COOPERATIVE STRATEGIES, ACHIEVABLE RATES AND RESOURCE ALLOCATION FOR OFDMA CHANNELS

İSMAİL SEZİ BAKIM

B.S., Electronics Engineering, Işık University, 2001

M.S., Electrical Engineering, University of New Mexico, 2003

Submitted to the Graduate School of Science and Engineering
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in
Electronics Engineering

IŞIK UNIVERSITY

2011

IŞIK UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

COOPERATIVE STRATEGIES, ACHIEVABLE RATES AND RESOURCE ALLOCATION FOR OFDMA CHANNELS

İSMAİL SEZİ BAKIM

APPROVED BY:

Assist. Prof. Onur Kaya (Thesis Supervisor)	Işık University	_____
Assoc. Prof. Hasan F. Ateş	Işık University	_____
Prof. Hakan A. Çırpan	İstanbul Technical University	_____
Assist. Prof. Mutlu Koca	Boğaziçi University	_____
Assoc. Prof. Mengüç Öner	Işık University	_____

APPROVAL DATE: 28/07/2011

COOPERATIVE STRATEGIES, ACHIEVABLE RATES AND RESOURCE ALLOCATION FOR OFDMA CHANNELS

Abstract

The design of next generation wireless communication systems brings along new challenges, since the degrading factors such as fading and multi-user interference become harder to deal with as the number of users and the bandwidth requirements increase. Orthogonal Frequency Division Multiple Access (OFDMA) is a multiple accessing technique which provides a solution to both of the problems above: it provides a relatively simple way of assigning available bandwidth to users, while avoiding interference; and at the same time, it converts a frequency selective fading channel, to parallel flat fading subchannels, hence reducing the effects of intersymbol interference. However, in wireless channels, what is traditionally considered as interference is in fact side information, and combined with the diversity created by the orthogonal subchannels in OFDMA, this side information can be carefully taken advantage of to increase the rates achievable by the users. In our thesis, without imposing any prior constraints on subchannel allocation, we investigate cooperation strategies, achievable rates and resource allocation for OFDMA channels.

We propose new cooperative encoding strategies for wireless communication networks over OFDMA channels. We particularly focus on a two user cooperative OFDMA system, based on block Markov superposition encoding (BMSE). We obtain expressions of the resulting achievable rate regions for all proposed cooperative encoding strategies. We show that, by allowing for re-partitioning and re-encoding of the cooperative messages across subchannels, it is possible to better exploit the diversity created by OFDMA, and higher rates can be achieved.

In order to take full advantage of the diversity created by OFDMA, we then introduce a channel adaptive cooperation strategy for OFDMA, and optimize the transmit powers as a function of the channel states. We provide the optimality conditions that need to be satisfied by the powers associated with the users' codewords and derive the closed form expressions for the optimal powers. We

propose two algorithms that can be used to optimize the powers to achieve any desired rate point on the rate region boundary. We observe that, utilization of power control to take advantage of the diversity offered by the cooperative OFDMA system, not only leads to a remarkable improvement in achievable rates, but also may help determine how the subchannels have to be instantaneously allocated to various tasks in cooperation.

DFBÇE KANALLARI İÇİN İŞBİRLİKÇİ STRATEJİLER, ERİŞİLEBİLİR VERİ HIZLARI VE KAYNAK TAHSİSİ

Özet

Yeni nesil kablosuz haberleşme sistemlerinde, kullanıcı sayısı ve bantgenişliği talebindeki artış sebebiyle, sönümlenme ve çoklu-kullanıcı karışımı gibi sorunlarla başa çıkılması ve uygun protokollerin tanımlaması gittikçe zorlaşmaktadır. Dik Frekans Bölmeli Çoklu Erişim (DFBÇE) yukarıdaki iki probleme de çözüm sağlayan bir çoklu erişim tekniğidir: Kullanıcılara sahip olunan frekans tayfını dik olarak dağıtmak suretiyle karışımı engellerken aynı zamanda frekans seçici sönümlenmeli bir kanalı, paralel düz sönümlü alt kanallara çevirir, böylelikle semboller arası karışımın etkilerini azaltır. Hâlbuki kablosuz kanallarda, karışım olarak adlandırılan şey aslında yan bilgidir ve bu yan bilgi DFBÇE'nin ortogonal alt kanallarının çeşitliliği ile uygun bir şekilde birleştirildiğinde, kullanıcıların erişebilecekleri veri hızlarını arttırmak için kullanılabilir. Tezimizde, kanal ataması ile ilgili herhangi bir önkoşul koymadan, DFBÇE kanalları için işbirlikçi stratejileri, erişilebilir veri hızlarını ve kaynak tahsis tekniklerini inceliyoruz.

Önce, DFBÇE altyapısı kullanan kablosuz haberleşme şebekeleri için yeni işbirlikçi kodlama stratejileri öneriyoruz. Özellikle, blok Markov bindirmeli kodlama uygulayan iki kullanıcılı DFBÇE sistemine yoğunlaşıyoruz. Önerdiğimiz tüm işbirlikçi kodlama stratejileri için erişilebilir veri hızı bölgelerini elde ediyoruz. Mesajların alt kanallara yeniden bölüştürülmesi ve yeniden kodlanmasına müsaade edilmesi ile, DFBÇE'nin beraberinde getirdiği çeşitlilikten daha fazla faydalanılabileceğini ve daha yüksek veri hızlarına ulaşılabileceğini gösteriyoruz.

Daha sonra, DFBÇE'nin oluşturduğu çeşitliliğin getirdiği faydalardan tam olarak yararlanmak amacıyla, DFBÇE için kanal uyarlamalı bir işbirlikçi strateji öneriyoruz ve çıkış güçlerini kanal durumlarının fonksiyonu olarak eniyiliyoruz. Kullanıcıların kodkelimelerine ilişkin güçlerinin sağlaması gereken eniyilik koşullarını buluyoruz ve en iyi güç değerleri için kapalı form denklemlerini türetiyoruz. Veri hızı bölgesinin sınırında herhangi bir noktaya ulaşılmasını sağlayan güçleri

eniyilemek için kullanılabilecek iki algoritma öneriyoruz. İşbirlikçi DFBÇE sisteminin sunduğu çeşitlilikten yararlanan güç kontrolünü uygulamanın sadece erişilebilir veri hızlarında önemli bir iyileşme sağlamakla kalmadığını, aynı zamanda anlık olarak hangi alt kanalın hangi amaç için kullanılması gerektiğini belirlemede de fayda sağlayabileceğini kanıtlıyoruz.

Acknowledgements

I would like to thank my mother, Zinet Demirağ Bakım for always being there with me, always showing me the right way and her endless love and support. I am grateful to my sister, Sevi Bakım who has always had confidence in me and helped me so much. I thank my father, Mustafa Bakım.

If there is a success in this thesis for my side, that belongs to my family. I dedicate this thesis to my mother, Zinet Demirağ Bakım.

I am so glad that I had a chance to work with my advisor, Assist. Prof. Onur Kaya. He has always been an advisor much better than I could imagine to have. I have been able to write this thesis with his very valuable guidance and his generosity in sharing his deep knowledge in my research area. I also would like to thank Assist. Prof. Onur Kaya's family who always showed great patience and understanding during our long research and discussion hours at night after work and in the weekends.

I would like to thank Assoc. Prof. Hasan Fehmi Ateş, Prof. Hakan Ali Çırpan, Assist. Prof. Mutlu Koca and Assoc. Prof. Mustafa Mengüç Öner for being in my PhD thesis defense committee and for their very valuable guidance and suggestions.

I would like to thank Prof. Edl Schamiloglu, Prof. Sıddık Yarman, Prof. Erdal Panayırcı, Prof. Mithat İdemen, Prof. Vural Altın, Prof. Chaouki Abdallah, Prof. Christos Christodoulou, Prof. Yorgo İstefanopulos, Prof. Tolga Yarman, Semiral Özdemir and my all teachers, instructors who have taught me so much.

I am also grateful to my company, Turkcell, my managers, Ekrem Özorbeyi, Bahadır Üçer, Eyüp Dilaverler, Cem Tığcı, Yüksel Yılmaz and my colleagues at Turkcell, who always supported me during my entire PhD. Without their support and understanding, I would not be able to achieve the PhD degree during my full-time work at Turkcell.

Finally, I would like to acknowledge the Scientific and Technological Research Council of Turkey (TÜBİTAK) for the grant, 108E208 and the Domestic PhD Scholarship Program, 2211.

To my mother, Zinet Demirağ Bakım...

Table of Contents

List of Figures

List of Symbols

$x^{(i)}$	Value of variable, x on subchannel (i)
$A^{(i)}$	First sum rate inequality corresponding to the coherent combining gains in the minimum of sum rates operation
$B^{(i)}$	First part of the second sum rate inequality in the minimum of sum rates operation corresponding to the direct links
$C^{(i)}$	Second part of the second sum rate inequality in the minimum of sum rates operation corresponding to the cooperative links
$C(.)$	Channel capacity for the AWGN channel
$E[.]$	Expected value operation
E_{bP}	Event that the power constraint is violated by the codewords of some user
$E^{(i)}_{b\hat{w}_{kj}}$	Event that the received codewords at user j in block b are jointly typical with the codewords corresponding to message $\hat{w}^{(i)}_{kj}$
$F^{(i)}_{b\tilde{v}_{12},\tilde{v}_{21},\tilde{w}_{10},\tilde{w}_{20}}$	Event that the codewords received by the receiver in block b are jointly typical with codewords corresponding to $\tilde{v}^{(i)}_{12}, \tilde{v}^{(i)}_{21}, \tilde{w}^{(i)}_{10}, \tilde{w}^{(i)}_{20}$
$h^{(i)}_{kj}$	Fading coefficient between nodes k and j
$\mathcal{I}_j$	Set of available subchannels used in half-duplex cooperative encoding
$\bar{p}_k$	Average power constraint at node k
$p^{(i)}_{k0}$	Power assigned to $X^{(i)}_{k0}$
$p^{(i)}_{kj}$	Power assigned to $X^{(i)}_{kj}$
$p^{(i)}_{U_k}$	Power assigned to $U^{(i)}_k$
$R^{(i)}_k$	Rate for user k
$s^{(i)}_{kj}$	Normalized power-fading coefficient between nodes k and j
$U^{(i)}_k$	Common codeword sent by both transmitters for the resolution of the remaining uncertainty from the previous block
$v^{(i)}_{kj}$	Re-encoded submessage of node k, intended for the receiver

w_k	Message of node k
$w_{k0}^{(i)}$	Submessage of node k intended to be decoded by the receiver
$w_{kj}^{(i)}$	Submessage of node k intended to be decoded by user j
$\hat{w}_{jk}^{(i)}$	Estimates of messages at the cooperative partner
$\tilde{w}_{jk}^{(i)}$	Estimates of messages at the receiver
$X_k^{(i)}$	Codeword transmitted by node k
$X_{k0}^{(i)}$	Codeword carrying fresh information intended for the receiver
$X_{kj}^{(i)}$	Codeword carrying information intended for transmitter j for cooperation in the next block
$Y_j^{(i)}$	Symbol received at node j
$Z_j^{(i)}$	Zero-mean additive white Gaussian noise at node j
$\epsilon_t^{(i)}(\mathbf{s})$	Lagrange multiplier corresponding to non-negativity constraints for each power
γ_k	Lagrange multiplier corresponding to sum rate inequality constraints for each power
λ_k	Lagrange multiplier corresponding to average power constraints for each power
μ_k	Variables used in maximizing the weighted sum of rates
ϵ	Upper bound for average probability that each event occurs
${\sigma_j^{(i)}}^2$	Variance of $Z_j^{(i)}$

List of Abbreviations

A/D	**A**nalog-to-**D**igital
AEP	**A**symptotic **E**quipartition **P**roperty
AF	**A**mplify and **F**orward
AWGN	**A**dditive **W**hite **G**aussian **N**oise
BMSE	**B**lock **M**arkov **S**uperposition **E**ncoding
CF	**C**ompress and **F**orward
CSI	**C**hannel **S**tate **I**nformation
CDMA	**C**ode **D**ivision **M**ultiple **A**ccess
CMAC	**C**ooperative **M**ultiple **A**ccess **C**hannel
D/A	**D**igital-to-**A**nalog
DC	**D**elay-**C**onstrained
DF	**D**ecode and **F**orward
DFT	**D**iscrete **F**ourier **T**ransfrom
DVB	**D**igital **V**ideo **B**roadcasting
FFT	**F**ast **F**ourier **T**ransform
i.i.d.	**I**ndependent and **I**dentically **D**istributed
IDFT	**I**nverse **D**iscrete **F**ourier **T**ransform
KKT	**K**arush-**K**uhn-**T**ucker
LTE	**L**ong **T**ime **E**volution
MAC	**M**ultiple **A**ccess **C**hannel
MAC-GF	**M**ultiple **A**ccess **C**hannel with **G**eneralized **F**eedback
MC	**M**ulti**c**arrier
MIMO	**M**ulti-**I**nput-**M**ulti-**O**utput
NAF	**N**on-orthogonal **A**mplify-and-**F**orward

NDC	**N**on-**D**elay-**C**onstrained
OFDM	**O**rthogonal **F**requency **D**ivision **M**ultiplexing
OFDMA	**O**rthogonal **F**requency **D**ivision **M**ultiple **A**ccess
QoS	**Q**uality **o**f **S**ervice
SCF	**S**imple **C**ompress and **F**orward
SINR	**S**ignal to **I**nterference and **N**oise **R**atio
SNR	**S**ignal to **N**oise **R**atio
TDMA	**T**ime **D**ivision **M**ultiple **A**ccess

Chapter 1

Introduction

The design of next generation wireless communication systems brings along new challenges, since the degrading factors such as fading and multi-user interference become harder to deal with as the number of users and the bandwidth requirements increase. Efficient utilization of the orthogonal structure of the vector multiple accessing techniques may be useful to overcome these challenges. In particular, OFDMA is a vector multiple accessing technique that provides a relatively simple way of assigning available bandwidth to users, while avoiding interference; and at the same time, it converts a frequency selective fading channel, to parallel flat fading subchannels, hence reducing the effects of intersymbol interference. With effective allocation of resources like power, bandwidth and time, we can overcome the mentioned problems and even increase the achievable rates by using diversity techniques employing the time and frequency varying structure of the channel.

One of the resources to be carefully allocated in a communication system, is the transmitter's power. Power allocation is one of the fundamental problems of wireless communications, different approaches have been provided for the related constraints to be satisfied and performance criteria. By looking at the most general case, we can say that there are two main approaches for power allocation. In the first approach, the aim is to guarantee quality of service (QoS) and to provide a minimum level of QoS by increasing the power level when the channel

states get worse [1, 2]. This type of power control is generally used for real-time applications sensitive to delay (guaranteed service) like voice and video. In the second approach of power allocation, the main goal is to maximize the information theoretical capacity over longer time intervals by using average rates as the objective functions. Power allocation protocols in this type are more profitable: when the channel states get worse, instead of increasing power, they save from the average transmitter power and when the channel conditions become better, they use more power and maximize the achievable rates in the long run. This type of power allocation is more suitable for scenarios requiring higher data rates, like data or file transfer (best-effort services), that are not as sensitive to delay as the real-time applications. Recently, higher demand for this type of applications, made information theoretical approach to power control more important. In this thesis, we investigate information theoretical approach to resource allocation for ergodic fading cooperative channels [3].

Power control for the fading point-to-point channels was first studied by Goldsmith and Varaiya [4]. There, for a system with one transmitter, one receiver and both with full channel state information (CSI), the optimal adaptive transmission scheme was shown to use water-pouring in time for power adaptation, and a variable-rate multiplexed coding scheme. It was proved that waterfilling algorithm is the optimum power allocation strategy that offers power allocation directly related to the channel states (quality). In this strategy, more power is used in good channel states, less power is used in bad channel states and no power is used in very bad channel states. For the multi-user system, power allocation for maximizing the overall system throughput was obtained by Knopp and Humblet [5]. The optimum power allocation strategy came up with the interesting result that at any time, only the user with the best channel state should transmit with the power level calculated by waterfilling algorithm as if it was the only user in the system, while the others should not transmit at that time and should wait. Therefore, power allocation problem actually also determines the optimum method for finding out who should access the system at each channel state. This

shows the importance of power allocation in system design. Rate region maximization problem for multi-user systems was solved by Tse and Hanly [6]. It was proved that the optimum codebook selection could also be done by multiplying the codewords in a certain codebook with the instantaneously varying power levels instead of a variable-rate multiplexed coding scheme. Thereby, it was shown that power control is not only necessary to achieve maximum achievable rate in the case of full CSI, but is also sufficient. In this thesis we utilize an approach similar to that of Tse and Hanly.

Power allocation for Gaussian multiple access channels (MAC) was investigated in [7] and it was shown that power allocation in each dimension can be obtained by iterative waterfilling algorithm over the users. Accordingly, in each power calculation, all users' power levels in all dimensions are fixed except for one user and power allocation for this user is performed, and it was shown that by performing this method iteratively over all users, the optimum resource allocation can be achieved for the system. Similar results were obtained for Code Division Multiple Access (CDMA), that is also an example of vector multiple access channels [8, 9].

In OFDMA systems, there are two major resource allocation problems: power allocation and subchannel allocation. Jang and Lee [10], solved the problem of power and subchannel allocation by dividing the problem into two phases: first of all, each subchannel is assigned to the user with the best channel state. Then, since the problem will be in orthogonal structure, optimum power allocation is solved with single user waterfilling algorithm in parallel Gaussian channels. Actually this two-phase solution is not necessary, the optimum power allocation solution actually gives the same result without forcing the constraint that every subchannel should be used by only one user. This kind of solution may be optimal for the network, but it may not be fair for the users. Solution of a similar problem with proportional fairness consideration was presented by Shen, Andrews and Evans [11]. In this paper, the authors criticized the high complexity of the waterfilling algorithm and instead of this method, they proposed a new algorithm with linear complexity for a two user system whose result is close to the optimum

solution. Gao and Cui dealt with the problem of maximizing the achievable sum rate while guaranteeing a minimum individual rate for each user [12]. Their method first considered the minimum individual rate criteria and then considered allocating the remaining subchannels by solving the joint power, subchannel and rate optimization problem together. A similar problem was solved in [13] by optimizing a utility function, which is assumed to be a function of the rates. There, Lagrange techniques and Karush-Kuhn-Tucker (KKT) conditions, which are widely used methods in convex optimization, were used and an algorithm with relatively low complexity was proposed.

Monahram and Bhashyam [14], solved the resource allocation problem in OFDMA downlink scenario by using joint subcarrier and power allocation for channel-aware queue-aware scheduling. Wong and Evans formulated and solved continuous and discrete ergodic weighted sum rate maximization by employing Lagrange techniques and offered algorithms with almost optimal rate in OFDMA downlink assuming the availability of perfect CSI [15]. Phischella and Belfiore, addressed resource allocation in multi-cell OFDMA networks with a heuristic method for distributed sum power minimization under target data rate requirements [16]. The proposed method establishes a set of target signal to interference plus noise ratio (SINR) per user and subcarrier, determined to reach the target data rate of each user, power control is then performed independently over each subcarrier. Resource allocation for downlink OFDMA systems that yields a multiuser waterfilling structure was presented in [17].

In a wireless communication system traffic does not have to be either delay-constrained (DC) data or non-delay-constrained (NDC) data. There are some heterogenous networks carrying both DC and NDC data (e.g. file transfer together with voice). Such a heterogeneous multiuser Orthogonal Frequency Division Multiplexing (OFDM) system with both types of traffic is investigated in [18] and it was proved that the optimal power allocation over subcarriers follows a multi-level waterfilling principle; moreover, the valid candidates competing for each subcarrier include only one NDC user but all DC users.

While orthogonal MAC techniques such as OFDMA traditionally use orthogonal transmission to mitigate interference, we direct our attention to an alternative approach to managing "interference" in wireless networks. In wireless channels, what is traditionally considered as interference, is in fact side information, and this side information can be carefully taken advantage of to increase the rates achievable by the users through intelligently designed protocols. When users can decode the other users' messages, they behave like a virtual multi-antenna system and therefore attain diversity gain through cooperative coding and obtain higher achievable rates. Although the idea of cooperation in wireless communication networks has become more important recently, it roots from quite old information-theoretic works. The simplest form of cooperation is the relay channel, which was first studied by van der Meulen [19]. Then, fundamental capacity and achievable rate theorems for the relay channel were proved, and several coding and decoding techniques were proposed by Cover and El Gamal in their groundbreaking paper [20]. In the relay channel, there is one source, one relay and one receiver. The only goal here, is to successfully transmit the message of the source to the receiver with the help of the relay. Most of the research in the literature, deals with this kind of one-sided cooperative relay model, but these strategies are not suitable for the systems where all cooperative users have their own information. On the other hand, next generation communication systems require to support high data rates for multi-users simultaneously.

The fundamentals of the mentioned real bidirectional cooperative systems, based on the two user problem, was solved in the beginning of 1980's [21]-[22]. Firstly, Cover and Leung [21], determined the achievable rate region for a discrete memoryless multiple access channel with feedback. The region found by Cover and Leung for the discrete memoryless multiple access channel with feedback to both encoders, is proved to be achievable also with feedback to only one encoder by Willems and van der Meulen [23]. Willems and van der Meulen introduced a communication situation in which the encoders of a MAC are partially cooperating where the encoders are connected by communication links with finite capacities

and established the capacity region of the MAC with partially cooperating encoders in [24]; later determined the capacity regions for various communication situations in which one or both encoders for a MAC crib from the other encoder [22]. This type of channel models are important to give upper bounds for the achievable rate regions for more realistic channels with limited cooperation opportunity.

Due to the information spreading property of wireless communication channels, these channels are best described using, the multiple access channel with generalized feedback (MAC-GF) proposed by Carleial [25] and later by Willems et al. [26]. According to this model, users are fedback two different channel outputs and these channel outputs represent the "overheard" information. The first achievable rate region for this model by Carleial [25] was improved by Willems et al. by utilizing block Markov superposition encoding (BMSE) and backward decoding [26]. More recently, in [27], the MAC-GF was used by Sendonaris, Erkip and Aazhang to model a fading cooperative additive white Gaussian noise (AWGN) channel, and the results therein offering high gains made cooperative communications very attractive. Meanwhile Laneman, Tse and Wornell provided outage analysis for a more practical cooperative system employing half-duplex operation [28].

Intensive research has been conducted and plenty of development has been achieved in the above mentioned fields, over the last ten years. An important part of this development has been for the case where the number of the cooperative users is more than two. In most of these works, cooperation is limited with dedicated relays assigned only to help other users or a single relay assigned to help multiple users. Examples of the most significant of these are; Multiple Access Relay Channels [29], Single Source Multiple Level Relay Channel [30] and Parallel Relay Networks [31]. More recently, some works on bidirectional cooperation for multi-user systems also started appearing in the literature, and it was shown that important rate gains can be achieved [32]. A large number of references on cooperative relay systems can be found in [33].

In cooperative communications, one of the critical issues is resource allocation, especially power allocation. Most the of work in this field, again focuses on the scenario of relay networks with only one-sided cooperation where relays do not have their own messages. Research on resource allocation for the bidirectional cooperation employing Decode and Forward (DF) approach (where encoding and decoding is much more complex) is quite limited. Works in this field show that utilizing optimum power control simplifies the cooperation strategy and increases the achievable rates significantly, just like in CDMA and OFDMA systems [34]. Summarizing all these, we can conclude that in multiple access systems (like OFDMA and CDMA) employing cooperation, resource allocation is both an important and a promising problem.

The use of cooperative protocols in OFDM systems was investigated extensively by many authors over the recent years. One of the first works that considers OFDM together with cooperation is [35]. There Yatawatta and Petropulu, divided the subchannels of OFDM into two, used Amplify and Forward (AF) cooperation scheme in orthogonal structure and showed the performance improvement in terms of channel capacity and diversity. In [36], Lin and Stefanov studied pairwise error probability for single antenna OFDM systems using cooperative convolutional codes.

Generic relay and subcarrier allocation schemes for multicarrier (MC) system with AF relays were proposed in [37]. The outage probability bounds were derived analytically for each scheme and the average best relay selection scheme, was shown to be best suited for practical implementation. In [38], an optimally joint subcarrier matching and power allocation scheme to further maximize the total channel capacity with the constrained total system power was proposed. First, the problem is formulated as a mixed binary integer programming problem. Second, by making use of the equivalent channel power gain for any matched subcarrier pair, a low complexity scheme is proposed.

Based on the above summary of results, we can conclude that, there has been

an intensive research on power control and subchannel allocation in OFDMA, on power control in cooperative communications, on relay selection in cooperative systems, and on resource allocation in one-sided cooperative OFDMA systems. However, the problems of achievable rates and optimum resource allocation in a mutually cooperative OFDMA system employing DF cooperation and design of cooperative protocols for OFDMA channels have not been investigated and solved yet. One of the basic advantages of OFDMA channels is that they allow simultaneous transmission over different subchannels, therefore it is clear that techniques will bring up new opportunities for exploiting the diversity via new cooperative strategies. In this thesis, without imposing any prior constraints on subchannel allocation or orthogonalization, we will investigate cooperation strategies, achievable rates and resource allocation for next generation wireless networks.

This thesis is organized as follows; we present some general information about the techniques used throughout the thesis in Chapter 2. In Chapter 3, the proposed cooperative strategies for OFDMA and achievable rates for the two user cooperative OFDMA channels are explained. We apply resource allocation for the two user cooperative OFDMA channels and describe the applied resource allocation in Chapter 4. Conclusion of the thesis is given in Chapter 5. The interested readers can find more detailed information related to the proves of the theorems and the lemma in the appendixes.

Chapter 2

Background

2.1 Overview

In this thesis, various topics and techniques from wireless communications, information theory and optimization are employed. These include OFDMA, channel capacity and achievable rates for wireless channels and convex optimization methods. In the following sections we provide some general information about the techniques used.

2.2 Orthogonal Frequency Division Multiplexing (OFDM)

OFDM is a technique employed by many telecommunication systems, such as Digital Video Broadcasting (DVB), local area networks and Long Term Evolution (LTE) [39]. OFDM is a MC modulation that divides a given bandwith into many parallel subchannels or subcarriers, so that multiple symbols can be sent in parallel.

The first OFDM schemes were presented by Chang [40] and Saltzberg [41]. OFDM became more practical through work of Chang and Gibby [42], Weinstein and Ebert [43], Peled and Ruiz [44] and Hirosaki [45]. OFDM using the Discrete Fourier transform (DFT) was developed by [43], OFDM with a cyclic prefix was developed by [44]. The DFT (implemented with a Fast Fourier Transform (FFT))

and the cyclic prefix have made OFDM both practical and attractive. OFDM is often attractive by its two main features: it is considered to be spectrally efficient and it offers a clever way to deal with equalization of dispersive slowly fading channels.

OFDM is a block transmission technique. In the baseband, complex-valued data symbols modulate a large number of carrier waveforms. The transmitted OFDM signal multiplexes several low-rate data streams, each data stream is carried by a given subcarrier. The main advantage of this concept in a radio environment is that each of the data streams is transmitted through an almost flat fading channel. Orthogonal subcarriers of OFDM signal, modulated by these parallel data streams are of the form:

$$\phi_k(t) = e^{j2\pi f_k t}, \tag{2.1}$$

where f_k is the frequency of the k^{th} subcarrier. One baseband OFDM symbol (without a cyclic prefix) multiplexes N modulated subcarriers:

$$s(t) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_k \phi_k(t), \qquad 0 < t < NT, \tag{2.2}$$

where x_k is the k^{th} complex data symbol, T is the sampling period and NT is the length of the OFDM symbol. The subcarrier frequencies, f_k are equally spaced

$$f_k = \frac{k}{NT}, \tag{2.3}$$

which makes the subcarriers $\phi_k(t)$ on $0 < t < NT$ orthogonal. The signal (2.2) separates data symbols in frequency by overlapping subcarriers, thus using the available spectrum in an efficient way. As the OFDM signal is the sum of a large number of independent and identically distributed (i.i.d.) components, its amplitude distribution becomes approximately Gaussian due to the central limit theorem.

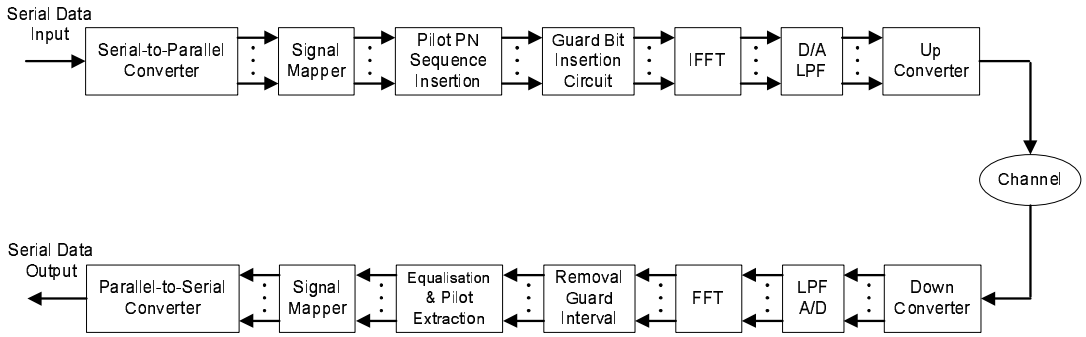


Figure 2.1: Block Diagram of an OFDM System Using FFT, Pilot PN Sequence and a Guard Bit Insertion [47].

T-spaced sampling of the in-phase and quadrature components of the OFDM symbol gives

$$s(nT) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_k e^{j2\pi \frac{nk}{N}}, \qquad 0 < t < NT, \tag{2.4}$$

which is the Inverse Discrete Fourier Transform (IDFT) of the constellation symbols [46]. Therefore, demodulation of the sampled data can be performed using DFT operation. This is one of the key properties of OFDM, first proposed by Weinstein and Ebert [43].

OFDM also has some disadvantages. Firstly, like all orthogonal transmission techniques, OFDM incurs some rate penalty. Because OFDM divides a given bandwidth into many narrow subcarriers each with relatively small carrier spacing, it is sensitive to carrier frequency errors. Furthermore, to preserve the orthogonality between subcarriers, the amplifiers need to be linear. OFDM systems also should have a high peak-to-average power ratio, which may require a large amplifier power back-off and a large number of bits in the analog-to-digital (A/D) and digital-to-analog (D/A) designs.

OFDM divides the spectrum into many parallel subchannels or subcarriers and is a special case of parallel Gaussian channels. In the next subsection, we provide

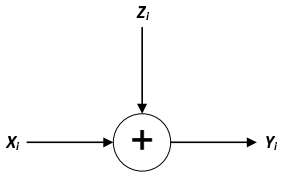


Figure 2.2: The Gaussian Channel.

an overview for the parallel Gaussian channel and the channel capacity for these channels.

2.2.1 Capacity of Parallel Gaussian Channel

Information capacity of a discrete memoryless channel is given by

$$C = \max_{p(x)} I(X;Y), \tag{2.5}$$

where the maximum is taken over all possible input distributions $p(x)$ [48]. Channel capacity is defined as the highest rate at which information can be transmitted with arbitrarily low probability of error through the channel. The Gaussian channel is the most commonly encountered continuous alphabet channel. In the Gaussian channel, the output Y_i at time i, is the sum of the input, X_i and the noise, Z_i, where the noise Z_i is zero-mean, i.i.d. from a Gaussian distribution with variance N, i.e., $Z_i \sim \mathcal{N}(0, N)$. The Gaussian channel, shown in Figure 2.2 is described by

$$Y_i = X_i + Z_i, \quad i = 1, ..., n. \tag{2.6}$$

Let us assume that for any codeword $(x_1, x_2, ..., x_n)$ transmitted over the channel, there is an average power constraint,

$$\frac{1}{n}\sum_{i=1}^{n} {x_i}^2 \leq P. \tag{2.7}$$

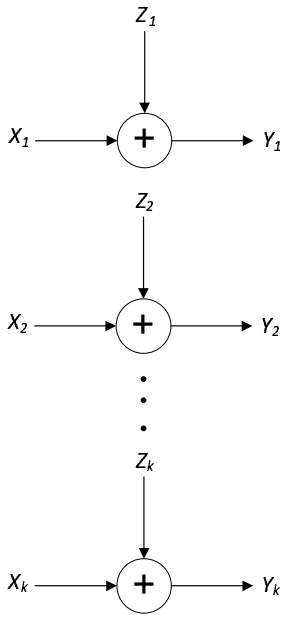


Figure 2.3: The Parallel Gaussian Channels.

The capacity of the Gaussian channel with power constraint P and noise variance N is

$$C = \frac{1}{2}\log\left(1 + \frac{P}{N}\right) \text{ bits per transmission.} \tag{2.8}$$

Assume that there are a set of Gaussian channels in parallel as shown in Figure 2.3, for each Gaussian channel j, the output is the sum of the input and the Gaussian noise

$$Y_j = X_j + Z_j, \quad j = 1, 2, ..., k, \tag{2.9}$$

where the noise, $Z_j \sim \mathcal{N}(0, N_j)$ is assumed to be independent from channel to channel. We assume that there is a common power constraint on total power used over all channels, i.e.,

$$E\sum_{j=1}^{n} {X_j}^2 \leq P. \tag{2.10}$$

Here, the goal is to maximize the total capacity by distributing power among the channels. The information capacity of the parallel Gaussian channel is

$$C = \max_{f(x_1,x_2,...,x_k):\sum EX_i^2 \leq P} I(X_1, X_2, ..., X_k; Y_1, Y_2, ..., Y_k). \tag{2.11}$$

Since Z_1, Z_2,...,Z_k are independent it can be shown that [48]

$$C = I(X_1, X_2, ..., X_k; Y_1, Y_2, ..., Y_k) \leq \frac{1}{2} \sum_{i=1}^{k} \log \left(1 + \frac{P_i}{N_i}\right). \tag{2.12}$$

From this point on, the problem becomes that of finding a power allocation policy that maximizes the capacity subject to the power constraint, i.e., $\sum P_i = P$. This is a standard optimization problem and can be solved by Lagrange multipliers. We can write the following Lagrangian function (see section 2.4).

$$\mathcal{L}(P_1, ..., P_k) = \frac{1}{2} \sum_{i=1}^{k} \log \left(1 + \frac{P_i}{N_i}\right) + \lambda(\sum_{i=1}^{k} P_i), \tag{2.13}$$

and differentiating with respect to P_i, we achieve

$$P_i = \nu - N_i. \tag{2.14}$$

P_i's must be non-negative, therefore the solution of the power level at each channel becomes

$$P_i = (\nu - N_i)^+, \tag{2.15}$$

where $(x)^+$ is $\max(x, 0)$. ν is chosen such that

$$\sum (\nu - N_i)^+ = P. \tag{2.16}$$

As a result, there is a certain power limit, ν in each channel. Power is distributed among the channels like the way water distributes itself in a vessel. This power allocation procedure is referred as "waterfilling".

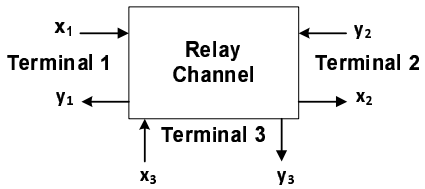


Figure 2.4: Three-Terminal Relay Channel.

In our thesis we employ a cooperative OFDMA system with parallel OFDMA subchannels and obtain optimal power distribution with the aim to maximize the achievable rate region of the system. Details on resource allocation for the cooperative OFDMA channels can be found in Chapter 4.

2.3 Cooperative Models

In the simplest form of communications, data is transmitted directly between the transmitter and the receiver. In these communication systems users do not help each other. Recently, networks with the so-called "cooperative" nodes, assisting other nodes have been introduced. Cooperative communication utilizes the fact that the transmitted messages are actually "overheard" by other users. This cooperation can be unidirectional or bidirectional. As a result of the cooperative communication, we can improve the achievable rates, coverage area and battery life [49].

Van der Meulen was the first to introduce the idea of cooperation [19] in information theory and established the roots of the relay channel. The general three-node relay channel, shown in Figure 2.4, consists of Terminal 1 (source), Terminal 3 (destination) and Terminal 2 (relay). Here, the goal is to have the highest rate form Terminal 1 to Terminal 3 with the support of Terminal 2. Terminal 2 does not have its own message, it just serves in order to improve the direct link's rate between Terminal 1 to Terminal 3.

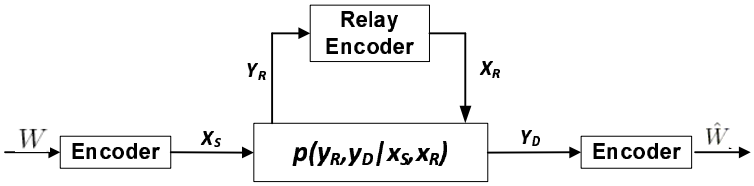


Figure 2.5: The Relay Channel.

Van der Meulen, calculated the capacity of the relay channel in [19] for some specific channels. In their seminal paper [20], Cover and El Gamal further developed communication/relaying methodologies in a general relay channel. They studied the memoryless relay channel and provided two relaying techniques in [20]; DF and Compress and Forward (CF). The capacity of the general relay channel has not been solved for over thirty years. Willems' MAC-GF model [26] was a milestone in mutual cooperation. The recent papers on cooperative communications like [27] and [28] made cooperative communications very attractive. The authors showed the advantages of mutual cooperation in fading environments in [27]. The offered gain improvements and the ability to apply these techniques made cooperative communications and related topic, one of the key research areas.

We now introduce the full-duplex relay channel to model unidirectional cooperation and describe the three fundamental relaying protocols AF, DF and CF. We then describe the MAC-GF model to explain mutual cooperation.

2.3.1 Relaying

The basic relay channel model is shown in Figure 2.5, where $p(y_R, y_D|x_S, x_R)$ is the discrete memoryless channel, W denotes the message, X_S and X_R are the transmitted signals at the source and the relay, Y_R and Y_D are the received signals at the relay and at the destination, and $\hat{W}$ is the destination's estimate of W.

If the channel is real AWGN, then we can write

$$Y_R = h_{SR}X_S + Z_R, \tag{2.17}$$

$$Y_D = h_{SD}X_S + h_{RD}X_R + Z_D, \tag{2.18}$$

where h_{SR}, h_{SD} and h_{RD} are respectively the channel gains between the source and the relay, the source and the destination, and the relay and the destination. The AWGN noise at the relay and destination are respectively denoted by Z_R and Z_D, which are assumed to be zero mean and unit variance. There are individual power constraints for both the source and the relay, P_S and P_R. The objective is to find the capacity, the maximum achievable rate beyond which reliable communication is not possible. In [48] it is proved that if R is an the achievable rate,

$$R \leq \max_{p(x_S,x_R)} \min \left\{ I(X_S; Y_R, Y_D); I(X_S, X_R; Y_D) \right\}, \tag{2.19}$$

which becomes

$$\begin{aligned} R \leq \max_{\rho\in[0,1]} \min \Bigg\{ & \frac{1}{2}\log(1+\rho h_{SR}^2 P_S + \rho h_{SD}^2 P_S), \\ & \frac{1}{2}\log(1+h_{SD}^2 P_S + h_{SD}^2 P_R + 2\sqrt{(1-\rho)h_{SD}^2 h_{RD}^2 P_S P_R}) \Bigg\}, \end{aligned} \tag{2.20}$$

for the Gaussian case where ρ denotes the correlation between source and relay signals.

AF, DF and CF are the relaying techniques proposed in [20] and [28] for the general relay channel. In AF, the relay amplifies the received signal, and forwards it to the destination. In DF, the relay decodes its received signal, re-encodes it, and forwards it to the destination. In CF, the relay first compresses its received signal and then forwards the compressed signal through the relay-destination channel. In the following section, we describe the general structure of these protocols.

2.3.1.1 Amplify and Forward

AF scheme, the relay only scales its received signal to satisfy its own power constraint and re-transmits to the destination [50]. This scheme uses the original message and the repeated message to the destination by the relay, but the repetition also amplifies its noise to the destination [28].

2.3.1.2 Decode and Forward

In the Decode and Forward (DF) protocol, the source, and the relay utilize BMSE. In the destination either sliding window or backward decoding can be used [33], [30], [51], [22]. In our thesis, we performed backward decoding, therefore we will explain the backward decoding here shortly. Transmission occurs in B blocks, where each block consists of N symbols, with N and B both large. The source transmits a length-N codeword $x_S(w_{b-1}, w_b)$ in each block b. The initial message and the message in the B^{th} block are equal to 1, i.e., $w_0 = 1$ $w_B = 1$. The relay decodes the source message correctly, with very high probability, i.e., $\hat{w_b} = w_b$, if the source rate, $R^{(DF)}$ satisfies the following condition:

$$R^{(DF)} \leq I(X_S; Y_R|X_R). \tag{2.21}$$

When this is satisfied, the relay finds an estimate $\hat{w_b}$ for w_b at the end of block b and sends $x_R(\hat{w_b})$ in block $b+1$. In the first block the relay transmits a predetermined codeword $x_R(1)$. In block $b+1$ the source repeats w_b, the relay and the source can perform the cooperation.

The destination starts decoding after all B blocks are received and moves backwards [33], [51], [22]. At block B, no fresh information is sent and the destination tries to decode w_{B-1}. Successful decoding occurs if

$$R^{(DF)} \leq I(X_S, X_R; Y_D). \tag{2.22}$$

As soon as the destination decodes w_{B-1}, it moves backwards to decode w_{B-2},...,w_1, performing similar decoding. Combining the constraints (2.21) and (2.22), we conclude that for a fixed input distribution $p(x_S, x_R)$,

$$R^{(DF)} \leq \min\left\{I(X_S; Y_R|X_R), I(X_S, X_R; Y_D)\right\}, \tag{2.23}$$

is achievable. Maximizing over all input distributions, we can write the maximum rate the DF protocol achieves is

$$R^{(DF)}_{max} = \max_{p(x_S,x_R)} \min\left\{I(X_S; Y_R|X_R), I(X_S, X_R; Y_D)\right\}. \tag{2.24}$$

2.3.1.3 Compress and Forward

In the DF protocol the relay decodes and makes a decision about the source message. However, letting relay decide about the source message may cause some decrease in the rates. Instead of this strategy, in Simple Compress and Forward (SCF) the relay compresses its received signal Y_R in block b to form $\hat{Y}_R$, maps this to the channel codeword X_R and transmits X_R to the destination in block $b+1$ [49]. The relay can listen and compress the message in the current block simultaneously while transmitting the compressed signal of the previous block, since the relay is full duplex. Therefore the received signal Y_D in the destination in block b is a function of both the source and the relay signals X_S and X_R transmitted in the same block.

The destination starts decoding after all B blocks are transmitted, performing backward decoding. It first decodes the relay signal X_R, treating the source signal as noise and recovers $\hat{Y}_R$. If the relay's compression rate is below the relay-to-destination achievable rate considering the source signal as noise, we say we have reliable transmission. For this, the following condition has to be satisfied for a fixed input distribution $p(x_S)p(x_R)p(\hat{y}_R|x_R, y_R)p(y_R, y_D|x_S, x_R)$

$$I(\hat{Y}_R; Y_R|X_R) \leq I(X_R; Y_D). \tag{2.25}$$

After decoding $\hat{Y}_R$ at the destination, $\hat{Y}_R$ and Y_D from the previous block are both used at the destination to determine the original source signal X_S. This can be done reliably if

$$R^{(SCF)} < I(X_S; \hat{Y}_R, Y_D | X_R). \tag{2.26}$$

If Wyner-Ziv type compression [52] is employed instead of simple compression, the performance of the CF can be improved. The Wyner-Ziv technique allows for lower compression rates if correlated side information is available at the decoder. In the relay channel, the direct signal from the source received at the destination in the previous block can be thought of as the side information. Therefore correlated side information is available at the decoder in the relay channel. The benefit of this in CF is that, this side information lowers the compression rate from $I(\hat{Y}_R; Y_R | X_R)$ to $I(\hat{Y}_R; Y_R | X_R, Y_D)$ yielding

$$I(\hat{Y}_R; Y_R | X_R, Y_D) \leq I(X_R; Y_D). \tag{2.27}$$

When Wyner-Ziv type compression is used the overall CF achievable rate becomes

$$R^{(CF)} = I(X_S; \hat{Y}_R Y_D | X_R), \tag{2.28}$$

subject to (2.27) where the joint probability distribution is $p(x_S)p(x_R)p(\hat{y}_R|x_R, y_R)$ $p(y_R, y_D|x_S, x_R)$. Achievable rates are potentially higher when Wyner-Ziv type compression is used, because (2.27) is looser than (2.25).

AF, DF and CF protocols model relaying or unidirectional cooperation. We now explain the MAC-GF model to describe mutual cooperation.

2.3.2 MAC with Generalized Feedback

Cooperation roots from the relay channel. In the relay channel, the relay node does not have its own information to send; it only serves to improve achievable

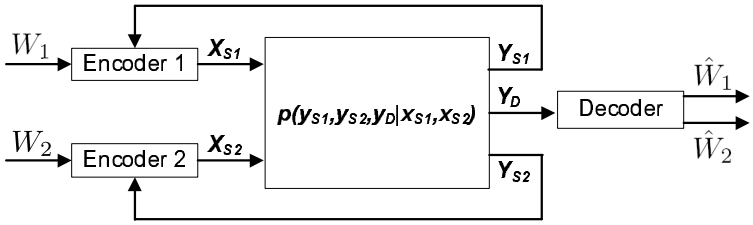


Figure 2.6: The Multiple Access Channel with Generalized Feedback.

communication rates. So, in relay channel, there is actually only unidirectional communication. In case of mutual cooperation in the MAC with generalized feedback (MAC-GF), there are two or more sources, all which have their own information and the sources communicate bidirectionally to attain better rates.

The two-user discrete-memoryless MAC-GF is shown in Figure 2.6, where W_1 and W_2 denote the messages of source 1 (S_1) and source 2 (S_2), X_{S_1} and X_{S_2} denote the source signals, $p(y_{S_1}, y_{S_2}, y_D|x_{S_1}, x_{S_2})$ is the channel, Y_{S_1}, Y_{S_2} and Y_D are respectively the received signals at S_1, S_2 and the destination, and $\hat{W}_1$ and $\hat{W}_2$ are the destination's estimate of original messages W_1 and W_2. Mutual cooperation is possible in the MAC-GF model as the user assist each other. The achievable rate region is the collection of all achievable rate pairs. In the MAC-GF, the cooperative partner (relay) is another source node, which has its own information to transmit.

Achievable rate region for partial DF based MAC-GF was developed in [26]. The relay only decodes part of the source message, and the remaining part is directly sent to the destination without the relay's help in partial DF. In each block, each source's messages are considered to be composed of three parts. The first part of the message is sent directly to the destination without other's assistance. The second part of S_1's message in block b is aimed to be decoded only at S_2 at the end of block b. Decoding this part of the message, S_2 forms the third part of its message in block $b+1$. Similarly, at the end of block b, S_1 decodes the second part of S_2's message, and re-encoding it forms the third part of its own message in

block $b+1$. This procedure enables bidirectional communication simultaneously. The achievable rate region analysis can be found in [26] and [53].

We propose two mutually cooperative encoding strategies using DF approach and based on BMSE and backward decoding, details of which can be found in Chapter 3.

The following subsection gives a review of basic optimization techniques that we utilize while dealing with power allocation to maximize the achievable rates of the proposed cooperative OFDMA systems.

2.4 Optimization

Most information theoretical capacity or achievable rate maximization problems are convex in nature. In our thesis we deal with convex optimization problems for achievable rate maximization.

An optimization problem, of the following form

$$\begin{aligned} &\textit{minimize} && f_0(x), \\ &\textit{subject to} && f_i(x) \leq 0, \quad i = 1, ..., m, \\ &&& h_i(x) = 0, \quad i = 1, ..., p, \end{aligned} \tag{2.29}$$

describes the problem of finding an x that minimizes $f_0(x)$ among all x that satisfy the conditions $f_i(x) \leq 0$, $i = 1, ..., m$, and $h_i(x) = 0$, $i = 1, ..., p$. Here, the vector $x = (x_1, ..., x_n)$, $x \in \mathbf{R}^n$ is the optimization variable of the problem, the function, $f_0 : \mathbf{R}^n \rightarrow \mathbf{R}$ is the objective function or cost function, the functions $f_i : \mathbf{R}^n \rightarrow \mathbf{R}$, $i = 1, ..., m$, are the inequality constraint functions and $h_i : \mathbf{R}^n \rightarrow \mathbf{R}$, $i = 1, ..., p$, are the equality constraint functions (or we can take $-f_0(x)$ as the the objective function and deal with maximizing the objective function). The domain is denoted by, $D = \bigcap_{i=0}^{m} \mathbf{dom}\ f_i \ \cap \ \bigcap_{i=1}^{p} \mathbf{dom}\ h_i$. p^* is

the optimal value of the problem (2.29) if it satisfies the following:

$$p^* = \inf \{ f_0(x) \mid f_i(x) \leq 0,\ i = 1, ...,\ h_i(x) = 0,\ i = 1, ..., p \}. \tag{2.30}$$

Let x^* be an optimal point or a solution for the problem (2.29). The optimal set is the set of all optimal points is denoted by

$$X_{opt} = \{ x \mid f_i(x) \leq 0,\ i = 1, ...,\ h_i(x) = 0,\ i = 1, ..., p,\ f_0(x) = p^* \}. \tag{2.31}$$

An important set of optimization problems are convex optimization problems. In a convex optimization problem, the objective and the inequality constraint functions must be convex, meaning they must satisfy the inequality [54]

$$f_i(\alpha x + \beta y) \leq \alpha f_i(x) + \beta f_i(y), \tag{2.32}$$

for all x, $y \in \mathbf{R}^n$ and all $\alpha,\ \beta \in \mathbf{R}$ with $\alpha + \beta = 1$, $\alpha \geq 0$, $\beta \geq 0$ and the equality constraint functions, $h_i(x) = a_i^T x - b_i$ must be affine. The problems we tackle in our thesis can be formed as convex optimization problems. Therefore, we present convex optimization techniques briefly.

2.4.1 The Lagrangian

Lagrangian duality is formed by adding the objective function and a weighted sum of the constraint functions in (2.29). The Lagrangian associated with the problem (2.29) is defined as, $L : \mathbf{R^n} \times \mathbf{R^m} \times \mathbf{R^p} \times \rightarrow \mathbf{R}$

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i, \tag{2.33}$$

where λ_i is the Lagrange multiplier associated with the ith inequality constraint, $f_i(x) \leq 0$ and ν_i is the Lagrange multiplier associated with the ith equality

constraint, $h_i(x) = 0$. λ and ν vectors are called the Lagrange multiplier or dual variables vector, associated with the problem (2.29).

The Lagrange dual function, $g : \mathbf{R^m} \times \mathbf{R^p} \rightarrow \mathbf{R}$ is defined as

$$g(\lambda, \nu) = \inf_{x \in D} L(x, \lambda, \nu) = \inf_{x \in D} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i \right). \tag{2.34}$$

2.4.2 KKT Optimality Conditions

If $f_0, ..., f_m, h_1, ..., h_p$ are differentiable, f_i are convex, h_i are affine and $\tilde{x}$, $\tilde{\lambda}$, $\tilde{\nu}$ are any points satisfying the following KKT conditions

$$f_i(\tilde{x}) \leq 0, \quad i = 1, ..., m, \tag{2.35}$$

$$h_i(\tilde{x}) = 0, \quad i = 1, ..., p, \tag{2.36}$$

$$\tilde{\lambda}_i \geq 0, \quad i = 1, ..., m, \tag{2.37}$$

$$\tilde{\lambda}_i f_i(\tilde{x}) = 0, \quad i = 1, ..., m, \tag{2.38}$$

$$\nabla f_0(\tilde{x}) + \sum_{i=1}^{m} \tilde{\lambda}_i \nabla f_i(\tilde{x}) + \sum_{i=1}^{p} \tilde{\nu}_i \nabla h_i(\tilde{x}) = 0, \tag{2.39}$$

then $\tilde{x}$, $(\tilde{\lambda}, \tilde{\nu})$ are primal and dual optimal, with zero duality gap. Here, (2.35) and (2.36) let us conclude that $\tilde{x}$ is primal feasible. Using (2.37), we see that $L(x, \tilde{\lambda}, \tilde{\nu})$ is convex in x. (2.39) shows that the gradient of $L(x, \tilde{\lambda}, \tilde{\nu})$ with respect to x vanishes at $x = \tilde{x}$, therefore $\tilde{x}$ minimizes $L(x, \tilde{\lambda}, \tilde{\nu})$ over x. In some cases, the KKT conditions can be solved analytically, giving the global optimal solution. In cases where they can not, these optimization problems can be solved using iterative algorithms. Gradient/subgradient algorithms are effective iterative algorithms that can be used to solve the constrainted optimization problems. In the next subsection, we describe such techniques that we utilized in our thesis in Chapter 4.

2.4.3 Gradient Algorithm

Descent methods produce a minimizing sequence $x^{(k)}$ where

$$x^{(k+1)} = x^{(k)} + t^{(k)} \triangle x^{(k)},$$

$\triangle x^{(k)}$ denote the step or search direction, $k = 1, ...$ denotes the iteration number, $t^{(k)} > 0$ (except for the optimal $x^{(k)}$) denote the step size at the k^{th} iteration.

In gradient algorithm (or gradient descent method) search direction is the negative gradient, $\triangle x^{(k)} = -\nabla f(x)$. Gradient algorithm is explained briefly below [54]:

Given a stating point $x \in dom\ f$.

Repeat

1. $\triangle x := -\nabla f(x)$.
2. Choose step size, t via exact or backtracking line search.
3. Update $x := x + t\triangle x$.

Until Stopping criterion.

The stopping criterion is usually of the form $\|\nabla f(x)\|_2 \leq \eta$, where η is small and positive. The main advantage of the gradient algorithm is its simplicity. The main disadvantage is, convergence can be very slow, even for problems with number of condition in the 100s. When the number of condition is larger, it may not be practical to use the gradient method. Another disadvantage is that the convergence depends on the choice of backtracking parameters α, β.

The subgradient methods and gradient methods are in fact very similar, however in subgradient method the objective functions not necessarily monotonically non-decreasing/non-increasing. Subgradient gives affine global underestimator of f. If f is convex, it has at least one subgradient at every point in **relint dom** f, if f is convex and differentiable, $\nabla f(x)$ is a subgradient of f at x [54].

Let us denote the objective function by $f(x)$. A subgradient is any vector $g \in R^n$, at $x \in \mathbf{dom}\ f$ if for all $z \in \mathbf{dom}\ f$ that satisfies

$$f(z) \geq f(x) + g^T(z - x). \tag{2.40}$$

The subgradient method uses the following iteration to minimize f

$$x^{(k+1)} = x^{(k)} - \alpha_k g^{(k)},$$

where $x^{(k)}$ is the kth iterate, $g^{(k)}$ is any subgradient of f at $x^{(k)}$, and $\alpha_k > 0$ is the kth step size. We actually take a step in the direction of a negative subgradient at each iteration of the subgradient method. There are some step size rules used in subgradient method. Three examples are given below:

1) First of all, the subgradient method may use a constant step size, i.e., $\alpha_k = h$ is a constant, independent of k. Step lengths are fixed ahead of time, instead of an exact or approximate line search as in the gradient method, i.e., $\alpha_k = h/\|g^{(k)}\|_2$, meaning $\|x^{(k+1)} - x^{(k)}\|_2 = h$.

2) Step sizes are square summable but not summable, i.e., $\sum_{k=1}^{\infty} \alpha_k^2 < \infty$, $\sum_{k=1}^{\infty} \alpha_k = \infty$. A common example of such a step size is $\alpha_k = a/(b + k)$, where $a > 0$ and $b > 0$.

3) Step sizes are nonsummable diminishing, i.e., $\lim_{k\to\infty} \alpha_k = 0$, $\sum_{k=1}^{\infty} \alpha_k = \infty$. A typical example is $\alpha_k = a/\sqrt{k}$, where $a > 0$.

The projected subgradient method is an extension of the subgradient method which solves the following constrained convex optimization problem

$$\begin{aligned} &\textit{minimize } f(x), \\ &\textit{subject to } x \in \mathcal{C}, \end{aligned} \tag{2.41}$$

where $\mathcal{C}$ is a convex set. The projected subgradient method is defined as

$$x^{(k+1)} = P\left(x^{(k)} - \alpha_k g^{(k)}\right), \tag{2.42}$$

where P is (Euclidean) projection on $\mathcal{C}$, and $g^{(k)}$ is any subgradient of f at $x^{(k)}$. The size step sizes rules for subgradient method also apply to the projected subgradient method.

In our thesis we utilize a projected subgradient algorithm, and an iterative water-filling-like algorithm based on KKT conditions on optimality to determine the optimal power distribution for the cooperative OFDMA system achieving an arbitrary rate point on the achievable rate region boundary.

Chapter 3

Cooperative Strategies And Achievable Rates For Two User OFDMA Channels

3.1 Introduction

The idea of cooperative communications roots from the simplest form of cooperative channel model: the relay channel. This channel was first studied in [19]. Later, in the seminal paper [20], fundamental capacity and achievable rate theorems for the relay channel were proved, and several coding and decoding techniques were proposed. The extension of the one-sided cooperative relay model to mutual cooperation was made possible by the introduction of MAC-GF [25, 26]. In [26], an achievable rate region, which was larger than that of [25], was obtained by utilizing BMSE and backward decoding. More recently, in [27], the MAC-GF was used to model a fading cooperative AWGN channel, and the results therein made cooperative communications very attractive.

The use of cooperative protocols in OFDM systems was investigated extensively by many authors over the recent years. In [36], the authors obtained bounds on pairwise error probability for single antenna OFDM systems employing cooperative convolutional codes. In [55], per subcarrier hybrid cooperation strategies were proposed, with the goal of minimizing the error probability. Methods for subcarrier selection in multihop OFDM systems were developed in [56].

The works in the literature related to cooperative OFDM, some examples of which are listed above, either consider a one-sided cooperation strategy, or a mutually cooperative strategy based on two parallel dedicated relay channels, or mutual cooperation based on a time division protocol. In this thesis, without imposing any prior constraints on which users will use which subchannels, we propose two full duplex cooperative encoding strategies: intra-subchannel and inter-subchannel cooperative encoding. These strategies use DF approach and are based on BMSE. Intra-subchannel cooperative encoding is an extension of the two user cooperative strategy in [26, 27] to OFDMA, and inter-subchannel cooperative encoding is a novel method which allows for re-partitioning and re-encoding of the cooperative messages across subchannels. We further propose another cooperative OFDMA system with half-duplex operation in each subchannel. We obtain the achievable rate regions for all three strategies, and compare them with the non-cooperative OFDMA capacity region. We demonstrate through simulations that, cooperative OFDMA may provide significant rate improvements over its non-cooperative counterpart.

3.2 System Model

We consider a fading two user cooperative OFDMA system with N subchannels. On each subchannel, unless otherwise stated, each user is capable of both transmitting and receiving signals. The system is illustrated in Figure 3.1, and is modelled by,

$$Y_0^{(i)} = h_{10}^{(i)} X_1^{(i)} + h_{20}^{(i)} X_2^{(i)} + Z_0^{(i)}, \tag{3.1}$$

$$Y_1^{(i)} = h_{21}^{(i)} X_2^{(i)} + Z_1^{(i)}, \tag{3.2}$$

$$Y_2^{(i)} = h_{12}^{(i)} X_1^{(i)} + Z_2^{(i)}, \tag{3.3}$$

where, for each subchannel $i \in \{1, \ldots, N\}$, $X_k^{(i)}$ is the symbol transmitted by node k, $Z_j^{(i)}$ is the zero-mean additive white Gaussian noise at node j, $h_{kj}^{(i)}$ is the

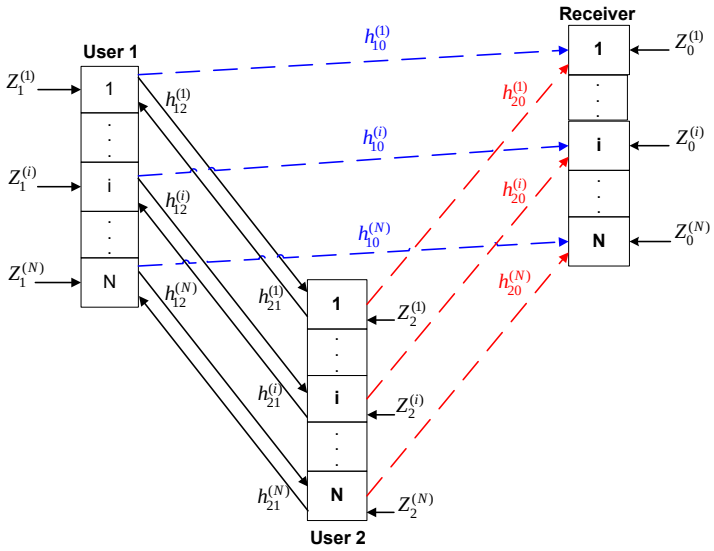


Figure 3.1: Gaussian Cooperative OFDMA Channel.

fading coefficient between nodes k and j, and $Y_j^{(i)}$ is the symbol received at node j; with $k \in \{1,2\}$, $j \in \{0,1,2\}$ and $k \neq j$. Here, the receiver is denoted by $j = 0$. The variance of $Z_j^{(i)}$ is given by ${\sigma_j^{(i)}}^2$. To simplify the notation throughout the thesis, we define the normalized power-fading coefficients $s_{kj}^{(i)} = \frac{(h_{kj}^{(i)})^2}{{\sigma_j^{(i)}}^2}$, and the Gaussian capacity function, $C(x) \triangleq \frac{1}{2}\log(1+x)$.

Note that, this model can be extended to a more general setting with many users, if we assign different subsets of available subchannels to separate pairs of users, and the results we obtain in this thesis will be readily extendable, thanks to the orthogonal nature of the channel.

3.3 Coding Techniques and Rate Regions for Cooperative OFDMA

The channel model given in (3.1)-(3.3) consists of N orthogonal two user cooperative multiple access channels (CMAC) in parallel, over which users 1 and 2 communicate with a common receiver. Therefore, the encoding and decoding techniques known for the two user CMAC [27] can be easily extended to this system with N subchannels in parallel, by simply dividing the total message to

be transmitted into smaller submessages, and encoding each submessage independently over the orthogonal cooperative channels. This approach, which we call **intra-subchannel cooperative encoding**, is relatively easy to implement, but it does not allow us to take full advantage of the diversity created across subchannels by the OFDM system. For instance, if on one subchannel, the inter-user link is very strong, but user-destination links are consistently very weak, the total data rate of the submessage on that link will be dictated by the weaker link. If, however, we allow submessages received by a user on each subchannel to be combined, re-encoded and forwarded to the intended receiver by a re-allocation of the received message onto the subchannels, we will potentially obtain better rates. Therefore, we propose an implementation of such an approach, and call it **inter-subchannel cooperative encoding**. Since both of the encoding policies rely on the same message generation process, we first discuss how the messages to be transmitted are formed. Next, the two encoding policies briefly explained above, will be presented in more detail and their corresponding achievable rate regions will be derived.

3.3.1 Message Generation

We assume that users 1 and 2 have independent messages, w_1 and w_2 respectively, to be conveyed to the receiver. These messages are first divided into two submessages, i.e., $w_1 = \{w_{10}, w_{12}\}$ and $w_2 = \{w_{20}, w_{21}\}$, as in [26, 27]. Here, w_{kj} is intended to be decoded by user j and the receiver, and w_{k0} is intended to be decoded by the receiver, where $j, k \in \{1, 2\}$ and $j \neq k$. These two submessages are further divided into N submessages each, to be separately transmitted over N subchannels:

$$\begin{aligned} w_{k0} &= \left\{ w_{k0}^{(1)}, ..., w_{k0}^{(N)} \right\}, \\ w_{kj} &= \left\{ w_{kj}^{(1)}, ..., w_{kj}^{(N)} \right\}, \end{aligned} \tag{3.4}$$

with the respective rates $\left\{ R_{k0}^{(1)}, ..., R_{k0}^{(N)} \right\}$ and $\left\{ R_{kj}^{(1)}, ..., R_{kj}^{(N)} \right\}$.

3.3.2 Intra-subchannel Cooperative Encoding

This encoding strategy is a rather trivial extension of the BMSE used in scalar channels [26, 27, 34], to OFDMA. The encoding in each subchannel is done independently, by mapping the submessages $w_{k0}^{(i)}$ and $w_{kj}^{(i)}$ onto codewords $X_k^{(i)}$ for each subchannel i. Following the notation in [27] and [34], and the message generation process described in (3.4), the transmitters divide their messages $w_k^{(i)}(b)$ in block $b = 1, \cdots, B$ into submessages, $\left\{w_{k0}^{(i)}(b), w_{kj}^{(i)}(b)\right\}_{i=1}^{N}$. These submessages are then encoded using block Markov encoding, i.e.,

$$X_{k0}^{(i)}\left(w_{k0}^{(i)}[b], u^{(i)}(w_{kj}^{(i)}[b-1], \hat{w}_{jk}^{(i)}[b-1])\right), \tag{3.5}$$

$$X_{kj}^{(i)}\left(w_{kj}^{(i)}[b], u^{(i)}(w_{kj}^{(i)}[b-1], \hat{w}_{jk}^{(i)}[b-1])\right), \tag{3.6}$$

$$U_k^{(i)}\left(w_{kj}^{(i)}[b-1], \hat{w}_{jk}^{(i)}[b-1]\right). \tag{3.7}$$

Here, $X_{k0}^{(i)}$ carries the fresh information intended for the receiver, $X_{kj}^{(i)}$ carries the information intended for transmitter j for cooperation in the next block and $U_k^{(i)}$ is the common information sent by both transmitters for the resolution of the remaining uncertainty from the previous block, all of which are transmitted over the ith subchannel and chosen from unit Gaussian distributions. The caret, as in $\hat{w}_{jk}^{(i)}[b-1]$, will be used to denote the estimates of messages in the cooperative partner throughout the thesis. Note that $\hat{w}_{kj}^{(i)}[b-1] = w_{kj}^{(i)}[b-1]$ when $u_1^{(i)}$ and $u_2^{(i)}$ are identical, i.e., $u_1^{(i)} = u_2^{(i)} = u^{(i)}$. Finally, user k's codeword in subchannel i is formed by the superposition of the codewords (3.5)-(3.7),

$$X_k^{(i)} = \sqrt{p_{k0}^{(i)}}X_{k0}^{(i)} + \sqrt{p_{kj}^{(i)}}X_{kj}^{(i)} + \sqrt{p_{U_k}^{(i)}}U_k^{(i)}, \tag{3.8}$$

for $j, k \in \{1, 2\}$, $k \neq j$. The codewords (3.5)-(3.7) in each subchannel, are assigned separate powers, which are required to satisfy the following average power constraints:

$$\sum_{i=1}^{N} p_{k0}^{(i)} + p_{kj}^{(i)} + p_{U_k}^{(i)} = \sum_{i=1}^{N} p_k^{(i)} \leq \bar{p}_k, \quad k = 1, 2. \tag{3.9}$$

Due to the orthogonal structure of OFDMA, the signals in separate subchannels may be independently decoded using backward decoding at the receiver, without interfering with each other. Therefore, by suitably extending the results in [27] to N parallel subchannels, it can be shown that the rate region given in the following theorem is achievable:

Theorem 3.1. *For a two user cooperative OFDMA channel which employs intra-subchannel cooperative encoding described by (3.5)-(3.8), an achievable rate region is given by the closure of the convex hull of all rate pairs (R_1, R_2) satisfying*

$$R_{12}^{(i)} < C_{12}^{(i)} \triangleq E\left[C\left(\frac{s_{12}^{(i)}p_{12}^{(i)}}{s_{12}^{(i)}p_{10}^{(i)}+1}\right)\right], \tag{3.10}$$

$$R_{21}^{(i)} < C_{21}^{(i)} \triangleq E\left[C\left(\frac{s_{21}^{(i)}p_{21}^{(i)}}{s_{21}^{(i)}p_{20}^{(i)}+1}\right)\right], \tag{3.11}$$

$$R_{10}^{(i)} < C_{10}^{(i)} \triangleq E\left[C\left(s_{10}^{(i)}p_{10}^{(i)}\right)\right], \tag{3.12}$$

$$R_{20}^{(i)} < C_{20}^{(i)} \triangleq E\left[C\left(s_{20}^{(i)}p_{20}^{(i)}\right)\right], \tag{3.13}$$

$$R_{10}^{(i)} + R_{20}^{(i)} < C_{0}^{(i)} \triangleq E\left[C\left(s_{10}^{(i)}p_{10}^{(i)} + s_{20}^{(i)}p_{20}^{(i)}\right)\right], \tag{3.14}$$

$$R_{1}^{(i)} + R_{2}^{(i)} < C_{s}^{(i)} \triangleq E\left[C\left(s_{10}^{(i)}p_{1}^{(i)} + s_{20}^{(i)}p_{2}^{(i)} + 2\sqrt{s_{10}^{(i)}s_{20}^{(i)}p_{U_1}^{(i)}p_{U_2}^{(i)}}\right)\right], \tag{3.15}$$

where $C(x) \triangleq \frac{1}{2}\log(1+x)$; $R_1 = \sum_{i=1}^{N} R_1^{(i)} \triangleq \sum_{i=1}^{N} R_{10}^{(i)} + R_{12}^{(i)}$, $R_2 = \sum_{i=1}^{N} R_2^{(i)} \triangleq \sum_{i=1}^{N} R_{20}^{(i)} + R_{21}^{(i)}$, and the convex hull is taken over all power allocation policies that satisfy (3.9).

It is more instructive to express the rate constraints in terms of R_1 and R_2 directly, rather than expressing them in terms of the rates of individual submessages as in Theorem 3.1. This is done in the following corollary.

Corollary 3.2. *The achievable rate region given in Theorem 3.1 is equivalent to the closure of the convex hull of all rate pairs (R_1, R_2) satisfying*

$$R_1 < \sum_{i=1}^{N} \min\left\{C_{12}^{(i)} + C_{10}^{(i)}, C_{s}^{(i)}\right\}, \tag{3.16}$$

$$R_2 < \sum_{i=1}^{N} \min\left\{C_{21}^{(i)} + C_{20}^{(i)}, C_s^{(i)}\right\}, \tag{3.17}$$

$$R_1 + R_2 < \sum_{i=1}^{N} \min\left\{C_s^{(i)}, C_{12}^{(i)} + C_{21}^{(i)} + C_0^{(i)}\right\}. \tag{3.18}$$

Proof. On each subchannel i, the rates of submessages $w_{12}^{(i)}$ and $w_{21}^{(i)}$ are subject to two constraints, one for the inter-user link and one for the user-receiver links. Therefore, the rate $R_k^{(i)}$ for user k on subchannel i is restricted by the minimum of $C_{kj}^{(i)} + C_{k0}^{(i)}$ and $C_s^{(i)}$ (note that unlike a traditional MAC with no cooperation, the sum rate bound required for error free decoding at the receiver may in fact be stricter than the single user bound, hence the need for the minimum operation). A similar argument is valid for the sum rate. Then, since the OFDMA channel consists of parallel channels over which independent messages are transmitted, the total rate R_k, $k = 1, 2$ and the sum rate $R_1 + R_2$, are simply bounded by the sum of their counterparts on each subchannel. □

The minimum operations, required over each subchannel as in (3.16)-(3.18), prevent us from efficiently exploiting the diversity offered by OFDMA. Therefore, we propose a new encoding strategy, explained in the following section, which removes the obligation to transmit the same submessages on user-user and user-receiver links over each subchannel, thereby moving the minimum operations in (3.16)-(3.18) outside of the summations.

3.3.3 Inter-subchannel Cooperative Encoding

The BMSE relies on users decoding part of each other's messages in each block, and re-encoding them in the next block. Although this can be done on a subchannel (and hence submessage) basis as described in Section 3.3.2, it can also be done by re-encoding the overall message received over all subchannels, by an appropriate re-partitioning of that message to subchannels, to be used in conjunction with the message generation process described in Section 3.3.1.

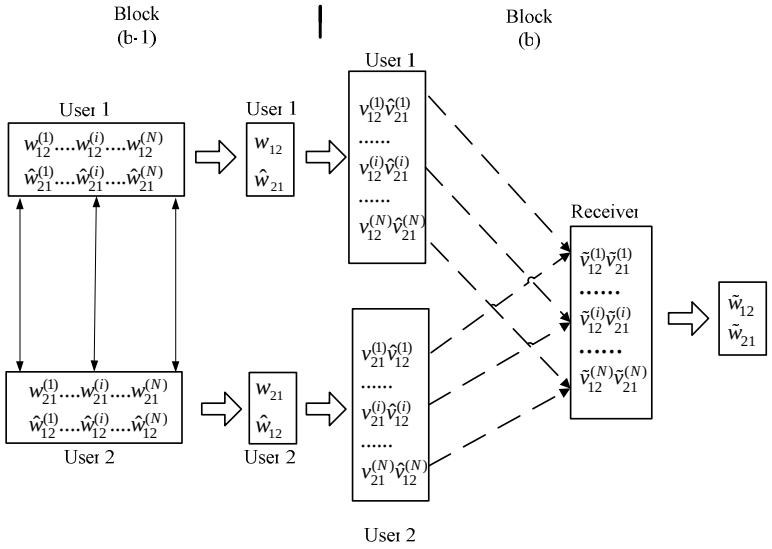


Figure 3.2: Re-partitioning of cooperative messages in inter-subchannel cooperative encoding. For the ease of demostration, the transmission to the receiver in block $b-1$, and the transmissions among the users in block b are not shown.

Let us denote, the estimate at user j, of the message transmitted in the previous block by user k over subchannel i, by $\hat{w}_{kj}^{(i)}$, $j \neq k$. Since user j's real objective is to decode the overall message w_{kj} and re-encode it, in the re-encoding process w_{kj} can be divided into new submessages with different rates:

$$w_{12} = \left\{ v_{12}^{(1)}, ..., v_{12}^{(N)} \right\}, \qquad w_{21} = \left\{ v_{21}^{(1)}, ..., v_{21}^{(N)} \right\}. \tag{3.19}$$

The corresponding submessages over each subchannel have the rates $\left\{ R_{12}^{'(1)}, ..., R_{12}^{'(N)} \right\}$ and $\left\{ R_{21}^{'(1)}, ..., R_{21}^{'(N)} \right\}$, respectively. The exchange of cooperative messages, their re-partitioning, retransmission and decoding is illustrated in Figure 3.2. It is assumed that a table to match the new cooperative submessages, $v_{kj}^{(i)}$, to the cooperative submessages received in the previous block, $w_{kj}^{(i)}$, is available at the users and the receiver. Hence, the receiver can obtain w_{kj} upon decoding $v_{kj}^{(i)}$ for all i.

It is important to note that, since $\left\{ w_{kj}^{(i)} \right\}_{i=1}^{N}$ and $\left\{ v_{kj}^{(i)} \right\}_{i=1}^{N}$, are just different partitionings of the same message w_{kj}, their total rates have to be the same, i.e.,

$$2^{nR_{12}} = 2^{nR_{12}^{(1)} + ... + nR_{12}^{(N)}} = 2^{nR_{12}^{'(1)} + ... + nR_{12}^{'(N)}}, \tag{3.20}$$

$$2^{nR_{21}} = 2^{nR_{21}^{(1)}+\ldots+nR_{21}^{(N)}} = 2^{nR_{21}^{'(1)}+\ldots+nR_{21}^{'(N)}}. \tag{3.21}$$

An achievable rate region obtained using the re-partitioning in (3.19) is given in the following theorem.

Theorem 3.3. *For a two user cooperative OFDMA channel, the rate region described by the closure of the convex hull of all rate pairs* (R_1, R_2) *satisfying*

$$R_1 = \sum_{i=1}^{N} R_{10}^{(i)} + R_{12}^{(i)} = \sum_{i=1}^{N} R_{10}^{(i)} + R_{12}^{'(i)}, \tag{3.22}$$

$$R_2 = \sum_{i=1}^{N} R_{20}^{(i)} + R_{21}^{(i)} = \sum_{i=1}^{N} R_{20}^{(i)} + R_{21}^{'(i)}, \tag{3.23}$$

with

$$R_{12}^{(i)} < C_{12}^{(i)}, \tag{3.24}$$

$$R_{21}^{(i)} < C_{21}^{(i)}, \tag{3.25}$$

$$R_{10}^{(i)} < C_{10}^{(i)}, \tag{3.26}$$

$$R_{20}^{(i)} < C_{20}^{(i)}, \tag{3.27}$$

$$R_{10}^{(i)} + R_{20}^{(i)} < C_{0}^{(i)}, \tag{3.28}$$

$$R_{12}^{'(i)} + R_{21}^{'(i)} + R_{10}^{(i)} + R_{20}^{(i)} < C_{s}^{(i)}, \tag{3.29}$$

is achievable; where the convex hull is taken over all power allocation policies that satisfy (3.9), and $C_{12}^{(i)}$, $C_{21}^{(i)}$, $C_{10}^{(i)}$, $C_{20}^{(i)}$, $C_{0}^{(i)}$ *and* $C_{s}^{(i)}$ *are defined in (3.10)-(3.15).*

Proof. Although the statement of the theorem is very similar to that of Theorem 3.1, there are now two sets of rates, $R_{kj}^{(i)}$ and $R_{kj}^{'(i)}$, arising from the modification we propose in block Markov encoding. The codebook generation, encoding, and decoding policies are obtained by an extension of the approach in [26], to accommodate multiple submessages, as described below.

Codebook Generation:

The following codebook generation procedure is repeated for each subchannel $i = 1, \cdots, N$.

- Generate $2^{n(R_{12}^{'(i)}+R_{21}^{'(i)})}$ length n sequences $u^{(i)}$ with i.i.d. unit Gaussian entries, and assign them to distinct message pairs $\{v_{12}^{(i)}, v_{21}^{(i)}\} \in \{1, \cdots, 2^{nR_{12}^{'(i)}}\} \times \{1, \cdots, 2^{nR_{21}^{'(i)}}\}$, to form $u^{(i)}(v_{12}^{(i)}, v_{21}^{(i)})$.

- For every $u^{(i)}(v_{12}^{(i)}, v_{21}^{(i)})$, generate $2^{nR_{12}^{(i)}}$ length n sequences $x_{12}^{(i)}$, and $2^{nR_{10}^{(i)}}$ length n sequences $x_{10}^{(i)}$ from independent unit Gaussian distributions, and assign them to distinct $w_{12}^{(i)} \in \{1, \cdots, 2^{nR_{12}^{(i)}}\}$ and $w_{10}^{(i)} \in \{1, \cdots, 2^{nR_{10}^{(i)}}\}$ respectively; to form $x_{12}^{(i)}(w_{12}^{(i)}, u^{(i)}(v_{12}^{(i)}, v_{21}^{(i)}))$ and $x_{10}^{(i)}(w_{10}^{(i)}, u^{(i)}(v_{12}^{(i)}, v_{21}^{(i)}))$.

- For every $u^{(i)}(v_{12}^{(i)}, v_{21}^{(i)})$, generate $2^{nR_{21}^{(i)}}$ length n sequences $x_{21}^{(i)}$ and $2^{nR_{20}^{(i)}}$ length n sequences $x_{20}^{(i)}$ from independent unit Gaussian distributions, and assign them to distinct $w_{21}^{(i)} \in \{1, \cdots, 2^{nR_{21}^{(i)}}\}$ and $w_{20}^{(i)} \in \{1, \cdots, 2^{nR_{20}^{(i)}}\}$ respectively; to form $x_{21}^{(i)}(w_{21}^{(i)}, u^{(i)}(v_{12}^{(i)}, v_{21}^{(i)}))$ and $x_{20}^{(i)}(w_{20}^{(i)}, u^{(i)}(v_{12}^{(i)}, v_{21}^{(i)}))$.

Encoding:

The way BMSE is executed is mostly similar to the case of intra-subchannel cooperative encoding, with the key difference that, in block b, the cooperative codewords $U_k^{(i)}$ of each user k are now assigned to $\hat{v}_{jk}^{(i)}[b-1]$ and $v_{kj}^{(i)}[b-1]$, which are re-partitionings of the cooperative messages exchanged in the previous block. Then, in each block $b = 1, \cdots, B$, appropriate codewords, which correspond to the messages from previous and current block, are selected from the randomly generated codebook above, i.e.,

$$X_{k0}^{(i)}\left(w_{k0}^{(i)}[b], u^{(i)}(v_{kj}^{(i)}[b-1], \hat{v}_{jk}^{(i)}[b-1])\right), \tag{3.30}$$

$$X_{kj}^{(i)}\left(w_{kj}^{(i)}[b], u^{(i)}(v_{kj}^{(i)}[b-1], \hat{v}_{jk}^{(i)}[b-1])\right), \tag{3.31}$$

$$U_k^{(i)}\left(v_{kj}^{(i)}[b-1], \hat{v}_{jk}^{(i)}[b-1]\right). \tag{3.32}$$

These codewords are superposed using (3.8), and are subject to the power constraint (3.9) as in Section 3.3.2.

Note that, in the first block, $b = 1$, cooperative messages are set to $\{v_{12}^{(i)}[1], v_{21}^{(i)}[1]\} = \{1, 1\}$; and in the last block, $b = B$, the fresh information is set to $\{w_{12}^{(i)}(B), w_{10}^{(i)}(B), w_{21}^{(i)}(B), w_{20}^{(i)}(B)\} = \{1, 1, 1, 1\}$, over each subchannel $i = 1, \cdots, N$.

Decoding:

For decoding, in each block b, and on each subchannel i, each user j uses joint typicality check to decode $\hat{w}_{kj}^{(i)}(b)$, using $X_{kj}^{(i)}$, and treating $X_{k0}^{(i)}$ as noise, leading to the constraints (3.34), (3.35). The receiver on the other hand uses backward decoding [26] to determine the transmitted messages. That is, in block B, the receiver decodes $\tilde{v}_{12}^{(i)}[B-1]$ and $\tilde{v}_{21}^{(i)}[B-1]$ over each subchannel i using joint typicality check, and therefore also knows $\tilde{w}_{12}[B-1]$ and $\tilde{w}_{21}[B-1]$. Then, in block $B-1$, it uses this information to jointly decode $\{\tilde{v}_{12}^{(i)}[B-2], \tilde{w}_{10}^{(i)}[B-1], \tilde{v}_{21}^{(i)}[B-2], \tilde{w}_{20}^{(i)}[B-1]\}$, based on which it may also deduce $\tilde{w}_{12}[B-2]$ and $\tilde{w}_{21}[B-2]$, and this process continues until the first block.

The decoding operation in an arbitrary block b for each subchannel $i = 1, ..., N$, is equivalent to finding $\tilde{v}_{12}^{(i)}[b-1], \tilde{v}_{21}^{(i)}[b-1], \tilde{w}_{10}^{(i)}[b]$ and $\tilde{w}_{20}^{(i)}[b]$ for which

$$\Big\{ y^{(i)}[b], u^{(i)}(\tilde{v}_{12}^{(i)}[b-1], \tilde{v}_{21}^{(i)}[b-1]), x_{12}^{(i)}(\tilde{w}_{12}^{(i)}[b], u^{(i)}(\tilde{v}_{12}^{(i)}[b-1], \tilde{v}_{21}^{(i)}[b-1])),$$
$$x_{21}^{(i)}(\tilde{w}_{21}^{(i)}[b], u^{(i)}(\tilde{v}_{12}^{(i)}[b-1], \tilde{v}_{21}^{(i)}[b-1])), x_{1}^{(i)}(\tilde{w}_{10}^{(i)}[b], \tilde{w}_{12}^{(i)}[b], u^{(i)}(\tilde{v}_{12}^{(i)}[b-1], \tilde{v}_{21}^{(i)}[b-1])),$$
$$x_{2}^{(i)}(\tilde{w}_{20}^{(i)}[b], \tilde{w}_{21}^{(i)}[b], u^{(i)}(\tilde{v}_{12}^{(i)}[b-1], \tilde{v}_{21}^{(i)}[b-1])) \Big\} \tag{3.33}$$

are jointly typical. The estimates of the re-partitioned cooperative messages $\tilde{v}_{kj}^{(i)}[b-1]$ are converted to estimates of the cooperative messages $\tilde{w}_{kj}^{(i)}[b-1]$ using the match-up table available at the users and the receiver.

Using this decoding strategy, and well known properties of jointly typical sequences [26, 48], we show in the Appendix A that for n large enough, the average error probability can be made arbitrarily close to zero provided that the rates

satisfy the following constraints:

$$R_{12}^{(i)} < I(X_{12}^{(i)}; Y_2^{(i)} | X_2^{(i)}, U^{(i)}), \tag{3.34}$$

$$R_{21}^{(i)} < I(X_{21}^{(i)}; Y_1^{(i)} | X_1^{(i)}, U^{(i)}), \tag{3.35}$$

$$R_{10}^{(i)} < I(X_1^{(i)}; Y^{(i)} | X_2^{(i)}, X_{12}^{(i)}, U^{(i)}), \tag{3.36}$$

$$R_{20}^{(i)} < I(X_2^{(i)}; Y^{(i)} | X_1^{(i)}, X_{21}^{(i)}, U^{(i)}), \tag{3.37}$$

$$R_{10}^{(i)} + R_{20}^{(i)} < I(X_1^{(i)}, X_2^{(i)}; Y^{(i)} | X_{12}^{(i)}, X_{21}^{(i)}, U^{(i)}), \tag{3.38}$$

$$R_{12}^{'(i)} + R_{21}^{'(i)} + R_{10}^{(i)} + R_{20}^{(i)} < I(X_1^{(i)}, X_2^{(i)}; Y^{(i)}). \tag{3.39}$$

Finally, evaluating (3.34)-(3.39) for Gaussian codewords, we obtain the desired result. □

The improvement in the set of achievable rate pairs (R_1, R_2), obtained by inter-subchannel cooperative encoding, becomes more apparent if we restate the achievable rates in Theorem 3.3 in terms of total rates of the users:

Corollary 3.4. *The achievable rate region given in Theorem 3.3 is equivalent to the closure of the convex hull of all rate pairs* (R_1, R_2) *satisfying*

$$R_1 < \sum_{i=1}^{N} C_{12}^{(i)} + C_{10}^{(i)}, \tag{3.40}$$

$$R_2 < \sum_{i=1}^{N} C_{21}^{(i)} + C_{20}^{(i)}, \tag{3.41}$$

$$R_1 + R_2 < \min \left\{ \sum_{i=1}^{N} C_s^{(i)}, \sum_{i=1}^{N} C_{12}^{(i)} + C_{21}^{(i)} + C_0^{(i)} \right\}. \tag{3.42}$$

Proof. Constraints (3.40)-(3.41) follow trivially from (3.22)-(3.28). Constraint (3.42) follows from (3.22)-(3.25), (3.28)-(3.29) and the fact that $R_{kj}^{(i)}$ and $R_{kj}^{'(i)}$ are constrained separately, and need not be equal on a given subchannel i; as long as their sum over all subchannels remains the same. □

Comparing the achievable rate region (3.40)-(3.42) of Corollary 2 with (3.16)-(3.18) of Corollary 1, we see that the minimum operations required for the individual rate constraints for each subchannel are removed, and the minimum operation required for the sum rate constraint (3.18) is taken outside the summation over the subchannels. This eliminates possible bottlenecks on achievable rates, caused by the per-subchannel constraints, and the rate region obtained by inter-subchannel cooperative encoding contains that obtained by intra-subchannel cooperation. The rate regions achievable by both strategies will be compared for some sample fading scenarios, in Section 3.4.

3.3.4 Half-duplex Cooperative Encoding

The two cooperative OFDMA models proposed so far assumed full-duplex operation in each subchannel. However, in practice, due to the vast difference between the transmitted and received signal strengths, it is not possible to transmit and receive simultaneously on the same band. Hence, in this section, we define a more practical orthogonal cooperative encoding strategy, with half-duplex operation in each subchannel. The available subchannels are first divided into three sets, i.e., $\mathcal{I} \triangleq \{1, \ldots, N\} = \{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3\}$. On subchannels $i \in \mathcal{I}_1$, user 1 transmits, while user 2 and the receiver listen, on subchannels $i \in \mathcal{I}_2$, user 2 transmits while user 1 and the receiver listen, and on $i \in \mathcal{I}_3$, both users transmit simultaneously, to obtain coherent combining gain, while only the receiver listens. The encoding is done across subchannels, allowing the messages received in the previous block to be re-partitioned in the next block. Therefore, our proposed half-duplex scheme is a special case of inter-subchannel cooperative encoding, with $p_{20}^{(i)} = p_{21}^{(i)} = p_{U_1}^{(i)} = p_{U_2}^{(i)} = 0$ for $i \in \mathcal{I}_1$, $p_{10}^{(i)} = p_{12}^{(i)} = p_{U_1}^{(i)} = p_{U_2}^{(i)} = 0$ for $i \in \mathcal{I}_2$ and $p_{12}^{(i)} = p_{21}^{(i)} = 0$ for $i \in \mathcal{I}_3$. Substituting these power levels in (3.40)-(3.42), we obtain the following theorem [57], which describes the achievable rate region for the half-duplex cooperative encoding strategy.

Theorem 3.5. *For a two user cooperative OFDMA channel which employs half-duplex cooperative encoding, an achievable rate region is given by the closure of the convex hull of all rate pairs* (R_1, R_2) *satisfying*

$$R_1 < \sum_{i\in\mathcal{I}_1} E\left[C\left(\frac{s_{12}^{(i)}p_{12}^{(i)}}{s_{12}^{(i)}p_{10}^{(i)}+1}\right)\right] + \sum_{i\in\{\mathcal{I}_1,\mathcal{I}_3\}} E\left[C\left(s_{10}^{(i)}p_{10}^{(i)}\right)\right], \tag{3.43}$$

$$R_2 < \sum_{i\in\mathcal{I}_2} E\left[C\left(\frac{s_{21}^{(i)}p_{21}^{(i)}}{s_{21}^{(i)}p_{20}^{(i)}+1}\right)\right] + \sum_{i\in\{\mathcal{I}_2,\mathcal{I}_3\}} E\left[C\left(s_{20}^{(i)}p_{20}^{(i)}\right)\right], \tag{3.44}$$

$$\begin{aligned} R_1+R_2 < \min\Bigg\{ & \sum_{i\in\mathcal{I}_3} E\left[C\left(s_{10}^{(i)}p_1^{(i)}+s_{20}^{(i)}p_2^{(i)}+2\sqrt{s_{10}^{(i)}s_{20}^{(i)}p_{U_1}^{(i)}p_{U_2}^{(i)}}\right)\right] \\ & + \sum_{i\in\mathcal{I}_1} E\left[C\left(s_{10}^{(i)}(p_{10}^{(i)}+p_{12}^{(i)})\right)\right] \\ & + \sum_{i\in\mathcal{I}_2} E\left[C\left(s_{20}^{(i)}(p_{20}^{(i)}+p_{21}^{(i)})\right)\right], \end{aligned} \tag{3.45}$$

$$\begin{aligned} & \sum_{i\in\mathcal{I}_1} E\left[C\left(\frac{s_{12}^{(i)}p_{12}^{(i)}}{s_{12}^{(i)}p_{10}^{(i)}+1}\right)+C\left(s_{10}^{(i)}p_{10}^{(i)}\right)\right] \\ & + \sum_{i\in\mathcal{I}_2} E\left[C\left(\frac{s_{21}^{(i)}p_{21}^{(i)}}{s_{21}^{(i)}p_{20}^{(i)}+1}\right)+C\left(s_{20}^{(i)}p_{20}^{(i)}\right)\right] \\ & + \sum_{i\in\mathcal{I}_3} E\left[C\left(s_{10}^{(i)}p_{10}^{(i)}+s_{20}^{(i)}p_{20}^{(i)}\right)\right]\Bigg\}. \end{aligned} \tag{3.46}$$

Proof. The proof relies on the key observation that our proposed half-duplex scheme may be viewed as a special case of inter-subchannel cooperative encoding, with $p_{20}^{(i)} = p_{21}^{(i)} = p_{U_1}^{(i)} = p_{U_2}^{(i)} = 0$ for $i \in \mathcal{I}_1$, $p_{10}^{(i)} = p_{12}^{(i)} = p_{U_1}^{(i)} = p_{U_2}^{(i)} = 0$ for $i \in \mathcal{I}_2$ and $p_{12}^{(i)} = p_{21}^{(i)} = 0$ for $i \in \mathcal{I}_3$. Substituting these power levels in (3.40)-(3.42), we obtain the desired result. □

3.4 Simulation Results

In this section we evaluate the achievable rate regions (3.40)-(3.42) for inter-subchannel cooperative encoding and (3.16)-(3.18) for intra-subchannel cooperative encoding, and compare them with the capacity region of a non-cooperative

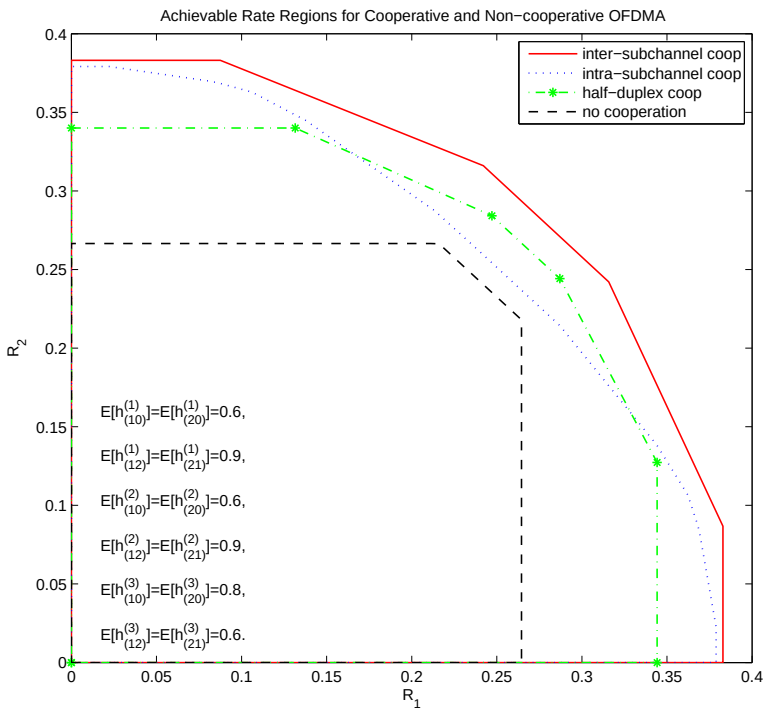


Figure 3.3: Achievable rate regions for fading scenario 1.

OFDMA system, as well as the achievable rate region of our half-duplex cooperative encoding strategy (3.43)-(3.46). We assume that, for all four protocols, the users are able to allocate their total power across subchannels and codewords, and the achievable rate regions are generated by taking the convex hull over all valid power allocation policies. Note however that adaptive power allocation in terms of instantaneous fading states is not considered, the power assigned to each subchannel and codeword remains the same throughout the transmission. The total power of each user and the noise variances are both set to unity (except for Figure 3.5 where Signal to Noise Ratio (SNR) is varied).

In Figure 3.3 the achievable rate/capacity regions are generated for fading scenario 1, where we choose channel gains from independent Rayleigh distributions, the means of which are shown in the figure. For half-duplex cooperation, we simply set $\mathcal{I}_1 = \{1\}$, $\mathcal{I}_2 = \{2\}$ and $\mathcal{I}_3 = \{3\}$, which is in fact optimal among all fixed subchannel allocations. We see that the single user achievable rates for both cooperative strategies are similar, but the gap between the achievable rates

of the two policies increase near the sum rate point. This can be explained as follows: the cooperative links on the first two subchannels are better on average, while the direct link is better on the third. Therefore, although the term $C_s^{(i)}$ is not very restrictive on the single user rates (3.16)-(3.17), it is restrictive for the sum rate (3.18). Since inter-subchannel cooperative coding can assign powers to cooperative codewords so that it mostly uses the third subchannel to send the cooperative codeword U, and the first two for cooperation among users, its advantage near the sum rate point is more pronounced. Under these channel conditions, we see that sum rates for both cooperative strategies are always higher than non-cooperative OFDMA while the sum rate of half-duplex cooperative encoding is higher than that of the intra-subchannel cooperative strategy since the direct links are stronger over the first and second subchannels and the cooperative link is stronger over the third subchannel, and with our selection of $\mathcal{I}_1$, $\mathcal{I}_2$ and $\mathcal{I}_3$, the half-duplex cooperative encoding uses subchannels 1 and 2 to create common information, and subchannel 3 for direct transmission to obtain coherent combining gain.

In Figure 3.4, we switch to the fading scenario 2, and the gap among the two cooperative policies becomes more apparent. Note that the direct link gain of user 1 is worse than user 2 on subchannel 1, and vice versa on subchannel 2. Therefore, the rate gains achievable in those subchannels are hindered by the per-subchannel constraints of intra-subchannel cooperation, and in fact around the sum rate point it achieves rates only slightly better than non-cooperative OFDMA, while inter-subchannel cooperation still gives large gains. The sum rate of half-duplex cooperation, which uses the same optimal channel assignment as above, is less than both full-duplex cooperative strategies due to the fact that the achievable rate of the half-duplex cooperative encoding relies heavily on the direct link gains on subchannel 3, that is relatively weak for this fading scenario.

In Figure 3.5, we provide further simulation results comparing the sum rates achievable by intra and inter-subchannel cooperation, as a function of the average transmit signal power to noise power ratio, for three other fading scenarios (3-5)

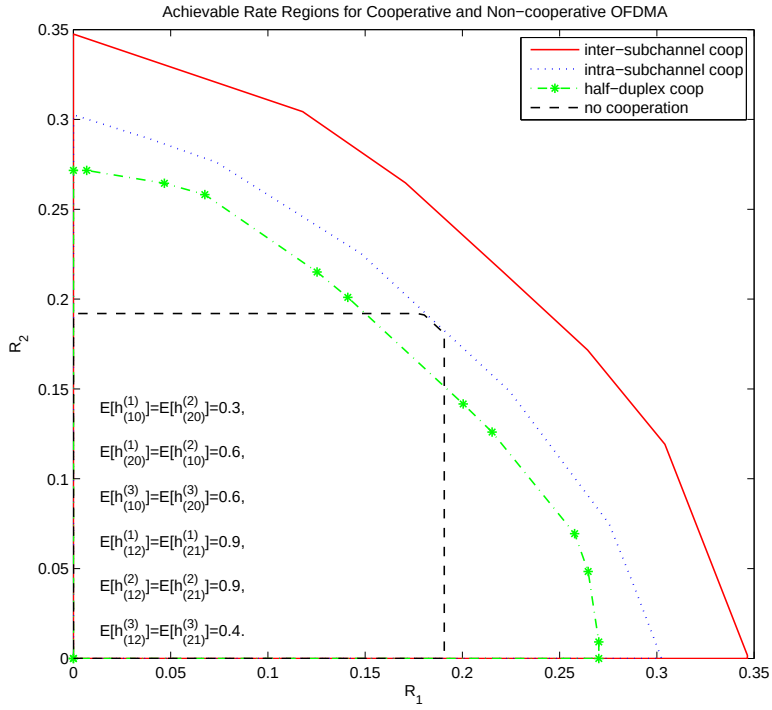


Figure 3.4: Achievable rate regions for fading scenario 2.

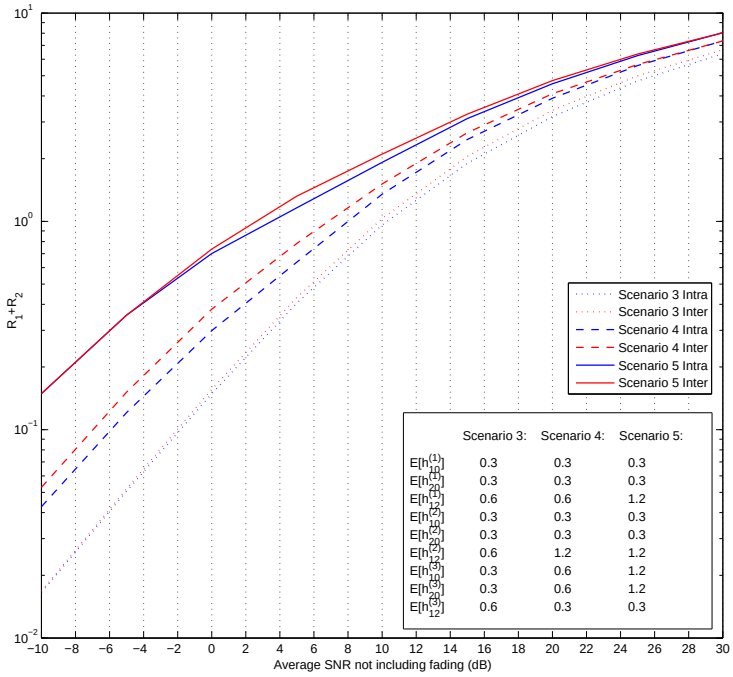


Figure 3.5: Comparison of sum rates achievable by intra and inter-subchannel cooperative encoding, as a function of SNR.

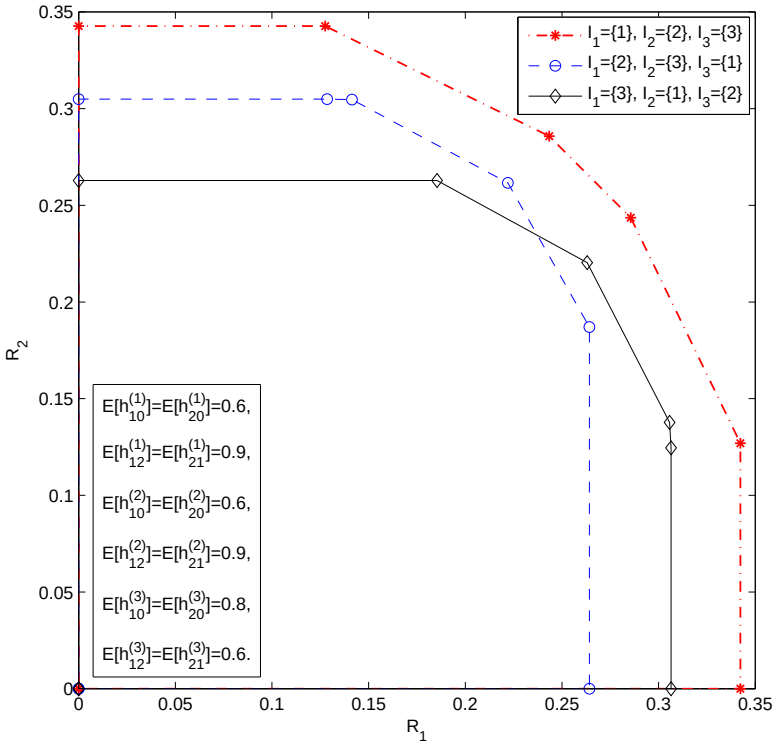


Figure 3.6: Comparison of varying subchannel assignments for a simple half-duplex setup with three subchannels.

described in the respective figures. We observe the relative behavior for varying SNR to depend highly on the fading statistics, but especially when the fading is asymmetric, the gain is more pronounced for low and moderate SNR values. The performance depends on relative average strengths of fading coefficients on subchannels, rather than SNR: in scenarios 4 and 5, there is nearly 2dB gain at low and moderate SNR values respectively, in the symmetric scenario 3 where no subchannel provides a relative advantage to either policy, inter-subchannel cooperation only slightly outperforms intra-subchannel cooperation over almost all SNR values.

While the general rate expressions for half-duplex strategy are derived for an arbitrary channel assignment, for the simple simulation setting with three subchannels allocated for the pair of users, we assigned the channels so that the first subchannel, with the best average channel gain from user 1 to user 2, is selected as $\mathcal{I}_1$, dedicated for user 1 to user 2 (and hence automatically for user 1 to the receiver) transmission, likewise subchannel two is selected as $\mathcal{I}_2$, and subchannel

three, with better direct link gains is selected as $\mathcal{I}_1$, to convey established common information to the receiver. In fact, this strategy is also optimal for the case in question, as it is rather easy to search over all possible subchannel allocations. To demonstrate the optimality, we performed a simple simulation, and obtained the rate regions achievable by other possible assignments, as given in Figure 3.6. Note that only those allocations leading to different rate regions are plotted.

3.5 Conclusion

In this chapter we introduced a two user cooperative OFDMA system, and we proposed two full duplex encoding strategies: intra-subchannel cooperative encoding and inter-subchannel cooperative encoding, based on BMSE. We derived rate region expressions for both encoding strategies and showed that re-partitioning and re-encoding of the cooperative messages across subchannels, i.e., inter-subchannel cooperative encoding, is always superior to intra-subchannel cooperative encoding, and provides significant rate gains. We further proposed a half-duplex implementation of cooperation based on inter-subchannel cooperative encoding, and showed that despite its orthogonal structure, it too may outperform the full duplex intra-subchannel cooperation. Achievable rates for all three proposed scenarios show that it is very advantageous to make use of overheard information, especially taking into account the added diversity created by the OFDMA subchannels. The results provided in this chapter were presented in part in [58], and were published in [57].

Chapter 4

Resource Allocation For Two User Cooperative OFDMA Channels

4.1 Introduction

The ability of OFDMA to cope with both intersymbol and interuser interference, combined with its low complexity of implementation, have made it a popular choice for the next generation wireless networks. As a result, the problem of resource allocation in OFDMA systems was studied extensively in the literature. One example is [59], where it was proved that in an OFDMA uplink system, allocating subcarriers to the users with the maximum marginal rate is a necessary condition for maximizing the system throughput. A similar problem was solved in [13] using KKT conditions, by optimizing a utility function which was assumed to be a function of the rates. In [12], a low-complexity algorithm for subcarrier, power, and rate allocation for OFDMA was proposed, to maximize the sum rate under individual rate constraints to guarantee fairness. The downlink ergodic sum rate maximization problem was considered in [15], where the authors developed a linear complexity subcarrier and power allocation algorithm. These works, as well as many others on OFDMA, naturally assume orthogonal multiple access, thereby choosing to avoid interference. However, like all orthogonal transmission techniques, OFDMA incurs some rate penalty. Moreover, "interference" in wireless channels is in fact free side information, and not ignoring it opens up the possibility of user cooperation. Therefore, in this chapter, we focus on resource

allocation for a two user OFDMA channel, which allows for mutual cooperation among the users over each subchannel, each taking into account the available side information.

The overheard information in a typical wireless MAC, is captured by modeling the system as a MAC-GF [26]. In [26], achievable rates for the MAC-GF were obtained based on BMSE and backwards decoding. In [27], these encoding and decoding techniques were applied to a Gaussian MAC in fading, and the resulting rate regions were characterized. In [34], power allocation policies that maximize the rates achievable by BMSE for the same model were obtained.

While the above works all deal with a scalar MAC-GF, some works on resource allocation for user cooperation in vector channels, specifically OFDMA, also exist. A cooperative OFDMA system where each user is allowed to transmit and receive at the same time, but necessarily on different subcarriers, was considered in [60]. Subcarrier and power allocation schemes for a time division duplex amplify and forward protocol were employed in [61] with the aim of maximizing system throughput and enhancing fairness in a cooperative OFDMA uplink system. Resource allocation and cooperative partner selection in cooperative OFDM networks was investigated with the objective of minimizing the overall power in [62]. However, these works consider either a one sided cooperation strategy, or a mutually cooperative strategy based on two parallel dedicated relay channels, or mutual cooperation based on a time division protocol. We introduced a more general cooperative OFDMA model based on parallel MAC-GFs, which does not make any prior assumptions about the way in which the subchannels are assigned to the users in Chapter 3, where two full-duplex cooperative encoding strategies, namely intra-subchannel cooperative encoding and inter-subchannel cooperative encoding were proposed. However, in Chapter 3 we did not apply power allocation as a function of fading states, which can further take advantage of the temporal diversity over each subchannel.

In this chapter, we extend the cooperative OFDMA model proposed in Chapter 3, to a channel adaptive scenario, and we solve the problem of power optimization with the goal of maximizing the achievable rates. We first obtain the properties of the power allocation policy that maximizes the sum rate of the cooperative OFDMA system employing intra-subchannel cooperative encoding and inter-subchannel cooperative encoding. We then focus on the inter-subchannel cooperative encoding, which provably outperforms intra-subchannel cooperative encoding and derive the expressions of the optimum power levels that maximize the rate region achievable by the inter-subchannel cooperative encoding strategy. Despite the complex re-encoding structure employed in inter-subchannel cooperative encoding, and the fact that the powers allocated to each subchannel have to satisfy a sum power constraint over subchannels, the achievable rate region turns out to be of a relatively similar form to its scalar counterpart, and we are able to extend some properties of the optimal power allocation derived in [34] for scalar cooperative MAC, to cooperative OFDMA. As a result, the weighted sum of rates, which can be used to obtain any point on the rate region boundary, becomes concave, and convex optimization techniques can be employed. To obtain an arbitrary rate point on the achievable rate region boundary, we first employ a projected subgradient algorithm that converges to the optimum and maximizes the achievable rate region. Next, we derive the optimality conditions, and closed form expressions for optimum powers analytically. We are then able to propose an alternative efficient iterative algorithm with a much lower complexity, to obtain the rate points on the achievable rate region boundary. This algorithm works by solving the KKT optimality conditions iteratively over the users, to obtain the optimal powers. As a result, we demonstrate that by jointly exploiting the diversity provided by OFDMA's parallel subchannels, and the temporal diversity created by the time varying channel, we obtain very promising gains in achievable rates. More interestingly, we observe that the optimal power allocation may automatically dictate that some subchannels are assigned exclusively to certain users/tasks, depending on the instantaneous channel state.

4.2 System Model

We consider a two user cooperative OFDMA system with N subchannels. On each subchannel, unless otherwise stated, each user is capable of both transmitting and receiving signals. The system model given in Chapter 3, Figure 3.1 is repeated here for convenience:

$$Y_0^{(i)} = h_{10}^{(i)} X_1^{(i)} + h_{20}^{(i)} X_2^{(i)} + Z_0^{(i)}, \tag{4.1}$$

$$Y_1^{(i)} = h_{21}^{(i)} X_2^{(i)} + Z_1^{(i)}, \tag{4.2}$$

$$Y_2^{(i)} = h_{12}^{(i)} X_1^{(i)} + Z_2^{(i)}. \tag{4.3}$$

For a real number x, we define $(x)^+ = \max(x, 0)$ and introduce the function

$$f(a, b, c) = \left(\frac{-b + \sqrt{b^2 - 4ac}}{2a} \right)^+, \tag{4.4}$$

which will be used extensively while characterizing the optimum power allocation, as some of the optimal powers turn out to be the non-negative roots of certain quadratic equations.

4.3 Long-Term Achievable Rates for Cooperative OFDMA

While the system model is the same as the system model in Chapter 3, we now assume that the users and the receiver have full CSI of both the cooperative links and the direct link. Therefore, differently from Chapter 3, the users can further adapt their transmitted symbols $X_k^{(i)}$ as a function of the joint fading state, $\mathbf{s}$, i.e., $X_k^{(i)}(\mathbf{s})$, and perform the channel adaptation through powers. There are two ways to view the channel adaptive transmission that maximizes the long term (ergodic) achievable rates: we can either use a variable power, variable rate codebook, as in [4], or we can use a single codebook, whose rate is supported by the channel in the long term, and perform the channel adaptation by simply multiplying entries from this codebook by channel adaptive powers, as in [6]. In our thesis, we employ the

latter approach, and propose a channel adaptive version of the encoding strategies in Chapter 3, where we scale each of the above codewords by variable powers,

$$X_k^{(i)} = \sqrt{p_{k0}^{(i)}(\mathbf{s})}X_{k0}^{(i)} + \sqrt{p_{kj}^{(i)}(\mathbf{s})}X_{kj}^{(i)} + \sqrt{p_{U_k}^{(i)}(\mathbf{s})}U_k^{(i)}, \tag{4.5}$$

where $k, j \in \{1, 2\}$, $k \neq j$, $i = 1, \cdots, N$. The powers are subject to the average power constraints,

$$\sum_{i=1}^{N} E\Big[p_{k0}^{(i)}(\mathbf{s}) + p_{kj}^{(i)}(\mathbf{s}) + p_{U_k}^{(i)}(\mathbf{s})\Big] \triangleq \sum_{i=1}^{N} E\Big[p_k^{(i)}(\mathbf{s})\Big] \leq \bar{p}_k. \tag{4.6}$$

The achievable rate regions for power controlled intra-subchannel cooperative encoding and inter-subchannel cooperative encoding are obtained by extending Theorem 3.1 and Theorem 3.3 respectively, using the new the channel adaptive encoding defined in (4.5). The resulting achievable rate region for intra-subchannel cooperative encoding with power control is given by the closure of the convex hull of all rate pairs (R_1, R_2) satisfying

$$\begin{aligned} R_1 < \sum_{i=1}^{N} \min \Bigg\{ & E\left[C\left(\frac{s_{12}^{(i)}p_{12}^{(i)}(\mathbf{s})}{s_{12}^{(i)}p_{10}^{(i)}(\mathbf{s})+1}\right) + C\left(s_{10}^{(i)}p_{10}^{(i)}(\mathbf{s})\right)\right], \\ & E\left[C\left(s_{10}^{(i)}p_1^{(i)}(\mathbf{s}) + s_{20}^{(i)}p_2^{(i)}(\mathbf{s}) + 2\sqrt{s_{10}^{(i)}s_{20}^{(i)}p_{U_1}^{(i)}(\mathbf{s})p_{U_2}^{(i)}(\mathbf{s})}\right)\right]\Bigg\}, \end{aligned} \tag{4.7}$$

$$\begin{aligned} R_2 < \sum_{i=1}^{N} \min \Bigg\{ & E\left[C\left(\frac{s_{21}^{(i)}p_{21}^{(i)}(\mathbf{s})}{s_{21}^{(i)}p_{20}^{(i)}(\mathbf{s})+1}\right) + C\left(s_{20}^{(i)}p_{20}^{(i)}(\mathbf{s})\right)\right], \\ & E\left[C\left(s_{10}^{(i)}p_1^{(i)}(\mathbf{s}) + s_{20}^{(i)}p_2^{(i)}(\mathbf{s}) + 2\sqrt{s_{10}^{(i)}s_{20}^{(i)}p_{U_1}^{(i)}(\mathbf{s})p_{U_2}^{(i)}(\mathbf{s})}\right)\right]\Bigg\}, \end{aligned} \tag{4.8}$$

$$\begin{aligned} R_1 + R_2 < \sum_{i=1}^{N} \min \Bigg\{ & E\left[C\left(s_{10}^{(i)}p_1^{(i)}(\mathbf{s}) + s_{20}^{(i)}p_2^{(i)}(\mathbf{s}) + 2\sqrt{s_{10}^{(i)}s_{20}^{(i)}p_{U_1}^{(i)}(\mathbf{s})p_{U_2}^{(i)}(\mathbf{s})}\right)\right], \\ & E\left[C\left(\frac{s_{12}^{(i)}p_{12}^{(i)}(\mathbf{s})}{s_{12}^{(i)}p_{10}^{(i)}(\mathbf{s})+1}\right) + C\left(\frac{s_{21}^{(i)}p_{21}^{(i)}(\mathbf{s})}{s_{21}^{(i)}p_{20}^{(i)}(\mathbf{s})+1}\right) + C\left(s_{10}^{(i)}p_{10}^{(i)}(\mathbf{s}) + s_{20}^{(i)}p_{20}^{(i)}(\mathbf{s})\right)\right]\Bigg\}, \end{aligned} \tag{4.9}$$

and the achievable rate region for inter-subchannel cooperative encoding with power control is given by the closure of the convex hull of all rate pairs (R_1, R_2)

satisfying

$$R_1 < \sum_{i=1}^{N} E\left[C\left(\frac{s_{12}^{(i)}p_{12}^{(i)}(\mathbf{s})}{s_{12}^{(i)}p_{10}^{(i)}(\mathbf{s})+1}\right)+C\left(s_{10}^{(i)}p_{10}^{(i)}(\mathbf{s})\right)\right], \tag{4.10}$$

$$R_2 < \sum_{i=1}^{N} E\left[C\left(\frac{s_{21}^{(i)}p_{21}^{(i)}(\mathbf{s})}{s_{21}^{(i)}p_{20}^{(i)}(\mathbf{s})+1}\right)+C\left(s_{20}^{(i)}p_{20}^{(i)}(\mathbf{s})\right)\right], \tag{4.11}$$

$$\begin{aligned} R_1+R_2 < \min\Bigg\{ & \sum_{i=1}^{N} E\left[C\left(s_{10}^{(i)}p_1^{(i)}(\mathbf{s})+s_{20}^{(i)}p_2^{(i)}(\mathbf{s})+2\sqrt{s_{10}^{(i)}s_{20}^{(i)}p_{U_1}^{(i)}(\mathbf{s})p_{U_2}^{(i)}(\mathbf{s})}\right)\right], \\ & \sum_{i=1}^{N} E\left[C\left(\frac{s_{12}^{(i)}p_{12}^{(i)}(\mathbf{s})}{s_{12}^{(i)}p_{10}^{(i)}(\mathbf{s})+1}\right)+C\left(\frac{s_{21}^{(i)}p_{21}^{(i)}(\mathbf{s})}{s_{21}^{(i)}p_{20}^{(i)}(\mathbf{s})+1}\right)+C\left(s_{10}^{(i)}p_{10}^{(i)}(\mathbf{s})+s_{20}^{(i)}p_{20}^{(i)}(\mathbf{s})\right)\right]\Bigg\}, \end{aligned} \tag{4.12}$$

where the convex hulls are taken over all valid power allocation policies. In the next section, we obtain the power allocation policies which achieve the rate tuples on the rate region boundary. To do this, we first derive a simplifying property of the optimal power allocation for both cooperative encoding strategies, and then we focus on inter-subchannel cooperative encoding that provides superior achievable rates.

4.4 Channel Adaptive Power Allocation

If we set $N = 1$ in (4.7)-(4.9) or (4.10)-(4.12), the problem reduces to a scalar cooperative MAC. In [34], it was shown for this scalar case that, based on the instantaneous channel state, the optimal power allocation dictates that each user either sends cooperative information, or fresh information, but not both. Although in OFDMA, there is a sum power constraint over the subchannels, and one would expect the power allocation over each subchannel to be dependent on the powers assigned to the other subchannels, we show that many properties of the optimal power allocation for the proposed cooperative OFDMA system remain surprisingly parallel to those in the scalar case [34], and the codewords that should be used over each subchannel are determined solely by the instantaneous

fading coefficients over that particular subchannel, as stated in the following theorem:

Theorem 4.1. *The power allocation policy that maximizes the sum rate of a cooperative OFDMA system using intra-subchannel cooperative encoding and inter-subchannel cooperative encoding should satisfy;*

1. $p_{10}^{(i)*}(\mathbf{s}) = p_{20}^{(i)*}(\mathbf{s}) = 0$, *if* $\mathbf{s} \in \mathcal{S}_1$,
2. $p_{10}^{(i)*}(\mathbf{s}) = p_{21}^{(i)*}(\mathbf{s}) = 0$, *if* $\mathbf{s} \in \mathcal{S}_2$,
3. $p_{12}^{(i)*}(\mathbf{s}) = p_{20}^{(i)*}(\mathbf{s}) = 0$, *if* $\mathbf{s} \in \mathcal{S}_3$,
4. $p_{12}^{(i)*}(\mathbf{s}) = p_{21}^{(i)*}(\mathbf{s}) = 0$ *or* $p_{10}^{(i)*}(\mathbf{s}) = p_{21}^{(i)*}(\mathbf{s}) = 0$ *or* $p_{12}^{(i)*}(\mathbf{s}) = p_{20}^{(i)*}(\mathbf{s}) = 0$, *if* $\mathbf{s} \in \mathcal{S}_4$,

where $\mathcal{S}_1 = \{\mathbf{s} : \ s_{12}^{(i)} > s_{10}^{(i)}, \ s_{21}^{(i)} > s_{20}^{(i)}\}$, $\mathcal{S}_2 = \{\mathbf{s} : \ s_{12}^{(i)} > s_{10}^{(i)}, \ s_{21}^{(i)} \leq s_{20}^{(i)}\}$, $\mathcal{S}_3 = \{\mathbf{s} : \ s_{12}^{(i)} \leq s_{10}^{(i)}, \ s_{21}^{(i)} > s_{20}^{(i)}\}$, $\mathcal{S}_4 = \{\mathbf{s} : \ s_{12}^{(i)} \leq s_{10}^{(i)}, \ s_{21}^{(i)} \leq s_{20}^{(i)}\}$.

Proof. Assume that we know the total optimal power $p_k^{(i)*}(\mathbf{s})$, allocated to each subchannel i at each channel state $\mathbf{s}$. Then, for intra-subchannel cooperative encoding, the sum rate (4.9) is maximized if each term in the summation is maximized. Since the total power allocated to each term is fixed, we have N independent optimization problems, and by [34, Proposition 1] the result follows. For inter-subchannel cooperative encoding, the sum rate (4.12) is maximized if each argument of the minimum operation is maximized. The first argument of (4.12) is insensitive to the choice of $p_{k0}^{(i)*}(\mathbf{s})$ or $p_{kj}^{(i)*}(\mathbf{s})$, as long as their sum is fixed; whereas the second argument is maximized if we separately maximize its summands for each i. The result follows by noting that this is also equivalent to N independent optimization problems, each yielding a scalar case, and [34, Proposition 1] holds, giving the desired result. A more detailed proof is given in Appendix B. □

An important observation is that, setting two of the powers equal to zero as suggested by Theorem 4.1, is also optimal for the entire rate region maximization, as the right hand sides of all three constraints, for both policies, are maximized by choosing the powers according to Theorem 4.1.[1] Therefore, from now on we focus only on policies that satisfy Theorem 4.1.

Note that, the bounds (4.10), (4.11) and (4.12) on R_1, R_2 and $R_1 + R_2$ respectively for inter-subchannel cooperative encoding are looser than the corresponding bounds (4.7), (4.8) and (4.9) for intra-subchannel cooperative encoding, as the minimum operations in (4.7), (4.8) are removed, and the minimum in (4.9) is taken outside the summation, to obtain (4.10), (4.11) and (4.12). As a result, the achievable rate region of inter-subchannel cooperative encoding contains that of intra-subchannel cooperative encoding. Hence, it is sufficient to limit our focus on the inter-subchannel cooperative encoding strategy, which results in a uniformly better rate region. Then, it is easy to check that the rate constraints in (4.10)-(4.12) now become concave in the power vector $\mathbf{p}(\mathbf{s}) = [p_{10}^{(i)*}(\mathbf{s}), p_{12}^{(i)*}(\mathbf{s}), p_{U_1}^{(i)*}(\mathbf{s}), p_{20}^{(i)*}(\mathbf{s}), p_{21}^{(i)*}(\mathbf{s}), p_{U_2}^{(i)*}(\mathbf{s}),\ i = 1, \ldots, N]$, lending themselves to well known techniques in convex optimization, which we discuss in the next sections.

4.4.1 Achievable Rate Maximization Using Projected Subgradient

Since all bounds of the achievable rate region are concave in powers, so is any weighted sum $\mu_1 R_1 + \mu_2 R_2$ at the corners. Moreover, it is easy to show that the rate region is strictly convex [6, 34]. Therefore, we can obtain points on the rate region boundary by maximizing $R_\mu = \mu_1 R_1 + \mu_2 R_2$, where $\{R_1, R_2\}$ is the corner of the pentagon obtained for a given power allocation policy, defined by (4.10)-(4.12). Assuming $\mu_1 > \mu_2$ without loss of generality, and employing Theorem 4.1 to simplify (4.10)-(4.12), the optimization problem can be stated as:

[1] We choose the first option for $\mathbf{s} \in \mathcal{S}_4$, which may cause a slight deviation from optimality for the sum rate. However, this case rarely occurs in practice, and this suboptimality can be ignored, as it has been done in [34].

$$
\begin{aligned}
\max_{\mathbf{p}(\mathbf{s})} \Bigg(& (\mu_1-\mu_2)\sum_{i=1}^{N}\left(E_{\mathcal{S}_1,\mathcal{S}_2}\left[C\left(s_{12}^{(i)}p_{12}^{(i)}(\mathbf{s})\right)\right]+E_{\mathcal{S}_3,\mathcal{S}_4}\left[C\left(s_{10}^{(i)}p_{10}^{(i)}(\mathbf{s})\right)\right]\right) \\
& +\mu_2\min\left\{\sum_{i=1}^{N}E\left[C\left(s_{10}^{(i)}p_{1}^{(i)}(\mathbf{s})+s_{20}^{(i)}p_{2}^{(i)}(\mathbf{s})+2\sqrt{s_{10}^{(i)}s_{20}^{(i)}p_{U_1}^{(i)}(\mathbf{s})p_{U_2}^{(i)}(\mathbf{s})}\right)\right],\right. \\
& \sum_{i=1}^{N}E_{\mathcal{S}_1}\left[C\left(s_{12}^{(i)}p_{12}^{(i)}(\mathbf{s})\right)+C\left(s_{21}^{(i)}p_{21}^{(i)}(\mathbf{s})\right)\right] \\
& +\sum_{i=1}^{N}E_{\mathcal{S}_2}\left[C\left(s_{12}^{(i)}p_{12}^{(i)}(\mathbf{s})\right)+C\left(s_{20}^{(i)}p_{20}^{(i)}(\mathbf{s})\right)\right] \\
& +\sum_{i=1}^{N}E_{\mathcal{S}_3}\left[C\left(s_{10}^{(i)}p_{10}^{(i)}(\mathbf{s})\right)+C\left(s_{21}^{(i)}p_{21}^{(i)}(\mathbf{s})\right)\right] \\
& \left.\left.+\sum_{i=1}^{N}E_{\mathcal{S}_4}\left[C\left(s_{10}^{(i)}p_{10}^{(i)}(\mathbf{s})+s_{20}^{(i)}p_{20}^{(i)}(\mathbf{s})\right)\right]\right\}\right), \qquad (4.13) \\
\text{s.t. } & \sum_{i=1}^{N}E\left[p_{k0}^{(i)}(\mathbf{s})+p_{kj}^{(i)}(\mathbf{s})+p_{U_k}^{(i)}(\mathbf{s})\right]\le\bar{p}_k, \\
& p_{k0}^{(i)}(\mathbf{s}),p_{kj}^{(i)}(\mathbf{s}),p_{U_k}^{(i)}(\mathbf{s})\ge 0,\ k,j\in\{1,2\},\ k\neq j,
\end{aligned}
$$

where $E_{\mathcal{S}_d}$ denotes the expectation over $\mathbf{s}\in\mathcal{S}_d$, $d=1,2,3,4$.

Due to the minimum operation in (4.13), the gradient of the objective function does not exist everywhere. In particular, there are two gradient vectors, depending on which argument of the minimum in (4.13) is active. Yet, these vectors may be viewed instead as subgradients, which makes it possible to employ the method of projected subgradients, for power optimization. Due to the convex nature of our constraints, this method is guaranteed to converge to the global optimum [63], with a diminishing stepsize normalized by the norm of the subgradient.

Since the calculation of the subgradients requires rather tedious formulas which give little insight, we will instead directly provide some examples of the achievable rate region, and the resulting power allocation policy, based on simulations in Section 4.5 instead. The major drawbacks of the subgradient algorithm are its slow rate of convergence, and complexity. As the number of subchannels increase, so does the size of the vector of power variables, making the process of computing the subgradients, and the projection operations formidable. Hence, we next

obtain the analytical expressions for weighted sum-rate optimal power control, and propose an alternative iterative implementation which converges much faster than the subgradient algorithm.

4.4.2 Iterative Achievable Rate Maximization Based on KKT Conditions

The problem (4.13), can be stated in an equivalent differentiable form:

$$\max_{\mathbf{p}(\mathbf{s})} \; R_\mu$$

$$\begin{aligned} s.t.\ R_\mu \leq\ & (\mu_1-\mu_2)\sum_{i=1}^{N}\left(E_{S_1,S_2}\left[C\left(s_{12}^{(i)}p_{12}^{(i)}(\mathbf{s})\right)\right]+E_{S_3,S_4}\left[C\left(s_{10}^{(i)}p_{10}^{(i)}(\mathbf{s})\right)\right]\right) \\ & +\mu_2\sum_{i=1}^{N}E\left[C\left(s_{10}^{(i)}p_1^{(i)}(\mathbf{s})+s_{20}^{(i)}p_2^{(i)}(\mathbf{s})+2\sqrt{s_{10}^{(i)}s_{20}^{(i)}p_{U_1}^{(i)}(\mathbf{s})p_{U_2}^{(i)}(\mathbf{s})}\right)\right], \end{aligned} \tag{4.14}$$

$$\begin{aligned} R_\mu \leq\ & (\mu_1-\mu_2)\sum_{i=1}^{N}\left(E_{S_1,S_2}\left[C\left(s_{12}^{(i)}p_{12}^{(i)}(\mathbf{s})\right)\right]+E_{S_3,S_4}\left[C\left(s_{10}^{(i)}p_{10}^{(i)}(\mathbf{s})\right)\right]\right) \\ & +\mu_2\sum_{i=1}^{N}\Big(E_{S_1}\left[C\left(s_{12}^{(i)}p_{12}^{(i)}(\mathbf{s})\right)+C\left(s_{21}^{(i)}p_{21}^{(i)}(\mathbf{s})\right)\right] \\ & +E_{S_2}\left[C\left(s_{12}^{(i)}p_{12}^{(i)}(\mathbf{s})\right)+C\left(s_{20}^{(i)}p_{20}^{(i)}(\mathbf{s})\right)\right] \\ & +E_{S_3}\left[C\left(s_{10}^{(i)}p_{10}^{(i)}(\mathbf{s})\right)+C\left(s_{21}^{(i)}p_{21}^{(i)}(\mathbf{s})\right)\right] \\ & +E_{S_4}\left[C\left(s_{10}^{(i)}p_{10}^{(i)}(\mathbf{s})+s_{20}^{(i)}p_{20}^{(i)}(\mathbf{s})\right)\right]\Big), \end{aligned} \tag{4.15}$$

$$\sum_{i=1}^{N}\left(E_{S_3,S_4}\left[p_{10}^{(i)}(\mathbf{s})\right]+E_{S_1,S_2}\left[p_{12}^{(i)}(\mathbf{s})\right]+E\left[p_{U_1}^{(i)}(\mathbf{s})\right]\right)\leq\bar{p}_1, \tag{4.16}$$

$$\sum_{i=1}^{N}\left(E_{S_2,S_4}\left[p_{20}^{(i)}(\mathbf{s})\right]+E_{S_1,S_3}\left[p_{21}^{(i)}(\mathbf{s})\right]+E\left[p_{U_2}^{(i)}(\mathbf{s})\right]\right)\leq\bar{p}_2, \tag{4.17}$$

$$p_{10}^{(i)}(\mathbf{s}),p_{12}^{(i)}(\mathbf{s}),p_{U_1}^{(i)}(\mathbf{s}),p_{20}^{(i)}(\mathbf{s}),p_{21}^{(i)}(\mathbf{s}),p_{U_2}^{(i)}(\mathbf{s})\geq 0,\quad \forall\mathbf{s}. \tag{4.18}$$

Note that (4.14)-(4.18) is a convex optimization problem, with differentiable constraints, and hence the KKT conditions are necessary and sufficient for optimality.

Assigning the Lagrange multipliers $\gamma_1, \gamma_2, \lambda_1, \lambda_2$ to the constraints (4.14)-(4.17), and $\epsilon_t^{(i)}(\mathbf{s})$, $t = 1, ..., 6$, to the positivity constraints (4.18), we obtain the conditions for optimality, given in the following lemma.

Lemma 4.2. *Define the variable $A^{(i)}$, $i = 1, \cdots, N$; and the indices m, n as follows:*

$$A^{(i)} = 1 + s_{10}^{(i)} p_1^{(i)}(\mathbf{s}) + s_{20}^{(i)} p_2^{(i)}(\mathbf{s}) + 2\sqrt{s_{10}^{(i)} s_{20}^{(i)} p_{U_1}^{(i)}(\mathbf{s}) p_{U_2}^{(i)}(\mathbf{s})}, \tag{4.19}$$

$$m = \begin{cases} 0, & \text{if } \mathbf{s} \in \mathcal{S}_3 \cup \mathcal{S}_4 \\ 2, & \text{if } \mathbf{s} \in \mathcal{S}_1 \cup \mathcal{S}_2 \end{cases}, \quad n = \begin{cases} 0, & \text{if } \mathbf{s} \in \mathcal{S}_2 \cup \mathcal{S}_4 \\ 1, & \text{if } \mathbf{s} \in \mathcal{S}_1 \cup \mathcal{S}_3 \end{cases}. \tag{4.20}$$

A power allocation policy $p_{1m}^{(i)}(\mathbf{s})$, $p_{2n}^{(i)}(\mathbf{s})$, $p_{U_1}^{(i)}(\mathbf{s})$, $p_{U_2}^{(i)}(\mathbf{s})$ is optimal for the problem (4.14)-(4.18), if and only if it satisfies, for $\mathbf{s} \in \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3 \triangleq \mathcal{S}_4^c$,

$$\lambda_1 \geq (\mu_1 - \mu_2 + \gamma_1\mu_2) \frac{s_{1m}^{(i)}}{1 + s_{1m}^{(i)} p_{1m}^{(i)}(\mathbf{s})} + \gamma_2\mu_2 \frac{s_{10}^{(i)}}{A^{(i)}}, \tag{4.21}$$

$$\lambda_2 \geq \gamma_1\mu_2 \frac{s_{2n}^{(i)}}{1 + s_{2n}^{(i)} p_{2n}^{(i)}(\mathbf{s})} + \gamma_2\mu_2 \frac{s_{20}^{(i)}}{A^{(i)}}, \tag{4.22}$$

$$\lambda_k \geq \gamma_2\mu_2 \frac{\sqrt{s_{k0}^{(i)} s_{j0}^{(i)} p_{U_j}^{(i)}(\mathbf{s})} + s_{k0}^{(i)} \sqrt{p_{U_k}^{(i)}(\mathbf{s})}}{A^{(i)} \sqrt{p_{U_k}^{(i)}(\mathbf{s})}}, \quad k \in \{1, 2\}, \tag{4.23}$$

and for $\mathbf{s} \in \mathcal{S}_4$,

$$\lambda_1 \geq (\mu_1 - \mu_2) \frac{s_{10}^{(i)}}{1 + s_{10}^{(i)} p_{1m}^{(i)}(\mathbf{s})} + \gamma_1\mu_2 \frac{s_{10}^{(i)}}{1 + s_{10}^{(i)} p_{1m}^{(i)}(\mathbf{s}) + s_{20}^{(i)} p_{2n}^{(i)}(\mathbf{s})} + \gamma_2\mu_2 \frac{s_{10}^{(i)}}{A^{(i)}}, \tag{4.24}$$

$$\lambda_2 \geq \gamma_1\mu_2 \frac{s_{20}^{(i)}}{1 + s_{20}^{(i)} p_{2n}^{(i)}(\mathbf{s}) + s_{10}^{(i)} p_{1m}^{(i)}(\mathbf{s})} + \gamma_2\mu_2 \frac{s_{20}^{(i)}}{A^{(i)}}, \tag{4.25}$$

$$\lambda_k \geq \gamma_2\mu_2 \frac{\sqrt{s_{k0}^{(i)} s_{j0}^{(i)} p_{U_j}^{(i)}(\mathbf{s})} + s_{k0}^{(i)} \sqrt{p_{U_k}^{(i)}(\mathbf{s})}}{A^{(i)} \sqrt{p_{U_k}^{(i)}(\mathbf{s})}}, \quad k,\ j \in \{1, 2\}, \tag{4.26}$$

where the Lagrange multipliers γ_1, $\gamma_2 = 1 - \gamma_1$, λ_1, and λ_2 are selected so that the constraints (4.14)-(4.17) are satisfied with equality. Each of the constraints (4.21), (4.22) and (4.23) when $\mathbf{s} \in \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3$ (correspondingly (4.24), (4.25)

and (4.26) when $\mathbf{s} \in \mathcal{S}_4$*) are satisfied with equality if and only if the respective power levels,* $p_{1m}^{(i)}(\mathbf{s})$*,* $p_{2n}^{(i)}(\mathbf{s})$ *or* $p_{U_k}^{(i)}(\mathbf{s})$ *are positive.*

Proof. See Appendix C. □

The optimality conditions given in Lemma 4.2 for each power component are heavily coupled, thereby making the computation of the optimal power allocation policy seemingly difficult. Yet, in the following theorem, we show that, after some non-trivial observations, the coupling among the constraints is partially removed, and as a result, we are able to provide closed form expressions for the optimal power levels.

Theorem 4.3. *For a cooperative OFDMA system employing inter-subchannel cooperative encoding, the optimal power allocation,* $p_{1m}^{(i)}(\mathbf{s})$*,* $p_{2n}^{(i)}(\mathbf{s})$*,* $p_{U_1}^{(i)}(\mathbf{s})$*,* $p_{U_2}^{(i)}(\mathbf{s})$*, that solves (4.14)-(4.18) is given by*

$$p_{U_k}^{(i)}(\mathbf{s}) = s_{k0}^{(i)} \frac{\frac{\mu_2(1-\gamma_1)}{\lambda_k}\left(s_{k0}^{(i)} + \frac{\lambda_k}{\lambda_j} s_{j0}^{(i)}\right) - \left(1 + s_{10}^{(i)} p_{1m}^{(i)} + s_{20}^{(i)} p_{2n}^{(i)}\right)}{\left(s_{k0}^{(i)} + \frac{\lambda_k}{\lambda_j} s_{j0}^{(i)}\right)^2}, \tag{4.27}$$

$$p_{1m}^{(i)}(\mathbf{s}) = \begin{cases} \left(\frac{(\mu_1-\mu_2+\gamma_1\mu_2)\left(\lambda_2 s_{10}^{(i)}+\lambda_1 s_{20}^{(i)}\right)}{\lambda_1^2 s_{20}^{(i)}} - \frac{1}{s_{1m}^{(i)}}\right)^+, & \text{if } \mathbf{s} \in \mathcal{S}_4^c \quad (4.28a) \\ f\Bigg(s_{10}^{(i)^2}, \frac{(\lambda_1 s_{20}^{(i)}+\lambda_2 s_{10}^{(i)})(\mu_1-\mu_2+\gamma_1\mu_2) s_{10}^{(i)^2} - {\lambda_1}^2 s_{20}^{(i)}(2 s_{10}^{(i)} + s_{10}^{(i)} s_{20}^{(i)} p_{2n}^{(i)}(\mathbf{s}))}{-{\lambda_1}^2 s_{20}^{(i)}}, \\ \frac{(\lambda_1 s_{20}^{(i)}+\lambda_2 s_{10}^{(i)})\left[(\mu_1-\mu_2+\gamma_1\mu_2)+(\mu_1-\mu_2) s_{20}^{(i)} p_{2n}^{(i)}(\mathbf{s})\right] s_{10}^{(i)} - {\lambda_1}^2 s_{20}^{(i)}(1+s_{20}^{(i)} p_{2n}^{(i)}(\mathbf{s}))}{-{\lambda_1}^2 s_{20}^{(i)}}\Bigg), & o.w. \quad (4.28b) \end{cases}$$

$$p_{2n}^{(i)}(\mathbf{s}) = \begin{cases} \left(\frac{\gamma_1\mu_2\left(\lambda_2 s_{10}^{(i)}+\lambda_1 s_{20}^{(i)}\right)}{\lambda_2^2 s_{10}^{(i)}} - \frac{1}{s_{2n}^{(i)}}\right)^+, & \text{if } \mathbf{s} \in \mathcal{S}_4^c \quad (4.29a) \\ \left(\frac{\gamma_1\mu_2\left(\lambda_2 s_{10}^{(i)}+\lambda_1 s_{20}^{(i)}\right)}{\lambda_2^2 s_{10}^{(i)}} - \frac{1}{s_{20}^{(i)}} - \frac{s_{10}^{(i)}}{s_{20}^{(i)}} p_{1m}^{(i)}(\mathbf{s})\right)^+, & \text{if } \mathbf{s} \in \mathcal{S}_4 \quad (4.29b) \end{cases}$$

if the powers obtained from (4.27) are positive, i.e., $p_{U_k}^{(i)}(\mathbf{s}) > 0$*; and*

$$p_{U_k}^{(i)}(\mathbf{s}) = 0, \tag{4.30}$$

$$p_{1m}^{(i)}(\mathbf{s}) = \begin{cases} f\Big(\lambda_1 s_{10}^{(i)} s_{1m}^{(i)}, -\mu_1 s_{10}^{(i)} s_{1m}^{(i)} + \lambda_1 (s_{10}^{(i)} + s_{1m}^{(i)} + s_{1m}^{(i)} s_{20}^{(i)} p_{2n}^{(i)}(\mathbf{s})), \lambda_1 (1 + s_{20}^{(i)} p_{2n}^{(i)}(\mathbf{s})) \\ -(\mu_1 - \mu_2 + \gamma_1 \mu_2)\, s_{1m}^{(i)} (1 + s_{20}^{(i)} p_{2n}^{(i)}(\mathbf{s})) - \mu_2 (1 - \gamma_1) s_{10}^{(i)} \Big), & \text{if } \mathbf{s} \in \mathcal{S}_4^c \quad (4.31a) \\ f\Big(\lambda_1 {s_{10}^{(i)}}^2, -\mu_1 {s_{10}^{(i)}}^2 + \lambda_1 s_{10}^{(i)} (2 + s_{20}^{(i)} p_{2n}^{(i)}(\mathbf{s})), \\ -\mu_1 s_{10}^{(i)} - (\mu_1 - \mu_2) s_{10}^{(i)} s_{20}^{(i)} p_{2n}^{(i)}(\mathbf{s}) + \lambda_1 (1 + s_{20}^{(i)} p_{2n}^{(i)}(\mathbf{s})) \Big), & \text{if } \mathbf{s} \in \mathcal{S}_4 \quad (4.31b) \end{cases}$$

$$p_{2n}^{(i)}(\mathbf{s}) = \begin{cases} f\left(\lambda_2 s_{20}^{(i)} s_{2n}^{(i)}, -\mu_2 s_{20}^{(i)} s_{2n}^{(i)} + \lambda_2 (s_{20}^{(i)} + s_{2n}^{(i)} + s_{10}^{(i)} s_{2n}^{(i)} p_{1m}^{(i)}(\mathbf{s})), \lambda_2 (1 + s_{10}^{(i)} p_{1m}^{(i)}(\mathbf{s}))\right. \\ \left. -\gamma_1 \mu_2 s_{2n}^{(i)} (1 + s_{10}^{(i)} p_{1m}^{(i)}(\mathbf{s})) - \mu_2 (1 - \gamma_1) s_{20}^{(i)} \right), & \text{if } \mathbf{s} \in \mathcal{S}_4^c \quad (4.32a) \\ \left(\frac{\mu_2}{\lambda_2} - \frac{1}{s_{20}^{(i)}} - \frac{s_{10}^{(i)}}{s_{20}^{(i)}} p_{1m}^{(i)}(\mathbf{s}) \right)^+, & \text{if } \mathbf{s} \in \mathcal{S}_4 \quad (4.32b) \end{cases}$$

otherwise, where γ_1, λ_1 and λ_2 are selected to satisfy the constraints (4.14)-(4.17) with equality, and $f(\cdot)$ is defined in (4.4).

Proof. We start by noting that, to obtain coherent combining gain, the optimal cooperative powers $p_{U_k}^{(i)}(\mathbf{s})$, $k = 1, 2$, over a given subchannel and given channel state $\mathbf{s}$, should either be both positive, or both zero. Let us first assume that both $p_{U_1}^{(i)}(\mathbf{s})$ and $p_{U_2}^{(i)}(\mathbf{s})$ are positive. Then, the constraints (4.23), (equivalently (4.26)) should be satisfied with equality, for $k = 1, 2$. Evaluating (4.23), (equivalently (4.26)) separately for $k = 1, 2$, and dividing the two resulting equalities, we get

$$\frac{\sqrt{s_{20}^{(i)}}\sqrt{p_{U_2}^{(i)}(\mathbf{s})} + \sqrt{s_{10}^{(i)}}\sqrt{p_{U_1}^{(i)}(\mathbf{s})}}{\sqrt{s_{10}^{(i)}}\sqrt{p_{U_1}^{(i)}(\mathbf{s})} + \sqrt{s_{20}^{(i)}}\sqrt{p_{U_2}^{(i)}(\mathbf{s})}} \frac{\sqrt{s_{10}^{(i)}}}{\sqrt{s_{20}^{(i)}}} \frac{\sqrt{p_{U_2}^{(i)}(\mathbf{s})}}{\sqrt{p_{U_1}^{(i)}(\mathbf{s})}} = \frac{\lambda_1}{\lambda_2}, \tag{4.33}$$

which yields

$$p_{U_1}^{(i)}(\mathbf{s}) = \frac{\lambda_2^2 s_{10}^{(i)}}{\lambda_1^2 s_{20}^{(i)}} p_{U_2}^{(i)}(\mathbf{s}). \tag{4.34}$$

Plugging (4.34) into (4.23) (equivalently (4.26)), we achieve the following crucial equality

$$\frac{\gamma_2 \mu_2}{A^{(i)}} = \frac{\lambda_1 \lambda_2}{\lambda_1 s_{20}^{(i)} + \lambda_2 s_{10}^{(i)}}. \tag{4.35}$$

The significance of (4.35) is that, its left hand size, which involves all power components through $A^{(i)}$, and appears in all of (4.21)-(4.26), can be replaced by

a term which depends only on the fixed Lagrange multipliers, λ_1 and λ_2, and the direct link gains, $s_{k0}^{(i)}$. Therefore, the optimality constraints for $p_{1m}^{(i)}(\mathbf{s})$ and $p_{2n}^{(i)}(\mathbf{s})$ can be rewritten independently of $p_{U_k}^{(i)}(\mathbf{s})$. For example, using (4.35) in (4.21), we get

$$(\mu_1 - \mu_2 + \gamma_1\mu_2)\,\frac{s_{1m}^{(i)}}{1 + s_{1m}^{(i)} p_{1m}^{(i)}(\mathbf{s})} \leq \frac{{\lambda_1}^2 s_{20}^{(i)}}{\lambda_1 s_{20}^{(i)} + \lambda_2 s_{10}^{(i)}}, \tag{4.36}$$

which yields the waterfilling solution, (4.28a). Similarly, using (4.35) in (4.22), (4.24) and (4.25), we obtain (4.29a), (4.28b) and (4.29b) respectively. The expression, (4.27), of optimal $p_{U_1}^{(i)}(\mathbf{s})$ follows from (4.23), (4.34) and (4.35).

Note however that, $p_{U_k}^{(i)}(\mathbf{s})$ obtained by (4.27) is not guaranteed to be positive. In case it is not, this means that (4.23) (equivalently (4.26)) is satisfied with strict inequality, the optimal solution for $p_{U_k}^{(i)}(\mathbf{s})$ should be set to 0 and (4.35) can no longer be used. Then, when $p_{U_k}^{(i)}(\mathbf{s}) = 0$, instead of (4.21)-(4.22) and (4.24)-(4.25) we have to apply the conditions:

$$(\mu_1 - \mu_2 + \gamma_1\mu_2)\,\frac{s_{1m}^{(i)}}{1 + s_{1m}^{(i)} p_{1m}^{(i)}(\mathbf{s})} + \gamma_2\mu_2 \frac{s_{10}^{(i)}}{1 + s_{10}^{(i)} p_{1m}^{(i)}(\mathbf{s}) + s_{20}^{(i)} p_{2n}^{(i)}(\mathbf{s})} \leq \lambda_1, \tag{4.37}$$

$$\gamma_1\mu_2 \frac{s_{2n}^{(i)}}{1 + s_{2n}^{(i)} p_{2n}^{(i)}(\mathbf{s})} + \gamma_2\mu_2 \frac{s_{20}^{(i)}}{1 + s_{10}^{(i)} p_{1m}^{(i)}(\mathbf{s}) + s_{20}^{(i)} p_{2n}^{(i)}(\mathbf{s})} \leq \lambda_2, \tag{4.38}$$

for $\mathbf{s} \in \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3$, and

$$(\mu_1 - \mu_2)\,\frac{s_{10}^{(i)}}{1 + s_{10}^{(i)} p_{1m}^{(i)}(\mathbf{s})} + \mu_2 \frac{s_{10}^{(i)}}{1 + s_{10}^{(i)} p_{1m}^{(i)}(\mathbf{s}) + s_{20}^{(i)} p_{2n}^{(i)}(\mathbf{s})} \leq \lambda_1, \tag{4.39}$$

$$\mu_2 \frac{s_{20}^{(i)}}{1 + s_{20}^{(i)} p_{2n}^{(i)}(\mathbf{s}) + s_{10}^{(i)} p_{1m}^{(i)}(\mathbf{s})} \leq \lambda_2, \tag{4.40}$$

for $\mathbf{s} \in \mathcal{S}_4$.

When $p_{U_k}^{(i)}(\mathbf{s}) = 0$, $k = 1, 2$; the powers $p_{1m}^{(i)}(\mathbf{s})$ and $p_{2n}^{(i)}(\mathbf{s})$ are automatically independent of $p_{U_k}^{(i)}(\mathbf{s})$. However, (4.37) and (4.38); (4.39) and (4.40) are coupled, and each should be solved by finding the positive roots of a quadratic equation. Since all power values are non-negative, i.e., $p_{1m}^{(i)}(\mathbf{s}) \geq 0$ and $p_{2n}^{(i)}(\mathbf{s}) \geq 0$, we can

achieve $p_{1m}^{(i)}(\mathbf{s})$ in (4.31a), $p_{2n}^{(i)}(\mathbf{s})$ in (4.32a) by solving (4.37) and (4.38). Similarly, $p_{1m}^{(i)}(\mathbf{s})$ in (4.31b) and $p_{2n}^{(i)}(\mathbf{s})$ in (4.32b) can be obtained using (4.39) and (4.40). γ_1, λ_1 and λ_2 are selected in such a way that, when the power levels in (4.27)-(4.32b) are used, the constraints (4.14)-(4.17) are satisfied. □

The power levels of the cooperative codewords on each subchannel, $p_{1m}^{(i)}(\mathbf{s})$ and $p_{2n}^{(i)}(\mathbf{s})$ in (4.28a) and (4.29a), have an interesting single user waterfilling type interpretation, as they solely depend on the channel gains of only that particular subchannel, and the fixed Lagrange multipliers. The water level is determined by the direct link gains. However, in (4.31a) and (4.32a) the power $p_{1m}^{(i)}(\mathbf{s})$ depends on $p_{2n}^{(i)}(\mathbf{s})$, and vice-versa: increasing one of the powers will decrease the other, should the constraints (4.37)-(4.40) be satisfied with equality, and we now have a multi-user waterfilling type solution. This is somewhat different than the observations in [34], which conjectured that a single user waterfilling type solution for cooperative powers would be sufficient in all scenarios, for the much simpler case of the scalar MAC, and sum rate maximization only.

At this point, it should be clear that although (4.28a)-(4.29b) and (4.31a)-(4.32b) do not explicitly depend on $p_{U_k}^{(i)}(\mathbf{s})$, the decision regarding which of these equations should be used while computing $p_{kj}^{(i)}(\mathbf{s})$ does. Likewise, $p_{U_k}^{(i)}(\mathbf{s})$ are clearly functions of $p_{kj}^{(i)}(\mathbf{s})$, which makes equations (4.28a)-(4.29b), (4.31a)-(4.32b) and (4.37)-(4.40) coupled. Note however that, the way we proved Theorem 1 automatically suggests a natural way of solving the KKT conditions iteratively. To this end, we propose an algorithm which performs updates on the powers of the users, one-user-at-a-time: given $p_{U_1}^{(i)}(\mathbf{s})$ and $p_{12}^{(i)}(\mathbf{s})$, it computes $p_{U_2}^{(i)}(\mathbf{s})$ and $p_{21}^{(i)}(\mathbf{s})$, and using these new values for user 2, it re-iterates the powers of user 1. The outline of the algorithm is given below:

Algorithm 1 Iterative Power Allocation Algorithm

for $\mu_2 = 0 : 1$ **do**
  while (4.14)-(4.15) are not satisfied **do**
    while (4.16) is not satisfied **do**
      Calculate $p_{1m}^{(i)}(\mathbf{s})$ using (4.28a)-(4.28b) and $p_{U_1}^{(i)}(\mathbf{s})$ using (4.27) assuming $p_{U_1}^{(i)}(\mathbf{s}) > 0$, $\forall i$
      while $\exists\ \mathbf{s}'$ s.t. $p_{U_1}^{(i)}(\mathbf{s}') < 0$ **do**
        Set $p_{U_1}^{(i)}(\mathbf{s}') = 0$ and re-calculate $p_{1m}^{(i)}(\mathbf{s}')$ using (4.31a)-(4.31b) and $p_{U_1}^{(i)}(\mathbf{s}')$ using (4.27)
      end while
      Update λ_1
    end while
    while (4.17) is not satisfied **do**
      Calculate $p_{2n}^{(i)}(\mathbf{s})$ using (4.29a)-(4.29b) and $p_{U_2}^{(i)}(\mathbf{s})$ using (4.27) assuming $p_{U_2}^{(i)}(\mathbf{s}) > 0$, $\forall i$
      while $\exists\ \mathbf{s}'$ s.t. $p_{U_2}^{(i)}(\mathbf{s}') < 0$ **do**
        Set $p_{U_2}^{(i)}(\mathbf{s}') = 0$ and re-calculate $p_{2n}^{(i)}(\mathbf{s}')$ using (4.32a)-(4.32b) and $p_{U_2}^{(i)}(\mathbf{s}')$ using (4.27)
      end while
      Update λ_2
    end while
    Update γ_1
  end while
end for

This algorithm simplifies the seemingly difficult task of obtaining the optimal powers from the coupled equations, and due to the convex nature of the problem, and the Cartesian nature of the constraints across users, it provably converges to the optimal solution, as at the end of the iterations, the KKT conditions will be satisfied.

Perhaps the most important feature of this algorithm is that, *regardless of the number of subchannels used*, we only need to solve for three Lagrange multipliers, which relate the powers allocated to the subchannels, to obtain the optimum power allocation. This reduces the complexity of the algorithm dramatically, and makes it scalable, compared to the subgradient algorithm. As a result, the convergence is much faster.

4.5 Simulation Results

In order to obtain the optimal power allocation policy, and the resulting achievable rate region, we implement the projected subgradient algorithm, and the iterative waterfilling-like algorithm based on KKT conditions on optimality, for a simple case with only three subchannels. The achievable rate region for the inter-subchannel cooperative encoding strategy is obtained by running this algorithm for varying priorities μ_k, and then by taking a convex hull over the resulting power optimized regions. In Figure 4.1, we compare the achievable rate region for power controlled cooperative OFDMA utilizing the projected subgradient algorithm and the iterative algorithm, with those for several enconding strategies without power control, from [57]. We assume that, for the channel non-adaptive protocols, the users are still able to allocate their total power across subchannels and codewords. The total power of each user and the noise variances are set to unity. The fading coefficients are chosen from independent Rayleigh distributions, the means of which are shown in Figure 4.1. We observe that, when the powers are chosen jointly optimally with inter-subchannel cooperative encoding, there is a major improvement in achievable rates. This unusually high gain from power control can be attributed to our ability to take advantage of the additional diversity created by OFDMA: power allocation not only allows us to use the subchannels at time varying instantaneous rates based on the channel qualities, but also to use them adaptively for varying purposes, i.e., cooperation, common message generation or direct transmission. The gain achieved by power control through the iterative algorithm always exceeds the projected subgradient algorithm, especially in the sum rate region. The main reason is that, the subgradient algorithm had still not fully converged, when it was stopped at 10000 iterations, while the iterative algorithm did fully converge to the optimal power allocation, and in a much shorter time.

In Figure 4.2, we compare the rate regions in a uniform fading environment with means expressed on the figure. Here we ensure $\mathbf{s} \in \mathcal{S}_1$, with the motivation of

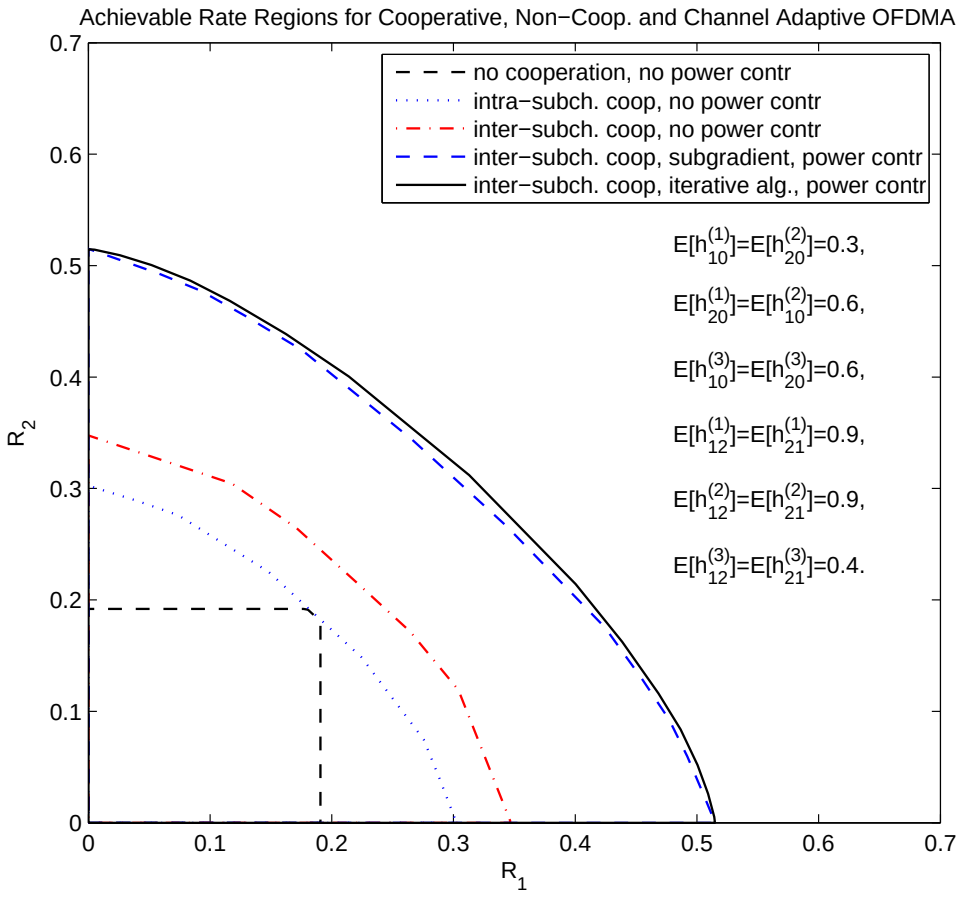


Figure 4.1: Achievable rate regions in Rayleigh fading.

obtaining a strictly optimal power allocation, and a simpler description of the power distributions. In this setting, since some of the power values are always zero, the number of power variables is less, and hence the subgradient algorithm nearly converges to the optimum within 10000 iterations, and the rate regions of subgradient and iterative algorithms nearly coincide. For this setting, we further analyze the optimal power distributions over the channel states, in Figures 4.3(a)-4.3(c), 4.4(a)-4.4(c) and 4.5(a)-4.5(c).

Figures 4.3(a)-4.3(c) and 4.4(a)-4.4(c) demonstrate the optimal powers allocated to subchannel 1, as functions of the inter-user link gains, when the direct link gains are fixed to two different sets specified on the figures. Powers $p_{U_2}^{(1)}$ are not shown, to save space, as they are identical to $p_{U_1}^{(1)}$ due to the symmetry in fading. In Figures 4.3(a)-4.3(c), the direct link gains are at their maximum, hence

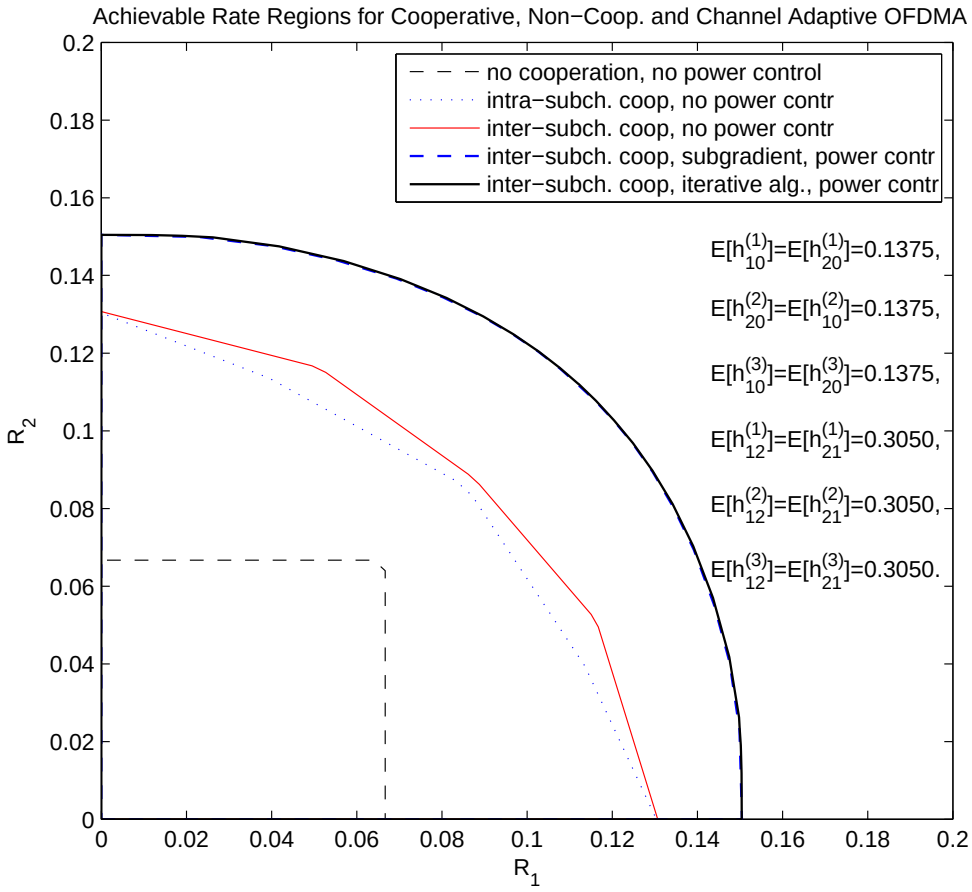
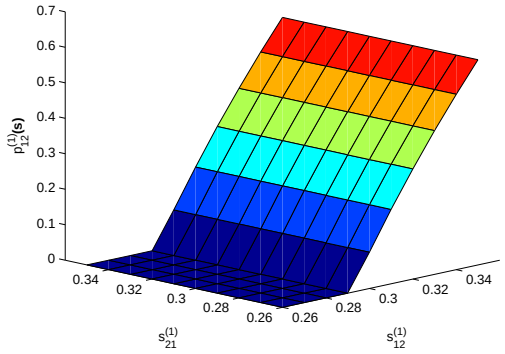
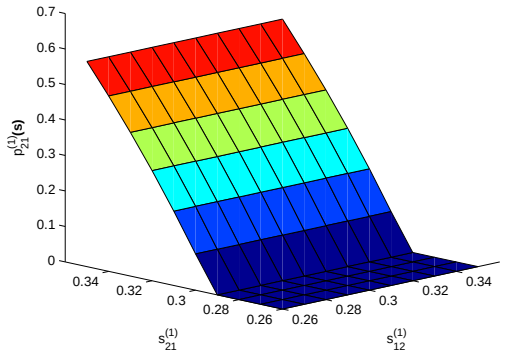


Figure 4.2: Achievable rate regions in uniform fading.

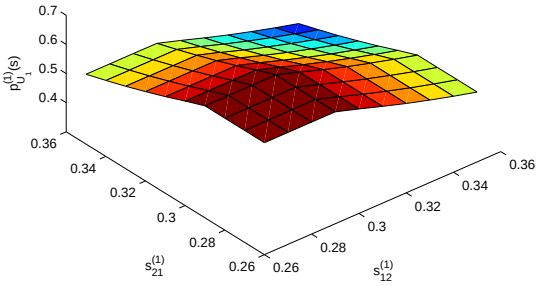
the cooperative powers, $p_{U_k}^{(1)}$, are always positive. In this case, we observe the expected single user waterfilling type behavior for the distributions of $p_{12}^{(1)}(\mathbf{s})$ and $p_{21}^{(1)}(\mathbf{s})$. In Figures 4.4(a)-4.4(c) however, when the direct links are moderate on the average, we have a more interesting scenario: when $s_{21}^{(1)}$ is significantly stronger instantaneously, only user 2 uses the subchannel. When both inter-user links are instantaneously strong, the users exchange information using simultaneous waterfilling, and set $p_{U_k}^{(1)}$ to zero. When both inter-user links are weak, the users use the subchannel solely to convey common information to the RX, by using only $p_{U_1}^{(1)}$ and $p_{U_2}^{(1)}$. An important observation is that, although we make no prior assumptions on subchannel allocation to users/codewords, the optimal powers sometimes dictate exclusive use of some subchannels for dedicated tasks. The resulting power distributions show that the KKT conditions are indeed satisfied at the fixed point of our iterative algorithm, verifying convergence.



(a) Power level, $p_{12}^{(1)}$

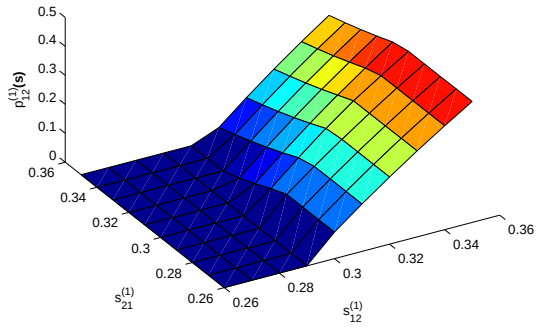


(b) Power level, $p_{21}^{(1)}$

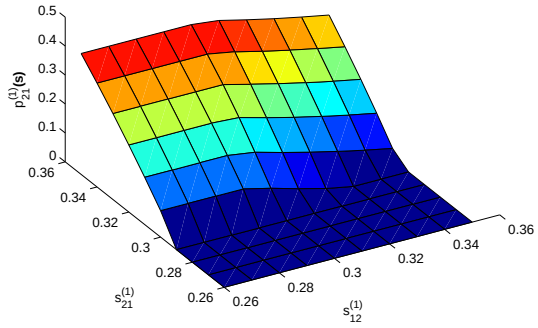


(c) Power level, $p_{U_1}^{(1)} = p_{U_2}^{(1)}$

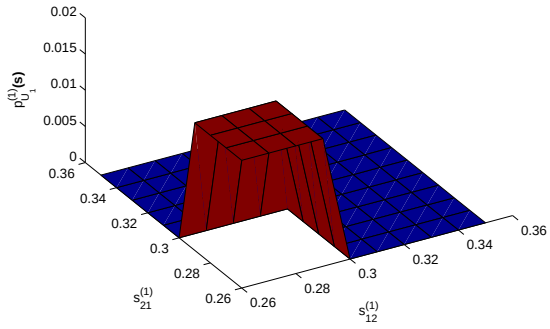
Figure 4.3: Optimal power allocation when $s_{10}^{(1)}$ and $s_{20}^{(1)}$ are maximum (*i.e.*, $s_{10}^{(1)} = s_{20}^{(1)} = 0.25$), fixed and always less than $s_{12}^{(1)}$ and $s_{21}^{(1)}$. $p_{U_k}^{(1)}$ are always positive, to take advantage of strong direct links. $p_{kj}^{(1)}$ obey single user waterfilling, as expected.



(a) Power level, $p_{12}^{(1)}$

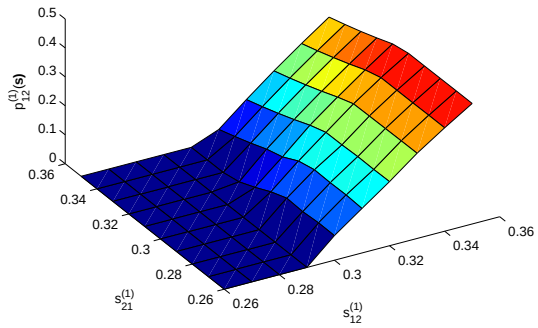


(b) Power level, $p_{21}^{(1)}$

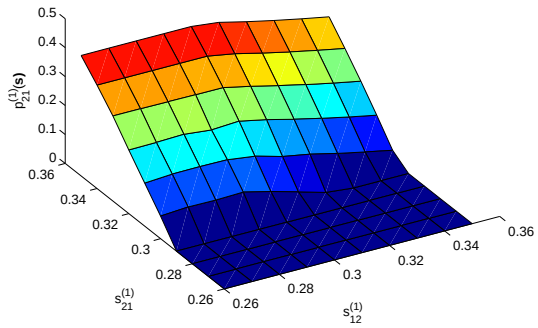


(c) Power level, $p_{U_1}^{(1)} = p_{U_2}^{(1)}$

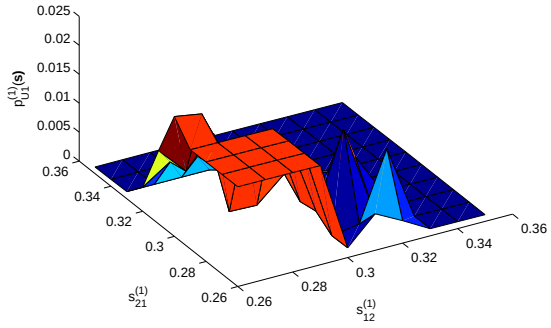
Figure 4.4: Optimal power allocation when $s_{10}^{(1)} = s_{20}^{(1)} = 0.15$, fixed and always less than $s_{12}^{(1)}$ and $s_{21}^{(1)}$. When $p_{U_k}^{(1)}$ is positive, $p_{kj}^{(1)}$ obey single user waterfilling. As the inter-user links get stronger, it becomes more profitable to create common information, $p_{U_k}^{(1)}$ become 0, and the users perform simultaneous waterfilling.



(a) Power level, $p_{12}^{(1)}$

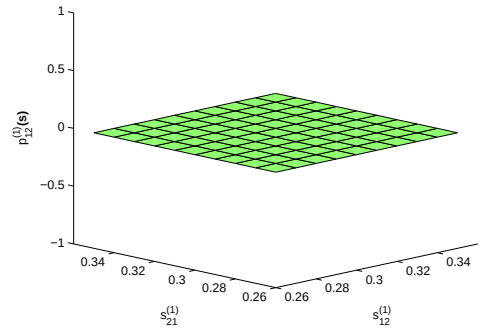


(b) Power level, $p_{21}^{(1)}$

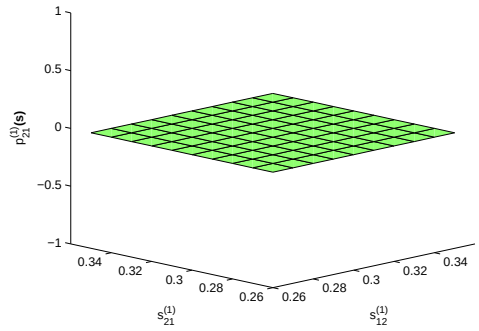


(c) Power level, $p_{U_1}^{(1)}$

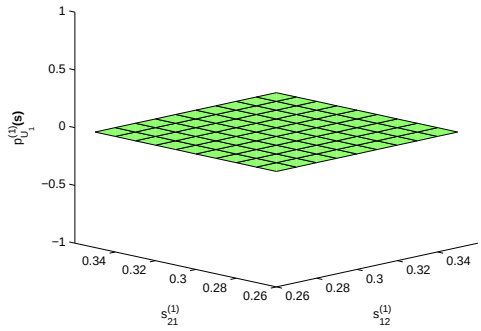
Figure 4.5: Power allocation obained after 10000 iterations of the subgradient algorithm, when $s_{10}^{(1)} = s_{20}^{(1)} = 0.15$, fixed and always less than $s_{12}^{(1)}$ and $s_{21}^{(1)}$. The algorithm has not yet converged to the optimum value, despite a much longer running time compared to the iterative algorithm. Achievable rates are nearly within 0.1% of the optimum value.



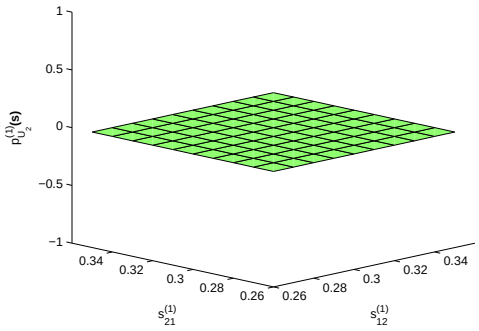
(a) Power levels, $p_{12}^{(1)}$



(b) Power levels, $p_{21}^{(1)}$



(c) Power levels, $p_{U_1}^{(1)} = p_{U_2}^{(1)}$



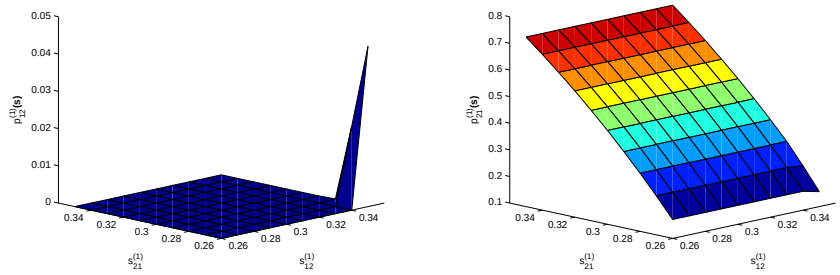
(d) Power levels, $p_{U_2}^{(1)}$

Figure 4.6: Results of power allocation when $s_{10}^{(1)}$ and $s_{20}^{(1)}$ are minimum (*i.e.*, $s_{10}^{(1)} = s_{20}^{(1)} = 0.025$), fixed and always less than $s_{12}^{(1)}$ and $s_{21}^{(1)}$.

In Figures 4.5(b)-4.5(c), we plot the power distributions obtained using the subgradient algorithm instead, for the same setting as in Figures 4.4(a)-4.4(c). The subgradient algorithm is terminated after 10000 iterations. It is observed that while the powers $p_{12}^{(1)}(\mathbf{s})$ and $p_{21}^{(1)}(\mathbf{s})$ seem to have nearly converged to the optimal values shown in Figures 4.4(a)-4.4(c) (only $p_{21}^{(1)}(\mathbf{s})$ is shown, as $p_{12}^{(1)}(\mathbf{s})$ is simply symmetrical), the cooperative power $p_{U_1}^{(1)}(\mathbf{s})$ has still not fully converged, though it is close to its optimal distribution. Note that, the effect of this is negligible on the rate regions, as was shown in Figure 4.2.

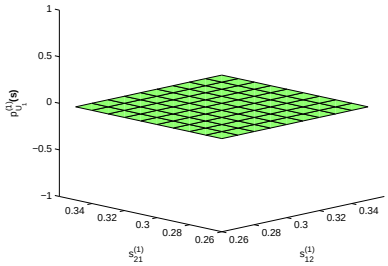
We observe that when both direct links are in deep fading, no power is allocated to any user for any purpose, as can be seen in Figures 4.6(a)-4.6(d).

Figures 4.7(a)-4.7(d) show an interesting result; although $s_{12}^{(1)}$ and $s_{21}^{(1)}$ have same

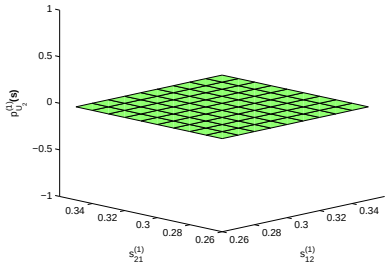


(a) Power level, $p_{12}^{(1)}$

(b) Power level, $p_{21}^{(1)}$



(c) Power level, $p_{U_1}^{(1)}$



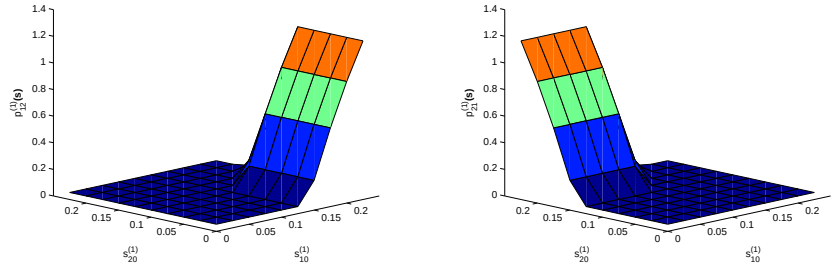
(d) Power level, $p_{U_2}^{(1)}$

Figure 4.7: Results of power allocation when $s_{10}^{(1)} = 0.125$, $s_{20}^{(1)} = 0.175$, fixed and always less than $s_{12}^{(1)}$ and $s_{21}^{(1)}$.

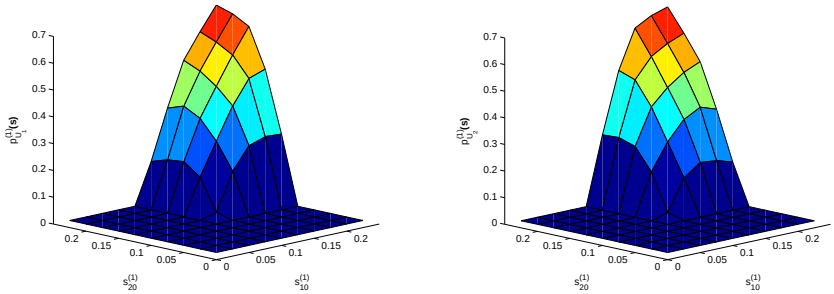
average channel states, more power is allocated to the $p_{21}^{(1)}$, since $s_{20}^{(1)} > s_{10}^{(1)}$ in this case. Only when $s_{12}^{(1)}$ is maximum, a little amount of power is reserved for $p_{12}^{(1)}$, other than that all power is used for $p_{21}^{(1)}$, like a single-user filling algorithm. In this case, no power is allocated to p_{U_k}.

For the case where the cooperative links have the least values, shown in Figures 4.8(a)-4.8(d), when the cooperative links are at their worst states, no power is allocated to any user, nor for cooperation, neither for direct transmission purpose. As both direct links' channel states get better and better, power is distributed to both users. But when only one user's direct link states achieve higher states, power distributed to this user via waterfilling algorithm.

At the situation where both cooperative links have the highest channel states, when both direct links have high channel states, power is distributed to both users for both purposes, cooperation and direct transmission. But when only one

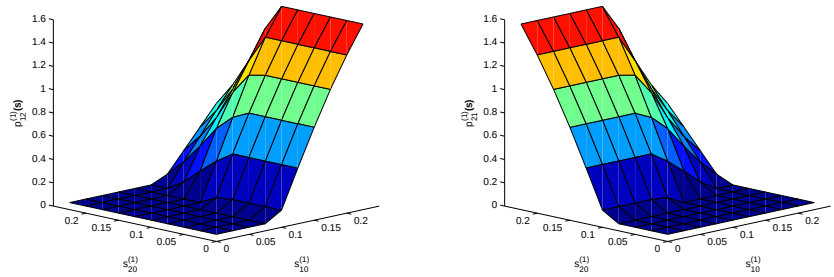


(a) Power level, $p_{12}^{(1)}$ (b) Power level, $p_{21}^{(1)}$

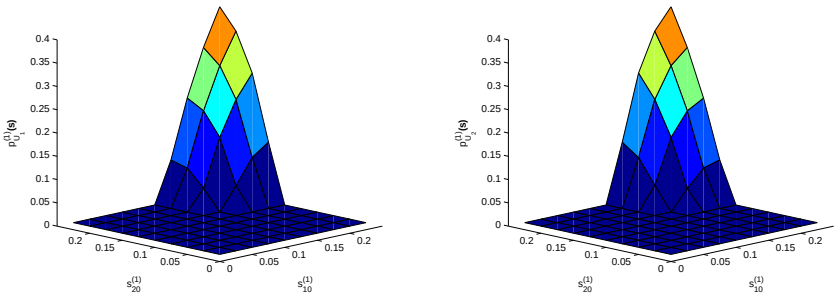


(c) Power level, $p_{U_1}^{(1)}$ (d) Power level, $p_{U_2}^{(1)}$

Figure 4.8: Results of power allocation when $s_{12}^{(1)}$ and $s_{21}^{(1)}$ are minimum (*i.e.*, $s_{12}^{(1)} = s_{21}^{(1)} = 0.26$), fixed and always more than $s_{10}^{(1)}$ and $s_{20}^{(1)}$.



(a) Power level, $p_{12}^{(1)}$ (b) Power level, $p_{21}^{(1)}$



(c) Power level, $p_{U_1}^{(1)}$ (d) Power level, $p_{U_2}^{(1)}$

Figure 4.9: Results of power allocation when $s_{12}^{(1)}$ and $s_{21}^{(1)}$ are maximum (*i.e.*, $s_{12}^{(1)} = s_{21}^{(1)} = 0.35$), fixed and always more than $s_{10}^{(1)}$ and $s_{20}^{(1)}$.

user has a good direct link, then power is allocated to that user through single user waterfilling, as shown in Figures 4.9(a)-4.9(d).

4.6 Conclusion

We obtained the optimum power allocation policies for a cooperative OFDMA channel employing intra-subchannel cooperative encoding and inter-subchannel cooperative encoding strategies. We developed a subgradient algorithm which converges to the optimum power allocation policies that maximize the achievable rate region, and a more efficient iterative algorithm which maximizes the achievable rate region. The number of iterations of the algorithm does not depend on the number of subchannels, which makes the algorithm scalable. We demonstrated that the optimal power allocation may also serve as a guideline for subchannel assignment to the users' cooperative codewords, and that power allocation for cooperative OFDMA provides significant rate improvements, due to its ability to exploit the diversity provided by OFDMA.

Chapter 5

Conclusion

5.1 Summary of the Results

We proposed and evaluated three encoding strategies for a two user cooperative OFDMA system, based on block Markov superposition encoding, by allowing pairs of users to share the subchannels. We utilized mutual cooperation in OFDMA channels. The inter-subchannel cooperative encoding strategy is a novel extension of block Markov encoding for MAC-GF. We obtained the expressions for the resulting achievable rate regions for all three encoding strategies. We show that, by allowing for re-partitioning and re-encoding of the cooperative messages across subchannels, it is possible to better exploit the diversity created by OFDMA, and higher rates can be achieved through the full-duplex inter-subchannel cooperative encoding strategy. We demonstrated potential rate gains attained by cooperative OFDMA, through simulations.

For a two user cooperative OFDMA system with full CSI, we obtained the optimal power allocation policies which maximize the rate region achievable by a channel adaptive implementation of inter-subchannel BMSE, used in conjunction with backwards decoding. We provided the optimality conditions that need to be satisfied by the powers associated with the users' codewords and derived the closed form expressions for the optimal powers. We proposed two algorithms that can be used to optimize the powers to achieve any desired rate pair on the rate region

boundary: a projected subgradient algorithm, and an iterative waterfilling-like algorithm based on KKT conditions for optimality, which operates one user at a time and converges much faster. We observed that, utilization of power control to take advantage of the diversity offered by the cooperative OFDMA system, not only leads to a remarkable improvement in achievable rates, but also may help determine how the subchannels have to be instantaneously allocated to various tasks in cooperation.

5.2 Future Directions

The main focus of our thesis is to develop cooperative strategies, determine the achievable rates and apply resource allocation to the cooperative OFDMA systems, but actually OFDMA is a specific case of the vector multiple access channels, with diagonal channel matrix. An interesting problem can be extending the cooperative encoding strategies and resource allocation methods proposed in our thesis to a more general condition, to vector multiple access channels where the channel matrix is not necessarily diagonal, especially to cooperative MIMO systems. Deriving achievable rate region expressions and investigating optimum power allocation protocols for the vector multiple access channels employing mutual cooperation can be a very challenging research topic. We hope that applying our proposed cooperative encoding strategies and resource allocation techniques to vector multiple access channels (specifically MIMO systems) can further increase the gains provided in our thesis to cooperative OFDMA channels.

References

[1] J. Zander. Performance of optimum transmitter power control in cellular radio systems. *IEEE Transactions on Vehicular Technology*, 41(1):57–62, Feb. 1992.

[2] R. D. Yates. A framework for uplink power control in cellular radio systems. *IEEE Journal on Selected Areas in Communications*, 13(7):1341–1347, Sep. 1995.

[3] E. Biglieri, J. Proakis, and S. Shamai. Fading channels: information-theoretic and communications aspects. *IEEE Transactions on Information Theory*, 44 (6):2619–2692, Oct. 1998.

[4] A. J. Goldsmith and P. P. Varaiya. Capacity of fading channels with channel side information. *IEEE Transactions on Information Theory*, 43(6):1986–1992, Nov. 1997.

[5] R. Knopp and P. A. Humblet. Information capacity and power control in single-cell multiuser communications. In *1995 IEEE International Conference on Communications, 1995. ICC '95 Seattle, 'Gateway to Globalization'*, volume 1, pages 331–335, Jun. 1995.

[6] D. N. C. Tse and S. V. Hanly. Multiaccess fading channels. i. polymatroid structure, optimal resource allocation and throughput capacities. *IEEE Transactions on Information Theory*, 44(7):2796–2815, Nov. 1998.

[7] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi. Iterative water-filling for gaussian vector multiple-access channels. *IEEE Transactions on Information Theory*, 50(1):145–152, Jan. 2004.

[8] O. Kaya and S. Ulukus. Optimum power control for cdma with deterministic sequences in fading channels. *IEEE Transactions on Information Theory*, 50 (10):2449–2462, Oct. 2004.

[9] O. Kaya and S. Ulukus. Achieving the capacity region boundary of fading cdma channels via generalized iterative waterfilling. *IEEE Transactions on Wireless Communications*, 5(11):3215–3223, Nov. 2006.

[10] J. Jang and K. B. Lee. Transmit power adaptation for multiuser ofdm systems. *IEEE Journal on Selected Areas in Communications*, 21(2):171–178, Feb. 2003.

[11] Z. Shen, J. G. Andrews, and B. L. Evans. Adaptive resource allocation in multiuser ofdm systems with proportional rate constraints. *IEEE Transactions on Wireless Communications*, 4(6):2726–2737, Nov. 2005.

[12] L. Gao and S. Cui. Efficient subcarrier, power, and rate allocation with fairness consideration for ofdma uplink. *IEEE Transactions on Wireless Communications*, 7(5):1507–1511, May 2008.

[13] C. Y. Ng and C. W. Sung. Low complexity subcarrier and power allocation for utility maximization in uplink ofdma systems. *IEEE Transactions on Wireless Communications*, 7(5):1667–1675, May 2008.

[14] C. Mohanram and S. Bhashyam. Joint subcarrier and power allocation in channel-aware queue-aware scheduling for multiuser ofdm. *IEEE Transactions on Wireless Communications*, 6(9):3208–3213, Sep. 2007.

[15] I. C. Wong and B. L. Evans. Optimal downlink ofdma resource allocation with linear complexity to maximize ergodic rates. *IEEE Transactions on Wireless Communications*, 7(3):962–971, Mar. 2008.

[16] M. Pischella and J.-C. Belfiore. Distributed resource allocation for rate-constrained users in multi-cell ofdma networks. *IEEE Communications Letters*, 12(4):250–252, Apr. 2008.

[17] A. Lozano, A. M. Tulino, and S. Verdii. Optimum power allocation for multiuser ofdm with arbitrary signal constellations. *IEEE Transactions on Communications*, 56(5):828–837, May 2008.

[18] M. Tao, Y.-C. Liang, and F. Zhang. Resource allocation for delay differentiated traffic in multiuser ofdm systems. *IEEE Transactions on Wireless Communications*, 7(6):2190–2201, Jun. 2008.

[19] E. C. van der Meulen. Three-terminal communication channels. *Advances in Applied Probability*, 3:120–154, 1971.

[20] T. Cover and A. E. Gamal. Capacity theorems for the relay channel. *IEEE Transactions on Information Theory*, 25(5):572–584, Sep. 1979.

[21] T. Cover and C. Leung. An achievable rate region for the multiple-access channel with feedback. *IEEE Transactions on Information Theory*, 27(3): 292–298, May 1981.

[22] F. Willems and E. van der Meulen. The discrete memoryless multiple-access channel with cribbing encoders. *IEEE Transactions on Information Theory*, 31(3):313–327, May 1985.

[23] F. Willems and E. van der Meulen. Partial feedback for the discrete memoryless multiple access channel. *IEEE Transactions on Information Theory*, 29(2):287–290, Mar. 1983.

[24] F. Willems. The discrete memoryless multiple access channel with partially cooperating encoders. *IEEE Transactions on Information Theory*, 29(3): 441–445, May 1983.

[25] A. Carleial. Multiple-access channels with different generalized feedback signals. *IEEE Transactions on Information Theory*, 28(6):841–850, Nov. 1982.

[26] F. M. J. Willems, E. C. van der Meulen, and J. P. M. Schalkwijk. An achievable rate region for the multiple access channel with generalized feedback. In

Proc. Allerton Conf. Communications, Control, and Computing, Monticello, IL, pages 284–292, Oct. 1983.

[27] A. Sendonaris, E. Erkip, and B. Aazhang. User cooperation diversity. part i. system description. *IEEE Transactions on Communications*, 51(11):1927–1938, Nov. 2003.

[28] J. N. Laneman, D. N. C. Tse, and G. W. Wornell. Cooperative diversity in wireless networks: Efficient protocols and outage behavior. *IEEE Transactions on Information Theory*, 50(12):3062–3080, Dec. 2004.

[29] L. Sankaranarayanan, G. Kramer, and N. B. Mandayam. Capacity theorems for the multiple-access relay channel. In *Proc. 42nd Annu. Allerton Conf. Communications, Control, and Computing, Monticello, IL*, pages 1872–1791, Sep. 2004.

[30] L.-L. Xie and P. R. Kumar. A network information theory for wireless communication: scaling laws and optimal operation. *IEEE Transactions on Information Theory*, 50(5):748–767, May 2004.

[31] B. Schein. *Distributed Coordination in Network Information Theory*. PhD thesis, MIT, MIT, Massachusetts, Aug. 2001.

[32] C. Edemen and O. Kaya. Achievable rates for the three user cooperative multiple access channel. In *IEEE Wireless Communications and Networking Conference, 2008. WCNC 2008*, pages 1507–1512, Apr. 2008.

[33] G. Kramer, M. Gastpar, and P. Gupta. Cooperative strategies and capacity theorems for relay networks. *IEEE Transactions on Information Theory*, 51 (9):3037–3063, Sep. 2005.

[34] O. Kaya and S. Ulukus. Power control for fading cooperative multiple access channels. *IEEE Transactions on Wireless Communications*, 6(8):2915–2923, Aug. 2007.

[35] S. Yatawatta and A. P. Petropulu. A multiuser ofdm system with user cooperation. In *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, 2004*, volume 1, pages 319–323, Nov. 2004.

[36] J. C. H. Lin and A. Stefanov. Coded cooperation for ofdm systems. In *2005 International Conference on Wireless Networks, Communications and Mobile Computing*, volume 1, pages 7–10, Jun. 2005.

[37] M. Kaneko, K. Hayashi, P. Popovski, K. Ikeda, H. Sakai, and R. Prasad. Amplify-and-forward cooperative diversity schemes for multi-carrier systems. *IEEE Transactions on Wireless Communications*, 7(5):1845–1850, May 2008.

[38] W. Wang, S. Yan, and S. Yang. Optimally joint subcarrier matching and power allocation in ofdm multihop system. *EURASIP J. Adv. Signal Process*, 2008, Jan. 2008.

[39] 3GPP. Evolved universal terrestrial radio access (e-utra) and evolved universal terrestrial radio access network (e-utran); overall description stage 2. *3GPP TS 36.300*, 8.12.0, 2010.

[40] R. W. Chang. Synthesis of band-limited orthogonal signals for multichannel data transmission. *Bell System Technical Journal*, 45:1775–1796, 1966.

[41] B. Saltzberg. Performance of an efficient parallel data transmission system. *IEEE Transactions on Communication Technology*, 15(6):805–811, Dec. 1967.

[42] R. Chang and R. Gibby. A theoretical study of performance of an orthogonal multiplexing data transmission scheme. *IEEE Transactions on Communication Technology*, 16(4):529–540, Aug. 1968.

[43] S. Weinstein and P. Ebert. Data transmission by frequency-division multiplexing using the discrete fourier transform. *IEEE Transactions on Communication Technology*, 19(5):628–634, Oct. 1971.

[44] A. Peled and A. Ruiz. Frequency domain data transmission using reduced computational complexity algorithms. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '80*, volume 5, pages 964–967, Apr. 1980.

[45] B. Hirosaki. An orthogonally multiplexed qam system using the discrete fourier transform. *IEEE Transactions on Communications*, 29(7):982–989, Jul. 1981.

[46] J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson. On channel estimation in ofdm systems. In *1995 IEEE 45th Vehicular Technology Conference*, volume 2, pages 815–819, Jul. 1995.

[47] W. Y. Zou and Y. Wu. Cofdm: an overview. *IEEE Transactions on Broadcasting*, 41(1):1–8, Mar. 1995.

[48] T. Cover and J. Thomas. *Elements of Information Theory*. New York: Wiley, 1991.

[49] M. Yuksel and E. Erkip. *Cooperative Communications for Improved Wireless Network Transmission: Framework for Virtual Antenna Array Applications*.

[50] E. Erkip, A. Sendonaris, A. Stefanov, and B. Aazhang. *Advances in Network Information Theory*. AMS DIMACS Series, 2004.

[51] C.-M. Zeng, F. Kuhlmann, and A. Buzo. Achievability proof of some multiuser channel coding theorems using backward decoding. *IEEE Transactions on Information Theory*, 35(6):1160–1165, Nov. 1989.

[52] A. Wyner and J. Ziv. The rate-distortion function for source coding with side information at the decoder. *IEEE Transactions on Information Theory*, 22(1):1–10, Jan. 1976.

[53] F. M. J. Willems. *Informationtheoretical Results for the Discrete Memoryless Multiple Access Channel*. PhD thesis, Katholieke Universiteit Leuven, Leuven, Belgium, Oct. 1982.

[54] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[55] B. Can, H. Yomo, and E. De Carvalho. Hybrid forwarding scheme for cooperative relaying in ofdm based networks. In *IEEE International Conference on Communications, 2006. ICC '06*, volume 10, pages 4520–4525, Jun. 2006.

[56] L. Dai, B. Gui, and L. J. Cimini. Selective relaying in ofdm multihop cooperative networks. In *IEEE Wireless Communications and Networking Conference, 2007. WCNC 2007*, pages 963–968, Mar. 2007.

[57] S. Bakim and O. Kaya. Cooperative strategies and achievable rates for two user ofdma channels. *IEEE Transactions on Wireless Communications*, 10 (12):4029–4034, Dec. 2011.

[58] S. Bakim and O. Kaya. Achievable rates for two user cooperative ofdma. In *2010 IEEE Global Telecommunications Conference. GLOBECOM 2010*, pages 1–5, Dec. 2010.

[59] K. Kim, Y. Han, and S.-L. Kim. Joint subcarrier and power allocation in uplink ofdma systems. *IEEE Communications Letters*, 9(6):526–528, Jun. 2005.

[60] L. Weng and R. D. Murch. Cooperation strategies and resource allocations in multiuser ofdma systems. *IEEE Transactions on Vehicular Technology*, 58(5):2331–2342, Jun. 2009.

[61] W. Shim, Y. Han, and S. Kim. Fairness-aware resource allocation in a cooperative ofdma uplink system. *IEEE Transactions on Vehicular Technology*, 59(2):932–939, Feb. 2010.

[62] Z. Han, T. Himsoon, W. P. Siriwongpairat, and K. J. R. Liu. Resource allocation for multiuser cooperative ofdm networks: Who helps whom and how to cooperate. *IEEE Transactions on Vehicular Technology*, 58(5):2378–2391, Jun. 2009.

[63] N. Z. Shor. *Minimization Methods for Non-Differentiable Functions.* Springer-Verlag, 1979.

Appendix A

Probability Of Error Analysis In The Achievability Proof

In decoding, an error occurs either if one or more of the codewords violate the power constraint, or if one or more of the decoded submessages in any of the B blocks of transmission are not equal to the transmitted submessages, i.e., $\{\tilde{v}_{kj}^{(i)}[b] \neq v_{kj}^{(i)}[b]\}$ or $\{\tilde{w}_{kj}^{(i)}[b] \neq w_{kj}^{(i)}[b]\}$, or $\{\tilde{w}_{k0}^{(i)}[b] \neq w_{k0}^{(i)}[b]\}$, for some $i \in \{1, ..., N\}$ and some $b \in \{1, ..., B\}$. The probability of error, averaged over all random codebooks and all messages [48], is given by

$$\overline{P_e^B} = \overline{Pr}\left\{\bigcup_{b=1}^{B}\left[\bigcup_{i=1}^{N}\Big(\{\tilde{v}_{12}^{(i)}[b] \neq v_{12}^{(i)}[b]\} \cup \{\tilde{v}_{21}^{(i)}[b] \neq v_{21}^{(i)}[b]\} \cup \{\hat{w}_{12}^{(i)}[b] \neq w_{12}^{(i)}[b]\} \right.\right.$$
$$\left.\left. \cup\{\hat{w}_{21}^{(i)}[b] \neq w_{21}^{(i)}[b]\} \cup \{\tilde{w}_{10}^{(i)}[b] \neq w_{10}^{(i)}[b]\} \cup \{\tilde{w}_{20}^{(i)}[b] \neq w_{20}^{(i)}[b]\}\Big) \cup E_{bP}\right]\right\}, \tag{A.1}$$

where, E_{bP} denotes the event that the power constraint is violated by the codewords of some user. The transmitters and the receiver all use joint typicality decoding. Let us define the following events to be used in the error probability calculation.

Event $E_{b\hat{w}_{kj}}^{(i)}$ represents the event that the received codewords at user j over the ith subchannel in block b are jointly typical with the codewords corresponding to message $\hat{w}_{kj}^{(i)}$, i.e.,

$$E_{b\hat{w}_{kj}}^{(i)} = \Big\{\Big(u^{(i)}(\hat{v}_{kj}^{(i)}[b-1], v_{jk}^{(i)}[b-1]), x_{kj}^{(i)}(\hat{w}_{kj}^{(i)}[b], u^{(i)}(\hat{v}_{kj}^{(i)}[b-1], v_{jk}^{(i)}[b-1])),$$

$$x_{jk}^{(i)}(w_{jk}^{(i)}[b], u^{(i)}(\hat{v}_{kj}^{(i)}[b-1], v_{jk}^{(i)}[b-1])),$$
$$x_j^{(i)}(w_{j0}^{(i)}[b], w_{jk}^{(i)}[b], u^{(i)}(\hat{v}_{kj}^{(i)}[b-1], v_{jk}^{(i)}[b-1])), y_j^{(i)}[b]\Big)$$
$$\in A_\varepsilon(U^{(i)}, X_{kj}^{(i)}, X_{jk}^{(i)}, X_j^{(i)}, Y_j^{(i)})\Big\}. \tag{A.2}$$

Event $F^{(i)}_{b\tilde{v}_{12},\tilde{v}_{21},\tilde{w}_{10},\tilde{w}_{20}}$ represents the event that the codewords received by the receiver over the ith subchannel in block b are jointly typical with codewords corresponding to $\tilde{v}_{12}^{(i)}, \tilde{v}_{21}^{(i)}, \tilde{w}_{10}^{(i)}, \tilde{w}_{20}^{(i)}$:

$$\begin{aligned} F^{(i)}_{b\tilde{v}_{12},\tilde{v}_{21},\tilde{w}_{10},\tilde{w}_{20}} = \Big\{\Big(&y^{(i)}[b], u^{(i)}(\tilde{v}_{12}^{(i)}[b-1], \tilde{v}_{21}^{(i)}[b-1]),\\ &x_{12}^{(i)}(\tilde{w}_{12}^{(i)}[b], u^{(i)}(\tilde{v}_{12}^{(i)}[b-1], \tilde{v}_{21}^{(i)}[b-1])),\\ &x_{21}^{(i)}(\tilde{w}_{21}^{(i)}[b], u^{(i)}(\tilde{v}_{12}^{(i)}[b-1], \tilde{v}_{21}^{(i)}[b-1])),\\ &x_{1}^{(i)}(\tilde{w}_{10}^{(i)}[b], \tilde{w}_{12}^{(i)}[b], u^{(i)}(\tilde{v}_{12}^{(i)}[b-1], \tilde{v}_{21}^{(i)}[b-1])),\\ &x_{2}^{(i)}(\tilde{w}_{20}^{(i)}[b], \tilde{w}_{21}^{(i)}[b], u^{(i)}(\tilde{v}_{12}^{(i)}[b-1], \tilde{v}_{21}^{(i)}[b-1]))\Big)\\ &\in A_\varepsilon(Y^{(i)}, U^{(i)}, X_{12}^{(i)}, X_{21}^{(i)}, X_1^{(i)}, X_2^{(i)})\Big\}. \end{aligned} \tag{A.3}$$

Using the above defined events, and the union bound, we have

$$\begin{aligned} \overline{P_e^B} \le &\sum_{b=1}^{B} Pr\{E_{bP}\} + \sum_{i=1}^{N}\Bigg(\sum_{b=1}^{B-1}\overline{Pr}\left\{\left(E^{(i)}_{b\hat{w}_{12}}\right)^c\right\} + \sum_{b=1}^{B-1}\overline{Pr}\left\{\left(E^{(i)}_{b\hat{w}_{21}}\right)^c\right\}\\ &+\sum_{b=1}^{B-1}\sum_{\hat{w}_{12}^{(i)}[b]\neq w_{12}^{(i)}[b]}\overline{Pr}\left\{E^{(i)}_{b\hat{w}_{12}}\right\} + \sum_{b=1}^{B-1}\sum_{\hat{w}_{21}^{(i)}[b]\neq w_{21}^{(i)}[b]}\overline{Pr}\left\{E^{(i)}_{b\hat{w}_{21}}\right\}\\ &+\sum_{b=1}^{B}\overline{Pr}\left\{\left(F^{(i)}_{b\tilde{v}_{12},\tilde{v}_{21},\tilde{w}_{10},\tilde{w}_{20}}\right)^c\right\}\\ &+\sum_{b=1}^{B}\sum_{(\tilde{v}_{12}^{(i)}[b],\tilde{v}_{21}^{(i)}[b],\tilde{w}_{10}^{(i)}[b],\tilde{w}_{20}^{(i)}[b])\neq(v_{12}^{(i)}[b],v_{21}^{(i)}[b],w_{10}^{(i)}[b],w_{20}^{(i)}[b])}\overline{Pr}\left\{F^{(i)}_{b\tilde{v}_{12},\tilde{v}_{21},\tilde{w}_{10},\tilde{w}_{20}}\right\}\Bigg). \end{aligned} \tag{A.4}$$

To give bounds on the average error probability, we assume without loss of generality, $\forall b \in \{1,\cdots,B\}$ and $\forall i \in \{1,\cdots,N\}$ that $(v_{12}^{(i)}[b-1], v_{21}^{(i)}[b-1], w_{12}^{(i)}[b], w_{21}^{(i)}[b], w_{10}^{(i)}[b], w_{20}^{(i)}[b]) = (1,1,1,1,1,1)$. Then, by dropping the block index, we redefine the events (A.2) and (A.3) as $E^{(i)}_{\hat{w}_{kj}}$, $F^{(i)}_{\tilde{v}_{12},\tilde{v}_{21},\tilde{w}_{10},\tilde{w}_{20}}$, and the average error

probability expression can be rewritten as

$$\begin{aligned}
\overline{P_e^B} \leq \sum_{i=1}^{N} \Bigg(& (B-1)\overline{Pr}\left\{\left(E_{\hat{w}_{12}}^{(i)}\right)^c\right\} + (B-1)\overline{Pr}\left\{\left(E_{\hat{w}_{21}}^{(i)}\right)^c\right\} \\
& + (B-1)\sum_{\hat{w}_{12}^{(i)}\neq 1} \overline{Pr}\left\{E_{\hat{w}_{12}}^{(i)}\right\} + (B-1)\sum_{\hat{w}_{21}^{(i)}\neq 1} \overline{Pr}\left\{E_{\hat{w}_{21}}^{(i)}\right\} + \overline{Pr}\left\{\left(F_{\tilde{v}_{12},\tilde{v}_{21},1,1}^{(i)}\right)^c\right\} \\
& + \sum_{\tilde{w}_{10}^{(i)}\neq 1} \overline{Pr}\left\{F_{1,1,\tilde{w}_{10},1}^{(i)}\right\} + \sum_{\tilde{w}_{20}^{(i)}\neq 1} \overline{Pr}\left\{F_{1,1,1,\tilde{w}_{20}}^{(i)}\right\} \\
& + \sum_{(\tilde{v}_{12}^{(i)},\tilde{v}_{21}^{(i)})\neq(1,1)} \overline{Pr}\left\{F_{\tilde{v}_{12},\tilde{v}_{21},1,1}^{(i)}\right\} + (B-2)\overline{Pr}\left\{\left(F_{\tilde{v}_{12},\tilde{v}_{21},\tilde{w}_{10},\tilde{w}_{20}}^{(i)}\right)^c\right\} \\
& + (B-2)\sum_{(\tilde{v}_{12}^{(i)},\tilde{v}_{21}^{(i)})\neq(1,1)} \overline{Pr}\left\{F_{\tilde{v}_{12},\tilde{v}_{21},\tilde{w}_{10},\tilde{w}_{20}}^{(i)}\right\} + (B-2)\sum_{\tilde{w}_{10}^{(i)}\neq 1,\tilde{w}_{20}^{(i)}\neq 1} \overline{Pr}\left\{F_{1,1,\tilde{w}_{10},\tilde{w}_{20}}^{(i)}\right\} \\
& + (B-2)\sum_{\tilde{w}_{10}^{(i)}\neq 1} \overline{Pr}\left\{F_{1,1,\tilde{w}_{10},1}^{(i)}\right\} + (B-2)\sum_{\tilde{w}_{20}^{(i)}\neq 1} \overline{Pr}\left\{F_{1,1,1,\tilde{w}_{20}}^{(i)}\right\} \\
& + \overline{Pr}\left\{\left(F_{1,1,\tilde{w}_{10},\tilde{w}_{20}}^{(i)}\right)^c\right\} + \sum_{\tilde{w}_{10}^{(i)}\neq 1,\tilde{w}_{20}^{(i)}\neq 1} \overline{Pr}\left\{F_{1,1,\tilde{w}_{10},\tilde{w}_{20}}^{(i)}\right\} \Bigg) + \sum_{b=1}^{B} Pr\{E_{bP}\}.
\end{aligned} \tag{A.5}$$

By the law of large numbers, $\exists n_1$ such that $\forall n > n_1$, $Pr\{E_{bP}\} < \epsilon$, $\forall \epsilon > 0$ [48]. Also, due to asymptotic equipartition property (AEP) [48], $\forall \epsilon > 0$, $\exists n_2$ such that $\forall n > n_2$, we have

$$\begin{aligned}
\overline{Pr}\left\{\left(E_{\hat{w}_{12}}^{(i)}\right)^c\right\} \leq \varepsilon, & \overline{Pr}\left\{\left(E_{\hat{w}_{21}}^{(i)}\right)^c\right\} \leq \varepsilon, \\
\overline{Pr}\left\{\left(F_{\tilde{v}_{12},\tilde{v}_{21},1,1}^{(i)}\right)^c\right\} \leq \varepsilon, & \overline{Pr}\left\{\left(F_{\tilde{v}_{12},\tilde{v}_{21},\tilde{w}_{10},\tilde{w}_{20}}^{(i)}\right)^c\right\} \leq \varepsilon, \quad \overline{Pr}\left\{\left(F_{1,1,\tilde{w}_{10},\tilde{w}_{20}}^{(i)}\right)^c\right\} \leq \varepsilon.
\end{aligned}$$

Note that, the average probability that event $E_{\hat{w}_{12}}^{(i)}$ occurs, $\forall i, i = 1, ..., N$ is;

$$\begin{aligned}
\overline{Pr}\left\{E_{\hat{w}_{12}}^{(i)}\right\} = & \sum_{(u^{(i)},x_{12}^{(i)},x_{21}^{(i)},x_2^{(i)},y_2^{(i)})\in A_\varepsilon(U^{(i)},X_{12}^{(i)},X_{21}^{(i)},X_2^{(i)},Y_2^{(i)})} P(u^{(i)})P(x_{12}^{(i)}|u^{(i)})P(x_{21}^{(i)}|u^{(i)})P(x_2^{(i)}|x_{21}^{(i)},u^{(i)})P(y_2^{(i)}|x_2^{(i)},x_{21}^{(i)},u^{(i)}), \\
\leq & 2^{n(H(U^{(i)},X_{12}^{(i)},X_{21}^{(i)},X_2^{(i)},Y_2^{(i)})+\varepsilon)} 2^{-n(H(U^{(i)},X_{12}^{(i)},X_{21}^{(i)},X_2^{(i)})-\varepsilon)} 2^{-n(H(Y_2^{(i)}|U^{(i)},X_{21}^{(i)},X_2^{(i)})-2\varepsilon)}, \\
= & 2^{-n(I(X_{12}^{(i)};Y_2^{(i)}|X_2^{(i)},X_{21}^{(i)},U^{(i)})-4\varepsilon)}, \\
= & 2^{-n(I(X_{12}^{(i)};Y_2^{(i)}|X_2^{(i)},U^{(i)})-4\varepsilon)} \quad \text{if} \quad \hat{w}_{12}^{(i)} \neq 1.
\end{aligned} \tag{A.6}$$

Similarly, using AEP, the probability of each event (A.5) can easily be derived, leading to the following bounds:

$$\overline{Pr}\left\{E^{(i)}_{\hat{w}_{21}}\right\} \leq 2^{-n(I(X^{(i)}_{21};Y^{(i)}_{1}|X^{(i)}_{1},U^{(i)})-4\varepsilon)} \qquad \text{if } \hat{w}^{(i)}_{21} \neq 1, \tag{A.7}$$

$$\overline{Pr}\left\{F^{(i)}_{\tilde{v}_{12},\tilde{v}_{21},1,1}\right\} \leq 2^{-n(I(X^{(i)}_{1},X^{(i)}_{2};Y^{(i)})-3\varepsilon)} \qquad \text{if } (\tilde{v}^{(i)}_{12},\tilde{v}^{(i)}_{21}) \neq (1,1), \tag{A.8}$$

$$\overline{Pr}\left\{F^{(i)}_{\tilde{v}_{12},\tilde{v}_{21},\tilde{w}_{10},\tilde{w}_{20}}\right\} \leq 2^{-n(I(X^{(i)}_{1},X^{(i)}_{2};Y^{(i)})-3\varepsilon)} \qquad \text{if } (\tilde{v}^{(i)}_{12},\tilde{v}^{(i)}_{21}) \neq (1,1), \tag{A.9}$$

$$\overline{Pr}\left\{F^{(i)}_{1,1,\tilde{w}_{10},\tilde{w}_{20}}\right\} \leq 2^{-n(I(X^{(i)}_{1},X^{(i)}_{2};Y^{(i)}|X^{(i)}_{12},X^{(i)}_{21},U_{(i)})-4\varepsilon)} \qquad \text{if } \tilde{w}^{(i)}_{10} \neq 1, \tilde{w}^{(i)}_{20} \neq 1, \tag{A.10}$$

$$\overline{Pr}\left\{F^{(i)}_{1,1,\tilde{w}_{10},1}\right\} \leq 2^{-n(I(X^{(i)}_{1};Y^{(i)}|X^{(i)}_{2},X^{(i)}_{12},U^{(i)})-4\varepsilon)} \qquad \text{if } \tilde{w}^{(i)}_{10} \neq 1, \tag{A.11}$$

$$\overline{Pr}\left\{F^{(i)}_{1,1,1,\tilde{w}_{20}}\right\} \leq 2^{-n(I(X^{(i)}_{2};Y^{(i)}|X^{(i)}_{1},X^{(i)}_{21},U^{(i)})-4\varepsilon)} \qquad \text{if } \tilde{w}^{(i)}_{20} \neq 1. \tag{A.12}$$

Plugging (A.6)-(A.12) into the average probability of error equation, for n large enough, we obtain

$$\begin{aligned}
\overline{P^{B}_{e}} \leq \sum_{b=1}^{B}\epsilon + \sum_{i=1}^{N}\Big[&(B-1)\Big(\varepsilon+\varepsilon+2^{nR^{(i)}_{12}}2^{-n(I(X^{(i)}_{12};Y^{(i)}_{2}|X^{(i)}_{2},U^{(i)})-4\varepsilon)} \\
&+ 2^{nR^{(i)}_{21}}2^{-n(I(X^{(i)}_{21};Y^{(i)}_{1}|X^{(i)}_{1},U^{(i)})-4\varepsilon)}\Big)+\varepsilon+2^{n(R^{'(i)}_{12}+R^{'(i)}_{21})}2^{-n(I(X^{(i)}_{1},X^{(i)}_{2};Y^{(i)})-3\varepsilon)} \\
&+ (B-2)\Big(\varepsilon+2^{n(R^{'(i)}_{12}+R^{'(i)}_{21}+R^{(i)}_{10}+R^{(i)}_{20})}2^{-n(I(X^{(i)}_{1},X^{(i)}_{2};Y^{(i)})-3\varepsilon)} \\
&+ 2^{n(R^{(i)}_{10}+R^{(i)}_{20})}2^{-n(I(X^{(i)}_{1},X^{(i)}_{2};Y^{(i)}|X^{(i)}_{12},X^{(i)}_{21},U_{(i)})-4\varepsilon)} \\
&+ 2^{nR^{(i)}_{10}}2^{-n(I(X^{(i)}_{1};Y^{(i)}|X^{(i)}_{2},X^{(i)}_{12},U_{(i)})-4\varepsilon)} \\
&+ 2^{nR^{(i)}_{20}}2^{-n(I(X^{(i)}_{2};Y^{(i)}|X^{(i)}_{1},X^{(i)}_{21},U_{(i)})-4\varepsilon)}\Big)+\varepsilon \\
&+ 2^{n(R^{(i)}_{10}+R^{(i)}_{20})}2^{-n(I(X^{(i)}_{1},X^{(i)}_{2};Y^{(i)}|X^{(i)}_{12},X^{(i)}_{21},U_{(i)})-4\varepsilon)} \\
&+ 2^{nR^{(i)}_{10}}2^{-n(I(X^{(i)}_{1};Y^{(i)}|X^{(i)}_{2},X^{(i)}_{12},U_{(i)})-4\varepsilon)} \\
&+ 2^{nR^{(i)}_{20}}2^{-n(I(X^{(i)}_{2};Y^{(i)}|X^{(i)}_{1},X^{(i)}_{21},U_{(i)})-4\varepsilon)}\Big].
\end{aligned} \tag{A.13}$$

Finally, note that $\exists n_3$ such that $\forall n > n_3$, all the terms that involve the rates

and mutual informations on the right hand side of (A.13) can be made arbitrarily small (say $< \epsilon$ each), if $R_{12}^{'(i)}$, $R_{21}^{'(i)}$, $R_{12}^{(i)}$, $R_{21}^{(i)}$, $R_{10}^{(i)}$, $R_{20}^{(i)}$ satisfy the following rate inequalities:

$$R_{12}^{(i)} \leq I(X_{12}^{(i)}; Y_2^{(i)} | X_2^{(i)}, U^{(i)}) - 4\varepsilon, \tag{A.14}$$

$$R_{21}^{(i)} \leq I(X_{21}^{(i)}; Y_1^{(i)} | X_1^{(i)}, U^{(i)}) - 4\varepsilon, \tag{A.15}$$

$$R_{10}^{(i)} \leq I(X_1^{(i)}; Y^{(i)} | X_2^{(i)}, X_{12}^{(i)}, U^{(i)}) - 4\varepsilon, \tag{A.16}$$

$$R_{20}^{(i)} \leq I(X_2^{(i)}; Y^{(i)} | X_1^{(i)}, X_{21}^{(i)}, U^{(i)}) - 4\varepsilon, \tag{A.17}$$

$$R_{10}^{(i)} + R_{20}^{(i)} \leq I(X_1^{(i)}, X_2^{(i)}; Y^{(i)} | X_{12}^{(i)}, X_{21}^{(i)}, U^{(i)}) - 4\varepsilon, \tag{A.18}$$

$$R_{12}^{'(i)} + R_{21}^{'(i)} + R_{10}^{(i)} + R_{20}^{(i)} \leq I(X_1^{(i)}, X_2^{(i)}; Y^{(i)}) - 3\varepsilon. \tag{A.19}$$

Therefore, taking $n > n_0 \triangleq \max\{n_1, n_2, n_3\}$, the average error probability is bounded by $\overline{P_e^B} \leq N(9B-8)\varepsilon + B\varepsilon$, and using standard arguments such as selecting the best codebook and throwing away the worst half of the codewords, the maximum probability of error per block can be made arbirtarily small.

Evaluating the rate constraint (A.14) for $R_{12}^{(i)}$ for Gaussian codewords, we have

$$\begin{aligned}
R_{12}^{(i)} &\leq I(X_{12}^{(i)}; Y_2^{(i)} | X_2^{(i)}, U^{(i)}), \\
&= h(Y_2^{(i)} | X_2^{(i)}, U^{(i)}) - h(Y_2^{(i)} | X_{12}^{(i)}, X_2^{(i)}, U^{(i)}), \\
&= h(X_{10}^{(i)} + X_{12}^{(i)} + U^{(i)} + Z_2^{(i)} | X_2^{(i)}, U^{(i)}) \\
&\quad - h(X_{10}^{(i)} + X_{12}^{(i)} + U^{(i)} + Z_2^{(i)} | X_{12}^{(i)}, X_2^{(i)}, U^{(i)}), \\
&= h(X_{10}^{(i)} + X_{12}^{(i)} + Z_2^{(i)} | X_2^{(i)}, U^{(i)}) - h(X_{10}^{(i)} + Z_2^{(i)} | X_{12}^{(i)}, X_2^{(i)}, U^{(i)}), \\
&= h(X_{10}^{(i)} + X_{12}^{(i)} + Z_2^{(i)}) - h(X_{10}^{(i)} + Z_2^{(i)}), \\
&= C\left(\frac{s_{12}^{(i)} p_{12}^{(i)}}{s_{12}^{(i)} p_{10}^{(i)} + 1}\right).
\end{aligned} \tag{A.20}$$

Likewise, the following rate constraints can be derived, thereby proving the theorem

$$R_{21}^{(i)} \leq C\left(\frac{s_{21}^{(i)} p_{21}^{(i)}}{s_{21}^{(i)} p_{20}^{(i)} + 1}\right), \tag{A.21}$$

$$R_{10}^{(i)} \leq C\left(s_{10}^{(i)} p_{10}^{(i)}\right), \tag{A.22}$$

$$R_{20}^{(i)} \leq C\left(s_{20}^{(i)} p_{20}^{(i)}\right), \tag{A.23}$$

$$R_{10}^{(i)} + R_{20}^{(i)} \leq C\left(s_{10}^{(i)} p_{10}^{(i)} + s_{20}^{(i)} p_{20}^{(i)}\right), \tag{A.24}$$

$$R_{12}^{'(i)} + R_{21}^{'(i)} + R_{10}^{(i)} + R_{20}^{(i)} \leq C\left(s_{10}^{(i)} p_{1}^{(i)} s_{20}^{(i)} p_{2}^{(i)} + 2\sqrt{s_{10}^{(i)} s_{20}^{(i)} p_{u_1}^{(i)} p_{u_2}^{(i)}}\right). \tag{A.25}$$

Appendix B

Proof Of Theorem 4.1 In Chapter 4

Let $p_k^{(i)}(\mathbf{s}) = p_{k0}^{(i)}(\mathbf{s}) + p_{kj}^{(i)}(\mathbf{s}) + p_{U_k}^{(i)}(\mathbf{s})$. To simplify the notation, let us drop the dependency of the powers on the channel states in each subchannel and define the following equations.

$$\begin{aligned}
A^{(i)} &= 1 + s_{10}^{(i)} p_1^{(i)} + s_{20}^{(i)} p_2^{(i)} + 2\sqrt{s_{10}^{(i)} s_{20}^{(i)} p_{U_1}^{(i)} p_{U_2}^{(i)}}, \\
B^{(i)} &= \frac{1 + s_{10}^{(i)} p_{10}^{(i)} + s_{20}^{(i)} p_{20}^{(i)}}{(1 + s_{12}^{(i)} p_{10}^{(i)})(1 + s_{21}^{(i)} p_{20}^{(i)})}, \\
C^{(i)} &= (1 + s_{12}^{(i)}(p_{10}^{(i)} + p_{12}^{(i)}))(1 + s_{21}^{(i)}(p_{20}^{(i)} + p_{21}^{(i)})).
\end{aligned}$$

Then the sum rate becomes:

$$R_{sum} = \sum_{i=1}^{N} \left(R_1^{(i)} + R_2^{(i)} \right) = \min \left\{ \frac{1}{2} \sum_{i=1}^{N} E\left[\log(A^{(i)}) \right], \frac{1}{2} \sum_{i=1}^{N} E\left[\log(B^{(i)} C^{(i)}) \right] \right\}.$$

Let us fix $p_k^{(i)}$, as well as $p_{U_k}^{(i)}$, then $A^{(i)}$ and $C^{(i)}$ will be fixed since $A^{(i)}$ is a function of $p_k^{(i)}$ and $p_{U_k}^{(i)}$, $C^{(i)}$ is a function of $\left(p_k^{(i)} - p_{U_k}^{(i)} \right)$. Only $B^{(i)}$ will have variables, so for maximizing the sum rate, we can focus on the maximization $B^{(i)}$.

$$\begin{gathered}
\max_{p_{10}^{(i)}, p_{20}^{(i)}} B^{(i)}(p_{10}^{(i)}, p_{20}^{(i)}), \\
s.t.\ p_{10}^{(i)} + p_{12}^{(i)} = p_1^{(i)} - p_{U_1}^{(i)} \ and \ p_{20}^{(i)} + p_{21}^{(i)} = p_2^{(i)} - p_{U_2}^{(i)}.
\end{gathered}$$

We take partial derivatives with respect to $p_{10}^{(i)}$ and $p_{20}^{(i)}$ to determine how to maximize R_{sum}:

$$\frac{\partial B^{(i)}}{\partial p_{10}^{(i)}} = \frac{s_{10}^{(i)} - s_{12}^{(i)}(1 + s_{20}^{(i)} p_{20}^{(i)})}{(1 + s_{12}^{(i)} p_{10}^{(i)})(1 + s_{21}^{(i)} p_{20}^{(i)})},$$
$$\frac{\partial B^{(i)}}{\partial p_{20}^{(i)}} = \frac{s_{20}^{(i)} - s_{21}^{(i)}(1 + s_{10}^{(i)} p_{10}^{(i)})}{(1 + s_{21}^{(i)} p_{20}^{(i)})(1 + s_{12}^{(i)} p_{10}^{(i)})}.$$

1. if $s_{12}^{(i)} > s_{10}^{(i)}$ and $s_{21}^{(i)} > s_{20}^{(i)} \Rightarrow \frac{\partial B^{(n)}}{\partial p_{10}^{(n)}} < 0$ and $\frac{\partial B^{(n)}}{\partial p_{20}^{(n)}} < 0$.
   Therefore $B^{(i)}(p_{10}^{(i)}, p_{20}^{(i)})$ is monotically decreasing in both $p_{10}^{(i)}$ and $p_{20}^{(i)}$.
   $R_{sum}\left(B^{(i)}(p_{10}^{(i)}, p_{20}^{(i)})\right)$ is maximized at $p_{10}^{(i)} = p_{20}^{(i)} = 0$.

2. if $s_{12}^{(i)} > s_{10}^{(i)}$ and $s_{21}^{(i)} \leq s_{20}^{(i)} \Rightarrow \frac{\partial B^{(i)}}{\partial p_{10}^{(i)}} < 0$ therefore $p_{10}^{(i)} = 0$, $\frac{\partial B^{(i)}}{\partial p_{20}^{(i)}} > 0$ when $p_{10}^{(i)} = 0$ and $\frac{\partial^2 B^{(i)}}{\partial {p_{20}^{(i)}}^2} = \frac{-2 s_{21}^{(i)}(s_{20}^{(i)} - s_{21}^{(i)})}{(1 + s_{21}^{(i)} p_{20}^{(i)})^3} < 0$ when $p_{10}^{(i)} = 0$.
   Therefore $p_{20}^{(i)}$ should take its maximum possible value, i.e., $p_{21}^{(i)} = 0$ ($p_{20}^{(i)} = p_2^{(i)} - p_{U_2}^{(i)}$).
   $R_{sum}\left(B^{(i)}(p_{10}^{(i)}, p_{20}^{(i)})\right)$ is maximized at $p_{10}^{(i)} = p_{21}^{(i)} = 0$.

3. if $s_{12}^{(i)} \leq s_{10}^{(i)}$ and $s_{21}^{(i)} > s_{20}^{(i)} \Rightarrow \frac{\partial B^{(i)}}{\partial p_{20}^{(i)}} < 0$ therefore $p_{20}^{(i)} = 0$, $\frac{\partial B^{(i)}}{\partial p_{10}^{(i)}} > 0$ when $p_{20}^{(i)} = 0$ and $\frac{\partial^2 B^{(i)}}{\partial {p_{10}^{(i)}}^2} = \frac{-2 s_{12}^{(i)}(s_{10}^{(i)} - s_{12}^{(i)})}{(1 + s_{12}^{(i)} p_{10}^{(i)})^3} < 0$ when $p_{20}^{(i)} = 0$.
   Therefore $p_{10}^{(i)}$ should take its maximum possible value, i.e., $p_{12}^{(i)} = 0$ ($p_{10}^{(i)} = p_1^{(i)} - p_{U_1}^{(i)}$).
   $R_{sum}\left(B^{(i)}(p_{10}^{(i)}, p_{20}^{(i)})\right)$ is maximized at $p_{12}^{(i)*} = p_{20}^{(i)*} = 0$,

4. if $s_{12}^{(i)} \leq s_{10}^{(i)}$ and $s_{21}^{(i)} \leq s_{20}^{(i)} \Rightarrow \frac{\partial^2 B^{(i)}}{\partial {p_{10}^{(i)}}^2} = \frac{-2 s_{12}^{(i)}(s_{10}^{(i)} - s_{12}^{(i)}(1 + s_{20}^{(i)} p_{20}^{(i)}))}{(1 + s_{12}^{(i)} p_{10}^{(i)})^3 (1 + s_{21}^{(i)} p_{20}^{(i)})} < 0$ and
   $\frac{\partial^2 B^{(i)}}{\partial {p_{20}^{(i)}}^2} = \frac{-2 s_{21}^{(i)}(s_{20}^{(i)} - s_{21}^{(i)}(1 + s_{10}^{(i)} p_{10}^{(i)}))}{(1 + s_{21}^{(i)} p_{20}^{(i)})^3 (1 + s_{12}^{(i)} p_{10}^{(i)})} < 0$.
   Determinant of the Hessian matrix < 0, therefore it is a saddle point.
   $R_{sum}\left(B^{(i)}(p_{10}^{(i)}, p_{20}^{(i)})\right)$ is maximized at one of the boundaries;
   $p_{12}^{(i)*} = p_{21}^{(i)*} = 0$ or $p_{10}^{(i)*} = p_{21}^{(i)*} = 0$ or $p_{12}^{(i)*} = p_{20}^{(i)*} = 0$.

Appendix C

Proof Of Lemma 4.2 In Chapter 4

Note that KKT conditions are necessary and sufficient for optimality. To obtain the KKT conditions we first assign the Lagrange multipliers γ_1, γ_2, λ_1 and λ_2 to the inequality constraints (4.14), (4.15), (4.16), (4.17) respectively, and we further assign $\epsilon_t^i(\mathbf{s})$, $t = 1, \ldots, 6$, $\forall \mathbf{s}$ to the positivity constraints (4.18), to obtain the Lagrangian

$$
\begin{aligned}
\mathcal{L} = R_\mu &+ \gamma_1 \Bigg[(\mu_1 - \mu_2) \left(\sum_{i=1}^{N} E_{S_1,S_2} \left[C(p_{12}^{(i)}(\mathbf{s}) s_{12}^{(i)}) \right] + \sum_{i=1}^{N} E_{S_3,S_4} \left[C(p_{10}^{(i)}(\mathbf{s}) s_{10}^{(i)}) \right] \right) \\
&+ \mu_2 \sum_{i=1}^{N} \Bigg(E_{S_1} \left[C(p_{12}^{(i)}(\mathbf{s}) s_{12}^{(i)}) + C(p_{21}^{(i)}(\mathbf{s}) s_{21}^{(i)}) \right] + E_{S_2} \left[C(p_{12}^{(i)}(\mathbf{s}) s_{12}^{(i)}) + C(p_{20}^{(i)}(\mathbf{s}) s_{20}^{(i)}) \right] \\
&+ E_{S_3} \left[C(p_{10}^{(i)}(\mathbf{s}) s_{10}^{(i)}) + C(p_{21}^{(i)}(\mathbf{s}) s_{21}^{(i)}) \right] + E_{S_4} \left[C(p_{10}^{(i)}(\mathbf{s}) s_{10}^{(i)} + p_{20}^{(i)}(\mathbf{s}) s_{20}^{(i)}) \right] \Bigg) - R_\mu \Bigg] \\
&+ \gamma_2 \Bigg[(\mu_1 - \mu_2) \left(\sum_{i=1}^{N} E \left[C(p_{1m}^{(i)}(\mathbf{s}) s_{1m}^{(i)}) \right] \right) + \mu_2 \sum_{i=1}^{N} E \Big[C \Big(s_{10}^{(i)} (p_{1m}^{(i)}(\mathbf{s}) + p_{U_1}^{(i)}(\mathbf{s})) \\
&+ s_{20}^{(i)} (p_{2n}^{(i)}(\mathbf{s}) + p_{U_2}^{(i)}(\mathbf{s})) + 2\sqrt{s_{10}^{(i)} s_{20}^{(i)} p_{U_1}^{(i)}(\mathbf{s}) p_{U_2}^{(i)}(\mathbf{s})} \Big) \Big] - R_\mu \Bigg] \\
&+ \lambda_1 \left(\bar{p}_1 - \sum_{i=1}^{N} \left(E_{S_3,S_4} \left[p_{10}^{(i)}(\mathbf{s}) \right] + E_{S_1,S_2} \left[p_{12}^{(i)}(\mathbf{s}) \right] + E \left[p_{U_1}^{(i)}(\mathbf{s}) \right] \right) \right) \\
&+ \lambda_2 \left(\bar{p}_2 - \sum_{i=1}^{N} \left(E_{S_2,S_4} \left[p_{20}^{(i)}(\mathbf{s}) \right] + E_{S_1,S_3} \left[p_{21}^{(i)}(\mathbf{s}) \right] + E \left[p_{U_2}^{(i)}(\mathbf{s}) \right] \right) \right) \\
&+ \epsilon_1^{(i)}(\mathbf{s}) p_{10}^{(i)}(\mathbf{s}) + \epsilon_2^{(i)}(\mathbf{s}) p_{12}^{(i)}(\mathbf{s}) + \epsilon_3^{(i)}(\mathbf{s}) p_{U_1}^{(i)}(\mathbf{s}) + \epsilon_4^{(i)}(\mathbf{s}) p_{20}^{(i)}(\mathbf{s}) + \epsilon_5^{(i)}(\mathbf{s}) p_{21}^{(i)}(\mathbf{s}) + \epsilon_6^{(i)}(\mathbf{s}) p_{U_2}^{(i)}(\mathbf{s}).
\end{aligned}
\tag{C.1}
$$

For $\mathbf{s} \in \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3$, we take partial derivatives of the Lagrangian function, $\mathcal{L}$ with respect to $p_{1m}^{(i)}(\mathbf{s})$, $p_{2n}^{(i)}(\mathbf{s})$, and $p_{U_k}^{(i)}(\mathbf{s})$, $\forall i$ and $\forall \mathbf{s}$, to obtain the respective conditions

$$\gamma_2\mu_2\left(\frac{s_{10}^{(i)}}{A^{(i)}}\right)+(\mu_1-\mu_2+\gamma_1\mu_2)\left(\frac{s_{1m}^{(i)}}{1+s_{1m}^{(i)}p_{1m}^{(i)}(\mathbf{s})}\right)-\lambda_1+\epsilon_{e_1}^{(i)}(\mathbf{s})=0, \quad \text{(C.2)}$$

$$\gamma_2\mu_2\left(\frac{s_{20}^{(i)}}{A^{(i)}}\right)+\gamma_1\mu_2\left(\frac{s_{2n}^{(i)}}{1+s_{2n}^{(i)}p_{2n}^{(i)}(\mathbf{s})}\right)-\lambda_2+\epsilon_{e_2}^{(i)}(\mathbf{s})=0, \quad \text{(C.3)}$$

$$\gamma_2\mu_2\left(\frac{\sqrt{s_{k0}^{(i)}s_{j0}^{(i)}p_{U_j}^{(i)}(\mathbf{s})}+s_{k0}^{(i)}\sqrt{p_{U_k}^{(i)}(\mathbf{s})}}{A^{(i)}\sqrt{p_{U_k}^{(i)}(\mathbf{s})}}\right)-\lambda_k+\epsilon_{e_3}^{(i)}(\mathbf{s})=0, \quad \text{(C.4)}$$

where $e_1 \in \{1,2\}$, $e_2 \in \{4,5\}$ and $e_3 \in \{3,6\}$ take their values based on with respect to which power the derivative is taken. Likewise, for $\mathbf{s} \in \mathcal{S}_4$, and the respective partial derivatives yield

$$\gamma_2\mu_2\left(\frac{s_{10}^{(i)}}{A^{(i)}}\right)+(\mu_1-\mu_2)\left(\frac{s_{10}^{(i)}}{1+s_{10}^{(i)}p_{1m}^{(i)}(\mathbf{s})}\right)$$
$$+\gamma_1\mu_2\left(\frac{s_{10}^{(i)}}{1+s_{10}^{(i)}p_{1m}^{(i)}(\mathbf{s})+s_{20}^{(i)}p_{2n}^{(i)}(\mathbf{s})}\right)-\lambda_1+\epsilon_{e_1}^{(i)}(\mathbf{s})=0, \quad \text{(C.5)}$$

$$\gamma_2\mu_2\left(\frac{s_{20}^{(i)}}{A^{(i)}}\right)+\gamma_1\mu_2\left(\frac{s_{20}^{(i)}}{1+s_{10}^{(i)}p_{1m}^{(i)}(\mathbf{s})+s_{2n}^{(i)}p_{20}^{(i)}(\mathbf{s})}\right)-\lambda_2+\epsilon_{e_2}^{(i)}(\mathbf{s})=0, \quad \text{(C.6)}$$

$$\gamma_2\mu_2\left(\frac{\sqrt{s_{k0}^{(i)}s_{j0}^{(i)}p_{U_j}^{(i)}(\mathbf{s})}+s_{k0}^{(i)}\sqrt{p_{U_k}^{(i)}(\mathbf{s})}}{A^{(i)}\sqrt{p_{U_k}^{(i)}(\mathbf{s})}}\right)-\lambda_k+\epsilon_{e_3}^{(i)}(\mathbf{s})=0. \quad \text{(C.7)}$$

Since the optimal power allocation policy should satisfy the complementary slackness constraints,

$$\begin{aligned} &p_{10}^{(i)}(\mathbf{s})\epsilon_1^{(i)}(\mathbf{s})=0, \quad p_{12}^{(i)}(\mathbf{s})\epsilon_2^{(i)}(\mathbf{s})=0, \quad p_{U_1}^{(i)}(\mathbf{s})\epsilon_3^{(i)}(\mathbf{s})=0, \\ &p_{20}^{(i)}(\mathbf{s})\epsilon_4^{(i)}(\mathbf{s})=0, \quad p_{21}^{(i)}(\mathbf{s})\epsilon_5^{(i)}(\mathbf{s})=0, \quad p_{U_2}^{(i)}(\mathbf{s})\epsilon_6^{(i)}(\mathbf{s})=0, \end{aligned} \quad \text{(C.8)}$$

we can either drop $\epsilon_t^{(i)}(\mathbf{s})$ in each of (C.2)-(C.7), if the corresponding power is positive; or we can replace the equality by a strict inequality, meaning that $\epsilon_t^{(i)}(\mathbf{s})$

is non-zero but its corresponding power is zero. Hence, using the relevant conditions from (C.8) in (C.2)-(C.7), and dropping the dependencies on $\epsilon_t^{(i)}(\mathbf{s})$, we write the conditions for optimality in terms of inequalities instead, which yield (4.21)-(4.26). The inequalities hold with equality if and only if the corresponding power level is positive, and with strict inequality of that power level is zero.

Partial derivatives with respect to the dual variables dictate that the conditions (4.14)-(4.17) are satisfied. Finally, partial derivatives with respect to R_μ yields $\gamma_1 + \gamma_2 = 1$, hence the condition $\gamma_1 = 1 - \gamma_2$.

Vita

Publications

[1] S. Bakim, C. T. Abdallah, E. Schamiloglu. Communicating With Microwave-Propelled Sails. *IEEE Antennas and Propagation Magazine*, 45(4): 111-122, Aug. 2003.

[2] S. Bakim, O. Kaya. Achievable Rates for Two User Cooperative OFDMA. *In 2010 IEEE Global Telecommunications Conference. GLOBECOM 2010*, pages 1-5, Dec. 2010.

[3] S. Bakim, O. Kaya. Cooperative Strategies and Achievable Rates for Two User OFDMA Channels. *IEEE Transactions on Wireless Communications*, 10(12): 4029-4034, Dec. 2011.

[4] S. Bakim, O. Kaya. Optimum Power Control for Transmitter Cooperation in OFDMA Based Wireless Networks. *In Proc. of 2011 IEEE Global*

Communications Conference, Workshop on Multicell Cooperation, Houston, TX, Dec. 2011.

[5] S. Bakim, O. Kaya. Power Control for Two User Cooperative OFDMA Channels. *Submitted to IEEE Transactions on Wireless Communications.*