Resource Allocation for Relay Based Green Communication Systems

Thesis submitted in accordance with the requirements of
the University of Liverpool for the degree of Doctor in Philosophy
by
Linhao Dong

September 2013

Declaration

The work in this thesis is based on research carried out at the University of Liverpool. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all my own work unless referenced to the contrary in the text.

Linhao Dong
Liverpool, United Kingdom

Abstract

The relay based cooperative network is one of the promising techniques for next generation wireless communications, which can help extend the cell coverage and enhance the diversity. To deploy relays efficiently with limited power and bandwidth under certain performance requirements, resource allocation (RA) plays an increasingly important role in the system design. In recent years, with the fast growth of the number of mobile phone users, great portion of CO_2 emission is contributed by wireless communication systems. The combination of relay techniques and RA schemes reveals the solution to green communications, which aims to provide high data rate with low power consumption. In this thesis, RA is investigated for next generation relay based green wireless systems, including the long-range cellular systems, and the short-range point-to-point (P2P) systems.

In the first contribution, an optimal asymmetric resource allocation (ARA) scheme is proposed for the decode-and-forward (DF) dual-hop multi-relay OFDMA cellular systems in the downlink. With this scheme, the time slots for the two hops via each of the relays are designed to be asymmetric, *i.e.*, with K relays in a cell, a total of $2K$ time slots may be of different durations, which enhances the degree of freedom over the previous work. Also, a destination may be served by multiple relays at the same time to enhance the transmission diversity. Moreover, closed-form results for optimal resource allocation are derived, which require only limited amount of feedback information. Numerical results show that, due to the multi-time and multi-relay diversities, the proposed ARA scheme can provide a much better performance than the scheme with symmetric time allocation, as well as the scheme with asymmetric time allocation for a cell composed of independent single-relay sub-systems, especially when the relays are relatively close to the source. As a result, with the optimal relay location, the system can achieve high throughput in downlink with limited transmit power.

In the second contribution, the power consumption in relay based 60 GHz cooperative networks is studied, which is based on three-terminal diversity amplify-and-forward (DAF) and diversity DF (DDF) relaying strategies. A total power consumption model including drive power, decoding power, and power consumption of power amplifier (PA) is proposed, excluding the transmit power, as it is relatively small compared to decoding power and PA power in the indoor environment. This model is formulated as a

function of drive power, which gives an easy access to the system level power allocation. To minimise the system total power consumption, the optimal drive power can be allocated to the source node by numerical searching method while satisfying the data rate requirement. The impact of relay locations on the total power consumption is also investigated. It is shown that, with the same data rate requirement, in the small source-relay separation case, DAF consumes slightly less power than DDF; while with larger source-relay separation, DAF consumes much more power than DDF.

In the future work, multiuser relay-based short-range communication systems will be considered for the 60 GHz communication in the fading channel scenario, which extends the proposed power consumption model in a more practical way. The power consumption model of other components, such as analog-to-digital converter, data buffer, modulation/demodulation could also be considered to provide more details about green P2P communications.

Acknowledgement

Throughout the passing four years, I have received loads of support and help from the following people who generously show their love and encouragement. Without their kindness, this thesis would not have been completed.

I would like to give my deepest gratitude to my supervisor Dr. Xu Zhu, who has been teaching me with great patient and guiding me from a graduate student to a mature researcher. This work no doubt benefits from her invaluable comments. I also appreciate the supervision upon my study from Prof. Yi Huang, who has contributed great comments in some of my publications.

I am grateful to Dr. Sumei Sun at Institute for Infocomm Research (I^2R), Singapore, where I spent six months as a postgraduate student attachment. During this period, Dr. Sun gave me lots of advices and guidance in my research work. I would also like to thank Dr. Yeow-Khiang Chia at I^2R, who helped me a lot with the information theory. Besides, the following researchers at I^2R deserve my gratitude for academical assistance and family-like atmosphere, and they are Dr. Koichi Adachi, Dr. Lin Shan, Dr. Xiaojuan Zhang, Mr. Gang Yang, Dr. Jingon Joung, Dr. Peng-Hui Tan, and Dr. Chi-Keong Ho. It is my pleasure to share this good time with you.

I would like to thank my (previous) colleagues in the Wireless Communication and Smart Grid group: Dr. Nan Zhou and Dr. Jingbo Gao for inspiring discussions in resource allocation schemes, Mr. Yufei Jiang for sharing views in the channel estimation theory, Mr. Teng Ma for the tutorials in channel measurement and body fitness, Mr. Chao Zhang and Yanghao Wang for discussions in smart grid and photography, Mr. Wenfei Zhu, Mr. Chaowei Liu, Mr. Jun Yin, Mr. Qinyuan Qian, and Mr. Yang Li for sharing a family-like atmosphere and delicious meals. I am so pleased to see this group now is growing up with the extra three new students: Mr. Mohammad Heggo, Mr. Kainan Zhu, and Mr. Zhongxiang Wei. Your guys are the finest!

I would also like to show my great appreciation to the Chinese congregations in Liverpool and Singapore of Jehovah's Witnesses. Thanks to your Christian love and spiritual support, I could get over the loneliness and find the meaning of real life.

Finally, my gratitude is dedicated to my parents. Without their support and love, I would have no opportunity to study abroad and pursue this degree.

This thesis is a humble gift to my grandfather Mr. Fu Hou, my friends Mrs. Chris-

tine Hobden and Mrs. Coi-Wan Tam. You are the true worriers fighting against cancer.

To God YHWH

[Prv 2:10] quia intrabit sapientia cor tuum, et scientia animae tuae placebit.

Contents

List of Figures

List of Tables

Nomenclature

3G third generation

3GPP third generation partnership project

4G fourth generation

5G fifth generation

AD analogue-to-digital

AF amplify-and-forward

AMPS advanced mobile phone system

ARA asymmetric resource allocation

AWGN additive white Gaussian noise

BEP bit error probability

BER bit error rate

BS base station

BSC binary symmetric channel

CDMA code-division multiple access

CEPT European Conference of Postal and Telecommunications Administrations

CIR channel impulse response

CNR channel-to-noise ratio

CP cyclic prefix

CSI channel state information

CoMP coordinated multipoint

DAF diversity amplify-and-forward

DDF diversity decode-and-forward

DF decode-and-forward

DxF diversity x forward

EM electromagnetic

ERP effective radiated power

EV-DO evolution-data optimized

FDMA frequency-division multiple access

FFT fast Fourier transform

GSM global system for mobile communications

HSPA high speed packet access

ICI inter-carrier interference

IFFT inverse fast Fourier transform

ISI inter-symbol interference

ITU International Telecommunication Union

IxF intersymbol interference x forward

KKT Karush-Kuhn-Tucker

KL Kullback-Leibler

LNA low noise amplifier

LOS line-of-sight

LS least square

LTE long term evolution

MAC medium access control

MIMO multiple-input and multiple-output

MMSE minimum mean squared error

MRC maximum ratio combining

MTS mobile telephone service

MxF multi-hop x forward

OFDM orthogonal frequency-division multiplexing

OFDMA orthogonal frequency-division multiple access

P2P point-to-point

PA power amplifier

PAE power added efficiency

PAPR peak-to-average power ratio

PE processor element

PL path loss

QoS quality of service

RA resource allocation

RF radio frequency

RMS root-mean-square

SA sub-carrier allocation

SCxF split-combine x forward

SNR signal-to-noise ratio

SRA symmetric resource allocation

TDMA time-division multiple access

UMTS universal mobile telecommunications system

VLSI very-large-scale integration

WiMAX worldwide interoperability for microwave access

WLAN wireless local area networks

WPAN wireless personal area networks

ZF zero-forcing

Chapter 1

Introduction

1.1 Motivation

Wireless communications have experienced a rapid development over the past decades [2] [3], and a surge of research activities have been carried out to provide a higher system capacity [4], a better quality of service (QoS) [5], lower power consumption [6], and more flexible system coverage [7]. However, according to a recent survey, 2% of the global CO_2 emission is contributed by the information and communication technology [8], and this figure is expected to increase each year. To provide the high data rate with low power consumption, the concept of 'green communications' attracts much attention [9].

Relay technologies, including amplify-and-forward (AF) [10] and decode-and-forward (DF) [11] relaying strategies, have been considered as an achievable solution for the green communication systems such as long term evolution advanced (LTE-Advanced) systems [12]. With the same cell coverage, cooperative networks can reduce the system power consumption with the same data rate requirement compared to the conventional networks [9]. Meanwhile, adaptive resource allocation (RA) plays an essential part in multi-carrier based energy efficient wireless systems [13]. Hence, the combination of relay techniques and RA schemes leads the development in green communications [14].

Green communications at 60 GHz attract considerable attention for the need of indoor high speed local area networks [15] [16] [17]. With the large bandwidth and low power consumption, 60 GHz green communications are able to support multi-Gbps wireless connections between personal electronic devices. However, one big challenge for 60 GHz wireless communications is the high attenuation in propagation [18]. To combat this problem with low transmit power, relay techniques are considered as promising solutions, which can improve the reliability of transmission and extend the coverage of signal source [19]. Hence, the power consumption of the relay based systems is less compared to the convectional P2P architectures, while satisfying the same high data rate requirement.

1.2 Research Contributions

This thesis first presents a joint ARA scheme for relay based long-range communication systems, where the RA consists of power allocation, sub-carrier allocation, and time duration allocation. By applying this RA scheme, the system throughput is maximised with limited transmit power, leading to a high energy efficiency. After that, a system power consumption model is proposed for relay based short-range green communication systems at 60 GHz, where the transmit power, PA power, and decoding power are considered as functions of the drive power of PA. By numerically searching method, the optimal drive power can be allocated at the source node to minimise the system power consumption while satisfying the data rate requirement.

The research during this PhD study has produced the following main contributions:

- An ARA scheme for DF dual-hop multi-relay OFDMA systems for long-range communication is proposed. This work is different in the following aspects. First, this is the first work to apply asymmetric time allocation to a general multi-relay multi-destination system. As a result, it enhances the degree of freedom for transmission over the previous symmetric RA (SRA) scheme. Second, an optimal algorithm is proposed to perform joint time, power and subcarrier allocation to obtain the global optimal results, with only limited amount of feedback information from relays and destinations. Compared to [20], the proposed work allows multiple relays to serve a single destination by using CoMP technique, which enhances the degrees of freedom. Numerical results show that, thanks to the multi-time and multi-relay diversities, the proposed ARA scheme outperforms the SRA algorithm in [21], as well as the ARA algorithm in [20], with higher achievable system throughput, especially when the relays are relatively close to the source. Moreover, impact of the relays' locations on the results of asymmetric time allocation is demonstrated.

- The power consumption in 60 GHz short-range communications with DAF and DDF relaying strategies. This work is different from others in the following aspects. First, a total power consumption model, which includes drive power, decoding power, and PA power is proposed. This model is a function of drive power, which makes access to the power allocation for the system. By searching for the optimal drive power, the minimum total power consumption can be achieved off-line. Second, this is the first work that applies the decoding power model from [22] in the relay systems, which gives a clear clue on the behaviour of decoding power against the drive power. Third, the parameters of the latest state-of-the-art 60 GHz PA and low noise amplifier (LNA) are applied into the system, which gives a meaningful guideline in power consumption since the transmit powers at source and relay are limited by the saturation of PA. The impact of relay locations

on the system performance is also investigated, and the comparison between the DAF and DDF relaying is made.

1.3 Thesis Organisation

The rest of this thesis is organised as follows. The wireless channels and systems are introduced in Chapter 2. Chapter 3 presents the literature survey on RA of broadband wireless communication systems. An ARA scheme for relay based long-range cellular systems is proposed in Chapter 4, where both optimal and suboptimal algorithms are considered. In Chapter 5, a power consumption model with its power allocation scheme is proposed for 60 GHz relay based short-range P2P communications, where the DAF and DDF relaying strategies are compared. Conclusions and future work are presented in the final chapter.

1.4 Publication List

A number of publications during the study of this research are listed below, which partially contribute to the thesis.

1. L. Dong, X. Zhu, Y. Jiang and Y. Huang "Optimal Asymmetric Resource Allocation for Dual-Hop Multi-Relay Based Downlink OFDMA Systems," *Mobile Computing*, vol. 2, no. 1, pp 1-8, Feb. 2013.

2. L. Dong, S. Sun, X. Zhu, and Y.-K. Chia, "Power Efficient 60 GHz Wireless Communication in Cooperative Networks," submitted to *IEEE Trans. Veh. Technol.*

3. L. Dong, X. Zhu, N. Zhou, and Y. Huang, "Asymmetric Resource Allocation for Decode-and-Forward Multi-Relay Systems," in *Proc. IEEE WiCOM'11*, Wuhan, China, Sep. 2011.

4. L. Dong, X. Zhu, and Y. Huang, "Optimal Asymmetric Resource Allocation for Multi-Relay Based LTE-Advanced Systems," in *Proc. IEEE GLOBECOM'11*, Houston, USA, Dec. 2011.

5. L. Dong, X. Zhu, and Y. Huang, "Optimal Asymmetric Resource Allocation for Dual-Hop Multi-Relay LTE-Advanced Systems in the Downlink," in *Proc. IEEE ICC'13*, Budapest, Hungary, Jun. 2013.

6. L. Dong, S. Sun, X. Zhu, and Y.-K. Chia, "Power Efficient 60 GHz Wireless Communication Networks with Relays," in *Proc. IEEE PIMRC'13*, London, UK, Sep. 2013.

Chapter 2

Research Overview

In this chapter, the wireless communication channel models and the channel capacity are described in Section 2.1 and 2.2, respectively. An overview of wireless communication systems is presented in Section 2.3. In Sections 2.4 and 2.5, the fundamentals of OFDM and relay techniques are introduced.

2.1 Wireless Communication Channel Models

The wireless channel is the air interface that links the transmitter and the receiver. Its properties dominate the information-theoretical capacity of communications, the performance limit of wireless systems, and the QoS level. Hence, it is essential to know the behaviour of wireless channels, which helps in the system design.

2.1.1 Large-Scale Path Loss

In the propagation of the electromagnetic (EM) wave through a certain medium, the energy of the carried signals is radiated out from a transmitter. If the receiver can 'pick

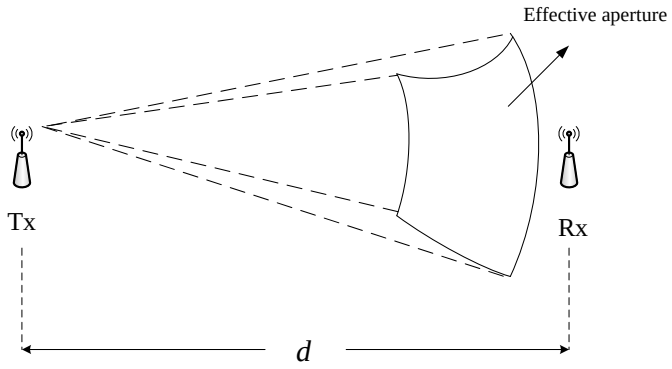

Figure 2.1: The propagation of the EM wave as a closed sphere surrounding the transmitter, (d – the distance between Tx and Rx)

up' part of this energy, it is possible that the information contained in the signals can be detected at the receiver, thus a data link is established.

If the radiation pattern of the antenna is assumed as isotropic, and the medium is as a vacuum environment, the EM wave propagates in a pattern of a sphere as shown in Fig. 2.1, where 'Tx' denotes the transmitter and 'Rx' denotes the receiver. The energy conservation law tells that the total energy of signals remains unchanged surrounding the closed free space of the transmitter. With the increasing of the distance that the EM wave travels, the total surface area of the sphere also grows correspondingly, which causes the reduction of the power at the unit area. As a result, the strength of the signals is attenuated. The concept of 'power' is used instead of 'energy' without the loss of generality. d is defined as the distance between Tx and Rx, P_T as the wavelength of the EM wave, and G_T as the gain of the transmitter antenna. The power density, or effective radiated power (ERP) [2], can be defined as

$$P_{eff} = \frac{G_T P_T}{4\pi d^2} \tag{2.1}$$

From (2.1), it can be observed that if the distance d is doubled, the ERP will decrease by around 6 dB.

At the Rx, G_R is defined as the antenna gain of the receiver. The effective aperture [23] can be simplified as

$$a_{eff} = \frac{\lambda^2}{4\pi} G_R \tag{2.2}$$

where λ is the wavelength of the EM wave at central frequency f_c. According to the Frii's equation [2], the received power can be written as

$$P_R = P_{eff} \times a_{eff} = P_T G_T G_R \left(\frac{\lambda}{4\pi}\right)^2 \left(\frac{1}{d}\right)^2 \tag{2.3}$$

Thus, the free-space path loss (PL) [24] is expressed as

$$A_{free}(d) = \frac{P_R}{P_T} = G_T G_R \left(\frac{\lambda}{4\pi}\right)^2 \left(\frac{1}{d}\right)^2 \tag{2.4}$$

If α is defined as the PL exponent, a general expression of PL can be written as

$$A(d) = \frac{P_R}{P_T} = G_T G_R \left(\frac{\lambda}{4\pi}\right)^2 \left(\frac{1}{d}\right)^\alpha \tag{2.5}$$

The value of α can be fitted by a huge number of measurements, and some example values of PL exponent are shown in Table 2.1.

In the practical usage, PL is usually given in the log form [25]. Therefore, (2.5) can be rewritten as

$$A_{\mathrm{dB}}(d) = A_0 + 10\alpha \log_{10} d \tag{2.6}$$

where $A_0 = 10 \log_{10} \frac{(4\pi)^2}{G_T G_R \lambda^2}$. Due to the fact that PL is mainly affected by the distance, PL plays a vital part in the wireless network design and performance analysis.

Table 2.1: PL exponent values in different scenarios

Scenarios	PL exponent α
Indoor line-of-sight (LOS)	$1.6 \sim 1.8$
Free space	2
Obstructed in factories	$2 \sim 3$
Urban area cellular	$2.7 \sim 3.5$
Shadowed urban cellular	$3 \sim 5$
Obstructed indoor	$4 \sim 6$

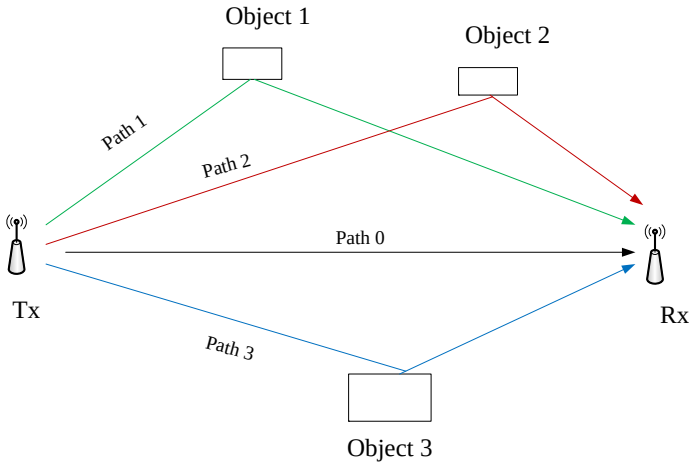

Figure 2.2: An simplified example of multipath effect of EM wave

2.1.2 Small-Scale Multipath Fading

Multipath fading is the rapid fluctuation of channel gain and change of phases over a short period of time or distance, which causes rapid small-scale variation of signals. In the urban area, the transmitted EM wave could be reflected, diffracted and scattered by objects such as buildings, vehicles, trees, and rough surfaces. Hence, the signal travels via different paths to the receiver. Fig. 2.2 shows the simplified model of the multipath effect, where path 0 indicates the LOS path. Each of the paths generates a unique wave of transmitted signal with the randomly distributed amplitude, phase, and delay. As a result, all these EM waves are combined at the receiver (as shown in Fig. 2.3), which causes the signal distortion and fading.

2.1.2.1 Parameters of Multipath Fading

Defining g_i and τ_i as the channel gain and delay for ith path, respectively, the *mean excess delay* [24] is the first moment of the power delay profile, and can be defined as $\bar{\tau} = \sum_i |g_i|^2 \tau_i / \sum_i |g_i|^2$. The root-mean-square (RMS) delay spread [24] is the second

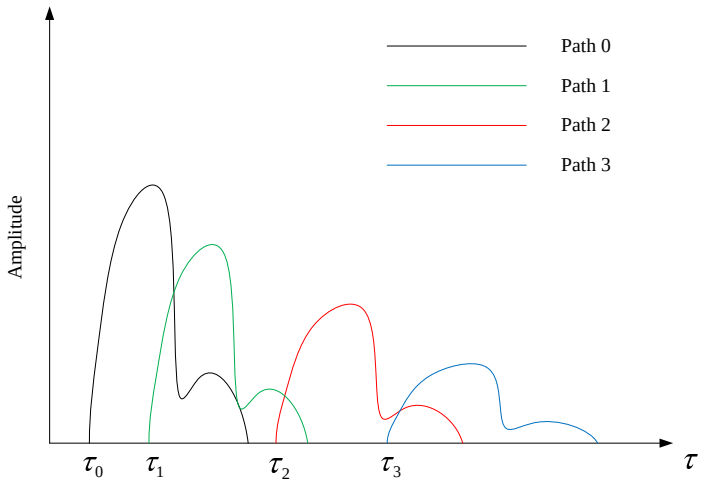

Figure 2.3: The received signals via different paths

central moment of the power delay profile, which can be defined as

$$\mu = \sqrt{\bar{\tau^2} - (\bar{\tau})^2} \tag{2.7}$$

where $\bar{\tau^2} = \sum_i |g_i|^2 \tau_i^2 / \sum_i |g_i|^2$. The RMS delay spread indicates the effect of multi-path. In other words, the higher the RMS delay spread is, the larger the effect will be.

In wireless communications, coherence bandwidth B_c is defined as the range of frequencies where two frequency components have a strong potential correlation for amplitude [3]. For instance, if 95% coherence bandwidth is 5 kHz, the behaviours of two frequency components, which are 5 kHz or less separated apart, are nearly the same. Roughly, coherence bandwidth can be approximated from RMS delay spread μ, *i.e.*, $B_{c,50\%} \approx 1/5\mu$, $B_{c,90\%} \approx 1/50\mu$.

When the mobile user moves with a relative velocity v with the base station as the inertial frame of reference, there is a shift on the frequency of signals. This effect is called *Doppler effect* [24], following the name of the Austrian physicist Christian Doppler. Defining θ as the angle between the direction of the received signal wave and the direction of the mobile user's motion, the Doppler shift f_d can be expressed as

$$f_d = \frac{v}{\lambda} \cos\theta \tag{2.8}$$

When $\theta = 0$, the direction of the received signal is in accordance with the direction of motion, where Doppler shift achieves the maximum as $f_{d\_\max} = v/\lambda$.

The coherence time T_c is referred to as the time duration over which the channel is considered to be not varying [3]. It is mainly dominated by the doppler shift, and can be approximated as $T_c \approx \frac{9}{16\pi f_{d\_\max}}$. In other words, the faster the mobile user moves, the more frequently the channel varies.

2.1.2.2 Types of Multipath Fading

Since the characteristics of the signal wave and channel are defined, the different types of small-scale multipath fading can be categorised by some thresholds. If T_s and B_{sg} are defined as the symbol period and signal bandwidth, respectively, there are four different types of multipath fading shown in Table 2.2.

Table 2.2: Types of multipath fading

Flat fading	$B_{sg} < B_{c,50\%}$, $T_s > 5\mu$
Frequency selective fading	$B_{sg} > B_{c,50\%}$, $T_s < 5\mu$
Slow fading	$T_s \ll T_c$
Fast fading	$T_s > T_c$

These types of fading are explained in the following details:

- **Flat fading**: When the symbol period is longer than the RMS delay spread (*i.e.*, $T_s > 5\mu$), or the bandwidth of the transmitted signal is less than the channel coherence bandwidth (*i.e.*, $B_{sg} < B_{c,50\%}$), all the frequency components within the transmitted bandwidth have nearly the same channel gain and linear phase. In this scenario, the channel is referred to as flat fading. Flat fading is the most common type described in the literature. The channel gains of flat fading vary in time randomly, and they follow different distributions such as Rayleigh fading, Rician fading, and Nakagami fading [2].

- **Frequency selective fading**: When the symbol period is shorter than the RMS delay spread of a wireless channel (*i.e.*, $T_s < 5\mu$), or the bandwidth of the transmitted signal is larger than the channel coherence bandwidth (*i.e.*, $B_{sg} > B_{c,50\%}$), previous transmitted symbols could easily cause interference to current transmitted symbols. This interference is called inter-symbol interference (ISI). In the frequency domain, the ISI is presented by a formation that frequency components of the received signal's spectrum undergo different amplitudes. Hence, this fading is referred to as frequency selective fading.

- **Slow fading**: The channel remains unchanged over one or several symbol periods is called slow fading channel, where the symbol period is much less than the coherence time (*i.e.*, $T_s \ll T_c$).

- **Fast fading**: In this scenario, the channel changes so fast that the signal passes through different channels even within one symbol period. In other words, the symbol period is longer than the coherence time (*i.e.*, $T_s > T_c$).

According to the definitions, the RMS delay spread and coherence time are two independent properties. Thus, there are four combinations of these fading channels, which are flat slow fading channel, flat fast fading channel, frequency selective slow

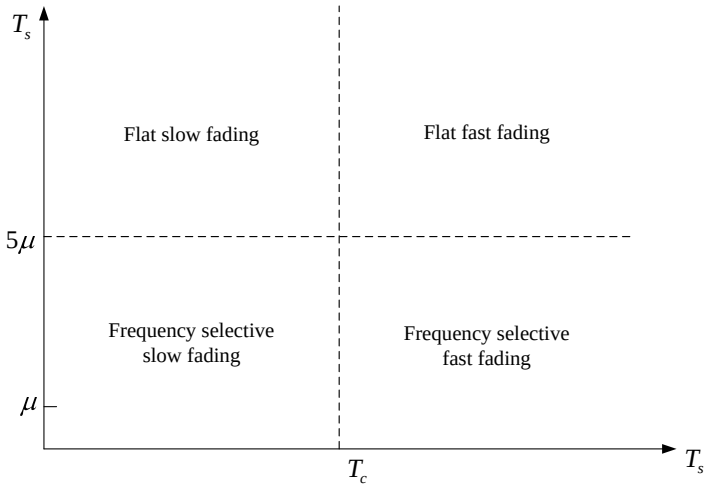

Figure 2.4: Fading types classified by symbol period

fading channel, and frequency selective fast fading channel. Fig. 2.4 illustrates the four multipath fading combinations with RMS delay spread and coherence time as thresholds. Although there are no clear boundaries amongst these types, the channel is still assumed as 'pure' flat slow fading or 'pure' frequency selective fast fading, etc. Throughout this thesis, frequency selective slow fading channels are focused on.

2.1.3 Fading Channel Model

In this subsection, the fading channel model is introduced. For a wireless communication system, each symbol is transmitted from the transmitter within signal bandwidth B_{sg}, and during symbol period T_s. Assuming there are N_p paths between the transmitter and receiver, the channel impulse response (CIR) can be written as

$$g(t) = \sum_{i=0}^{N_p-1} g_i \delta\left(t - \tau_i\right) \tag{2.9}$$

where $\delta(\cdot)$ is defined as the impulse function. Note that when $N_p = 1$, the channel is classified as flat fading.

It is assumed that each path is independent from another with different delays, and g_i is an independent zero mean complex Gaussian random variable. The variance of g_i follows the discrete exponential power delay profile as

$$\mathbb{E}\left\{|g_i|^2\right\} = b \cdot \exp\left(-\frac{\tau_i}{\mu}\right) \tag{2.10}$$

where b is the normalising factor. Usually, $|g_i|$ follows the Rayleigh distribution, of which the probability density function can be written as

$$p\left(|g_i|\right) = \frac{|g_i|}{V^2} \exp\left(-\frac{|g_i|^2}{2V^2}\right),\ 0 \leq |g_i| < \infty \tag{2.11}$$

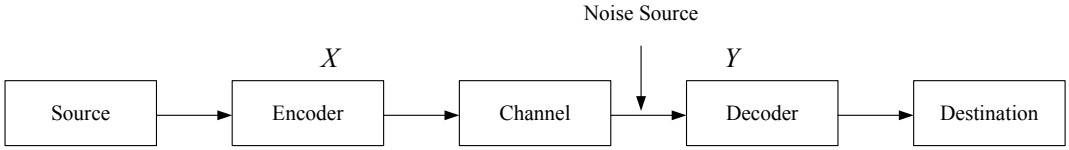

Figure 2.5: Block diagram of a communication system

where V is the RMS voltage of the received signal, and V^2 denotes the time average power of the received signal.

Letting $f(t)$ represent the pulse shape with the effects of the transmit and receive filters, the overall CIR is the convolution of the physical CIR $g(t)$ with $f(t)$, as

$$h(t) = f(t) * \sum_{i=0}^{N_p-1} g_i \delta\left(t - \tau_i\right) = \sum_{i=0}^{N_p-1} g_i f\left(t - \tau_i\right) \tag{2.12}$$

In the simulation of a wireless system, the raised-cosine filter, which is an implementation of a low-pass Nyquist filter, is widely applied to approximate the filter effect at transmitter and receiver [26]. Defining φ as the roll-off factor, $f(t)$ can be written as

$$f(t) = \operatorname{sinc}\left(\frac{t}{T_s}\right) \frac{\cos\left(\frac{\pi\varphi t}{T_s}\right)}{1 - \frac{4\varphi^2 t^2}{T_s^2}} \tag{2.13}$$

where T_s is the symbol period.

2.2 Channel Capacity

In the information theory, channel capacity of a communication channel is the tightest upper bound on the data rate that can be reliably transmitted [27]. As early as in 1948, Claude E. Shannon derived this upper bound for both discrete and continuous channels [28], which is called *Shannon capacity* nowadays. According to his results, capacity of a channel is given by the supremum of the average mutual information between the source and destination.

Now consider a simplest channel model as shown in Fig. 2.5, where X and Y are discrete random variable ensembles representing the input and output of the channel, respectively. $p_{X|Y}(x|y)$ is defined as the conditional probability density function of X given Y, which is determined by the channel. Defining $p_Y(y)$ as the marginal probability density of Y, and $p_{X,Y}(x,y)$ as the joint probability density, where

$$p_{X,Y}(x,y) = p_{X|Y}(x|y)p_Y(y) \tag{2.14}$$

Self-information indicates how much information is contained in one random variable as the outcome [29]. The conditional self-information of variable x given y can be

defined as

$$I_{X|Y}(x|y) = -\log p_{X|Y}(x|y) \tag{2.15}$$

Letting $p_X(x)$ denote the marginal probability density of X, the self-information of x can be written as

$$I_X(x) = -\log p_X(x) \tag{2.16}$$

Hence, the mutual information between X and Y can be written as the difference between (2.15) and (2.16), which is

$$I_{X;Y}(x,y) = I_X(x) - I_{X|Y}(x|y) = \log \frac{p_{X|Y}(x|y)}{p_X(x)} \tag{2.17}$$

From [29] it is known that the *entropy* is a measure of the uncertainty in random variable ensembles, and can be calculated as the expected value of self-information. $H(X|Y)$ is defined as the conditional entropy of X over joint XY ensemble, and it can be expressed as

$$H(X|Y) = -\sum_{x\in X, y\in Y} p_{X,Y}(x,y) \log p_{X|Y}(x|y) \tag{2.18}$$

Similarly, the entropy of X can be written as

$$H(X) = -\sum_{x\in X} p_X(x) \log p_X(x) \tag{2.19}$$

The average mutual information can be written as

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= \sum_{x\in X}\sum_{y\in Y} p_{X,Y}(x,y) \log \frac{p_{X|Y}(x|y)}{p_X(x)} \\ &= \sum_{x\in X}\sum_{y\in Y} p_{X,Y}(x,y) \log \frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)} \end{aligned} \tag{2.20}$$

If variables X and Y are continuous, (2.20) can be rewritten as

$$I(X;Y) = \int_X \int_Y p_{X,Y}(x,y) \log \frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)} dxdy \tag{2.21}$$

When the ensembles X and Y are mutually independent, the joint probability density can be written as $p_{X,Y}(x,y) = p_X(x)p_Y(y)$. As a result, (2.17) and (2.21) both become zero. This case indicates that if the channel quality is extremely low, it is impossible to transmit any information from the source to destination. In other words, mutual information describes how much information can be shared between two communication nodes. The channel capacity thus can be defined as

$$C = \sup_{p_X(x)} I(X;Y) \tag{2.22}$$

The rest of this section discusses different forms of average mutual information in binary symmetric channel (BSC), additive white Gaussian noise (AWGN) channel, and fading channel. Throughout this thesis, the capacity of AWGN channel and fading channel is considered.

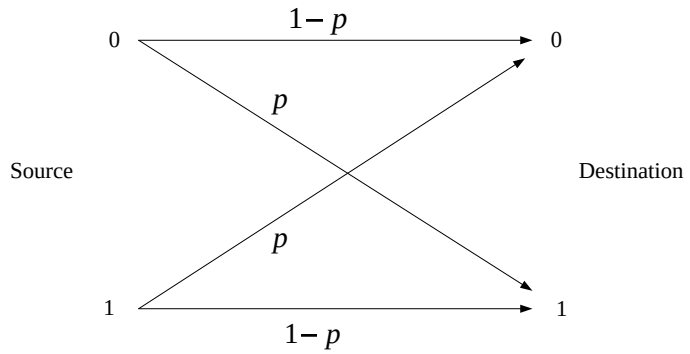

Figure 2.6: Binary symmetric channel

2.2.1 Binary Symmetric Channel

Since digital systems are most widely applied in communication network, the messages transmitted between the source and destination can be simply regarded as binary sequences composed by '0' and '1'. To describe the behaviour of the transmission, BSC model is introduced [29].

Fig. 2.6 illustrates the basic principle of BSC, where p is the flipping probability of binary symbol. When the source transmits only binary symbols, the destination can receive either 0 or 1. If the channel output is observed from the destination, the flipping probability can be expressed as

$$p_{X|Y}(x=0|y=0) = p_{X|Y}(x=1|y=1) = 1-p \tag{2.23}$$

$$p_{X|Y}(x=1|y=0) = p_{X|Y}(x=0|y=1) = p \tag{2.24}$$

The entropy $H(X)$ of ensemble X is a binary entropy function [27], as

$$\begin{aligned} H(X) &= -\sum_{X} p_X(x)\log_2 p_X(x) \\ &= -p\log_2 p - (1-p)\log_2(1-p) \end{aligned} \tag{2.25}$$

The capacity can be written as

$$C = 1 - H(X) \tag{2.26}$$

For instance, if the flipping probability is $p = 0.01$, the capacity is $C \approx 0.92$.

The idea of BSC is simple. While in the system design, the expression of BSC capacity is not practical since flipping probability cannot reveal much information about parameters setup, such as transmit power, bandwidth, etc. in the system. Therefore, a more practical form of channel capacity is in need.

2.2.2 Additive White Gaussian Noise Channel

For AWGN channel model, communication quality is only affected by an additive white random noise. This noise is caused by many sources such as the thermal vibration of

the particles contained in electronic components, black body radiation from the earth, and celestial sources, etc. The noise amplitude follows the Gaussian distribution, thus this kind of noise is named *Gaussian noise* [24].

In the measurement and research, Johnson-Nyquist noise [3] is usually approximated as AWGN. Defining T as the temperature in Kelvin, the single-sided noise power spectral density N_0 can be defined as

$$N_0 = k_B T \tag{2.27}$$

where k_B as the Boltzmann's constant, of which the value is 1.38×10^{-23} J/K. B is defined as the bandwidth of the noise, the noise power σ_0^2 thus can be written as $\sigma_0^2 = BN_0$. Defining z ($z \in Z$) as the AWGN noise, the probability density of z can be expressed as

$$p_Z(z) = \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left(-\frac{|z|^2}{2\sigma_0^2}\right) \tag{2.28}$$

Now consider a transmission where the source transmits the symbol x ($x \in X$, and $\mathbb{E}\{|x|^2\} = 1$) with the average power P_T over distance d between the source and destination. The noise z is added to the received symbols. At the destination, the received symbol y ($y \in Y$) can be expressed as

$$y = \sqrt{A(d)P_T}x + z \tag{2.29}$$

where $A(d)$ is the PL. From [28], the entropy of the noise Z is

$$\begin{aligned} H(Z) &= -\int_Z p_Z(z) \log p_Z(z) dz \\ &= \frac{1}{2} \log 2\pi e \sigma_0^2 \end{aligned} \tag{2.30}$$

It is assumed that the transmitted symbol ensemble X is independent from noise Z. Hence, the average power of Y can be written as

$$\begin{aligned} \mathbb{E}\{|y|^2\} &= \mathbb{E}\left\{|\sqrt{A(d)P_T}x + z|^2\right\} \\ &= A(d)P_T + \sigma_0^2 \end{aligned} \tag{2.31}$$

Since $y = \sqrt{A(d)P_T}x + z$, Y can be regarded as the variable which follows Gaussian distribution with variance $A(d)P_T + \sigma_0^2$. As a result, the entropy $H(Y)$ can be written as

$$H(Y) = \frac{1}{2} \log 2\pi e \left(A(d)P_T + \sigma_0^2\right) \tag{2.32}$$

In [28], Shannon proved that the mutual information $I(X;Y)$ for AWGN channel can be expressed as

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Z) \tag{2.33}$$

As a result, substituting (2.30) and (2.32) into (2.33), the capacity for AWGN channel can be approximated as

$$C = \frac{1}{2} \log\left(1 + \frac{A(d)P_T}{\sigma_0^2}\right), \text{ bit/s/Hz} \tag{2.34}$$

Note that if the modulation is complex, there is no pre-log factor 1/2.

Now consider a complex modulation in the continuous-time AWGN channel with bandwidth B Hz. By the passband-baseband conversion [24], there should be at least B complex samples in one second. Hence, the channel capacity can be written as

$$C = B \log \left(1 + \frac{A(d) P_T}{B N_0} \right), \text{ bit/s} \tag{2.35}$$

where the received SNR is $A(d)P_T/BN_0$. Equation (2.34) and (2.35) with their alternative forms are frequently quoted throughout this thesis.

2.2.3 Fading Channel

In this scenario, it can be assumed that the complex channel h_s has the same amplitude for all the frequency components within channel state s ($s \in S$), where the expected value over all the states is $\mathbb{E}\{|h_s|^2\} = A(d)$. C_s is defined as the channel capacity of h_s, and $p(s)$ as the probability that the channel remains state s over all the time period. The channel capacity for fading channel can be expressed as [30]

$$C = \sum_S C_s p(s) \tag{2.36}$$

Defining $x[n]$ as the transmitted symbol at discrete time n, and $1/B$ as the sample rate, the instantaneous symbol at receiver can be written as

$$y[n] = h_s[n] \sqrt{P_T} x[n] + z[n] \tag{2.37}$$

where P_T is the transmit power, and $z[n]$ is the sampled AWGN noise. The instantaneous channel-to-noise ratio (CNR) at receiver is defined as $\gamma[n] = |h_s[n]|^2/BN_0$, of which the expected value is $\gamma = A(d)/BN_0$. Letting $p_\gamma(\gamma[n]) = p(\gamma[n] = \gamma)$ denote the probability distribution, (2.36) can be rewritten as

$$C = \int_\gamma B \log \left(1 + P_T \gamma[n]\right) p_\gamma(\gamma[n]) d\gamma[n] \tag{2.38}$$

where the channel state set S is assumed continuous. According to Jensen's equality, $\mathbb{E}_S\{C_S(\gamma[n])\} \leq C_S(\mathbb{E}_S\{\gamma[n]\})$ always holds [2], which shows the fading channel capacity is no larger than the AWGN capacity with the same average transmit power.

If the receiver can give the feedback of the full channel state information (CSI), it is possible to adjust the transmit power $P_T(\gamma[n])$ at the source with $\gamma[n]$. Thus, the channel capacity becomes

$$C = \max_{P_T(\gamma[n])} \int_\gamma B \log \left(1 + P_T(\gamma[n]) \gamma[n]\right) p_\gamma(\gamma[n]) d\gamma[n] \tag{2.39}$$

subject to:

$$P_T(\gamma[n]) \leq P_T \tag{2.40}$$

where (2.40) limits the transmit power by P_T. More details about power allocation is discussed in Chapter 3.

2.3 Overview of Wireless Communication Systems

Nowadays, wireless communication is so common that it enables people to get internet access anywhere with mobile devices. However, the radio transmission has a long history of development. When the days back to 1860s, it was mathematically and experimentally proved by James Clerk Maxwell that EM waves have the ability of propagating in a certain medium [31]. After more than ten years of early attempts with inventions, in 1879 David Edward Hughes discovered that sparks would excite signals that can be detected in a telephone receiver, and he demonstrated the transmission of Mores code with his invention called 'spark-gap transmitter' [32]. In the late 1880s, Heinrich Rudolf Hertz observed sparks can be transformed back from electric waves by a slotted metallic circle (like a loop antenna) from a modified spark-gap transmitter, which is the first time of wireless transmission throughout history [33]. Around 1893, Nikola Tesla suggested the information could be transmitted without wires, and proposed that the radiation property in radio frequency wave might be utilised in telecommunication. In 1897, Guglielmo Marconi established the 'Wireless Telegraph Trading Signal Company', and built a radio station at Isle of Wright to start the experimental usage of wireless transmission. Following his series of demonstrations on the radio transmission at different places, the commercialisation of radio was motivated.

After a few pioneering attempts since 1918, the first commercialised mobile telephone service (MTS) was commenced by AT&T in 1948, which provided services for 5,000 customers [34]. In the late 1960s, the early model of modern cellular network was proposed by Richard H. Frenkiel [35], which includes the basic idea of frequency reuse and handoff. Since then, the cellular network architecture has been regarded as the standard structure of multiuser wireless communication networks. The first cellular network was deployed in North America around 1978, which is called advanced mobile phone system (AMPS). In this system, frequency-division multiple access (FDMA) scheme is accepted as a solution to spectrum sharing. However, due to the disadvantage of the unencrypted analogue signals, the voice messages were easily scanned. With the revolution of the semiconductor and very-large-scale integration (VLSI) technologies, the wireless technology stepped into the digital age in the early 1990s. The first second generation (2G) standard network, global system for mobile communications (GSM) system, was launched in 1991 in Finland. Until 2007, GSM had provided its service to over 2 billion customers around the world [36].

The early 2000s have witnessed the breakthrough of the development in internet. At nearly the same time, two 3G standards, universal mobile telecommunications system (UMTS) [37] and code division multiple access 2000 (CDMA2000) [38] were deployed. One of the outstanding features of 3G standard is the internet service that allows the connection between the customers' handsets and the internet servers. However, UMTS and CDMA2000 meet their bottleneck to fully satisfy the dramatic increasing demands

from customers for the data services. Hence, to meet this data rate requirement, evolution-data optimized (EV-DO) [39] and high speed packet access (HSPA) [40], which are regarded as 3.5G standards, were launched in 2003 and 2005, respectively. The maximum download speeds for HSPA and EV-DO are 14.4 Mbit/s and $4.9 \times N$ Mbit/s, respectively, and their maximum upload speeds are 5.76 Mbit/s and $1.9 \times N$ Mbit/s, respectively, where N is the number of 1.25 MHz spectrum chunk used in EV-DO systems. The updated version of HSPA, which is called as HSPA+ [41], was proposed to improve the system capacity of the early version. The downlink and uplink peak data rates of HSPA+ are $42 \times N$ and $11 \times N$, respectively, where N is the number of 5 MHz carriers employed.

Recently, the markets in North America, Asia and Europe have launched the 3.9G standards, worldwide interoperability for microwave access (WiMax) and long term evolution (LTE) systems. Based on the IEEE 802.16 protocol, WiMAX can provide broadband access with downlink and uplink speeds up to 75 Mbit/s and 25 Mbit/s, respectively. Meanwhile, the current LTE, which belongs to 3rd generation partnership project (3GPP) release 8, offers up to 326.4 Mbit/s and 86.4 Mbit/s as the peak downlink and uplink data rates. After 30 years of development from 1G, it is so surprising to see that the data rate of LTE is 5,800 times of the peak data rate of AMPS in downlink (the peak data rate of AMPS is 56 kbit/s)! While, the pursuing of higher data rate in wireless communication never stops. Developed from WiMAX and LTE, WiMAX 2 and LTE-Advanced were proposed in 2011 as the 4G standards to support peak downlink data rate up to 1 Gbit/s and peak uplink data rate up to 500 Mbit/s [42]. In LTE-Advanced systems, relay and coordinated multipoint (CoMP) techniques are adopted to help improve the cell-edge performance [12] [43]. How about the 5G network? It is believed that millimetre-wave communication is one of the promising solutions to the future 5G standard [44] [45], while there are still many issues need to be addressed. Chapter 5 introduces some technical details about power consumption of millimeter-wave communication systems. All the aforementioned wireless communication techniques are summarised in Table A.1 of Appendix A. As shown in the table, it can be observed that in the early days of wireless communication, each country had her own standards. While with the tendency of globalisation, the boundaries of standards between countries have been merging.

2.4 OFDM Techniques

In recent years, the revolution of smart mobile devices has excited the need for high data rate transmission. These smart phones and tablet computers, such as iPhone, Android phone, iPad, etc., are all equipped with wireless communication components. Due to this fact, a variety of online services can be provided to the customers. For instance, the online video clips need high requirement for the QoS, such as low packet latency,

and acceptable resolution of images. While, the higher the transmit data rate, the larger the impact of frequency selective fading is, which prevents the system satisfying the QoS of online video.

One of the existing solutions is deploying high speed Wi-Fi access points to provide internet connection at the local area. This communication protocol is based on the IEEE 802.11 series. In 1999, 802.11a (now is specified as clause 18 [46]) was proposed to use orthogonal frequency-division multiplexing (OFDM) as the modulation scheme to combat with the impact of frequency selective fading. Since then, OFDM has been effectively applied in the successive protocols such as 802.11g/n, and the peak downlink data rate can achieve up to 150 Mbit/s for one stream [46]. Compared to other techniques such as time-division multiple access (TDMA) and CDMA, OFDM requires less complex equalisation filters, has higher spectral efficiency, and is robust against the ISI. However, some disadvantages of OFDM limit the system performance in some particular situations. One is that OFDM signals have high peak-to-average power ratio (PAPR), which causes poor power efficiency by applying linear amplification circuity at transmitter [47]. The solutions of PAPR problem are discussed in Section 3.1. Another is that OFDM signal transmission is sensitive to Doppler shift [26], which causes inter-carrier interference (ICI). A variety of schemes are proposed in [48] [49] [50] to cancel the ICI. Hence, OFDM has been selected as one of the key techniques in 4G standard, such as LTE-Advanced [51], after weighting the pros and cons.

2.4.1 OFDM

OFDM uses a large number of closely-spaced orthogonal sub-carriers to carry data [52]. The data is reformed into several parallel streams, with each stream on one sub-carrier. Then each sub-carrier is modulated with a conventional modulation scheme at a low symbol rate, maintaining total data rates similar to conventional single-carrier modulation schemes in the same bandwidth. The OFDM streams can be viewed as many slowly-modulated narrow-band signals rather than one rapidly-modulated wide-band signal, which allows low-complex equalisation design.

Fig. 2.7 shows the basic structure of OFDM in a communication system, where $x[n]$ $(n = 0, ..., N-1)$ are the source symbols, and N is the number of sub-carriers. The cyclic prefix (CP) is designed to eliminate the multipath effect when the OFDM symbols arrive at the receiver. Defining T_s as the symbol period, the transmitted symbols after the inverse fast Fourier transform (IFFT) can be expressed as

$$X(t) = \frac{1}{N} \sum_{i=0}^{N-1} x[i] e^{j\frac{2\pi}{T_s} it}, \ 0 \leq t \leq T_s \tag{2.41}$$

where $e^{j\frac{2\pi}{T_s} it}$ is the ith sub-carrier out of N sub-carriers. With the sample rate of T_s/N,

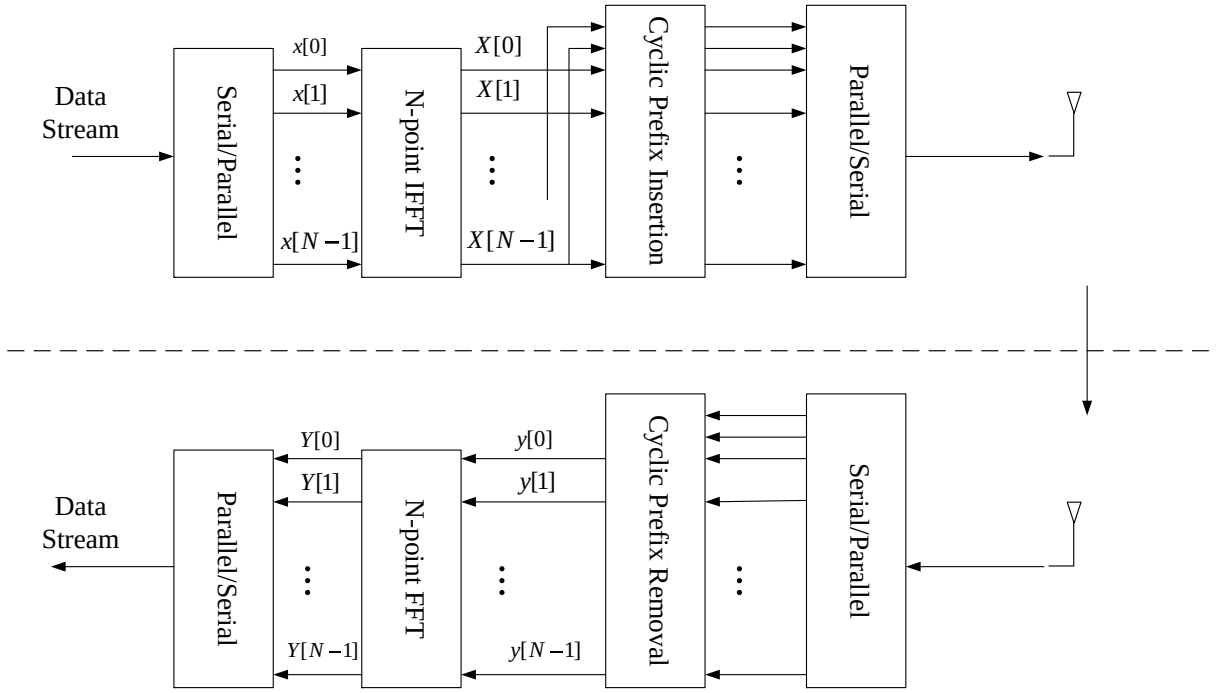

Figure 2.7: Simplified block diagram of the OFDM system

the discrete OFDM symbol can be sampled as $X[n] = X(nT_s/N)$, which is

$$X[n] = \frac{1}{N} \sum_{i=0}^{N-1} x[i] e^{j\frac{2\pi}{N} in},\ n = 0, ..., N-1 \tag{2.42}$$

From Fig. 2.7, it is known that CP is normally the prefixing of an OFDM symbol with the repetition of the end. To cover the whole information, only the part from 0 to $N-1$ is considered for simplicity. It can be also observed that each OFDM symbol contains the partial information from all N original symbols $\{x[n]\}_{n=0}^{N-1}$. Hence the symbol period spreads as NT_s, which leads to a mitigated ISI.

At the receiver, the signals are

$$y(t) = h(t) \otimes \sqrt{p_n} X[n] + z[n],\ n = 0, ..., N-1 \tag{2.43}$$

where $\otimes$ denotes the circular convolution, $h(t)$ is the channel coefficient, $z[n]$ is the AWGN noise, and p_n is the transmit power for symbol $X[n]$. N_c is defined as the 'channel memory', which is the maximum expected value of the delay taps [24]. After sampling, the received symbols can be expressed as

$$y[n] = \sqrt{p_n} \sum_{i=0}^{N_c} h[i] X[i-n] + z[n],\ n = 0, ..., N-1 \tag{2.44}$$

The CP is removed after the received serial symbols sequence being converted into parallel vector. Then the fast Fourier transform (FFT) is applied to the sampled

received symbols, which yields

$$\begin{aligned} Y[n] &= \text{FFT}\left\{\sqrt{p_n}\sum_{i=0}^{N_c} h[i]X[i-n] + z[n]\right\} \\ &= H_n\sqrt{p_n}x[n] + Z[n],\ n = 0, ..., N-1 \end{aligned} \tag{2.45}$$

where H_n and $Z[n]$ are the frequency domain CIR and noise on the nth sub-carrier in the frequency domain. Note that each sub-carrier is flat fading to the OFDM symbol, therefore each sub-carrier can be processed independently. As long as the perfect or imperfect CSI is obtained by the receiver, the channel can be estimated as $\hat{H}_n$. The original symbols thus can be detected by various channel equalisation method, such as zero-forcing (ZF) or minimum mean squared error (MMSE) [53]. For instance, the result from ZF equalisation can be written as

$$\begin{aligned} \lim_{\hat{H}_n \to H_n} \tilde{x}[n] &= \lim_{\hat{H}_n \to H_n} \left\{\frac{H_n}{\hat{H}_n\sqrt{p_n}}\sqrt{p_n}x[n] + \frac{Z[n]}{\hat{H}_n\sqrt{p_n}}\right\} \\ &= x[n] + \frac{Z[n]}{H_n\sqrt{p_n}},\ n = 0, ..., N-1 \end{aligned} \tag{2.46}$$

From (2.46), the symbol error of $\tilde{x}[n]$ depends on the accuracy of estimated channel information $\hat{H}_n$ and the weighted noise $Z[n]/\hat{H}_n\sqrt{p_n}$. Note that the error may be caused by the 'amplified' noise, so ZF is not an optimal equalisation method.

As each sub-carrier is flat faded, the capacity of OFDM system is the summation of the channel capacity of all the sub-carriers [25], as

$$C = \max_{p_n} B\sum_{n=0}^{N-1}\log_2\left(1 + \frac{p_n|H_n|^2}{BN_0}\right) \tag{2.47}$$

subject to

$$\sum_{n=0}^{N-1} p_n \leq P_T \tag{2.48}$$

where B is the total bandwidth for transmission, N_0 is the single-sided noise power spectral density. and P_T is the limited transmit power.

2.4.2 OFDMA

OFDM technique is not only a modulation technique, but can also be extended to a multiple access scheme, which is orthogonal frequency-division multiple access (OFDMA) [54], for multiuser systems. Fig. 2.8 shows a simple OFDMA system where two users, D_1 and D_2 share one channel for transmission in the downlink. The base station assigns sub-carriers to D_1 and D_2 after receiving the transmission requirements. As shown in Fig. 2.8, the red arrows represent the sub-carriers assigned to D_1, and the blue ones are for D_2. The sub-carrier allocation (SA) information is included in the header of each

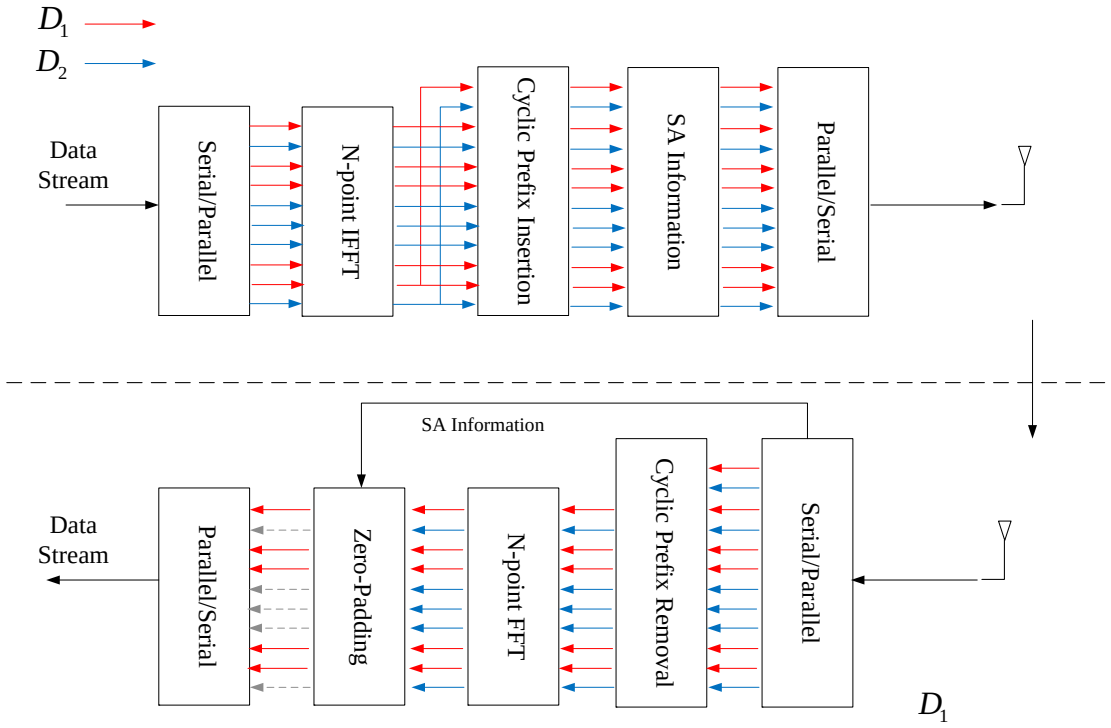

Figure 2.8: Block diagram of an OFDMA system in downlink

packet, which is also transmitted to all users. When one of the user, *i.e.* D_1, receives the signals, the SA information tells D_1 the indexes of wanted sub-carriers.

In the 4G standard, OFDMA is used in several protocols such as 802.16 WiMAX [55] and 3GPP LTE. A few literatures have investigated the OFDMA system in cellular networks. To minimise the total transmit power, the joint RA algorithm with different constraints are proposed in [6] [56]. The authors in [57] [58] and [59] proposed a chunk-based RA algorithm to maximise the system throughput under the average bit-error-rate (BER) and total power constraints. As green radio has a huge potential of saving energy, the trade-off between energy efficiency and spectral efficiency was studied in [60], and a low-complexity RA algorithm was also proposed for practical application. However, the aforementioned works only investigated the single-cell network with direct links. The RA for relay-based OFDMA systems was studied in [61] [62] [63] [64]. An efficient joint RA scheme was proposed in [61] to maximise the weighted data rate for a two-way relay based OFDMA system. Focusing on the weighted sum data rate maximisation as [61] did, the authors in [62] investigated the multiple DF relaying system with optimal RA algorithm. In [63], a joint optimisation problem of RA and scheduling was formulated in multiple-input and multiple-output (MIMO) OFDMA systems, where the full-duplex relays are considered. An efficient radio resource management scheme was proposed in [64], where dynamic routing, fair scheduling, and

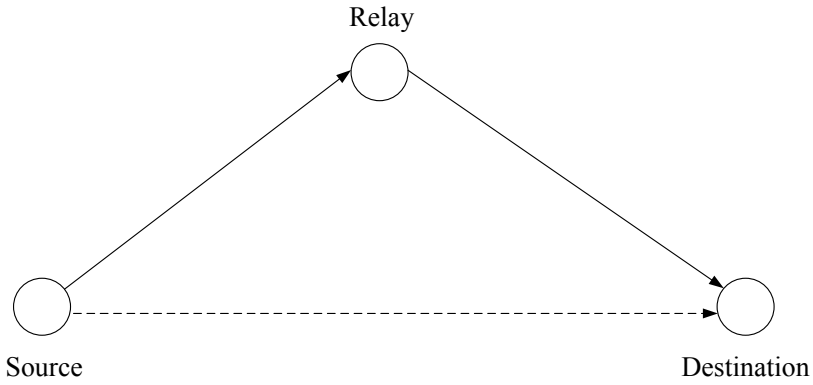

Figure 2.9: A three-terminal relay network

load balancing amongst cells are considered. While in this thesis, the OFDMA cellular systems with dual-hop DF relaying is considered in Chapter 4.

2.5 Relay Techniques

In 1971, Van Der Meulen proposed a three-terminal communication channel model in [65], which set the basic concept of cooperative networks. According to his theory, it was proved that an additional node is able to enhance the communication quality of a P2P network. In 1979, Thomas Cover *et al* derived the channel capacity of cooperative memoryless channels with different scenarios. Nicholas Laneman *et al* proposed and analysed low-complexity cooperative protocols in [66], which includes most attractive AF and DF relaying strategies, and the outage behaviour was developed as a characteristic of system performance. Since then, relay technique has been considered as one of promising solutions to future generation green networks [67].

Fig. 2.9 illustrates a one-way three-terminal relay network, which consists of the source node, the relay node, and the destination. The good direct link between the source to destination cannot be guaranteed all the time due to the obstacles, or the unfulfilled QoS requirement. Hence, the dashed line indicates this uncertainty of channel state. When there is a direct link, the relay node helps to improve the diversity of the system [68] [69]. However, when the direct link is obstructed or 'weak', the indirect link can be supported by the relay node.

A few strategies were proposed in recent years [66] [70] [71], amongst of which AF and DF relaying strategies are most well investigated. In the rest of this section, these two strategies are demonstrated in a dual-hop relay network without the loss of generality, where the direct link from the source to destination is disabled [2].

2.5.1 Amplify-and-Forward Relaying

Now $\{b_S^i\}$ is defined as the information bits transmitted from the source, and $\{x_S^n\}$ is defined as the coded symbols with code rate $R_c = i/n$. In the first hop transmission, the received symbols at the relay node can be simplified as

$$y_R^n = h_{SR}\sqrt{P_S}x_S^n + z_R \tag{2.49}$$

where h_{SR} is the CIR from the source to relay, P_S is the limited transmit power at the source, and z_R is the AWGN noise at the receiver of the relay node. The main function of the relay node in AF relaying is to amplify the received signal by a certain factor β in amplitude. Due to the fact that no detection or decoding is involved in the relaying, this strategy is also called non-regenerative relaying [72], where the signals are amplified as analogue waveforms.

One of the main issues for AF relaying is the design of amplification factor β. Defining $A_{SR} = \mathbb{E}\{|h_{SR}|^2\}$ as the PL of the source-relay link, the average power of y_R can be written as

$$\mathbb{E}\{|y_R^n|^2\} = A_{SR}P_S + \sigma_0^2 \tag{2.50}$$

where σ_0^2 is the noise variance. The idea behind amplification factor is to equalise the effect of PL and noise, keeping the transmit power under the limit P_R at relay. Hence, the factor β can be designed as

$$\beta = \sqrt{\frac{P_R}{A_{SR}P_S + \sigma_0^2}} \tag{2.51}$$

Note that (2.51) contains the large-scale PL, which only depends on the distance between the source and relay, so (2.51) is called *fixed gain* amplification factor [73]. If the instantaneous CSI can be estimated by the relay, the factor can be redefined as

$$\beta = \sqrt{\frac{P_R}{|h_{SR}|^2P_S + \sigma_0^2}} \tag{2.52}$$

which is called *variable gain* amplification factor [2]. Then the symbols transmitted from the relay are $\{\beta y_R^n\}$.

In the second hop, h_{RD} is defined as the CIR of relay-destination link, and z_D is defined as the AWGN noise. At the destination node, the received symbols can be written as

$$\begin{aligned} y_D^n &= h_{RD}\beta y_R^n + z_D \\ &= h_{SR}h_{RD}\beta\sqrt{P_S}x_S^n + (h_{RD}\beta z_R + z_D) \end{aligned} \tag{2.53}$$

Note that the noise at relay is amplified by $h_{RD}\beta$. Although $h_{RD}\beta$ may be less than 1, the overall noise is still greater than z_D alone. After the symbols $\{y_D^n\}$ received, the coded symbols from the source can be detected at the destination as $\{\tilde{x}_S^n\}$, from which

the information bits can be decoded as $\{\tilde{b}_S^i\}$. Γ_{eq} is defined as the equivalent SNR of the AF relay channel, which is

$$\Gamma_{eq} = \frac{|h_{SR}|^2|h_{RD}|^2|\beta|^2 P_S}{(|h_{RD}|^2|\beta|^2+1)\,\sigma_0^2} = \frac{\frac{|h_{SR}|^2 P_S}{\sigma_0^2}\frac{|h_{RD}|^2 P_R}{\sigma_0^2}}{\frac{|h_{SR}|^2 P_S}{\sigma_0^2}+\frac{|h_{RD}|^2 P_R}{\sigma_0^2}+1} \tag{2.54}$$

If $\Gamma_{SR} = |h_{SR}|^2 P_S/\sigma_0^2$ and $\Gamma_{RD} = |h_{RD}|^2 P_R/\sigma_0^2$ are defined as the 'per-hop' SNRs in the source-relay and relay-destination links, respectively, (2.54) can be rewritten as

$$\Gamma_{eq} = \frac{\Gamma_{SR}\Gamma_{RD}}{\Gamma_{SR}+\Gamma_{RD}+1} \tag{2.55}$$

Therefore, the channel capacity is

$$C = \frac{1}{2}\log_2(1+\Gamma_{eq}) \tag{2.56}$$

where 1/2 indicates the half-duplex mode.

Since the relay's duty is to amplify the analogue waveforms and forward them to the destination, there are only handful components inside the node, such as LNA at receiving end, PA for transmission, demodulation/modulation block, *etc.* [25]. Hence, one of the advantages of AF relaying is low cost with simple structure. Due to no decoding involved, the relay node is not able to send retransmission request back to the source before forwarding messages. Therefore, the speed of forwarding is high without delay in decoding, which is another advantage of AF relaying. However, the noise at relay is amplified and forwarded to the destination as well, which causes the accumulated potential error for decoding. In other words, AF is suitable for low BER requirement, high delay tolerance in communication, such as voice transmission.

2.5.2 Decode-and-Forward Relaying

In the dual-hop DF relaying strategy, it is assumed that three nodes in the network share the same codebook. In the first hop, after receiving the symbols from the source, the relay node detects the coded symbols as $\{\tilde{x}_S^n\}$ from $\{y_R^n\}$ (same with (2.49)). Afterwards, the information bits $\{\tilde{b}_S^i\}$ are decoded from the detected symbols. If the information bits are correctly decoded, the relay uses $\{\tilde{b}_S^i\}$ to encode another set as $\{x_R^n\}$, and then transmits them to the destination. $\Gamma_{SR} = |h_{SR}|^2 P_S/\sigma_0^2$ is recalled as the received SNR at the relay, the channel capacity in the first hop can be written as

$$C_{SR} = \log_2(1+\Gamma_{SR}) \tag{2.57}$$

In the second hop, with the same definition of h_{RD}, P_R, and z_D as in AF relaying, the received signal at the destination can be defined as

$$y_D^n = h_{RD}\sqrt{P_R}x_R^n + z_D \tag{2.58}$$

After detection, the destination yields the estimated coded symbols $\{\tilde{x}_R^n\}$. As a result, the information bits can be decoded as $\{\check{b}_S^n\}$. The received SNR at the destination can be recalled as $\Gamma_{RD} = |h_{RD}|^2 P_R/\sigma_0^2$, thus the channel capacity of the second hop can be expressed as

$$C_{RD} = \log_2(1 + \Gamma_{RD}) \tag{2.59}$$

The overall channel capacity is determined by the minimum of C_{SR} and C_{RD} [2], which is

$$C = \frac{1}{2}\min\{C_{SR}, C_{RD}\} \tag{2.60}$$

where 1/2 is also the half-duplex factor.

Since the decoding is processed in the relay node, the error after decoding at the relay can be mitigated by a certain coding scheme or the retransmitted symbols from the source. Hence the relay can forward nearly original information, which largely reduces the end-to-end error. However, there is a need for at least another two components, analogue-to-digital (AD) convertor and decoding/encoding block, apart from other ones that AF relay has [25]. The cost therefore is higher than AF relay. On the other hand, decoding and retransmission need extra time, which causes delay in the communication. Consequently, DF relay is suitable for the high BER requirement, low delay tolerance in communication.

2.5.3 Relaying Protocols

Table 2.3: Types of relaying protocols

	Phase 1			Phase 2		
	Source	Relay	Destination	Source	Relay	Destination
MxF	Tx	Rx	-	-	Tx	Rx
SCxF	Tx	Rx	-	Tx	Tx	Rx
DxF	Tx	Rx	Rx	-	Tx	Rx
IxF	Tx	Rx+Tx	Rx	Tx	Tx+Rx	Rx

During the past a few years, different relaying strategies have been proposed and studied in order to improve the performance [2]. Some well-investigated protocols are concluded as below, where 'x' stands for 'amplify', 'compress', 'decode', *etc.*

- **Multi-hop x Forward (MxF)**: In the first hop, the source transmits, and only the relay listens. In the second hop, only the relay transmits, and the destination listens. If there are N relays, the number of hops are $N + 1$ between the source and destination. Dual-hop relaying is a special case of MxF.

- **Split-Combine x Forward (SCxF)**: In the first phase (timeslot), the source transmits and only the relay is listening. In the second phase, both source and relay transmits, and the destination listens.

- **Diversity x Forward (DxF)**: In the first phase, the source transmits, and both relay and destination are listening. In the second phase, only the relay transmits and the destination listens. Note that the destination can receive two 'copies' of the original information block. This protocol utilise the broadcast nature of wireless transmission.

- **Intersymbol Interference x Forward (IxF)**: If the relay node works in full-duplex mode, IxF is applicable. In the first hop, the source transmits one information block to the relay and destination. In the second phase, the relay transmits this block to destination, while receiving another block from the source. As a result, the destination is continuously listening, and in each phase hears a combination of the current block from the source, and the previous block from the relay simultaneously.

The above mentioned protocols can be summarised in Table 2.3. Throughout this thesis, MxF (dual-hop case) and DxF are studied in Chapters 4 and 5.

Chapter 3

Overview of Radio Resources and Their Allocation Schemes

This chapter presents a literature review on RA for broadband wireless communication systems. The radio resources are introduced in Section 3.1. The existing RA schemes are described in Section 3.2 for multi-carrier based communication systems.

3.1 Radio Resources in Wireless Communications

For any of the wireless communication systems mentioned in the Chapter 2, EM wave is the solely material to carry signals. Therefore, radio resources are usually referred to the following prime elements: bandwidth, carrier frequency, and transmit power.

3.1.1 Bandwidth and Carrier Frequency

The bandwidth of a communication system is the spectrum between the upper and lower cutoff frequencies. For a baseband signal, the bandwidth equals to the upper cut-off frequency. To avoid interferences suffering from the bandwidth overlap, the bandwidth of each communication standard and protocol is regulated by some forums such as International Telecommunication Union (ITU), European Conference of Postal and Telecommunications Administrations (CEPT), *etc.*. For instance, the GSM-Railway system in UK is licensed in two operating bands: 876-880 MHz for uplink, and 921-925 MHz for downlink [74]. While for an IMT system (UMTS-FDD system, regulated under Ofcom in UK), the uplink and downlink bands are 1,920-1,980 MHz and 2,110-2,170 MHz, respectively. However, the limited bandwidth in each standard may not satisfy the demands of data service from the growing number of customers in the future, and this concern is called 'spectrum scarcity' [75]. While, there are still a few bands which are not allocated to any commercial or governmental usage, such as 47-47.2 GHz, and 77.5-78 GHz [74]. Hence, some literatures put the focuses on the application of unlicensed bands [76] [77] [78] [79]. In [76], the authors analysed the performance of a modification of the multiple frequency shift system in the unlicensed 2.4 GHz band. To

avoid the interference with other sources and keep low power density, direct sequence and frequency hopping were considered in the application. In [77], a method was developed for examining the wireless services coexistence between wireless local area network (WLAN) and WPAN, where the closed-form solution for the probability of collision was derived. A spectrum sharing problem was studied in [78] for the coexistence of multiple systems, where the self-forcing protocol must correspond to an equilibrium of a game. It was proved that the repeated game is more appropriate than the one shot game to model the interaction. In [79], a unified framework was presented for interference characterisations in the unlicensed frequency bands, where a new spatial-spectral interference model was introduced. In this model, interferences can be any power spectral density and are distributed according to Poisson process in space and frequency domains. In Chapter 5, the unlicensed 60 GHz wireless communication in a short-range scenario is considered.

At the transmission end, the length of the antenna should be comparable with the wavelength in order to deliver the EM wave efficiently, *i.e.*, a quarter of the wavelength [80]. While in the early wireless communication system, human voice is the only information source needed to be transmitted. The spectrum of human voice in telephony system is usually limited between 300 and 3,000 Hz [81], so the minimum size of the antenna should be around 25 km, which is incredibly long. To solve this problem, the baseband signal needs to be up-converted to a radio frequency (RF) waveform [24], which allows the contained information to be transmitted out by small-sized antennas. Fig. 3.1 shows the relationship between the passband and its baseband of a complex modulated signal in the frequency domain, where B is the bandwidth, f_c is the central carrier frequency. Defining $X_b(f)$ as the frequency response of the baseband signal $x_b(t)$, the passband signal $x(t)$ [24] can be written as

$$x(t) = \sqrt{2}\Re\left[x_b(t)e^{j2\pi f_c t}\right] \tag{3.1}$$

where $\Re[\cdot]$ is the real part of a complex number, and $e^{j2\pi f_c t}$ is the carrier waveform. Applying Fourier transform $\mathfrak{F}\{\cdot\}$, the frequency response of $x(t)$ can be defined as

$$X(f) = \mathfrak{F}\{x(t)\} = \frac{1}{\sqrt{2}}\left\{X_b(f - f_c) + X_b^*(-f - f_c)\right\} \tag{3.2}$$

Since the frequency response of the carrier $e^{j2\pi f_c t}$ is an impulse function, $X(f)$ keeps the same shape of $X_b(f)$ when $f > 0$ as shown in Fig. 3.1. Hence, all the information of the source is preserved in the RF waveform $x(t)$ without any loss, which largely reduces the length of antenna. To avoid the distortion by the overlap of frequency bands between the positive and negative spectrum parts of the passband signal, the carrier frequency is required as $B < 2f_c$ at least. For instance, the central frequency of IMT standard in downlink is around 2,140 MHz, while the bandwidth is only 60 MHz.

If the transmit power and the distance are fixed, the data rate is determined by the bandwidth and central carrier frequency. According to (2.5), it is obvious that

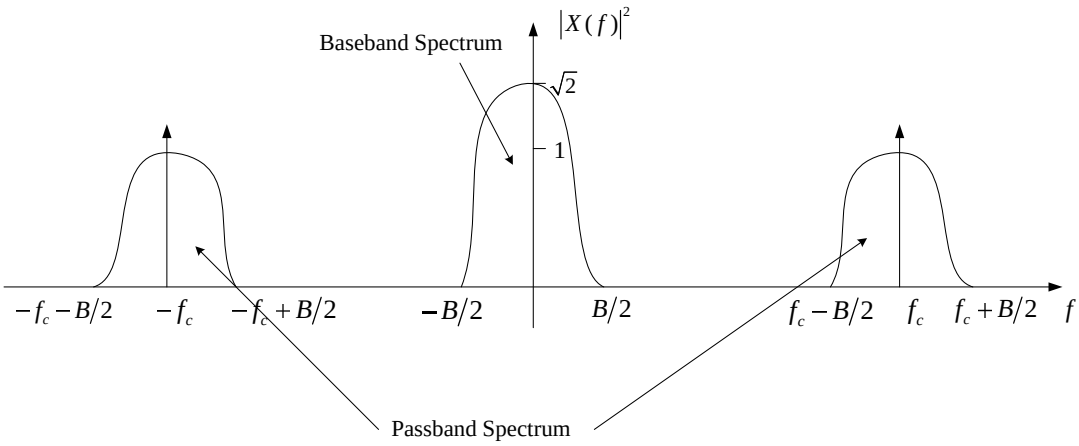

Figure 3.1: Illustration of the relationship between a passband spectrum and its baseband equivalent

the carrier central frequency affects the PL, and further affects the received SNR. The higher f_c is, the lower the received SNR will be. From (2.39) it is easy to observe that once the f_c is fixed, the data rate for transmission is mainly determined by the bandwidth. In other words, ultra high frequency transmission can provide ultra wide band, but it also brings serious attenuation on the strength of signals.

3.1.2 Transmit Power

In order to combat with the attenuation from PL and contamination from the noise, signals need to maintain a certain level of average power at the receiving end. Therefore, the transmit power should be large enough. From (2.39), it is clear that the channel capacity is in proportion to the transmit power. Ideally, the power is expected as large as possible. However, this assumption is not realistic in the practical scenario, since there are some limitations upon the transmit power.

At the transmitter, the output power of PA that is connected to the transmit antenna provides the power of RF passband signals. Usually, this output power equals to the average power of the passband waveform. Fig. 3.2 illustrates the amplification model of the PA with a fixed gain, where P_{in} is the input drive power, P_{out} is the output power, and P_{1dB} is the 1 dB compression point. It can be seen that there is a limit on the linear amplification of the PA. If P_{in} is greater than certain thresholds, PA will work in the nonlinear or saturation regions. If P_{in} is even greater than the output value of P_{sat}, it is meaningless to apply this PA. Hence, to achieve the maximum output power efficiency, PAs are usually operated near or in the saturation region [82]. However, when the waveform has a high PAPR, *i.e.*, the OFDM signal [47], the PA should work with low value of P_{in} to guarantee that the peaks stay away from the saturation region.

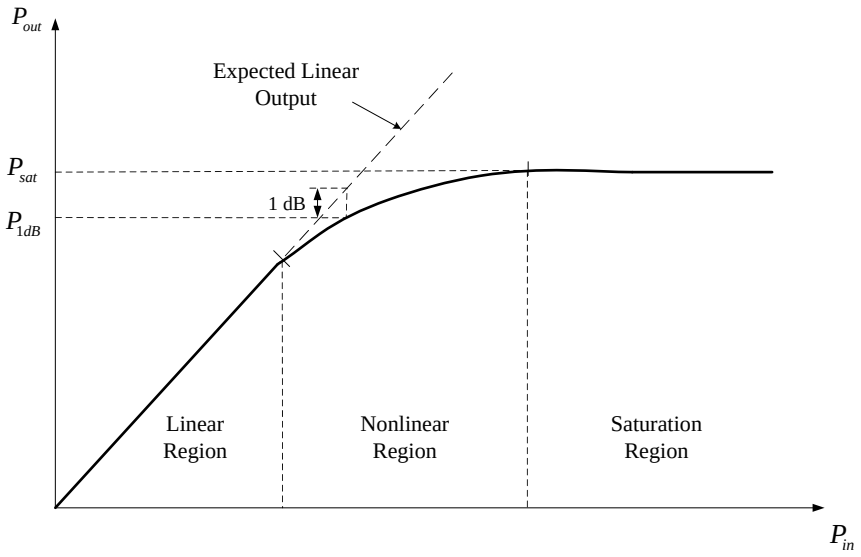

Figure 3.2: The amplification model of PA

Although this method can reduce the distortion of the whole waveform, the PA cannot provide sufficient transmit power, which affects the system performance in terms of data rate and coverage. Therefore, several methods were proposed to reduce the PAPR of OFDM signals [83] [84] [85] [86]. In [83], a PAPR reduction technique was developed using a hybrid scheme of the label-inserted encoder and soft amplitude limiter, which provides 5.5 dB PAPR reduction in OFDM systems. In [84], three PAPR reduction methods were proposed with a non-redundant linear precoding scheme in OFDM MIMO systems. Due to the incorporation between PAPR reduction and precoding, no spectral overhead was introduced compared to the conventional PAPR reduction schemes. A modified companding algorithm was proposed for the PAPR reduction in [85], which improves the BER and minimises the out-of-band interference while reducing PAPR effectively. In [86], a shift scheme of null sub-carriers amongst data sub-carriers was proposed with low complexity, which is compatible with the current commercial systems and can incorporate with most other PAPR reduction methods. Throughout this thesis, it is assumed that the PAs are operated in the linear region for both long-range and short-range communications.

Another limitation on transmit power is the regulation in communication standards. By evaluating the hardware, modulation methods, the QoS requirement, the interference between cells and the hazard of EM radiation comprehensively, proposers are likely to put limits on the transmit power of base stations and mobile devices. For instance, the transmit power of relay nodes in 802.16 standard varies from 9 dBm to 37 dBm with 2 dB intervals [55], which indicates the maximum transmit power is 37 dBm (5.09 W). The question is how to make full use of the transmit power to achieve

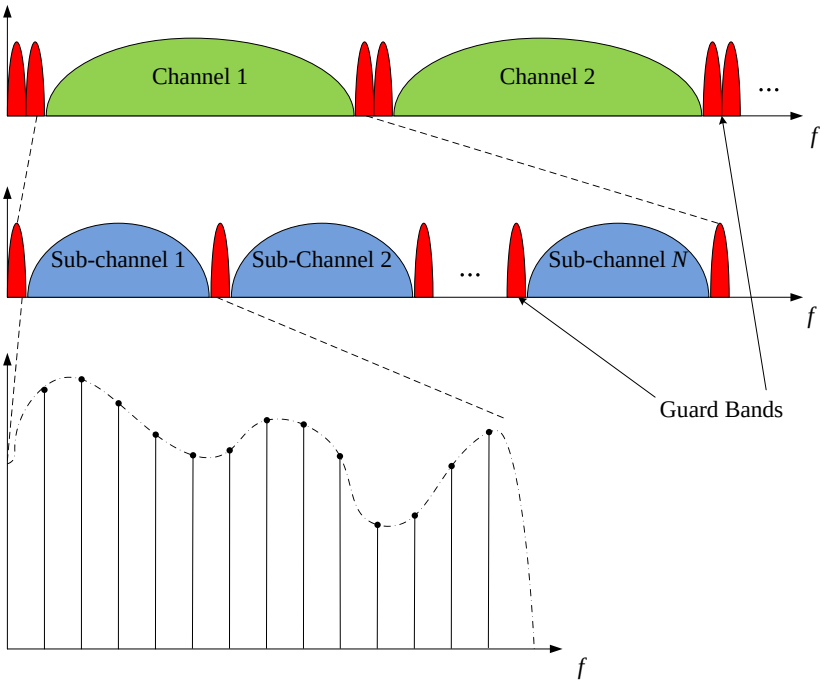

Figure 3.3: Illustration of the division of channels in a multi-carrier system, where f represents the frequency , and the red portions are guard bands

as high throughput as possible in green communications. In the following section, the problem of power allocation is discussed.

3.2 Resource Allocation in Multi-Carrier Based Systems

From Section 3.1 it is clear that the system performance is determined by the three prime radio resources, bandwidth, carrier central frequency, and transmit power. Since the values of these resources are regulated by different standards, RA throughout this thesis focuses on the multi-carrier based systems, where the frequency diversity and multiuser diversity are utilised to provide an enhanced system performance.

3.2.1 Sub-carrier Allocation

In this subsection, the SA techniques in multi-carrier systems are viewed. Due to the frequency selective fading property in a broadband channel, some parts of the spectrum are in deep fading while some parts enjoy good channel qualities [3]. To produce frequency diversity, multi-carrier systems, *i.e.*, FDMA and OFDMA, divide the whole channel into several sub-channels, and each of the sub-channels is divided into a number of narrow-band sub-carriers. Therefore, each of the sub-carrier can be treated as flat fading [2]. Fig. 3.3 illustrates the spectrum of the division of channels in a multi-carrier

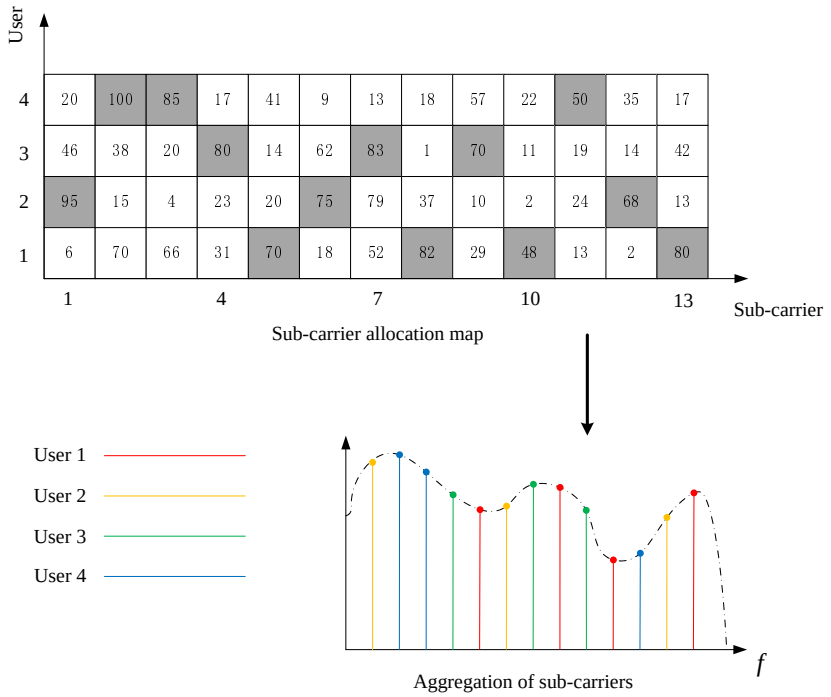

Figure 3.4: An example of sub-carrier allocation, where the numbers in the cube are the relative values of SNR, and the grey parts indicate the maximum SNR amongst users

system, where the sub-carriers experience different degrees of fading. The exploration of the frequency diversity leads to a high degree of freedom for the transmission, which results in a better performance. For instance, employing the suitable modulation type in each frequency component helps with increasing the system capacity meanwhile decreasing the BER. In [30], the variable rate and power MQAM modulation scheme was proposed for a single-user wireless communication system. The RA for single-user OFDM systems was investigated in [87].

Another property of wireless communication channels is time variation, which refers to the fading changes over time [24]. Due to this fact, the performance of systems may vary in a wide range with time periods [88]. In multiuser wireless communication systems, the channels are independent across different users. For instance, in period 1, a channel is in deep fading for user 1, while it is in good condition for user 2; in period 2, the case is just the opposite. Therefore, the system can always choose the user with a higher SNR (or CNR) to transmit, which leads to a higher system capacity. Carrier aggregation is one of the most distinct features in the future network such as 3GPP LTE-Advanced standard [89], which allows several continuous or non-continuous carriers to be aggregated to support high data rate transmission over wide bandwidth [90]. However, the aggregation of large portion of continuous carriers is hardly available,

so the aggregation of multiple non-continuous carriers has become popular in practice [91]. Hence, it is efficient to allocate a number of sub-carriers to different users. In Fig. 3.4, an example of SA was illustrated. As shown, the system obtains the CSI of sub-carriers from all users first, which forms a SA map. Afterwards, each sub-carrier is assigned to the user with highest SNR amongst all in one transmission period, and all these sub-carriers are aggregated as one sub-channel together. Across the whole time periods, each user can get a different number of sub-carriers according to the channel conditions. Intuitively, when the number of users increases, the SNR for transmission also augments, since there is more degrees of freedom due to scheduling of different users. Hence, the system throughput benefits more from the multiuser diversity with the increase of the number of users.

3.2.2 Power Allocation

In broadband wireless communication systems, since the system bandwidth is larger than the channel coherence bandwidth, different frequency bands usually have different channel characteristics, such as shadowing, short term Rayleigh fading and noise [2]. Hence, uniform power allocation for the whole bandwidth will cause degradations of the system throughput BER. When multi-carrier systems, *i.e.*, OFDMA, are applied, power on each sub-carrier can vary [92] [87], since all sub-carriers are mutually independent. The adaptive power [93] on each sub-carrier enhances the flexibility, which provides better data rate and robustness of data transmission [94].

For power allocation, the water-filling algorithm [95] is the most widely applied in RA schemes. The basic idea of the water-filling can be illustrated in Fig. 3.5: limited volume of water is poured into a pool with the bottom that has different level, where the blue parts (water) denote the powers on each sub-carrier, and the green bars (bottom) are determined by the reciprocal of CNR. As shown, the system can set a unified 'water level' Λ after receiving the CSI from the destination, and then calculate the value of transmit power on each sub-carrier. Λ usually denotes the transmit power budget that is determined by the limit of PA. The larger the total transmit power is, the higher the water level will be. Defining γ_n as the CNR of sub-carrier n, the allocated power p_n can be calculated as

$$p_n = \left[\Lambda - \frac{1}{\gamma_n}\right]^+, \ n = 0, ..., N \tag{3.3}$$

where $[\cdot]^+$ rounds the negative values up to zero. Fig. 3.5 shows the water-filling algorithm in a 9-carrier system. It can be observed that when the channel condition is good, *i.e.*, for sub-carrier 8, the allocated power is high, while when sub-carrier is in bad condition, *i.e.*, for sub-carrier 3, little or no power is allocated. In a multiuser system, the water-filling algorithm is usually jointly applied with certain SA schemes [4] [13]. In [4], the optimal SA and power allocation schemes were proposed for OFDM systems. By allocating each sub-carrier to the user with the best channel quality on it, the power is

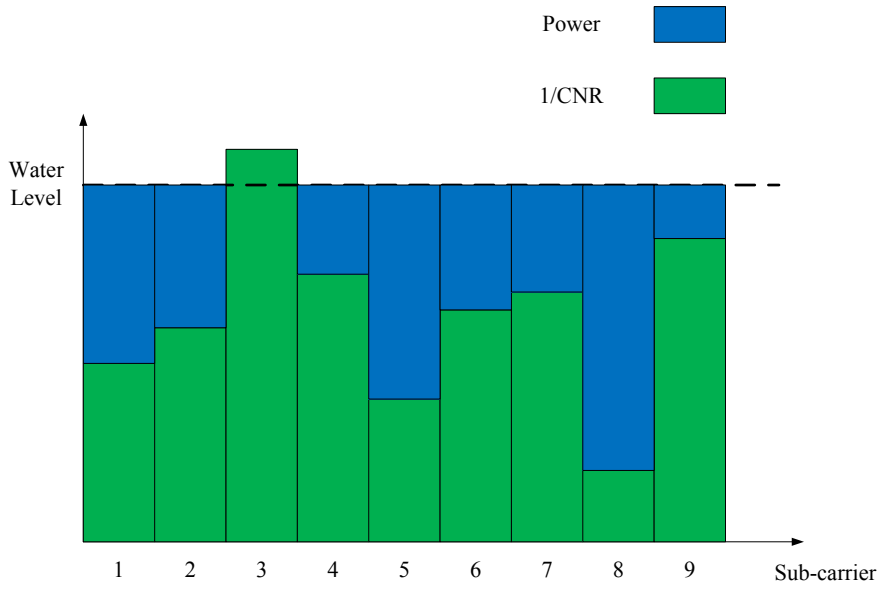

Figure 3.5: Illustration of the water-filling algorithm

allocated by the water-filling algorithm, which provides the maximum system throughput. In [13], to guarantee users' proportional data rates with fairness-awareness while maximising the system throughput, suboptimal SA and power allocation algorithms were proposed.

From the literature in this section, RA plays an important part in multiuser OFDM systems for throughput enhancement. Throughout this thesis, Chapter 4 focuses on the multiuser multi-carrier OFDM systems, while Chapter 5 studies the P2P communication with single carrier. Consequently, both water-filling power allocation and sub-carrier allocation schemes are discussed in Chapter 4, but only power allocation is studied and applied in the systems of Chapter 5 with searching method.

Chapter 4

Asymmetric Resource Allocation for Long-Range Multi-Relay Based OFDMA Systems

Relay technologies, including AF [10] and DF [11], have been considered as key architectures of the 4G communication systems such as LTE-Advanced systems [12]. The relay networks are designed to enlarge the coverage of the base station, as well to satisfy the data rate requirement of the users at the cell-edge. Meanwhile, relay networks can be easily adapted into small-cell layouts, which can help the system enhance its energy efficiency [8] and reduce the CO_2 emission.

Adaptive RA [13] plays an essential part in relay systems. In [96], sub-carrier pairing and power allocation were conducted successively in a single-relay system for both AF and DF relaying to maximise the system capacity. In [97], the optimal joint sub-carrier matching and power allocation were proposed for DF relay systems with single destination. However, it requires the bits to be grouped in relaying, moreover, the complexity of sub-carrier pairing is too high to guarantee that the channel state remains unchanged during this process. In [98] an optimal RA algorithm was proposed to maximise the bandwidth efficiency of the DF relay system, however, it assumes fixed geometric locations of the relays, which is merely practical in the real urban environment. While in [21], the authors considered a multi-relay system with random relay locations, and also proposed an optimal RA algorithm. However, in most previous work, the time slot durations for the two hops were designed to be symmetric. In [99], a two-way relay channel model was investigated, and an asymmetric time allocation scheme was proposed, where the two consecutive time slot durations were designed to be asymmetric. However, it is only assumed that a uniform power allocation is conducted across all sub-carriers, which is not efficient. In [20], an optimal ARA scheme was proposed for an OFDM based DF relay system, however, it can only be applied to a scenario where a cell is divided into multiple single-relay sub-systems with RA conducted in each of individual sub-systems independently. Hence, the ARA scheme is

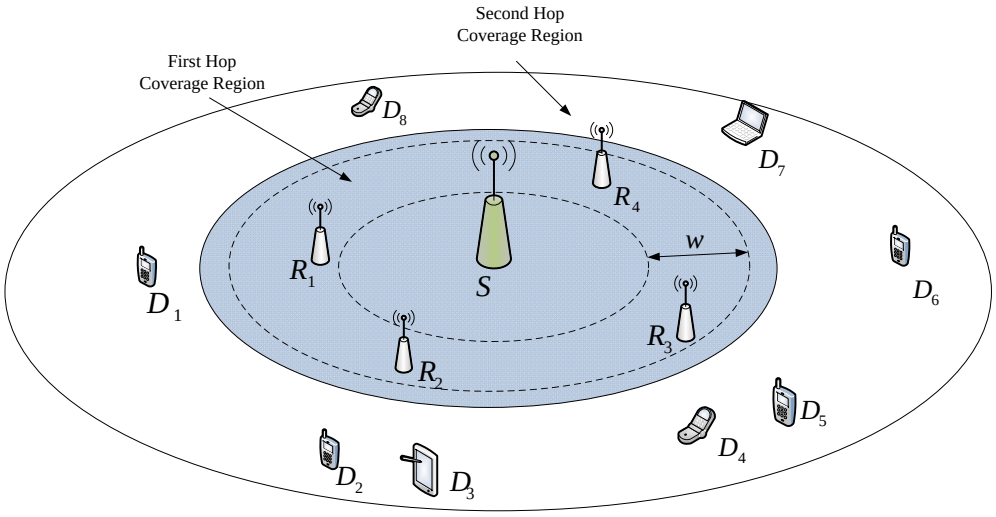

Figure 4.1: A multi-relay based OFDMA cellular system

not readily extendable to a multi-relay multi-destination system.

In this chapter, an ARA scheme for DF dual-hop multi-relay based OFDMA systems is investigated. This work is different in the following aspects. First, this is the first work to apply asymmetric time allocation to a general multi-relay multi-destination system, where the time slots for the two hops via each of the relays are designed to be asymmetric, *i.e.*, with K relays in the cell, a total of $2K$ time slots may be of different durations. As a result, it enhances the degree of freedom for transmission over the previous symmetric RA (SRA) scheme. Second, an optimal algorithm is proposed to perform joint time, power and sub-carrier allocation to obtain the global optimal results, with only limited amount of feedback information from relays and destinations. The CoMP technique [43] that allows multiple relays to serve single destination is applied at relays, which enhances the degrees of freedom. Numerical results show that, thanks to the multi-time and multi-relay diversities, the proposed ARA scheme outperforms the SRA algorithm in [21], as well as the ARA algorithm in [20], with higher achievable system throughput, especially when the relays are relatively close to the source. Moreover, the impact of the relay locations on the results of asymmetric time allocation is demonstrated.

The rest of this chapter is organised as the follows. Section 4.1 presents the system model, and Section 4.2 describes the problem formulation. In Section 4.3, the optimal ARA algorithm is proposed. Simulation results are shown in Section 4.4, and the summary of this chapter is presented in Section 4.5.

4.1 System Model

As illustrated in Fig. 4.1, a DF relay based OFDMA cellular system is considered in the downlink, with two hops and K relays. In Fig. 4.1, S denotes the source node (BS), $\{D_l\}$ $(l = 1, ..., L)$ denote the different destination nodes (mobile users), and $\{R_k\}$ $(k = 1, ..., K)$ denote the multiple relay nodes. The K relays are uniformly distributed within the range of width w in Fig. 4.1 (the width between two dashed circles). Single-cell scenario is analysed through this chapter. For the resource allocation, single-cell system can be regarded as an individual sub-section of a multi-cell system [100]. Also, it is easy to derive the optimal results of resources under the single-cell system assumption. However, the inter-cell interference caused by the frequency reuse in a multi-cell network is not considered. It is assumed that all relays are connected with the BS by optical fibers for information exchange in backhaul. d_{SR} is defined as the distance from the source to the mid position between the two dashed circles, and d_k is defined as the distance between the source and relay k. Without the loss of generality, it is assumed that all the destinations have the approximately same distance from the source, which is denoted by d_{SD}, and they are uniformly distributed along a circle. This assumption allows the system throughput to be examined as a matter of distance from BS. In other words, the average throughput of one user can be estimated by the values of d_{SD} and L. However, the simultaneous throughput of one user cannot be given since not all users are located on the same circle at the same time. Similar to d_k, $d_{k,l}$ denotes the distance between the relay k to the destination l. The hop from source to relays is the first hop, and the hop from the relays to destinations is the second hop. Now $\{1, ..., N\}$ is defined as the orthogonal sub-carriers set, where N is the total number of sub-carriers. The total transmit time duration is denoted as T. The time slot durations via the relay k for the first and the second hops are $T_{1,k}$ and $T_{2,k}$, respectively.

OFDMA technique allows symbols for different users to be modulated on different sub-carriers, and each sub-carrier experiences flat fading for those symbols, *i.e.*, (2.45) in Section 2.4. In the first hop, Assuming perfect channel estimation at both relay nodes and destinations, $H_{1,k}^{(n)}$ is defined as the frequency domain CIR on sub-carrier n between the source and relay k, which follows the complex Gaussian distribution as $H_{1,k}^{(n)} \sim CN(0, A_{1,k}(d_k))$, where $A_{1,k}$ is the PL over distance d_k. The average power of $H_{1,k}^{(n)}$ is written as $\mathbb{E}\{|H_{1,k}^{(n)}|^2\} = A_{1,k}$. Defining $p_{1,k}^{(n)}$ as the transmit power for the source to relay k on sub-carrier n, $x_{1,k}^{(n)}$ as the symbols from the source to relay k on sub-carrier n, $Z_k^{(n)}$ as the noise on relay k, the received symbols $Y_k^{(n)}$ on sub-carrier n is defined as

$$Y_k^{(n)} = H_{1,k}^{(n)} \sqrt{p_{1,k}^{(n)}} x_{1,k}^{(n)} + Z_k^{(n)} \tag{4.1}$$

Hence, The instantaneous CNR is $\gamma_{1,k}^{(n)} = N|H_{1,k}^{(n)}|^2/(BN_0)$ in the first hop, where N_0 is defined as the single-sided power spectral density of the AWGN. The bandwidth

for each sub-carrier is B/N. The maximum achievable throughput for relay k can be written as the summation of the throughput from the allocated sub-carriers as:

$$C_{1,k} = \frac{B}{N} \sum_{n=1}^{N} \xi_{1,k}^{(n)} \log_2 \left(1 + p_{1,k}^{(n)} \gamma_{1,k}^{(n)}\right) \tag{4.2}$$

where $\xi_{1,k}^{(n)} \in \{0, 1\}$ denotes SA indicator for sub-carrier n in the first hop. If sub-carrier n is allocated to relay k, $\xi_{1,k}^{(n)} = 1$, otherwise, $\xi_{1,k}^{(n)} = 0$.

Once receiving the symbols from the source, relays decode these codewords first, and then encode them with the same or a different code-book. Afterwards, relays need to regroup these codewords symbols according to the sub-carrier allocation information for different destinations. To improve the throughput of the users which are on the cell-edge, joint transmission [43] scheme of CoMP is applied in the system. If one user is on the coverage edge of two relays, it can be served by these two relays simultaneously. Here $x_{2,k,l}^{(n)}$ is defined as the symbol transmitted from relay k to the destination l on sub-carrier n. In the second hop, $H_{2,k,l}^{(n)}$ is defined as the frequency domain CIR between relay k and destination l, which follows the complex Gaussian distribution as $H_{2,k,l}^{(n)} \sim CN(0, A_{2,k,l}(d_{k,l}))$, where $A_{2,k,l}$ is the PL over distance $d_{k,l}$. The average power of $H_{2,k,l}^{(n)}$ thus can be written as $\mathbb{E}\{|H_{2,k,l}^{(n)}|^2\} = A_{2,k,l}$. Similar to $p_{1,k}^{(n)}$, $p_{2,k,l}^{(n)}$ is defined as the transmit power for the transmission between relay k and destination l on sub-carrier n. For simplicity, it is assumed that there is no intra-cell interference between the transmission nodes. Defining $Z_l^{(n)}$ as the noise at the destination l, the received symbols $Y_l^{(n)}$ on sub-carrier n is defined as

$$Y_l^{(n)} = H_{2,k,l}^{(n)} \sqrt{p_{2,k,l}^{(n)}} x_{2,k,l}^{(n)} + Z_l^{(n)} \tag{4.3}$$

The instantaneous CNR is $\gamma_{2,k,l}^{(n)} = N|H_{2,k,l}^{(n)}|^2/(BN_0)$ for sub-carrier n. The maximum achievable throughput of the second hop thus can be written as:

$$C_{2,k} = \frac{B}{N} \sum_{l=1}^{L} \sum_{n=1}^{N} \xi_{2,k,l}^{(n)} \log_2 \left(1 + p_{2,k,l}^{(n)} \gamma_{2,k,l}^{(n)}\right) \tag{4.4}$$

Similarly, $\xi_{2,k,l}^{(n)}$ is denoted as the SA indicator in the second hop.

The system allows the transmission in the first hop to take up $T_{1,k}/T$ of time duration, and the transmission in the second hop to take up $T_{2,k}/T$, where $T = T_{1,k} + T_{2,k}$. Note that $T_{1,k}$ and $T_{2,k}$ are not necessarily equal. As a result, the overall maximum achievable throughput via relay k can be written as:

$$C_k = \frac{1}{2} \min \left\{ \frac{T_{1,k}}{T} C_{1,k}, \frac{T_{2,k}}{T} C_{2,k} \right\} \tag{4.5}$$

where $1/2$ is the half-duplex factor. Hence, the overall system maximum throughput can be written as

$$C = \sum_{k=1}^{K} C_k \tag{4.6}$$

4.2 Problem Formulation

In this section, the RA problem of the investigated system is formulated. Now define P_T as the total transmit power of the system in the downlink, P_S as the transmit power of the source node, and P_R as the total transmit power of all the relay nodes. It is also assumed that each sub-carrier is occupied by only one node in a time slot. Unlike SRA in [21], where the time slot durations for the two hops are identical, the proposed ARA scheme allows all the $2K$ time slots to be of different durations. Also, it allows a certain sub-carrier to be allocated to any of the K relays in the first hop, while in [20], N sub-carriers are divided into K groups and each group serves a single-relay sub-system. In the second hop, the proposed system allows multiple relays to serve one destination, which is different from the multiple single-relay sub-systems, where each destination can only be served by the single relay by CoMP technique. The objective is to maximise the achievable system throughput, which is formulated by:

$$\max_{p_{1,k}^{(n)},\, p_{2,k,l}^{(n)},\, T_{1,k},\, T_{2,k}} C \tag{4.7}$$

subject to:

$$T_{1,k} + T_{2,k} = T,\ (\forall k) \tag{4.8}$$

$$\sum_{k=1}^{K}\sum_{n=1}^{N} p_{1,k}^{(n)} \le P_S,\ \sum_{k=1}^{K}\sum_{l=1}^{L}\sum_{n=1}^{N} p_{2,k,l}^{(n)} \le P_R,\ \text{where } p_{1,k}^{(n)} \ge 0,\ p_{2,k,l}^{(n)} \ge 0,\ (\forall k,l,n) \tag{4.9}$$

$$P_S + P_R = P_T \tag{4.10}$$

$$\sum_{k=1}^{K} \xi_{1,k}^{(n)} = 1,\ \sum_{k=1}^{K}\sum_{l=1}^{L} \xi_{2,k,l}^{(n)} = 1,\ (\forall n) \tag{4.11}$$

Constraint (4.8) implies that the time slot durations for both hops via each relay may be asymmetric, while in [21], they are fixed to be equal, *i.e.*, $T_{1,k} = T_{2,k} = T/2$ $(\forall k)$. Constraint (4.11) indicates the exclusivity of sub-carriers allocated to nodes. Constraint (4.9) and (4.10) limit the transmit power of BS and relay nodes. Note that the time delay tolerance and the minimum data rate are not set for each individual destination as QoS requirements, because the proposed ARA algorithm focuses on the system throughput maximisation. On the other hand, the derivation of closed-form results of resources is not easy to be conducted with these two constraints [20].

4.3 Asymmetric RA Algorithm

In this section, both optimal and suboptimal algorithms for ARA are proposed. In Subsection 4.3.1, the optimal ARA algorithm jointly allocated the sub-carrier, power and time duration with an iterative method, where not all relays are necessarily involved in forwarding messages in each time slot. While in Subsection 4.3.2, the SA is conducted first, and afterwards the power and time are allocated using the closed-form results.

4.3.1 Optimal ARA Algorithm

In this subsection, the optimal solution to the objective function (4.7) is proposed. It can be derived that (4.7) is optimised when $T_{1,k}C_{1,k} = T_{2,k}C_{2,k}$ $(\forall k)$ [20]. Also using (4.6), (4.7) can be rewritten as

$$\max_{p_{1,k}^{(n)},\, p_{2,k,l}^{(n)},\, T_{1,k},\, T_{2,k}} \sum_{k=1}^{K} \frac{1}{2}\left(\frac{T_{1,k}}{T}C_{1,k} + \frac{T_{2,k}}{T}C_{2,k}\right) \tag{4.12}$$

It can be shown in the Appendix B that (4.12) is a convex function. Thus (4.12) can be solved using the Karush-Kuhn-Tucker (KKT) conditions [101]. Using (4.12), and (4.8)-(4.11), the Lagrangian function can be built as,

$$\begin{aligned}
L = &\sum_{k=1}^{K}\frac{1}{2}\left[\frac{T_{1,k}}{T}\sum_{n=1}^{N}\xi_{1,k}^{(n)}\log_2\left(1+p_{1,k}^{(n)}\gamma_{1,k}^{(n)}\right) + \frac{T_{2,k}}{T}\sum_{l=1}^{L}\sum_{n=1}^{N}\xi_{2,k,l}^{(n)}\log_2\left(1+p_{2,k,l}^{(n)}\gamma_{2,k,l}^{(n)}\right)\right] \\
&- \sum_{k=1}^{K}\lambda_k\left[\frac{T_{1,k}}{T}\sum_{n=1}^{N}\xi_{1,k}^{(n)}\log_2\left(1+p_{1,k}^{(n)}\gamma_{1,k}^{(n)}\right) - \frac{T_{2,k}}{T}\sum_{l=1}^{L}\sum_{n=1}^{N}\xi_{2,k,l}^{(n)}\log_2\left(1+p_{2,k,l}^{(n)}\gamma_{2,k,l}^{(n)}\right)\right] \\
&- \mu_1\left(\sum_{k=1}^{K}\sum_{n=1}^{N}p_{1,k}^{(n)} - P_S\right) - \mu_2\left(\sum_{k=1}^{K}\sum_{l=1}^{L}\sum_{n=1}^{N}p_{2,k,l}^{(n)} - P_R\right) - \sum_{k=1}^{K}\sigma_k(T_{1,k}+T_{2,k}-T) \\
&- \sum_{n=1}^{N}\delta_1^{(n)}\left(\sum_{k=1}^{K}\xi_{1,k}^{(n)} - 1\right) - \sum_{n=1}^{N}\delta_2^{(n)}\left(\sum_{k=1}^{K}\sum_{l=1}^{L}\xi_{2,k,l}^{(n)} - 1\right)
\end{aligned} \tag{4.13}$$

where constant B/N is omitted in order to simplify the derivation. λ_k $(k = 1, ..., K)$, μ_1, μ_2, σ_k $(k = 1, ..., K)$, $\delta_1^{(n)}$, and $\delta_2^{(n)}$ $(n = 1, ..., N)$ are defined as Lagrange multipliers.

Assuming λ_k^*, μ_1^*, μ_2^*, and σ_k^* are the optimal Lagrange multipliers, the resulting optimal power and time slot durations can be derived from $\partial L/\partial p_{1,k}^{*(n)} = 0$, $\partial L/\partial p_{2,k,l}^{*(n)} = 0$, $\partial L/\partial T_{1,k}^* = 0$, and $\partial L/\partial T_{2,k}^* = 0$, respectively, which are:

$$\begin{aligned}
&\left.\frac{\partial L}{\partial p_{1,k}^{(n)}}\right|_{T_{1,k}=T_{1,k}^*, T_{2,k}=T_{2,k}^*, \xi_{1,k}^{(n)}=\xi_{1,k}^{*(n)}, \xi_{2,k,l}^{(n)}=\xi_{2,k,l}^{*(n)}, p_{1,k}^{(n)}=p_{1,k}^{*(n)}, p_{2,k,l}^{(n)}=p_{2,k,l}^{*(n)}} \\
&= \left(\frac{1}{2} - \lambda_k^*\right)\frac{T_{1,k}^*\xi_{1,k}^{*(n)}\gamma_{1,k}^{(n)}}{\ln 2T\left(1+p_{1,k}^{*(n)}\gamma_{1,k}^{(n)}\right)} - \mu_1^* \\
&\begin{cases} < 0, & p_{1,k}^{*(n)} = 0; \\ = 0, & p_{1,k}^{*(n)} > 0 \end{cases}
\end{aligned} \tag{4.14}$$

$$\left.\frac{\partial L}{\partial p_{2,l,k}^{(n)}}\right|_{T_{1,k}=T_{1,k}^*,T_{2,k}=T_{2,k}^*,\xi_{1,k}^{(n)}=\xi_{1,k}^{*(n)},\xi_{2,k,l}^{(n)}=\xi_{2,k,l}^{*(n)},p_{1,k}^{(n)}=p_{1,k}^{*(n)},p_{2,k,l}^{(n)}=p_{2,k,l}^{*(n)}}$$

$$= \left(\frac{1}{2}+\lambda_k^*\right)\frac{T_{2,k}^*\xi_{2,k,l}^{*(n)}\gamma_{2,k,l}^{(n)}}{\ln 2T\left(1+p_{2,k,l}^{*(n)}\gamma_{2,k,l}^{(n)}\right)}-\mu_2^* \tag{4.15}$$

$$\begin{cases} <0, & p_{2,k,l}^{*(n)}=0; \\ =0, & p_{2,k,l}^{*(n)}>0 \end{cases}$$

$$\left.\frac{\partial L}{\partial T_{1,k}}\right|_{T_{1,k}=T_{1,k}^*,T_{2,k}=T_{2,k}^*,\xi_{1,k}^{(n)}=\xi_{1,k}^{*(n)},\xi_{2,k,l}^{(n)}=\xi_{2,k,l}^{*(n)},p_{1,k}^{(n)}=p_{1,k}^{*(n)},p_{2,k,l}^{(n)}=p_{2,k,l}^{*(n)}}$$

$$= \frac{\left(\frac{1}{2}-\lambda_k^*\right)}{T}\sum_{n=1}^{N}\xi_{1,k}^{*(n)}\log_2\left(1+p_{1,k}^{*(n)}\gamma_{1,k}^{(n)}\right)-\sigma_k^* \tag{4.16}$$

$$\begin{cases} <0, & T_{1,k}^*=0; \\ =0, & T_{1,k}^*>0 \end{cases}$$

and

$$\left.\frac{\partial L}{\partial T_{2,k}}\right|_{T_{1,k}=T_{1,k}^*,T_{2,k}=T_{2,k}^*,\xi_{1,k}^{(n)}=\xi_{1,k}^{*(n)},\xi_{2,k,l}^{(n)}=\xi_{2,k,l}^{*(n)},p_{1,k}^{(n)}=p_{1,k}^{*(n)},p_{2,k,l}^{(n)}=p_{2,k,l}^{*(n)}}$$

$$= \frac{\left(\frac{1}{2}+\lambda_k^*\right)}{T}\sum_{l=1}^{L}\sum_{n=1}^{N}\xi_{2,k,l}^{*(n)}\log_2\left(1+p_{2,k,l}^{*(n)}\gamma_{2,k,l}^{(n)}\right)-\sigma_k^* \tag{4.17}$$

$$\begin{cases} <0, & T_{2,k}^*=0; \\ =0, & T_{2,k}^*>0 \end{cases}$$

By applying $T_{1,k}C_{1,k} = T_{2,k}C_{2,k}$ $(\forall k)$, (4.16) and (4.17), the time duration can be derived as:

$$T_{1,k}^* = \left(\frac{1}{2}-\lambda_k^*\right)T \tag{4.18}$$

$$T_{2,k}^* = \left(\frac{1}{2}+\lambda_k^*\right)T \tag{4.19}$$

By substituting (4.18) and (4.19) into (4.14) and (4.15), respectively, the transmit power can be expressed as:

$$p_{1,k}^{*(n)} = \xi_{1,k}^{*(n)}\left[\frac{\left(\frac{1}{2}-\lambda_k^*\right)^2}{\mu_1^*\ln 2}-\frac{1}{\gamma_{1,k}^{(n)}}\right]^+ \tag{4.20}$$

$$p_{2,k,l}^{*(n)} = \xi_{2,k,l}^{*(n)}\left[\frac{\left(\frac{1}{2}+\lambda_k^*\right)^2}{\mu_2^*\ln 2}-\frac{1}{\gamma_{2,k,l}^{(n)}}\right]^+ \tag{4.21}$$

where $[\cdot]^+ = \max\{\cdot, 0\}$.

Since $\xi_{1,k}^{(n)}$ and $\xi_{2,k,l}^{(n)}$ are binary indicators, to apply KKT condition, both $\xi_{1,k}^{(n)}$ and $\xi_{2,k,l}^{(n)}$ need to be relaxed as real values [98]. Thus, there are

$$\begin{aligned}
&\left.\frac{\partial L}{\partial \xi_{1,k}^{(n)}}\right|_{T_{1,k}=T_{1,k}^*,T_{2,k}=T_{2,k}^*,\xi_{1,k}^{(n)}=\xi_{1,k}^{*(n)},\xi_{2,k,l}^{(n)}=\xi_{2,k,l}^{*(n)},p_{1,k}^{(n)}=p_{1,k}^{*(n)},p_{2,k,l}^{(n)}=p_{2,k,l}^{*(n)}} \\
&= \left(\frac{1}{2}-\lambda_k^*\right)\frac{T_{1,k}^*}{T}\left[\log_2\left(1+p_{1,k}^{*(n)}\gamma_{1,k}^{(n)}\right)-\frac{p_{1,k}^{*(n)}\gamma_{1,k}^{(n)}}{\ln 2\left(1+p_{1,k}^{*(n)}\gamma_{1,k}^{(n)}\right)}\right]-\delta_1^{*(n)} \\
&= \left(\frac{1}{2}-\lambda_k^*\right)^2\log_2\left[\frac{\left(\frac{1}{2}-\lambda_k^*\right)^2}{\mu_1^*\ln 2}\right]-\left[\frac{\left(\frac{1}{2}-\lambda_k^*\right)^2}{\ln 2}-\frac{\mu_1^*}{\gamma_{1,k}^{(n)}}\right]^+-\delta_1^{*(n)} \\
&\begin{cases} <0, & \xi_{1,k}^{*(n)}=0; \\ =0, & 0<\xi_{1,k}^{*(n)}<1; \\ >0, & \xi_{1,k}^{*(n)}=1 \end{cases}
\end{aligned} \tag{4.22}$$

and

$$\begin{aligned}
&\left.\frac{\partial L}{\partial \xi_{2,k,l}^{(n)}}\right|_{T_{1,k}=T_{1,k}^*,T_{2,k}=T_{2,k}^*,\xi_{1,k}^{(n)}=\xi_{1,k}^{*(n)},\xi_{2,k,l}^{(n)}=\xi_{2,k,l}^{*(n)},p_{1,k}^{(n)}=p_{1,k}^{*(n)},p_{2,k,l}^{(n)}=p_{2,k,l}^{*(n)}} \\
&= \left(\frac{1}{2}+\lambda_k^*\right)\frac{T_{2,k}^*}{T}\left[\log_2\left(1+p_{2,k,l}^{*(n)}\gamma_{2,k,l}^{(n)}\right)-\frac{p_{2,k,l}^{*(n)}\gamma_{2,k,l}^{(n)}}{\ln 2\left(1+p_{2,k,l}^{*(n)}\gamma_{2,k,l}^{(n)}\right)}\right]-\delta_2^{*(n)} \\
&= \left(\frac{1}{2}+\lambda_k^*\right)^2\log_2\left[\frac{\left(\frac{1}{2}+\lambda_k^*\right)^2}{\mu_2^*\ln 2}\right]-\left[\frac{\left(\frac{1}{2}+\lambda_k^*\right)^2}{\ln 2}-\frac{\mu_2^*}{\gamma_{2,k,l}^{(n)}}\right]^+-\delta_2^{*(n)} \\
&\begin{cases} <0, & \xi_{2,k,l}^{*(n)}=0; \\ =0, & 0<\xi_{2,k,l}^{*(n)}<1; \\ >0, & \xi_{2,k,l}^{*(n)}=1 \end{cases}
\end{aligned} \tag{4.23}$$

When (4.22) and (4.23) equal to zero, $\delta_1^{(n)}$ and $\delta_2^{(n)}$ can be viewed as functions of other Lagrange multipliers for each sub-carrier, which can be expressed as:

$$\delta_1^{(n)}\left(\lambda_k,\mu_1\right)=\left(\frac{1}{2}-\lambda_k\right)^2\log_2\left[\frac{\left(\frac{1}{2}-\lambda_k\right)^2}{\mu_1\ln 2}\right]-\left[\frac{\left(\frac{1}{2}-\lambda_k\right)^2}{\ln 2}-\frac{\mu_1}{\gamma_{1,k}^{(n)}}\right]^+,\ (\forall k) \tag{4.24}$$

$$\delta_2^{(n)}\left(\lambda_k,\mu_2\right)=\left(\frac{1}{2}+\lambda_k\right)^2\log_2\left[\frac{\left(\frac{1}{2}+\lambda_k\right)^2}{\mu_2\ln 2}\right]-\left[\frac{\left(\frac{1}{2}+\lambda_k\right)^2}{\ln 2}-\frac{\mu_2}{\gamma_{2,k,l}^{(n)}}\right]^+,\ (\forall k,l) \tag{4.25}$$

Hence, $\delta_1^{*(n)}$ and $\delta_2^{*(n)}$ indicate the outputs from the optimal input Lagrange multipliers.

Each sub-carrier is allocated to one node exclusively. In the first hop, the sub-carrier

n is allocated to relay node k^* when k^* satisfies the following criteria as:

$$\begin{aligned} k^* &= \arg\max_k \delta_1^{(n)}(\lambda_k, \mu_1) \\ &= \arg\max_k \left(\frac{1}{2} - \lambda_k\right)^2 \log_2\left[\frac{\left(\frac{1}{2}-\lambda_k\right)^2}{\mu_1 \ln 2}\right] - \left[\frac{\left(\frac{1}{2}-\lambda_k\right)^2}{\ln 2} - \frac{\mu_1}{\gamma_{1,k}^{(n)}}\right]^+, \ (\forall k) \end{aligned} \tag{4.26}$$

In the second hop, the sub-carrier n is allocated to the relay-destination pair $\{k', l'\}$ which satisfies the same criteria as:

$$\begin{aligned} \{k', l'\} &= \arg\max_{\{k,l\}} \delta_2^{(n)}(\lambda_k, \mu_2) \\ &= \arg\max_{\{k,l\}} \left(\frac{1}{2} + \lambda_k\right)^2 \log_2\left[\frac{\left(\frac{1}{2}+\lambda_k\right)^2}{\mu_2 \ln 2}\right] - \left[\frac{\left(\frac{1}{2}+\lambda_k\right)^2}{\ln 2} - \frac{\mu_2}{\gamma_{2,k,l}^{(n)}}\right]^+, \ (\forall k, l) \end{aligned} \tag{4.27}$$

Note that k^* and k' may denote different relays. Thus, the optimal results can be written as:

$$p_{1,k}^{*(n)} = \begin{cases} \left[\frac{\left(\frac{1}{2}-\lambda_k^*\right)^2}{\mu_1^* \ln 2} - \frac{1}{\gamma_{1,k}^{(n)}}\right]^+, & k = k^*; \\ 0, & k \neq k^* \end{cases} \tag{4.28}$$

$$p_{2,k,l}^{*(n)} = \begin{cases} \left[\frac{\left(\frac{1}{2}+\lambda_k^*\right)^2}{\mu_2^* \ln 2} - \frac{1}{\gamma_{2,k,l}^{(n)}}\right]^+, & \{k,l\} = \{k', l'\}; \\ 0, & \{k,l\} \neq \{k', l'\} \end{cases} \tag{4.29}$$

$$T_{1,k}^* = \left(\frac{1}{2} - \lambda_k^*\right) T$$

$$T_{2,k}^* = \left(\frac{1}{2} + \lambda_k^*\right) T$$

From the optimal results, it is clear that $\mu_{1,2} > 0$, and $\lambda_k (k = 1, ..., K) \in (-\frac{1}{2}, \frac{1}{2})$. Note that when symmetric time allocation [21] is applied, *i.e.*, $T_{1,k}^* = T_{2,k}^* = T/2$, the optimal power allocation results are the same as (4.28)-(4.19), except that $(1/2 - \lambda_k^*)^2$ in (4.28) and $(1/2 + \lambda_k^*)^2$ in (4.29) are replaced with $(1/4 - \lambda_k^*)$ and $(1/4 + \lambda_k^*)$, respectively. Hence, the proposed ARA algorithm has a similar complexity to the SRA algorithm. In addition, the value of these multipliers are the limited information which need to be fed back from the relay nodes to BS. The amount of feedback information increases linearly with the number of relays in the cell.

The following $(K + 2)$ equations are to be solved to obtain the optimal Lagrangian multipliers:

$$U = P_S - \sum_{k=1}^{K} \sum_{n=1}^{N} \xi_{1,k}^{(n)} \left[\frac{\left(\frac{1}{2}-\lambda_k\right)^2}{\mu_1 \ln 2} - \frac{1}{\gamma_{1,k}^{(n)}}\right]^+ = 0 \tag{4.30}$$

$$V = P_R - \sum_{k=1}^{K} \sum_{l=1}^{L} \sum_{n=1}^{N} \xi_{2,k,l}^{(n)} \left[\frac{\left(\frac{1}{2}+\lambda_k\right)^2}{\mu_2 \ln 2} - \frac{1}{\gamma_{2,k,l}^{(n)}}\right]^+ = 0 \tag{4.31}$$

$$
\begin{aligned}
W =&(1+2\lambda_k)\sum_{l=1}^{L}\sum_{n=1}^{N}\xi_{2,k,l}^{(n)}\left[\log_2\left(\frac{\left(\frac{1}{2}+\lambda_k\right)^2}{\mu_2\ln 2}\gamma_{2,k,l}^{(n)}\right)\right]^{+}\\
&-(1-2\lambda_k)\sum_{n=1}^{N}\xi_{1,k}^{(n)}\left[\log_2\left(\frac{\left(\frac{1}{2}-\lambda_k\right)^2}{\mu_1\ln 2}\gamma_{1,k}^{(n)}\right)\right]^{+}=0,\ (\forall k)
\end{aligned}
\tag{4.32}
$$

The sub-gradient searching algorithm [98], can be applied to solve (4.30)-(4.32) with an iterative manner. In the first hop, $\boldsymbol{\gamma_1}$ is defined as the CNR map, which can be written as the following matrix:

$$
\boldsymbol{\gamma_1}=\begin{pmatrix}
\gamma_{1,1}^{(1)} & \gamma_{1,1}^{(2)} & \cdots & \gamma_{1,1}^{(N)}\\
\gamma_{1,2}^{(1)} & \gamma_{1,2}^{(2)} & \cdots & \gamma_{1,2}^{(N)}\\
\vdots & \vdots & \ddots & \vdots\\
\gamma_{1,K}^{(1)} & \gamma_{1,K}^{(2)} & \cdots & \gamma_{1,K}^{(N)}
\end{pmatrix}
\tag{4.33}
$$

Similarly, in the second hop, $\boldsymbol{\gamma_2}$ is defined as the CNR map, and can be written as:

$$
\boldsymbol{\gamma_2}=\left(\boldsymbol{\gamma_2^{(1)}},\ldots,\boldsymbol{\gamma_2^{(n)}},\ldots,\boldsymbol{\gamma_2^{(N)}}\right)
\tag{4.34}
$$

and for each sub-carrier, there is:

$$
\boldsymbol{\gamma_2^{(n)}}=\begin{pmatrix}
\gamma_{2,1,1}^{(n)} & \gamma_{2,1,2}^{(n)} & \cdots & \gamma_{2,1,L}^{(n)}\\
\gamma_{2,2,1}^{(n)} & \gamma_{2,2,2}^{(n)} & \cdots & \gamma_{2,2,L}^{(n)}\\
\vdots & \vdots & \ddots & \vdots\\
\gamma_{2,K,1}^{(n)} & \gamma_{2,K,2}^{(n)} & \cdots & \gamma_{2,K,L}^{(n)}
\end{pmatrix}
\tag{4.35}
$$

$\gamma_{1,\min}=\min\{\boldsymbol{\gamma_1}\}$ and $\gamma_{2,\min}=\min\{\boldsymbol{\gamma_2}\}$ denote the minimum value amongst all CNRs in the first and second hop, respectively. Defining t as tth iteration, $\epsilon_{1,2,3}$ as arbitrary small non-zero values, and $\Delta_{1,2,3}$ as the step size for each updating of Lagrange multipliers, the proposed ARA procedure can be summarised as follow steps:

S1 Initialise $\mu_1^{(t)}=\gamma_{1,\min}$, $\mu_2^{(t)}=\gamma_{2,\min}$, and $\lambda_k^{(t)}=0,\ (k=1,...,K)$;

S2 In the first hop, for each sub-carrier n, calculate $\delta_1^{(n)}\left(\mu_1^{(t)},\lambda_k^{(t)}\right),\ (k=1,\ldots,K)$, and find relay k^* that satisfy $k^*=\arg\max_k\delta_1^{(n)}$; in the second hop, for each sub-carrier n, calculate $\delta_2^{(n)}\left(\mu_2^{(t)},\lambda_k^{(t)}\right),\ (k=1,\ldots,K,l=1,\ldots,L)$, and find relay-destination pair $\{k',l'\}$ that satisfy $\{k',l'\}=\arg\max_{\{k',l'\}}\delta_2^{(n)}$; return the binary indicators $\xi_{1,k}^{(n)}=\begin{cases}1, & k=k^*;\\ 0, & k\neq k^*\end{cases}$ and $\xi_{2,k,l}^{(n)}=\begin{cases}1, & \{k,l\}=\{k',l'\};\\ 0, & \{k,l\}\neq\{k',l'\}\end{cases}$

S3 Substitute $\mu_1^{(t)}$, $\mu_2^{(t)}$, $\lambda_k^{(t)}$, $\xi_{1,k}^{(n)}$ and $\xi_{2,k,l}^{(n)}$ into U, V, and W; if $|U|<\epsilon_1$, $|V|<\epsilon_2$ and $|W|<\epsilon_3$ are satisfied simultaneously, stop and return the optimal results with $\mu_1^*=\mu_1^{(t)}$, $\mu_2^*=\mu_2^{(t)}$, and $\lambda_k^*=\lambda_k^{(t)}$, otherwise update the Lagrange multipliers as $\mu_1^{(t+1)}=\mu_1^{(t)}+\Delta_1U$, $\mu_2^{(t+1)}=\mu_2^{(t)}+\Delta_2V$, and $\lambda_k^{(t+1)}=\lambda_k^{(t)}+\Delta_3W$, then repeat **S1**-**S3**.

It is possible that two relay-destination pairs have the identical destination, *i.e.* $\xi_{2,k_1,l}^{(n_1)} = \xi_{2,k_2,l}^{(n_2)} = 1$, during the iterative RA procedure. Hence, CoMP at relays is able to be applied after the SA. Note that $\Delta_{1,2,3}$ determine the speed of convergence, and they can be selected from either constant or variable values (typically functions of t). For instance, if Δ_1 is a constant value, the convergence speed of $|U|$ to ϵ_1 will increase with the augment of iterations. On the contrary, if Δ_1 is a variable step for each iteration t, say $\frac{1}{\sqrt{1+t}}$, the convergence will eventually slow down. In other words, sub-gradient searching method is very dynamic. However, this method takes up many iterations for slow convergence to obtain accurate results, which is not suitable in practical scenario where the channel information changes fast.

Another characteristic of the proposed optimal ARA algorithm is that the result of SA can be updated at each iteration due to the updating of the Lagrange multipliers. Meanwhile, the results of transmit power and time duration are also updated with SA result in the system. Hence, sub-carrier, transmit power and time duration are allocated jointly with optimal results.

4.3.2 Suboptimal ARA Algorithm

To ensure a good tradeoff between the performance and the complexity, the optimisation problem (4.7) can be decomposed into two steps. First, SA is conducted. Second, joint time and power allocation is performed, where the time slot durations in the two consecutive slots are also designed to be asymmetric.

In the first hop, K relays feedback the CSI of all sub-carriers to the BS, and BS obtains a $K \times N$ CNR matrix $\{\gamma_{1,k}^{(n)}\}$ $(k = 1, .., K$ and $n = 1, ..., N)$ with perfect channel estimation. Defining $\Omega_{1,k}$ $(k = 1, ...K)$ as the sub-carrier set of relay k, which is an integer set representing the number of sub-carriers allocated to relay k, the following steps describe the SA in the first hop:

S1 Initialise the sub-carrier set $\Omega_{1,k} = \emptyset$, and initialise the transmit power on each sub-carrier as $p_{1,k}^{(n)} = P_S/N$ $(\forall k, n)$

S2 For $k = 1, ..., K$, find the maximal CNR $\gamma_{1,k}^{(n^*)}$, then update $\Omega_{1,k} = \{n^*\} \bigcup \Omega_{1,k}$, which is equivalent as $\xi_{1,k}^{(n^*)} = 1$.

S3 Calculate the throughput of each relay node by using (4.2). Find the minimal value of capacity C_{1,k^*} amongst K relays, then give the priority to k^* to select the maximal CNR $\gamma_{1,k^*}^{(n')}$, update $\Omega_{1,k^*} = \{n'\} \bigcup \Omega_{1,k^*}$, and $\boldsymbol{N} = \boldsymbol{N} - \{n'\}$. Repeat **S3** until $\boldsymbol{N} = \emptyset$.

After SA, the throughput of relay k can be rewritten from (4.2) as

$$C_{1,k} = \frac{B}{N} \sum_{n \in \Omega_{1,k}} \log_2 \left(1 + p_{1,k}^{(n)} \gamma_{1,k}^{(n)}\right) \tag{4.36}$$

In the second hop, L destination nodes feedback the CSI of all sub-carrier to each of the K relays, and the relays pass these channel information to the BS. As a result, the BS obtains a $K \times N \times L$ 3-dimensional CNR matrix $\{\gamma_{2,k,l}^{(n)}\}$ ($k = 1, ..., K$,$l = 1, ..., L$ and $n = 1, ..., N$). For each individual relay, there is a $N \times L$ CNR matrix, and the SA is based on the sub-carrier-destination pair. Defining $\Omega_{2,k}$ ($k = 1, ...K$) as the set of sub-carrier-destination pairs, *i.e.*, (n, l), the SA can be described as the following steps:

S1 Initialise the sub-carrier set $\Omega_{2,k} = \emptyset$, and initialise the transmit power on each sub-carrier as $p_{2,k,l}^{(n)} = P_R/N$ $(\forall k, n)$

S2 For $k = 1, ..., K$, find the maximal CNR $\gamma_{2,k,l^*}^{(n^*)}$, then update $\Omega_{2,k} = \{(n^*, l^*)\} \bigcup \Omega_{2,k}$.

S3 Calculate the throughput of each relay node by using (4.4). Find the minimal value of capacity C_{2,k^*} amongst K relays, then give the priority to k^* to select the maximal CNR $\gamma_{2,k^*,l'}^{(n')}$, update $\Omega_{2,k^*} = \{(n', l')\} \bigcup \Omega_{2,k^*}$, and $\boldsymbol{N} = \boldsymbol{N} - \{n'\}$. Repeat **S3** until $\boldsymbol{N} = \emptyset$.

Similar to the first hop, after SA, the throughput of relay k can be rewritten from (4.4) as

$$C_{2,k} = \frac{B}{N} \sum_{n \in \Omega_{2,k}} \log_2 \left(1 + p_{2,k,l}^{(n)} \gamma_{2,k,l}^{(n)}\right) \tag{4.37}$$

From the previous steps, it can be seen that binary values are given to the variables $\xi_{1,k}^{(n)}$ and $\xi_{2,k,l}^{(n)}$ after SA. Hence, the objective function for throughput maximisation can be formulated as

$$\max_{p_{1,k}^{(n)},\, p_{2,k,l}^{(n)},\, T_{1,k},\, T_{2,k}} C \tag{4.38}$$

subject to:

$$T_{1,k} + T_{2,k} = T, \ (\forall k) \tag{4.39}$$

$$\sum_{k=1}^{K} \sum_{n \in \Omega_{1,k}} \leq P_S, \ \sum_{k=1}^{K} \sum_{n \in \Omega_{2,k}} p_{2,k,l}^{(n)} \leq P_R, \text{ where } p_{1,k}^{(n)} \geq 0, \ p_{2,k,l}^{(n)} \geq 0, \ (\forall k, l, n) \tag{4.40}$$

$$P_S + P_R = P_T \tag{4.41}$$

The closed-form results of suboptimal power and time duration can be obtained by Lagrange function as well, which have the same forms as (4.28) (4.19). However, the equations for Lagrange multipliers become

$$P_S - \sum_{k=1}^{K} \sum_{n \in \Omega_{1,k}} \left[\frac{\left(\frac{1}{2} - \lambda_k\right)^2}{\mu_1 \ln 2} - \frac{1}{\gamma_{1,k}^{(n)}} \right]^+ = 0 \tag{4.42}$$

$$P_R - \sum_{k=1}^{K} \sum_{n \in \Omega_{2,k}} \left[\frac{\left(\frac{1}{2} + \lambda_k\right)^2}{\mu_2 \ln 2} - \frac{1}{\gamma_{2,k,l}^{(n)}} \right]^+ = 0 \tag{4.43}$$

$$
\begin{aligned}
&(1+2\lambda_k)\sum_{n\in\Omega_{2,k}}\left[\log_2\left(\frac{\left(\frac{1}{2}+\lambda_k\right)^2}{\mu_2\ln 2}\gamma_{2,k,l}^{(n)}\right)\right]^+ \\
&-(1-2\lambda_k)\sum_{n\in\Omega_{1,k}}\left[\log_2\left(\frac{\left(\frac{1}{2}-\lambda_k\right)^2}{\mu_1\ln 2}\gamma_{1,k}^{(n)}\right)\right]^+ = 0,\ (\forall k)
\end{aligned}
\tag{4.44}
$$

where $\xi_{1,k}^{(n)}$ and $\xi_{2,k,l}^{(n)}$ are replaced by the sub-carrier sets $\Omega_{1,k}$ and $\Omega_{2,k}$, respectively.

4.4 Numerical Results

Table 4.1: The parameter setup in the multi-relay based OFDMA cellular system

Carrier central frequency f_c	5 GHz
Bandwidth in downlink B	50 MHz
Number of sub-carriers N	256
Number of channel path	6
RMS delay of channel	0.5 μs
Width of band region of relays w	20 m
AWGN Noise spectral density N_0	-174 dBm/Hz
Total transmit power P_T	40 dBm
Transmit power P_S, P_R	38 dBm
Total transmit time duration T	5 ms
PL exponent in the first hop	4.5
PL exponent in the second hop	3.5

In this section, numerical results demonstrate the performance of ARA in terms of system maximum achievable throughput and time allocation. Table 4.1 presents the parameter setup of the system. Except for Figs. 4.3 and 4.4, there are $K = 4$ relays, and $L = 8$ destinations in the system, and all relays are uniformly distributed within a band region of width w, as shown in Fig. 4.1. Apply the Type-D (Roof-to-Roof) PL model [102] in the first hop as $A_1(d_{SR}) = 2.05 f_c^{2.6} \times 10^{-26}$. Applying the Type-E (Roof-to-Ground) PL model [102] in the second hop, the average PL is as $A_2(d_{RD}) = 38.4$ dB. In all figures except Fig. 4.5, the distance between source and destinations is fixed at $d_{SD} = 2,000$ m, where the system performance at the cell-edge of the second hop coverage region is focused on. The distance between the source and the mid circle of the relays d_{SR} varies between 600 m and 1,800 m. The ARA approach in [20] is implemented for comparison, where the whole cellular system is divided into $K = 4$ single-relay sub-systems.

Fig. 4.2 illustrates the impact of the locations of relays on the maximum achievable system throughput. The proposed ARA scheme outperforms the single-relay based ARA scheme in [20] and the multi-relay based SRA scheme [21], especially when the relays are relatively close to the source. For instance, when $d_{SR} = 600$ m, the performance of the proposed ARA is around 46.9% and 70.5% higher than that of single-relay

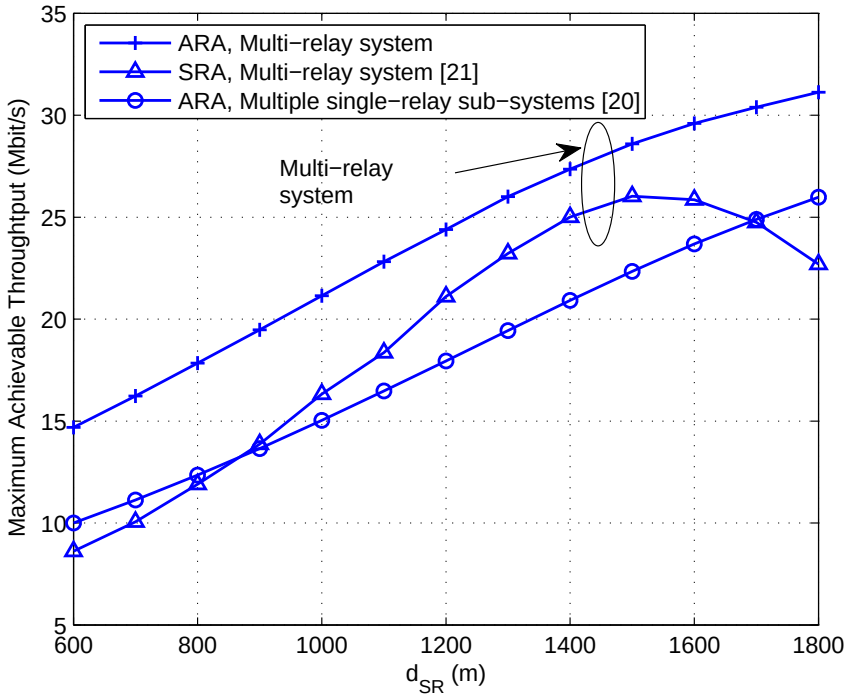

Figure 4.2: Impact of distance between the source and the mid circle of relays on maximum achievable throughput, with $K = 4$ relays and $L = 8$ destinations

ARA and SRA, respectively. It can also be observed that the performances of proposed ARA and single-relay ARA have the same tendency of increasing when d_{SR} is in the range of 600 m to 1800 m. When the d_{SR} further increases, the performance of SRA deteriorates rapidly, as the symmetric time allocation cannot compensate for the performance imbalance caused by different levels of PL attenuation in the two hops.

In Fig. 4.3, the impact of the number of relays on the maximum achievable throughput is illustrated, where $d_{SR} = 1,200$ m. The number of destinations is fixed as $L = 16$. It is shown that when the number of relays increases from 1 to 16, the system throughput of proposed ARA is enlarged by 65.6% due to multi-relay diversity, while the throughput of the single-relay based ARA scheme [20] decreases by 9.6% due to lack of multi-relay diversity. When the number of relays is 1, the proposed ARA scheme is degraded to the scheme in [20].

With the same configuration as Fig. 4.3, Fig. 4.4 illustrates the impact of the number of destinations on the maximum achievable system throughput, with $K = 4$ relays. It can be observed that when the number of destinations increases from 8 to 24, the system throughput has little increase, which demonstrates that the multi-user diversity has less impact on performance compared with multi-time diversity and multi-relay diversity, as shown in Figs. 4.2 and 4.3.

Defining d_{SR}/d_{SD} as the ratio of the distance between source and the mid circle

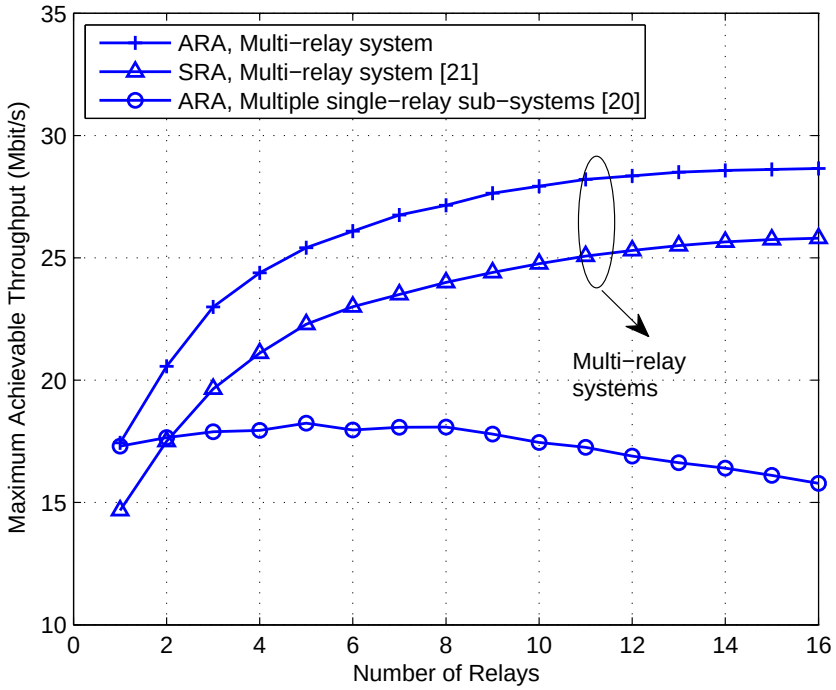

Figure 4.3: Impact of the number of relays on maximum achievable throughput, with $L = 16$ destinations

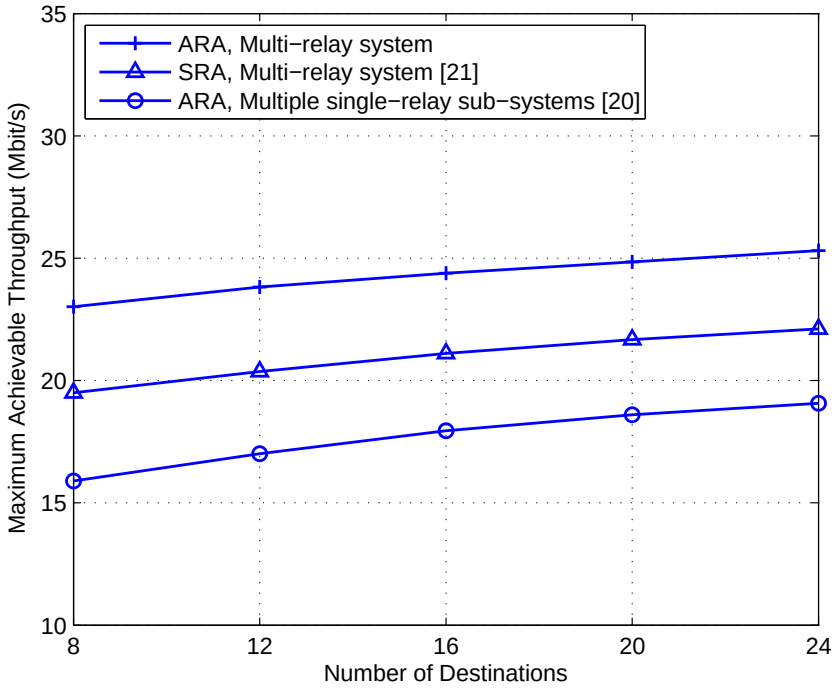

Figure 4.4: Impact of the number of destinations on maximum achievable throughput, with $K = 4$ relays

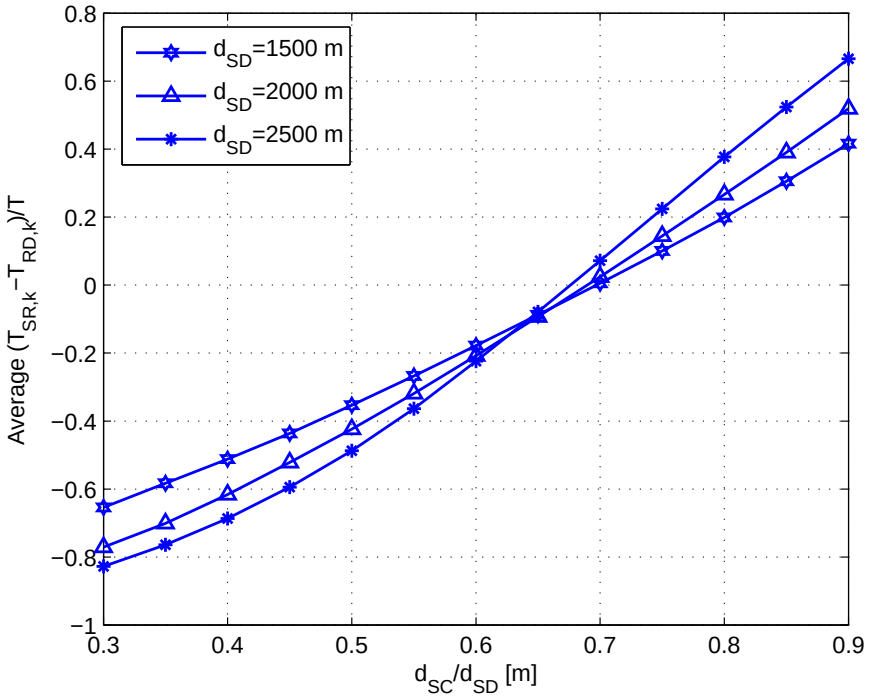

Figure 4.5: Impact of normalised distance between source and destinations on the average time difference between two slots, with $K = 4$ relays and $L = 8$ destinations

of the relays band to the total distance between source and destinations, the impact of d_{SR}/d_{SD} on ΔT_{ave} is demonstrated in Fig. 4.5, where $\Delta T_{ave} = \sum_{k=1}^{K} E\{(T_{1,k} - T_{2,k})/T\}/K$ denotes the average normalised time difference between two hops. It can be observed that when relays are closer to source than destinations, the optimal time duration in the first hop is smaller than that in the second hop. For instance, when $d_{SR}/d_{SD} = 0.3$, $\Delta T_{ave} = -0.78$ for $d_{SD} = 2,000$ m, which indicates the time duration allocated in the second hop can combat the higher level of PL attenuation caused by longer distance to achieve the same throughputs with the first hop. It is also observed that the longer the distances between source and destinations are, the steeper the curve is. Therefore, when destinations are further away from source, asymmetric time allocation plays a more important role in system performance. It can be deducted that symmetric time allocation may be a preferable alternative at $d_{SR}/d_{SD} = 0.65 \sim 0.7$ (irrespective of the value of d_{SD}), due to easier synchronisation.

Although the numerical results demonstrates and compares the maximal system throughput achieved by different RA algorithms, these results are mainly obtained from theoretical calculation, and have not been validated by simulation using realistic models.

4.5 Summary

In this chapter, the ARA has been investigated for DF dual-hop multi-relay based OFDMA cellular systems in the downlink. The proposed ARA scheme performs joint time, power and sub-carrier allocation to achieve an enhanced degree of freedom from the multi-time, multi-relay, and multi-user diversities.

Numerical results show that asymmetric time allocation plays an important role in the system performance. As a result, the proposed ARA scheme can outperform the SRA scheme [21] significantly due to multi-time diversity, at a similar complexity, especially when relays are relatively close to the source. Also, when the distances between source and destinations are relatively large, the impact of asymmetric allocation is more significant. Furthermore, the proposed optimal multi-relay ARA scheme outperforms the single-relay based ARA scheme in [20], thanks to the multi-relay diversity from CoMP technique. Numerical results also show that the multi-user diversity plays a less significant role than the multi-relay and multi-time diversities in the system performance.

Chapter 5

Power Allocation for Short-Range 60 GHz Relay Based Systems

Recently, communications at 60 GHz have attract considerable attention for the need of indoor high speed local area networks [15] [16] [17]. With the large bandwidth (usually more than 1 GHz), 60 GHz communications are able to support multi-Gbps wireless connections between personal devices, such as the wireless transmission of realtime HD video stream between the smart phone and digital TV set. The development in the design of circuits and PAs [103] [1] [104] also motivates the standardisation, such as the IEEE 802.11ad WiGig [105] and the IEEE 802.15.3c [106] task groups.

One big challenge for 60 GHz wireless communication is the high attenuation in propagation. To combat this problem, relay techniques, such as AF relaying and DF relaying [2] are considered as promising solutions, which can improve the reliability of transmission and extend the coverage of signal source. In [107], a cross-layer design was presented for the multi-Gbps indoor communications, with the consideration that the millimetre (mm) wave is highly directional and susceptible to blockage. To maintain network connectivity, a medium access control (MAC) protocol for multi-hop relay transmission was proposed. In [19], the authors proposed a 'best fit deflection routing (BFDR)' scheme to find the optimal relay path in DF relay systems with least interference, which improves the effective throughput in the mm-wave wireless personal area networks (WPANs). The non-linear distortion of PA and phase lock loop (PLL) was also considered in this work. A multi-hop concurrent transmission scheme was proposed in [18] to exploit the spatial capacity of mm-wave WPANs, where the time slot utilisation is improved in a time division multiplexing. However, most of these works focused on the MAC layer protocol design, and none of them considered the decoding power. In [108], Cui *et al.* proved that in short-range communications, total energy consumption is determined by both transmit power and decoding power. Furthermore, in [22], based on Thompson's very-large-scale integration (VSLI) model [109], the asymptotic

lower bound of power consumption on encoding/decoding error-correction codes was derived in the system level. It is associated with the average received power, target bit error probability (BEP), and data rate. But the results are only applied in short range conventional P2P communications with ideal transmit power setup.

In this chapter, the power consumption in 60 GHz communications with DAF and DDF relaying strategies is investigated. These systems mainly consist of mobile devices and are applied in the indoor environment, so the power consumption is limited by the batteries. Hence, to minimise the total power consumption while satisfying certain QoS requirements is in need. The proposed work is different from others in the following aspects. First, a total power consumption model is proposed, which includes drive power, decoding power, and PA power. By searching method, the optimal drive power can be allocated to minimise the total power consumption off-line. Second, this is the first work that applies the decoding power model from [22] in the relay systems, which gives a clear clue on the behaviour of decoding power against the drive power. Third, the parameters of the latest state-of-the-art 60 GHz PA and LNA are applied into the system, which gives a meaningful guideline of power consumption since the transmit powers at source and relay are limited by the saturation of PA. The impact of relay locations on the system performance is investigated, and the comparison between the DAF and DDF relaying is made. From the simulation results, it can be concluded that in the small source-relay separation case, the difference of total power consumption between DAF and DDF is small, and DAF consumes slightly less power. While with larger source-relay separation, the difference gets bigger, and DDF is much more power saving. This is different from the common intuition that DDF is more power consuming than DAF due to the extra decoding power consumption at relay, which is due to the fact that the large source-relay separation limits the effective destination SNR in DAF, leading to more substantial decoding power consumption in the many more decoding iterations than DDF.

The rest of the chapter is organised as follows. The system model is presented in Section 5.1. The power consumption model is introduced in Section 5.2. The problem formulation of DAF and DDF is described in Section 5.3, respectively. Section 5.4 demonstrates the simulation results, and Section 5.5 gives the summary.

5.1 System Model

A three-terminal relaying network in half-duplex mode for indoor short-range communications is considered, as shown in Fig. 5.1, where the destination can directly hear messages from the source. 'S' represents the signal source. 'R' represents the relay node, and 'D' the destination. The diversity-x(amplify/decode)-and-forward (DxF) relaying protocol [2] is considered in this chapter. In the first phase, the source broadcasts messages to both relay and destination. In the second phase, only the relay

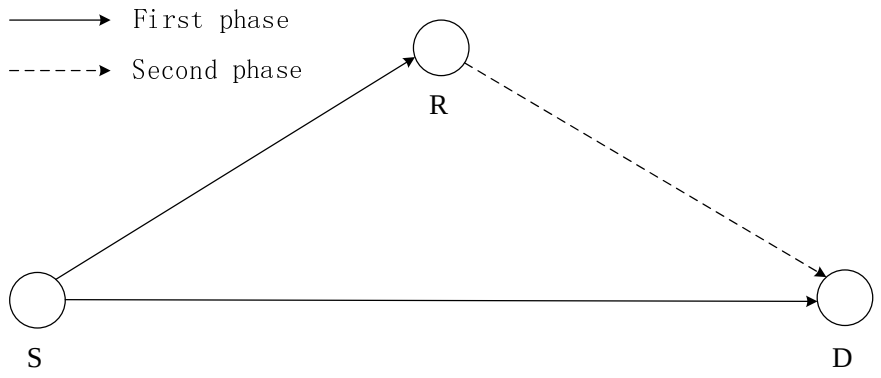

Figure 5.1: Illustration of a three-terminal DxF relay network

transmits messages to the destination. As a result, the destination receives two 'copies' of the original messages through the relay channel. Now $A_{SD}(d_{SD})$, $A_{SR}(d_{SR})$, and $A_{RD}(d_{RD})$ are defined as the PL in the source-destination link, source-relay link and relay-destination link, respectively. Defining an i-bit stream as an information block b_1^i that is encoded into n-bit codewords x_S^n at the source, it is assumed the code-book is identical for both DAF and DDF relaying. Now consider the following strategies:

5.1.1 Diversity AF Relaying

With the DAF relaying strategy, the relay node only amplifies the received waveforms of signals without decoding, and then it forwards waveforms to the destination. Now P_{in} is defined as the drive power of the PA at the source, of which the value is variable. β is defined as the amplification factor of PA in voltage, and $|\beta|^2$ is defined as the amplification gain in power. Correspondingly, the transmit power at source can be written as

$$P_S = \begin{cases} |\beta|^2 P_{in}, & P_{in} \leq P_{in\_max}; \\ P_{sat}, & P_{in} > P_{in\_max} \end{cases} \tag{5.1}$$

where P_{sat} is the output saturation power, and P_{in_max} is the upper bound of drive power. (5.1) indicates when the drive power $P_{in} > P_{in\_max}$, the linearity of the PA cannot be maintained. h_{SR} and h_{SD} are defined as the channel responses from the source to relay, and from the source to destination, respectively. These channel responses follow the complex normal distribution as $h_{SR} \sim CN(0, A_{SR})$ and $h_{SD} \sim CN(0, A_{SD})$, respectively. The average gains of the channel are as $\mathbb{E}\{|h_{SR}|^2\} = A_{SR}$ and $\mathbb{E}\{|h_{SD}|^2\} = A_{SD}$. In the first phase, the received symbols at relay and destination can be written as

$$y_{SR}^n = h_{SR}\sqrt{P_S}x_S^n + z_R \tag{5.2}$$

$$y_{SD}^n = h_{SD}\sqrt{P_S}x_S^n + z_D \tag{5.3}$$

respectively, where z_R and z_D are the complex AWGN noise at relay and destination, respectively, which follow the same complex normal distribution $CN(0, \sigma_0^2)$ without loss

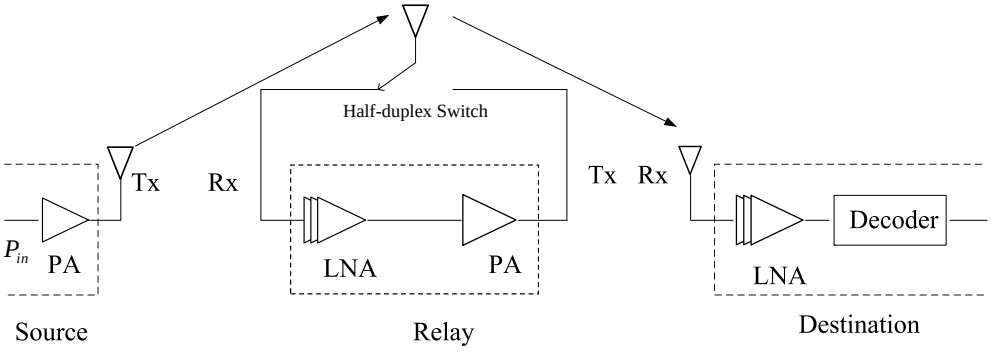

Figure 5.2: The simplified block diagram of an DAF relaying

of generality. The average powers of noise are as $\mathbb{E}\{|z_R|^2\} = \sigma_0^2$ and $\mathbb{E}\{|z_D|^2\} = \sigma_0^2$, where σ_0^2 is the noise-variance.

In Fig. 5.2, a simplified block diagram of a DAF relay node is shown, where Rx represents the receiving end, and Tx represents the transmitting end. After picked up by Rx, the received symbols are amplified first by LNA, of which the amplification gain on power is G, and the noise figure is F. For the multi-cascaded LNAs, the overall noise figure can be calculated by Frii's formula. Assuming that LNAs at each stage are identical, and n_R and n_D are defined as the number of cascaded LNAs at the relay and destination, respectively. Thus, the overall noise figure can be written as:

$$F_{total\_R,D} = F + (F-1)\,\frac{1 - G^{1-n_{R,D}}}{G-1} \tag{5.4}$$

After the LNAs at relay and destination, the signals are amplified as

$$y'^{n}_{SR} = h_{SR}\sqrt{P_S G^{n_R}}\,x^n_S + \sqrt{G^{n_R} F_{total\_R}}\,z_R \tag{5.5}$$

$$y'^{n}_{SD} = h_{SD}\sqrt{P_S G^{n_D}}\,x^n_S + \sqrt{G^{n_D} F_{total\_D}}\,z_D \tag{5.6}$$

In the second phase, the source is switched to silence, and only relay transmits the amplified symbols. z_D is defined as the noise at the destination. Hence, the received symbols at the destination can be written as

$$\begin{aligned} y^n_{RD} &= h_{RD}\beta y'^{n}_{SR} + z_D \\ &= h_{RD}\beta\left(h_{SR}\sqrt{P_S G^{n_R}}\,x^n_S + \sqrt{G^{n_R} F_{total\_R}}\,z_R\right) + z_D \\ &= h_{RD}h_{SR}\beta\sqrt{P_S G^{n_R}}\,x^n_S + \left(h_{RD}\beta\sqrt{G^{n_R} F_{total\_R}}\,z_R + z_D\right) \end{aligned} \tag{5.7}$$

After amplified by the LNAs at the destination, the symbols are

$$\begin{aligned} y'^{n}_{RD} = {} & h_{RD}h_{SR}\beta\sqrt{P_S G^{n_R} G^{n_D}}\,x^n_S \\ & + \left(h_{RD}\beta\sqrt{G^{n_R} G^{n_D} F_{total\_R} F_{total\_D}}\,z_R + \sqrt{G^{n_D} F_{total\_D}}\,z_D\right) \end{aligned} \tag{5.8}$$

Note that the signals from the relay contain the amplified noise from the relay. For DAF relaying, the destination can add up the signals from both relay and source before decoding, resulting in an improved SNR by maximum ratio combining (MRC) at destination [110]. w_{SD}, and w_{RD} are defined as the weights multiplied to y'^{n}_{SD} and y'^{n}_{RD}, respectively. In order to obtain

$$x^n_D = \arg\min_{x^n} \Big| w_{SD} y'^{n}_{SD} + w_{RD} y'^{n}_{RD} - \left(w_{SD} h_{SD} \sqrt{P_S G^{n_D}} x^n_S + w_{RD} h_{RD} h_{SR} \beta \sqrt{P_S G^{n_R} G^{n_D}} x^n_S \right) \Big|^2 \quad (5.9)$$

The weights are designed as

$$w_{SD} = \frac{h^*_{SD}}{\left(|h_{SD}|^2 + |\beta|^2 |h_{SR}|^2 |h_{RD}|^2 \sqrt{F_{total\_D} G^{n_D}} \right)} \quad (5.10)$$

$$w_{RD} = \frac{\beta \sqrt{G^{n_R}} h^*_{SR} h^*_{RD}}{\left(|h_{SD}|^2 + |\beta|^2 |h_{SR}|^2 |h_{RD}|^2 \sqrt{F_{total\_D} G^{n_D}} \right)} \quad (5.11)$$

where * is the conjugation. As a result, the equivalent SNR Γ_{MRC_DAF} at the destination can be written as

$$\Gamma_{MRC\_DAF} = \frac{\left(A_{SD} + |\beta|^2 G^{n_R} A_{SR} A_{RD} \right)^2 P_S}{F_{total\_D} \left(A_{SD} + |\beta|^2 G^{n_R} A_{SR} A_{RD} + |\beta|^4 G^{2n_R} A_{SR} A^2_{RD} F_{total\_R} \right) \sigma^2_0} \quad (5.12)$$

The power variance of equivalent noise is defined as

$$\sigma^2_{MRC\_DAF} = F_{total\_D} \left(A_{SD} + |\beta|^2 G^{n_R} A_{SR} A_{RD} + |\beta|^4 G^{2n_R} A_{SR} A^2_{RD} F_{total\_R} \right) \sigma^2_0 \quad (5.13)$$

By applying combining, detection and demodulation at the destination, the estimated symbols can be obtained as $\hat{x}^n_S$. Afterwards, the estimated bit streams can be obtained from decoding. The overall normalised achievable data rate can be defined as $C_{DAF} = \frac{1}{2} \log_2(1 + \Gamma_{MRC\_DAF})$.

5.1.2 Diversity DF Relaying

With DDF relaying strategy, the repetition coding strategy [2] is applied in the system, which allows the relay to use the same code-book as the source. Consequently, the destination adds up the received symbols before decoding, which yields an improved SNR. Unlike the DAF relaying, the relay decodes the messages from the source first, and then re-encodes them with the same code-book. Hence, the destination receives signals without amplified noise from the relay. P_{in_S} is defined as the drive power of PA at source. In the linear region, $P_S = |\beta|^2 P_{in\_S}$ is the transmit power. Fig. 5.3 shows the simplified block structure of DDF relay node, where the source and destination are not shown. In the first phase, the received symbols at relay and destination are:

$$y^n_{SR} = h_{SR} \sqrt{P_S} x^n_S + z_R \quad (5.14)$$

$$y^n_{SD} = h_{SD} \sqrt{P_S} x^n_S + z_D \quad (5.15)$$

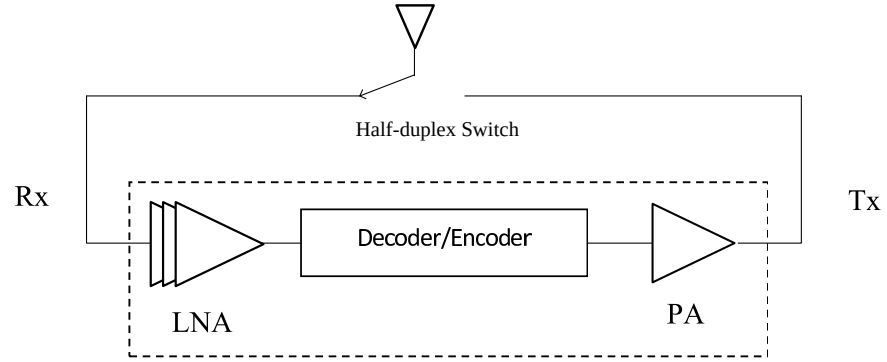

Figure 5.3: The simplified block diagram of DDF relay node, where the source and destination are omitted due to same structures with DAF

The symbols of (5.14) and (5.15) are amplified by LNAs, which yields

$$y'^{n}_{SR} = h_{SR}\sqrt{P_S G^{n_R}}x^n_S + \sqrt{F_{total\_R}G^{n_R}}z_R \tag{5.16}$$

$$y'^{n}_{SD} = h_{SD}\sqrt{P_S G^{n_D}}x^n_S + \sqrt{F_{total\_D}G^{n_D}}z_D \tag{5.17}$$

Hence, the SNR before decoding at relay is $\Gamma_{SR} = A_{SR}P_S/F_{total\_R}\sigma_0^2$. After LNAs, y'^{n}_{SR} pass through the detector, and yield the estimated codewords $\hat{y}'^{n}_{SR}$. These codewords are decoded first, and then encoded into symbols x^n_R for the second phase transmission. After encoding at the baseband, relay provides a new drive power P_{in_R} for PA, which is independent from the received signals at Rx. $P_R = |\beta|^2 P_{in\_R}$ is defined as the transmission power in the linear region. It should be noticed that the two drive powers P_{in_S} and P_{in_R} vary separately. At the destination in the second phase, the received symbols can be written as

$$y^n_{RD} = h_{RD}\sqrt{P_R}x^n_S + z_R \tag{5.18}$$

After amplified by LNAs, the symbols become

$$y'^{n}_{RD} = h_{RD}\sqrt{P_R G^{n_D}}x^n_S + \sqrt{F_{total\_D}G^{n_D}}z_D \tag{5.19}$$

To achieve an improved SNR at destination, the MRC is applied before decoding [111]. Now define w_{SD} and w_{RD} as the weights for the symbols received from the source and relay, respectively. To obtain

$$x^n_D = \arg\min_{x^n} \left| w_{SD}y'^{n}_{SD} + w_{RD}y'^{n}_{RD} - \left(w_{SD}h_{SD}\sqrt{P_S G^{n_D}}x^n_S + w_{RD}h_{RD}\sqrt{P_R G^{n_D}}x^n_R \right) \right|^2 \tag{5.20}$$

the weights are designed as

$$w_{SD} = \frac{h^*_{SD}}{(|h_{SD}|^2 + |h_{RD}|^2)\sqrt{F_{total\_D}G^{n_D}}} \tag{5.21}$$

$$w_{RD} = \frac{\rho h^*_{RD}}{(|h_{SD}|^2 + |h_{RD}|^2)\sqrt{F_{total\_D}G^{n_D}}} \tag{5.22}$$

where the factor $\rho = \sqrt{P_R/P_S}$. It is assumed that when the target BEP at relay node is low enough, the approximation as $x_S^n \doteq x_R^n$ holds. Hence, the equivalent SNR can be written as

$$\begin{aligned}\Gamma_{MRC\_DDF} &= \frac{\mathbb{E}\left\{\left(w_{SD}h_{SD}\sqrt{P_S G^{n_D}}x_S^n + w_{RD}h_{RD}\sqrt{P_S G^{n_D}}x_R^n\right)^2\right\}}{\mathbb{E}\left\{\left(w_{SD}\sqrt{F_{total\_D}G^{n_D}}z_D + w_{RD}\sqrt{F_{total\_D}G^{n_D}}z_D\right)^2\right\}} \\ &\doteq \frac{\mathbb{E}\left\{\left(w_{SD}h_{SD}\sqrt{P_S G^{n_D}} + w_{RD}h_{RD}\sqrt{P_S G^{n_D}}\right)^2\right\}\mathbb{E}\left\{|x_S^n|^2\right\}}{\mathbb{E}\left\{\left(w_{SD}\sqrt{F_{total\_D}G^{n_D}}z_D + w_{RD}\sqrt{F_{total\_D}G^{n_D}}z_D\right)^2\right\}} \\ &\doteq \frac{A_{SD}P_S + A_{RD}P_R}{F_{total\_D}\sigma_0^2} \end{aligned} \tag{5.23}$$

After combining the received symbols from the relay and the source, the destination node can obtain the estimated symbols $\hat{x}_S^n$. $\hat{x}_S^n$ are finally decoded into a bit stream block. The overall normalised achievable throughput of DDF relaying can be defined as $C_{DDF} = \frac{1}{2}\min\{C_{SR}, C_{MRC\_DDF}\}$ [112], where $C_{SR} = \log_2(1 + \Gamma_{SR})$ is the achievable data rate for source-relay link, and $C_{MRC\_DDF} = \log_2(1+\Gamma_{MRC\_DDF})$ is the achievable data rate at the destination with MRC.

5.2 Power Consumption Model

Unlike long-range communications described in Chapter 4, the decoding power should also be considered in 60 GHz short-range communications [108]. In this section, a dynamic model of decoding power is introduced, which can be transformed to a function of the drive power. In the second part, the power consumption of PA is abstracted as the function of drive power as well. Since the total power model includes the decoding power and PA power, it gives a means to allocate an optimal drive power which help the system minimise the total power consumption.

5.2.1 Decoding Power

Grover *et al.* in [22] obtained the lower bound of decoding power consumption in AWGN channel based on the abstracted VLSI model [109]. This model is a message-passing topology where a number of processor elements (PE) are connected to each other by wires in a binary tree pattern. It is assumed that each PE communicates at most ζ other PEs at each of the l iterations, which is shown in Fig. 5.4. Since each iteration consumes a certain amount of power to decode message bits, the decoding power for the whole VLSI decoder is determined by the lower bound on the total number of iterations, which is directly related to the *maximum neighbourhood size* k [22]. Hence, the decoding power is the function of k. Define $C\left(\sigma_G^2\right) = \log_2\left(1 + P_{rec}/\sigma_G^2\right)$ as the normalised AWGN channel capacity over the hypothetical channel, where σ_G^2 is the noise-variance, and P_{rec} as the received average signal power. For decoding at destination, the following

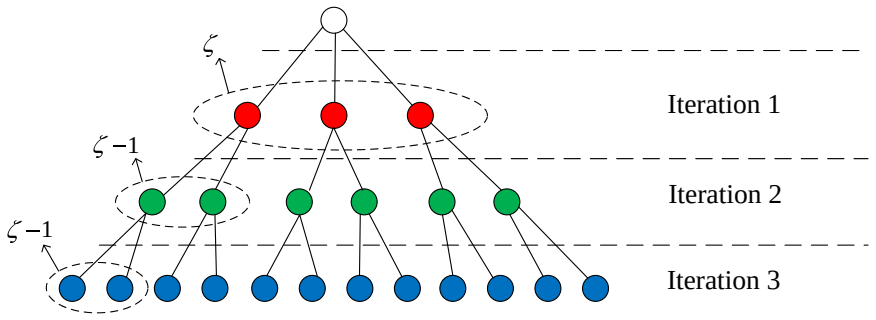

Figure 5.4: Illustration of a message node and its decoding neighbours, where $\zeta = 3$ and the total number of iterations is 3

lower bound holds on the average BEP:

$$P_e \geq \sup_{\sigma_G^2 : C(\sigma_G^2) < R_{ch}} \frac{h_b^{-1}\left(\delta\left(\sigma_G^2\right)\right)}{2} \exp\Bigg(- kD\left(\sigma_G^2 || \sigma_{0\_eq}^2\right) - \sqrt{k}\left(\frac{3}{2} + 2\ln\left(\frac{2}{h_b^{-1}\left(\delta\left(\sigma_G^2\right)\right)}\right)\right)\left(\frac{\sigma_G^2}{\sigma_{0\_eq}^2} - 1\right)\Bigg) \quad (5.24)$$

where $\delta(\sigma_G^2) = 1 - C(\sigma_G^2)/R_{ch}$, the Kullback-Leibler (KL) divergence $D(\sigma_G^2||\sigma_{0\_eq}^2) = \frac{1}{2}\left(\frac{\sigma_G^2}{\sigma_{0\_eq}^2} - 1 - \ln\frac{\sigma_G^2}{\sigma_{0\_eq}^2}\right)$, and the inverse binary entropy function $h_b^{-1}(\cdot) \in [0, 1/2]$. Note that $\sigma_{0\_eq}^2$ indicates the equivalent noise. For instance, in DAF relaying, there is $\sigma_{0\_eq}^2 = \sigma_{MRC\_DAF}^2$. Inequality (5.24) can be approximately rewritten as

$$kD\left(\sigma_G^2 || \sigma_{0\_eq}^2\right) + \sqrt{k}\left[\frac{3}{2} + 2\ln\left(\frac{2}{h_b^{-1}\left(\delta\left(\sigma_G^2\right)\right)}\right)\right] \times \left(\frac{\sigma_G^2}{\sigma_{0\_eq}^2} - 1\right) + \ln\frac{2P_e}{h_b^{-1}\left(\delta\left(\sigma_G^2\right)\right)} \geq 0 \quad (5.25)$$

Note that (5.25) is a quadratic inequality of $\sqrt{k}$. By solving (5.25) (demonstrated in Appendix C), the lower bound on k can be expressed as:

$$k \geq \max\left[1, \sup_{\sigma_G^2 : C(\sigma_G^2) < R_{ch}} K\left(P_{rec}, \sigma_G^2\right)\right] \quad (5.26)$$

where $K\left(P_{rec}, \sigma_G^2\right)$ is the function of the received power and hypothetical noise.

Defining E_{node} as the energy consumed by each message node, R_{dec} as the decoding rate, and R_{ch} as the channel rate, and the lower bound of decoding power P_{dec} can be written as:

$$P_{dec} = \frac{E_{node} R_{dec}}{R_{ch}} \cdot l \geq \frac{E_{node} R_{dec}}{R_{ch}} \cdot \left\lceil \frac{\log_2\left(\frac{\zeta-2}{\zeta}k + \frac{2}{\zeta}\right)}{\log_2(\zeta - 1)} \right\rceil \quad (5.27)$$

where $\lceil \cdot \rceil$ rounds up the value inside to the nearest integer. From (5.27) it can be deduced that with the same target BEP, the higher the received SNR is, the less decoding power is consumed.

5.2.2 PA Power

Defining P_{PA_S} and P_{PA_R} as the PA powers at source and relay, respectively, and $P_{PA} = P_{PA\_S} + P_{PA\_R}$ is defined as the total PA power. Defining $P_{LNA\_R,D} = n_{R,D} P_{LNA\_each}$ as the total cascaded LNAs power at relay and destination, where P_{LNA_each} is the dissipated power of each LNA stage, and the total LNA power is $P_{LNA} = P_{LNA\_R} + P_{LNA\_D}$. Now P_{dec} denotes the decoding power.

The power added efficiency (PAE) of the PA is one parameter that describes how much DC power is transformed to the output power. η is defined as the PAE. From [82], it is known that $\eta = (P_{out} - P_{in})/P_{PA}$, where P_{out} is the output power of a PA. Since both source and relay apply the same PAs, the power of PA can be written in the form of transmit power, as:

$$P_{PA\_S,R} = \frac{|\beta|^2 - 1}{\eta} P_{in\_S,R} \tag{5.28}$$

From the data in [1], PAE can be extracted as the following function by least square (LS) method:

$$\eta\left(P_{in}\right) = -6.8663 \times 10^5 \cdot \left(P_{in} - 5.89 \times 10^{-4}\right)^2 + 0.2249, \ |\beta|^2 P_{in} \leq P_{\text{1dB}} \tag{5.29}$$

where P_{1dB} is the output 1dB compression point. Hence, the power consumption of PA can be extracted as a function of the drive power. The procedure of curve fitting is shown in the Appendix D. Intuitively, the higher the transmit power is, the more PA power is consumed. Hence, when combining with the decoding power, there must be a trade-off point of power consumption with the augment of transmit power (or drive power of PAs).

5.3 Problem Formulation

In this section, the total power consumption model for both DAF and DDF relaying strategies is considered. This model includes the drive power, PA power, decoding power, and LNA power. For simplicity, the power of LNA is modeled as a fixed value. The objective of the system is to minimise the total power consumption with the data rate requirement, by allocating an optimal drive power at the source and the relay (for DDF) nodes. In this case, the source can transmit high data rate streams with low power consumption.

5.3.1 Problem Formulation for DAF Relaying

As shown in Fig. 5.2, one portion of the power is consumed at the PA at the source and relay nodes, another portion is consumed at the LNAs at both the relay and destination nodes, and the last portion is the decoding power at the destination. Defining P_{total_DAF} as the total power consumption of DAF relaying, it can be written as a function of the drive power at the source P_{in_S} as:

$$P_{total\_DAF} = P_{in\_S} + P_{PA\_S}\left(P_{in\_S}\right) + P_{PA\_R}\left(P_{in\_S}\right) + P_{dec}\left(P_{in\_S}\right) + P_{LNA\_S} + P_{LNA\_R} \quad (5.30)$$

Defining R_{data} as the transmit data rate, and B as the bandwidth, the objective function of DAF relaying can be formulated as:

$$\min_{P_{in\_S}} P_{total\_DAF} = \left(P_{in\_S} + P_{PA}\left(P_{in\_S}\right) + P_{dec}\left(P_{in\_S}\right)\right) + P_{LNA} \quad (5.31)$$

subject to:

$$C\left(\sigma_0^2\right) > R_{ch} \quad (5.32)$$

$$\frac{\left(2^{2R_{data}/B} - 1\right)\sigma^2_{MRC\_DAF}}{\left(A_{SD} + |\beta|^2 G^{n_R} A_{SR} A_{RD}\right)^2 |\beta|^2} < P_{in} < \frac{P_{1\text{dB}}}{|\beta|^2} \quad (5.33)$$

where constraint (5.32) indicates that the minimum data rate is selected as one of QoS requirements. Constraint (5.33) is set to maintain the PA in the linear amplification region as well as to satisfy the data rate requirement. $P_{in\_S}(\min) = \frac{\left(2^{2R_{data}/B}-1\right)\sigma^2_{MRC\_DAF}}{\left(A_{SD}+|\beta|^2 G^{n_R} A_{SR} A_{RD}\right)^2 |\beta|^2}$ is set as the minimum drive power that meets the data rate requirement. Due to the complexity of the calculation of lower bound k, it is difficult to find the closed-form result of the optimal P_{in_S}. However, line searching method can be applied by adjusting P_{in_S} in the range of (5.33) with appropriate resolution to achieve the approximated optimal value for the minimum P_{total_DAF}. t $(t = 1, 2, ...)$ denotes the number of iteration for finding the minimal total power consumption, and $\delta_S = \left|P_{in\_S}(\min) - \frac{P_{1\text{dB}}}{|\beta|^2}\right| / N_1$ is defined as the resolution of the variable range (5.33), where N_1 can be any real number that is greater than 1. The searching method can be summarised as below:

S1 Initialise the drive power as $P_{in\_S} = P_{in\_S}(\min) + \delta_S$.

S2 Calculate the maximum neighbourhood size k by golden section search [113]:

- **s1** Calculate the current equivalent noise variance $\sigma^2_{MRC\_DAF}$ by (5.12); set $\sigma^2_{G1} = \sigma^2_{MRC\_DAF}$ and $\sigma^2_{G2} = N_2\sigma^2_{MRC\_DAF}$ as the lower and upper bounds of hypothetical noise variance, where N_2 is a real value that is greater than 1.
- **s2** Set σ^2_{G3} as $\sigma^2_{G1} < \sigma^2_{G3} < \sigma^2_{G2}$, which satisfies $|\sigma^2_{G3} - \sigma^2_{G1}| < |\sigma^2_{G2} - \sigma^2_{G3}|$ and $|\sigma^2_{G3} - \sigma^2_{G1}|/|\sigma^2_{G2} - \sigma^2_{G3}| = \varphi$, where $\varphi = (-1 + \sqrt{5})/2$.

s3 Set a probe point σ^2_{G4} between σ^2_{G3} and σ^2_{G2}, and let $|\sigma^2_{G3}-\sigma^2_{G4}|/|\sigma^2_{G4}-\sigma^2_{G2}| = \varphi$.

s4 If $K(P_{in\_S}, \sigma^2_{G4}) \geq K(P_{in\_S}, \sigma^2_{G3})$ (see (C.5) for the detail of function K), replace σ^2_{G3} by σ^2_{G1} and σ^2_{G4} by σ^2_{G3}, respectively, and delete the original σ^2_{G1}, then repeat **s3**; if $K(P_{in\_S}, \sigma^2_{G4}) < K(P_{in\_S}, \sigma^2_{G3})$, replace σ^2_{G1} by σ^2_{G2} and σ^2_{G4} by σ^2_{G1}, respectively, and delete the original σ^2_{G2}, then repeat s3.

s5 Continue s4 until $|\sigma^2_{G3} - \sigma^2_{G4}|/|\sigma^2_{G3}| < \epsilon$, where ϵ ($\epsilon \in \mathbb{R}$ and $\epsilon > 0$) is the tolerance that can be set arbitrary small; set the optimal noise variance σ^2_{G*} as $\sigma^2_{G*} = \sigma^2_{G3}$, and obtain k by (C.5).

S3 Use k from s5 to calculate P_{dec} by (5.27); get the total power $P^{(t)}_{total\_DAF}$ by (5.30).

S4 Update $P_{in\_S} = P_{in\_S} + \delta_S$, and calculate the total power $P^{(t+1)}_{total\_DAF}$ through **S2**-**S3**; compare $P^{(t)}_{total\_DAF}$ with $P^{(t+1)}_{total\_DAF}$: if $P^{(t)}_{total\_DAF} \leq P^{(t+1)}_{total\_DAF}$, $P^{(t)}_{total\_DAF}$ is the minimal total power and $P^*_{in\_S} = P_{in\_S} - \delta_S$ is the optimal drive power; if $P^{(t)}_{total\_DAF} > P^{(t+1)}_{total\_DAF}$, update $P_{in\_S} = P_{in\_S} + \delta_S$ and repeat S2 - S3 until $P_{in\_S} = \frac{P_{1\mathrm{dB}}}{|\beta|^2}$, then the minimal total power is $P^{(t+1)}_{total\_DAF}$ and optimal drive power is $P^*_{in\_S} = \frac{P_{1\mathrm{dB}}}{|\beta|^2}$.

Note that σ^2_{G*} is determined by the values of N_2 and ϵ. The appropriate value of N_2 allows the optimum to be included in the range $[\sigma^2_{G1}, \sigma^2_{G2}]$, and (5.32) tells that N_2 must be greater than 1. While the value of ϵ controls the speed of converge as well as the accuracy.

5.3.2 Problem Formulation for DDF Relaying

In DDF relaying, decoding is conducted at both relay and destination, hence P_{dec_R} and P_{dec_D} are defined as the two parts of decoding powers at the relay and the destination, respectively. Setting P_{e_R} and P_{e_D} as the target BEPs at relay and destination respectively, the overall target BEP for the DDF relaying can be expressed as $P_e = 1-(1-P_{e\_R})(1-P_{e\_D})$ approximately. The total power consumption model can be written as:

$$P_{total\_DDF} = (P_{in\_S} + P_{PA\_S}(P_{in\_S}) + P_{dec\_R}(P_{in\_S})) + (P_{in\_R} + P_{PA\_R}(P_{in\_R}) + P_{dec\_D}(P_{in\_R})) + P_{LNA} \quad (5.34)$$

where two drive powers P_{in_S} and P_{in_R} determine the total power consumption.

Similar to DAF relaying, the objective function of DDF relaying is formulated to minimise the total power consumption, which is:

$$\min_{P_{in\_S}, P_{in\_R}} P_{total\_DDF}(P_{in\_S}, P_{in\_R}) \quad (5.35)$$

subject to:

$$C_{SR}\left(\sigma_{0\_R}^2\right), C_{MRC\_DDF}\left(\sigma_{0\_D}^2\right) > R_{ch} \tag{5.36}$$

$$\frac{\left(2^{2R_{data}/B}-1\right)F_{total\_R}\sigma_0^2}{A_{SR}|\beta|^2} < P_{in\_S} < \frac{P_{\text{1dB}}}{|\beta|^2} \tag{5.37}$$

$$\frac{\left(2^{2R_{data}/B}-1\right)F_{total\_D}\sigma_0^2 - A_{SD}|\beta|^2 P_{in\_S}}{A_{RD}|\beta|^2} < P_{in\_R} < \frac{P_{\text{1dB}}}{|\beta|^2} \tag{5.38}$$

where constraint (5.36) also indicates that the minimum data rate is selected as one of QoS requirements in both first and second hops, which is similar to DAF. Constraints (5.37) and (5.38) limit the PAs in the linear amplification region, and satisfy the data rate requirement in two hops as well. Note that the lower bound of (5.38) depends on the value of P_{in_S}, which indicates the allocation of P_{in_S} is prior to that of P_{in_R}. $P_{in\_S}(\min) = \frac{\left(2^{2R_{data}/B}-1\right)F_{total\_R}\sigma_0^2}{A_{SR}|\beta|^2}$ and $P_{in\_R}(\min) = \frac{\left(2^{2R_{data}/B}-1\right)F_{total\_D}\sigma_0^2 - A_{SD}|\beta|^2 P_{in\_S}}{A_{RD}|\beta|^2}$ are set as the lower bounds of drive power for the source and relay, respectively. Similar to DAF, it is not easy to get the closed-form results of the optimal values of P_{in_S} and P_{in_R}. Hence, the grid searching method is applied to find the approximated optimal value by adjusting both P_{in_S} and P_{in_R} in the ranges of (5.37) and (5.38) with appropriate resolutions, respectively. The same definition of t and δ_S are kept, meanwhile, $\delta_R = \left|P_{in\_R}(\min) - \frac{P_{\text{1dB}}}{|\beta|^2}\right| / N_1$ is defined as the resolution of range (5.38). Note that (5.38) has the same size of resolution as (5.37) in order to make fair comparison between DAF and DDF. The following steps summarises the grid searching method:

S1 Initialise the drive power at source as $P_{in\_S} = P_{in\_S}(\min) + \delta_S$.

S2 Substitute P_{in_S} into (5.38); initialise the drive power at relay as $P_{in\_R} = P_{in\_R}(\min) + \delta_R$, then search for the minimal value of total power with fixed P_{in_S}:

- **s1** Use P_{in_S} and P_{in_R} to calculate the maximum neighbourhood size at relay k_R and at destination k_D from (C.5) by golden section search, respectively; get the total power consumption $P_{total\_DF}^{(t)}$ by (5.34).
- **s2** Update $P_{in\_R} = P_{in\_R} + \delta_R$, and calculate the total power $P_{total\_DF}^{(t+1)}$ as in **s1**; compare $P_{total\_DF}^{(t)}$ with $P_{total\_DF}^{(t+1)}$: if $P_{total\_DF}^{(t)} \leq P_{total\_DF}^{(t+1)}$, $P_{total\_DF}^{(t)}$ is the minimal total power; if $P_{total\_DF}^{(t)} > P_{total\_DF}^{(t+1)}$, update $P_{in\_R} = P_{in\_R} + \delta_R$ and repeat **s1** until $P_{in\_R} = \frac{P_{\text{1dB}}}{|\beta|^2}$, then the minimal total power is $P_{total\_DF}^{(t+1)}$.
- **s3** Return P_{total_DF}; set $t = 0$.

S3 Update $P_{in\_S} = P_{in\_S} + \delta_S$, and calculate the total power $P_{total\_DF}^{(t+1)}$ through **S2**; compare $P_{total\_DF}^{(t)}$ with $P_{total\_DF}^{(t+1)}$: if $P_{total\_DF}^{(t)} \leq P_{total\_DF}^{(t+1)}$, $P_{total\_DF}^{(t)}$ is the minimal total power, and $P_{in\_S}^* = P_{in\_S} - \delta_S$ and $P_{in\_R}^* = P_{in\_R} - \delta_R$ are the optimal drive powers; if $P_{total\_DF}^{(t)} > P_{total\_DF}^{(t+1)}$, update $P_{in\_S} = P_{in\_S} + \delta_S$ and repeat **S2** until $P_{in\_S} = \frac{P_{\text{1dB}}}{|\beta|^2}$, then the minimal total power is $P_{total\_DF}^{(t+1)}$ and optimal drive power at source is $P_{in\_S}^* = \frac{P_{\text{1dB}}}{|\beta|^2}$.

The golden section search in **S2** is the same as in Section 5.3.1, hence the details of this method is omitted in this section.

5.4 Numerical Results

In this section, some numerical results are shown to demonstrate the performance of DAF and DDF relaying in terms of power consumption. In Table 5.1, the parameters are setup for the system, where the data of LNA is from [104]. Amongst all, the required data rate R_{data} and target BEP P_e are the QoS requirements, which are applied at the destination node wherever it is in the system.

Table 5.1: The parameter setup in the relay based short-range communication systems

Carrier central frequency f_c	60 GHz
Bandwidth B	3 GHz
Required data rate R_{data}	1.5 Gb/s [22]
Energy consumption on each message node E_{node}	10 pJ [22]
Maximum number of ζ	4 [22]
AWGN Noise spectral density σ_0^2	-174 dBm/Hz
Amplification factor $\|\beta\|^2$ of PA	16 dB
Output 1 dB compression point of PA $P_{1\text{dB}}$	12.7 dBm
Power gain of each LNA G	10.5 dB
Noise figure of each LNA F_{each}	5.5 dB
Dissipated power of each LNA P_{LNA_each}	4.4 mW
Cascaded number of LNAs	6
Target BEP P_e	10^{-6}
Antenna gain G_T, G_R	1 [18]
PL exponent α	3 [114]

In Fig. 5.5, the impact of drive power on the total power consumption in DAF relaying is shown, where $d_{SD} = d_{SR} = d_{RD} = 6$ m. It can be seen that the total power decreases first with a 'stairway' shape dramatically, and then increases slightly with the increasing of drive power. When the drive power is 0.41 mW, the minimum total power consumption can be achieved as 225.4 mW. It can be also observed that the decoding power keeps decreasing when the drive power increases. However, the value of decoding power varies with discrete intervals, which are 30 mW power per iteration, and when $P_{in\_S} > 0.41$ mW, the destination needs only 1 iteration. While the total power does not remains as flat as the decoding power, since the PA power at the source keeps increasing. As a result, the minimum total power consumption is the trade-off point between the decoding power and PA power.

With the same configuration of Fig. 5.5, in Fig. 5.6 demonstrates the impact of drive power at relay and source on the total power consumption in DDF relaying. It can be seen that the minimum value can be found with the global optimum. When $P_{in\_S} = 0.4$ mW, $P_{in\_R} = 0.22$ mW, the minimum total power consumption is 257.7 mW, which is

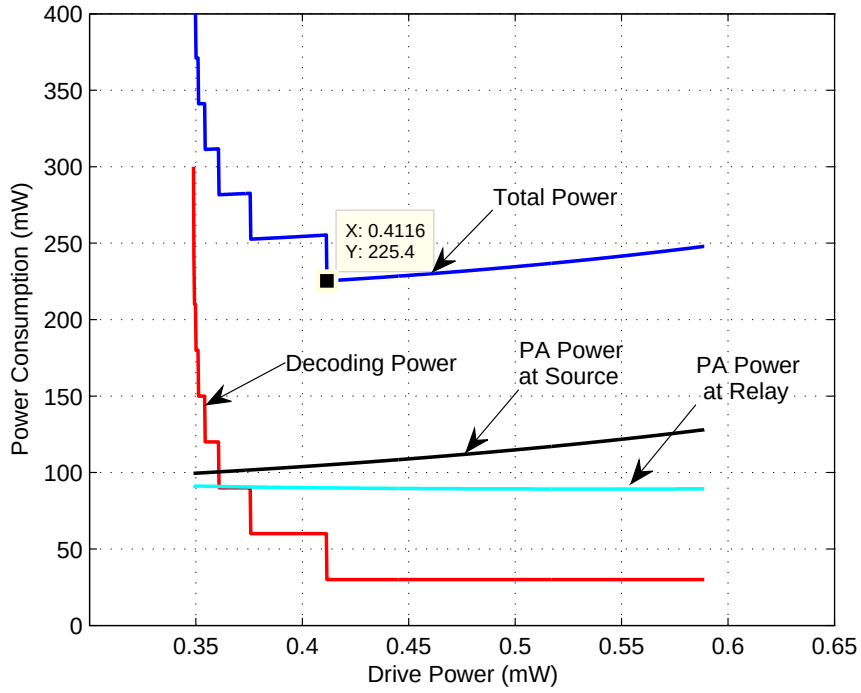

Figure 5.5: The impact of dive power on the total power consumption in DAF relaying, where $d_{SD} = d_{SR} = d_{RD} = 6$ m

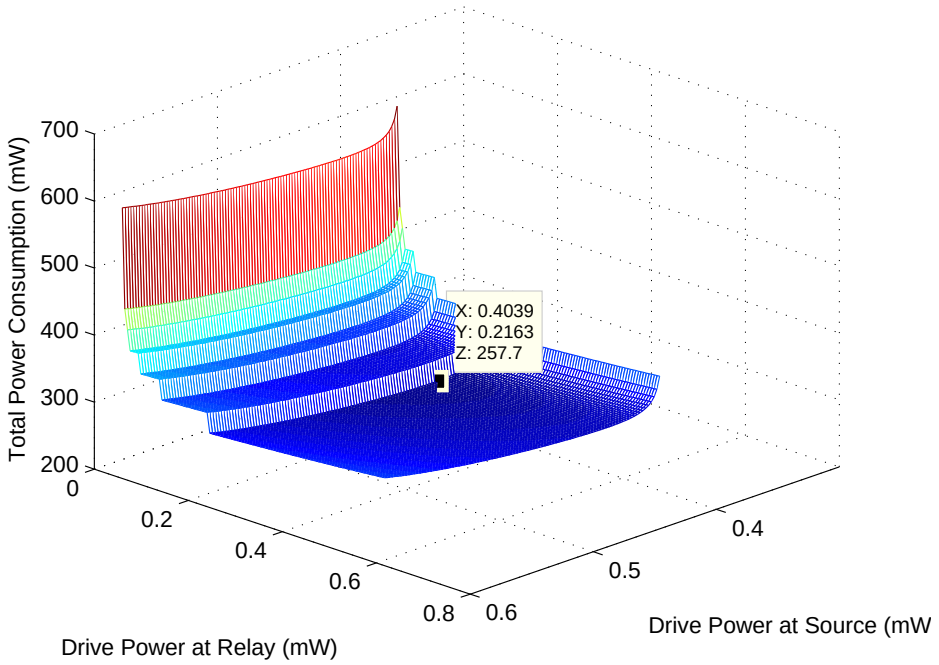

Figure 5.6: The impact of dive power on the total power consumption in DDF relaying, where $d_{SD} = d_{SR} = d_{RD} = 6$ m

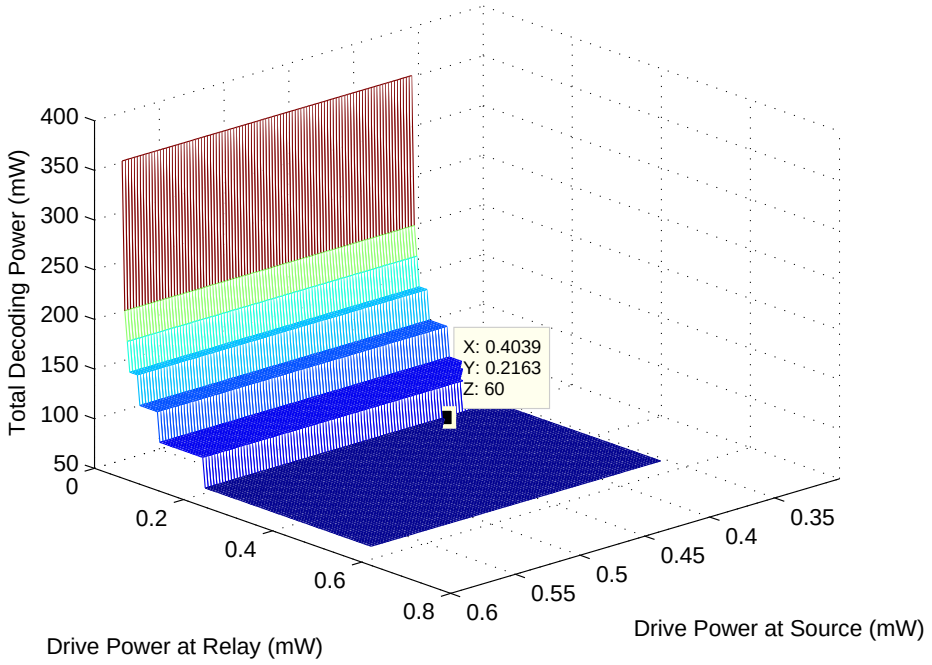

Figure 5.7: The impact of drive power on the total decoding power in DDF relaying

more than that in DAF. Because both DAF and DDF use the same resolution for drive power adjusting, and only AWGN channel capacity is considered, the values of total power consumption for both DAF and DDF are statistically meaningful. The total power value also shows a 'stairway' tendency when the drive power has small values. When the drive power is small enough, the decoding power weights most portion of the total power consumption.

From Fig. 5.7, the variation of decoding power is shown on the impact of the drive power. It can be observed that, when P_{in_R} increases, the decoding power decreases and converges to 60 mW, where 1 iteration is consumed at both relay and destination nodes. It can also be observed that when the drive powers at source and relay are at the global optimum point, the decoding power is at its minimum value 60 mW, which is two times the decoding power in DAF.

In Fig. 5.8, it can be observed that when the drive powers are at optimum point, the total PA power is not at the minimum value, which means the minimum total power is a trade-off between the decoding power and PA power. If simply assuming that the transmit power has a linear relation with the PA power, it can also be concluded that the more the transmit power is, the less the decoding power will be, and the trade-off between these two makes the minimal total power consumption.

In Fig. 5.9, a scenario where the source, relay and destination nodes are at the vertexes of an isosceles right-angle triangle, with $d_{SR} = d_{RD} = d_{SD}/\sqrt{2}$. All the dashed lines represent the DAF relaying, and the solid lines represent the DDF relaying. The

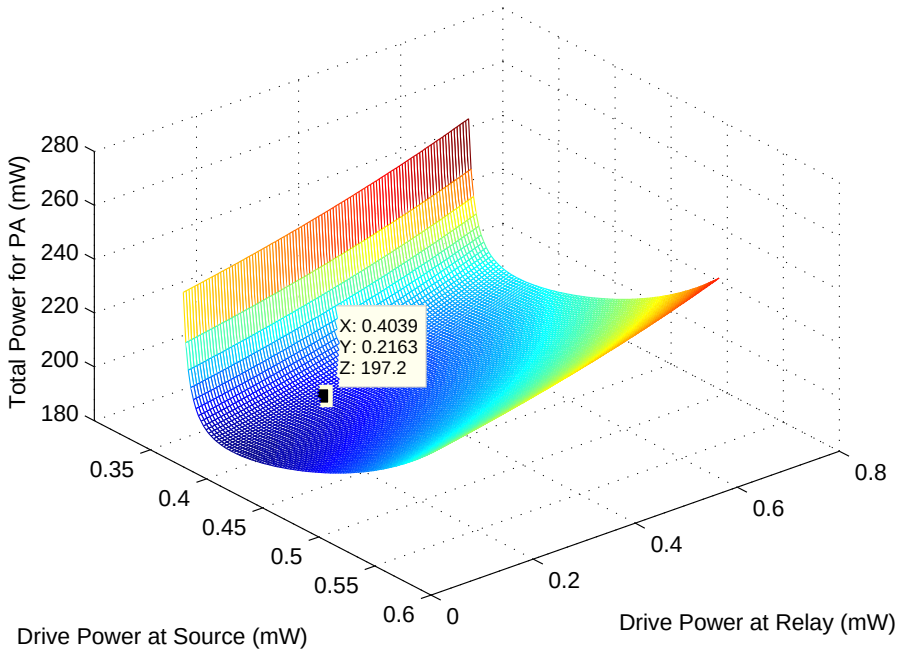

Figure 5.8: The impact of drive power on total PA power in DDF relaying

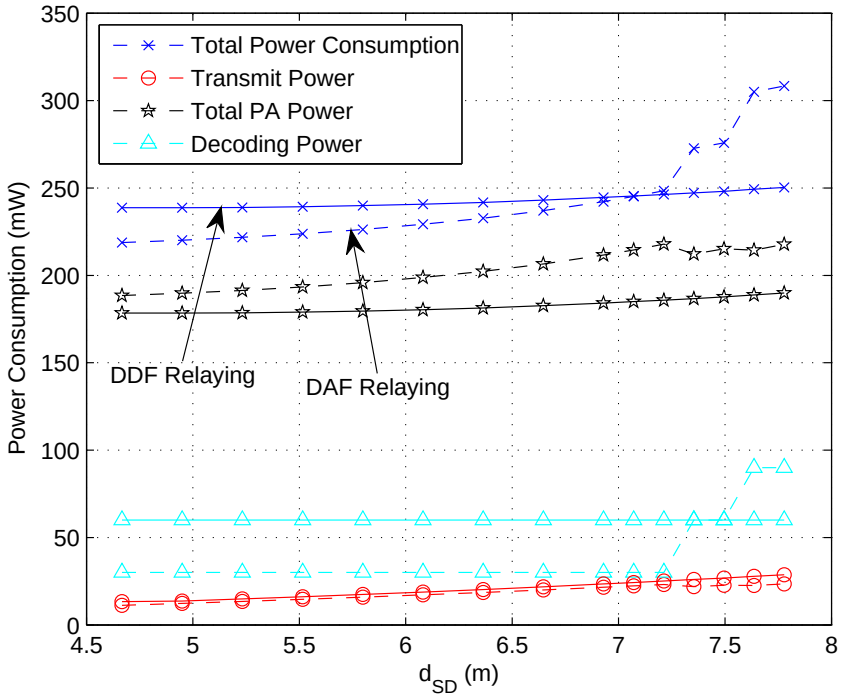

Figure 5.9: The impact of the distances between each nodes on the power consumption, where $d_{SR} = d_{RD} = d_{SD}/\sqrt{2}$. Solid lines: DDF; dashed lines: DAF

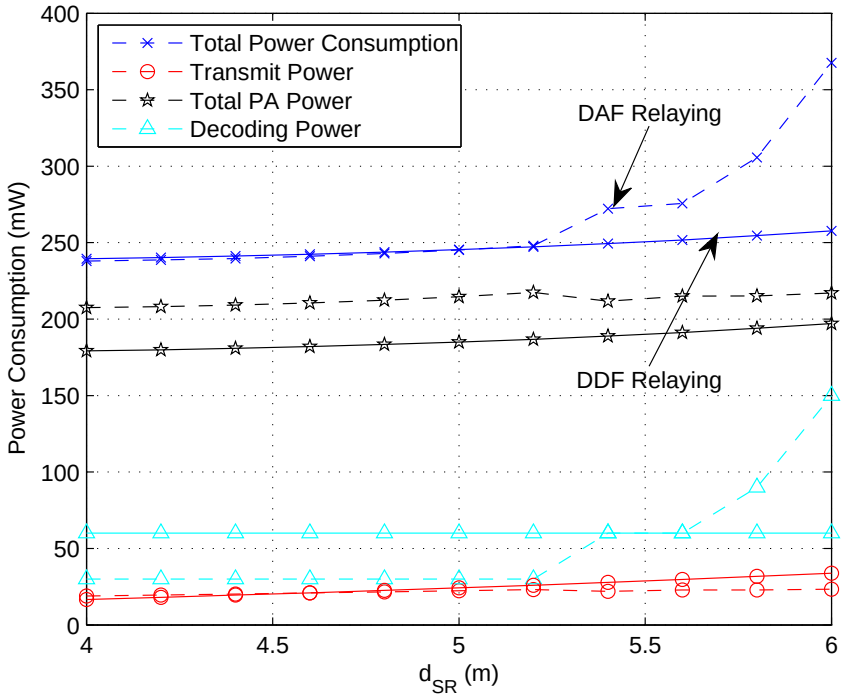

Figure 5.10: The impact of relay's location on the power consumption, where $d_{SD} = 7.07$ m. Solid lines: DDF; dashed lines: DAF

impact of the triangle area on the power consumption is demonstrated. With the increasing of d_{SD} (augment of the area) from 4.67 to 7.78 m (d_{SR} and d_{RD} are from 3.3 to 5.5 m), the total powers of both DAF and DDF increase. When $d_{SD} < 7.07$ m, the total power of DAF is less than DDF. When $d_{SD} > 7.07$ m, both decoding power and total power of DAF increase dramatically. For DAF relaying, the equivalent received SNR is easily affected by the distances between nodes. When d_{SD} is close to 7.07 m, the transmit power of the source is driven towards $P_{1\text{dB}}$ of PA. When $d_{SD} > 7.07$ m, the transmit power is fixed at around 12.7 dBm, and the equivalent received SNR decreases dramatically with the continuous increasing of d_{SD}. Hence, more iterations are used for decoding to compensate the loss in the received SNR, which results in more decoding power consumed at the destination. For DDF relaying however, since the amplified noise at the relay is not added to the destination, the equivalent received SNR is higher than that in DAF relaying. As a result, the transmit powers for both source and relay nodes are much less than $P_{1\text{dB}}$. Hence, both the relay and destination have less burden for decoding, and consume less decoding power.

In Fig. 5.10, a scenario where the relay's locations are on the ellipse orbit is considered, where $\alpha_{SD} = 2.17$. The distance d_{SD} is fixed at 7.07 m, and $d_{SR} + d_{RD} = 10$ m is fixed as well (the source and destination nodes are on the position of two focuses of the ellipse). With the augment of d_{SR} from 4 to 6 m, the total powers of both DAF

and DDF increase. When $d_{SR} > 5.2$ m, the total power of DAF increases dramatically with 'steps', which is caused by the increasing of decoding power and fixed maximal transmit power. This phenomenon results from the degradation of the effective SNR at the destination when the source-relay separation is large. It can be also observed that the DDF has a good robustness against the relay's location.

Although some realistic PA models are used, the results are mainly obtained from theoretical calculation, and have not been validated by simulation. However, these results can give meaningful lower bounds of the system power consumption for both DAF and DDF relaying.

5.5 Summary

In this chapter, the relay networks in 60 GHz short-range communications are investigated, which includes the DAF relaying and DDF relaying. The power consumption model including the drive power, PA power and decoding power is proposed in Section 5.2. The parameters from the latest PA and the decoding power model have been applied to support the proposed model in this work.

From the numerical results, it can be observed that the decoding power cannot be omitted in this 60 GHz short-range communication system, and it has a minimum value when only 1 iteration conducted. From Fig. 5.5 to Fig. 5.8, it can be observed that the minimum total power consumption can be achieved by allocating the optimal drive power to the source and relay. Due to the complexity of (5.30) and (5.34), only searching methods are used for power allocation in this chapter, which makes the off-line system design possible. Furthermore, the total power consumption between DAF and DDF relaying is compared. It has been found that when the separation between source and relay is small, DAF is preferable than DDF in terms of power saving; while with larger separation between source and relay, DDF saves much more power than DAF, and is preferable to be deployed in the system. This is different from the common intuition that DDF is more power consuming than DAF due to the extra decoding power consumption at relay, which is due to the fact that the large source-relay separation limits the effective destination SNR in DAF, leading to more substantial decoding power consumption in the many more decoding iterations than DDF.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

In this thesis, resource allocation in the relay based communication networks has been investigated, including the long-range OFDMA cellular networks and short-range P2P networks with 60 GHz communication.

In Chapter 4, the ARA has been investigated for DF dual-hop multi-relay based OFDMA green cellular systems in the downlink. The proposed ARA scheme allocates the time duration, transmit power and sub-carrier jointly to achieve an enhanced degree of freedom from the multi-time, multi-relay, and multi-user diversities. Simulation results show that asymmetric time allocation plays an important role in the system performance. As a result, the proposed ARA scheme can outperform the SRA scheme [21] significantly due to the multi-time diversity in terms of maximising system throughput with limited transmit power. Also, when the distances between source and destinations are relatively large, the impact of asymmetric allocation is more significant. Furthermore, the proposed optimal multi-relay ARA scheme outperforms the single-relay based ARA scheme in [20], thanks to the multi-relay diversity. Simulation results also show that the multi-user diversity plays a less significant role than the multi-relay and multi-time diversities in the system performance.

In Chapter 5, the relay networks in 60 GHz short-range communications have been investigated with the DAF and DDF relaying strategies. The power consumption model including the drive power, PA power and decoding power is proposed in Section 5.2. From the simulation results, it can be observed that the decoding power plays an important role in this 60 GHz short-range communication system, and it has a minimum value when only 1 iteration conducted. In Fig. 5.5 and Fig. 5.8, it can be observed that the minimum total power consumption can be achieved by allocating the optimal drive power to the source and relay. Due to the complexity of (5.30) and (5.34), only searching methods are used for power allocation at this stage, which makes the off-line system design possible. Furthermore, the total power consumption between DAF and DDF relaying is compared. It has been found that when the separation between source

and relay is small, DAF is preferable than DDF in terms of power saving; while with larger separation between source and relay, DDF saves much more power than DAF, and is preferable to be deployed in the system.

6.2 Future Work

In the research of short-range 60 GHz communication, AWGN channel has been considered only, because the lower bound of decoding power is proposed for BSC and AWGN channels [22]. In practice however, the fading channel model can help to achieve more accurate power allocation results. Furthermore, due to no specified modulation has been considered, the power consumption of AD converter, modulator/demodulator is omitted in this power consumption model. Hence, the future research topics are summarised as the following:

1. To modify Grover's decoding power model with the flat fading channel and frequency selective channels. Less complex expression with only SNR and target BER is needed for power allocation with closed-form results.

2. A modulation method will be selected in 60 GHz communications, which helps to propose the power consumption models of AD converter and modulator/demodulator, and to establish a total power consumption model with full components.

3. Multiuser scenario will be considered in the short-range 60 GHz communications. In addition, the cross-layer design of this system can be also investigated.

4. To apply the energy efficiency as the objective functions in the green wireless system, which considers the trade-off between the total power consumption and the system throughput.

Appendix A

Time Line of Wireless Communications

Generation		Launch Year	Standards	Areas
0G		1948	MTS	USA
		1952	A-Netz	F. R. Germany[1]
		1959	System 1	UK
		1962	IMTS[2]	USA
		1963	Altai	USSR
		1971	ARP[3]	Finland
		1971	AMR[4]	Czechoslovakia
		1972	B-Netz	F. R. Germany
1G		1978	AMPS	USA
		1986		Isreal
		1987		Australia
		1979	TZ-801 TZ-801 TZ-803	Japan
		1981	NMT	Nordic countries[5] Switzerland Netherlands Eastern Europe Russia
		1983	TACS[6]	UK Ireland Hong Kong
		1985	C450 (C-Netz)	F. R. Germany Portugal South Africa
2G	2G	1991	GSM	The World
		1993	PDC[7]	Japan
		1995	cdmaOne (IS-95)	Americas parts of Asia
	2.5G	1999	GPRS[8]	The World
	2.75G	2003	EDGE[9]	The World

3G	3G	2001	UMTS	Europe
				Japan
				China
		2002	CDMA2000	North America
				South Korea
	3.5G	2010	HSPA	The World
		2006	EV-DO	The World
	3.9G	2009	LTE	-
		2010	WiMAX	-
4G		-	LTE-Advanced	-
		-	WiMAX 2	-

Table A.1: Time line of evolution of mobile telephony standards

[1]Federal Republic of Germany, better known as West Germany, which existed from 1949 to 1990

[2]The abbreviation of 'Improved Mobile Telephone Service'.

[3]The abbreviation of 'Autoradiopuhelin'.

[4]The abbreviation of 'Automatizovaný městský radiotelefon'.

[5]The territorial concept of Northern Europe and the North Atlantic which consists of Denmark, Finland, Iceland, Norway and Sweden.

[6]The abbreviation of 'Total Access Communication System'.

[7]The abbreviation of 'Personal Digital Cellular'.

[8]The abbreviation of 'General Packet Radio Service'.

[9]The abbreviation of 'Enhanced Data rates for GSM Evolution'.

Appendix B

The Proof of Convexity of The Objective Function (4.12)

It can be proved that the objective function (4.12) is a convex optimisation problem, by proving the convexity of each C_k in (4.12). For the purpose of simplicity in the proof, $q_{1,k}^{(n)} = p_{1,k}^{(n)}/\xi_{1,k}^{(n)}$ and $q_{2,k,l}^{(n)} = p_{2,k,l}^{(n)}/\xi_{2,k,l}^{(n)}$ are defined, where $\xi_{1,k}^{(n)}$ and $\xi_{2,k,l}^{(n)}$ are relaxed to continuous variables which are in the interval of $[0, 1]$. Using (4.2), (4.4), (4.8) and $T_{1,k}C_{1,k} = T_{2,k}C_{2,k}$, it can be derived that

$$\frac{T_{1,k}}{T_{2,k}} = \frac{\sum_{l=1}^{L}\sum_{n=1}^{N}\xi_{2,k,l}^{(n)}\log_2\left(1+q_{2,k,l}^{(n)}\gamma_{2,k,l}^{(n)}\right)}{\sum_{n=1}^{N}\xi_{1,k}^{(n)}\log_2\left(1+q_{1,k}^{(n)}\gamma_{1,k}^{(n)}\right)} \tag{B.1}$$

And using (4.2), (4.4), (B.1) and $C_k = (T_{1,k}C_{1,k} + T_{2,k}C_{2,k})/2T$, the maximum achievable system throughput via relay k can be expressed as

$$C_k = \frac{B}{N}\left[\frac{1}{\sum_{n=1}^{N}\xi_{1,k}^{(n)}\log_2\left(1+q_{1,k}^{(n)}\gamma_{1,k}^{(n)}\right)} + \frac{1}{\sum_{l=1}^{L}\sum_{n=1}^{N}\xi_{2,k,l}^{(n)}\log_2\left(1+q_{2,k,l}^{(n)}\gamma_{2,k,l}^{(n)}\right)}\right] \tag{B.2}$$

The concavity of $\sum_{n=1}^{N}\xi_{1,k}^{(n)}\log_2\left(1+q_{1,k}^{(n)}\gamma_{1,k}^{(n)}\right)$ and $\sum_{l=1}^{L}\sum_{n=1}^{N}\xi_{2,k,l}^{(n)}\log_2\left(1+q_{2,k,l}^{(n)}\gamma_{2,k,l}^{(n)}\right)$ are proved first. Note that $\sum_{n=1}^{N}\xi_{1,k}^{(n)}\log_2\left(1+q_{1,k}^{(n)}\gamma_{1,k}^{(n)}\right)$ can be decomposed into N components, and the nth component is $\xi_{1,k}^{(n)}\log_2\left(1+q_{1,k}^{(n)}\gamma_{1,k}^{(n)}\right)$. Defining a function $f\left(\xi_{1,k}^{(n)}, p_{1,k}^{(n)}\right) = \xi_{1,k}^{(n)}\log_2\left(1+p_{1,k}^{(n)}\gamma_{1,k}^{(n)}/\xi_{1,k}^{(n)}\right)$, the Hessian matrix of $f\left(\xi_{1,k}^{(n)}, p_{1,k}^{(n)}\right)$ can be written as

$$\boldsymbol{H}(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial(\xi_{1,k}^{(n)})^2} & \frac{\partial^2 f}{\partial\xi_{1,k}^{(n)}\partial p_{1,k}^{(n)}} \\ \frac{\partial^2 f}{\partial p_{1,k}^{(n)}\partial\xi_{1,k}^{(n)}} & \frac{\partial^2 f}{\partial(p_{1,k}^{(n)})^2} \end{bmatrix} = \frac{1}{\ln 2}\begin{bmatrix} -\frac{(p_{1,k}^{(n)})^2}{\xi_{1,k}^{(n)}(\xi_{1,k}^{(n)}+p_{1,k}^{(n)})^2} & \frac{p_{1,k}^{(n)}}{(\xi_{1,k}^{(n)}+p_{1,k}^{(n)})^2} \\ \frac{p_{1,k}^{(n)}}{(\xi_{1,k}^{(n)}+p_{1,k}^{(n)})^2} & -\frac{\xi_{1,k}^{(n)}}{(\xi_{1,k}^{(n)}+p_{1,k}^{(n)})^2} \end{bmatrix} \tag{B.3}$$

Defining $\boldsymbol{r} = [r_1, r_2]$ as a non-zero real valued number vector, there is

$$\boldsymbol{r}\boldsymbol{H}(f)\boldsymbol{r}^T = -\frac{1}{\ln 2(\xi_{1,k}^{(n)}+p_{1,k}^{(n)})^2}\left(\frac{p_{1,k}^{(n)}}{\sqrt{\xi_{1,k}^{(n)}}}r_1 - \sqrt{\xi_{1,k}^{(n)}}r_2\right)^2 \tag{B.4}$$

It is obvious that when $0 < \xi_{1,k}^{(n)} < 1$ and $p_{1,k}^{(n)} > 0$, $\boldsymbol{r}\boldsymbol{H}(f)\boldsymbol{r}^T < 0$ always holds, which means that $\boldsymbol{H}(f)$ is a negative definite matrix. Thus, it can be concluded that $f\left(\xi_{1,k}^{(n)}, p_{1,k}^{(n)}\right)$ is concave. Accordingly, $\sum_{n=1}^{N} \xi_{1,k}^{(n)} \log_2\left(1 + p_{1,k}^{(n)}\gamma_{1,k}^{(n)}\right)$ is concave. Similarly, $\sum_{l=1}^{L}\sum_{n=1}^{N} \xi_{2,k,l}^{(n)} \log_2\left(1 + p_{2,k,l}^{(n)}\gamma_{2,k,l}^{(n)}\right)$ is concave. Therefore, (B.2) is convex, and (4.12) is convex objective function, which means that the globe optimal solution to (4.12) can be obtained, as given by (4.28)-(4.19).

Appendix C

Solution to the Quadratic Inequality (5.25)

In this part, the derivation of k is demonstrated. Substituting $x = \sqrt{k}$ into (5.25), the quadratic inequality can be rewritten as:

$$ax^2 + bx + c \geq 0 \tag{C.1}$$

where

$$a = D\left(\sigma_G^2 || \sigma_{0\_eq}^2\right) = \frac{1}{2}\left(\frac{\sigma_G^2}{\sigma_{0\_eq}^2} - 1 - \ln\frac{\sigma_G^2}{\sigma_{0\_eq}^2}\right) \tag{C.2}$$

$$b = \left[\frac{3}{2} + 2\ln\frac{2}{h_b^{-1}\left(\delta\left(\sigma_G^2\right)\right)}\right]\left(\frac{\sigma_G^2}{\sigma_{0\_eq}^2} - 1\right) \tag{C.3}$$

$$c = \ln\frac{2P_e}{h_b^{-1}\left(\delta\left(\sigma_G^2\right)\right)} \tag{C.4}$$

When $b^2 - 4ac \geq 0$, the lower bound of k can be written as

$$k \geq \max\left[1, \left(\frac{-b + \sqrt{b^2 - 4ac}}{2a}\right)^2\right] = \max\left[1, \sup_{\sigma_G^2 : C(\sigma_G^2) < R_{ch}} K\left(P_{rec}, \sigma_G^2\right)\right] \tag{C.5}$$

Due to the negativity of $-b/2a$, the regime $k \leq (-b - \sqrt{b^2 - 4ac})/2a$ should not be selected. As a result, it can be seen that k is a function of P_{rec} and σ_G^2.

Appendix D

Curve Fitting of the PA

It is assumed that the output power is less than $P_{1\text{dB}}$, which indicates the PA works under the linear amplification region as $P_{out} = |\beta|^2 P_{in}$. From [1], it can be observed that the PAE varies with the input power. From Fig. 15 in [1], the data of PAE and P_{in} are estimated when 1.8 V DC voltage is applied, which is shown in Table D.1.

Table D.1: PAE vs. P_{in}, at 1.8 V DC voltage

P_{in} (dBm)	P_{in} (mW)	PAE (%)
-18	0.0158	0.7
-17	0.0200	0.8
-16	0.0251	1
-15	0.0316	1.3
-14	0.0398	1.8
-13	0.0501	2.1
-12	0.0631	2.8
-11	0.0794	3.6
-10	0.1000	4.8
-9	0.1259	6.9
-8	0.1585	10.2
-7	0.1995	13.9
-6	0.2512	16.2
-5	0.3162	17.5
-4	0.3981	19.5
-3	0.5012	21
-2	0.6310	22.5

Since no explicit function of PAE with the variable of the input power is available, LS method is applied to get an estimated empirical formula of PAE. x_i $(i = 0, ..., 16)$ is defined as each input power from Table D.1 in watt, and y_i $(i = 1, ..., 16)$ is defined as each PAE value. From the shape of curve in Fig. 15 of [1], it is predicted that the empirical formula as $y = a(x - 5.89 \times 10^{-4})^2 + b$ where a and b are non-zero real numbers. r_i $(i = 1, ..., 16)$ is defined as the residual between each x_i and y_i, and there

is

$$r_i = y_i - a\left(x_i - 5.89 \times 10^{-4}\right)^2 - b, \forall\left(x_i, y_i\right) \tag{D.1}$$

The predicted empirical should have the minimum sum of the squared residuals, and the objective function can be written as:

$$\min_{a,b} S = \sum_{i=0}^{16} |r_i|^2 \tag{D.2}$$

subject to:

$$a > 1, \text{and } 0 < b < 1 \tag{D.3}$$

where constraint (D.3) can be observed from the figure comparing with $y = a(x - 5.89 \times 10^{-4})^2 + b$. $f_1 = \partial S/\partial a$ and $f_2 = \partial S/\partial b$ are defined as the partial derivatives of a and b, respectively. To get the minimum value, $\partial S/\partial a^*|_{a=a^*} = 0$ is set and along with $\partial S/\partial b^*|_{b=b^*} = 0$, where a^* and b^* are the optimal values. The equations can be written as:

$$\begin{cases} f_1 = -2\sum_{i=0}^{16}[y_i - a\left(x_i - 5.89 \times 10^{-4}\right)^2 - b]\left(x_i - 5.89 \times 10^{-4}\right)^2 = 0 \\ f_2 = -2\sum_{i=0}^{16}[y_i - a\left(x_i - 5.89 \times 10^{-4}\right)^2 - b] = 0 \end{cases} \tag{D.4}$$

The sub-gradient methods is utilised to solve equations (D.4). t is defined as the index of iteration, and $a^{(t)}$ and $b^{(t)}$ are defined as the tth trials. Now update a and b as the following:

$$\begin{cases} a^{(t+1)} = a^{(t)} - \Delta a f_1|_{a=a^{(t)}} \\ b^{(t+1)} = b^{(t)} - \Delta b f_2|_{b=b^{(t)}} \end{cases} \tag{D.5}$$

until $|f_1| < \varepsilon_1$ and $|f_2| < \varepsilon_2$, where ε_1 and ε_2 are non-zero criteria. Δa and Δb are the step sizes. Table D.2 shows the setup of parameters of the LS method.

Table D.2: Sub-gradient parameters setup

Initial value $a^{(1)}$	-1
Initial value $b^{(1)}$	1
Δa	$(t+1)^2$
Δb	10^{-4}
ε_1	10^{-8}
ε_2	10^{-8}

The approximated optimal values are obtained as $a^* = -6.8663 \times 10^5$, and $b^* = 0.2249$. Hence the approximated empirical formula can be written as:

$$\text{PAE} = -6.8663 \times 10^5 \cdot \left(P_{in} - 5.89 \times 10^{-4}\right)^2 + 0.2249, \ |\beta|^2 P_{in} \leq P_{\text{1dB}} \tag{D.6}$$

Fig. D.1 shows the curve fitting with the data from [1]. It can be seen from Fig. D.1 that the fitting is close to the real data, and (5.29) can be applied in the system objective function.

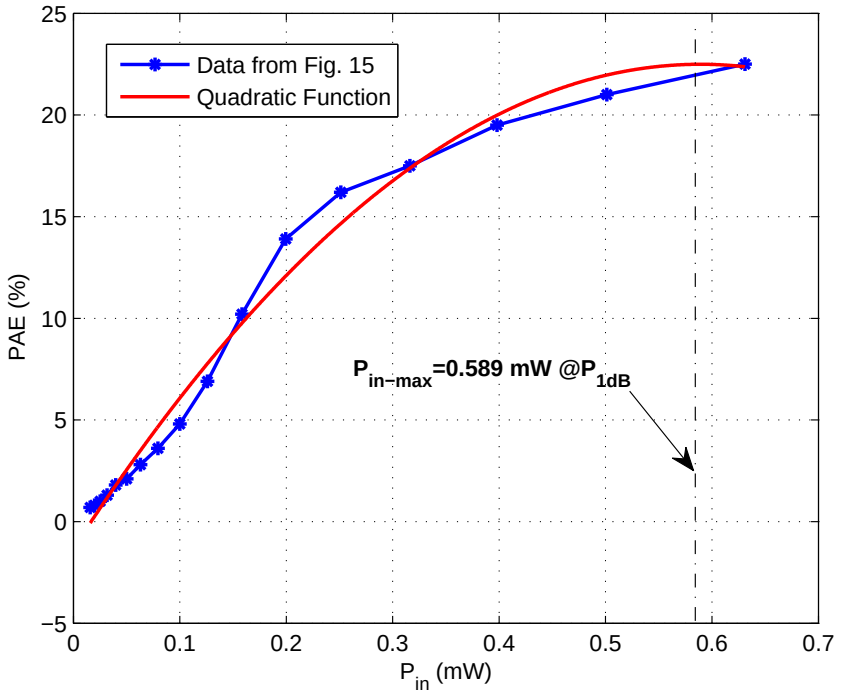

Figure D.1: The curve fitting vs. data from [1]

Bibliography

[1] A. Siligaris, Y. Hamada, C. Mounet, C. Raynaud, B. Martineau, N. Deparis, N. Rolland, M. Fukaishi, and P. Vincent, "A 60 GHz Power Amplifier With 14.5 dBm Saturation Power and 25% Peak PAE in CMOS 65 nm SOI," *IEEE J. Solid-State Circuits*, vol. 45, no. 7, pp. 1286–1294, Jul. 2010.

[2] A. F. Molisch, *Wireless Communications, 2nd Edition.* John Wiley & Sons Ltd., 2011.

[3] T. S. Rappaport, *Wireless Communications.* Prentice Hall, 2002.

[4] J. Jang and K. B. Lee, "Transmit Power Adaptation for Multiuser OFDM Systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 2, pp. 171–178, Feb. 2003.

[5] X. Gao, G. Wu, and T. Miki, "End-to-End QoS Provisioning in Mobile Heterogeneous Networks," *IEEE Wireless Commun. Mag.*, vol. 11, no. 3, pp. 24–34, Jun. 2004.

[6] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with Adaptive Subcarrier, Bit, and Power Allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.

[7] T. C.-Y. Ng and W. Yu, "Joint Optimization of Relay Strategies and Resource Allocations in Cooperative Cellular Networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 328–339, Feb. 2007.

[8] Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green Cellular Networks: A Survey, Some Research Issues and Challenges," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 4, pp. 524–540, Fourth Quarter 2011.

[9] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental Trade-offs on Green Wireless Networks," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 30–37, Jun. 2011.

[10] S. Yang and J.-C. Belfiore, "Towards the Optimal Amplify-and-Forward Cooperative Diversity Scheme," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3114–3126, Sep. 2008.

[11] T. T. Kim, G. Caire, and M. Skoglund, "Decode-and-Forward Relaying With Quantized Channel State Feedback: An Outage Exponent Analysis," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4548–4564, Oct. 2008.

[12] M. Iwamura, H. Takahashi, and S. Nagata, "Relay Technology in LTE-Advanced," *NTT DOCOMO Technical Journal*, vol. 12, no. 2, pp. 29–36, Feb. 2010.

[13] Z. Shen, J. G. Andrews, and B. L. Evans, "Adaptive Resource Allocation in Multiuser OFDM Systems With Proportional Rate Constraints," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2726–2737, Nov. 2005.

[14] R. Devarajan, A. Punchihewa, and V. K. Bhargav, "Energy-Aware Power Allocation in Cooperative Communication Systems with Imperfect CSI," *IEEE Trans. Commun.*, vol. 61, no. 5, pp. 1633–1639, May 2013.

[15] C. Park and T. S. Rappaport, "Short-Range Wireless Communications for Next-Generation Networks: UWB, 60 GHz Millimeter-Wave WPAN, And ZigBee," *IEEE Wireless Commun. Mag.*, vol. 14, no. 4, pp. 70–78, Aug. 2007.

[16] R. C. Daniels and R. W. Heath, "60 GHz Wireless Communications: Emerging Requirements and Design Recommendations," *IEEE Veh. Technol. Mag.*, vol. 2, no. 3, pp. 41–50, Sep. 2007.

[17] J. Foerster, J. Lansford, J. Laskar, T. S. Rappaport, and S. Kato, "Realizing Gbps Wireless Personal Area Networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1313–1317, Oct. 2009.

[18] J. Qiao, L. X. Cai, X. S. Shen, and J. W. Mark, "Enabling Multi-Hop Concurrent Transmissions in 60 GHz Wireless Personal Area Networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3824–3833, Nov. 2011.

[19] Z. Lan, C.-S. Sum, J. Wang, T. Baykas, F. Kojima, H. Nakase, and H. Harada, "Relay with Deflection Routing for Effective Throughput Improvement in Gbps Millimeter-Wave WPAN Systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1453–1465, Oct. 2009.

[20] N. Zhou, X. Zhu, and Y. Huang, "Optimal Asymmetric Resource Allocation and Analysis for OFDM-Based Multidestination Relay Systems in the Downlink," *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 1307–1312, Mar. 2011.

[21] M. Salem, A. Adinoyi, M. Rahman, H. Yanikomeroglu, D. Falconer, Y.-D. Kim, E. Kim, and Y.-C. Cheong, "An Overview of Radio Resource Management in Relay-Enhanced OFDMA-Based Networks," *IEEE Commun. Surveys Tuts.*, vol. 12, no. 3, pp. 442–438, Third Quarter 2010.

[22] P. Grover, K. Woyach, and A. Sahai, "Towards a Communication-Theoretic Understanding of System-Level Power Consumption," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1744–1755, Sep. 2011.

[23] T. T. Chen, "On the Definition of the Effective Aperture of Antennas," *IRE Trans. Antenna and Propagation*, vol. 9, no. 2, pp. 224–225, Mar. 1961.

[24] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[25] M. Dohler and Y. Li, *Cooperative Communications - Hardware, Channel & PHY.* John Wiley & Sons, Inc., 2010.

[26] J. Proakis, *Digital Communications, 3rd Edition.* McGraw-Hill Inc., 1995.

[27] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms.* Cambridge University Press, 2005.

[28] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, Jul. Oct. 1948.

[29] R. G. Gallager, *Information Theory and Reliable Communication.* John Wiley & Sons, Inc., 1968.

[30] A. J. Goldsmith and P. P. Varaiya, "Capacity of Fading Channels with Channel Side Information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.

[31] J. C. Maxwell, "A Dynamical Theory of the Electromagnetic Field," *Philosophical Trans. of the Royal Society of London*, vol. 155, pp. 459–512, 1865.

[32] J. J. Fahie, *A History of Wireless Telegraphy, 2nd Edition, revised.* William Blackwood & Sons, 1901.

[33] H. Hertz and D. E. Jones, *Electric Waves: Being Researches on the Propagation of Electric Action with Finite Velocity Through Space.* MacMillan and Co., 1900.

[34] (29 Jun. 2013) 1946: First Mobile Telephone Call. [Online]. Available: http://www.corp.att.com/attlabs/reputation/timeline/46mobile.html

[35] R. H. Frenkiel, "A High-Capacity Mobile Radiotelephone System Model Using a Coordinated Small-Zone Approach," *IEEE Trans. Veh. Technol.*, vol. 19, no. 2, pp. 173–177, May 1970.

[36] (27 Sep. 2007) Subscriber Statistics end Q1 2007. [Online]. Available: http://www.gsmworld.com/news/statistics/pdf/

[37] E. Berruto, M. Gudmundson, R. Menolascino, W. Mohr, and M. Pizarroso, "Research Activities on UMTS Radio Interface, Network Architectures, and Planning," *IEEE Commun. Mag.*, vol. 36, no. 2, pp. 82–95, Feb. 1998.

[38] D. Knisely, S. Kumar, S. Laha, and S. Nanda, "Evolution of Wireless Data Services: IS-95 to cdma2000," *IEEE Commun. Mag.*, vol. 36, no. 10, pp. 140–149, Oct. 1998.

[39] R. Attar, D. Ghosh, C. Lott, M. Fan, P. Black, R. Rezaiifar, and P. Agashe, "Evolution of cdma2000 Cellular Networks: Multicarrier EV-DO," *IEEE Commun. Mag.*, vol. 44, no. 3, pp. 46–53, Mar. 2006.

[40] H. Holma, A. Toskala, K. Ranta-aho, and J. Pirskanen, "High-Speed Packet Access Evolution in 3GPP Release 7," *IEEE Commun. Mag.*, vol. 45, no. 12, pp. 29–35, Dec. 2007.

[41] K. Johansson, J. Bergman, D. Gerstenberger, M. Blomgren, and A. Wallén, "Multi-Carrier HSPA Evolution," in *Proc. IEEE VTC Spring'09*, Barcelona, Spain, Apr. 2009.

[42] "Overview of 3GPP Release 10 V0.1.8," Tech. Rep., Mar. 2013. [Online]. Available: http://www.3gpp.org/ftp/Information/

[43] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, "Coordinated Multipoint Transmission and Reception in LTE-Advanced: Deployment Scenarios and Operational Challenges," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 148–155, Feb. 2012.

[44] Z. Pi and F. Khan, "An Introduction to Millimeter-Wave Mobile Broadband Systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.

[45] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!" *IEEE Access*, vol. 1, pp. 335–349, May. 2013.

[46] "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification," IEEE Std. 802.11, 2012.

[47] T. Jiang and Y. Wu, "An Overview: Peak-to-Average Power Ratio Reduction Techniques for OFDM Signals," *IEEE Trans. Broadcast.*, vol. 54, no. 2, pp. 257–268, Jun. 2008.

[48] J. Zhang, H. Rohling, and P. Zhang, "Analysis of ICI Cancellation Scheme in OFDM Systems with Phase Noise," *IEEE Trans. Broadcast.*, vol. 50, no. 2, pp. 97–106, Jun. 2004.

[49] W.-S. Hou and B.-S. Chen, "ICI Cancellation for OFDM Communication Systems in Time-Varying Multipath Fading Channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2100–2110, Sep. 2005.

[50] M.-X. Chang, "A Novel Algorithm of Inter-Subchannel Interference Self-Cancellation for OFDM Systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2881–2893, Aug. 2007.

[51] A. Ghosh, R. Patasuk, B. Mondal, N. Mangalvedhe, and T. Thomas, "LTE-Advanced: Next-Generation Wireless Broadband Technology," *IEEE Wireless Commun. Mag.*, vol. 17, no. 3, pp. 11–22, Jun. 2010.

[52] R. W. Chang, "Synthesis of Band-Limited Orthogonal Signals for Multichannel Data Transmission," *Bell System Tech. J.*, vol. 46, pp. 1775–1796, Dec. 1966.

[53] X. Zhu, "Wireless MIMO Antenna Systems for Frequency Selective Fading Channels," PhD thesis, Hong Kong University of Science and Technology, Apr. 2003.

[54] H. Sari, Y. Levy, and G. Karam, "An Analysis Of Orthogonal Frequency-Division Multiple Access," in *Proc. IEEE GLOBECOM'97*, vol. 3, Phoenix, USA, Nov. 1997, pp. 1635–1639.

[55] "IEEE Standard for Air Interface for Broadband Wireless Access Systems," IEEE Std. 802.16, 2012.

[56] D. Kivanc, G. Li, and H. Liu, "Computationally Efficient Bandwidth Allocation and Power Control for OFDMA," *IEEE Trans. Wireless Commun.*, vol. 2, no. 6, pp. 1150–1158, Nov. 2003.

[57] H. Zhu and J. Wang, "Chunk-Based Resource Allocation in OFDMA Systems - Part I: Chunk Allocation," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2734–2744, Sep. 2009.

[58] ——, "Chunk-Based Resource Allocation in OFDMA Systems - Part II: Joint Chunk, Power and Bit Allocation," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 499–509, Feb. 2012.

[59] V. D. Papoutsis and S. A. Kotsopoulos, "Chunk-Based Resource Allocation in Multicast OFDMA Systems with Average BER Constraint," *IEEE Commun. Lett.*, vol. 15, no. 5, pp. 551–553, May 2011.

[60] C. Xiong, G. Y. Li, S. Zhang, Y. Chen, and S. Xu, "Energy- and Spectral-Efficiency Tradeoff in Downlink OFDMA Networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3874–3886, Nov. 2011.

[61] K. Jitvanichphaibool, R. Zhang, and Y.-C. Liang, "Optimal Resource Allocation for Two-Way Relay-Assisted OFDMA," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3311–3321, Sep. 2009.

[62] T. Wang and L. Vandendorpe, "WSR Maximized Resource Allocation in Multiple DF Relays Aided OFDMA Downlink Transmission," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3964–3976, Aug. 2011.

[63] D. W. K. Ng, E. S. Lo, and R. Schober, "Dynamic Resource Allocation in MIMO-OFDMA Systems with Full-Duplex and Hybrid Relaying," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1291–1304, May 2012.

[64] M. Salem, A. Adinoyi, M. Rahman, H. Yanikomeroglu, D. Falconer, and Y.-D. Kim, "Fairness-Aware Radio Resource Management in Downlink OFDMA Cellular Relay Networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1628–1639, May 2010.

[65] E. C. V. D. Meulen, "Three-Terminal Communication Channels," *Adv. Appl. Prob.*, vol. 3, no. 1, pp. 120–154, Spring 1971.

[66] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[67] C. Politis, T. Oda, S. Dixit, A. Schieder, H.-Y. Lach, M. I. Smirnov, S. Uskela, and R. Tafazolli, "Cooperative Networks for the Future Wireless World," *IEEE Commun. Mag.*, vol. 42, no. 9, pp. 70–79, Sep. 2004.

[68] A. Sendonaris, E. Erkip, and B. Aazhang, "User Cooperation Diversity - Part I: System Description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[69] ——, "User Cooperation Diversity - Part II: Implementation Aspects and Performance Analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[70] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative Strategies and Capacity Theorems for Relay Networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[71] M. Benjillali and L. Szczecinski, "A Simple Detect-and-Forward Scheme in Fading Channels," *IEEE Commun. Lett.*, vol. 13, no. 5, pp. 309–311, May 2009.

[72] L. Sanguinetti, A. A. D'Amico, and Y. Rong, "A Tutorial on the Optimization of Amplify-and-Forward MIMO Relay Systems," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1331–1346, Sep. 2012.

[73] K. Ishibashi and H. Ochiai, "Performance Analysis of Amplify and Forward Cooperation over Peak-Power Limited Channels," in *Proc. IEEE ICC'11*, Kyoto, Japan, Jun. 2011.

[74] "United Kingdom Frequency Allocation Table," the National Frequency Planning Group, 2010.

[75] F. K. Jondral, "Cognitive Radio: A Communications Engineering View," *IEEE Wireless Commun. Mag.*, vol. 14, no. 4, pp. 28–33, Aug. 2007.

[76] S. Glisic, Z. Nikolic, N. Milosevic, and A. Pouttu, "Advanced Frequency Hopping Modulation for Spread Spectrum WLAN," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 1, pp. 16–29, Jan. 2000.

[77] I. Howitt, "WLAN and WPAN Coexistence in UL Band," *IEEE Trans. Veh. Technol.*, vol. 50, no. 4, pp. 1114–1124, Jul. 2001.

[78] R. Etkin, A. Parekh, and D. Tse, "Spectrum Sharing for Unlicensed Bands," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 517–528, Apr. 2007.

[79] Y. M. Shobowale and K. A. Hamdi, "A Unified Model for Interference Analysis in Unlicensed Frequency Bands," *IEEE Trans. Wireless Commun.*, vol. 9, no. 8, pp. 4004–4013, Aug. 2009.

[80] R. S. Elliott, *Antenna Theory and Design*. John Wiley & Sons, Inc., 2003.

[81] J. J. R. Deller, J. H. L. Hansen, and J. G. Proakis, *Discrete-Time Processing of Speech Signals*. John Wiley & Sons, Inc., 2000.

[82] S. C. Cripps, *RF Power Amplifiers for Wireless Communications, 2nd Edition*. Artech House, Inc., 2006.

[83] G. Yue and X. Wang, "A Hybrid PAPR Reduction Scheme for Coded OFDM," *IEEE Trans. Wireless Commun.*, vol. 5, no. 10, pp. 2712–2722, Oct. 2006.

[84] J. Gao, X. Zhu, and A. K. Nandi, "Non-Redundant Precoding and PAPR Reduction in MIMO OFDM Systems with ICA Based Blind Equalization," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 3038–3049, Jun. 2009.

[85] Y. Jiang, "New Companding Transform for PAPR Reduction in OFDM," *IEEE Commun. Lett.*, vol. 14, no. 4, pp. 282–284, Apr. 2010.

[86] B. Wang, P.-H. Ho, and C.-H. Lin, "OFDM PAPR Reduction by Shifting Null Subcarriers Among Data Subcarriers," *IEEE Commun. Lett.*, vol. 16, no. 9, pp. 1377–1379, Sep. 2012.

[87] T. J. Willink and P. H. Wittke, "Optimization and Performance Evaluation of Multicarrier Transmission," *IEEE Trans. Inf. Theory*, vol. 43, no. 8, pp. 426–440, Aug. 1997.

[88] G. Song, "Cross-Layer Resource Allocation and Scheduling in Wireless Multi-Carrier Networks," PhD thesis, Georgia Institute of Technology, Aug. 2005.

[89] M. Iwamura, K. Etemad, M.-H. Fong, R. Nory, and R. Love, "Carrier Aggregation Framework in 3GPP LTE-Advanced," *IEEE Commun. Mag.*, vol. 48, no. 8, pp. 60–67, Aug. 2010.

[90] G. Yuan, X. Zhang, W. Wang, and Y. Yang, "Carrier Aggregation for LTE-Advanced Mobile Communication Systems," *IEEE Commun. Mag.*, vol. 48, no. 2, pp. 88–93, Feb. 2010.

[91] Z. Shen, A. Papasakellariou, J. Montojo, D. Gerstenberger, and F. Xu, "Overview of 3GPP LTE-Advanced Carrier Aggregation for 4G Wireless Communications," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 122–130, Feb. 2012.

[92] K.-P. Ho and J. M. Kahn, "Transmission of Analog Signals Using Multicarrier Modulation: A Combined Source-Channel Coding Approach," *IEEE Trans. Commun.*, vol. 44, no. 11, pp. 1432–1443, Nov. 1996.

[93] I. Kalet, "The Multitone Channel," *IEEE Trans. Commun.*, vol. 37, no. 2, pp. 119–124, Feb. 1989.

[94] P. W. C. Chan, E. S. Lo, R. R. Wang, E. K. S. Au, V. K. N. Lau, R. S. Cheng, W. H. Mow, R. D. Murch, and K. B. Letaief, "The Evolution Path of 4G Networks: FDD or TDD?" *IEEE Commun. Mag.*, vol. 40, no. 12, pp. 42–50, Dec. 2006.

[95] P. He, L. Zhao, S. Zhou, and Z. Niu, "Water-Filling: A Geometric Approach and its Application to Solve Generalized Radio Resource Allocation Problems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3637–3647, Jul. 2013.

[96] Y. Li, W. Wang, J. Kong, W. Hong, X. Zhang, and M. Peng, "Power Allocation and Subcarrier Pairing in OFDM-Based Relaying Networks," in *Proc. IEEE ICC'08*, Beijing, China, May 2008.

[97] W. Y. Wang, S. F. Yan, and S. F. Yang, "Optimally Joint Subcarrier Matching and Power Allocation in OFDM Multihop System," *EURASIPJ. Adv. Signal Process.*, vol. 2008, pp. 1–8, Jan. 2008.

[98] W. Nam, W. Chang, S.-Y. Chung, and Y. H. Lee, "Transmit Optimization for Relay-based Cellular OFDMA Systems," in *Proc. IEEE ICC'07*, Glasgow, U.K., Jun. 2007.

[99] A. Agustin, J. Vidal, and O. Muñoz, "Protocols and Resource Allocation for the Two-Way Relay Channel with Half-Duplex Terminals," in *Proc. IEEE ICC'09*, Dresden, Germany, Jun. 2009.

[100] J. Joung and S. Sun, "Power Efficient Resource Allocation for Downlink OFDMA Relay Cellular Networks," *IEEE Signal Process. Lett.*, vol. 60, no. 5, pp. 2447–2459, May 2012.

[101] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge Univiversity Press, 2004.

[102] (4 Apr. 2006) Multihop Relay System Evaluation Methodology. [Online]. Available: http://ieee802.org/16/relay/docs/

[103] H. Zirath, T. Masuda, R. Kozhuharov, and M. Ferndahl, "Development of 60-GHz Front-End Circuits for a High-Data-Rate Communication System," *IEEE J. Solid-State Circuits*, vol. 39, no. 10, pp. 1640–1649, Oct. 2004.

[104] P.-Y. Chang, S.-H. Su, S. S. H. Hsu, W.-H. Cho, and J.-D. Jin, "An Ultra-Low-Power Transformer-Feedback 60 GHz Low-Noise Amplifier in 90 nm CMOS," *IEEE Microw. Wireless Compon. Lett.*, vol. 22, no. 4, pp. 197–199, Apr. 2012.

[105] (20 Apr. 2013) IEEE 802.11 ad Wireless Gigabit Alliance Task Group. [Online]. Available: http://wirelessgigabitalliance.org/specifications/

[106] (20 Apr. 2013) IEEE 802.15 WPAN Millimeter Wave Alternative PHY Task Group 3c (TG3c.). [Online]. Available: http://www.ieee802.org/15/pub/TG3c.html

[107] S. Singh, F. Ziliotto, U. Madhow, E. Belding, and M. Rodwell, "Blockage and Directivity in 60 GHz Wireless Personal Area Networks: From Cross-Layer Model to Multihop MAC Design," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1400–1413, Oct. 2009.

[108] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-Constrained Modulation Optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sep. 2005.

[109] C. D. Thompson, "Area-Time Complexity for VLSI," in *Proc. ACM STOC'79*, Atlanta, USA, May 1979.

[110] Q. Liu, W. Zhang, X. Ma, and G. T. Zhou, "Designing Peak Power Constrained Amplify-and-Forward Relay Networks with Cooperative Diversity," *IEEE Trans. Wireless Commun.*, vol. 11, no. 5, pp. 1733–1743, May 2012.

[111] T. Wang, A. Cano, G. B. Giannakis, and J. N. Laneman, "High-Performance Cooperative Demodulation With Decode-and-Forward Relays," *IEEE Trans. Wireless Commun.*, vol. 55, no. 7, pp. 1427–1438, Jul. 2007.

[112] S. Sun, "Wireless Network Coding for Cooperative Communications," in *Proc. IEEE ICCS'10*, Singapore, Nov. 2010.

[113] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing, Third Edition.* New York: Cambridge University Press, 2007.

[114] S. Geng, J. Kivinen, X. Zhao, and P. Vainikainen, "Millimeter-Wave Propagation Channel Characterization for Short-Range Wireless Communications," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1.