UNIVERSITY OF SOUTHAMPTON

Faculty of Engineering and Physical Science
School of Electronics and Computer Science

Optimization of Multicarrier-Division Duplex Wireless Systems

by

Bohan Li

ORCiD: 0000-0001-7686-8605

A thesis for the degree of
Doctor of Philosophy

Supervisor:

Prof. Lie-Liang Yang

and Prof. Robert G. Maunder

September 2022

Dedicated to my beloved parents and wife

University of Southampton

Abstract

Faculty of Engineering and Physical Science
School of Electronics and Computer Science

Doctor of Philosophy

Optimization of Multicarrier-Division Duplex Wireless Systems

by Bohan Li

Owing to the potentials of enabling to double spectral efficiency (SE) and reduce network latency, in-band full duplex (IBFD) has drawn a lot of research so as to substitute the conventional half duplex (HD) of time-division duplex (TDD) and frequency-division duplex (FDD). However, the self-interference (SI) problem has hindered IBFD from practical deployment. Inspired by the principles of full duplex (FD), in this thesis, a multicarrier-division duplex (MDD) scheme is introduced, which is capable of combining the advantages of IBFD and HD, while simultaneously circumventing their drawbacks.

Firstly, in order to make MDD feasible for the operation in large-scale multiple-input multiple-output (MIMO) systems, the thesis commences with addressing the SI problem in propagation-domain with considering the SI cancellation (SIC) requirement of practical analog-to-digital converter (ADC) at receiver. Then, the channel estimation (CE) in MDD MIMO systems is proposed by exploiting the reciprocity and correlation existing between the uplink (UL) and downlink (DL) subchannels. Secondly, the potential of MDD with resource allocation (RA) is first demonstrated, when an unfair greedy algorithm is applied for RA in the multiuser single-input single-output (MU-SISO) systems. Then, a suboptimal algorithm is proposed for MDD millimeter-wave (mmWave) MIMO systems to jointly maximize the sum-rate and achieve the proportional fairness among DL and UL mobile stations (MSs). Two mainstream hybrid precoding strategies are evaluated in the proposed RA scheme and the impact of insufficient SIC on RA is also studied. Thirdly, upon taking the advantages of the flexible time-frequency resource scheduling provided by MDD, two types of frame structures are designed to relieve the channel aging problem in high-mobility communication scenarios. Correspondingly, two Wiener-filtering based predictors (WPs) are introduced under the proposed frame structures for comparing the performance between MDD and TDD, when both CE and residual SI errors are invoked. Moreover, the closed-form expressions for approximating the lower bounded average sum rates of both MDD and TDD systems are derived, when the zero-forcing (ZF) precoding and maximal ratio combining (MRC) are respectively assumed for DL transmission and UL detection.

Following the above studies in the context of cellular systems, the synergies between MDD and cell-free massive MIMO (CF-mMIMO) networks are focused. Firstly, a distributed MDD-CF scheme is introduced to enable the FD-style operation but with reduced inter-AP (access point) interference (IAI) and inter-MS interference (IMI). Then, two optimization cases of MDD-CF systems are analyzed by considering simultaneously AP-selection, power- and subcarrier-allocation, under the constraints of individual MSs' quality of services (QoSs). Specifically, in the first optimization case considering one coherence time (CT) interval, a quadratic transform with successive convex approximation (QT-SCA) algorithm is proposed to achieve the SE maximization. By contrast, in the second optimization case on the basis of one radio frame, a two-phase CT (TPCT) interval is designed for MDD-CF systems to guarantee the robust performance over time-varying channels. Correspondingly, a two-step iterative optimization algorithm aided by bisection method is proposed for SE maximization.

Our studies show that the proposed QT-SCA algorithm is capable of converging and achieving reliable performance within a few of iterations. However, its complexity increases exponentially with the sizes of CF networks, determined by the numbers of APs, UL/DL users, subcarriers, etc. With this regard and to attain the dynamic power allocation at reduced overhead, a heterogeneous graph neural network (HGNN) is specifically designed for CF networks, which is named as CF-HGNN. The CF-HGNN consists of the adaptive node embedding layer, message passing layer, attention layer and the downstream power allocation layer. Our studies show that CF-HGNN is scalable to the MDD-CF networks with various numbers of nodes and subcarriers. Furthermore, when assisted by the proposed user clustering, the CF-HGNN trained based on a small CF network can be applied to the large-scale MDD-CF networks, which may cover large area, have a big number of subcarriers and/or simultaneously support a big number of nodes.

Contents

Declaration of Authorship

I, Bohan Li, declare that the thesis entitled **Optimization of Multicarrier-Division Duplex Wireless Systems** and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. 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;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. 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;
5. I have acknowledged all main sources of help;
6. 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;
7. Parts of this work have been published as shown in the list of publications.

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

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Professor Lie-Liang Yang, for all the support and encouragement he gave me. Without his consistent guidance and extraordinary insights this PhD would not have been achievable. I will never forget his hundreds of pages of handwritten manuscripts, which constantly polish my written skills and strengthen my research capability. He is not only my research tutor but also an incredible role model in my life. My sincere gratitude also extends to my supervisor, Professor Robert G. Maunder, for his outstanding supervision and invaluable support. I still remember the elaborate notes he took for me during my first viva, which warm my heart and provide me with precious advice.

I am also thankful to Professor Pei Xiao from University of Surrey, Professor Songlin Sun from Beijing University of Posts and Telecommunications, for their immense support and contribution to my PhD study. Many thanks to my friends and colleagues in the University of Southampton, for their long-term company and invaluable advice. The financial support from China Scholarship Council (CSC) is also gratefully acknowledged.

Lastly and most importantly, I am deeply indebted to the greatest love from my parents. I could not have undertaken this wonderful journey without their support. Finally, to my beloved wife, Min Ye, it is her selfless love and devotion that motivate me to face any challenges and difficulties in the last four years.

List of Publications

1. B. Li, L.-L. Yang, R. G. Maunder, and S. Sun, "Self-interference cancellation and channel estimation in multicarrier-division duplex systems with hybrid beamforming," IEEE Access, vol. 8, pp. 160 653–160 669, 2020.

2. B. Li, L.-L. Yang, R. G. Maunder, and S. Sun, "Resource allocation in millimeter-wave multicarrier-division duplex systems with hybrid beamforming," IEEE Transactions on Vehicular Technology, vol. 70, no.9, pp. 7921-7935, 2021.

3. B. Li, L.-L. Yang, R. G. Maunder, P. Xiao, and S. Sun, "Multicarrier-division duplex: A duplexing technique for the shift to 6G wireless communications," IEEE Vehicular Technology Magazine, vol. 16, no. 4, pp. 57–67, 2021.

4. B. Li, L.-L. Yang, R. G. Maunder, P. Xiao, and S. Sun, "Multicarrier-Division Duplex for Solving the Channel Aging Problem in Massive MIMO Systems," IEEE Transactions on Vehicular Technology, Accepted, 2022.

5. B. Li, L.-L. Yang, R. G. Maunder, P. Xiao, and S. Sun, "Spectral-Efficiency of Cell-Free Massive MIMO with Multicarrier-Division Duplex," IEEE Transactions on Vehicular Technology, Major revision, 2022.

6. B. Li, L.-L. Yang, R. G. Maunder, P. Xiao, and S. Sun, "Heterogeneous Graph Neural Network for Power Allocation in Multicarrier-Division Duplex Cell-Free Massive MIMO Systems," IEEE Transactions on Wireless Communications, Submitted, 2022

Glossary

Abbreviation

1G	First Generation
2G	Second Generation
3G	Third Generation
4G	Fourth Generation
5G	Fifth Generation
ADC	Analog-to-Digital Converter
AoA	Angle of Arrival
AoD	Angle of Departure
BS	Base Station
C-RAN	Cloud Radio Access Network
CCD	Cyclic Coordinate Descent
CCFP	Concave-Convex Fractional Programming
CDMA	Code Division Multiple Access
CE	Channel Estimation
CF-HGNN	CF-Heterogeneous GNN
CF	Cell-Free
CIR	Channel Impulse Response
CLI	Cross-Link Interference
CoMP	Coordinated Multipoint
CPU	Central Processor Unit
CP	Cyclic Prefix

CSI	Channel State Information
CT	Coherence Time
D2D	Device-to-Device
DAC	Digital-to-Analog Converter
DA	Direct Approach
DCNN	Deep Convolutional Neural Network
DD-WP	Decision-Directed Wiener Predictor
DL	Downlink
DNN	Deep Neural Network
DoF	Degrees-of-Free
DS-CDMA	Direct Sequence Code Division Multiple Access
DTDD	Dynamic Time-Division Duplex
EE	Energy Efficiency
ENOB	Effective Number of Bits
FDD	Frequency-Domain Duplex
FDPS	Frequency-Domain Pilot Sequences
FD	Full Duplex
FFT	Fast Fourier Transform
GBD	Generalized Benders Decomposition
GNN	Graph Neural Network
Gpbs	Gigabits Per Second
HD	Half Duplex
IAI	Inter-AP Interference
IBFD	In-Band Full Duplex
ICI	Intercarrier Interference
IFG	Improved Fair Greedy
IMI	Inter-MS Interference
IoT	Internet of Things
KKT	Karush-Kuhn-Tucker

LDPC	Low-Density Parity-Check
LMMSE	Linear Minimum Mean Square Error
LNA	Low-Noise Amplifier
LoS	Line-of-Sight
LSAS	Large-Scale Antenna Systems
LS	Least Squares
LTE	Long Term Evolution
MIMO	Multi-Input Multi-Output
MLP	Multi-Layer Perceptron
ML	Machine Learning
mMIMO	Massive Multi-Input Multi-Output
MMSE	Minimum Mean Square Error
mMTC	Massive Machine Type Communications
mmWave	Millimeter Wave
Mpbs	Megabits Per Second
MRC	Maximum Ratio Combining
MSE	Mean Square Error
MU-MIMO	Multi-User Multi-Input Multi-Output
MUD	Multiuser Detection
MUG	Modified Unfair Greedy
MUI	Multiuser Interference
NAFD	Network-Assisted Full Duplex
NLoS	Non-Line-of-Sight
NOMA	Non-Orthogonal Multiple Access
OBFD	Out-of-Band Full Duplex
OFDM	Orthogonal Frequency Division Multiplexing
OTFS	Orthogonal Time Frequency Space
PAPR	Peak-to-Average Power Ratio
PA	Power Allocation

PGD	Projected Gradient Descent
PS	Pilot Symbol
QoS	Quality of Service
QT	Quadratic Transform
RA	Resource Allocation
RF	Radio Frequency
RIS	Reconfigurable Intelligent Surface
SA	Subcarrier Allocation
SC-FDMA	Single-Carrier Frequency-Division Multiple Access
SCA	Successive Convex Approximation
SE	Spectral Efficiency
SIC	Machine Learning
SINR	Signal-to-Interference-Plus-Noise Ratio
SISO	Single-Input Single-Output
SQNR	Signal to Quantization Noise Ratio
TD-SCDMA	Time Division-Synchronous Code Division Multiple Access
TDD	Time-Domain Duplex
TFPMC	Time-Frequency-Packed Multicarrier
TPCT	Two-Phase CT
UAV	Unmanned Aerial Vehicle
UG	Unfair Greedy
ULA	Uniformly Spaced Linear Antenna Array
UL	Uplink
URLLC	Ultra-Reliable and Low Latency Communications
WCDMA	Wideband Code Division Multiple Access
WCGCN	Wireless Channel Graph Convolution Network
WMMSE	Weighted MMSE
WP	Wiener Predictor

ZF	Zero-Forcing

General Notation

a	Standard lowercase letters denote scalars
$\boldsymbol{a}$	Boldface lowercase letters denote column vectors
$(\boldsymbol{a})_i$	The i element of $\boldsymbol{a}$
Diag$\{\boldsymbol{a}\}$	A diagonal matrix with $\boldsymbol{a}$ being its diagonal elements
$\boldsymbol{A}$	Boldface uppercase letters denote matrices
$(\boldsymbol{A})_{i,j}$	The (i, j)-th element of $\boldsymbol{A}$
$(\boldsymbol{A})^{(i,:)}$	The i-th row of $\boldsymbol{A}$
$(\boldsymbol{A})^{(:,j)}$	The j-th column of $\boldsymbol{A}$
Diag$\{\boldsymbol{A}\}$	A vector comprised of the diagonal elements of the matrix $\boldsymbol{A}$
Tr$(\boldsymbol{A})$	Trace of matrix $\boldsymbol{A}$
$\boldsymbol{A}^*$	Complex conjugate of $\boldsymbol{A}$
$\boldsymbol{A}^T$	Transpose of $\boldsymbol{A}$
$\boldsymbol{A}^{-1}$	Inverse of $\boldsymbol{A}$
$\boldsymbol{A}^H$	Hermitian transpose of $\boldsymbol{A}$
$\boldsymbol{A}^\dagger$	Moore-Penrose inverse of $\boldsymbol{A}$
$\lvert\boldsymbol{A}\rvert$	The determinant of $\boldsymbol{A}$
arg$(\boldsymbol{A})$	Phases of all elements of $\boldsymbol{A}$
$\boldsymbol{I}_N$	A $(N \times N)$ identity matirx
$\mathbb{E}[\cdot]$	Expectation operator
$\log(\cdot)$	Logarithmic operator
$\ln(\cdot)$	Natural logarithmic operator
$\max(\cdot)$	Maximization operator
$\min(\cdot)$	Minimization operator
$\lVert\cdot\rVert_2$	The Euclidean norm of a vector
$\lVert\cdot\rVert_F$	The Frobenius norm of a matrix
argmax$(\cdot)$	Finds the arguments that maximize the target function

$\operatorname{argmin}(\cdot)$	Finds the arguments that minimize the target function
$\exp(\cdot)$	Exponential operator
$\otimes$	The Kronecker product
$\mathbb{R}$	The field of real numbers
$\mathbb{C}$	The field of complex numbers
$\mathcal{A}$	Set
$\lvert\mathcal{A}\rvert$	The cardinality of set
$\mathcal{CN}(\mathbf{0}, \boldsymbol{A})$	The zero-mean complex Gaussian with covariance matrix $\boldsymbol{A}$
σ^2	The power of AWGN per dimension
N_0	Power spectrum density of AWGN

Chapter 1

Introduction

In the last few years, the revolution of mobile devices and applications have been stimulating the unprecedented demands on data rates. With this regard, correspondingly, wireless networks have evolved from the fourth generation (4G) to the current fifth generation (5G) with the aid of several key techniques, such as massive multi-input muliti-output (mMIMO), millimeter-wave (mmWave), etc. Specifically, mMIMO can provide large beamforming gain and spatial multiplexing gain, thereby enabling more users to simultaneously transmit or receive signals at large signal-to-interference-plus-noise ratio (SINR). On the other hand, mmWave exploits high frequency band to support communications, where available bandwidth is abundant. However, although the peak rate of 5G can theoretically reach to 10 Gigabits per second (Gbps), the average rate in practice is usually well below par, due to the fact that the prevalently commercial base stations (BSs) can rarely be equipped with tens of antennas and the mmWave can only cover short range communiocation. The 5G performance report provided by RootMetrics [1] showed that the 5G median download rate in central London varies from operator to operator in terms of BS deployment and bandwidth usage. In particular, Vodafone has the highest data rate, clocking at 180 Megabits per second (Mbps), while Everything Everywhere (EE) only provides rate of 110 Mbps. We should mention that these measurements were based on a fact that only eight percent of smartphone users had 5G service in the UK until March 2021 [2]. Hence, it can be predicted that with the significant increase of 5G users in the next few years, the existing techniques will hardly achieve the goals of Gbps coverage and less than 1 ms end-to-end latency, especially, in the high-density spots like stadiums and concert halls. To this end, several techniques originally expected to be exploited in the early-stage 5G, such as reconfigurable intelligent surfaces (RIS), machine learning (ML), in-band full duplex (IBFD), etc., are still the potential candidates for the evolution of the 5G and beyond systems.

Among the aforementioned potential techniques, IBFD, as a promising duplex mode leveraging whole-time/frequency resources, explicitly has a high priority for application in any scenarios for the sake of high special efficiency and low latency. However,

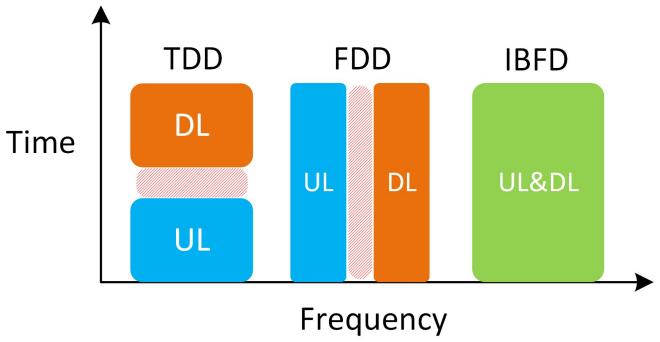


FIGURE 1.1: The comparison diagram of TDD, FDD and IBFD

IBFD has been presented and researched since 1990's, but has not been widely used in the current wireless networks due to the paucity of efficient and low-complexity self-interference cancellation (SIC) techniques. According to [3, 4], IBFD systems are extremely susceptible to the SIC capability. Insufficient SIC can cause significant degradation of system performance, rendering IBFD to lose its advantages over the conventional half duplex (HD) counterparts modes. Therefore, although the HD modes of time-division duplex (TDD) and frequency-domain duplex (FDD) are unable to exploit the full time-frequency resources, they are still the two mainstream duplexing modes introduced in 5G networks. Naturally, some questions arise as to whether the hardware/software technologies fail to provide the desired SIC for the practical implementation of IBFD in near future, *is there a transitional alternative from HD to IBFD, which can not only inherit the main merits of IBFD but also relieve the stringent SIC requirements of IBFD?* Before answering these questions, let us briefly introduce the research development of TDD, FDD and IBFD and their unique characteristics.

1.1 Overview of Mainstream Duplex Modes

As shown in Fig. 1.1, TDD and FDD are two fundamental HD modes in wireless communication systems, relying on separate time slots and frequency bands, respectively, to achieve bidirectional communications. To avoid cross-link interference (CLI), guard time period is required by TDD, while guard band is required by FDD. By contrast, IBFD mode supports simultaneous downlink (DL) and uplink (UL) transmissions over the same band under no constraint of guard limitation, but it suffers from the serious CLI problem caused by the own transmitter or neighbor transmitters.

1.1.1 Time-Division Duplex and Frequency-Division Duplex

Looking back in history, FDD dominated in the first generation (1G) and second generation (2G) wireless networks, while TDD was initially applied in 2G in 1990s as the

digital local loop [5]. Followed by 2G, the third generation (3G) wireless networks were standardized in the early 2000s, and the FDD-relied code division multiple access (CDMA) and wideband CDMA (WCDMA) were defined in two major standards and the TDD-relied time division-synchronous code division multiple access (TD-SCDMA) was only suggested in China and received little interest elsewhere [6]. However, it was TD-SCDMA that showed the potential advantages of the TDD mode, including large frequency spectrum, flexible resource allocation (RA), supporting adaptive antennas, etc. [7]. After the successful application of TD-SCDMA, in the 4G long-term evolution (LTE) era, TDD has become equally important as FDD in the commercial wireless communication networks, as the rise of orthogonal frequency division multiplexing (OFDM) and MIMO techniques [8]. Then, TD-SCDMA has been gradually substituted by TD-LTE. Table 1.1 summarizes the evolution of the standards from 1G to 4G wireless networks.

Wireless Networks	Standards	
	TDD	FDD
1G	×	AMPS
2G	×	GSM, IS95
3G	TD-SCDMA	WCDMA, CDMA2000
4G	TD-LTE	LTE FDD

TABLE 1.1: History development of TDD and FDD

When wireless communications come to 5G and towards 5G beyond (5G+), mMIMO and ultra-dense networks have been widely considered, and TDD has shown the superiority to FDD [9, 10]. The most important reason pushing this to happen is that compared with FDD, TDD has much less overhead of channel state information (CSI) acquisition by exploiting channel reciprocity existing between UL and DL, or outgoing and incoming channels [7]. More specifically, in TDD systems, since the same frequency band is used, UL and DL channel impulse responses (CIRs) are reciprocal to each other. Hence, once the CSI is estimated from UL pilots, it can be directly exploited for assisting DL transmission. By contrast, in FDD systems, as DL and UL transmissions are on different frequency bands that are far apart, the channels need to be estimated at both directions, and the estimated CSI has to be quantized and feedback to the transmitters for implementing pre-processing [9]. Moreover, the advantage of TDD in channel estimation (CE) can be extended to benefit mMIMO systems, where the number of antennas equipped at BSs is assumed to be much larger than the number of users supported [11]. To elucidate this point, let us make a brief example of the CSI acquisition overhead. We assume that the length of the mutually orthogonal UL pilot sequences in a TDD system is equal to the number of users D, and the channel hardening eliminates the needs for DL pilots. Then, if one coherence time (CT) period can accommodate T_c number of symbols, the pilot overhead in this TDD system is $\frac{D}{T_c}$. By contrast, the training pilots of two directions are necessary in a corresponding FDD system, leading to the $(D + N_t)$

length of pilots and the pilot overhead is hence $\frac{(D+N_t)}{T_c}$, where N_t is the number of antennas employed at BSs. In mMIMO systems, N_t can be very large, which may result in that FDD fails to provide a desirable spectral efficiency with a huge amount of resource for CE [9]. Recently, there are some researches having focused on the management of the CSI acquisition overhead in FDD mMIMO systems, such as [12,13]. However, these proposed FDD-based schemes are only feasible, when the channels are sparse with a strong direct or specular path, which may be possibly satisfied at high frequencies like mmWave.

However, we should mention that although TDD systems can leverage the channel reciprocity to obtain CSI without invoking any feedback, the estimated CSI may be outdated due to the UL training, guard period and long-lasting DL transmission, especially when communication happens over fast time-varying channels. For example, the rapid development of unmanned aerial drones and autonomous vehicles requires future wireless networks to support low-latency connection and provide sufficient spectral efficiency in high-mobility scenarios. For this sake, to mitigate the channel aging problem in TDD systems, several techniques of channel prediction have been presented in the literature, like Wiener filter, Kalman filter and deep learning methods [14–16]. However, subject to the structure of TDD, where DL and UL transmissions have to be sequentially implemented, these methods are unable to provide stable performance. Hence, more advanced duplexing mode that are feasible for supporting communications in high-mobility scenarios is demanded. We will analyze the channel aging issue more deeply in Chapter 5.

It is well-recognized that TDD has another advantage over FDD, which is that TDD can support dynamic asymmetrical communications [7]. This is also paramount in future wireless communication systems. Unlike the FDD systems operated on predetermined DL/UL bands, TDD systems can redistribute the resources between UL and DL by altering their time slots. For example, when the demand of DL transmission increases while that of UL traffic remains low, more time slot resources can be reallocated for DL transmissions until the UL traffic is increased.

1.1.2 In-Band Full Duplex

IBFD has been proposed to significantly increase the spectral efficiency, and resolve the sore points of the traditional HD modes, such as the inefficiency of time-frequency resource usage and the high latency in both physical and network layers [17]. The concept has been first introduced to radar systems since the 1940s at least [18], but it was rarely mentioned in wireless communications until 2000s [19]. Since then, there have been growing research on the potential applications of IBFD in cellular networks. However, to date, IBFD has still not been introduced to the standards for commercial use. The

most obvious reason is self-interference (SI) problem, which refers to the interference that a transmitting IBFD BS or AP causes significant interference to itself.

1.1.2.1 Self-Interference Problem

Generally, SI signal, leaking out from a transmitter to a receiver at the same side, includes mainly three components [20], namely the linear components, non-linear components and the transmitter noise. The linear components include the combination of the various delayed copies of the transmit signal. The non-linear part accounts for the non-linearly distorted components of the transmit signal due to the radio frequency (RF) imperfection. The transmitter noise consists of the channel noise and phase noise generated by local oscillators [21]. In order to mitigate SI, we usually resort to three types of SIC methods operated respectively in the propagation, analog and digital domains, all of which are indispensable to enable an IBFD system to work efficiently. For instance, as shown in Fig. 1.2, let us assume that a femto BS operated in the IBFD mode transmits at 30 dBm with a receiver noise floor of -90 dBm. Let us further assume a 30 dB of passive propagation-domain SIC, which can be achieved in practical wireless systems by the path-loss, antenna separation, cross polarization, etc [21]. Suppose that the 12-bit analog-to-digital converter (ADC), whose effective number of bits (ENOB) is 9, are used by receiver, yielding an effective dynamic range of $6.02(\text{ENOB} - 2) \approx 42$ dB [22]. Then, to allow the ADCs work in an efficient way and avoid large quantization noise, the maximum input power to the ADCs is limited to -48 dBm. In other words, the digital-domain SIC can only suppress SI up to the effective range of the ADCs. As shown in Fig. 1.2, once the passive propagation- and digital-domain SIC are fixed, which are assumed to be 30 dB and 42 dB, respectively, the demand on the active propagation- and analog-domain SIC can be obtained, which is 48 dB in this example.

In order to meet the aforementioned SIC requirements, various methods have been proposed and investigated in recent years. Below are some mainstream propagation-, analog- and digital-domain methods.

- **Propagation-domain SIC**. The Propagation-domain SIC can be divided into passive and active methods [23]. Antenna separation is the simplest passive SIC mechanism and some researchers have explored antenna placement as a practicable cancellation technique. Antenna separation uses the fact that the distance between transmit and receive antennas naturally provides SI reduction due to signal attenuation [24]. However, the authors of [25] found that the system can benefit from antenna separation ($>$40 dB) only when the distance between Tx/Rx is above 150 mm. Obviously, this is impractical for the size of current devices like mobile phones. To address the SIC in small devices, a method of using nulling transmit antenna was proposed in [26]. The nulling direction can be controlled,

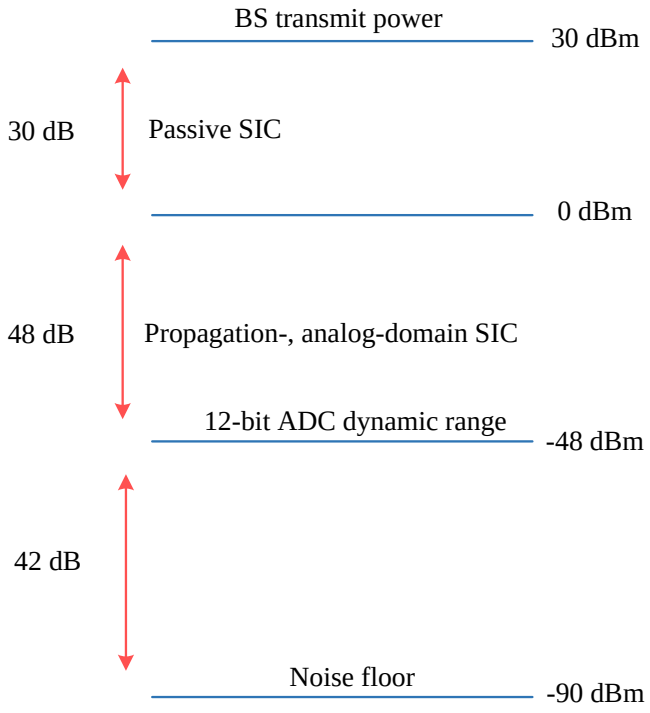


FIGURE 1.2: Overall SIC requirements

and we can position this direction to match the position of the receive antenna. It was shown that this method allows 25 dB-30 dB of interference in the nulling direction. However, it may result in the reduction of the desired signal upto 25 dB lower, which arises from the nulling direction. Another antenna design for SIC was presented in [25]. Different from the other references using at least two antennas [27–29], they applied a dual-port polarized antenna combined with a self-tunable cancellation circuit, which can provide nearly 75 dB of SI reduction over a bandwidth of 10 MHz at 2.45 GHz. This kind of antenna only occupies about $90mm \times 90mm$ space. In order to further reduce the dimension of antennas, they proposed a design by applying a tunable electrical balance isolator in combination with a single-port antenna, which offers 50 dB of SI mitigation.

In the context of active cancellation, beamforming technique is the mainstream SIC method in separate-antenna MIMO systems. Beamforming SIC, at the expense of degree-of-freedom in design precoding/combining vector for desired transmission, can null the radiation pattern at the receive antennas. However, it potentially has two challenges: 1) the exact channel knowledge is required for beamforming and 2) the desired communication may be affected by the process of SI suppression. Considering that the antenna arrays are separated, the authors of [30] presented a digital beamforming scheme, which is called SoftNull method, to suppress SI in mMIMO networks. Based on the measurements obtained from a 72-element antenna array in both indoor and outdoor environment, the IBFD with SoftNull was shown to perform better than the TDD in small cells. Furthermore,

in [31], the SoftNull method was optimized to propose a new method called the JointNull. There are two significant differences between SoftNull and JointNull. The first one is that the JointNull combines beamforming with circuitry rather than totally getting rid of the SIC at RF level. The second one is that the SoftNull decouples the SI suppression precoder design from the DL precoding, while the JointNull jointly optimizes SI suppression along with the DL precoding. In addition to the precoding approach, the authors of [32] also studied the SIC in the mMIMO FD systems, and presented a new method to suppress SI, named as the beam-domain SIC. With this method, a low-dimension beam-domain channel matrix was substituted for the traditional mMIMO channel matrix, based on which a SIC strategy was proposed.

- **Analog-domain SIC**. Analog-domain SIC has also received a lot of research attention. Bozidar et al [33] built a prototype to implement the analog-domain SIC based on the Quellan QHx220 noise cancellers, which can handle a dynamic range of up to 30 dB. Another promising method is using the knowledge about the transmission to cancel SI in the RF signal, before it is digitized [24, 34, 35]. For example, the authors of [24] used a second transmit chain to create an analog cancellation signal from a digital estimate of the SI, canceling 33 dB of SI over a 625 KHz bandwidth signal. However, this approach increases the cost of system, since it requires extra transmit chain. In addition, the authors of [36] proposed a novel mechanism, called as the balun cancellation, which is depended on the signal inversion implemented by a balun circuit. The studies show that this approach can cancel an arbitrarily high transmit power. Melissa et al. in [37] designed a FD analog cancellation architecture consisting of one digital-to-analog converter (DAC), one transmitter radio, one RF attenuator, and one RF adder. The SIC formula can be expressed as

$$y_{\mathrm{SI}}[n,f] = (h_{\mathrm{SI}}[f] - h_z[f]\hat{k}[f])x[n,f] \tag{1.1}$$

where $x[n,f]$ denotes the n-th symbol transmitted from BS during frame f, $h_{\mathrm{SI}}[f]$ denotes the SI channel, which is assumed to remain constant during the transmission of frame f, $h_Z[f]$ denotes the magnitude and phase change generated when a signal passes through an attenuator and a wire to the RF adder, while $\hat{k}[f]$ is the estimate of $(h_{\mathrm{SI}}[f]/h_z[f])$, including both the estimation errors of $h_{\mathrm{SI}}[f]$ and $h_z[f]$. They also studied the SI channel modeling, which is supposed to follow a Ricean distribution with an appropriate K-factor. They found that the K-factor for the SI channel reduces due to active cancellation and the amount of reduction increases as the capability of SIC increases. In [38], the authors proposed a two-stage analog cancellation architecture with the aid of the conventional RF-tapping and digital-tapping approaches. According to their experiments, the proposed method may find different applications, by flexibly reconstructing and canceling the multi-path SI signals and the transmitter noise. Note that the above-mentioned three

methods belong to the single-tap RF SIC. The key drawback to these approaches is that they are only suitable to narrowband systems.

In order to mitigate SI in wideband domain, multi-tap RF cancelers were proposed, for example, in [21, 39, 40]. Specifically, in [21], the authors designed the dynamic algorithms to estimate the distortions introduced by analog circuits and to accurately model the SI being experienced by the received signal. Then, they designed a novel programmable analog cancellation circuit using off-the-shelf components, which implement the algorithm in analog domain to dynamically cancel SI. According to their experiments, this analog circuit and tuning algorithm can provide at least 60 dB of SI cancellation. In [39], a four-tap RF canceler with non-uniform pre-weighted taps was presented. It was shown that in OFDM systems, the scheme can achieve 78 dB and 70 dB of analog-domain SIC over the 20 MHz and 120 MHz bandwidths, respectively. In [40], the authors firstly developed a closed-form expression to calculate the residual power after analog SIC under the assumption of imperfect CSI. They found that the optimal number of taps should be equal to the least number of taps that covers the bandwidth of the received SI. In [41], a novel method, namely the self-mixed analog SIC method, was proposed by taking the advantage of the characteristics of SI signal, which is obtained by multiplying the transmitted signal with the received composite signal.

- **Digital-domain SIC**. The objective of digital SIC is to sweep away possibly all the remaining SI after propagation-domain and analog-domain SIC [42]. In general, the digital-domain SI consists of both linear component and non-linear component. The linear SI component is attributed to the residual SI signal after the analog domain SIC, as well as to the delayed reflections of the transmitted signal. To implement the digital domain SIC, CE is the first and also the most essential step [21, 36, 43, 44]. Since all the channel fading matrices can be obtained through the proper CE and feedback methods [45–47], below we assume that perfect CSI can be obtained before signal transmission. Then, the final received signal for data detection can be expressed as

$$\boldsymbol{y} = \boldsymbol{H}_{\mathrm{UL}}\boldsymbol{x}_{\mathrm{UL}} + (\boldsymbol{H}_{\mathrm{SI}} - \hat{\boldsymbol{H}}_{\mathrm{SI}})\boldsymbol{x}_{\mathrm{DL}} + \boldsymbol{n} \tag{1.2}$$

where $\boldsymbol{H}_{\mathrm{UL}}$, $\boldsymbol{H}_{\mathrm{SI}}$ and $\hat{\boldsymbol{H}}_{\mathrm{SI}}$ represent UL channel, SI channel and the estimated SI channel respectively, $\boldsymbol{x}_{\mathrm{UL}}$ is the UL signal vector, while $\boldsymbol{x}_{\mathrm{DL}}$ is the DL signal vector. In (1.2), $\boldsymbol{H}_{\mathrm{SI}}$ includes not only the wired channel between transmitter and receiver, but also the transceiver impairments, such as ADC quantization noise, transceiver phase noise, etc. [43, 48]. In [49], several SIC algorithms were proposed to mitigate transceiver phase noise based on the separate oscillators. In [50], the closed-form expressions for digital SIC were derived in terms of the common phase error, which shows that the FD-based systems with either lower

transmission power or stronger analog-domain SIC capability is less vulnerable to phase noise. In [43], ADC's quantization noise was modeled in detail, and an algorithm was presented to mitigate the impact of quantization noise in FD systems.

By contrast, the suppression of the non-linear SI component is much more complicated [51–53]. The experimental and analytical results in [54–56] indicated that the nonlinearities play an important role for the achievable performance of the FD-based systems. In practical wireless systems, the main sources of nonlinearity are the power amplifier at the transmitter side, and the low-noise amplifier (LNA) at the receiver side. In principle, for any nonlinear block, the output signal y can be written as a polynomial function of the input signal g in digital domain as [43]

$$y = \sum_{m=1,3,\ldots}^{M} \alpha_m g \left|g\right|^{m-1} \tag{1.3}$$

where g and y denote the digital-domain input signal and the output of the non-linear block, respectively. Note that in (1.3), a limited number of orders contribute to the major distortion and the higher orders could be neglected. In practical systems, the nonlinearity is typically characterized by the third-order intercept point [57], which is defined as the point at which the power of the third harmonic is equal to the power of the first harmonic. In order to cancel the non-linear part of the digital-domain SI, the authors of [58] used least squares (LS) estimator to obtain the nonlinearity coefficients, based on which the nonlinear SI signal is reconstructed and subtracted from the received signal. In [43], the authors proposed the nonlinearity estimation and suppression technique based on the model (1.3), and the simulation results show that the proposed method is able to achieve upto 23 dB SI reduction.

1.1.2.2 Cross-link Interference Problem

In TDD- or FDD-based systems, the interference is only from the same type of signals, i.e., DL only experience interference from DL signals and UL only experience interference from UL signals, provided that the synchronization among APs or the guard band for DL/UL separation is appropriately implemented. This can in general guarantee the interference power to be significantly lower than the signal power and enable wireless communications in low interference scenarios. By contrast, the IBFD-based systems have the severe inter-AP interference (IAI) and inter-MS interference (IMI), which are known as the CLI, as shown in Fig. 1.3. The CLI may become even worse, when many APs and MSs form an ultra dense network or cell-free network.

The methods of suppressing the CLI in IBFD systems can be mainly classified into beamforming and power control. The most challenging problem of the beamforming

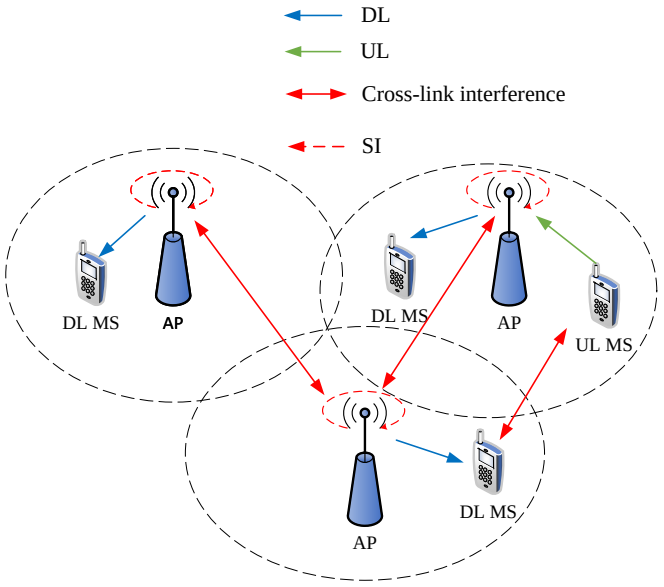


FIGURE 1.3: The IBFD system with different types of links.

method is to attain a good trade-off between interference suppression and the performance of the desired signal transmissions. Specifically, for the IAI cancellation, in [59], the authors demonstrated that when APs are featured with a large number of antennas, the IAI can be suppressed to a very low level. In [60], two IAI suppression methods, namely the centralized and decentralized beamfoming, were proposed, which aim to minimize the AP's transmit power to make IAI below an acceptable level, while meeting the QoS required by DL MSs. In [61], the authors proposed the interference alignment to point the interfering signals from different APs to a signal subspace, so that the desired signals can be received from another subspace without much interference. By contrast, the cancellation of IMI is much harder than that of IAI, as MSs can not accommodate large-scale antenna arrays and often change their positions. To this end, the authors of [62] proposed a table showing the IMI range based on the transmitter/receiver beam index of the AP. Through sharing the table between APs, scheduling and RA can be carried out to efficiently manage the IMI.

In comparison with the beamforming method, the power-control is a more straightforward method, which mitigates CLI via adjusting the transmit power of interfering sources. A side effect of this method is that, reducing the transmission power may also decrease the power of the desired signal. Hence, the implementation of the power-control methods should consider trade-off between the performance degradation of the desired signal and the CLI suppression. In literature, a range of methods have been proposed. For example, in [63], the authors applied an adaptive DL power-control to adjust the AP's transmit power by dividing the distance between the serving AP and the interfering AP into multiple sections. In [64], the authors proposed the joint DL and UL power-control, with the aim to mitigate the impact of IAI. On top of power-control, user scheduling can also be used for CLI mitigation. In [65], the authors proposed

a scheduling strategy to suppress IMI, showing that the new designed reverse frame structure can be used to manage IMI. In [66], the authors introduced a joint scheduling and transceiver design-based CLI mitigation method to avoid the APs transmitting signal to the DL MSs, which experience severe interference from the UL MSs in the neighboring cells.

1.2 Multicarrier-Division Duplex

1.2.1 Characteristics of MDD

According to [3,4], IBFD systems are extremely susceptible to the SIC capability. Insufficient SIC can cause significant degradation of system performance, rendering IBFD to lose its advantages over the conventional HD counterparts. Moreover, the applications of ultra-mMIMO and terahertz (THz) communications in 6G may impose further burden on SIC, resulting in new challenges for the practical deployment of IBFD. Additionally, the authors of [67] demonstrated that even when SI is perfectly canceled, the IAI and IMI problem may still hinder IBFD from outperforming HD. To this end, the multicarrier-division duplex (MDD), an out-of-band FD (OBFD) scheme originally proposed in [68] and recently studied in [4, 69], is capable of enabling the concurrent DL and UL transmissions within the same time slot and the same frequency band, but on different subcarriers. To achieve these, in MDD systems, the subcarriers of one band are divided into two mutually exclusive subsets, namely a DL subcarrier subset and a UL subcarrier subset, to support DL and UL transmissions, respectively, as shown in Fig. 1.4.

Due to this feature, the MDD is suitable for operation with different multicarrier schemes, including orthogonal frequency-division multiple access (OFDMA), filter-bank multicarrier, single-carrier frequency-division multiple access (SC-FDMA), etc. It is flexible for integration of different multicarrier schemes, which makes it highly feasible for the evolution of the 4G LTE/LTE-A systems to the 5G and beyond systems. In addition, MDD can aggregate the resources for UL and DL communications, which are conventionally separately used. Hence, in the MDD systems, joint RA can be implemented among all the users communicating on the UL and DL. By operation in such a way, the MDD can not only improve the efficiency of resource usage, but also significantly enhance the efficiency of communications via multiuser diversity and frequency diversity, and hence, substantially increase the capacity of wireless systems. Furthermore, the MDD is convenient for implementation of the non-orthogonal transmission schemes [70], which have been demonstrated to have the spectral and energy efficiency advantages over the orthogonal transmission schemes, and have been considered for the 5G deployment. The MDD is also flexible for implementation of the sparsely spread CDMA, in order for the 5G systems to aggregate and exploit the discontinues spectrum

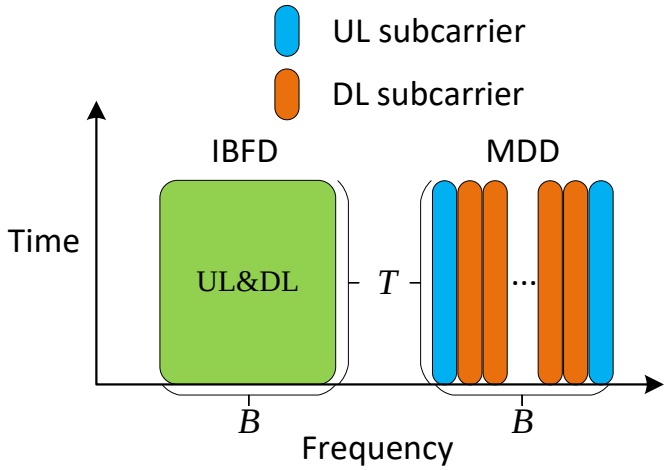


FIGURE 1.4: The comparison of MDD and IBFD modes.

resources, maximize the spectral and energy efficiency, and achieve near-optimum detection performance with the relatively low-complexity detection algorithms, such as, the brief propagation algorithm developed along with the low-density parity-check (LDPC) codes [71].

In a nutshell, the MDD is capable of overcoming some drawbacks of the IBFD and conventional HD modes without compromising their key benefits. The comparison of MDD, HD and IBFD are summarized in the following subsections.

1.2.1.1 MDD *Vs* IBFD

Since IBFD fully exploits the spectral resources, the spectral efficiency achieved by IBFD can be twice as much as that achieved by MDD in ideal scenarios, where SI is perfectly canceled in both analog and digital domains [72]. However, in practice, suppressing all SI is complex and power-consuming, which impede the implementation of IBFD in commercial networks. In order to keep the key advantages of IBFD, such as, the improved bandwidth efficiency and no use of guard interval or band, and lower the overhead of SIC at the same time, MDD constitutes a promising choice to make a good trade-off between IBFD and HD.

Fig. 1.5 illustrates the SIC process in MDD systems. From this figure, it can be observed that the DL/UL transmissions relying on different sets of subcarriers within one band can occur simultaneously within the same OFDM symbol duration. However, as the receiver is very close to the transmitter at BS side, the UL signal at receiver antennas is overwhelmed by the DL signal, the power of which may significantly exceed the available dynamic ranges of the ADCs. Hence, in order to enable ADCs to work efficiently and avoid large quantization noise, MDD systems have to resort to the

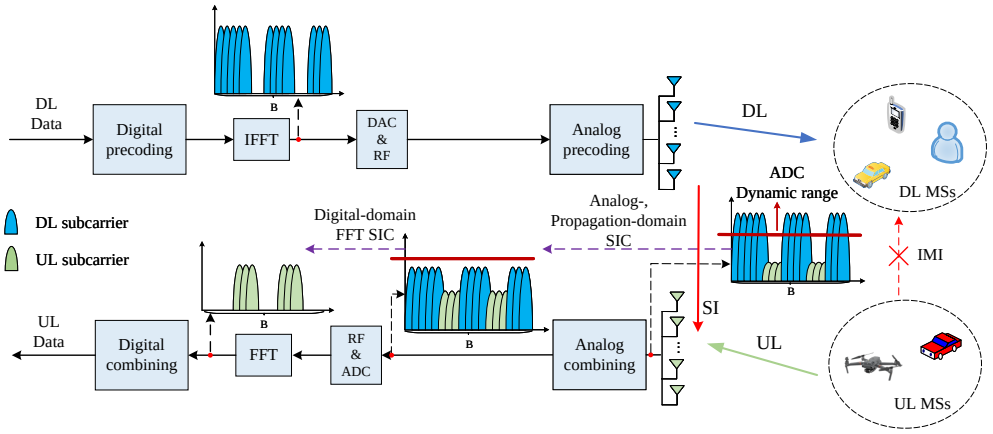


FIGURE 1.5: Illustration of SIC processing in MDD systems.

propagation/analog-domain SIC, which are similar to the IBFD systems, so that the signal prior to entering ADCs is limited within their efficient dynamic ranges. After the ADC processing, the conventional IBFD systems have to further mitigate the residual digital-domain SI by first reconstructing the digital-domain signal between BS's DAC and ADC, and then subtracting it from the received signal [23]. As mentioned previously in Section 1.1.2.1, the performance of the existing digital-domain SIC largely depend on the estimation of SI channel and the reconstruction of SI signal, which are power-consuming and infeasible for implementation in the green communication networks. Furthermore, they are impractical to the small access points (APs) powered by batteries, such as, unmanned aerial vehicle (UAV), Internet of things (IoT) gateway, etc. By contrast, as its counterpart, MDD systems only rely on fast Fourier transform (FFT) operation to easily remove the residual SI in digital domain. In [73], the simulation results show that, subject to the higher power consumption caused by the implementation of more data streams and digital-domain SIC, IBFD was demonstrated to be defeated by MDD in terms of energy efficiency.

Besides, it is noteworthy that the IMI imposed by the UL MSs on the DL MSs in MDD systems can be significantly mitigated, as seen in Fig. 1.5. This is because the IMI and the desired signal received by a DL MS are on the mutually orthogonal subcarriers, which is similar to the SI at BS. Hence, as the result of the large-scale fading of the MS-MS link and of the MS's relatively low transmit power, the IMI in MDD systems belongs to the low-power SI. Consequently, when all MSs and BS are assumed to be synchronized within an allowable time-window, the residual IMI can be ideally canceled with the aid of the FFT operation in digital domain. Analogously, in the dense networks with multiple APs, e.g., in the cell-free networks, the APs operated in MDD mode can also be largely free from IAI in digital domain. By contrast, IBFD systems suffer from severe CLI in both analog and digital domains, as mentioned in Section 1.1.2.2.

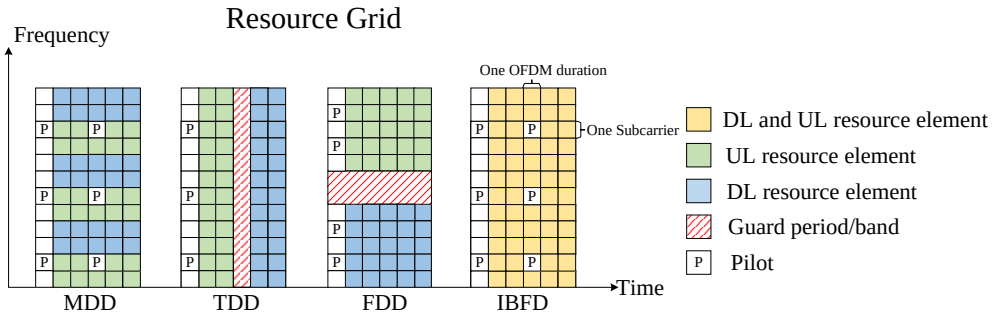


FIGURE 1.6: The resource grid in MDD, TDD and FDD systems.

1.2.1.2 MDD *Vs* HD

Straightforwardly, IBFD employs the capability of integrating the time/frequency resources for supporting simultaneous DL/UL transmissions. However, as analyzed above, before IBFD is made practically implementable with low power consumption and low complexity SIC mechanisms, MDD can be a promising transitional technique from HD to IBFD. Nevertheless, people may argue that *MDD system separates the spectrum into two groups of subcarriers to render FD operation. However, it is unable to double the SE as IBFD system can potentially do. Instead, it experiences SI in analog domain, leading to the extra complexity burden when compared with TDD or FDD-based HD systems. Whether is it worth the effort?* To address this concern, next we will briefly present several potentials of MDD over HD of the FDD/TDD modes.

In comparison of MDD and FDD, both of them can be categorized into the OBFD schemes, but the difference is that the DL and UL transmissions occur in the same band in MDD mode, while different bands are provided to support DL and UL in FDD mode. This feature makes CE distinctive between these two modes. It is well-known that TDD outperforms FDD in mMIMO networks, as the cost of UL training in TDD systems is much smaller than that in FDD systems, due to the reciprocity between the DL and UL channels in TDD systems [9]. In FDD systems, DL training is indispensable and the overhead of training is proportional to the size of the antenna arrays at BS. Although the method like beamsteering codebooks can reduce the complexity of CE in FDD systems, the requirement of CSI feedback from MSs to BS still costs significant system resources. By contrast, MDD can leverage the reciprocity and correlation between DL and UL channels to make CE more efficiently. Furthermore, as shown in Fig. 1.4, a sufficiently wide guard band is required in FDD systems for the sake of avoiding power leakage between the DL and UL, which results in a waste of the valuable spectral resources.

Similar to the FDD, TDD also requires a guard interval in the time domain to separate the UL and DL, leading to the decrease of data symbols in one frame. By comparison, MDD can retain the IBFD's advantages of eliminating the guard interval and hence can

transmit more symbols in the same time slot. TDD has the same feature as IBFD and MDD, where the DL channel information can be obtained by UL training owing to the UL/DL reciprocity. However, with the rapidly growing demand for high-mobility services in, such as, ground vehicles, high-speed trains, UAVs, etc., the reliability of CSI is no longer dominated by pilot contamination but by channel aging. That is to say that in high-mobility scenarios, as the result of fast time-varying, the channel estimated during UL training may become outdated at the time when it is applied, resulting in significant performance degradation. Moreover, in the future wireless systems, both BS and MS terminals are expected to be equipped with the large-scale antenna arrays for attaining the required capacity, which pose further challenges on the CSI acquisition, especially, in the fast time-varying communication environments. As the only feasible duplex mode considered so far for practical mMIMO systems, TDD can perform well in low-mobility scenarios. However, it deteriorates quickly with the increase of mobility even when highly efficient channel prediction approaches are introduced [16, 74]. The main reason behind is that in TDD-based systems, the pilots transmitted via UL at the beginning of a transmission period are used to serve the subsequent detection of UL data and DL transmission. Under this design, if channel varies fast, it becomes increasingly difficult to be predicted with an acceptable accuracy, when DL/UL symbols are far away from pilots. To overcome this problem, more pilots may be inserted during data transmission so as to renew the CSI. However, doing so would introduce more switching intervals and also increases system complexity, leading to the decrease of both SE and EE. Due to the above-mentioned issues, it has been recognized that TDD is not suitable for supporting high-mobility wireless communications, especially, in mMIMO systems [16].

By contrast, in MDD systems, DL/UL data and pilots can be simultaneously transmitted within one OFDM symbol duration. Hence, the CSI for UL detection and DL transmission can be refreshed whenever needed. Owing to this feature, MDD can be a promising alternative for supporting wireless communications over fast time-varying channels. We may argue that in MDD systems, pilots are transmitted at the sacrifice of the available subcarriers for DL transmissions. In this regard, IBFD can be a better option if sufficient SI suppression can be provided by low-cost and low-complexity SIC methods. However, to date, the existing SIC approaches for IBFD are unable to meet the demand. Needless to mention that in the high-mobility scenarios, SIC becomes even more challenging. Therefore, when taking the trade-off among all the related factors into account, MDD may be the best option for the near-future wireless systems, before a practical IBFD system is available for high-mobility communications.

Moreover, in multicarrier OFDM systems, pilots are usually equally distributed in the frequency band, as shown in Fig. 1.6, and the estimated CSI is then used to serve for a coherence block, whose size is about $B_c T_c$, where B_c and T_c denote the coherence bandwidth and coherence time, respectively. Currently, the TDD mode, which leverages the

channel reciprocity between DL and UL, dominates in mMIMO systems. However, as shown in Fig. 1.6, the requirement of switching periods may largely limit the system performance of TDD-relied systems. This can be seen that when either B_c or T_c decreases due to environment changing, it causes the size of coherence block to shrink. As the result, the CE needs to be operated more frequently to update CSI, which increases the switching frequency and hence reduces the sum rate in the long run. In the extreme case when channel varies too fast, making B_cT_c be too small, TDD mode may even become invalid to provide reliable CSI for the operation of mMIMO systems. As for the FDD mode, although DL/UL can be operated concurrently, the channel reciprocity is hard to achieve, since DL band is usually separated far away from UL band. Hence, in FDD systems, using CSI for DL transmission imposes a large burden on the MSs for CE and the backhaul resources for feeding back CSI to BS. Due to these considerations, the feasibility of FDD in mMIMO and ultra-mMIMO based systems is highly dubious [10].

On the other side, the MDD mode can take the advantages of the correlation existing among the DL/UL subcarriers to implement CE like TDD. Moreover, it enables DL, UL and pilots to be located within one OFDM symbol duration but on different subcarriers, as shown by the example in Fig. 1.6. Owing to this configuration, the CSI for data transmission is possible to be updated whenever it is needed, at no expense of increasing the switching period, as the TDD mode. Furthermore, at the network level, a MDD-relied node is allowed to receive signals from user plane and control plane concurrently, which is capable of decreasing the access latency and boosting the operation speed of the overall network.

Additionally, we can know from Fig. 1.6 that the asymmetric communications can be easily achieved in MDD systems by allocating different numbers of subcarriers to DL and UL. In comparison with the TDD systems that control the DL/UL capacity ratio through time-slot allocation, MDD systems may attain a more stable capacity ratio, as its asymmetry is implemented at the OFDM symbol level and not affected by the time-varying channels. By contrast, FDD systems are usually more suitable for supporting symmetrical UL/DL traffics, whereas it cannot be dynamically configured for obtaining different capacity ratios, due to the fixed assignment of UL/DL frequency bands.

The overall comparison of aforementioned duplex modes are presented in Table 1.2, where the explicit advantages are highlighted in green.

1.2.2 Implementation Challenges

MDD employs a range of merits, as analyzed above, which provide us opportunities to meet the requirements of the 5G+ and 6G wireless systems. However, opportunities often lead to challenges in practical implementation. MDD is specifically designed for

	TDD	FDD	IBFD	MDD
Simultaneous DL/UL	None	Exist	Exist	Exist
Guard interval/band	Required	Required	None	None
UL/DL spectrum	Combined	Separated	Combined	Combined
UL/DL channel reciprocity	Exist	None	Exist	Exist
Synchronization	Required	None	Required	Required
Asymmetric traffic	Flexible	Symmetric	Symmetric	Flexible
OFDM PAPR	Classic	Classic	Classic	Reduced
Time-frequency efficiency	Low	Low	Very high	High
High-mobility communications	Classic	Classic	Improved	Improved
SIC	×	×	Hard	Easy
SIC overhead	×	×	High	Low
CLI suppression	Easy	Easy	Hard	Easy

TABLE 1.2: Comparison of duplexing modes.

multicarrier systems. Hence, depending on the specific multicarrier scheme employed, such as, OFDM, SC-FDMA, FBMC, time-frequency-packed multicarrier (TFPMC), etc., it may have the common technical challenges, such as the effect of frequency offset, time offset, etc. Furthermore, as MDD is an OBFD scheme for multicarrier communications, the following direct challenges should be addressed in implementation.

First, similar to the TDD systems, the MDD systems are also expected to have a stringent time synchronization of all the terminals, including all BSs and other wireless terminals. Otherwise, the relative delays existing between the signals sent by different terminals may generate intercarrier interference (ICI) or/and multiuser interference (MUI), which need to be handled by the employment of advanced signal processing approaches. However, future wireless systems are expected to be ultra-dense deployment and in these ultra-dense wireless systems, communication terminals are expected to be close to each other in order to maximize the spectral- and energy efficiency. When MDD is operated in the ultra-dense wireless systems, synchronization may be achieved in a geographically distributed way, to make the wireless terminals close to each other be well synchronized. In this way, the interference from nearby wireless terminals may be efficiently avoided, while the interference from the far away wireless terminals may not generate noticeable effect. The demand on synchronization in the MDD systems may be relaxed, when the techniques robust to synchronization are introduced. For example, if time-domain direct-sequence spreading is introduced to form the multicarrier DS-CDMA [68], the interference generated by the non-ideal synchronization may be significantly reduced after the despreading operation at receiver. Furthermore, when the estimations of relative delays and other CSI are available, the employment of advanced multiuser detection (MUD) can also relax the constraint on synchronization.

Second, although MDD is free from SI in digital domain with appropriate synchronization, it still experiences severe SI in propagation- and analog-domain like IBFD. These

interference may significantly degrade the achievable performance of MDD-based multicarrier systems, if they are not efficiently mitigated. Fortunately, with the research and development of the multicarrier systems and the IBFD schemes, many approaches, as introduced in Section 1.1.2.1, have now been in place for mitigation of the expected interference in the MDD-based multicarrier systems.

1.3 Cell-Free Massive MIMO

1.3.1 Introduction of Cell-Free Massive MIMO

In the current cellular networks where a BS is located at the center of a cell, the cell-center users can usually obtain much higher data rates than the cell-edge users. This performance variation causes the unreliable service quality, which is devastating to the real-time high-demanding services, such as the extended reality and autonomous vehicles. Hence, in order to guarantee the stable and consistent data rates in most of the geographical coverage area, the user-centric cell-free (CF)-mMIMO, blending the advantages of cellular mMIMO, coordinated multipoint (CoMP) transmission and ultra-dense network [75], has attracted significant research interests in wireless communications.

The conception of CF was firstly presented in [76], where the CF large-scale antenna systems (LSAS), as a special example of LSAS (The conception LSAS in [76] consists of three types, i.e., single cell, multi cell and CF LSAS), was studied. The combination of CF and mMIMO was initially researched in [77], and the authors demonstrated that CF-mMIMO systems can significantly outperform the small-cell mMIMO systems in terms of the 95%-likely per-user throughput. Similar to the conventional mMIMO, where BS has more antennas than the number of MSs to be served, the 'massive' in CF systems refers to the wireless scheme, where the number of APs is larger than that of MSs. In general, a CF-mMIMO system includes L geographically distributed APs jointly serving D geographically distributed MSs, where $L \gg D$. Each AP equipped with $N \geq 1$ antennas is connected with a central processor unit (CPU) through fronthaul links, which are used to transmit/receive the DL/UL data to/from APs. Alternatively, all APs can be connected together via fronthaul links so as to share CSI and conduct synchronization.

According to the operation manner of CPU and APs, CF-mMIMO systems can be classified into the centralized and distributed CF-mMIMO systems. Specifically, in centralized CF-mMIMO, CPU is expected to conduct CE, beamforming and signal detection, while APs only need to transmit the coded signal received from the CPU to MSs, or pass the received UL signal from MSs to the CPU. In this case, the system behaves like a co-located mMIMO with a large number of distributed antennas, where the coordinated

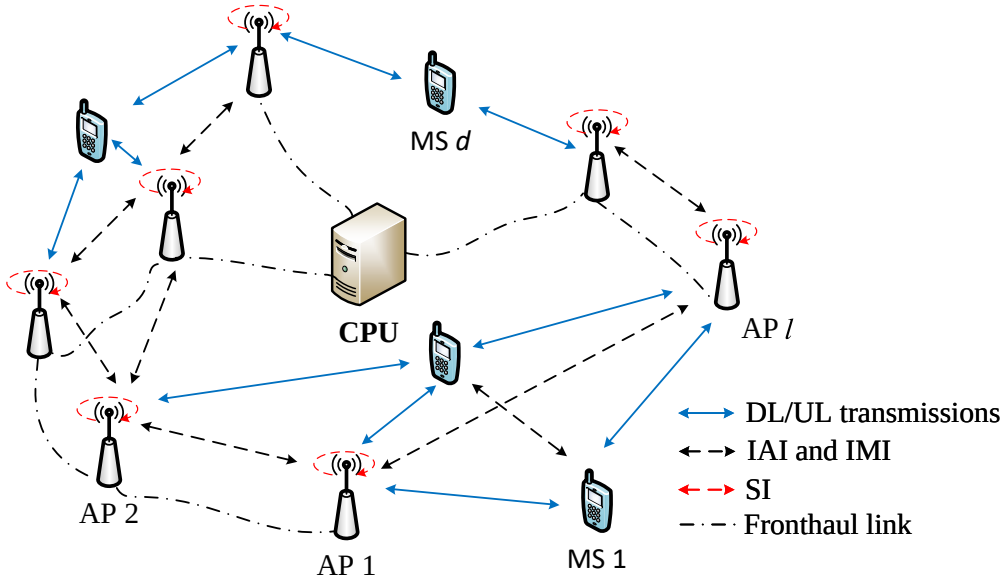


FIGURE 1.7: An illustration of FD-style CF-mMIMO systems.

beamforming can be implemented among APs. To date, most of the existing research in open literature is about the centralized CF-mMIMO systems [78–82]. For instance, in [78], the centralized CF system implementing the coordinated conjugate beamforming can achieve a better SE performance and is also more robust to shadow fading correlation than the corresponding small-cell systems. In [79], the authors demonstrated that the centralized CF-mMIMO with optimal minimum mean square error (MMSE) processing can not only achieve a better SE performance but also largely reduce the fronthaul signaling, compared with the distributed implementation.

However, the centralized implementation inevitably increases the computational complexity at CPU, especially when the APs employing multiple anntennas are densely distributed. In this regard, if each AP is equipped with a powerful baseband processor and multiple antennas, the distributed approach can be a better option for the implementation of CF-mMIMO systems, when considering the scalability and system overhead [75]. In [83], the authors proposed a user-centric distributed CF-mMIMO systems, where each AP uses the locally estimated CSI to send data only to a subset of MSs. In [84], by considering the significant amounts of control signaling and the limited capacity of fronthaul links, the authors presented a distributed CF-mMIMO systems, which is capable of achieving full scalability at the cost of a modest performance loss, when compared with the fully-centralized CF-mMIMO systems. In [85], a reduced-complexity distributed CF-mMIMO system was proposed, where the MMSE-based combining was derived relying only on the locally estimated CSI at APs. It is shown that the proposed scheme can provide unbounded capacity with linear-growing complexity in the high-pilot-contaminated scenarios.

1.3.2 FD-Style Cell-Free Massive MIMO

In order to fully exploit the time-frequency resource and decrease the system latency, it is not too far-fetched to integrate IBFD technique into CF-mMIMO systems, named IBFD-CF mMIMO, as shown in Fig. 1.7, where both APs and MSs are assumed to operate in FD mode. It is expected to reap all the merits of IBFD and CF-mMIMO, thereby enabling a paradigm shift in the future wireless systems. Despite the obvious benefits of the IBFD-CF mMIMO, there are some critical challenges on practical implementation. First, as previously mentioned, SI problem is still the blockage in any of IBFD-relied systems, as DL and UL signals reside in the same resource block. However, thanks to the fact that the AP-MS links are relatively short and LoS-dominant in CF systems, the transmit power of APs is not as large as that of BS in the co-located mMIMO systems, which leads to the reduced requirement for SIC. Second, the densely distributed APs concurrently transmit and receive signals, and therefore cause the serious CLI problem by the IAI and IMI. According to [67], if SI can be appropriately mitigated, CLI becomes the most important limiting factor in multi-cell systems, which is unfortunately also the case in CF systems. Third, the large numbers of APs and MSs result in the extremely complicated optimization problem, which involves AP-MS association, power allocation (PA) and subcarrier assignment.

In [86], the authors studied the centralized IBFD-CF mMIMO systems. With the aid of the coordinated precoding and successive interference cancellation at APs' receivers, the IBFD-CF was shown to outperform the TDD-CF in terms of SE and EE. However, the proposed digital-domain IAI cancellation methods can only provide a small amount of mitigation at APs' receiver. In this case, the AP have to cut down its transmit power to decrease the influence of IAI on neighboring APs, which may lower the system performance. In fact, in comparison with the propagation/analog-domain IAI mitigation, the digital-domain IAI suppression is much harder to achieve in IBFD-CF systems. This is because the former can be achieved with the aid of the power-control and some passive approaches, such as path-loss, cross-polarization, antenna-directionality, etc. [20,23], while the latter for cell-based systems, such as MMSE and successive interference cancellation, is usually not efficient for mitigating the interference generated by a large number of APs and coming from various directions. Noticeably, in comparison with IAI, the IMI problem is even more difficult to handle, as the baseband processors in MSs are not as powerful as those in APs. On the bright side, since the transmit power of MSs is relatively small, the IMI can be largely mitigated in analog domain by power-control and user scheduling [87].

Aiming at relieving SI and IAI, the authors of [88,89] proposed a network-assisted FD (NAFD)-CF mMIMO system, where IAI and SI can be significantly mitigated at CPU by assuming that the CPU knows all the transmitted symbols and the channel knowledge of IAI among APs. Simulation results show that the NAFD-CF can offer a higher SE

than both the TDD-CF and the IBFD-CF. However, we should mention that the suppression of the digital-domain IAI and SI in NAFD-CF is heavily relied on the centralized processing and relatively precise CSI or the IAI channels, which unavoidably impose a heavy burden on the fronthaul and CPU overhead. On top of that, the analog-domain CLI was not appropriately considered in [88,89]. As we elaborated in Fig. 1.2, when two APs or MSs locate close to each other, the CLI in analog domain must be canceled to a certain level. Otherwise, the dominant interference may overwhelm the desired signal at receiver and exceed the effective dynamic range of ADC, leading to the unprecedented quantization error. In this case, the following digital-domain suppression can not obtain the desired signal [72]. Furthermore, a dynamic TDD (DTDD)-CF mMIMO system was proposed in [90], where the majority of interference is also suppressed at CPU relying on the estimation of interference channels. Hence, it confronts with all the aforementioned challenges met in NAFD- and IBFD-CF systems.

Against the background, as seen in Fig. 1.5, since the subcarriers of IAI/IMI are orthogonal to that of the desired UL/DL signal in MDD systems, after the CLI is sufficiently mitigated in the analog domain by the large-scale fading and power-control, the residual digital-domain CLI can be efficiently suppressed with the aid of FFT operation at nearly no extra expense. In this regard, integrating MDD with CF-mMIMO to form the MDD-CF mMIMO may envisage a potential application in the future wireless networks. The comparison of TDD-CF and the other FD-style CF schemes is presented in Table 1.3, where the advantages are highlighted in green. In this thesis, the MDD-CF mMIMO systems will be comprehensively explored in Chapter 6.

1.3.3 Resource Allocation in Cell-Free Massive MIMO

Feasible and practical RA is paramount to achieve the desired performance in CF-mMIMO systems, considering that the densely distributed APs may introduce a large number of variables and cause severe IAI. The optimization of RA can target different QoS metrics, such as the maximization of SE/EE under a power budget constraint, or the minimization of total power under the required sum rate constraint.

In [81], the max-min PA was investigated in TDD-CF mMIMO systems. Simulation results show that under the max-min PA, most of APs in the CF-mMIMO systems are found not to transmit at their full power, and the CF-mMIMO systems can provide five- to ten-fold of improvement in the 95%-likely per-user SE over small-cell systems. In [83], the authors studied the DL/UL SE maximization under the power constraints, when the fairness is or is not considered. To tackle the non-convex objective functions, an approach named as the successive lower-bound maximization, was proposed to efficiently solved the optimization problems via the alternating optimization and sequential convex programming. In [91], the DL SE maximization under the max-min fairness problems was considered, which is solved by two alternating optimization algorithms

	TDD-CF	DTDD-CF	NAFD-CF	IBFD-CF	MDD-CF
Simultaneous DL/UL	None	Exist	Exist	Exist	Exist
Time-frequency efficiency	Low	Medium	Very high	Very high	High
Synchronization	Required	Required	Required	Required	Required
Operation manner	Any	Centralized only	Centralized only	Centralized only	Any
Network latency	high	Medium	low	low	low
High-mobility communications	Classic	Classic	Improved	Improved	Improved
SI suppression	None	None	Required	Required	Required
IAI suppression	None	Easy	Easy	Hard	Easy
IAI CSI	None	Required	Required	Required	None
IMI suppression	Classic	Classic	Classic	Classic	Improved
Fronthaul requirement	Classic	High	High	High	Classic

TABLE 1.3: Comparison of different CF schemes.

based on the weighted MMSE (WMMSE) and fractional programming, respectively. In [92], a joint PA and user grouping problem was formulated under the constraints of users' QoS requirements. In order to solve the optimization problem, a so-called generalized benders decomposition (GBD) method was proposed to minimize the transmit power. Furthermore, the max-min PA optimization problem, subject to power constraints, was studied in [93], where the first order method was applied to reduce the computational complexity.

In the context of FD-style CF-mMIMO systems, in [86], the joint AP-MS association and AP selection were optimized as a PA problem to maximize the SE/EE of the IBFD-CF mMIMO systems, which is solved by an algorithm combining the Dinkelbach method with the successive convex approximation (SCA). In [94], by assuming the limited-capacity of fronthaul in the IBFD-CF mMIMO systems, the weighted sum EE optimization associated with DL/UL PA was carried out by a two-layered approach. In [88], the user selection and SE maximization in NAFD-CF mMIMO systems were formulated, and a low-complexity SCA-based iterative algorithm was proposed to solve the highly complex non-convex problem. Furthermore, in [95], the authors introduced a mixed-integer optimization problem to maximize the SE of the NAFD-CF mMIMO systems, where a two-stage strategy, including iterative SCA and binary relaxations, was proposed to achieve the optimization.

In CF systems, as the APs with one or multiple antennas are densely distributed, the optimization of SE/EE or of user clustering is of highly computational complexity. This becomes even more complicated in the FD-style CF-mMIMO systems, where the DL and UL optimizations are coupled in one problem to optimize. To solve this kind of optimization problems but at a reduced system overhead, the learning-based methods have been often taken into consideration. For example, in [96], the authors conducted PA to accomplish the max-min fairness in TDD-CF mMIMO systems, on the basis of a fully-connected deep neural network (DNN). This DNN-based algorithm was shown to be able to achieve approximately the same performance as the state-of-the-art methods but at the reduced computational complexity. In [97], the authors proposed a DNN architecture consisting of two convolutional layers and four fully-connected layers to solve PA problem to achieve max-min fairness in TDD-CF mMIMO systems. In [98], the authors applied the DNN with the fully-connected layers to implement the SE maximization and max-min fairness PA for the UL of the TDD-CF mMIMO systems. In [99], a deep convolutional neural network (DCNN) relying only on the knowledge about large-scale fading was proposed and the simulation results show that the DCNN can efficiently exploit the large-scale fading for near-optimal PA. Additionally, in [100], a deep reinforcement learning algorithm was proposed to solve the max-min fairness PA, showing that it is capable of achieving the same SE as the optimal solution obtained by the conventional convex solvers.

However, Although the aforementioned learning-based PA methods are shown to prevail over the traditional methods, they are not designed to exploit the structures of the wireless communication networks. Hence, they are not scalable in terms of networks and can not be generalized to the unseen scenarios, such as various networks with different cell sizes or/and AP/MS densities. Furthermore, in these papers, the PA optimization is implemented in a supervised manner, which requires a lot of labeled data that are obtained by the conventional convex optimization algorithms, leading to significantly increased system overhead. Moreover, to the best of our knowledge, there are rare research articles on the PA in the FD-style wireless systems using the learning-based methods, probably due to the extremely hard network training. With this regard, in this thesis, a cutting-edge learning-based technique, namely graph neural network (GNN), will be designed and studied in the context of our MDD-assisted CF-mMIMO systems, as detailed in Chapter 6.

1.4 Summary of Main Contributions

The main contributions and research aspects provided by this thesis are summarized as follows:

- An MDD-based MIMO system operated in FD mode is proposed and investigated, which is endowed with the advantages of both TDD and FDD. It also shares some merits of FD and allows to be free of SI in digital domain, but faces the same challenge of SI as the FD in analog domain. With this regard, an adaptive beamforming-assisted SI cancellation scheme based on cyclic coordinate descent (CCD) algorithm is first proposed. Taking into account the practical requirement of ADC, the proposed approach is shown capable of jointly coping with the desired signals' transmission and SI suppression. Then, the CE in MDD multi-user MIMO system is proposed by exploiting the reciprocity between the UL and DL subcarrier channels that is provided by MDD. Correspondingly, the orthogonality-achieving pilot symbols are designed, and the LS-assisted CE as well as the linear minimum mean-square error (LMMSE)-assisted CE are derived. Furthermore, the performance of MDD-based multi-user MIMO systems employing the proposed SIC method is investigated, with respect to the SI cancellation capability, sum-rate potential, CE performance, and the effect of CE on the achievable performance.
- To demonstrate the advantages of MDD over HD in terms of time-frequency resource usage, the upper bound performance of MDD and HD is compared by assuming the unfair greedy RA. Then, considering a more complicated application scenario of the mmWave with hybrid beamforming, the RA optimization with the

QoS constraints on both DL and UL MSs is proposed. This non-convex RA problem is then divided into a suboptimal subcarrier allocation (SA) problem that is solved by the proposed improved fair greedy algorithm and a convex PA problem. Furthermore, two general hybrid precoders based on matrix factorization and direct approach are designed. Our studies show that the proposed RA algorithm is capable of achieving the performance near the upper bound achieved by the unfair greedy algorithm, while concurrently guaranteeing the proportional fairness among all DL/UL MSs. The performance of the two precoding schemes is depended on the number of radio frequency chains supported. Furthermore, with appropriate antenna deployment, the proposed SIC algorithm is able to provide sufficient SI mitigation, and can be implemented without impacting the RA operation.

- To combat the channel aging problem over fast time-varying channels and fully exploit the time-domain resource, an MDD mMIMO scheme and two types of well-designed frame structures for operation are proposed. In comparison with TDD, the corresponding frame structures related to 3GPP standards and their variant forms are presented. The MDD-specific general Wiener predictor (WP) and decision-directed Wiener predictor (DDWP) are introduced to predict the channel state information, respectively, in the time domain based on UL pilots and in the frequency domain based on the detected UL data, considering the impact of residual SI. Moreover, by applying the zero-forcing (ZF) precoding and maximum ratio combining, the closed-form approximations for the lower bounded rate achieved by TDD and MDD systems over time-varying channels are derived. Simulation results show that MDD, endowed with the capability of full-duplex but less demand on SI cancellation than IBFD, outperforms both the conventional TDD and IBFD in combating channel aging.

- An MDD-based CF scheme, namely MDD-CF, is proposed, which enables DL data and UL data (or pilots) to be concurrently transmitted on mutually orthogonal subcarriers in the distributed CF-mMIMO systems. To demonstrate the advantages of MDD-CF, the SE performance is studied in terms of one coherence interval when AP selection, PA and SA are considered. However, the SE optimization is a mixed-integer non-convex problem that is NP-hard. Hence, to solve this problem, the inherent association between involved variables are executed to transform the problem into a continuous-integer convex-concave problem. Then, a quadratic transform based SCA (QT-SCA) iterative algorithm is proposed to achieve the SE maximization. Next, we extend our study to the scenario with one radio frame consisting of several CT intervals. Correspondingly, a novel two-phase CT interval scheme is designed to not only improve the SE in radio frame but also provide consistent data transmissions over fast time-varying channels. Correspondingly, to facilitate the optimization, we propose a two-step iterative

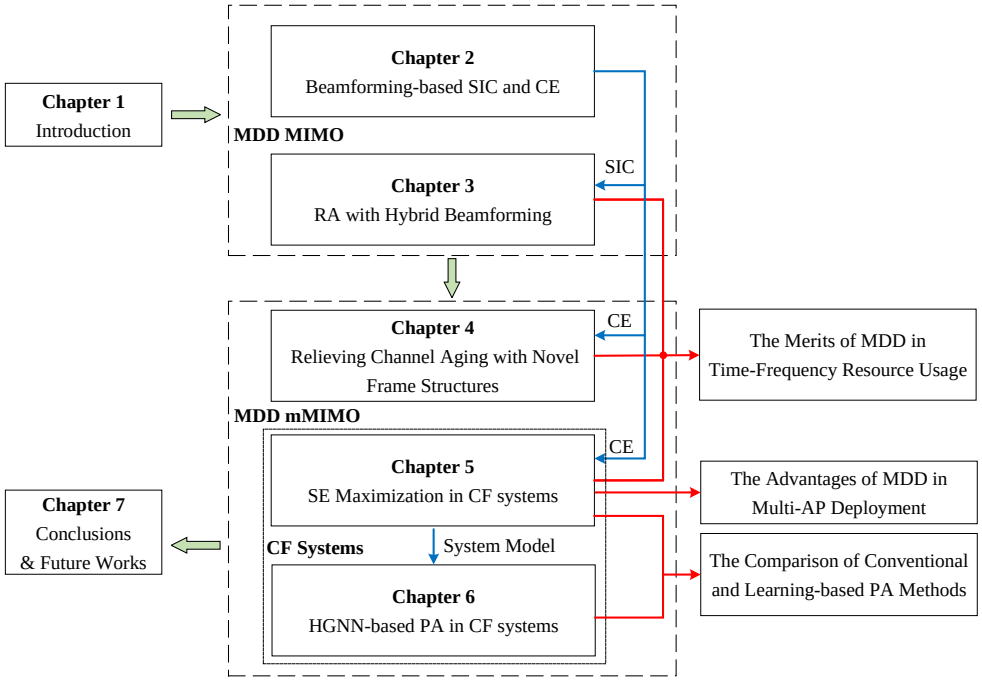


FIGURE 1.8: The outline of the thesis.

algorithm by building the connections between two phases in TPCT through an iteration factor. Simulation results show that, MDD-CF can significantly outperform IBFD-CF due to the efficient interference management. Furthermore, compared with TDD-CF, MDD-CF is more robust to high-mobility scenarios and achieves better SE performance.

- In order to maximize the SE with the less computational complexity in the distributed MDD-CF mMIMO systems, we present heterogeneous graph neural network specific for CF systems, namely CF-HGNN. The proposed CF-HGNN consists of an adaptive node embedding layer, meta-path based message passing, meta-path based attention and downstream PA learning. In particular, the adaptive node embedding layer can handle the varying number of APs, MSs and subcarriers, and the involved attention mechanism enables each AP/MS node in CF-HGNN to aggregate the information from interfering path and communication path with different priorities. The proposed CF-HGNN is capable of using 10^4 times less operation time to achieve the 99% performance of the SE of QT-SCA. Additionally, CF-HGNN also significantly outperforms unfair greedy method in terms of SE performance. Furthermore, CF-HGNN exhibits good adaptivity to varying number of nodes and subcarriers, and also generalization ability to different sizes of CF network.

1.5 Thesis Outline

The outline of this thesis is highlighted in Fig. 1.8.

- **Chapter 2**: In Chapter 2, we introduce a novel duplexing mode, namely MDD, to the MU-MIMO systems with hybrid beamforming. The motivation of suppressing SI in MDD MIMO is presented in Section 2.1. Then, the basic models of DL/UL communications, SI channel and ADC dynamic range in MDD MIMO are introduced in Section 2.2. In Section 2.3, on the basis of CCD algorithm, an adaptive beamforming-based SIC scheme via analog precoder/combiner design is proposed, which aims to provide sufficient SIC in propagation/analog-domain to let ADC work within an effective dynamic range. Furthermore, the CE methods with well-designed frequency-domain pilots is specifically proposed for MDD MIMO systems in Section 2.4. In Section 2.5, the performance and computational complexity of the proposed SIC approach is comprehensively investigated and compared with the conventional beamforming-based method. Section 2.6 concludes this chapter. Overall, Chapter 2 comprised of robust SIC and efficient CE can be deemed as the fundamentals of the following research on MDD-based MIMO systems.

- **Chapter 3**: Further to Chapter 2, the RA problem consisting of subcarrier assignment and power allocation in MDD MIMO systems is studied. The brief review of RA in FD OFDMA systems and the motivation of studying RA in MDD MIMO are provided in Section 3.1. The DL/UL communication models and the RA optimization problem considering the proportional data rate constraints are presented in Section 3.2. To demonstrate the performance advantages of MDD in terms of resource usage, a MU-SISO system using unfair greedy RA algorithm is firstly investigated in Section 3.3. Then, in order to solved the NP-hard and nonconvex problem presented in Section 3.2, the suboptimal SA and PA algorithms are proposed in Section 3.4. For SA, an improved fair greedy algorithm is proposed, which consists of three stages to concurrently achieve the sum-rate maximization, coarse fairness and spatial multiplexing gain. For PA, after the implementation of SA, the PA problem is transformed into a convex form by leveraging the full-digital ZF precoding, which is then solved with the aid of Karush-Kuhn-Tucker (KKT) conditions. Furthermore, in Section 3.5, two hybrid precoding schemes are proposed based on matrix factorization and direct approach, respectively, for conducting RA under the fairness constraint, and also a novel combining design incorporating an adaptive SIC method originally proposed in Chapter 2. In Section 3.6, the performance results in terms of RA and SIC are presented. Section 3.7 concludes this chapter.

- **Chapter 4**: In Chapter 4, we explore the advantages of MDD in dealing with channel aging problem. The brief review of the existing techniques of performing communications over fast time-varying channels and the motivation of introducing MDD to relieve channel aging problem are presented in Section 4.1. The DL/UL communication models and channel aging model are presented in Section 4.2. In Section 4.3, two types of frame structures dedicated for MDD are proposed to mitigate the channel aging problem. Then, WP and DDWP are proposed in Section 4.4 to supporting consistent communications over two types of frame structures, which predict time-domain CSI and frequency-domain CSI, respectively. The estimated CSI, which are taken as the input of the two predictors, can be obtained based on the method proposed in Chapter 2. Furthermore, in Section 4.5, the closed-form expressions for approximating the lower bounded average sum rates of both TDD and MDD systems are derived using the ZF precoding and the maximum ratio combining. The numerical results for the performance comparison of the MDD and TDD systems with two types of frame structures are presented in Section 4.6. Lastly, we conclude this chapter in Section 4.7.

- **Chapter 5**: In Chapter 5, we study the application of MDD in CF-mMIMO. In Section 5.1, we review the most of the HD- and IBFD-based CF-mMIMO systems and present the potentials of MDD-CF in the distributed CF-mMIMO. In Section 5.2, a distributed MDD-CF scheme is proposed with the consideration of AP-selection, PA, SA as well as MSs' QoS constraints. In addition, the impact of SI, IAI and IMI are practically modeled. In Section 5.3, the case of DL/UL transmissions in one CT interval is first considered. To solve the corresponding optimization problem, a QT-SCA is proposed to achieve the SE maximization. Then, in Section 5.4, the case of DL/UL transmissions in one radio frame including multiple CTs is considered, and the CE is implemented based on the method proposed in in Chapter 2. In order to maximize the SE in radio frame, a two-phase CT interval is specifically designed for FD-style CF systems. Due to the fact that the two phases are tightly coupled, we introduce an iterative factor to build the connection between them, and then the original optimization problem in radio frame is transformed into a two-step iterative optimization, which can be solved with the aid of QT-SCA and bisection methods. Furthermore, a comprehensive comparison between IBFD-CF, TDD-CF and MDD-CF is implemented in Section 5.5. Section 5.6 concludes this chapter.

- **Chapter 6**: As the proposed QT-SCA in Chapter 5 for PA in the distributed MDD-CF systems impose heavy burden on system overhead with the increased APs, MSs and subcarriers, we further propose a heterogeneous graph neural network (HGNN)-based approach to efficiently allocate power among all of APs and MSs over multiple subcarriers in Chapter 6. In Section 6.1, we present the motivation of using GNN in implementing PA in CF-mMIMO. In Section 6.2, the basic

system model of distributed MDD-CF within one CT interval is inherited from Section 5.2. We then reformulate the PA problem in MDD-CF mMIMO under the MSs' QoS constraints, while the AP-seletion and SA are omitted. Thereafter, in Section 6.3, the CF-HGNN consisting of four fundamental modules, i.e., adaptive node embedding layer, meta-path based message passing, meta-path based attention and downstream PA learning, is proposed. Furthermore, in Section 6.4, we compare the performance achieved by QT-SCA, unfair greedy method and CF-HGNN. Section 6.5 concludes this chapter.

Chapter 2

Beamforming-Based Self-Interference Cancellation and Channel Estimation in MDD MIMO

This chapter endeavors to solve two fundamental challenges in multicarrier-division duplex (MDD) multi-input multi-output (MIMO) systems, namely self-interference (SI) problem and channel estimation (CE), relying on the designs of hybrid beamforming and frequency-domain pilots, respectively.

2.1 Introduction

As we discussed in Chapter 1, owing to the fact that the SI and desired uplink (UL) signal occur over mutually orthogonal subcarriers, the digital-domain SI can be easily removed by fast Fourier transform (FFT) operation in MDD systems, while the propagation/analog-domain SI still require the appropriate SIC methods to ensure the efficient work of analog-to-digital converter (ADC) at receiver, which is similar to the case in in-band full duplex (IBFD) systems. In this regard, aiming at providing the sufficient self-interference cancellation (SIC) prior to ADC, and hence enabling MDD systems to perform in an efficient way without SI, we will study the beamforming-based SIC technique in MDD MIMO systems, which can not only suppress SI to a very low level, but guarantee the performance of desired UL transmission. The particular motivation for addressing this specific problem arises from the following observations.

From literature [34, 101–105] we know that in the context of the conventional IBFD MIMO systems, various SIC methods in the propagation and analog domains have

been proposed. These SIC methods may also be introduced to the MDD MIMO systems. However, when MDD MIMO system is in the large-scale, explained by the number of transmit/receive antennas, the traditional SIC approaches for MIMO may not be suitable, due to the consequence of huge overhead and complexity. Fortunately, in this case, the big number of antennas can be leveraged for SI suppression. To this end, beamforming-based SIC has become one of the most important methods for SI reduction. Specifically in [30], a full-digital precoder has been designed to point SI signals to the null space of desired received signals and, thereby, cancel the SI signals in analog domain. By properly designing the full-digital precoders, SI can also be suppressed along with the maximization of sum-rate [106, 107]. Instead of full-digital precoding, as shown in [108], hybrid precoding is capable of achieving the SI reduction of upto 30 dB. However, to the best of our knowledge, on the joint design of desired signal transmission and SIC, all the beamforming-based SIC methods presented so far only provide a fixed amount of SI reduction. This may not satisfy the different requirements of propagation/analog-domain SI reduction in practice and, consequently, causes large quantization noise after ADC, if SI reduction is insufficient. Furthermore, in these references, only the point-to-point single-carrier MIMO communication scenarios have been considered. Additionally, in [4], the MDD-assisted point-to-point multicarrier MIMO system employing full-digital precoder/combiner has been proposed and studied, demonstrating that MDD-mode is capable of outperforming half duplex (HD) and IBFD modes in some communication scenarios. However, the existing SIC methods relying on full-digital beamformers are no longer feasible in large-scale MIMO systems, as the full-digital beamformers are expected to be substituted with the hybrid ones for the sake of efficiency. Hence, to satisfy the requirement of ADC in various communication environments, the study on the dynamic SI reduction in the hybrid beamforming assisted IBFD or MDD systems is highly important, but has not been considered in open literature.In addition to the SI problem, we will focus on the CE in MDD systems, which exhibits the merits of channel reciprocity and channel correlation between downlink (DL) and UL channels, as we mentioned in Section 1.2.1.

Against the background, in this chapter, the novelty of our work is compared with the related works in Table 2.1. Note that in addition to the differences stated in the table, the CE in MDD-based systems was not considered in [4]. The contributions of this chapter can be summarized as follows:

Firstly, the SIC requirement for ADC to efficiently operate in its dynamic range is modeled. Then, to make MDD-mode feasible in large-scale multiuser (MU) MIMO systems, we propose an adaptive beamforming based SIC method via analog precoder/combiner design in order to dynamically suppress SI, which has not been investigated in the open literature. Various design options are addressed by taking account of the trade-off between performance and complexity. Furthermore, zero-forcing (ZF) and minimum mean square error (MMSE) algorithms are respectively introduced to design the digital

precoder and combiner. Our studies show that the proposed beamforming SIC method are capable of simultaneously suppressing the SI while maintaining the performance of desired DL/UL communications, when appropriate initialization is applied in the proposed algorithms. Furthermore, our proposed method is capable of providing SIC over a big dynamic range upto 300 dB, which is achieved via applying different system configurations and an appropriate number of SIC iterations.

Secondly, we address the CE in MDD MIMO systems, and consider the estimation of both UL and DL channels by exploiting the correlation existing among subcarriers and the reciprocity existing between the UL and DL channels. To be more specific, we first design a set of frequency-domain orthogonality-achieving pilot symbols (PSs) for estimating the UL/DL channels of all mobile stations (MSs). Based on these orthogonality-achieving PSs, the least square (LS) CE is implemented to obtain the time-domain UL CIR, from which the UL/DL CIRs of all subcarriers are derived with the aid of the above-mentioned subcarrier correlation and reciprocity of UL/DL channels. Furthermore, we consider the CE in the scenario where employing orthogonality-achieving PSs is impossible due to a big number of MSs and/or randomly distributed UL subcarriers. Correspondingly, the CE is accomplished in the principles of linear minimum mean square error (LMMSE).

Finally, we investigate and compare the performance of the MDD MIMO systems. First, we demonstrate the performance of the proposed SIC schemes. Then, the spectral-efficiency performance of MDD MIMO systems with various beamforming aided SIC options is investigated and compared. Furthermore, the performance of the proposed CE schemes is demonstrated, and the impact of CE on the achievable spectral-efficiency is studied and compared. Our studies show that the LS CE relying on the orthogonality-achieving PSs enables the MDD MIMO system to achieve nearly the same sum-rate as the counterpart system with perfect channel state information (CSI). In the case that orthogonality-achieving PSs are impossible, the LMMSE CE is promising, allowing to achieve the sum-rate close to that attained under the assumption of perfect CSI.

The remainder of this chapter is outlined as follows. In Section 2.2, we address the modeling in MDD MIMO systems, including transmitter, channel and receiver models, as well as the modeling of ADC. In Section 2.3, the hybrid beamforming is designed with the objective to maximize sum-rate while simultaneously meet the SIC target. CE is studied in Section 2.4, while simulation results are demonstrated in Section 2.5. Finally, in Section 2.6, we summarize the findings from research and consider their implications.

TABLE 2.1: Comparison of beamforming based SIC methods.

	System model	Operation mode	SIC method	Feature
This Chapter	Large-scale MU-MIMO	MDD	Hybrid beamforming SIC; Digital-domain SIC by FFT operation	SIC requirement modeling of ADC; Provide SI reduction upto 300 dB; Joint design of desired transmission and SIC
[31], [30]	MU-MIMO	IBFD	Full-digital beamforming SIC	Different antenna array deployments
[106]	Point-to-point MIMO	IBFD	Full-digital beamforming SIC	Joint design of desired transmission and SIC
[107]	MU-MIMO	IBFD	Full-digital beamforming SIC	Joint design of desired transmission and SIC
[108]	Point-to-point MIMO	IBFD	Hybrid beamforming SIC	Provide SI reduction upto 30 dB
[4]	Point-to-point MIMO	MDD	Full-digital beamforming SIC; Digital-domain SIC by FFT operation	Consideration of ADC requirement

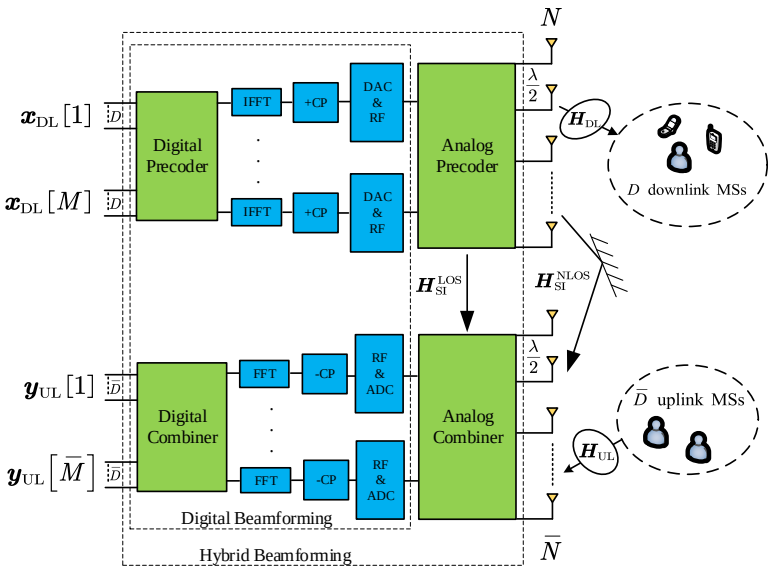


FIGURE 2.1: Schematic diagram of the MDD MIMO system.

2.2 System Model

We consider an MDD MIMO system, where a N-element transmit antenna array at BS uses N_{RF} RF chains to serve D DL MSs and simultaneously, a $\bar{N}$-element receive antenna array also at BS uses $\bar{N}_{\text{RF}}$ radio frequency (RF) chains to serve $\bar{D}$ UL MSs. Hence, the total number of MSs is $D_{\text{sum}} = D + \bar{D}$. All MSs are equipped with single antenna. In our proposed system, the BS works in MDD mode, while MSs are operated in TDD mode. As shown in Fig. 2.1, we assume that the transmitter and receiver at BS are reasonably separated in space, and both of them are equipped with the uniformly spaced linear antenna array (ULA). We assume OFDM-assisted data transmission, that the channels between BS and MSs experience frequency selective fading, and that sufficient cyclic prefix (CP) is invoked to avoid inter-symbol interference (ISI). Furthermore, following the principles of MDD [4], subcarriers are divided into two mutually exclusive subsets, namely a DL subcarrier subset $\mathcal{M}$ with M subcarriers, and an UL subcarrier subset $\bar{\mathcal{M}}$ with $\bar{M}$ subcarriers, i.e., $|\mathcal{M}| = M$ and $|\bar{\mathcal{M}}| = \bar{M}$. The total number of subcarriers is expressed as $M_{\text{sum}} = M + \bar{M}$. We assume that UL/DL MSs are scheduled in such a way, so that the interference generated by UL MSs on a DL MS is sufficiently low, without distorting the operation of the DL MSs' receiver ADCs.

2.2.1 Communications Channel Model

The channel between the n-th transmit antenna at BS and the d-th DL MS is modeled by a L-tap frequency-selective fading channel, with the time-domain CIR (TDCIR) expressed as [109]

$$\boldsymbol{g}_{n,d}^{\text{DL}} = [g_{n,d}[1], ..., g_{n,d}[l], ..., g_{n,d}[L]]^T, \tag{2.1}$$

where $g_{n,d}^{\mathrm{DL}}[l] = \alpha_{n,d,l}^{\mathrm{DL}}$ follows a complex Gaussian distribution of $\alpha_{n,d,l}^{\mathrm{DL}} \sim \mathcal{CN}(0, 1/L)$. Similarly, the UL channel between the $\bar{d}$-th UL MS and the $\bar{n}$-th receive antenna at BS is defined as the L-tap TDCIR expressed as $\boldsymbol{g}_{\bar{n},\bar{d}}^{\mathrm{UL}}$. Furthermore, when the same frequency band is considered, which is the case in MDD mode, we have $\boldsymbol{g}_{n,d}^{\mathrm{DL}} = \boldsymbol{g}_{n,d}^{\mathrm{UL}}$, meaning that there is no distinction between UL and DL channels, i.e., they are reciprocal. Hence, when there is no confusion, the notations 'DL' and 'UL' are removed.

According to the principles of OFDM [68], the frequency-domain CIR (FDCIR) $\boldsymbol{h}_{n,d}$ can be obtained as

$$\boldsymbol{h}_{n,d} = \boldsymbol{F}\boldsymbol{\Psi}\boldsymbol{g}_{n,d}, \tag{2.2}$$

where $\boldsymbol{F} \in \mathbb{C}^{M_{\mathrm{sum}} \times M_{\mathrm{sum}}}$ is the FFT matrix, $\boldsymbol{\Psi} \in \mathbb{C}^{M_{\mathrm{sum}} \times L}$ is constructed by the first L columns of $\boldsymbol{I}_{M_{\mathrm{sum}}}$. Furthermore, the DL subchannels $\boldsymbol{h}_{n,d}^{\mathrm{DL}}$ and UL subchannels $\boldsymbol{h}_{n,d}^{\mathrm{UL}}$ between the n-th antenna at BS and the d-th MS can be obtained from (2.2), which can be expressed as

$$\boldsymbol{h}_{n,d}^{\mathrm{DL}} = \boldsymbol{\Phi}_{\mathrm{DL}}^{T}\boldsymbol{h}_{n,d} = [h_{n,d}[1], ..., h_{n,d}[m], ..., h_{n,d}[M]]^{T}, \tag{2.3}$$

$$\boldsymbol{h}_{n,d}^{\mathrm{UL}} = \boldsymbol{\Phi}_{\mathrm{UL}}^{T}\boldsymbol{h}_{n,d} = [h_{n,d}[1], ..., h_{n,d}[\bar{m}], ..., h_{n,d}[\bar{M}]]^{T}, \tag{2.4}$$

where $\boldsymbol{\Phi}_{\mathrm{DL}}^{T} \in \mathbb{C}^{M \times M_{\mathrm{sum}}}$ and $\boldsymbol{\Phi}_{\mathrm{UL}}^{T} \in \mathbb{C}^{\bar{M} \times M_{\mathrm{sum}}}$ are the mapping matrices, constructed from $\boldsymbol{I}_{M_{\mathrm{sum}}}$ by choosing the columns corresponding to the particular subcarriers assigned to DL and UL, respectively.

2.2.2 Self-Interference Channel Model

Since both UL and DL are operated in the same frequency band based on MDD, the MDD MIMO system experiences SI, as shown in Fig.2.1. We assume that the SI channel experiences Rician fading, constituting both LoS component and NLOS component, which is expressed as [108]

$$\boldsymbol{H}_{\mathrm{SI}} = \sqrt{\frac{\kappa}{\kappa+1}}\boldsymbol{H}_{\mathrm{SI}}^{\mathrm{LoS}} + \sqrt{\frac{1}{\kappa+1}}\boldsymbol{H}_{\mathrm{SI}}^{\mathrm{NLOS}}, \tag{2.5}$$

where κ is the Rician factor. As the transmit and receiver antennas at BS are close to each other, $\boldsymbol{H}_{\mathrm{SI}}^{\mathrm{LoS}}$ denotes the LoS near-field flat fading channel, with the (i,j)-th element expressed as [110]

$$\left(\boldsymbol{H}_{\mathrm{SI}}^{\mathrm{LoS}}\right)_{i,j} = \frac{\rho}{r_{ij}}\exp\left(-j2\pi\frac{r_{ij}}{\lambda}\right), \tag{2.6}$$

where ρ is the power normalization constant making $\mathbb{E}[||\boldsymbol{H}_{\mathrm{SI}}^{\mathrm{LoS}}||_F^2] = N\bar{N}$, r_{ij} is the distance between the i-th element of transmit array and the j-th element of receive array, and λ is the wavelength. By contrast, $\boldsymbol{H}_{\mathrm{SI}}^{\mathrm{NLOS}}$ in (2.5) denotes the NLOS channel, which is assumed to follow the flat Rayleigh fading, with its elements obeying the distribution of $\left(\boldsymbol{H}_{\mathrm{SI}}^{\mathrm{NLOS}}\right)_{i,j} \sim \mathcal{CN}(0,1)$.

2.2.3 Requirement of Self-Interference Cancellation

In full-duplex systems, ADC is the most critical component determining the system operability and achievable performance. A practical ADC has only limited dynamic range and resolution. Hence, if the input signal to an ADC is beyond a particular level, the signal will be distorted, yielding large quantization noise and non-linear distortion, which would further decrease the performance of the following digital-domain signal processing [21]. Specifically, when assuming a Q-bit ADC, the signal to quantization noise ratio (SQNR) is about $6.02Q$ [111]. When given the bandwidth of B (Hz), the noise floor at receiver is given by [23] $-174 + k_N + 10\log(B)$ dBm, where k_N is noise factor. Hence, the maximum input signal power to the receiver is

$$s_{\max}(dBm) = -174 + k_N + 10\log(B) + 6.02Q, \tag{2.7}$$

Therefore, for an SI contaminated signal to pass an ADC without distortion, the propagation- and analog-domain SI cancellation should provide the SIC of at least [21]

$$C_{\mathrm{SI}}^{\mathrm{Demand}}(dB) = P_{\mathrm{DL}}(dBm) - s_{\max}(dBm) + 10(dB), \tag{2.8}$$

where P_{DL} is the transmit power of DL, while 10 dB is added to account for the peak-to-average power ratio (PAPR), as an OFDM signal's power may rise upto 10 dB above the average power [112].

After the propagation- and analog-domain SIC, the SI input to ADC of BS receiver is $P_{\mathrm{SI}} = P_{\mathrm{DL}} - C_{\mathrm{SI}}$, where C_{SI} is the total SI reduced in the propagation- and analog-domains. Hence with the aid of (2.8), we know that when given P_{DL}, the SIC should satisfy $P_{\mathrm{SI}} \leq P_{\mathrm{DL}} - C_{\mathrm{SI}}^{\mathrm{Demand}}$, yielding the SIC requirement of

$$\frac{P_{\mathrm{DL}}}{P_{\mathrm{SI}}} \geq 10^{\frac{C_{\mathrm{SI}}^{\mathrm{Demand}}}{10}}, \tag{2.9}$$

Above we have provided the channel models in the MDD MIMO systems, and analyzed the target for SI cancellation. Below we start considering the transceiver design for the MDD MIMO systems.

2.2.4 Transmitter Model

For DL transmission, let the symbol vector transmitted by BS on the DL subcarriers be expressed as $\boldsymbol{x}_{\mathrm{DL}}[m] \in \mathbb{C}^{D\times 1}$, which is normalized to satisfy $\mathbb{E}\left[\boldsymbol{x}_{\mathrm{DL}}[m]\boldsymbol{x}_{\mathrm{DL}}[m]^H\right] = \boldsymbol{I}_D$. In order to mitigate SI and attain beamforming gain, at transmitter, $\boldsymbol{x}[m]$ is processed by a precoder $\boldsymbol{F}[m] \in \mathbb{C}^{N\times D}$, with the constraint of $\|\boldsymbol{F}[m]\|_F^2 \leq P_m$, where P_m is the maximum transmit power of the m-th DL subcarrier. The total transmit power of DL

satisfies $\sum_{m=1}^{M} P_m \leq P_{\mathrm{DL}}$. As shown in Fig. 2.1, the transmitter precoder constitutes a digital precoder $\boldsymbol{F}_{\mathrm{BB}}[m]$ for each individual DL subcarrier and an analog precoder $\boldsymbol{F}_{\mathrm{RF}}$ that is common to all DL subcarriers. Hence, the overall precoding for a DL subcarrier can be expressed as $\boldsymbol{F}[m] = \boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}}[m]$, where $\boldsymbol{F}_{\mathrm{RF}} \in \mathbb{C}^{N \times N_{\mathrm{RF}}}$, and $\boldsymbol{F}_{\mathrm{BB}}[m] \in \mathbb{C}^{N_{\mathrm{RF}} \times D}$. The baseband signal transmitted on the m-th subcarrier can be expressed as

$$\boldsymbol{s}_{\mathrm{DL}}[m] = \boldsymbol{F}[m]\boldsymbol{x}_{\mathrm{DL}}[m],\ m = 1, 2, \ldots, M, \tag{2.10}$$

where $\boldsymbol{s}[m] \in \mathbb{C}^{N \times 1}$.

2.2.5 Receiver Model

When given the transmitted signal as shown in (2.10), the received signals by the D MSs from the m-th subcarrier can be expressed as

$$\boldsymbol{y}_{\mathrm{DL}}[m] = \boldsymbol{H}_{\mathrm{DL}}[m]\boldsymbol{F}[m]\boldsymbol{x}_{\mathrm{DL}}[m] + \boldsymbol{z}_{\mathrm{DL}}[m], \tag{2.11}$$

where $\boldsymbol{H}_{\mathrm{DL}}[m] \in \mathbb{C}^{D \times N}$ and $\boldsymbol{z}[m] \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_D)$ are the DL channel matrix and additive Gaussian noise corresponding to the m-th DL subcarrier, respectively. It is noteworthy that for simplicity we ignore the interference from UL MSs to DL MSs in (2.11), so that we can focus on the SIC and CE in MDD MIMO systems[1].

On the other side, the signals transmitted at BS also propagate to its receive antenna array for the UL. This SI signal at m-th subcarrier after the analog combining can be expressed as

$$\boldsymbol{y}_{\mathrm{SI}}[m] = \boldsymbol{W}_{\mathrm{RF}}^{H}\boldsymbol{H}_{\mathrm{SI}}\boldsymbol{F}[m]\boldsymbol{x}_{\mathrm{DL}}[m], \tag{2.12}$$

In (2.12), $\boldsymbol{W}_{\mathrm{RF}} \in \mathbb{C}^{\bar{N} \times \bar{N}_{\mathrm{RF}}}$ is the analog combiner, which is regarded as the pseudo-identity matrix in digital beamforming system, i.e. $\boldsymbol{W}_{\mathrm{RF}} = \boldsymbol{I}_{\bar{N} \times \bar{N}_{\mathrm{RF}}}$ [12]. Based on (2.12), the total SI power entering the ADC in the BS receiver is given by $P_{\mathrm{SI}} = \sum_{m=1}^{M} \mathbb{E}\, \|\boldsymbol{y}_{\mathrm{SI}}[m]\|_2^2$.

Note that the distance between the transmitter and receiver arrays at BS is much smaller than the communication links from MSs to BS, which leads to SI signals to be 50-100 dB stronger than the desired signals received from UL. This means that although DL and UL are operated on different subcarriers, prior to digitization, UL signals would be overwhelmed by SI signals in ADC, if propagation- and analog-domain SIC cannot provide sufficient SI reduction. In this case, quantization noise may be significant and unable to be mitigated by any digital-domain signal processing techniques. Therefore, a certain amount of SI reduction has to be achieved to satisfy (2.9) prior to the ADC at receiver. On the other side, provided that the constraint of (2.9) is satisfied, as shown in

[1] Note furthermore that the interference from UL MSs to DL MSs in FD-based systems can be effectively mitigated via the scheduling at medium access control (MAC) layer [113–115]. However, the in-depth research in the context of MDD-based systems is required, which will be addressed in our future work.

Fig. 2.1, the received digital signals after RF processing and ADC can be expressed as

$$\boldsymbol{y}_{\mathrm{UL}}[\bar{m}] = \boldsymbol{W}_{\mathrm{RF}}^{H}\boldsymbol{H}_{\mathrm{UL}}[\bar{m}]\boldsymbol{x}_{\mathrm{UL}}[\bar{m}] + \sum_{m\in\mathcal{M}} \boldsymbol{W}_{\mathrm{RF}}^{H}\boldsymbol{H}_{\mathrm{SI}}\boldsymbol{F}[m]\boldsymbol{x}_{\mathrm{DL}}[m] + \boldsymbol{W}_{\mathrm{RF}}^{H}\boldsymbol{z}_{\mathrm{UL}}[\bar{m}], \tag{2.13}$$

where $\boldsymbol{x}_{\mathrm{UL}}[\bar{m}]$ denotes the UL data symbols satisfying $\mathbb{E}\left[\bar{\boldsymbol{x}}_{\mathrm{DL}}[m]\bar{\boldsymbol{x}}_{\mathrm{DL}}[m]^{H}\right] = \boldsymbol{I}_{\bar{D}}$, $\boldsymbol{z}_{\mathrm{UL}}[\bar{m}]$ is the complex Gaussian noise obeying the distribution of $\mathcal{CN}(0, \sigma^2\boldsymbol{I}_{\bar{N}})$. Further, after CP removing and digital-domain SIC by FFT operation, the final signal for UL detection can be expressed as

$$\boldsymbol{y}_{\mathrm{UL}}[\bar{m}] = \boldsymbol{W}^{H}[\bar{m}]\boldsymbol{H}_{\mathrm{UL}}[\bar{m}]\boldsymbol{x}_{\mathrm{UL}}[\bar{m}] + \boldsymbol{W}^{H}[\bar{m}]\boldsymbol{z}_{\mathrm{UL}}[\bar{m}] \tag{2.14}$$

for $\bar{m} = 1, 2, \ldots, \bar{M}$, where $\boldsymbol{W}[\bar{m}] = \boldsymbol{W}_{\mathrm{RF}}\boldsymbol{W}_{\mathrm{BB}}[\bar{m}]$, $\boldsymbol{W}_{\mathrm{BB}}[\bar{m}] \in \mathbb{C}^{\bar{N}_{\mathrm{RF}}\times\bar{D}}$.

2.3 Adaptive Beamforming-Based Self-Interference Cancellation

In this section, we address the beamforming-aided SIC implemented via the design of hybrid precoder/combiner. The objective of SIC is to make the SIC requirement of (2.9) be satisfied.

According to (2.12), the power of the SI signals before ADC can be evaluated as

$$P_{\mathrm{SI}} = \sum_{m=1}^{M} \mathbb{E}\left\{\mathrm{Tr}\left(\boldsymbol{y}_{\mathrm{SI}}[m]^{H}\boldsymbol{y}_{\mathrm{SI}}[m]\right)\right\} = \mathrm{Tr}\left(\boldsymbol{F}_{\mathrm{RF}}^{H}\boldsymbol{H}_{\mathrm{SI}}^{H}\boldsymbol{W}_{\mathrm{RF}}\boldsymbol{W}_{\mathrm{RF}}^{H}\boldsymbol{H}_{\mathrm{SI}}\boldsymbol{F}_{\mathrm{RF}}\sum_{m=1}^{M}\boldsymbol{F}_{\mathrm{BB}}[m]\boldsymbol{F}_{\mathrm{BB}}^{H}[m]\right). \tag{2.15}$$

From this equation we can know that P_{SI} is proportional to $P_{\mathrm{SI}}^{'} = \mathrm{Tr}(\boldsymbol{F}_{\mathrm{RF}}^{H}\boldsymbol{H}_{\mathrm{SI}}^{H}\boldsymbol{W}_{\mathrm{RF}}\boldsymbol{W}_{\mathrm{RF}}^{H}\boldsymbol{H}_{\mathrm{SI}}\boldsymbol{F}_{\mathrm{RF}})$. Hence, P_{SI} can be reduced via minimizing $P_{\mathrm{SI}}^{'}$. Therefore, according to the structure of $P_{\mathrm{SI}}^{'}$, we consider two design options for SIC, which are depended on the optimization of $\boldsymbol{F}_{\mathrm{RF}}$ and $\boldsymbol{W}_{\mathrm{RF}}$, respectively, in analog domain.

2.3.1 Precoding Optimization Aided Self-Interference Cancellation

The first design option assumes that SI suppression is solely dependent on the design of $\boldsymbol{F}_{\mathrm{RF}}$ at transmitter. Thereby, the design of combiner at receiver only focuses on the UL transmissions without considering the impact of SI. In this chapter, we consider the MMSE method for UL combining [116], and the full-digital combiner can be expressed as [117]

$$\boldsymbol{W}_{\mathrm{MMSE}}[\bar{m}] = \left(\boldsymbol{H}_{\mathrm{UL}}[\bar{m}]\boldsymbol{H}_{\mathrm{UL}}^{H}[\bar{m}] + \sigma^2\boldsymbol{I}\right)^{-1}\boldsymbol{H}_{\mathrm{UL}}[\bar{m}],$$
$$\bar{m} = 1, 2, \ldots, \bar{M}. \tag{2.16}$$

Then, the hybrid UL combiners can be designed to approximate $\boldsymbol{W}_{\mathrm{MMSE}}[\bar{m}]$, formulated as

$$\begin{aligned} &\arg \min_{\boldsymbol{W}_{\mathrm{RF}},\boldsymbol{W}_{\mathrm{BB}}[1],\ldots,\boldsymbol{W}_{\mathrm{BB}}[\bar{M}]} \sum_{\bar{m}=1}^{\bar{M}} \|\boldsymbol{W}_{\mathrm{MMSE}}[\bar{m}] - \boldsymbol{W}_{\mathrm{RF}}\boldsymbol{W}_{\mathrm{BB}}[\bar{m}]\|_F^2 \\ &\text{subject to } |(\boldsymbol{W}_{\mathrm{RF}})_{i,j}| = 1, \forall i,j. \end{aligned} \tag{2.17}$$

The optimization problem of (2.17) is a typical one in the design of hybrid beamforming, which can be solved by different algorithms, such as that in [118–122]. Specifically in the performance study in Section 2.5 of this chapter, we introduce the projected gradient descent (PGD) algorithm [121, 122]. Readers interesting in the details of the algorithm are referred to these references.

After $\boldsymbol{W}_{\mathrm{RF}}$ is obtained, the SI suppression can be executed based on the optimization of

$$\arg \min_{\boldsymbol{F}_{\mathrm{RF}}} \ f(\boldsymbol{F}_{\mathrm{RF}}) \tag{2.18a}$$

$$\text{subject to } |(\boldsymbol{F}_{\mathrm{RF}})_{i,j}| = 1, \forall i,j, \tag{2.18b}$$

$$\frac{P_{\mathrm{DL}}}{P_{\mathrm{SI}}} \geq 10^{\frac{C_{\mathrm{SI}}^{\mathrm{Demand}}}{10}}. \tag{2.18c}$$

where $f(\boldsymbol{F}_{\mathrm{RF}}) = \mathrm{Tr}(\boldsymbol{H}_{\mathrm{w}}\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{RF}}^H\boldsymbol{H}_{\mathrm{w}}^H)$ with $\boldsymbol{H}_{\mathrm{w}} = \boldsymbol{W}_{\mathrm{RF}}^H\boldsymbol{H}_{\mathrm{SI}}$, which is equal to the $P_{\mathrm{SI}}^{'}$ in (2.15). However, this optimization problem is non-convex and hard to solve. To circumvent this dilemma and reduce the computational complexity, we propose an adaptive algorithm based on the CCD algorithm [123–126], so as to dynamically suppress the SI. It is noteworthy that the performance of the CCD algorithm is sensitive to the initialization [127]. Hence, in order to suppress SI while simultaneously maintaining the required performance of DL, the initialization of $\boldsymbol{F}_{\mathrm{RF}}$ is very important in our algorithm. In Section 2.5.1, we will investigate the effect of the initialization of $\boldsymbol{F}_{\mathrm{RF}}$ on the achievable performance. Furthermore, to calculate P_{SI} during the optimization process, the digital precoder for the m-th subcarrier is assumed to be in ZF principle, which yields

$$\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{ZF}}[m] = \boldsymbol{H}_{\mathrm{eq}}^H[m]\left(\boldsymbol{H}_{\mathrm{eq}}[m]\boldsymbol{H}_{\mathrm{eq}}^H[m]\right)^{-1}\boldsymbol{P}^{\frac{1}{2}}[m],\ m \in \mathcal{M}, \tag{2.19}$$

where $\boldsymbol{H}_{\mathrm{eq}}[m] = \boldsymbol{H}_{\mathrm{DL}}[m]\boldsymbol{F}_{\mathrm{RF}}$ is the resultant channel after the analog precoding, $\boldsymbol{P}[m] = \mathrm{diag}(p_1[m], p_2[m], ..., p_D[m])$ is obtained from the water-filling algorithm [128].

To solve the optimization problem (2.18) iteratively, the element $(\boldsymbol{F}_{\mathrm{RF}})_{i,j}$ is firstly optimized by assuming that all the other elements are fixed. In this case, (2.18a) can be simplified to

$$f(\boldsymbol{F}_{\mathrm{RF}}) = \mathrm{Tr}(\boldsymbol{A}_j) + \boldsymbol{F}_{\mathrm{RF}}^{(:,j)^H}\boldsymbol{H}_{\mathrm{w}}^H\boldsymbol{H}_{\mathrm{w}}\boldsymbol{F}_{\mathrm{RF}}^{(:,j)} = \mathrm{Tr}(\boldsymbol{A}_j) + \zeta_{ij}^{\tilde{H}} + 2\mathrm{Re}\{(\boldsymbol{F}_{\mathrm{RF}}^*)_{i,j}\,\eta_{ij}^{\tilde{H}}\}, \tag{2.20}$$

where $\zeta_{ij}^{\tilde{H}} = \left(\tilde{\boldsymbol{H}}_j\right)_{i,i} + 2\mathrm{Re}\{\sum_{m\neq i, n\neq i}\left(\boldsymbol{F}_{\mathrm{RF}}^{*}\right) m, j\left(\tilde{\boldsymbol{H}}_j\right)_{m,n}\left(\boldsymbol{F}_{\mathrm{RF}}\right)_{n,j}\}$, $\eta_{ij}^{\tilde{H}} = \sum_{m\neq i}\left(\tilde{\boldsymbol{H}}_j\right)_{i,m}\left(\boldsymbol{F}_{\mathrm{RF}}\right)_{m,j}$, associated with $\tilde{\boldsymbol{H}}_j = \boldsymbol{H}_{\mathrm{w}}^{H}\boldsymbol{H}_{\mathrm{w}}$, and $\boldsymbol{A}_j = \boldsymbol{H}_{\mathrm{w}}\bar{\boldsymbol{F}}_{\mathrm{RF}}^{j}(\bar{\boldsymbol{F}}_{\mathrm{RF}}^{j})^{H}\boldsymbol{H}_{\mathrm{w}}^{H}$, where $\bar{\boldsymbol{F}}_{\mathrm{RF}}^{j}$ denotes a matrix obtained from $\boldsymbol{F}_{\mathrm{RF}}$ with its j-th column $\boldsymbol{f}_{\mathrm{RF}}^{(j)}$ removed.

Since all the elements in $\boldsymbol{F}_{\mathrm{RF}}$ other than $\left(\boldsymbol{F}_{\mathrm{RF}}\right)_{i,j}$ are fixed, $\boldsymbol{A}_j$, $\zeta_{ij}^{\tilde{H}}$ and $\eta_{ij}^{\tilde{H}}$ seen in (2.20) are complex constants. Furthermore, under the modulus constraint of the analog precoder, i.e., $\left(\boldsymbol{F}_{\mathrm{RF}}\right)_{i,j} = e^{-j\theta_{ij}}$, (2.20) can be re-stated as

$$g(\theta_{ij}) = \mathrm{Tr}(\boldsymbol{A}_j) + \zeta_{ij}^{\tilde{H}} + 2\mathrm{Re}\{\eta_{ij}^{\tilde{H}} e^{j\theta_{ij}}\} = \mathrm{Tr}(\boldsymbol{A}_j) + \zeta_{ij}^{\tilde{H}} + \eta_{ij}^{\tilde{H}} e^{j\theta_{ij}} + (\eta_{ij}^{\tilde{H}})^{*} e^{-j\theta_{ij}}. \tag{2.21}$$

Now the optimization is converted to an extreme-value problem, which can be readily solved. In detail, upon taking the derivative of $g(\theta_{ij})$ with respect to θ_{ij}, we obtain $\frac{\partial g(\theta_{ij})}{\partial \theta_{ij}} = j\eta_{ij}^{\tilde{H}} e^{j\theta_{ij}} - j(\eta_{ij}^{\tilde{H}})^{*} e^{-j\theta_{ij}} = 0$, which is equivalent to

$$\mathrm{Re}\{\eta_{ij}^{\tilde{H}}\}\sin(\theta_{ij}) + \mathrm{Im}\{\eta_{ij}^{\tilde{H}}\}\cos(\theta_{ij}) = 0. \tag{2.22}$$

Let us represent it in the form of

$$|\eta_{ij}^{\tilde{H}}|\sin(\theta_{ij} + \phi_{ij}) = 0, \tag{2.23}$$

where

$$\phi_{ij} = \begin{cases} \sin^{-1}\left(\frac{\mathrm{Im}\{\eta_{ij}^{\tilde{H}}\}}{|\eta_{ij}^{\tilde{H}}|}\right), & \text{if } \mathrm{Re}\{\eta_{ij}^{\tilde{H}}\} \geq 0, \\ \pi - \sin^{-1}\left(\frac{\mathrm{Im}\{\eta_{ij}^{\tilde{H}}\}}{|\eta_{ij}^{\tilde{H}}|}\right), & \text{if } \mathrm{Re}\{\eta_{ij}^{\tilde{H}}\} < 0. \end{cases} \tag{2.24}$$

Solving (2.24) under the constraint of $\theta_{ij} \in (0, 2\pi)$, we obtain

$$\begin{aligned} \theta_{ij}^{(1)} &= -\phi_{ij}, \\ \theta_{ij}^{(2)} &= \pi - \phi_{ij}. \end{aligned} \tag{2.25}$$

However, there is only one solution yielding a minimum value of $g(\theta_{ij})$, i.e., the optimum solution. Hence, the final solution to θ_{ij} is given by

$$\theta_{ij}^{(\mathrm{opt})} = \underset{\theta_{ij}^{(1)}, \theta_{ij}^{(2)}}{\arg\min}\left(g(\theta_{ij}^{(1)}), g(\theta_{ij}^{(2)})\right). \tag{2.26}$$

The above optimization process is repeated with respect to each of the elements in $\boldsymbol{F}_{\mathrm{RF}}$, and the elements of $\boldsymbol{F}_{\mathrm{RF}}$ are iteratively optimized until the cost function converges to a local minimum. Hence, before reaching the minimum, SI can be gradually suppressed with the increase of the number of iterations. A shortcoming of the CCD algorithm is that convergence is usually slow and dependent on the cost function of (2.18a), which in turn related to the SI channel and antenna configurations [111]. Nevertheless, once the

condition of (2.18c) is satisfied, meaning that the SI reduction provided by beamforming is sufficient to make ADC work efficiently, more iterations for further SI reduction is no longer necessary. Hence, once the constraint of (2.18c) is met, the process of SI suppression can be terminated to save time. In Section 2.5.1, we will demonstrate the convergence performance of the CCD algorithm.

In summary, the first design option is stated as Algorithm 1. For UL reception, when given $\boldsymbol{W}_{\mathrm{MMSE}}[\bar{m}]$, the hybrid combiner is obtained by the PGD algorithm. For SI suppression, after the initialization of $\boldsymbol{F}_{\mathrm{RF}}$, the analog precoder is iteratively updated to reduce the SI power based on the CCD algorithm, until the SIC meets the requirement. During every iteration, the digital precoder $\boldsymbol{F}_{\mathrm{BB}}[m]$ is derived based on ZF method and water-filling algorithm.

Algorithm 1: Precoding optimization aided SIC (Option 1)

Input : P_m, P_{DL}, $C_{\mathrm{SI}}^{\mathrm{Demand}}$

Compute $\boldsymbol{W}_{\mathrm{BB}}[\bar{m}]$ and $\boldsymbol{W}_{\mathrm{RF}}$ by the PGD method after obtaining $\boldsymbol{W}_{\mathrm{MMSE}}[\bar{m}]$ in (2.16);
Initialize $\boldsymbol{F}_{\mathrm{RF}}$;
for $1 \leq j \leq N_{RF}$ **do**
  for $1 \leq i \leq N$ **do**
    Calculate $\eta_{ij}^{\tilde{H}}$ and $g(\theta_{ij})$ in (2.20) and (2.21), respectively;
    Derive $\theta_{ij}^{(1)}$ and $\theta_{ij}^{(2)}$ via minimizing $g(\theta_{ij})$;
    Find $\theta_{ij}^{(\mathrm{opt})} := \arg\min\left(g(\theta_{ij}^{(1)}), g(\theta_{ij}^{(2)})\right)$;
    $\boldsymbol{F}_{\mathrm{RF}}(i,j) = e^{-j\theta_{ij}^{(\mathrm{opt})}}$
  end for
end for
for $1 \leq m \leq M$ **do**
  Derive $\boldsymbol{P}[m]$ based on the water-filling algorithm;
  Compute $\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{ZF}}[m] = \boldsymbol{H}_{\mathrm{eq}}^{H}[m](\boldsymbol{H}_{\mathrm{eq}}[m]\boldsymbol{H}_{\mathrm{eq}}^{H}[m])^{-1} \times \boldsymbol{P}^{\frac{1}{2}}[m]$, where $\boldsymbol{H}_{\mathrm{eq}}[m] = \boldsymbol{H}_{\mathrm{DL}}[m]\boldsymbol{F}_{\mathrm{RF}}$;
end for
Calculate the power of SI: $P_{\mathrm{SI}} = \sum_{m=1}^{M} \mathbb{E}\,\|\boldsymbol{y}_{\mathrm{SI}}[m]\|_2^2$;
if $\frac{P_{DL}}{P_{SI}} \geq 10^{\frac{C_{SI}^{Demand}}{10}}, \forall m$ **then**
  store $\boldsymbol{F}_{\mathrm{RF}}$ and $\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{ZF}}[m]$;
else
  go to Step 3.
end if

2.3.2 Combining Optimization Aided Self-Interference Cancellation

In the context of the second design option, we assume that $\bar{N} > N$. In this scenario, the analog precoder $\boldsymbol{F}_{\mathrm{RF}}$ is derived via maximizing the DL spectral efficiency without considering the effect of the SI on the UL receiving. Instead, SI suppression is only

attempted by the design of $\boldsymbol{W}_{\mathrm{RF}}$. Therefore, the design of the hybrid combiner in Option 2 is similar to the design of the hybrid precoder in Option 1, except that there is no PA in combiner's design. Furthermore, the analog and digital precoders in Option 2 can be designed by employing the PGD algorithm, when an overall precoder in the form of (2.19) is prepared.

It can be argued that the design in Option 2 has lower complexity than that in Option 1. The reason is that in Option 1, the digital precoder and analog precoder need to be iteratively updated, so that the SI on the UL receiving is gradually reduced to an allowed value. By contrast, in Option 2, once the hybrid DL precoder is obtained, the SI on UL receiver becomes stable. Hence, no iteration is required between the design of the digital combiner and that of the analog combiner. In other words, the analog combiner can be firstly designed to suppress SI to a sufficient level. Then, digital combiner can be derived for a fixed analog combiner. In summary, the design in Option 2 is stated as Algorithm 2.

It is worth noting that following Options 1 and 2, there is a third option for the design, which optimizes $\boldsymbol{W}_{\mathrm{RF}}$ and $\boldsymbol{F}_{\mathrm{RF}}$ jointly. However, it can be shown that the SIC performance is mainly determined by the DL transmitter or UL receiver, depending on which of them has more antenna elements. As demonstrated in Section 2.5, the side (either DL transmitter or UL receiver) with less antenna elements can hardly provide any gain for SIC. Owing to this, the third design option is not further considered in this chapter.

Algorithm 2: Combining Optimization Aided SIC (Option 2)

Input : P_m, P_{DL}, $C_{\mathrm{SI}}^{\mathrm{Demand}}$

Derive full-digital precoder $\boldsymbol{F}^{\mathrm{ZF}}[m]$ and divide it into $\boldsymbol{F}_{\mathrm{RF}}$ and $\boldsymbol{F}_{\mathrm{BB}}[m]$ based on PGD;

Initialize $\boldsymbol{W}_{\mathrm{RF}}$ via maximizing the UL spectral-efficiency in [123];

Design $\boldsymbol{W}_{\mathrm{RF}}$ by following Step 3 to Step 10 in Algorithm 1;

Calculate the power of SI: $P_{\mathrm{SI}} = \sum_{m=1}^{M} \mathbb{E} \left\| \boldsymbol{y}_{\mathrm{SI}}[m] \right\|_2^2$;

if $\frac{P_{DL}}{P_{SI}} \geq 10^{\frac{C_{SI}^{Demand}}{10}}, \forall m$ **then**

  store $\boldsymbol{W}_{\mathrm{RF}}$;

else

  go to Step 3.

end if

for $1 \leq m \leq \bar{M}$ **do**

  Compute $\boldsymbol{W}_{\mathrm{BB}}[\bar{m}] = \left(\boldsymbol{H}_{\mathrm{eq}}[\bar{m}] \boldsymbol{H}_{\mathrm{eq}}^{H}[\bar{m}] + \sigma^2 \boldsymbol{I} \right)^{-1}$

  $\times \boldsymbol{H}_{\mathrm{eq}}[\bar{m}]$, where $\boldsymbol{H}_{\mathrm{eq}}[\bar{m}] = \boldsymbol{W}_{\mathrm{RF}}^{H} \boldsymbol{H}_{\mathrm{UL}}[\bar{m}]$;

end for

2.4 Channel Estimation in MDD MIMO

As argued in Section 1.2.1, in MDD systems, the DL channels can be estimated based on the observations received from the UL channels by exploiting the reciprocity existing between the DL and UL subchannels, which is generated by the frequency-domain correlated fading. In this section, we consider the CE in MDD MIMO systems. We first consider the CE based on orthogonal transmission and focus on the design of frequency-domain pilot symbol (PS) vectors and the associated conditions. Then, the CE in the scenario of non-orthogonal transmission is considered. Note that, below we only consider the CE of communication channels. The SI channel can be estimated by the various approaches proposed in references, e.g., in [129–132].

By observing (2.2), (2.3) and (2.4), we can know that the CE can be initialized with the UL training in frequency-domain. Using the frequency-domain training, the time-domain channel $\boldsymbol{g}_{n,d}$ can be estimated. Then, from $\boldsymbol{g}_{n,d}$ both the DL and UL channels of the subcarriers can be obtained with the aid of (2.3) and (2.4).

Let us assume that all MSs synchronously transmit their PSs. For CE, we assume that each MS transmits PSs on all the UL subcarriers. The fading of channels is assumed to be slow enough for making use of the reciprocity for UL/DL processing. Then, consider that the PSs transmitted by the d-th MS are given by $\boldsymbol{x}_d = [x_d[1], ..., x_d[\bar{m}], ..., x_d[\bar{M}]]^T$, where d is either a UL MS or a DL MS. The received signal from the $\bar{m}$-th subcarrier by the n-th antenna (either the transmit or receive antenna at BS) can be expressed as

$$y_n[\bar{m}] = \sqrt{\rho_r} \sum_{d=1}^{D_{\text{sum}}} h_{n,d}[\bar{m}] x_d[\bar{m}] + z_n[\bar{m}],\ \bar{m} = 1, ..., \bar{M}, \tag{2.27}$$

where $z_n[\bar{m}] \sim \mathcal{CN}(0, \sigma^2)$ is the additive complex Gaussian noise, ρ_r represents the average transmit power per symbol per MS. Let $\boldsymbol{y}_n = [y_n[1], \ldots, y_n[\bar{M}]]^T$. Then, we can have an expression of

$$\boldsymbol{y}_n = \sqrt{\rho_r} \sum_{d=1}^{D_{\text{sum}}} \boldsymbol{X}_d \boldsymbol{h}_{n,d}^{\text{UL}} + \boldsymbol{z}_n = \sqrt{\rho_r} \sum_{d=1}^{D_{\text{sum}}} \boldsymbol{X}_d \boldsymbol{\Phi}_{\text{UL}}^T \boldsymbol{F} \boldsymbol{\Psi} \boldsymbol{g}_{n,d} + \boldsymbol{z}_n, \tag{2.28}$$

where $\boldsymbol{X}_d = \text{diag}\{\boldsymbol{x}_d\}$, $\boldsymbol{z}_n \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_{\bar{M}})$, $\boldsymbol{h}_{n,d}^{\text{UL}}$ is given by (2.4), and with the aid of (2.2), i.e., $\boldsymbol{h}_{n,d}^{\text{UL}} = \boldsymbol{\Phi}_{\text{UL}}^T \boldsymbol{F} \boldsymbol{\Psi} \boldsymbol{g}_{n,d}$, $\boldsymbol{y}_n$ is directly expressed in terms of the TDCIR $\boldsymbol{g}_{n,d}$.

From (2.28), we can see that if the PSs can be designed to satisfy

$$\begin{cases} \left(\boldsymbol{X}_i \boldsymbol{\Phi}_{\text{UL}}^T \boldsymbol{F} \boldsymbol{\Psi}\right)^H \left(\boldsymbol{X}_i \boldsymbol{\Phi}_{\text{UL}}^T \boldsymbol{F} \boldsymbol{\Psi}\right) = \varrho \boldsymbol{I}_L, \\ \left(\boldsymbol{X}_i \boldsymbol{\Phi}_{\text{UL}}^T \boldsymbol{F} \boldsymbol{\Psi}\right)^H \left(\boldsymbol{X}_j \boldsymbol{\Phi}_{\text{UL}}^T \boldsymbol{F} \boldsymbol{\Psi}\right) = \boldsymbol{0}_L,\ \forall\, i \neq j, \end{cases} \tag{2.29}$$

where ϱ is a constant, then the TDCIR $\boldsymbol{g}_{n,i}$ from MS i can be readily estimated by the LS method, given by

$$\hat{\boldsymbol{g}}_{n,i} = \frac{1}{\varrho\sqrt{\rho_r}} \left(\boldsymbol{X}_i \boldsymbol{\Phi}_{\mathrm{UL}}^T \boldsymbol{F} \boldsymbol{\Psi} \right)^H \boldsymbol{y}_n. \tag{2.30}$$

Furthermore, if MS i is a UL user and Antenna n is the receive antenna at BS, then BS uses (2.4) to obtain the frequency-domain channel gains for UL detection. By contrast, if MS i is a DL user and Antenna n is the transmit antenna of BS, then BS uses (2.3) to obtain the frequency-domain channel gains for DL precoding.

The set of PSs satisfying (2.30) are referred to as the orthogonality-achieving PSs. With the aid of the approach proposed in [133], it can be shown that if $\bar{M} \geq L$ and the $\bar{M}$ UL subcarriers are evenly distributed among the M_{sum} subcarriers, the set of PSs given by

$$\begin{aligned} \boldsymbol{x}_d = & \left[1, e^{j2\pi \frac{(d-1)\xi}{\bar{M}}}, e^{j2\pi \frac{2(d-1)\xi}{\bar{M}}}, \cdots, e^{j2\pi \frac{(\bar{M}-1)(d-1)\xi}{\bar{M}}} \right]^T, \\ & 1 \leq d \leq D_{\mathrm{sum}}, \end{aligned} \tag{2.31}$$

where $\xi = \left\lfloor \frac{\bar{M}}{D_{\mathrm{sum}}} \right\rfloor$, are orthogonality-achieving PSs. Hence, we have the following Proposition.

Proposition 1. When $\xi = \left\lfloor \frac{\bar{M}}{D_{\mathrm{sum}}} \right\rfloor \geq L$, and the $\bar{M}$ number of UL subcarriers are evenly distributed with a spacing of $l = M_{\mathrm{sum}}/\bar{M}$, the D_{sum} PSs given in (2.31) make the matrices $\boldsymbol{P}[d] = \boldsymbol{X}_d \boldsymbol{\Phi}_{\mathrm{UL}}^T \boldsymbol{F} \boldsymbol{\Psi}$ for $d = 1, ..., D_{\mathrm{sum}}$ mutually orthogonal.

Proof. See Appendix A. □

However, if the conditions stated in Proposition 1 are not satisfied, or more random PSs are used, orthogonality-achieving PSs may not be available. In this case, we can write (2.28) as

$$\boldsymbol{y}_n = \boldsymbol{Q}_i \boldsymbol{g}_{n,i} + \boldsymbol{T}_i + \boldsymbol{z}_n, \tag{2.32}$$

where by definition, $\boldsymbol{Q}_i = \sqrt{\rho_r} \boldsymbol{X}_i \boldsymbol{\Phi}_{\mathrm{UL}}^T \boldsymbol{F} \boldsymbol{\Psi}$, and $\boldsymbol{T}_i = \sqrt{\rho_r} \sum_{d=1, d \neq i}^{D_{\mathrm{sum}}} \boldsymbol{X}_d \boldsymbol{\Phi}_{\mathrm{UL}}^T \boldsymbol{F} \boldsymbol{\Psi} \boldsymbol{g}_{n,d}$ is the interference signal from other MSs. In order to suppress the interference from the other MSs, let us introduce the LMMSE estimator for CE. Then, the estimate to $\boldsymbol{g}_{n,i}$ in (2.28) can be expressed as $\hat{\boldsymbol{g}}_{n,i} = \boldsymbol{A}_{n,i}^H \boldsymbol{y}_n$, where $\boldsymbol{A}_{n,i}$ achieving MMSE can be obtained from the minimization problem formulated as

$$\boldsymbol{A}_{n,i}^{\mathrm{LMMSE}} = \arg\min_{\boldsymbol{A}} E \left[\left\| \boldsymbol{g}_{n,i} - \boldsymbol{A}^H \boldsymbol{y}_n \right\|^2 \right]. \tag{2.33}$$

Assume that the TDCIRs from MSs are uncorrelated, i.e., $E[\boldsymbol{g}_{n,d} \boldsymbol{g}_{n,d'}^H] = \boldsymbol{0}, \forall d \neq d'$, which is usually satisfied as MSs are in general well separated in space. Then, the

solution to (2.33) is

$$\boldsymbol{A}_{n,i}^{\mathrm{LMMSE}} = \left(\frac{1}{L}\boldsymbol{Q}_i\boldsymbol{Q}_i^H + E[\boldsymbol{T}_i\boldsymbol{T}_i^H] + \sigma^2\boldsymbol{I}_{\bar{M}}\right)^{-1}\left(\frac{1}{L}\boldsymbol{Q}_i\right), \tag{2.34}$$

where $E[\boldsymbol{T}_i\boldsymbol{T}_i^H] = \frac{\rho_r}{L}\sum_{d=1,d\neq i}^{D_{\mathrm{sum}}}\boldsymbol{X}_d\boldsymbol{\Phi}_{\mathrm{UL}}^T\boldsymbol{F}\boldsymbol{\Psi}\boldsymbol{\Psi}^H\boldsymbol{F}^H\boldsymbol{\Phi}_{\mathrm{UL}}^*\boldsymbol{X}_d^H$. Correspondingly, the estimate $\hat{\boldsymbol{g}}_{n,i}$ is given by

$$\hat{\boldsymbol{g}}_{n,i} = \frac{1}{L}\boldsymbol{Q}_i^H\left(\frac{1}{L}\boldsymbol{Q}_i\boldsymbol{Q}_i^H + E[\boldsymbol{T}_i\boldsymbol{T}_i^H] + \sigma^2\boldsymbol{I}_{\bar{M}}\right)^{-1}\boldsymbol{y}_n. \tag{2.35}$$

It is well-known that LMMSE yields biased estimation. To attain an unbiased estimator, we can form the estimate as

$$\check{\boldsymbol{g}}_{n,i} = \left[\left(\boldsymbol{A}_{n,i}^{\mathrm{LMMSE}}\right)^H\boldsymbol{Q}_i\right]^{-1}\hat{\boldsymbol{g}}_{n,i}. \tag{2.36}$$

Again, after obtaining the estimation of $\check{\boldsymbol{g}}_{n,i}$, if MS i is a UL user and Antenna n is the UL receive antenna at BS, BS uses (2.4) to obtain the frequency-domain channel for UL detection. By contrast, if MS i is a DL user and Antenna n is the DL transmit antenna at BS, BS uses (2.3) to obtain the frequency-domain channel for DL precoding.

2.5 Performance Results

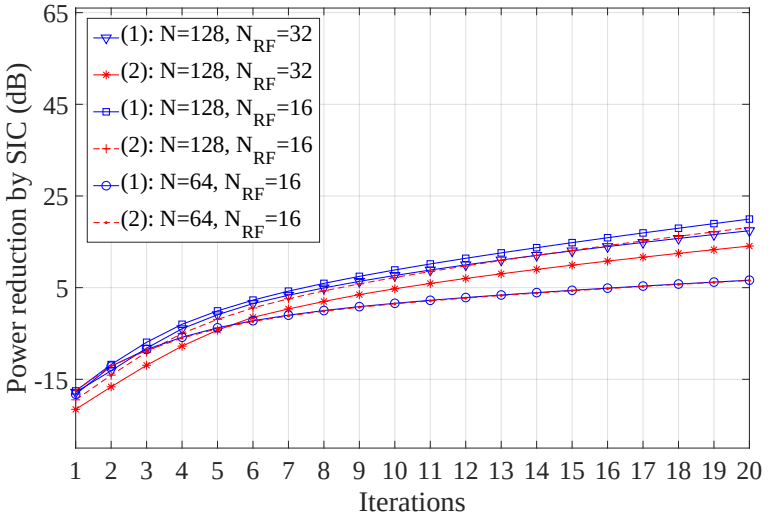
In this section, we first evaluate the SIC performance. Then, under the assumption of ideal CE, the performance results for the MDD MIMO systems are depicted and discussed, when the precoder and combiner designed in Section 2.3 are employed. Then, the performance of the CE method introduced in Section 2.4 is investigated. Finally, we compare the MDD and IBFD hybrid beamforming systems employed with the proposed SIC algorithm, in terms of energy efficiency.

In our simulations, we assume the channel model as presented in Section 2.2.1, and the ULA at BS with the half-wavelength spacing between two adjacent antenna elements. The distance r_{ij} between the i-th element of transmitter and the j-th element of receiver is set according to [106] and the default angle between the transmitter array and receiver array is $\varphi = 120^\circ$. We further assume that the number of CIR taps for communications channel is $L = 4$. For the SI channel, we set $\kappa = 20$ dB as a default value.

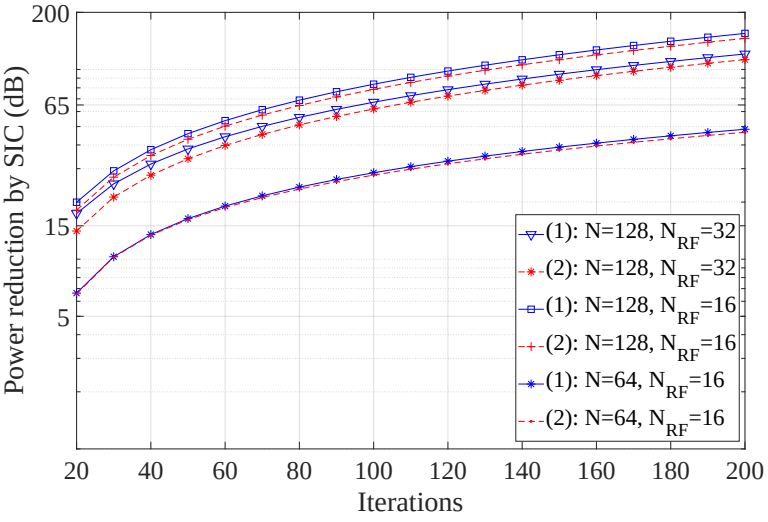
2.5.1 Performance of Self-Interference Cancellation

According to [134], we assume that the transmit power of BS and the signaling bandwidth are $P_{\mathrm{DL}} = 30$ dBm and $B = 20$ MHz, respectively. The total transmit power is

uniformly allocated to the DL subcarriers. We further assume the 12-bit ADCs used by the UL receiver at BS. Then, the maximum input power to the UL receiver and the demand of SIC can be found from (2.7) and (2.8), which are $s_{\max} \approx -25$ dBm, $C_{\text{SI}}^{\text{Demand}} = 65$ dB, respectively. In other words, the system needs to achieve at least 65 dB of SI reduction, so that the UL receiver at BS can work efficiently.



(A) Iterations: 1-20



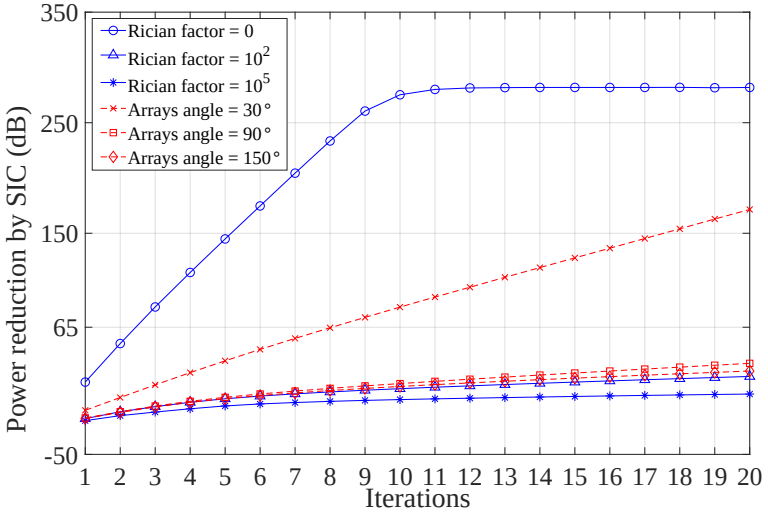
(B) Iterations: 20-200

FIGURE 2.2: SI reduction versus number of iterations with $P_{\text{DL}} = 30$ dBm, $B = 20$ MHz, $\varphi = 120^\circ$, $\kappa = 20$ dB. (1) Optimized initial analog precoder; (2) Random initial analog precoder.

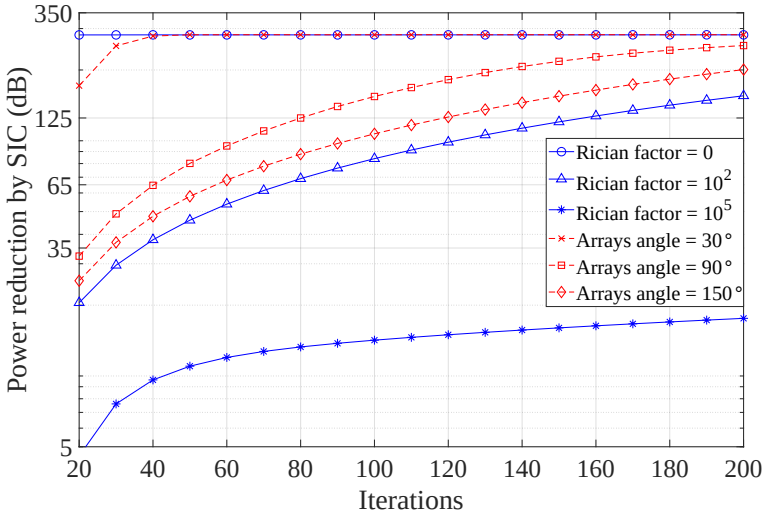
In the first experiment, we demonstrate the SIC performance in the MDD MIMO systems with the transceivers designed under Option 1 of Section 2.3. In this study, we assume that the number of antennas and RF chains at UL receiver are $\bar{N} = 32$ and $\bar{N}_{\text{RF}} = 8$, respectively. The other parameters are detailed under the caption of Fig. 2.2. In this figure, we compare the SIC performance of the proposed iterative coordinate descent algorithm, when the analog precoder is either randomly initialized or initialized

via optimizing the DL sum-rate as in equation (19) in [123], which are referred to as the 'Random initial' and 'Optimized initial' analog precoders, respectively, in the figure. Note that the random initial analog precoder is obtained by firstly extracting the angle information from the null space matrix of $\boldsymbol{H}_{\mathrm{SI}}$, i.e., from $\boldsymbol{V}_{(N-\mathrm{rank}(\boldsymbol{H}_{\mathrm{SI}})):\mathrm{end}}$, and then, N_{RF} columns of $\boldsymbol{V}_{(N-\mathrm{rank}(\boldsymbol{H}_{\mathrm{SI}})):\mathrm{end}}$ are randomly selected to construct F_{RF}. Note furthermore that the SIC behavior from the 1st to the 20th iterations are separately depicted in Fig. 2.2(A) to highlight the relatively sharp changes. As shown in Fig. 2.2, explicitly, the SI reduces with the increase of the number of iterations. Specifically, when BS employs $N = 128$ DL transmit antennas and $N_{\mathrm{RF}} = 16$ DL RF chains, 65 dB of SIC reduction can be achieved after 80 iterations, when either the random initial or optimized initial analog precoder is employed. The performance achieved by the random initial analog precoder and optimized initial analog precoder is very similar. Fig. 2.2 illustrates that employing more DL transmit antennas allows more SI reduction for a given number of iterations. For $N_{\mathrm{RF}} = 16$ and when the number of DL transmit antennas is decreased from 128 to 64, the SI reduction capability is reduced from about 65 dB to about 25 dB after 80 iterations. Furthermore, Fig. 2.2 shows that when the number of RF chains is reduced from 32 to 16 while keeping the number of DL transmit antennas fixed, more than 2.5 dB of SI reduction can be obtained. Note that, the above observations can be similarly obtained from the systems operated under Option 2, when the above-stated DL transmit antennas is changed to the UL receive antennas.

The second experiment considers the effect of the Rician factor κ and the arrays angle φ on the performance of SIC. Again, we assume that the system is operated under Option 1 with the parameters as shown in the caption of Fig. 2.3, where the range from the 1st to the 20th iterations are separately depicted in Fig. 2.3(A) for highlighting the behavior. It can be observed that the Rician factor and arrays angle yield big impacts on the SIC performance. Specifically in terms of the Rician factor, it can be seen that when the LoS component becomes more dominant, the proposed SIC method becomes less efficient. However, we should note that while the beamforming methods work inefficiently with the LoS component, other easy-implementing approaches, such as, of adding blockage between transmitter and receiver [29,35,39,72,135], may be employed to significantly mitigate the LoS SI. In terms of the angle between the transmit and receive antenna arrays, Fig. 2.3 shows that the narrower angel is set, the more SI can be reduced. This is because when the angle is narrower, the SI power imposed by the SI signal of a given transmit array element on all the receiver array elements is nearly same, which is beneficial to SIC by using beamforming technique. By contrast, if the angle between transmit/receive arrays is wide, the distances from a given transmit array element to all the receive antenna array elements can be very different. Hence, the SI power from a given transmit array element to the receive array elements is very different. Consequently, it is difficult for the beamformer to simultaneously suppress them efficiently.



(A) Iterations: 1-20



(B) Iterations: 20-200

FIGURE 2.3: SIC versus number of iterations performance with respect to different angles between the transmit and receive antenna arrays, as well as to different Rician factors of SI channel, when assuming $P_{\mathrm{DL}} = 30$ dBm, $B = 20$ MHz, $N = 128$, $\bar{N} = 32$, $N_{\mathrm{RF}} = 16$ and $\bar{N}_{\mathrm{RF}} = 8$.

Additionally, from Figs. 2.2 and 2.3 we can be implied that the SI reducing rate and the SIC potential provided by the proposed CCD algorithm are dependent on the antenna configuration and SI channel's characteristics. For instance, if the SI channel only has NLOS components, the CCD algorithm makes the SI approach a fixed value after about 10 iterations, yielding a SI reduction of about 300 dB, which is much larger than $C_{\mathrm{SI}}^{\mathrm{Demand}}$ of required. As shown in Figs. 2.2 and 2.3, in some cases, the SI reducing rate is relatively small, but the CCD algorithm can still allow to achieve the required SIC. For example, when the Rician factor is 100, the algorithm is able to provide about 65 dB SI reduction after about 80 iterations. In some other cases, such as, when the Rician factor is 10^5, the SI reducing rate is very small and the SIC requirement of (2.18b) is hard to meet, even after many iterations. However, it is worth noting that in this case,

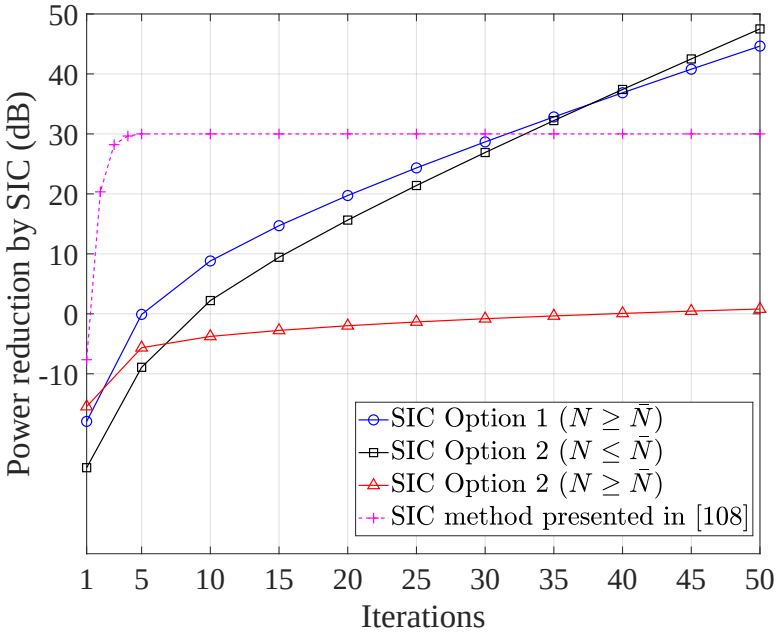


FIGURE 2.4: SIC performance comparison of the proposed methods with Option 1 and Option 2 and the method presented in [108], when $P_{\mathrm{DL}} = 30$ dBm, $B = 20$ MHz, $\varphi = 120°$, and $\kappa = 20$ dB.

the LoS propagation is dominant and the SI can be efficiently suppressed by the other approaches of, such as, using blockage. In summary, from the results of Figs. 2.2 and 2.3 we can know that to attain a good performance of SI reduction, we can increase the number of antennas at the side implementing SIC and/or reduce the angle between TX/RX antenna arrays.

Finally, we compare the SIC performance of the MDD MIMO systems with the transceivers designed by Option 1 and Option 2, respectively, in Fig. 2.4. The beamforming based SIC algorithm presented in [108] is shown as the benchmark. In this study, we set the parameters to $N = 128$, $\bar{N} = 32$, $N_{\mathrm{RF}} = 16$ and $\bar{N}_{\mathrm{RF}} = 8$ for Option 1, for Option 2 in the case of $N \geq \bar{N}$ and also for the algorithm presented in [108]. For the Option 2 in the case of $N < \bar{N}$, we set $N = 32, \bar{N} = 128, N_{\mathrm{RF}} = 8$, and $\bar{N}_{\mathrm{RF}} = 16$. From Fig. 2.4 we observe that Option 1 and Option 2 are equally efficient for SI mitigation, when the same number of antennas used for SI suppression is assumed. Moreover, as shown in Fig. 2.4, in the case of $N \geq \bar{N}$, if Option 2 is employed to mitigate SI, i.e., SI is suppressed by receive antenna array, the SIC gain is very limited. Therefore, when given the deployment of $N \geq \bar{N}$ or $N \leq \bar{N}$, there is not much meaning to implement the joint transmit/receive beamforming for SI mitigation. This is because in contrast to using either transmit beamforming in the case of $N \geq \bar{N}$ or receive beamforming under $N \leq \bar{N}$, the SIC gain provided by the joint transmit/receive beamforming is marginal, while the increase of implementation complexity is significant. Additionally, when comparing the SIC Option 1 with Fig. 2.2(b) in the case of (1): $N = 128$, $N_{\mathrm{RF}} = 16$, both cases use the same parameters, we can see that the proposed algorithm is capable

TABLE 2.2: Comparison between the proposed and competing beamforming aided SIC methods

Method	Constraint	Complexity	SIC Capability
Option 1	$N \geq \bar{N}$	$\mathcal{O}(\bar{N}^3 + N\bar{N}\bar{N}_{\mathrm{RF}})$	high
Option 2	$N \leq \bar{N}$	$\mathcal{O}(N\bar{N}\bar{N}_{\mathrm{RF}})$	high
Option 2	$N \geq \bar{N}$	$\mathcal{O}(N\bar{N}\bar{N}_{\mathrm{RF}})$	low
[108]	$N \geq \bar{N}$	$\mathcal{O}(\bar{N}^3 + NN_{\mathrm{RF}}^2)$	medium

of providing 65 dB of required SI reduction after 80 iterations. By contrast, although the SIC algorithm presented in [108] can provide 30 dB SI reduction after 5 iterations, it however gets saturated at this value and is unable to achieve the required SI reduction in analog-domain, no matter how many iterations of the algorithm are executed.

In Table 2.2, we summarize the comparison between the proposed SIC methods and the method presented in [108] in terms of their complexity and SIC capability. The complexity shown in Table 2.2 includes both the complexity for SI reduction and that for digital precoder/combiner, when the number of antenna elements and that of RF chains at transmitter or receiver are given. First, regarding to the SIC performance, as we stated in Section 2.3.2, both Option 1 and Option 2 ($N \leq \bar{N}$) have the highest SIC capability. However, as the SIC is independent of the design of precoder, Option 2 ($N \leq \bar{N}$) demands lower computational complexity than Option 1. As shown in Table 2.2, Option 2 ($N \geq \bar{N}$) also has low computational complexity, but it is unable to provide sufficient SI suppression due to the constraint on antenna arrays, as shown in Fig. 2.4. As for the method presented in [108], it can provide upto 30 dB SI reduction after about 5 iterations, but no further SI reduction is available, no matter how many iterations are executed. Hence, when taking into account the SIC capability and the required complexity, as shown in Table 2.2, Option 2 with $N \leq \bar{N}$ constitutes the most desirable SIC method.

2.5.2 Performance of MDD MIMO with Self-Interference Cancellation

We now demonstrate the achievable performance of the hybrid MDD MIMO systems with SIC. For this purpose, we consider a MDD MIMO system, where BS employs $N = 128$ transmit antennas and $N_{\mathrm{RF}} = 16$ DL RF chains to support $D = 6$ DL MSs, as well as $\bar{N} = 32$ receive antennas and $\bar{N}_{\mathrm{RF}} = 8$ UL RF chains to serve $\bar{D} = 4$ UL MSs. Unless explicitly noted, the transmit power of BS and MS is set to $P_{\mathrm{DL}} = 30$ dBm and $P_{\mathrm{MS}} = 20$ dBm, respectively [134, 136]. Furthermore, at BS, the transmit power is evenly assigned to the M DL subcarriers, while the transmit power of a DL subcarrier is assigned to the D number of DL MSs based on the water-filling principle. The total bandwidth is assumed to be 20 MHz and the number of DL and UL subcarriers are $M = 64$ and $\bar{M} = 32$, respectively. The MSs are assumed to be uniformly distributed

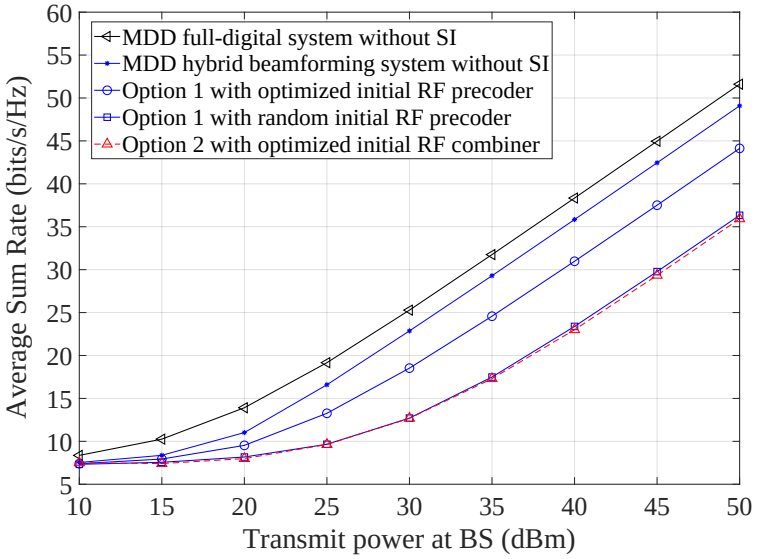


FIGURE 2.5: Average sum-rate of the MDD MIMO systems with different beamforming/initialization schemes, when $N = 128$, $N_{\text{RF}} = 16$, $\bar{N} = 32$, $\bar{N}_{\text{RF}} = 8$, $M = 64$, $\bar{M} = 32$, $D = 6$, $\bar{D} = 4$ and $L = 4$.

within a circular area of radius $R = 60$ m. Furthermore, the pathloss for a MS with distance d from BS is modeled as $P_L(\text{dB}) = 72 + 29.2 \log_{10}(d)$. Additionally, the power spectral density of noise is -173 dBm/Hz. In the following figures, the average sum rate denotes the total rate of a system, including both DL and UL, which is formulated as

$$R_{\text{MDD}} = \frac{1}{M_{\text{sum}}} \left(\sum_{m=1}^{M} \sum_{d=1}^{D} R_d[m] + \sum_{\bar{m}=1}^{\bar{M}} \sum_{\bar{d}=1}^{\bar{D}} R_{\bar{d}}[\bar{m}] \right). \tag{2.37}$$

Note that the comprehensive comparison of MDD with TDD/IBFD in mmWave environment can be found in [4]. More general comparison of MDD with FDD/TDD can be found in [137]. Readers interested in the more details about the comparison are referred to these references.

Fig. 2.5 demonstrates the sum-rate performance of the MDD MIMO systems with various beamforming schemes. In this study, the transmit power at BS varies from 10 dBm to 50 dBm, while the transmit power of individual MSs is 20 dBm. From the results of Fig. 2.5, we observe that the ideal full-digital MDD MIMO system without SI provides the upper-bound performance. Note that, here 'without SI' means that the required SIC can be achieved by the ideal SIC approaches. However, when hybrid beamforming assisted SIC is considered, as Fig. 2.5 shows, using different initial analog precoders under Option 1 yield a big impact on the achievable sum-rate. The optimized initial

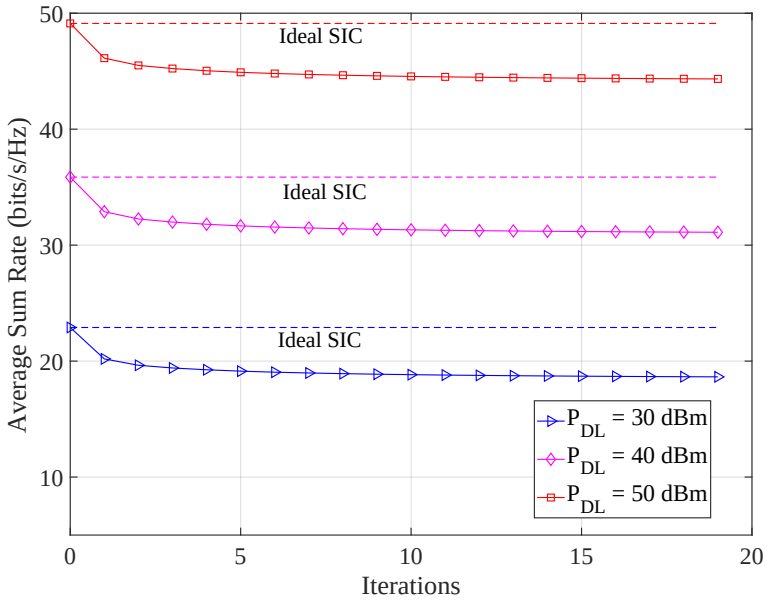


FIGURE 2.6: Average sum-rate versus the number of iterations of the analog precoder operated during the SIC process, when $N = 128$, $N_{\text{RF}} = 16$, $\bar{N} = 32$, $\bar{N}_{\text{RF}} = 8$, $M = 64$, $\bar{M} = 32$, $D = 6$, $\bar{D} = 4$ and $L = 4$.

analog precoder achieves much higher sum-rate than the random initial analog precoder. By contrast, when operated under Option 2, the optimized initial analog combiner can only achieve the similar performance of the Option 1 with random initial analog precoder. The reason behind is that in this study, we assumed $\bar{N} << N$, and in this case, as argued in Section 2.3, the performance of SIC should be dominated by the transmitter-based SIC, i.e., Option 1 design.

Fig. 2.6 shows the impact of the number of iterations used by the analog precoder during SIC on the average sum-rate achieved. The SIC is achieved by the analog precoder operated under Option 1, and the parameters are detailed with the figure. Note that the 'Ideal SIC' refers to that the remaining SI is ideally suppressed by the other SIC methods without using analog precoder. Therefore, as shown in Fig. 2.6, the highest sum-rate is observed before the analog precoder starts operating. When the analog precoder is operated with more iterations, the achieved sum-rate reduces, as the result that more degrees-of-freedom provided by transmit antennas are used for SIC. However, the achievable sum-rate becomes steady after only about 3 or 4 iterations, yielding the sum-rate cost for SIC. By contrast, as shown in Fig. 2.2, the amount of SI compressed monotonically increases with the increase of the number of iterations.

The impact of the number of DL transmit antennas N and the number of DL RF chains N_{RF} on the sum-rate performance of MDD MIMO systems is shown in Fig. 2.7, when assuming that SI is suppressed using the Option 1 method. Explicitly, for a given number of DL transmit antennas, the achievable sum-rate increases, as the number of RF chains increases, which is at the cost of the increase of implementation complexity. By

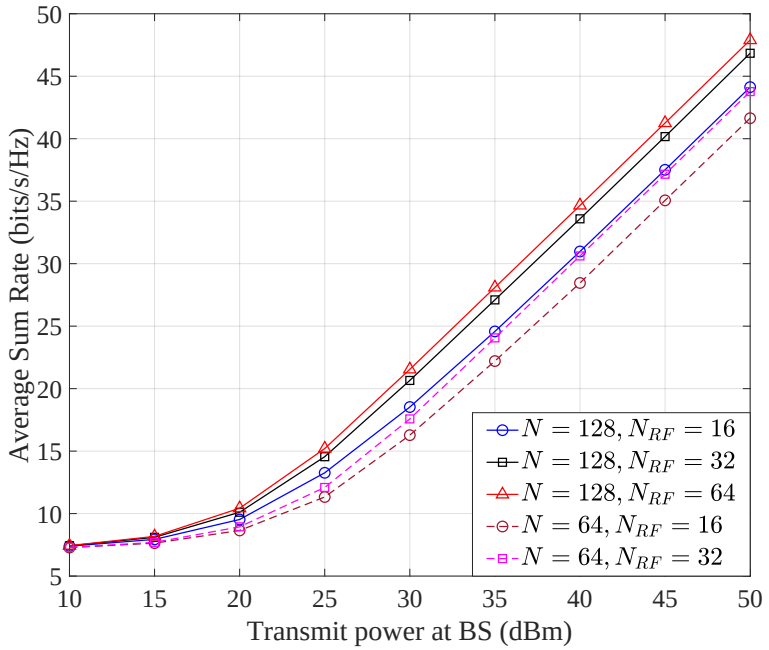


FIGURE 2.7: Average sum-rate versus BS's transmit power for the MDD MIMO systems employing different numbers of DL transmit antennas and RF chains.

contrast, when N_{RF} is fixed, the sum-rate increases, as N increases, owing to the improved SI capability.

Fig. 2.8 depicts the effect of the numbers of DL/UL subcarriers, i.e., M and $\bar{M}$, on the sum-rate performance of MDD MIMO systems, when the total number of subcarriers is set to $M + \bar{M} = 128$, and the other parameters are set as default values. As each UL MS has fixed transmit power, while the total BS transmit power is shared by all DL MSs, the total throughput of MDD MIMO system is dominated by the UL, when the transmit power of BS is relatively low. In this case, when the UL employs more subcarriers, the total throughput of MDD MIMO system is higher. By contrast, when BS's transmit power is sufficiently high, the system's throughput will become DL dominant. Correspondingly, employing more DL subcarriers provides higher throughtput by the MDD MIMO system.

By contrast, in Fig. 2.9, we investigate the effect of the numbers of DL/UL MSs, i.e., D and $\bar{D}$, on the sum-rate performance of MDD MIMO systems, while the other parameters are set to default values. As seen in Fig. 2.9, first, supporting more DL or/and UL MSs in general improves the total throughput of MDD MIMO systems. Second, when D is fixed, the total throughput of MDD MIMO systems increases, as the value of $\bar{D}$ increases. Finally, in the case that $\bar{D}$ is fixed, we observe that the total throughput achieved by $D = 12$ is slightly lower than that obtained by $D = 6$, when BS's transmit power is relatively low. However, when BS's transmit power is relatively high, the observation reverses, i.e., the total throughput achieved by $D = 12$ is higher than that obtained by $D = 6$. The reason behind this observation is that when at relatively low

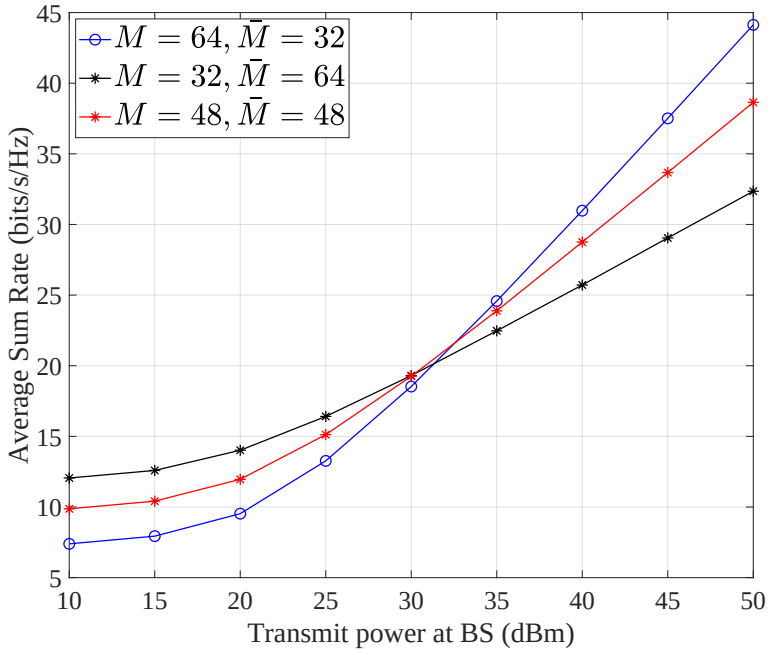


FIGURE 2.8: Average sum-rate versus BS's transmit power for the MDD MIMO systems employing different numbers of DL/UL subcarriers.

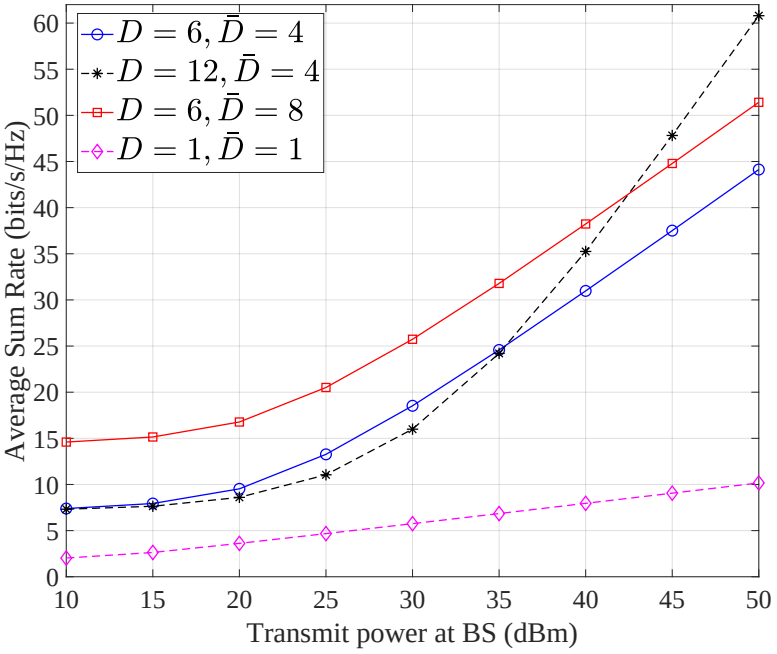


FIGURE 2.9: Average sum-rate versus BS's transmit power for the MDD MIMO systems employing different numbers of DL/UL users.

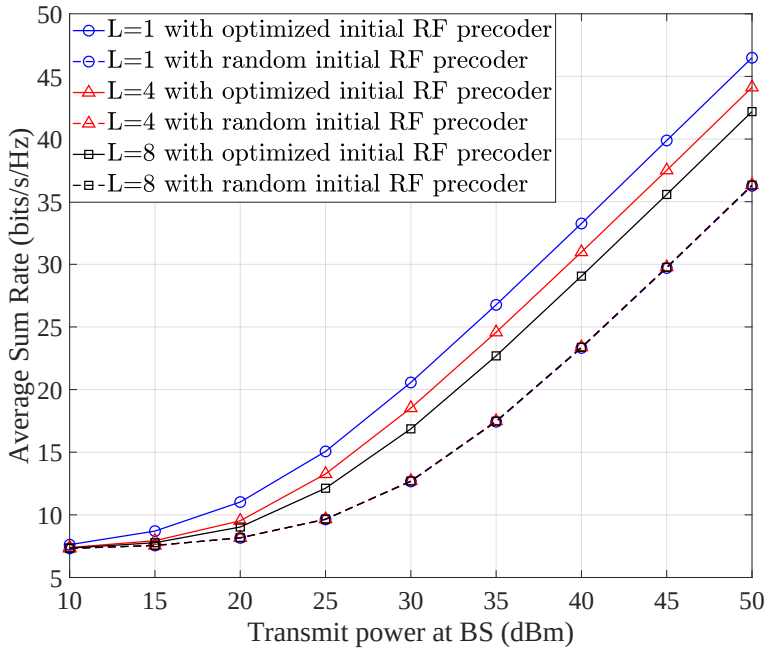


FIGURE 2.10: Average sum-rate performance of MDD MIMO systems experiencing the frequency-selective fading channels with different taps.

transmit power, the systems' total throughput is dominated by the UL MSs. By contrast, at relatively high transmit power, the systems' total throughput is dominated by the DL MSs, owing to the contribution from the joint PA among the DL MSs.

Fig. 2.10 shows the effect of the severity of frequency-selective fading channel on the achievable sum-rate performance of MDD MIMO systems, where increasing L means the channel becomes more frequency-selective. As shown in Fig. 2.10, when the optimized initial analog precoder is employed, the achieved sum-rate reduces with the increase of L. This is because when L increases, the fading experienced by different subcarriers becomes more random, making the optimized initial analog precoder appear more like the random initial analog precoder. By contrast, when the random initial analog precoder is employed, the achieved sum-rate is very similar, regardless of the values of L. Hence, while the random initial analog precoder usually achieves lower sum-rate than the optimized initial analog precoder, it is more robust to the time-varying of channels.

2.5.3 Performance of Channel Estimation

In order to investigate the performance of CE, we consider a MDD MIMO system where a BS employs a 128×32 antenna array, and the number of DL and UL subcarriers are $M = 128$ and $\bar{M} = 64$, respectively. Specifically in Fig. 2.11, we compare the MSE performance of CE in three cases. In the first case, the LS CE with orthogonality-achieving PSs, i.e., with the settings satisfying Proposition 1 is considered. In the second case,

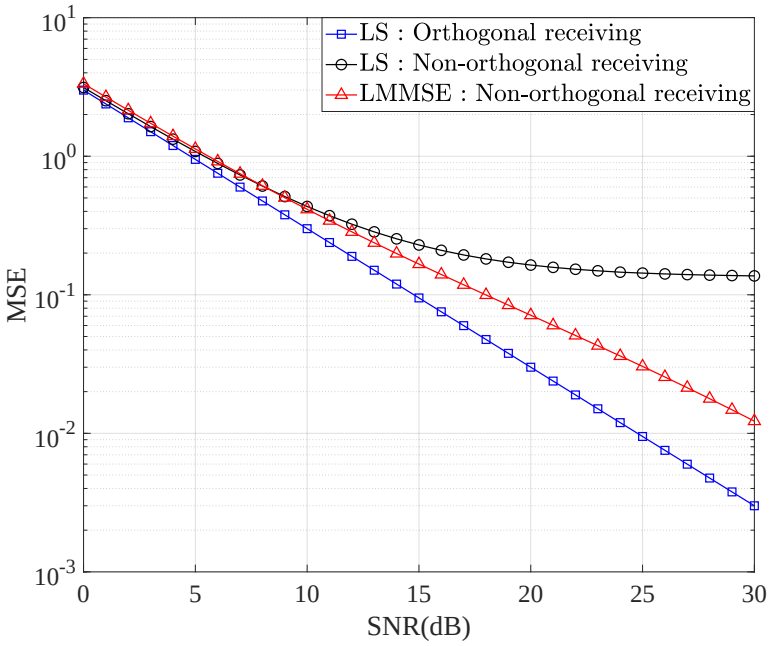


FIGURE 2.11: Mean-square error (MSE) performance of CE in MDD MIMO systems, when $N = 128$, $\bar{N} = 32$, $M = 128$, $\bar{M} = 64$ and $L = 4$, and for LS $D = 6$ and $\bar{D} = 4$, while for LMMSE $D = 12$ and $\bar{D} = 6$.

we also assume the LS CE but with the UL subcarriers randomly selected. Therefore, the orthogonality-achieving PSs cannot be guaranteed. Finally, in the case of LMMSE CE, the UL subcarriers are also randomly distributed. From Fig. 2.11 we observe that the LS CE with orthogonality-achieving PSs achieves better MSE performance than the other two CEsts, while the LS CE with randomly selected UL subcarriers achieves the worst MSE performance. Furthermore, the LS CE with random UL subcarriers experiences the interference, which is unable to be suppressed by the method, hence yielding MSE floor. By contrast, the LMMSE CE is capable of efficiently suppressing the interference and removing the MSE floor. We should note that, although the LS CE with orthogonality-achieving PSs outperforms the LMMSE CE in terms of MSE performance, it has the disadvantages of, such as, low number of orthogonality-achieving PSs due to the constraint of Proposition 1.

Finally, in Fig. 2.12 we compare the achievable sum-rate of MDD MIMO systems, when the channel knowledge is obtained by the LS CE with orthogonality-achieving PSs or the LMMSE CE with random UL subcarriers. Furthermore, the case of ideal CE is included as a benchmark. Explicitly, both CE schemes work efficiently over the SNR range considered. The sum-rate gap between that achieved by assuming ideal channel knowledge and that achieved by practical CE is marginal. When comparing the LS and LMMSE methods, we find that the LS CE allows to transmit an extra of 2.5 bits/s/Hz beyond the sum-rate achieved by the LMMSE CE. However, we should remember the LS CE with orthogonality-achieving PSs is limited by Proposition 1.

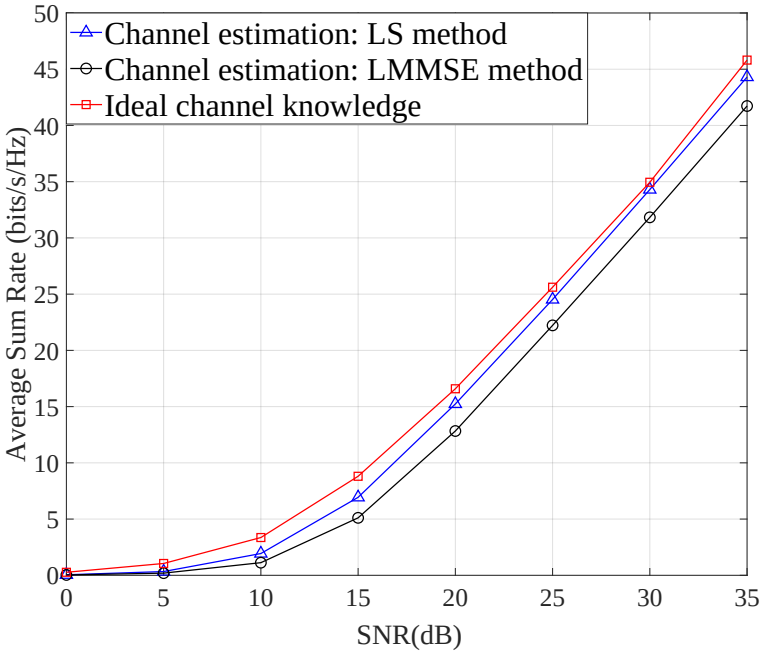


FIGURE 2.12: Average sum-rate versus SNR performance of MDD MIMO systems with channel knowledge provided by different methods. Parameters: $M = 128$, $\bar{M} = 64$, $L = 4$, $D = 8$ and $\bar{D} = 4$.

2.6 Chapter Summary and Conclusions

In this chapter, we have investigated the adaptive SIC methods based on the hybrid beamforming design of both transmitter and receiver as well as the CE in MDD MIMO systems. We have first highlighted the design of the hybrid transmitter precoder for DL transmission and the hybrid receiver combiner for UL signal detection, both of which can simultaneously suppress SI to a desired level. Our studies show that SI can be dynamically mitigated by employing either the analog transmitter precoding or the analog receiver combining. Furthermore, it is shown that SIC should be implemented at BS transmitter or at BS receiver depending on which of them has more antenna elements. Then, the CE in MDD MIMO systems has been designed by exploiting the reciprocity existing between UL/DL channels resulted from the correlated fading of subcarriers. Our studies reveal that first, when the number of MSs is relatively small and the PSs are evenly arranged across the subcarriers, the design of orthogonality-achieving PSs is available. In this case, a LS CE achieves optimum performance. However, when randomly distributed PSs have to be used due to, such as, a big number of MSs, the performance of LS CE degrades significantly. Instead, a LMMSE CE is near-optimum, and is capable of significantly improving the performance over the LS CE. Finally, the performance of the MDD MIMO systems with our proposed hybrid beamforming SIC method has been investigated, when CSI is provided by our proposed CE, showing that the achievable performance can be close to that with ideal CSI.

Chapter 3

Resource Allocation with Hybrid Beamforming Design in MDD MIMO

This chapter investigates the advantages of multicarrier-division duplex (MDD) over the conventional frequency-division duplex (FDD) and time-division duplex (TDD) in terms of resource allocation (RA). Furthermore, the RA with the advanced hybrid beamforming in MDD multi-input multi-output (MIMO) is studied.

3.1 Introduction

With the increasing demand of high capacity and SE for the next-generation wireless networks, two significant techniques have come into our vision, namely millimeter wave (mmWave) and in-band full duplex (IBFD) [72,138]. On one side, mmWave communication makes use of the unexploited spectrum resource with GHz of bandwidth to meet the requirements of high-rate applications. On the other side, owing to the downlink (DL) and uplink (UL) transmissions concurrently occurring on the same frequency band, IBFD has the potential to double the spectral efficiency that is achievable by the FDD and TDD modes [72]. However, as analyzed in Section 1.2.1 and Chapter 2, IBFD suffers from serious self-interference (SI) problem and significant system complexity, whereas MDD, exhibiting its natural advantage of SI suppression in digital domain aided by fast Fourier transform (FFT) at receiver, can be a transitional technique from half duplex (HD) to IBFD. In light of this, MDD systems require promising RA strategies to fully exploit the available time-frequency resources. Compared with HD systems, since DL and UL signals in MDD systems are transmitted on the different subcarriers of an orthogonal frequency division multiplexing (OFDM) block, there are

more degrees-of-free (DoF) for dynamically allocating subcarriers to DL and UL, which further enhances spectral efficiency (SE). However, similar to IBFD systems, when considering the influence of residual SI, the implementation of RA is also challenging in MDD systems.

In Chapter 2, we conduct beamforming from a self-interference cancellation (SIC) perspective, which pays more attention to the performance of SI mitigation. On the contrary, in order to fully take the advantages of MDD, this chapter will focus on the joint design of RA and hybrid bearmformers for SE maximization, with the consideration of quality of service (QoS) constraints as well as practical SI suppression. In the following, we will briefly overview some related works to introduce our motivations of studying RA in MDD MIMO mmWave systems.

As the increasing interest in MIMO systems, the joint design of beamformers and RA has become more and more attractive. In [139–141], the authors studied the channel and power allocation (PA) with digital beamforming and QoS constraints. However, the proposed approaches cannot be directly applied in the large-scale MIMO mmWave systems with hybrid beamforming. Furthermore, the algorithms with fairness constraint are unable to guarantee the proportional fairness exactly, leading to the performance fluctuation among mobile stations (MSs). The authors of [121] studied the RA to maximize the sum-rate of the mmWave orthogonal frequency division multiplexing access (OFDMA) systems with hybrid beamforming under the power constraint of BS, where PA, precoder design and subcarrier assignment are concurrently implemented, causing a high complexity. Different from [121], the authors of [142] jointly optimized the hybrid beamformers and RA in a distributed way to minimize the transmission latency in mobile edge computing mmWave networks. However, the QoS requirements of MSs were not considered in both these references. Furthermore, in [143], a proportional fairness constrained hybrid beamforming algorithm was proposed for the DL transmission in mmWave MU-MIMO systems, in which the constant modulus constraints on analog beamformers are ignored for simplicity. Furthermore, the proposed algorithm fails to achieve the expected fairness, when the inter-MS distances are large. We should note that in the RA of mmWave systems, the QoS requirements of MSs are important, as MSs suffer from large path-losses, in particular, when they are at the cell edges. To the best of our knowledge, there is no work in the open literature considered the RA in the hybrid beamforming-assisted systems with the constraints on individual MSs' QoS.

On the other hand, there are limited works having so far studied the RA in FD OFDMA systems. In [144], the authors considered to maximize the sum-rate of a multiuser single-input single-output (MU-SISO) IBFD system by jointly optimizing the subcarrier assignment and PA. In [145], the joint MS pairing and RA problem for MU-SISO IBFD system were investigated by proposing a low-complexity near-optimal matching

algorithm. Moreover, the authors of [146] developed an optimal iterative RA algorithm, aiming at minimizing the total power consumption via jointly optimizing DL beamformer, UL transmit power and antenna selection. In [147], the authors jointly optimized the UL transmit power and DL beamforming to minimize the long-term transmit power under delay constraints. Although all the above-mentioned works endeavor to exploit the resources provided by FD, the SI problem, which is regarded as the biggest obstacle to achieve high-efficiency FD communications, has not been assumed in a practical way. More specifically, the SI was directly assumed as Gaussian noise in [144], while in [145], SI was treated as a coefficient in the range of 70 dB to 110 dB. In [146] and [147], SI was assumed to be perfectly mitigated by a separate SIC subsystem before the beamforming design. To the authors' best knowledge, all the existing studies on the RA optimization in FD systems ignore the practical implementation of SIC, but assume a near-optimal SIC. This assumption may be feasible in a small-scale MIMO system equipped with a small number of antennas or radio frequency (RF) chains supported by the conventional SIC approaches proposed, such as, in [25,39,148]. However, these SIC methods may cause significant power consumption and complexity in large-scale mmWave MIMO systems. Hence, a feasible RA algorithm for FD systems should be designed by considering a practically efficient SIC method, which is capable of providing sufficient SI suppression at the lowest possible cost of system resources.

To fill the research gap, in this chapter, we propose and study the RA in MDD MIMO mmWave systems for not only leveraging the DL/UL resources but also relieving the impact of SI, via designing hybrid beamforming and SIC schemes, while simultaneously meeting the fairness constraints near-perfectly. Our contributions can be summarized as follows:

Firstly, considering a MU-SISO system with unfair greedy RA, we demonstrate the performance advantages of MDD against the HD modes of FDD and TDD.

Secondly, we study the RA in MDD MIMO mmWave systems. Since MDD enables system to jointly allocate resources among DL/UL, which leads to a NP-hard problem, we divide the joint UL/DL RA problem into the suboptimal subcarrier allocation (SA) and PA. Specifically for SA, an improved fair greedy (IFG) algorithm is proposed, which consists of three stages to concurrently achieve sum-rate maximization, coarse fairness and spatial multiplexing gain. After SA and applying the full-digital zero-forcing (ZF) precoding, PA becomes a convex solvable problem. Our results show that the proposed RA algorithms allow to attain the performance near the upper bound that is only achievable by the conventional unfair greedy algorithm. Simultaneously, our algorithms guarantee the near-accurate proportional fairness among all the DL/UL MSs, even when some MSs are located far away from BS.

Thirdly, we design two hybrid precoding schemes for the implementation of RA based on the matrix factorization and direct approach, respectively, under the fairness constraint. Our studies show that although the matrix factorization method performs well in the systems without considering RA [149], it cannot perform as well as the direct approach, when the number of RF chains is relatively low. It is shown that this is because the difference between a full digital precoder and the hybrid precoder designed from it by using matrix factorization is magnified after RA, which leads to the increase of inter-user interference.

Furthermore, we propose an adaptive SIC algorithm in conjunction with the design of hybrid combiner at BS receiver. To the best of our knowledge, this is the first attempt to jointly consider SIC and RA in FD systems. It is shown that our SIC algorithm is capable of providing the flexible levels of SI suppression over a big range and decoupled with the implementation of RA.

The remainder of this chapter is outlined as follows. In Section 3.2, we address the modeling in MDD MIMO mmWave hybrid beamforming system and introduce the formulation of RA problem. In Section 3.3, we briefly compare the MDD mode with HD modes in the MU-SISO system applying unfair greedy RA algorithm. In Section 3.4, suboptimal subcarrier and PA are presented to jointly consider the maximization of sum rate and fairness among DL and UL MSs. The two precoding strategies are proposed for RA scheme and an adaptive SIC method based on the combiner is presented in Section 3.5. Finally, several simulation results are presented in Section 3.6.

3.2 System Model

3.2.1 System Architecture

Consider an MDD MIMO mmWave system, where an $N_{\text{sum}}^{\text{BS}}$-element antenna BS communicates with D downlink single-antenna MSs and $\bar{D}$ uplink single-antenna MSs, where $D + \bar{D} = D_{\text{sum}}$. As shown in Fig. 3.1, we assume that BS implements the separate antenna configuration working in MDD mode, in which the Tx and Rx are equipped with N and $\bar{N}$ antenna elements, respectively. By contrast, MSs work in HD mode. We further assume that the transmission channels between BS and MSs are frequency selective. In order to deal with small-scale fading, OFDM modulation is applied with the cyclic prefix introduced to avoid inter-symbol interference. In our proposed MDD system, all MSs choose the subcarriers from a set $\mathcal{M}$, with the size of $|\mathcal{M}| = M$. Denoting $\alpha_{m,d}$ and $\alpha_{m,\bar{d}}$ the two indicators of subcarrier assignments, which take values as

$$\alpha_{m,d} \text{ (or } \alpha_{m,\bar{d}}) = \begin{cases} 1, \text{if } d \text{ (or } \bar{d}) \text{ assigned subcarrier } m, \\ 0, \text{otherwise}, \end{cases} \tag{3.1}$$

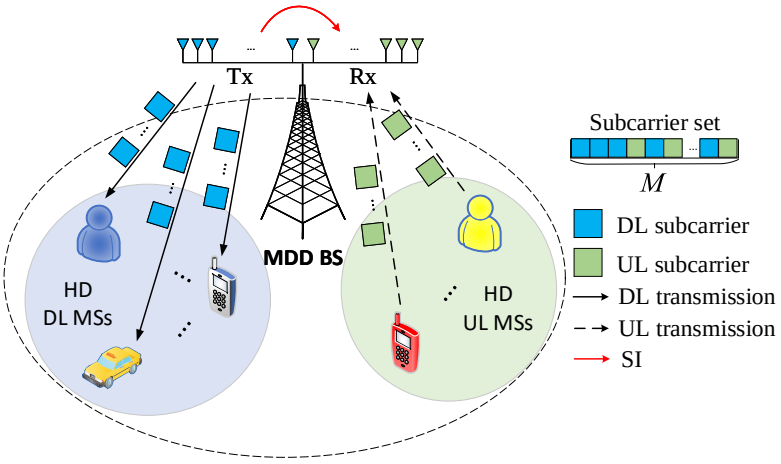


FIGURE 3.1: System model: MDD MIMO mmWave network.

where d and $\bar{d}$ are DL MS and UL MS, respectively. Note that a subcarrier can only be assigned either to DL MSs or to UL MSs, not to both simultaneously. After SA, we obtain two mutually exclusive subsets, namely downlink subset $\mathcal{M}_{\mathrm{DL}}$ and uplink subset $\mathcal{M}_{\mathrm{UL}}$. The size of the two sets are pre-defined considering the QoS requirements for both DL and UL MSs, and we have $|\mathcal{M}_{\mathrm{DL}}| \leq M_{\mathrm{DL}}, |\mathcal{M}_{\mathrm{UL}}| \leq M_{\mathrm{UL}}$ and $M_{\mathrm{DL}} + M_{\mathrm{UL}} \leq M$. Let the DL signal transmitted to d or the UL signal transmitted by $\bar{d}$ on the m-th subcarrier be denoted as $z_d[m] \in \mathbb{C}$ or $z_{\bar{d}}[m] \in \mathbb{C}$. They satisfy $\mathbb{E}\left[|z_d[m]|^2\right] = \mathbb{E}\left[|z_{\bar{d}}[m]|^2\right] = 1$ and $\mathbb{E}\left[z_i[m]z_j^*[m]\right] = 0, \forall\, i \neq j$.

Although DL and UL are operated on different subcarriers, MDD-based system still suffers from SI in analog domain at BS, as IBFD systems [4]. If SI is not sufficiently suppressed in analog domain, the BS transmit signal may overwhelm the ADC of the BS receiver and cause large unexpected quantization noise. Therefore, the effect of SI and its suppression will be considered in this chapter. However, the inter-user interference from the UL MSs on the DL MSs is ignored. This is because in general, MSs are distributed from each other with relatively big distances, which results in big SI reduction due to propagation pathloss [72][1]. Then, when all MSs and BS are assumed to be synchronized within an allowable limit, the interference among different subcarriers can be ideally canceled with the aid of the FFT operation in digital domain, as assumed in the references on IBFD OFDMA systems [144, 145].

[1]Note that, in the case where a UL MS is close to a DL MS, we can leverage the user scheduling operated at medium access control (MAC) layer to minimize the interference between these two MSs [113]. This is however out of the scope of this chapter.

3.2.2 Channel Model

3.2.2.1 Communication Channel Model

According to [122], we present a T-delay channel model combined with limited scattering characteristics of mmWave. As the antenna arrays at BS are separated, below we only introduce the wideband channel model between the BS transmit array and the d-th DL MS. The UL transmission channels have the same expression. The t-th delay tap of the DL channel, $\boldsymbol{g}_d[t] \in \mathbb{C}^N$, can be written as [122]

$$\boldsymbol{g}_d[t] = \beta \sum_{l=1}^{L_d} v_{l,d} p_{\mathrm{rc},d}(tT_s - \tau_{l,d}) \alpha_{\mathrm{MS},d}(\theta_{l,d}) \boldsymbol{\alpha}_{\mathrm{BS}}(\phi_{l,d}), \tag{3.2}$$

where L_d is the number of propagation paths, $\beta = \sqrt{N/(L_d P_L)}$, P_L, $v_{l,d}$, $\tau_{l,d}$, $\theta_{l,d}$, $\phi_{l,d}$ are the pathloss, complex gain, delay, angle of arrival (AoA) and angel of departure (AoD) of the l-th path, respectively, $1/T_s$ is the sampling frequency, $p_{\mathrm{rc},d}(\tau)$ is the raised cosine (RC) pulse evaluated at the time of τ. Finally, $\alpha_{\mathrm{MS},d}(\theta_{l,d})$ and $\boldsymbol{\alpha}_{\mathrm{BS}}(\phi_{l,d})$ are the array response vectors of the receiver and the transmitter, respectively. Correspondingly, the m-th subcarrier channel can be expressed as

$$\boldsymbol{h}_d[m] = \sum_{t=1}^{T} \boldsymbol{g}_d[t] e^{-j2\pi t \frac{m}{M}}, m = 1, 2, ..., M. \tag{3.3}$$

3.2.2.2 Self-Interference Channel Model

Here, we have the similar assumption of SI channel (i.e., $\boldsymbol{H}_{\mathrm{SI}}$) with (2.5), and the $\boldsymbol{H}_{\mathrm{LoS}}$ can also be found in (2.6). Moreover, since the distance between the BS transmit array and the BS receive array is small, the NLoS components can be well modeled to experience flat fading, having an expression in the form of

$$\boldsymbol{H}_{\mathrm{SI}}^{\mathrm{NLoS}} = \beta_{\mathrm{SI}} \sum_{l=1}^{L_n} v_l \boldsymbol{\alpha}_{\mathrm{R}}(\theta_l) \boldsymbol{\alpha}_{\mathrm{T}}^{\mathrm{H}}(\phi_l), \tag{3.4}$$

where L_n is the number of propagation paths between the BS transmit array and the BS receive array, while $\beta_{\mathrm{SI}} = \sqrt{N\bar{N}/L_n}$. Additionally, v_l, $\boldsymbol{\alpha}_{\mathrm{R}}(\theta_l)$ and $\boldsymbol{\alpha}_{\mathrm{T}}^{\mathrm{H}}(\phi_l)$ have the similar definitions as that in (3.2).

3.2.3 Downlink Transmission

For DL transmission, the signal $z_d[m]$ is first processed by a digital precoder $\boldsymbol{f}_{\mathrm{BB}}^d[m] \in \mathbb{C}^{N_{\mathrm{RF}}}$, where N_{RF} denotes the number of RF chains at BS transmitter. Then, after DAC operation and RF processing, the time-domain signal is further processed by an analog

precoder $\boldsymbol{F}_{\mathrm{RF}} \in \mathbb{C}^{N \times N_{\mathrm{RF}}}$. As analog beamformers are implemented using unit-modulus phase shifters, each element in $\boldsymbol{F}_{\mathrm{RF}}$ should meet the constraint $\left|(\boldsymbol{F}_{\mathrm{RF}})_{i,j}\right| = 1$. Finally, the signal $\boldsymbol{x}_d[m]$ transmitted on the m-th subcarrier by the BS transmit antennas can be formulated as

$$\boldsymbol{x}_d^{\mathrm{DL}}[m] = \alpha_{m,d}\sqrt{p_{m,d}}\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}_{\mathrm{BB}}^{d}[m]z_d[m],\ m = 1, ..., M, \tag{3.5}$$

where $p_{m,d}$ denotes the transmit power for MS d over subcarrier m, which satisfies $\sum_{d=1}^{D}\sum_{m=1}^{M} p_{m,d} \leq P_{\mathrm{DL}}$ with P_{DL} denoting the total power of DL transmission. Moreover, we normalize $\|\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}_d[m]\|_2^2 = 1$ to ensure that $\left\|\boldsymbol{x}_d^{\mathrm{DL}}[m]\right\|_2^2 = \alpha_{m,d}p_{m,d}$. Given the m-th subchannel vector $\boldsymbol{h}_d[m] \in \mathbb{C}^N$ to the d-th MS, the received signal at MS d is given by

$$y_d^{\mathrm{DL}}[m] = \boldsymbol{h}_d^H[m]\sum_{d=1}^{D}\boldsymbol{x}_d^{\mathrm{DL}}[m] + n_d[m], \tag{3.6}$$

where $n_d[m] \sim \mathcal{CN}(0, \sigma^2)$ is the complex Gaussian noise added to the m-th subcarrier, which has the power of σ^2.

It is assumed that coherent detection is employed at MSs. The SINR of the m-th subcarrier at the d-th DL MS can be obtained from (3.6) and is given by

$$\mathrm{SINR}_{m,d} = \frac{\alpha_{m,d}p_{m,d}\left|\boldsymbol{h}_d^H[m]\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}_{\mathrm{BB}}^{d}[m]\right|^2}{\sum_{i\neq d}^{D}\alpha_{m,i}p_{m,i}\left|\boldsymbol{h}_d^H[m]\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}_{\mathrm{BB}}^{i}[m]\right|^2 + \sigma^2}. \tag{3.7}$$

Furthermore, when Gaussian distributed transmit signals are assumed, the achievable sum-rate of the d-th DL MS is given by $R_d = \sum_{m=1}^{M}\log(1 + \mathrm{SINR}_{m,d})$.

3.2.4 Uplink Transmission

For UL transmission, the multiple-access signal received from the m-th subcarrier by BS receiver can be expressed as

$$\boldsymbol{y}_{\mathrm{UL}}[m] = \underbrace{\sum_{\bar{d}=1}^{\bar{D}}\boldsymbol{x}_{\bar{d}}^{\mathrm{UL}}[m]}_{\text{Desired signal}} + \underbrace{\boldsymbol{H}_{\mathrm{SI}}\boldsymbol{x}_{\mathrm{DL}}}_{\mathrm{SI}} + \boldsymbol{n}[m], \tag{3.8}$$

where $\boldsymbol{x}_{\mathrm{DL}} = \sum_{d=1}^{D}\sum_{m=1}^{M}\boldsymbol{x}_d^{\mathrm{DL}}[m]$ generates SI, the desired UL signal from MS $\bar{d}$ is $\boldsymbol{x}_{\bar{d}}^{\mathrm{UL}}[m] = \sqrt{p_{m,\bar{d}}}\alpha_{m,\bar{d}}\boldsymbol{h}_{\bar{d}}[m]z_{\bar{d}}[m]$, where $p_{m,\bar{d}}$ denotes the transmit power on subcarrier m of the $\bar{d}$-th UL MS. The transmit power of a UL MS satisfies the constraint of $\sum_{m=1}^{M}\alpha_{m,\bar{d}}p_{m,\bar{d}} \leq P_{\bar{d}}$. Finally, $\boldsymbol{n}[m] \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2\boldsymbol{I}_{\bar{N}})$ is the Gaussian noise.

At the BS receiver, the received signal of (3.8) is first processed by an analog combiner $\boldsymbol{W}_{\mathrm{RF}} \in \mathbb{C}^{\bar{N}\times\bar{N}_{\mathrm{RF}}}$. As shown in (3.8), the received signal includes both SI and desired UL signal. In theory, the SI can be removed in digital domain by FFT operation, as UL and DL signals are transmitted on different subcarriers. However, the power of SI may be much larger than the power of desired signal. Hence, SI may overwhelm the limited dynamic range of the ADC at BS receiver, introducing significant quantization noise. Therefore, after the transmitter preprocessing seen in (3.5) and the receiver combining using $\boldsymbol{W}_{\mathrm{RF}}$, SI should be sufficiently suppressed to make the received desired signal pass the ADC with high efficiency. As done in [148], we model the residual SI[2], by a vector $\boldsymbol{r}_{\mathrm{SI}} \sim \mathcal{CN}\left(\boldsymbol{0}, \xi \mathrm{diag}\left(\mathbb{E}\left[\boldsymbol{W}_{\mathrm{RF}}^{H}\boldsymbol{H}_{\mathrm{SI}}\boldsymbol{x}_{\mathrm{DL}}\boldsymbol{x}_{\mathrm{DL}}^{H}\boldsymbol{H}_{\mathrm{SI}}^{H}\boldsymbol{W}_{\mathrm{RF}}\right]\right)\right)$, where $0 < \xi \leq 1$ denotes the SIC capability provided by the SIC techniques implemented in propagation- and analog-domain, such as those considered in [39, 72, 150]. Consequently, after ADC conversion, FFT operation and digital combining, the final signal for detecting the information of MS $\bar{d}$ can be expressed as

$$y_{\bar{d}}[m] = \boldsymbol{w}_{\mathrm{BB}}^{\bar{d},H}[m]\boldsymbol{W}_{\mathrm{RF}}^{H}\sum_{\bar{d}=1}^{\bar{D}}\boldsymbol{x}_{\bar{d}}^{\mathrm{UL}}[m] + \boldsymbol{w}_{\mathrm{BB}}^{\bar{d},H}[m]\boldsymbol{W}_{\mathrm{RF}}^{H}\boldsymbol{n}[m] + \boldsymbol{w}_{\mathrm{BB}}^{\bar{d},H}[m]\boldsymbol{r}_{\mathrm{SI}}, \tag{3.9}$$

where $\boldsymbol{w}_{\mathrm{BB}}^{\bar{d}}[m] \in \mathbb{C}^{\bar{N}_{\mathrm{RF}}}$ is the digital combiner for detecting the mth subcarrier of the $\bar{d}$-th UL MS. Note furthermore that $\boldsymbol{W}_{\mathrm{RF}}$ satisfies $\left|(\boldsymbol{W}_{\mathrm{RF}})_{i,j}\right| = 1$, as it operates the analog beamforming.

Similar to the DL detection, the SINR achieved by the mth subcarrier of the $\bar{d}$-th UL MS is given by

$$\mathrm{SINR}_{m,\bar{d}} = \frac{\alpha_{m,\bar{d}}p_{m,\bar{d}}\left|\boldsymbol{w}_{\mathrm{BB}}^{\bar{d},H}[m]\boldsymbol{W}_{\mathrm{RF}}^{H}\boldsymbol{h}_{\bar{d}}[m]\right|^{2}}{\sum\limits_{j\neq\bar{d}}^{\bar{D}}\alpha_{m,j}p_{m,j}\left|\boldsymbol{w}_{\mathrm{BB}}^{\bar{d},H}[m]\boldsymbol{W}_{\mathrm{RF}}^{H}\boldsymbol{h}_{j}[m]\right|^{2} + \sigma^{2}\left\|\boldsymbol{w}_{\mathrm{BB}}^{\bar{d},H}[m]\boldsymbol{W}_{\mathrm{RF}}^{H}\right\|_{2}^{2} + \left|\boldsymbol{w}_{\mathrm{BB}}^{\bar{d},H}[m]\boldsymbol{r}_{\mathrm{SI}}\right|_{2}^{2}}. \tag{3.10}$$

Correspondingly, the data-rate achieved by the $\bar{d}$-th UL MS is given by $\bar{R}_{\bar{d}} = \sum\limits_{m=1}^{M}\log(1+\mathrm{SINR}_{m,\bar{d}})$.

[2]Note that the residual SI consists of the combined effect of the additive noise introduced by automatic gain control (AGC), non-linearity of ADC and the phase noise generated by oscillator due to RF imperfection [148].

3.2.5 Problem Formulation

To design an optimum adaptive RA scheme achieving proportional data rate constraints in MDD MIMO system with hybrid beamforming, we have to solve a non-convex problem described as

$$\max_{\alpha_{m,d},\alpha_{m,\bar{d}},\boldsymbol{F}_{\mathrm{RF}},\boldsymbol{W}_{\mathrm{RF}},\boldsymbol{f}_{\mathrm{BB}}^{d}[m],\boldsymbol{w}_{\mathrm{BB}}^{\bar{d}}[m],p_{m,d},p_{m,\bar{d}}} \quad \Lambda \tag{3.11a}$$

$$\text{s.t.} \quad \sum_{d=1}^{D}\sum_{m=1}^{M}\alpha_{m,d}p_{m,d} \le P_{\mathrm{DL}}, \tag{3.11b}$$

$$\sum_{m=1}^{M}\alpha_{m,\bar{d}}p_{m,\bar{d}} \le P_{\bar{d}}, \ \ \forall \bar{d}, \tag{3.11c}$$

$$\alpha_{m,d} + \alpha_{m,\bar{d}} \le 1, \ \ \forall m, d, \bar{d}, \tag{3.11d}$$

$$\sum_{d=1}^{D}\alpha_{m,d} \le N_{\mathrm{RF}}, \ \ \forall m, \tag{3.11e}$$

$$\sum_{\bar{d}=1}^{\bar{D}}\alpha_{m,\bar{d}} \le \bar{N}_{\mathrm{RF}}, \ \ \forall m, \tag{3.11f}$$

$$R_1 : R_2 : ... : R_D = \gamma_1 : \gamma_2 : ... : \gamma_D, \tag{3.11g}$$

$$\bar{R}_1 : \bar{R}_2 : ... : \bar{R}_{\bar{D}} = \eta_1 : \eta_2 : ... : \eta_{\bar{D}}, \tag{3.11h}$$

$$\left\|\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}_{\mathrm{BB}}^{d}[m]\right\|_2^2 = 1, \ \forall m, d, \tag{3.11i}$$

$$\left|(\boldsymbol{F}_{\mathrm{RF}})_{i,j}\right| = \left|(\boldsymbol{W}_{\mathrm{RF}})_{i,j}\right| = 1, \ \ \forall i, j, \tag{3.11j}$$

$$|\mathcal{M}_{\mathrm{DL}}| \le M_{\mathrm{DL}}, |\mathcal{M}_{\mathrm{UL}}| \le M_{\mathrm{UL}}, \tag{3.11k}$$

where $\Lambda = \sum_{\forall m,d,\bar{d}} \log_2\left(1+\mathrm{SINR}_{m,d}\right) + \log_2\left(1+\mathrm{SINR}_{m,\bar{d}}\right)$, (3.11b) and (3.11c) impose the constraint on the transmit power at BS and individual UL MSs, respectively, (3.11d) explains that one subcarrier can only be assigned either to DL or to UL, (3.11e) and (3.11f) state that the number of MSs allocated to the same subcarrier does not exceed the number of RF chains of the DL and UL, respectively, $\{\gamma_i\}_{i=1}^{D}$ in (3.11g) and $\{\eta_i\}_{i=1}^{\bar{D}}$ in (3.11h) are given to ensure the proportional fairness among DL and UL MSs[3]. Finally, (3.11k) means that the number of DL and UL subcarriers have their specific constraints.

3.3 Potential of MDD with Resource Allocation

Before attempting to solve the optimization problem (3.11), let us first demonstrate the potential of MDD over the HD modes of FDD and TDD, when the unfair greedy (UG)

[3]Note that the exact proportional fairness shown in (3.11h) cannot be guaranteed among UL MSs in principle. This is because UL MSs work in the distributed manner and carry out PA separately. Therefore, the proportional fairness of UL can only be implemented via SA.

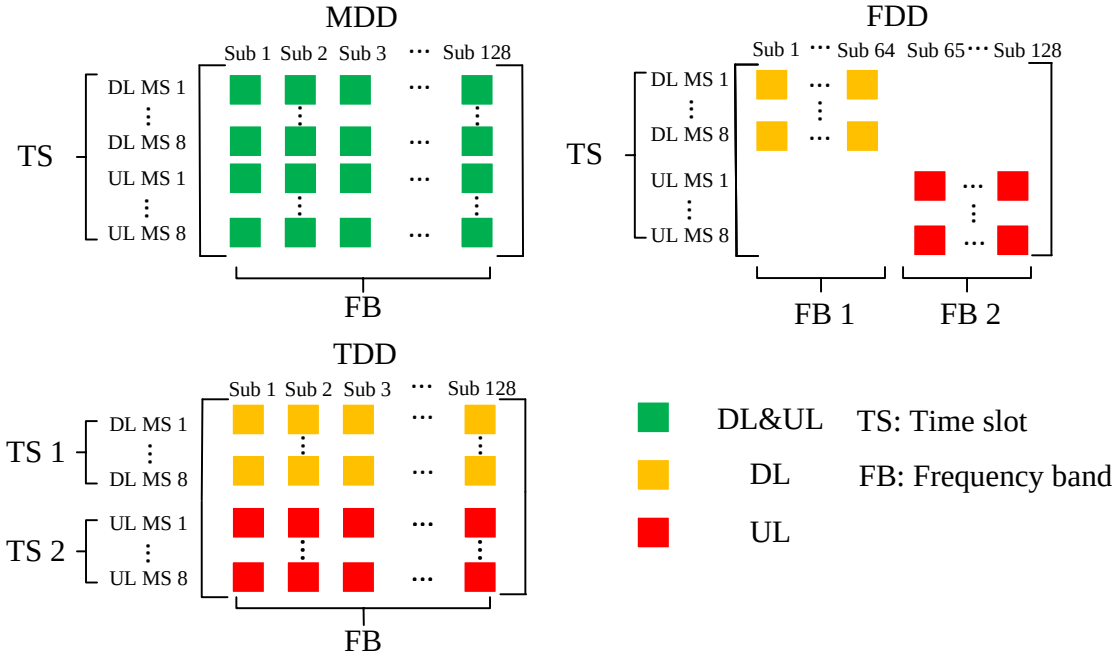


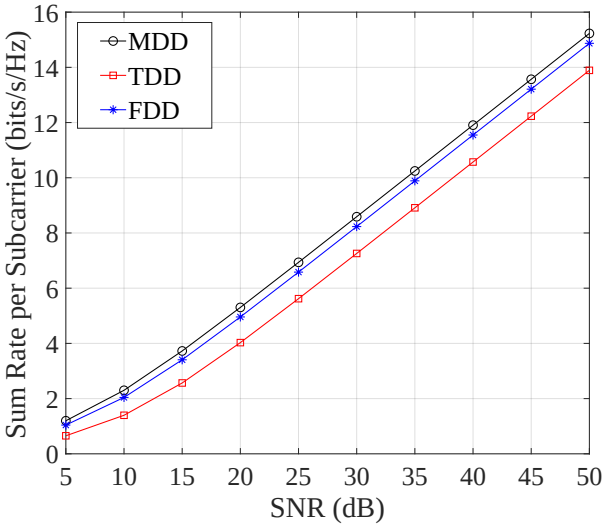
FIGURE 3.2: Illustration of subcarrier resources in the MDD, TDD and FDD modes, where green squares denote the subcarriers available for both DL and UL allocations, yellow squares denote the subcarriers only for DL allocations, while red squares represent the subcarriers only for UL allocations.

RA algorithm is employed[4]. To make the comparison straightforward, here we assume a MU-SISO system, where both BS and MS terminals employ single antenna, and the channel models of (3.2) with $P_L = 1$. We assume that the transmit power of BS is $P_{\text{DL}} = D$, and that of each UL MS is $P_{\vec{d}} = 1$. The noise variance for both DL and UL is $\sigma^2 = 1/\text{SNR}$. Furthermore, as we compare the upper-bound of the three schemes, we assume that there is no SI in MDD system or in other words, ideal SIC is implemented.

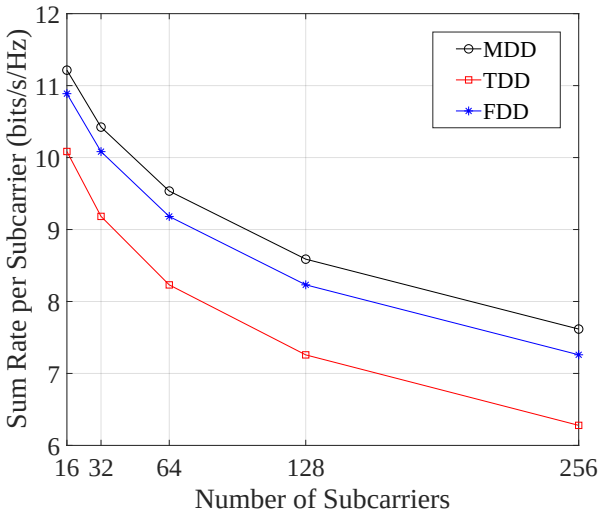
Fig. 3.2 illustrates the subcarriers available for UL/DL allocation, when MDD, FDD and TDD are respectively employed. We can observe that as the UL/DL in MDD mode are operated within the same frequency band at the same time, a subcarrier can be assigned either to DL or to UL. By contrast, in FDD mode (or TDD mode), DL and UL MSs are constrained on different frequency bands (or time slots). Hence, a subcarrier is pre-defined either for DL or for UL, and cannot be allocated by jointly considering both the UL and DL MSs.

To compare these modes, we assume that both DL and UL are assigned an equal resource in TDD and FDD modes, in terms of the number of time slots and subcarriers, respectively. To match this, when MDD mode is considered, we assume that DL and UL have the same number of subcarriers. Hence, the spectral efficiency of the MDD-,

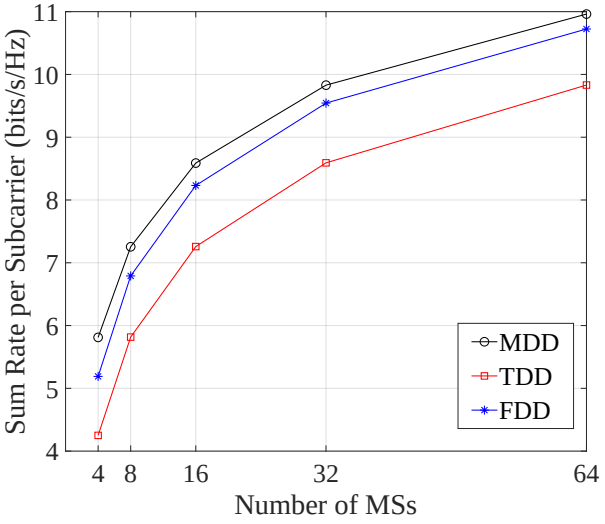
[4]To demonstrate the upper bound performance, the UG RA algorithm is employed, which maximizes the sum-rate of OFDM systems [151].



(A) Sum rate vs. SNR



(B) Sum rate vs. number of subcarriers



(C) Sum rate vs. number of MSs

FIGURE 3.3: Sum rate achieved by the MDD-, FDD- and TDD-based MU-SISO systems. The experimental parameters in three subfigures are: (a) $M = 128$, $D = 8$ and $\bar{D} = 8$; (b) SNR= 30 dB, $D = 8$ and $\bar{D} = 8$; (c) SNR= 30, $M = 128$ and $D = \bar{D}$

FDD- and TDD-based MU-SISO systems can be expressed as

$$\begin{aligned} R_{\mathrm{MDD}} &= \frac{\alpha_{\mathrm{MDD}}}{M} \sum_{\tilde{m}=1}^{M} R_{\tilde{m}}, \\ R_{\mathrm{FDD}} &= \frac{\alpha_{\mathrm{FDD}}}{M} \left(\sum_{m=1}^{M/2} R_m + \sum_{\bar{m}=M/2}^{M} R_{\bar{m}} \right), \\ R_{\mathrm{TDD}} &= \frac{\alpha_{\mathrm{TDD}}}{M} \left(\sum_{m=1}^{M} R_m + \sum_{\bar{m}=1}^{M} R_{\bar{m}} \right), \end{aligned} \tag{3.12}$$

where m and $\bar{m}$ denote the corresponding subcarriers assigned to DL and UL, respectively, $\tilde{m} \in \{m, \bar{m}\}$. Note that, as both DL and UL are always active in MDD and FDD modes while in TDD mode, DL and UL become active alternatively, we have $\alpha_{\mathrm{MDD}} = \alpha_{\mathrm{FDD}} = 2\alpha_{\mathrm{TDD}} = 1$. In our simulations for Fig. 3.3, we set $M = 128, D = \bar{D} = 8$.

From Fig. 3.3 we can observe that MDD is capable of outperforming both TDD and FDD. The reason behind is that, in MDD mode, each subcarrier has the largest spatial-domain degree-of-freedoms (DoFs) (i.e., 16) to be allocated either to a DL or a UL MS. By contrast, in TDD or FDD mode, as the 8 UL MSs and the 8 DL MSs cannot jointly share the resources, each subcarrier has only 8 DoFs to be allocated. It is noteworthy that the performance advantage of MDD is attained under the assumption of the shared-antenna configuration, which may be degraded in the separate-antenna configuration resulted from the loss of antenna gain, when MDD and TDD/FDD modes assume the same number of antenna elements. Nevertheless, the studies show that the separate-antenna configuration is more desirable in FD systems, since it has the potential to suppress SI at lower cost than the shared-antenna configuration [23]. On the other side, when the large antenna arrays satisfying ($N + \bar{N} \gg D_{\mathrm{sum}}$) are considered, we have the massive MIMO structure that is envisioned for future wireless systems. In this scenario, the loss of antenna gain and spatial multiplexing gain in the MDD systems resulted from separate-antenna configuration can be negligible [152].

In addition, we should also note that the implementation of RA at BS requires near accurate CSI and it is usually a knotty challenge in high-mobility scenarios. For instance, when a car moving at a speed of 36 km/h communicates within 30 GHz mmWave band, the normalized Doppler spread is 0.07 (when assuming 15 kHz inter-subcarrier spacing). Consequently, channels vary fast from symbol to symbol, according to the Jakes autocorrelation model [153], which may lead to the outdated CSI for DL signal preprocessing after the UL training, when TDD systems are considered. To mitigate this problem, extra training pilots have to be added to the phase of data transmissions. However, for a given coherence period, increasing UL pilots not only directly results in the decrease of UL data rate, but also the degradation of DL data rate, as UL training has to be executed more frequently. As for FDD, it is well-known that it is infeasible for operation in massive MIMO systems, due to the extra complexity of CSI acquisition

and the extreme overhead introduced by DL training and the CSI feedback from MSs to BS, not to mention the channel aging resulted from CSI estimation and feedback.

By contrast, in MDD-based systems, both DL and UL are operated within the same time slot and also the same frequency band. Hence, after the normal (initial) UL training, BS can concurrently update the CSI with the aid of either the pilots sent along with the UL data or the decision-directed CE using the detected UL data [154]. Therefore, MDD systems are capable of implementing the DL transmission based on the CSI of nearly real-time. In other words, MDD can effectively avoid the channel aging problem. Note that above is just a brief comparison between MDD and TDD/FDD, as it is beyond the scope of this chapter. A comprehensive comparison of MDD and TDD in fast fading environments will be provided in Chapter 4.

3.4 Subcarrier and Power Allocation

In order to approximately solve the non-convex and NP-hard problem described in (3.11), we seek the suboptimal methods to divide the optimization into SA and PA. To be more specific, the variables $\{\alpha_{m,d}, \alpha_{m,\bar{d}}\}$ in (3.11a) are first obtained from the SA based on an IFG algorithm, followed by the PA, for which the values of $p_{m,d}$ are calculated by solving a convex optimization problem while that of $p_{m,\bar{d}}$ are obtained by the water-filling algorithm. Note that in this section, we assume that the system is nearly free from SI (i.e., $\xi \ll 1$), as done in the other IBFD RA systems [145, 146]. However, different from these references, in this chapter, we will propose an adaptive beamforming based SIC method in Section 3.5, which guarantees that SI has little impact on the performance of the RA in this section[5].

3.4.1 Subcarrier Allocation

Let $\mathcal{M}_d$ and $\mathcal{M}_{\bar{d}}$ denote the sets of subcarriers assigned to the d-th DL MS and $\bar{d}$-th UL MS, respectively. Hence, we have $\mathcal{M}_{\mathrm{DL}} = \mathcal{M}_1 \cup ... \cup \mathcal{M}_d ... \cup \mathcal{M}_D$ and $\mathcal{M}_{\mathrm{UL}} = \mathcal{M}_1 \cup ... \cup \mathcal{M}_{\bar{d}} ... \cup \mathcal{M}_{\bar{D}}$. Let $\mathcal{D}_m$ (or $\bar{\mathcal{D}}_m$) denotes the set of DL MSs (or UL MSs), which are allocated the m-th subcarrier. $\mathcal{D}_m$ (or $\bar{\mathcal{D}}_m$) satisfies $|\mathcal{D}_m| \leq N_{\mathrm{RF}}$ (or $|\bar{\mathcal{D}}_m| \leq \bar{N}_{\mathrm{RF}}$). Note again that the same subcarrier m can only be assigned either to DL or to UL, not to both simultaneously. During the SA stage, we assume that the maximal-ratio transmission (MRT) assisted precoding is employed by BS for its computational simplicity [152]. Then, the SNR for the DL and UL MSs can be expressed as $\mathrm{SNR}_{m,d} = \frac{\|\boldsymbol{h}_d[m]\|_2^2}{\sigma^2}$ and $\mathrm{SNR}_{m,\bar{d}} = \frac{\|\boldsymbol{h}_{\bar{d}}[m]\|_2^2}{\sigma^2}$, respectively. Correspondingly, the achievable rates of the DL MS

[5]To the best of our knowledge, there are no SIC methods in hybrid beamforming systems, which can provide more than 100 dB of SIC at low complexity and low power consumption. Furthermore, we found that there are some beamforming based SIC approaches, such as that studied in [30, 108], which are coupled with RA. With these approaches, the results of RA change during the process of SI suppression.

d and UL MS $\bar{d}$ can be temporarily expressed as $\tilde{R}_d = \sum_{m\in\mathcal{M}_d} \log(1+\mathrm{SNR}_{m,d})$ and $\tilde{R}_{\bar{d}} = \sum_{m\in\mathcal{M}_{\bar{d}}} \log(1+\mathrm{SNR}_{m,\bar{d}})$.

Then, the IFG SA algorithm is operated, which enables MSs to choose the best possible subchannels and exploit the space degree-of-freedom in MIMO systems while under the constraint of fairness. The proposed IFG SA algorithm is stated as Algorithm 3, which is divided into three stages. In view of the practice that DL transmission usually needs more resources than UL transmission, the algorithm commences allocating subcarriers to DL MSs during all of the three stages. Specifically, during the first stage one subcarrier is initially assigned to each MS by employing the UG SA algorithm. During the second stage, subcarriers are allocated with considering the fairness until the constraint (3.11k) is satisfied. Finally, the third stage leverages the spatial diversity provided by the MIMO system to accomplish the rest allocation. The details of the three stages of Algorithm 3 are stated as follows:

- During the first stage, the channel gain matrix, $\boldsymbol{G} = [\boldsymbol{G}_1; \boldsymbol{G}_2]$, is initialized to $g_{ij} = \|\boldsymbol{h}_i[j]\|_2$, where $\boldsymbol{G}_1 \in \mathbb{C}^{D\times M}$, $\boldsymbol{G}_2 \in \mathbb{C}^{\bar{D}\times M}$. The candidate subcarrier set is initialized to $\mathcal{T}_{\mathrm{DL}} = \mathcal{M}$ for DL and to $\mathcal{T}_{\mathrm{UL}} = \mathcal{M}$ for UL. Furthermore, the iteration index is set to $k = 1$. During the k-th iteration, the largest element $(\boldsymbol{G})_{d,m}$ of the remaining candidates of $\boldsymbol{G}$ is identified. Depending on MS d being a DL MS or a UL MS, there are two options to update the related sets, as shown in line 9 and line 15. Note that, since each MS is only assigned one subcarrier during the first stage, the d-th row is removed from $\boldsymbol{G}$ after the d-th MS is assigned a subcarrier during the k-th iteration. Simultaneously, under the constraint of (3.11d) that a subcarrier assigned to DL (or UL) can no longer be assigned to UL (or DL), the m-th column is also removed from $\boldsymbol{G}_2$ (or $\boldsymbol{G}_1$), after the m-th subcarrier is assigned during the kth iteration. Furthermore, to meet the constraints of (3.11e) and (3.11f), if a subcarrier has already been assigned to N_{RF} DL MSs (or $\bar{N}_{\mathrm{RF}}$ UL MSs), the subcarrier and its corresponding column are removed from both $\mathcal{T}_{\mathrm{DL}}$ (or $\mathcal{T}_{\mathrm{UL}}$) and $\boldsymbol{G}_1$ (or $\boldsymbol{G}_2$).

- During the second stage, subcarriers are allocated with the consideration of fairness. For this purpose, during an allocation iteration, the d-th DL MS with the lowest ratio of $\tilde{R}_d/\gamma_d$ first chooses the best subcarrier m from the candidate set $\mathcal{T}_{\mathrm{DL}}$, as described in line 21. Then, $\tilde{R}_d$ is updated before going to the next iteration. This procedure is repeated until the condition $|\mathcal{M}_{\mathrm{DL}}| = M_{\mathrm{DL}}$ is satisfied. Analogously, UL allocation can be executed.

- After the second stage of allocation, all MSs (including both DL and UL MSs) have obtained their subcarriers. However, the conditions of (3.11e) and (3.11f) with equality might not be reached. This implies that there are still spatial degrees-of-freedom that can be exploited. Therefore, during the third stage, SA following the

Algorithm 3: IFG Subcarrier Allocation Algorithm for MDD MIMO System

Input : Channel gains $\boldsymbol{h}_d[m]$ and $\boldsymbol{h}_{\bar{d}}[m], \forall d = 1, ..., D, \bar{d} = 1, ..., \bar{D}, m \in \mathcal{M}$
Output: $\mathcal{M}_{\mathrm{DL}}, \mathcal{M}_{\mathrm{UL}}, \mathcal{M}_d, \mathcal{M}_{\bar{d}}, \mathcal{D}_m, \bar{\mathcal{D}}_m$

1 **[The First Stage]**
2 **Initialization:**
3 $\boldsymbol{G} = [\boldsymbol{G}_1; \boldsymbol{G}_2], \boldsymbol{G}_1 \in \mathbb{C}^{D \times M}, \boldsymbol{G}_2 \in \mathbb{C}^{\bar{D} \times M}, \mathcal{T}_{\mathrm{DL}} = \mathcal{T}_{\mathrm{UL}} = \mathcal{M}$;
4 $\mathcal{M}_{\mathrm{DL}} = \mathcal{M}_{\mathrm{UL}} = \emptyset, \mathcal{M}_d = \mathcal{M}_{\bar{d}} = \emptyset, \mathcal{D}_m = \bar{\mathcal{D}}_m = \emptyset$;
5 **end**
6 **for** $k = 1 \rightarrow D_{sum}$ **do**
7 Calculate $(d, m) = \arg\max_{(i,j)} (\boldsymbol{G})_{i,j}$;
8 **if** $1 \leq d \leq D$ **then**
9 Update $\boldsymbol{G}_1^{(d,:)} = \boldsymbol{0}, \boldsymbol{G}_2^{(:,m)} = \boldsymbol{0}$ in $\boldsymbol{G}$;
10 $\mathcal{D}_m = \mathcal{D}_m \cup \{d\}, \mathcal{M}_{\mathrm{DL}} = \mathcal{M}_{\mathrm{DL}} \cup \{m\}, \mathcal{M}_d = \mathcal{M}_d \cup \{m\}, \mathcal{T}_{\mathrm{UL}} = \mathcal{T}_{\mathrm{UL}} \backslash \{m\}$;
11 **if** $|\mathcal{D}_m| = N_{RF}$ **then**
12 $\boldsymbol{G}_1^{(:,m)} = \boldsymbol{0}, \mathcal{T}_{\mathrm{DL}} = \mathcal{T}_{\mathrm{DL}} \backslash \{m\}$;
13 **end if**
14 **else**
15 Update $\boldsymbol{G}_2^{(d-D,:)} = \boldsymbol{0}, \boldsymbol{G}_1^{(:,m)} = \boldsymbol{0}$ in $\boldsymbol{G}$, then $\bar{\mathcal{D}}_m, \mathcal{M}_{\mathrm{UL}}, \mathcal{M}_{\bar{d}}, \mathcal{T}_{\mathrm{DL}}, \boldsymbol{G}_2$ and $\mathcal{T}_{\mathrm{UL}}$ are further updated by following the same rules from line 10 to line 13;
16 **end if**
17 **end for**
18 **end**
19 **[The Second Stage]**
20 **repeat**
21 Compute the rate $\tilde{R}_d, d = 1, ..., D$, and find the d-th DL MS satisfying $\frac{\tilde{R}_d}{\gamma_d} \leq \frac{\tilde{R}_i}{\gamma_i}$ for all $1 \leq i \leq D, i \neq d$;
22 Assign subcarrier m to the d-th MS that satisfies $\|\boldsymbol{h}_d[m]\|_2 \geq \|\boldsymbol{h}_d[j]\|_2$, $m, j \in \mathcal{T}_{\mathrm{DL}}$, and execute the same operations in line 10;
23 **if** $|\mathcal{D}_m| = N_{RF}$ **then**
24 $\mathcal{T}_{\mathrm{DL}} = \mathcal{T}_{\mathrm{DL}} \backslash \{m\}$;
25 **end if**
26 **until** $|\mathcal{M}_{DL}| = M_{DL}$;
27 Similarly, the UL allocation can be executed by following the operations from line 20 to line 26 to obtain the updated $\mathcal{M}_{\bar{d}}, \mathcal{M}_{\mathrm{UL}}, \bar{\mathcal{D}}_m, \mathcal{T}_{\mathrm{DL}}$ and $\mathcal{T}_{\mathrm{UL}}$;
28 **end**
29 **[The Third Stage]**
30 **repeat**
31 Carry out the operations described from line 20 to line 26, with only $\mathcal{D}_m$ and $\mathcal{M}_d$ shown in line 10 being updated;
32 **until** $\mathcal{T}_{DL} = \emptyset$;
33 Similarly, the UL allocation can be done by following the operations from line 20 to line 26 to further update $\mathcal{M}_{\bar{d}}$ and $\bar{\mathcal{D}}_m$;
34 **end**

second stage is continued, until the constraints of (3.11e) and (3.11f) with equality are reached.

It can be shown that after SA, the fairness constraints of (3.11g) and (3.11h) can be coarsely accomplished. Next, PA is executed to attain the proportional rate constraints of (3.11g) and (3.11h), while maximizing the total rate of the system, as detailed below.

3.4.2 Power Allocation

After SA, the optimization problem in (3.11) can be divided into the DL PA and UL PA, which can be executed independently.

Specifically, for DL, the PA problem can now be stated as

$$\max_{\boldsymbol{F}_{\mathrm{RF}},\boldsymbol{f}_{\mathrm{BB}}^{d}[m],p_{m,d}} \sum_{m\in\mathcal{M}_{\mathrm{DL}}} \sum_{d\in\mathcal{D}_m} \log_2\left(1+\frac{p_{m,d}\left|\boldsymbol{h}_d^H[m]\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}_{\mathrm{BB}}^{d}[m]\right|^2}{\sum\limits_{i\in\mathcal{D}_m,i\neq d} p_{m,i}\left|\boldsymbol{h}_d^H[m]\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}_{\mathrm{BB}}^{i}[m]\right|^2+\sigma^2}\right) \tag{3.13a}$$

$$\text{s.t. } (3.11g),(3.11i),(3.11j),$$

$$\sum_{d=1}^{D}\sum_{m\in\mathcal{M}_d} p_{m,d}\le P_{\mathrm{DL}}. \tag{3.13b}$$

By contrast, for UL, the PA problem can be formulated as

$$\max_{p_{m,\bar{d}}} \sum_{m\in\mathcal{M}_{\mathrm{UL}}} \sum_{\bar{d}\in\bar{\mathcal{D}}_m} \log_2\left(1+\frac{p_{m,\bar{d}}\left|\boldsymbol{w}_{\mathrm{BB}}^{\bar{d},H}[m]\boldsymbol{W}_{\mathrm{RF}}^H\boldsymbol{h}_{\bar{d}}[m]\right|^2}{\sum\limits_{j\in\bar{\mathcal{D}}_m,j\neq\bar{d}} p_{m,j}\left|\boldsymbol{w}_{\mathrm{BB}}^{\bar{d},H}[m]\boldsymbol{W}_{\mathrm{RF}}^H\boldsymbol{h}_j[m]\right|^2+IN_{m,\bar{d}}}\right) \tag{3.14a}$$

$$\text{s.t. } \sum_{m\in\mathcal{M}_{\bar{d}}} p_{m,\bar{d}}\le P_{\bar{d}},\ \forall\bar{d}. \tag{3.14b}$$

Observing from (3.13) and (3.14), we can know that DL PA is under the constraint of the total power of all DL MSs. By contrast, each UL MS carries out PA separately under the constraint of individual MS's power. Below we first consider DL PA.

3.4.2.1 Downlink Power Allocation

The optimization problem in (3.13) is the sum-rate maximization problem coupled with fairness constraints, which is still a non-convex problem. To simplify it, we first apply the ZF precoding with assuming a fixed power assignment to all MSs so as to obtain a full-digital precoder. This will allow us to transfer the non-convex problem of (3.13)

to a convex optimization problem. To this objective, the full-digital precoder can be expressed as

$$\boldsymbol{F}_{\mathrm{ZF}}[m] = \boldsymbol{H}_{\mathrm{DL}}^{H}[m]\left(\boldsymbol{H}_{\mathrm{DL}}[m]\boldsymbol{H}_{\mathrm{DL}}^{H}[m]\right)^{-1}, \tag{3.15}$$

where $\boldsymbol{H}_{\mathrm{DL}}[m] \in \mathbb{C}^{|\mathcal{D}_m|\times N} = \left[\boldsymbol{h}_1[m],...,\boldsymbol{h}_d[m],...,\boldsymbol{h}_{|\mathcal{D}_m|}[m]\right]^H$, and $\boldsymbol{F}_{\mathrm{ZF}}[m] \in \mathbb{C}^{N\times|\mathcal{D}_m|} = \left[\boldsymbol{f}_{\mathrm{ZF}}^{1}[m],...,\boldsymbol{f}_{\mathrm{ZF}}^{d}[m],...,\boldsymbol{f}_{\mathrm{ZF}}^{|\mathcal{D}_m|}[m]\right]$. It can be shown that owing to the power constraint in (3.13d), $\boldsymbol{f}_{\mathrm{ZF}}^{d}[m]$ is in fact a normalized version, given as $\boldsymbol{f}_{\mathrm{ZF}}^{d}[m] \leftarrow \boldsymbol{f}_{\mathrm{ZF}}^{d}[m] / \left\|\boldsymbol{f}_{\mathrm{ZF}}^{d}[m]\right\|_2$. Consequently, upon replacing $\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}_{\mathrm{BB}}^{d}[m]$ in (3.13) by $\boldsymbol{f}_{\mathrm{ZF}}^{d}[m]$, the DL PA problem can now be formulated as

$$\begin{aligned}\max_{p_{m,d}} \sum_{d=1}^{D}\sum_{m\in\mathcal{M}_d} \log_2\left(1+p_{m,d}J_{m,d}\right) \\ \text{s.t. } (3.11g), (3.13b),\end{aligned} \tag{3.16}$$

where

$$J_{m,d} = \frac{\left|\frac{\boldsymbol{h}_d^H[m]\boldsymbol{f}_{\mathrm{ZF}}^{d}[m]}{\left\|\boldsymbol{f}_{\mathrm{ZF}}^{d}[m]\right\|_2}\right|^2}{\left(\sum\limits_{i\in\mathcal{D}_m, i\neq d} p_{m,i}\left|\frac{\boldsymbol{h}_d^H[m]\boldsymbol{f}_{\mathrm{ZF}}^{i}[m]}{\left\|\boldsymbol{f}_{\mathrm{ZF}}^{i}[m]\right\|_2}\right|^2 + \sigma^2\right)}. \tag{3.17}$$

It can be readily shown that this is a convex optimization problem[6]. Therefore, by introducing the Lagrange multiplier, the equivalent cost function can be formulated as

$$\begin{aligned}L = &\sum_{d=1}^{D}\sum_{m\in\mathcal{M}_d}\log_2\left(1+p_{m,d}J_{m,d}\right) + \lambda_1\left(\sum_{d=1}^{D}\sum_{m\in\mathcal{M}_d} p_{m,d} - P_{\mathrm{DL}}\right) \\ &+ \sum_{d=2}^{D}\lambda_d\left(\sum_{m\in\mathcal{M}_1}\log_2\left(1+p_{m,1}J_{m,1}\right) - \frac{\gamma_1}{\gamma_d}\sum_{m\in\mathcal{M}_d}\log_2\left(1+p_{m,d}J_{m,d}\right)\right).\end{aligned} \tag{3.18}$$

In order to obtain the optimal $\{p_{m,d}\}$ from (3.18), by following [155], the PA can be divided into two parts, namely the PA among the subcarriers of a DL MS for given power P_d of this MS and the PA among the DL MSs. Specifically, when allocating the power P_d of the d-th DL MS to its $|\mathcal{M}_d|$ subcarriers, we assume without any loss of generality that $J_{1,d} \leq J_{2,d} \leq ... \leq J_{|\mathcal{M}_d|,d}$. Then, it can be shown that with the aid of the KKT conditions [156], we have

$$\begin{aligned}p_{m,d} &= p_{1,d} + \frac{J_{m,d}-J_{1,d}}{J_{m,d}J_{1,d}}, \text{ and} \\ P_d &= |\mathcal{M}_d|\, p_{1,d} + \sum_{m=2}^{|\mathcal{M}_d|}\frac{J_{m,d}-J_{1,d}}{J_{m,d}J_{1,d}}\end{aligned} \tag{3.19}$$

[6]Note that, as the result of ZF precoding, the interference imposed by the other MSs on the d-th DL MS is approximately zero, i.e., we have $\sum\limits_{i\in\mathcal{D}_m, i\neq d}\left|\frac{\boldsymbol{h}_d^H[m]\boldsymbol{f}_{\mathrm{ZF}}^{i}[m]}{\left\|\boldsymbol{f}_{\mathrm{ZF}}^{i}[m]\right\|_2}\right|^2 \approx 0$. Therefore, $H_{m,d}$ is irrelevant to the variables $\left\{p_{m,i}\right\}_{i\neq d}$.

for $m \in \mathcal{M}_d$ and $d = 1, 2, ..., D$. The PA of (3.19) implies that the optimal power assignment among the different subcarriers of a DL MS follows the water-filling principles [157]. Hence, it may happen that some subcarriers of a DL MS may be allocated negative power. If this happens, those subcarriers with negative power are removed from the subcarrier set of the MS.

In the context of the PA among different DL MSs, the solutions can be obtained via solving the two sets of nonlinear equations by the Newton-Raphson method [155], expressed as,

$$\begin{aligned} &\frac{\gamma_1}{\gamma_d} = \frac{|\mathcal{M}_1| \left(\log_2 \left(1 + J_{1,1} \frac{P_1 - V_1}{|\mathcal{M}_1|} \right) + \log_2 G_1 \right)}{|\mathcal{M}_d| \left(\log_2 \left(1 + J_{1,d} \frac{P_d - V_d}{|\mathcal{M}_d|} \right) + \log_2 G_d \right)}, \text{ and} \\ &\sum_{d=1}^{D} P_d = P_{\mathrm{DL}} \end{aligned} \tag{3.20}$$

for $d = 1, 2, .., D$, where $V_d = \sum_{m=2}^{|\mathcal{M}_d|} \frac{J_{m,d} - J_{1,d}}{J_{m,d} J_{1,d}}$ and $G_d = \left(\prod_{m=2}^{|\mathcal{M}_d|} \frac{J_{m,d}}{J_{1,d}} \right)^{\frac{1}{|\mathcal{M}_d|}}$. To summarize, the PA of DL is stated in Algorithm 4.

Algorithm 4: Power Allocation Algorithm for DL

Input : $\mathcal{M}_{\mathrm{DL}}, \mathcal{M}_d, \mathcal{D}_m, J_{m,d}, P_{\mathrm{DL}}, \boldsymbol{H}_{\mathrm{SI}}$
Output: $\mathcal{M}_{\mathrm{DL}}, \mathcal{M}_d, \mathcal{D}_m, p_{m,d}, P_d, \boldsymbol{W}_{\mathrm{RF}}$
Initialization: $P_d = \frac{P_{\mathrm{DL}}}{D}$, for $\forall d$;
[Power allocation among DL MSs]
  repeat
    Solve two sets of nonlinear equations in (3.20) by the Newton-Raphson method to obtain $\{P_d\}_{d=1}^{D}$;
    for $d = 1 \rightarrow D$ **do**
      if $P_d \leq V_d$ **then**
        Calculate $\{p_{m,d}\}_{m \in \mathcal{M}_d}$ based on (3.19);
        For any $p_{m,d} \leq 0$, the corresponding subcarrier index is removed from $\mathcal{M}_d$ and $\mathcal{M}_{\mathrm{DL}}$;
        Update $\mathcal{D}_m$, $\boldsymbol{f}_{\mathrm{ZF}}^{d}[m]$, and $J_{m,d}$;
      else
        Continue;
      end if
    end for
  until $P_d > V_d$, *for* $\forall d$;
end
[Power allocation among subcarriers of individual MSs]
  for $d = 1 \rightarrow D$ **do**
    Apply the water-filling algorithm to obtain $\{p_{m,d}\}_{m \in \mathcal{M}_d}$ based on (3.19);
  end for
end

3.4.2.2 Uplink Power Allocation

For UL, PA is only among the subcarriers of a single UL MS. As there is no fairness issue, the objective during this stage is solely to maximize data rate. Hence, the water-filling PA can be executed based on the similar equation in (3.19) but under the constraint of the total power of an individual MS.

3.5 Hybrid Beamformer Design Associated with Resource Allocation

In this section, we first propose two approaches to design the hybrid precoder associated with RA. Then, an adaptive SIC algorithm is introduced for practical FD systems via the hybrid combiner design.

3.5.1 Hybrid Precoder

3.5.1.1 Factorization Approach

In order to obtain $\boldsymbol{F}_{\mathrm{RF}}$ and $\boldsymbol{f}_{\mathrm{BB}}^{d}[m]$, $m \in \mathcal{M}_{\mathrm{DL}}$, $d \in \mathcal{D}_m$, the factorization approach for designing the hybrid precoder from $\boldsymbol{f}_{\mathrm{ZF}}^{d}[m]$ can be formulated as

$$\begin{aligned} \min_{\boldsymbol{F}_{\mathrm{RF}},\left\{\boldsymbol{f}_{\mathrm{BB}}^{d}[m]\right\}} \; & \sum_{m\in\mathcal{M}_{\mathrm{DL}}} \sum_{d\in\mathcal{D}_m} \left\| \boldsymbol{f}_{\mathrm{ZF}}^{d}[m] - \boldsymbol{F}_{\mathrm{RF}} \boldsymbol{f}_{\mathrm{BB}}^{d}[m] \right\|_2^2 \\ & \text{s.t. } (3.11i), (3.11j). \end{aligned} \tag{3.21}$$

It is noteworthy that our factorization approach differs from the one considered in [122] and [118], in which the power constraint is set to $\|\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}}[m]\|_F^2 = |\mathcal{D}_m|$ with $\boldsymbol{F}_{\mathrm{BB}}[m] = \left[\boldsymbol{f}_{\mathrm{BB}}^{1}[m], \boldsymbol{f}_{\mathrm{BB}}^{2}[m], ..., \boldsymbol{f}_{\mathrm{BB}}^{|\mathcal{D}_m|}[m]\right]$. In our optimization, as shown in (3.21), the more strict constraints of $\left\|\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}_{\mathrm{BB}}^{d}[m]\right\|_2^2 = 1$, $\forall m, d$ are imposed. In order to solve the problem of (3.21), first, we minimize $\left\|\boldsymbol{f}_{\mathrm{ZF}}^{d}[m] - \boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}_{\mathrm{BB}}^{d}[m]\right\|_2^2$ for a given $\boldsymbol{F}_{\mathrm{RF}}$. This gives a least square (LS) solution to $\boldsymbol{f}_{\mathrm{BB}}^{d}[m]$, i.e., $\hat{\boldsymbol{f}}_{\mathrm{BB}}^{d}[m] = \boldsymbol{F}_{\mathrm{RF}}^{\dagger}\boldsymbol{f}_{\mathrm{ZF}}^{d}[m]$, where $\boldsymbol{F}_{\mathrm{RF}}^{\dagger} = \left(\boldsymbol{F}_{\mathrm{RF}}^{H}\boldsymbol{F}_{\mathrm{RF}}\right)^{-1}\boldsymbol{F}_{\mathrm{RF}}^{H}$, as $N > N_{\mathrm{RF}}$ and $\boldsymbol{F}_{\mathrm{RF}}$ is assumed to be a full column rank matrix. After the power normalization to meet $\left\|\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}_{\mathrm{BB}}^{d}[m]\right\|_2^2 = 1$, we obtain

$$\tilde{\boldsymbol{f}}_{\mathrm{BB}}^{d}[m] = \frac{\hat{\boldsymbol{f}}_{\mathrm{BB}}^{d}[m]}{\left\|\boldsymbol{F}_{\mathrm{RF}}\hat{\boldsymbol{f}}_{\mathrm{BB}}^{d}[m]\right\|_2}. \tag{3.22}$$

Then, substituting $\tilde{\boldsymbol{f}}_{\mathrm{BB}}^{d}[m]$ for $\boldsymbol{f}_{\mathrm{BB}}^{d}[m]$ into (3.21), the cost function of the problem can be rewritten as

$$\Phi\left(\boldsymbol{F}_{\mathrm{RF}}\right)=\sum_{m\in\mathcal{M}_{\mathrm{DL}}}\sum_{d\in\mathcal{D}_{m}}\left\|\boldsymbol{f}_{\mathrm{ZF}}^{d}[m]-\boldsymbol{F}_{\mathrm{RF}}\tilde{\boldsymbol{f}}_{\mathrm{BB}}^{d}[m]\right\|_{2}^{2}. \tag{3.23}$$

According to the PGD principle [158], the local minimum can be reached via updating $\boldsymbol{F}_{\mathrm{RF}}$ along the negative conjugate gradient direction and then, projecting the solution to the feasible set where the elements of matrix have unit modulus. Hence, after obtaining the conjugate gradient as shown in Appendix B, the above-mentioned two steps can be expressed as

$$\begin{aligned}\boldsymbol{F}_{\mathrm{RF}}^{(i+1)}&=\boldsymbol{F}_{\mathrm{RF}}^{(i)}-l\frac{\partial(\Phi)}{\partial\boldsymbol{F}_{\mathrm{RF}}^{*(i)}},\\ \boldsymbol{F}_{\mathrm{RF}}^{(i+1)}&=\arg(\boldsymbol{F}_{\mathrm{RF}}^{(i+1)}),\end{aligned} \tag{3.24}$$

where l denotes a positive decreasing step size towards the local optimum. Once the optimal $\boldsymbol{F}_{\mathrm{RF}}$ is obtained, $\boldsymbol{f}_{\mathrm{BB}}^{d}[m]$ can be derived as

$$\boldsymbol{f}_{\mathrm{BB}}^{d}[m]=\left(\boldsymbol{F}_{\mathrm{RF}}^{(i)}\right)^{\dagger}\boldsymbol{f}_{\mathrm{ZF}}^{d}[m]/\left\|\boldsymbol{F}_{\mathrm{RF}}^{(i)}\left(\boldsymbol{F}_{\mathrm{RF}}^{(i)}\right)^{\dagger}\boldsymbol{f}_{\mathrm{ZF}}^{d}[m]\right\|_{2},\ \forall d,m. \tag{3.25}$$

3.5.1.2 Direct Approach

Although the PGD algorithm with a decreasing step size guarantees the convergence to a local minimum, the difference between $\boldsymbol{f}_{\mathrm{ZF}}^{d}[m]$ and $\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}_{d}[m]$ may be increased by RA, especially when the number of RF chains is small, which results in that the interference in (3.17) from other MSs cannot be approximated to zero. In this case, Algorithm 4 may be hard to achieve the proportional rates among DL MSs at a sufficient accuracy, as the result that $H_{m,d}$ in (3.18) also includes $p_{m,i}, i\neq d$, in addition to the desired $p_{m,d}$. Consequently, due to the deviation introduced by the hybrid precoder, the problem of (3.16) is no longer a convex optimization problem, which is hence prohibitive from being solved by an efficient approach. Therefore, for comparison purposes, below we present a so-called Direct Approach (DA) to derive $\boldsymbol{F}_{\mathrm{RF}}$ and $\boldsymbol{f}_{d}[m]$, which is capable of providing the precise proportions of fairness.

To begin with, according to [123], suboptimal $\boldsymbol{F}_{\mathrm{RF}}$ can be obtained by solving the following optimization problem :

$$\begin{aligned}&\max_{\boldsymbol{F}_{\mathrm{RF}}}\ \log_{2}\left|\boldsymbol{I}+\frac{P_{\mathrm{DL}}}{\sigma^{2}NN_{\mathrm{RF}}}\boldsymbol{F}_{\mathrm{RF}}^{H}\tilde{\boldsymbol{H}}\boldsymbol{F}_{\mathrm{RF}}\right|\\ &\text{s.t. }\ (3.11j),\end{aligned} \tag{3.26}$$

where $\tilde{\boldsymbol{H}} = \frac{1}{|\mathcal{M}_{\mathrm{DL}}|} \sum_{m \in \mathcal{M}_{\mathrm{DL}}} \left(\boldsymbol{H}_{\mathrm{DL}}[m]^H \boldsymbol{H}_{\mathrm{DL}}[m] \right)$. Next, the digital ZF precoders $\boldsymbol{F}_{\mathrm{ZF}}^{\mathrm{BB}}[m]$ are obtained by replacing $\boldsymbol{H}_{\mathrm{DL}}[m]$ with $\boldsymbol{H}_{\mathrm{eq}}[m]$ in (3.15), where $\boldsymbol{H}_{\mathrm{eq}}[m] = \boldsymbol{H}_{\mathrm{DL}}[m]\boldsymbol{F}_{\mathrm{RF}}$. Then, after normalization, we can solve the optimization problem by letting $\boldsymbol{F}_{\mathrm{ZF}}[m]$ in (3.15) equal to $\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{ZF}}^{\mathrm{BB}}[m]$. Since the optimization (3.16) is convex, which enables us to derive the solutions directly using Algorithm 4, which yields the precise proportions of fairness. The details can be found in Algorithm 5 as shown below.

Algorithm 5: Hybrid Precoder Design: Direct Approach

Input : $\mathcal{M}_{\mathrm{DL}}, \mathcal{M}_d, \mathcal{D}_m$ (these three sets are obtained by SA), P_{DL}
Output: $\mathcal{M}_{\mathrm{DL}}, \mathcal{M}_d, \mathcal{D}_m, p_{m,d}, P_d, \boldsymbol{F}_{\mathrm{RF}}, \boldsymbol{f}_d[m]$
1 Initialization: $P_d = \frac{P_{\mathrm{DL}}}{D}, \forall d$;
2 Compute $\boldsymbol{F}_{\mathrm{RF}}$ and $\boldsymbol{F}_{\mathrm{ZF}}^{\mathrm{BB}}[m]$ using (3.26) and (3.15), respectively ;
3 Calculate $J_{m,d}$ in (3.17) by letting $\boldsymbol{F}_{\mathrm{ZF}}[m]$ equal to $\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{ZF}}^{\mathrm{BB}}[m]$, $\forall m \in \mathcal{M}_{\mathrm{DL}}, d \in \mathcal{D}_m$;
4 Execute step 2 to step 19 in Algorithm 4

It is noteworthy that both the proposed hybrid beamforming algorithms have the same overall computation complexity, expressed as $\mathcal{O}(NN_{\mathrm{RF}}^2)$. However, it can be observed from Algorithm 5 that the DA method is coupled with PA. This means that the hybrid precoders in the DA method need to be continuously updated during PA, as shown in Algorithm 4 (line 9), which leads to an increased complexity. By contrast, the hybrid beamformers in the PGD method are independent of the PA, which are hence simpler for practical implementation.

3.5.2 Hybrid Combiner with Self-Interference Cancellation

During RA, the effect of SI is ignored by assuming perfect SI suppression. However, existing SIC methods are unable to provide sufficient SI reduction while keeping a low complexity and overhead at the same time. Hence in this section, we introduce an adaptive beamforming based SIC algorithm by exploiting the large number of antennas employed in mmWave systems.

First, according to (3.9), the total power of the SI signal $\boldsymbol{r}_{\mathrm{SI}}$ after analog combining can be expressed as

$$\mathrm{Tr}\left(\mathrm{Cov}\left(\boldsymbol{r}_{\mathrm{SI}}\right)\right) = \xi \mathrm{Tr}\left(\boldsymbol{F}_{\mathrm{RF}}^H \boldsymbol{H}_{\mathrm{SI}}^H \boldsymbol{W}_{\mathrm{RF}} \boldsymbol{W}_{\mathrm{RF}}^H \boldsymbol{H}_{\mathrm{SI}} \boldsymbol{F}_{\mathrm{RF}} \times \left(\sum_{d=1}^{D} \sum_{m \in \mathcal{M}_d} p_{m,d} \boldsymbol{f}_{\mathrm{BB}}^d[m] \boldsymbol{f}_{\mathrm{BB}}^{d,H}[m] \right) \right), \tag{3.27}$$

from which we can observe that the SI power is proportional to $\left\| \boldsymbol{W}_{\mathrm{RF}}^H \boldsymbol{H}_{\mathrm{SI}} \boldsymbol{F}_{\mathrm{RF}} \right\|_F^2$. Since $\boldsymbol{F}_{\mathrm{RF}}$ has already been determined by DL PA and is fixed, the suppression of SI can only be depended on the design of the analog combiner $\boldsymbol{W}_{\mathrm{RF}}$. Therefore, we have the

optimization problem stated as

$$\min_{\boldsymbol{W}_{\mathrm{RF}}} \left\| \boldsymbol{W}_{\mathrm{RF}}^{H} \boldsymbol{H}_{\mathrm{SI}} \boldsymbol{F}_{\mathrm{RF}} \right\|_{F}^{2} \quad \text{s.t. } (3.11j). \tag{3.28}$$

The problem (3.28) can be solved by the CCD algorithm via iteratively optimizing the individual elements of $\boldsymbol{W}_{\mathrm{RF}}$, the details of which can be found in [69] and are not repeated for brevity. After obtaining the analog combiner for SIC, the optimal digital combiner for a subcarrier can be derived based on conventional MMSE approach [159].

Note that our proposed SIC is operated independently from RA. Hence, the process of SIC does not have any impact on the operations of RA, which were considered in the previous sections.

3.6 Performance Results

In this section, we firstly present the simulation results for the MDD MIMO mmWave systems with the proposed RA algorithms, when the QoS requirements of both UL and DL MSs are imposed. Then, to tackle the problem of SI in MDD RA systems, the performance of our proposed SI reduction method is investigated in Section 3.6.2.

3.6.1 Performance of Resource Allocation

First, let us demonstrate the performance of the MDD MIMO mmWave systems with our proposed RA algorithms. To begin with, we show the performance upper-bound by following [151] without considering the fairness constraints, when SA is achieved using the UG algorithm, while transmit power is assigned by the water-filling algorithm. However, we should note that in [151], all MSs and BS are equipped with single antenna and are operated in HD mode. Hence, to make it compatible with our MDD MIMO system, we consider a modified UG (MUG) algorithm for the scheme of [151] to operate the second and third stages of SA but keep the water-filling algorithm for PA. In the MUG algorithm, we assume that inter-user interference is mitigated by a ZF precoder. Then, each subcarrier may be assigned to multiple MSs with the best channel gains under the constraint of $|\mathcal{D}_m| \leq N_{\mathrm{RF}}$. Afterwards, BS allocates the transmit power to all DL subcarriers based on the water-filling principle, while UL power assignment is implemented via water-filling for each UL MS's subcarriers. Additionally, we also compare another RA method, namely MUG-PowAve. With the MUG-PowAve, SA is the same as the MUG algorithm, but DL power is firstly evenly distributed to all DL subcarriers, and then water-filling is used to assign the power of a subcarrier to the

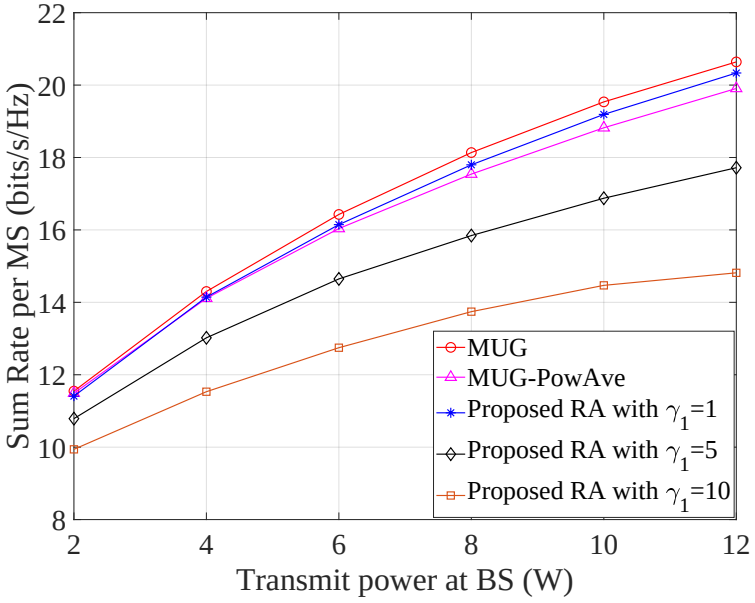


FIGURE 3.4: Performance comparison of the MUG algorithm and the proposed algorithm with different fairness constraints, when $\bar{N} = 128$ $\bar{N}_{\mathrm{RF}} = 16$, $N = 32$, $N_{\mathrm{RF}} = 16$, $D = 20$, $\bar{D} = 10$.

MSs sharing this subcarrier. Note that this RA approach has been considered in many references, such as, in [122, 159].

In this subsection, we consider a MDD MIMO mmWave system operated in 28 GHz with the SIC capability of $\xi = -100$ dB, where BS employs $N = 32$ transmit antennas and $N_{\mathrm{RF}} = 16$ DL RF chains to support $D = 20$ DL MSs, and $\bar{N} = 128$ receive antennas and $\bar{N}_{\mathrm{RF}} = 16$ UL RF chains to serve $\bar{D} = 10$ UL MSs. All MSs are randomly distributed with their distances from BS being between 50 m to 200 m, i.e., $d_m \in [50, 200]$. The path-loss is $P_L(\mathrm{dB}) = 72 + 29.2 \cdot \log10(d_m)$. The total number of subcarriers are $M = 64$, in which the number of DL and UL subcarriers used for transmissions satisfy $|\mathcal{M}_{\mathrm{DL}}| = |\mathcal{M}_{\mathrm{UL}}| \leq \frac{M}{2} = 32$. The transmit power of BS transmitter is $P_{\mathrm{DL}} = 12$ W, and all UL MSs have the same transmit power of $P_{\bar{d}} = P_{\mathrm{DL}}/D = 0.6$ W. All the above-mentioned parameter values are default values, unless they are stated specifically. The default proportional fairness among DL and UL MSs are $\{\gamma_i = 1\}_{i=1}^{D}$ and $\{\eta_i = 1\}_{i=1}^{\bar{D}}$, respectively. Additionally, we will consider the other two fairness constraints of $\gamma_1 = 5$ and $\gamma_1 = 10$, respectively, while the other values remain the same as the default value of 1. Besides, we assume that the transmit and receive arrays at BS are the ULA arrays with the half-wavelength spacing between two adjacent antenna elements. Both DL and UL channels are assumed to have $T = 6$ delay taps and $L_d = 4$ paths. The channel gains obey the distribution $v_{l,d} \in \mathcal{CN}(0, 1)$, and the AoA/AoD azimuth angles are assumed to be uniformly distributed in $[0, 2\pi]$. Furthermore, the noise variance σ^2 is set to 10^{-13} W. The results are obtained via average of 100 channel realizations.

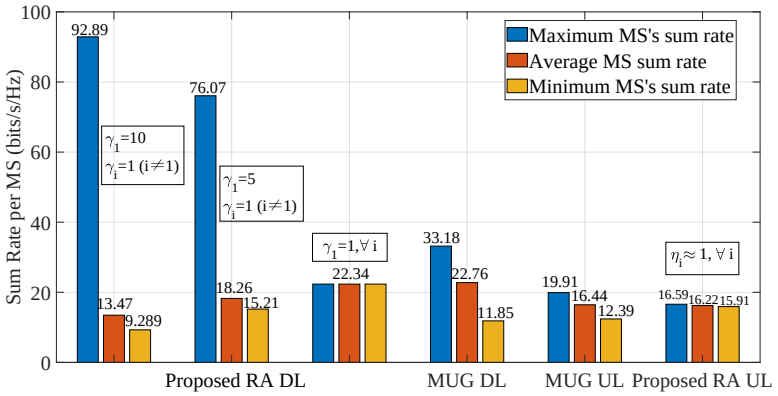


FIGURE 3.5: Fairness comparison of MUG algorithm and proposed RA algorithm, when $P_{\text{DL}} = 12\ W$, $\bar{N} = 128$ $\bar{N}_{\text{RF}} = 16$, $N = 32$, $N_{\text{RF}} = 16$, $D = 20$, $\bar{D} = 10$.

Firstly in Fig. 3.4, we demonstrate the performance of the MDD MIMO mmWave systems with our proposed RA algorithm having different fairness constraints. Here, it is assumed that full-digital precoders are employed at BS transmitter. Fig. 3.4 demonstrates that the proposed RA with $\{\gamma_i = 1\}_{i=1}^{D}$ slightly outperforms the MUG-PowAve method but slightly worse than the MUG method in terms of the sum rate per MS. As γ_1 increases, the average sum rate per MS achieved by the proposed RA becomes lower. The reason behind is that all MSs in the system experience large-scale fading. Due to the high level path-loss in mmWave communications, if one MS far away from the BS imposes the highest demand for a data rate, which corresponds to $\gamma_1 = 5$ or 10, the BS has to assign most of the transmit power to this MS, making the PA inefficient and hence the degradation of overall data rate.

Fig. 3.5 shows the fairness comparison between the MUG algorithm and the proposed RA algorithm in different cases. Explicitly, our proposed RA method can achieve the near-accurate proportions of rates for the DL MSs, when γ_1 has different values. Fig. 3.5 also shows the corresponding results of the MUG algorithm, when $\{\gamma_i = 1\}_{i=1}^{D}$ and $\{\eta_i = 1\}_{i=1}^{\bar{D}}$. We can see that the MUG algorithm slightly outperforms the proposed RA algorithm in terms of the average sum rate per DL or UL MS. However, it causes a big difference between the maximum and minimum rates (about 21 bits/s/Hz), which explains the possible poor performance attained by the MSs located at cell edge. By contrast, for the proposed RA algorithm, the proportional fairness of all DL MSs can be guaranteed. However, as the fairness in UL case is only considered via SA, the achieved rates of UL MSs have small fluctuation. Nevertheless, they are still much more stable than that obtained by the MUG algorithm.

Next, we compare the complexity of our proposed and the competing RA methods. In all these schemes, RA is divided into two stages, SA and PA. Both MUG and MUG-PowAve algorithms use the same SA method, which has the complexity of $\mathcal{O}(M)$, while the proposed method needs to meet the fairness constraint and has the complexity of $\mathcal{O}\left(D_{\text{sum}}M\right)$. For PA, both MUG and MUG-PowAve apply the water-filling

TABLE 3.1: Complexity comparison of the proposed and competing RA methods in MDD systems

Method	Complexity	Sum Rate	Feature
MUG	$\mathcal{O}(2M)$	High	Optimal RA method
MUG-PowAve	$\mathcal{O}(N_{\text{RF}}M + M)$	Medium	Easy-to-implement
Proposed ($\gamma_1 = 1$)	$\mathcal{O}(D_{\text{sum}}M + D^3)$	High	Fairness guaranteed

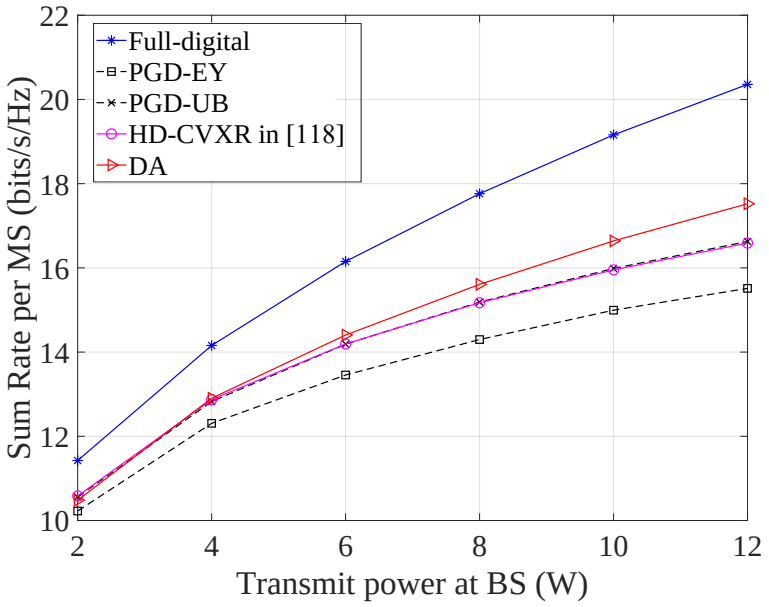


FIGURE 3.6: Sum rate versus transmit power of BS when different hybrid precoders are employed and when assuming $\bar{N} = 128$, $\bar{N}_{\text{RF}} = 16$, $N = 32$, $N_{\text{RF}} = 16$, $D = 20$, $\bar{D} = 10$, $\{\gamma_i = 1\}_{i=1}^{D}$ and $\{\eta_i = 1\}_{i=1}^{\bar{D}}$.

algorithm, whose complexity is linearly proportional to the number of variables used in the optimization problem [160]. Therefore, these two methods have the complexity of $\mathcal{O}(M)$ and $\mathcal{O}(N_{\text{RF}}M)$, respectively. By contrast, the proposed algorithm needs to solve the nonlinear equations and calculate the matrix inversion when applying the New-Raphson method. These lead to a complexity of $\mathcal{O}(D^3)$. Overall, the computational complexity of the considered RA algorithms is summarized in Table 3.1. In a nutshell, although our proposed algorithm has higher complexity than the other two UG algorithms, when the number of MSs becomes large, this extra complexity investment lead to the promising performance and at the same time, the near-accurate fairness defined for different MSs.

Fig. 3.6 and 3.7 show the sum rate performance of the MDD MIMO mmWave systems, when various hybrid precoders are employed. Here PGD-EY indicates the hybrid precoder with $\boldsymbol{F}_{\text{RF}}$ initialized by the Eckart-Young theorem [122]. PGD-UB indicates the hybrid precoder with $\boldsymbol{F}_{\text{RF}}$ initialized by maximizing the upper bound of DL's spectral efficiency for a given $\boldsymbol{F}_{\text{BB}}$ [159]. Additionally, the hybrid precoder design by convex

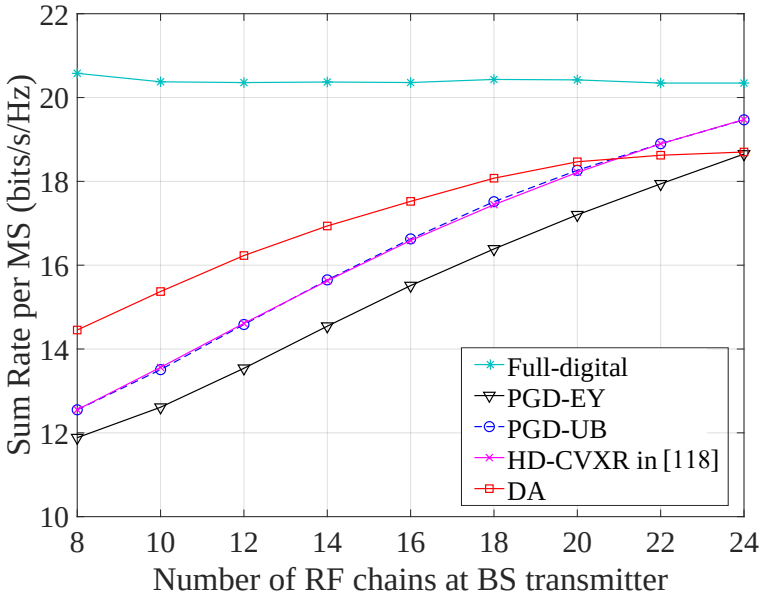


FIGURE 3.7: Sum rate versus the number of RF chains at BS transmitter, when different hybrid precoder designs are employed and when assuming $P_{\mathrm{DL}} = 12\ W$, $\bar{N} = 128$, $\bar{N}_{\mathrm{RF}} = 16$, $N = 32$, $D = 20$, $\bar{D} = 10$, $\{\gamma_i = 1\}_{i=1}^{D}$ and $\{\eta_i = 1\}_{i=1}^{\bar{D}}$.

relaxation (HD-CVXR) presented in [118] is compared, which is demonstrated to provide a high spectral efficiency in the conventional hybrid beamforming systems. Fig. 3.6 shows that when $N_{\mathrm{RF}} = 16$, DA is capable of achieving the highest sum rate among the considered schemes, which is only 3 bits/s/Hz lower than that obtained by the full-digital ZF solution. The reason behind is that the factorization of the full-digital ZF precoder using the approach, such as PGD and HD-CVXR, introduces inter-user interference within a subcarrier. This interference is further amplified by PA. The merit of DA is that the inter-user interference can be efficiently mitigated by applying ZF processing for $\boldsymbol{F}_{\mathrm{BB}}$. As shown in Fig. 3.6, the initialization of RF precoder by the PGD method generates a big impact on the sum rate performance, and the one from [159] is better for our system than the other one from [122].

From Fig. 3.7, we observe that both the PGD-UB and HD-CVXR algorithms surpass the DA, when the number of RF chains is big, which converge to the upper bound performance achieved by the full-digital ZF. This can be explained as follows. When the number of RF chains is relatively big, the difference between the full-digital ZF and the hybrid precoders becomes sufficiently small, making the inter-user interference yields little influence on the sum rate even after PA.

Finally, Fig. 3.8 illustrates the rates achieved by individual DL MS on one channel realization. Compared with Fig. 3.7 where the methods of PGD-UB, DA and HD-CVXR achieve nearly the same sum rate within 100 channel realizations when $N_{\mathrm{RF}} = 20$ and the fairness constraint on DL MSs is $\{\gamma_i = 1\}_{i=1}^{D}$, Fig. 3.8 shows that, as analyzed in Section 3.5, the DA with higher implementation complexity can achieve the exact

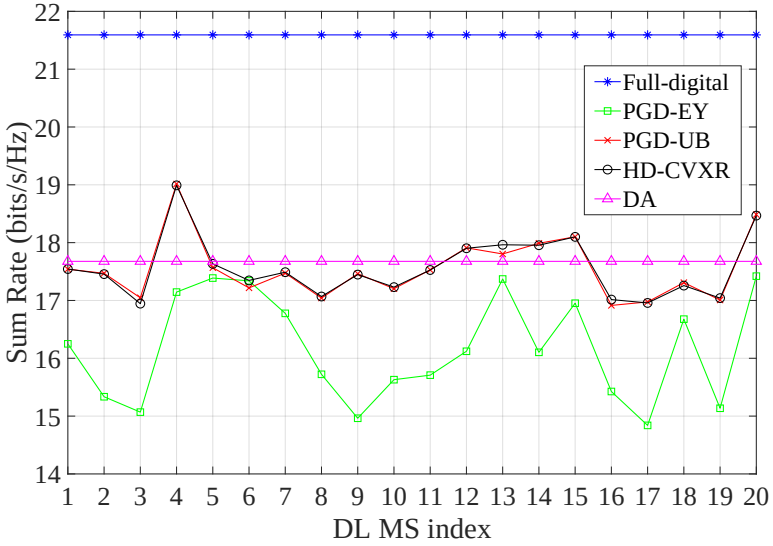


FIGURE 3.8: Rates of individual DL MSs achieved by different hybrid precoder designs on one channel realization, when assuming $P_{\mathrm{DL}} = 12$ W, $\bar{N} = 128$, $\bar{N}_{\mathrm{RF}} = 16$, $N = 32$, $N_{\mathrm{RF}} = 20$, $D = 20$ and $\{\gamma_i = 1\}_{i=1}^{D}$.

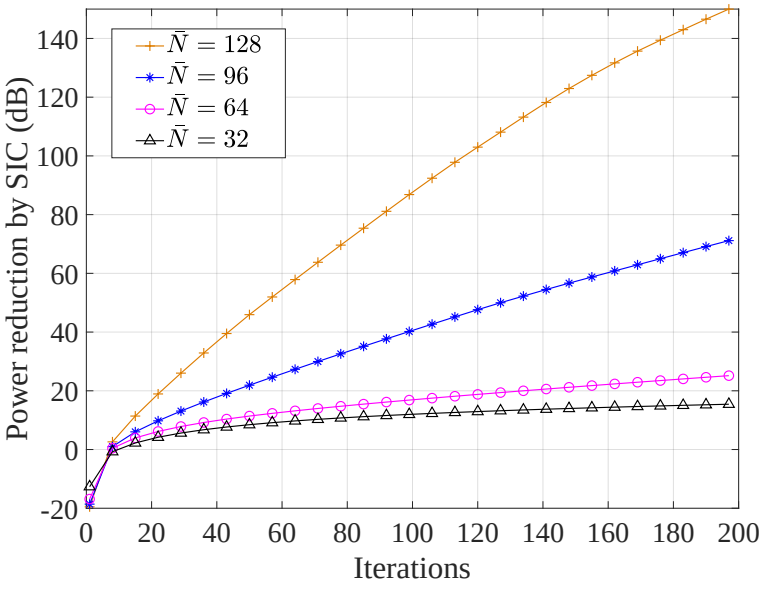


FIGURE 3.9: SIC performance with respect to different number of antenna elements at receiver, when assuming $P_{\mathrm{DL}} = 12$ W, $\bar{N}_{\mathrm{RF}} = 16$, $N = 32$, $N_{\mathrm{RF}} = 16$, $\varphi = 120^{\circ}$, $\kappa = 10$.

fairness constraints and stable sum rate among DL MSs by solving the optimization problem in (3.13) on every single channel realization, while the factorization of the full-digital precoder (e.g., PGD and HD-CVXR) introduces inter-user interference and destroys the equal rate constraint, leading to the fluctuation of performance among DL MSs.

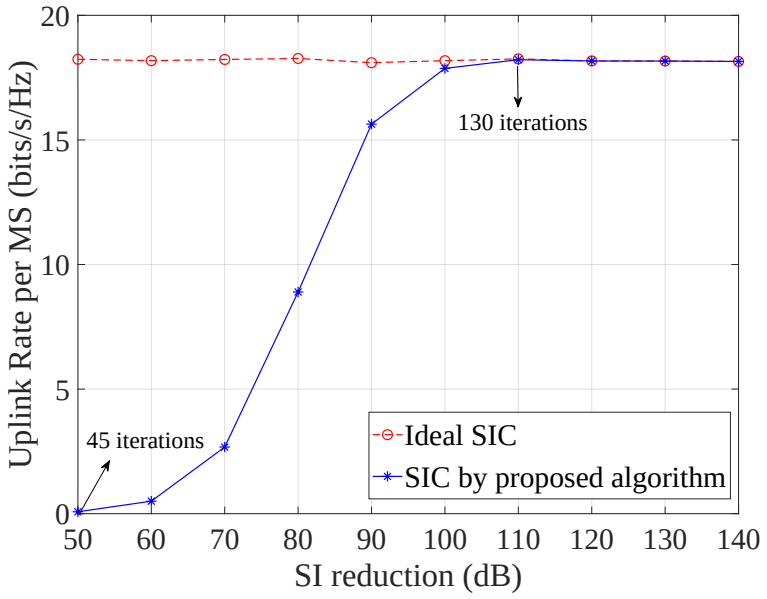


FIGURE 3.10: The uplink performance with the assumption of real and ideal SIC, when assuming $P_{\mathrm{DL}} = 12\ W$, $\bar{N} = 128$, $\bar{N}_{\mathrm{RF}} = 16$, $N = 32$, $N_{\mathrm{RF}} = 16$, $\varphi = 120^{\circ}$, $\kappa = 10$.

3.6.2 Performance of Self-Interference Cancellation

In Section 3.6.1, we assume the system with ideal SI suppression (i.e., $\xi = -100$ dB) provided by existed approaches such as [39, 72, 150, 161]. However, to the best of our knowledge, none of existed SIC methods can provide 100 dB SI reduction in FD hybrid beamforming MIMO systems. In this section, we assume that there is no other SIC methods (i.e., $\xi = 0$ dB) and all the SI reduction depends on our proposed method. Apart from the default values used in Section 3.6.1, for the LoS component of the SI channel, we assume that the distance r_{ij} between the i-th element of transmit array and the j-th element of receive array is set according to [106] and the default angle between transmit array and receive array is $\varphi = 120^{\circ}$. For the NLoS components, we assume that the channel has $L_n = 6$ paths and the complex gain α_l, AoA and AoD have the same features as the communication channels stated in Section 3.6.1.

In Fig. 3.9, we plot the SIC performance versus the number of iterations operated by the proposed SIC algorithm. It is observed that for a given transmit antenna array and a given number of iterations, more SI can be mitigated, when BS is equipped with more receive antenna elements. The reason behind is that as the number of receive antenna elements increases, the null space of $\boldsymbol{H}_{\mathrm{SI}}^{H}$ has a bigger rank. Therefore, the receiver is more capable of suppressing SI signals.

Fig. 3.10 studies the influence of the SI suppression provided by our proposed SIC algorithm on the desired UL transmissions, where ideal SI cancellation means that $\boldsymbol{r}_{\mathrm{SI}} = \boldsymbol{0}$ in (3.9) during signal detection. From Fig. 3.10, we can observe that to attain the same performance as the ideal system without SI, a practical system should provide

at least 110 dB SI reduction, which requires our proposed SIC method to operate 130 iterations. The insufficient SI suppression may lead to large quantization noise and make RA inefficient in FD systems. It is worth noting that in practice, the SI channel varies very slowly due to long coherence time, as both transmitter and receiver are fixed at BS. Hence, the SIC operations are only required to be executed once after a relatively long time period. Furthermore, an SIC updating process can be started at a nearly converged state, which allows a significantly shorter time than 130 iterations to reach the required SIC. Therefore, the proposed SIC algorithm is efficient for practical operation.

3.7 Chapter Summary and Conclusions

In this chapter, we have studied the RA problem in the large-scale MDD MIMO mmWave systems with the consideration of user fairness, SI suppression and hybrid precoder/-combiner design. To demonstrate the advantage of MDD, we have first considered the RA based on the unfair greedy algorithm in MU-SISO systems. The studies show that MDD mode has the potential to outperform HD modes, owing to that the MDD mode enables UL and DL to jointly share the available resource. Then, the suboptimal RA algorithms have been proposed for the MDD MIMO systems to maximize the sum-rate and simultaneously to meet the fairness requirements among DL and UL MSs. Furthermore, two approaches, namely the matrix factorization approach and DA, have been proposed for designing the hybrid precoder associated with the RA schemes. Our studies reveal that when the number of RF chains is relatively small, the DA outperforms the matrix factorization approach. By contrast, when the number of RF chains is relatively big, the matrix factorization approach is a better option for designing precoders. Furthermore, to address the SI problem in MDD systems, we have modeled the influence of SI and proposed an adaptive beamforming aided SIC algorithm embedded in combiner design. Our simulation results demonstrate that the proposed SIC algorithm enable the RA operations free from SI with appropriate antenna deployment, while imposing no impact on the desired UL transmissions.

Chapter 4

MDD for Mitigation of Channel Aging Problem in Massive MIMO

This chapter investigates the performance of multicarrier-division duplex (MDD) in high-mobility massive multiple-input multiple-output (mMIMO) scenarios, showing the MDD's potential for mitigating the channel aging problem. Furthermore, to facilitate the robust communications over fast time-varying channels, the custom-tailored frame structures and channel predictors are designed and studied. Additionally, the performance of MDD and time-division duplex (TDD) systems is compared in high mobility communications scenarios.

4.1 Introduction

The mMIMO has been regarded as one of the most promising technologies for significantly improving the spectral efficiency (SE) in the 5G wireless systems [45]. However, due to the large number of antennas equipped at base station (BS), the acquisition of channel state information (CSI) becomes highly challenging. As the result, TDD is deemed to be the only feasible mode in practical mMIMO systems to date, which exploits the channel reciprocity to acquire downlink (DL) CSI from uplink (UL) pilots [162]. As its counterpart, in frequency-division duplex (FDD) systems, as elaborated in Section 1.1.1, the overhead for training/feedback is proportional to the size of the antenna arrays at BS, resulting in less resources for sending desired data in one coherence block [10]. On the other hand, with the growing demand on the mobile services operated in ground vehicles, high-speed trains, unmanned aerial vehicle (UAVs), etc., the performance of TDD systems is no longer dominated by pilot contamination but the channel aging problem. In other words, due to fast time-varying, the channel at the time when it is estimated via UL training may be very different from that at the time when it is applied for DL preprocessing and UL detection. Moreover, the UL/DL

switching intervals required by TDD systems exacerbate the situation. This CSI mismatch may lead to significant degradation of the SE of both UL and DL.

In order to support the robust communications over fast time-varying channels, channel prediction based on the Wiener filter or Kalman filter has been studied. In [74], the authors considered the channel aging problem in mMIMO systems and applied the Wiener predictor (WP) to predict future CSI. The results show that using channel prediction can partially relieve the channel aging effect and render an increase of sum rate. The authors of [14] analyzed the effect of WP for time-varying channels on the system performance. However, in these two papers, the one-step WP relies on the latest channels estimated using the UL pilots that are always available in the front of the symbol to be predicted. Unfortunately, in practical TDD systems, UL pilots are unavailable during DL transmission and hence, the accuracy of channel prediction degrades with time, causing the performance degradation. For this sake, in [15], the authors considered a more practical scenario, where no UL pilots are sent during the Kalman filter assisted channel prediction for DL transmission. The results show that the Kalman predictor is capable of improving performance when compared with the case without prediction. However, due to the time-varying effect, the achievable DL rate reduces quickly with time as the result of the fact that the predicted CSI becomes less and less accurate. Recently in [16], a learning-based approach was introduced, showing that it can outperform the conventional non-linear Kalman predictor in dealing with the channel aging problem. Again, the prediction accuracy of CSI deteriorates as the interval between the channel being predicted and the pilot symbols becomes larger.

In contrast to the half duplex (HD) mode of TDD, in-band full duplex (IBFD) has the inherent advantage for solving the channel aging problem. This is because in IBFD systems, DL and UL occur concurrently, which enables to acquire the latest CSI during data transmission without invoking DL/UL switching. However, to the best of our knowledge, there are no open references which have considered the IBFD-relied transmission design for the multicarrier mMIMO systems communicating over fast time-varying channels. One conceivable reason may be that the SI problem in IBFD-relied systems is still intractable. This becomes even more challenging in the high-mobility communication scenarios. Note that, although channel aging is not considered, in [163], the channel acquisition relying on a hybrid IBFD and TDD mode was demonstrated to be efficient. With this hybrid scheme, partial DL transmissions are activated during UL training, leading to an increased SE. Moreover, the effect of self-interference cancellation (SIC) on the performance of channel estimation (CE) was studied with the IBFD systems, showing that if SIC is insufficient, IBFD is unable to outperform the conventional TDD. Similar observations were also obtained in [4].

To overcome the weakness of both IBFD and TDD over fast time-varying communication channels, inspired by [68], we propose the MDD associated with the dedicated frame design so as to relieve the channel aging problem, which constitutes the main

motivation of this chapter. The rationale can be briefly explained as follows. On the one hand, in MDD systems, both DL and UL transmissions can occur concurrently within the same frequency band but on different subcarriers. Hence, when needed, UL pilots can be continuously (or frequently) transmitted during DL transmissions, as shown in Figure 1.6. Therefore, CSI can be updated in time and does not become outdated as in TDD systems. On the other hand, MDD is capable of circumventing the stringent requirement for SIC in IBFD systems. According to Section 1.1.2.1, in practical IBFD systems, in addition to the SIC in the propagation- and analog-domain, the digital-domain SIC has to cancel the main SI signal component by at least 30 dB. Achieving this is very power-consuming and technically demanding, especially for the relatively small-sized MSs, e.g., UAVs and smartphones. By contrast, in the MDD systems, this amount of SIC can be attained at nearly no extra cost of system resources, thanks to the embedded fast Fourier transform (FFT) operation [4], which allows to ideally separate UL signals from DL signals in the digital-domain. Owing to the aforementioned merits of MDD, it is important to investigate the MDD's implementation in fast time-varying scenarios, and to unlock its advantages over the conventional TDD and IBFD systems. Therefore, our main contributions of this chapter are summarized as follows:

Firstly, a model for the MDD-based mMIMO systems communicating over time-varying channels is presented. To alleviate the channel aging problem, we propose two types of frame structures dedicated for MDD, which enable UL pilots (or UL data transmissions) to occur concurrently with DL transmissions, so that CSI can be promptly updated whenever needed. For comparison, the corresponding TDD frame structures related to the 3GPP standards [136, 164] as well as their modified forms for supporting different mobility scenarios are introduced.

Secondly, to operate with the proposed frame structures, we introduce two finite impulse response WPs that consider both CE and residual SI errors, namely the general WP and the decision-directed WP (DD-WP), dedicated to MDD systems. To be more specific, the general WP directly predicts the time-domain CSI based on the observations collected from the UL CE. By contrast, the DD-WP leverages both the UL pilots and the detected UL data symbols to predict the frequency-domain CSI. Along with these WPs, the impact of residual SI and the order of WPs on the performance of MDD systems are studied and compared.

Thirdly, by assuming the zero-forcing (ZF) precoding for DL transmission and the maximum ratio combining (MRC) for UL detection, the closed-form expressions for approximating the lower bounded average sum rates of both the TDD and MDD systems are derived, when the proposed WPs are operated with two general types of frame structures.

Our studies and simulation results show that TDD systems suffer from the channel aging problem, whose performance degrades significantly with time, when channels vary

fast. By contrast, the MDD systems endowed with the FD capability can effectively mitigate the channel aging problem and hence, are capable of significantly outperforming their TDD counterparts, when communicating over fast time-varying channels. Moreover, the studies demonstrate that the SIC in FD systems plays a paramount role in channel prediction. In the case of imperfect SIC, MDD becomes more competent than IBFD for operation in high-mobility communications scenarios.

The rest of the chapter is organized as follows. In Section 4.2, the model for the MDD-based system is described. In Section 4.3, two general frame structures for TDD and MDD systems are introduced. Section 4.4 presents the principles of CE and two approaches for channel prediction. Section 4.5 analyzes the lower bounded average sum rates of MDD and TDD systems. Section 4.6 provides the performance results to compare TDD, MDD and IBFD systems. Finally, conclusions are drawn in Section 4.7.

4.2 System Model

Consider a single-cell mMIMO orthogonal frequency division multiplexing (OFDM) system having an BS equipped with N antennas and D single-antenna MSs randomly distributed in the cell. The system is operated in the MDD mode, allowing DL and UL to communicate concurrently in the same frequency band but on different subcarriers. Following the concept of mMIMO, the number of antennas at BS is assumed to be much larger than the number of served MSs [10], i.e., $N \gg D$. Furthermore, based on the principles of MDD [4], subcarriers are divided into two mutually exclusive subsets, namely a DL subcarrier subset $\mathcal{M}$ with M subcarriers, and a UL subcarrier subset $\bar{\mathcal{M}}$ with $\bar{M}$ subcarriers, i.e., $|\mathcal{M}| = M$ and $|\bar{\mathcal{M}}| = \bar{M}$. The total number of subcarriers is expressed as $M_{\text{sum}} = M + \bar{M}$. Similar to the other works on RA, e.g., [144, 165, 166], the DL/UL subcarrier allocation (SA) is carried out at BS based on the CSI and the requirements of MSs' quality of services. The allocation results are informed to MSs through control channels during the initial access procedure. Below we assume that the DL/UL SA results have been obtained, and that each MS knows its assigned subcarriers. The feasible SA algorithm in MDD systems can be found in Section 3.4.

4.2.1 Channel Model

We assume that the channels are frequency-selective in terms of the M_{sum} subcarriers but each subcarrier experiences flat fading. To consider the influence of channel aging, we assume that the channel coefficients do not change within one OFDM symbol, but vary from one symbol to the next. Therefore, the time-domain CSI of the L-tap channel between the d-th MS and the n-th BS antenna over the i-th OFDM symbol duration can

be expressed as

$$\boldsymbol{g}_{n,d}[i] = [g_{n,d}[i,1], ..., g_{n,d}[i,l], ..., g_{n,d}[i,L]]^T, \tag{4.1}$$

where $g_{n,d}[i,l] = \nu_{n,d}[i,l]\sqrt{\beta_d}$ and $\nu_{n,d}[i,l] \sim \mathcal{CN}(0,1/L)$ is the small-scale fading, while β_d represents the large-scale fading, which only depends on the distance between MS d and BS and is assumed to remain constant over one communication frame. For any user-antenna pair, the channels of different taps are assumed to be independent. Hence we have $\boldsymbol{R}_g^d = \mathbb{E}\left[\boldsymbol{g}_{n,d}[i]\boldsymbol{g}_{n,d}^H[i]\right] = \frac{\beta_d}{L}\boldsymbol{I}_L$.

According to the principles of OFDM [68], the frequency-domain CSI $\boldsymbol{h}_{n,d}[i]$ over the i-th OFDM symbol duration can be obtained as

$$\boldsymbol{h}_{n,d}[i] = \boldsymbol{F}\boldsymbol{\Psi}\boldsymbol{g}_{n,d}[i], \tag{4.2}$$

where $\boldsymbol{F} \in \mathbb{C}^{M_{\text{sum}} \times M_{\text{sum}}}$ is the FFT matrix with $(\boldsymbol{F})_{p,q} = \frac{1}{\sqrt{M_{\text{sum}}}} e^{-j2\pi(p-1)(q-1)/M_{\text{sum}}}$, $\boldsymbol{\Psi} \in \mathbb{C}^{M_{\text{sum}} \times L}$ is constructed by the first L columns of $\boldsymbol{I}_{M_{\text{sum}}}$. Furthermore, the DL subchannels $\boldsymbol{h}_{n,d}^{\text{DL}}[i]$ and UL subchannels $\boldsymbol{h}_{n,d}^{\text{UL}}[i]$ between the n-th antenna at BS and the d-th MS over the i-th OFDM symbol can be obtained from (4.2), which can be expressed as

$$\boldsymbol{h}_{n,d}^{\text{DL}}[i] = \boldsymbol{\Phi}_{\text{DL}}\boldsymbol{h}_{n,d}[i] = [h_{n,d}[i,1], ..., h_{n,d}[i,m], ..., h_{n,d}[i,M]]^T, \tag{4.3}$$

$$\boldsymbol{h}_{n,d}^{\text{UL}}[i] = \boldsymbol{\Phi}_{\text{UL}}\boldsymbol{h}_{n,d}[i] = [h_{n,d}[i,1], ..., h_{n,d}[i,\bar{m}], ..., h_{n,d}[i,\bar{M}]]^T, \tag{4.4}$$

where $\boldsymbol{\Phi}_{\text{DL}} = \boldsymbol{I}_{M_{\text{sum}}}^{(\mathcal{M},:)}$ and $\boldsymbol{\Phi}_{\text{UL}} = \boldsymbol{I}_{M_{\text{sum}}}^{(\bar{\mathcal{M}},:)}$ are the mapping matrices, constructed from $\boldsymbol{I}_{M_{\text{sum}}}$ by choosing its rows corresponding to the particular subcarriers assigned to DL and UL, respectively. Due to the short separation between transmitter and receiver, the SI channels are assumed to be flat fading and identical to all subcarriers, and also quasi-static within one frame. Specifically, the SI channels at BS and MSs follow the distributions of $(\boldsymbol{H}_{\text{SI}})_{i,j} \sim \mathcal{CN}(0,1)$, where $\boldsymbol{H}_{\text{SI}} \in \mathbb{C}^{N \times N}$, and $h_{\text{SI}} \sim \mathcal{CN}(0,1)$, respectively.

4.2.2 Channel Aging

As we stated in Section 4.1, the relative mobility between BS and MSs leads to time-varying channels, causing that the channel varies between the time when it is estimated and the time when the estimated CSI is applied for transmissions. This is the well-known channel aging problem. To model the channel aging, we introduce the Jakes model, which has the normalized discrete-time autocorrelation function at the BS given by [16]

$$R[k] = J_0(2\pi f_D T_s |k|), \tag{4.5}$$

where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, T_s is the OFDM symbol duration, f_D is the maximum Doppler frequency shift, and $|k|$ is the delay in terms of the number of symbols.

Furthermore, for the sake of analyzing MDD frame structures with WP and DD-WP in the following sections, we adopt the autoregressive model of order 1, denoted as AR(1), for approximating the temporal correlation between adjacent symbols. In this case, the channel coefficient $\boldsymbol{g}_{n,d}[i]$ of (4.1) can be formulated as [153]

$$\boldsymbol{g}_{n,d}[i] = \alpha \boldsymbol{g}_{n,d}[i-1] + \boldsymbol{v}_{n,d}[i], \tag{4.6}$$

where $\alpha = J_0(2\pi f_D T_s)$, which is assumed to be perfectly known at BS, $\boldsymbol{v}_{n,d}[i]$ is an uncorrelated channel error vector due to channel aging, which can be modeled as a stationary Gaussian random process with i.i.d. entries and the distribution of $\mathcal{CN}\left(\boldsymbol{0}, (1-\alpha^2)\boldsymbol{R}_g^d\right)$. According to the channel model of (4.6), $\boldsymbol{g}_{n,d}[i]$ is a stationary Markov random process, and has the auto-correlation of $\mathbb{E}\left[\boldsymbol{g}_{n,d}[i-q]\boldsymbol{g}_{n,d}^H[i-p]\right] = \alpha^{|p-q|}\boldsymbol{R}_g^d$. Moreover, according to (4.3) and (4.4), we can obtain the DL subchannels satisfying

$$h_{n,d}[i,m] = \alpha h_{n,d}[i-1,m] + \underbrace{\boldsymbol{\psi}_m \boldsymbol{v}_{n,d}[i]}_{\tilde{v}_{n,d}^m[i]}, \tag{4.7}$$

where $\boldsymbol{\psi}_m = \boldsymbol{I}_{\text{sum}}^{(m,:)}\boldsymbol{F}\boldsymbol{\Psi}$ and $\tilde{v}_{n,d}^m[i] \in \mathcal{CN}\left(0, (1-\alpha^2)R_h^d\right)$ with $R_h^d = \mathbb{E}\left[h_{n,d}[i,m]h_{n,d}^H[i,m]\right] = \beta_d / M_{\text{sum}}$. Similarly, the UL subchannels $h_{n,d}[i,\bar{m}]$ can be generated.

4.2.3 Downlink Transmission

For DL transmission, the signal $\boldsymbol{s}_{\text{DL}}[i,m] \in \mathbb{C}^{N\times 1}$ transmitted by BS on the m-th subcarrier in the i-th OFDM symbol duration is

$$\boldsymbol{s}_{\text{DL}}[i,m] = \sqrt{p_{\text{DL}}}\boldsymbol{F}^{\text{ZF}}[i,m]\boldsymbol{x}[i,m], \tag{4.8}$$

where p_{DL} is the power budget per DL subcarrier and the total BS transmit power is $P_{\text{DL}} = p_{\text{DL}}M$, $\boldsymbol{x}[i,m] = [x_1[i,m], ..., x_D[i,m]]^T$ contains the information symbols normalized to satisfy $\mathbb{E}\left[\boldsymbol{x}[i,m]\boldsymbol{x}[i,m]^H\right] = \boldsymbol{I}_D$. We assume that the ZF precoding is applied at transmitter, giving $\boldsymbol{F}^{\text{ZF}}[i,m] = \boldsymbol{H}_{\text{DL}}^H[i,m]\left(\boldsymbol{H}_{\text{DL}}[i,m]\boldsymbol{H}_{\text{DL}}^H[i,m]\right)^{-1}$, where $\boldsymbol{F}^{\text{ZF}}[i,m] = [\boldsymbol{f}_1[i,m], ..., \boldsymbol{f}_D[i,m]]$ with the constraint of $\|\boldsymbol{f}_d[i,m]\|_2 = 1/\sqrt{D}$, such that $\mathbb{E}\left[\|\boldsymbol{s}_{\text{DL}}[i,m]\|_2^2\right] = p_{\text{DL}}$, and $\boldsymbol{H}_{\text{DL}}[i,m] = [\boldsymbol{h}_1[i,m], ..., \boldsymbol{h}_d[i,m], ...\boldsymbol{h}_D[i,m]]^H$, where $\boldsymbol{h}_d[i,m] = [h_{1,d}[i,m], h_{2,d}[i,m], ..., h_{N,d}[i,m]]^T$. It is noteworthy that the analysis in this chapter is based on the equal PA among MSs. Intuitively, an appropriate PA method, e.g., that in Section 3.4, will increase the achievable sum rates, owing to both user and subcarrier diversity being exploited. The received signal at the d-th MS can be expressed as

$$\begin{aligned} y_d[i,m] = \sqrt{p_{\text{DL}}}\boldsymbol{h}_d^H[i,m]\boldsymbol{f}_d[i,m]x_d[i,m] + \sqrt{p_{\text{DL}}}\sum_{k\neq d}\boldsymbol{h}_d^H[i,m]\boldsymbol{f}_k[i,m]x_k[i,m] + z_d^{\text{SI}}[i] \\ + z_d[i,m], \end{aligned} \tag{4.9}$$

where $z_d^{\mathrm{SI}}[i] \sim \mathcal{CN}(0, \xi_{\mathrm{MS}} p_{\mathrm{UL}} \bar{M})$ denotes the residual SI[1] signal imposed by the UL transmission of MS d with p_{UL} being the power budget per UL subcarrier, such that the total MS transmit power is $P_{\mathrm{UL}} = p_{\mathrm{UL}} \bar{M}$. Note that ξ_{MS} in (4.9) and ξ_{BS} in (4.10) denote the SIC capability that is provided by the existing SIC techniques, such as, the antenna circulator [21], spatial beamforming [69], dual-port polarized antenna [25], multi-tap RF canceller [39], etc., implemented in propagation- and analog-domain, as well as the FFT operation operated in digital-domain. In (4.9), the covariance of $z_d^{\mathrm{SI}}[i]$ is calculated in detail in Appendix C. In addition, $z_d[i,m] \sim \mathcal{CN}(0, \sigma^2)$ denotes the additive white Gaussian noise. Note also that, as we mentioned in Equation 1.2.1.1, since the inter-MS interference (IMI) and the desired DL signal are located on the mutually orthogonal subcarriers, with the aid of approriate power control and large-scale fading, the IMI can be easily suppressed to the effective dynamic level of ADC at recevier and then canceled by FFT operation in digital domain. To avoid the deviation of the core of this chapter, we assume that the interference caused by the other MSs' UL transmission on the DL reception is perfectly suppressed.

4.2.4 Uplink Transmission

For the UL transmission, the signal received at BS over the $\bar{m}$-th UL subcarrier and i-th OFDM symbol duration can be expressed as

$$\boldsymbol{s}_{\mathrm{UL}}[i,\bar{m}] = \sum_{d=1}^{D} \boldsymbol{h}_d[i,\bar{m}] \sqrt{p_{\mathrm{UL}}} x_d[i,\bar{m}] + \boldsymbol{z}^{\mathrm{SI}}[i] + \boldsymbol{z}[i,\bar{m}], \tag{4.10}$$

where $\boldsymbol{h}_d[i,\bar{m}] = [h_{1,d}[i,\bar{m}], h_{2,d}[i,\bar{m}], ..., h_{N,d}[i,\bar{m}]]^T$, $\boldsymbol{z}[i,\bar{m}] \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_N)$. $\boldsymbol{z}^{\mathrm{SI}}[i]$ is the SI generated by BS transmissions, which is modeled as $\boldsymbol{z}^{\mathrm{SI}}[i] \sim \mathcal{CN}(\boldsymbol{0}, \xi_{\mathrm{BS}} p_{\mathrm{DL}} M \boldsymbol{I}_N)$. The covariance of $\boldsymbol{z}^{\mathrm{SI}}[i]$ is derived in detail in Appendix C. In this chapter, we assume that the MRC is used for signal detection. Then we have $\boldsymbol{W}^{\mathrm{MRC}}[i,\bar{m}] = \boldsymbol{H}_{\mathrm{UL}}[i,\bar{m}]$, where $\boldsymbol{H}_{\mathrm{UL}}[i,\bar{m}] = [\boldsymbol{h}_1[i,\bar{m}], ..., \boldsymbol{h}_D[i,\bar{m}]]$. Let us write $\boldsymbol{W}^{\mathrm{MRC}}[i,\bar{m}] = [\boldsymbol{w}_1[i,\bar{m}], ..., \boldsymbol{w}_D[i,\bar{m}]]$. Then the decision variable for MS d can be expressed as

$$\begin{aligned} y_d[i,\bar{m}] = {} & \sqrt{p_{\mathrm{UL}}} \boldsymbol{w}_d^H[i,\bar{m}] \boldsymbol{h}_d[i,\bar{m}] x_d[i,\bar{m}] + \sqrt{p_{\mathrm{UL}}} \sum_{k \neq d} \boldsymbol{w}_d^H[i,\bar{m}] \boldsymbol{h}_k[i,\bar{m}] x_k[i,\bar{m}] \\ & + \boldsymbol{w}_d^H[i,\bar{m}] \boldsymbol{z}^{\mathrm{SI}}[i] + \boldsymbol{w}_d^H[i,m] \boldsymbol{z}[i,\bar{m}], \end{aligned} \tag{4.11}$$

where the second term at the right-hand side is multiuser interference, while the third term is due to SI.

[1]The residual SI consists of the combined effect of the additive noise introduced by automatic gain control (AGC), non-linearity of analog-to-digital converter (ADC) and the phase noise generated by oscillator due to radio frequency (RF) imperfection [148].

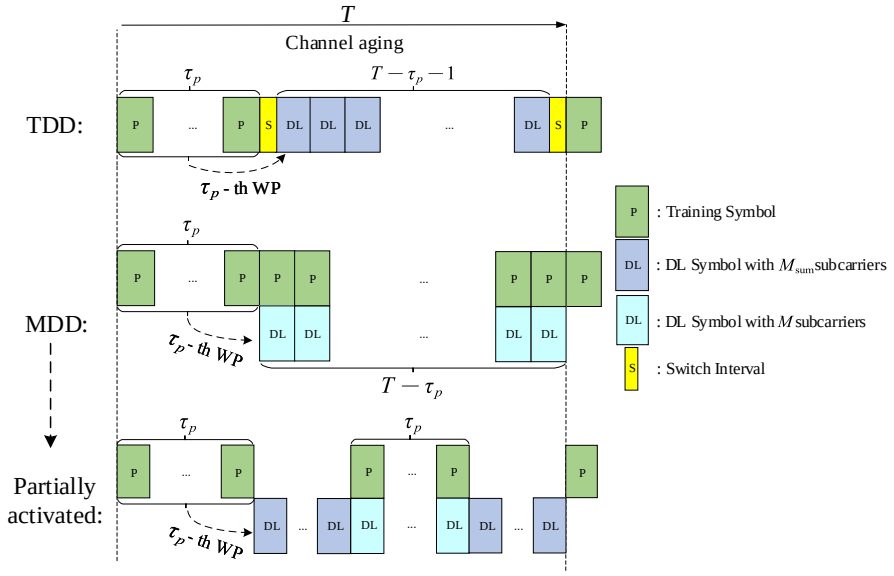


FIGURE 4.1: Type I frame structure for the conventional TDD and proposed MDD systems.

4.3 Frame Structure Design

Due to the effect of channel aging, channel coefficients may vary from symbol to symbol, leading to poor system performance in high-mobility communication scenarios [74]. To overcome the time-varying fading problem, in conventional TDD mode, channel prediction is usually implemented after the training phase to reduce CSI error. However, when the relative velocity between transmitter and receiver increases, the channel becomes less correlated, increasing the prediction error, which in turn renders the later detected symbols less reliable. One way to mitigate this problem is to employ extra training symbols for updating the CSI during data transmission. Following this observation, if the channel aging problem becomes severe, more training symbols are inserted, which results in more frequent switching between UL training and DL data transmission, as well as the decrease of useful data rate. To this end, our main motivation of this chapter is to demonstrate that the MDD mode is capable of efficiently solving the channel aging problem in fast fading communication scenarios. For demonstration and comparison, in the following subsections, we will present two types of frame structures for coping with the channel aging problem and also for analyzing the performance of MDD systems in terms of sum-rate.

4.3.1 Frame Structure: Type I

In this subsection, we consider a simple scenario, where the frame structure only includes UL training and DL transmission, as shown in Fig. 4.1. The length of one frame[2] is assumed to be equal to T OFDM symbol durations.

4.3.1.1 TDD

As shown in Fig. 4.1, in the TDD mode, a frame consists of τ_p training symbols transmitted by UL and $T - \tau_p - 1$ symbols used for DL transmission. The switching from UL training to DL data transmission or from DL data transmission to the next training phase incurs a fixed cost, which is assumed to be half of an OFDM symbol.[3] To cope with channel aging, the channel prediction is implemented by a τ_p-th order WP[4] after the training phase, as shown in Fig. 4.1, and all the precoding/detection are operated at BS based on the predicted CSI. It is noteworthy that in practice, the τ_p training symbols can be first disassembled into several groups, which are then evenly distributed over one frame, so as to relieve the effect of accumulated prediction errors. However, the accompany of the increased switching intervals results in the reduced time for DL transmission and hence the decreased system's efficiency.

4.3.1.2 MDD

In the context of MDD, as UL/DL transmissions are on different subcarriers, CSI estimation and prediction can be implemented in the FD mode. That is to say, UL pilots can be kept active to update CSI in real time within one OFDM symbol. In detail, as shown in Fig. 4.1, UL training is always activated and hence, the channels for DL transmission can be predicted by a WP based on the newest τ_p pilots. However, in low-speed scenarios, the excessive transmission of UL pilots inevitably increases energy consumption and system complexity, which is unnecessary as the channels change slowly. Moreover, in the case where only the DL transmission is needed, continuous UL training on $\bar{M}$ subcarriers leads to only M of the M_{sum} subcarriers being used for the DL transmission. Therefore, in practice, the pilots within one MDD frame can be partially transmitted (as illustrated by the example presented in Fig. 4.1) so that some UL subcarriers can be 'borrowed' for DL data transmission to reduce the overhead of SI

[2]Note that one frame defined in this chapter may include several subframes used in 4G long-term evolution (LTE) systems or slots in 5G new radio (NR) systems, where each subframe or slot contains 14 OFDM symbols. Specifically in 4G LTE, one frame consists of 10 subframes and two of them are used for training [136]. The frame structure in 5G NR is more flexible than that in 4G LTE [164].

[3]The length of the switching period in LTE can be a different value, such as 1, 2, 3, 4, 9 or 10 OFDM symbols [136]. We choose 0.5 to minimize the influence of the switching interval on the evaluation of TDD system's performance.

[4]To ease analysis and facilitate comparison, only WP is considered in this chapter. The Kalman predictor and the state-of-the-art deep learning approaches will be studied in our future work.

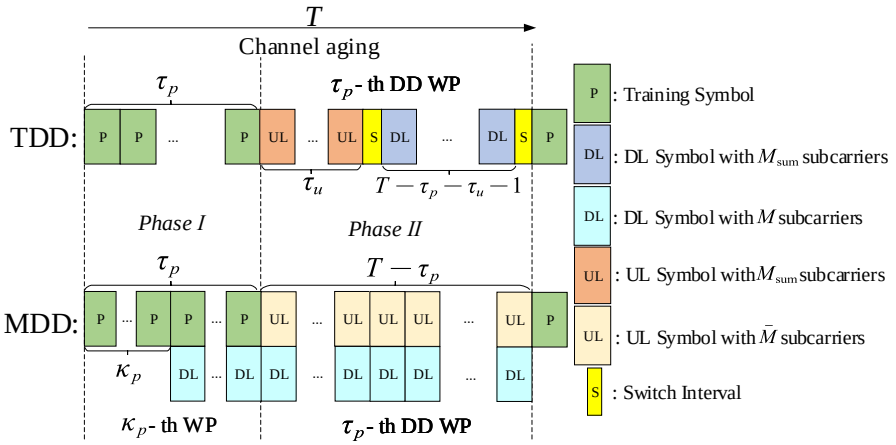


FIGURE 4.2: Type II frame structure for conventional TDD and proposed MDD systems.

mitigation. In Section 4.6, the TDD/MDD frame structures with different distributions of pilots will be investigated.

4.3.2 Frame Structure: Type II

A more general frame structure consisting of the UL training followed by both UL transmission and DL transmission is considered for both TDD and MDD, as shown in Fig. 4.2.

4.3.2.1 TDD

In the TDD mode, as shown in Fig. 4.2, τ_p pilot symbols are transmitted in Phase I for initializing the transmission. During Phase II, the detected UL data symbols and possibly a part of pilots are used to predict channels by a τ_p-th order DD-WP. With this setting, during the UL transmission, the channel knowledge can be continuously updated with the DD prediction principle. By contrast, during the DL transmission, channel updating is unavailable as there is no UL transmission. Consequently, the transmissions of the $T - \tau_p - \tau_u - 1$ DL symbols are based on the outdated CSI, which is similar to the situation with the Type I frame structure.

4.3.2.2 MDD

As shown in Fig. 4.2, the Type II frame structure in MDD systems can be designed in a very efficient way. Firstly, for fair comparison with TDD, in Phase I, τ_p pilot symbols are still transmitted for the implementation of a τ_p-th order DD-WP in Phase II. However, different from TDD, during the transmission of pilot symbols, DL symbols can also be

transmitted with the aid of the CSI predicted by a κ_p-th order WP, where $1 \leq \kappa_p \leq \tau_p$. In the extreme case, when $\kappa_p = 1$, $\tau_p - 1$ more DL symbols than TDD can be transmitted in Phase I, which may significantly increase the SE. During Phase II, both DL and UL can transmit $T - \tau_p$ symbols. Furthermore, a continuously updated τ_p-th order DD-WP can be implemented to predict the channels using the pilots and the detected UL symbols.

4.4 Channel Estimation and Prediction in MDD systems

In this section, we study the CE and prediction in MDD systems, when the proposed frame structures in Section 4.3 are introduced.

4.4.1 Channel Estimation

Let us assume that all MSs synchronously transmit their frequency-domain pilot sequences (FDPS) over $\bar{M}$ UL subcarriers. Specifically, the FDPS transmitted by the d-th MS is expressed as $\boldsymbol{x}_{p,d} = \left[x_{p,d}[1], ..., x_{p,d}[\bar{m}], ..., x_{p,d}[\bar{M}]\right]^T$. Then, the received signal at the n-th antenna of BS corresponding to the i-th symbol is given by

$$\begin{aligned} \boldsymbol{y}_{p,n}[i] &= \sqrt{p_{\mathrm{UL}}} \sum_{d=1}^{D} \boldsymbol{X}_{p,d} \boldsymbol{h}_{n,d}^{\mathrm{UL}}[i] + \boldsymbol{z}_n^{\mathrm{SI}}[i] + \boldsymbol{z}_{p,n}[i] \\ &= \sqrt{p_{\mathrm{UL}}} \sum_{d=1}^{D} \boldsymbol{X}_{p,d} \boldsymbol{\Phi}_{\mathrm{UL}} \boldsymbol{F} \boldsymbol{\Psi} \boldsymbol{g}_{n,d}[i] + \boldsymbol{z}_n^{\mathrm{SI}}[i] + \boldsymbol{z}_{p,n}[i], \end{aligned} \tag{4.12}$$

where $\boldsymbol{X}_{p,d} = \mathrm{diag}\left\{\boldsymbol{x}_{p,d}\right\}$, $\boldsymbol{z}_{p,n}[i] \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_{\bar{M}})$. Note that, in MDD systems, BS may experience SI during UL training, if DL symbols are transmitted at the same time. For example, in the frame structure as shown in Fig. 4.1, when $1 \leq i \leq \tau_p$, we have $\boldsymbol{z}_n^{\mathrm{SI}}[i] = 0$, as there is no DL transmissions. By contrast, when $\tau_p < i \leq T$, $\boldsymbol{z}_n^{\mathrm{SI}}[i] \sim \mathcal{CN}\left(0, \xi_{\mathrm{BS}} p_{\mathrm{DL}} M \boldsymbol{I}_{\bar{M}}\right)$.

According to the CE method proposed in Section 2.4, the noisy observation of $\boldsymbol{g}_{n,d}[i]$ can be obtained as

$$\tilde{\boldsymbol{y}}_{p,n}^{d}[i] = \boldsymbol{P}_d^H \boldsymbol{y}_{p,n}[i] = \frac{\sqrt{p_{\mathrm{UL}}}\bar{M}}{M_{\mathrm{sum}}} \boldsymbol{g}_{n,d}[i] + \boldsymbol{P}_d^H (\boldsymbol{z}^{\mathrm{SI}}[i] + \boldsymbol{z}_{p,n}[i]). \tag{4.13}$$

Correspondingly, the MMSE estimate to $\boldsymbol{g}_{n,d}[i]$ is given by

$$\hat{\boldsymbol{g}}_{n,d}[i] = \frac{\frac{\sqrt{p_{\mathrm{UL}}}\bar{M}}{M_{\mathrm{sum}}} \frac{\beta_d}{L}}{\frac{p_{\mathrm{UL}}\bar{M}^2}{M_{\mathrm{sum}}^2} \frac{\beta_d}{L} + \frac{\xi_{\mathrm{BS}} p_{\mathrm{DL}} M \bar{M}}{M_{\mathrm{sum}}} + \frac{\bar{M}}{M_{\mathrm{sum}}} \sigma^2} \tilde{\boldsymbol{y}}_{p,n}^{d}[i]. \tag{4.14}$$

4.4.2 Channel Prediction

4.4.2.1 General Wiener Predictor

In this subsection, we focus on the prediction of $\boldsymbol{g}_{n,d}[i+1]$ based on the current and previous received training signals, which are expressed as $\bar{\boldsymbol{y}}_{p,n}^{d}[i] = \left[\tilde{\boldsymbol{y}}_{p,n}^{d,H}[i], \tilde{\boldsymbol{y}}_{p,n}^{d,H}[i-1], ..., \tilde{\boldsymbol{y}}_{p,n}^{d,H}[i+1-\tau_p]\right]^H$, where $\tilde{\boldsymbol{y}}_{p,n}^{d}[i]$ is given by (4.13). We first assume a τ_p-th order WP, with the weights $\boldsymbol{V}_{n,d} = \left[\boldsymbol{V}_{n,d}^{1}, ..., \boldsymbol{V}_{n,d}^{q}, ..., \boldsymbol{V}_{n,d}^{\tau_p}\right]$, $\boldsymbol{V}_{n,d}^{q} \in \mathbb{C}^{L\times L}$. Then, the predictor can be formulated as

$$\check{\boldsymbol{g}}_{n,d}^{\mathrm{WP}}[i+1] = \sum_{q=1}^{\tau_p} \boldsymbol{V}_{n,d}^{q} \tilde{\boldsymbol{y}}_{p,n}^{d}[i+1-q] = \boldsymbol{V}_{n,d} \bar{\boldsymbol{y}}_{p,n}^{d}[i]. \tag{4.15}$$

According to the principles of Wiener filter [74, 167], the solution to $\boldsymbol{V}_{n,d}$ is

$$\boldsymbol{V}_{n,d} = \boldsymbol{R}_{g\bar{y}}^{d}[1] \left(\boldsymbol{R}_{\bar{y},i}^{d}[0]\right)^{-1}, \tag{4.16}$$

where the cross-correlation between the real channel of the next symbol and the observations is

$$\boldsymbol{R}_{g\bar{y}}^{d}[1] = \mathbb{E}\left[\boldsymbol{g}_{n,d}[i+1]\bar{\boldsymbol{y}}_{p,n}^{d,H}[i]\right] = \frac{\sqrt{p_{\mathrm{UL}}}\bar{M}}{M_{\mathrm{sum}}}[\boldsymbol{\delta}(\tau_p, \alpha) \otimes \boldsymbol{R}_{g}^{d}], \tag{4.17}$$

where $\boldsymbol{\delta}(\tau_p, \alpha) = [\alpha, \alpha^2, ..., \alpha^{\tau_p}]$. In (4.16), the autocorrelation matrix of the training signals is

$$\boldsymbol{R}_{\bar{y},i}^{d}[0] = \mathbb{E}\left[\bar{\boldsymbol{y}}_{p,n}^{d}[i]\bar{\boldsymbol{y}}_{p,n}^{d,H}[i]\right] = \boldsymbol{\xi}(\tau_p, \alpha) \otimes \left(\frac{p_{\mathrm{UL}}\bar{M}^2}{M_{\mathrm{sum}}^2}\boldsymbol{R}_{g}^{d}\right) + \frac{\bar{M}\sigma^2}{M_{\mathrm{sum}}}\boldsymbol{I}_{L\tau_p} + \boldsymbol{\Xi}_{\mathrm{SI}}[i], \tag{4.18}$$

where

$$\boldsymbol{\xi}(\tau_p, \alpha) = \begin{bmatrix} 1 & \alpha & \cdots & \alpha^{\tau_p-1} \\ \alpha & 1 & \cdots & \alpha^{\tau_p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{\tau_p-1} & \alpha^{\tau_p-2} & \cdots & 1 \end{bmatrix}, \tag{4.19}$$

It can be observed that in the Type I frame structure for fast-fading channels, the amount of SI is dependent on the observation time, which leads to

$$\boldsymbol{\Xi}_{\mathrm{SI}}[i] = \begin{cases} \boldsymbol{0}_{L\tau_p\times L\tau_p}, & i = \tau_p, \\ \begin{bmatrix} \frac{\xi_{\mathrm{BS}}p_{\mathrm{DL}}M\bar{M}}{M_{\mathrm{sum}}}\boldsymbol{I}_{(i-\tau_p)L} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}, & \tau_p < i \le 2\tau_p, \\ \frac{\xi_{\mathrm{BS}}p_{\mathrm{DL}}M\bar{M}}{M_{\mathrm{sum}}}\boldsymbol{I}_{\tau_p L}, & i > 2\tau_p. \end{cases} \tag{4.20}$$

According to the properties of WP, $\boldsymbol{g}_{n,d}[i+1]$ can be orthogonally decomposed into

$$\boldsymbol{g}_{n,d}[i+1] = \check{\boldsymbol{g}}^{\mathrm{WP}}_{n,d}[i+1] + \check{\boldsymbol{v}}_{n,d}[i+1], \tag{4.21}$$

where $\check{\boldsymbol{v}}_{n,d}[i+1] \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}^d_g - \boldsymbol{\Upsilon}_{d,i})$ is the uncorrelated channel prediction error vector with $\boldsymbol{\Upsilon}_{d,i} = \boldsymbol{R}^d_{g\bar{y}}[1]\boldsymbol{R}^{d}_{\bar{y},i}{}^{-1}[0]\boldsymbol{R}^{d,H}_{g\bar{y}}[1]$.

Similar to (4.7), the predicted DL subchannels can be attained, which are expressed as

$$h_{n,d}[i+1,m] = \check{h}^{\mathrm{WP}}_{n,d}[i+1,m] + \underbrace{\boldsymbol{\psi}_m \check{\boldsymbol{v}}_{n,d}[i+1]}_{\check{v}^{\mathrm{WP}}_{n,d}[i+1,m]}, \; m = 1, ..., M, \tag{4.22}$$

where $\check{h}^{\mathrm{WP}}_{n,d}[i+1,m] \sim \mathcal{CN}\left(0, \sigma^2_{\check{h}^{\mathrm{WP}}_{d,i,m}}\right)$ with $\sigma^2_{\check{h}^{\mathrm{WP}}_{d,i,m}} = \boldsymbol{\psi}_m \boldsymbol{\Upsilon}_{d,i} \boldsymbol{\psi}^H_m$, and $\check{v}^{\mathrm{WP}}_{n,d}[i+1,m] \sim \mathcal{CN}(0, \sigma^2_{\check{v}^{\mathrm{WP}}_{d,i,m}})$ with $\sigma^2_{\check{v}^{\mathrm{WP}}_{d,i,m}} = R^d_h - \sigma^2_{\check{h}^{\mathrm{WP}}_{d,i,m}}$. Moreover, as the channels of different antennas are assumed to be independent, the channels between BS antennas and the d-th MS can be collected to a vector as

$$\check{\boldsymbol{h}}^{\mathrm{WP}}_d[i+1,m] = \boldsymbol{h}_d[i+1,m] + \check{\boldsymbol{v}}^{\mathrm{WP}}_d[i+1,m], \tag{4.23}$$

where $\check{\boldsymbol{h}}^{\mathrm{WP}}_d[i+1,m] \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2_{\check{h}^{\mathrm{WP}}_{d,i,m}} \boldsymbol{I}_N)$ and $\check{\boldsymbol{v}}^{\mathrm{WP}}_d[i+1,m] \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2_{\check{v}^{\mathrm{WP}}_{d,i,m}} \boldsymbol{I}_N)$.

4.4.2.2 Decision-Directed Wiener Predictor

In the proposed Type II frame structure, as UL transmission is activated, channels can be predicted in the DD principles [168]. Assume that a τ_p-th order DD-WP is performed based on the τ_p latest detected symbols received by BS. According to (4.10), the received signal by the n-th antenna of BS over the $\bar{m}$-th UL subcarrier and the i-th symbol duration is

$$s_n[i,\bar{m}] = \sqrt{p_{\mathrm{UL}}}\boldsymbol{x}^T[i,\bar{m}]\boldsymbol{h}_n[i,\bar{m}] + z^{\mathrm{SI}}[i] + z[i,\bar{m}], \tag{4.24}$$

where $\boldsymbol{x}[i,\bar{m}] = [x_1[i,\bar{m}], ..., x_D[i,\bar{m}]]^T$, $\boldsymbol{h}_n[i,\bar{m}] = [h_{n,1}[i,\bar{m}], ..., h_{n,D}[i,\bar{m}]]^T$ and $z^{\mathrm{SI}}[i] \sim \mathcal{CN}(0, \zeta_{\mathrm{BS}} p_{\mathrm{DL}} M)$.

Based on the DD signals of (4.24), a τ_p-th order DD-WP is employed to predict the channels $\boldsymbol{h}_n[i+1,\bar{m}]$ as $\check{\boldsymbol{h}}^{\mathrm{DD}}_n[i+1,\bar{m}] = \boldsymbol{V}_{\mathrm{DD}}\bar{\boldsymbol{s}}_n[i,\bar{m}]$, where $\bar{\boldsymbol{s}}_n[i,\bar{m}] = [s_n[i,\bar{m}], ..., s_n[i+1-\tau_p,\bar{m}]]^T$, and $\boldsymbol{V}_{\mathrm{DD}}$ can be expressed as

$$\boldsymbol{V}_{\mathrm{DD}} = \boldsymbol{R}_{h\bar{s}}[1,\bar{m}]\boldsymbol{R}^{-1}_{\bar{s},i}[0,\bar{m}] \tag{4.25}$$

with $\boldsymbol{R}_{h\bar{s}}[1,\bar{m}] = \mathbb{E}\left[\boldsymbol{h}_n[i+1,\bar{m}]\bar{\boldsymbol{s}}^H_n[i,\bar{m}]\right]$ and $\boldsymbol{R}_{\bar{s},i}[0,\bar{m}] = \mathbb{E}\left[\bar{\boldsymbol{s}}_n[i,\bar{m}]\bar{\boldsymbol{s}}^H_n[i,\bar{m}]\right]$, which are detailed below.

First, the cross-correlation matrix in (4.25) can be obtained with the aid of (4.7) and (4.24), which is

$$\begin{aligned}\boldsymbol{R}_{h\bar{s}}[1,\bar{m}] &= \mathbb{E}\left[\boldsymbol{h}_n[i+1,\bar{m}]\bar{\boldsymbol{s}}_n^H[i,\bar{m}]\right] = \sqrt{p_{\mathrm{UL}}}\left[\alpha\boldsymbol{R}_h\boldsymbol{x}^*[i,\bar{m}],...,\alpha^{\tau_p}\boldsymbol{R}_h\boldsymbol{x}^*[i+1-\tau_p,\bar{m}]\right] \\ &= \sqrt{p_{\mathrm{UL}}}\boldsymbol{A}(\tau_p,\alpha)\boldsymbol{Q}(\tau_p,D)\boldsymbol{B}^H(\tau_p,D),\end{aligned} \tag{4.26}$$

where

$$\begin{aligned}&\boldsymbol{R}_h = \mathrm{diag}(R_h^1,...,R_h^D),\ \boldsymbol{Q}(\tau_p,D) = \boldsymbol{R}_h\otimes\boldsymbol{I}_{\tau_p},\ \boldsymbol{A}(\tau_p,\alpha) = \boldsymbol{I}_D\otimes\boldsymbol{\delta}(\tau_p,\alpha), \\ &\boldsymbol{B}(\tau_p,D) = [\mathrm{diag}(\hat{\boldsymbol{x}}_1[\bar{m}]),...,\mathrm{diag}(\hat{\boldsymbol{x}}_d[\bar{m}]),...,\mathrm{diag}(\hat{\boldsymbol{x}}_D[\bar{m}])]\end{aligned} \tag{4.27}$$

and in $\boldsymbol{B}(\tau_p,D)$, $\hat{\boldsymbol{x}}_d[\bar{m}] = \left[\hat{x}_d[i,\bar{m}],...,\hat{x}_d[i+1-\tau_p,\bar{m}]\right]^T$ are UL detected symbols. To obtain (4.26), each element in $\hat{\boldsymbol{x}}_d[\bar{m}]$ needs to be correctly detected. This can be achieved by measuring the reliability of each detection in terms of the likelihood value of a detected symbol. If the likelihood value is larger than a preset threshold resulting in high reliability, the symbol is then included for DD prediction. Otherwise, the symbol will not be used for DD channel prediction.

The auto-correlation matrix $\boldsymbol{R}_{\bar{s},i}[0,\bar{m}]$ in (4.25) is given by

$$\begin{aligned}\boldsymbol{R}_{\bar{s},i}[0,\bar{m}] &= \mathbb{E}\left[\bar{\boldsymbol{s}}_n[i,\bar{m}]\bar{\boldsymbol{s}}_n^H[i,\bar{m}]\right] \\ &= p_{\mathrm{UL}}\boldsymbol{B}(\tau_p,D)\boldsymbol{C}(\tau_p,D)\boldsymbol{Q}(\tau_p,D)\boldsymbol{B}^H(\tau_p,D) + \sigma^2\boldsymbol{I}_{\tau_p} + \boldsymbol{\Xi}_{\mathrm{SI}}^{\mathrm{DD}}[i],\end{aligned} \tag{4.28}$$

where $\boldsymbol{C}(\tau_p,D) = \boldsymbol{I}_D\otimes\boldsymbol{\xi}(\tau_p,\alpha)$. When the proposed Type II frame structure is employed, SI is also present at the observation time, which leads to

$$\boldsymbol{\Xi}_{\mathrm{SI}}^{\mathrm{DD}}[i] = \begin{cases}\begin{bmatrix}\xi_{\mathrm{BS}}p_{\mathrm{DL}}M\boldsymbol{I}_{\tau_p-1} & \boldsymbol{0}\\ \boldsymbol{0} & 0\end{bmatrix}, & i=\tau_p,\\ \xi_{\mathrm{BS}}p_{\mathrm{DL}}M\boldsymbol{I}_{\tau_p}, & i>\tau_p.\end{cases} \tag{4.29}$$

With the aid of the orthogonal decomposition, we have the expression

$$\boldsymbol{h}_n[i+1,\bar{m}] = \check{\boldsymbol{h}}_n^{\mathrm{DD}}[i+1,\bar{m}] + \check{\boldsymbol{e}}_n^{\mathrm{DD}}[i+1,\bar{m}], \tag{4.30}$$

where $\check{\boldsymbol{e}}_n^{\mathrm{DD}}[i+1,\bar{m}] \sim \mathcal{CN}(\boldsymbol{0},\boldsymbol{R}_h-\boldsymbol{\Theta}_i[\bar{m}])$ is the uncorrelated channel prediction error vector, and $\boldsymbol{\Theta}_i[\bar{m}] = \boldsymbol{R}_{h\bar{s}}[1,\bar{m}]\boldsymbol{R}_{\bar{s},i}^{-1}[0,\bar{m}]\boldsymbol{R}_{h\bar{s}}^H[1,\bar{m}]$, which is the covariance matrix of $\check{\boldsymbol{h}}_n^{\mathrm{DD}}[i+1,\bar{m}]$. Furthermore, we can write $\check{\boldsymbol{h}}_{n,d}^{\mathrm{UL}}[i+1] = \left[\check{h}_{n,d}^{\mathrm{DD}}[i+1,1],...,\check{h}_{n,d}^{\mathrm{DD}}[i+1,\bar{M}]\right]^T$, which follows the distribution of $\check{\boldsymbol{h}}_{n,d}^{\mathrm{UL}}[i+1] \sim \mathcal{CN}(\boldsymbol{0},\boldsymbol{\Gamma}_{d,i})$, with $\boldsymbol{\Gamma}_{d,i}$ derived in detail in Appendix D.

Having obtained the frequency-domain prediction of $\boldsymbol{h}_{n,d}[i+1]$ over UL subcarriers, the corresponding time-domain CSI can be found as

$$\check{\boldsymbol{g}}_{n,d}^{\mathrm{DD}}[i+1] = \boldsymbol{J}\check{\boldsymbol{h}}_{n,d}^{\mathrm{UL}}[i+1] \tag{4.31}$$

provided that $\bar{M} \geq L$. In (4.31), $\boldsymbol{J} = \left((\boldsymbol{\Phi}_{\mathrm{UL}}\boldsymbol{F}\boldsymbol{\Psi})^H(\boldsymbol{\Phi}_{\mathrm{UL}}\boldsymbol{F}\boldsymbol{\Psi})\right)^{-1}(\boldsymbol{\Phi}_{\mathrm{UL}}\boldsymbol{F}\boldsymbol{\Psi})^H$. Finally, upon following the same approach described in (4.21)-(4.22), the DL channel vector between BS and the d-th MS with respect to the m-th DL subcarrier and the $(i+1)$ symbol duration can be predicted as $\check{\boldsymbol{h}}_d^{\mathrm{DD}}[i+1,m]$, which follows

$$\check{\boldsymbol{h}}_d^{\mathrm{DD}}[i+1,m] = \boldsymbol{h}_d[i+1,m] + \check{\boldsymbol{v}}_d^{\mathrm{DD}}[i+1,m],\ m = 1,...,M, \tag{4.32}$$

where $\check{\boldsymbol{h}}_d^{\mathrm{DD}}[i+1,m] \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2_{\check{h}^{\mathrm{DD}}_{d,i,m}}\boldsymbol{I}_N)$, $\sigma^2_{\check{h}^{\mathrm{DD}}_{d,i,m}} = \boldsymbol{\psi}_m\boldsymbol{J}\boldsymbol{\Gamma}_{d,i}\boldsymbol{J}^H\boldsymbol{\psi}_m^H$, and $\check{\boldsymbol{v}}_d^{\mathrm{DD}}[i+1,m] \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2_{\check{v}^{\mathrm{DD}}_{d,i,m}}\boldsymbol{I}_N)$, $\sigma^2_{\check{v}^{\mathrm{DD}}_{d,i,m}} = R_h^d - \sigma^2_{\check{h}^{\mathrm{DD}}_{d,i,m}}$.

4.5 Performance Analysis over Fast-Fading Channels

In this section, we derive the ergodic rates attainable by the MDD and TDD systems with the proposed channel acquisition, where the two types of frame structures introduced in Section 4.3 are respectively applied.

4.5.1 Type I Frame Structure

The ergodic achievable rate of the MDD systems with the Type I frame structure shown in Fig. 4.1 can be expressed as

$$R^{\text{MDD-1}}(\tau_p) = \frac{1}{TM_{\mathrm{sum}}}\sum_{d=1}^{D}\sum_{i=\tau_p+1}^{T}\sum_{m=1}^{M} R_{d,i,m}^{\text{MDD-1}}, \tag{4.33}$$

where $R_{d,i,m}^{\text{MDD-1}}$ denotes the rate obtained by the d-th MS over the m-th DL subcarrier during the i-th symbol duration. Note that since no DL pilots are employed, only the expected effective channel gain is available at MSs. Hence, the mean effective channel gain is treated as the CSI for signal detection at MSs' sides. The effective channel gain is $\omega_{d,d}^{\mathrm{ZF}}[i,m] = \boldsymbol{h}_d^H[i,m]\check{\boldsymbol{f}}_d[i,m]$, where $\check{\boldsymbol{f}}_d[i,m]$ is derived by the ZF principle with the predicted channel $\check{\boldsymbol{h}}_d^{\mathrm{WP}}[i+1,m]$ obtained in (4.23). Then, we can derive the lower bounded

achievable ergodic rate of $R_{d,i,m}^{\text{MDD-1-LB}}$ as

$$\begin{aligned} R_{d,i,m}^{\text{MDD-1}} &\geq R_{d,i,m}^{\text{MDD-1–LB}} \\ &= \log_2\left(1+\frac{p_{\text{DL}}\left|\mathbb{E}\left[\omega_{d,d}^{\text{ZF}}[i,m]\right]\right|^2}{p_{\text{DL}}\text{var}\left\{\omega_{d,d}^{\text{ZF}}[i,m]\right\}+p_{\text{DL}}\sum\limits_{k=1,k\neq d}^{D}\mathbb{E}\left[\left|\omega_{d,k}^{\text{ZF}}[i,m]\right|^2\right]+\text{var}\left\{z_d^{\text{SI}}\right\}+\sigma^2}\right). \end{aligned} \tag{4.34}$$

According to Lemma 4 in [169], the expectation and variance of the effective channel gain $\omega_{d,d}^{\text{ZF}}[i,m]$ can be approximated as

$$\mathbb{E}\left[\omega_{d,d}^{\text{ZF}}[i,m]\right]=\mathbb{E}\left[1/(\sqrt{D}\left\|\check{\boldsymbol{f}}_d[i,m]\right\|)\right]\approx\sqrt{\frac{N-D+1}{D}}\sqrt{\sigma^2_{\check{h}_{d,i,m}^{\text{WP}}}} \tag{4.35}$$

and

$$\text{var}\left\{\omega_{d,d}^{\text{ZF}}[i,m]\right\}\approx\frac{1}{D}\left(\frac{1}{4}\sigma^2_{\check{h}_{d,i,m}^{\text{WP}}}+\sigma^2_{\check{v}_{d,i,m}^{\text{WP}}}\right), \tag{4.36}$$

respectively, where $\check{\boldsymbol{f}}_d[i,m]$ is the unnormalized column of the matrix $\check{\boldsymbol{F}}^{\text{ZF}}[i,m]$ seen in (4.8). For the interference signal from the other MSs, as seen in the denominator of (4.34), we have

$$\begin{aligned} \mathbb{E}\left[\left|\omega_{d,k}^{\text{ZF}}[i,m]\right|^2\right] &= \mathbb{E}\left[\left|\check{\boldsymbol{h}}_d^{\text{WP},H}[i,m]\check{\boldsymbol{f}}_k[i,m]+\check{\boldsymbol{v}}_d^{\text{WP},H}[i+1,m]\check{\boldsymbol{f}}_k[i,m]\right|^2\right] \\ &= \mathbb{E}\left[\left|\check{\boldsymbol{v}}_d^{\text{WP},H}[i+1,m]\check{\boldsymbol{f}}_k[i,m]\right|^2\right] \\ &= \frac{1}{D}\sigma^2_{\check{v}_{d,i,m}^{\text{WP}}}. \end{aligned} \tag{4.37}$$

Consequently, upon substituting (4.22), (4.35), (4.37) and (C.1) into (4.34), the approximation to the lower bounded achievable ergodic rate of the d-th MS on the m-th DL subcarrier within the i-th OFDM symbol duration can be expressed as

$$\begin{aligned} &R_{d,i,m}^{\text{MDD-1-LB}} \\ &\approx\log_2\left(1+\frac{p_{\text{DL}}\left(N-D+1\right)\boldsymbol{\psi}_m\boldsymbol{\Upsilon}_{d,i}\boldsymbol{\psi}_m^H}{0.25p_{\text{DL}}\boldsymbol{\psi}_m\boldsymbol{\Upsilon}_{d,i}\boldsymbol{\psi}_m^H+p_{\text{DL}}D\left(\frac{\beta_d}{M_{\text{sum}}}-\boldsymbol{\psi}_m\boldsymbol{\Upsilon}_{d,i}\boldsymbol{\psi}_m^H\right)+\xi_{\text{MS}}p_{\text{UL}}D\bar{M}+D\sigma^2}\right). \end{aligned} \tag{4.38}$$

Finally, $R^{\text{MDD-1-LB}}(\tau_p)$ is obtained by substituting (4.38) into (4.33).

On the other hand, the achievable ergodic rate of the TDD systems with the Type I frame structure shown in Fig. 4.1 can be written as

$$R^{\text{TDD-1}}(\tau_p) = \frac{1}{TM_{\text{sum}}} \sum_{d=1}^{D} \left(\sum_{i=\tau_p+2}^{T-1} \sum_{m=1}^{M_{\text{sum}}} R_{d,i,m}^{\text{TDD-1}} + \frac{1}{2} \sum_{m=1}^{M_{\text{sum}}} \left(R_{d,\tau_p+1,m}^{\text{TDD-1}} + R_{d,T,m}^{\text{TDD-1}} \right) \right). \quad (4.39)$$

Furthermore, the approximated lower bounded achievable ergodic rate of TDD systems can be derived by following nearly the same steps as above for the MDD systems, except only the following two points. The first one is that all the channels within one frame are predicted relying on the first τ_p training symbols sent at the start of the frame, yielding $\boldsymbol{\delta}(\tau_p, \alpha)$ in (4.17) being changed to $[\alpha^{i-\tau_p}, \alpha^{i+1-\tau_p}, ..., \alpha^{i-1}]$. The second point is that the TDD systems do not experience SI. Therefore, we have $\Xi_{\text{SI}}[i] = 0, \forall i$ in (4.20).

4.5.2 Type II Frame Structure

The ergodic achievable rate of the MDD systems with the Type II frame structure shown in Fig. 4.2 can be expressed as

$$R^{\text{MDD-2}}(\tau_p.\kappa_p) = \frac{1}{TM_{\text{sum}}} \sum_{d=1}^{D} \left(\sum_{i=\kappa_p+1}^{\tau_p} \sum_{m=1}^{M} R_{d,i,m}^{\text{MDD-1}} + \sum_{i=\tau_p+1}^{T} \sum_{m=1}^{M} R_{d,i,m}^{\text{MDD-2-DL}} + \sum_{i=\tau_p+1}^{T} \sum_{\bar{m}=1}^{\bar{M}} R_{d,i,\bar{m}}^{\text{MDD-2-UL}} \right), \quad (4.40)$$

where $R_{d,i,m}^{\text{MDD-1}}$ denotes the rate attained during Phase I, $R_{d,i,m}^{\text{MDD-2-DL}}$ and $R_{d,i,\bar{m}}^{\text{MDD-2-UL}}$ denote respectively the DL and UL rates achievable during Phase II. To compute the lower bounded rate $R^{\text{MDD-2-LB}}(\tau_p.\kappa_p)$, in (4.40), the lower bound for $R_{d,i,m}^{\text{MDD-1}}$ is given by (4.38), while the lower bound for $R_{d,i,m}^{\text{MDD-2-DL}}$ can be obtained from the approach for computing $R_{d,i,m}^{\text{MDD-1-LB}}$ by replacing $\check{h}_{d,i,m}^{\text{WP}}$ using $\check{h}_{d,i,m}^{\text{DD}}$. Therefore, we only need to derive the lower bound for the achievable UL ergodic rate, i.e., $R_{d,i,\bar{m}}^{\text{MDD-2-UL}}$. Denote the effective channel gain as $\omega_{d,d}^{\text{MRC}}[i,\bar{m}] = \check{\boldsymbol{w}}_d^H[i,\bar{m}]\boldsymbol{h}_d[i,\bar{m}]$. Then, we have

$$\begin{aligned}
R_{d,i,\bar{m}}^{\text{MDD-2-UL}} &\geq R_{d,i,\bar{m}}^{\text{MDD-2-UL-LB}} \\
&= \log_2\Big(1 + \frac{p_{\text{UL}} \left| \mathbb{E}\left[\omega_{d,d}^{\text{MRC}}[i,\bar{m}]\right] \right|^2}{p_{\text{UL}} \sum_{k=1,k\neq d}^{D} \mathbb{E}\left[\left| \omega_{d,k}^{\text{MRC}}[i,\bar{m}] \right|^2 \right] + \left(\text{var}\left\{\boldsymbol{z}^{\text{SI}}[i]\right\} + \sigma^2\right) \mathbb{E}\left[\|\check{\boldsymbol{w}}_d[i,\bar{m}]\|_2^2 \right]} \Big) \\
&\overset{(a)}{=} \log_2\Big(1 + \frac{p_{\text{UL}} \left| \mathbb{E}\left[\check{\boldsymbol{h}}_d^{\text{DD},H}[i,\bar{m}]\boldsymbol{h}_d[i,\bar{m}]\right] \right|^2}{p_{\text{UL}} \sum_{k=1,k\neq d}^{D} \mathbb{E}\left[\left| \check{\boldsymbol{h}}_d^{\text{DD},H}[i,\bar{m}]\boldsymbol{h}_k[i,\bar{m}] \right|^2 \right] + \left(\xi_{\text{BS}} p_{\text{DL}} M + \sigma^2\right) \mathbb{E}\left[\left\| \check{\boldsymbol{h}}_d^{\text{DD}}[i,\bar{m}] \right\|_2^2 \right]} \Big),
\end{aligned} \quad (4.41)$$

where (a) is obtained by using $\check{\boldsymbol{w}}_d[i,\bar{m}] = \check{\boldsymbol{h}}_d^{\mathrm{DD}}[i,\bar{m}]$ to implement the MRC receiver. In particular, the expectation term in the numerator of (4.41) can be derived as

$$\begin{aligned}\left|\mathbb{E}\left[\check{\boldsymbol{h}}_d^{\mathrm{DD},H}[i,\bar{m}]\boldsymbol{h}_d[i,\bar{m}]\right]\right|^2 &= \left|\mathbb{E}\left[\check{\boldsymbol{h}}_d^{\mathrm{DD},H}[i,\bar{m}]\left(\check{\boldsymbol{h}}_d^{\mathrm{DD}}[i,\bar{m}]+\check{\boldsymbol{e}}_d^{\mathrm{DD}}[i,\bar{m}]\right)\right]\right|^2 \\ &\overset{(b)}{=} \left|\mathbb{E}\left[\check{\boldsymbol{h}}_d^{\mathrm{DD},H}[i,\bar{m}]\check{\boldsymbol{h}}_d^{\mathrm{DD}}[i,\bar{m}]\right]\right|^2 \\ &= \left|\mathbb{E}\left\|\check{\boldsymbol{h}}_d^{\mathrm{DD}}[i,\bar{m}]\right\|_2^2\right|^2, \end{aligned} \tag{4.42}$$

where (b) is due to the fact that $\check{\boldsymbol{v}}_d^{\mathrm{DD}}[i,m]$ has zero mean and is independent of $\check{\boldsymbol{h}}_d[i,\bar{m}]$. The first expectation in the denominator of (4.41) can be derived as

$$\begin{aligned}\mathbb{E}\left[\left|\check{\boldsymbol{h}}_d^{\mathrm{DD},H}[i,\bar{m}]\boldsymbol{h}_k[i,\bar{m}]\right|^2\right] &= \mathbb{E}\left[\left|\check{\boldsymbol{h}}_d^{\mathrm{DD},H}[i,\bar{m}]\check{\boldsymbol{h}}_k^{\mathrm{DD}}[i,\bar{m}]+\check{\boldsymbol{h}}_d^{\mathrm{DD},H}[i,\bar{m}]\check{\boldsymbol{e}}_k^{\mathrm{DD}}[i,\bar{m}]\right|^2\right] \\ &\overset{(c)}{\leq} \mathbb{E}\left[\left|\check{\boldsymbol{h}}_d^{\mathrm{DD},H}[i,\bar{m}]\check{\boldsymbol{h}}_k^{\mathrm{DD}}[i,\bar{m}]\right|^2\right]+\mathbb{E}\left[\left|\check{\boldsymbol{h}}_d^{\mathrm{DD},H}[i,\bar{m}]\check{\boldsymbol{e}}_k^{\mathrm{DD}}[i,\bar{m}]\right|^2\right] \\ &\overset{(d)}{=} R_h^k\mathbb{E}\left[\left\|\check{\boldsymbol{h}}_d^{\mathrm{DD}}[i,\bar{m}]\right\|_2^2\right], \end{aligned} \tag{4.43}$$

where (c) and (d) follow from the fact that $\check{\boldsymbol{h}}_d^{\mathrm{DD}}[i,\bar{m}]$, $\check{\boldsymbol{h}}_k^{\mathrm{DD}}[i,\bar{m}]$ and $\check{\boldsymbol{e}}_k^{\mathrm{DD}}[i,\bar{m}]$ are independent vectors and $\check{\boldsymbol{v}}_k^{\mathrm{DD}}[i,\bar{m}]$ has zero mean. Consequently, upon substituting (4.30), (4.42) and (4.43) into (4.41), we can obtain the approximated lower-bounded UL rate, which is given by

$$R_{d,i,\bar{m}}^{\text{MDD-2-UL-LB}} \approx \log_2\left(1+\frac{p_{\mathrm{UL}}N\left(\boldsymbol{\Theta}_i[\bar{m}]\right)_{d,d}}{p_{\mathrm{UL}}\sum\limits_{k=1,k\neq d}^{D}\frac{\beta_k}{M_{\mathrm{sum}}}+\xi_{\mathrm{BS}}p_{\mathrm{DL}}M+\sigma^2}\right). \tag{4.44}$$

In the context of TDD, the achievable ergodic rate in the case of the Type II frame structure shown in Fig. 4.2 can be expressed as

$$\begin{aligned}R^{\text{TDD-2}}(\tau_p,\tau_u) = \frac{1}{TM_{\mathrm{sum}}}\sum_{d=1}^{D}\Bigg(&\sum_{i=\tau_p+1}^{\tau_p+\tau_u}\sum_{m=1}^{M_{\mathrm{sum}}}R_{d,i,m}^{\text{TDD-2-UL}}+\sum_{i=\tau_p+\tau_u+2}^{T}\sum_{m=1}^{M_{\mathrm{sum}}}R_{d,i,m}^{\text{TDD-2-DL}} \\ &+\frac{1}{2}\sum_{m=1}^{M_{\mathrm{sum}}}\left(R_{d,\tau_p+\tau_u+1,m}^{\text{TDD-2-UL}}+R_{d,T,m}^{\text{TDD-2-DL}}\right)\Bigg). \end{aligned} \tag{4.45}$$

Based on (4.45), the approximated lower bound $R^{\text{TDD-2-LB}}(\tau_p,\tau_u)$ can be obtained in the way similar to that for MDD systems. Therefore, the derivation is omitted to avoid redundancy.

TABLE 4.1: Simulation parameters

Default parameters	Value
Distance between BS and MSs (D_d)	$[50, 100]$ m
Number of BS antennas (N)	32
Number of MSs (D)	8
Number of DL/UL subcarriers for TDD (M_{sum})	96
Number of DL/UL subcarriers for MDD ($M/\bar{M}$)	64 / 32
BS transmit power ($p_{\text{DL}}M$)	30 dBm
MS transmit power ($p_{\text{UL}}\bar{M}$)	20 dBm
Noise power	-94 dBm
Large-scale fading (β_d)	$(D_d)^{3.8}$
Delay taps (L)	4
Carrier center frequency	5 GHz
Subcarrier bandwidth	15 kHz
OFDM symbol duration (T_s)	66.67 μs
Total symbols in one frame (T)	28
Switching interval in TDD	0.5
Digital modulation	16 QAM
SIC capability at BS (ξ_{BS})	130 dB
SIC capability at MS (ξ_{MS})	120 dB

4.6 Simulation Results and Discussion

In this section, we present the numerical results for the performance comparison of the MDD and TDD systems with two types of frame structures, as shown in Fig. 4.1 and Fig. 4.2, respectively. We also compare the performance of IBFD and MDD systems with the proposed frame structures, when SIC capability is limited. Furthermore, we validate the analytical results derived in Section 4.5 by Monte Carlo simulations over 1000 channel realizations. Throughout this section, we assume a single-cell network with one BS and multiple single-antenna MSs. Some key parameters are listed in Table 4.1. It is noteworthy that to demonstrate the tendency of the rate change within one frame, the switching intervals in the TDD mode are not considered, while the X-axis represents the index of OFDM symbol. By contrast, the switching intervals are taken into account in evaluating the average sum rate over one frame, as shown in (4.39) and (4.45).

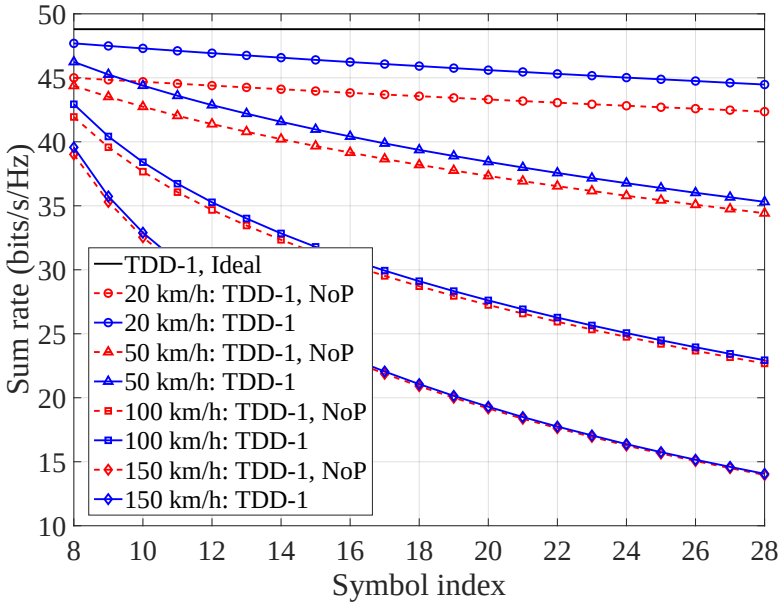
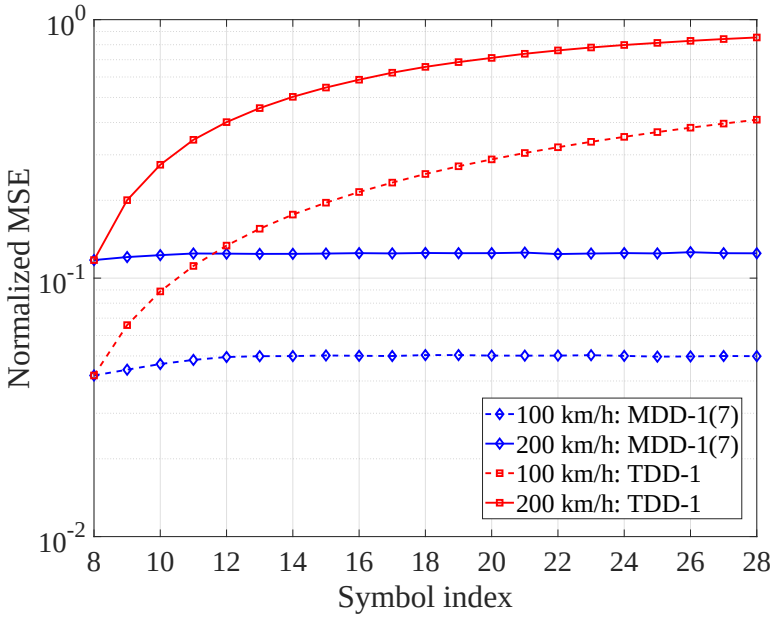


FIGURE 4.3: Sum rate versus OFDM symbol index of the TDD systems with the different assumptions about channel acquisition and relative velocity.

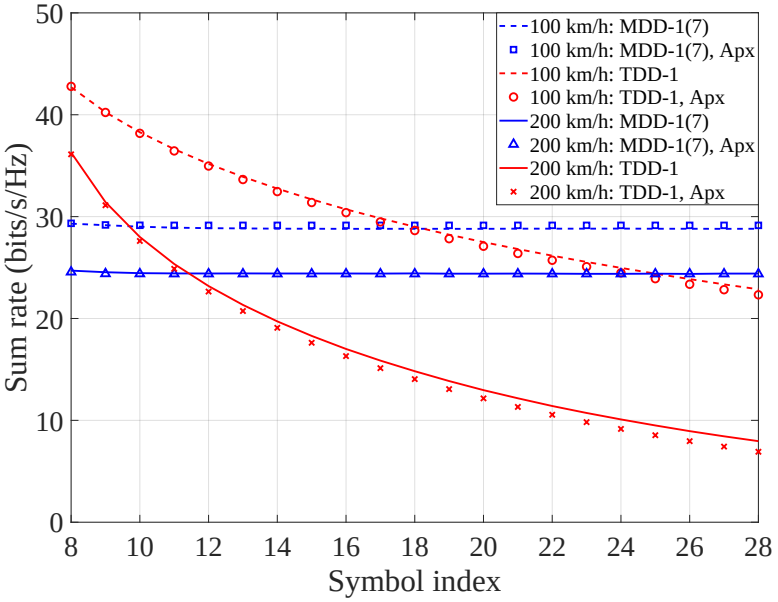
4.6.1 Type I Frame Structure: TDD Vs MDD

In this subsection, TDD and MDD systems are compared under the Type I frame structure. For TDD systems, a total of 7 pilots are transmitted within one frame. By contrast, for MDD systems, UL pilots can always be transmitted or be transmitted in part. Note that, in the following figures, 'TDD-1' denotes the scheme that all pilots are consecutively transmitted at the beginning of the frame and a 7-th order WP is applied. The notations of 'TDD-1, Ideal' and 'TDD-1, NoP' indicate respectively that the channel is ideal without the channel aging problem and that the CSI acquisition does not rely on channel prediction. 'MDD-1(z)' denotes the scheme that UL pilots are continuously sent by MSs and a z-th order WP is applied. Furthermore, the results of 'TDD-1' and 'MDD-1(z)' are plotted by using the lower-bounded rates given in (4.33), (4.39), (4.40) and (4.45), while the results of 'TDD-1, Apx' and 'MDD-1(z), Apx' represent the correspondingly approximated rates derived in Section 4.5.

In Fig. 4.3, we evaluate the performance of the TDD systems operated in different mobility scenarios with or without channel prediction. We observe that when the relative velocity is 20 km/h, the TDD system with WP outperforms the TDD system without channel prediction. However, as the symbol index increases, meaning that CSI becomes more outdated, the performance gap between the above two strategies becomes narrower. Therefore, in the TDD mode, the latter symbols benefit less from channel prediction, as only 7 symbols are used for prediction and the prediction error increases over time. Additionally, as seen in Fig. 4.3, as the relative speed increases, the advantage of employing channel prediction diminishes, and the performance drops faster and becomes significantly worse than the system with precise CSI. The results in Fig.



(A) Normalized MSE



(B) Sum rate

FIGURE 4.4: Performance comparison of the MDD and TDD systems, when Type I frame structure and 7-th order WP are employed.

4.3 explicitly indicate that the traditional TDD mode is not feasible for communication in high-mobility scenarios, even when channel prediction is employed.

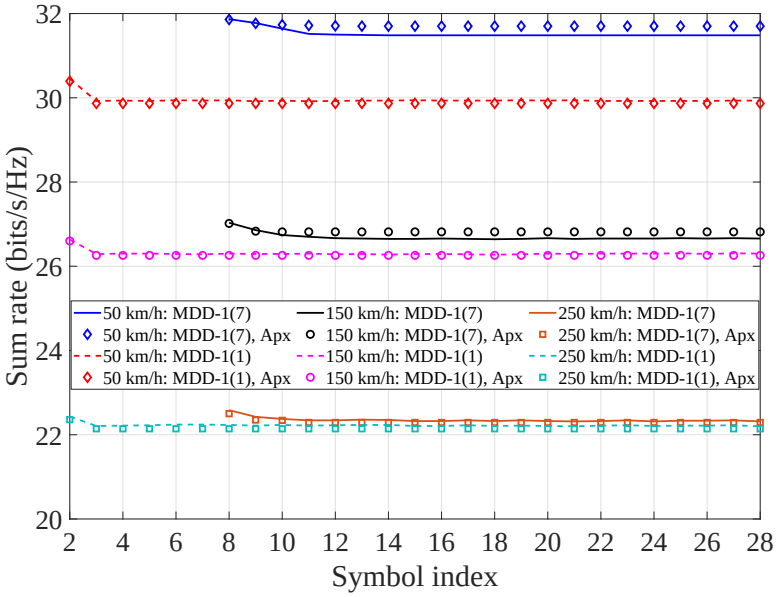
Based on the Type I frame structure presented in Fig. 4.1, we demonstrate the effect of the time-varying channels on the performance of TDD and MDD systems in Fig. 4.4. It can be observed from Fig. 4.4(A) that the accuracy of channel prediction in MDD systems is nearly constant after 7 pilot symbols, regardless of the relative traveling velocity of 100 km/h or 200 km/h. This is because in MDD systems, the UL pilots can continuously provide BS with the most updated CSI for DL transmission, which hence

guarantees a stable system sum rate, as shown in Fig. 4.4(B). Moreover, the difference between the sum rates achieved by the MDD systems at 100 km/h and 200 km/h is relatively small, meaning that MDD systems are robust to time-varying fading channels. By contrast, although the sum rate of TDD systems is higher than that of MDD systems within the first several symbols, for both the velocities considered, it reduces quickly with time in terms of the OFDM symbol indices, due to the reduced accuracy of channel prediction with time. Furthermore, when the velocity is changed from 100 km/h to 200 km/h, the sum rate deteriorates significantly. Therefore, the performance of TDD systems is sensitive to the mobility of wireless channels. The rationale is that in TDD systems, UL training and DL have to be performed in an alternative way. Once it is switched to DL transmission, the CSI has to be predicted based on the old data, which may be outdated, if mobility is high. Additionally, in Fig. 4.4, it is demonstrated that the approximations of the lower bounded ergodic rates derived in Section 4.5 match closely to the lower bounded sum rates attained via Monte Carlo simulations.

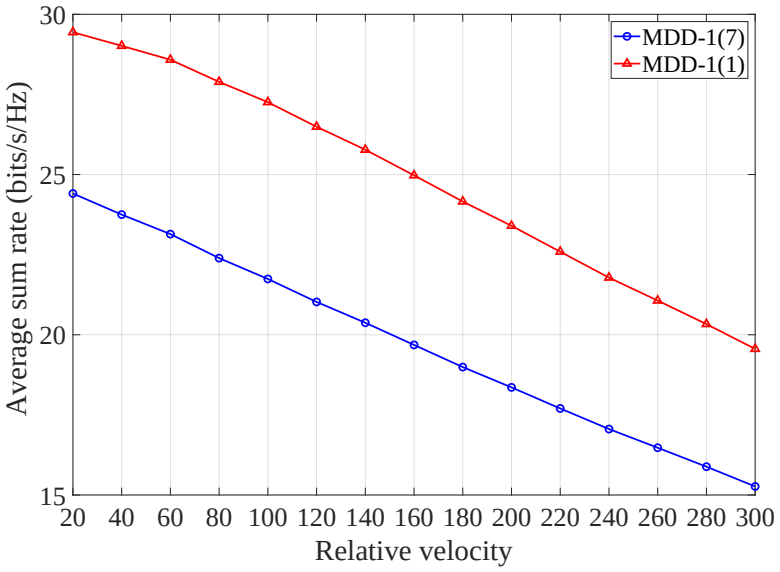
In Fig. 4.5, we investigate the performance of MDD systems, when Type I frame structure and different orders of WPs are employed. According to Fig. 4.1, a higher order WP uses more channel observations received in the past to predict channels. It is observed from Fig. 4.5(A) that at the relative speed of 50 km/h, the 7-th order WP in general significantly outperforms the 1-st order WP. However, when the speed is increased to 250 km/h, the 1-st order WP achieves nearly the same performance as the 7-th order WP. The reason behind is that when channel varies fast, the prediction of the current channel is more dependent on the nearest pilot symbols and the observations in the past become less correlated to the prediction. Fig. 4.5(B) further shows that when pilots are continuously transmitted under the Type I frame structure, the 1-st order WP attains a better performance averaged over one frame, owing to the fact that in this case, 6 more DL symbols can be transmitted, even though the slightly reduced prediction accuracy results in some performance loss.

Next, to comprehensively compare the performance of the TDD and MDD systems employing the Type I frame, we consider two more pilot distribution methods for TDD, namely the TDD-1-ES and TDD-1-TG, as shown in Fig. 4.6. Specifically, with the TDD-1-ES, 7 training symbols are evenly distributed within one frame. By contrast, with the TDD-1-TG, 7 training symbols are first divided into two groups, which are then inserted into the frame. Note that the TDD-1-TG is similar to the frame structure applied in 3GPP 36.211, where two pilot subframes are evenly distributed within one frame [170]. For MDD systems, we also introduce an alternative Type I frame structure denoted as MDD-1-PA, as depicted in Fig. 4.1. With the MDD-1-PA, $\tau_p = 1$, and a total of 7 pilots are activated, which are evenly distributed within one frame.

In Fig. 4.7, we compare the average sum rate of the MDD and TDD systems with the various Type I frame structures, and when two frame lengths are considered. Note that,



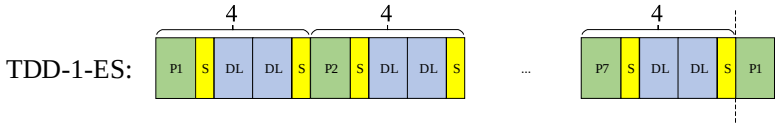
(A) Sum rate versus OFDM symbol index.



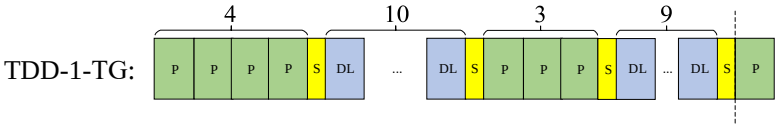
(B) Average sum rate within one frame versus relative velocity.

FIGURE 4.5: Performance comparison of the MDD systems with Type I frame structure and different orders of WPs.

as the frame length increases, the number of pilots is set to remain the same in MDD-1-PA and three TDD structures, while the number of DL symbols increases. It can be observed from Fig. 4.7 that MDD-1-PA outperforms all the other arrangements, when the relative speed is lower than 280 km/h in Fig. 4.7(A) and 210 km/h in Fig. 4.7(B). This is because when only a fraction of pilots are transmitted, the saved UL resources (subcarriers) can be exploited to transmit DL data, provided that the CSI updating is enough for protecting the DL transmission. By contrast, MDD-1(1) is superior to the MDD-1-PA in the high-speed case. In this case, although there are only M subcarriers used for DL transmission, the continuously transmitted pilots on the $\bar{M}$ UL subcarriers



(A) Evenly spaced pilots in TDD Type I frame structure.



(B) Two groups of pilots in TDD Type I frame structure.

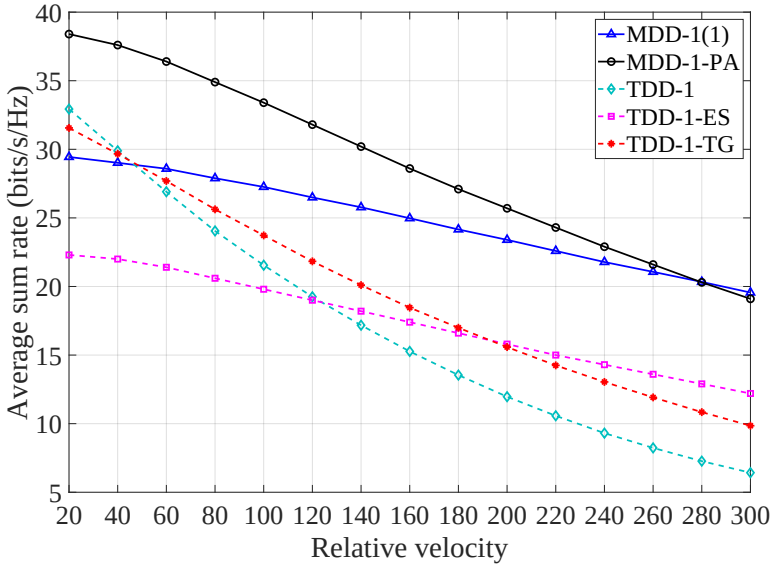
FIGURE 4.6: Alternative pilot distribution methods for the Type I frame structure in TDD mode, when the total number of pilots is 7.

guarantees the reliable channel prediction, which the MDD-1-PA structure is however unable to provide. With respect to TDD, the advantages of the three frame structures are dependent on the relative velocity and the length of frames. As shown in Fig. 4.7(A) and 4.7(B), in general, TDD-1 and TDD-TG have relatively better performance, when the relative velocity is low, owing to the use of less switching intervals. By contrast, TDD-1-ES benefits from the frequent transmission of training symbols within one frame and hence performs better in the high-speed long-frame case.

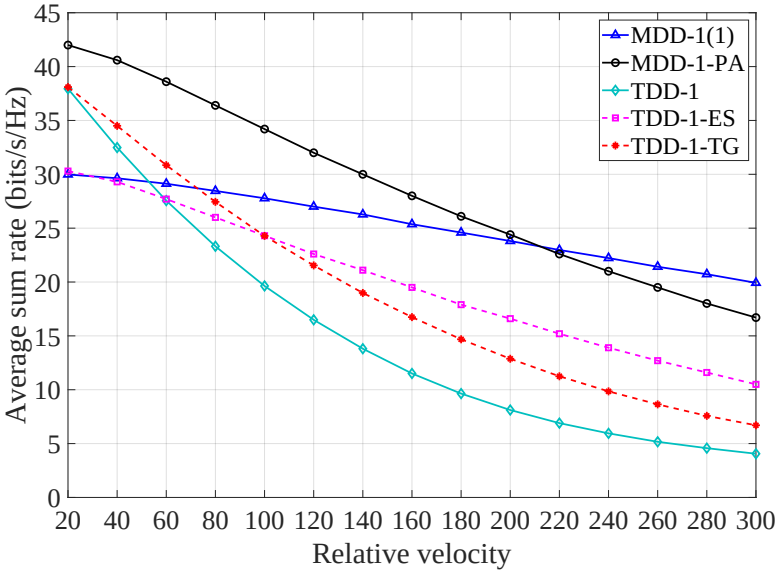
4.6.2 Type II Frame Structure: TDD Vs MDD

In this subsection, we compare the performance of MDD and TDD systems operated with the Type II frame structure. In the following figures, 'TDD-2' and 'MDD-2' denote respectively the TDD and MDD systems employing the proposed frame structure as shown in Fig. 4.2 with $\tau_p = 7$, $\tau_u = 7$ and $\kappa_p = 1$.

The performance comparison of MDD and TDD with respect to symbol index is plotted in Fig. 4.8, when Type II frame structure is considered. In conjunction with Fig. 4.2, we can see that, as the MDD system is able to operate DL and UL simultaneously during Phase I, DL data symbols can be transmitted while collecting the channel observations for prediction in Phase II. From Fig. 4.8(A) we observe that the WP in Phase I yields less prediction errors when compared with the DD-WP in Phase II. This is as expected, since the pilots used for WP in Phase I are known to BS. In Phase II, although the performance of channel prediction is better without the effect of SI, TDD system only activates the UL transmission from the 8-th to the 14-th symbol, leading to lower sum rate when compared with the MDD system, which is capable of transmitting data over both DL and UL. At the 15-th symbol, both MDD and TDD systems apply the 7-th DD-WP and the TDD system attains a higher sum rate, as it uses all the subcarriers for DL transmission. However, similar to the Type I frame structure, new UL data symbols are not available for the TDD system to update channel causing the degraded accuracy of channel prediction. Consequently, its sum rate decreases, as symbol index increases.



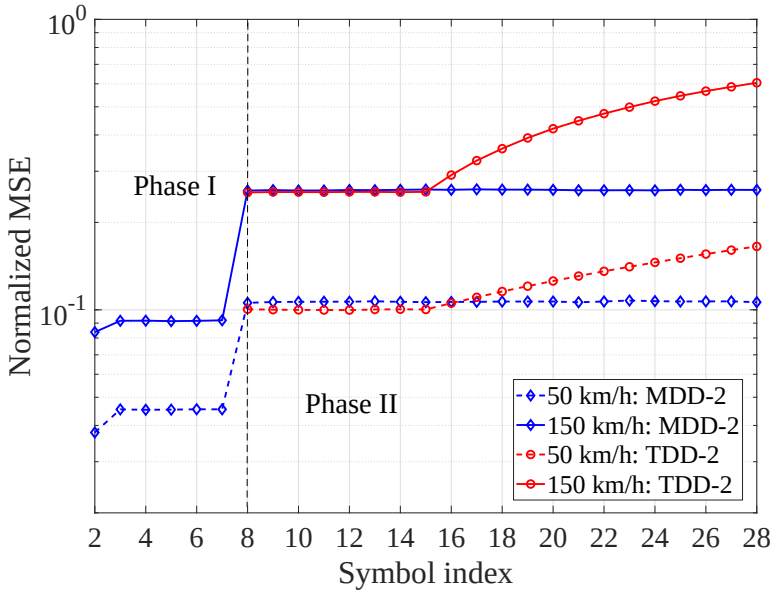
(A) $T = 28$



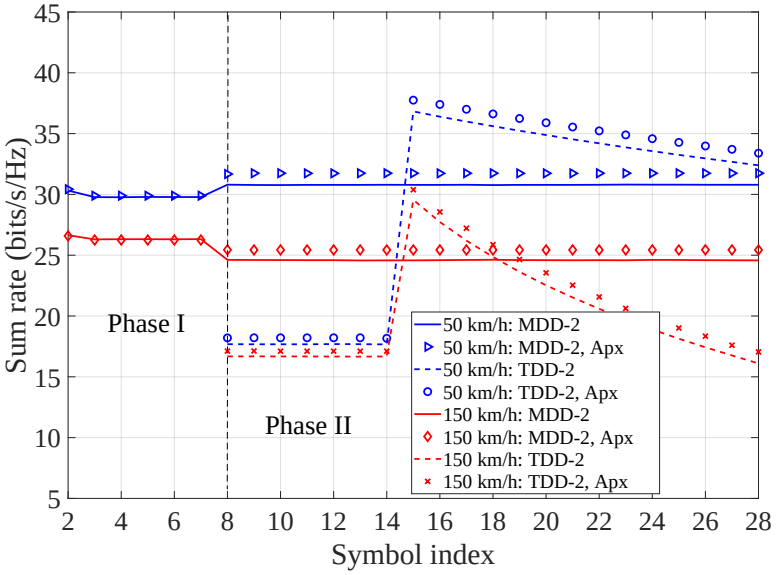
(B) $T = 56$

FIGURE 4.7: Average sum rate versus relative velocity, when the Type I frame structure is used.

In Fig. 4.9, we compare the average sum rate versus the relative speed of the MDD and TDD systems, when the Type II frame structure is employed. Here, we also consider an optional frame structure for TDD systems, which is shown in Fig. 4.10, where two groups of pilots are separately inserted into one frame. We can observe from Fig. 4.9 that owing to the DL transmission occurred in both Phase I and the start of Phase II, MDD significantly outperforms TDD over the whole velocity, ranging from 20 km/h to 300 km/h, for both the frame-length considered. Specifically, for the TDD in low-mobility case, the rate achieved by the TDD-2-TG is considerably lower than that achieved by the TDD-2. However, benefiting from the extra training in the middle of the frame, the TDD-2-TG finally surpasses the TDD-2 at 240 km/h when $T = 28$,



(A) Normalized MSE



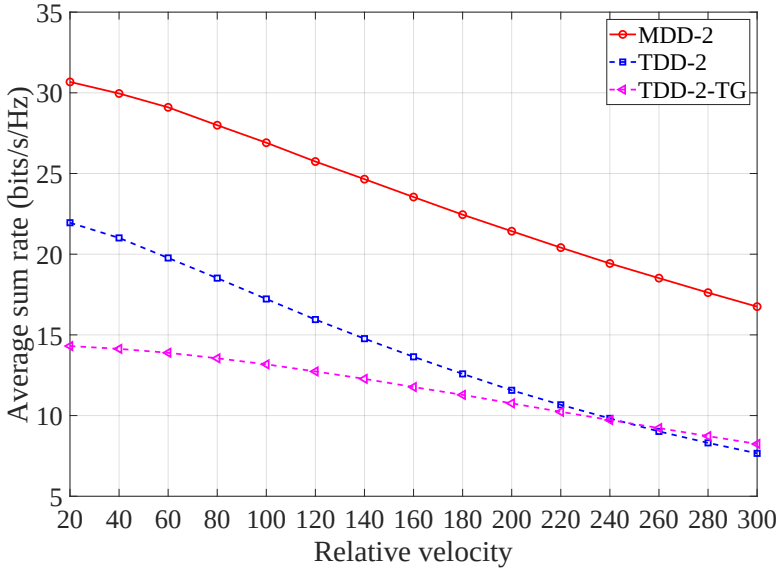
(B) Sum rate

FIGURE 4.8: Performance comparison of the MDD and TDD systems, when Type II frame structure is applied.

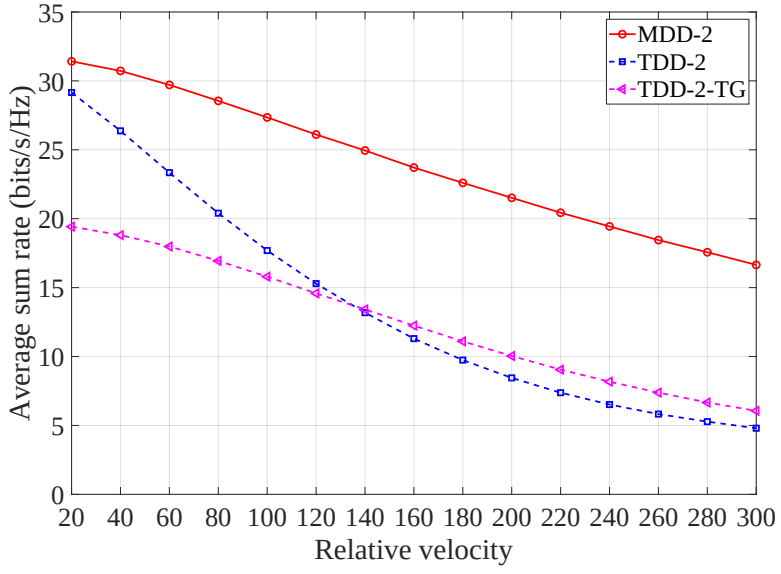
and at 140 km/h when $T = 56$. The results in Fig. 4.9 imply that in the general communication scenarios where UL training, UL transmission and DL transmissions are all required at the same time, MDD shows significant advantages over TDD, when the channel aging problem is encountered.

4.6.3 Influence of Self-Interference Cancellation on IBFD and MDD

In this part, we demonstrate the impact of SIC on the performance of MDD and IBFD systems. To this end, we consider the conventional IBFD with the Type I frame structure



(A) $T = 28$



(B) $T = 56$

FIGURE 4.9: Average sum rate versus the relative velocity, when the Type II frame structure is used.

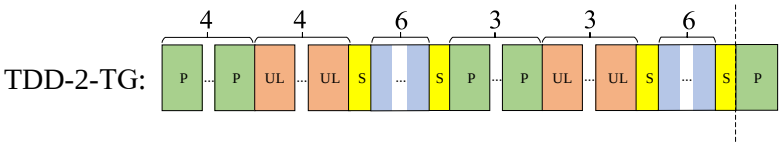


FIGURE 4.10: Distribution of two groups of pilots in TDD Type II frame structure, when the total number of pilots is 7.

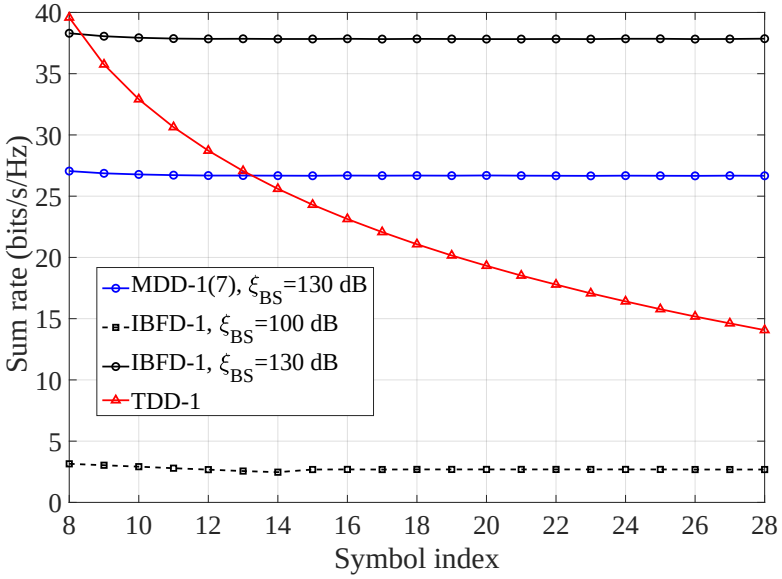


FIGURE 4.11: Sum rate comparison of MDD, IBFD and TDD with respect to OFDM symbol index, when Type I frame structure and the relative velocity of 150 km/h are assumed.

denoted as 'IBFD-1', where UL pilots are unremittingly transmitted over $\bar{M}$ subcarriers and a 7-th order WP is applied for channel prediction. Note that, unlike the MDD-1 that only uses M subcarriers for DL transmission, as stated in Section 4.6.1, IBFD-1 uses all the M_{sum} subcarriers for the DL transmission, which imposes an extra SIC burden in digital-domain. According to [21], the digital-domain SIC in IBFD systems is expected to cancel the main linear signal component by more than 30 dB. By contrast, MDD systems are almost free from linear SI in the digital-domain owing to the FFT operation, which separates the UL subcarriers from the DL subcarriers. Based on this observation, it is sensible to assume that with the same amount of system resources used for SI suppression, MDD can always achieve 30 dB more SIC than IBFD. From Fig. 4.11, it is shown that when both IBFD and MDD systems have the same SI mitigation capability[5], IBFD apparently outperforms both TDD and MDD. However, if IBFD can only mitigate the SI by 100 dB[6], the achievable sum rate of IBFD systems drops dramatically, and is much lower than that of MDD systems.

4.7 Chapter Summary and Conclusions

An MDD-assisted mMIMO has been studied and comprehensively compared with the TDD-relied mMIMO over fast time-varying channels. Two types of MDD frame structures incorporated respectively with WP and DD-WP have been proposed to combat

[5]This means IBFD costs more computation effort and system complexity than MDD for SIC, and hence achieves lower energy efficient [73].

[6]Note that 100 dB of SI is often assumed in IBFD systems, which is regarded as the bottleneck value based on the existing SIC approaches [145, 146].

channel aging. For comparison, the frames in TDD mMIMO have been structured according to the 3GPP standards or improved for handling the channel aging problem. MDD has also been compared with the IBFD by considering their SIC capabilities. Moreover, the approximated lower bounds for the achievable rates of the MDD and TDD mMIMO systems communicating based on the proposed frame structures have been derived. According to our studies and performance comparison, we have the following main observations. First, when the Type I frame structure is employed, benefiting from the instantly available CSI updating while without involving UL/DL switching, MDD outperforms TDD over time-varying channels, which is more significant when mobility becomes higher. Second, the above performance advantage of MDD over TDD retains under the Type II frame structure, owing to the MDD's merits that it can proactively start DL transmission in Phase I, while simultaneously updating the CSI with the aid of the UL transmission in Phase II. Third, it is shown that SIC is critical to IBFD systems. By taking the advantage of near-free SI in the digital-domain, MDD is capable of achieving better performance than IBFD in the presence of channel aging, if SIC in IBFD systems is imperfect. Additionally, the approximated lower bounded rates derived in this chapter have been validated by the Monte-Carlo simulations, showing that they agree well with each other.

Chapter 5

MDD Assisted Cell-Free Massive MIMO

This chapter studies the spectral efficiency (SE) of multicarrier-division duplex assisted cell-free (MDD-CF) massive multiple-input multiple-output (mMIMO), with the emphasis on the advantages of MDD over time-division duplex (TDD) and in-band full duplex (IBFD) when they are operated in CF systems.

5.1 Introduction

As one of the promising techniques integrating the advantages of cloud radio access network (C-RAN), mMIMO and coordinated multipoint (COMP), CF-mMIMO has attracted the growing attention from both academia and industry in recent years [75]. The studies have demonstrated that CF-mMIMO can efficiently suppress intra-cell interference with the aid of the physical layer techniques, such as beamforming in the collocated mMIMO. Furthermore, the inter-cell interference, which imposes the main limit on the performance of the cell-edge mobile stations (MSs) in cellular networks, is no longer intractable in the CF-mMIMO systems, as the result of no cell boundaries, and hence leading to the possible seamless coverage [78].

As it is evolved from the collocated mMIMO, CF-mMIMO systems still consider TDD as the dominant duplexing mode [10,171]. However, TDD has its limitation, especially, when implemented in fast time-varying environment. First, in TDD, the guard period (GP) between downlink (DL) and uplink (UL) transmissions is indispensable. In this regard, when communication channels change quickly in high-mobility scenarios, UL training has to be frequently performed due to the shorter coherence time (CT) interval, and hence the portion of time for data transmission decreases [73]. Second, the successive transmission of DL and UL represses the explosive demand of asymmetric

communications [87]. For instance, several users having high demand on DL data may have to wait for a single user with a low demand on UL data, which in turn causes the low utilization efficiency of the time-frequency resources. To mitigate the disadvantages of TDD-CF, several full duplex (FD)-style CF systems have been proposed in the last few years, such as IBFD-CF, dynamic time-division duplex (DTDD)-CF and network-assisted full duplex (NAFD)-CF. However, according to our brief overview and analysis in Section 1.3.2, these CF schemes more or less have some defects, which hinder them from practical implementation. For instance, none of them appropriately considered the cross-link interference (CLI) suppression in analog domain. Besides, to the authors' best knowledge, all the existing FD-style CF schemes operate in a centralized manner, where the central processor unit (CPU) conducts all signal processing tasks, while access points (APs) only transmit precoded DL signals from the CPU or pass the received UL signal to the CPU. In this case, the suppression of CLI imposes heavy burden on the CPU as well as fronthaul capability. Despite the better scalability and less dependence on fronthaul, there are few research about the distributed CF systems. The reason behind is that the suppression of CLI can be highly intractable when APs work independently without sharing channel information with the CPU.

To alleviate the problem of CLI including inter-AP interference (IAI) and inter-MS interference (IMI), but maximize the time-frequency resource usage at the same time, in this chapter, we propose an MDD-CF scheme. According to the principles of MDD, DL and UL transmissions take place at the same time but on different subcarriers. Therefore, the IAI/IMI are mutually orthogonal with the desired UL/DL signals in the digital domain, and can be easily removed during reception without any additional system overhead. On the other hand, owing to the FD-style operation, MDD allows to concurrently implement DL transmission and UL training, which improves the SE even over fast time-varying channels. To leverage these advantages, however, one of the challenges is how to efficiently allocate the power and time-frequency resources among multiple APs and MSs for simultaneously supporting data transmission or training in two directions. To the best of our knowledge, there is no study in the open literature that has considered this kind of optimization problem. Therefore, against the background, we comprehensively study the AP-selection, power allocation (PA) and subcarrier allocation (SA) issues in the MDD-CF, with the motivation to achieve the optimal SE performance. In summary, the novelties and contributions of this chapter can be briefly described as follows:

Firstly, in order to mitigate the interference to enable the FD-style operation in distributed CF-mMIMO, the MDD-CF scheme is proposed, where the effect of SI, IAI and IMI are practically modeled. Additionally, to study the optimization of AP-selection, PA and SA under the constraint of MSs' quality of service (QoS), two application scenarios, which assume that DL/UL transmissions occur in one CT interval or in one radio frame, are considered.

Secondly, in the case of DL/UL transmissions in one CT interval, we firstly leverage the inner association between the continuous variables for PA, and binary variables for AP-selection and SA, to change the mixed-integer optimization into a continuous-integer convex-concave problem. Then, a quadratic transform with successive convex approximation (QT-SCA) iterative algorithm is proposed to achieve the SE maximization in MDD-CF scheme.

Thirdly, in the case of DL/UL transmissions in one radio frame, we consider imperfect CE. A two-phase CT (TPCT) interval is designed for the CF systems operated in the MDD and IBFD modes. Since the two phases are tightly coupled and both of them support simultaneous transmissions in two directions, they lead to a very intricate formulation. Correspondingly, we introduce an iteration factor to build the connection between the two phases, and transform the original problem to a two-step iterative optimization with the aid of the bisection method.

Finally, we comprehensively compare MDD-, IBFD- and TDD-CF in distributed CF-mMIMO systems under the practical network settings. Our simulation results demonstrate the superiority of MDD-CF over IBFD-CF, due to more effective suppression of IAI and IMI in the digital domain. Furthermore, the well-designed TPCT interval with the proposed two-step iterative algorithm enables MDD-CF to achieve much higher SE than TDD-CF in high-mobility communication scenarios.

The remainder of this chapter is organized as follows. In Section 5.2, our MDD-CF scheme is proposed. In Section 5.3, we maximize the SE over one CT interval, considering AP-section, PA and SA. In Section 5.4, we extend our analysis to the case of radio frame and propose a novel algorithm to optimize the SE performance. The simulation results are presented in Section 5.5 and chapter is concluded in Section 5.6.

5.2 System Model

We consider an MDD-CF scheme as shown in Fig. 5.1, where the set $\mathcal{D} = \{1, ..., d, ..., D\}$ of single-antenna MSs and the set $\mathcal{L} = \{1, ..., l, ..., L\}$ of APs of each with N antennas are operated in the MDD mode that rely on the mutually orthogonal subcarrier sets [69], namely $\mathcal{M} = \{1, ..., m, ..., M\}$ with $|\mathcal{M}| = M$ and $\bar{\mathcal{M}} = \{1, ..., \bar{m}, ..., \bar{M}\}$ with $|\bar{\mathcal{M}}| = \bar{M}$, for DL and UL, respectively. The total number of subcarriers is $M_{\text{sum}} = M + \bar{M}$. Furthermore, we assume that the CF system is operated in a distributed way, where CPU offloads most of the tasks to APs to relieve its computation burden, and only sends coded data to APs for DL transmissions or integrates the received UL data from APs via fronthaul links without any knowledge of channel information. We assume that pilots and DL/UL data are transmitted in a CT interval T_{c} in terms of OFDM symbols. As shown in Fig. 5.2, in the TDD mode, which is deemed as the mainstream duplex mode in mMIMO systems [10], signal transmissions are performed in sequence, where

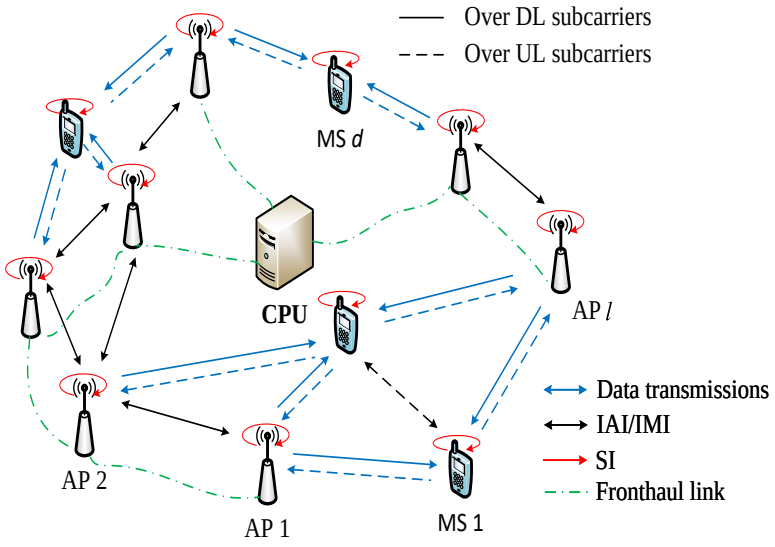


FIGURE 5.1: Illustration of MDD-CF scheme.

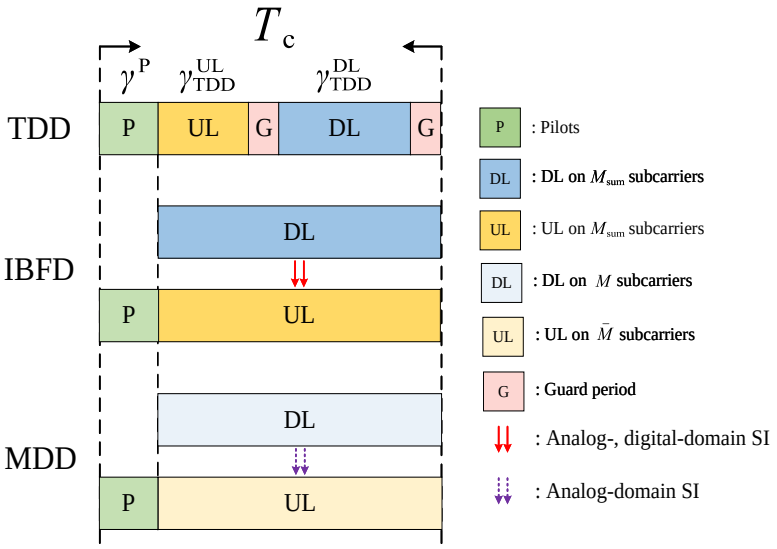


FIGURE 5.2: Transmission in a coherence time interval of TDD-, IBFD- and MDD-schemes.

UL training, GP and DL/UL transmissions require γ^{P}, γ^{G}, $\gamma^{\mathrm{DL}}_{\mathrm{TDD}}$ and $\gamma^{\mathrm{UL}}_{\mathrm{TDD}}$ symbol durations, respectively. Note that, in the TDD mode, the GP between UL and DL is indispensable. By contrast, in both IBFD and MDD modes, after the same training time γ^{P} as that in the TDD mode, DL and UL data transmission time are $T_{\mathrm{c}} - \gamma^{\mathrm{P}}$, which can be much higher than that in the TDD mode.

5.2.1 Channel Model

For the convenience of notation, we denote the SI channel at the l-th AP and d-th MS by $\boldsymbol{H}_{ll} \in \mathbb{C}^{N\times N}$ and h_{dd}, respectively. These two SI channels are modeled as

$$\begin{aligned}(\boldsymbol{H}_{ll})_{i,j} &= \sqrt{\xi_l^{\mathrm{SI}}}\,\alpha_s,\\ h_{dd} &= \sqrt{\xi_d^{\mathrm{SI}}}\,\alpha_s,\end{aligned} \tag{5.1}$$

where $\alpha_s \sim \mathcal{CN}(0,1)$ is the small-scale fading, ξ_l^{SI} and $\xi_d^{\mathrm{SI}} \in (0,1]$ denote the residual SI levels at AP and MS receivers, respectively. We assume that $\xi_l^{\mathrm{SI}} < \xi_d^{\mathrm{SI}}$, meaning that the SI at AP receiver can be mitigated to a lower level than that at MS receiver. This is because APs are capable of employing more complicated circuits and higher resource budget for SI suppression than MSs.

Remark 5.1. We consider two assumptions for the SI channels. 1) Since the SI link is relatively short in comparison to the AP-MS link, we assume SI link is of single-path without suffering from the large-scale fading; 2) The residual SI level ($\xi_l^{\mathrm{SI}}/\xi_d^{\mathrm{SI}}$) relies on the SIC capability, which accounts for the propagation/analog- and digital-domain SIC methods[1], such as dual-port polarized antenna [25], antenna circulator [21] and multi-tap RF canceller [39].

Furthermore, we denote the time-domain CIRs of the communication channels between the d-th MS and the n-th antenna at the l-th AP, the IAI channels between the n-th antenna at the l-th AP and the n'-th antenna at the l'-th AP, and the IMI channels between the d-th MS and d'-th MS by $\boldsymbol{g}_{ld}^{n} \in \mathbb{C}^{U\times 1}$, $\boldsymbol{g}_{ll'}^{nn'} \in \mathbb{C}^{U\times 1}$ and $\boldsymbol{g}_{dd'} \in \mathbb{C}^{U\times 1}$, respectively, where U is the number of taps of multipath channels. Specifically, the u-th tap of these channels can be generally modeled as $(\boldsymbol{g})_u = \sqrt{\beta/U}\,\alpha_s$ with $\boldsymbol{g} \in \left\{\boldsymbol{g}_{ld}^{n}, \boldsymbol{g}_{ll'}^{nn'}, \boldsymbol{g}_{dd'}\right\}$, where $\beta \in \{\beta_{ld}, \beta_{ll'}, \beta_{dd'}\}$ accounts for the large-scale fading of path loss and shadowing. The channels of different taps are assumed to be independent. Additionally, the spatial correlation among the antennas of AP is not considered in this chapter, which will be addressed in the future research.

Given the time-domain CIRs, the frequency-domain channels can be obtained as $\boldsymbol{h} = \boldsymbol{F}\boldsymbol{\Psi}\boldsymbol{g}$ with $\boldsymbol{h} \in \left\{\boldsymbol{h}_{ld}^{n}, \boldsymbol{h}_{ll'}^{nn'}, \boldsymbol{h}_{dd'}\right\}$, where $\boldsymbol{F} \in \mathbb{C}^{M_{\mathrm{sum}}\times M_{\mathrm{sum}}}$ is the FFT matrix, $\boldsymbol{\Psi} \in \mathbb{C}^{M_{\mathrm{sum}}\times U}$ is constructed by the first U columns of $\boldsymbol{I}_{M_{\mathrm{sum}}}$. Moreover, the single DL/UL subcarrier channel can be expressed as $h[m] = \boldsymbol{\phi}_{\mathrm{DL}}^{T}\boldsymbol{h}$ and $h[\bar{m}] = \boldsymbol{\phi}_{\mathrm{UL}}^{T}\boldsymbol{h}$, respectively, where $\boldsymbol{\phi}_{\mathrm{DL}} = \boldsymbol{I}_{M_{\mathrm{sum}}}^{(:,m)}$ and $\boldsymbol{\phi}_{\mathrm{UL}} = \boldsymbol{I}_{M_{\mathrm{sum}}}^{(:,\bar{m})}$ are the mapping vectors. Note that here $h[m]$ or $h[\bar{m}]$ denotes the point-to-point subcarrier channel, which will be further integrated into the vector

[1] As we mentioned before, MDD systems can be nearly free from the SI in digital domain due to the FFT operation. By contrast, in order to mitigate the digital-domain SI in IBFD-based systems, the receiver has to estimate the channel between DAC and ADC, and then reconstruct the transmitted signal, which is subsequently subtracted from the received signal [42, 43]. This process incurs large overhead, especially when the system is operated with a large number of subcarriers.

and matrix for the AP-AP channel (i.e., $\boldsymbol{H}_{ll'}[m \text{ or } \bar{m}] \in \mathbb{C}^{N\times N}$), AP-MS channel (i.e., $\boldsymbol{h}_{ld}[m] \in \mathbb{C}^{N\times 1}$ and $\boldsymbol{h}_{ld}[\bar{m}] \in \mathbb{C}^{N\times 1}$), respectively, as shown later in Section 5.2.B.

5.2.2 Downlink Transmission

Within each CT interval, the data transmitted on the m-th DL subcarrier for the d-th MS is denoted by $x_d[m]$, which satisfies $\mathbb{E}\left\{|x_d[m]|^2\right\} = 1$. The transmitted signal on the m-th DL subcarrier by the l-th AP is given by

$$\boldsymbol{s}_l[m] = \sum_{d\in\mathcal{D}} \lambda_{ld}\mu_{ldm}\sqrt{p_{ldm}}\boldsymbol{f}_{ld}[m]x_d[m], \tag{5.2}$$

where the binary variable λ_{ld} denotes the association relationship between the l-th AP and the d-th MS, with $\lambda_{ld} = 1$ expressing that the d-th MS is associated with the l-th AP and $\lambda_{ld} = 0$, otherwise. The binary variable μ_{ldm} explains the operation status of MS d on the m-th DL subcarrier, with $\mu_{ldm} = 1$ implying that the m-th DL subcarrier is activated by the l-th AP for DL transmission to MS d and $\mu_{ldm} = 0$, otherwise. In (5.2), $\boldsymbol{f}_{ld}[m] \in \mathbb{C}^{N\times 1}$ is the precoding vector with $\|\boldsymbol{f}_{ld}[m]\|_2^2 = 1$, and p_{ldm} is the power allocated to the m-th subcarrier of the d-th MS by the l-th AP. The total power budget at the l-th AP is expressed as P_l, satisfying $\sum_{m\in\mathcal{M}}\sum_{d\in\mathcal{D}} \lambda_{ld}\mu_{ldm}p_{ldm} \leq P_l$.

The signal received from the m-th DL subcarrier at the d-th MS can be expressed as

$$y_d[m] = \underbrace{\sum_{l\in\mathcal{L}} \boldsymbol{h}_{ld}^H[m]\boldsymbol{s}_l[m]}_{\text{Desired signal + MUI}} + z_d^{\text{SI}} + z_d^{\text{IMI}} + n_d, \tag{5.3}$$

where $n_d \sim \mathcal{CN}(0,\sigma^2)$ is the additive white Gaussian noise. According to [146, 148], the residual interference in digital domain arising from SI and IMI $\left(\text{i.e., } z_d^{\text{SI}} \text{ and } z_d^{\text{IMI}} \text{ in } (5.3)\right)$ are modeled as Gaussian noise [148]. Specifically, $z_d^{\text{SI}} \sim \mathcal{CN}(0, \mathbb{E}\left[\bar{z}_d^{\text{SI}}\left(\bar{z}_d^{\text{SI}}\right)^*\right])$ with $\bar{z}_d^{\text{SI}} = h_{dd}\sum_{\bar{m}\in\bar{\mathcal{M}}}\mu_{d\bar{m}}\sqrt{p_{d\bar{m}}}x_d[\bar{m}]$, where $x_d[\bar{m}]$ denotes the data transmitted on the $\bar{m}$-th UL subcarrier by the d-th MS, $p_{d\bar{m}}$ denotes the transmitted power. $z_d^{\text{IMI}} \sim \mathcal{CN}(0, \xi_d^{\text{IMI}}\mathbb{E}\left[\bar{z}_d^{\text{IMI}}\left(\bar{z}_d^{\text{IMI}}\right)^*\right])$ with $\bar{z}_d^{\text{IMI}} = \sum_{d'\in\mathcal{D}\setminus\{d\}}\sum_{\bar{m}\in\bar{\mathcal{M}}}\mu_{d'\bar{m}}\sqrt{p_{d'\bar{m}}}h_{dd'}[\bar{m}]x_{d'}[\bar{m}]$, where ξ_d^{IMI} denotes the residual IMI level at MS d.

Based on (5.3), it can be shown that the received SINR on the m-th DL subcarrier at the d-th MS is given by

$$\text{SINR}_{d,m} = \frac{\left|\sum_{l\in\mathcal{L}}\lambda_{ld}\mu_{ldm}\sqrt{p_{ldm}}\boldsymbol{h}_{ld}^H[m]\boldsymbol{f}_{ld}[m]\right|^2}{\text{MUI}_{d,m} + \text{var}\left\{z_d^{\text{SI}}\right\} + \text{var}\left\{z_d^{\text{IMI}}\right\} + \sigma^2}, \tag{5.4}$$

where

$$\text{MUI}_{d,m} = \sum_{l\in\mathcal{L}}\sum_{d'\in\mathcal{D}\setminus\{d\}} \lambda_{ld'}\mu_{ld'm}p_{ld'm}\left|\boldsymbol{h}_{ld}^H[m]\boldsymbol{f}_{ld'}[m]\right|^2. \tag{5.5}$$

5.2.3 Uplink Transmission

The received UL signal by the l-th AP from the $\bar{m}$ subcarrier of MS d can be expressed as

$$\boldsymbol{y}_l[\bar{m}] = \underbrace{\sum_{d\in\mathcal{D}} \mu_{d\bar{m}}\sqrt{p_{d\bar{m}}}\boldsymbol{h}_{ld}[\bar{m}]x_d[\bar{m}]}_{\text{Desired signal + MUI}} + \boldsymbol{z}_l^{\text{SI}} + \boldsymbol{z}_l^{\text{IAI}} + \boldsymbol{n}_l, \tag{5.6}$$

where $p_{d\bar{m}}$ denotes the power allocated by MS d to the $\bar{m}$-th UL subcarrier, which satisfies $\sum_{\bar{m}\in\mathcal{M}} \mu_{d\bar{m}} p_{d\bar{m}} \leq P_d$. Similar to the received signals at MSs, the residual interference due to the SI and IAI are modeled as Gaussian noise, where

$$\begin{aligned}
\boldsymbol{z}_l^{\text{SI}} &\sim \mathcal{CN}\left(0, \text{diag}\left(\mathbb{E}\left[\bar{\boldsymbol{z}}_l^{\text{SI}}\left(\bar{\boldsymbol{z}}_l^{\text{SI}}\right)^H\right]\right)\right), \\
\boldsymbol{z}_l^{\text{IAI}} &\sim \mathcal{CN}\left(0, \xi_l^{\text{IAI}}\text{diag}\left(\mathbb{E}\left[\bar{\boldsymbol{z}}_l^{\text{IAI}}\left(\bar{\boldsymbol{z}}_l^{\text{IAI}}\right)^H\right]\right)\right)
\end{aligned} \tag{5.7}$$

with $\bar{\boldsymbol{z}}_l^{\text{SI}} = \boldsymbol{H}_{ll}\sum_{m\in\mathcal{M}}\boldsymbol{s}_l[m]$ and $\bar{\boldsymbol{z}}_l^{\text{IAI}} = \sum_{l'\in\mathcal{L}\setminus\{l\}}\sum_{m\in\mathcal{M}}\boldsymbol{H}_{ll'}[m]\boldsymbol{s}_{l'}[m]$, where ξ_l^{IAI} denotes the residual IAI level at the l-th AP.

Remark 5.2. $\boldsymbol{z}_d^{\text{IMI}}$ and $\boldsymbol{z}_l^{\text{IAI}}$ in (5.3) and (5.6) are mainly attributed to the large-scale fading of the interfering links, PA and the residual IAI/IMI after the supplementary mitigation of IAI/IMI in propagation/analog- and digital-domain. In distributed CF systems, since each AP works independently, the propagation/analog-domain IAI suppression can hardly rely on the coordinated transmit beamforming and successive interference cancellation [86, 172]. Furthermore, the passive methods, such as antenna cross-polarization, beam separation and absorber can only provide a small part of IAI suppression [72]. Additionally, due to the employment of single-antenna, the MSs with less powerful baseband processor may fail to handle any IMI in the propagation/analog domain. However, as seen in (5.3) and (5.6), in our porposed MDD-CF scheme, the IAI and IMI are mutually orthogonal to the desired UL and DL signals, respectively, in digital domain. Therefore, aided by the large-scale fading and PA to limit the received signal within the effective dynamic range of ADC, the followed FFT operation in digital domain can provide extra mitigation of IAI/IMI [23].

We should mention that like any other multicarrier systems, the FFT operation for the mitigation of residual SI and IAI/IMI requires accurate time synchronization [68]. As we know, it is relatively easy to achieve time synchronization between transmitter and receiver at one side or between different APs via low-latency fronthauls [173], while the synchronization between MSs is challenging and the excessive time synchronization error beyond the allowable time window may degrade the performance of the FFT-relied interference cancellation. However, this issue is beyond the scope of this chapter, and will be studied in our future work.

Due to the feature of distributed operation in our proposed system, each AP firstly processes the received signals from MSs using the local combining vectors, yielding $\tilde{y}_l[\bar{m}] = \boldsymbol{w}_{ld}^H[\bar{m}]\boldsymbol{y}_l[\bar{m}]$, where $\boldsymbol{w}_{ld}[\bar{m}]$ denotes the local combining vector of AP l for detecting MS d. Then, the local estimated data by all APs are further collected by the CPU for final processing, which can be expressed as $y_{\text{cpu}}[\bar{m}] = \sum_{l\in\mathcal{L}} \tilde{y}_l[\bar{m}]$. The SINR obtained by the CPU for detecting the data transmitted on the UL subcarrier $\bar{m}$ of MS d can be expressed as

$$\text{SINR}_{d,\bar{m}} = \frac{\mu_{d\bar{m}} p_{d\bar{m}} \left|\tilde{\boldsymbol{w}}_d[\bar{m}]\tilde{\boldsymbol{h}}_d[\bar{m}]\right|^2}{\text{MUI}_{d,\bar{m}} + \text{SI}_{d,\bar{m}} + \text{IAI}_{d,\bar{m}} + \sigma^2 \left\|\tilde{\boldsymbol{w}}_d[\bar{m}]\right\|^2}, \tag{5.8}$$

where $\tilde{\boldsymbol{w}}_d[\bar{m}] = \left[\boldsymbol{w}_{1d}^H[\bar{m}], ..., \boldsymbol{w}_{Ld}^H[\bar{m}]\right] \in \mathbb{C}^{1\times NL}$, $\tilde{\boldsymbol{h}}_d[\bar{m}] = [\boldsymbol{h}_{1d}^H[\bar{m}], ..., \boldsymbol{h}_{Ld}^H[\bar{m}]]^H \in \mathbb{C}^{NL\times 1}$, $\text{MUI}_{d,\bar{m}} = \sum_{d'\in\mathcal{D}\setminus\{d\}} \mu_{d'\bar{m}} p_{d'\bar{m}} \left|\tilde{\boldsymbol{w}}_d[\bar{m}]\tilde{\boldsymbol{h}}_{d'}[\bar{m}]\right|^2$, $\text{SI}_{d,\bar{m}} = \sum_{l\in\mathcal{L}} \mathbb{E}\left[\left\|\boldsymbol{w}_{ld}^H[\bar{m}]\boldsymbol{z}_l^{\text{SI}}\right\|^2\right]$, $\text{IAI}_{d,\bar{m}} = \sum_{l\in\mathcal{L}} \mathbb{E}\left[\left\|\boldsymbol{w}_{ld}^H[\bar{m}]\boldsymbol{z}_l^{\text{IAI}}\right\|^2\right]$.

5.2.4 Beamforming Strategy

In this chapter, the zero-forcing (ZF) beamforming strategy is chosen for transmitting and receiving at APs. Generally speaking, MMSE beamforming outperforms ZF beamforming when perfect CSI is available, but when considering the multi-MS interfernece suppression, computation complexity as well as concise formulation, ZF is applied in the following analysis, and it can be easily substituted by MMSE in our proposed system. Based on the ZF principle [174], the precoder/combiner at the l-th AP, i.e., $\boldsymbol{F}_l^{\text{ZF}}[m] = \left[\boldsymbol{f}_{l1}^{\text{ZF}}[m], ..., \boldsymbol{f}_{lD}^{\text{ZF}}[m]\right]$ and $\boldsymbol{W}_l^{\text{ZF}}[\bar{m}] = \left[\boldsymbol{w}_{l1}^{\text{ZF}}[\bar{m}], ..., \boldsymbol{w}_{lD}^{\text{ZF}}[\bar{m}]\right]$, can be derived as $\boldsymbol{F}_l^{\text{ZF}}[m] = \boldsymbol{H}_l^H[m]\left(\boldsymbol{H}_l[m]\boldsymbol{H}_l^H[m]\right)^{-1}$ and $\boldsymbol{W}_l^{\text{ZF}}[\bar{m}] = \boldsymbol{H}_l[\bar{m}]\left(\boldsymbol{H}_l^H[\bar{m}]\boldsymbol{H}_l[\bar{m}]\right)^{-1}$, respectively, where $\boldsymbol{H}_l[m] = [\boldsymbol{h}_{l1}[m], ..., \boldsymbol{h}_{lD}[m]]^H$, $\boldsymbol{H}_l[\bar{m}] = [\boldsymbol{h}_{l1}[\bar{m}], ..., \boldsymbol{h}_{lD}[\bar{m}]]$. Note that, in order to ensure that the MUI is fully suppressed, the implementation of ZF beamforming should adhere to the constraint of $N \geq D$.[2] In this case, the MUI terms in (5.4) and (5.8) are equal to zero. Therefore, the $\text{SINR}_{d,m}$ and $\text{SINR}_{d,\bar{m}}$ can be rewritten as follows

$$\begin{aligned}
\text{SINR}_{d,m} &= \frac{\left|\sum_{l\in\mathcal{L}} \lambda_{ld}\mu_{ldm}\sqrt{p_{ldm}}\omega_{ldm}\right|^2}{\xi_d^{\text{SI}}\Theta_{\text{DL}} + \sigma^2}, \\
\text{SINR}_{d,\bar{m}} &= \frac{\mu_{d\bar{m}} p_{d\bar{m}} L^2}{\sum_{l\in\mathcal{L}} v_{ld\bar{m}}\left(\xi_l^{\text{SI}}\Theta_{\text{UL}} + \sigma^2\right)},
\end{aligned} \tag{5.9}$$

[2]For the sake of convenience, we assume that each AP is employed with sufficient antennas so as to suppress the interference that itself generates. Although an AP is expected to be equipped with a small number of antennas in CF systems, our assumption is still practical, as each AP can be treated as a secondary central unit controlling N single-antenna APs operated in a centralized mode through fronthaul connections.

where

$$\begin{aligned}
\tilde{\boldsymbol{w}}_d[\bar{m}] &= \left[\boldsymbol{w}_{1d}^H[\bar{m}], ..., \boldsymbol{w}_{Ld}^H[\bar{m}]\right] \in \mathbb{C}^{1\times NL}, \\
\tilde{\boldsymbol{h}}_d[\bar{m}] &= [\boldsymbol{h}_{1d}^H[\bar{m}], ..., \boldsymbol{h}_{Ld}^H[\bar{m}]]^H \in \mathbb{C}^{NL\times 1}, \\
\mathrm{MUI}_{d,\bar{m}} &= \sum_{d'\in\mathcal{D}\setminus\{d\}} \mu_{d'\bar{m}} p_{d'\bar{m}} \left|\tilde{\boldsymbol{w}}_d[\bar{m}]\tilde{\boldsymbol{h}}_{d'}[\bar{m}]\right|^2, \\
\mathrm{SI}_{d,\bar{m}} &= \sum_{l\in\mathcal{L}} \mathbb{E}\left[\left\|\boldsymbol{w}_{ld}^H[\bar{m}]\boldsymbol{z}_l^{\mathrm{SI}}\right\|^2\right], \\
\mathrm{IAI}_{d,\bar{m}} &= \sum_{l\in\mathcal{L}} \mathbb{E}\left[\left\|\boldsymbol{w}_{ld}^H[\bar{m}]\boldsymbol{z}_l^{\mathrm{IAI}}\right\|^2\right].
\end{aligned} \tag{5.10}$$

For the details of simplification, please refer to Appendix E.

Consequently, the average SE of the MDD-CF scheme in nats/s/Hz can be expressed as

$$\Lambda_{\mathrm{SE}} = \left(1 - \frac{\gamma^{\mathrm{P}}}{T_{\mathrm{c}}}\right) \frac{1}{M_{\mathrm{sum}}} \sum_{d\in\mathcal{D}} \Big(\sum_{m\in\mathcal{M}} R(\mathrm{SINR}_{d,m}) + \sum_{\bar{m}\in\bar{\mathcal{M}}} R(\mathrm{SINR}_{d,\bar{m}}) \Big), \tag{5.11}$$

where $R(x) \triangleq ln(1+x)$.

5.3 Spectral Efficiency Optimization within Single Coherence Time Interval

In this section, we aim to maximize the SE over one CT interval in the MDD-CF scheme, as shown in Fig. 5.2. Given the ZF beamforming, the optimization problem can be stated as:

$$\max_{\lambda_{ld},\mu_{ldm},\mu_{d\bar{m}},p_{ldm},p_{d\bar{m}}} \quad \Lambda_{\mathrm{SE}} \tag{5.12a}$$

$$\text{s.t.} \quad \lambda_{ld} \in \{0,1\}, \ \forall l \in \mathcal{L}, d \in \mathcal{D}, \tag{5.12b}$$

$$\mu_{ldm} \in \{0,1\}, \ \forall l \in \mathcal{L}, d \in \mathcal{D}, m \in \mathcal{M}, \tag{5.12c}$$

$$\mu_{d\bar{m}} \in \{0,1\}, \ \forall d \in \mathcal{D}, \bar{m} \in \bar{\mathcal{M}}, \tag{5.12d}$$

$$\sum_{m\in\mathcal{M}} \sum_{d\in\mathcal{D}} \lambda_{ld}\mu_{ldm}p_{ldm} \le P_l, \ \forall l \in \mathcal{L}, \tag{5.12e}$$

$$\sum_{\bar{m}\in\bar{\mathcal{M}}} \mu_{d\bar{m}}p_{d\bar{m}} \le P_d, \ \forall d \in \mathcal{D}, \tag{5.12f}$$

$$\sum_{m\in\mathcal{M}} R(\mathrm{SINR}_{d,m}) \ge \chi_{\mathrm{DL}}, \ \forall d \in \mathcal{D}, \tag{5.12g}$$

$$\sum_{\bar{m}\in\bar{\mathcal{M}}} R(\mathrm{SINR}_{d,\bar{m}}) \ge \chi_{\mathrm{UL}}, \ \forall d \in \mathcal{D}, \tag{5.12h}$$

where the constraints of (5.12g) and (5.12h) are applied to guarantee the MS's QoS requirements for DL and UL so as to avoid unbalanced greedy resource-allocation among MSs. It can be observed that (5.12) is an optimization problem of AP-selection and resource-allocation, which is hard to solve because the binary variables $\big($i.e., $\boldsymbol{\lambda} = \{\lambda_{ld}\}, \boldsymbol{\mu} = (\{\mu_{ldm}\}, \{\mu_{d\bar{m}}\})\big)$ are tightly coupled with the continuous variables $\boldsymbol{p} = (\{p_{ldm}\}, \{p_{d\bar{m}}\})$. To circumvent this problem, in what follows, we first focus on the reduction and approximation of the involved binary variables.

5.3.1 Reduction of Binary Variables

5.3.1.1 Reduction of μ_{ldm} and $\mu_{d\bar{m}}$

Let us first consider the case of μ_{ldm}. According to (5.4), the relationship between μ_{ldm} and p_{ldm} at the l-th AP can be provided by the following lemma.

Lemma 5.3. *For the potentially optimal solution of* (5.12)*, the only feasible combinations of* μ_{ldm} *and* p_{ldm} *are* $(\mu^*_{ldm}, p^*_{ldm}) \in \{(0,0), (1, \tilde{p}_{ldm})\}$*, where* $\tilde{p}_{ldm} \neq 0$*.*

Proof. For the potentially optimal solution of (5.12), the possible combinations of μ_{ldm} and p_{ldm} are $(\mu^*_{ldm}, p^*_{ldm}) \in \{(0,0), (0, \tilde{p}_{ldm}), (1,0), (1, \tilde{p}_{ldm})\}$, where $\tilde{p}_{ldm} \neq 0$. Then, it can be easily found that $\Lambda_{\mathrm{SE}}\big((\lambda, \boldsymbol{\mu}, \boldsymbol{p}) \,|\, \mu_{ldm} = 0 \,\&\, p_{ldm} = 0\big) = \Lambda_{\mathrm{SE}}\big((\lambda, \boldsymbol{\mu}, \boldsymbol{p}) \,|\, \mu_{ldm} p_{ldm} = 0 \,\&\, \mu_{ldm} + p_{ldm} \neq 0\big)$, since μ_{ldm} and p_{ldm} are tightly coupled in (5.9). When μ_{ldm} or p_{ldm} is equal to 0, $\mathrm{SINR}_{d,m}$ and $\mathrm{SINR}_{d,\bar{m}}$ remain unchanged as $\mu_{ldm} p_{ldm} = 0$. In conclusion, due to the special relationship between μ_{ldm} and p_{ldm}, these two variables in the optimal solution of (5.12) can only be either $(0,0)$ or $(1, \tilde{p}_{ldm})$. This completes the proof. □

Based on Lemma 5.3, after the optimization problem (5.12) is solved, all μ_{ldm} can be subsequently obtained from the optimal PA, i.e., p^*_{ldm}, which can be given as

$$\mu_{ldm} = \begin{cases} 0, & \frac{p^*_{ldm}}{P_l} < \kappa \\ 1, & \frac{p^*_{ldm}}{P_l} \geq \kappa \end{cases}, \tag{5.13}$$

where κ is a very small number, implying that a small value of p^*_{ldm} can be deemed as zero. Analogously, we can apply Lemma 5.3 and (5.13) to derive $p_{d\bar{m}}$ and $\mu_{d\bar{m}}$.

5.3.1.2 Reduction of λ_{ld}

The binary variable λ_{ld} denotes the association status between AP l and MS d. Intuitively, once any of the DL subcarriers is activated at AP l for transmitting data to MS d,

i.e., $\exists m \in \mathcal{M}, \mu_{ldm} = 1$, the communication link between AP l and MS d is established. Therefore, λ_{ld} can be obtained as

$$\lambda_{ld} = \max\left\{\mu_{ldm} | m \in \mathcal{M}\right\}, \ \forall l \in \mathcal{L}, \ d \in \mathcal{D}. \tag{5.14}$$

5.3.2 Maximization of Spectral Efficiency Based on Quadratic Transform

After the reduction of the binary variables, as shown in Section 5.3.A, the SE optimization problem can be transformed to a relatively simple form as

$$\max_{\boldsymbol{p}} \ \Lambda_{\mathrm{SE}} \tag{5.15a}$$

$$\text{s.t.} \quad (5.12\text{b}), (5.12\text{c}), (5.12\text{d}), (5.12\text{e}), (5.12\text{f}), \tag{5.15b}$$

$$\mathrm{SINR}_{d,m} \geq e^{\frac{\chi_{\mathrm{DL}}}{M_d}} - 1, \ \forall d \in \mathcal{D}, m \in \mathcal{M}, \tag{5.15c}$$

$$\mathrm{SINR}_{d,\bar{m}} \geq e^{\frac{\chi_{\mathrm{UL}}}{\bar{M}_d}} - 1, \ \forall d \in \mathcal{D}, \bar{m} \in \bar{\mathcal{M}}, \tag{5.15d}$$

where M_d and $\bar{M}_d$ denote the numbers of DL and UL subcarriers assigned to MS d, respectively. During the optimization, the values of $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ are initialized to $\mathbf{1}$ and then iteratively updated with the results of PA. Furthermore, the ZF precoder/combiner are also re-computed during each iteration according to the updated results of AP-MS connection and SA. Note that, the original constraints (5.12g) and (5.12h) contain a sum of M and $\bar{M}$ nonconvex components for each MS d, respectively, resulting in extremely high complexity. Hence, to make the optimization tractable, these two constraints are simplified to (5.15c) and (5.15d).

It can be shown that the scaled-down objective function (5.15a) belongs to the general multiple-ratio concave-convex fractional programming (CCFP) problem [175]. However, the constraints of (5.15c) and (5.15d) are still nonconvex, which have to be approximated by the convex ones.

Specifically, based on (5.9), the constraint (5.15c) can be equivalently written as

$$(5.15\text{c}) \Longleftrightarrow \begin{cases} \mathrm{SINR}_{d,m} \triangleq \varpi_{d,m}^2 / \psi_{d,m} \geq e^{\frac{\chi_{\mathrm{DL}}}{M_d}} - 1, & (5.16\text{a}) \\ 0 < \varpi_{d,m} \leq \sum_{l \in \mathcal{L}} \sqrt{p_{ldm}} \omega_{ldm}, & (5.16\text{b}) \\ \psi_{d,m} \geq \zeta_d^{\mathrm{SI}} \Theta_{\mathrm{DL}} + \sigma^2, & (5.16\text{c}) \end{cases}$$

where $\varpi_{d,m}$ and $\psi_{d,m}$ are new variables, while (5.16b) and (5.16c) are linear constraints. For (5.16a), since the function $f_{\mathrm{ca}}(\varpi_{d,m}, \psi_{d,m}) \triangleq \varpi_{d,m}^2 / \psi_{d,m}$ with $(\varpi_{d,m}, \psi_{d,m}) \in \mathbb{R}_{++}^2$ is convex, it can be approximated using the successive convex approximation (SCA)

properties as [176]

$$f_{\text{ca}}(\varpi_{d,m},\psi_{d,m}) \geq \frac{2\varpi_{d,m}^{(t)}}{\psi_{d,m}^{(t)}}\varpi_{d,m} - \frac{(\varpi_{d,m}^{(t)})^2}{(\psi_{d,m}^{(t)})^2}\psi_{d,m} := f_{\text{ca}}^{(t)}(\varpi_{d,m},\psi_{d,m}), \tag{5.17}$$

where $(\varpi_{d,m}^{(t)},\psi_{d,m}^{(t)})$ is the feasible point obtained at the t-th iteration. Therefore, (5.16a) can be substituted by the new constraint given by

$$f_{\text{ca}}^{(t)}(\varpi_{d,m},\psi_{d,m}) \geq e^{\frac{\chi_{\text{DL}}}{M_d}} - 1,\ \forall d \in \mathcal{D}, m \in \mathcal{M}. \tag{5.18}$$

Following the same spirit, (5.15d) can be replaced by the following convex constraints:

$$(5.15\text{d}) \Longleftrightarrow \begin{cases} f_{\text{ca}}^{(t)}(\sqrt{\varpi_{d,\bar{m}}},\psi_{d,\bar{m}}) \geq e^{\frac{\chi_{\text{UL}}}{M_d}} - 1, & (5.19\text{a}) \\ 0 < \varpi_{d,\bar{m}} \leq p_{d\bar{m}}L^2, & (5.19\text{b}) \\ \psi_{d,m} \geq \sum_{l\in\mathcal{L}} v_{ld\bar{m}}\left(\xi_l^{\text{SI}}\Theta_{\text{UL}} + \sigma^2\right). & (5.19\text{c}) \end{cases}$$

To this point, all the constraints in (5.15) are convex, and we can now apply the QT to deal with (5.15). According to [175, Corollary 2], the sum-of-functions-of-ratio problem in (5.15a), i.e.,

$$\max_{\boldsymbol{p}} \frac{1}{M_{\text{sum}}}\sum_{d=1}^{D}\Big(\sum_{m=1}^{M} R_{d,m}\Big(\frac{A_{d,m}(\boldsymbol{p})}{B_{d,m}(\boldsymbol{p})}\Big) + \sum_{\bar{m}=1}^{\bar{M}} R_{d,\bar{m}}\Big(\frac{A_{d,\bar{m}}(\boldsymbol{p})}{B_{d,\bar{m}}(\boldsymbol{p})}\Big)\Big) \tag{5.20a}$$

$$\text{s.t. } (5.15\text{b}), (5.16\text{b}), (5.16\text{c}), (5.18), (5.19), \tag{5.20b}$$

can be equivalently described as

$$\max_{\boldsymbol{p}} \frac{1}{M_{\text{sum}}}\sum_{d=1}^{D}\Bigg(\sum_{m=1}^{M} R_{d,m}\big(2z_{dm}\sqrt{A_{d,m}(\boldsymbol{p})} - z_{dm}^2 B_{d,m}(\boldsymbol{p})\big) + \sum_{\bar{m}=1}^{\bar{M}} R_{d,\bar{m}}\big(2z_{d\bar{m}}\sqrt{A_{d,\bar{m}}(\boldsymbol{p})} - z_{d\bar{m}}^2 B_{d,\bar{m}}(\boldsymbol{p})\big)\Bigg) \tag{5.21a}$$

$$\text{s.t. } z_{dm} \in \mathbb{R},\ \forall d \in \mathcal{D}, m \in \mathcal{M}\ ,\ z_{d\bar{m}} \in \mathbb{R},\ \forall d \in \mathcal{D}, \bar{m} \in \bar{\mathcal{M}}, \tag{5.21b}$$

$$(5.15\text{b}), (5.16\text{b}), (5.16\text{c}), (5.18), (5.19), \tag{5.21c}$$

where $A(\boldsymbol{p})$ and $B(\boldsymbol{p})$ denote the numerator and denominator of (5.9), respectively. Since $R(x)$ for all d, m and $\bar{m}$ is non-decreasing and concave, and for a given $\boldsymbol{p}$, $\frac{A(\boldsymbol{p})}{B(\boldsymbol{p})}$ is in the concave-convex form, the optimal $\boldsymbol{z} = (\{z_{dm}\},\{z_{d\bar{m}}\})$ can be obtained as $\boldsymbol{z}^* = \frac{\sqrt{A(\boldsymbol{p})}}{B(\boldsymbol{p})}$. Then, for a given $\boldsymbol{z}$, the problem (5.21) is a concave maximization problem over $\boldsymbol{p}$. Therefore, the overall problem can essentially be solved by a block coordinate ascent algorithm, with $\boldsymbol{z}$ and $\boldsymbol{p}$ iteratively optimized, until the optimization converges to a local optimum. Note that our QT-SCA algorithm mainly relies on the QT and SCA methods, and the details for the proof of their convergence can be found in [175] and

[176]. Furthermore, the convergence properties and the complexity of the algorithm will be studied in Section 5.5. In summary, the overall optimization algorithm for SE maximization is stated as Algorithm 6.

Algorithm 6: QT-SCA Algorithm for SE maximization in MDD-CF scheme

Initialization:
Set $\boldsymbol{\lambda} = \mathbf{1}, \boldsymbol{\mu} = \mathbf{1}$;
Compute $\{\omega_{ldm}\}, \{v_{ld\bar{m}}\}, \forall l, d, m, \bar{m}$;
Set $t = 0$, $\boldsymbol{p}^{(0)}$ to a feasible value, $\varpi_{d,m}^{(0)} = 1, \varpi_{d,\bar{m}}^{(0)} = 1, \psi_{d,m}^{(0)} = 1, \psi_{d,\bar{m}}^{(0)} = 1, \forall d, m, \bar{m}$;
[QT Iteration]
  repeat
    Compute $\boldsymbol{z}^{(t)}$ using $\boldsymbol{z}^{(t)} = \frac{\sqrt{A(\boldsymbol{p}^{(t)})}}{B(\boldsymbol{p}^{(t)})}$, for a fixed $\boldsymbol{p}^{(t)}$;
    Update $\boldsymbol{p}^{(t+1)}$ by solving (5.21), for a fixed $\boldsymbol{z}^{(t)}$;
    Update $\varpi_{d,m}^{(t+1)}, \varpi_{d,\bar{m}}^{(t+1)}, \psi_{d,m}^{(t+1)}, \psi_{d,\bar{m}}^{(t+1)}$;
    Set $t = t + 1$;
  until Convergence;
end
Update $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ using (5.13) and (5.14), respectively;
Repeat Step 3 to Step 13 until $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ are stable, and obtain the optimal SE, i.e., $\Lambda_{\mathrm{SE}}^{*}$.
Output: $\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{p}, \Lambda_{\mathrm{SE}}^{*}$

5.4 Spectral Efficiency Optimization within Radio Frame

The optimization problem addressed in the previous section only considered the single CT interval and assumed perfect CSI. In this section, we extend our studies by considering a more complicated scenario, where the radio frame with imperfect CE is assumed. More specifically, as shown in Fig. 5.3, in our proposed frame structure, starting from the second CT interval of T_{c}^{2}, IBFD- and MDD-CF schemes can exploit an extra γ^{P} of time for DL transmission at the beginning of the interval during UL training, owing to their FD feature. Therefore, for the convenience of illustration, we assume that after T_{c}^{1}, each TPCT interval consists of two transmission phases, namely Phase I with γ^{P} of supplementary DL transmission and UL training, and Phase II with $(T_{\mathrm{c}} - \gamma^{\mathrm{P}})$ of DL/UL simultaneous transmissions. By contrast, in the counterpart TDD-CF scheme, since pilots and data are sequentially transmitted, the SE performance may deteriorate quickly in some cases. For example, when the relative velocity between MS and AP increases, UL training has to be implemented more frequently as the result of the shorter CT intervals, which unavoidably leads to the SE degradation.

However, in order to unleash the full advantages of FD, two paramount challenges need to be addressed in this case. Firstly, since the channel varies continuously from one CT interval to the next, the performance of the supplementary DL transmission in

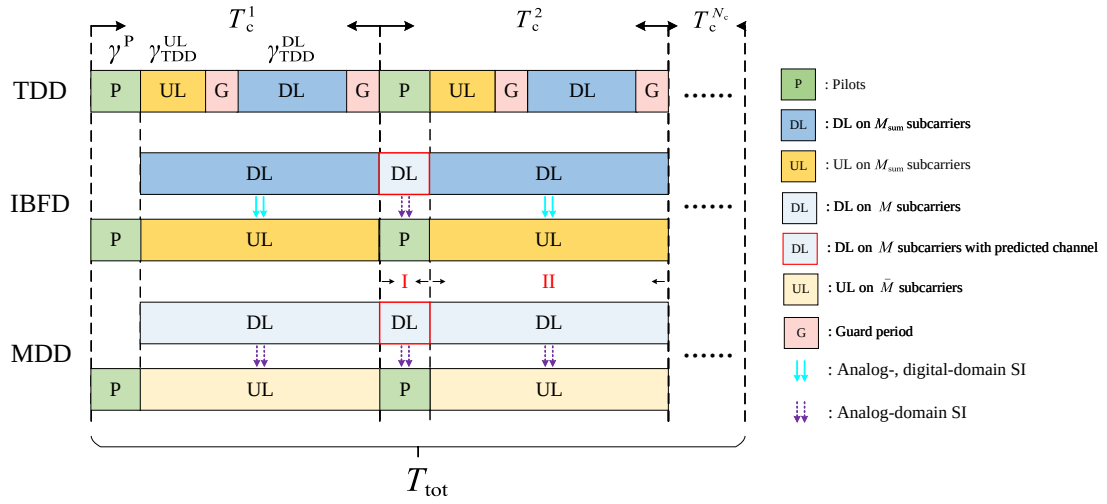


FIGURE 5.3: Structure of radio frame of TDD-, IBFD- and MDD- schemes, where N_c denotes the number of CT intervals within one radio frame.

Phase I hinges on the predicted channel. For this problem, in literature, there are various channel prediction methods, e.g., Wiener filter [74], Kalman filter [15] and deep learning [16], which can be applied to predict the required channels with high accuracy. Secondly, as depicted in Fig. 5.3, although the extra DL transmission during Phase I can increase SE due to the added transmission time, it may cause interference on the receiving of UL pilots, and hence affect the CE accuracy, which in turn leads to degraded performance in Phase II. Therefore, there exists a trade-off between the accuracy of channel acquisition and the SE provided by the supplementary DL transmission. To this end, we shall focus on this trade-off problem in the sequel under the assumption that the CSI used for the supplementary DL transmission is predicted using the Wiener filter.

5.4.1 Channel Estimation in MDD Cell-Free Scheme

Let us assume that all MSs synchronously transmit their frequency-domain pilot sequences (FDPS) over the $\bar{M}$ UL subcarriers, while all APs transmit DL data to MSs over the DL subcarriers. Let the FDPS transmitted by the d-th MS be expressed as $\boldsymbol{x}_d^{\mathrm{p}} = \left[x_d^{\mathrm{p}}[1], ..., x_d^{\mathrm{p}}[\bar{m}], ..., x_d^{\mathrm{p}}[\bar{M}]\right]^T$. Then, the received training signal at the n-th antenna of AP l can be written as

$$\boldsymbol{y}_l^n = \sum_{d=1}^{D} \sqrt{p_d^{\mathrm{p}}} \boldsymbol{X}_d^{\mathrm{p}} \boldsymbol{\Phi}_{\mathrm{UL}} \boldsymbol{F} \boldsymbol{\Psi} \boldsymbol{g}_{ld}^n + \boldsymbol{z}_l^{n,\mathrm{SI}} + \boldsymbol{z}_l^{n,\mathrm{IAI}} + \boldsymbol{n}_l^n, \tag{5.22}$$

where p_d^{p} denotes the power assigned for sending pilots on each UL subcarrier by MS d, $\boldsymbol{\Phi}_{\mathrm{UL}} = \boldsymbol{I}_{M_{\mathrm{sum}}}^{(\bar{\mathcal{M}},:)}$, $\boldsymbol{X}_d^{\mathrm{p}} = \mathrm{diag}\left\{\boldsymbol{x}_d^{\mathrm{p}}\right\}$, $\boldsymbol{n}_l^n \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_{\bar{M}})$. Note that the transmit power of

pilots is assumed to be larger than that of UL data, which is fixed during communication. According to Appendix E, we can obtain $\boldsymbol{z}_l^{n,\mathrm{SI}} \sim \mathcal{CN}\left(\boldsymbol{0}, \left(\mathrm{cov}\left\{\boldsymbol{z}_l^{\mathrm{SI}}\right\}\right)_{n,n} \boldsymbol{I}_{\bar{M}}\right)$ and $\boldsymbol{z}_l^{n,\mathrm{IAI}} \sim \mathcal{CN}\left(\boldsymbol{0}, \left(\mathrm{cov}\left\{\boldsymbol{z}_l^{\mathrm{IAI}}\right\}\right)_{n,n} \boldsymbol{I}_{\bar{M}}\right)$.

Based on the CE approach proposed in Section 2.4, the noisy observation of $\boldsymbol{g}_{n,d}[i]$ can be formed as

$$\tilde{\boldsymbol{y}}_{ld}^n = \boldsymbol{J}_d^H \boldsymbol{y}_l^n = \frac{\sqrt{p_d^{\mathrm{P}}}\bar{M}}{M_{\mathrm{sum}}} \boldsymbol{g}_{ld}^n + \boldsymbol{J}_d^H (\boldsymbol{z}_l^{n,\mathrm{SI}} + \boldsymbol{z}_l^{n,\mathrm{IAI}} + \boldsymbol{n}_l^n). \tag{5.23}$$

Correspondingly, the MMSE estimate to $\boldsymbol{g}_{ld}^n$ is given by

$$\hat{\boldsymbol{g}}_{ld}^n = \frac{\frac{\beta_{ld}\sqrt{p_d^{\mathrm{P}}}}{U}}{\frac{\beta_{ld} p_d^{\mathrm{P}} \bar{M}}{U M_{\mathrm{sum}}} + \left(\xi_l^{\mathrm{SI}} I_l + \sigma^2\right)} \tilde{\boldsymbol{y}}_{ld}^n, \tag{5.24}$$

where $I_l = \sum_{m\in\mathcal{M}} \sum_{d\in\mathcal{D}} \lambda_{ld}^{\mathrm{I}} \mu_{ldm}^{\mathrm{I}} p_{ldm}^{\mathrm{I}} + \frac{\xi_l^{\mathrm{IAI}}}{\xi_l^{\mathrm{SI}}} \sum_{l'\in\mathcal{L}\setminus\{l\}} \sum_{m\in\mathcal{M}} \sum_{d\in\mathcal{D}} \frac{\beta_{ll'}}{M_{\mathrm{sum}}} \lambda_{l'd}^{\mathrm{I}} \mu_{l'dm}^{\mathrm{I}} p_{l'dm}^{\mathrm{I}}$ denotes the summation of SI and IAI. Here the superscript 'I' means that the interference arises from the supplementary DL transmission of Phase I, as shown in Figure 5.3. Then, according to the properties of MMSE, $\boldsymbol{g}_{ld}^n$ can be orthogonally decomposed into $\boldsymbol{g}_{ld}^n = \hat{\boldsymbol{g}}_{ld}^n + \boldsymbol{e}_{ld}^n$, where $\boldsymbol{e}_{ld}^n$ is the CE error vector uncorrelated with $\hat{\boldsymbol{g}}_{ld}^n$, which has the covariance matrix

$$\Xi_e = \mathrm{cov}\left\{\boldsymbol{e}_{ld}^n\right\} = \left(\frac{\beta_{ld}}{U} - \frac{\frac{\beta_{ld}^2 p_d^{\mathrm{P}} \bar{M}}{U^2 M_{\mathrm{sum}}}}{\frac{\beta_{ld} p_d^{\mathrm{P}} \bar{M}}{U M_{\mathrm{sum}}} + \left(\xi_l^{\mathrm{SI}} I_l + \sigma^2\right)}\right) \boldsymbol{I}_U. \tag{5.25}$$

5.4.2 Spectral Efficiency Maximization within Two-Phase Coherence Time Interval

As the case of the SE optimization in T_{c}^1 has already been addressed in Section 5.4, here we only study the case of the new designed TPCT interval, i.e., T_{c}^n, $n = 2, ..., N_c$, as shown in Fig. 5.3.

To begin with, the average SE in Phase I can be expressed as

$$\Lambda_{\mathrm{SE}}^{\mathrm{I}} = \frac{\gamma^{\mathrm{P}}}{T_{\mathrm{c}}} \frac{1}{M_{\mathrm{sum}}} \sum_{d\in\mathcal{D}} \sum_{m\in\mathcal{M}} R(\check{\mathrm{SINR}}_{d,m}), \tag{5.26}$$

where $\check{\mathrm{SINR}}_{d,m}$ is different from the previous SINR, which depends on the predicted CSI by Wiener filter. Then, the maximization of the average SE in a TPCT interval

amounts to the following optimization problem:

$$\max_{\boldsymbol{p}^{\mathrm{I}},\boldsymbol{p}^{\mathrm{II}},\boldsymbol{\lambda}^{\mathrm{I}},\boldsymbol{\mu}^{\mathrm{I}},\boldsymbol{\lambda}^{\mathrm{II}},\boldsymbol{\mu}^{\mathrm{II}}} \Lambda_{\mathrm{SE}}^{\mathrm{TPCT}} = \Lambda_{\mathrm{SE}}^{\mathrm{I}} + \Lambda_{\mathrm{SE}}^{\mathrm{II}} \tag{5.27a}$$

$$\text{s.t.} \quad (5.12\text{b}), (5.12\text{c}), (5.12\text{d}), (5.12\text{e}), (5.12\text{f}), (5.15\text{b}), (5.15\text{c}), \tag{5.27b}$$

$$\lambda_{ld}^{\mathrm{I}} \in \{0,1\}, \ \forall l \in \mathcal{L}, d \in \mathcal{D}, \tag{5.27c}$$

$$\mu_{ldm}^{\mathrm{I}} \in \{0,1\}, \ \forall l \in \mathcal{L}, d \in \mathcal{D}, m \in \mathcal{M}, \tag{5.27d}$$

$$\sum_{m\in\mathcal{M}} \sum_{d\in\mathcal{D}} \lambda_{ld}^{\mathrm{I}} \mu_{ldm}^{\mathrm{I}} p_{ldm}^{\mathrm{I}} \leq P_l, \ \forall l \in \mathcal{L}, \tag{5.27e}$$

where $\boldsymbol{p}^{\mathrm{I/II}} = \left\{p_{ldm}^{\mathrm{I/II}}\right\}$, $\boldsymbol{\lambda}^{\mathrm{I/II}} = \left\{\lambda_{ld}^{\mathrm{I/II}}\right\}$, $\boldsymbol{\mu}^{\mathrm{I/II}} = \left\{\mu_{ldm}^{\mathrm{I/II}}\right\}$. The $\Lambda_{\mathrm{SE}}^{\mathrm{I}}$ is only related to $\boldsymbol{p}^{\mathrm{I}}$, while the $\Lambda_{\mathrm{SE}}^{\mathrm{II}}$ depends on both $\boldsymbol{p}^{\mathrm{I}}$ and $\boldsymbol{p}^{\mathrm{II}}$. (5.12b)-(5.12f) and (5.27c)-(5.27e) explain the power constraints and binary selections for Phase II and Phase I, respectively. (5.15b) and (5.15c) are the QoS constraints on Phase II, meaning that the objective in Phase I is to increase the overall SE as much as possible, while the QoS constraints are only imposed on Phase II. However, as shown in (5.24) and (5.25), $\boldsymbol{p}^{\mathrm{I}}$ in Phase I is included in the error vector of the time-domain estimated CSI. Hence, it implicitly affects the following DL/UL transmission in Phase II, which makes the above optimization problem intractable to obtain $\boldsymbol{p}^{\mathrm{I}}$ and $\boldsymbol{p}^{\mathrm{II}}$ at the same time.

In order to circumvent this problem, we introduce a two-step iterative algorithm to split the original optimization problem of (5.27) into two sub-problems, namely the optimizations of $\Lambda_{\mathrm{SE}}^{\mathrm{I}}$ and $\Lambda_{\mathrm{SE}}^{\mathrm{II}}$, which are expressed as follows:

$$\text{Sub 1:} \quad \max_{\boldsymbol{p}^{\mathrm{I}},\boldsymbol{\lambda}^{\mathrm{I}},\boldsymbol{\mu}^{\mathrm{I}}} \Lambda_{\mathrm{SE}}^{\mathrm{I}} \tag{5.28a}$$

$$\text{s.t.} \quad (5.27\text{c}), (5.27\text{d}), (5.27\text{e}), \tag{5.28b}$$

$$\xi_l^{\mathrm{SI}} I_l \leq \eta_l \sigma^2, \ \forall l, \tag{5.28c}$$

and

$$\text{Sub 2:} \quad \max_{\boldsymbol{p}^{\mathrm{II}},\boldsymbol{\lambda}^{\mathrm{II}},\boldsymbol{\mu}^{\mathrm{II}}} \Lambda_{\mathrm{SE}}^{\mathrm{II}} \tag{5.29a}$$

$$\text{s.t.} \quad (5.27\text{b}). \tag{5.29b}$$

Note that, since the above two sub-problems are strongly correlated, we introduce (5.28c) in the first sub-problem to build the connection between them, where η_l is a configurable factor for controlling the performance of CE, which is elaborated in the following remark.

Remark 5.4. The value of η_l directly influences the accuracy of CE, and further affects the DL/UL transmissions in Phase II. For instance, as η_l decreases, to meet the constraint of (5.28c), APs have to decrease their transmit power so as to reduce I_l at the

cost of compromised SE in Phase I. By contrast, a smaller I_l results in the smaller estimation error in (5.25), which in turn results in a higher SE in Phase II. Hence, iteratively updating η_l is essential for the proposed optimization problem. In our implementation, we resort to the bisection method to find the optimal η_l in the range of $[0, \eta^{\max}]$, where 0 implies that the supplementary DL transmission in Phase I is not activated. In this case, there is no extra burden on the CE in the current CT interval. By contrast, $\eta^{\max} = \frac{\xi_l^{\mathrm{SI}} I_l^{\max}}{\sigma^2}$ denotes the maximum effective value, where $I_l^{\max}$ means that all APs transmit at their highest possible power for the supplementary DL transmission in Phase I.

Therefore, in the proposed algorithm, we first set $\eta_l = 0$, and obtain the total initial SE of $\Lambda_{\mathrm{SE}}^{\mathrm{TPCT}(0)} = \Lambda_{\mathrm{SE}}^{\mathrm{I}(0)} + \Lambda_{\mathrm{SE}}^{\mathrm{II}(0)}$, where $\Lambda_{\mathrm{SE}}^{\mathrm{I}(0)} = 0$. Then, η_l is iteratively updated based on the bisection method until convergence is achieved. The overall algorithm is summarized in Algorithm 7.

Algorithm 7: SE Maximization within TPCT Interval

Initialization:
Set $t = 0$, $t' = 0$ $\eta_l \in [\eta^a, \eta^b]$, $\eta^a = \eta_l^{(0)} = 0, \forall l, \eta^b = \eta^{\max}$;
Solve Sub 1 and Sub 2 problems in sequence using the QT-SCA method, and obtain $\Lambda_{\mathrm{SE}}^{\mathrm{I}(t)}$ and $\Lambda_{\mathrm{SE}}^{\mathrm{II}(t)}$;
Compute $\Lambda_{\mathrm{SE}}^{\mathrm{TPCT}(t)} = \Lambda_{\mathrm{SE}}^{\mathrm{I}(t)} + \Lambda_{\mathrm{SE}}^{\mathrm{II}(t)}$;
[Bisection Iteration of η_l]
  repeat
    Set $t = t + 1$, and update $\eta_l^{(t)} = \frac{\eta^a + \eta^b}{2}$;
    Implement Step 3 and Step 4;
    if $\Lambda_{\mathrm{SE}}^{\mathrm{TPCT}(t)} < \Lambda_{\mathrm{SE}}^{\mathrm{TPCT}(t')}$ **then**
      $\eta^b = \eta_l^{(t)}$;
    else
      $\eta^a = \eta_l^{(t)}$, $t' = t$;
    end if
  until Convergence;
end
Output: $\boldsymbol{p}^{\mathrm{I}}, \boldsymbol{\lambda}^{\mathrm{I}}, \boldsymbol{\mu}^{\mathrm{I}}, \boldsymbol{p}^{\mathrm{II}}, \boldsymbol{\lambda}^{\mathrm{II}}, \boldsymbol{\mu}^{\mathrm{II}}, \Lambda_{\mathrm{SE}}^{\mathrm{TPCT}*}$

5.5 Simulation Results and Discussion

In this section, we provide the numerical results for comparison of MDD-, IBFD- and TDD-CF in distributed CF-mMIMO systems in terms of SE. All the simulations are implemented on MATLAB using the CVX tool [177].

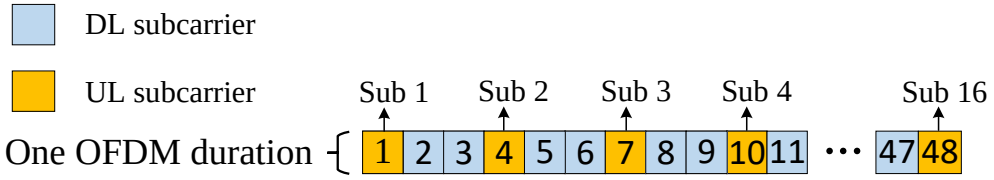


FIGURE 5.4: The subcarriers arrangement in MDD-CF scheme.

5.5.1 Parameters and Setup

The following results are obtained based on either one CT interval or one radio frame consisting of one CT interval and $(N_c - 1)$ TPCT intervals. In our studies, we assume that the subcarrier spacing is 15 KHz (with the central carrier frequency of 5 GHz), and each OFDM symbol with cyclic prefix spans $t_s = 71.35\mu$s [164]. We assume that all MSs move at a relative speed of $v = 5$ km/h, while APs are stationary. Hence, the coherence time is $t_{ct} \approx 21.6$ ms. In this case, one CT interval can accommodate about $T_c = t_{ct}/t_s \approx 300$ OFDM symbols. Regarding the frame structure [164,170], we assume that for all systems, the pilot transmission requires 15 OFDM symbol durations within one CT (i.e., $\gamma^P = 15$, as shown in Fig. 5.2 and Fig. 5.3), while for TDD systems, the GP lasts for 15 OFDM symbol durations in a TDD radio frame (i.e., $\gamma^G = 15$).

Assuming the delay spread of 40 ns [178], the coherence bandwidth is $B_c \approx 4.2$ MHz. To make subcarrier signals experience flat fading, we assume that the total number of subcarriers is $M_{sum} = 48$ for all systems, while in MDD systems, the numbers of DL and UL subcarriers are $M = 32$ and $\bar{M} = 16$, respectively. On the basis of Proposition 1, to guarantee the orthogonal receiving among different MSs during channel estimation, the 16 UL subcarriers using for UL/pilot transmissions are equally selected from the 48 subcarriers, while the remaining 32 subcarriers are used for DL transmissions, as shown in Fig. 5.4. This subcarriers arrangement keeps constant during one CT interval and radio frame. For the sake of fair comparison, in TDD systems, the ratio of DL/UL transmission times is set to the same ratio of DL/UL subcarrier numbers in MDD-CF, i.e., $\gamma^{DL}_{TDD} = 180$ and $\gamma^{UL}_{TDD} = 90$, if the CT interval has 300 symbols, while the other 30 symbols are pilot symbols and GP intervals. Then, the total SE of TDD-CF within one CT or one radio frame can be expressed as

$$\Lambda^{TDD}_{SE} = \frac{1}{M_{sum}} \sum_{d \in \mathcal{D}} \Big(\sum_{m=1}^{M_{sum}} \frac{\gamma^{DL}_{TDD}}{T_c} R(\text{SINR}_{d,m}) + \sum_{\bar{m}=1}^{M_{sum}} \frac{\gamma^{UL}_{TDD}}{T_c} R(\text{SINR}_{d,\bar{m}}) \Big). \tag{5.30}$$

Owing to the DL transmission in Phase I, as shown in Figure 5.3, the total SE achieved by the MDD-CF in one radio frame is different from that in one CT interval, which can

be expressed as

$$\Lambda_{\mathrm{SE}}^{\text{MDD-RF}} = \frac{1}{N_c}\Lambda_{\mathrm{SE}} + \frac{N_c}{N_c - 1}\Lambda_{\mathrm{SE}}^{\text{TPCT}}, \tag{5.31}$$

where Λ_{SE} and $\Lambda_{\mathrm{SE}}^{\text{TPCT}}$ are derived based on (5.11) and (5.27), respectively. The SE computation of IBFD-CF within one radio frame is similar to that of MDD-CF.

We assume that APs in both IBFD-CF and MDD-CF schemes with distributed operation are capable of providing 30 dB IAI suppression in the propagation/analog domain with the existing approaches as mentioned previously. Provided that the 12-bit ADC is employed, MDD-CF can suppress IAI up to 72 dB (i.e., $\xi_l^{\text{MDD-IAI}} = -72$ dB, $\forall l$), of which 42 dB is attributed to the digital cancellation by FFT. On the contrary, as IBFD-CF can hardly cope with the IAI in digital domain, we assume that it can provide no more than 10 dB of digital-domain IAI mitigation[3], yielding $\xi_l^{\text{IBFD-IAI}} = -40$ dB, $\forall l$. In the context of the IMI suppression, since the propagation/analog-domain IMI methods are relatively complicated to implement at the single-antenna MS, it is assumed that $\xi_d^{\text{IBFD-IMI}} = 0$ dB, $\forall d$, while $\xi_d^{\text{MDD-IMI}} = -42$ dB, $\forall d$, owing to the FFT-assisted suppression.

We assume that all APs and MSs are uniformly and randomly distributed within a square of size $(S_{\mathrm{D}} \times S_{\mathrm{D}})$ m^2. The large-scale fading coefficients $\beta \in \{\beta_{ld}, \beta_{ll'}, \beta_{dd'}\}$ is given by $\beta[\text{dB}] = \text{PL} + \sigma_{\text{sh}} z$, where the shadowing is characterized by $\sigma_{\text{sh}} z$ with a standard deviation of $\sigma_{\text{sh}} = 4$ dB and $z \sim \mathcal{N}(0,1)$. The PL exponent is assumed to be -3.8 [75]. Unless otherwise noted, the other parameters are listed in Table 5.1. The empirical cumulative distribution function (CDF) of per-MS SE and the average per-MS SE in the following simulations are obtained based on 1000 random realizations of AP and MS positions.

5.5.2 Case of One Coherence Time Interval

Given a randomly generated network layout, Fig. 5.5 shows the AP-selection and PA results attained by Algorithm 6 over one CT interval. The numbers within the blue and pink squares denote the AP and MS's transmit power, respectively. The black arrowed lines denote the DL links with the transmit power larger than 1 W. It can be observed from the figure that, subject to the limited SIC capability, the transmit power of a MS for UL transmission is much less than the budget power. Moreover, some APs, such as AP 1, AP 2 and AP 9, located far away from MSs are controlled by the algorithm to reduce their transmit power to avoid IAI on the other APs. Furthermore, in Fig. 5.6, as an example, we show the detailed PA and SA results of AP 11 to MS 3. Explicitly, based on (5.13), AP 11 allocates different power to the 32 DL subcarriers, where some

[3]In fact, authors in [86] presented several digital-domain methods to suppress IAI, which end up with only providing about 10 dB of IAI mitigation. The study implies that the IAI suppression in IBFD-CF scheme with centralized operation is very challenging, not to mention the CF systems operated in a distributed way.

TABLE 5.1: Simulation parameters

Default parameters	Value
Number of APs, MSs (L, D)	(12, 4)
Number of antennas per AP (N)	6
AP's and MS's Power budget $(P_l, P_d, \forall l, d)$	$(10, 1)$ W
UL pilot power $(p_d^{\mathrm{p}}, \forall d)$	0.6 W
MSs' QoS requirements $(\chi_{\mathrm{DL}}, \chi_{\mathrm{UL}})$	(0.5, 0.1)
Noise power (σ^2)	-94 dBm
Delay taps (U)	4
Number of CT intervals within one frame (N_c)	10
Residual SI level at AP $(\xi_l^{\text{MDD-SI}}, \xi_l^{\text{IBFD-SI}}, \forall l)$	-130 dB
Residual SI level at MS $(\xi_d^{\text{MDD-SI}}, \xi_d^{\text{IBFD-SI}}, \forall d)$	-120 dB
Cell length (S_{D})	400 m

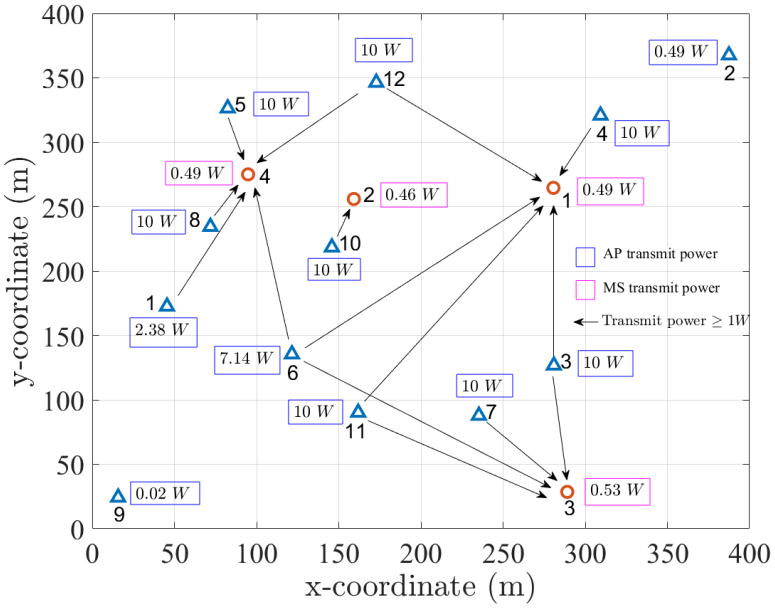


FIGURE 5.5: AP-selection and PA results obtained by Algorithm 1 in MDD-CF scheme.

subcarriers, namely subcarriers 11, 12, 22, 23 and 24, with very small power are not assigned to MS 3.

In the following simulations, we focus on the performance comparison of MDD-, TDD- and IBFD-CF schemes over one CT interval, as shown in Fig. 5.2. First, the performance of three schemes is presented in Fig. 5.7. From the results, when $v = 5$ km/h (corresponding to $T_c = 300$ OFDM symbols), TDD-CF slightly outperforms MDD-CF in terms of the 95%-likely and mean per-MS SE as the result of the IAI and IMI being not perfectly canceled in MDD-CF. Although IBFD-CF provides substantially higher SE for the strongest MSs, its 95%-likely performance is only about 0.5 nats/s/Hz, which

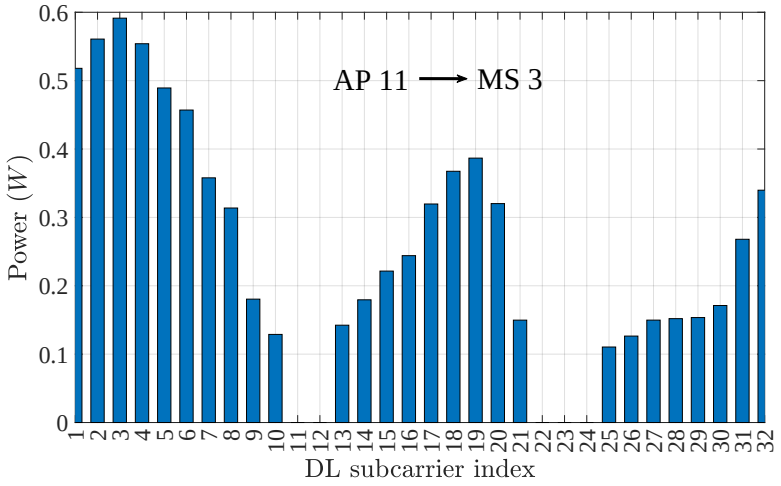


FIGURE 5.6: PA and SA results of AP 11 to MS 3 achieved by Algorithm 1.

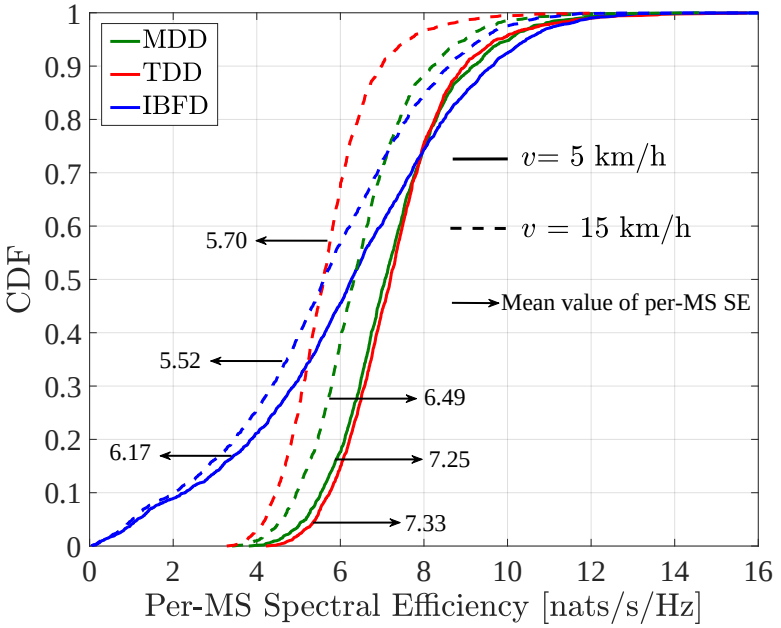
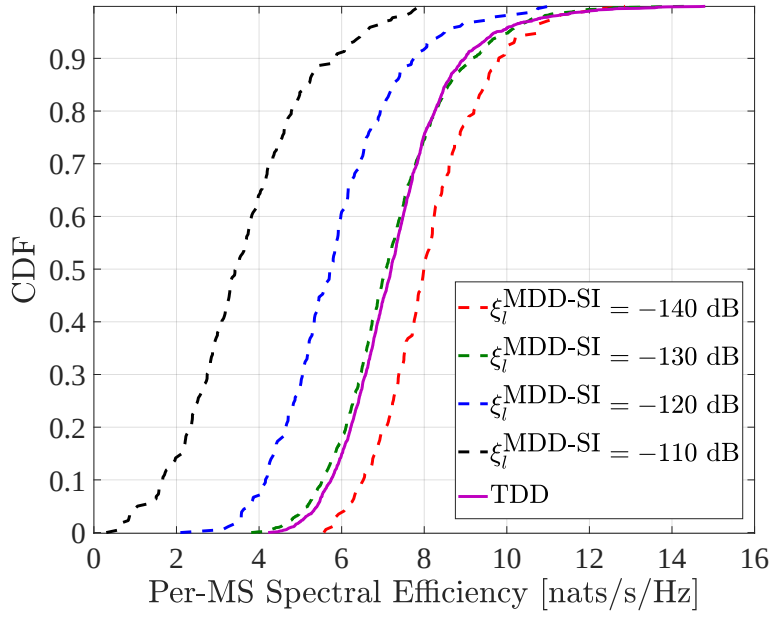
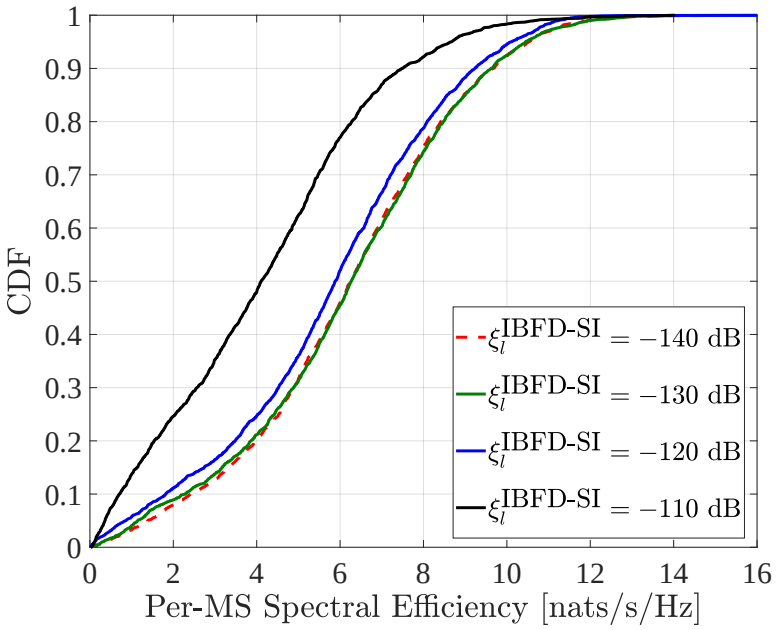


FIGURE 5.7: Cumulative distribution of the per-MS in different CF schemes.

is 10 times lower than MDD- and TDD-CF (5 nats/s/Hz). The rationale is that there are usually only a small number of MSs that can benefit from IBFD mode. These MSs should be located far away from the neighboring MSs and their serving APs are also not close to each other. Otherwise, to meet the QoS requirements, see (5.12g) and (5.12h), the DL/UL transmit power has to be decreased so as to reduce the negative effect of IAI/IMI. Furthermore, the performance of these schemes is very different when the relative speed between MSs and APs is increased to 15 km/h, which corresponds to 100 OFDM symbols transmitted in one CT interval. Fig. 5.7 clearly shows that the mean per-MS SE of TDD-CF deteriorates to 5.70 nats/s/Hz, as the relative speed increases. This is because the length of GP is constant, but its proportion within one CT interval becomes larger when the relative speed goes up, which therefore leads to the reduced



(A)



(B)

FIGURE 5.8: Cumulative distribution versus Per-MS SE in MDD- and IBFD-CF schemes with different SIC.

SE.

The influence of the SIC capability at APs and MSs in the MDD- and IBFD-CF schemes is demonstrated in Fig. 5.8. Note that the value of $\xi_d^{\text{MDD-SI}}$ is not shown in the figures, which is always set to be 10 dB lower than that of $\xi_l^{\text{MDD-SI}}$. From Fig. 5.8(a), it can be seen that the higher SIC capability the APs and MSs have, the better performance the MDD-CF can attain. Specifically, when $\xi_l^{\text{MDD-SI}} = -140$ dB, MDD-CF outperforms TDD-CF with a higher performance upper bound. This is because MDD mode is free of

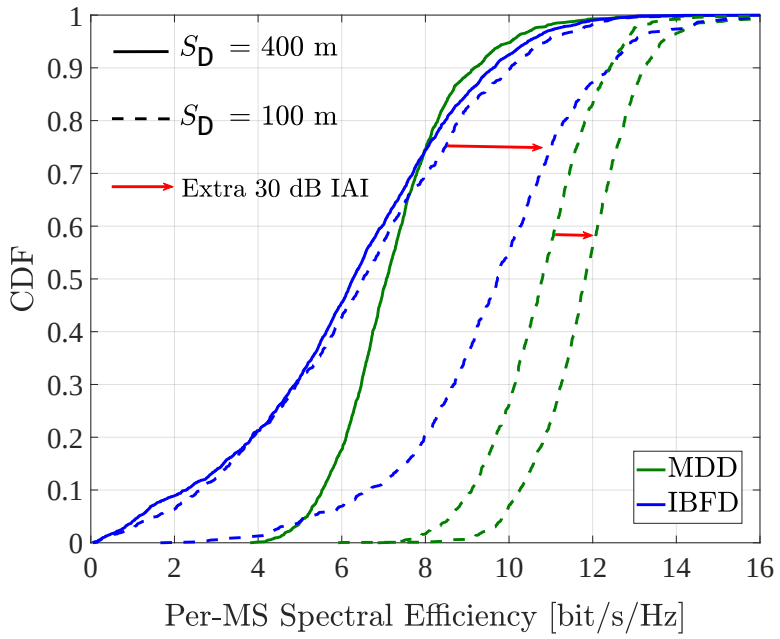


FIGURE 5.9: Cumulative distribution versus Per-MS SE in MDD- and IBFD-CF schemes with different cell size.

GPs, and hence have more time resource for data transmissions. On the contrary, from Fig. 5.8(b), the interesting result is that when the $\xi_l^{\text{IBFD-SI}}$ reaches at 120 dB, the further increased SIC capability can hardly improve the performance of IBFD-CF. This implies that the IBFD-CF is mainly IAI/IMI-limited.

Fig. 5.9 evaluates the influence of cell size on the performance of MDD- and IBFD-CF schemes. It can be clearly seen that as the cell size decreases, meaning that the distribution of MS and AP become denser, the 95%-likely per-MS SE of the MDD-CF is increased by 4 nats/s/Hz, while the IBFD-CF fails to obtain the significant gain from the denser network deployment. The reason is that, in IBFD-CF, the denser distribution of APs and MSs also means the shorter interference links of AP-AP and MS-MS. In this case, the large-scale fading is unable to provide enough IAI and IMI mitigation, consequently, both the APs and MSs have to decrease their transmit power so as to control the level of interferences. By contrast, the FFT-assisted IAI and IMI cancellation in MDD-CF can efficiently mitigate the interference in digital domain. Hence, MDD-CF can benefit significantly from the short-distance communications. Moreover, to further illustrate the effect of IAI, we assume that the APs in both MDD-CF and IBFD-CF schemes have an extra 30 dB of IAI suppression capability, when $S_D = 100$ m. Then, it can be seen that the IBFD-CF has a significant improvement in terms of both the 95%-likely and the median per-MS SE. However, the 95%-likely per-MS SE of the IBFD-CF with extra IAI suppression is still 4 nats/s/Hz lower than that of the MDD-CF without extra IAI suppression, due to the less IMI mitigation. We can conclude from Fig. 5.8 and Fig. 5.9 that the IBFD is not a desirable mode for FD-style operation in distributed CF-mMIMO systems, while the MDD mode is more promising.

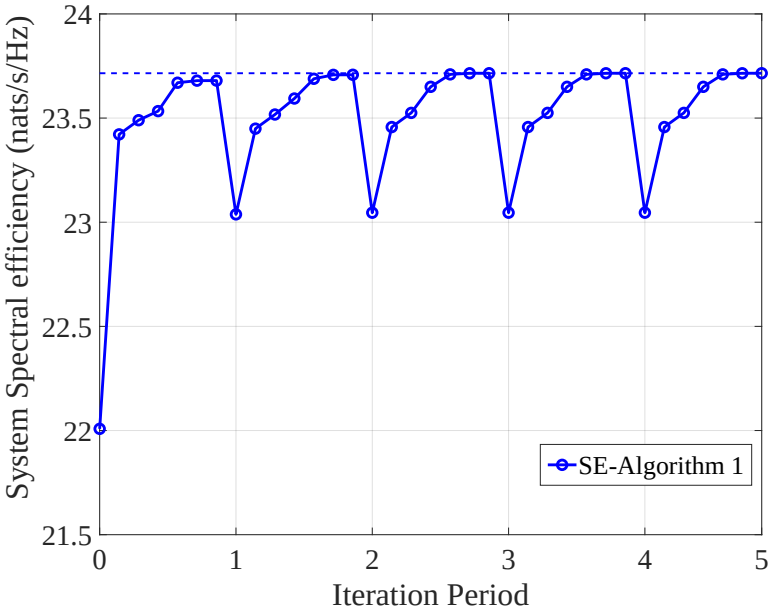


FIGURE 5.10: SE convergence behavior of Algorithm 6 in MDD-CF scheme.

The convergence behavior of the Algorithm 6 in MDD-CF is shown in Fig. 5.10, where each iteration period denotes one cycle, which includes step 3 - step 13 in the Algorithm. The lowest starting point at the beginning of each iteration period is due to the reinitialization at step 4. From the figure, it can be easily seen that Algorithm 6 can reach convergence only after two iteration periods. Concerning the complexity, Algorithm 6 is mainly attributed to solving the optimization problem (5.21), regardless of the computation of the ZF matrices and the initialization of the involved variables. Hence, according to [179], the approximated complexity of Algorithm 6 is $\mathcal{O}\big((LDM + 2DM + 3D\bar{M})^2(L + D + 3DM_{\text{sum}})^{2.5} + (L + D + 3DM_{\text{sum}})^{3.5}\big)$ per iteration. From these results, one can see that the numbers of MSs and subcarriers are the two dominant factors limiting the scalability of MDD-CF, where the MSs are densely distributed with a large number of available subcarriers. To this end, the more scalable approaches for the MDD-CF scheme deserve further investigation in the future.

5.5.3 Case of One Radio Frame

In this subsection, we demonstrate the advantages of MDD-CF, which can fully exploit the time resource within one radio frame with the aid of Algorithm 7 to improve the SE performance, as detailed below.

Fig. 5.11 evaluates the per-MS SE of the three different schemes over a range of relative speeds. Explicitly, the SE performance of TDD-CF deteriorates quickly, as the relative speed increases. The explanation is as follows. In TDD mode, DL/UL transmissions can only be implemented in a sequential manner. When the relative speed between MSs and APs increases, the reduced CT interval leads to the less time for data transmission.

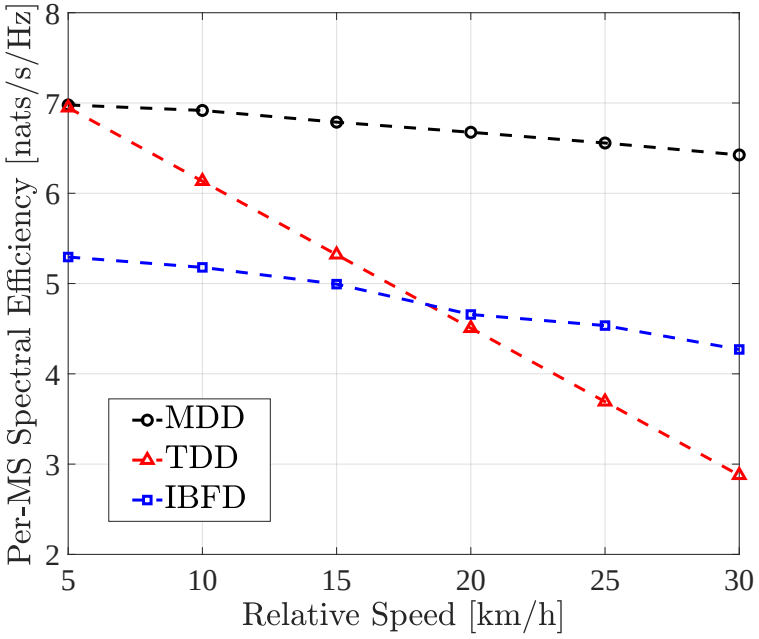


FIGURE 5.11: Per-MS SE of one radio frame versus the relative speed.

By contrast, with the aid of our proposed Algorithm 7, both MDD-CF and IBFD-CF are robust to the high-mobility scenarios, as expected. In particular, as MDD-CF can benefit from the lower IAI/IMI as well as the FD operation during both Phase I and II, it doubling the per-MS SE compared to TDD-CF, at the speed of 30 km/h. As shown in Fig. 5.11, the IBFD-CF is also outperformed by the MDD-CF, due to the larger residual IAI/IMI, which not only affect the UL and DL communications in Phase II, but also the CE in Phase I. However, it can eventually surpass the TDD-CF at about 18 km/h and achieve 1 nats/s/Hz higher per-MS SE than the TDD-CF at 30 km/h, due to the fact that Phase I becomes increasingly paramount in TPCT interval.

The performance of Phase I and that of the combination of Phase I and Phase II are demonstrated in Fig. 5.12 and Fig. 5.13, respectively. Due to the limited capability of IAI/IMI management in IBFD-CF, in order to guarantee the accuracy of CE in Phase I, the APs have to cut down the DL's transmit power so as to limit the interference imposed on the pilot receiving. Consequently, as seen in Fig. 5.12, the IBFD-CF attains a much poorer performance in Phase I than the MDD-CF. Furthermore, Fig. 5.13 shows that when the relative speed is increased to 30 km/h, the 95%-likely per-MS SE of the MDD-CF is nearly two times higher than that of the TDD-CF. Additionally, as seen in Fig. 5.13, although the IBFD-CF lags behind the TDD-CF in terms of the 95%-likely per-MS SE, its median performance is higher than that of the TDD-CF at 30 km/h.

Finally, the convergence of Algorithm 7 is shown in Fig. 5.14. Note that η marked in the figure denotes the average value of η_l, i.e., $\eta = \sum \eta_l / L$. It can be seen from the results that when $t = 0$ and $\eta = 0$, only Phase II is activated and there is no data transmission in Phase I. Then, as η iterates in line 7 of Algorithm 7, the general observation is that if

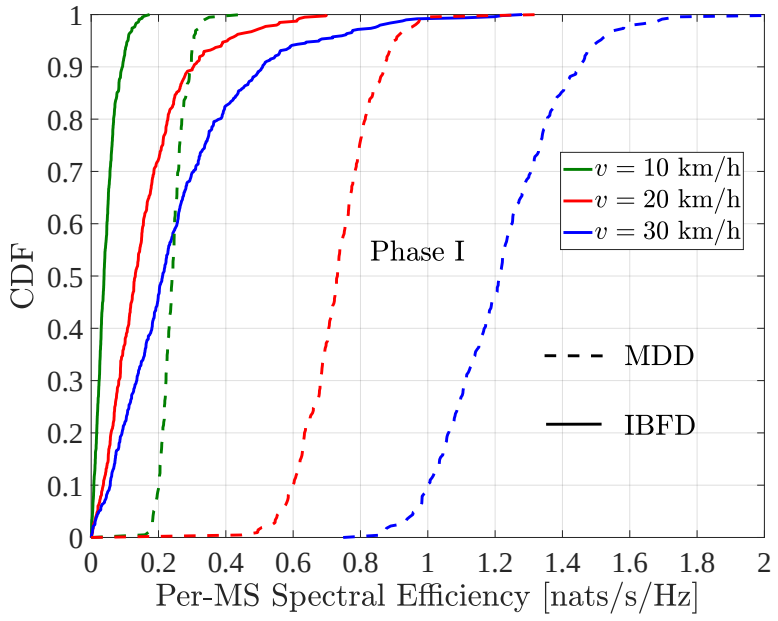


FIGURE 5.12: Cumulative distribution versus per-MS SE in Phase I.

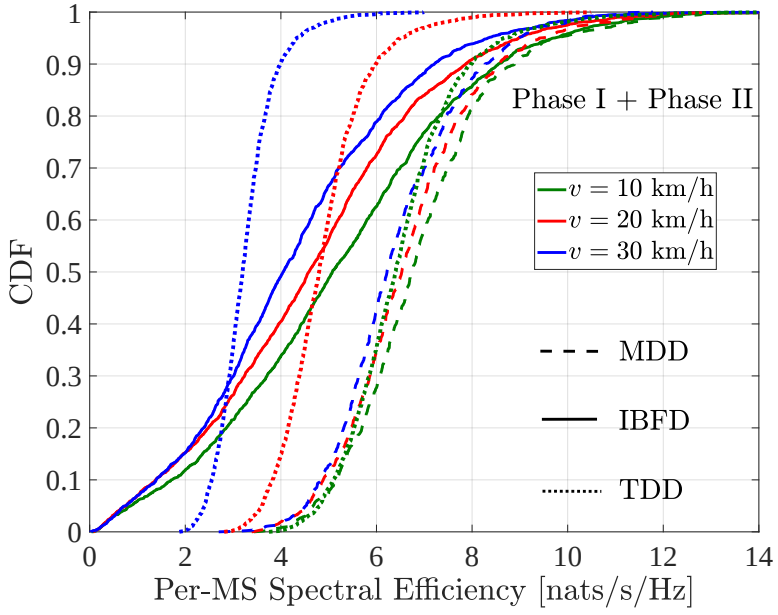


FIGURE 5.13: Cumulative distribution versus per-MS SE in the case of combined Phase I and Phase II.

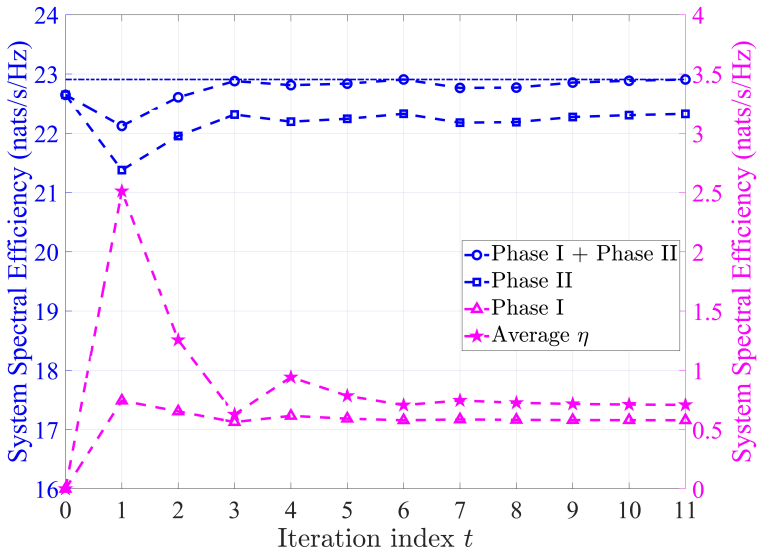


FIGURE 5.14: SE convergence behavior of Algorithm 3 in MDD-CF scheme.

the performance of one phase improves, the performance of the other phase degrades. The reason is that according to (5.28c), if more power is allocated for the DL transmission in Phase I, it results in not only the SE increase but also a larger CE error, which leads to the SE degradation in Phase II, and vice versa. The results demonstrate that our proposed algorithm is capable of attaining the 99% system performance after 6 iterations, and fully converging within about 10 iterations. The complexity of Algorithm 7 is mainly attributed by the optimization of (5.27), which has two phases. However, as indicated by (5.28) and (5.29), the complexity of Phase I is much lower than that of Phase II, as there are less variables to consider and also less constraints involved in the optimization of Phase I. Consequently, the complexity of Algorithm 7 is dominated by the operations in Phase II, which is nearly the same as Algorithm 6.

5.6 Chapter Summary and Conclusions

The MDD-CF scheme has been proposed and comprehensively studied in terms of SE performance, when AP-selection, PA and SA are taken into consideration. Firstly, the SE optimization problem has been studied in one single CT interval. We have exploited the interdependence of the involved variables and transformed the mixed-integer optimization to a continuous-integer convex-concave problem. In order to efficiently solve the problem, we have proposed a QT-SCA iterative algorithm, which is capable of quickly achieving a local optimum. The simulation results show that in distributed CF-mMIMO systems, MDD-CF can significantly outperform IBFD-CF due to the more efficient IAI and IMI mitigation. Secondly, the case of one radio frame with imperfect CE has been studied. To fully take the advantages of FD operation in radio frame, we

have designed a TPCT interval comprised of two tightly coupled phases. Then, a two-step iterative algorithm based on bisection method has been proposed to maximize the SE in TPCT interval. According to the simulation results, with the aid of TPCT interval and Algorithm 7, MDD-CF is more robust to high-mobility scenarios, while the performance of TDD-CF degrades quickly with the increase of the relative speed between APs and MSs.

Chapter 6

MDD Assisted Cell-Free Massive MIMO with Heterogeneous Graph Neural Network for Power Allocation

This chapter describes a graph learning based method, in comparison with the classic method proposed in Chapter 5, for the implementation of power allocation (PA) in multicarrier-division duplex assisted cell-free (MDD-CF) massive multiple-input multiple-output (mMIMO) systems. Its achievable performance and scalability will be demonstrated.

6.1 Introduction

In Chapter 5, we proposed a CF scheme driven by MDD mode, and the quadratic transform with successive convex approximation (QT-SCA) algorithm was presented to deal with the PA problem in MDD-CF mMIMO. However, according to the analysis of computational complexity in Fig. 5.9, despite the Algorithm 6 can quickly converge to local optimum, its complexity will exponentially grow as the increase of network size, which is impractical for timely application over each coherence time (CT) interval. In this regard, the explosive research of machine learning (ML) sheds some light on a new way to resolving the PA problem in CF systems. As we mentioned in Section 1.3.3, there have been several papers studying PA based on the state-of-the-art neural networks, e.g., deep neural network (DNN), reinforcement deep Q-learning network, etc. Nevertheless, most of the existing learning-based PA approaches was shown to prevail over the traditional methods, they fail to exploit the structure of the wireless communication

networks, and hence can not be generalized to unseen scenarios such as the networks with various cell sizes or access point (AP) / mobile station (MS) densities. An ongoing idea of enabling a scalable and generalizable model is to integrate the structure of the objective task with the neural network architecture.

To this end, the convolutional neural network (CNN) assisted models in computer vision fully exploiting the neighbor pixels of an image were judged a great success [180]. This idea has also been brought to study optimization problems in wireless communication. In [181], authors exploits the users' geographical location in a 2D Euclidean space, and proposed spatial convolution, which can deal with the device-to-device (D2D) networks with thousands of users but cannot handle fading channels. Then, authors in [182] regarded fading channel matrix as the adjacency matrix of graph and proposed a random edge graph neural network (GNN), which inhibits a good generalization with varying numbers of users. Since then, GNN reaping the advantages of scalability, generalization and parallel execution has attracted significant interests from researchers working in wireless communications. In [183], a wireless channel graph convolution network (WCGCN) was proposed for dealing with PA and bearmforming problem in D2D networks. Authors showed that the WCGCN trained on small size system can be generalized to a large system with higher density of MS and larger cell sizes. Additionally, [184] studied the PA in ad hoc networks based on GNN. However, above papers only exploited the homogeneous GNN in ad hoc or D2D networks with only one type of nodes, which cannot meet the demand of more complicated scenario like full duplex (FD)-style CF networks, where AP and MS are two different node types employed with different node features, and each of node is connected with two different types of node via communication path and interference path.

Against this background, we propose a CF-heterogeneous GNN (HGNN) network, namely CF-HGNN, to solve the PA problem in MDD-CF scheme, which is to the best of our knowledge the first HGNN-based network to handle the PA in multicarrier FD-style mMIMO systems. The major contributions of this chapter are summarized as follows:

Firstly, for the sake of generalization and less complexity, we propose CF-HGNN to deal with the PA problem in MDD-CF mMIMO. The CF-HGNN can handle varying number of APs, MSs and subcarriers through an adaptive node embedding layer. Furthermore, the message passing algorithm combined with the meta-path attention mechanism is introduced in CF-HGNN, which is able to learn the importance of interfering path and data transmission path during information aggregation at each AP/MS node.

Secondly, we conduct comprehensive experiments to evaluate the effectiveness of CF-HGNN for PA problem in MDD-CF mMIMO. The training of CF-HGNN is implemented in an unsupervised way with unlabeled data. Simulation results show that

CF-HGNN is capable of learning nearly the same PA strategy with QT-SCA algorithm and significantly outperforms the traditional unfair greedy method, without any prior information. Additionally, CF-HGNN is adaptive to varying number of nodes and subcarriers, and also generalization ability to different sizes of CF network. Finally, it is demonstrated that CF-HGNN has an incredible advantage over QT-SCA in terms of computational complexity and operation time.

The rest of the chapter is organized as follows. In Section II, we reformulate the PA problem in MDD-CF mMIMO, which is similar to (5.15). In Section III, the CF-HGNN is proposed to solve the formulated problem in MDD-CF mMIMO systems from a learning perspective. Simulation results are presented in Section IV, and finally, conclusions are drawn in Section V.

6.2 Problem Formulation

To begin with, the channel models, DL/UL transmissions as well as beamforming strategy are kept nearly the same with Section 5.2, while the subcarrier allocation (SA) and AP-selection indicators are omitted. Besides, we only consider the one CT interval scenario, as shown in Fig. 5.2, in this chapter. For the sake of convenience, here we reformulate the PA problem, which is given by

$$\max_{p_{ldm},p_{d\bar{m}}} \Lambda_{\mathrm{SE}} = \frac{1}{M_{\mathrm{sum}}} \sum_{d\in\mathcal{D}} \Big(\sum_{m\in\mathcal{M}} \ln(1+\mathrm{SINR}_{d,m}) + \sum_{\bar{m}\in\bar{\mathcal{M}}} \ln(1+\mathrm{SINR}_{d,\bar{m}}) \Big) \tag{6.1a}$$

$$\text{s.t.} \quad \sum_{m\in\mathcal{M}} \sum_{d\in\mathcal{D}} p_{ldm} \le P_l,\ \forall l \in \mathcal{L}, \tag{6.1b}$$

$$\sum_{\bar{m}\in\bar{\mathcal{M}}} p_{d\bar{m}} \le P_d,\ \forall d \in \mathcal{D}, \tag{6.1c}$$

$$\sum_{m\in\mathcal{M}} \ln(1+\mathrm{SINR}_{d,m}) \ge \chi_{\mathrm{DL}},\ \forall d \in \mathcal{D}, \tag{6.1d}$$

$$\sum_{\bar{m}\in\bar{\mathcal{M}}} \ln(1+\mathrm{SINR}_{d,\bar{m}}) \ge \chi_{\mathrm{UL}},\ \forall d \in \mathcal{D}, \tag{6.1e}$$

where all the involved variables are defined as Section 5.2. It can be observed that (6.1) is a NP-hard nonconvex problem with nonconvex constraints as shown in (6.1d) and (6.1e). Generally, to solve this kind of complicated problem, we have to transform the main function and constraints into the approximated convex ones and solve them in an iterative way. With this regard, we proposed a QT-SCA in Section 5.3.2. However, the complexity of QT-SCA significantly increases as the size of network becomes large, especially in multicarrier FD-like systems. Next, we will resort to graph neural network to efficiently deal with this optimization problem.

6.3 Graph Learning in MDD-CF networks

In order to leverage the structural information of MDD-CF network to solve the complicated problem, as that formulated in (6.1), the heterogeneous graph learning based PA scheme is introduced. We aim to learn a scalable and transferable HGNN to efficiently distribute both APs' and MSs' transmit power to maximize the SE of the distributed MDD-CF systems.

6.3.1 Definition of Heterogeneous Graph

A heterogeneous graph can be represented as $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of nodes, and $\mathcal{E}$ is the set of edges. The heterogeneous graph has a node type mapping function $\phi : \mathcal{V} \longrightarrow \mathcal{Q}$ and an edge type mapping function $\psi : \mathcal{E} \longrightarrow \mathcal{P}$, where $\mathcal{Q}$ and $\mathcal{P}$ denote the sets of predefined node types and link types, $|\mathcal{Q}| + |\mathcal{P}| > 2$ [185]. Explicitly, we can write $\mathcal{Q} = \{Q_1, Q_2, ...\}$ and $\mathcal{P} = \{P_1, P_2, ...\}$, where Q_i and P_j are the i-th node type and j-th link type. Let $\boldsymbol{v}_i \in \mathbb{R}^{F_v \times 1}$ denote a node with F_v-dimensional features and $\boldsymbol{e}_{i,j} \in \mathbb{R}^{F_e \times 1}$ denote an edge pointing from $\boldsymbol{v}_j$ to $\boldsymbol{v}_i$, which has F_e-dimensional features. Given the mapping function ϕ and ψ, each node belongs to a particular node type of $\phi(\boldsymbol{v}) \in \mathcal{Q}$ and each edge belongs to a specific relation of $\psi(\boldsymbol{e}) \in \mathcal{P}$. The neighborhood of a node $\boldsymbol{v}_i$ is defined as $\mathcal{N}_i = \{\boldsymbol{v}_j \in \mathcal{V} | \boldsymbol{e}_{i,j} \in \mathcal{E}\}$. Furthermore, in a heterogeneous graph, two nodes may be connected via different semantic paths. For example, an AP and an MS can be connected via two paths, namely the AP-DL-MS and MS-UL-AP links. Therefore, we introduce the concept of meta-path as in [186]. In detail, a meta-path Φ is defined as a path in the form of $Q_1 \xrightarrow{P_1} Q_2 \xrightarrow{P_2} \cdots \xrightarrow{P_n} Q_{n+1}$, which defines a composite relation $P = P_1 \circ P_2 \circ \dots P_n$ from node type Q_1 to node type Q_{n+1}, where $\circ$ denotes the composition operator on relations. Once a meta-path Φ is given, the specific neighbors $\mathcal{N}_i^{\Phi}$ of node $\boldsymbol{v}_i$ can be obtained, which are given by the set of nodes connected with $\boldsymbol{v}_i$ via meta-path Φ.

6.3.2 Heterogeneous Graph for MDD-CF Networks

Intuitively, as shown in Fig. 6.1, it is straightforward to model an MDD-CF network as a heterogeneous graph. It can be observed from the figure that there are two types of nodes, namely AP and MS, each of which is associated with two meta-paths. Specifically, for AP l, all MSs are connected with it via the meta-path Φ_1^{I} (MS-UL-AP), and the remaining APs are connected with it via the meta-path Φ_2^{I} (AP-IAI-AP). Note that, the SI caused by AP l itself is classified into Φ_2^{I} by adding a self-loop. Similarly, the meta-paths Φ_1^{II} and Φ_2^{II} associated with MS d are AP-DL-MS and MS-IMI-MS, respectively.

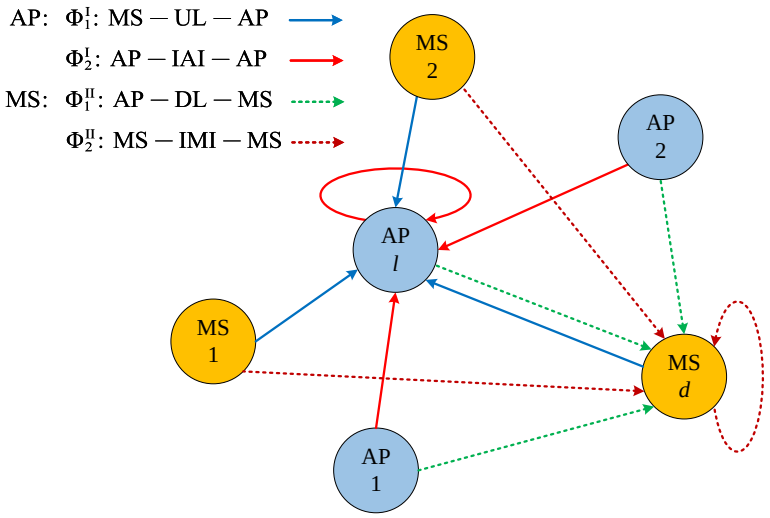


FIGURE 6.1: An exemplified graph of CF network.

In the heterogeneous graph of MDD-CF, the node feature vectors of AP l and MS d are defined as

$$\begin{aligned}\boldsymbol{v}_l \in \mathbb{R}^{(DM+3)\times 1} &= \left[\boldsymbol{\omega}_{l1}^T ... \boldsymbol{\omega}_{lM}^T, P_l, \xi_l^{\mathrm{SI}}, \xi_l^{\mathrm{IAI}}\right]^T, \\ \boldsymbol{v}_d \in \mathbb{R}^{(L\bar{M}+3)\times 1} &= \left[\boldsymbol{v}_{1d}^T ... \boldsymbol{v}_{\bar{M}d}^T, P_d, \xi_d^{\mathrm{SI}}, \xi_d^{\mathrm{IMI}}\right]^T,\end{aligned} \tag{6.2}$$

respectively, where $\boldsymbol{\omega}_{lm} = [\omega_{l1m}, ..., \omega_{lDm}]^T$ is the m-th DL equivalent subchannel gains between AP l and all the D MSs, and $\boldsymbol{v}_{\bar{m}d} = [v_{1d\bar{m}}, ..., v_{Ld\bar{m}}]^T$ is the $\bar{m}$-th UL equivalent subchannel gains between MS d and all the L APs, when the ZF beamforming as in (5.9) is applied. Moreover, to simplify the model for the sake of reducing complexity, the attribute of edges is assumed to be the Euclidean distance between any two nodes, expressed as $e_{i,j} = d_{i,j}, \forall i, j \in \mathcal{V}$. If $i = j, e_{i,j} = 0$ denotes the edge feature of self-loop.

Remark 6.1. In the distributed MDD-CF systems, each AP equipped with a baseband processor is able to independently implement DL/UL beamforming, and the CPU only needs to collect the processed signals from APs and then accomplish the final data detection. Hence, in order to fully exploit the APs' computational potentials while avoiding using long-stacked channel vectors as node features, we assume that, each AP first computes the equivalent DL/UL subchannel gains based on the estimated CSI. Then, the equivalent DL/UL subchannel gains are transmitted to the CPU as the AP/MS node features. All these are done during the offline graph training[1]. Note that in some

[1]In principle, apart from ZF, the beamforming methods, such as MMSE or matched filtering, may also be employed. If the number of antennas at individual AP and the numbers of MSs and subcarriers are small, it is possible to represent the AP and MS node features by their channel vectors, which have the size of $(2DNM + 3)$ and $(2LN\bar{M} + 3)$, respectively, when both the real and imaginary parts of the complex vectors are considered. Correspondingly, the beamforming vectors associated with PA can be directly learned at the final output layer, rather than choosing one of the known beamforming schemes to generate the network inputs. However, if the numbers of antennas at each AP, MSs and subcarriers are large, the huge dimensions of node features will make the model extremely hard to train.

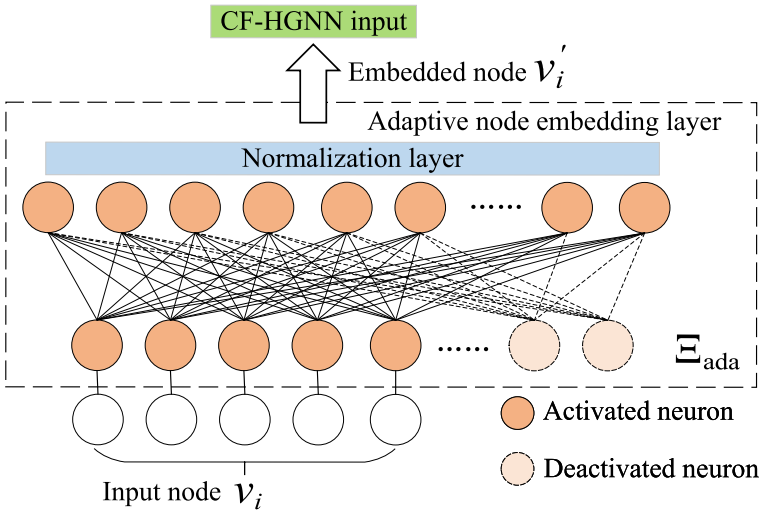


FIGURE 6.2: Adaptive node embedding layer.

references, such as [181], the authors used the geographic location information (GLI) as the input of the learning-based network to reduce the training overhead. This approach has the advantage that the low-dimensional coordinate values of AP/MS can be directly used as node features without requiring to computing beamformers. However, in MDD-based systems, power needs to be allocated among different subcarriers, while the GLI lacks not only the small-scale information of the involved communication channels but also the multi-antenna characteristics of MIMO systems. Explicitly, the GLI is unable to provide the required information and hence, is infeasible in MDD-based systems.

6.3.3 Heterogeneous Graph Learning Assisted Power Allocation in MDD-CF Networks

In this subsection, we formally present the CF-HGNN to solve the PA problem in MDD-CF networks. The architecture of CF-HGNN consists of four components: 1) adaptive node embedding; 2) meta-path based message passing; 3) meta-path based attention; 4) downstream PA learning. The overall CF-HGNN is type-specific and the parameters for processing AP and MS nodes are not shared.

6.3.3.1 Adaptive Node Embedding

In general, in order to guarantee the scalability of GNN, the size of node feature should be irrelevant to the number of nodes involved. As we can see in (6.2), the feature dimensions of input nodes are related to L and D, which vary with respect to the scales of CF networks. Hence, they are not feasible to exhibit the scalability and transferability. To tackle this problem, we propose an adaptive node embedding layer, which can

handle graph nodes with varying input feature dimensions through adaptively activating or deactivating the neurons in the multi-layer perceptron (MLP), i.e., $\Xi_{\text{ada}}(\boldsymbol{v}_l)$ and $\Xi_{\text{ada}}(\boldsymbol{v}_d)$, as depicted in Fig. 6.2. Specifically, before the normalization layer, AP and MS nodes are transformed by two embedding matrices, which are expressed as

$$\begin{aligned} \boldsymbol{v}_l^{'} &= \boldsymbol{W}_{\text{AP}}^{(:,1:DM+3)} \boldsymbol{v}_l, \\ \boldsymbol{v}_d^{'} &= \boldsymbol{W}_{\text{MS}}^{(:,1:L\bar{M}+3)} \boldsymbol{v}_d, \end{aligned} \tag{6.3}$$

where $\boldsymbol{W}_{\text{AP}} \in \mathbb{R}^{F_{\text{AP}}^{'} \times F_{\text{AP}}}$ and $\boldsymbol{W}_{\text{MS}} \in \mathbb{R}^{F_{\text{MS}}^{'} \times F_{\text{MS}}}$ map different AP and MS input nodes into two feature domains with predefined sizes of $F_{\text{AP}}^{'}$ and $F_{\text{MS}}^{'}$. Note that, as the ZF beamforming is applied, the maximum input size of the embedding layer for AP nodes is subject to the number of antennas employed at individual AP, having the relationship of $F_{\text{AP}} = NM + 3$. In the context of MS, we assume $F_{\text{MS}} = L^{'}\bar{M} + 3$, where $L^{'}$ denotes the maximum number of APs that can be deployed in a CF network within a certain area.

Remark 6.2. In Section 5.2.4, we assumed that, in our proposed distributed MDD-CF system, each AP having $N \geq D$ antennas individually communicates with all the D MSs using ZF beamforming. In this case, the above-mentioned embedding layer imposes strict limitation on the maximum number of MSs, which should not exceed the number of antennas configured at each AP. However, if MSs are densely distributed or each AP is only equipped a small number of antennas, the AP node features can no longer be set as (6.2), as the ZF beamforming cannot be achieved in the case of $D > N$ [2]. Moreover, in the above model, the number of APs is restricted to $L^{'}$, meaning that the CF-HGNN cannot be generalized to the CF networks with more than $L^{'}$ APs [3]. In these cases, to guarantee the scalability of CF-HGNN and reduce the computational complexity, one possible approach is to transform the dense graph into the sparse graph with the aid of user-centric clustering [75], where each AP only serves a certain number of MSs. For example, if ZF beamforming is used, each AP can serve up to N MSs depending on the channel conditions. Given this constraint, the dimensions of the AP node features are only related to the number of antennas, while that of the MSs node features are only relied on a fraction of APs. Hence, our node feature definition of (6.2) is still applicable. In Section 6.4.4, we will present an example accounting for the user-centric clustering.

[2] Although the MMSE beamforming can be applied in the case of $D > N$, the increasing number of MSs leads to large multiuser interference and high dimensions of AP node features.

[3] In principle, $L^{'}$ can be predefined to a sufficiently large number so as to make CF-HGNN scalable to the network with densely distributed APs. However, in doing so, it will significantly increase the training overhead of the CF-HGNN.

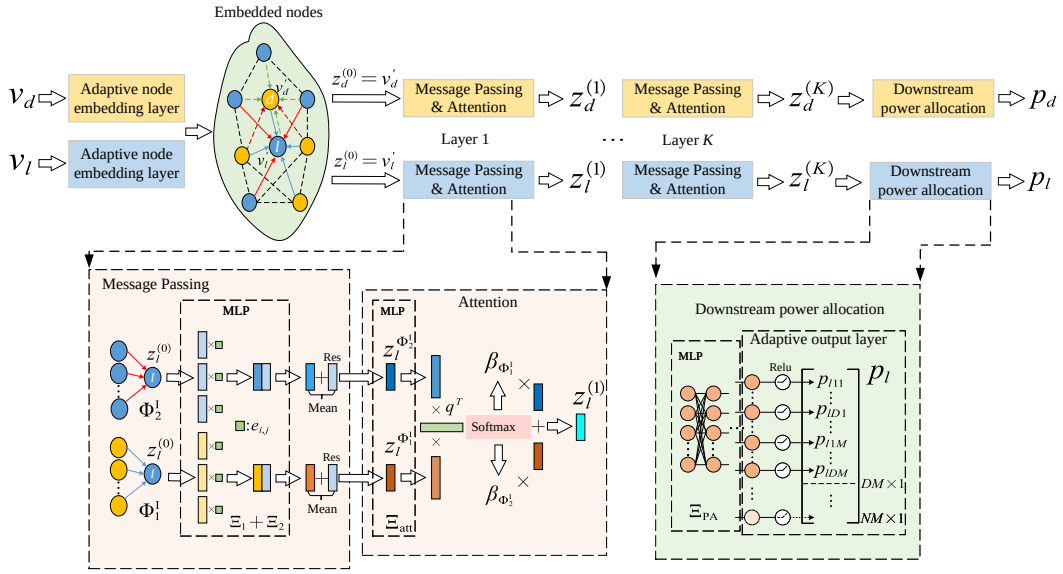


FIGURE 6.3: The overall architecture of CF-HGNN network.

6.3.3.2 Meta-Path Based Message Passing

According to the example shown in Fig. 6.1, each AP/MS node connects with its neighboring nodes via two meta-paths, and all the meta-paths are mutually independent. Hence, within each meta-path Φ, the process of message passing from neighboring nodes to itself is formulated as

$$\boldsymbol{z}_i^{\Phi} = \text{Mean}\Big(\Xi_2\Big(\boldsymbol{v}_i^{'} \oplus \frac{1}{|\mathcal{N}_i^{\Phi}|}\Xi_1\big(\sum_{j\in\mathcal{N}_i^{\Phi}} \boldsymbol{v}_j^{'} \cdot e_{i,j}\big)\Big) + \underbrace{\boldsymbol{v}_i^{'}}_{\text{Res}}\Big), \tag{6.4}$$

where Ξ_1 and Ξ_2 represent two different MLPs, with each including the fully-connected, activation and normalization layers. The Res $\boldsymbol{v}_i^{'}$ term denotes the residual connection, which can maintain the original node information after the multi-layer message passing. $\oplus$ is the concatenation operation.

6.3.3.3 Meta-Path Based Attention

In general cases, the information update of node i is the summation of the message collected from all the meta-paths with the same coefficients, expressed as $\boldsymbol{z}_i = \sum_{\Phi} \boldsymbol{z}_i^{\Phi}$. However, in MDD-CF networks, since AP/MS nodes receive information via both the interfering path and data transmission path, it is intuitive that these two meth-paths may has quite different impact on the information integration. For instance, if an AP node is closely surrounded by MS nodes but the other AP nodes are far away from it, the meta-path Φ_1^{I} should be more important than Φ_2^{I}. Based on this observation, we propose the meta-based attention in CF-HGNN to enable the GNN to automatically learn the importance of the two meta-paths. An example of calculating the attention

vector of an AP node is as follows [187]:

$$\begin{aligned}\alpha_{\Phi_1^{\mathrm{I}}} &= \frac{1}{L}\sum_{l\in\mathcal{L}} \boldsymbol{q}^T \Xi_{\mathrm{att}}(\boldsymbol{z}_l^{\Phi_1^{\mathrm{I}}}),\\ \beta_{\Phi_1^{\mathrm{I}}} &= \frac{\exp(\alpha_{\Phi_1^{\mathrm{I}}})}{\exp(\alpha_{\Phi_1^{\mathrm{I}}})+\exp(\alpha_{\Phi_2^{\mathrm{I}}})},\end{aligned} \tag{6.5}$$

where Ξ_{att} is the MLP layer for attention, $\boldsymbol{q}$ is the learnable attention vector, $\boldsymbol{z}_l^{\Phi_1^{\mathrm{I}}}$ denotes the aggregated information via meta-path Φ_1^{I}. Then, the final node representation of AP l is $\boldsymbol{z}_l = \beta_{\Phi_1^{\mathrm{I}}}\boldsymbol{z}_l^{\Phi_1^{\mathrm{I}}} + \beta_{\Phi_2^{\mathrm{I}}}\boldsymbol{z}_l^{\Phi_2^{\mathrm{I}}}$, where $\boldsymbol{z}_l \in \mathbb{R}^{F'_{\mathrm{AP}}\times 1}$. Similarly, $\boldsymbol{z}_d \in \mathbb{R}^{F'_{\mathrm{MS}}\times 1}$ at MS d can be obtained. It is noteworthy that the process of message passing plus attention can be iteratively implemented for K times by initializing $\boldsymbol{z}_l^{(0)} = \boldsymbol{v}'_l$ and $\boldsymbol{z}_d^{(0)} = \boldsymbol{v}'_d$, so as to collect the high-hop neighbors. Owing to this, our proposed model can also be termed as the K-layer CF-HGNN.

6.3.3.4 Downstream Power Allocation Learning

After the K-th iteration, the final representation of AP and MS nodes, i.e., $\boldsymbol{z}_l^{(K)}$ and $\boldsymbol{z}_d^{(K)}$, are used for the downstream PA learning, which can be expressed as:

$$\begin{aligned}\boldsymbol{p}_l \in \mathbb{R}^{DM\times 1} &= \mathrm{Relu}\Big(\Xi_{\mathrm{PA}}(\boldsymbol{z}_l^{(K)})^{(1:DM)}\Big),\\ \boldsymbol{p}_d \in \mathbb{R}^{\bar{M}\times 1} &= \mathrm{Relu}\Big(\Xi_{\mathrm{PA}}(\boldsymbol{z}_d^{(K)})\Big),\end{aligned} \tag{6.6}$$

where Ξ_{PA} denotes the MLP for PA learning, and $\mathrm{Relu}(\cdot)$ is used to constrain the power to be positive. Corresponding to the adaptive node embedding considered in part 1), in order to make it feasible for the network with different number of MSs, the adaptive output layer can support the PA for up to NM elements by activating or deactivating the neurons. By contrast, for UL PA, provided that the number of available UL subcarriers is fixed, the size of the output layer for MS node remains to $\bar{M}$, as each MS only needs to allocate UL power.

The overall architecture of the CF-HGNN for an AP node is shown in Fig. 6.3, which is the same for MS nodes, except the different size of output layer as above-mentioned. Note that all the MLPs shown in Fig. 6.3 are AP-specific, which means that the MLPs for MSs are different from that for APs.

Finally, to perform the CF-HGNN training in an unsupervised way, we define the loss function as

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}) = \mathbb{E}\Bigg[-\Lambda_{\mathrm{SE}} + \sum_{d=1}^{D}\Bigg(\kappa_1 \mathrm{Relu}\big(\chi_{\mathrm{DL}} - \sum_{m\in\mathcal{M}} \ln(1+\mathrm{SINR}_{d,m})\big) \\ + \kappa_2 \mathrm{Relu}\big(\chi_{\mathrm{UL}} - \sum_{\bar{m}\in\bar{\mathcal{M}}} \ln(1+\mathrm{SINR}_{d,\bar{m}})\big) \\ + \kappa_3 \mathrm{Relu}\big(\sum_{\bar{m}\in\bar{\mathcal{M}}} p_{d\bar{m}} - P_d\big)\Bigg) + \sum_{l=1}^{L} \kappa_4 \mathrm{Relu}\big(\sum_{m\in\mathcal{M}}\sum_{d\in\mathcal{D}} p_{ldm} - P_l\big)\Bigg],\end{aligned} \tag{6.7}$$

where $\boldsymbol{\theta}$ denotes all the parameters of the neural network, κ_i denotes the weighted factor, and the expectation is taken with respect to the channel realizations. In (6.7), each ReLU penalty term has a positive value, only if the DL/UL QoS requirements and transmit power budgets are not satisfied. This will enforce the training process towards satisfying the given requirements. Additionally, in (6.7), the positive parameters κ_i give different priorities to the penalty terms.

6.4 Simulation Results and Analysis

Let us now demonstrate the achievable SE performance of the distributed MDD-CF systems supported by our proposed CF-HGNN.

6.4.1 Simulation Setup

The large-scale fading model is given by [75]:

$$\beta[\mathrm{dB}] = -30.5 - 36.7\log_{10}(d) + \sigma_{\mathrm{sh}} z, \tag{6.8}$$

where d denotes the distance between any two nodes, $\sigma_{\mathrm{sh}} z$ is the shadowing fading with a standard deviation of $\sigma_{\mathrm{sh}} = 4$ dB and $z \sim \mathcal{N}(0,1)$. Furthermore, we assume that APs are capable of providing 30 dB of IAI suppression in the propagation/analog domain by employing the existing approaches [72]. Then, assume that the 12-bit ADCs are applied, the MDD system can suppress IAI up to 72 dB (i.e., $\xi_l^{\text{MDD-IAI}} = -72$ dB, $\forall l$, in (5.6)), of which 42 dB is attributed to the digital-domain cancellation by FFT [4]. By contrast, as MSs are of lightweight equipments with single antenna and can hardly share channel knowledge with other MSs, they can not actively suppress IMI. However, similar to APs, MSs are able to cancel 42 dB IMI in the digital domain (i.e., $\xi_d^{\text{MDD-IMI}} =$

[4]The 12-bit ADC has a maximum dynamic range of 42dB, which means it can accommodate up to 42 dB of power of IAI, and transform it into digital signal without extra quantization noise. Then, in the digital domain, since the interference signal is transmitted over DL subcarriers, which is mutually orthogonal to the desired UL signal, MDD systems supported by the FFT operation can then remove the remaining IAI.

TABLE 6.1: Simulation parameters

Default parameters	Value
Cell area ($S_{\mathrm{D}} \times S_{\mathrm{D}}$)	$(400 \times 400)\mathrm{m}^2$
Number of antennas per AP (N)	8
Number of DL/UL subcarriers ($M, \bar{M}$)	(4, 2)
AP and MS power ($P_l, P_d, \forall l, d$)	$(40, 30)$ dBm
QoS requirements ($\chi_{\mathrm{DL}}, \chi_{\mathrm{UL}}$)	(0.5, 0.1) nats/s/Hz
Noise power (σ^2)	-94 dBm
Delay taps (U)	4
Residual SI level at AP ($\xi_l^{\mathrm{SI}}, \forall l$)	-120 dB
Residual SI level at MS ($\xi_d^{\mathrm{SI}}, \forall d$)	-110 dB

-42 dB, $\forall d$, in (5.3)) with the aid of the FFT operation [148]. Unless otherwise noted, the simulation parameters of the MDD-CF network are listed in Table 6.1.

As for the settings of the neural network, we adopt a 2-layer CF-HGNN based on Pytorch Geometric [188]. The general MLPs Ξ_1 and Ξ_2 in (6.4) during the message passing stage contains multiple fully-connected linear layers followed by the LeakyRelu activation layer and batch normalization layer. By contrast, the Ξ_{PA} in (6.6) is employed with multiple fully-connected linear layers without batch normalization layer. Moreover, the Ξ_{ada} in (6.3) is a single-layer MLP with one fully-connected linear layer followed by one batch normalization layer, while Ξ_{att} in (6.5) is a single-layer MLP with only one fully-connected linear layer. As formulated in (6.7), the overall learning is unsupervised without any ground truth. To optimize the CF-HGNN, we adopt the Adam optimizer with a learning rate of 0.001 [189]. Furthermore, we empirically set κ_i in (6.7) as $\{0.1, 1, 0.1, 0.1\}$ during training. For the training data, we randomly generate 10000 and 1000 CF network layout samples for training and testing, respectively, under the assumption that APs and MSs are uniformly distributed within a square area of $(S_D \times S_D)\mathrm{m}^2$. The batch size for training is 64, and the network parameters are only updated during training, which stay constant during testing. The CF-HGNN is run on a GeForce GTX laptop 3080Ti, while the other algorithms are implemented on the 12th Gen Intel(R) Core(TM) i7-12700H 2.70 GHz.

For comparison, we introduce QT-SCA as a benchmark, which is implemented using the CVX tool during simulations [177]. In addition, our proposed CF-HGNN is also compared with the greedy unfair allocation method [151], where the water-filling algorithm is assumed by each AP/MS to distribute their power over the DL/UL subcarriers, regardless of the QoS constraints.

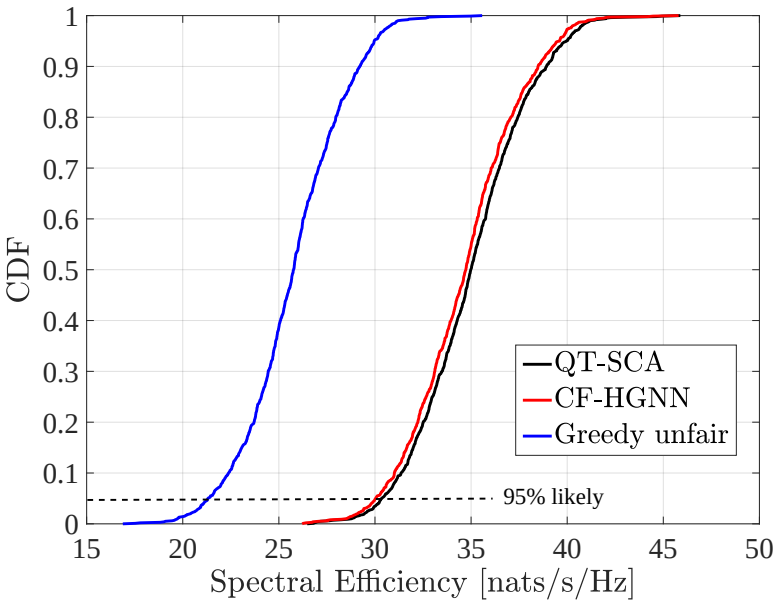


FIGURE 6.4: Cumulative distribution of SE, when the MDD-CF network has $L = 24$ APs and $D = 6$ MSs.

6.4.2 Performance Comparison

In this section, we make a comprehensive performance comparison between different PA methods in MDD-CF networks, where the numbers of MSs and APs are fixed during training and testing, and are set to $L = 24, D = 6$. In this case, the embedding matrices in Ξ_{ada} and Ξ_{PA} act as identical matrices and are not learnable.

Firstly, we compare the QT-SCA, CF-HGNN and the greedy unfair methods in terms of the SE distribution obtained from 1000 testing CF network layouts. As shown in Fig. 6.4, the CF-HGNN achieves nearly the same performance as the QT-SCA in terms of the 95%-likely SE. In more detail, the SE performance gaps between the QT-SCA and CF-HGNN with regard to these 1000 network layouts are depicted in Fig. 6.5. It can be observed that the absolute SE gaps between these two methods are lower than 1.5 nats/s/Hz. Furthermore, there are several layouts, where the CF-HGNN outperforms the QT-SCA, which reflects that the CF-HGNN is capable of learning the near-optimal solutions for PA in MDD-CF networks. As shown in Fig. 6.4, both the QT-SCA and CF-HGNN significantly outperform the greedy unfair method. The rationale behind is that although the greedy unfair method can maximize the SE of the classic HD-based multiuser OFDM systems [151], it lacks the capability to manage the complicated interference in the CF FD-like systems, hence leading to poor performance at both APs and MSs.

Secondly, we randomly select one of the CF network layouts used in the testing samples, as shown in Fig. 6.6, to deeply investigate the PA attained by the QT-SCA and CF-HGNN. To make the drawing clear in Fig. 6.6, the DL connections with transmit

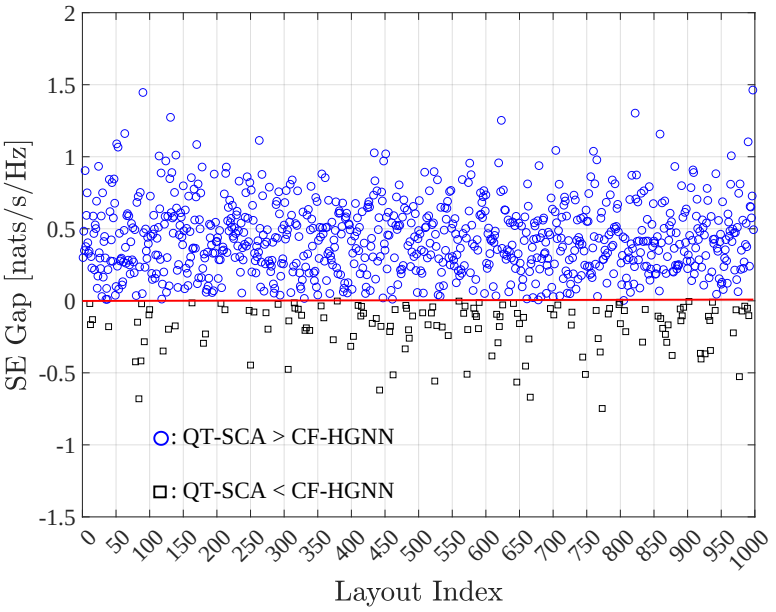


FIGURE 6.5: SE performance gap between QT-SCA and CF-HGNN, when the MDD-CF network uses $L = 24$ APs to support $D = 6$ MSs.

power less than $2W$ are omitted, and different colored lines are used to denote the DL connections obtained by either methods or by both. It is not surprise that, as Fig. 6.6 shows, except MS 2, both the CF-HGNN and QT-SCA yield the same subset of the major serving APs for each of MSs, and also obtain the similar results of UL PA. As for MS 2, since it locates relatively far away from APs, more APs are required to transmit signal to it in order to meet the demand of DL's QoS. Although different DL connections are obtained by the two methods, there is in fact no difference between them. For example, apart from the common connections, the CF-HGNN picks AP 3 and 9 to serve MS 2, while the QT-SCA chooses AP 8 located between AP 3 and 9 to serve MS 2. Also, AP 20 and 17 with the similar distances from MS 2 are selected by the CF-HGNN and QT-SCA, respectively.

It is worth mentioning that the DL power (and UL power) used for Fig. 6.6 is the sum of power allocated to DL subcarriers (and UL subcarriers). In Table 6.2, an example showing the detailed PA to subcarriers at AP 1, MS 2 and MS 6 are presented. It can be observed that the CF-HGNN and greedy unfair methods yield the more consistent allocation among the subcarreirs, which means that the power distribution is largely relied on large-scale fading, while the QT-SCA has better capability of exploiting the small-scale fading. Moreover, from the results of the CF-HGNN and QT-SCA, we can see that MS 6 has a lower UL transmit power than MS 2. The reason behind is that MS 6 is in the close proximity of MS 1, the increase of its UL transmit power may cause not only more SI on the DL reception, but also larger IMI on MS 1, leading to the degradation of SE. By contrast, the greedy unfair scheme allocates power depending only on the quality of communication channels regardless of the effect of interference.

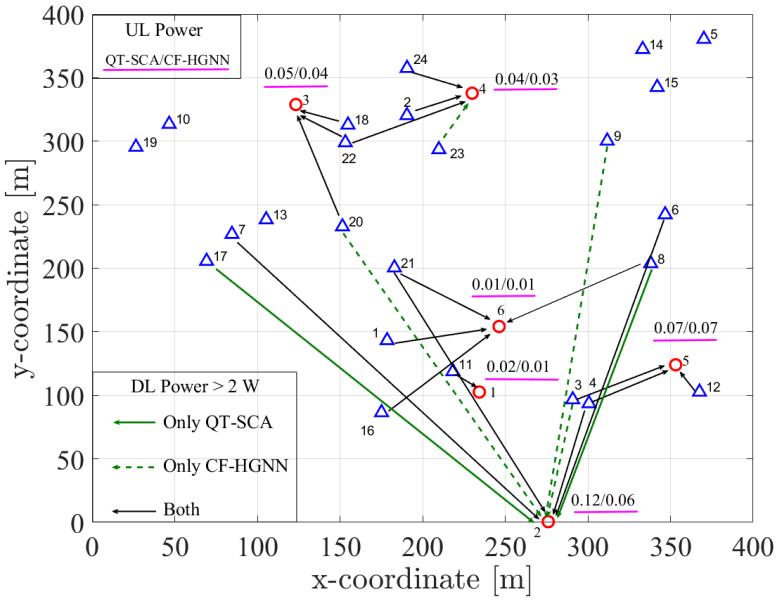


FIGURE 6.6: An MDD-CF network topology with $L = 24$ APs and $D = 6$ MSs.

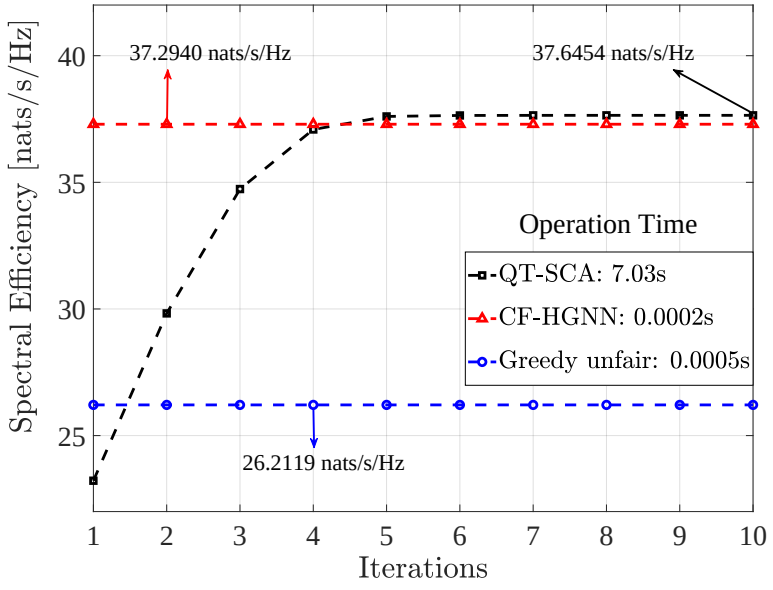


FIGURE 6.7: SE convergence behavior and operation time of different methods, where $L = 24, D = 6$.

TABLE 6.2: PA among Subcarriers at AP 1, MS 2 and MS 6.

Connections	Method	DL subcarriers' power (*W*)				DL power (*W*)	UL subcarriers' power (*W*)		UL power (*W*)
AP 1→MS 6, MS 6→APs	QT-SCA	1.89	0.81	1.90	1.89	**6.49**	0.006	0.008	**0.014**
	CF-HGNN	1.77	1.74	1.78	1.79	**7.08**	0.003	0.010	**0.013**
	Greedy unfair	1.26	1.19	1.25	1.25	**4.95**	0.499	0.501	**1.000**
AP 1→MS 2, MS 2→APs	QT-SCA	0.75	0.33	0.49	0.31	**1.88**	0.026	0.038	**0.064**
	CF-HGNN	0.25	0.23	0.22	0.27	**0.97**	0.000	0.117	**0.117**
	Greedy unfair	0.00	0.00	0.00	0.00	**0.00**	0.000	1.000	**1.000**

TABLE 6.3: Generalization to Different Cell Sizes. The SE of the CF-HGNN are normalized by that of the QT-SCA.

Cell length (S_{D})	QT-SCA (nats/s/Hz)	DCF-HGNN	GCF-HGNN
350	37.85	96.64%	95.19%
300	40.64	98.08%	93.13%
250	43.76	98.42%	91.84%
200	46.84	99.91%	91.78%
150	50.10	101.42%	92.79%

The performance of different methods in terms of SE, convergence behavior and operation time is plotted in Fig. 6.7. As expected, the QT-SCA and CF-HGNN achieve the comparable SE, with a performance gap of only 0.35 nats/s/Hz. However, the QT-SCA is an iterative algorithm with very high computational complexity, which converges within 6 iterations and takes 7.03s operation time in total. On the contrary, with the aid of the parallel computation of GPU, the CF-HGNN is capable of reaching 99% of the SE achieved by the QT-SCA but using only 10^{-4} times of its operation time. As to the greedy unfair method, although it has a similar operation time as the CF-HGNN, due to the fact that only the water-filling algorithm is applied at both APs and MSs, its SE performance is significantly worse than that of the other two methods.

Next, we evaluate the generalization performance of the proposed CF-HGNN in MDD-CF networks with different cell sizes. As shown in Table 6.3, the results under the Generalized CF-HGNN (GCF-HGNN) are obtained by the CF-HGNN model trained on the specific network with $S_{\mathrm{D}} = 400$, while the results under the Dedicated CF-HGNN (DCF-HGNN) are obtained by the CF-HGNN models trained correspondingly using different cell sizes. It can be observed from the table that as the cell size reduces, the performance of the GCF-HGNN degrades slightly, but is still in the acceptable range. Note that, the DCF-HGNN finally outperforms the QT-SCA when $S_{\mathrm{D}} = 150$, which means that the DCF-HGNN is capable of obtaining a better PA strategy in dense networks.

The comparison between QT-SCA and CF-HGNN in terms of computational complexity and operation time are then analyzed. According to [179], the approximated computational complexity of the QT-SCA method is $\mathcal{O}\big((LDM + 2DM + 3D\bar{M})^2(L + D + 3DM_{\mathrm{sum}})^{2.5} + (L + D + 3DM_{\mathrm{sum}})^{3.5}\big)$ per iteration. As to the CF-HGNN, the computational complexity is mainly due to the matrix computation, as shown in (6.4)-(6.6). Table 6.4 summarizes the results for the different values of L, D, M and $\bar{M}$, showing that when doubling the numbers of APs, MSs and subcarriers, the computational complexity of the QT-SCA increases much more quickly than that of the CF-HGNN, which is about three orders for the QT-SCA versus one order for the CF-HGNN. The operation

TABLE 6.4: Computational complexity comparison

APs (L)	MSs (D)	DL/UL subcarriers ($M/\bar{M}$)	Methods	
			QT-SCA	CF-HGNN
6	2	4/2	4.48×10^8	6.83×10^6
12	4	8/4	4.78×10^{11}	2.75×10^7
24	8	16/8	6.79×10^{14}	1.48×10^8

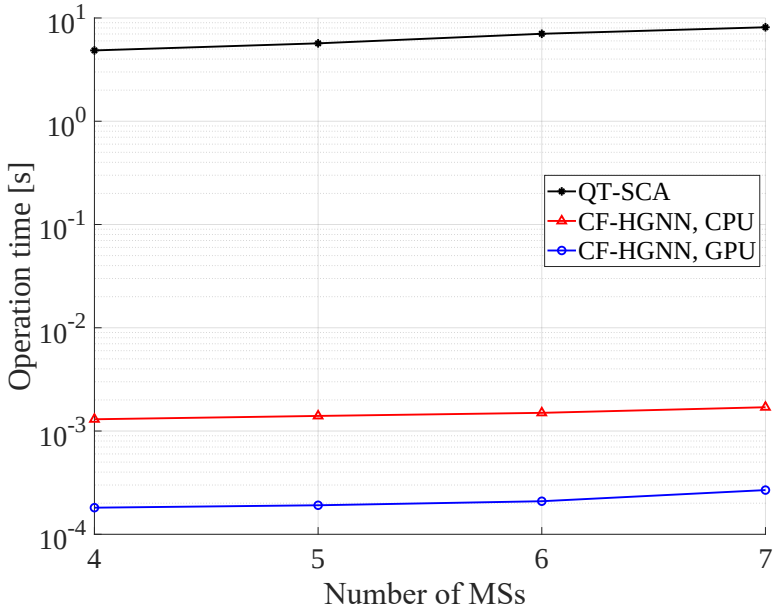


FIGURE 6.8: Comparison of operation time between the QT-SCA and CF-HGNN methods.

time cost by the QT-SCA and CF-HGNN is plotted in Fig. 6.8. Explicitly, the CF-HGNN spends much less time than the QT-SCA to accomplish the PA. Moreover, thanks to the parallel computation of GPU, the CF-HGNN trained on GPU has the lowest operation time.

6.4.3 Scalability: Case 1

The simulation results in Section IV-B demonstrated that the CF-HGNN is able to achieve the similar performance as the QT-SCA. However, the CF-HGNN considered so far can only be applied to the specific MDD-CF networks with the fixed numbers of nodes and subcarriers, as the fully-connected liner layers in Ξ_{ada} and Ξ_{PA} act as identical matrices. With this regard, to enable the CF-HGNN to be adaptive to the varying MDD-CF networks, we now invoke the adaptive node embedding layer and adaptive output layer. In particular, we consider two cases regarding to the scalability of CF-HGNN. In Case 1, we still assume that $L < L' = 24$ and $N = 8$, and all MSs are served by all APs using the ZF beamforming under the constraint of $N \geq D$. Specifically,

10000 MDD-CF network samples with the specific numbers of nodes and subcarriers are leveraged to train the DCF-HGNN. For fair comparison, to train the Adaptive CF-HGNN (ACF-HGNN), 10000 MDD-CF network samples consisting of various numbers of nodes and subcarriers are used.

Table 6.5 shows that by employing the adaptive layers, the ACF-HGNN can attain the relatively stable performance, when dealing with the various networks. Hence, it is feasible for implementing dynamic PA in MDD-CF networks. On the contrary, the QT-SCA and DCF-HGNN have to re-solve the complicated optimization problem or re-train the CF-HGNN, once the CF network has some changes in terms of the numbers of APs, MSs and subcarriers. Note that, in the table, the SE decreases with the increase of the number of subcarriers. This is because the total power allocated to subcarriers is constrained, while the SE is normalized by M_{sum}, as shown in (6.1). Apparently, the total SE obtained by multiplying M_{sum} with the value in the table always increases with the increase of the number of subcarriers, which is attributed to the subcarrier diversity.

6.4.4 Scalability: Case 2

Although the proposed ACF-HGNN can handle the PA in various MDD-CF networks, it can hardly cut the mustard when the numbers of APs and MSs become too large, i.e., when $L \gg L'$ and $D \gg N$, as explained in Remark 6.2. Therefore, in Case 2, we resort to the user-centric clustering strategy to transform the dense graph into the sparse graph, thereby maintaining the scalability of the CF-HGNN for operation in the large-scale MDD-CF networks. For the sake of explanation, here we consider an extreme scenario, where each AP is only equipped with one antenna and serves only one MS using the ZF beamforming. Correspondingly, the clustering can be achieved in two steps[5]: 1) Initialize the set of APs serving MS d as $\mathcal{L}_d = \emptyset, \forall d = 1, ..., D$, and the set of MSs assigned to AP l as $\mathcal{D}_l = \emptyset, \forall l = 1, ..., L$. To guarantee a non-zero SE, each MS d is firstly assigned to a master AP by following the optimization of

$$l = \underset{l \in \{1,...,L\}, m \in \{1,...,M\}}{\operatorname{argmax}} \omega_{ldm}, \ \mathcal{L}_d = \{l\}, \ \mathcal{D}_l = \{d\} . \tag{6.9}$$

After this optimization, we have $|\mathcal{L}_d| = 1, \forall d = 1, ..., D$. 2) If there are still idle APs, which satisfy $\{l' | \mathcal{D}_{l'} = \emptyset\}$, allocate them to MSs according to the optimization of

$$d = \underset{d \in \{1,...,D\}, m \in \{1,...,M\}}{\operatorname{argmax}} \omega_{l'dm}, \ \mathcal{D}_{l'} = \{d\}, \ \mathcal{L}_d = \mathcal{L}_d \cup \{l'\} . \tag{6.10}$$

[5]Note that the considered two-stage clustering approach is suboptimal, and becomes less efficient in the case of $N > 1$, where each AP can simultaneously serve more than one MS using the increased spatial degrees of freedom. To this end, our future work motivates to propose a GNN-assisted deep Q-learning network so as to obtain the optimal user grouping and PA at the same time.

TABLE 6.5: Generalization to the different numbers of APs, MSs and subcarriers. The SE of the CF-HGNN and greedy unfair algorithm are normalized by that of the QT-SCA.

APs (L)	MSs (D)	DL/UL subcarriers ($M/\bar{M}$)	QT-SCA (nats/s/Hz)	DCF-HGNN	ACF-HGNN	Greedy unfair
24	4	4/2	26.42	98.07%	94.14%	74.45%
	5		31.34	98.02%	94.51%	74.06%
	6		34.23	97.78%	94.77%	73.06%
	7		36.24	98.18%	95.45%	73.68%
20	6	4/2	33.04	98.09%	97.22%	72.79%
16			29.99	97.70%	95.07%	71.16%
12			25.69	98.52%	94.08%	69.95%
8			21.13	97.40%	94.46%	68.67%
16	6	32/16	9.26	98.81%	91.58%	72.35%
		16/8	15.53	99.03%	91.50%	68.45%

To train the Large-Scale ACF-HGNN (LSACF-HGNN), we have 10000 training samples collected from the networks with $L = 12$ single-antenna APs and $D = 6$ MSs uniformly distributed within an area of $(200\sqrt{2} \times 200\sqrt{2})\text{m}^2$. We then increase the number of APs and MSs in testing samples while the densities of AP and MSs are fixed. Note that all the involved training and testing samples have been pre-processed by the proposed clustering algorithm. As shown in Table 6.6, although the performance of the LSACF-HGNN is much better than that of the greedy unfair method, the gap between the QT-SCA and the LSACF-HGNN becomes larger as the network size increases. The rationale behind is possibly because the larger CF network has much more complicated CLI problem than the smaller one. Hence, the LSACF-HGNN may require extra training or an improved model structure to manage the CLI. The detailed analysis is out of the scope of this chapter and is left for the future research. However, we should mention that, as shown in Table 6.4, the QT-SCA relied optimization is extremely time-consuming, when it is applied to the large-scale CF networks, while the LSACF-HGNN is capable of accomplishing PA with high-efficiency.

6.5 Chapter Summary and Conclusions

In this chapter, we proposed a distributed MDD-CF system, and investigated the PA at both AP and MS sides for SE maximization under the constraints of QoS requirements. In order to solve the non-convex and NP-hard PA problem, we first proposed the QT-SCA algorithm, which achieves the optimization in the classic way. Then, the CF-HGNN was proposed to solve the optimization from a learning perspective. The CF-HGNN was trained in an unsupervised fashion with unlabeled data, which significantly reduces the system overhead. Our studies and numerical results show that the CF-HGNN is capable of achieving comparable SE performance to the QT-SCA but demanding much less operation time and computational complexity. The CF-HGNN significantly outperforms the greedy unfair method in terms of SE performance. Furthermore, with the aid of adaptive node embedding layer and adaptive output layer, the CF-HGNN can implement PA in the MDD-CF networks with various numbers of APs, MSs and subcarriers. Additionally, with the aid of user clustering, the CF-HGNN trained based on a relatively small-scale MDD-CF network can be generalized for operation in the large-scale MDD-CF networks.

TABLE 6.6: Generalization to the Large-scale MDD-CF networks. The SE of the CF-HGNN and greedy unfair algorithm are normalized by that of the QT-SCA.

Cell length (S_D)	APs (L)	MSs (D)	QT-SCA (nats/s/Hz)	LSACF-HGNN	Greedy unfair
400	24	12	22.93	93.68%	65.68%
450	30	15	28.28	92.33%	65.38%
500	38	19	35.18	91.42%	65.66%
$400\sqrt{2}$	48	24	43.93	88.05%	65.65%
$400\sqrt{3}$	72	36	63.92	82.09%	65.35%
800	96	48	82.10	80.58%	67.23%

Chapter 7

Conclusions and Future Research

In this chapter, we first present the overall summary and conclusions of this thesis in Section 7.1. Then, several potential research topics for future research are discussed in Section 7.2.

7.1 Thesis Summary and Conclusions

- **Chapter 2**: This chapter proposed an adaptive beamforming assisted SIC scheme for both IBFD- and MDD-based MIMO hybrid beamforming systems, where either analog precoder at transmitter or analog combiner at receiver can be leveraged to mitigate SI, depending on which side has more antenna elements. The numerical results show that, on the basis of the CCD algorithm, the proposed hybrid beamforming scheme is capable of providing a big range of SI reduction, while at only little cost of system performance with the aid of ZF digital precoding and MMSE digital combining. Moreover, the Rician factor of SI channel, the number of antennas and the angle between transmitter and receiver antenna arrays may significantly affect the SIC performance. Furthermore, by exploiting the time-domain reciprocity between DL and UL channels and the correlation existing among subcarriers, a set of frequency-domain orthogonality-achieving PSs has been designed for the CE in the MDD MIMO systems. Based on the well-designed PSs, both the LS and LMMSE approaches were introduced for CE. Our simulation results demonstrate that when the number of MSs is relatively small and the PSs are evenly distributed over the subcarriers, the LS approach outperforms the LMMSE approach in terms of the MSE and the average sum rate. However, when the Proposition 1 in this chapter is no longer satisfied due to a large number of MSs, the LMMSE approach is near-optimum and shows significant advantage over the LS approach.

- **Chapter 3**: This chapter studied the RA optimization, including user fairness, SI suppression and hybrid beamforming design, in the large-scale MDD MIMO mmWave systems. To solve the NP-hard problem, the joint DL/UL RA optimization problem was divided into the suboptimal SA and PA problems. Specifically, for the SA, an IFG algorithm was proposed, which can fully exploit the spatial multiplexing gain to attain the sum-rate maximization and coarse fairness. For the PA, the DL PA problem was transformed to a convex solvable problem with the aid of the full-digital ZF precoding, which is then solved by the New-Raphson method. By contrast, the UL PA was solved by the conventional water-filling algorithm. Our studies and simulation results show that the proposed RA algorithm is capable of attaining the near upper bound performance that is only achievable by the conventional unfair greedy algorithm. Furthermore, to reduce the complexity of the full-digital precoder/combiner, two hybrid precoding schemes, namely matrix factorization and direct approach, and one hybrid combining scheme were proposed for the implementation of RA. The studies reveal that the direct approach is capable of achieving not only higher sum rates but also attaining a more robust fairness among the DL MSs than the matrix factorization, when the number of RF chains is relatively low. Additionally, the proposed combining scheme exhibits the SIC capability, which enables to enable the RA operations to be free from SI. Lastly, in order to demonstrate the potentials of MDD against the HD modes of FDD and TDD, we presented an example of the MU-SISO system assisted by the unfair greedy RA. The results reveal that MDD can benefit from a larger multiuser diversity and hence, outperform both TDD and FDD in terms of the time-frequency usage.

- **Chapter 4**: This chapter explored the application of MDD in high-mobility scenarios. In order to combat the channel aging problem, two types of well-designed MDD frame structures and the corresponding WPs, namely the general WP and DD-WP, were proposed and studied. For comparison, the TDD frame structures related to the 3GPP standards and their variant forms were presented. The studies and simulation results demonstrate that the MDD employed with the proposed frame structures can significantly relieve the channel aging problem. In particular, when the Type I frame structure is implemented, as the UL pilots and DL symbols can be transmitted concurrently, MDD allows to achieve the consistent prediction accuracy. Hence, its achievable performance is not largely affected by the different time-varying fading situations. It was also shown that when the relative speed between BS and MSs is not too high, the MDD with partially activated pilots can achieve the highest sum rates within one frame. On the other hand, when the Type II frame structure is considered, since the received UL symbols can be used as the pilots for channel prediction in the principle of DD-WP, the DL and UL transmissions in MDD still perform better than that in TDD. Moreover, owing to the advantages of SIC in digital domain, MDD can also outperform

IBFD over fast time-varying channels, provided that the SI mitigation in IBFD is imperfect. Lastly, the approximated lower bounds for the achievable rates of the proposed MDD and TDD systems were derived, which were shown to agree closely with the results obtained from the Monte-Carlo simulations.

- **Chapter 5**: This chapter studied the MDD assisted CF-mMIMO systems. To maximize the SE of MDD-CF systems, the optimization of AP-selection, SA and PA were proposed under the constraints of MSs' QoS. Two application scenarios were considered, which assume that DL/UL transmissions occur in one CT interval or in one radio frame, respectively. In the case of one CT interval, to facilitate the mixed-integer optimization, the inner association between the continuous variables for PA, and the binary variables for AP-selection and SA were leveraged to transform the original problem into a continuous-integer convex-concave problem. Then, a QT-SCA iterative algorithm was proposed to achieve the SE maximization within one CT interval. Simulation results demonstrate that the MDD-CF can attain nearly the same performance as the TDD-CF, but significantly outperforms the IBFD-CF owing to the advantages of CLI cancellation. In the case of one radio frame, in light of the capability of simultaneously transmitting pilots and DL symbols, a TPCT interval was specifically designed for MDD-CF and IBFD-CF systems to transmit extra information. Then, to achieve the SE maximization within one radio frame, a two-step iterative algorithm based on bisection method was proposed. According to the simulation results, both MDD-CF and IBFD-CF schemes are capable of taking the advantage of Phase I in TPCT to achieve higher sum rates than TDD-CF in high-mobility scenarios.

- **Chapter 6**: This chapter solved the PA problem in the MDD-CF mMIMO networks from a learning perspective. As the computational complexity of the QT-SCA method proposed in Chapter 5 grows exponentially with the numbers of APs, MSs and subcarriers in MDD-CF increases, an HGNN assisted PA method, named as CF-HGNN, was proposed, so as to solve the PA more efficiently. The CF-HGNN consists of adaptive node embedding, meta-path based message passing, meta-path based attention and downstream PA learning, and was trained with unlabeled data in an unsupervised fashion. Our studies and simulation results show that, in a specific MDD-CF network, the CF-HGNN is capable of using 10^4 times less of operation time to achieve 99% of the SE achievable by the QT-SCA. Furthermore, the scalability of CF-HGNN was explored. The studies show that the CF-HGNN employed with adaptive node embedding and adaptive output layers can adapt to the various MDD-CF networks with different numbers of APs, MSs and subcarriers. Additionally, with the aid of user clustering, the CF-HGNN can deal with the PA in large-scale MDD-CF networks, where the cell size and the numbers of APs and MSs may be very large.

In conclusion, the research results of the thesis demonstrate that MDD employs a range of merits to enable it for possible applications in the future wireless communications systems. MDD is enabled with the advantages of TDD, FDD and IBFD schemes, while allowing to circumvent their main shortcomings. For example, the more flexible time-frequency resource usage than TDD and FDD and the more efficient management of SI and CLI than IBFD make MDD have the potential to achieve high spectral efficiency. Like IBFD systems, in MDD systems, UL pilots for channel estimation can be transmitted along with the DL data transmission. Therefore, employing MDD can fundamentally solve the channel aging problem in high mobility communications scenarios, while avoiding the SI problem in IBFD systems. Furthermore, owing to its better interference management capability than IBFD, MDD can be deployed with the CF wireless networks to achieve high efficiency wireless networking. Owing to its advantages as demonstrated in this thesis, MDD can be envisaged as a duplexing scheme, which may be combined with the other multicarrier-based techniques, such as non-orthogonal multiple access (NOMA), orthogonal time frequency space (OTFS), filter OFDM, etc., to support the various types of services, including massive machine type communications (mMTC), ultra-reliable and low latency communications (URLLC) and ultra-high-speed communications, in 5G+/6G networks.

7.2 Future Research

Based on the study of this thesis, a few of further research issues are identified, which are as follows.

- **Deep learning aided frame structure design and channel prediction for application in high-mobility scenarios**: In Chapter 4, we proposed two basic frame structures as well as two Wiener filter related channel prediction approaches, namely general WP and DD-WP, for TDD and MDD systems so as to relieve channel aging problem. However, as shown in Fig. 4.7 and Fig. 4.9, the performance gaps between various TDD and MDD frame structures differ from velocity to velocity. For instance, as shown in Fig. 4.7, the MDD with partially activated pilots (i.e., MDD-1-PA) outperforms that with fully activated pilots (i.e., MDD-1(1)), when the relative velocity is higher than 280 km/h in the case of $T = 28$ or 220 km/h in the case of $T = 56$. Analogously, the pilot distribution in TDD systems also significantly affects the system performance. Hence, there should exist an optimal frame structure corresponding to a given mobility scenario. On the other side, the proposed channel prediction techniques of general WP and DD-WP are only feasible for the model-based channel aging problem. In practice, the simplified time-correlated channel fading model cannot capture the precise variations of CSI, while the complex model allows better achievable performance but

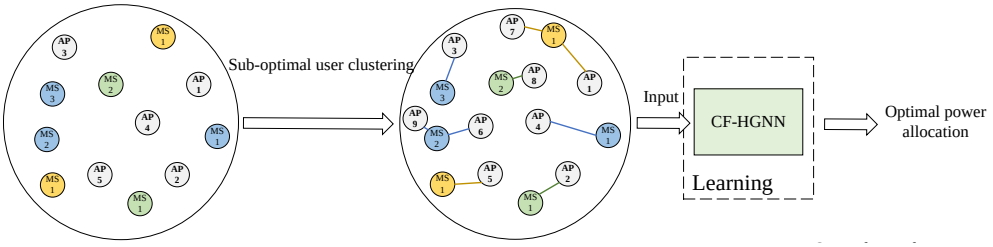


FIGURE 7.1: The process of CF-HGNN for large-scale MDD-CF networks.

at the expense of high complexity. Furthermore, the model-based methods are depended highly on the channel statistics and Doppler estimation. However, in the high-mobility communication scenarios, the Doppler estimation is very challenging. With this regard, due to the possible unstationary factors, the learning-based methods may be introduced for the optimal designs of both the frame structure and channel predictor, so that near-optimum performance can be retained in any cases. Specifically, the combination of long short term memory (LSTM) and deep reinforcement learning (DRL) may be a promising solution. This is because the frame structure design in dynamic vehicular environments can be deemed as the stochastic game described by a finite-state Markov decision process, while LSTM can be used to capture the CSI variations [190, 191].

- **Combination of MDD and orthogonal time frequency space for ultra high-mobility scenarios**: In Chapter 4 we have assumed the time-domain correlation between different OFDM symbols, as shown in Equation 4.6, and that the multipath fading within each symbol duration is time-invariant. The reason for us to use the above assumption is presented as follows. According to 5G NR standards [164], in an OFDM system having a carrier frequency of 5 GHz and a subcarrier spacing of 15 kHz and supporting a relative velocity of 500 km/h, the maximum Doppler shift is $f_D = 2.315$ kHz and hence the channel's coherence interval is about $1/(4f_D) = 108$ μs [192], which is much larger than one OFDM symbol duration equaling to about 80 μs including cyclic prefix. With this regard and considering that we have mainly focused on the channel prediction between adjacent OFDM symbols in Chapter 4, the multipath channel within one OFDM symbol duration can be deemed as time-invariant, which allows to ignore the inter-carrier interference (ICI) is neglected.

  However, 6G networks are expected to support wireless communications in the ultra high-mobility scenarios with the velocity up to 1000km/h [193]. In this case, the doubly-selective channel must be considered even within one OFDM symbol duration. Consequently, when considering the severe ICI problem, OFDM will be no longer the feasible modulation. Instead, the OTFS can be deemed as a potential candidate for gaining benefit from both the delay and Doppler domains. The recent studies show that OTFS modulates information in the delay-Doppler

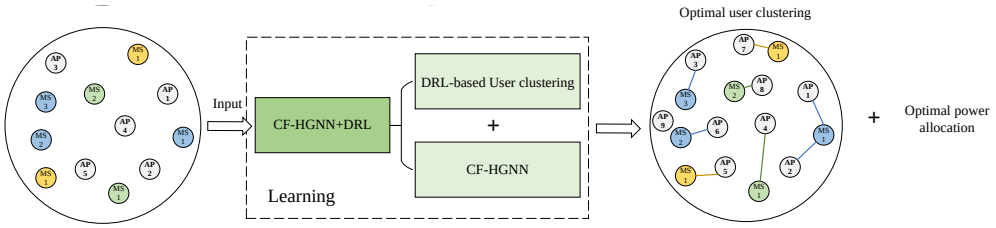


FIGURE 7.2: The process of CF-HGNN+DRL for large-scale MDD-CF networks.

domain, which is capable of providing a strong delay-resilience and Doppler-resilience [192]. Therefore, by introducing our proposed MDD to OTFS, we may substantially solve the channel aging and ICI problems to support reliable communications in ultra high-mobility scenarios, in addition to the benefits provided by OTFS.

- **DRL aided HGNN for user clustering and power allocation in MDD-CF networks**: According to our studies in Chapter 6, to apply the CF-HGNN in a large-scale MDD-CF network, where the cell size and the numbers of APs and MSs are very large, the proposed two-stage user clustering algorithm has to transform the dense graph into a sparse graph for generating training or testing samples. Then, the sparse graph is taken as the input of the CF-HGNN to implement PA. The overall process can be depicted in Fig. 7.1. However, the proposed user clustering algorithm is suboptimal and also time-consuming when the size of CF network becomes large. To this end, we may enhance the design by integrating user clustering into the learning network, as shown in Fig. 7.2, to form the CF-HGNN+DRL, which implements user clustering and PA at the same time. Although the DRL has been demonstrated to be efficient for solving the user clustering problem [194], how DRL and HGNN work cooperatively in CF-HGNN+DRL deserves the in-depth research in the future.

- **Resource allocation with integrated sensing and communication in MDD-CF networks**: With the increase of the requirements of sensing in wireless communications, the integrated sensing and communication (ISAC) in CF networks will be a promising application scenario in 5G beyond and 6G networks. Specifically, in MDD-CF networks, each AP equipped with multiple antennas can be deemed as an individual ISAC transceiver to communicate with MSs or sense targets. Furthermore, it can also cooperate with neighboring APs via fronthaul links to provide enhanced sensing and communication (S&C) services. In this kind of networks, RA is of paramount importance to S&C, as power and bandwidth allocation are almost related to almost all the performance metrics of S&C, such as the SE and EE in communications, as well as the range resolution and estimation accuracy in sensing [195]. Therefore, it is intuitive and essential to carry out the research of RA with ISAC in MDD-CF networks.

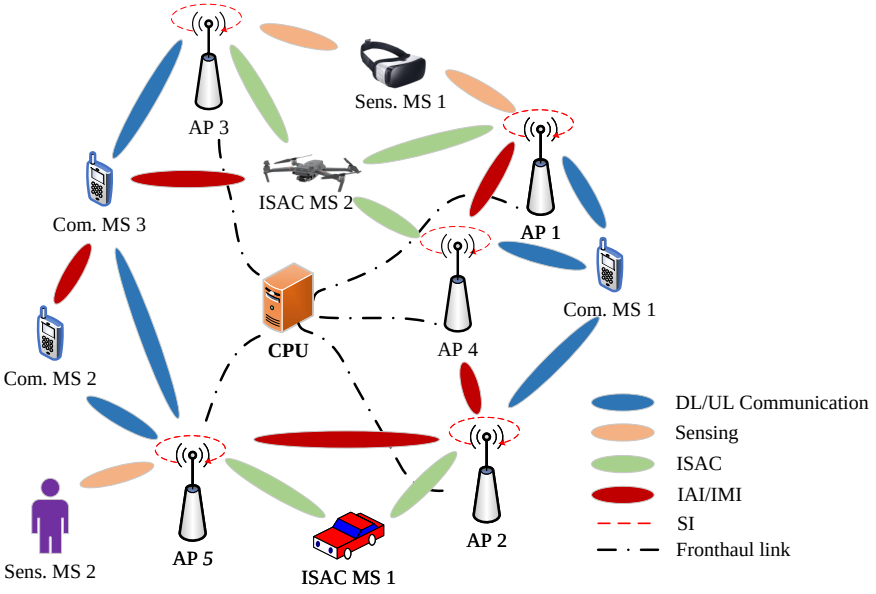


FIGURE 7.3: Illustration for the application of ISAC in MDD-CF networks

A possible scenario for researching the ISAC in MDD-CF networks is shown in Fig. 7.3. The APs operated in MDD mode serve three types of MSs, including the DL/UL communication MSs, sensing MSs and ISAC MSs. The DL/UL communication MSs, such as smartphones require stable and high-quality DL/UL communication services. The sensing MSs like the extended reality devices and humans, which can be either device-free or device-based, require the sensing services consisting of detection, localization and tracking. Finally, the ISAC MSs are mainly device-based, such as UAVs and vehicles relying on both S&C services, which is particularly the case when the ISAC MSs do not transmit UL signals to APs. Explicitly, each type of the above-mentioned MSs have different QoS requirements. For instance, the DL/UL communication MSs emphasize SE, while the sensing MSs demand on the localization accuracy. However, all the QoS requirements can only be met via allocating the corresponding transmitted power and bandwidth [196]. In order to maximize the performance of ISAC in MDD-CF networks and provide services tailored for different QoS requirements, the RA should be comprehensively investigated.

Appendix A

The Proof of Proposition 1

Given that $\xi = \left\lfloor \frac{\bar{M}}{D_{\text{sum}}} \right\rfloor \geq L$, and the UL subcarriers are uniformly distributed with spacing of $l = M_{\text{sum}}/\bar{M}$ number of subcarriers between two adjacent UL subcarriers, we need to prove that

$$\boldsymbol{P}[m]^H \boldsymbol{P}[n] = \begin{cases} \mathbf{0}_L, & \text{if } m \neq n, \\ \varrho \boldsymbol{I}_L, & \text{if } m = n, \end{cases} \tag{A.1}$$

where $\mathbf{0}_L$ is a $(L \times L)$ all-zero matrix, and $\varrho = \bar{M}/M_{\text{sum}}$ is constant.

Since the UL subcarriers are uniformly-spaced with a spacing of l number of subcarriers between two adjacent UL subcarriers, it can be shown that the matrix $\boldsymbol{\Phi}_{\text{UL}}^T \boldsymbol{F} \boldsymbol{\Psi}$ is given by

$$\begin{aligned} \boldsymbol{G}_{\text{UL}} &= \boldsymbol{\Phi}_{\text{UL}}^T \boldsymbol{F} \boldsymbol{\Psi} \\ &= \frac{1}{M_{\text{sum}}} \begin{bmatrix} 1 & e^{-j2\pi \frac{(\bar{m}_1 - 1)}{M_{\text{sum}}}} & \cdots & e^{-j2\pi \frac{(\bar{m}_1 - 1)(L-1)}{M_{\text{sum}}}} \\ 1 & e^{-j2\pi \frac{(\bar{m}_2 - 1)}{M_{\text{sum}}}} & \cdots & e^{-j2\pi \frac{(\bar{m}_2 - 1)(L-1)}{M_{\text{sum}}}} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & e^{-j2\pi \frac{(\bar{m}_{\bar{M}} - 1)}{M_{\text{sum}}}} & \cdots & e^{-j2\pi \frac{(\bar{m}_{\bar{M}} - 1)(L-1)}{M_{\text{sum}}}} \end{bmatrix}_{\bar{M} \times L}, \end{aligned} \tag{A.2}$$

where $\bar{m}_1, \bar{m}_2, \cdots, \bar{m}_{\bar{M}}$ are the indices of UL subcarriers. Note that, to obtain (A.2), the relationships of $\bar{m}_2 - \bar{m}_1 = ... = \bar{m}_{\bar{M}} - \bar{m}_{\bar{M}-1} = l$ are used. Let

$$\tilde{\boldsymbol{P}} = \boldsymbol{P}[m]^H \boldsymbol{P}[n] = \left(\boldsymbol{X}_m \boldsymbol{G}_{\text{UL}}\right)^H \left(\boldsymbol{X}_n \boldsymbol{G}_{\text{UL}}\right), \; 1 \leq m, n \leq D_{\text{sum}}. \tag{A.3}$$

Then, upon applying (2.31) and with the aid of (A.2), we can show that the (u, v)-th element of $\tilde{\boldsymbol{P}}$ is

$$\begin{aligned}
\left(\tilde{\boldsymbol{P}}\right)_{u,v} &= \frac{1}{M_{\text{sum}}} \sum_{t=1}^{\bar{M}} (\boldsymbol{X}_m)^*_{t,t} (\boldsymbol{X}_n)_{t,t} e^{\frac{j2\pi(\bar{m}_t-1)(u-v)}{M_{\text{sum}}}} \\
&= \frac{1}{M_{\text{sum}}} \sum_{t=1}^{\bar{M}} e^{\frac{j2\pi(t-1)(n-m)\xi}{\bar{M}}} e^{\frac{j2\pi(\bar{m}_t-1)(u-v)}{M_{\text{sum}}}} \\
&\overset{(a)}{=} \frac{1}{M_{\text{sum}}} \sum_{t=1}^{\bar{M}} e^{\frac{j2\pi(t-1)(n-m)\xi}{\bar{M}}} e^{\frac{j2\pi(t-1)(u-v)}{\bar{M}}} \\
&= \frac{1}{M_{\text{sum}}} \sum_{t=1}^{\bar{M}} e^{\frac{j2\pi(t-1)[(n-m)\xi+u-v]}{\bar{M}}} \\
&= \frac{\bar{M}}{M_{\text{sum}}} \delta[\Delta_{uv} - (m-n)\xi],
\end{aligned} \tag{A.4}$$

where $\Delta_{uv} = u - v$. In (A.4), and Equation (a) is because of $l = M_{\text{sum}}/\bar{M}$.

Explicitly, $\left(\tilde{\boldsymbol{P}}\right)_{u,v} = \bar{M}/M_{\text{sum}}$ provided that $\Delta_{uv} - (m-n)\xi = 0$. Hence, when $m = n$, we have $\left(\tilde{\boldsymbol{P}}\right)_{u,v} = \bar{M}/M_{\text{sum}}$ and $\left(\tilde{\boldsymbol{P}}\right)_{u,v} = 0$ for all $u \neq v$. Therefore, $\boldsymbol{P}[m]^H \boldsymbol{P}[m] = \varrho \boldsymbol{I}_L$ with $\varrho = \bar{M}/M_{\text{sum}}$.

By contrast, in the case of $m \neq n$, since $\tilde{\boldsymbol{P}}$ is a $(L \times L)$ matrix, we have $-(L-1) \leq \Delta_{uv} \leq L-1$. Hence, $\Delta_{uv} - (m-n)\xi \neq 0$, provided that $\xi \geq L$. Therefore, we have $\boldsymbol{P}[m]^H \boldsymbol{P}[n] = \boldsymbol{0}_L$ for all $m \neq n$, provided that $\xi \geq L$.

In summary, when given $\xi \geq L$ and that the UL subcarriers are uniformly arranged with a constant spacing of l between two adjacent UL subcarriers, the set of PSs given by (2.31) satisfy (A.1), i.e., they are the orthogonality-achieving PSs.

Appendix B

Conjugate Gradient Computation

According to (3.23), we have

$$\begin{aligned}\Phi\left(\boldsymbol{F}_{\mathrm{RF}}\right) &= \sum_{m\in\mathcal{M}_{\mathrm{DL}}}\sum_{d\in\mathcal{D}_m}\left\|\boldsymbol{f}^{d}_{\mathrm{MMSE}}[m]-\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}^{\tilde{d}}_{\mathrm{BB}}[m]\right\|_2^2\\ &= \sum_{m\in\mathcal{M}_{\mathrm{DL}}}\sum_{d\in\mathcal{D}_m}\left\|\boldsymbol{f}^{d}_{\mathrm{MMSE}}[m]\right\|_2^2\\ &\quad-\mathrm{Tr}\left(\boldsymbol{f}^{d}_{\mathrm{MMSE}}[m]\boldsymbol{f}^{\tilde{d},H}_{\mathrm{BB}}[m]\boldsymbol{F}^{H}_{\mathrm{RF}}+\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}^{\tilde{d}}_{\mathrm{BB}}[m]\boldsymbol{f}^{d,H}_{\mathrm{MMSE}}[m]\right)+1.\end{aligned} \tag{B.1}$$

Then, to further simplify the second term in above equation based on (3.22), we have

$$\begin{aligned}\text{The second term} &= \frac{2\mathrm{Tr}\left(\boldsymbol{f}^{d}_{\mathrm{MMSE}}[m]\boldsymbol{f}^{d,H}_{\mathrm{MMSE}}[m]\boldsymbol{F}_{\mathrm{RF}}\left(\boldsymbol{F}^{H}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{RF}}\right)^{-1}\boldsymbol{F}^{H}_{\mathrm{RF}}\right)}{\left\|\boldsymbol{F}_{\mathrm{RF}}\left(\boldsymbol{F}^{H}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{RF}}\right)^{-1}\boldsymbol{F}^{H}_{\mathrm{RF}}\boldsymbol{f}^{d}_{\mathrm{MMSE}}[m]\right\|_2}\\ &= 2\sqrt{\mathrm{Tr}\left(\left(\boldsymbol{F}^{H}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{RF}}\right)^{-1}\boldsymbol{F}^{H}_{\mathrm{RF}}\boldsymbol{f}^{d}_{\mathrm{MMSE}}[m]\boldsymbol{f}^{d,H}_{\mathrm{MMSE}}[m]\boldsymbol{F}_{\mathrm{RF}}\right)}\\ &= 2\left\|\boldsymbol{b}_d[m]\right\|_2,\end{aligned} \tag{B.2}$$

where $\boldsymbol{b}_d[m] = \left(\boldsymbol{F}^{H}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{RF}}\right)^{-1/2}\boldsymbol{F}^{H}_{\mathrm{RF}}\boldsymbol{f}^{d}_{\mathrm{MMSE}}[m]$. Hence, the cost function can be transformed to

$$\Phi\left(\boldsymbol{F}_{\mathrm{RF}}\right) = \sum_{m\in\mathcal{M}_{\mathrm{DL}}}\sum_{d\in\mathcal{D}_m}\left\|\boldsymbol{f}^{d}_{\mathrm{MMSE}}[m]\right\|_2^2-2\left\|\boldsymbol{b}_d[m]\right\|_2+1. \tag{B.3}$$

Then, upon taking the derivative of the (B.3) with respected to $\boldsymbol{F}^{*}_{\mathrm{RF}}$, we obtain

$$\begin{aligned}\frac{\partial(\Phi)}{\partial\boldsymbol{F}^{*}_{\mathrm{RF}}} = \sum_{m\in\mathcal{M}_{\mathrm{DL}}}\sum_{d\in\mathcal{D}_m}\frac{1}{\left\|\boldsymbol{b}_d[m]\right\|_2}\Big(&\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}^{H}_{\mathrm{RF}}\boldsymbol{f}^{d}_{\mathrm{MMSE}}[m]\boldsymbol{f}^{d,H}_{\mathrm{MMSE}}[m]\boldsymbol{F}_{\mathrm{RF}}\left(\boldsymbol{F}^{H}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{RF}}\right)^{-2}-\\ &\boldsymbol{f}^{d}_{\mathrm{MMSE}}[m]\boldsymbol{f}^{d,H}_{\mathrm{MMSE}}[m]\boldsymbol{F}_{\mathrm{RF}}\left(\boldsymbol{F}^{H}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{RF}}\right)^{-1}\Big).\end{aligned} \tag{B.4}$$

Appendix C

Calculation of Self-Interference Power

C.1 Self-Interference in Equation (4.9)

According to the model for residual SI [148], the covariance of $z_d^{\mathrm{SI}}[i]$ is

$$\begin{aligned}\mathrm{cov}\left\{z_d^{\mathrm{SI}}[i]\right\} &= \mathbb{E}\left[h_{\mathrm{SI}} x_d[i] x_d^H[i] h_{\mathrm{SI}}^H\right] \\ &= \xi_{\mathrm{MT}} p_{\mathrm{UL}} \bar{M},\end{aligned} \tag{C.1}$$

when using $\mathbb{E}\left[h_{\mathrm{SI}} h_{\mathrm{SI}}^H\right] = 1$. (C.1) is obtained by substituting $x_d[i] = \sqrt{p_{\mathrm{UL}}} \sum_{\bar{m}=1}^{\bar{M}} x_d[i, \bar{m}]$.

C.2 Self-Interference in Equation (4.10)

To compute the covariance of $z^{\mathrm{SI}}[i]$ in (4.10), we again refer to the residual model in [148] and have

$$\mathrm{cov}\left\{\boldsymbol{z}^{\mathrm{SI}}[i]\right\} = \xi_{\mathrm{BS}} \mathrm{diag}\left(\mathbb{E}\left[\boldsymbol{H}_{\mathrm{SI}} \boldsymbol{s}_{\mathrm{DL}}[i] \boldsymbol{s}_{\mathrm{DL}}^H[i] \boldsymbol{H}_{\mathrm{SI}}^H\right]\right). \tag{C.2}$$

Upon substituting $\boldsymbol{s}_{\mathrm{DL}}[i] = \sum_{m=1}^{M} \sqrt{p_{\mathrm{DL}}} \boldsymbol{F}^{\mathrm{ZF}}[i,m]\boldsymbol{x}[i,m]$ into (C.2), we can obtain the SI power at the n-th receive element as

$$\begin{aligned}
\mathrm{cov}\left\{\boldsymbol{z}^{\mathrm{SI}}[i]\right\}_{n,n} &= \xi_{\mathrm{BS}}\mathbb{E}\left[\boldsymbol{H}_{\mathrm{SI}}^{(n,:)}\boldsymbol{s}_{\mathrm{DL}}[i]\boldsymbol{s}_{\mathrm{DL}}^{H}[i]\boldsymbol{H}_{\mathrm{SI}}^{(n,:),H}\right] \\
&= \xi_{\mathrm{BS}}\mathbb{E}\left[\mathrm{Tr}\left(p_{\mathrm{DL}}\sum_{m=1}^{M}\boldsymbol{H}_{\mathrm{SI}}^{(n,:)}\boldsymbol{F}^{\mathrm{ZF}}[i,m]\boldsymbol{F}^{\mathrm{ZF},H}[i,m]\boldsymbol{H}_{\mathrm{SI}}^{(n,:),H}\right)\right] \\
&\overset{(a)}{=} \xi_{\mathrm{BS}}p_{\mathrm{DL}}\mathbb{E}\left[\sum_{m=1}^{M}\left\|\boldsymbol{F}^{\mathrm{ZF}}[i,m]\right\|_{F}^{2}\right] \\
&\overset{(b)}{=} \xi_{\mathrm{BS}}p_{\mathrm{DL}}M, \qquad \text{(C.3)}
\end{aligned}$$

where (a) is obtained using $\mathbb{E}[\boldsymbol{H}_{\mathrm{SI}}^{(n,:),H}\boldsymbol{H}_{\mathrm{SI}}^{(n,:)}] = \boldsymbol{I}_N$ and (b) is due to the power normalization $\left\|\boldsymbol{F}^{\mathrm{ZF}}[i,m]\right\|_F^2 = 1$. Consequently, $\mathrm{cov}\left\{\boldsymbol{z}^{\mathrm{SI}}[i]\right\} = \xi_{\mathrm{BS}}p_{\mathrm{DL}}M\boldsymbol{I}_N$.

Appendix D

Covariance matrix of predicted UL subcarrier channels in (4.31)

The covariance matrix $\mathbf{\Gamma}_{d,i}$ in (4.31) can be written as

$$\begin{aligned}\mathbf{\Gamma}_{d,i} &= \mathbb{E}\left[\check{\boldsymbol{h}}_{n,d}^{\mathrm{UL}}[i+1]\check{\boldsymbol{h}}_{n,d}^{\mathrm{UL},H}[i+1]\right] \\ &= \begin{bmatrix} \mathbb{E}\left[\check{h}_{n,d}^{\mathrm{DD}}[i+1,1]\check{h}_{n,d}^{\mathrm{DD},H}[i+1,1]\right] & \cdots & \mathbb{E}\left[\check{h}_{n,d}^{\mathrm{DD}}[i+1,1]\check{h}_{n,d}^{\mathrm{DD},H}[i+1,\bar{M}]\right] \\ \vdots & \ddots & \vdots \\ \mathbb{E}\left[\check{h}_{n,d}^{\mathrm{DD}}[i+1,\bar{M}]\check{h}_{n,d}^{\mathrm{DD},H}[i+1,1]\right] & \cdots & \mathbb{E}\left[\check{h}_{n,d}^{\mathrm{DD}}[i+1,\bar{M}]\check{h}_{n,d}^{\mathrm{DD},H}[i+1,\bar{M}]\right] \end{bmatrix}.\end{aligned} \tag{D.1}$$

As for the diagonal elements in $\mathbf{\Gamma}_{d,i}$, according to (4.30), we have

$$h_{n,d}[i+1,\bar{m}] = \check{h}_{n,d}^{\mathrm{DD}}[i+1,\bar{m}] + \check{e}_{n,d}^{\mathrm{DD}}[i+1,\bar{m}]. \tag{D.2}$$

Hence, the variance of $\check{h}_{n,d}^{\mathrm{DD}}[i+1,\bar{m}]$ is $(\mathbf{\Theta}_i[\bar{m}])_{d,d}$, given in (4.30). For the off-diagonal elements of $\mathbf{\Gamma}_{d,i}$, we have

$$\begin{aligned}&\mathbb{E}\left[\check{h}_{n,d}^{\mathrm{DD}}[i+1,\bar{m}_1]\check{h}_{n,d}^{\mathrm{DD},H}[i+1,\bar{m}_2]\right] \\ &= \mathbb{E}\left[(h_{n,d}[i+1,\bar{m}_1] - \check{e}_{n,d}^{\mathrm{DD}}[i+1,\bar{m}_1])(h_{n,d}^{H}[i+1,\bar{m}_2] - \check{e}_{n,d}^{\mathrm{DD},H}[i+1,\bar{m}_2])\right] \\ &\overset{(a)}{=} \mathbb{E}\left[h_{n,d}[i+1,\bar{m}_1]h_{n,d}^{H}[i+1,\bar{m}_2]\right] \\ &\overset{(b)}{=} \mathbb{E}\left[\boldsymbol{\psi}_{\bar{m}_1}\boldsymbol{g}_{n,d}[i+1]\boldsymbol{g}_{n,d}^{H}[i+1]\boldsymbol{\psi}_{\bar{m}_2}^{H}\right] \\ &= \frac{\beta_d}{L}\boldsymbol{\psi}_{\bar{m}_1}\boldsymbol{\psi}_{\bar{m}_2}^{H},\ \bar{m}_1 \neq \bar{m}_2,\end{aligned} \tag{D.3}$$

where (a) is obtained based on the fact that the prediction error of the $\bar{m}_1$-th subcarrier channel is independent of the $\bar{m}_2$-th subcarrier channel and vice versa, while (b) can

be derived from the FFT operation. In summary, the covariance matrix $\mathbf{\Gamma}_{d,i}$ can be expressed as

$$\mathbf{\Gamma}_{d,i} = \begin{bmatrix} (\mathbf{\Theta}_i[1])_{d,d} & \cdots & \frac{\beta_d}{L}\boldsymbol{\psi}_1\boldsymbol{\psi}_{\bar{M}}^H \\ \vdots & \ddots & \vdots \\ \frac{\beta_d}{L}\boldsymbol{\psi}_{\bar{M}}\boldsymbol{\psi}_1^H & \cdots & (\mathbf{\Theta}_i[\bar{M}])_{d,d} \end{bmatrix}. \tag{D.4}$$

Appendix E

The simplification of $\mathrm{SINR}_{d,m}$ and $\mathrm{SINR}_{d,\bar{m}}$ in (5.9)

For $\mathrm{SINR}_{d,m}$, since the ZF precoder is employed with the constraint of power normalization, the numerator of (5.4) can be transformed to $\left|\sum_{l\in\mathcal{L}} \lambda_{ld}\mu_{ldm}\sqrt{p_{ldm}}\omega_{ldm}\right|^2$, where $\omega_{ldm} = \frac{1}{\left\|\boldsymbol{f}_{ld}^{\mathrm{ZF}}[m]\right\|_2}$. In the denominator, $\mathrm{MUI}_{d,m} \approx 0$ due to the ZF precoding, when $N \geq D$. Furthermore, $\mathrm{var}\left\{z_d^{\mathrm{SI}}\right\}$ and $\mathrm{var}\left\{z_d^{\mathrm{IMI}}\right\}$ can be obtained as follows

$$\begin{aligned}
\mathrm{var}\left\{z_d^{\mathrm{SI}}\right\} &= \mathbb{E}\left[\bar{z}_d^{\mathrm{SI}}\left(\bar{z}_d^{\mathrm{SI}}\right)^H\right] \overset{(a)}{=} \xi_d^{\mathrm{SI}} \sum_{\bar{m}\in\bar{\mathcal{M}}} \mu_{d\bar{m}} p_{d\bar{m}}, \\
\mathrm{var}\left\{z_d^{\mathrm{IMI}}\right\} &= \xi_d^{\mathrm{IMI}}\mathbb{E}\left[\bar{z}_d^{\mathrm{IMI}}\left(\bar{z}_d^{\mathrm{IMI}}\right)^H\right] \overset{(b)}{=} \xi_d^{\mathrm{IMI}} \sum_{d'\in\mathcal{D}\setminus\{d\}} \sum_{\bar{m}\in\bar{\mathcal{M}}} \frac{\beta_{dd'}}{M_{\mathrm{sum}}} \mu_{d'\bar{m}} p_{d'\bar{m}},
\end{aligned} \tag{E.1}$$

where (a) is derived using $\mathbb{E}\left[h_{dd}h_{dd}^H\right] = \xi_d^{\mathrm{SI}}$ according to (5.1), and (b) is obtained using $\mathbb{E}\left[h_{dd'}[\bar{m}]h_{dd'}^H[\bar{m}]\right] = \mathbb{E}\left[(\boldsymbol{\phi}_{\mathrm{UL}}^T\boldsymbol{F}\boldsymbol{\Psi}\boldsymbol{g}_{dd'})(\boldsymbol{\phi}_{\mathrm{UL}}^T\boldsymbol{F}\boldsymbol{\Psi}\boldsymbol{g}_{dd'})^H\right] = \frac{\beta_{dd'}}{M_{\mathrm{sum}}}$.

For the $\mathrm{SINR}_{d,\bar{m}}$, since $(\boldsymbol{w}_{ld}^{\mathrm{ZF}}[\bar{m}])^H\boldsymbol{h}_{ld}^{\mathrm{ZF}}[\bar{m}] \approx 1$, the term in the numerator of (5.8) can be changed to $\mu_{d\bar{m}}p_{d\bar{m}}L^2$. In the denominator, $\mathrm{MUI}_{d,\bar{m}} \approx 0$, while the second term can be obtained as

$$\begin{aligned}
&\sum_{l\in\mathcal{L}} \mathbb{E}\left[\left\|\left(\boldsymbol{w}_{ld}^{\mathrm{ZF}}[\bar{m}]\right)^H \boldsymbol{z}_l^{\mathrm{SI}}\right\|^2\right] = \sum_{l\in\mathcal{L}} \mathrm{Tr}\left[\left(\boldsymbol{w}_{ld}^{\mathrm{ZF}}[\bar{m}]\right)^H \mathrm{diag}\left(\mathrm{cov}\left\{\bar{\boldsymbol{z}}_l^{\mathrm{SI}}\right\}\right)\boldsymbol{w}_{ld}^{\mathrm{ZF}}[\bar{m}]\right] \\
&\overset{(a)}{=} \xi_l^{\mathrm{SI}} \sum_{l\in\mathcal{L}} \left\|\boldsymbol{w}_{ld}^{\mathrm{ZF}}[\bar{m}]\right\|_2^2 \sum_{m\in\mathcal{M}} \sum_{d\in\mathcal{D}} \lambda_{ld}\mu_{ldm}p_{ldm},
\end{aligned} \tag{E.2}$$

where (a) holds since

$$\begin{aligned}\left(\operatorname{cov}\left\{\bar{\boldsymbol{z}}_l^{\mathrm{SI}}\right\}\right)_{i,i} &= \sum_{m\in\mathcal{M}}\sum_{d\in\mathcal{D}} \lambda_{ld}\mu_{ldm}p_{ldm}\mathbb{E}\left[\left|\boldsymbol{H}_{ll}^{(i,:)}\boldsymbol{f}_{ld}^{\mathrm{ZF}}[m]x_d[m]\right|^2\right] \\ &\overset{(b)}{=} \sum_{m\in\mathcal{M}}\sum_{d\in\mathcal{D}} \lambda_{ld}\mu_{ldm}p_{ldm}\mathrm{Tr}\left\{\boldsymbol{f}_{ld}^{\mathrm{ZF}}[m]\left(\boldsymbol{f}_{ld}^{\mathrm{ZF}}[m]\right)^H\mathbb{E}\left[\left(\boldsymbol{H}_{ll}^{(i,:)}\right)^H\boldsymbol{H}_{ll}^{(i,:)}\right]\right\} \\ &\overset{(c)}{=} \xi_l^{\mathrm{SI}}\sum_{m\in\mathcal{M}}\sum_{d\in\mathcal{D}} \lambda_{ld}\mu_{ldm}p_{ldm}, \end{aligned} \tag{E.3}$$

where we have (b) due to that the ZF precoder only focuses on the desired signal, and hence the ZF precoding vector is uncorrelated with the SI channel, we have(c) according to the assumption that $\left\|\boldsymbol{f}_{ld}^{\mathrm{ZF}}[m]\right\|_2^2 = 1$ and $\mathbb{E}\left[\left(\boldsymbol{H}_{ll}^{(i,:)}\right)^H\boldsymbol{H}_{ll}^{(i,:)}\right] = \xi_l^{\mathrm{SI}}\boldsymbol{I}_N$.

Similarly, the third term in the denominator of (5.8) can be obtained as

$$\sum_{l\in\mathcal{L}}\mathbb{E}\left[\left\|\boldsymbol{w}_{ld}^H[\bar{m}]\boldsymbol{z}_l^{\mathrm{IAI}}\right\|^2\right] = \xi_l^{\mathrm{IAI}}\sum_{l\in\mathcal{L}} v_{ld\bar{m}}\sum_{l'\in\mathcal{L}\setminus\{l\}}\sum_{m\in\mathcal{M}}\sum_{d\in\mathcal{D}}\frac{\beta_{ll'}}{M_{\mathrm{sum}}}\lambda_{l'd}\mu_{l'dm}p_{l'dm}, \tag{E.4}$$

where $v_{ld\bar{m}} = \left\|\boldsymbol{w}_{ld}^{\mathrm{ZF}}[\bar{m}]\right\|_2^2$. Consequently, $\mathrm{SINR}_{d,m}$ and $\mathrm{SINR}_{d,\bar{m}}$ can be simplified to the desired form as shown in (5.9).

References

[1] RootMetrics, "The state of 5G in the UK - 2H 2021," Tech. Rep., 2021.

[2] OFCOM, "OFCOM technology tracker 2021," Tech. Rep., 2021.

[3] R. Li, Y. Chen, G. Y. Li, and G. Liu, "Full-duplex cellular networks," *IEEE Communications Magazine*, vol. 55, no. 4, pp. 184–191, 2017.

[4] R. Rajashekar, C. Xu, N. Ishikawa, L.-L. Yang, and L. Hanzo, "Multicarrier division duplex aided millimeter wave communications," *IEEE Access*, vol. 7, pp. 100 719–100 732, 2019.

[5] V. K. Garg, *Wireless network evolution: 2G to 3G*. Prentice Hall PTR, 2001.

[6] B. Li, D. Xie, S. Cheng, J. Chen, P. Zhang, W. Zhu, and B. Li, "Recent advances on TD-SCDMA in China," *IEEE Communications Magazine*, vol. 43, no. 1, pp. 30–37, 2005.

[7] P. W. Chan, E. S. Lo, R. R. Wang, E. K. Au, V. K. Lau, R. S. Cheng, W. H. Mow, R. D. Murch, and K. B. Letaief, "The evolution path of 4G networks: FDD or TDD?" *IEEE Communications Magazine*, vol. 44, no. 12, pp. 42–50, 2006.

[8] P. Zhang, X. Tao, J. Zhang, Y. Wang, L. Li, and Y. Wang, "A vision from the future: beyond 3G TDD," *IEEE Communications Magazine*, vol. 43, no. 1, pp. 38–44, 2005.

[9] J. Flordelis, F. Rusek, F. Tufvesson, E. G. Larsson, and O. Edfors, "Massive MIMO performance—TDD versus FDD: What do measurements say?" *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2247–2261, 2018.

[10] L. Sanguinetti, E. Björnson, and J. Hoydis, "Toward massive MIMO 2.0: Understanding spatial correlation, interference suppression, and pilot contamination," *IEEE Transactions on Communications*, vol. 68, no. 1, pp. 232–257, 2019.

[11] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE transactions on wireless communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

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

[13] A. Adhikary, E. Al Safadi, M. K. Samimi, R. Wang, G. Caire, T. S. Rappaport, and A. F. Molisch, "Joint spatial division and multiplexing for mm-wave channels," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1239–1255, 2014.

[14] A. K. Papazafeiropoulos and T. Ratnarajah, "Deterministic equivalent performance analysis of time-varying massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 10, pp. 5795–5809, 2015.

[15] S. Kashyap, C. Mollén, E. Björnson, and E. G. Larsson, "Performance analysis of (TDD) massive MIMO with kalman channel prediction," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 3554–3558.

[16] 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.

[17] D. Kim, H. Lee, and D. Hong, "A survey of in-band full-duplex transmission: From the perspective of PHY and MAC layers," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2017–2046, 2015.

[18] M. B. Steer, *Microwave and RF design: a systems approach*. SciTech Pub., 2010.

[19] H. Ju, E. Oh, and D. Hong, "Improving efficiency of resource usage in two-hop full duplex relay systems based on resource sharing and interference cancellation," *IEEE Transactions on Wireless Communications*, vol. 8, no. 8, pp. 3933–3938, 2009.

[20] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 680–694, 2014.

[21] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *ACM SIGCOMM computer communication review*, vol. 43, no. 4. ACM, 2013, pp. 375–386.

[22] J. A. Wepman, "Analog-to-digital converters and their applications in radio receivers," *IEEE Communications magazine*, vol. 33, no. 5, pp. 39–45, 1995.

[23] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on selected areas in communications*, vol. 32, no. 9, pp. 1637–1652, 2014.

[24] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers*. IEEE, 2010, pp. 1558–1562.

[25] B. Debaillie, D.-J. van den Broek, C. Lavin, B. van Liempd, E. A. Klumperink, C. Palacios, J. Craninckx, B. Nauta, and A. Pärssinen, "Analog/RF solutions enabling compact full-duplex radios," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1662–1673, 2014.

[26] S. Nikolaou, R. Bairavasubramanian, C. Lugo, I. Carrasquillo, D. C. Thompson, G. E. Ponchak, J. Papapolymerou, and M. M. Tentzeris, "Pattern and frequency reconfigurable annular slot antenna using pin diodes," *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 2, pp. 439–448, 2006.

[27] A. K. Khandani, "Two-way (true full-duplex) wireless," in *2013 13th Canadian Workshop on Information Theory*. IEEE, 2013, pp. 33–38.

[28] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," in *Proceedings of the sixteenth annual international conference on Mobile computing and networking*. ACM, 2010, pp. 1–12.

[29] M. A. Khojastepour, K. Sundaresan, S. Rangarajan, X. Zhang, and S. Barghi, "The case for antenna cancellation for scalable full-duplex wireless communications," in *Proceedings of the 10th ACM Workshop on Hot Topics in Networks*. ACM, 2011, p. 17.

[30] E. Everett, C. Shepard, L. Zhong, and A. Sabharwal, "Softnull: Many-antenna full-duplex wireless via digital beamforming," *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 8077–8092, 2016.

[31] N. M. Gowda and A. Sabharwal, "Jointnull: Combining partial analog cancellation with transmit beamforming for large-antenna full-duplex wireless systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 2094–2108, 2018.

[32] X. Xia, K. Xu, D. Zhang, Y. Xu, and Y. Wang, "Beam-domain full-duplex massive MIMO: realizing co-time co-frequency uplink and downlink transmission in the cellular system," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 10, pp. 8845–8862, 2017.

[33] B. Radunovic, D. Gunawardena, P. Key, A. Proutiere, N. Singh, V. Balan, and G. Dejean, "Rethinking indoor wireless mesh design: Low power, low frequency, full-duplex," in *2010 Fifth IEEE Workshop on Wireless Mesh Networks*. IEEE, 2010, pp. 1–6.

[34] Y. Hua, P. Liang, Y. Ma, A. C. Cirik, and Q. Gao, "A method for broadband full-duplex MIMO radio," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 793–796, 2012.

[35] J.-H. Lee, "Self-interference cancelation using phase rotation in full-duplex wireless," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4421–4429, 2013.

[36] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proceedings of the 17th annual international conference on Mobile computing and networking*. ACM, 2011, pp. 301–312.

[37] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4296–4307, 2012.

[38] X. Quan, Y. Liu, W. Pan, Y. Tang, and K. Kang, "A two-stage analog cancellation architecture for self-interference suppression in full-duplex communications," in *2017 IEEE MTT-S International Microwave Symposium (IMS)*. IEEE, 2017, pp. 1169–1172.

[39] K. E. Kolodziej, J. G. McMichael, and B. T. Perry, "Multitap RF canceller for in-band full-duplex wireless communications," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 4321–4334, 2016.

[40] D. Liu, Y. Shen, S. Shao, Y. Tang, and Y. Gong, "On the analog self-interference cancellation for full-duplex communications with imperfect channel state information," *IEEE Access*, vol. 5, pp. 9277–9290, 2017.

[41] H. Lu, S. Shao, K. Deng, and Y. Tang, "Self-mixed self-interference analog cancellation in full-duplex communications," *Science China Information Sciences*, vol. 59, no. 4, p. 042303, 2016.

[42] M. Yilan, O. Gurbuz, and H. Ozkan, "Integrated linear and nonlinear digital cancellation for full duplex communication," *IEEE Wireless Communications*, vol. 28, no. 1, pp. 20–27, 2021.

[43] E. Ahmed and A. M. Eltawil, "All-digital self-interference cancellation technique for full-duplex systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3519–3532, 2015.

[44] D. Korpi, L. Anttila, V. Syrjälä, and M. Valkama, "Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1674–1687, 2014.

[45] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *arXiv preprint arXiv:1304.6690*, 2013.

[46] C.-K. Wen, S. Jin, K.-K. Wong, J.-C. Chen, and P. Ting, "Channel estimation for massive MIMO using Gaussian-mixture Bayesian learning," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1356–1368, 2015.

[47] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, 2014.

[48] T. Schenk, *RF imperfections in high-rate wireless systems: impact and digital compensation*. Springer Science & Business Media, 2008.

[49] E. Ahmed and A. M. Eltawil, "On phase noise suppression in full-duplex systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1237–1251, 2014.

[50] X. Quan, Y. Liu, S. Shao, C. Huang, and Y. Tang, "Impacts of phase noise on digital self-interference cancellation in full-duplex communications," *IEEE Transactions on Signal Processing*, vol. 65, no. 7, pp. 1881–1893, 2017.

[51] L. Anttila, D. Korpi, V. Syrjälä, and M. Valkama, "Cancellation of power amplifier induced nonlinear self-interference in full-duplex transceivers," in *2013 Asilomar conference on signals, systems and computers*. IEEE, 2013, pp. 1193–1198.

[52] D. Korpi, Y.-S. Choi, T. Huusari, L. Anttila, S. Talwar, and M. Valkama, "Adaptive nonlinear digital self-interference cancellation for mobile inband full-duplex radio: Algorithms and rf measurements," in *2015 IEEE global communications conference (GLOBECOM)*. IEEE, 2015, pp. 1–7.

[53] M. S. Sim, M. Chung, D. Kim, J. Chung, D. K. Kim, and C.-B. Chae, "Nonlinear self-interference cancellation for full-duplex radios: From link-level and system-level performance perspectives," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 158–167, 2017.

[54] D. W. Bliss, T. Hancock, and P. Schniter, "Hardware phenomenological effects on cochannel full-duplex MIMO relay performance," in *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE, 2012, pp. 34–39.

[55] E. Ahmed, A. M. Eltawil, and A. Sabharwal, "Rate gain region and design tradeoffs for full-duplex wireless communications," *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3556–3565, 2013.

[56] A. Sahai, G. Patel, C. Dick, and A. Sabharwal, "On the impact of phase noise on active cancelation in wireless full-duplex," *IEEE transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4494–4510, 2013.

[57] B. Razavi, *Design of analog CMOS integrated circuits*. Tata McGraw-Hill Education, 2002.

[58] E. Ahmed, A. M. Eltawil, and A. Sabharwal, "Self-interference cancellation with nonlinear distortion suppression for full-duplex systems," in *2013 Asilomar Conference on Signals, Systems and Computers*. IEEE, 2013, pp. 1199–1203.

[59] Y. Huang, B. Jalaian, S. Russell, and H. Samani, "Reaping the benefits of dynamic TDD in massive MIMO," *IEEE Systems Journal*, vol. 13, no. 1, pp. 117–124, 2018.

[60] E. de Olivindo Cavalcante, G. Fodor, Y. C. Silva, and W. C. Freitas, "Distributed beamforming in dynamic TDD MIMO networks with BS to BS interference constraints," *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 788–791, 2018.

[61] C. Na, X. Hou, and H. Jiang, "Interference alignment based dynamic TDD for small cells," in *2014 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2014, pp. 700–705.

[62] "Cross-link interference management based on coordinated beamforming," in *R1-1701674, TSG RAN WG1 Meeting #88*. 3GPP, 2017.

[63] H. Lee and D.-H. Cho, "Combination of dynamic-TDD and static-TDD based on adaptive power control," in *2008 IEEE 68th Vehicular Technology Conference*. IEEE, 2008, pp. 1–5.

[64] M. Ding, D. L. Pérez, A. V. Vasilakos, and W. Chen, "Dynamic TDD transmissions in homogeneous small cell networks," in *2014 IEEE International Conference on Communications Workshops (ICC)*. IEEE, 2014, pp. 616–621.

[65] H. Kim, K. Lee, H. Wang, and D. Hong, "Cross link interference mitigation schemes in dynamic TDD systems," in *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*. IEEE, 2019, pp. 1–5.

[66] Z. Huo, N. Ma, and B. Liu, "Joint user scheduling and transceiver design for cross-link interference suppression in MU-MIMO dynamic TDD systems," in *2017 3rd IEEE International Conference on Computer and Communications (ICCC)*. IEEE, 2017, pp. 962–967.

[67] J. M. B. da Silva, G. Wikström, R. K. Mungara, and C. Fischione, "Full duplex and dynamic TDD: Pushing the limits of spectrum reuse in multi-cell communications," *IEEE Wireless Communications*, vol. 28, no. 1, pp. 44–50, 2021.

[68] L.-L. Yang, *Multicarrier communications*. John Wiley & Sons, 2009.

[69] B. Li, L.-L. Yang, R. G. Maunder, and S. Sun, "Self-interference cancellation and channel estimation in multicarrier-division duplex systems with hybrid beamforming," *IEEE Access*, vol. 8, pp. 160 653–160 669, 2020.

[70] Z. Xiao, L. Zhu, Z. Gao, D. O. Wu, and X.-G. Xia, "User fairness non-orthogonal multiple access (NOMA) for millimeter-wave communications with analog beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 7, pp. 3411–3423, 2019.

[71] T. Kim, J. Kim, J. G. Andrews, and T. S. Rappaport, "Multi-code multicarrier cdma: Performance analysis," in *2004 IEEE International Conference on Communications*, 2004, pp. 973–977.

[72] K. E. Kolodziej, B. T. Perry, and J. S. Herd, "In-band full-duplex technology: Techniques and systems survey," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 7, pp. 3025–3041, 2019.

[73] B. Li, L.-L. Yang, R. G. Maunder, P. Xiao, and S. Sun, "Multicarrier-division duplex: A duplexing technique for the shift to 6G wireless communications," *IEEE Vehicular Technology Magazine*, vol. 16, no. 4, pp. 57–67, 2021.

[74] K. T. Truong and R. W. Heath, "Effects of channel aging in massive MIMO systems," *Journal of Communications and Networks*, vol. 15, no. 4, pp. 338–351, 2013.

[75] Ö. T. Demir, E. Björnson, and L. Sanguinetti, "Foundations of user-centric cell-free massive MIMO," *arXiv preprint arXiv:2108.02541*, 2021.

[76] H. Yang and T. L. Marzetta, "Capacity performance of multicell large-scale antenna systems," in *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2013, pp. 668–675.

[77] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO: Uniformly great service for everyone," in *2015 IEEE 16th international workshop on signal processing advances in wireless communications (SPAWC)*. IEEE, 2015, pp. 201–205.

[78] 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.

[79] 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, 2019.

[80] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4247–4261, 2020.

[81] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4445–4459, 2017.

[82] Z. H. Shaik, E. Björnson, and E. G. Larsson, "Mmse-optimal sequential processing for cell-free massive MIMO with radio stripes," *IEEE Transactions on Communications*, vol. 69, no. 11, pp. 7775–7789, 2021.

[83] S. Buzzi, C. D'Andrea, A. Zappone, and C. D'Elia, "User-centric 5G cellular networks: Resource allocation and comparison with the cell-free massive MIMO approach," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 1250–1264, 2019.

[84] G. Interdonato, P. Frenger, and E. G. Larsson, "Scalability aspects of cell-free massive MIMO," in *ICC 2019-2019 IEEE International Conference on Communications (ICC)*. IEEE, 2019, pp. 1–6.

[85] A. Á. Polegre, L. Sanguinetti, and A. G. Armada, "Pilot decontamination processing in cell-free massive MIMO," *IEEE Communications Letters*, vol. 25, no. 12, pp. 3990–3994, 2021.

[86] H. V. Nguyen, V.-D. Nguyen, O. A. Dobre, S. K. Sharma, S. Chatzinotas, B. Ottersten, and O.-S. Shin, "On the spectral and energy efficiencies of full-duplex cell-free massive MIMO," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1698–1718, 2020.

[87] H. Kim, J. Kim, and D. Hong, "Dynamic TDD systems for 5G and beyond: A survey of cross-link interference mitigation," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2315–2348, 2020.

[88] X. Xia, P. Zhu, J. Li, H. Wu, D. Wang, Y. Xin, and X. You, "Joint user selection and transceiver design for cell-free with network-assisted full duplexing," *IEEE Transactions on Wireless Communications*, vol. 20, no. 12, pp. 7856–7870, 2021.

[89] D. Wang, M. Wang, P. Zhu, J. Li, J. Wang, and X. You, "Performance of network-assisted full-duplex for cell-free massive MIMO," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1464–1478, 2019.

[90] A. Chowdhury, R. Chopra, and C. R. Murthy, "Can dynamic TDD enabled half-duplex cell-free massive MIMO outperform full-duplex cellular massive MIMO?" *arXiv preprint arXiv:2110.09968*, 2021.

[91] S. Chakraborty, Ö. T. Demir, E. Björnson, and P. Giselsson, "Efficient downlink power allocation algorithms for cell-free massive MIMO systems," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 168–186, 2020.

[92] F. Guo, H. Lu, and Z. Gu, "Joint power and user grouping optimization in cell-free massive MIMO systems," *IEEE Transactions on Wireless Communications*, 2021.

[93] L. Du, L. Li, H. Q. Ngo, T. C. Mai, and M. Matthaiou, "Cell-free massive MIMO: Joint maximum-ratio and zero-forcing precoder with power control," *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 3741–3756, 2021.

[94] S. Datta, D. N. Amudala, E. Sharma, R. Budhiraja, and S. S. Panwar, "Full-duplex cell-free massive MIMO systems: Analysis and decentralized optimization," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 31–50, 2021.

[95] X. Xia, P. Zhu, J. Li, H. Wu, D. Wang, and Y. Xin, "Joint optimization of spectral efficiency for cell-free massive MIMO with network-assisted full duplexing," *Science China Information Sciences*, vol. 64, no. 8, pp. 1–16, 2021.

[96] S. Chakraborty, E. Björnson, and L. Sanguinetti, "Centralized and distributed power allocation for max-min fairness in cell-free massive MIMO," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2019, pp. 576–580.

[97] Y. Zhao, I. G. Niemegeers, and S. H. De Groot, "Power allocation in cell-free massive MIMO: A deep learning method," *IEEE Access*, vol. 8, pp. 87 185–87 200, 2020.

[98] C. D'Andrea, A. Zappone, S. Buzzi, and M. Debbah, "Uplink power control in cell-free massive MIMO via deep learning," in *2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*. IEEE, 2019, pp. 554–558.

[99] M. Bashar, A. Akbari, K. Cumanan, H. Q. Ngo, A. G. Burr, P. Xiao, M. Debbah, and J. Kittler, "Exploiting deep learning in limited-fronthaul cell-free massive MIMO uplink," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1678–1697, 2020.

[100] L. Luo, J. Zhang, S. Chen, X. Zhang, B. Ai, and D. W. K. Ng, "Downlink power control for cell-free massive MIMO with deep reinforcement learning," *IEEE Transactions on Vehicular Technology*, 2022.

[101] T. Riihonen, A. Balakrishnan, K. Haneda, S. Wyne, S. Werner, and R. Wichman, "Optimal eigenbeamforming for suppressing self-interference in full-duplex MIMO relays." in *CISS*, 2011, pp. 1–6.

[102] I. Atzeni and M. Kountouris, "Full-duplex MIMO small-cell networks with interference cancellation," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 8362–8376, 2017.

[103] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 5983–5993, 2011.

[104] P. Lioliou, M. Viberg, M. Coldrey, and F. Athley, "Self-interference suppression in full-duplex MIMO relays," in *2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers*. IEEE, 2010, pp. 658–662.

[105] T. Riihonen, S. Werner, and R. Wichman, "Residual self-interference in full-duplex MIMO relays after null-space projection and cancellation," in *Proc. 44th Annual Asilomar Conference on Signals, Systems, and Computers*, 2010.

[106] X. Liu, Z. Xiao, L. Bai, J. Choi, P. Xia, and X.-G. Xia, "Beamforming based full-duplex for millimeter-wave communication," *Sensors*, vol. 16, no. 7, p. 1130, 2016.

[107] S. Huberman and T. Le-Ngoc, "MIMO full-duplex precoding: A joint beamforming and self-interference cancellation structure," *IEEE Transactions on Wireless Communications*, vol. 14, no. 4, pp. 2205–2217, 2014.

[108] K. Satyanarayana, M. El-Hajjar, P.-H. Kuo, A. Mourad, and L. Hanzo, "Hybrid beamforming design for full-duplex millimeter wave communication," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1394–1404, 2018.

[109] X. Guo, J. Zhang, S. Chen, X. Mu, and L. Hanzo, "Two-stage time-domain pilot contamination elimination in large-scale multiple-antenna aided and TDD based OFDM systems," *IEEE Access*, vol. 5, pp. 8629–8641, 2016.

[110] A. Yaghjian, "An overview of near-field antenna measurements," *IEEE Transactions on antennas and propagation*, vol. 34, no. 1, pp. 30–45, 1986.

[111] A. V. Oppenheim, *Discrete-time signal processing*. Pearson Education India, 1999.

[112] S. Gokceli, T. Levanen, T. Riihonen, M. K. Renfors, and M. Valkama, "Frequency-selective PAPR reduction for OFDM," *IEEE Transactions on Vehicular Technology*, 2019.

[113] A. Tang and X. Wang, "A-duplex: Medium access control for efficient coexistence between full-duplex and half-duplex communications," *IEEE Transactions on Wireless Communications*, vol. 14, no. 10, pp. 5871–5885, 2015.

[114] K. C.-J. Lin, K.-C. Hsu, and H.-Y. Wei, "Inter-client interference cancellation for full-duplex networks with half-duplex clients," *IEEE/ACM Transactions on Networking*, vol. 27, no. 5, pp. 2150–2163, 2019.

[115] A. C. Cirik and Y. Hua, "Interference management in full-duplex cellular networks," in *Full-Duplex Communications for Future Wireless Networks*. Springer, 2020, pp. 137–163.

[116] J. Hoydis, S. Ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE Journal on selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, 2013.

[117] K. Guo, Y. Guo, G. Ascheid *et al.*, "Uplink power control with MMSE receiver in multi-cell MU-massive-MIMO systems," in *2014 IEEE International Conference on Communications (ICC)*. IEEE, 2014, pp. 5184–5190.

[118] C. Rusu, R. Mendez-Rial, N. González-Prelcic, and R. W. Heath, "Low complexity hybrid precoding strategies for millimeter wave communication systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 8380–8393, 2016.

[119] H. Ghauch, T. Kim, M. Bengtsson, and M. Skoglund, "Subspace estimation and decomposition for large millimeter-wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 528–542, 2016.

[120] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE transactions on wireless communications*, vol. 13, no. 3, pp. 1499–1513, 2014.

[121] V. N. Ha, D. H. Nguyen, and J.-F. Frigon, "Subchannel allocation and hybrid precoding in millimeter-wave OFDMA systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 5900–5914, 2018.

[122] J. P. González-Coma, J. Rodriguez-Fernandez, N. González-Prelcic, L. Castedo, and R. W. Heath, "Channel estimation and hybrid precoding for frequency selective multiuser mmwave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 2, pp. 353–367, 2018.

[123] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 501–513, 2016.

[124] M. Wu, C. Dick, J. R. Cavallaro, and C. Studer, "High-throughput data detection for massive MU-MIMO-OFDM using coordinate descent," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 63, no. 12, pp. 2357–2367, 2016.

[125] H. Wei, D. Wang, J. Wang, and X. You, "TDD reciprocity calibration for multi-user massive MIMO systems with iterative coordinate descent," *Science China information sciences*, vol. 59, no. 10, p. 102306, 2016.

[126] Z. Pi, "Optimal transmitter beamforming with per-antenna power constraints," in *2012 IEEE International Conference on Communications (ICC)*. IEEE, 2012, pp. 3779–3784.

[127] S. J. Wright, "Coordinate descent algorithms," *Mathematical Programming*, vol. 151, no. 1, pp. 3–34, 2015.

[128] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European transactions on telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[129] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE transactions on signal processing*, vol. 54, no. 3, pp. 884–893, 2006.

[130] Y. G. Li, J. H. Winters, and N. R. Sollenberger, "MIMO-OFDM for wireless communications: signal detection with enhanced channel estimation," *IEEE Transactions on communications*, vol. 50, no. 9, pp. 1471–1477, 2002.

[131] C. Shin, R. W. Heath, and E. J. Powers, "Blind channel estimation for MIMO-OFDM systems," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 2, pp. 670–685, 2007.

[132] J. Yue, K. J. Kim, J. D. Gibson, and R. A. Iltis, "Channel estimation and data detection for MIMO-OFDM systems," in *GLOBECOM'03. IEEE Global Telecommunications Conference (IEEE Cat. No. 03CH37489)*, vol. 2. IEEE, 2003, pp. 581–585.

[133] Y. Li *et al.*, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Transactions on wireless communications*, vol. 1, no. 1, pp. 67–75, 2002.

[134] L. ETSI, "Evolved universal terrestrial radio access (E-UTRA); base station (BS) radio transmission and reception (3GPP TS 36.104 version 14.3. 0 release 14)," *ETSI Online, Mar*, 2017.

[135] A. Sahai, G. Patel, and A. Sabharwal, "Pushing the limits of full-duplex: Design and real-time implementation," *arXiv preprint arXiv:1107.0607*, 2011.

[136] T. ETSI, "136 213 v11. 3.0 (3GPP TS 36.213 version 11.3. 0 release 11)," *LTE.,(Jul. 2013)*, vol. 178.

[137] L. L. Yang, "Multicarrier-division duplexing (MDD): A duplexing scheme whose time has come," *IEEE COMSOC MMTC E-Lett.*, vol. 10, no. 1, pp. 35–38, 2015.

[138] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave cellular wireless networks: Potentials and challenges," *arXiv preprint arXiv:1401.2560*, 2014.

[139] W. W. Ho and Y.-C. Liang, "Optimal resource allocation for multiuser MIMO-OFDM systems with user rate constraints," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 3, pp. 1190–1203, 2008.

[140] T. F. Maciel and A. Klein, "On the performance, complexity, and fairness of suboptimal resource allocation for multiuser MIMO-OFDMA systems," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 1, pp. 406–419, 2009.

[141] X. Wu, Z. Ma, X. Chen, F. Labeau, and S. Han, "Energy efficiency-aware joint resource allocation and power allocation in multi-user beamforming," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4824–4833, 2019.

[142] C. Zhao, Y. Cai, A. Liu, M. Zhao, and L. Hanzo, "Mobile edge computing meets mmwave communications: Joint beamforming and resource allocation for system delay minimization," *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2382–2396, 2020.

[143] I. Ahmed, H. Khammari, and A. Shahid, "Resource allocation for transmit hybrid beamforming in decoupled millimeter wave multiuser-MIMO downlink," *IEEE Access*, vol. 5, pp. 170–182, 2016.

[144] C. Nam, C. Joo, and S. Bahk, "Joint subcarrier assignment and power allocation in full-duplex OFDMA networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3108–3119, 2015.

[145] B. Di, S. Bayat, L. Song, Y. Li, and Z. Han, "Joint user pairing, subchannel, and power allocation in full-duplex multi-user OFDMA networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 8260–8272, 2016.

[146] D. W. K. Ng, Y. Wu, and R. Schober, "Power efficient resource allocation for full-duplex radio distributed antenna networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 2896–2911, 2016.

[147] M. T. Kabir and C. Masouros, "Delay-constrained beamforming and resource allocation in full duplex systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 3476–3480, 2020.

[148] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 8, pp. 1541–1553, 2012.

[149] J. Jin, Y. R. Zheng, W. Chen, and C. Xiao, "Hybrid precoding for millimeter wave MIMO systems: A matrix factorization approach," *IEEE Transactions on Wireless Communications*, vol. 17, no. 5, pp. 3327–3339, 2018.

[150] A. T. Le, X. Huang, Y. J. Guo *et al.*, "Beam-based analog self-interference cancellation in full-duplex MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2460–2471, 2020.

[151] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE Journal on selected areas in communications*, vol. 21, no. 2, pp. 171–178, 2003.

[152] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE communications magazine*, vol. 52, no. 2, pp. 186–195, 2014.

[153] K. E. Baddour and N. C. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1650–1662, 2005.

[154] J. Gao and H. Liu, "Decision-directed estimation of MIMO time-varying rayleigh fading channels," *IEEE Transactions on wireless communications*, vol. 4, no. 4, pp. 1412–1417, 2005.

[155] Z. Shen, J. G. Andrews, and B. L. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE transactions on wireless communications*, vol. 4, no. 6, pp. 2726–2737, 2005.

[156] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[157] G. Scutari, D. P. Palomar, and S. Barbarossa, "The MIMO iterative waterfilling algorithm," *IEEE Transactions on Signal Processing*, vol. 57, no. 5, pp. 1917–1935, 2009.

[158] J. Lee and S. Leyffer, *Mixed integer nonlinear programming*. Springer Science & Business Media, 2011, vol. 154.

[159] F. Sohrabi and W. Yu, "Hybrid analog and digital beamforming for mmwave OFDM large-scale antenna arrays," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1432–1443, 2017.

[160] D. P. Palomar and J. R. Fonollosa, "Practical algorithms for a family of waterfilling solutions," *IEEE transactions on Signal Processing*, vol. 53, no. 2, pp. 686–695, 2005.

[161] Z. Zhang, X. Chai, K. Long, A. V. Vasilakos, and L. Hanzo, "Full duplex techniques for 5G networks: self-interference cancellation, protocol design, and relay selection," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 128–137, 2015.

[162] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, 2016.

[163] J. Mirza, G. Zheng, K.-K. Wong, S. Lambotharan, and L. Hanzo, "On the performance of multiuser MIMO systems relying on full-duplex CSI acquisition," *IEEE Transactions on Communications*, vol. 66, no. 10, pp. 4563–4577, 2018.

[164] 3GPP, "NR; physical channels and modulation (Release 15), TS 38.211."

[165] X. Li, X. Ge, X. Wang, J. Cheng, and V. C. Leung, "Energy efficiency optimization: Joint antenna-subcarrier-power allocation in OFDM-DASs," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7470–7483, 2016.

[166] Y. J. Zhang and K. B. Letaief, "Multiuser adaptive subcarrier-and-bit allocation with adaptive cell selection for OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 3, no. 5, pp. 1566–1575, 2004.

[167] S. S. Haykin, *Adaptive Filter Theory*. Pearson Education India, 2005.

[168] D. Schafhuber and G. Matz, "MMSE and adaptive prediction of time-varying channels for OFDM systems," *IEEE Transactions on wireless communications*, vol. 4, no. 2, pp. 593–602, 2005.

[169] A. Khansefid and H. Minn, "Achievable downlink rates of MRC and ZF precoders in massive MIMO with uplink and downlink pilot contamination," *IEEE Transactions on Communications*, vol. 63, no. 12, pp. 4849–4864, 2015.

[170] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation (Release 8), TS 36.211."

[171] G. Interdonato, E. Björnson, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Ubiquitous cell-free massive MIMO communications," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–13, 2019.

[172] S. Kusaladharma, W.-P. Zhu, W. Ajib, and G. A. A. Baduge, "Achievable rate characterization of NOMA-aided cell-free massive MIMO with imperfect successive interference cancellation," *IEEE Transactions on Communications*, vol. 69, no. 5, pp. 3054–3066, 2021.

[173] S. Jeong, A. Farhang, F. Gao, and M. F. Flanagan, "Frequency synchronisation for massive MIMO: a survey," *IET Communications*, vol. 14, no. 16, pp. 2639–2645, 2020.

[174] Y. Jiang, M. K. Varanasi, and J. Li, "Performance analysis of ZF and MMSE equalizers for MIMO systems: An in-depth study of the high SNR regime," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2008–2026, 2011.

[175] K. Shen and W. Yu, "Fractional programming for communication systems—part i: Power control and beamforming," *IEEE Transactions on Signal Processing*, vol. 66, no. 10, pp. 2616–2630, 2018.

[176] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operations research*, vol. 26, no. 4, pp. 681–683, 1978.

[177] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[178] X. Zhao, J. Kivinen, P. Vainikainen, and K. Skog, "Propagation characteristics for wideband outdoor mobile communications at 5.3 GHz," *IEEE Journal on selected areas in communications*, vol. 20, no. 3, pp. 507–514, 2002.

[179] D. Peaucelle, D. Henrion, Y. Labit, and K. Taitz, "User's guide for sedumi interface 1.04," *LAAS-CNRS, Toulouse*, 2002.

[180] S. Khan, H. Rahmani, S. A. A. Shah, and M. Bennamoun, "A guide to convolutional neural networks for computer vision," *Synthesis Lectures on Computer Vision*, vol. 8, no. 1, pp. 1–207, 2018.

[181] W. Cui, K. Shen, and W. Yu, "Spatial deep learning for wireless scheduling," *ieee journal on selected areas in communications*, vol. 37, no. 6, pp. 1248–1261, 2019.

[182] M. Eisen and A. Ribeiro, "Optimal wireless resource allocation with random edge graph neural networks," *ieee transactions on signal processing*, vol. 68, pp. 2977–2991, 2020.

[183] Y. Shen, Y. Shi, J. Zhang, and K. B. Letaief, "Graph neural networks for scalable radio resource management: Architecture design and theoretical analysis," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 1, pp. 101–115, 2020.

[184] A. Chowdhury, G. Verma, C. Rao, A. Swami, and S. Segarra, "Unfolding WMMSE using graph neural networks for efficient power allocation," *IEEE Transactions on Wireless Communications*, vol. 20, no. 9, pp. 6004–6017, 2021.

[185] Y. Sun and J. Han, "Mining heterogeneous information networks: a structural analysis approach," *Acm Sigkdd Explorations Newsletter*, vol. 14, no. 2, pp. 20–28, 2013.

[186] X. Wang, H. Ji, C. Shi, B. Wang, Y. Ye, P. Cui, and P. S. Yu, "Heterogeneous graph attention network," in *The world wide web conference*, 2019, pp. 2022–2032.

[187] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," *arXiv preprint arXiv:1710.10903*, 2017.

[188] M. Fey and J. E. Lenssen, "Fast graph representation learning with pytorch geometric," *arXiv preprint arXiv:1903.02428*, 2019.

[189] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[190] N. Zhao, Y.-C. Liang, D. Niyato, Y. Pei, M. Wu, and Y. Jiang, "Deep reinforcement learning for user association and resource allocation in heterogeneous cellular networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5141–5152, 2019.

[191] Y. Liao, Y. Hua, and Y. Cai, "Deep learning based channel estimation algorithm for fast time-varying MIMO-OFDM systems," *IEEE Communications Letters*, vol. 24, no. 3, pp. 572–576, 2019.

[192] Z. Wei, W. Yuan, S. Li, J. Yuan, G. Bharatula, R. Hadani, and L. Hanzo, "Orthogonal time-frequency space modulation: A promising next-generation waveform," *IEEE wireless communications*, vol. 28, no. 4, pp. 136–144, 2021.

[193] A. H. Sodhro, S. Pirbhulal, Z. Luo, K. Muhammad, and N. Z. Zahid, "Toward 6G architecture for energy-efficient communication in IoT-enabled smart automation systems," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5141–5148, 2020.

[194] Y. Al-Eryani, M. Akrout, and E. Hossain, "Multiple access in cell-free networks: Outage performance, dynamic clustering, and deep reinforcement learning-based design," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 4, pp. 1028–1042, 2020.

[195] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Towards dual-functional wireless networks for 6G and beyond," *IEEE Journal on Selected Areas in Communications*, 2022.

[196] F. Dong, F. Liu, Y. Cui, W. Wang, K. Han, and Z. Wang, "Sensing as a service in 6G perceptive networks: A unified framework for ISAC resource allocation," *arXiv preprint arXiv:2202.09969*, 2022.