with the increase of the number of receive antennas deployed at AP. Furthermore, for the considered systems and settings, the BER performance achieved by the estimated channels is near that obtained by the ideal channels, explaining that the channel estimation is efficient. This is more declared, when the SNR is sufficiently high, as the result that more reliable channel estimation can be attained at higher SNR.

In Fig. 4.14, we investigate the effect of pilot density on the BER performance. For this purpose, given the total number of pilots of $N_P = 16$, we assume two frame lengths of $N_F = 128$ and $N_F = 1024$. Hence, the pilot ratios are $16/128 = 12.5\%$ and $16/1024 = 1.6\%$, respectively. The results show that the BER performance in both cases is very similar. This is mainly because our channels are estimated in the time-domain, which is mainly depended on the number of pilot symbols and the number of subcarriers. The same is the observation from the results shown in Fig. 4.13, increasing the number of receive

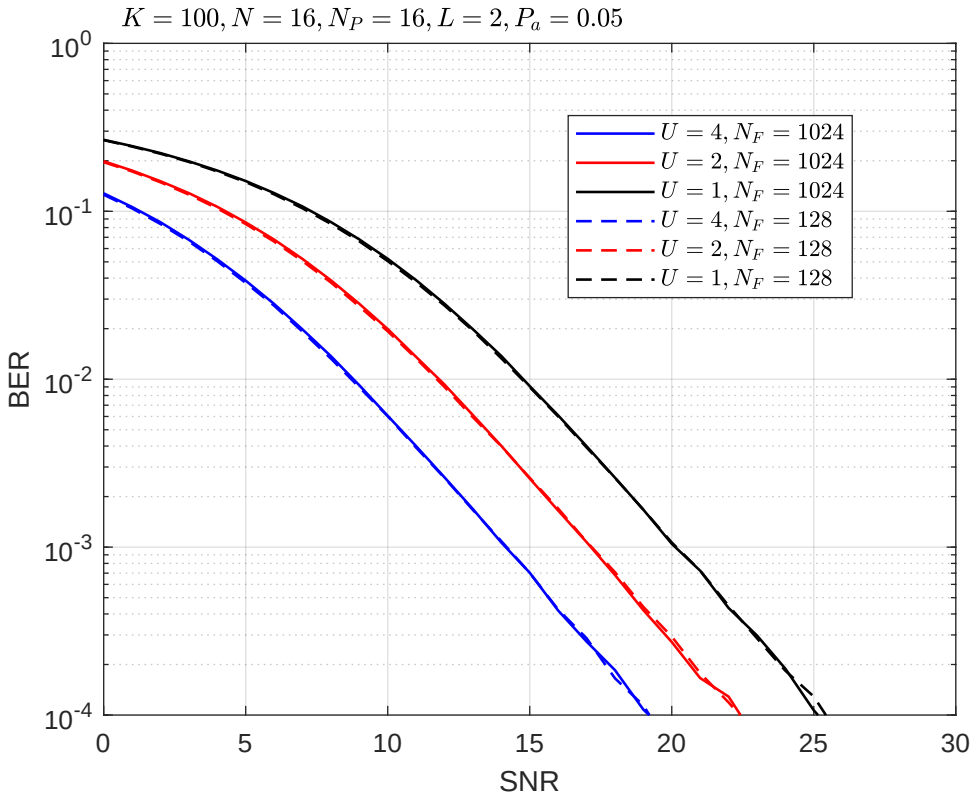


Figure 4.14: BER performance of MRA/MC-mGFMA systems employing SIC-MMSE detection.

antennas of AP can significantly improve the BER performance of the MRA/MC-mGFMA systems.

In Fig. 4.15, we demonstrate the BER performance of the MRA/MC-mGFMA systems employing the SIC-MMSE detection, when channels are estimated using the approaches in Section 3.3. As shown in the figure, we assume that the activation probability of each potential UE is $P_a = 0.05$ or $P_a = 0.15$. From the results we can have the following observations. First, when more receive antennas are employed, the MRA/MC-mGFMA is capable of attaining better BER performance. Second, when $P_a = 0.15$, it can be observed that the BER performance curves present error floors, which can be explained as follows. As the number of multipaths is $L = 2$ and the number of potential UEs is $K = 100$. The average variables to be estimated is $100 \times 0.15 \times 2 = 30$, which is nearly twice of the spreading factor $N = 16$. Furthermore, since each UE activates independently, there is a big probability that the number of active UEs is larger than 15. In these cases, the channel estimation becomes not very reliable, which results in low reliability of detection. By contrast, when $P_a = 0.05$, we can see that the BER decreases as expected.

Fig. 4.16 shows the achievable BER performance of the MRA/MC-mGFMA systems employing the proposed signature codes by comparing it with the systems employing ran-

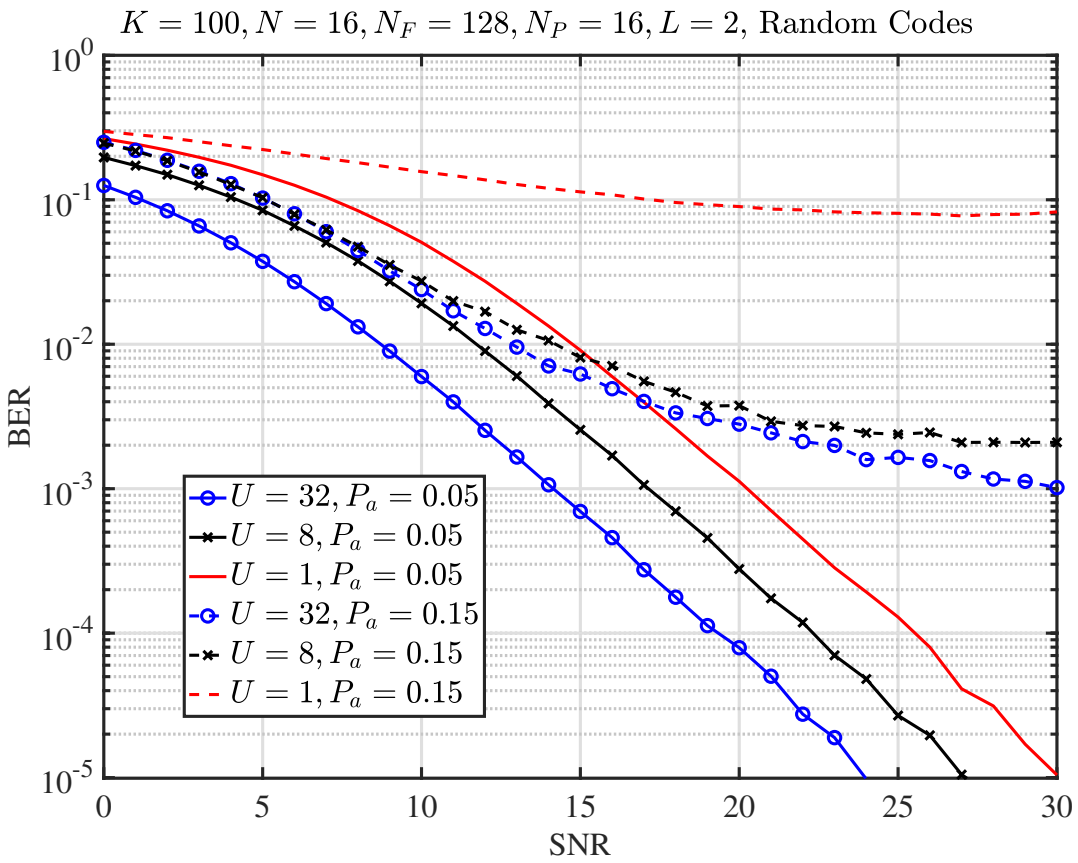


Figure 4.15: BER performance of MRA/MC-mGFMA systems employing SIC-MMSE detection.

dom signature codes. Again, the SIC-MMSE detection is employed and the channels are assumed to be ideal. The results show that the proposed set of signature codes outperform the random signature codes, and much better BER performance can be achieved in both the cases of $L = 2$ and $L = 8$. Furthermore, the BER performance in the case of $L = 8$ is better than that in the case of $L = 2$. This is because under the assumption of ideal channel knowledge, the case of $L = 8$ can provide the detection with more frequency diversity than the case of $L = 2$.

Fig. 4.17 also compares the BER performance achieved by the proposed signature codes and the random codes. We assume the MRA/MC-mGFMA systems employing the SIC-MMSE detection, and when the channels are estimated using $N_P = 32$ pilot symbols in the frames of length $N_F = 128$. Explicitly, the proposed signature codes allow the MRA/MC-mGFMA systems to achieve better BER performance than the random signature codes. Furthermore, deploying more receive antennas at AP is capable of attaining better BER performance.

Fig. 4.18 compares the BER versus UE activity performance of the MRA/MC-mGFMA systems employing the conventional MMSE detection and the SIC-MMSE detection for

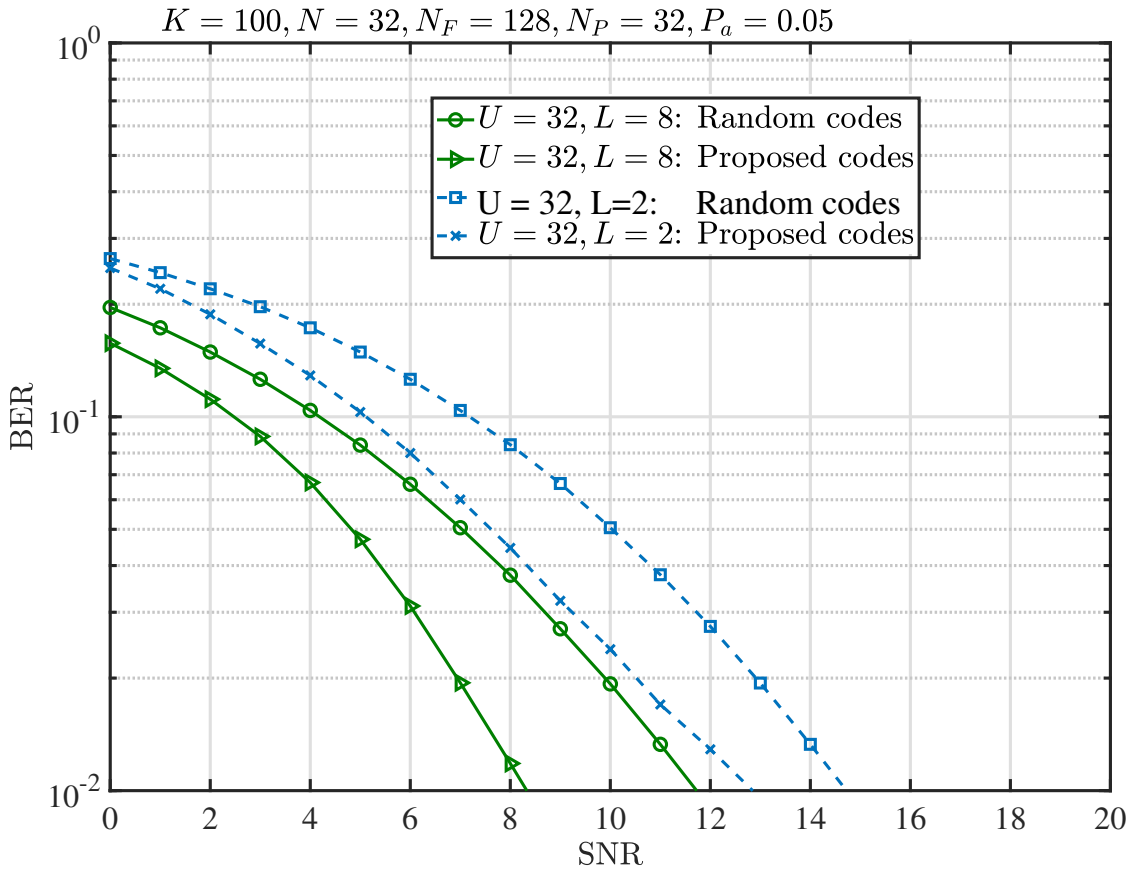


Figure 4.16: BER performance of MRA/MC-mGFMA systems employ SIC-MMSE detection when using random codes and proposed codes respectively.

two different SNR values, namely $SNR = 15$ dB and 20 dB. Note that, the channel knowledge required for the conventional MMSE and SIC-MMSE detection is acquired using the channel estimation provided in Section 4.3.2. From the results shown in Fig. 4.18 we can know that while the SIC-MMSE detector is capable of achieving very promising BER performance even when only 8 pilots are used for channel estimation, the conventional MMSE detector's performance is quite poor. The BER performance achieved by the conventional MMSE detector increases quickly with the increase of the UE's activation probability. These observations explain that the conventional MMSE detector is not efficient for operation in the dynamic GFMA systems. By contrast, with the aid of the embedded SIC, the SIC-MMSE detector is robust to the dynamics of GFMA systems, making them achieve promising BER performance.

Finally, the set of results shown in Figs. 4.19 and 4.20 are for the SIC-MMSE-JCUD described by Algorithm-4.6. Specifically, in Figs. 4.19a and 4.19b, the probabilities of false-alarm and miss are respectively depicted, while the BER of the identified active UEs is shown in Fig. 4.20. From the results shown in Figs. 4.19a and 4.19b, we may have the following observations. First, for a given U of the number of receive antennas at AP

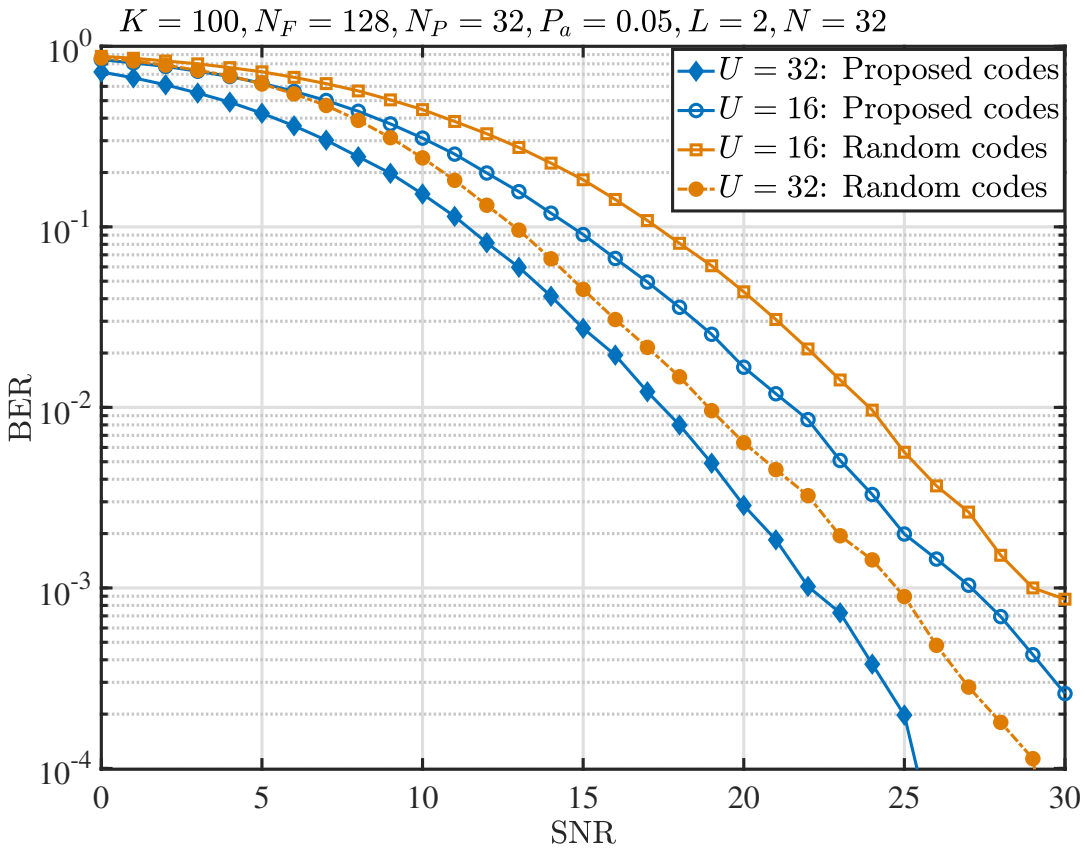


Figure 4.17: SIC Detection for operation of proposed MRA/MC-mGFMA.

and at a given SNR, when N_c is changed from $N_c = 2$ to $N_c = 4$, as shown in the figures, the false-alarm probability increases while the miss probability decreases. This is understandable because N_c is the number of erroneously detected pilots allowed, the false-alarm probability should increase and the miss probability should decrease, when N_c is a bigger value. Second, for a given value of N_c and at a given SNR, both the false-alarm probability and the miss probability decreases, as more receive antennas are employed by AP. This observation follows what we attained in the previous of results in the section. Again, the reason behind is that employing more receive antennas at AP is beneficial to the distinction between active and inactive UEs.

In Fig. 4.20 showing the BER of the active UEs in the MRA/MC-mGFMA systems employing SIC-MMSE-JCUD algorithm, we consider only the case of $N_c = 4$. Explicitly, the BER performance improves, as the number of receive antennas at AP increases. From the parameters shown on the top of the figure, we can know that the average number of active UEs per time-slot is 10, and there is still a certain probability that the number of active UEs per time-slot is larger than $N = 16$. However, from the BER performance results, we can be implied that the SIC-MMSE-JCUD works efficiently, as there is no error-floor observed, even when $U = 1$, i.e., when AP has only one receive antenna. Therefore,

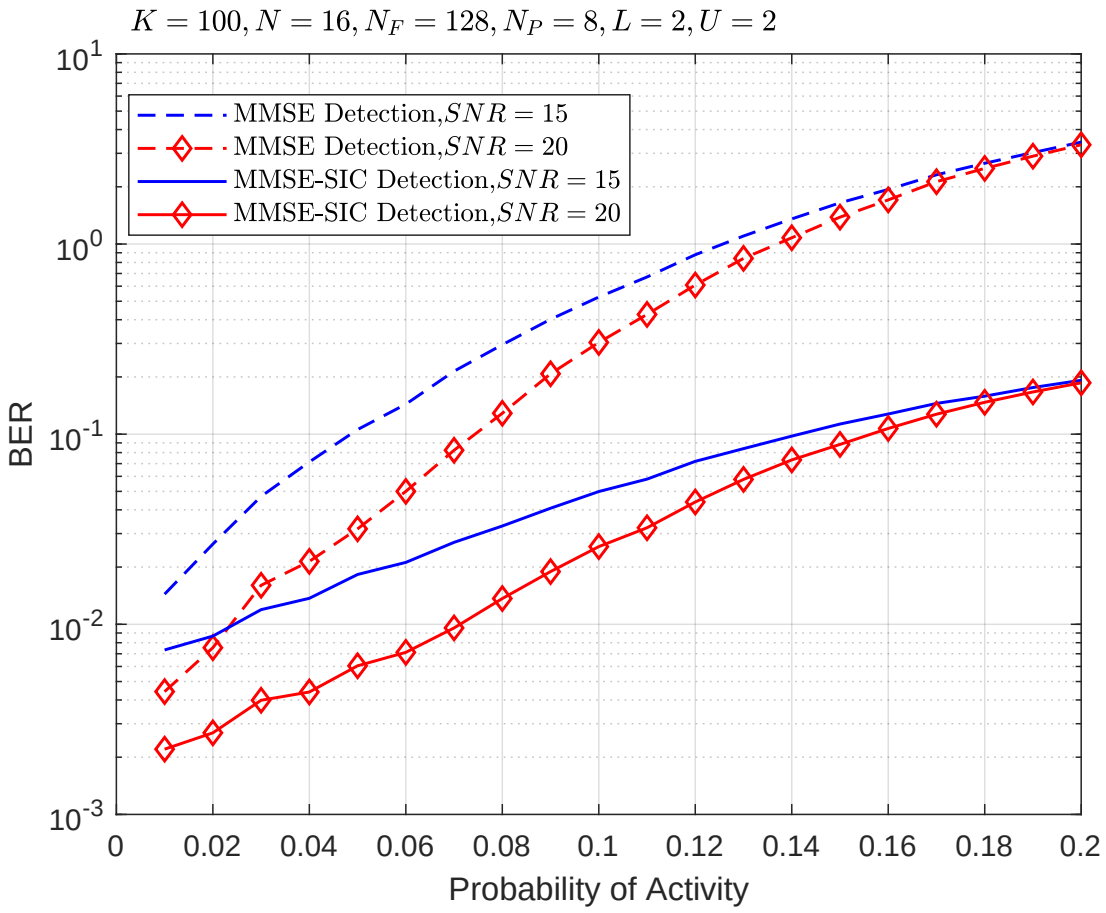
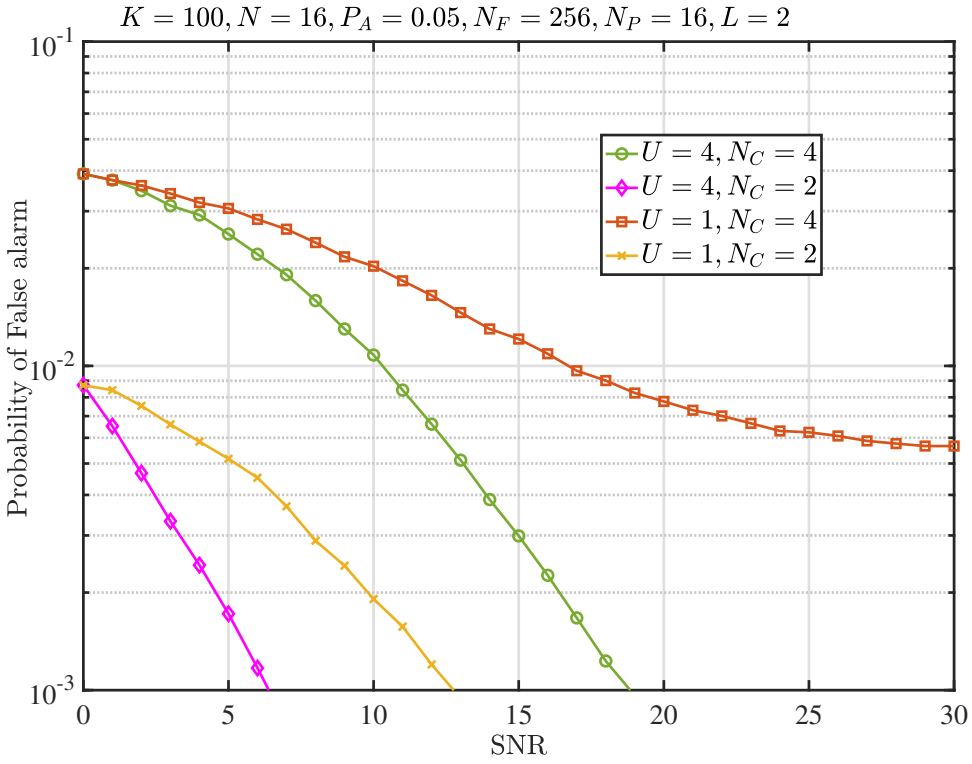


Figure 4.18: BER versus UE activation probability of the MRA/MC-mGFMA systems employing respectively the SIC-MMSE and conventional MMSE detection.

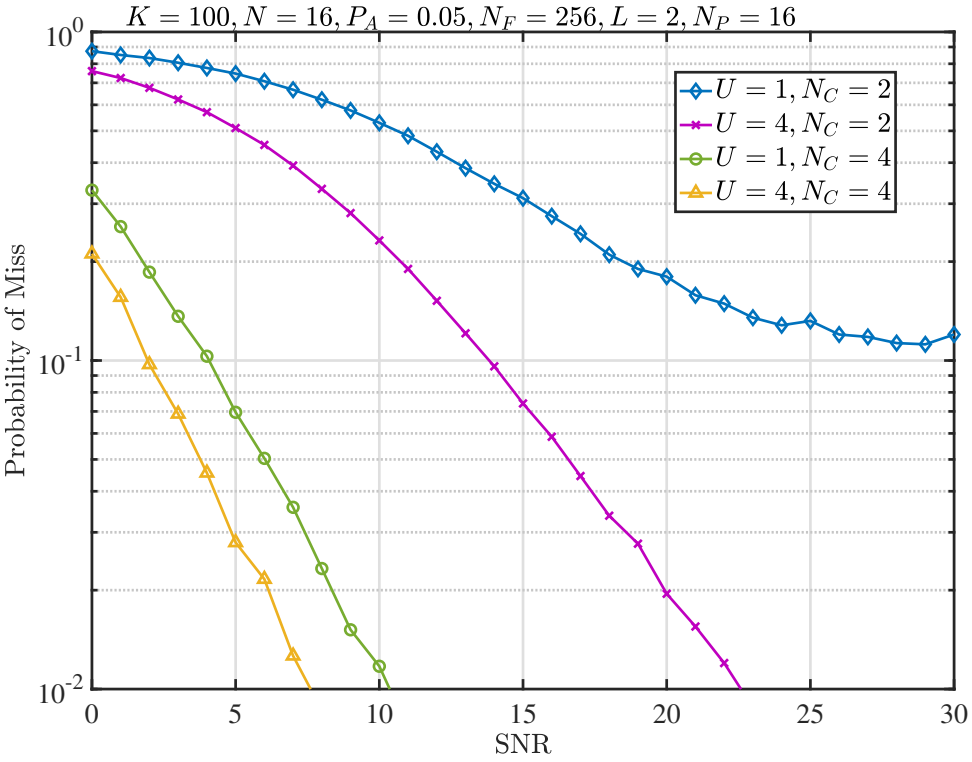
from the results of Figs. 4.19 and 4.20 we can know that the SIC-MMSE-JCUD is a highly efficient algorithm, where channel estimation, UAI and data detection can enhance each other.

4.8 Chapter Conclusions

We have investigated a MRA/MC-mGFMA system, where each AP has multiple receive antennas, for a massive number of UEs to access wireless network without requiring grant scheduling. First, the channel estimation in the principle of MMSE has been considered, when AP is assumed to have ideal knowledge, no knowledge or partial knowledge about the active UEs. Then, based on channel estimation, a low-complexity TB-UAI scheme has been proposed for UAI. Finally, AP carries out information detection of active UEs in the principles of SIC-MMSE, after the channel estimation and UAI. Furthermore, a joint channel estimation, UAI and information detection that is relied on the SIC-MMSE, referred to as the SIC-MMSE-JCUD, has been proposed and studied. We have demonstrated the im-



(a) Prob. of false-alarm



(b) Prob. of miss

Figure 4.19: Probabilities of false alarm and miss, when UEs are identified using the SICD-UAI algorithm.

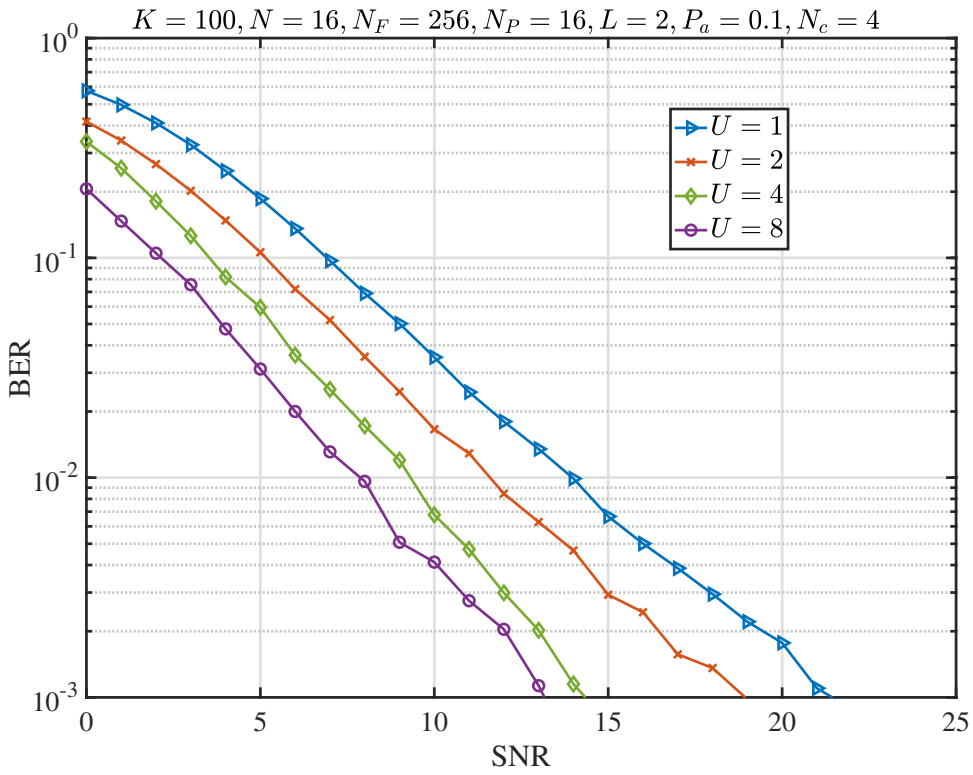


Figure 4.20: BER of active UEs in the MRA/MC-mGFMA systems employing SIC-MMSE-JCUD.

pact of multiple receive antennas on the system design and achievable performance, showing that significant performance improvement is available for channel estimation, UAI and information detection, when AP is deployed with multiple receive antennas. Furthermore, when AP is equipped with multiple receive antennas, a low-complexity TB-UAI is highly effective for achieving promising performance, and the associated threshold is robust for practical setting. Specifically for information detection, in this chapter, two detection schemes have been considered, which are the SIC-MMSE and SIC-MMSE-JCUD, both of them rely on the SIC-MMSE principle. Our studies and performance results show that these SIC-MMSE-relied detection schemes are highly efficient for signal detection in mGFMA systems. It is shown that their performance is much better than the benchmark MMSE detector. More importantly, they are flexible for operation in the mGFMA systems where active UEs and the number of them are highly dynamic. Specifically in the SIC-MMSE-JCUD, channel estimation, UAI and data detection can iteratively enhance each other to attain the promising performance. Additionally, by combining the ZC-sequences and Gold-sequences, we have proposed a class of signature sequences for supporting the massive number of UEs. Our studies demonstrate that the performance achieved by this class of sequences is better than that attainable by binary random sequences.

Chapter 5

Channel Estimation, User Activity Identification and Data Detection in Distributed Antenna mGFMA Systems

Distributed MIMO can significantly improve system performance, as it can bring BS antennas close communication terminals, such as, UEs, making wireless communication user centric. Following our studies in the previous chapters, in this chapter, we study the massive distributed grant-free multiple-access (MDR-GFMA) with respect to the channel estimation, UAI and data detection. We assume a MDR-GFMA system where remote radio heads (RRHs) or APs, or simply distributed antennas (DAs) are randomly distributed in a given area based on the point Poisson (PP) distribution, while UEs are uniformly distributed. We assume that signals transmitted by UEs experience both the large-scale fading of propagation path-loss and shadowing and the small-scale Rayleigh fading. Signals received by different RRHs/APs are forwarded to a so-called signal processing central unit (SPCU), where channel estimation, UAI and data detection are carried out. In terms of signal processing at SPCU, channel estimation is achieved in the principle of minimum mean-square error (MMSE). Following channel estimation, an orthogonal matching pursuit (OMP) relied algorithm is implemented to attain initial UAI, which is enhanced with the aid of the pilot detection of each initially identified active UE. Finally, the data sent by active UEs are detected using either MMSE detection or the successive interference cancellation assisted MMSE (MMSE-SIC) detection. The performance of the above-mentioned schemes is studied with the aid of Monte-Carlo simulations. Our studies conceive that the proposed algorithms are effective, and are capable of achieve promising performance in the MDR-GFMA systems with various dynamics, including active UEs and the number of them, locations of DAs and the number of them serving different UEs, geographically

resulted large-scale fading of propagation path-loss and shadowing.

5.1 Introduction

The rapid growth of applications imposes the challenges on the 5G beyond (5G+) and the 6G wireless systems to have higher throughput, wider coverage, lower latency, further enhanced reliability, etc., than what the 5G systems can provide [1]. To meet these expectations, numerous novel techniques have been proposed and investigated, which allow to flexibly and collaboratively to work for various application scenarios, such as, hotspots [222], in order to attain high-efficiency. Furthermore, the concept of integration architectures for future wireless networks has been proposed in 3GPP [223], which motivates to reduce the propagation delay, so as to establish the new radio (NR) massive sensing networks friendly. Additionally, in terms of the resource allocation in future wireless networks, reduced cell size and adaptive coordination have been studied in a collaborative manner among different types of remote radio heads (RRHs) or access points (APs). Correspondingly, the user-centric heterogeneous network [224] has been expected to be implemented for providing services in the internet of things (IoT) applications.

In this chapter, we consider the GFMA in the distributed RRUs/APs systems, where the RRHs/APs are randomly distributed, such as following a homogeneous Poisson Point Process (PPP), in the system's coverage and are connected via feeder cables/links to a signal processing centre unit (SPCU). Hence, the RRHs/APs can work cooperatively for different purposes, such as for signal transmission/receiving on downlink/uplink. Furthermore, in the considered system, we assume that there are potentially a massive number of user equipments (UEs), but each of which only has a small probability to become active to transmit at a time. Therefore, we refer to the considered system as the massive distributed antenna GFMA (MRD-GFMA) system for convenience of description. In this chapter, our research focus is on the uplink transmission, with the emphasis on the channel estimation, UE activity identification (UAI) and the data detection of active UEs.

Fundamentally, MDA systems belong to the family of the distributed antenna (DAS) systems [225]. Along with the research and development of MIMO systems [226–228], cell-free (CF) systems [229–232], ets., MDA systems have attracted wide attention, owing to its potential to provide the technical support for the user-centric systems where every UE can be served by a few of RRHs/APs around it. Hence, in a user-centric system, each service is centralized by a UE associated with a predominately formed cell including the RRHs/APs supporting the communications of the UE. In literature, this approach has been studied in the context of the open radio access network (O-RAN) [233], cloud radio access

network (C-RAN) [234], large-scale DAS [235] and cell-free distributed MIMO (CF-D-MIMO) [236], etc., with the motivation to meet the challenges of 5G+ and 6G systems. In brief, MDA systems can have the following advantages. First, owing to the massive number of antennas distributed in the network, MDA systems can usually have extremely high spectral-efficiency. Second, as each UE only communicates with the RRHs/APs close to it, MDA systems can be highly energy efficient. Third, owing to the fact that each UE has some RRHs/APs close to it and at these RRHs/UEs, this UE is a strong UE, no power-control is needed in terms of the near-far problem. Furthermore, handover in MDA systems is relatively easy to implement, as SPCU manages the communication between a UE and the RRHs/APs close to it. The SPCU can readily associate/de-associate a RRH/AP with/from a UE.

On the other hand, various GFMA schemes suitable for short and sporadic traffic have been proposed [237]. However, as there is no handshaking between UE and BS/AP, signal detection in GFMA systems is more challenging. The BS/AP has to first carry out the user activity identification (UAI) to know which UEs are active before it execute the data detection, or carry out joint UAI and data detection [238]. In [116], the authors enhanced the CS-based approach by proposing a greedy algorithm relying on the maximum a posteriori (MAP) criterion, which performs UAI and data detection jointly by exploiting the a-posteriori probabilities of each other. In [115], the authors introduced a two-stage cooperative MUD designed for the uplink (UL) GFMA systems. In this method, the m-sequence in time-domain and the training sequence in frequency-domain are utilized respectively to achieve synchronization and gain the priori information about UEs' activities. The authors in [117] considered the detection scenario where the sparsity is time-varying and correspondingly, a switching algorithm relying on the sparsity level was proposed so as to enjoy the merits of both the CS-based and conventional detection schemes. Similarly, by considering the time-varying characteristics of sparsity in GFMA systems, the dynamic low-complexity CS-based MUD was studied in [118, 119]. With this method, the active UE set in the current time-slot is used to estimate the active UEs in the following time-slot. It was shown that the method can improve the successful data recovery rate. Moreover, in [239], a block-sparsity model was proposed and based on which the cross-validation aided block sparsity adaptive subspace pursuit (CVA-BSASP) algorithm and the threshold-aided block sparsity adaptive subspace pursuit (TA-BSASP) algorithm were introduced to detect the information in GFMA systems. Very recently, the machine learning methods have been conceived more frequently for UAI and information detection in GFMA systems [162, 240, 241]. Specifically, in [162], the authors considered the deep neural network (DNN) algorithms for detecting the active UEs in GFMA systems. In [241], the authors proposed a method to jointly carry out the sparse signal recovery

and sparse support recovery under the multiple measurement vector (MMV) models for GFMA-MIMO systems. Furthermore, in [242], the authors considered an uplink grant-free sparse coded multiple access (GF-SCMA) system and correspondingly, a generalized likelihood ratio based convolutional neural network (CNN) algorithm was proposed for UAI.

From the above studies we learned that most CS-based approaches are efficient, when the sparsity level is low, i.e., satisfying the RIP [243]. However, in practical massive GFMA systems, the number of UEs may be huge, UEs may be distributed with ultra-high density, while the active UEs and the number of them are highly dynamic. Therefore, the RIP for sparsity may not be satisfied, making the CS-based approached inefficient. Furthermore, in this chapter, we consider the MDR-GFMA, where RRHs/APs are distributed in the communications area. Each RRH/AP only serves the UEs around it, while each UE is only connected with a few of RRHs/APs around it. Furthermore, the signals received by an RRH/AP from different UEs have different power. All the above-mentioned settings/assumptions impose new challenges to the design of the algorithms for the UAI and data detection in MDR-GFMA systems. Hence, in this chapter, we motivate to study the low-complexity while still efficient methods for the channel estimation, UAI and data detection in the MDR-GFMA systems. The main contributions and novelties can be summarized as follows:

- A MDA-GFMA system is proposed and investigated, where the RRHs/APs (or simply distributed antennas (DAs)) are distributed based on the Poisson point (PP) distribution and UEs are randomly and uniformly distributed in the coverage area. In the considered MDA-GFMA system, RRHs/APs are assumed to be connected to a SPCU, where signal processing is carried out. UEs become active randomly with a small activation probability. Each of the active UEs is only connected to a few of the DAs close to the UE. Hence, no path-loss resulted power-control is needed, the operation is UE centric and the system is cell-free.

- In terms of signal processing, first, the MMSE-assisted channel estimation is implemented to estimate the channels of UEs, regardless of them being active or inactive, when assuming that the SPCU has the different levels of knowledge about the large-scale fading, including propagation path-loss and shadowing, experienced by the UEs. Furthermore, the characteristics of the estimated channels for active and inactive UEs are analyzed and demonstrated.

- Second, based on the MMSE-relied channel estimation, an orthogonal matching pursuit assisted UAI, referred to as the OMP-MMSE-UAI, algorithm is proposed and investigated, when each of UEs is assumed to be served by the RRHs/APs within a

given area centred at the UE. The OMP-MMSE-UAI provides the initial UAI, which is further enhanced via the detection of the pilot symbols expected for different UEs.

- Third, after the channel estimation and UAI, the data transmitted by active UEs are detected. Both MMSE detection and the successive interference cancellation assisted MMSE (MMSE-SIC) detection are introduced, again, by assuming that each active UE is served by a few of RRHs/APs around the UE.

- The performance of channel estimation, UAI and data detection in MDR-GFMA systems is investigated with the aid of Monte-Carlo simulations. The results demonstrate that the proposed algorithms are effective. The MMSE-based channel estimation is low-complexity and is capable of achieving promising performance, which significantly improves with the increase of the density of the distributed RRHs/APs. The OMP-MMSE-UAI algorithm can effectively limit the miss of active UEs, while avoiding the false-alarms of inactive UEs. Furthermore, the MMSE-SIC is highly efficient for operation in the MDR-GFMA systems, where different dynamics exist, including active UEs and the number of them, the number of RRHs/APs that a UE connects, and the geographically resulted large-scale fading.

The remainder of this chapter is outlined as follows. In Section 5.2, we describe the MDA-GFMA system model and define the main system parameters. Section 5.3 addresses the MMSE-assisted channel estimation. UAI is considered in Section 5.4, while data detection is addressed in Section 5.5. Performance results and discussion are in Section 5.6. Finally, the main conclusions derived from the research are summarized in Section 5.7.

5.2 Description of Massive Distributed Antenna GFMA System

We consider a massive distributed antenna GFMA (MDA-GFMA) system, as shown in Fig. 5.1 in which many remote radio heads (RRHs) and APs are distributed. Note here that RRHs are just the extensions of APs for collecting/sending signals from different locations. These RRHs/APs are assumed to be connected using optical fiber with a signal processing centra unit (SPCU), which is also referred to as a baseband unit (BBU). Each RRH/AP is assumed to have one antenna for signal receiving. Hence, in the forthcoming discourses, we may exchangeably use the terms of distributed antenna (DA) and RRH/AP without notation. We assume that the RRHs/APs are assumed to be well synchronized. Furthermore, to improve connectivity, we assume a user-centric scenario where the density of distributed antennas is high. In this MDA-GFMA system, each antenna may support all

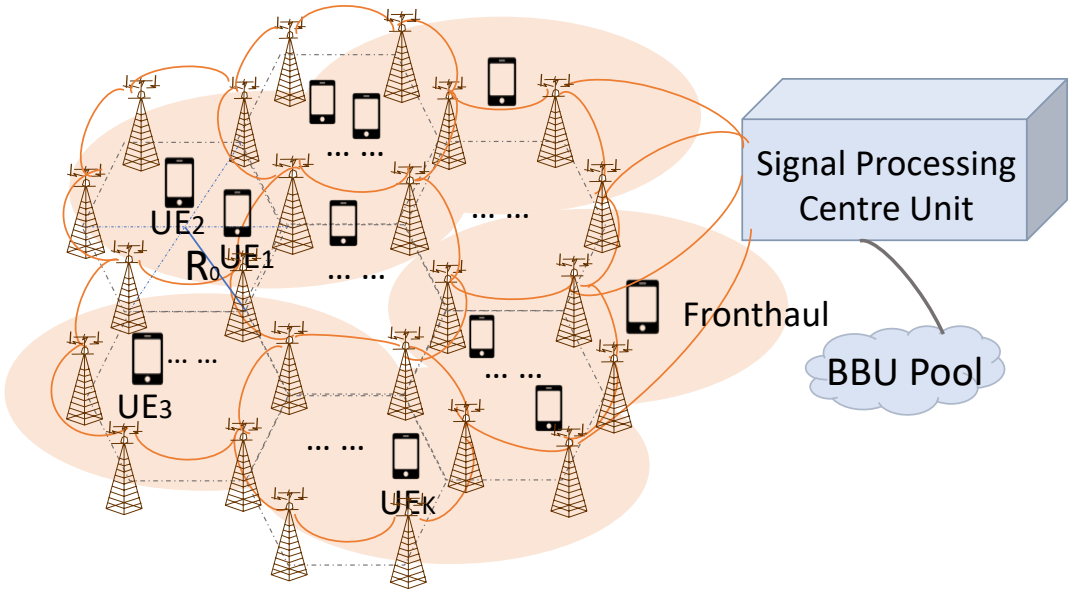


Figure 5.1: Illustration of system structure for MDA-GFMA system.

the UEs in the system or the part of UEs around it. UEs are assumed to be randomly (uniformly) distributed in the system.

Below we describe the MDA-GFMA system in terms of its transmitter, channel model and receiver.

5.2.1 Transmission Modelling

As in the previous chapter, we assume that in the MDA-GFMA system, there are K potential UEs randomly accessing the network. Each of them has a small probability P_a to become active to transmit its information at a time-slot. Hence, during a time-slot, the number of active UEs, expressed as K_A, is a random variable with its average satisfying $\bar{K}_A \ll K$. Once a UE becomes active, it starts transmitting information at the start of the next time-slot, one frame per time-slot. Since the IoT devices having low-rate requirement are concerned, we assume that the low-complexity binary phase-shifting keying (BPSK) based modulation is employed for UEs to send their data. Hence, the transmitted signal by a UE can be written as

$$s_k(t) = \sqrt{2P_t} b_k(t) \cos(2\pi f_c t + \phi_k) \tag{5.1}$$

where P_t is the transmit power, f_c is carrier frequency and ϕ_k is the initial phase introduced by carrier modulation. Additionally, $b_k(t) = \sum_f b_k^{(f)} \psi_{T_b}(t - fT_b)$ is the data waveform, with T_b being the bit (symbol) duration, $\psi_{T_b}(t)$ the transmitted waveform, and $b_k^{(f)} \in \{+1, -1\}$.

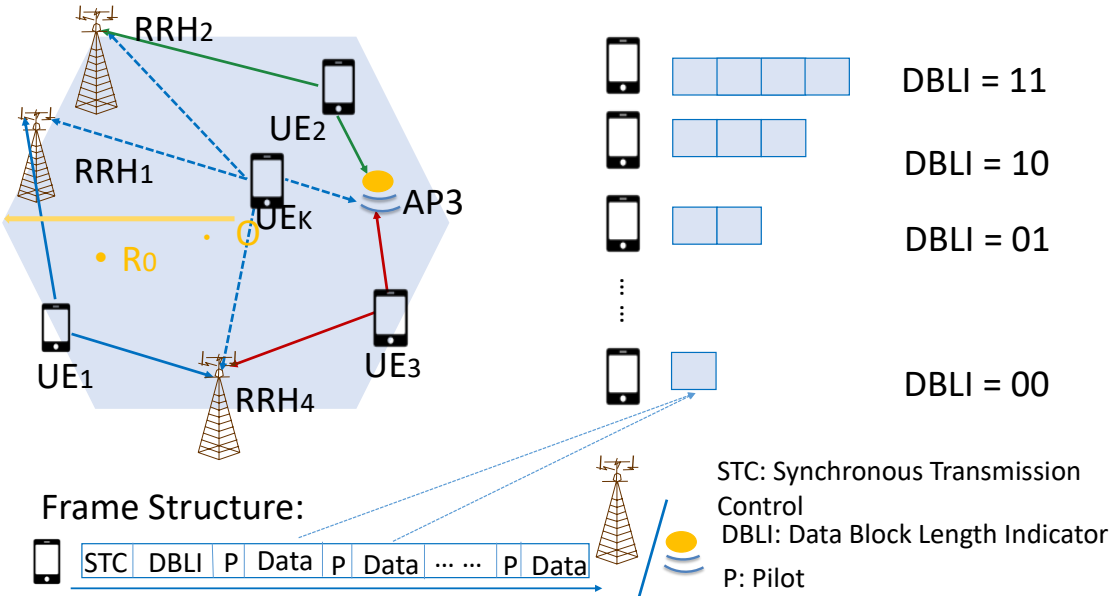


Figure 5.2: Illustration of an example for UEs' actions based on MDA-GFMA principles.

Note that BPSK can be readily extended to the other quadrature amplitude modulation (QAM), when higher data rate is demanded.

As above-mentioned, data of each UE is sent in frame and one frame per time-slot. We assume that each time-slot has relatively short duration, so that the channel of a UE over one time-slot is nearly constant. Specifically, the frame is structured as shown in Fig. 5.2, which has the length expressed as $N_F = N_I + N_P + N_L$, where N_I denotes the number of data bits per frame, N_P is the number of pilot bits per frame, inserted for channel estimation and UE activity identification (UAI), while the N_L bits are used to inform SPCU the number of frames that a UE transmits continuously. Hence, these bits are marked as the data block length indicator (DBLI), as shown in Fig. 5.2. Note that, in our performance study, we typically set $N_L = 2$, meaning that at most 4 frames can be transmitted once a UE activates.

In Fig. 5.2, we also illustrate a specific example with active UEs being associated with different RRHs/APs. Due to the propagation path-loss and a UE having different distances from different RRHs/APs, as shown in Fig. 5.2, at a certain time, UE_1's transmitted signals may only be received by RRH_1 and RRH_4, UE_2's transmitted signals are only received by RRH_2 and AP_3, UE_3's transmitted signals are only received by AP_3 and RRH_4, while UE_k's transmitted signals are simultaneously received by RRH_1, RRH_2, RRH_4 and AP_3. Certainly, before UAI, a RRH/AP is unable to know from which UE it receives signals. Instead, after the RRHs/APs receive the signals of a frame, they forward their received signals to the SPCU (see Fig. 5.1), where channel estimation, UAI and data detection

are operated. Furthermore, from Fig. 5.2 we can be implied that when all the observations sampled from the RRHs/APs are considered, the channel matrix (or access matrix) is sparse, as each UE is only associated with a small number of RRHs/APs. Let us now describe the channel model in the next subsection.

5.2.2 Channel Modelling

Since we assume that in the considered MDA-GFMA system, antennas are distributed with relatively high density, it can be convinced that at any location of a UE, there are always one to several antennas around the UE, which receive the relatively high power from this UE. Hence, in the chapter, we do not assume power control. Instead, we assume that all UEs transmit at the same power expressed as P_t. Furthermore, for measurement and performance illustration, we define the signal-to-noise ratio at transmit, expressed as SNR_t, which is in fact the transmit power normalized by the noise's power of σ^2, i.e., $\text{SNR}_t = P_t/\sigma^2$.

In this chapter, we assume a composite shadowing fading propagation model [244], which includes the large-scale propagation path-loss and shadowing as well as the small-scale fast fading. Specifically, when we assume a distance d between a UE and an antenna at RRH/AP, the path-loss expressed in dB can be expressed as

$$\Gamma(d) = 10\alpha \log_{10}(d/d_0)\ (dB),\ d \geq d_0 \tag{5.2}$$

where α is the path-loss exponent usually taking a value in the range between 2 and 6, while d_0 is a reference distance, below which the path-loss is constant.

Assume that the transmit power is P_t. Then, the average received power P_r from a transmission distance d is given by

$$P_r = \frac{P_t}{10^{\Gamma(d)/10}}. \tag{5.3}$$

Overall, let the channel gain between the kth UE and the mth DA be expressed as $g_{k,m}$. It can be written as

$$g_{k,m} = \sqrt{\alpha_{k,m}} h_{k,m} \tag{5.4}$$

where $h_{k,m}$ represents the small-scale fast fading, which is assumed to follow the complex Gaussian, i.e., Rayleigh fading, distribution with unit power [245], [246], expressed as $\mathcal{CN}(0,1)$, while $\alpha_{k,m}$ denotes the large-scale fading, following the lognormal distribution with the probability density function (PDF) of

$$p_{\alpha_{k,m}}(x) = \frac{\xi}{\sqrt{2\pi}\sigma_\alpha x} \exp\left[-\frac{(10\log_{10} x - \Gamma_{k,m})^2}{2\sigma_\alpha^2}\right] \tag{5.5}$$

where $\xi = 10/\ln 10 = 4.3429$ [245], and $\Gamma_{k,m}$ (dB) and σ_α (dB) represent the mean and the standard deviation of $10\log_{10}\alpha_{k,m}$. Here, the mean of $\Gamma_{k,m}$ is given by the propagation path-loss of (5.2). However, we assume a two-slope path-loss model [247], which can be formulated as

$$\Gamma_{k,m}(d) = \begin{cases} 0, & d_{k,m} < d_0 \\ -10\log_{10}\left[d_{k,m}^{\gamma_1}\left(d_0 + \frac{d_{k,m}}{\delta}\right)^{\gamma_2}\right] + C, & d_{k,m} \geqslant d_0 \end{cases} \tag{5.6}$$

where $d_{k,m}$ is the distance between the kth UE and the mth DA, γ_1 usually taking a value of 2 is the basic path-loss exponent, γ_2 is the additional path-loss exponent, and δ is defined as the break point of the path-loss curve. Additionally, C is added to make sure that $\Gamma_{k,m}(d_{k,m} = d_0) = 0$ dB.

5.2.3 Representation of Received Signals from Distributed Antennas

When the transmitted signal in the form of (5.7) is transmitted over the channel described in Section 5.2.2, the received signal corresponding to one bit (symbol) by the mth DA can be expressed in baseband discrete form as

$$r_m^{(f)} = \sum_{k=1}^{K} I_k g_{k,m} b_{k,f} + n_m^{(f)},\ f = 1,2,\ldots,N_f;\ m = 1,2,\ldots,M \tag{5.7}$$

where $I_k = 1$ or 0 indicates whether UE k is active or inactive, N_f is the frame length in bits, and M is the total number of DAs in the system. In (5.7), $n_m^{(f)}$ is the Gaussian noise, which has zero means and the variance of $1/\mathrm{SNR}_t$ after normalized by the transmit power P_t.

Let us define $\boldsymbol{y}_f = [y_1^{(f)}, y_2^{(f)}, \cdots, y_M^{(f)}]^T$ and $\boldsymbol{n}_f = [n_1^{(f)}, n_2^{(f)}, \cdots, n_M^{(f)}]^T$ for $f = 1,2,\cdots,N_F$. Let the channel gains of a UE to the M DAs be collected into a vector as

$$\boldsymbol{g}_k = [g_{k,1}, g_{k,2}, \cdots, g_{k,M}]^T,\ k = 1,2,\cdots,K \tag{5.8}$$

When explicitly show the large- and small-scale fading components, we can express $\boldsymbol{g}_k$ as

$$\begin{aligned} \boldsymbol{g}_k &= \begin{pmatrix} \sqrt{\alpha_{k,1}}h_{k,1} \\ \sqrt{\alpha_{k,2}}h_{k,2} \\ \vdots \\ \sqrt{\alpha_{k,M}}h_{k,M} \end{pmatrix} = \begin{pmatrix} \sqrt{\alpha_{k,1}} & & & \\ & \sqrt{\alpha_{k,2}} & & 0 \\ & & \ddots & \\ & 0 & & \sqrt{\alpha_{k,M}} \end{pmatrix} \begin{pmatrix} h_{k,1} \\ h_{k,2} \\ \vdots \\ h_{k,M} \end{pmatrix} \\ &= \boldsymbol{\alpha}_k \boldsymbol{h}_k \end{aligned} \tag{5.9}$$

With the above definitions, it can be shown that we have the observation vector corresponding to the fth transmitted bits (symbols) expressed as

$$\boldsymbol{y}_f = \sum_{k=1}^{K} I_k \boldsymbol{g}_k b_{k,f} + \boldsymbol{n}_f,\ f = 1, 2, \cdots, N_F \tag{5.10}$$

Let $\boldsymbol{b}_f = [I_1 b_{1,f}, I_2 b_{2,f}, \cdots, I_k b_{K,f}]^T$. Then, we can write (5.10) in impact form as

$$\boldsymbol{y}_f = \boldsymbol{G}\boldsymbol{b}_f + \boldsymbol{n}_f,\ f = 1, 2, \cdots, N_F \tag{5.11}$$

where $\boldsymbol{G} = [\boldsymbol{g}_1, \boldsymbol{g}_2, \cdots, \boldsymbol{g}_K]$. From (5.11) we can gain the following insights. First, as each UE has only a small probability to become active during a time-slot, recovering $\boldsymbol{b}_f$ is a sparse recovery problem. Hence, the well developed compressive-sensing (CS) algorithms can be employed to achieve UAI and data detection. Second, as each UE is only connected with a small fraction of the total number of DAs in the MDA-GFMA system, $\boldsymbol{G}$ itself is a sparse matrix, which is beneficial to designing the algorithms for channel estimation, UAI and data detection. Below we will exploit some of these merits to design the UAI and data detection algorithms.

5.3 Channel Estimation in MDA-GFMA Systems

In this section we address the channel estimation in the MDA-GFMA system. Following the studies in the previous chapters, we assume that the SPCU either has no knowledge about the active UEs or has partial knowledge about the active UEs. Here the partial knowledge is obtained when an active UE has several frames to transmit after an activation. Again, the channel estimation is in MMSE principle under two assumption scenarios. The first one assumes that SPCU employs the knowledge of the large-scale fading with respect to all UEs. By contrast, the second scenario assumes that SPCU only knows the positions of UEs and, hence, only the propagation path-loss from different UEs. Let us first consider the case when SPCU has no knowledge about active UEs.

5.3.1 SPCU Has No Knowledge about Active UEs

When SPCU has no knowledge about the active UEs, it has to carry out the channel estimation in a symbol-by-symbol way, as analyzed in Chapter 3.3. Hence, following our analysis in Section 3.3.2 of Chapter 3.3, the channel of UE i can be estimated as

$$\hat{\boldsymbol{g}}_i = \frac{1}{N_P} \sum_{p=1}^{N_P} \boldsymbol{W}_{i,p}^H \boldsymbol{y}_p,\ i = 1, 2, \cdots, K \tag{5.12}$$

when without considering the unbiased estimation issue. In (5.12), N_p is the number of pilot symbols, and $\boldsymbol{y}_p$ is given by (5.10) or (5.11) corresponding to the pth pilot symbol. In (5.12), the weight matrix derived in MMSE principle is given by

$$\boldsymbol{W}_{i,p} = \boldsymbol{R}_{\boldsymbol{y}_p}^{-1} \boldsymbol{R}_{\boldsymbol{y}_{p,i}} \tag{5.13}$$

where $\boldsymbol{R}_{\boldsymbol{y}_p}$ is the auto-correlation matrix of $\boldsymbol{y}_p$, which however can be estimated using all the observations of a frame and can be expressed as

$$\boldsymbol{R}_{\boldsymbol{y}_p} = \boldsymbol{R}_a = \frac{1}{N_F} \sum_{f=1}^{N_F} \boldsymbol{y}_f \boldsymbol{y}_f^H \tag{5.14}$$

where $\boldsymbol{y}_f$ is given by (5.10) or (5.11). In (5.13), $\boldsymbol{R}_{\boldsymbol{y}_{p,i}}$ is the cross-correlation matrix between $\boldsymbol{y}_p$ and $\boldsymbol{g}_i$, i.e., we have

$$\begin{aligned} \boldsymbol{R}_{\boldsymbol{y}_{p,i}} =& E[\boldsymbol{y}_p \boldsymbol{g}_i^H] \\ =& E\left[\left(\sum_{k=1}^{K} I_k b_{k,p} \boldsymbol{g}_k + \boldsymbol{n}_p \right) \boldsymbol{g}_i^H \right] \end{aligned} \tag{5.15}$$

Since the objective is to estimate the channel of UE i, SPCU has to assume that UE i is active, we can simplify (5.15) to

$$\boldsymbol{R}_{\boldsymbol{y}_{p,i}} = b_{i,p} E[\boldsymbol{g}_i \cdot \boldsymbol{g}_i^H] \tag{5.16}$$

Below we consider two scenarios to further simplify (5.16).

First, when SPCU has the knowledge about the large-scale fading of all UEs, meaning that SPCU knows $\boldsymbol{\alpha}_k$, $k = 1, 2, \ldots, K$, in (5.9), we can readily derive that

$$\boldsymbol{R}_{\boldsymbol{y}_{p,i}} = b_{i,p} \boldsymbol{A}_i \tag{5.17}$$

where $\boldsymbol{A}_i = \boldsymbol{\alpha}_i^2$.

Second, if SPCU only knows the positions of all the UEs as well as the statistics of the shadowing fading, meaning that SPCU only knows the second moments of $\boldsymbol{g}_i$, we can obtain

$$\begin{aligned} \boldsymbol{R}_{\boldsymbol{y}_{p,i}} =& b_{i,p} E[\boldsymbol{A}_i] \\ =& b_{i,p} \boldsymbol{B}_i \end{aligned} \tag{5.18}$$

where $\boldsymbol{B}_i = E[\boldsymbol{A}_i]$ and is given by

$$\boldsymbol{B}_i = \text{diag}\{10^{(\Gamma_{i,1}+\sigma_{i,1}^2/2)/10}, 10^{(\Gamma_{i,2}+\sigma_{i,2}^2/2)/10}, \cdots, 10^{(\Gamma_{i,M}+\sigma_{i,M}^2/2)/10}\} \tag{5.19}$$

Hence, when SPCU has the knowledge of $\boldsymbol{\alpha}_k$ for all UEs, it can construct the corresponding $\boldsymbol{R}_{\boldsymbol{y}_{p,i}}$ based on (5.17), and consequently obtain $\boldsymbol{W}_{i,p}$, $i = 1, 2, \cdots, K$, based on

(5.13), (5.14) and (5.17). By contrast, if SPCU only knows the positions of UEs and the statistics of shadowing fading, i.e., only knows $\boldsymbol{B}_k$ for all UEs, it can construct the corresponding $\boldsymbol{R}_{\boldsymbol{y}_{p,i}}$ based on (5.18). Then, $\boldsymbol{W}_{i,p}$, $i = 1, 2, \cdots, K$, are formed based on (5.13), (5.14) and (5.18).

In terms of the MSE performance, in the case of that SPCU knows $\boldsymbol{\alpha}_k$ for all UEs, the MSE of the estimation can be derived as

$$MSE_i = ||I_i \boldsymbol{g}_i - \hat{\boldsymbol{g}}_i||^2. \tag{5.20}$$

Hence, when UE i is active, it can be shown that the MSE is

$$\begin{aligned} MSE_i =& Tr(E[\boldsymbol{g}_i \boldsymbol{g}_i^H] - \boldsymbol{R}_{\boldsymbol{y}_{p,i}}^H \boldsymbol{W}_{i,p}) \\ =& Tr(\boldsymbol{A}_i - \boldsymbol{R}_{\boldsymbol{y}_{p,i}}^H \boldsymbol{W}_{i,p}) \qquad (5.21) \\ =& \sum_{m=1}^{M} \alpha_{i,m}^2 - Tr(\boldsymbol{R}_{\boldsymbol{y}_{p,i}}^H \boldsymbol{W}_{i,p}) \qquad (5.22) \end{aligned}$$

where $\boldsymbol{R}_{\boldsymbol{y}_{p,i}}^H \boldsymbol{W}_{i,p}$ can also be written as $\boldsymbol{R}_{\boldsymbol{y}_{p,i}}^H \boldsymbol{R}_{\boldsymbol{y}_p}^{-1} \boldsymbol{R}_{\boldsymbol{y}_{p,i}}$. By contrast, if UE i is inactive, the MSE of estimation is

$$\begin{aligned} MSE_i =& Tr(E[\boldsymbol{W}_{i,p}^H \boldsymbol{y}_p \boldsymbol{y}_p^H \boldsymbol{W}_{i,p}]) \\ =& Tr(\boldsymbol{W}_{i,p}^H \boldsymbol{R}_{\boldsymbol{y}_p} \boldsymbol{W}_{i,p}) \\ =& Tr(\boldsymbol{R}_{\boldsymbol{y}_{p,i}}^H \boldsymbol{R}_{\boldsymbol{y}_p}^{-1} \boldsymbol{R}_{\boldsymbol{y}_{p,i}}). \qquad (5.23) \end{aligned}$$

In the second case when SPCU only knows the positions of UEs and the statistics of shadowing fading, it can be shown that the MSE of estimation for an active UE is

$$\begin{aligned} MSE_i =& Tr(\boldsymbol{B}_i - \boldsymbol{R}_{\boldsymbol{y}_{p,i}}^H \boldsymbol{W}_{i,p}) \\ =& \sum_{m=1}^{M} 10^{(\Gamma_{i,m} + \sigma_{i,m}^2/2)/10} - Tr(\boldsymbol{R}_{\boldsymbol{y}_{p,i}}^H \boldsymbol{W}_{i,p}) \qquad (5.24) \end{aligned}$$

Otherwise, when UE i is inactive, the MSE of estimation has the same expression of (5.23), provided that $\boldsymbol{A}_i$ is replaced by $\boldsymbol{B}_i$ in $\boldsymbol{R}_{\boldsymbol{y}_{p,i}}$.

5.3.2 SPCU Has Partial Knowledge about Active UEs

If we collect the observations corresponding to all the N_p pilot symbols to a vector $\boldsymbol{y} = \left[\boldsymbol{y}_1^T, \boldsymbol{y}_2^T, \ldots, \boldsymbol{y}_{N_p}^T\right]^T$, which is a MN_p-length vector, then, we have

$$\boldsymbol{y} = \sum_{k=1}^{K} I_k (\boldsymbol{p}_k \otimes \boldsymbol{I}_M) \boldsymbol{g}_k + \boldsymbol{n} \tag{5.25}$$

where $\boldsymbol{p}_k = [b_{k,1}, b_{k,2}, \cdots, b_{k,N_P}]^T$ and $\boldsymbol{n}$ is a MN_P-length Gaussian noise vector. Based on (5.25), we can develop a channel estimator in the MMSE principle as shown below.

Assume for the moment that SPCU knows the active UEs. Then, the MMSE-based channel estimator can be expressed as

$$\hat{\boldsymbol{g}}_i = \boldsymbol{W}_i^H \boldsymbol{y},\ i = 1, 2, \ldots, K \tag{5.26}$$

where

$$\boldsymbol{W}_i = \boldsymbol{R}_{\boldsymbol{y}}^{-1} \boldsymbol{R}_{\boldsymbol{y},i} \tag{5.27}$$

Here the autocorrelation matrix is

$$\begin{aligned}
\boldsymbol{R}_{\boldsymbol{y}} =& E[\boldsymbol{y}\boldsymbol{y}^H] \\
=& \sum_{k=1}^{K} I_k (\boldsymbol{p}_k \otimes \boldsymbol{I}_M) E[\boldsymbol{g}\boldsymbol{g}^H] (\boldsymbol{p}_k \otimes \boldsymbol{I}_M)^H + \frac{1}{\mathrm{SNR}_t} \boldsymbol{I}_{N_p M} \\
=& \sum_{k=1}^{K} I_k (\boldsymbol{p}_k \otimes \boldsymbol{I}_M) \boldsymbol{A}_k (\boldsymbol{p}_k \otimes \boldsymbol{I}_M)^H + \frac{1}{\mathrm{SNR}_t} \boldsymbol{I}_{N_p M} \\
=& \sum_{k=1}^{K} I_k (\boldsymbol{p}_k \otimes \boldsymbol{A}_k) (\boldsymbol{p}_k \otimes \boldsymbol{I}_M)^H + \frac{1}{\mathrm{SNR}_t} \boldsymbol{I}_{N_p M} \\
=& \sum_{k=1}^{K} I_k (\boldsymbol{p}_k \boldsymbol{p}_k^H \otimes \boldsymbol{A}_k) + \frac{1}{\mathrm{SNR}_t} \boldsymbol{I}_{N_p M}
\end{aligned} \tag{5.28}$$

when SPCU employs the knowledge about the large-scale fading of all the K UEs. By contrast, when SPCU only knows the path-loss and the statistics of shadowing fading, $\boldsymbol{R}_{\boldsymbol{y}}$ is also given by (5.28) but with $\boldsymbol{A}_k$ replaced by $\boldsymbol{B}_k$, where $\boldsymbol{A}_k$ and $\boldsymbol{B}_k$ are previously given in Section 5.3.1. The same is in the following derivation, which will not be repeatedly stated.

On the other side, from the symbol-based observation of (5.10), we can readily know that

$$\begin{aligned}
\boldsymbol{R}_{\boldsymbol{y}_f} =& E[\boldsymbol{y}_f \boldsymbol{y}_f^H] \\
=& \sum_{k=1}^{K} \boldsymbol{A}_k + \frac{1}{\mathrm{SNR}_t} I_M \\
=& \boldsymbol{R}_a
\end{aligned} \tag{5.29}$$

which can be estimated using the observations of SPCU based on (5.14). Therefore, (5.28)

is represented in detail as

$$\boldsymbol{R}_y = \begin{pmatrix} \boldsymbol{R}_a & \sum_{k=1}^{K} I_k b_{k,1} b_{k,2} \boldsymbol{A}_k & \cdots & \sum_{k=1}^{K} I_k b_{k,1} b_{k,N_p} \boldsymbol{A}_k \\ \sum_{k=1}^{K} I_k b_{k,2} b_{k,1} \boldsymbol{A}_k & \boldsymbol{R}_a & \cdots & \sum_{k=1}^{K} I_k b_{k,2} b_{k,N_p} \boldsymbol{A}_k \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{K} I_k b_{k,N_p} b_{k,1} \boldsymbol{A}_k & \sum_{k=1}^{K} I_k b_{k,N_p} b_{k,2} \boldsymbol{A}_k & \cdots & \boldsymbol{R}_a \end{pmatrix} \tag{5.30}$$

Above $\boldsymbol{R}_y$ is obtained by assuming that SPCU knows all the active UEs. If SPCU has no knowledge about the active UEs, the matrix of $\boldsymbol{R}_y$ will only have the diagonal elements of $\boldsymbol{R}_a$ and correspondingly, the estimator reduces to the one derived in Section 5.3.1. However, if SPCU knows a part of the active UEs, it can construct a $\boldsymbol{R}_y$ in the form of (5.30), where all the diagonal elements are given by $\boldsymbol{R}_a$ estimated from (5.14), while the non-diagonal elements include only those active UEs known to the SPCU.

In (5.27), the cross-correlation matrix is

$$\boldsymbol{R}_{\boldsymbol{y},i} = E[\boldsymbol{y}\boldsymbol{g}_i^H] = (\boldsymbol{p}_i \otimes \boldsymbol{I}_M)\boldsymbol{A}_i \tag{5.31}$$

which the SPCU can construct, as it knows the pilots sent by different UEs. Consequently, the weight matrix in MMSE sense for estimating the channels of UE i is given by

$$\boldsymbol{W}_i = \boldsymbol{R}_y^{-1}(\boldsymbol{p}_i \otimes \boldsymbol{I}_M)\boldsymbol{A}_i \tag{5.32}$$

Furthermore, the MSE of estimation for active UEs is

$$\begin{aligned} MSE_i =& Tr(\boldsymbol{A}_i - \boldsymbol{R}_{\boldsymbol{y},i}^H \boldsymbol{W}_i) \\ =& \sum_{m=1}^{M} \sigma_{i,m}^2 - Tr(\boldsymbol{R}_{\boldsymbol{y},i}^H \boldsymbol{W}_i) \\ =& \sum_{m=1}^{M} \sigma_{i,m}^2 - Tr(\boldsymbol{R}_{\boldsymbol{y},i}^H \boldsymbol{R}_{\boldsymbol{y}}^{-1} \boldsymbol{R}_{y,i}) \end{aligned} \tag{5.33}$$

while for inactive UEs is

$$\begin{aligned} MSE_i =& Tr(\boldsymbol{W}_i^H \boldsymbol{R}_{\boldsymbol{y}} \boldsymbol{W}_i) \\ =& Tr(\boldsymbol{R}_{\boldsymbol{y},i}^H \boldsymbol{R}_{\boldsymbol{y}}^{-1} \boldsymbol{R}_{y,i}) \end{aligned} \tag{5.34}$$

After the channel estimation, let us now consider UAI.

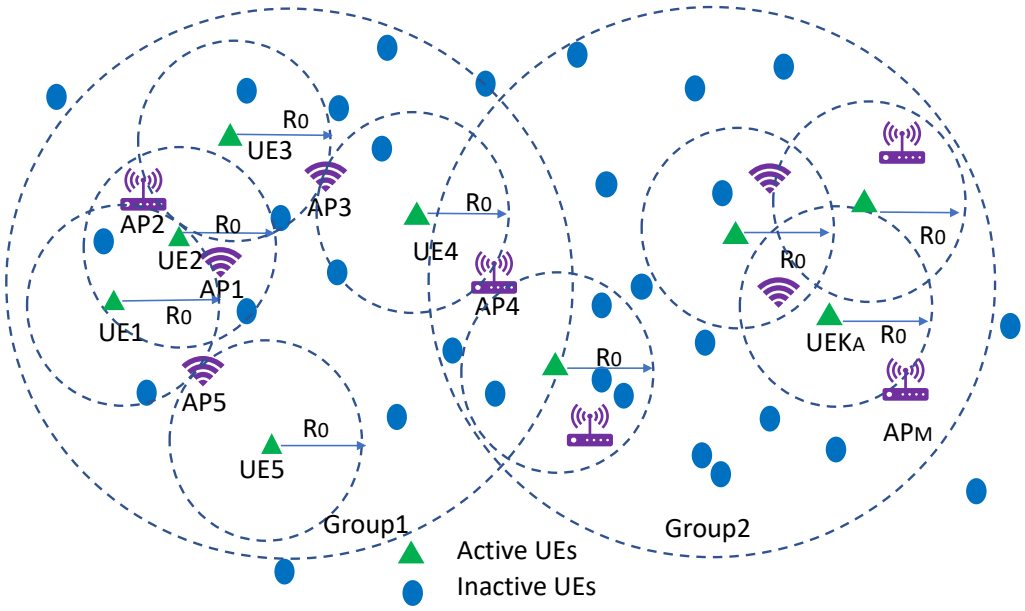


Figure 5.3: Illustration of "first-layer" RRHs/APs and groups in MDA-GFMA systems.

5.4 User Activity Identification in MDA-GFMA Systems

In this section, we propose a UAI algorithm relying on the MMSE-based channel estimation and the orthogonal matching pursuit (OMP) signal recovery, which is referred to as the OMP-MMSE-UAI for convenience of description. For this algorithm, we define the "first-layer" DAs (RRHs/APs) of a UE, as shown in Fig. 5.3, as the DAs located within the area of $R \leq R_0$ from the UE. The reason for this is that due to the propagation path-loss in distributed antenna systems, the signal received by a DA from a far away UE may be too weak to be useful, which may only result in performance degradation. Hence, we set a range R_0 relative to a UE, so that all the DAs within the range can provide meaningful contribution to the considered UE. By contrast, for a given UE, all the other DAs not in the first-layer are classified as the "second-layer" DAs. As shown in Fig. 5.3, different UEs may share some first-layer DAs, as the result that they are close to each other. Additionally, as shown in Fig. 5.3, we may group some UEs and their DAs as indicated by the big dashed-line circle, which will be addressed during the data detection in Section 5.5.

In Section 5.2, we have represented the observations obtained by the M DAs for the f-th bits of a frame as (5.11). Specifically for the pilot bits, the observations can be expressed as

$$\boldsymbol{y}_p = \boldsymbol{G}\boldsymbol{b}_p + \boldsymbol{n}_p, \ p = 1, 2, \cdots, N_P \tag{5.35}$$

where p indicates the pth pilot, the total number of pilots per frame is N_p, and $\boldsymbol{G} = [\boldsymbol{g}_1, \boldsymbol{g}_2, \cdots, \boldsymbol{g}_K]$. However, because of the above definition for the first-layer DAs, we

can know that $\boldsymbol{G}$ can be approximated by a sparse matrix expressed as $\bar{\boldsymbol{G}}$, where only the channel elements related to the DAs in the first-layers of UEs are retained, while all the other channel elements are set to zeros. Note that, $\bar{\boldsymbol{G}} = [\bar{\boldsymbol{g}}_1, \bar{\boldsymbol{g}}_2, \ldots, \bar{\boldsymbol{g}}_K]$ can be constructed using the UEs' channels estimated in Section 5.3 by considering the locations of UEs and DAs. In (5.35), $\boldsymbol{b}_p = [I_1 b_{1,p}, I_2 b_{2,p}, \cdots, I_k b_{K,p}]^T$, whose elements are either the pilots sent by active UEs, which are known to SPCU, or zeros, if the corresponding UEs are inactive. Therefore, $\boldsymbol{b}_p$ is a sparse vector as the fact that the activation probability P_a of UEs is small. Hence, the UAI in MDA-GFMA system is a sparse recovery problem that can be solved using various compressive sensing algorithms [248]. In this chapter, we specifically consider the low-complexity OMP-assisted algorithm for UAI, referred to as the OMP-MMSE-UAI, which is described as follows.

Algorithm 5.4 UE activity identification based on MMSE channel estimation and orthogonal matching pursuit (OMP-MMSE-UAI).

Inputs: $\boldsymbol{y}_p$ and $\boldsymbol{b}_p$ for $p = 1, 2, \ldots, N_P$, $\boldsymbol{y}_f$ corresponding to data, 'Stop-condition', N_c.

Channel Estimation: based on an approach in Section 5.3 to obtain $\boldsymbol{G}$ and construct $\bar{\boldsymbol{G}}$.

Initial UE activity identification by OMP:

for $p = 1, 2, \ldots, N_P$, **do**

1. Initialization: $\boldsymbol{r}_0 = \boldsymbol{y}_p$, $\mathcal{K}_A^{(0)} = \emptyset$, $\hat{\boldsymbol{b}}_p = \boldsymbol{0}$, iteration number $t = 1$
2. Find a possible active UE: $u_i = \arg \max_{j \notin \mathcal{K}_A^{(t-1)}} \left| \bar{\boldsymbol{G}}^H \boldsymbol{r}_{t-1} \right|$;
3. Update active UE set: $\mathcal{K}_A^{(t)} = \mathcal{K}_A^{(t-1)} \cup u_i$;
4. Estimate $\boldsymbol{b}_p^{(t)}$: $\boldsymbol{b}_p^{(t)} = \left[(\bar{\boldsymbol{G}}(\mathcal{K}_A^{(t)}))^H \bar{\boldsymbol{G}}(\mathcal{K}_A^{(t)}) \right]^{-1} (\bar{\boldsymbol{G}}(\mathcal{K}_A^{(t)}))^H \boldsymbol{y}_p$;
5. Update residual: $\boldsymbol{r}_t = \boldsymbol{y}_p - \bar{\boldsymbol{G}}(\mathcal{K}_A^{(t)}) \boldsymbol{b}_p^{(t)}$;
6. $t \leftarrow t + 1$. Then, return 2) if 'Stop-condition' not met.

Decision making: $\hat{\boldsymbol{b}}_p \left(\mathcal{K}_A^{(t_{end})} \right) = \text{sgn} \left(\Re\{ \boldsymbol{b}_p^{(t_{end})} \} \right)$, where t_{end} is the t index in the last iteration, $\hat{\boldsymbol{b}}_p \left(\mathcal{K}_A^{(t_{end})} \right)$ considers only those bits defined by the support $\mathcal{K}_A^{(t_{end})}$.

end for

Further UE activity identification:

Initializations: $\mathcal{K}_A = \emptyset$.

for $k = 1, 2, \ldots, K$, **do**

1. Count the number of correctly estimated pilot bits: $\bar{N}_c(k) = \sum_{p=1}^{N_P} 1(\hat{\boldsymbol{b}}_p(k) = \boldsymbol{b}_p(k))$, where $1(\hat{\boldsymbol{b}}_p(k) = \boldsymbol{b}_p(k))$ equals 1, if $\hat{\boldsymbol{b}}_p(k) = \boldsymbol{b}_p(k)$ and 0, otherwise;
2. If $\bar{N}_c \geq N_c$, UE k is identified to be active and set $\mathcal{K}_A = \mathcal{K}_A \cup k$.

end for

Output: $\mathcal{K}_A$.

To describe the OMP-MMSE-UAI algorithm of Algorithm 5.4, we define $\mathcal{K}_A^{(t)}$ as a support set containing the indices of the possible active UEs identified upto the tth iteration. $\mathcal{K}_A$ is the set to store the finally identified active UEs. $\bar{\boldsymbol{G}}(\mathcal{K}_A^{(t)})$ is a submatrix of $\bar{\boldsymbol{G}}$ with the columns defined by $\mathcal{K}_A^{(t)}$. The minimum number of correctly detected pilot bits of a UE should be at least N_c and otherwise, the UE is assumed inactive or meaningless due to the low-reliability of data detection. This is reasonable, as the detected data of an unreliable UE is no use, if there are too many detection errors. $\hat{\boldsymbol{b}}_p$ is used to store the detected bits, having the value $+1$ or -1 if a UE is active, or 0 if the UE is inactive.

As shown in the description of Algorithm 5.4, the OMP-MMSE-UAI algorithm is divided into two stages. During the first stage, the OMP algorithm [249] is operated with respect to each of the N_p pilot symbols to provide the initial UAI. Furthermore, the pilot bits of all the identified active UEs are estimated. However, at this point, first, the identified active UEs may be different for different pilot bits. Second, the 'Stop-condition' needs to be set so that the miss probability is sufficiently small, which however may result in some false-alarm. Therefore, during the second stage, an enforcement identification stage is executed. During this stage, the detected pilot bits from the first stage are compared with the pilot bits of the identified active UEs. If the number of detected pilot bits is at least N_c, the corresponding UE is confirmed to be active. Otherwise, if the number of detected pilot bits is less than N_c, the corresponding UE is changed to the state of inactive, no matter whether it is actually active or inactive. In this case, even the UE is active, the data detection may be low-reliability, and the detected errors may be too many to be corrected by the embedded error-correction decoding.

In terms of the 'Stop-condition' for the OMP-MMSE-UAI algorithm, various criteria

as shown in [249] may be used. Considering the speciality of our MDA-GFMA systems, it can be set as the relative residual energy, formulated as $\|\boldsymbol{r}_t\|^2/\|\boldsymbol{y}_p\|^2 < \xi$, where ξ is a SNR-depended small constant. Additionally, when given the number of potential UEs K, the activation probability of P_a, an integer K_T can be found so that the probability $P(K_A > K_T)$ is a small, such as 10^{-5}, value. In this case, the miss probability after the first stage of UAI should be small, but the false-alarm probability can be significant. Fortunately, the false-alarm resulted active UEs will be most likely removed during the second enforcement UAI stage, as the data detection of the false-alarm UEs should be low-reliability, resulting in significant detection errors of the pilot bits.

5.5 Data Detection of Active UEs in MDA-GFMA Systems

So far we have considered the channel estimation and UAI in MDA-GFMA systems. Let us now address the data detection of the active UEs by specifically considering the data detection of the ith (which is assumed to be active) UE in MMSE principle. Furthermore, we consider the detection of a group of UEs in the principle of MMSE-SIC.

In order to reduce the detection complexity, for detecting UE i, let us define a reference cell with the group of active UEs forming a set $\mathcal{S}_0$, which includes UE i and some other active UEs closer to UE i, as shown in Fig. 5.3. All the other active UEs are collected to a set $\mathcal{S}_I$. Correspondingly, all the DAs having their distances from any of the UEs in $\mathcal{S}_0$ less than R_0 are collected to form a set $\mathcal{M}_0$. Hence, the number of antennas considered for detecting UE i is given by $M_0 = |\mathcal{M}_0|$. Then, from (5.10) and upon considering only those observations from the reference cell, we can obtain an observation equation of

$$\bar{\boldsymbol{y}}_f = \sum_{k\in\mathcal{S}_0} \bar{\boldsymbol{g}}_k b_{k,f} + \sum_{k'\in\mathcal{S}_I} \bar{\boldsymbol{g}}_{k'} b_{k',f} + \bar{\boldsymbol{n}}_f,\ f = 1,2,\cdots,N_f \tag{5.36}$$

where $\bar{\boldsymbol{y}}_f$ is M_0-length, containing the M_0 observations from the M_0 DAs in the reference cell. Note furthermore that $\bar{\boldsymbol{g}}_k$ is M_0-length, only including the channels related to the DAs in the reference cell.

Having obtained the reduced representation for detecting UE i, $i \in \mathcal{S}_0$, the MMSE detector is straight forward, which can be expressed as

$$z_{i,f} = \boldsymbol{w}_i^H \bar{\boldsymbol{y}}_f,\ i \in S_0,\ f = 1,2,\ldots,N_F \tag{5.37}$$

where

$$\boldsymbol{w}_i = \boldsymbol{R}_{\bar{\boldsymbol{y}}_f}^{-1} \boldsymbol{r}_{\bar{\boldsymbol{y}}_{f,i}} \tag{5.38}$$

with

$$\boldsymbol{R}_{\bar{\boldsymbol{y}}_f} = E[\bar{\boldsymbol{y}}_f \bar{\boldsymbol{y}}_f^H] \tag{5.39}$$

$$= \sum_{k \in \mathcal{S}_0} \bar{\boldsymbol{g}}_k \bar{\boldsymbol{g}}_k^H + \sum_{k' \in \mathcal{S}_I} \bar{\boldsymbol{g}}_{k'} \bar{\boldsymbol{g}}_{k'}^H + \frac{1}{SNR_t} \boldsymbol{I}_{M_0} \tag{5.40}$$

$$\boldsymbol{r}_{\boldsymbol{y}_{f,i}} = \bar{\boldsymbol{g}}_i \tag{5.41}$$

Note that $\boldsymbol{R}_{\bar{\boldsymbol{y}}_f}$ can be directly estimated based on (5.39) using all the observations obtained by the DAs in the reference cell over one frame duration, i.e.,

$$\boldsymbol{R}_{\bar{\boldsymbol{y}}_f} = \frac{1}{N_F} \sum_{f=1}^{N_F} \bar{\boldsymbol{y}}_f \bar{\boldsymbol{y}}_f^H \tag{5.42}$$

By contrast, $\bar{\boldsymbol{g}}_i = \hat{\bar{\boldsymbol{g}}}_i$ can be obtained from the estimated $\hat{\boldsymbol{g}}_i$ in Section 5.3. Hence, an estimate to $\boldsymbol{w}_i$ of (5.38) can be obtained.

Algorithm 5.5 Successive interference cancellation assisted MMSE detection (MMSE-SIC).

Initialization: $\boldsymbol{y}_f, f = 1, 2, \ldots, N_F$, large-scale fading or propagation path-loss plus statistics of shadowing.

Channel estimation: based on an approach in Section 5.3 to give the estimated channel set $\{\hat{\boldsymbol{h}}_i\}$.

UE activity identification: based on the approach in Section 5.4 to obtain the active UE set $\mathcal{K}_A$.

Grouping of active UEs: active UEs are grouped in the principle as shown in Fig. 5.3, to form G groups. Assume that Group b, $b = 1, 2, \ldots, G$ has K_{A_b} active UEs.

For Group $b = 1, 2, \ldots, G$[1], and for $f = 1, 2, \ldots, N_F$ (except the pilot bits), execute:

- **Initialization**:

  1) Get $\bar{\boldsymbol{y}}_f$ from $\boldsymbol{y}_f$ and set $\bar{\boldsymbol{y}}_f^{(0)} = \bar{\boldsymbol{y}}_f$;

  2) Estimate the autocorrelation matrix using (5.42) and set $\boldsymbol{R}_{\bar{\boldsymbol{y}}_f}^{(0)} = \boldsymbol{R}_{\bar{\boldsymbol{y}}_f}$;

  3) Prepare the channel vectors for all the active UEs in Group b with the support set $\mathcal{S}_0$, forming a set $\{\bar{\boldsymbol{g}}_k, k \in \mathcal{S}_0\}$.

- **For** $s = 0, 1, \ldots, K_{A_b} - 1$**, execute**:

Step 1 *Form the decision variables for the* $(K_{A_b} - s)$ *active UEs in* S_0 *that have not been detected as:*
$z_{f,k} = \boldsymbol{w}_k^H \bar{\boldsymbol{y}}_f^{(s)}$, where $\boldsymbol{w}_k = \left(\boldsymbol{R}_{\bar{\boldsymbol{y}}_f}^{(s)}\right)^{-1} \hat{\tilde{\boldsymbol{g}}}_k$.

Step 2 *Measure the reliabilities of the* $K_{A_b} - s$ *UEs according to* (5.43) *and find the most reliable one as:* $k^{(s)} = \arg\max_{k'} \{L_{k'1}, L_{k'2}, \ldots, L_{k'_{|K_{A_b}|-s}}\}$.

Step 3 *Detect the most reliable UE in* S_0 *as* $\hat{b}_{k^{(s)}} = sgn(z_{f,k^{(s)}})$.

Step 4 *Interference cancellation and update*:
$\boldsymbol{y}_f^{(s)} = \boldsymbol{y}_f^{(s-1)} - \hat{\tilde{\boldsymbol{g}}}_{k^{(s)}} \hat{b}_{k^{(s)}}$, $\boldsymbol{R}_{\bar{\boldsymbol{y}}_f}^{(s)} = \boldsymbol{R}_{\bar{\boldsymbol{y}}_f}^{(s-1)} - \hat{\tilde{\boldsymbol{g}}}_{k^{(s)}} \left(\hat{\tilde{\boldsymbol{g}}}_{k^{(s)}}\right)^H$.

Outputs: Detected data of all active UEs.

Above the MMSE-based detection has been considered. It is well-known that the MMSE detection is capable of achieving good performance, when the number of variables (i.e., K_A, the number of active UEs) to be detected is less than the number of observations (i.e., M_0, the number of antennas in a reference cell) by at least 2 [127]. However, in the MDA-GFMA system, the number of active UEs in a reference cell is dynamic, which has certain probability to be close or even larger than the number of observations. In this case, good detection performance of MDA-GFMA systems cannot be guaranteed. In order for the detection in MDA-GFMA systems to attain the required performance, below we propose the MMSE-SIC detection, which is summarized as Algorithm 5.5.

As described in Algorithm 5.5, the algorithm uses the channels estimated in Section 5.3, while the active UEs are determined in Section 5.4. Then, the identified active UEs are divided into G groups in the principle as shown in Fig. 5.3. Then, for each of the G groups, the MMSE-SIC is operated to detect the data of the active UEs in the group from the most reliable UE to the least reliable UE in the principle of MMSE-SIC, as described previously in Chapter 2. During this process, the reliabilities of UEs are measured as

$$\begin{aligned} L_{i,f} =& (1 + \tilde{\gamma}_i) z_{i,f} \\ =& \left(\frac{1}{1 - \hat{\tilde{\boldsymbol{g}}}_i^H \hat{\boldsymbol{R}}_{\bar{\boldsymbol{y}}_f}^{-1} \hat{\tilde{\boldsymbol{g}}}_i} \right) z_{i,f}, f = 1, 2, \ldots, N_F;\ i = 1, 2, \ldots, K_A \end{aligned} \tag{5.43}$$

where $\tilde{\gamma}_i = \hat{\tilde{\boldsymbol{g}}}_i^H \hat{\boldsymbol{R}}_{\bar{\boldsymbol{y}}_f}^{-1} \hat{\tilde{\boldsymbol{g}}}_i / (1 - \hat{\tilde{\boldsymbol{g}}}_i^H \hat{\boldsymbol{R}}_{\bar{\boldsymbol{y}}_f}^{-1} \hat{\tilde{\boldsymbol{g}}}_i)$ is the estimate to the signal-to-interference-plus-noise ratio (SINR) for detecting UE i.

First, we note that during a detection stage, the number of groups is dependent on the number of active UEs as well as their locations. When these change, the number of

groups may be different. The detection of different groups can be implemented in parallel, which shortens the detection latency. However, to minimize the inter-group interference, it is desirable that different groups have the least possible connections. In other words, it is desirable that different groups of active UEs are geographically isolated, so that the propagation path-loss between groups can sufficiently reduce the inter-group interference.

Second, we note that the detection of the data bits of a frame can also be implemented in parallel for reducing the detection latency. Furthermore, although we use the indices $f = 1, 2, \ldots, N_F$ for convenience, the pilot bits do not need to be detected.

Based on the above discussion, we can know that detection latency is only proportional to the numbers of active UEs in the groups, and the overall detection delay is determined by the maximum number of active UEs in the groups. Therefore, in order to reduce the detection delay, it can be expected that groups are divided in such a way that all groups have a similar number of active UEs.

We note furthermore that in the algorithm, interference cancellation is only implemented within each of the groups. However, as shown in Fig. 5.3, different groups may have some overlapping, making some DAs belong to different groups. Furthermore, some UEs near the border may impose strong interference on the UEs in another group. Based on these observations, if the detection of different group can cooperate, further performance improvement is available. For example, when SPCU detects an active UE of a group and finds it imposes relatively strong interference on another group, it can also cancel the effect of this detected UE from the other group, so as to improve its detection performance. Correspondingly, the interference cancellation operation as seen in Algorithm 5.5 at the sth iteration can be modified to

$$\boldsymbol{y}_f^{(s)} = \boldsymbol{y}_f^{(s-1)} - \hat{\tilde{\boldsymbol{g}}}_{k^{(s)}} \hat{b}_{k^{(s)}} - \sum_b \hat{\tilde{\boldsymbol{g}}}_{k_b^{(s)}} \hat{b}_{k_b^{(s)}} \tag{5.44}$$

where $k^{(s)}$ is the detected UE in the reference group, $\sum_b$ means all the groups other than the reference group, and $k_b^{(s)}$ indicates the detected UE in the bth group.

5.6 Simulation Results and Analysis

Below we provide a range of simulation results to demonstrate the performance of channel estimation, UAI and data detection in the MDR-GFMA system. In our simulations, some common parameters used are as follows, unless they are specifically notified. We typically assume a normalized area (with respect to the reference distance d_0) of size 50×50. The radius defining the first-layer of a UE is $R_0 = 15$. The propagation parameters for the large-scale fading in (5.6) are set to $\gamma_1 = 2$, $\gamma_2 = 3$ and $\delta = 5$. In the MDR-GFMA

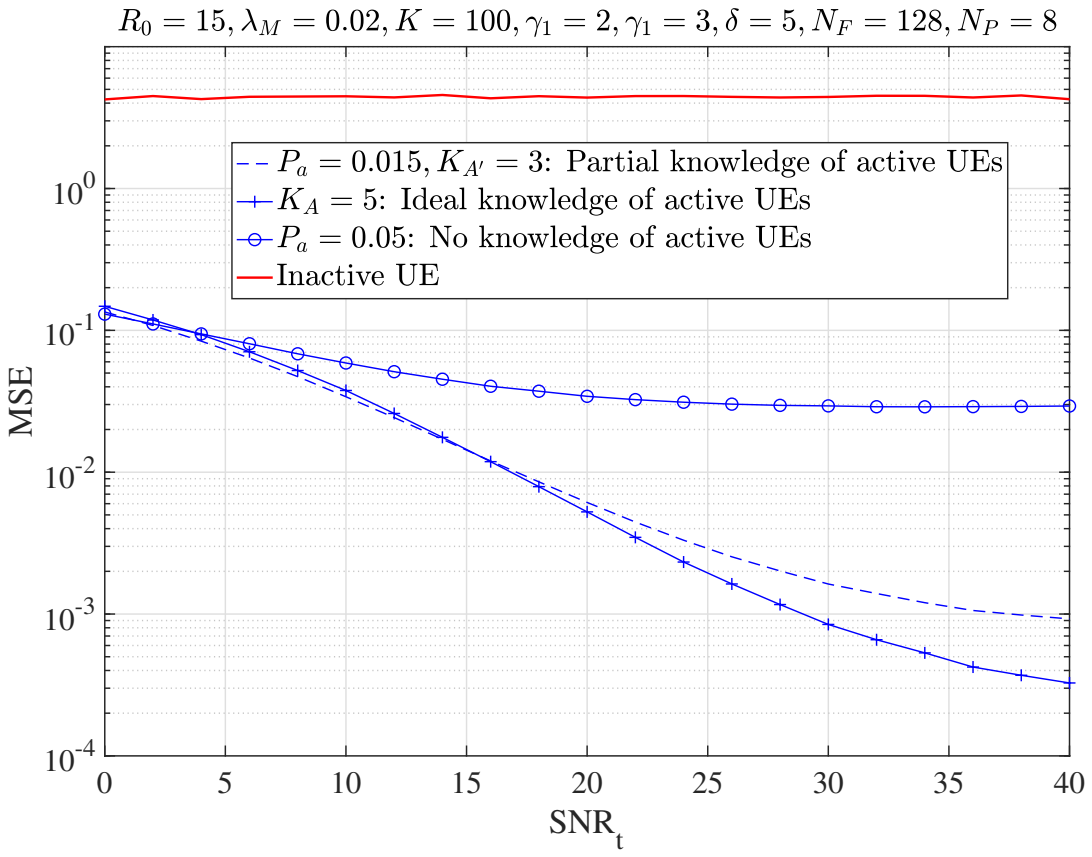


Figure 5.4: MSE versus SNR$_t$ performance of the MMSE-assisted channel estimation in the MDA-GFMA systems where SPCU employs the knowledge of all UEs' large-scale fading.

system, the RRUs/APs are distributed following the Point Poisson (PP) distribution with a density of λ_M [250]. We assume that one UE is served by no more than 4 DAs at one time. UEs are assumed to be randomly (uniformaly) distributed in the simulation area. During a time-slot, each UE has the same probability of P_a to become active. Additionally, we assume that the frame length is $N_F = 128$ bits, among which there are $N_P = 8$ pilot bits.

In Fig.5.4 and Fig. 5.5, we demonstrate the mean-square error (MSE) performance of channel estimation, where for Fig.5.4 we assume that the SPCU has the knowledge of large-scale fading including both the propagation path-loss and shadowing slow fading, while for Fig. 5.5 we assume that the SPCU knows the propagation path-loss but only the statistics of shadowing slow fading. Furthermore, for both figures, we assume that there are 100 UEs randomly distributed in the simulation area of 50×50, and each UE has the activation probability of $P_a = 0.015$. From the results we can learn that for the active UEs, in general, the performance of the case that SPCU has ideal knowledge of active UEs is better than that of the case that SPCU has partial knowledge of active UEs, and both of them have better performance than the case that SPCU has no knowledge about active

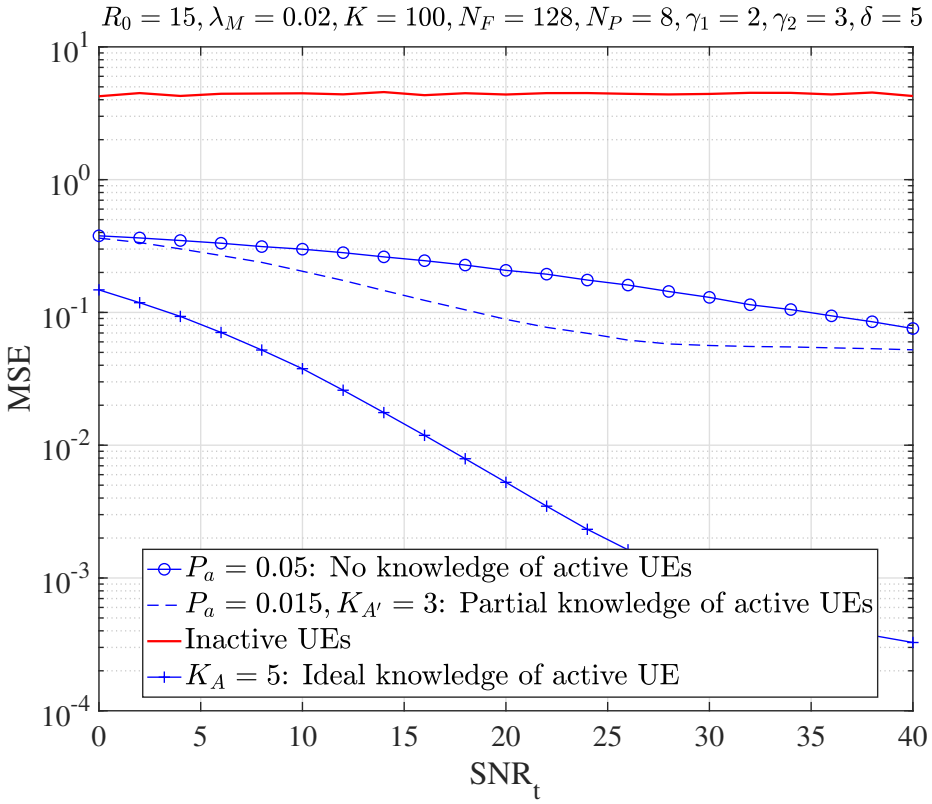


Figure 5.5: MSE versus SNR$_t$ performance of the MMSE-assisted channel estimation in the MDA-GFMA systems, where SPCU knows the positions of UEs and the statistics of shadowing.

UEs. Note that, when the transmit power is low, resulting in low SNR$_t$, we see that the MSE performance of the partial knowledge case can be slightly better than that of the ideal knowledge. The reason for this observation is that the activation probability is $P_a = 0.015$, yielding an average number of active UEs of 4.5 for the partial knowledge case. The results of Fig. 5.4 show that the estimation MSE for an inactive UE is always much higher in the considered SNR$_t$ range. Furthermore, when SNR$_t$ is high, the MSE performance appears floor, which is due to the fact that a UE's signal usually only reaches the several DAs that are close to the UE, all the other DAs' received signals from the UE are too weak to be estimated. Consequently, when SNR$_t$ is high, the MSE performance will be dominated by the estimation errors of these weak signals.

The performance shown in Fig.5.5 has similar tendency as that shown in Fig.5.4 for the different cases considered. However, as the result that SPCU only exploits the knowledge of propagation path-loss and the statistics of shadowing slow fading, the MSE performance of both the partial knowledge case and the no knowledge case is much worse than the corresponding MSE performance shown in Fig.5.4. By contrast, when SPCU has ideal knowledge of active UEs, the MSE performance in both figures is similar, due to the

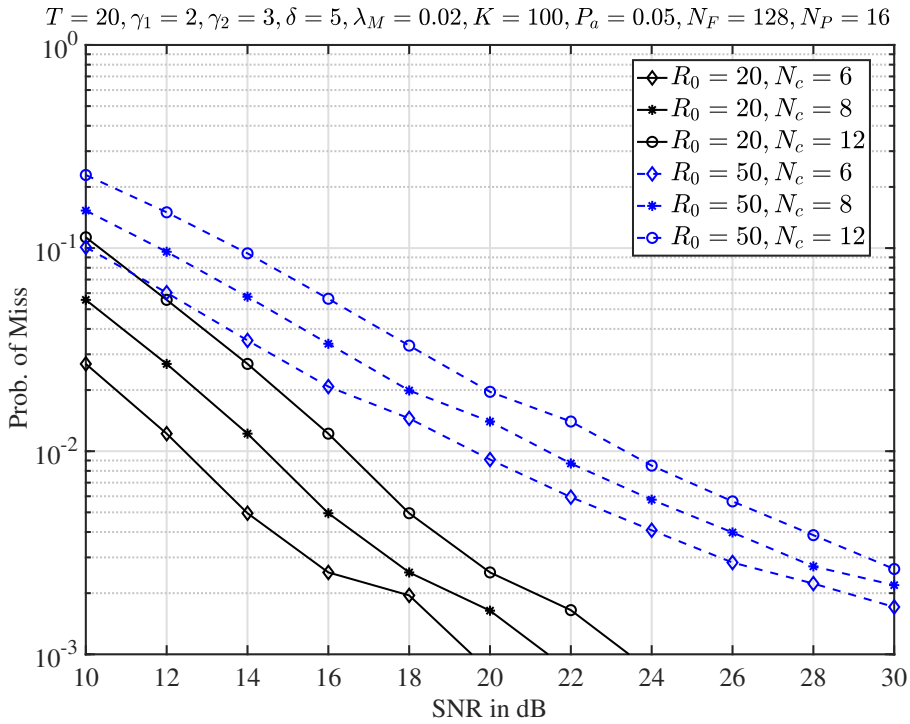


Figure 5.6: Miss probability of UAI, when the 'Stop-condition' for OMP is set to $T = 20$ iterations.

number of active UEs considered being small.

In Figs. 5.6 and 5.7, the probability of miss and that of false-alarm of the OMP-MMSE-UAI algorithm in Section 5.4 are demonstrated when different values of R_0 and different values of N_C are applied. First, for given R_0 and when N_C increases, the miss-probability increases, while the false-alarm probability decreases. This can be readily understood, as increasing N_C makes the conditions of a UE being identified active more strict. Second, for a given N_C and when R_0 increases, the probabilities of both miss and false-alarm increases, as the result that more UEs, which may include some weak UEs, are associated with an AP.

Figs. 5.8 and 5.9 show the effect of the number of allowed iterations T in the OMP-MMSE-UAI algorithm on the performance of UAI. Explicitly, as the number of allowed iterations is increased, the performance in terms of false-alarm and miss is improved. However, we should note that when the number of allowed iterations T increases, the computational complexity of the OMP-MMSE-UAI algorithm increases.

From Fig. 5.10 to Fig. 5.13, we show the performance of the signal detection in MDR-GFMA systems where UEs become active for data transmission based on an activation probability P_a, when various issues are considered. Specifically, in Fig. 5.10, we demonstrate the impact of the DA density (λ_M) on the BER performance of the RMD/GFMA

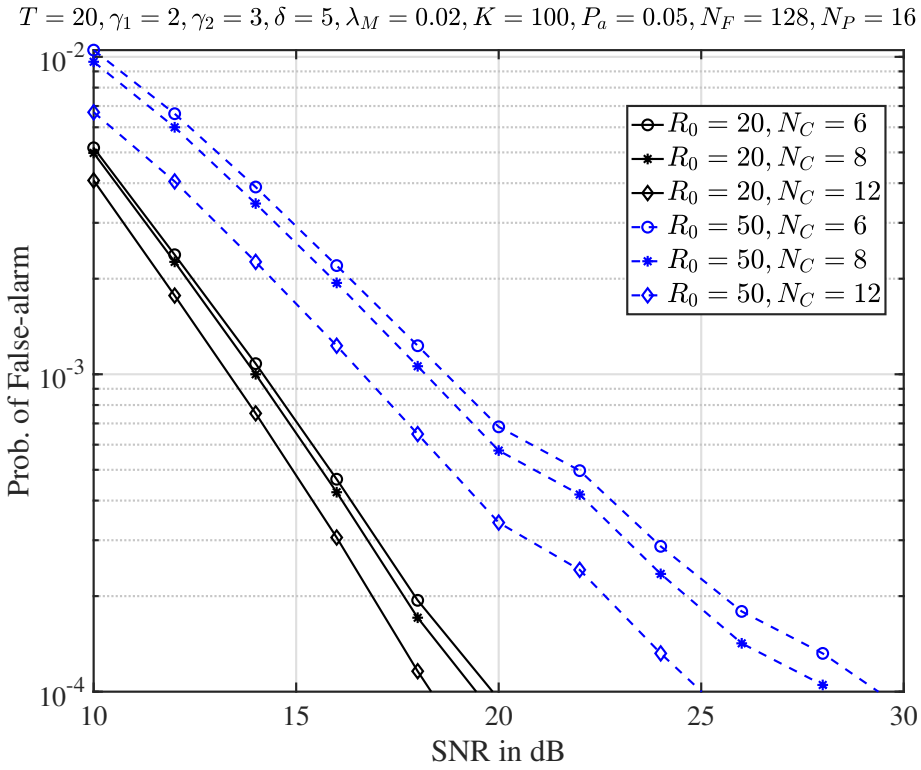


Figure 5.7: False-alarm probability of UAI, when the 'Stop-condition' for OMP is set to $T = 20$ iterations.

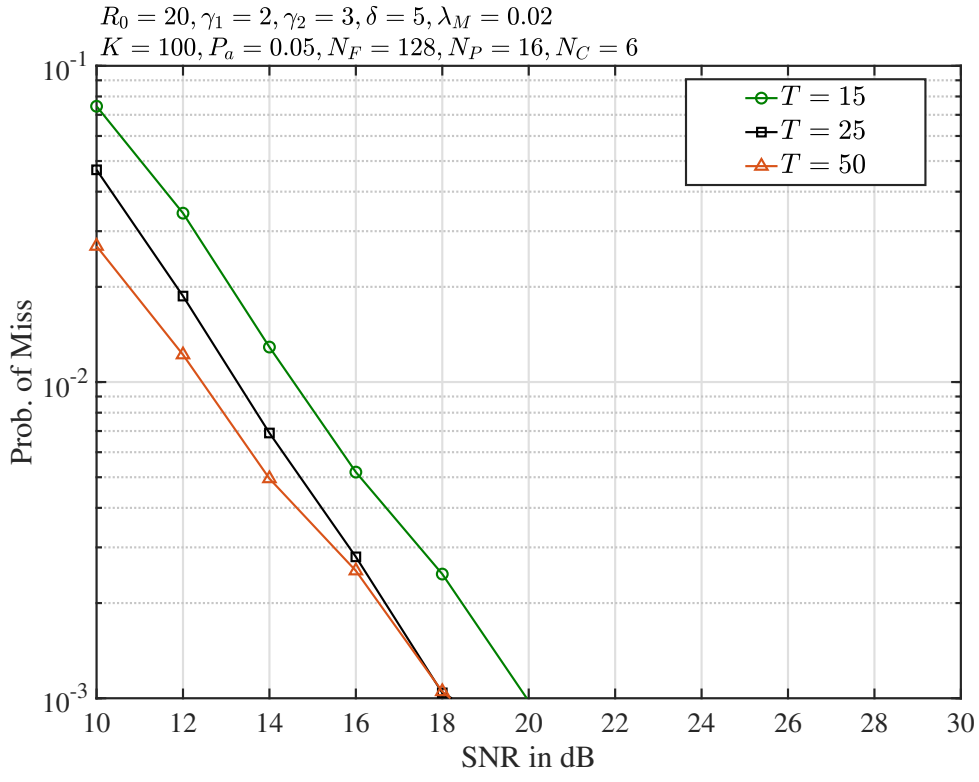


Figure 5.8: Miss probability of UAI, when the 'Stop-condition' for OMP are set to $T = 15$, 25 and 50 iterations, respectively.

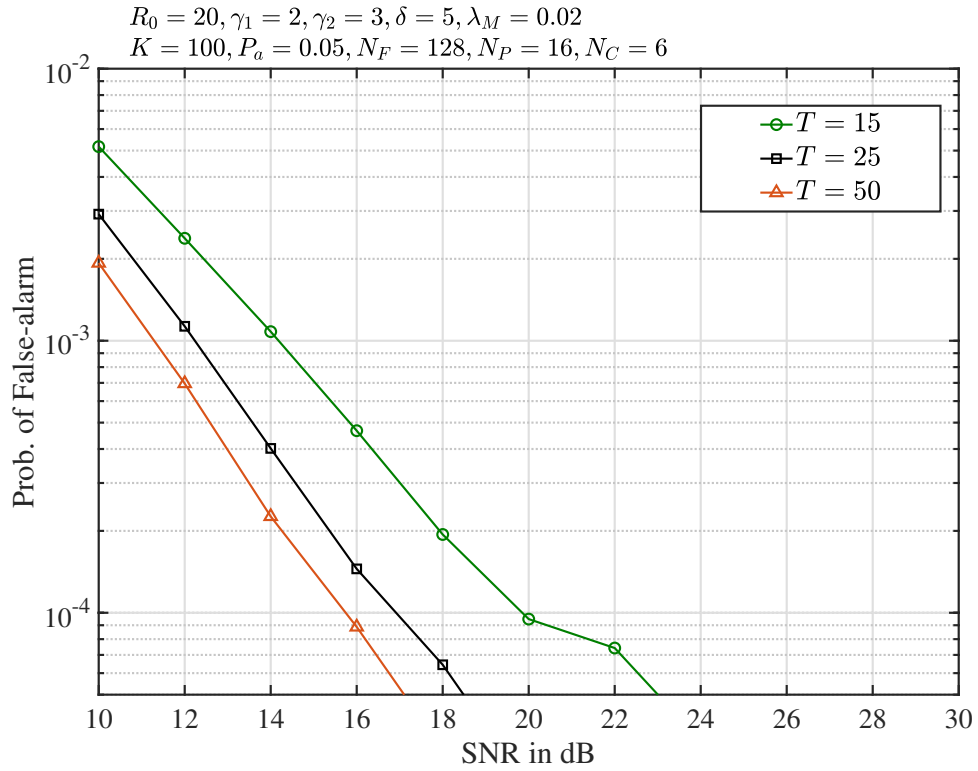


Figure 5.9: False-alarm probability of UAI, when the 'Stop-condition' for OMP are set to $T = 15$, 25 and 50 iterations, respectively.

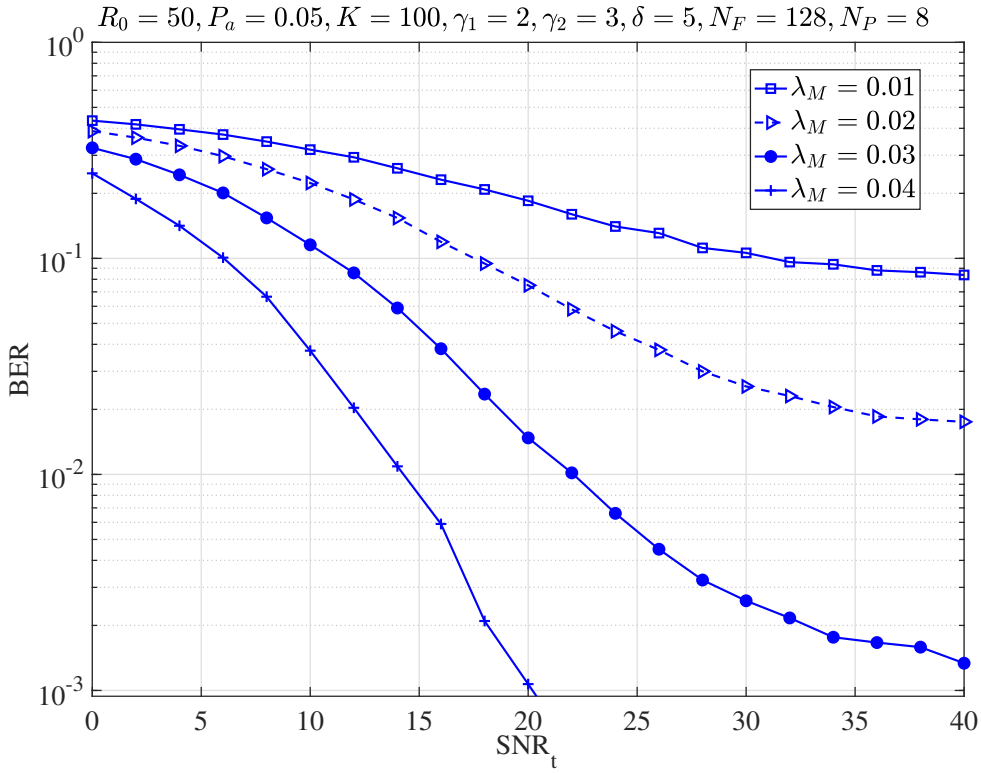


Figure 5.10: BER versus SNR$_t$ performance of the MDA-GFMA systems, when MMSE-SIC detection is employed.

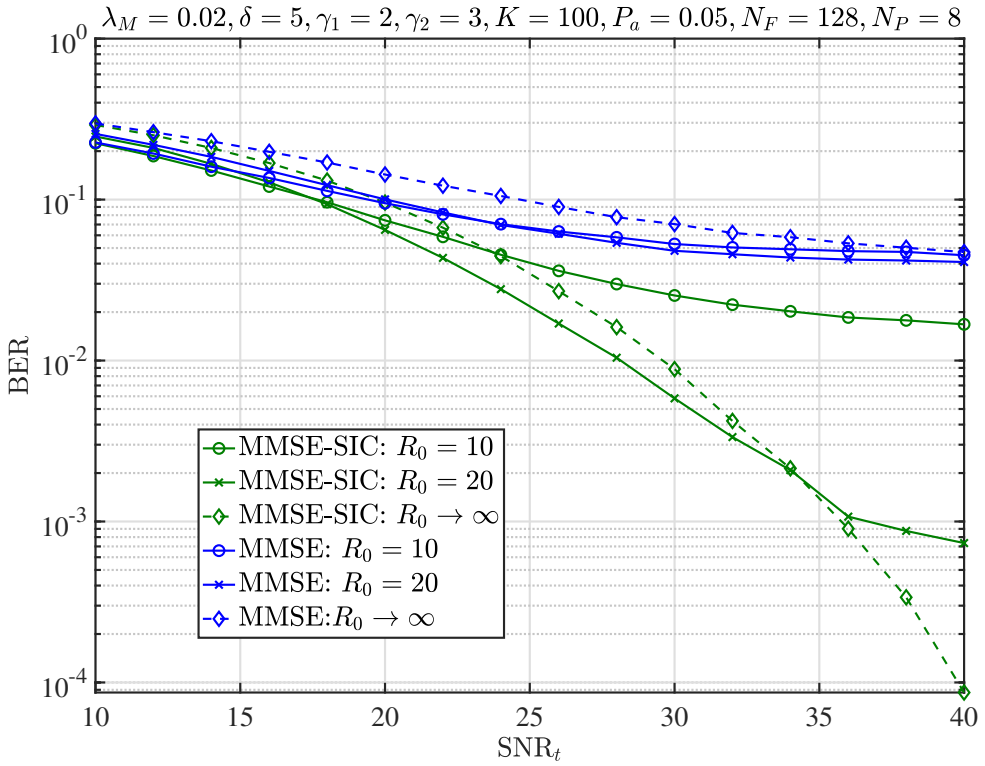


Figure 5.11: BER versus SNR$_t$ performance of the MDR-GFMA systems, when MMSE and MMSE-SIC are respectively employed.

systems employing the MMSE-SIC detection, where the systems support $K = 100$ potential UEs of each having the activation probability of $P_a = 0.05$. As we can see from Fig.5.10, the BER performance is significantly improved, when λ_M is increased. This is because, first, when setting $R_0 = 50$, the number of DAs serving a UE increases, as λ_M is increased. Furthermore, when λ_M is increased, in probability, a UE will have more close DAs. Hence, the propagation path-loss becomes smaller, resulting in the higher received SNR for a given transmit power.

In Fig.5.11, we investigate the effect of R_0 on the BER performance of MDR-GFMA systems, when either MMSE or MMSE-SIC detection is employed. Main parameters are detailed on the top of the figure, in addition to some common parameters mentioned at the beginning of this section. From the results we have the following observations. First, the MMSE-SIC detector significantly outperforms the MMSE detector. This again verified that the MMSE-SIC is efficient for operation in the dynamic systems, resulted not only from the UEs' activity but also from the randomness of DAs. Second, for the MMSE-SIC detection, when the SNR$_t$ is sufficiently high, the BER performance improves, as SNR$_t$ increases. However, SNR$_t$ is low, this may not be the case and using a bigger value of R_0 may result in worse BER performance. By contrast, for the MMSE detection, the BER performance typically degrades, as the value of R_0 increases. The reason behind the above

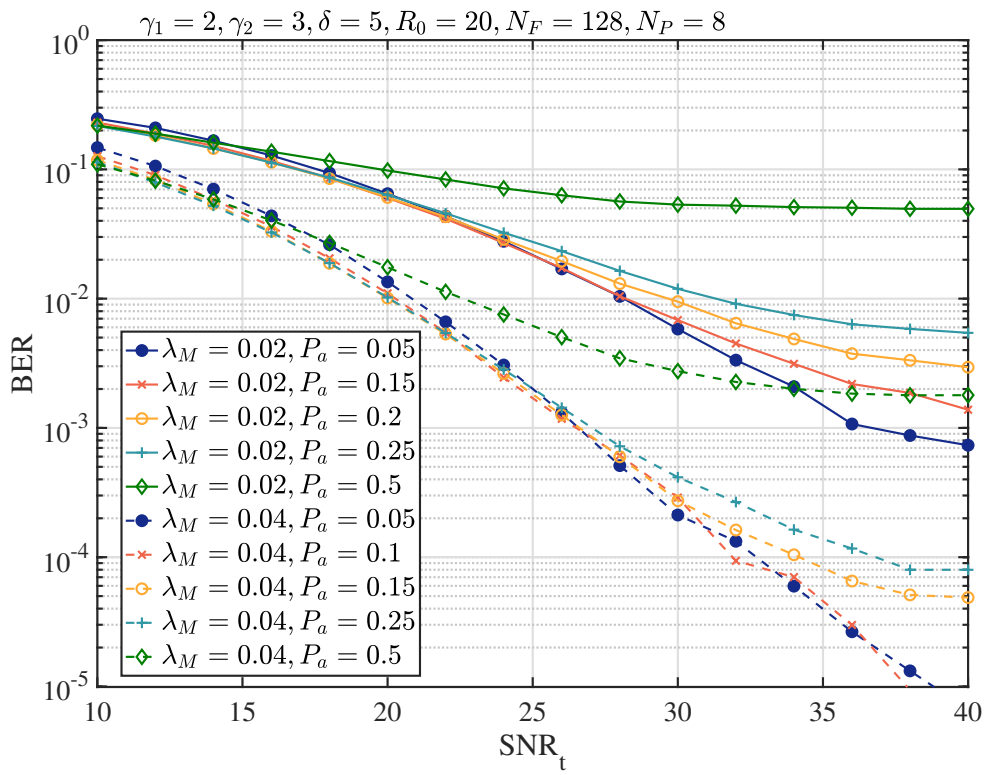


Figure 5.12: BER versus SNR$_t$ performance of the MDA-GFMA systems employing MMSE-SIC detection.

observation is that if SNR$_t$ is small and R_0 is big, there is possibility that the detection is based on the weak signals collected by the DAs far away from the UE, which results in degraded BER performance. From above observations we can gain the insight that when SNRt or transmit power is given, there is a given range for R_0 so that the MDR-GFMA system can achieve the best possible BER performance.

In Fig.5.12, the impact of both DA density and UE activation probability on the BER performance of MDR-GFMA systems is investigated. Again, common parameters are used with some are shown on the top of the figure. The observations are as follows. First, as that in Fig. 5.10, BER performance improves significantly, when DA density is increased from 0.02 to 0.04. Second, BER performance degrades, as P_a increases, as the result of the increased number of UEs in average. However, we should note that even when P_a is as high as $P_a = 0.25$, meaning that there are in average 25 active UEs per time-slot, the MMSE-SIC detector still works efficiently. This is declared, especially, when the DA density is $\lambda_M = 0.04$.

Finally, in Fig.5.13 we investigate the effect of service area and the size of reference cell on the BER performance of MDR-GFMA systems. Here we emphasize that the DA density is $\lambda_M = 0.04$ and the number of potential UEs is $K = 100$. Therefore, when the service area becomes larger, the UE density becomes lower. In this case, the probability

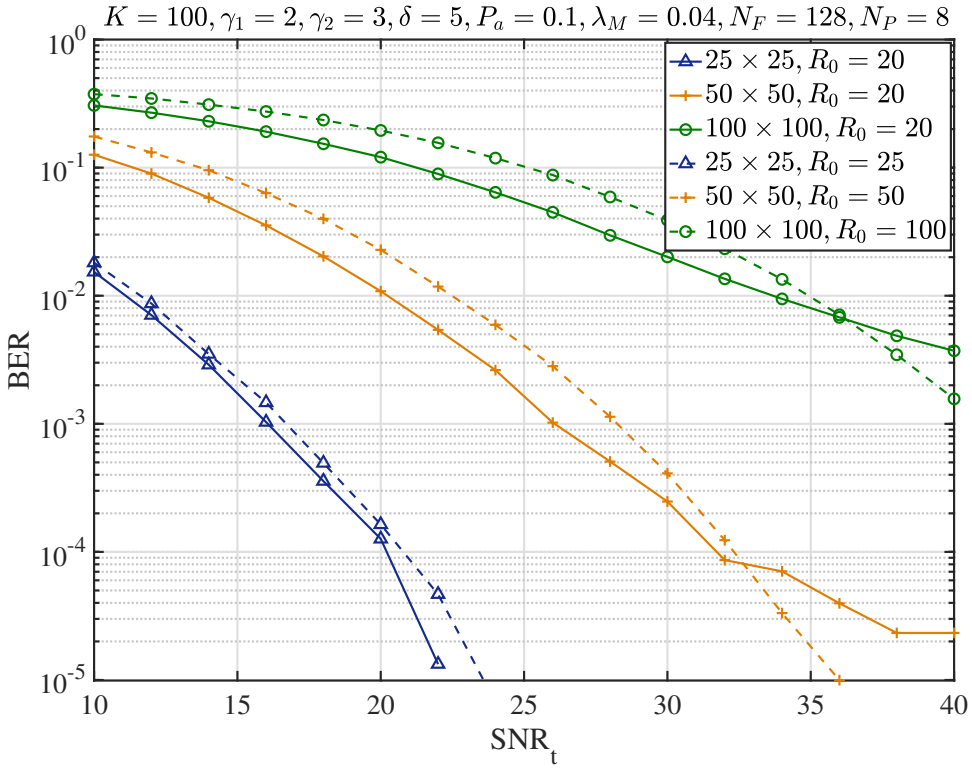


Figure 5.13: BER versus SNR$_t$ performance of the MDR-GFMA systems employing the MMSE-SIC detection.

that a UE has relatively big distance from DAs is relatively bigger. Therefore, as seen in Fig. 5.13, when the service becomes larger, from 25×25 to 50×50 and then to 100×100, the BER performance drops significantly. In terms of the effect of R_0, which determines the size of reference cell, when the service area is 25×25, the BER performance attained at $R_0 = 20$ is slightly better than that achieved at $R_0 = 25$, meaning that the whole service area is the reference cell. By contrast, for both the cases of 50×50 and 100×100, we can see that when the SNR$_t$ is beyond a given value, the case with the reference cell covering the whole service area obtains better BER performance than the case of $R_0 = 20$. Before this SNR$_t$ value, the case of $R_0 = 20$ achieve better BER performance. Again, this is because when R_0 is big, some DAs far away from a UE only receive the weak signals, which are unable to provide reliable channel estimation, making the received weak signals yield the negative effect on the BER performance of MDR-GFMA systems.

In Fig. 5.14 we compare the BER versus SNR$_t$ performance of the MDA-GFMA in Chapter 5 and the MRA/MC-GFMA systems in Chapter 4 when the MMSE-SIC detection is that there are 2 or 6 receiving antennas in the MRA/MC-GFMA systems. For the MDA-GFMA system, we assume that the density of antennas is $\lambda_M = 0.02$ or $\lambda_M = 0.06$ in a $R_0 = 10$ area. Furthermore, we assume that the number of potential UEs is $K = 100$, the number of subcarriers in MRA/MC-CDMA systems is 32. 8 pilots are used to estimate

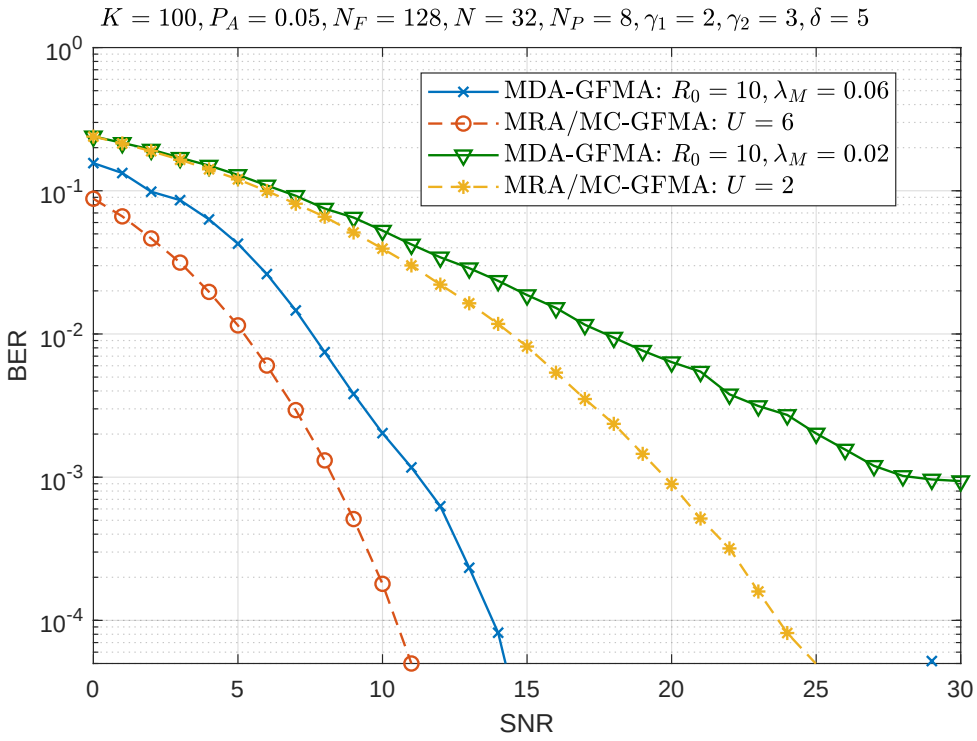


Figure 5.14: BER versus SNR_t performance employing MMSE-SIC detection in MDA-GFMA and MRA/MC-GFMA systems respectively.

channels. When $R_0 = 10$, there are direct channels existing between the UEs and APs. When the density of antenna equals $\lambda_M = 0.02$, in average there are 5 active UEs in the system, some of them may not be served by the antennas. By contrast, when the antenna density is $\lambda_M = 0.06$, the BER performance is improved significantly. However, the UEs in the MDA-GFMA systems are supposed to be randomly distributed, as the number of the antennas severing an active UE is dynamic. By contrast, the number of antennas severing a UE in the MRA/MC-GFMA systems is fixed.

5.7 Chapter Conclusions

In this chapter, we have studied the massive GFMA with distributed antenna MIMO, i.e., the MDR-GFMA system, when DAs follow a PP distribution and UEs follow uniform distribution in a studied area. Without imposing power control but considering the large-scale fading of propagation path-loss and shadowing, the signal sent by a UE can only be received by a small number of DAs close to the UE. Hence, when all the DAs in the system are considered, the channel matrix is sparse. Furthermore, as each of the UEs in the system has only a small probability to become active, the UAI and data detection in the MDR-GFMA system can be represented as the sparse recovery problems. In this chap-

ter, the channel estimation, UAI and data detection for MDR-GFMA systems have been studied. Specifically, the channels of all UEs are first estimated in the principle of MMSE. Then, considering that UAI is a sparse recovery problem but CS-based signal recovery experiences the constraint of RIP, we have proposed an OMP-MMSE-UAI algorithm, which makes use of the pilot symbols to enhance the UAI performance. Finally, the data of active UEs are detected in the principle of MMSE detection and MMSE-SIC detection.

From our studies and performance results we can have the following observations. First, the MMSE-assisted channel estimation is efficient for operation in the MDR-GFMA systems. Explicitly, the estimated channels of active UEs and inactive UEs have significant difference in terms of the MSE performance. This means that the estimated channels for a UE may be exploited to identify whether the UE is active. Second, in our OMP-MMSE-UAI algorithm, the embedded OMP algorithm is capable of providing an initial UAI with a small miss probability, while the followed UAI enhancement aided by the pilot detection can efficiently remove the false-alarmed UEs. Furthermore, concerning the two detection schemes considered, the studies show that the MMSE-SIC detector is highly efficient for data detection in MDR-GFMA systems, which experience a range of dynamics explained, for example, by the active UEs and the number of them, the geography resulted randomness, and the number of DAs serving a UE.

Explicitly, there are various future research issues possible in terms of the system/channel modeling, signal transmission and detection, network optimization, etc. A straightforward research issue on the current MDR-GFMA system is the joint channel estimation, UAI and data detection by designing the low-complexity and high-efficiency algorithms, which may take the advantages of both the CS-based approaches and the conventional approaches.

Chapter 6

Conclusions and Future Works

6.1 Conclusions

This thesis focuses on mGFMA systems, which is assumed to support a massive number of UEs of each with a small probability to become active to transmit its information to BS or AP. After receiving the signals from UEs, the tasks of BS or AP include UAI, channel estimation and the information detection of active UEs. In this thesis, UEs are always assumed to employ a single antenna. By contrast, at the BS or AP side, three scenarios have been considered, including the single-antenna BS (AP) in Chapters 2 and 3, multiple receive antenna BS (AP) in Chapter 4 and the BS (AP) with distributed antennas in Chapter 5.

To be more specific, first, in **Chapter 1**, we provided a motivation discussions about the research as well as the research background. Furthermore, we provided some comparative discussion about the GMA and GFMA, explaining that GFMA is a more efficient multiple-access scheme for providing the uplink transmissions in the mMTC networks.

Then, **Chapter 2** presented a DyDS-CDMA and a DyMC-CDMA schemes for supporting ultra-densely deployed wireless UEs, which result in that each BS may need to support a big number of potential UEs, while each of the UEs activates to transmit during a time-slot with a small probability. Hence, for a given time-slot, the active UEs and the number of active UEs simultaneously transmit information are dynamic. In this chapter, we assumed the ideal identification of active UEs and also assumed the ideal channel estimation of the active UEs, with the motivation to demonstrate the potentially achievable performance of the low-complexity MMSE-SICD operated in the DyDS-CDMA and DyMC-CDMA systems. Therefore, based on the Monte-Carlo simulations, we investigated the BER performance of the DyDS-CDMA and DyMC-CDMA systems employing

the MMSE-SICD, when different settings were considered. Furthermore, the MMSE-SICD was compared with the MMSE-MUD from different perspectives. Our studies demonstrate that the MMSE-SICD is a highly-efficient and robust detection scheme for operation in the DyDS-CDMA and DyMC-CDMA systems operated in mGFMA environments. It is capable of achieving near single-user BER performance, even when the average number of active UEs per time-slot reaches two times of N of the systems' degrees-of-freedom. With the aid of MMSE-SICD, the DyMC-CDMA system employing dense spreading codes is feasible for achieving the frequency diversity, when communicating over frequency-selective fading channels. In order to reduce the PAPR problem, we also studied the BER performance of the sparse DyMC-CDMA systems employing the MMSE-SICD. It is shown that the near-full gain of frequency diversity may only be achievable, when the sparsity of spreading sequences is sufficiently high. Specifically, a typical sparsity for achieving the near-full gain of frequency diversity is the one that is slightly larger than the number of time-domain resolvable paths of the frequency-selective fading channels.

Then, **Chapter 3** investigated the joint channel estimation and UAI in MC-mGFMA systems. Firstly, the channel estimation was addressed with assumptions that AP has the full knowledge, no knowledge or the partial knowledge about active UEs. The studies allowed us to reveal the fact that any added knowledge about active UEs can be exploited for enhancing the channel estimation and furthermore, for the design of novel UAI algorithms. Then, the statistics of the estimated channels of a UE was studied on the condition that the UE is active or inactive. It is shown that the estimated channels of active UEs and inactive UEs present distinctive features, explaining that the estimated channels of UEs can be utilized for UAI. Finally, based on the studies of channel estimation and their statistics, five UAI algorithms were proposed and their performance was studied. Our studies show that while TB-UAI has the lowest complexity, it yields a sharp trade-off between false-alarm and miss probabilities, making it hard to be implemented in the practical high-dynamic mGFMA environments. By contrast, the other four UAI algorithms are capable of providing significant performance enhancements in comparison with TB-UAI. Furthermore, the EAE-UAI and SIC-UAIa algorithms impose no trade-off between false-alarm and miss probabilities and the SIC-UAIb algorithm only makes little trade-off between false-alarm and miss probabilities. Although the AME-UAI algorithm brings in some trade-off between false-alarm and miss probabilities, the related parameter N_c can be relatively easily set in mGFMA systems.

In comparison with the existing approaches, the UAI algorithms proposed in this chapter do not experience the RIP constraint of CS and do not depend on the factor graphs required by the MPA-based approaches. Instead, the signatures considered are random

sequences and the number of active UEs can be highly dynamic, making the actual number of active UEs sometimes be significantly higher than the number of resource units of the system. Therefore, it can be expected that our proposed approaches are robust for operation in mGFMA systems.

In the above two chapters, BS or AP was assumed to employ single receive antenna. By contrast, in **Chapter 4**, we investigated a MRA/MC-mGFMA system, where each AP has multiple receive antennas, for a massive number of UEs to access wireless network without requiring grant scheduling. Firstly, the channel estimation in the principle of MMSE was considered, when AP was assumed to have ideal knowledge, no knowledge or partial knowledge about the active UEs. Then, based on channel estimation, a low-complexity TB-UAI scheme was proposed for UAI. Finally, AP carries out information detection of active UEs in the principles of SIC-MMSE, after the channel estimation and UAI. Furthermore, a joint channel estimation, UAI and information detection that is relied on the SIC-MMSE, referred to as the SIC-MMSE-JCUD, was proposed and studied. We demonstrated the impact of multiple receive antennas on the system design and achievable performance, showing that significant performance improvement is available for channel estimation, UAI and information detection, when AP is deployed with multiple receive antennas. Furthermore, when AP is equipped with multiple receive antennas, a low-complexity TB-UAI is highly effective for achieving promising performance, and the associated threshold is robust for practical settings. Specifically for information detection, in this chapter, two detection schemes were considered, which are the SIC-MMSE and SIC-MMSE-JCUD, both of them rely on the SIC-MMSE principle. Our studies and performance results show that these SIC-MMSE-relied detection schemes are highly efficient for signal detection in mGFMA systems. It is shown that their performance is much better than the benchmark MMSE detector. More importantly, they are flexible for operation in the mGFMA systems where active UEs and the number of them are highly dynamic. Specifically in the SIC-MMSE-JCUD, channel estimation, UAI and data detection can iteratively enhance each other to attain the promising performance. Additionally, by combining the ZC-sequences and Gold-sequences, we proposed a class of signature sequences for supporting the massive number of UEs. Our studies demonstrate that the performance achieved by this class of sequences is better than that attainable by binary random sequences.

Finally in **Chapter 5**, we studied the mGFMA with distributed antenna MIMO, i.e., the MDR-GFMA system, when DAs follow a PP distribution and UEs follow uniform distribution in a studied area. Without imposing power control but considering the large-scale fading of propagation path-loss and shadowing, the signal sent by a UE can only be received by a small number of DAs close to the UE. Hence, when all the DAs in the

system are considered, the channel matrix is sparse. Furthermore, as each of the UEs in the system has only a small probability to become active, the UAI and data detection in the MDR-GFMA system can be represented as the sparse recovery problems. In this chapter, the channel estimation, UAI and data detection for MDR-GFMA systems were studied. Specifically, the channels of all UEs were first estimated in the principle of MMSE. Then, considering that UAI is a sparse recovery problem but CS-based signal recovery experiences the constraint of RIP, we proposed an OMP-MMSE-UAI algorithm, which makes use of the pilot symbols to enhance the UAI performance. Finally, the data of active UEs were detected in the principle of MMSE detection and MMSE-SIC detection.

From our studies and performance results we can have the following observations. Firstly, the MMSE-assisted channel estimation is efficient for operation in the MDR-GFMA systems. Explicitly, the estimated channels of active UEs and inactive UEs have significant difference in terms of the MSE performance. This means that the estimated channels for a UE may be exploited to identify whether the UE is active. Secondly, in our OMP-MMSE-UAI algorithm, the embedded OMP algorithm is capable of providing an initial UAI with a small miss probability, while the followed UAI enhancement aided by the pilot detection can efficiently remove the false-alarmed UEs. Furthermore, concerning the two detection schemes considered, the studies show that the MMSE-SIC detector is highly efficient for data detection in MDR-GFMA systems, which experience a range of dynamics explained, for example, by the active UEs and the number of them, the geography resulted randomness, and the number of DAs serving a UE.

6.2 Suggested Future Work

There are many potential research issues in mGFMA systems. Below are a few of them closely related to the research of this thesis.

1. **Low-Complexity High-Efficiency Detection Algorithms**:

In this thesis, for the sake of low-complexity, we only considered the low-complexity MMSE and MMSE-SIC detection schemes. Although our studies show that the MMSE-SIC detection is highly effective for operation in the dynamic mGFMA systems and achieving promising performance, it however has the detection delay that is proportional to the number of active UEs. This property may make it not suitable for implementation with the systems requiring ultra-low latency information delivery. Therefore, it is important to study the low-complexity and low-latency signal detection schemes. One possible option might be the hybrid parallel-serial SICs operated simultaneously so as to reduce the processing time. However, this may result in a trade-off of the reliability performance.

Therefore, further research is needed to find the required answers.

2. **GFMA in the Ultradensy Distributed Antenna (UDA) Systems**

Explicitly, there are various research issues possible in terms of the system/channel modelling, signal transmission and detection, network optimization, etc. A straightforward research issue on the current MDR-GFMA system is the joint channel estimation, UAI and data detection by designing the low-complexity and high-efficiency algorithms, which may take the advantages of both the CS-based approaches and the conventional approaches.

Additionally, considering the UDA under the concept of the traditional MIMO, we can extend the studies to UDA systems in the millimetre wave (mmWave) domain, as mmWave has been proposed for various future wireless systems.

Pilot design for operation in UDA GFMA systems is another research issue. It is important to design the pilots that are feasible for separating the UEs at the stage of UAI and enhancing the performance of channel estimation and data detection.

Furthermore, UAI in random access UDA systems is also a research issue. For example, we may consider the MMSE channel estimation and the MMSE-SIC assisted signal detection, while studying the performance of UAI, when different system configurations are considered.

3. **mGFMA with Intelligence**

The machine learning (ML) methods have been introduced for channel estimation, UAI and information detection. However, the studies so far rely mainly on the traditional channel models [251–253]. In practice, wireless channels are usually time-variant and most environments may not be possibly explained by the existing channel models. Hence, the data based ML-optimization is highly desirable. Therefore, the mGFMA systems with unknown channel models need to be better investigated by taking the advantage that ML algorithms are capable of turning the optimization parameters on the fly. To this objective, ML algorithms may be introduced to associate UEs with BSs/APs, to optimize UE signature distribution, to synchronize the network and distributed UEs, to optimize signal transmission, to model the fading and interference channel, to carry out UAI, to detect UE's information, etc. Furthermore, some of the above mentioned, such as, channel modelling, UAI and information, may be jointly achieved by a ML algorithm.

Related to the research of this thesis, ML may be implemented with the mGFMA systems to achieve different tasks, including transceiver design and optimization, UE signature design and distribution, network cooperation and optimization, channel modelling, data relied optimizations, etc.

Bibliography

[1] T. Wild, V. Braun, and H. Viswanathan, "Joint design of communication and sensing for beyond 5G and 6G systems," *IEEE Access*, vol. 9, pp. 30845–30857, 2021.

[2] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The next generation wireless access technology*. Academic Press, 2020.

[3] D. Chandramouli, R. Liebhart, and J. Pirskanen, *Access Control and Mobility Management*, pp. 225–282. 2019.

[4] H. Takao, K. Sawada, and M. Ishida, "Multifunctional smart tactile-image sensor with integrated arrays of strain and temperature sensors on single air-pressurized silicon diaphragm," in *Solid-State Sensors, Actuators and Microsystems, 2005. Digest of Technical Papers.TRANSDUCERS'05. The 13th International Conference on*, vol. 1, pp. 45–48, IEEE, 2005.

[5] A. Whitmore, A. Agarwal, and L. Da Xu, "The internet of things—A survey of topics and trends," *Information Systems Frontiers*, vol. 17, no. 2, pp. 261–274, 2015.

[6] M. B. Shahab, R. Abbas, M. Shirvanimoghaddam, and S. J. Johnson, "Grant-free non-orthogonal multiple access for IoT: A survey," *IEEE Communications Surveys Tutorials*, vol. 22, no. 3, pp. 1805–1838, 2020.

[7] 3GPP, "Sa1-22.826: Study on communication services for critical medical applications," *3GPP Technical Report (TR)*, vol. R17, no. v1.0.0, 2019.

[8] W. Obile, "Ericsson mobility report," *Nov*, 2016.

[9] H. Al-Hamadi and R. Chen, "Trust-based decision making for health IoT systems," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1408–1419, 2017.

[10] J. Wang, C. Jiang, Z. Han, Y. Ren, R. G. Maunder, and L. Hanzo, "Taking drones to the next level: Cooperative distributed unmanned-aerial-vehicular networks for

small and mini drones," *IEEE vehIcular technology magazIne*, vol. 12, no. 3, pp. 73–82, 2017.

[11] J. Wang, C. Jiang, Z. Bie, T. Q. Quek, and Y. Ren, "Mobile data transactions in device-to-device communication networks: Pricing and auction," *IEEE Wireless Communications Letters*, vol. 5, no. 3, pp. 300–303, 2016.

[12] C. Pan, H. Mehrpouyan, Y. Liu, M. Elkashlan, and A. Nallanathan, "Joint pilot allocation and robust transmission design for ultra-dense user-centric TDD C-RAN with imperfect CSI," *IEEE Transactions on Wireless Communications*, vol. 17, pp. 2038–2053, Mar. 2018.

[13] 3GPP, "Nr prach preamble design," *3GPP R1-1611904*, vol. TSG-RAN, no. WG1, 2016.

[14] G. D. Jo, "Compact design of random access channel in 5g new radio," in *2020 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 828–830, 2020.

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

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

[17] Y.-M. Chen, C. D. S. González, P.-H. Wang, and K.-P. Chen, "Reinforcement learning-based scma codebook design for uplink rayleigh fading channels," *IEEE Wireless Communications Letters*, pp. 1–1, 2021.

[18] A. T. Abebe and C. G. Kang, "Grant-free uplink transmission with multi-codebook-based sparse code multiple access (MC-SCMA)," *IEEE Access*, vol. 7, pp. 169853–169864, 2019.

[19] Z. Yang, J. Cui, X. Lei, Z. Ding, P. Fan, and D. Chen, "Impact of factor graph on average sum rate for uplink sparse code multiple access systems," *IEEE Access*, vol. 4, pp. 6585–6590, 2016.

[20] D. Sobiski and J. Thorp, "Pdma-1: chaotic communication via the extended kalman filter," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 45, no. 2, pp. 194–197, 1998.

[21] D. Sobiski and J. Thorp, "Pdma-2: the feedback kalman filter and simultaneous multiple access of a single channel," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 45, no. 2, pp. 142–149, 1998.

[22] H. Ding, Q. Liu, Z. Ding, J. Xu, Z. Zhang, W. Liu, and X. Chen, "Low complexity iterative receiver with lossless information transfer for non-binary ldpc coded pdma system," *IEEE Access*, vol. 8, pp. 150964–150973, 2020.

[23] B. Ren, Y. Wang, S. Kang, X. Dai, X. Yue, W. Tang, and K. Niu, "Link performance estimation technique for pdma uplink system," *IEEE Access*, vol. 5, pp. 15571–15581, 2017.

[24] M. Moghimi, A. Zakeri, M. R. Javan, N. Mokari, and D. W. K. Ng, "Joint radio resource allocation and cooperative caching in pd-noma-based hetnets," *IEEE Transactions on Mobile Computing*, pp. 1–1, 2020.

[25] Y. V. Kryukov, D. A. Pokamestov, E. V. Rogozhnikov, A. Y. Demidov, and Y. S. Gromova, "Experimental research of pd/noma," in *2018 19th International Conference of Young Specialists on Micro/Nanotechnologies and Electron Devices (EDM)*, pp. 176–179, 2018.

[26] A. Mokdad, P. Azmi, N. Mokari, M. Moltafet, and M. Ghaffari-Miab, "Cross-layer energy efficient resource allocation in pd-noma based h-crans: Implementation via gpu," *IEEE Transactions on Mobile Computing*, vol. 18, no. 6, pp. 1246–1259, 2019.

[27] Q. Luo, P. Gao, Z. Liu, L. Xiao, Z. Mheich, P. Xiao, and A. Maaref, "An error rate comparison of power domain non-orthogonal multiple access and sparse code multiple access," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 500–511, 2021.

[28] Z. Liu, P. Xiao, and Z. Mheich, "Power-imbalanced low-density signatures (lds) from eisenstein numbers," in *2019 IEEE VTS Asia Pacific Wireless Communications Symposium (APWCS)*, pp. 1–5, 2019.

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

[30] R. Razavi, R. Hoshyar, M. A. Imran, and Y. Wang, "Information theoretic analysis of lds scheme," *IEEE Communications Letters*, vol. 15, no. 8, pp. 798–800, 2011.

[31] A. Masmoudi, F. Bellili, S. Affes, and A. Ghrayeb, "Maximum likelihood time delay estimation from single- and multi-carrier dsss multipath mimo transmissions for future 5g networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 4851–4865, 2017.

[32] R. Razavi, M. AL-Imari, M. A. Imran, R. Hoshyar, and D. Chen, "On receiver design for uplink low density signature ofdm (lds-ofdm)," *IEEE Transactions on Communications*, vol. 60, no. 11, pp. 3499–3508, 2012.

[33] L. Wen, R. Razavi, M. A. Imran, and P. Xiao, "Design of joint sparse graph for ofdm system," *IEEE Transactions on Wireless Communications*, vol. 14, no. 4, pp. 1823–1836, 2015.

[34] B. Fontana da Silva, C. A. Azurdia Meza, D. Silva, and B. F. Uchôa-Filho, "Exploiting spatial diversity in overloaded mimo lds-ofdm multiple access systems," in *2017 IEEE 9th Latin-American Conference on Communications (LATINCOM)*, pp. 1–6, 2017.

[35] L. Meylani, I. Hidayat, A. Kurniawan, and M. S. Arifianto, "Power allocation for group lds-ofdm in underlay cognitive radio," in *2019 11th International Conference on Information Technology and Electrical Engineering (ICITEE)*, pp. 1–5, 2019.

[36] J. Zhang, X. Wang, X. Yang, and H. Zhou, "Low density spreading signature vector extension (lds-sve) for uplink multiple access," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pp. 1–5, 2017.

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

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

[39] B. I. Ahmed and F. Jabeen, "Blind adaptive beamforming simulation using ncma for smart antenna," in *2017 IEEE 7th International Advance Computing Conference (IACC)*, pp. 492–495, 2017.

[40] H. Pan, L. Lu, and S. C. Liew, "Network-coded multiple access with high-order modulations," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 9776–9792, 2017.

[41] B. Xu, Z. Xiang, P. Ren, and X. Guo, "Outage performance of downlink full-duplex network-coded cooperative noma," *IEEE Wireless Communications Letters*, vol. 10, no. 1, pp. 26–29, 2021.

[42] M. Attar, Y. Chen, S. A. Cheema, T. Wild, and M. Haardt, "Noca versus idma using ufmc for 5g multiple access," in *WSA 2018; 22nd International ITG Workshop on Smart Antennas*, pp. 1–5, 2018.

[43] S. Tachikawa and T. Ueki, "Characteristics of ofdm/cdma system using frequency and time domain spreading," in *2006 IEEE Ninth International Symposium on Spread Spectrum Techniques and Applications*, pp. 148–152, 2006.

[44] P. Frenger, P. Orten, and T. Ottosson, "Code-spread cdma using maximum free distance low-rate convolutional codes," *IEEE Transactions on Communications*, vol. 48, no. 1, pp. 135–144, 2000.

[45]

[46] T. Nonaka, T. Fujii, O. Takyu, and M. Ohta, "Adapting the number of replicas in the e-irsa system using the power control," in *2020 International Conference on Information Networking (ICOIN)*, pp. 787–792, 2020.

[47] 3GPP, "Low code rate and signature based multiple access scheme for new radio," *3GPP R1-164869*, vol. TSG-RAN, no. WG1, 2016.

[48] Z. Yang, M. Chen, W. Saad, W. Xu, and M. Shikh-Bahaei, "Sum-rate maximization of uplink rate splitting multiple access (rsma) communication," *IEEE Transactions on Mobile Computing*, pp. 1–1, 2020.

[49] Z. Wang, D. Zhang, K. Xu, W. Xie, J. Xu, and X. Li, "Beam-superposition-based multi-beam rsma for hybrid mmwave systems," in *2019 IEEE 19th International Conference on Communication Technology (ICCT)*, pp. 1086–1092, 2019.

[50] A. F. Klaib and H. Osborne, "Rsma matching algorithm for searching biological sequences," in *2009 International Conference on Innovations in Information Technology (IIT)*, pp. 195–199, 2009.

[51] Y. Liu, X. Xiong, and Z. Luo, "Effect of carrier frequency offsets on ofdm-idma systems," in *2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet)*, pp. 299–302, 2012.

[52] 3GPP, "Non-orthogonal multiple access candidate for nr," *3GPP R1-163992*, vol. TSG-RAN, no. WG1, 2016.

[53] A. I. Sulyman and M. Hefnawi, "Adaptive mimo beamforming algorithm based on gradient search of the channel capacity in ofdm-sdma systems," *IEEE Communications Letters*, vol. 12, no. 9, pp. 642–644, 2008.

[54] Z. Doukha and S. Moussaoui, "An sdma-based mechanism for accurate and efficient neighborhood-discovery link-layer service," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 2, pp. 603–613, 2016.

[55] B. Clerckx, Y. Mao, R. Schober, and H. V. Poor, "Rate-splitting unifying sdma, oma, noma, and multicasting in miso broadcast channel: A simple two-user rate analysis," *IEEE Wireless Communications Letters*, vol. 9, no. 3, pp. 349–353, 2020.

[56] C.-L. Wang and T.-H. Liou, "Improved modulation schemes for lattice-partition-based downlink non-orthogonal multiple access systems," *IEEE Wireless Communications Letters*, vol. 9, no. 12, pp. 2130–2134, 2020.

[57] M. Qiu, Y.-C. Huang, S.-L. Shieh, and J. Yuan, "A lattice-partition framework of downlink non-orthogonal multiple access without sic," *IEEE Transactions on Communications*, vol. 66, no. 6, pp. 2532–2546, 2018.

[58] M. Qiu, Y.-C. Huang, J. Yuan, and C.-L. Wang, "Lattice-partition-based downlink non-orthogonal multiple access without sic for slow fading channels," *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1166–1181, 2019.

[59] M. A. Naim, J. P. Fonseka, and E. M. Dowling, "A building block approach for designing multilevel coding schemes," *IEEE Communications Letters*, vol. 19, no. 1, pp. 2–5, 2015.

[60] J. Kim, G. Lee, S. Kim, T. Taleb, S. Choi, and S. Bahk, "Two-step random access for 5g system: Latest trends and challenges," *IEEE Network*, vol. 35, no. 1, pp. 273–279, 2021.

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

[62] J. Zhang, L. Lu, Y. Sun, Y. Chen, J. Liang, J. Liu, H. Yang, S. Xing, Y. Wu, J. Ma, *et al.*, "PoC of SCMA-based uplink grant-free transmission in UCNC for 5G," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1353–1362, 2017.

[63] M. B. Shahab, R. Abbas, M. Shirvanimoghaddam, and S. J. Johnson, "Grant-free non-orthogonal multiple access for iot: A survey," *IEEE Communications Surveys & Tutorials*, 2020.

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

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

[66] R. Stoica, G. T. F. De Abreu, T. Hara, and K. Ishibashi, "Massively concurrent non-orthogonal multiple access for 5g networks and beyond," *IEEE Access*, vol. 7, pp. 82080–82100, 2019.

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

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

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

[70] M. E. Tarerefa, O. Falowo, and N. Ventura, "Energy efficient user association scheme in cellular machine to machine communications," in *AFRICON, 2017 IEEE*, (Cape Town, South Africa), pp. 192–197, IEEE, Sep. 2017.

[71] G. J. Sutton, R. P. Liu, and Y. J. Guo, "Harmonising coexistence of machine type communications with Wi-Fi data traffic under frame-based LBT," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 4000–4011, 2017.

[72] S. Hu, H. Guo, C. Jin, Y. Huang, B. Yu, and S. Li, "Frequency-domain oversampling for cognitive CDMA systems: Enabling robust and massive multiple access for internet of things," *IEEE Access*, vol. 4, pp. 4583–4589, 2016.

[73] A. Kiani and N. Ansari, "Edge computing aware NOMA for 5G networks," *IEEE Internet of Things Journal*, 2018.

[74] H. Huawei, "R1-164036 multiple access for UL small packets transmission," in *3GPP TSG RAN WG1 Meeting 85*, 3GPP, 2016.

[75] M. Vaezi, G. A. Aruma Baduge, Y. Liu, A. Arafa, F. Fang, and Z. Ding, "Interplay between noma and other emerging technologies: A survey," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 4, pp. 900–919, 2019.

[76] M. Mohammadkarimi, M. A. Raza, and O. A. Dobre, "Signature-based nonorthogonal massive multiple access for future wireless networks: Uplink massive connectivity for machine-type communications," *IEEE Vehicular Technology Magazine*, vol. 13, no. 4, pp. 40–50, 2018.

[77] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A survey of non-orthogonal multiple access for 5g," *IEEE Communications Surveys Tutorials*, vol. 20, no. 3, pp. 2294–2323, 2018.

[78] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, "Concept and practical considerations of non-orthogonal multiple access (noma) for future radio access," in *2013 International Symposium on Intelligent Signal Processing and Communication Systems*, pp. 770–774, 2013.

[79] C. Yan, A. Harada, A. Benjebbour, Y. Lan, A. Li, and H. Jiang, "Receiver design for downlink non-orthogonal multiple access (noma)," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, pp. 1–6, 2015.

[80] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5g systems with randomly deployed users," *IEEE Signal Processing Letters*, vol. 21, no. 12, pp. 1501–1505, 2014.

[81] M. B. Shahab, M. F. Kader, and S. Y. Shin, "A virtual user pairing scheme to optimally utilize the spectrum of unpaired users in non-orthogonal multiple access," *IEEE Signal Processing Letters*, vol. 23, no. 12, pp. 1766–1770, 2016.

[82] M. Al-Imari, P. Xiao, M. A. Imran, and R. Tafazolli, "Uplink non-orthogonal multiple access for 5g wireless networks," in *2014 11th International Symposium on Wireless Communications Systems (ISWCS)*, pp. 781–785, 2014.

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

[84] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh, "Scma codebook design," in *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, pp. 1–5, 2014.

[85] Y. Wu, S. Zhang, and Y. Chen, "Iterative multiuser receiver in sparse code multiple access systems," in *2015 IEEE International Conference on Communications (ICC)*, pp. 2918–2923, 2015.

[86] Y. Du, B. Dong, Z. Chen, J. Fang, and X. Wang, "A fast convergence multiuser detection scheme for uplink scma systems," *IEEE Wireless Communications Letters*, vol. 5, no. 4, pp. 388–391, 2016.

[87] H. Mu, Z. Ma, M. Alhaji, P. Fan, and D. Chen, "A fixed low complexity message pass algorithm detector for up-link scma system," *IEEE Wireless Communications Letters*, vol. 4, no. 6, pp. 585–588, 2015.

[88] A. Brencic and S. Lory, "Determination of the regulon and identification of novel mRNA targets of pseudomonas aeruginosa RsmA," *Molecular microbiology*, vol. 72, no. 3, pp. 612–632, 2009.

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

[90] A.-L. S. B. Nokia and L. S. Bell, "Performance of interleave division multiple access (IDMA) in combination with OFDM family waveforms," in *R1-165021, 3GPP TSG RAN WG1 Meeting*, vol. 85, 2016.

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

[92] P. Wang, J. Xiao, and L. Ping, "Comparison of orthogonal and non-orthogonal approaches to future wireless cellular systems," *IEEE Vehicular Technology Magazine*, vol. 1, no. 3, pp. 4–11, 2006.

[93] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, I. Chih-Lin, and H. V. Poor, "Application of non-orthogonal multiple access in lte and 5g networks," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 185–191, 2017.

[94] J. Kim, K. Lee, J. Kim, H. Wang, M. Na, and D. Hong, "A novel SCMA system for coexistence of active users and inactive users," *IEEE Communications Letters*, vol. 21, no. 12, pp. 2730–2733, 2017.

[95] G. Ma, B. Ai, F. Wang, and Z. Zhong, "Coded tandem spreading for grant-free random access system with massive connections," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*, pp. 1–6, IEEE, 2017.

[96] A. T. Abebe and C. G. Kang, "Comprehensive grant-free random access for massive & low latency communication," in *Communications (ICC), 2017 IEEE International Conference on*, pp. 1–6, IEEE, 2017.

[97] V. W. Wong, R. Schober, D. W. K. Ng, and L.-C. Wang, "Overview of new technologies for 5G systems," *Key Technologies for 5G Wireless Systems*, p. 1, 2017.

[98] J. Choi, N. Lee, S. Hong, and G. Caire, "Joint user selection, power allocation, and precoding design with imperfect CSIT for multi-cell MU-MIMO downlink systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 162–176, 2019.

[99] M. Caus, A. I. Pérez-Neira, J. Bas, and L. Blanco, "New satellite random access preamble design based on pruned DFT-spread FBMC," *IEEE Transactions on Communications*, 2020.

[100] A. Venkatasubramanian, V. Krithika, and B. Partibane, "Channel estimation for a multi-user MIMO-OFDM-IDMA system," in *2017 International Conference on Communication and Signal Processing (ICCSP)*, pp. 1823–1827, IEEE, 2017.

[101] C. Xu, Y. Hu, C. Liang, J. Ma, and L. Ping, "Massive MIMO, non-orthogonal multiple access and interleave division multiple access," *IEEE Access*, vol. 5, pp. 14728–14748, 2017.

[102] M. B. Shahab, M. Irfan, M. F. Kader, and S. Young Shin, "User pairing schemes for capacity maximization in non-orthogonal multiple access systems," *Wireless Communications and Mobile Computing*, vol. 16, no. 17, pp. 2884–2894, 2016.

[103] M. B. Shahab, M. F. Kader, and S. Y. Shin, "On the power allocation of non-orthogonal multiple access for 5g wireless networks," in *2016 International Conference on Open Source Systems Technologies (ICOSST)*, pp. 89–94, IEEE, 2016.

[104] J. Hong, W. Choi, and B. D. Rao, "Sparsity controlled random multiple access with compressed sensing," *IEEE Transactions on Wireless Communications*, vol. 14, no. 2, pp. 998–1010, 2015.

[105] F. Fazel, M. Fazel, and M. Stojanovic, "Random access compressed sensing over fading and noisy communication channels," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 2114–2125, 2013.

[106] N. Y. Yu, "Multiuser activity and data detection via sparsity-blind greedy recovery for uplink grant-free noma," *IEEE Communications Letters*, vol. 23, no. 11, pp. 2082–2085, 2019.

[107] S. Zhuang, W. Zhao, R. Wang, Q. Wang, and S. Huang, "New measurement algorithm for supraharmonics based on multiple measurement vectors model and or-

thogonal matching pursuit," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 6, pp. 1671–1679, 2019.

[108] L. Wang, D. Wang, and C. Hao, "A multiple-measurement vectors reconstruction method for low snr scenarios," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 4, pp. 785–789, 2020.

[109] F. Monsees, C. Bockelmann, and A. Dekorsy, "Reliable activity detection for massive machine to machine communication via multiple measurement vector compressed sensing," in *2014 IEEE Globecom Workshops (GC Wkshps)*, pp. 1057–1062, IEEE, 2014.

[110] F. Monsees, M. Woltering, C. Bockelmann, and A. Dekorsy, "Compressive sensing multi-user detection for multicarrier systems in sporadic machine type communication," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, pp. 1–5, IEEE, 2015.

[111] A. T. Abebe and C. G. Kang, "Compressive sensing-based random access with multiple-sequence spreading for mtc," in *2015 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6, IEEE, 2015.

[112] G. Chen, J. Dai, K. Niu, and C. Dong, "Sparsity-inspired sphere decoding (SI-SD): A novel blind detection algorithm for uplink grant-free sparse code multiple access," *IEEE Access*, vol. 5, pp. 19983–19993, 2017.

[113] J. Liu, G. Wu, S. Li, and O. Tirkkonen, "Blind detection of uplink grant-free scma with unknown user sparsity," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2017.

[114] A. T. Abebe and C. G. Kang, "Comprehensive grant-free random access for massive low latency communication," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2017.

[115] J. Tan, W. Ding, F. Yang, C. Pan, and J. Song, "Compressive sensing based time-frequency joint non-orthogonal multiple access," in *2016 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pp. 1–4, IEEE, 2016.

[116] B. K. Jeong, B. Shim, and K. B. Lee, "MAP-based active user and data detection for massive machine-type communications," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8481–8494, 2018.

[117] B. Shim and B. Song, "Multiuser detection via compressive sensing," *IEEE Communications Letters*, vol. 16, no. 7, pp. 972–974, 2012.

[118] B. Wang, L. Dai, Y. Zhang, T. Mir, and J. Li, "Dynamic compressive sensing-based multi-user detection for uplink grant-free NOMA," *IEEE Communications Letters*, vol. 20, pp. 2320–2323, Nov 2016.

[119] J. Xiao, G. Deng, G. Nie, H. Tian, and J. Jin, "Dynamic adaptive compressive sensing-based multi-user detection in uplink urllc," in *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1–6, IEEE, 2018.

[120] N. Vaswani and J. Zhan, "Recursive recovery of sparse signal sequences from compressive measurements: A review," *IEEE Transactions on Signal Processing*, vol. 64, no. 13, pp. 3523–3549, 2016.

[121] H. S. Jang, B. C. Jung, T. Q. S. Quek, and D. K. Sung, "Resource-hopping-based grant-free multiple access for 6g-enabled massive iot networks," *IEEE Internet of Things Journal*, vol. 8, no. 20, pp. 15349–15360, 2021.

[122] D. M. Amitu, R. N. Akol, and P. Nakeba, "Qos-aware splitting and radio resource allocation for machine type communications," in *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 941–947, IEEE, 2018.

[123] U. Tefek and T. J. Lim, "Clustering and radio resource partitioning for machine-type communications in cellular networks," in *2016 IEEE Wireless Communications and Networking Conference*, pp. 1–6, IEEE, 2016.

[124] N. M. Ahmed and N.-E. Rikli, "Qos-based data aggregation and resource allocation algorithm for machine type communication devices in next-generation networks," *IEEE Access*, vol. 9, pp. 119735–119754, 2021.

[125] Y. Polyanskiy, "A perspective on massive random-access," in *2017 IEEE International Symposium on Information Theory (ISIT)*, pp. 2523–2527, IEEE, 2017.

[126] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6463–6486, 2011.

[127] L.-L. Yang, *Multicarrier communications*. John Wiley & Sons, 2009.

[128] J. Zhang, P. Pan, L.-L. Yang, and R. G. Maunder, "Channel estimation and user activity identification in massive grant-free multiple-access," *IEEE Open Journal of Vehicular Technology*, vol. 1, pp. 296–316, 2020.

[129] Z. Zhang, X. Wang, Y. Zhang, and Y. Chen, "Grant-free rateless multiple access: A novel massive access scheme for internet of things," *IEEE Communications Letters*, vol. 20, no. 10, pp. 2019–2022, 2016.

[130] M. Cheraghchi, A. Hormati, A. Karbasi, and M. Vetterli, "Group testing with probabilistic tests: Theory, design and application," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 7057–7067, 2011.

[131] A. Zulfiqar, I. Rashid, F. Akam, and S. Rabab, "Analysis of phase transition using deterministic matrix in compressed sensing," in *Applied Sciences and Technology (IBCAST), 2017 14th International Bhurban Conference on*, pp. 333–336, IEEE, 2017.

[132] E. D. Di Claudio, R. Parisi, and G. Jacovitti, "Space time music: Consistent signal subspace estimation for wide-band sensor arrays," *arXiv preprint arXiv:1711.01631*, 2017.

[133] N. Soudani and R. Bouallegue, "Comparative study of phase jitter models for uplink mc-cdma and uplink mc-ds-cdma," in *2006 International Conference on Wireless and Mobile Communications (ICWMC'06)*, pp. 37–37, 2006.

[134] Y. Lee, "Performance analysis of ds-cdma and mc-cdma systems," in *2018 9th IEEE Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*, pp. 842–846, 2018.

[135] S. Mathur, S. S. Sagari, S. O. Amin, R. Ravindran, D. Saha, I. Seskar, D. Raychaudhuri, and G. Wang, "Demo abstract: CDMA-based IoT services with shared band operation of LTE in 5G," *arXiv preprint arXiv:1705.06968*, 2017.

[136] J. Sayritupac, E. Vanet, R. Caire, and C. Larios, "Behaviour analysis of an operational planning tool facing activation probabilities, for near optimal operation of smart grids," *CIRED-Open Access Proceedings Journal*, vol. 2017, no. 1, pp. 2127–2130, 2017.

[137] L. Liu and W. Yu, "Massive connectivity with massive MIMO - part I: Device activity detection and channel estimation," *IEEE Transactions on Signal Processing*, 2018.

[138] G. Berardinelli, N. H. Mahmood, R. Abreu, T. Jacobsen, K. Pedersen, I. Z. Kovács, and P. Mogensen, "Reliability analysis of uplink grant-free transmission over shared resources," *IEEE Access*, 2018.

[139] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*, vol. 2. prentice hall PTR New Jersey, 1996.

[140] A. C. Cirik, N. Mysore Balasubramanya, and L. Lampe, "Multi-user detection using ADMM-based compressive sensing for uplink grant-free NOMA," *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 46–49, 2018.

[141] T. Al-Khasib, L. Lampe, and H. M. Alnuweiri, "Uplink multiple-user v-blast optimal detection ordering with service differentiation," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2308–2319, 2010.

[142] L.-L. Yang, "Receiver multiuser diversity aided multi-stage minimum mean-square error detection for heavily loaded DS-CDMA and SDMA systems," *IEEE Transactions on communications*, vol. 58, no. 12, pp. 3397–3404, 2010.

[143] J. G. Proakis, *Digital Communications*. Mcgraw Hill, 4 ed., 2000.

[144] M. Masoudi, A. Azari, E. A. Yavuz, and C. Cavdar, "Grant-free radio access IoT networks: Scalability analysis in coexistence scenarios," in *2018 IEEE International Conference on Communications (ICC)*, pp. 1–7, IEEE, 2018.

[145] B. Wang, L. Dai, T. Mir, and Z. Wang, "Joint user activity and data detection based on structured compressive sensing for noma," *IEEE Communications Letters*, vol. 20, no. 7, pp. 1473–1476, 2016.

[146] Y. Du, B. Dong, Z. Chen, X. Wang, Z. Liu, P. Gao, and S. Li, "Efficient multi-user detection for uplink grant-free NOMA: Prior-information aided adaptive compressive sensing perspective," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 12, pp. 2812–2828, 2017.

[147] F. Wei, W. Chen, Y. Wu, J. Ma, and T. A. Tsiftsis, "Message-passing receiver design for joint channel estimation and data decoding in uplink grant-free SCMA systems," *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 167–181, 2019.

[148] Wei-Chiang Wu and Kwang-Cheng Chen, "Identification of active users in synchronous CDMA multiuser detection," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1723–1735, Dec 1998.

[149] A. Haghighat and M. R. Soleymani, "A subspace scheme for blind user identification in multiuser ds-cdma," in *2003 IEEE Wireless Communications and Networking, 2003. WCNC 2003.*, vol. 1, pp. 688–692 vol.1, 2003.

[150] A. Haghighat and M. R. Soleymani, "A blind music-based algorithm for user identification in multiuser ds-cdma," in *GLOBECOM '03. IEEE Global Telecommunications Conference (IEEE Cat. No.03CH37489)*, vol. 5, pp. 2447–2452 vol.5, 2003.

[151] D. D. Lin and Teng Joon Lim, "Subspace-based active user identification for a collision-free slotted ad hoc network," *IEEE Transactions on Communications*, vol. 52, no. 4, pp. 612–621, 2004.

[152] G. Hannak, M. Mayer, A. Jung, G. Matz, and N. Goertz, "Joint channel estimation and activity detection for multiuser communication systems," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, pp. 2086–2091, 2015.

[153] C. Wei, H. Liu, Z. Zhang, J. Dang, and L. Wu, "Approximate message passing-based joint user activity and data detection for NOMA," *IEEE Communications Letters*, vol. 21, no. 3, pp. 640–643, 2017.

[154] O. O. Oyerinde, "Compressive sensing algorithms for multiuser detection in uplink grant free NOMA systems," in *2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*, pp. 1–6, 2019.

[155] Y. Zhang, Q. Guo, Z. Wang, J. Xi, and N. Wu, "Block sparse bayesian learning based joint user activity detection and channel estimation for grant-free NOMA systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9631–9640, 2018.

[156] Y. Du, B. Dong, W. Zhu, P. Gao, Z. Chen, X. Wang, and J. Fang, "Joint channel estimation and multiuser detection for uplink grant-free noma," *IEEE Wireless Communications Letters*, vol. 7, no. 4, pp. 682–685, 2018.

[157] J. Zhang, Y. Pan, and J. Xu, "Compressive sensing for joint user activity and data detection in grant-free NOMA," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 857–860, 2019.

[158] A. T. Abebe and C. G. Kang, "Joint channel estimation and MUD for scalable grant-free random access," *IEEE Communications Letters*, vol. 23, no. 12, pp. 2229–2233, 2019.

[159] T. Ding, X. Yuan, and S. C. Liew, "Sparsity learning-based multiuser detection in grant-free massive-device multiple access," *IEEE Transactions on Wireless Communications*, vol. 18, no. 7, pp. 3569–3582, 2019.

[160] G. Chen, Y. Cui, H. V. Cheng, F. Yang, and L. Ding, "Analysis and optimization of successful symbol transmission rate for grant-free massive access with massive MIMO," *IEEE Communications Letters*, vol. 23, no. 12, pp. 2381–2385, 2019.

[161] X. Miao, D. Guo, and X. Li, "Grant-free NOMA with device activity learning using long short-term memory," *IEEE Wireless Communications Letters*, vol. 9, no. 7, pp. 981–984, 2020.

[162] W. Kim, Y. Ahn, and B. Shim, "Deep neural network-based active user detection for grant-free NOMA systems," *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2143–2155, 2020.

[163] S. Theodoridis, *Machine Learning: A Bayesian and Optimization Perspective*. London, UK: Academic Press, 2015.

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

[165] W. Yuan, N. Wu, J. A. Zhang, X. Huang, Y. Li, and L. Hanzo, "Iterative receiver design for ftn signaling aided sparse code multiple access," *IEEE Transactions on Wireless Communications*, 2019.

[166] Z. Pan, J. Lei, W. Liu, J. Luo, and C. Tang, "Grant-free rateless SCMA for cellular internet of things networks," *IEEE Access*, vol. 7, pp. 147954–147961, 2019.

[167] Z. Ding, R. Schober, P. Fan, and H. V. Poor, "Simple semi-grant-free transmission strategies assisted by non-orthogonal multiple access," *IEEE Transactions on Communications*, vol. 67, no. 6, pp. 4464–4478, 2019.

[168] J. Ahn, B. Shim, and K. B. Lee, "EP-based joint active user detection and channel estimation for massive machine-type communications," *IEEE Transactions on Communications*, vol. 67, no. 7, pp. 5178–5189, 2019.

[169] H. Jiang, D. Qu, J. Ding, and T. Jiang, "Multiple preambles for high success rate of grant-free random access with massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 18, no. 10, pp. 4779–4789, 2019.

[170] J. Ding and J. Choi, "Comparison of preamble structures for grant-free random access in massive MIMO systems," *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 166–170, 2020.

[171] J. Ding, D. Qu, H. Jiang, and T. Jiang, "Success probability of grant-free random access with massive MIMO," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 506–516, 2019.

[172] H. Han, Y. Li, W. Zhai, and L. Qian, "A grant-free random access scheme for M2M communication in massive MIMO systems," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3602–3613, 2020.

[173] J. Wang, Z. Zhang, and L. Hanzo, "Joint active user detection and channel estimation in massive access systems exploiting Reed-Muller sequences," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 3, pp. 739–752, 2019.

[174] S. Kim, H. Kim, H. Noh, Y. Kim, and D. Hong, "Novel transceiver architecture for an asynchronous grant-free IDMA system," *IEEE Transactions on Wireless Communications*, vol. 18, no. 9, pp. 4491–4504, 2019.

[175] J. Fu, G. Wu, Y. Zhang, L. Deng, and S. Fang, "Active user identification based on asynchronous sparse bayesian learning with SVM," *IEEE Access*, vol. 7, pp. 108116–108124, 2019.

[176] Z. Zhang, Y. Li, C. Huang, Q. Guo, C. Yuen, and Y. L. Guan, "DNN-aided block sparse bayesian learning for user activity detection and channel estimation in grant-free non-orthogonal random access," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 12, pp. 12000–12012, 2019.

[177] H. Yu, Z. Fei, Z. Zheng, and N. Ye, "Finite-alphabet signature design for grant-free NOMA: A quantized deep learning approach," *IEEE Transactions on Vehicular Technology*, pp. 1–1, 2020.

[178] T. Kim and B. C. Jung, "Performance analysis of grant-free multiple access for supporting sporadic traffic in massive IoT networks," *IEEE Access*, vol. 7, pp. 166648–166656, 2019.

[179] B. Wang, L. Dai, T. Mir, and Z. Wang, "Joint user activity and data detection based on structured compressive sensing for NOMA," *IEEE Communications Letters*, vol. 20, no. 7, pp. 1473–1476, 2016.

[180] Y. Du, B. Dong, Z. Chen, X. Wang, Z. Liu, P. Gao, and S. Li, "Efficient multi-user detection for uplink grant-free NOMA: Prior-information aided adaptive compressive sensing perspective," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 12, pp. 2812–2828, 2017.

[181] J. Zhang, Y. Pan, and J. Xu, "Compressive sensing for joint user activity and data detection in grant-free NOMA," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 857–860, 2019.

[182] Y. Du, B. Dong, W. Zhu, P. Gao, Z. Chen, X. Wang, and J. Fang, "Joint channel estimation and multiuser detection for uplink grant-free NOMA," *IEEE Wireless Communications Letters*, vol. 7, no. 4, pp. 682–685, 2018.

[183] B. K. Jeong, B. Shim, and K. B. Lee, "MAP-based active user and data detection for massive machine-type communications," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8481–8494, 2018.

[184] G. Chen, Y. Cui, H. V. Cheng, F. Yang, and L. Ding, "Analysis and optimization of successful symbol transmission rate for grant-free massive access with massive MIMO," *IEEE Communications Letters*, vol. 23, no. 12, pp. 2381–2385, 2019.

[185] T. Ding, X. Yuan, and S. C. Liew, "Sparsity learning-based multiuser detection in grant-free massive-device multiple access," *IEEE Transactions on Wireless Communications*, vol. 18, no. 7, pp. 3569–3582, 2019.

[186] C. Wei, H. Liu, Z. Zhang, J. Dang, and L. Wu, "Approximate message passing-based joint user activity and data detection for NOMA," *IEEE Communications Letters*, vol. 21, no. 3, pp. 640–643, 2016.

[187] Y. Zhang, Q. Guo, Z. Wang, J. Xi, and N. Wu, "Block sparse bayesian learning based joint user activity detection and channel estimation for grant-free NOMA systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9631–9640, 2018.

[188] A. T. Abebe and C. G. Kang, "Grant-free uplink transmission with multi-codebook-based sparse code multiple access (MC-SCMA)," *IEEE Access*, vol. 7, pp. 169853–169864, 2019.

[189] E. Dahlman, S. Parkvall, and J. Skold, *4G LTE/LTE-Advanced for Mobile Broadband*. Oxford, UK: Academic Press Inc, 2 ed., 2013.

[190] S. S. Haykin, *Adaptive filter theory*. Pearson Education India, 2008.

[191] H. V. Henderson and S. R. Searle, "The vec-permutation matrix, the vec operator and kronecker products: A review," *Linear and multilinear algebra*, vol. 9, no. 4, pp. 271–288, 1981.

[192] J. G. Proakis, *Digital signal processing: principles algorithms and applications*. Pearson Education India, 2001.

[193] R. A. Brualdi, H. J. Ryser, *et al.*, *Combinatorial matrix theory*. No. 39, Cambridge University Press, 1991.

[194] S. Lin and D. J. Costello, *Error control coding*, vol. 2. Prentice hall Scarborough, 2001.

[195] G. H. Golub and G. Meurant, "Matrices, moments and quadrature ii; how to compute the norm of the error in iterative methods," *BIT Numerical Mathematics*, vol. 37, no. 3, pp. 687–705, 1997.

[196] W. H. Press, W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, and W. T. Vetterling, *Numerical recipes in Pascal: the art of scientific computing*, vol. 1. Cambridge university press, 1989.

[197] F. Schoute, "Dynamic frame length aloha," *IEEE Transactions on communications*, vol. 31, no. 4, pp. 565–568, 1983.

[198] F. Baccelli, B. Blaszczyszyn, and P. Muhlethaler, "Stochastic analysis of spatial and opportunistic aloha," *IEEE journal on selected areas in communications*, vol. 27, no. 7, pp. 1105–1119, 2009.

[199] S.-J. Ku, C.-L. Wang, and C.-H. Chen, "A reduced-complexity PTS-based PAPR reduction scheme for OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 9, no. 8, pp. 2455–2460, 2010.

[200] R. Huang, V. W. Wong, and R. Schober, "Throughput optimization for grant-free multiple access with multiagent deep reinforcement learning," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 228–242, 2021.

[201] T. Jiang, Y. Shi, J. Zhang, and K. B. Letaief, "Joint activity detection and channel estimation for iot networks: Phase transition and computation-estimation tradeoff," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6212–6225, 2018.

[202] X. Shao, X. Chen, and R. Jia, "A dimension reduction-based joint activity detection and channel estimation algorithm for massive access," *IEEE Transactions on Signal Processing*, vol. 68, pp. 420–435, 2019.

[203] E. Olfat and M. Bengtsson, "Joint channel and clipping level estimation for OFDM in iot-based networks," *IEEE Transactions on Signal Processing*, vol. 65, no. 18, pp. 4902–4911, 2017.

[204] M. Alam and Q. Zhang, "Qos-aware noma with sequence block compressed sensing multiuser detection," in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, 2019.

[205] C.-B. Song and S.-T. Xia, "Sparse signal recovery by ¡formula formulatype="inline"¿ ¡tex notation="tex"¿ℓ_q¡/tex¿¡/formula¿ minimization under restricted isometry property," *IEEE Signal Processing Letters*, vol. 21, no. 9, pp. 1154–1158, 2014.

[206] Y. Zhang, Z. Yuan, Q. Guo, Z. Wang, J. Xi, and Y. Li, "Bayesian receiver design for grant-free noma with message passing based structured signal estimation," *IEEE Transactions on Vehicular Technology*, 2020.

[207] Z. Pan, J. Lei, W. Liu, J. Luo, and C. Tang, "Grant-free rateless scma for cellular internet of things networks," *IEEE Access*, vol. 7, pp. 147954–147961, 2019.

[208] X. Ma, L. Yang, Z. Chen, and Y. Siu, "Low complexity detection based on dynamic factor graph for scma systems," *IEEE Communications Letters*, vol. 21, no. 12, pp. 2666–2669, 2017.

[209] M. Jia, L. Wang, Q. Guo, X. Gu, and W. Xiang, "A low complexity detection algorithm for fixed up-link scma system in mission critical scenario," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3289–3297, 2017.

[210] Y. Yu, M. Pischella, and D. Le Ruyet, "Distributed antenna selection with message passing algorithm for mimo d2d communications," in *2017 International Symposium on Wireless Communication Systems (ISWCS)*, pp. 454–458, IEEE, 2017.

[211] T. Wei, Z. Xingcheng, L. Yuxiang, G. Shaoxiang, and F. Wenjiang, "Reduced-complexity mpa decoder based on multi-level dynamic input thresholds," *The Journal of China Universities of Posts and Telecommunications*, vol. 24, no. 1, pp. 40–46, 2017.

[212] K. Jiang, H. Wang, and B. He, "Fault location in active distribution networks using block sparse bayesian learning," in *2019 IEEE Sustainable Power and Energy Conference (iSPEC)*, pp. 1279–1284, IEEE, 2019.

[213] S. Sharma, S. Chaudhury, *et al.*, "Variational bayes block sparse modeling with correlated entries," in *2018 24th International Conference on Pattern Recognition (ICPR)*, pp. 1313–1318, IEEE, 2018.

[214] Y. Takaoka, M. Miyoshi, K. Sakaguchi, S. Morisada, K. Ohto, and H. Kawakita, "Recovery of filtered graphene oxide residue using elastic gel packed in a column by cross flow," *Processes*, vol. 6, no. 5, p. 43, 2018.

[215] P. Pan and L.-L. Yang, "Spatially modulated code-division multiple-access for high-connectivity multiple access," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 4031–4046, 2019.

[216] Z. Wei, F. Wang, G. Ma, B. Ai, and Z. Zhong, "Tandem spreading multiple access with successive interference cancellation," in *2019 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 427–431, IEEE, 2019.

[217] Z. Xinyu, "Analysis of m-sequence and gold-sequence in cdma system," in *2011 IEEE 3rd International Conference on Communication Software and Networks*, (Xi'an, China), pp. 466–468, IEEE, 2011.

[218] J. Kim, K. S. Kim, *et al.*, "User detection performance analysis for grant-free uplink transmission in large-scale antenna systems," in *2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN)*, pp. 352–354, IEEE, 2018.

[219] S. M. Kay, *Fundamentals of statistical signal processing*, vol. 2. Prentice Hall PTR, 1993.

[220] J. Zhang, P. Pan, and L.-L. Yang, "Dynamic ds-cdma aided by successive interference cancellation for massive grant-free multiple-access," in *2018 10th International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1–7, 2018.

[221] J. Benesty, Y. Huang, and J. Chen, "A fast recursive algorithm for optimum sequential signal detection in a BLAST system," *IEEE Transactions on Signal Processing*, vol. 51, pp. 1722 – 1730, July 2003.

[222] Z. Miao, Y. Wang, and Z. Han, "Nucleolus-based profit sharing for wireless small cells in content centric networks," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2019.

[223] 3GPP, "NR; Physical layer procedures for control," Technical Specification (TS) 38.213, 3rd Generation Partnership Project (3GPP), 08 2017. Version Realease15.

[224] 3GPP, "Study on Narrow-Band Internet of Things (NB-IoT) / enhanced Machine Type Communication (eMTC) support for Non-Terrestrial Networks (NTN)," Technical Report (TR) 36.763, 3rd Generation Partnership Project (3GPP), 04 2020. Version Realease17.

[225] H. Zhu, "Performance comparison between distributed antenna and microcellular systems," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 6, pp. 1151–1163, 2011.

[226] Y. Cao, W. Su, and S. N. Batalama, "A novel receiver design and maximum-likelihood detection for distributed MIMO systems in presence of distributed frequency offsets and timing offsets," *IEEE Transactions on Signal Processing*, vol. 66, no. 23, pp. 6297–6309, 2018.

[227] L. Hanzo, M. El-Hajjar, and O. Alamri, "Near-capacity wireless transceivers and cooperative communications in the MIMO era: Evolution of standards, waveform design, and future perspectives," *Proceedings of the IEEE*, vol. 99, no. 8, pp. 1343–1385, 2011.

[228] A. S. Lioumpas, P. S. Bithas, and A. Alexiou, "Partitioning of distributed MIMO systems based on overhead considerations," *IEEE Wireless Communications Letters*, vol. 2, no. 6, pp. 579–582, 2013.

[229] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.

[230] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 77–90, 2020.

[231] J. Qiu, K. Xu, X. Xia, Z. Shen, and W. Xie, "Downlink power optimization for cell-free massive MIMO over spatially correlated rayleigh fading channels," *IEEE Access*, vol. 8, pp. 56214–56227, 2020.

[232] A. Abdallah and M. M. Mansour, "Efficient angle-domain processing for fdd-based cell-free massive mimo systems," *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2188–2203, 2020.

[233] S. K. Singh, R. Singh, and B. Kumbhani, "The evolution of radio access network towards Open-RAN: Challenges and opportunities," in *2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pp. 1–6, IEEE, 2020.

[234] S. Costanzo, I. Fajjari, N. Aitsaadi, and R. Langar, "DEMO: SDN-based network slicing in C-RAN," in *2018 15th IEEE Annual Consumer Communications Networking Conference (CCNC)*, pp. 1–2, IEEE, 2018.

[235] X. Xia, P. Zhu, J. Li, D. Wang, Y. Xin, and X. You, "Joint sparse beamforming and power control for a large-scale DAS with network-assisted full duplex," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 7, pp. 7569–7582, 2020.

[236] S. Buzzi and C. D'Andrea, "Cell-free massive MIMO: User-centric approach," *IEEE Wireless Communications Letters*, vol. 6, no. 6, pp. 706–709, 2017.

[237] H. S. Jang, B. C. Jung, T. Q. Quek, and D. K. Sung, "Resource-hopping based grant-free multiple access for 6g-enabled massive iot networks," *IEEE Internet of Things Journal*, 2021.

[238] A. Wall and I. Agrafiotis, "A bayesian approach to insider threat detection," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 12, no. 2, 2021.

[239] Y. Du, C. Cheng, B. Dong, Z. Chen, X. Wang, J. Fang, and S. Li, "Block-sparsity-based multiuser detection for uplink grant-free noma," *IEEE Transactions on Wireless Communications*, vol. 17, no. 12, pp. 7894–7909, 2018.

[240] E. Recayte, A. Munari, and F. Clazzer, "Grant-free access: Machine learning for detection of short packets," in *2020 10th Advanced Satellite Multimedia Systems Conference and the 16th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, pp. 1–7, IEEE, 2020.

[241] Y. Cui, S. Li, and W. Zhang, "Jointly sparse signal recovery and support recovery via deep learning with applications in mimo-based grant-free random access," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 788–803, 2021.

[242] A.-Y. Lin, P.-N. Chen, S.-L. Shieh, and Y.-C. Huang, "Generalized likelihood-ratio enabled machine learning for ue detection over grant-free scma," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, pp. 1–6, IEEE, 2020.

[243] S. Brugiapaglia, "A compressive spectral collocation method for the diffusion equation under the restricted isometry property," in *Quantification of Uncertainty: Improving Efficiency and Technology*, pp. 15–40, Springer, 2020.

[244] M. K. Simon and M. Alouini, *Digital communication over fading channels*, vol. 95. John Wiley & Sons, 2005.

[245] P. Pan, J. Zhang, and L. L. Yang, "Massive distributed antenna systems: Channel estimation and signal detection," *IEEE Access*, vol. 8, pp. 186055–186070, 2020.

[246] M. S. Zia and S. A. Hassan, "On the impacts of composite fading on large-scale multi-user mimo systems," *IEEE Communications Letters*, vol. 19, no. 12, pp. 2286–2289, 2015.

[247] S. K. Kalyankar, Y. H. Lee, and Y. S. Meng, "Two-slope path loss model for curved-tunnel environment with concept of break point," *IEEE Transactions on Intelligent Transportation Systems*, 2020.

[248] R. Huang, V. W. Wong, and R. Schober, "Throughput optimization for grant-free multiple access with multiagent deep reinforcement learning," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 228–242, 2020.

[249] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4680–4688, 2011.

[250] B. Blaszczyszyn and D. Yogeshwaran, "Connectivity in sub-poisson networks," in *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 1466–1473, IEEE, 2010.

[251] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in ofdm systems," *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 114–117, 2018.

[252] H. Kim, J. Kim, W. Shin, H. Yang, N. Lee, S. J. Kim, and J. Lee, "On the design of tailored neural networks for energy harvesting broadcast channels: A reinforcement learning approach," *IEEE Access*, vol. 8, pp. 179678–179691, 2020.

[253] J. Yuan, H. Q. Ngo, and M. Matthaiou, "Machine learning-based channel prediction in massive mimo with channel aging," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 2960–2973, 2020.