CROSS-LAYER OPTIMIZATION FOR SPECTRAL AND ENERGY EFFICIENCY

A Dissertation
Presented to
The Academic Faculty

By

Guowang Miao

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
in
Electrical and Computer Engineering

School of Electrical and Computer Engineering
Georgia Institute of Technology
Dec. 2009

CROSS-LAYER OPTIMIZATION FOR SPECTRAL AND ENERGY EFFICIENCY

Approved by:

Dr. Geoffrey Ye Li, Advisor
Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Yorai Wardi
Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Gordon Stüber
Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Xingxing Yu
Professor, School of Mathematics
Georgia Institute of Technology

Dr. Xiaoli Ma
Assistant Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: Sep. 29 2009

To Shunlin Miao – my father.

ACKNOWLEDGMENTS

I'm most grateful to my dissertation advisor, Professor Geoffrey Ye Li, who has taught me a lot and improved my research capabilities significantly. I appreciate his care and concern for my intellectual and personal growth. I would like to thank him for his support, encouragement, and invaluable advice during the course of my Ph.D. study.

I have been most fortunate to have been working with Dr. Nageen Himayat, my Intel supervisor. It was her insight that helped me to initiate many topics in this research. My PhD research has significantly benefited from her valuable guidance and constant encouragement. I would say personally that she is my Ph.D. co-advisor.

Dr. Ananthram Swami has my sincerest gratitude for having taught me so much about doing research and pursuing high-quality academic writings.

I would like to thank Professor Xiaoli Ma, Professor Gordon Stüber, Professor Yorai Wardi, and Professor Xingxing Yu for serving in my dissertation committee. Their broad perspectives and suggestions helped me a lot in refining this dissertation.

I would also like to thank Professor Zhisheng Niu at Tsinghua University for initializing my interest in wireless communications and shaping my research capabilities.

I am thankful for all my labmates at the Information Transmission and Processing Laboratory and all friends at Centergy building for inspiring discussions and helps. I thank all friends at Georgia Institute of Technology. You make this place vivid, warm, and more attractive.

Finally, I would like to thank my parents and my wife. I could not have completed the PhD work without the love and support of Mom and Dad. Dad always educates me to do better, whatever I am working on. I have known from early childhood that there's no best, only better. My wife, Ting Ren, always stands behind me and encourages me to pursue my goal. This dissertation is dedicated to them.

This work was supported by Intel Corporation and the U.S. Army Research Laboratory

under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-20-0011.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The future success of communication networks hinges on the ability to overcome the mismatch between requested *quality of service* (QoS) and limited network resources. Spectrum is a natural resource that cannot be replenished and therefore must be used efficiently. On the other hand, *energy efficiency* (EE) is also becoming increasingly important as battery technology has not kept up with the growing requirements stemming from ubiquitous multimedia applications. This thesis focuses on improving both spectral and energy efficiency from different perspectives. Specifically, because of fading, the qualities of wireless channels vary with both time and user. We use *channel state information* (CSI) to dynamically assign wireless resources to users to improve spectral and energy efficiency.

We first investigate a series of general treatments of exploiting CSI in a distributed way to control the medium access to maximize spectral efficiency for networks with arbitrary topologies and traffic distributions. As the first step, we propose *decentralized optimization for multichannel random access* (DOMRA), which uses local CSI and two-hop static neighborhood information to improve slotted Aloha. DOMRA adapts to the inhomogeneous spatial traffic distribution and achieves performance comparable with the global optimum, which can only be obtained using complete network knowledge. The generic framework developed in DOMRA proved to be very useful in improving cellular networks as well. We develop cochannel interference avoidance (CIA) medium access control (MAC), which is optimized by DOMRA, to mitigate the downlink severe cochannel interference that is usually experienced by cell-edge users. Aloha-based schemes have low channel utilization efficiency because of the collision of entire data frames. Hence we further develop *channel-aware distributed MAC* (CAD-MAC), which avoids collision through signaling negotiation ahead of data transmission. With CAD-MAC, users with better channel states are scheduled in a distributed way. This scheme completely resolves contention of networks with arbitrary topologies. Besides, it achieves throughput close to that using centralized schedulers

and is robust to any channel uncertainty.

Then we address energy-efficient wireless communications while emphasizing *orthogonal frequency multiple access* (OFDMA) systems. We first discover the global optimal energy-efficient link adaptation in frequency-selective channels using the strict quasiconcavity of energy efficiency functions. This link adaptation optimally balances the power consumption of electronic circuits and that of data transmission on each subchannel. The global optimal energy-efficient transmission can be obtained using iterative operations, which may be complex to be implemented in a practical system. Besides, running iterative algorithms consumes additional energy. Hence, using a locally linear approximation, we further develop a closed-form link adaptation scheme, which performs close to the global optimum. Besides, since subchannel allocation in OFDMA systems determines the energy efficiency of all users, we develop closed-form resource allocation approaches that achieve near-optimal performance too. In an interference-free environment, a tradeoff between EE and *spectral efficiency* (SE) exists, as increasing transmit power always improves SE but not necessarily EE. We continue the investigation in interference-limited scenarios and show that since increased transmit power also brings higher interference to the network, SE is not necessarily higher and the tradeoff is reduced. Especially, in interference-dominated regimes, e.g., local area networks, both spectral- and energy-efficient communications desire optimized time-division protocols and the proposed DOMRA, CIA-MAC, and CAD-MAC can be used to improve both spectral and energy efficiency.

CHAPTER 1

INTRODUCTION

1.1 Motivation

Wireless communication systems have experienced tremendous growth, which continues unabated worldwide. The future success of wireless networks hinges on the ability to overcome the mismatch between the requested *quality of service* (QoS) and limited network resources. Spectrum is a natural resource that cannot be replenished and therefore must be used efficiently; that is where the significance of *spectral efficiency* (SE) lies. On the other hand, *energy efficiency* (EE) is also becoming increasingly important for small form factor mobile devices, as battery technology has not kept up with the growing requirements stemming from ubiquitous multimedia applications [1].

Spectral and energy efficiency is affected by all layers of system design, ranging from silicon to applications. The traditional layer-wise approach leads to independent design of different layers and results in high design margins. Cross-layer approaches exploit interactions between different layers and can significantly improve system performance as well as adaptability to service, traffic, and environment dynamics. Cross-layer optimization for throughput improvement has been a popular research theme [2, 3, 4]. Recent efforts have also been undertaken to tackle energy consumption at all layers of communication systems, from architectures [5, 6, 7] to algorithms [8, 9, 10].

The *physical* (PHY) layer plays a very important role in wireless communications due to the challenging nature of the communication medium. The PHY layer deals with data transmission over wireless channels and consists of *radio frequency* (RF) circuits, modulation, power control, channel coding units, etc. Traditional wireless systems are built to operate on a fixed set of operating points [11], e.g., no power adaptation. This results in excessive energy consumption or pessimistic data rate for peak channel conditions. Hence, a set of PHY parameters should be adjusted to adapt the actual user requirements (e.g.,

throughput and delay) and environments (such as shadowing and frequency selectivity) to trade off energy efficiency and spectral efficiency. As wireless is a shared medium, communication performance and energy consumption are affected not only by the layers comprising the point-to-point communication link, but also by the interaction between the links in the entire network. Hence, a system approach is required. The *medium access control* (MAC) layer ensures that wireless resources are efficiently allocated to maximize network-wide performance metrics while maintaining user QoS requirements. Here, pessimistic medium access strategies that allocate wireless resources to assure worst-case QoS may hurt network spectral and energy efficiency. In distributed access schemes, MAC should be improved to reduce the number of wasted transmissions that are corrupted by interference from other users, while in centralized access schemes, efficient scheduling algorithms should exploit the variations across users to maximize the overall network performance. The MAC layer manages wireless resources for the PHY layer and they both directly impact overall network performance and energy consumption. We focus on cross-layer optimization across the PHY and MAC layers to improve wireless spectral and energy efficiency. Furthermore, *orthogonal frequency division multiplexing* (OFDM) becomes a key modulation scheme for next-generation broadband wireless standards [12, 13]. OFDM-based systems are traditionally used for combating frequency-selective fading. From a resource allocation perspective, multiple channels in OFDM systems have the potential for more efficient MAC design since subcarriers can be assigned to different users [14]. Furthermore, adaptive power allocation on each subcarrier can be applied for further improvement [15]. Hence, we will emphasize the cross-layer design for OFDM systems to enhance spectral and energy efficiency.

1.2 Literature Review

In this section, we review state-of-the-art techniques for cross-layer optimizations of wireless networks to improve spectral and energy efficiency, including channel-aware medium

access control, energy-efficient link adaptation, energy-efficient resource allocation, and interference-aware energy-efficient communications.

1.2.1 Cross-Layer Optimization for Spectral-Efficient Communications

Because of fading, the quality of a wireless channel varies with both time and user. Wireless is a shared medium and communication performance is affected not only by individual communication links but also by the interaction among the links in the entire network. To fully exploit network resources, channel-aware medium access schemes have been proposed to adaptively transmit data and dynamically assign wireless resources based on CSI. The key idea of channel-aware medium access control is to schedule a user with favorable channel conditions to transmit with optimized link adaptation according to CSI [16, 14, 15, 17]. By exploiting the channel variations across users, channel-aware medium access control substantially improves network performance through multiuser diversity, whose gain increases with the number of users [17, 16].

1.2.1.1 Channel-Aware Random Access

Medium access can be either centralized or distributed. With a central controller, the best performance is obtained by scheduling the user with the best channel state [14, 15, 18]. However, CSI feedback incurs huge overhead, especially for networks with a large number of users at high mobility, which results in poor network scalability. To reduce CSI feedback, distributed approaches are preferred.

Random access algorithms provide the means to share network resources among users under distributed control. Traditional contention based random access methods include pure, slotted, and reservation Aloha schemes, *carrier sense multiple access* (CSMA) and CSMA with collision avoidance schemes, *multiple access with collision avoidance for wireless* (MACAW) schemes, and so on [19, 20]. These MAC approaches do not use CSI. Hence, when MAC decides to transmit a frame, the channel may be in a deep fade. On the other hand, MAC may not transmit even though the channel is in a good state, which

wastes channel resources. Recently, opportunistic random access schemes have been studied in [21, 22, 23, 24, 25, 26, 27, 28, 29, 30] and the references therein to use CSI for performance improvement. With opportunistic random access, each user exploits its own CSI to decide the contention behavior and users with better channel states have higher contention probabilities. A channel-aware Aloha is proposed in [21] to improve the uplink access contention for cellular type networks; users transmit data whenever their channel gains are above pre-determined thresholds. Since the channel state is random, the transmission is randomized. This scheme is then further studied in [22, 23, 24, 25] in different scenarios. In [26], users and the *base station* (BS) negotiate through mini time slots before the data transmission period such that the user with the best channel condition always wins the contention and transmits data. A channel-aware multicarrier random access scheme has been proposed in [23], where each user selects some subchannels with the best channel power gains for data transmission. Inspired by [21], it is proposed in [27] that each user in a cellular network sends request packets when the channel fading level exceeds a predetermined threshold, after which the BS processes downlink transmissions. Although the thresholds in [27] are chosen to optimize downlink throughput, the proposed scheme actually reduces uplink request collisions and hence also deals with random access. In [28], based on decentralized CSI, a general expression for the transmission probability that may depend on the channel and the physical layer implementation is given, and the transmission probability is optimized to achieve maximum stable throughput in the MAC layer. Through slotted Aloha, a reservation-based MAC scheme is found in [29] to maximize the overall throughput. The capacity of slotted Aloha is analyzed and the optimal transmission probabilities are obtained for a multi-packet reception MAC model [30]. One observation is that all these opportunistic random access schemes are for wireless networks where users transmit to a common receiver, e.g., a BS. However, this scenario does not fit many wireless communication environments, such as sensor [31], *ad hoc* [32], and mesh networks [33].

1.2.1.2 Cochannel Interference

Cellular networks are becoming increasingly interference limited as more users need to share the same spectrum to achieve high-rate multimedia communication. In typical cellular systems, *cochannel interference* (CCI) is one of the major factors limiting system spectral efficiency, especially as these systems move toward aggressive frequency reuse scenarios [12, 34]. While the overall system spectral efficiency may improve with aggressive frequency reuse, the performance of cell-edge users degrades substantially.

A commonly used method to avoid CCI is to assign different sets of channels to neighboring cells [35, 36] and a good summary of channel assignment can be found in [37]. One recent popular approach to reducing interference for cell-edge users is through fractional frequency reuse (FFR) [38, 12, 39, 40, 41]. With FFR, a lower frequency reuse is specified for users at cell edges, while full frequency reuse is applied for those at cell centers. This improves the throughput of cell-edge users since they experience lower levels of interference. To further improve frequency reuse efficiency, CCI can be mitigated by advanced digital signal processing techniques [42,43,44,45]. For example, in [45], various multiuser detection and CCI suppression schemes have been studied when *mobile terminals* (MTs) are equipped with multiple antennas. However, these techniques have high complexity and therefore result in high costs for MTs. For downlink transmission, CCI can be mitigated by joint pre-processing and encoding techniques among BSs [46], [47], or avoided by using cooperative scheduling among BSs [48], both of which require a lot of instantaneous information exchange. Recently, contention-based schemes have also been developed for CCI avoidance in addition to an intracellular centralized MAC protocol. In [49], each MT or BS keeps on broadcasting busy-tone signals located at the mini-slot of every data frame to prevent potential interferers from transmitting, and every BS or MT must listen to the mini-slots before transmission. This scheme effectively avoids CCI without considering fairness among users, and a group of greedy users may keep on broadcasting busy-tone signals, which always prevent others from transmitting.

1.2.2 Cross-Layer Optimization for Energy-Efficient Communications

As wireless is a shared medium, device energy efficiency is affected not only by the layers composing the point-to-point communication link, but also by the interaction between the links in the entire network. Hence, a systematic approach, including both transmission and multi-user resource management, is required for energy-efficient wireless communications.

1.2.2.1 Energy-Efficient Transmission

The quality of wireless channel varies with time and frequency. Therefore, link adaptation can be used to improve transmission performance. With link adaptation, modulation order, coding rate, and transmit power can be selected according to CSI.

Earlier research on link adaptation focuses on power allocation to improve channel capacity. Optimal power allocation for frequency-selective channels has been implied [50]. Here, the highest data rate on a bandlimited channel is achieved when the total received signal power at each frequency, consisting of channel noise and desired signal component, is a constant. Power allocation and bit-loading algorithms for OFDM are summarized in Chapter 3 of [51]. The terminology, adaptive modulation, was first used in [52] even though work on adaptive modulation had been reported before in [53].

In addition to spectral efficiency improvement, energy efficiency is becoming increasingly important for mobile communications due to the slow progress of battery technology [1] and the growing requirements of anytime and anywhere multimedia applications. With sufficient battery power, link adaptation can be geared toward peak performance delivery. However, with limited battery capacity, link adaptation could be adapted toward energy conservation to minimize battery drain. Energy-efficient communication also has the desirable benefit of reducing interference to other co-channel users as well as lessening environmental impacts, e.g., heat dissipation and electronic pollution.

Information theorists have studied energy-efficient transmission for at least two decades [54, 55]. The work in [54] defines reliable communication under a finite energy constraint in terms of the capacity per unit energy, which is the maximum number of bits that can be

transmitted per unit energy. This definition ensures that for any transmission rates below the capacity per unit energy, error probability decreases exponentially with the total energy. It is also shown that the capacity per unit energy is achieved using an unlimited number of degrees of freedom per information bit, e.g., with infinite bandwidth [56] or long-duration regime communications [57]. For example, the lowest order modulation should always be used while accommodating the delay constraint [57] to minimize energy consumption. The information-theoretic results derived in [56, 57] focus only on transmit power when considering energy consumption during transmission. Typically, a device will incur additional circuit power during transmission, which is relatively independent of the transmission rate [58, 59, 60]. Thus, a fixed cost of transmission is incurred that must be accounted for in optimizing energy consumption. Figure 1.1 shows the transmit and circuit energy consumptions when different time durations are used for transmitting one bit. The method to transmit with the longest duration is no longer the best since circuit energy consumption increases with transmission duration. Considering the impact of circuit power, the focus will shift toward using optimization theory framework for determining energy-optimal link settings.

The energy dissipation consisting of both transmitter electronics and RF output is studied in [58], and several energy-minimization techniques, including modulation and multiple access protocols, are derived for short-range asymmetric micro-sensor systems. It is shown that a high order modulation may enable energy savings compared with binary modulation for some short-range applications by decreasing the transmission time. In [59], these ideas are extended to a detailed energy consumption analysis specifically for both uncoded and coded *M-ary quadrature amplitude modulation* (M-QAM) and *multiple frequency shift keying* (MFSK) in additive white Gaussian channels. Therefore, energy-efficient transmission is formulated to find a trade-off among transmission energy, circuit energy, and transmission time. Similarly, a steepest descent gradient algorithm is designed in [61] to search the optimal rate that minimizes the average power consumption subject to a constraint on

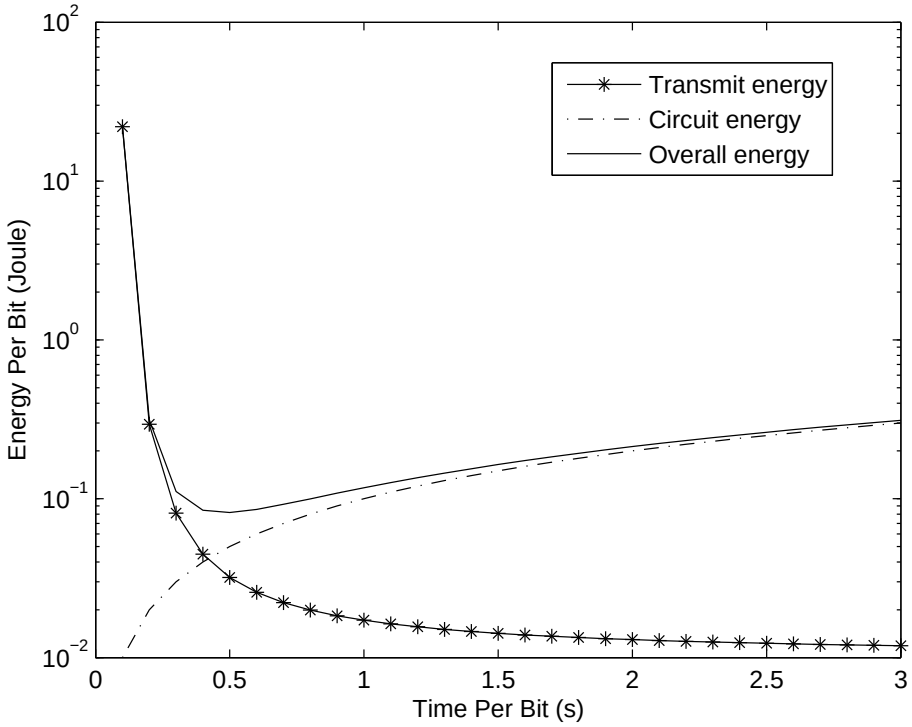


Figure 1.1. Relationship between energy consumption and symbol duration

average throughput.

1.2.2.2 Energy-Efficient Resource Management

Due to limited wireless resources, intricate performance trade-offs exist between an individual user and the whole network. The exploitation of diversity across all users will further reduce overall network energy consumption. Wireless resources can be managed in different domains to improve network energy efficiency.

In the time domain, e.g., in a *time-division multiple access* (TDMA) network, the channel medium is shared through time division. Each user tends to extend its transmission time to save energy and contradicts the intention of energy savings of other users. Thus the allocation of time duration among all users is critical in determining network energy efficiency. As the modulation order determines data rate and thus the time for transmitting a certain amount of information, finding the optimal slot length for each user is thus equivalent to determining its corresponding constellation size [62]. A centralized resource allocation

scheme is considered in [63]. The scheme assigns time slots to all users and specifies the transmission parameters of each user for energy-efficient communications. To make the resource management scheme applicable, the scheduling is partitioned into a design phase and a run-time phase. In the design-time phase, energy-performance representation can be derived for each user to capture the relevant energy and performance trade-offs. In the run-time phase, a fast greedy algorithm is used to tune the operating points to further improve energy efficiency.

Since wireless is broadcast, the transmission of one user interferes with neighboring users and reduces their energy efficiency. However, users can gain in energy efficiency if cooperation among neighboring users is allowed. Hence, spatial-domain resource management is important to manage the behaviors of users at different spatial locations. On the other hand, cooperation requires signalling overhead and consumes additional energy. Cooperation can also cause transmission delay that may impact throughput adversely and thus hurt energy efficiency. However, delay can be exploited for energy-efficient link adaptation, as extending transmission duration may improve energy efficiency. It has been observed that significant energy savings can be achieved and the savings grow almost linearly with distance when either transmitter or receiver cooperation is allowed [64]. Furthermore, it is also observed that cooperation can even reduce delay within a certain transmission ranges since cooperation enables higher order modulation and increases data rate [64]. Similarly, receiver cooperation is exploited in [65] and significant energy savings can be observed. Besides transmitter and receiver cooperation, relay cooperation is also effective in improving network energy efficiency. Since the energy for reliable data transmission grows exponentially with distance [37], it is more energy efficient to send data using several shorter intermediate hops than using a long hop if the energy to compute the route is negligible [66]. However, relay incurs delay and energy consumption of relay nodes. Therefore, in some scenarios, it is advantageous to use longer hops [67]. Hence, the optimal selection of relay nodes is a trade-off between source-node performance and relay cost to enhance

overall network energy efficiency.

While extensive efforts have been undertaken to improve energy-efficient resource management in both the spatial and time domains, little effort has been devoted to frequency domains. In the frequency domain, while increasing transmission bandwidth improves energy efficiency, the entire system bandwidth can not be allocated exclusively to one user in a multi-user system since this may hurt the energy efficiency of other users as well as that of the overall network. Hence, frequency-domain resource management is critical in determining overall network energy efficiency. Frequency selectivity of wideband wireless channels further accentuates this necessity.

1.2.2.3 Interference-Aware Energy-Efficient Communications

As more users need to share the same spectrum for wideband multimedia communications and cellular networks move toward aggressive full-frequency reuse scenarios [12], the performance of wireless cellular networks is heavily impaired by interference. This motivates the use of multi-cell power control optimization for interference management [12, 68, 69, 70]. Meanwhile, power optimization is also important for extending the battery life of mobile devices. Although power optimization plays a pivotal role in both interference management and energy utilization, little research has addressed their joint interaction. An implicit discussion can be found in [71], which summarizes existing approaches that address either throughput or energy efficiency separately in the context of power control for CDMA networks. In our work, we will address this joint limitation and investigate energy-efficient power optimization for OFDM communications in interference-limited environments.

1.3 Our Approaches and Thesis Outline

The major goal of this research is to investigate novel cross-layer transmission and resource management algorithms to significantly improve user experience, system spectral efficiency, and energy efficiency. By exploiting CSI of different users, this research leads

to integrated algorithms to utilize the spectrum and energy resources fairly and efficiently.

As the first step, we study an optimal channel-aware slotted Aloha in Chapter 2. A user transmits a packet when its channel gain is above a threshold and the threshold controls the contention probability. In this way, users with better channel states have higher probabilities of contention success. We provide a methodology guiding the selection of the threshold and link transmission to maximize the network throughput while assuring proportional fairness among all users. Using this methodology, we design decentralized optimization for multichannel random access (DOMRA). DOMRA uses local CSI and two-hop static neighborhood information to adapt to the inhomogeneous spatial distribution of traffic flows. It achieves performance comparable with the global optimum, which can only be obtained using complete network knowledge.

The generic framework developed for medium access control in DOMRA proved to be very useful in improving cellular networks as well. We developed *cochannel interference avoidance MAC* (CIA-MAC) in Chapter 3 to deal with the downlink severe cochannel interference that is usually experienced by cell-edge users. Here BSs producing severe interference transmit randomly and the randomization is optimized by DOMRA. One major issue is the detection of severe interference. We design two simple detectors for different implementation requirements and both distinguish severe interferers effectively. CIA-MAC requires only minor changes of existing cellular protocols to earn significant performance gain.

Aloha-based schemes have low channel utilization efficiency because of the collision of entire data frames. To further improve performance, in Chapter 4, we develop *channel-aware distributed MAC* (CAD-MAC), which avoids collision through signaling negotiation ahead of data transmission. We notice that the backoff-after-collision approach in traditional multiple access schemes like CSMA/CA ignore channel variations and deferring transmission without considering the variations might result in data transmission in deep

fades. Hence, in our design, each frame is divided into contention and transmission periods and conflicts are optimally resolved in the Aloha-based contention period using the methodology of DOMRA. With this design, users with better channel states have higher probabilities of contention success and the data transmission experiences better channels. Further proof shows that this scheme completely resolves the contention of networks with arbitrary topologies. Besides, it achieves throughput close to that using centralized schedulers and is robust to any channel uncertainty.

Shannon theory indicates that it is desirable to transmit a packet over a longer period of time to save transmit energy. However, when circuit energy, the one consumed by electronic components except amplifiers, is considered, this is no longer the case since the circuit energy increases with transmitting duration. Hence, an optimal transmitting duration, determined by the selected modulation order and power allocation, needs to be found to balance the transmit and circuit energy consumption. The frequency selectively in OFDM systems further complicates the problem since different modulation orders and amounts of power can be applied on different subcarriers. In Chapter 5, we show that there exists a unique global optimal energy-efficient OFDM link adaptation because of the strict quasi-concavity of energy efficiency function. A subcarrier is used only if it improves the overall energy efficiency. Using first order information, the optimal power allocation is found to be a dynamic water-filling scheme that adjusts both the overall transmit power and its allocation according to the circuit power and the states of all subchannels to minimize the overall energy consumption.

The approaches in Chapter 5 require iterative operations to obtain the global optimal link adaptation. They may be complex to be implemented in a practical system. Besides, running iterative algorithms consumes additional energy. We notice that if the energy efficiency function can be linearized, the solution would be simpler. Hence, in Chapter 6, we measure energy efficiency as a function of average user throughput and power consumption rather than as a function of instantaneous rates and power. Then using a locally linear

approximation, we develop an instantaneous link adaptation scheme in a closed form that performs close to the global optimum. Besides, since subcarrier allocation in an *orthogonal frequency division multiple access* (OFDMA) system determines the energy efficiency of all users, we further develop closed-form resource allocation approaches that also achieve near-optimal performance.

In Chapter 5, we have also observed that in an interference-free environment, a tradeoff between *energy efficiency* (EE) and *spectral efficiency* (SE) exists, as increasing transmit power always improves SE but not necessarily EE. What will happen in a multi-user environment where each user interferers with all others? In Chapter 7, we continue the investigation in interference-limited scenarios and design both cooperative and noncooperative energy-efficient power optimizations. According to our study, since increased transmit power also brings higher interference to the network, SE is not necessarily higher and the tradeoff is reduced. Especially, in interference-dominated regimes, e.g., local area networks, both spectral- and energy-efficient communications desire optimized time-division protocols and the proposed DOMRA, CIA-MAC, and CAD-MAC can be used to improve both spectral and energy efficiency.

CHAPTER 2

DECENTRALIZED OPTIMIZATION FOR MULTICHANNEL RANDOM ACCESS

In this chapter, we consider schemes for decentralized cross-layer optimization of multichannel random access by exploiting local channel state and traffic information. In the network we are considering, users are not necessarily within the transmission ranges of all others; therefore, when a user is transmitting, it may only interfere with some users, which is different from most existing channel aware Aloha schemes. Besides, we also consider complicated traffic distribution, e.g. each user may choose to send packets to or receive packets from different users simultaneously. We develop *decentralized optimization for multichannel random access* (DOMRA). DOMRA consists of three steps: neighborhood information collection, transmission control of the MAC layer based on the instantaneous channel state information, and power allocation for each traffic flow on each subchannel.

First we introduce the physical and MAC layers of the system in Section 4.1. In Section 2.2, we describe the transmission policy and formulate the problem. Then in Section 2.3, we decompose the cross-layer optimization into two sub-problems and provide suboptimal solutions. Finally, we demonstrate the performance improvement of the proposed scheme by computer simulations in Section 3.4 and summarize this chapter in Section 3.5.

2.1 System Description

Consider multichannel wireless networks. The whole band is divided into K subchannels. All channels between pairs of users are assumed to be reciprocal, i.e. when no interference exists, User A can receive signal from User B if and only if User B can receive signal from User A with the same channel gain. However, the interference environments at Users A and B may be different since they are at different locations. Each user has knowledge of its own CSI and makes independent transmission control decisions, including whether to transmit

given the CSI, what data rate to use and where to transmit, etc. Each user applies the same transmission control policy. In order to avoid onerous signalling burden, no communication pair has instantaneous cooperation, such as exchange of CSI, transmit power, or subchannel selections.

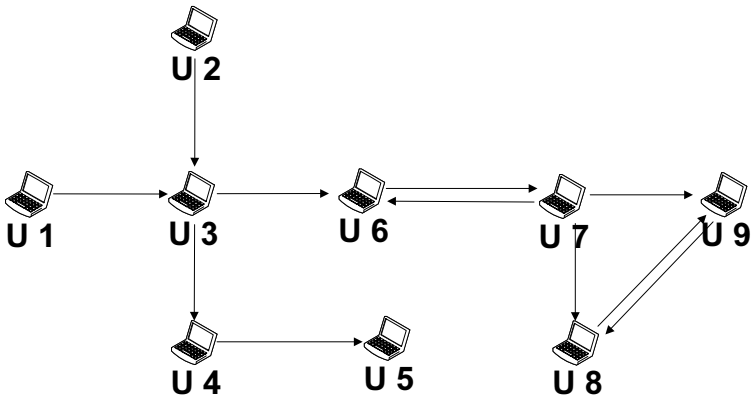


Figure 2.1. Network architecture example

All users are not necessarily within the transmission ranges of the others, which means that some users may not be able to receive packets from others due to weak received signal power. For simplicity, we assume those that can communicate with each other experience isotropic channels, i.e. channel power gains of different links are independent and identically distributed with probability density function, $f(h)$, and distribution function, $F(h)$. No capture is assumed for signal reception, i.e., the receiver cannot receive any signal successfully if any of its interfering neighbors, which are within the transmission range of the receiver, is transmitting simultaneously. A user can not transmit and receive simultaneously on the same subchannel; however, it may transmit on a set of subchannels and receive on a different set of subchannels at the same time. Each user may choose to send packets to or receive packets from different users on different channels, and we assume that the links that carry traffic are backlogged, i.e., they always have packets to transmit.

During transmission, each user is subject to both average and instantaneous power constraints [72]. The average power constraint is due to heat accumulation and overall power consumption, while the instantaneous power constraint comes from the limited linear range

of amplifiers. Two power allocation policies will be considered. In the first one, called channel inversion, each user transmits with just sufficient power to keep the received power constant so that the signal can be reliably detected. In the second, called adaptive modulation and power allocation, each user can vary both the modulation and transmit power during each transmission time slot to maximize throughput.

2.2 Problem Formulation

In this section, we describe our wireless network model, and propose a channel aware multi-channel random access scheme. The characteristics of the proposed scheme are analyzed, after which a criterion for cross-layer design is provided.

Denote the wireless network as a directed graph $G(\mathcal{V}, \mathcal{E}, \mathcal{L})$, where $\mathcal{V}$, $\mathcal{E}$, and $\mathcal{L}$ are the set of active users, the set of all links over all K subchannels, and the set of links available for communication. We denote $\mathcal{N}_i$ as the interfering neighbor set of User i. Each user may choose to send packets to or receive packets from several users, and $\mathcal{T}_i$ denotes the set of users receiving packets from i and $\mathcal{S}_j$ the set of users sending packets to j.

Figure 2.1 shows an example topology of a wireless network. The users are on a grid with unit spacing, and the transmission range is $\sqrt{2}$. The set of links available for communication is $\mathcal{L} = \{(1,3), (1,2), (1,4), (2,3), (2,6), (3,4), (3,5), (3,6), (4,6), (4,5), (5,6), (5,7), (5,8), (6,7), (6,8), (7,8), (7,9), (8,9)\}$. The arrows show the traffic flows in the network. For example, since $(4,6) \in \mathcal{L}$, any transmission by Users 4 or 6 will be received by the other though they may not have packets to send to each other. So Users 4 and 6 constitute an interfering pair and they interferer packet reception of each other. Observing User 3, it is easy to see that $\mathcal{T}_3 = \{4,6\}$, $\mathcal{S}_3 = \{1,2\}$, while $\mathcal{N}_3 = \{1,2,4,5,6\}$.

Slotted Aloha is a typical random access scheme. In slotted Aloha, the MAC layer makes transmission decisions based on the buffer occupancy and QoS requirement, and does not utilize the knowledge of the physical layer at all. Hence, when the MAC decides to transmit a frame, the channel may be in deep fade, but the physical layer still carries

out the transmission, and causes a waste of power. On the other hand, the MAC layer may decide not to transmit even though the channel power gain is high, because it does not have this information from the PHY layer; this leads to wasted opportunity. With channel knowledge, the sender will transmit only when the channel power gain is above a certain threshold [1]. Therefore, we propose the following *decentralized optimization for multichannel random access* (DOMRA).

DOMRA: *User i ($i \in \mathcal{V}$) decides to send packets to User j on subchannel k when the following conditions are satisfied: 1). User i has packets to send to j, $j \in \mathcal{T}_i$; 2). on subchannel k, link (i, j) has the best channel power gain, $h_{(i,j)_k} = \max_{l \in \mathcal{T}_i}\{h_{(i,l)_k}\}$; 3). the channel power gain is above a threshold, $h_{(i,j)_k} \geq \overline{H}_{(i,j)_k}$, where $\overline{H}_{(i,j)_k}$ is predetermined for link $(i, j)_k$. The transmission is then optimized according to $\overline{H}_{(i,j)_k}$, CSI and capability constraints.*

In DOMRA, each user transmits on the link with the best channel power gain provided that the gain is above a predetermined threshold. Proper choice of thresholds $\{\overline{H}_{(i,j)_k} | (i, j)_k \in \mathcal{E}\}$ and data transmission rates of all traffic flows, i.e. power allocation, will be determined in the following paragraphs so that overall network performance is optimized from certain perspectives.

As pointed out before, while many existing channel aware schemes such as [21, 26, 22, 23], assume that each user has only one traffic flow to send and is within the transmission range of all other users, the DOMRA will provide solutions to networks in which users are not necessarily within the transmission ranges of all other users, and each user could send packets to or receive packets from different users simultaneously on different subchannels.

[1] Channel gains may be inferred either due to CSI feedback or via channel reciprocity.

2.2.1 MAC Layer Analysis

According to the above transmission policy and the homogeneity assumption, the probability of a transmission on link $(i, j)_k \in \mathcal{E}$ is given by

$$p_{(i,j)_k} = \frac{1}{|\mathcal{T}_i|}\left(1 - F^{|\mathcal{T}_i|}(\overline{H}_{(i,j)_k})\right) \tag{2.1}$$

where $|\cdot|$ denotes the number of elements in the respective set. The proof of (2.1) is given in Appendix A.1.

The probability that User i transmits on subchannel k is

$$p_{i_k} = \sum_{j\in\mathcal{T}_i} p_{(i,j)_k} = \sum_{j\in\mathcal{T}_i} \frac{1}{|\mathcal{T}_i|}\left(1 - F^{|\mathcal{T}_i|}(\overline{H}_{(i,j)_k})\right). \tag{2.2}$$

Hence, the throughput on link $(i, j)_k$ is

$$T_{(i,j)_k} = R_{(i,j)_k} p_{(i,j)_k}(1 - p_{j_k}) \prod_{a\in\mathcal{N}_j, a\neq i}(1 - p_{a_k}), \tag{2.3}$$

where $R_{(i,j)_k}$ is the average data rate given that the user has decided to transmit on link $(i, j)_k$, and depends on the modulation and power allocation policy. $(1 - p_{j_k}) \prod_{a\in\mathcal{N}_j, a\neq i}(1 - p_{a_k})$ is the probability that neither user j nor its neighboring users except user i will transmit on subchannel k, which means successful transmission on link $(i, j)_k$.

For example, in Figure 2.1, the transmission from User 3 to User 6 on subchannel k succeeds only when neither User 6 nor his neighbors excluding User 3, i.e., users in $\mathcal{N}_6\backslash\{3\} = \{2, 4, 5, 7, 8\}$, transmit. Hence, the throughput from User 3 to 6 on subchannel k is $T_{(3,6)_k} = p_{(3,6)_k}(1 - p_{6_k})(1 - p_{2_k})(1 - p_{4_k})(1 - p_{5_k})(1 - p_{7_k})(1 - p_{8_k})R_{(i,j)_k}$.

2.2.2 Physical Layer Analysis

The average transmit power on link $(i, j)_k$ is the average of transmit power over all time slots, whether or not transmission happens on this link. According to the ergodicity of the channel, it is the average of transmit power over all channel states. Hence, we have

$$\mathbf{E}\{P_{(i,j)_k}\} = \int_0^\infty \Pr\{H_{(i,j)_k} = h, \text{and User } i \text{ transmits on } (i, j)_k\} P_{(i,j)_k}(h)\mathrm{d}h$$

$$= \int_0^\infty \Pr\{H_{(i,j)_k} = h\} \Pr\{\text{ User } i \text{ transmits on } (i,j)_k | H_{(i,j)_k} = h\} P_{(i,j)_k}(h) \mathrm{d}h$$
$$= \int_{\overline{H}_{(i,j)_k}}^\infty f(h) F^{|\mathcal{T}_i|-1}(h) P_{(i,j)_k}(h) \mathrm{d}h = \frac{1}{|\mathcal{T}_i|} \int_{\overline{H}_{(i,j)_k}}^\infty P_{(i,j)_k}(h) \mathrm{d}F^{|\mathcal{T}_i|}(h), \tag{2.4}$$

where $\mathbf{E}\{\}$ denotes expectation, $P_{(i,j)_k}(h)$ is the transmit power on link $(i,j)_k$ when the channel has power gain h and it depends on modulation and power allocation policy. For example, in order to achieve a constant *signal-to-noise ratio* (SNR) at the receiver, $P_{(i,j)_k}(h)$ is allocated such that $P_{(i,j)_k}(h) = \frac{P_r}{h}$, where P_r is the received power level satisfying the SNR requirement. According to the average power constraint, we have

$$\sum_{j \in \mathcal{T}_i, k=1,\ldots,K} \mathbf{E}\{P_{(i,j)_k}\} \leq P_a, \qquad \forall i, j \in \mathcal{V}. \tag{2.5}$$

In existing channel access protocols, there are usually several subchannels to be selected for utilization. For example, the IEEE 802.11b physical layer [73] has 14 subchannels, 5 MHz apart in frequency, all of which have the same transmission capability. However, typically there is only one single RF chain, and the peak constraint on the instantaneous transmit power has to be satisfied for the total combined transmission. We have the instantaneous power constraint

$$\sum_k \left(\max_{h,j} P_{(i,j)_k}(h) \right) \leq P_m, \qquad \forall i, j \in \mathcal{V}. \tag{2.6}$$

Given power allocation $P_{(i,j)_k}(h)$, the achieved average data rate given that a user has decided to transmit on link $(i,j)_k$ is

$$\begin{aligned} R_{(i,j)_k} =& \mathbf{E}\Big\{ R(\eta(h)) | \text{User } i \text{ transmits on } (i,j)_k \Big\} \\ =& \int_{\overline{H}_{(i,j)_k}}^\infty R(\eta(h)) \Pr\{H_{(i,j)_k} = h | \text{User } i \text{ transmits on } (i,j)_k\} \mathrm{d}h \\ =& \int_{\overline{H}_{(i,j)_k}}^\infty R(\eta(h)) \frac{A}{B} \mathrm{d}h, \end{aligned} \tag{2.7}$$

where

$$A = \Pr\{H_{(i,j)_k} = h, \text{User } i \text{ transmits on } (i,j)_k\} = f(h) F^{|\mathcal{T}_i|-1}(h), \tag{2.8}$$

and

$$\begin{aligned} B =& \Pr\{\text{User } i \text{ transmits on } (i,j)_k\} \\ =& \int_{\overline{H}_{(i,j)_k}}^{\infty} \Pr\{H_{(i,j)_k} = g, \text{User } i \text{ transmits on } (i,j)_k\} \mathrm{d}g \\ =& 1 - F^{|\mathcal{T}_i|}(\overline{H}_{(i,j)_k}). \end{aligned} \tag{2.9}$$

Hence,

$$R_{(i,j)_k} = \frac{\int_{\overline{H}_{(i,j)_k}}^{\infty} R(\eta(h)) \mathrm{d}F^{|\mathcal{T}_i|}(h)}{1 - F^{|\mathcal{T}_i|}(\overline{H}_{(i,j)_k})}, \tag{2.10}$$

where $\eta(h) = \frac{hP_{(i,j)_k}(h)}{N_o W/K}$ is the received SNR, N_o is noise spectral density, W is the total system bandwidth, and $R(\eta)$ is the instantaneous data rate when channel has SNR η.

If channel capacity is achieved in *additive white Gaussian noise* (AWGN) channels [2], $R(\eta) = W \log_2(1 + \eta)$. Assuming continuous-rate M-QAM and given the *bit-error rate* (BER) requirement, $R(\eta)$ can be expressed as $R(\eta) = W \log_2(1 + \frac{3\eta}{-2\ln(5BER)})$ according to [74]. It is easy to see that in both cases, $R(\eta)$ is strictly concave in η. In general, we assume that $R(\eta)$ is continuously differentiable with first order derivative $R'(\eta)$ positive and strictly decreasing in η.

2.2.3 Criterion for Cross-Layer Design

When optimizing multi-user networks, we have to take both overall network throughput and fairness into consideration. A very commonly discussed fairness criterion is max-min fairness [75]. When max-min fairness is achieved, the throughput of a certain link can not be increased without simultaneously decreasing the throughput of another link which already has smaller throughput. Usually, max-min fairness just implies to equal sharing of channel resources on each link, which compromises the overall throughput of the wireless network a lot since different links usually have different transmission conditions. Hence, we consider proportional fairness, the objective of which is to maximize the product of

[2] In slow fading channels, channel varies slightly within each packet. With sufficiently long packet length, ideal coding can be applied to achieve channel capacity

throughput of all links, or the geometric average [76]. As pointed out in [77], a vector of throughputs $T = (T_1, T_2, \cdots, T_n)$ is proportionally fair if it satisfies required constraints, and for any other feasible vector $\overline{T}$, the aggregate of proportional changes is non-positive, i.e. $\sum_{i=1}^{n} \frac{\overline{T}_i - T_i}{T_i} \leq 0$. Some analysis has been given in [76] from a game-theoretic standpoint and it is shown that a strategy achieving proportional fairness satisfies certain axioms of fairness and is a Nash arbitration strategy [78]. With proportional fairness, the network will be operated at a Pareto equilibrium, which corresponds to the situation where no user can improve its throughput without affecting at least one user adversely.

Denote transmission control of the whole network as $C = \{\overline{\mathcal{H}}, \mathcal{P}\}$, where $\overline{\mathcal{H}}$ is the set of predetermined channel power gain thresholds and $\mathcal{P}$ is the set of power allocation policies. With the constraints in (2.5) and (2.6), the optimal configuration of the whole network, $C^* = \{\overline{\mathcal{H}}^*, \mathcal{P}^*\}$, that achieves proportional fairness among all subchannels carrying traffic flows will be

$$C^* = \arg\max_{\{\overline{\mathcal{H}}, \mathcal{P}\}} \sum_{(i,j)_k \in \mathcal{E}, j \in \mathcal{T}_i} \ln(T_{(i,j)_k}), \tag{2.11a}$$

subject to

$$\sum_{j \in \mathcal{T}_i, k=1,\ldots,K} \frac{1}{|\mathcal{T}_i|} \int_{\overline{H}_{(i,j)_k}}^{\infty} P_{(i,j)_k}(h) \mathrm{d}F^{|\mathcal{T}_i|}(h) \leq P_a, \tag{2.11b}$$

and

$$\sum_{k} \left(\max_{h,j} P_{(i,j)_k}(h) \right) \leq P_m, \tag{2.11c}$$

where throughput $T_{(i,j)_k}$ is given by (2.3). Denote utility $U_{(i,j)_k} = \ln(T_{(i,j)_k})$. Problem (2.11) aims to maximize overall network utility subject to individual power limits.

2.3 Decentralized Optimization

In the previous section, we have discussed a criterion for cross-layer design. The optimization of (2.11) depends on the threshold configuration, $\overline{\mathcal{H}}$, power allocation, $\mathcal{P}$, and modulation policy. The global optimization of the problem is difficult and computationally expensive, and requires complete network knowledge for each user. Therefore, in

this section, we find a suboptimal solution, which only needs decentralized neighborhood information.

From (2.11), we have

$$C^* = \arg\max_{\{\overline{\mathcal{H}},\mathcal{P}\}} \sum_{(i,j)_k\in\mathcal{E}, j\in\mathcal{T}_i} \left(\ln\left(p_{(i,j)_k}(1-p_{j_k}) \prod_{a\in\mathcal{N}_j, a\neq i}(1-p_{a_k})\right) + \ln\left(R_{(i,j)_k}\right)\right). \tag{2.12}$$

(2.12) reveals two ways to improve the overall system performance. One way is to reduce the probability of collisions in the whole network, whose effect is captured by the term $p_{(i,j)_k}(1-p_{j_k})\prod_{a\in\mathcal{N}_j,a\neq i}(1-p_{a_k})$. The other is to allocate power properly so that the achieved data rate of each individual user can be maximized. Hence, we decompose it into two related problems, and find a suboptimal transmission control policy. The solution to find optimal MAC layer transmission control $\overline{\mathcal{H}}^*$ to resolve collisions in the whole network while guaranteeing proportional fairness can be formulated by

$$\overline{\mathcal{H}}^* = \arg\max_{\overline{\mathcal{H}}} \sum_{(i,j)_k\in\mathcal{E}, j\in\mathcal{T}_i} \left(\ln\left(p_{(i,j)_k}(1-p_{j_k}) \prod_{a\in\mathcal{N}_j, a\neq i}(1-p_{a_k})\right)\right). \tag{2.13}$$

Given MAC transmission decision, in order to maximize the mean physical layer throughput within power capability, the optimal power allocation $\mathcal{P}_i^*$ of User i is formulated by

$$\mathcal{P}_i^* = \arg\max_{\mathcal{P}_i} \sum_{j\in\mathcal{T}_i,k} R_{(i,j)_k}, \tag{2.14a}$$

subject to (2.11b)

$$\sum_{j\in\mathcal{T}_i,k=1,\ldots,K} \frac{1}{|\mathcal{T}_i|} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} P_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) \leq P_a, \tag{2.14b}$$

and (2.11c)

$$\sum_k \left(\max_{h,j} P_{(i,j)_k}(h)\right) \leq P_m, \tag{2.14c}$$

where $\{\overline{H}^*_{(i,j)_k}\}$ is the solution of (2.13) and $R_{(i,j)_k}$ is given by (2.10). Although problem (2.11) has been decomposed into (2.13) and (2.14) to resolve network collisions and improve individual transmission capability respectively, these two problems are closely coupled through $\overline{\mathcal{H}}^*$.

2.3.1 MAC Layer Transmission Control

When optimizing the network with proportional fairness in (2.13), all users are assumed to transmit at the same data rate once the channel power gain is above a certain threshold. Problem (2.13) turns out to be similar with the problem of finding distributed access control strategy to achieve proportional fairness in traditional Aloha networks [79] and [80]. By applying techniques used in [79] and [80], the optimal transmission probability is readily achieved,

$$p^*_{(i,j)_k} = \frac{1}{|\mathcal{S}_i| + \sum_{m\in\mathcal{N}_i} |\mathcal{S}_m|}. \tag{2.15}$$

Combining (2.1) and (2.15), Theorem 2.3.1 follows immediately, and the proof is omitted.

Theorem 2.3.1 *The optimal predetermined channel power gain threshold for any link* $(i, j)_k \in \mathcal{E}$ *where* $j \in \mathcal{T}_i$, $\overline{H}^*_{(i,j)_k}$, *as defined in (2.13), is given by*

$$\overline{H}^*_{(i,j)_k} = F^{-1}\left[(1 - \frac{|\mathcal{T}_i|}{|\mathcal{S}_i| + \sum_{m\in\mathcal{N}_i} |\mathcal{S}_m|})^{\frac{1}{|\mathcal{T}_i|}}\right]. \tag{2.16}$$

From threshold (2.16), the optimal threshold of User i is independent of the receiver j but depends on the neighborhood information of User i itself, including the number of users receiving packets from User i, $|\mathcal{T}_i|$, the number of users sending packets to User i, $|\mathcal{S}_i|$, and the total number of users sending packets to the interfering neighbors of User i, $\sum_{m\in\mathcal{N}_i} |\mathcal{S}_m|$. The first two are local information while $|\mathcal{S}_m|'s, m \in \mathcal{N}_i$, is information about interfering neighbors. The number of flows each interfering neighbor receives, i.e. $|\mathcal{S}_m|$ for all $m \in \mathcal{N}_i$, can be obtained through broadcasting of the interfering neighbor whenever this numbers changes. Since this knowledge needs to be broadcast to notify the interfering neighbors, we call it *two-hop knowledge*. The broadcasting of this two-hop knowledge incurs only trivial signalling overhead since only when either a traffic session or the network topology varies will this broadcasting be triggered. Besides, some form of two-hop knowledge is typical in many protocols, like routing information discovery in mobile *ad hoc* networks [81, 82]. Hence, it can be easily obtained.

Consider User 7 in Figure 2.1. It is easy to see that $|\mathcal{T}_7| = 3$, $|\mathcal{S}_7| = 1$, and the two-hop knowledge of interfering neighbors $|\mathcal{S}_5| = 1$, $|\mathcal{S}_6| = 2$, $|\mathcal{S}_8| = 2$, and $|\mathcal{S}_9| = 2$. Hence, for all $j \in \mathcal{T}_7$ and $k = 1, \cdots, K$, $\overline{H}^*_{(7,j)_k} = F^{-1}\left[(1 - \frac{3}{1+7})^{1/3}\right] = F^{-1}(0.855)$. If the channel is Rayleigh with average power gain h_a, $\overline{H}^*_{(7,j)_k} = 1.931 h_a$. Hence, since there are many traffic flows in the neighborhood of User 7, it transmits only when the channel has very good condition.

As we can see above, the optimal threshold can be obtained through two-hop knowledge. In the following, we consider two special applications.

1. Transmission Control with One-Hop Knowledge

   To avoid signalling broadcast, assume no user has two-hop knowledge, and it needs to be estimated to get approximation of the optimal thresholds. Since the transmission of each interfering neighbor $j \in \mathcal{N}_i$ can be detected by User i, $|\mathcal{T}_j|$ is available. User i can approximate $|\mathcal{S}_i| + \sum_{m \in \mathcal{N}_i} |\mathcal{S}_m|$, the total number of received traffic flows within the interfering range of User i, to be $|\mathcal{T}_i| + \sum_{m \in \mathcal{N}_i} |\mathcal{T}_m|$, the total number of transmitted traffic flows User i can detect. Hence, instead of (2.16), the transmission threshold with one-hop knowledge, i.e. local knowledge, is

   $$\overline{H}^*_{(i,j)_k} = F^{-1}\left[(1 - \frac{|\mathcal{T}_i|}{|\mathcal{T}_i| + \sum_{j \in \mathcal{N}_i} |\mathcal{T}_j|})^{\frac{1}{|\mathcal{T}_i|}}\right]. \tag{2.17}$$

   Since the approximation in (2.17) is not always accurate, there might be some performance degradation. Approximation error happens when there exists undetectable traffic flows that are sent either into or out of the interfering range of User i.

2. Transmission Control for One-Hop Networks

   Assume that all users are within the transmission range of each other, i.e., this is a one-hop network. A simple example is the uplink transmissions of different users to the access point in *wireless local-area network* (WLAN), and at most one traffic flow within the network can succeed in transmission in one transmission slot on one

subchannel. Denote $n = |\mathcal{S}_i| + \sum_{m \in \mathcal{N}_i} |\mathcal{S}_m|$ for any User i, then n is the same for all users and represents the total number of traffic flows in the network. During any time slot on each subchannel, at most one traffic flow within the network can send data successfully. The transmission threshold is given by

$$\overline{H}^*_{(i,j)_k} = F^{-1}\left[(1 - \frac{|\mathcal{T}_i|}{n})^{\frac{1}{|\mathcal{T}_i|}}\right]. \tag{2.18}$$

If each user has only one traffic flow to send, i.e. $|\mathcal{T}_i| = 1$, the transmission threshold is

$$\overline{H}^*_{(i,j)_k} = F^{-1}\left[(1 - \frac{1}{n})\right], \tag{2.19}$$

which is the same as the transmission control in [21]. [21] has demonstrated that the total throughput for such a system achieves a fraction, $(1 - \frac{1}{n})^{n-1}$, of its counterpart's throughput with an optimum centralized scheduler. The throughput reduction is due to the inherent contention in random access.

2.3.2 Physical Layer Optimization with Channel Inversion

Consider a simple transmitter adaptation technique, channel inversion [83], which maintains a constant received power level so that the signals can be reliably received during each traffic session. Once the MAC decides to transmit with channel power gain h, the transmit power is directly given by $P_t = P_r/h$, where P_r is the received power level. Different traffic flows may have different received power levels, P_r, according to the power allocation strategy. The reliable transmission data rate is given by $R(P_r)$. According to the assumption in Section 2.2.2, $R(P_r)$ is strictly concave in P_r since the average noise power is constant on each subchannel.

From (2.4), the average transmit power on link $(i, j)_k$ is

$$\mathbf{E}\{P_{(i,j)_k}\} = \frac{1}{|\mathcal{T}_i|} \int_{\overline{H}_{(i,j)_k}}^{\infty} \frac{P_{r(i,j)_k}}{h} \mathrm{d}F^{|\mathcal{T}_i|}(h). \tag{2.20}$$

Hence, the instantaneous received power is

$$P_{r(i,j)_k} = |\mathcal{T}_i|\mathbf{E}\{P_{(i,j)_k}\}\left(\int_{\overline{H}_{(i,j)_k}}^{\infty} \frac{\mathrm{d}F^{|\mathcal{T}_i|}(h)}{h}\right)^{-1}. \tag{2.21}$$

Denote by $\mathcal{P}_{ri} = \{P_{r(i,j)_k} | (i,j)_k \in \mathcal{E}, j \in \mathcal{T}_i\}$ the set of the received power configuration of User i. According to (2.10), the average data rate is $R_{(i,j)_k} = \mathbf{E}\{R(\eta(h))\} = R(P_{r(i,j)_k})$. The problem in (2.14) is equivalent to

$$\mathcal{P}^*_{ri} = \arg\max_{\mathcal{P}_{ri}} \sum_{j\in\mathcal{T}_i,k} R(P_{r(i,j)_k}), \tag{2.22a}$$

subject to

$$\sum_{j\in\mathcal{T}_i,k} \frac{1}{|\mathcal{T}_i|} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} \frac{P_{r(i,j)_k}}{h} \mathrm{d}F^{|\mathcal{T}_i|}(h) \le P_a, \tag{2.22b}$$

and

$$\sum_k \left(\max_j \frac{P_{r(i,j)_k}}{\overline{H}^*_{(i,j)_k}} \right) \le P_m. \tag{2.22c}$$

The above power allocation problem is solved by Theorem 2.3.2, which is proved in Appendix A.2.

Theorem 2.3.2 *Assuming the strict concavity of the data rate function $R(P_r)$, (2.22) has unique globally optimal reception power levels $P^*_{r(i,j)_k}$ on any link $(i,j)_k \in \mathcal{E}$ where $j \in \mathcal{T}_i$*

$$P^*_{r(i,j)_k} = \min\left(\frac{P_a}{K} \left(\int_{\overline{H}^*_{(i,j)_k}}^{\infty} \frac{1}{h} \mathrm{d}F^{|\mathcal{T}_i|}(h) \right)^{-1}, \frac{P_m \overline{H}^*_{(i,j)_k}}{K} \right), \tag{2.23}$$

in which $\overline{H}^*_{(i,j)_k}$ is determined by Theorem 2.3.1.

Whenever MAC decides to transmit, the physical layer always execute the transmission. However, when $\overline{H}^*_{(i,j)_k}$ is very small, (3.4) turns out to be very small and the physical layer has extremely low throughput due to the penalty of allowing transmission on deeply faded channels. Hence, $\overline{H}^*_{(i,j)_k}$ should be further modified by the physical layer to avoid transmitting on deeply faded channels. Observing (2.15), $p^*_{(i,j)_k}$ can be $1, \frac{1}{2}, \frac{1}{3}$, etc.. Assuming Rayleigh channel with average power gain h_a and one traffic flow is carried, the corresponding thresholds are $0, 0.69h_a, 1.10h_a$, etc.. Hence, transmission on deeply faded channels is possible only when $p^*_{(i,j)_k} = 1$. Thus, define $\overline{H}_o$ as

$$\overline{H}_o = \arg\max_{\overline{H}} R\left(\frac{P_a}{K \int_{\overline{H}}^{\infty} \frac{1}{h} f(h) \mathrm{d}h} \right) (1 - F(\overline{H})), \tag{2.24}$$

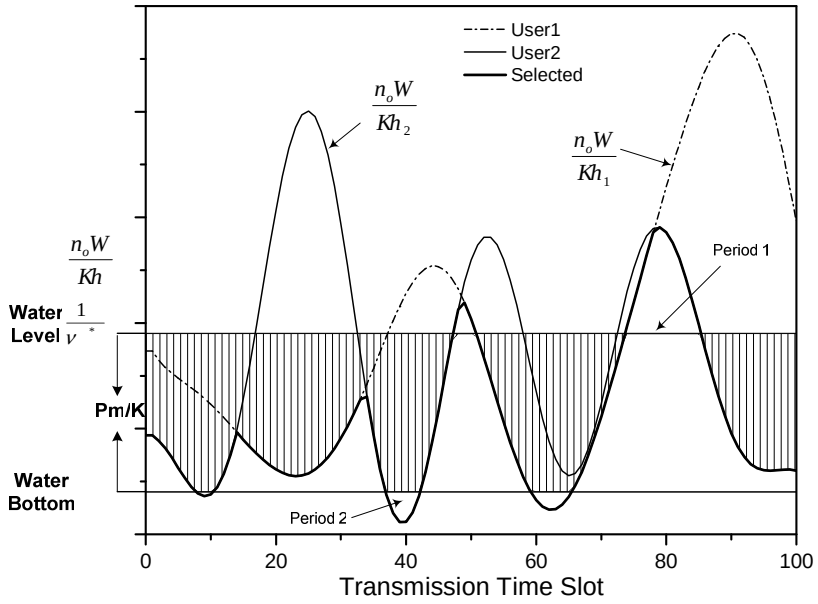


Figure 2.2. Capability limited water-filling over time

which leads to the maximum physical layer throughput when the physical layer is required to transmit under any channel conditions. If $\overline{H}^*_{(i,j)_k}$ determined by Theorem 2.3.1 is less than $\overline{H}_o$, then substitute it with $\overline{H}_o$. This lowers $p_{(i,j)_k}$ slightly since the channel is not deeply faded most of the time. The revision effectively improves link performance but impacts trivially overall network performance, and we do not need to further improve the thresholds of other users to adapt to this change for the sake of optimality in (2.13), which otherwise incurs additional signalling overhead.

With channel inversion, the instantaneous transmit power allocation $\mathcal{P}^*$ is:

$$P^*_{(i,j)_k}(h) = \begin{cases} \frac{P^*_{r(i,j)_k}}{h} & h \geq \overline{H}^*_{(i,j)_k} \\ 0 & otherwise \end{cases} . \tag{2.25}$$

2.3.3 Physical Layer Optimization with Adaptive Modulation and Power Allocation

Consider ideal physical layer transmissions. Each user can vary both the transmit power and rate to achieve the best transmission performance. According to (2.10) and (2.14), the power allocation strategy can be formulated by

$$\mathcal{P}^*_i = \arg\max_{\mathcal{P}_i} \sum_{j \in \mathcal{T}_i, k} \frac{\int_{\overline{H}^*_{(i,j)_k}}^{\infty} R(\eta(h)) \mathrm{d}F^{\mathcal{T}_i}(h)}{(1 - F^{\mathcal{T}_i}(\overline{H}^*_{(i,j)_k}))} \tag{2.26a}$$

subject to (2.11b)

$$\sum_{j\in\mathcal{T}_i,k} \frac{1}{|\mathcal{T}_i|} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} P_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) \le P_a, \tag{2.26b}$$

and (2.11c)

$$\sum_k \left(\max_{h,j} P_{(i,j)_k}(h)\right) \le P_m. \tag{2.26c}$$

The optimal solution of (2.26) is given in Theorem 2.3.3, which is proved in Appendix A.3.

Theorem 2.3.3 *Assume the data rate function $R(\eta)$ to be continuously differentiable and the first order derivative $R'(\eta)$ is positive and strictly decreasing. For any link $(i,j)_k \in \mathcal{E}$ where $j \in \mathcal{T}_i$, (2.26) has a unique globally optimal power allocation given by: if $P_m < \frac{P_a}{1-F^{|\mathcal{T}_i|}(\overline{H}^*_{(i,j)_k})}$, $P^*_{(i,j)_k}(h) = \frac{P_m}{K}$ for $h \ge \overline{H}^*_{(i,j)_k}$; otherwise,*

$$P^*_{(i,j)_k}(h) = \begin{cases} \frac{P_m}{K} & \nu^* < R'\left(\frac{hP_m}{n_oW}\right)\frac{hK}{n_oW}, \\ 0 & \nu^* \ge R'(0)\frac{hK}{n_oW}, \\ R'^{-1}\left(\frac{\nu^* n_o W}{hK}\right)\frac{n_oW}{Kh} & \text{otherwise,} \end{cases} \tag{2.27}$$

for $h \ge \overline{H}^*_{(i,j)_k}$. $R'^{-1}()$ is the inverse function of $R'()$. $\nu^* \ge 0$ is uniquely determined by

$$\int_{\overline{H}^*_{(i,j)_k}}^{\infty} P^*_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) = \frac{P_a}{K}, \tag{2.28}$$

where $\overline{H}^*_{(i,j)_k}$ is given by Theorem 2.3.1.

Observing (2.27), when $\nu^* \ge R'(0)\frac{hK}{n_oW}$, the channel is deeply faded and although the MAC layer decides to transmit, the physical layer further optimizes the transmission and decides not to transmit.

For example, assume the data rate function to be $R(\eta) = \frac{W}{K}\ln(1+\eta)$. The power allocation when $P_m \ge \frac{P_a}{1-F^{|\mathcal{T}_i|}(\overline{H}^*_{(i,j)_k})}$ is given by

$$P^*_{(i,j)_k}(h) = \begin{cases} \frac{P_m}{K} & \frac{1}{\nu^*} - \frac{n_oW}{Kh} > \frac{P_m}{K} \\ 0 & \frac{1}{h} \ge \frac{K}{\nu^* n_o W} \\ \frac{1}{\nu^*} - \frac{n_oW}{Kh} & \text{otherwise} \end{cases} \tag{2.29}$$

for $h \geq \overline{H}^*_{(i,j)_k}$, which is similar to the well-known water-filling power allocation scheme [84, 85, 86]. Since the proposed power allocation scheme has maximum instantaneous power constraint, we call it capability-limited water filling.

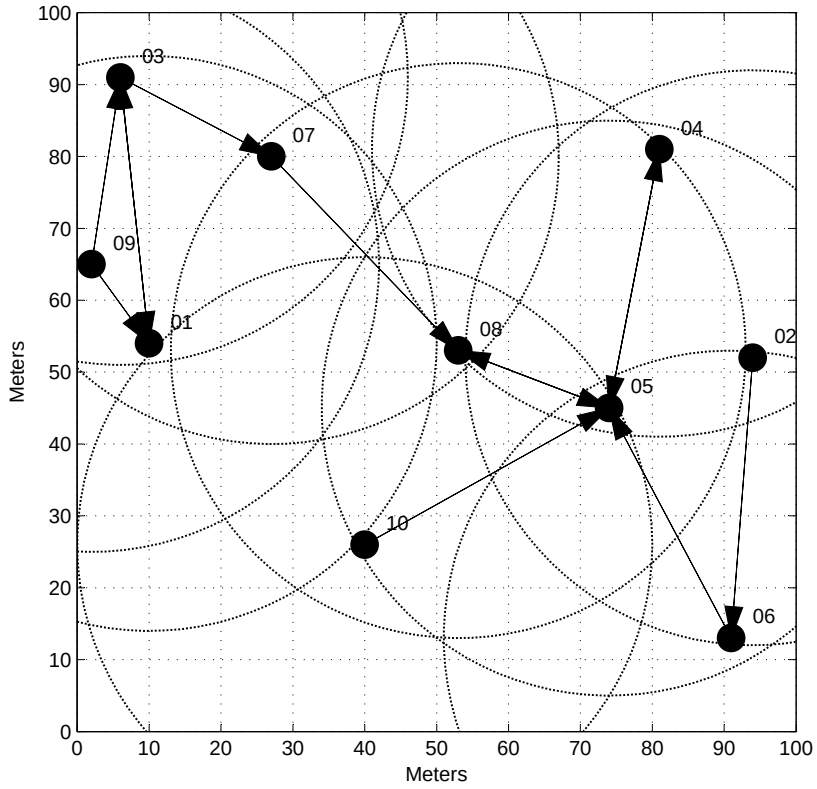


Figure 2.3. Random network topology.

According to (2.27), power will be optimally distributed over both time and all subchannels. Figure 2.2 illustrates the capability-limited power allocation of a user that is transmitting data to User 1 and 2 on a subchannel by using (2.29), and the striped parts in the figure represent the amount of power allocated. The power allocation during 100 transmission time slots is shown. We assume that $\frac{\nu^* n_o W}{K} > \overline{H}^*_{(i,j)_k}$ here. According to the transmission policy, the user always selects the destination with better channel power gains. As indicated by "Period 1" in Figure 2.2 there are no transmissions when subchannels of both User 1 and 2 are deeply faded. In "Period 2" in Figure 2.2, although the channel conditions are so good that higher data rates can be achieved, the actual data rate is limited by the instantaneous transmission capability P_m.

When $\frac{\nu^* n_o W}{K} < \overline{H}^*_{(i,j)_k}$, since the MAC decides to transmit only when $h > \overline{H}^*_{(i,j)_k}$, the physical layer will always transmit when the MAC wants to transmit according to (2.29). Assuming large P_m, the power allocation is always $\frac{1}{\nu^*} - \frac{n_o W}{Kh}$. Then according to (2.28), the water level is

$$\frac{1}{\nu^*} = \frac{\frac{n_o W}{K} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} \frac{1}{h} \mathrm{d}F^{|\mathcal{T}_i|}(h) + \frac{P_a}{K}}{1 - F^{|\mathcal{T}_i|}(\overline{H}^*_{(i,j)_k})}. \tag{2.30}$$

We can always use (2.30) to approximate the water level since with large probability, most transmissions will fall within the normal working ranges of the transmitter.

2.4 Simulation Results

In this section, we first demonstrate DOMRA performance in a network with random topologies. Then we further show how closely DOMRA performs to the globally optimum solution.

2.4.1 Network Performance Improvement

Consider a network with random topologies and compare the average performance of all simulation trials. In each simulation trial, users are randomly dropped and uniformly distributed in a square area with side length one hundred meters. Each user has a transmission range of forty meters and selects neighboring users randomly for data transmission. Figure 2.3 illustrates a network topology in one trial, where arrows indicate traffic flows and circles transmission ranges of different users. Different schemes will be implemented to provide detailed performance comparisons.

2.4.1.1 Single-channel network

Assume that the network operates with one channel. For simplicity, assume Rayleigh fading channel and $R(P) = W \ln(1 + \frac{hP}{WN_0})$. We will compare the performance of the proposed cross-layer transmission policy with the channel-aware Aloha in [21], and the optimal traditional Aloha in [79], which does not consider cross-layer optimizations. For

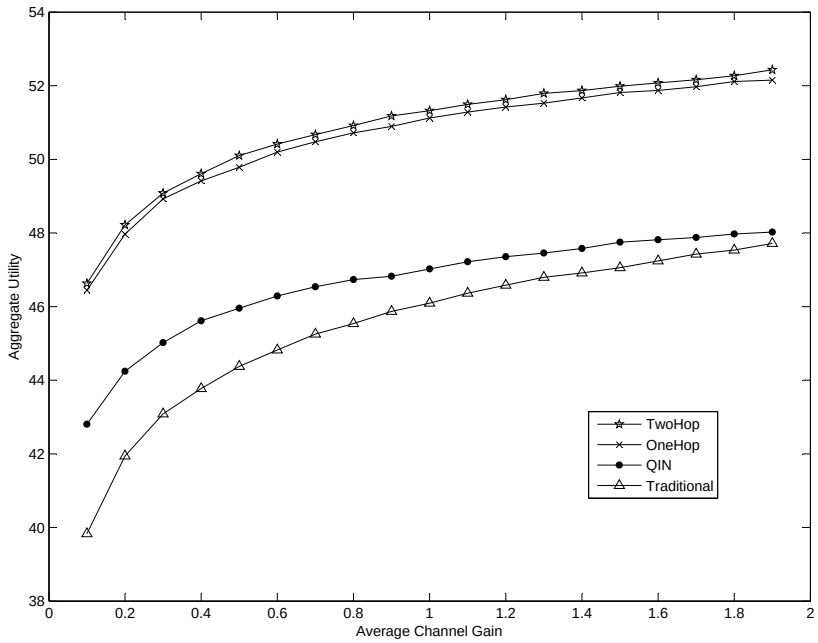


Figure 2.4. Network aggregate utility comparison. $P_m = 50$ **dBm,** $P_a = 43$ **dBm,** $W = 100$Hz**, and** $N_o = 0.001$W/Hz.

traditional Aloha transmissions, in order to make the comparison meaningful, the same average power constraint and instantaneous power constraint are enforced. Since there is no cooperation between MAC and the physical layer, the physical layer assumes that it keeps on transmitting except when the channel is deeply faded. In order to satisfy power constraints, the transmission threshold is chosen so that the average data rate is maximized, i.e. $\overline{H} = \arg\max_H (1 - F(H))R(P_r)$ subject to the instantaneous power constraint (2.6), and P_r is given by (2.21). The threshold is found through linear search.

Figure 2.4 shows the aggregate utility comparison of the whole network when the channel has different average channel gains. The "TwoHop" curve represents the result of DOMRA when each user has two-hop information of the neighboring users while the "OneHop" curve represents the result when each user has only one-hop information. As we can see, with only one-hop knowledge, the system has slight performance degradation as compared with the transmissions when two-hop knowledge is available. Curve "QIN" shows the performance of [21], which assumes that each user has the knowledge of how many users there are in the whole network. Curve "Traditional" shows the result using

the traditional optimal Aloha. As shown in Figure 2.4, with the advantage of cross layer design, the proposed scheme outperforms traditional optimal Aloha greatly. In addition, by exploiting the neighborhood information of each user, the proposed method also outperforms the existing channel-aware Aloha in [21]. This is due to the consideration of the inhomogeneous traffic spatial distribution in the proposed scheme and the channels are better utilized.

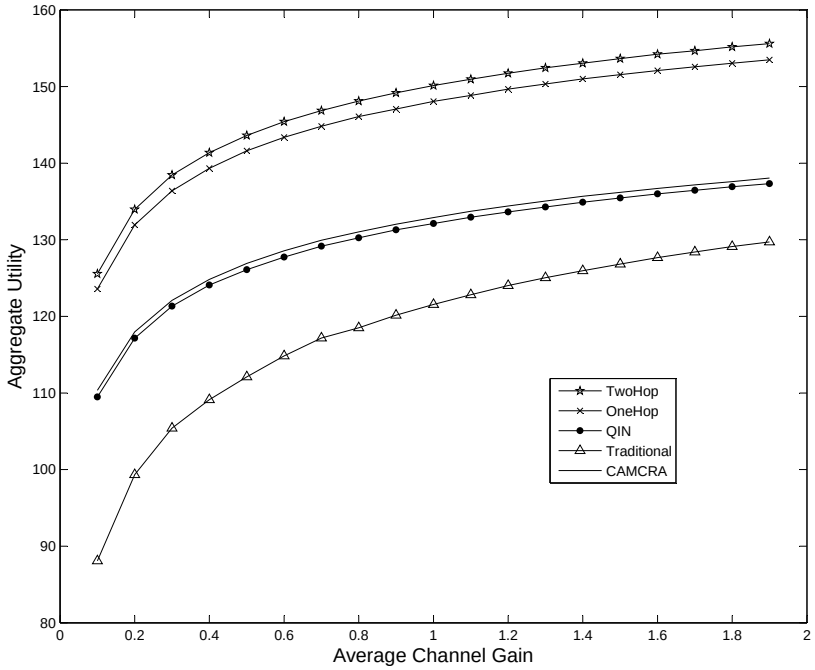


Figure 2.5. Five channel network aggregate utility comparison. P_m = 50dBm, P_a = 43dBm**,** W = 100Hz**,** N_o = 0.001W/Hz.

2.4.1.2 Multichannel network

Consider the same wireless network configurations as those in the single-channel network scenario except that there are five subchannels. Besides implementing schemes in the single-channel network scenario for multichannel environment, we also run the CAMCRA proposed in [23]. During each each transmission slot, CAMCRA chooses c subchannels with the c most significant gains, where $c = \max\left(1, \lfloor \frac{\text{subchannel number}}{\text{user number}} \rfloor\right)$. Then the method in ([21]) is applied on each subchannel given that each user knows how many users are using the subchannel. Since the number of users in each subchannel is a random variable,

it is proposed in [23] to use $\max\left(1, \frac{\text{user number}}{\text{subchannel number}}\right)$ as an estimate. As shown in Figure 2.5, the CAMCRA in [23] has slight performance improvement as compared with the channel-aware Aloha in [21] because of exploitation of multichannel diversity. However, these two schemes do not perform good when the network has arbitrary spatial traffic distribution. Our proposed DOMRA with either two-hop or one-hop information significantly outperforms these existing schemes due to exploitation of multiuser diversity and proper adaptive transmission settings and power allocation according to inhomogeneous traffic spatial distribution in the network.

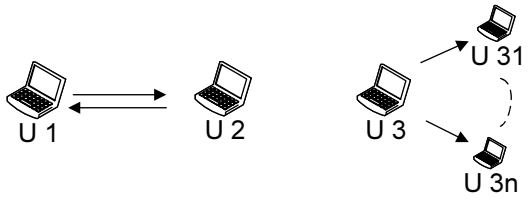


Figure 2.6. A simple network topology

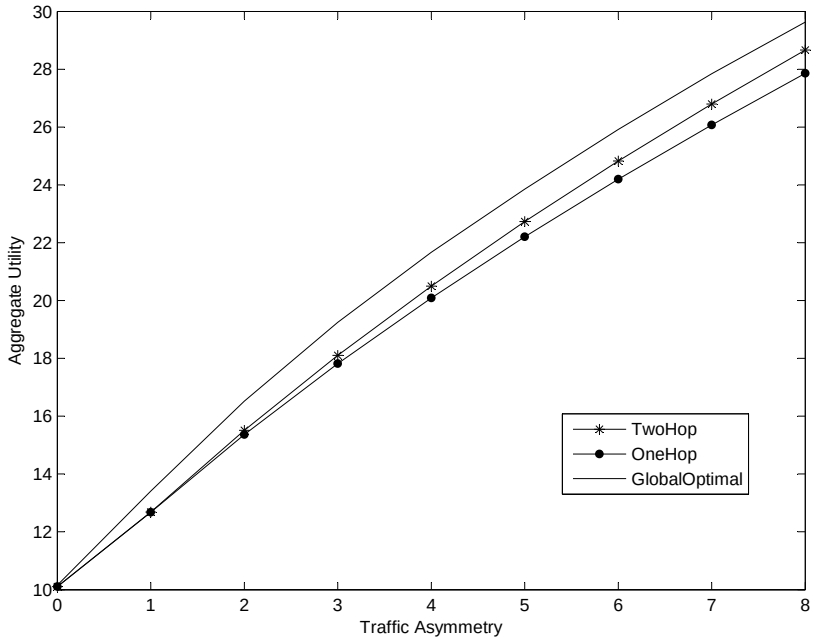


Figure 2.7. Aggregate utility gap to the global optimum. P_m = 50dBm, P_a = 43dBm**,** W = 100Hz**,** N_o = 0.001W/Hz.

2.4.2 Suboptimality Gap

Problem (2.11) is decomposed into subproblems (2.13) and (2.14) to obtain feasible suboptimal control policy. In order to show the suboptimality gap, we exhaustively search for the global optimum in (2.11), and run a simple network topology to reduce search complexity. As shown in Figure 2.6, arrows indicate traffic flows. User 3 is sending traffic to n receivers, who are all out of the transmission ranges of Users 1 and 2. User 1 can communicate with 2, but not 3, while User 2 can communicate with both. When n is zero, the traffic distribution is symmetric in the network. The larger the number n, the more asymmetric the traffic distribution is. We call *n traffic asymmetry*, and vary it from 0 to 8. Figure 2.7 compares network aggregate utility and shows the suboptimality gap. While the global optimum can only be obtained through floods of broadcast of complete network knowledge, our decomposition technique yields a feasible suboptimal decentralized solution, which requires limited (two-hop knowledge case), or no (one-hop knowledge case) signalling overhead. Besides, the proposed scheme performs closely to the global optimum, and even reaches the global optimum when the traffic is symmetric.

2.5 Summary

We have proposed a joint physical-MAC layer optimization policy for multichannel Aloha random access in wireless networks in which all users are not necessarily within the transmission range of each other and each user may have packets to send to or receiver from different users. The joint physical-MAC layer optimization policy exploits decentralized CSI, and achieves multi-user diversity through cross-layer design. System performance is optimized while proportional fairness is obtained with the consideration of the inhomogeneous characteristics of the traffic spatial distribution. Simulation results show that the proposed scheme significantly outperforms existing channel aware Aloha schemes. The generality of the design in this chapter will allow its applications in different types of wireless networks to fully exploit the system capacity. The scheme presented here is simple but

gives guidelines for decentralized cross-layer optimization in practical wireless networks. The methodology provided can be easily adapted to improve the performance of different wireless networks. For example, in networks based on 802.11 standards, besides using the backoff window technology, the transmission of RTS to compete for channel access can also be designed according to the proposed DOMRA to further decrease the collision probability and allow larger successful probability of users with better channel power gain. In the following two Chapters, we will show the application of DOMRA.

CHAPTER 3

COCHANNEL INTERFERENCE AVOIDANCE MAC

Severe cochannel interference in wireless cellular networks significantly affects users at cell edges. In this chapter, we develop a cost-effective *cochannel interference avoidance MAC* (CIA-MAC) to deal with the downlink transmission experiencing severe CCI, especially for the users at cell edges. The proposed CIA-MAC scheme maintains backward compatibility and requires only minor changes to existing BSs, while no improvement is necessary for MTs. Low overhead is added as the scheme requires only limited signaling coordination among BSs at a semi-static level. Only occasional cooperation is required when the network topology is changed and the instantaneous coordination at the packet level is not required. The proposed scheme is novel in its use of randomization by a BS for controlling the level of interference and it provides fair transmission opportunities for the users affected by severe CCI. Although the use of randomization for collision avoidance is used extensively for uplink random access channels and WLAN systems, we are first using the principle for automatically controlling the level of downlink interference per link in cellular networks. In the following, the system will be briefly described in Section 6.1. In Section 3.2, we present CIA-MAC with details. In Section 3.3, we address the conditions for triggering CIA-MAC and obtain two simple trigger mechanisms. The performance improvement is demonstrated through the simulation in Section 3.4. Finally, the conclusion and future work are given in Section 3.5.

3.1 Network with CIA

We only consider downlink transmission since complicated multiuser detection and CCI cancelation algorithms can be implemented at the BS for uplink CCI mitigation. The MTs at cell edges not only face the weakest signals but also suffer the largest amount of interference from neighboring cells. The CIA-MAC scheme targets the performance improvement

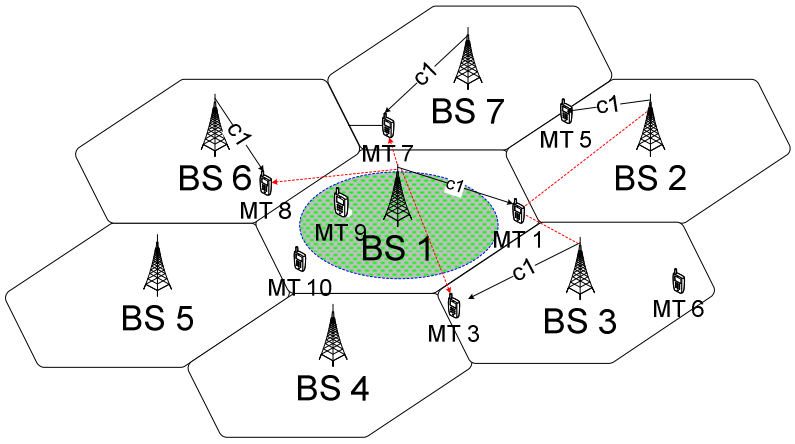


Figure 3.1. Cochannel interference in cellular networks with a reuse factor of one

of these users. In general, users are categorized into two classes: those experiencing no or slight CCI and those suffering from severe CCI. The first class will be scheduled by the traditional centralized MAC and the second will be first accepted by the traditional call admission control policies and then scheduled by the proposed CIA-MAC.

Each MT measures the average *interference-to-carrier ratio* (ICR) of neighboring BS k, which is defined to be

$$ICR_k = \frac{E(h_k P_k)}{E(hP)}, \tag{3.1}$$

where h and P are the channel power gain and the transmit power of the desired link while h_k and P_k correspond to these of the interfering link from BS k. $E()$ is the average over a sliding window of the past data and tracks slow fading, i.e. it is a local mean and averages the effect of fast fading [37]. This definition of average will also apply in the following paragraphs.

Severe Interferer: *If the ICR from neighboring BS k satisfies $ICR_k \geq \Gamma_m$, the transmission of BS k always causes the failure of packet reception, where Γ_m, called trigger, is a predetermined severe interference threshold. BS k is called a severe cochannel interferer.*

In Section 3.3, we will discuss the trigger selection and present examples.

If all BSs causing severe interference keep on transmitting, the packet receptions of the interfered MTs always fail. If BSs can collaborate, the interfering BSs may transmit in turn.

However, this incurs huge signaling overhead. If there is no collaboration among BSs, we let BSs transmit randomly when identified as severe interferers. Their transmission should be managed such that the overall network performance as well as the fairness among all users are jointly optimized. Hence, besides traditional MAC, a complementary MAC is used for optimizing the randomized transmission of the severe interferers. The new complementary MAC aims at improving cellular throughput through *cochannel interference avoidance*, and is, therefore, called CIA-MAC.

In the following, we only consider BSs and MTs controlled by CIA-MAC and CIA-MAC will be optimized by cross-layer design to fully exploit the system capacity and multiuser diversity while maintaining fairness.

The following definitions will be used in the subsequent discussion:

- $\mathcal{B} = \{1, 2, \cdots, M\}$: set of BSs.
- $\mathcal{M} = \{1, 2, \cdots, N\} = \bigcup_{i \in \mathcal{B}} \mathcal{M}_i$: set of MTs. $\mathcal{M}_i$ is the set of MTs in the cell of BS i. Obviously, $\mathcal{M}_i \bigcup \mathcal{M}_j = \emptyset, \forall i \neq j$.
- $\mathcal{E} = \{(i, j) | i \in \mathcal{B}, j \in \mathcal{M}_i\}$: set of transmission links; (i, j) denotes the link from BS i to MT j.
- $\mathcal{N}_m = \{(i, j) | \forall (i, j) \in \mathcal{E}$, transmission at link (i, j) causes severe interference to MT $m\}$: set of links whose transmission will bring severe interference to MT m;
- $\mathcal{T}_{(i,j)} = \{m | \forall m \in \mathcal{M}$, MT m is severely interfered by the transmission at link $(i, j)\}$: set of MTs severely interfered by transmission at link (i, j).

Figure 3.1 demonstrates an example. The solid lines represent data transmission links and the dashed ones links from severe interferers. We have $\mathcal{B} = \{1, 2, \cdots, 7\}$, $\mathcal{M} = \{1, 3, 5, 7, 8, 9, 10\}$, and $\mathcal{E} = \{(1, 1), (2, 5), (3, 3), (6, 8), (7, 7)\}$. The transmission from BS 1 to MT 1 on channel 1 is severely interfered by transmission from BS 2 to MT 5. Meanwhile, BS 1 also causes severe CCI to MTs 3, 7, and 8. Hence, $\mathcal{N}_1 = \{(2, 5), (3, 3)\}$ and

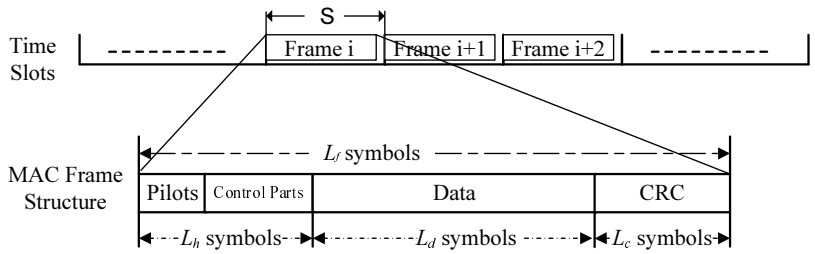


Figure 3.2. MAC transmission and frame structure

$\mathcal{T}_{(1,1)} = \{3, 7, 8\}$.

3.2 Principle of CIA-MAC

In this section, we describe the principle of CIA-MAC. Figure 3.2 shows the MAC transmission and the frame structure. Transmission time is divided into slots with length S. The MAC layer of each link independently sends a MAC frame at the beginning of each slot with probability p. Complete *channel state information* (CSI) is known and used to determine the MAC contention and the PHY link adaptation. To obtain CSI at a transmitter [87], the CSI can be estimated through pilots at the receiver and sent to the transmitter or CSI is already available at the transmitter whenever the channel is reciprocal, such as in a time-division duplex system. Incomplete CSI results in some performance loss and the study on its impact is out of the scope of this chapter. Furthermore, assume ideal *cyclic redundancy check* (CRC). Any error inside a frame will result in the drop of the frame. Errors are uniformly and independently distributed. Each frame has L_f symbols, of which L_d symbols carry data. Once the MAC layer makes a decision to transmit a frame, the frame will be continuously transmitted by the *physical* (PHY) layer until the frame is sent out.

In a traditional network, MAC makes transmission decision based on buffer status and *quality of service* (QoS) requirements and does not use PHY knowledge at all. When MAC decides to transmit, the physical channel may be in a deep fade, which wastes bandwidth and power resources. Alternatively, MAC may decide on no transmission while the channel is experiencing high gain. With cross-layer design, MAC decides whether to transmit or not according to channel information. We assume a block fading channel [88], that is, the

channel state remains constant during each MAC frame. If the channel power gain at a time slot, h, is above a predetermined threshold $\overline{H}$, MAC sends a frame. As wireless channels are inherently random, the MAC transmission is also randomized and the threshold $\overline{H}$ determines the transmission probability. The thresholds $\overline{H}$ and the PHY transmission should be jointly optimized for all BSs subject to their power constraints. Each BS maximizes its throughput with both average power constraint P_a and instantaneous power constraint P_m while assuring fairness to the users in other cells.

DOMRA will be used to optimize the operations of this network. Denote the probability cumulative distribution function of channel power gain as $F(h)$ and the cardinality of $\mathcal{T}_{(i,j)}$ as $|\mathcal{T}_{(i,j)}|$. From Theorem 2.3.1, the optimal channel gain threshold for any link $(i, j) \in \mathcal{E}$, $\overline{H}^*_{(i,j)}$, that also assures proportional fairness among all CIA-MAC links, is

$$\overline{H}^*_{(i,j)} = F^{-1}\left(\frac{|\mathcal{T}_{(i,j)}|}{1 + |\mathcal{T}_{(i,j)}|}\right), \tag{3.2}$$

and the corresponding transmission probability is

$$p^*_{(i,j)} = \frac{1}{|\mathcal{T}_{(i,j)}| + 1}, \tag{3.3}$$

where $F^{-1}(\cdot)$ denotes the inverse function of $F(\cdot)$. From (3.2), the optimal threshold of link (i, j) depends only on $|\mathcal{T}_{(i,j)}|$, the number of MTs severely interfered by transmission at link (i, j). This knowledge can be shared by the BSs of neighboring cells. $|\mathcal{T}_{(i,j)}|$ changes only when the severely-interfered MTs have large status variations that result in the obvious changes of the ICR in (3.1) to go across the trigger. For example, an existing traffic session ends, a new one starts, or the movements of MTs make either a new MT severely interfered or an existing MT no longer severely interfered. These variations will trigger the update of $|\mathcal{T}_{(i,j)}|$.

In the PHY layer, consider channel inversion [83] and each BS allocates transmit power to maintain a constant received power level so that signals can be reliably detected. Once MAC decides to transmit, the transmit power is given by $P_{(i,j)}(h) = P_r/h$, where P_r is the

power level for reliable receiver detection. Given $\overline{\mathcal{H}}^*$ in (3.2), the received power level is optimized by Theorem (2.3.2)

$$P^*_{r(i,j)} = \min\left(\frac{P_a}{\int_{\overline{H}^*_{(i,j)}}^{\infty} \frac{1}{h} f(h)\mathrm{d}h}, P_m \overline{H}^*_{(i,j)}\right). \tag{3.4}$$

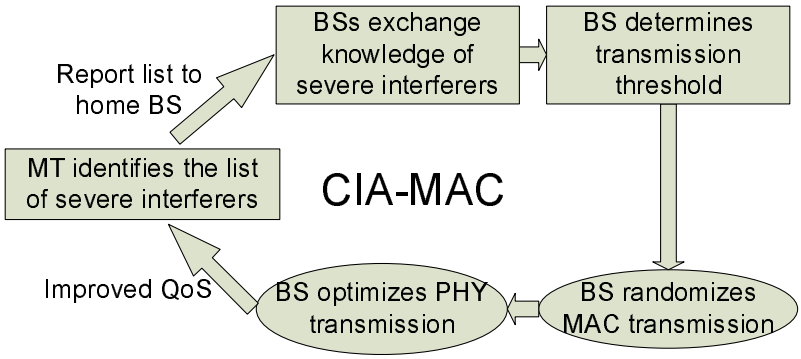


Figure 3.3. CIA-MAC flowchart

Figure 3.3 illustrates the flowchart of CIA-MAC. Each MT identifies the list of neighboring BSs causing severe interference by comparing their ICRs with the trigger Γ_m and reports the list to its home BS. The home BS communicates the list to other BSs and each BS knows the links on which it needs to randomize transmission. Then each BS determines a channel threshold $\overline{H}_{(i,j)}$ per CIA-MAC link (i, j) based on the number of links affected by the transmission on that particular link. A BS transmits on a CIA-MAC link only when the channel gain on the link exceeds its channel threshold and thus randomizes the transmission. The transmission power and modulation are optimized on each link separately. As shown in Figure 3.3, the operations of CIA-MAC are classified into two parts. The operations in the rectangles are semi-static and take place only when the severely-interfered MTs have large status variation. With this trivial cost, the operations in the ovals automatically improve the QoS of all MTs experiencing severe interference.

3.3 Trigger for CIA-MAC

In this section, we discuss the selection of the trigger.

3.3.1 Trigger Selection

Each MT measures the ICR of each neighboring BS and CIA-MAC is used at that BS when the corresponding ICR is above the trigger, Γ_m. Hence, the trigger determines severe interferers and relates to the performance of CIA-MAC. We will choose the trigger to maximize the throughput of the whole network rather than that of any individual link.

From (2.3), the throughput of link (i, j) can be expressed as

$$T_{(i,j)} = p_{(i,j)} \prod_{l \in \mathcal{N}_j} (1 - p_l) \frac{RL_d}{S} \cdot (1 - p_F) \tag{3.5}$$

where p_l is the transmission probability on link l, $\prod_{l \in \mathcal{N}_j}(1 - p_l)$ is the probability that none of the severe interferers of MT j transmit, R is the average transmitted bits per symbol when the MAC of BS i decides to transmit at link (i, j) and depends on the modulation and power allocation policy, and p_F is the *frame-error rate* (FER) when no severe interferers transmit.

Optimum triggers are different for MTs with different interference scenarios. Figure 3.1 illustrates an example. MT 3 wants to judge BS 1. If BS 1 is judged to be a severe interferer, it will transmit with lower probability according to Equation (3.3). However, the impact of this variation to the packet receptions of MTs 1, 7, and 8 is unknown to MT 3. Hence, it is difficult for each MT to evaluate the variation of the overall network throughput that results from the judgement of severe interferers. Even assume that these knowledge can be shared, different MTs have different interfering scenarios, and judging severe interferers and exchanging signaling would be daunting tasks across the whole network.

We will get one trigger for all to simplify the calculations. Consider a network in which each MT is severely interfered by K neighboring BSs on average. The BS of each MT also brings severe CCI to K MTs in the neighboring cells on average. Empirical values can be assigned to K, e.g. $K = 3$ is a good choice based on our simulation observations for a reuse one network. From equation (3.3), the transmission probability of each BS is $\frac{1}{1+K}$ and the

throughput is

$$T = \frac{1}{1+K}(1-\frac{1}{1+K})^K \frac{RL_d}{S}(1-p_F) = \frac{K^K}{(1+K)^{K+1}}(1-p_b)^{RL_d}\frac{RL_d}{S}, \tag{3.6}$$

where the frame error rate is approximated by $p_F = 1-(1-p_b)^{L_d R}$ according to the assumption of uniform and independent error distribution in Section 3.2. The BER is approximated by $p_b = P_e(\eta)$, where η is the average SNR and the BER function $P_e()$ depends on the modulation and coding. For example, the BER for coherently detected M-QAM with Gray mapping over an additive white Gaussian noise channel can be well approximated by [89]

$$P_e(\eta) \approx 0.2\exp\left(-\frac{1.5G_c\eta}{M-1}\right), \tag{3.7}$$

where G_c is the coding gain and M is the modulation order.

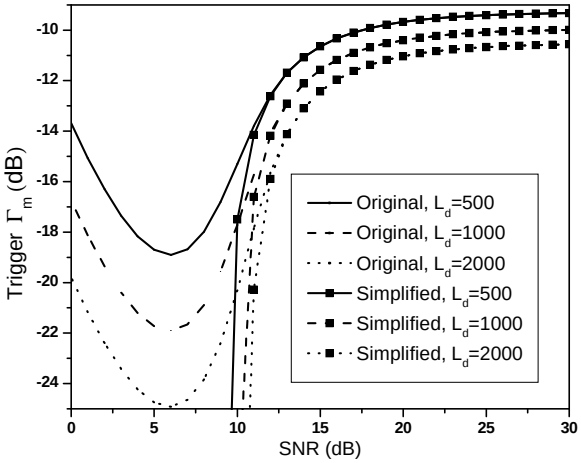


Figure 3.4. Trigger of severe cochannel interferer

With the traditional MAC, all BSs keep on transmitting. Each link has throughput

$$\widehat{T} = \frac{L_d\widehat{R}}{S}\cdot(1-\widehat{p}_F) = \frac{L_d\widehat{R}}{S}\cdot(1-\widehat{p}_b)^{L_d\widehat{R}}, \tag{3.8}$$

where $\widehat{p}_F$ and $\widehat{p}_b$ are the average frame and bit error rates, and $\widehat{R}$ is the average number of bits transmitted per symbol in this mode. $\widehat{p}_b$, $\widehat{p}_F$, and $\widehat{R}$ are different from those in (3.6) since BSs have different transmission durations and signal receptions are with different interference scenarios, which result in different power and modulation allocation approaches.

CIA-MAC is triggered when it achieves better throughput, i.e. $T > \widehat{T}$, or

$$\frac{K^K}{(1+K)^{K+1}}(1-p_b)^{L_d R}\frac{L_d R}{S} > (1-\widehat{p_b})^{L_d \widehat{R}}\frac{L_d \widehat{R}}{S}, \tag{3.9}$$

Then we have

$$\widehat{p_b} > 1 - \left(\frac{K^K}{(1+K)^{K+1}}(1-p_b)^{L_d R}\frac{R}{\widehat{R}}\right)^{\frac{1}{L_d \widehat{R}}}. \tag{3.10}$$

Since $\widehat{p_b} = P_e(\widehat{\eta})$, where $\widehat{\eta}$ is the average *signal-to-interference-plus-noise ratio* (SINR),

$$\widehat{\eta} < P_e^{-1}\left[1 - \left(\frac{K^K}{(1+K)^{K+1}}(1-p_b)^{L_d R}\frac{R}{\widehat{R}}\right)^{\frac{1}{L_d \widehat{R}}}\right]. \tag{3.11}$$

$SINR$ and ICR follow the relationship

$$ICR = \frac{1}{\widehat{\eta}} - \frac{1}{\eta}. \tag{3.12}$$

The trigger Γ_m follows immediately

$$\Gamma_m = \frac{1}{P_e^{-1}\left[1 - \left(\frac{K^K}{(1+K)^{K+1}}(1-P_e(\eta))^{L_d R}\frac{R}{\widehat{R}}\right)^{\frac{1}{L_d \widehat{R}}}\right]} - \frac{1}{\eta}. \tag{3.13}$$

Γ_m depends on SNR and BER function, both of which are known to each MT. Hence, Γ_m can be easily calculated for judgement of severe interferers.

For fixed modulation, $R = \widehat{R}$. The trigger is

$$\Gamma_m = \frac{1}{P_e^{-1}\left[1 - \left(\frac{K^K}{(1+K)^{K+1}}\right)^{\frac{1}{L_d R}}(1-P_e(\eta))\right]} - \frac{1}{\eta}. \tag{3.14}$$

For normal data transmission, SNR is high and $P_e(\eta) \ll 1$, thus (3.14) is further simplified to be

$$\Gamma_m = \frac{1}{P_e^{-1}\left[1 - \left(\frac{K^K}{(1+K)^{K+1}}\right)^{\frac{1}{L_d R}}\right]} - \frac{1}{\eta}. \tag{3.15}$$

3.3.2 An Alternate Trigger Mechanism Using Location Knowledge

In the flowchart of Figure 3.3, each MT determines the list of severe interferers and reports the list to the home BS, which requires additional improvement of MTs. In the following,

we show how to enable BSs to determine the severe interferers to avoid the necessity of MT improvement.

We assume that BSs have the position knowledge of MTs in both the home cells and the neighboring cells. Note that a large quantities of positioning techniques have been proposed in cellular networks [90, 91, 92, 93]. Hence, it is practical to obtain the position knowledge of each MT and this knowledge can be shared among neighboring BSs. Besides, assume that each BS knows the average received signal power at a desired MT, which can be obtained through feedback or observation of link power control. This knowledge will also be shared among neighboring BSs.

We have shown that the optimal threshold (3.2) for each BS depends on the number of MTs severely interfered by its transmission. This number can be obtained through co-operation among BSs. A BS located at coordinate (x_b, y_b) needs to determine whether it brings severe interference to the neighboring-cell MT at coordinate (x_m, y_m). The distance between them is

$$d_I = \sqrt{(x_b - x_m)^2 + (y_b - y_m)^2}, \tag{3.16}$$

which results in path loss $L(d_I)$. The average received signal power at the MT is P_s while the interfering BS has the average transmit power P_I. According to (3.1), the average interference to carrier ratio is

$$ICR = \frac{P_I h_I}{P_s}, \tag{3.17}$$

where h_I is the average channel power gain of this interfering link. However, h_I is unknown to the interfering BS and needs to be estimated. Radio propagation is characterized by three nearly independent phenomena: path loss variation with distance, slow log-normal shadowing, and fast multipath fading [37]. Similar to the ICR in (3.1), h_I tracks slow fading, i.e. it is a local mean and averages the effect of fast multipath fading. Hence, we consider only the path loss and shadowing. Shadow represents the error between the actual and estimated path loss [37]. While the estimated path loss is determined by the radio path distance d, the shadowing/estimation error has been observed to be nearly independent of

d and we assume the independence. Hence, we model the estimated average interference channel gain by two parts: the estimated path loss determined by path loss model $L(d)$ and the estimation error determined by the shadowing model. Shadows are generally modeled as being log-normally distributed and $10\log_{10}(h_I)$ has normal distribution with mean $-10\log_{10}(L(d))$ and standard deviation σ, where σ is independent of the radio path length d and typically ranges from 5 to 12 dB [37]. To ensure a detection probability β of severe interferers, the BS determines that the MT is severely interfered when the probability of severe interference is above β, i.e.

$$Prob[\text{severe interference}] = Prob\left[ICR \geq \Gamma_m\right] \geq \beta, \tag{3.18}$$

which is equivalent to

$$Prob\left[10\log_{10}(h_I) \geq 10\log_{10}\left(\frac{\Gamma_m P_s}{P_I}\right)\right] \geq \beta. \tag{3.19}$$

Solving (3.19) yields the detector of severe interference as follows

$$\frac{P_I}{L(d_I)P_s} \geq \Gamma_m 10^{-\frac{\sigma Q^{-1}(\beta)}{10}} \triangleq \Gamma_b, \tag{3.20}$$

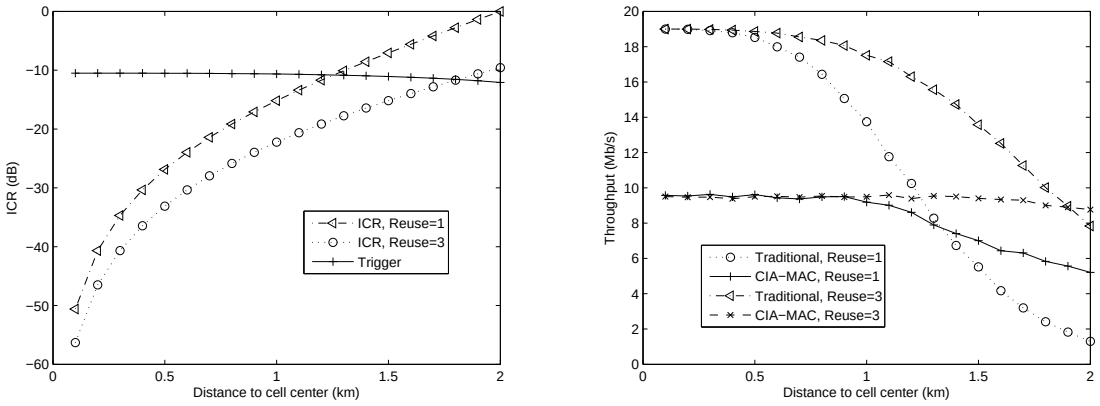
where $Q(x)$ is the right-tail probability of the standard normal distribution, that is

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt. \tag{3.21}$$

With (3.20), each BS can detect how many MTs are severely interfered and determine its transmission probability (3.3) as well as the threshold (3.2). In this case, no MT improvement is necessary for the functioning of CIA-MAC. Note that the selection of β determines how pessimistically or optimistically a severe interferer is judged.

3.4 Numerical and Simulation Results

In this section, we show the relationship between the trigger and SNR, verify the effectiveness of the trigger, and demonstrate the performance of CIA-MAC in a cellular network through comparison with the traditional MAC and a static FFR approach.



(a) (b)

Figure 3.5. Trigger effect.

In the cellular system simulated, the radius of each cell is 2 km and no sectoring is used. The thermal noise power is -104 dBm over the whole bandwidth. The carrier frequency is 900 MHz. BSs are 100 meters high with 8.2 dB antenna gain while MTs have height 1.5 meters with 2.2 dB antenna gain. Path loss is given by the urban-area Hata-Okumura model. Log-normal shadowing and Rayleigh fading are applied. Each MAC frame consists of 1000 symbols, in which 900 carry payload.

3.4.1 Relationship of Trigger and SNR

Considering uncoded 4-QAM modulation, the relationship between the trigger Γ_m and SNR when L_d has different values is illustrated by Figure 3.4. The amount of bits transmitted per MAC frame varies and is usually very large to fully exploit link capacity. For example, in 802.16e [12], each frame has a maximum length of 2048 bytes of payload followed by one CRC verification, i.e. 16384 bits of payload per frame. In *high-speed downlink packet access* (HSDPA) transmission of *universal mobile telecommunications system* (UMTS) [34], the size of a transport block followed by one CRC verification ranges widely from 15890 bits to 204000 bits. In general, we assume large L_dR and illustrate the cases when L_d is 500, 1000, and 2000 respectively. The curves without markers are calculated through (3.14) while those with markers through (3.15). Figure 3.4 clearly shows that (3.15) is a

good approximation of the threshold for high SNR. In the high-SNR region, the receiver can bear higher interference when signal power increases, yielding increasing trend of the curve. In the low-SNR region, when SNR goes lower, that is, noise power goes higher, interference needs to have stronger power to impact more on frame reception than noise. This indicates the increasing trend of the curve as SNR goes lower in the low-SNR region. This also indicates that in the noise-dominated region, it is better to ignore interference, as suppressing it will not provide much advantage.

3.4.2 Effect of the Trigger

Consider a simplified cellular network. Each BS serves one user on a channel. Compare the performance of CIA-MAC and the traditional MAC that keeps on transmitting. All MTs are located at the same distance away from their corresponding home BSs and the distance goes from 0 to the cell radius. Each MT is severely interfered by one neighboring BS and each BS causes severe CCI to a MT in a neighboring cell. Assume uncoded 4-QAM modulation and a transmit power of 43 dBm. Figure 3.5(a) shows the relationship between the ICR and the trigger when the network has a reuse factor of either one or three. Figure 3.5(b) shows the throughput of both CIA-MAC and the traditional MAC. Since the trigger depends on signal and noise powers but not interference, it is independent of the network reuse type, which determines interference environment. As a MT moves closer to a cell edge, the trigger decreases because of decreasing signal power while the ICR increases because of growing interference power. From Figure 3.5, CIA-MAC is triggered when the ICR is above the trigger and achieves better throughput than the traditional MAC through interference avoidance. The network performance is improved for MTs experiencing severe interference without equipping them with the ability to mitigate interference. We also see that the performance improvement of CIA-MAC is low when the reuse factor is three and will be even more trivial with higher reuse factors. This is because a higher reuse factor results in less severe interference and interference avoidance is less desired. Therefore, CIA-MAC is good for networks with low reuse factors. Note that network deployment

with lower frequency reuse factors is a popular trend in the next-generation communication systems [12,34] for achieving higher spectrum efficiency and reducing network deployment cost.

3.4.3 Network Performance Improvement

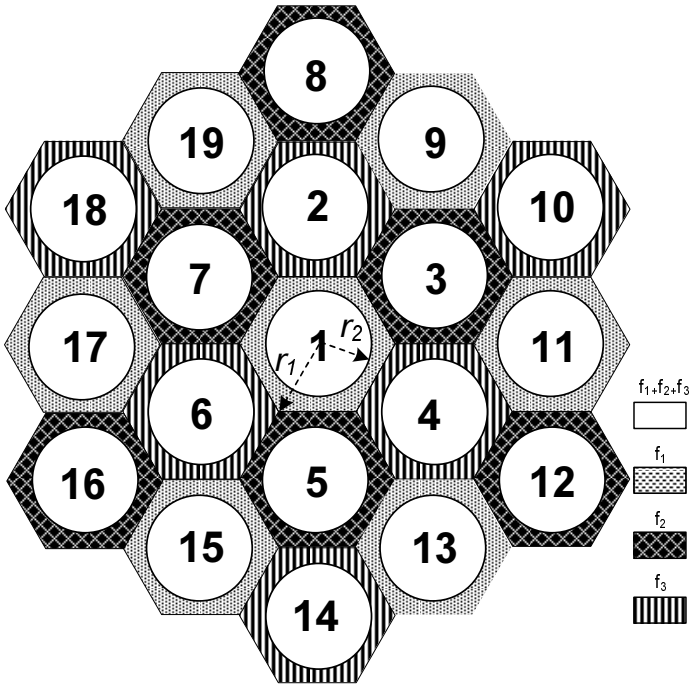


Figure 3.6. Cellular networks with fractional frequency reuse

In this section, we demonstrate the performance of CIA-MAC in a nineteen-cell cellular network. Each BS serves one user on a channel. Users are randomly dropped and uniformly distributed in each cell for each simulation trial.

CIA-MAC is implemented either with or without cross-layer design. The one with cross-layer design follows what we have discussed in this chapter and transmission happens only when channel power gain is above the threshold in Equation (3.2). For CIA-MAC without cross-layer design, each BS transmits randomly with probability given by Equation (3.3) and independently of channel states. We implement both trigger mechanisms and in the second one, the detection probability is set to be 0.9. We compare CIA-MAC with the traditional MAC and a static FFR and the overall system bandwidth is the same for

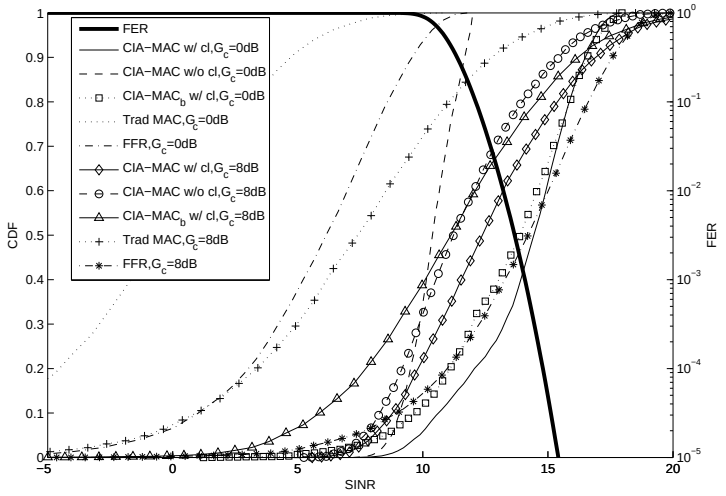


Figure 3.7. Cumulative distribution function of SINR.

all of them. The network has a reuse factor of one for both CIA-MAC and the traditional MAC. The traditional MAC keeps all BSs transmitting and users experience interference from all neighboring BSs. The static FFR reduces cell-edge interference through low frequency reuse at cell edges [38, 12, 39, 40, 41]. Figure 3.6 illustrates the network frequency deployment of FFR in our simulation. The radius of each cell is $r_1 = 2$ km and the cell-center users, located within $r_2 = 2r_1/3$ from the BS, will transmit over the whole frequency band. For cell-edge communications, the whole frequency band is equally divided into three subbands, f_1, f_2, and f_3, and users at cell edges are assigned one of them according to the frequency deployment in Figure 3.6. There are two interfering circumstances for FFR users. Cell-center users experience interference from all neighboring BSs. Neighboring BSs that use only one of the three subbands produce $1/3$ interference power since interference power distributes only in $1/3$ of the whole signal bandwidth. Cell-edge users experience interference from both the first-tier cells that are using the whole bandwidth and the second-tier cells that are using the same frequency subband as the center cell.

Note that the interference in Cell 1 that comes from the first-tier cells dominates the interference power and determines the performance of both the traditional MAC and FFR. Furthermore, the performance of CIA-MAC depends on the interference environments in

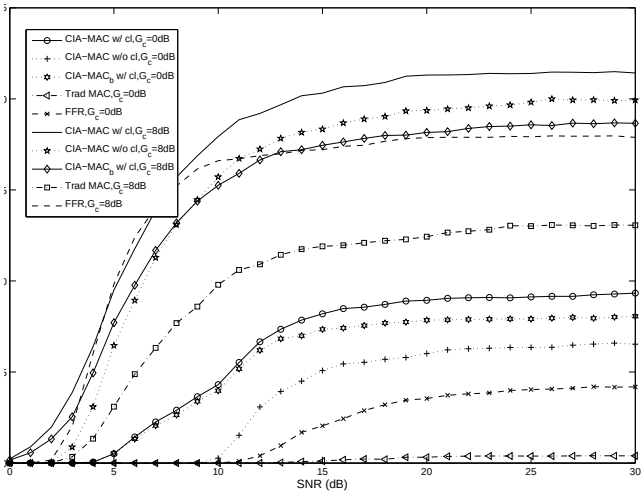


Figure 3.8. Throughput comparison

both the first-tier and second-tier cells according to Equations (3.2) and (3.5). Hence, the performances of CIA-MAC, traditional MAC, and FFR in Cell 1 are representative for their corresponding per-cell performance in general multi-cell cellular networks and we focus on the performance of cell 1. In Figures 3.7, 3.8, and 3.9, we compare the SINR and throughput of different schemes in Cell 1. Either uncoded 4-QAM or coded 4-QAM with different coding gains is used. The coding gain can be obtained through exploitation of receiver diversity or channel coding. In each trial, all BSs allocate the transmit power to maintain a constant received signal power level, i.e. keep a fixed SNR.

In Figure 3.7, we fix the SNR to be 12 dB and compare the *cumulative distribution functions* (CDFs) of average SINR of each trial. It includes statistics over $50,000$ trials. SINR is the equivalent value after decoding when coded 4-QAM is used. For reference, the relationship between FER and SINR is also plotted with a bold curve. Observing the FER curve, when SINR is lower than 10 dB, most decoded frames have at least one bit in error and do not pass CRC, resulting in transmission outage. In Figure 3.7, we compare the schemes with either uncoded 4-QAM or coded 4-QAM that has 8 dB coding gain. For CIA-MAC, only SINRs when no severe interferes transmit are averaged to produce the

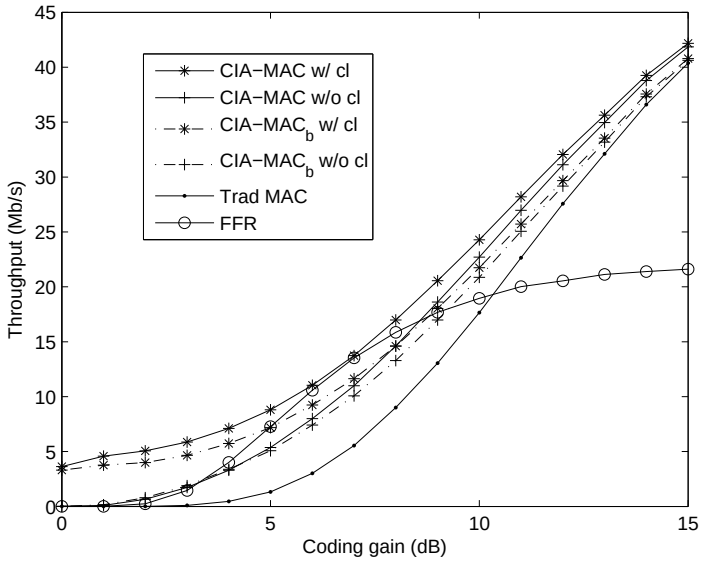


Figure 3.9. Throughput comparison (SNR=9 dB)

CDF curves. Curves with legend $CIA - MAC_b$ correspond to the performance of CIA-MAC when severe interferers are determined by the BSs according to (3.20). We can see that $CIA - MAC_b$ performs closely to $CIA - MAC$ in both cases and while significantly reducing the improvement cost, BS judgement is effective in detecting severe interferers using position knowledge. Without coding, the traditional MAC suffers strong interference and the average SINRs of all simulation trials falls far below 10 dB. In this case, the network is completely in outage. With FFR, the average SINRs of all trials are significantly improved. However, interference from neighboring cells still affects the SINRs and most SINRs are less than 10 dB since the target SNR is only 12 dB. Amazingly, CIA-MAC has better SINR distribution than FFR, even the one without cross-layer design. This is because we only average SINRs when no severe interferes transmit, i.e. when transmission succeeds in CIA-MAC. Furthermore, CIA-MAC with cross-layer design achieves very high average SINR. This is because the BS in Cell 1 also brings severe interference to many MTs in the neighboring cells, e.g. an average of 3.52 MTs in the simulation, resulting in a very high threshold in (3.2) and thus high average SINR. Now observe the performance of schemes with coded 4-QAM that has 8 dB coding gain. We can see that the SINR performance of

both FFR and the traditional MAC are improved by around 8 dB. Since MTs can mitigate a large amount of interference with coding, CIA-MAC judges much less severe interferers, e.g. 0.2 on average in our simulation. Hence, CIA-MAC finds no severe interferers in most cases and BSs simply keep on transmitting, as what the traditional MAC does. However, CIA-MAC still outperforms the traditional MAC because of avoidance of severe interference whenever it exists. Note that with high coding gain, CIA-MAC with cross-layer design has lower average SINR as compared with uncoded cases. This is because with 8 dB coding gain, only few severely interfered MTs are judged in the most trials and the BS is allowed to transmit in most channel conditions rather than very good ones. However, SINR performance does not solely determine the network performance, which also relate to spectral reuse efficiency, transmission probability, and so on. For example, compare the performance of CIA-MAC with cross-layer design when either coded or uncoded 4-QAM is used. The SINR performance with uncoded 4-QAM even outperforms the coded one with 8 dB coding gain. However, it is at the price of extremely low probability to transmit with only peak channel conditions. With high coding gain, BSs produce interference bearable to neighboring-cell MTs and are allowed to transmit at high probability. Both SINR and transmission probability impact the throughput, which is the performance we desire to improve. In the following, we further compare the network throughput.

In Figure 3.8, we compare the average throughput when the target SNR has different values. For each SNR, the throughput is the average over $50,000$ trials. From the figure, for a system with uncoded 4-QAM, the traditional MAC has almost no throughput improvement when SNR is increased. This is because with increased SNR, transmit powers of neighboring BSs also increase, resulting in stronger interference and trivial SINR improvement. This indicates the necessity of tackling CCI for this highly aggressive frequency reuse scenario. Through frequency reuse at cell edges, FFR successfully reduces interference. With FFR, the average SINR improves as SNR increases and thus higher throughput is obtained. However, in the low-SNR region, interference significantly affects

frame reception due to weak signal power and even with frequency reuse at cell edges, FFR still suffers from neighboring-cell interference. We can see that CIA-MAC schemes significantly outperforms all other schemes. This is due to the intelligent interference avoidance and full frequency reuse in CIA-MAC. With cross-layer design, transmission happens only with high channel power gain, which further improves throughput as compared with the one without cross-layer design. Note that the CIA-MAC with BS determining severe interferers performs closely to the one with MT, indicating good detection capability of BSs. Similar to the traditional MAC, CIA-MAC also suffers a saturation effect in the high-SNR region, i.e. increasing transmit power does not necessarily result in improved throughput because of proportionally increased interference power. With 8 dB coding gain, mobiles can mitigate a large portion of interference and all schemes have significant performance enhancement. We note that our proposed schemes, both with and without cross-layer design, still outperform the traditional MAC comprehensively due to the intelligent recognition of severe interferers for interference avoidance. The throughput of FFR increases with SNR in the low-SNR region due to increased SINR. In the high-SNR region, FFR throughput does no improve with SNR since all frames have been correctly received. However, in the high-SNR region, FFR performs not as well as the proposed CIA-MAC due to its low spectrum reuse efficiency at cell edges. Figure 3.9 further demonstrates this point. In Figure 3.9, each BS allocates power to maintain a 9 dB received SNR and the coding gains of all MTs are increased from 0 dB to 15 dB. We observe that CIA-MAC always outperforms the traditional MAC. With higher and higher coding gain, interference has less and less impact on frame reception and thus fewer and fewer severe interferers are judged. Hence, with high coding gain, the performance of CIA-MAC and traditional MAC tend to be the same. The static FFR suffers performance loss for low frequency-reuse efficiency at cell edges in the high-coding-gain region. Dynamic FFR schemes [39, 40, 41] can be used to further improve frequency-reuse efficiency at the cost of higher network deployment complexity. However, we should note that CIA-MAC improves the network performance in the

most cost-effective way and even the static FFR implemented here requires much higher deployment cost than CIA-MAC.

3.5 Summary

This chapter provides a low-cost solution to improve the performance for cell-edge MTs that are experiencing severe CCI in wireless cellular networks. The proposed CIA-MAC requires semi-static information exchange among BSs and automatically randomizes transmission to improve QoS for severely interfered MTs. The principle for triggering CIA-MAC is investigated and two simple trigger mechanisms are described. The proposed scheme significantly improves communication performance for MTs experiencing severe CCI because of intelligent recognition of severe interferers and the corresponding interference avoidance.

CHAPTER 4

CHANNEL AWARE DISTRIBUTED MEDIUM ACCESS CONTROL

Both DOMRA and CIA-MAC are Aloha based and once a collision happens, the entire data frame has to be dropped. The performance can be further improved through scheduling users in a distributed way to avoid the collision of data transmission. Hence, in this chapter, we continue investigating *channel-aware distributed medium access control* (CAD-MAC) for wireless networks with arbitrary topologies and traffic distributions, where users can receive traffic from or send traffic to different users and different communication links may interfere with each other. We consider heterogeneous channels, where the random channel gains of different links may have different distributions. To resolve the network contention in a distributed way, each frame is divided into contention and transmission periods. The contention period is used to resolve conflicts while the transmission period is used to send payload in collision-free scenarios. We design a channel-aware Aloha scheme for the contention period to enable users with relatively better channel states to have higher probability of contention success while assuring fairness among all users. With this approach users are scheduled in a distributed way. We show analytically that the proposed scheme completely resolves network contention and achieves throughput close to that using centralized schedulers. Besides, this scheme is also robust to any channel uncertainty. Simulation results demonstrate that the proposed scheme significantly improves network performance. Same as DOMRA, the proposed random access approach can be applied to different wireless networks, such as cellular, sensor, and mobile *ad hoc* networks, to improve quality of service.

The rest of this chapter is organized as follows. First we describe the system in Section 4.1. In Section 4.2, we design the channel aware distributed medium access scheme. Then in Section 4.3, we optimize the operation of CAD-MAC. The robustness of CAD-MAC is analyzed in Section 4.4. Finally, we demonstrate the performance improvement with simulations in Section 4.5 and conclude this chapter in Section 4.6.

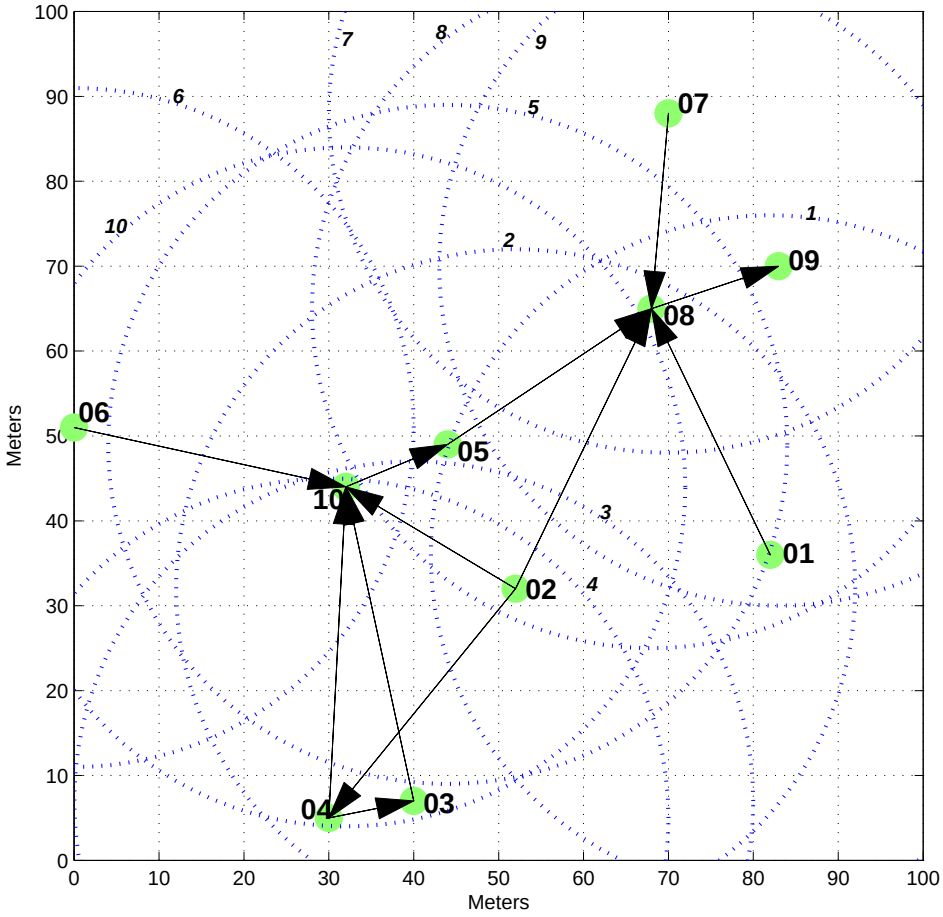


Figure 4.1. A network example.

4.1 System Description

Consider a network where users are not necessarily within the transmission ranges of all others, that is, some users may not be able to receive packets from others due to weak received signal power. All channels are assumed to be reciprocal when there is no interference. Each user has knowledge of its own CSI and makes an independent decision on its transmission. A receiver cannot decode any packet successfully if the channel is simultaneously used by another user within the transmission range of the receiver, i.e., a collision

happens. Each user may choose to send packets to or receive packets from different users. An example is illustrated in Figure 4.1, where arrows indicate traffic flows and dashed circles, marked by italic numbers, denote the transmission ranges of the corresponding users[1].

The process of the proposed channel-aware random access is illustrated in Figure 4.2. Each user has a queue with an infinite length for each traffic flow that needs to be sent and we assume the queue always has packets to be delivered. A dequeue controller fetches a desired amount of data and send it to the transmitter following the order of the medium access controller. The medium access controller collects information on channel states and decides when and how to transmit.

The backoff-after-collision approach in traditional CSMA can resolve contention. However, it ignores channel and multiuser diversity in wireless communications and deferring transmission without considering channel variations may result in data communications in deep fades. To fully exploit network diversity, the contention should be designed such that users with favorable channel conditions have higher probability of accessing the channels and the transmission should follow immediately after the contention resolution as otherwise the channel may change to an unfavorable state. Considering this, we design a new distributed random access scheme in the following sections. Since this novel scheme uses channel knowledge to improve network performance, we call it *channel-aware distributed medium access control* (CAD-MAC).

4.2 Channel-Aware Medium Access Control

As shown in Figure 4.2, the channel access time is divided into frame slots of length, T_f, and each slot consists of both contention and transmission periods. Block fading is assumed [88], that is, the channel state remains constant within each frame slot and is independent from one to another. The contention period is further divided into a maximum of $\widehat{K}$ *contention resolution slots* (CRSs) of length T_c, each for one contention resolution.

[1]Without loss of generality, we assume that the transmission and interference ranges are identical.

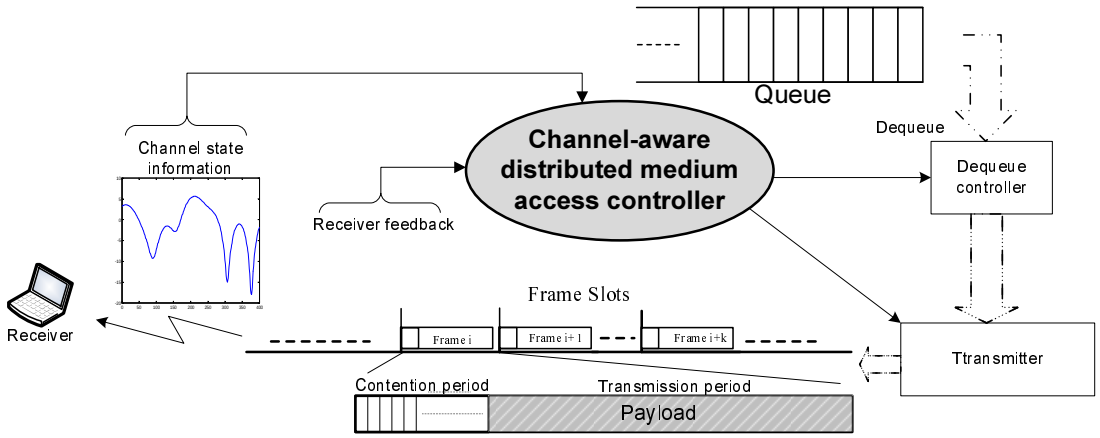


Figure 4.2. Traffic, energy, and channel aware medium access.

Users failing in all CRSs will be idle in the current frame slot. Users that succeed in any CRS will send data in that frame slot with optimized link adaptation. The actual number of CRSs may vary from frame to frame, depending on the contention results. The objective of the contention design is to select users with relatively better channel conditions for payload transmission and the selection should also assure fairness among all users. In this way, the network diversity can be exploited sufficiently. We use CSI to control the access contention and the contention is randomized because wireless channels are inherently random. In the following, let h_{ij} be the channel gain of Link (i, j), the one from User i to j, with probability density function $f_{ij}(h)$ and distribution function $F_{ij}(h)$. Both $f_{ij}(h)$ and $F_{ij}(h)$ are assumed to be continuous to facilitate our discussion. Here we assume that the channel gains of different links are independent but not necessarily identically distributed.

There are two types of contention. We denote Type-I and Type-II to be those among links with the same transmitter and with different transmitters, respectively. For example, the contention between Links $(2, 4)$, $(2, 8)$, and $(2, 10)$ in Figure 4.1 is Type-I and the contention between Links $(2, 4)$ and $(4, 3)$ is Type-II. Here we do not consider the case that two users are sending traffic to each other since the reciprocal channel between them is always the same for their transmission and they can negotiate easily to share the channel, e.g. in a time division fashion.

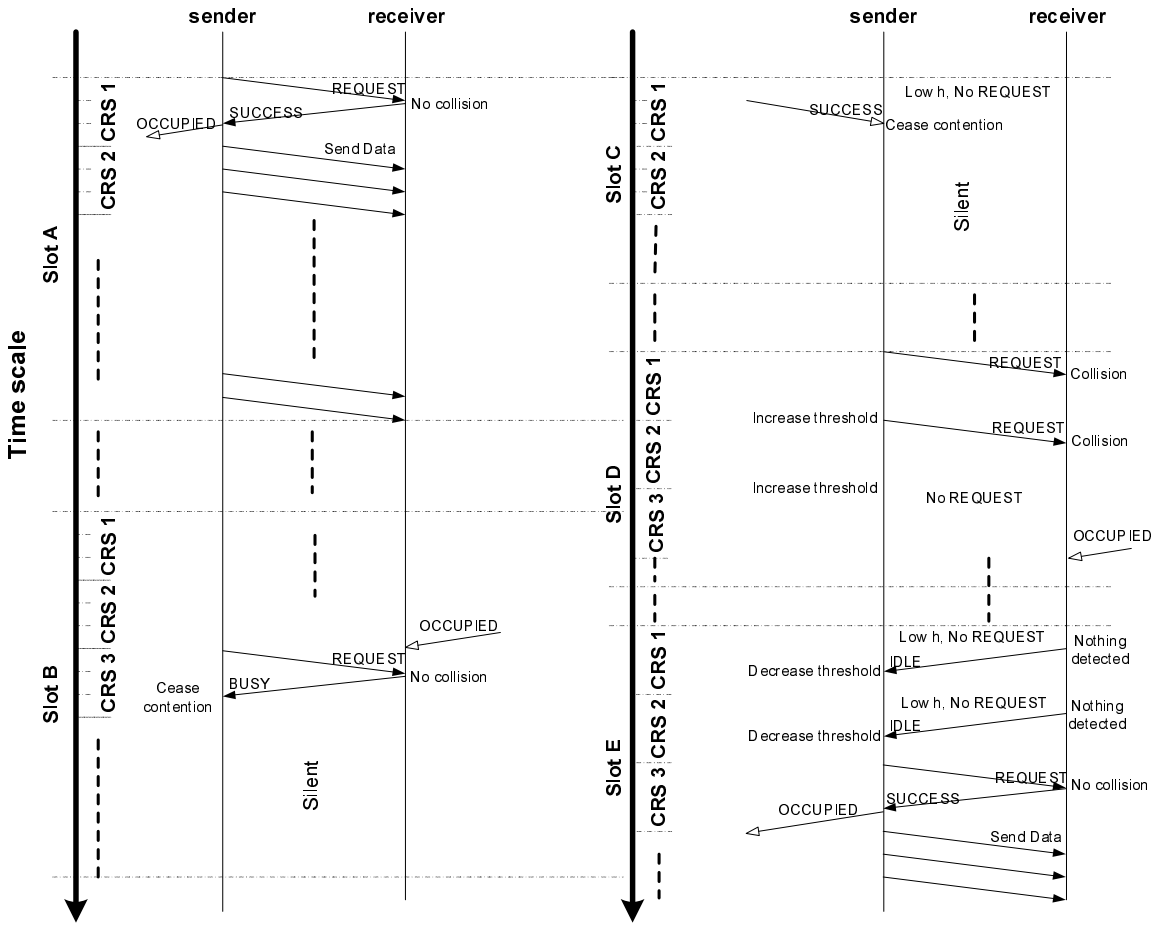


Figure 4.3. Flowcharts of typical access contention.

The Type-I contention can be easily resolved by the transmitter as it has CSI of all links and choosing the one with the best CSI will result in the best system performance while assuring fairness, i.e., User i chooses Link (i, j) that satisfies

$$j = \arg\max_{l} F_{il}(h_{il}). \tag{4.1}$$

Note that $F_{il}(h_{il})$ is the probability that the channel gain of Link (i, l) is worse than h_{il}. The link with the highest $F_{il}(h_{il})$ is the one with the best instantaneous channel condition relatively and criterion (4.1) effectively exploits the instantaneous multiuser diversity. Furthermore, $F_{il}(h_{il})$ is uniformly distributed between 0 and 1 for all (i, l). Hence, these links have the same probability of being scheduled and the scheme is fair.

We focus on resolving Type-II contention. Random access is needed and a link with a

better channel state should have a higher probability of success. The contention period is used to resolve this type of contention. The basic idea is to resolve the contention from one CRS to another and in each CRS, links with higher gains are selected in a distributed way to continue the following contention. Finally, only one link is selected within each local area and all interferers are informed that they should not send any data in the current frame slot. To facilitate the discussion of Type-II contention, REQUEST, BUSY, SUCCESS, IDLE, and OCCUPIED signals are defined as follows.

- REQUEST: sent by a transmitter to request access;
- BUSY: sent by a receiver to deny access;
- SUCCESS: sent by a receiver to allow access;
- IDLE: sent by a receiver to petition for access;
- OCCUPIED: sent by a transmitter to prevent neighbors from data reception.

Each CRS consists of the following three steps.

1. *Transmitters send REQUEST*: If User i has neither received a BUSY signal from j nor detected a SUCCESS signal destined to others, and

$$h_{ij} > \widehat{H}_{ij}[k], \tag{4.2}$$

   where $\widehat{H}_{ij}[k]$ is a predetermined threshold that is adjusted CRS-by-CRS, then it sends REQUEST to User j.

2. *Receivers notify BUSY, SUCCESS, IDLE*:

   - BUSY: User j responds BUSY if it receives REQUEST correctly and has received OCCUPIED in the previous CRSs.
   - SUCCESS: User j responds SUCCESS if the REQUEST is received correctly and no OCCUPIED signals received in the previous CRSs.

- IDLE: User j broadcasts IDLE to all users that want to send traffic to User j if no OCCUPIED signals received in the previous CRSs and no signals detected at Step 1.

Note that the BUSY or SUCCESS feedback is sent only when there is no collision, i.e., the contention succeeds.

3. *Transmitters broadcast OCCUPIED and start sending data*: If User i has received SUCCESS, it goes to the win state and broadcasts OCCUPIED to notify those within its transmission range that they should not receive data in this frame slot.

Five typical contention processes have been illustrated in Figure 4.3, where the solid arrows indicate signals between the observed pair of users and the empty arrows indicate signals sent from or detected by the interfering neighbors. As an example, observe the contention among only Links $(6, 10)$, $(10, 5)$, and $(8, 9)$ in Figure 4.1. If all the three links have good channel gains and send REQUEST in CRS 1, only User 9 receives REQUEST without collision and it sends back SUCCESS to User 8 at the second step while Users 5 and 10 remain silent. At the third step, User 8 broadcasts OCCUPIED. Then CRS 2 starts. Users 6 and 10 may still send REQUEST, depending on the adjusted threshold. Suppose both send and only User 5 receives a collision-free REQUEST. At the second step, User 5 responds BUSY to User 10. Nothing happens at Step 3. In CRS 3, only User 6 may still send REQUEST and User 10 will respond BUSY to prevent subsequent contention behaviors.

Remark 1: At Step 2, the BUSY or SUCCESS signals can always be received by User i correctly. This can be justified as follows. Suppose Links (i_1, j_1) and (i_2, j_2) succeed in their contention and Users j_1 and j_2 are sending BUSY or SUCCESS signals to i_1 and i_2 respectively. User i_1 does not interfere with j_2 and hence it can not receive any signal from j_2 since the channel is assumed to be reciprocal. Hence, User i_1 can receive the BUSY or SUCCESS signal without interference from j_2. Similarly User i_2 also receives the BUSY

or SUCCESS signals correctly. On the other hand, The IDLE signals from different links may collide. Since only IDLE signals may collide at Step 2, users can detect them if they are neither BUSY nor SUCCESS signals. In the following, we assume the IDLE signals are received correctly.

Remark 2: At Step 3, the OCCUPIED signals may collide. However, as only the OCCUPIED signals are broadcasted and if any signal is detected, it will be the OCCUPIED signal.

4.3 Access Optimization

In this section, we optimize the access parameters. The following notations are used. All links carrying traffic are denoted by set $\mathcal{L}[1] = \{(i, j)\}$. Denote the interfering neighbor set of User i by $\mathcal{N}_i$. Each user may choose to send packets to or receive packets from several users, with $\mathcal{T}_i$ the set of users receiving packets from i and $\mathcal{S}_j$ the set of users sending packets to j. For example, $\mathcal{N}_4 = \{2, 3, 10\}$, $\mathcal{T}_4 = \{3, 10\}$, and $\mathcal{S}_4 = \{2\}$ in Figure 4.1.

We desire to optimize the throughputs of all users in the network. The arithmetic-mean metric leads to the design for sum throughput maximization, but assures no fairness since some users may have zero throughput. The geometric-mean metric takes both throughput and fairness among all users [94] into consideration. Therefore, we will find the thresholds in (4.2) to maximize the geometric mean of the throughputs of all links, i.e.,

$$\{\widehat{H}^*_{ij}[k]\} = \underset{\{\widehat{H}_{ij}[k]\}}{\operatorname{argmax}} \prod_{(i,j)} T_{ij} = \underset{\{\widehat{H}_{ij}[k]\}}{\operatorname{argmax}} \sum_{(i,j)} \log(T_{ij}), \tag{4.3}$$

where T_{ij} is the average throughput of Link (i, j).

It is not feasible to globally optimize (4.3) because after the contention in each CRS, new local knowledge is collected according to receiver feedback and the detection of signals broadcasted from neighboring users. This knowledge is generally different from one CRS to another and can not be obtained in advance. To fully exploit this knowledge, the contention will be optimized sequentially, i.e., in a CRS-by-CRS way, and use newly collected knowledge to improve the contention behaviors afterward.

In the following, denote the probability that User i sends a REQUEST to User j in CRS k by $p_{ij}[k]$. The overall probability that User i sends REQUESTs to other users in CRS k is

$$p_i[k] = \sum_{j \in \mathcal{T}_i} p_{ij}[k]. \tag{4.4}$$

4.3.1 CRS 1

We first optimize CRS 1. The throughput on Link (i, j) out of CRS 1 is

$$T_{ij}[1] = R_{ij} p_{ij}[1](1 - p_j[1]) \prod_{m \in \mathcal{N}_j, m \neq i} (1 - p_m[1]), \tag{4.5}$$

where R_{ij} is the average data rate of payload transmission; $(1 - p_j[1]) \prod_{m \in \mathcal{N}_j, m \neq i}(1 - p_m[1])$ is the probability that neither user j nor its neighboring users except user i transmits, which means the successful contention of Link (i, j) in CRS 1. In Figure 4.1, the transmission from User 2 to User 4 succeeds only when neither User 4 nor its neighbors excluding User 2, i.e., users in $\mathcal{N}_4 \backslash \{2\} = \{3, 10\}$, transmit. Hence, $T_{2,4}[1] = R_{2,4} p_{2,4}[1](1 - p_4[1])(1 - p_3[1])(1 - p_{10}[1])$.

The contention probability for CRS 1 is given by

$$\{p_{ij}^*[1]\} = \arg \max_{\{p_{ij}[1]\}} \sum_{(i,j) \in \mathcal{L}[1]} \log(T_{ij}[1]). \tag{4.6}$$

Both $\log(p_{ij}[1])$ and $\log(1 - p_i[1]) = \log(1 - \sum_{j \in \mathcal{T}_i} p_{ij}[1])$ are strictly concave functions of $p_{ij}[1]$. Hence $\sum_{(i,j) \in \mathcal{L}[1]} \log(T_{ij}[1])$ is strictly concave in $\{p_{ij}[1]\}$ and a unique global optimal $\{p_{ij}^*[1]\}$ can be determined by setting the first-order derivative of the objective function to be zero. The optimal contention probability can be readily obtained after some mathematical manipulations and

$$p_{ij}^*[1] = \frac{1}{|\mathcal{S}_i| + \sum_{m \in \mathcal{N}_i} |\mathcal{S}_m|}, \tag{4.7}$$

which is the inverse of the total number of received traffic flows within the interference range of User i. Intuitively, $p_{ij}^*[1]$ says that as the interference footprint (number of affected users) increases, the contention probability of User i should decrease.

The threshold should be chosen to satisfy the contention probability in (4.7). According to Section 4.2, the contention probability of Link (i, j) is

$$\begin{aligned} p_{ij}[1] &= \mathtt{Pr}\{(i,j)\text{is chosen}; h_{ij} > \widehat{H}_{ij}[1]\} \\ &= \int_{\widehat{H}_{ij}[1]}^{\infty} f_{ij}(h)\mathrm{Pr}(j = \arg\max_{l\in\mathcal{T}_i} F_{il}(h_{il}))\mathrm{d}h \\ &= \int_{\widehat{H}_{ij}[1]}^{\infty} \mathrm{Pr}(F_{il}(h_{il}) < F_{ij}(h_{ij}) : l \neq j)\mathrm{d}F_{ij}(h) \\ &= \frac{1}{|\mathcal{T}_i|}\left(1 - F_{ij}^{|\mathcal{T}_i|}(\widehat{H}_{ij}[1])\right), \end{aligned} \tag{4.8}$$

where $|\cdot|$ denotes the number of elements in the set.

From (4.8) and (4.7), the optimal threshold is

$$\widehat{H}_{ij}^{*}[1] = F_{ij}^{-1}\left[\left(1 - \frac{|\mathcal{T}_i|}{|\mathcal{S}_i| + \sum_{m\in\mathcal{N}_i}|\mathcal{S}_m|}\right)^{\frac{1}{|\mathcal{T}_i|}}\right]. \tag{4.9}$$

The optimal threshold (4.9) depends on the number of users receiving packets from User i, $|\mathcal{T}_i|$, the number of users sending packets to User i, $|\mathcal{S}_i|$, and the total number of users sending packets to the interfering neighbors of User i, $\sum_{m\in\mathcal{N}_i}|\mathcal{S}_m|$. The first two require only local knowledge while the third can be obtained through signalling exchange. This exchange incurs only trivial signalling overhead since it will be triggered only when either a traffic session or the network topology changes sufficiently. Besides, this type of knowledge is typical in many protocols, such as routing discovery in mobile *ad hoc* networks [81, 82]. Hence, it can be readily obtained. Consider User 4 in Figure 4.1. $|\mathcal{T}_4| = 2$, $|\mathcal{S}_4| = 1$, $|\mathcal{S}_2| = 0$, $|\mathcal{S}_3| = 1$, and $|\mathcal{S}_{10}| = 4$. Hence, $\widehat{H}_{4,3}^{*}[1] = F_{4,3}^{-1}\left[(1 - \frac{2}{1+1+4})^{1/2}\right] = F_{4,5}^{-1}(0.667)$. If Link $(4, 3)$ experiences Rayleigh fading with average gain h_a, $\widehat{H}_{4,5}^{*}[1] = 1.1h_a$.

4.3.2 CRS k, $k > 1$

In the following CRSs, links whose transmitters have not been notified SUCCESS or BUSY continue the contention. The new threshold is chosen such that the contention probability is $p_{ij}[k]$. There are three possibilities adjusting the threshold.

- *Adjustment* (AD) I: If in the previous CRS, User i sent a REQUEST and no feedback is received, indicating a collision, all links involved in this collision should increase

their thresholds to reduce the probability of collision. From previous knowledge, $h_{ij} > \widehat{H}^*_{ij}[k-1]$ and $h_{ij} < \widehat{H}^M_{ij}$, where $\widehat{H}^M_{ij}$ is the minimum threshold in all the previous CRSs such that $h_{ij} < \widehat{H}^M_{ij}$ and initially $\widehat{H}^M_{ij} = \infty$. The new threshold satisfies

$$\Pr\left(h_{ij} > \widehat{H}^*_{ij}[k] \middle| \; h_{ij} > \widehat{H}^*_{ij}[k-1], h_{ij} < \widehat{H}^M_{ij}\right) = p_{ij}[k]. \tag{4.10}$$

Solving Equation (4.10) for $\widehat{H}^*_{ij}[k]$, we have

$$\begin{aligned} \widehat{H}^*_{ij}[k] = & F^{-1}_{ij}\Big((1 - p_{ij}[k])F_{ij}(\widehat{H}^M_{ij}) \\ & + p_{ij}[k] \cdot F_{ij}(\widehat{H}^*_{ij}[k-1])\Big). \end{aligned} \tag{4.11}$$

- AD II: If User i applied AD I or II, did not send REQUEST, and received IDLE from j in the previous CRS, indicating User i is still contending and all other contending users, if any, have channel states below their thresholds, User i should decrease the threshold. Similar to the first case, the new threshold satisfies

$$\Pr\left(h_{ij} > \widehat{H}^*_{ij}[k] \middle| \, h_{ij} < \widehat{H}^*_{ij}[k-1]; h_{ij} > \widehat{H}^m_{ij}\right) = p_{ij}[k], \tag{4.12}$$

where $\widehat{H}^m_{ij}$ is the maximum threshold in all the previous CRSs such that $h_{ij} > \widehat{H}^m_{ij}$ and initially $\widehat{H}^m_{ij} = 0$. Solving equation (4.12), we have

$$\begin{aligned} \widehat{H}^*_{ij}[k] = & \; F^{-1}_{ij}\Big(p_{ij}[k] \cdot F_{ij}(\widehat{H}^m_{ij}) \\ & + (1 - p_{ij}[k])F_{ij}(\widehat{H}^*_{ij}[k-1])\Big). \end{aligned} \tag{4.13}$$

- AD III: In other cases, the threshold is kept the same, i.e.,

$$\widehat{H}^*_{ij}[k] = \widehat{H}^*_{ij}[k-1]. \tag{4.14}$$

This usually happens when no REQUEST was sent and no IDLE was received in a previous CRS and User i temporarily quits the contention. In this case, User i would contend again only if it receives IDLE in the future CRSs.

Denote all the competing links in CRS k by $\mathcal{L}[k]$. With the same approach as in CRS 1, the optimal contention probability for $(i, j) \in \mathcal{L}[k]$ is

$$p^*_{ij}[k] = \frac{1}{|\mathcal{S}_i[k]| + \sum_{m \in \mathcal{N}_i[k]} |\mathcal{S}_m[k]|}, \tag{4.15}$$

where $\mathcal{S}_n[k]$ and $\mathcal{N}_n[k]$ are users that can contend in CRS k. A user may contend if and only if its threshold will be changed as in ADs I or II. However, who will adjust their thresholds is unknown to others and $p^*_{ij}[k]$ cannot be determined locally. Instead, we give a suboptimal approach as follows

$$p_{ij}[k] = \begin{cases} \frac{1}{2}, & \text{AD I,} \\ p_{ij}[k-1], & \text{AD II.} \end{cases} \tag{4.16}$$

Here we assign one half for AD I because after the selection in CRS 1, it is most likely that only one other link is contending with Link (i, j) if a collision happens. For AD II, an IDLE signal most likely indicates that the contention scenario is not changed and $p_{ij}[k]$ keeps the same.

4.4 Robustness Analysis

In this section, we analyze the robustness of CAD-MAC. We say a link wins the contention if it transmits data in the transmission period in the following.

The complete resolution of network contention is defined as follows.

Definition 4.4.1 *The contention of a network is completely resolved if*

1. *all links that have won the contention can transmit without collision;*
2. *if any additional link that has not won the contention transmits, it will collide with at least one link that has won the contention.*

Thus, complete resolution results are states in which the network capacity is fully exploited. The following theorem states that CAD-MAC can completely resolve the network contention and is proved in Appendix B.1.

Theorem 4.4.2 *With probability one, the contention of networks with any topology can be completely resolved by CAD-MAC if sufficient CRSs are allowed.*

One example that CAD-MAC fails to resolve the contention is shown in Figure 4.4 where the two channels are *independent and identically distributed* (i.i.d.). When $h_{12} = h_{34}$, Users 1 and 3 have the same update of the thresholds and their REQUESTs always collide. However, the probability that $h_{12} = h_{34}$ is zero because the two channels are independently fading with continuous probability distribution function $F_{ij}(h)$.

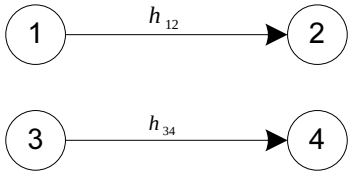


Figure 4.4. A network in which all interfere with others.

Theorem 4.4.2 indicates that CAD-MAC achieves performance comparable to that of a centralized scheduler. Compared to the centralized scheduler, CAD-MAC loses throughput due to the CRSs used for resolving network contention. Denote the throughputs of CAD-MAC and the centralized scheduler by $T_{CAD-MAC}$ and $T_{Centralized}$, respectively. Then we define the efficiency, γ, of CAD-MAC as follows,

$$\gamma = \frac{T_{CAD-MAC}}{T_{Centralized}} = 1 - \frac{\overline{K}T_c}{T_f}, \tag{4.17}$$

where $\overline{K}$ is the average number of CRSs necessary for completely resolving the network contention. In the following, we show that $\overline{K}$ is bounded regardless of the network type and size. To simplify the analysis, we assume in the following that a link contends again only if all neighbors of the receiver have resolved their contention and the receiver sends IDLE to the receiver since it can still receive data. Besides, assume sufficient CRSs.

First, consider the case that each link interferes with all others and only one link wins the contention in each frame slot, such as in a network where all users send traffic to a common receiver or a small-scale *ad hoc* network where each user is within the transmission range of all others. For a network with N traffic flows, each interfering with all others, an upper bound of $\overline{K}_N$ is given by the following theorem, which is proved in Appendix B.2.

Theorem 4.4.3 *For a network with N links, each interfering with all others, the average number of CRSs necessary to completely resolve the network contention satisfies*

$$\overline{K}_N \leq \frac{\widehat{M}_N}{1-(1-\frac{1}{N})^N} + \frac{(1-\frac{1}{N})^N}{(1-(1-\frac{1}{N})^N)^2}, \tag{4.18}$$

where $\widehat{M}_N = \sum_{n=1}^{N} \binom{N}{n} (\frac{1}{N})^n (1-\frac{1}{N})^{N-n} (\log_2(n)+1)$. *Furthermore,*

$$\overline{K}_N < \overline{K}_\infty \leq 2.43. \tag{4.19}$$

Based on Theorem 4.4.3, the following theorem gives a general upper bound of $\overline{K}$ for any type of networks and is proved in Appendix B.3.

Theorem 4.4.4 *For any type and size of network, the average number of CRSs necessary to completely resolve the contention satisfies*

$$\overline{K} < \frac{2.43 \cdot \overline{L}}{\beta}, \tag{4.20}$$

where the transmission coexistence factor, $\overline{L}$*, is the average number of links that win the contention in one frame slot and the contention coexistence factor,* β*, is the average number of simultaneous resolutions in each CRS.*

Remark 1: In Theorem 4.4.4, the contention coexistence factor β indicates how many simultaneous resolutions occur in each CRS. Here one resolution is the process that all links, among whom only one link will win, adjust their thresholds using ADs I or II. Since multiple links may win in one frame slot, the resolutions that lead to the win of these links may happen in the same CRS, and β characterizes this overlap. Readers are referred to Appendix B.3 for the strict definition of β. Obviously, both $\overline{L}$ and β depend on the distribution density and transmission range of all users.

For example, if each user interferes with all others and only one link wins, then $L = 1$, $\beta = 1$, and $\overline{K} < 2.425$ as in Theorem 4.4.3. If a network consists of 2 groups of users and the communication within different groups does not interfere with each other, then these two groups can resolve their contention within themselves to produce the two winners.

Consequently, $L = 2$. For example, denote the CRSs for a two-cell cellular network using different frequency sets in the two cells to resolve the contention to be $\mathcal{K}_1 = \{1, 2, \cdots, k_1\}$ and $\mathcal{K}_2 = \{1, 2, \cdots, k_2\}$, respectively, where k_1 and k_2 are random and vary from one frame to another. Then the resolution overlaps from CRSs 1 to $\min\{k_1, k_2\}$ and there is only one resolution from CRSs $\min\{k_1, k_2\} + 1$ to $\max\{k_1, k_2\}$. Consequently, according to Appendix B.3,

$$\beta = \frac{\mathbf{E}(k_1 + k_2)}{\mathbf{E}(\max\{k_1, k_2\})}. \tag{4.21}$$

and

$$\overline{K} < \frac{4.86 \cdot \mathbf{E}(\max\{k_1, k_2\})}{\mathbf{E}(k_1 + k_2)}. \tag{4.22}$$

From Theorems 4.4.2 and 4.4.4, we have the following proposition.

Proposition 4.4.5 *The efficiency of CAD-MAC satisfies*

$$\gamma > 1 - \frac{2.43 \cdot \overline{L} T_c}{\beta T_f}. \tag{4.23}$$

For a network where each user interferer with all others, the efficiency is

$$\gamma > 1 - \frac{2.43 \cdot T_c}{T_f}. \tag{4.24}$$

T_c and T_f are determined by the round-trip time of signal propagation and the channel coherence time respectively. If $T_f \gg T_c$ as in slow-fading channels, CAD-MAC performs almost the same as the centralized scheduler, which is generally impractical because of poor scalability and the huge overhead of CSI collection. For example, it is shown in [95] that the round trip time for 802.11 WLAN is within 10 μs and for cellular networks, with 6 km radius, is within 50μs. On the other hand, the channel coherence time is hundreds of milliseconds in indoor office or home environment and tens of milliseconds in cellular networks with 900 MHz carrier frequency and user speed 72 km/h [96]. Hence in both WLAN and cellular networks, the efficiency of CAD-MAC is close to unity.

Now suppose that that all users have imperfect channel state information $\{\widetilde{h}_{ij}\}$ and $\{\widetilde{F}_{ij}()\}$ and control the medium access. From the proofs of Theorems 4.4.2, 4.4.3, and 4.4.4, we

see that they are independent of the channel distribution of any user. Hence, they also hold for the operations of CAD-MAC based on $\{\widetilde{h}_{ij}\}$ and $\{\widetilde{F}_{ij}()\}$. Besides, suppose the centralized scheduler compared in (4.17) has the same imperfect channel knowledge. Then the efficiency of CAD-MAC is still given by (4.17). Therefore we have the following theorem about the robustness of CAD-MAC.

Theorem 4.4.6 *The conclusions in Theorems 4.4.2, 4.4.3, and 4.4.4 and Proposition 4.4.5 hold when all users have imperfect channel knowledge and CAD-MAC is robust to any channel uncertainty.*

4.5 Simulation Experiments

In this section, we demonstrate the performance of CAD-MAC in a network with random topologies. First we illustrate how CAD-MAC operates given a network instance. Then we show the cumulative distribution function of the number of CRSs that are used to completely resolve the network contention. Finally we compare the performance of CAD-MAC with the Aloha-based *decentralized optimization for multichannel random access* (DOMRA) scheme in [97], which also uses channel gains to optimize the access contention while assuring proportional fairness for the type of networks considered in this paper.

In each simulation trial, users are randomly dropped and uniformly distributed in a square area with side length of 100 meters. Each user has a transmission range of 40 meters and selects neighboring users randomly for data transmission. The number of selected receivers is uniformly distributed between 1 and half of the number of neighboring users. A network topology in one trial has been illustrated in Figure 4.1. Rayleigh block fading channel with the average fading level, h_o, is assumed. Hence, $F(h) = 1 - e^{-\frac{h}{h_o}}$. The data rate in each frame is given by $R(h) = W \ln(1 + \frac{hP}{N_o})$, where $W = 100$ KHz, is the system bandwidth, $P = 0.01$ watt, is the transmit power, and $N_o = 0.0001$ watt, is the noise power. The channel gains are independent with either the same or different averages. For homogeneous channels, $h_o = 1$ and for heterogeneous ones, h_o is uniformly distributed

between 0.5 and 1.5. The length of each frame slot is 20 ms and the CRS is 0.2 ms each.

First, consider the network topology given in Figure 4.1 and let $h_o = 1$. The contention process for a set of channel states in a frame slot is illustrated by Table 4.1, where blanks indicate no values or no actions. Each user chooses a receiver with the best channel gain, e.g., User 2 selects User 4. In the first CRS, Links $(4, 3)$ gets access. Users $2, 3$, and 10 detect SUCCESS and decide to stop contention since some neighboring users will receive data in this frame slot. In the second CRS, only Users 1, 5, 7, and 8 contend but none send REQUEST even their thresholds are lowered. In the third CRS, only User 7 sends REQUEST and wins the contention. Hence, three CRSs completely resolve the Type-II contention and Links $(4, 3)$ and $(7, 8)$ will send data in this frame slot. Note that this result also fully exploits the network capacity as transmission of any other users will produce interference and reduce the network throughput.

Table 4.1. Contention process for a set of channel states in Figure 4.1

User i		1	2	3	4	5	6	7	8	9	10
Receivers		8	4;8;10	10	3;10	8	10	8	9		5
Channel gains h		**0.66**	**1.36**;0.63;0.61	**0.91**	**2.98**;1.36	**0.49**	**1.33**	**0.94**	**0.23**		**0.11**
Selected receiver j		8	4	10	3	8	10	8	9		5
CRS 1	$H_{ij}[1]$	1.61	2.30	1.79	1.70	2.20	1.39	1.61	1.79		1.95
	Step 1				REQ						
	Step 2	IDL	TKN	TKN	SUC	IDL		IDL	IDL		TKN
	Step 3				OCP						
CRS 2	$H_{ij}[2]$	1.02				1.56	1.39	1.02	1.19		
	Step 1										
	Step 2	IDL				IDL		IDL	IDL		
	Step 3										
CRS 3	$H_{ij}[3]$	0.72				1.21	1.39	0.72	0.86		
	Step 1							REQ			
	Step 2	TKN				TKN		SUC	TKN		
	Step 3							OCP			

TKN: detect SUCCESS of others and stop contention; REQ: send REQUEST; SUC: feed back SUCCESS; OCP: broadcast OCCUPIED; IDL: send IDLE to transmitter; BSY: feed back BUSY.

Figure 4.5 shows the probability density function of the number of CRSs for completely resolving contention. To verify the impact of network load, we run simulations with 5, 10, 15, or 20 randomly distributed users, respectively. For each case, we run 1000 trials, each of which contains transmission of 5000 frame slots. We can see that heavier network load requires only slightly more CRSs. The average numbers of CRSs in these four cases are 2.35, 3.74, 4.92, and 6.00, while the corresponding standard deviations are 1.66, 2.39, 3.11,

and 4.00, respectively.

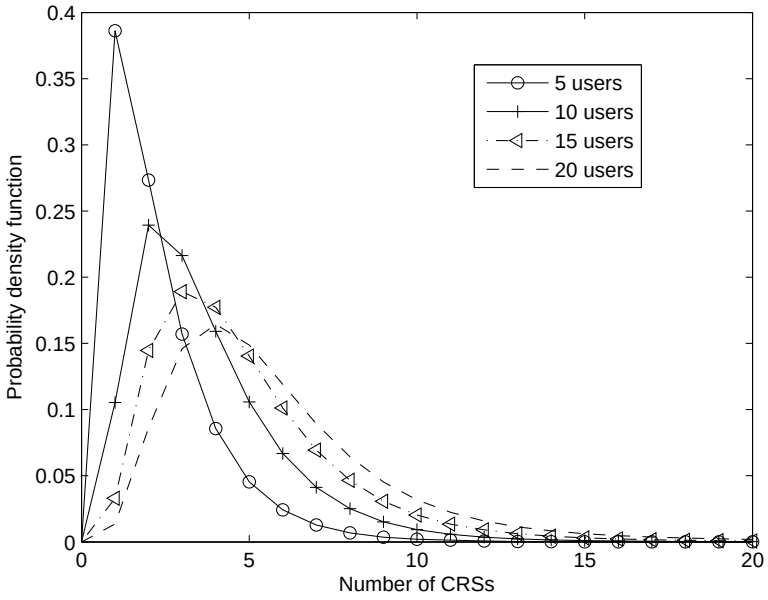


Figure 4.5. Probability density function of the number of CRSs necessary for complete contention resolution.

Figure 4.6 compares the throughput of the proposed CAD-MAC scheme and the DOMRA scheme in [97] when there are different numbers of active users. Again for each number of users, we run 1000 trials of simulation, each of which contains transmission of 5000 frame slots. Significant performance improvement can be observed. When there are 15 active users, the throughput of CAD-MAC outperforms DOMRA by approximately 50% because of the separate design of signalling contention and data transmission.

4.6 Conclusion and Future Work

We have designed a distributed channel-aware random access scheme without making any assumption on network topology and traffic distribution. In the proposed scheme, each frame is divided into contention and transmission periods. The contention period is used to resolve the conflicts of all users while the transmission period is used to send payload in collision-free scenarios. The proposed scheme can completely resolve network contention at a trivial signaling cost and performs closely to the centralized scheduler. Besides, it

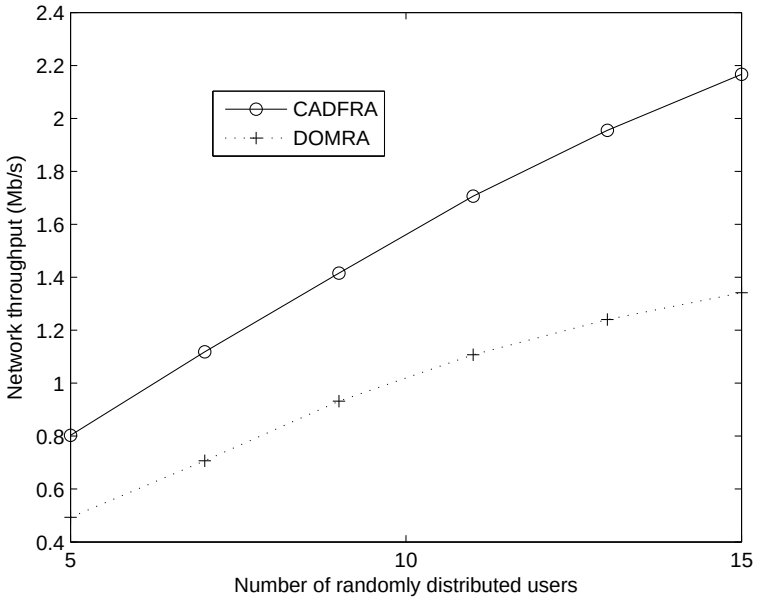


Figure 4.6. Throughput comparison of CAD-MAC and DOMRA.

is also robust to any channel uncertainty. Simulation results have demonstrated that the proposed scheme significantly improves network performance as compared with existing schemes. The generality of the design allows its application in different types of wireless networks, such as cellular networks, sensor networks, and mobile *ad hoc* networks.

In this research, we have not considered traffic characteristics, which influence MAC buffer status and thus its transmission probability. Hence, the contention needs to be improved to incorporate traffic characteristics in our future research. Furthermore, multichannel extensions of CAD-MAC are also desirable to exploit the diversity among different subchannels.

CHAPTER 5

ENERGY-EFFICIENT LINK ADAPTATION IN FREQUENCY-SELECTIVE CHANNELS

Energy efficiency is becoming increasingly important for small form factor mobile devices, as battery technology has not kept up with the growing requirements stemming from ubiquitous multimedia applications. This chapter addresses link adaptive transmission for maximizing energy efficiency, as measured by the "throughput per Joule" metric. In contrast to the existing water-filling power allocation schemes that maximize throughput subject to a fixed overall transmit power constraint, our scheme maximizes energy efficiency by adapting both overall transmit power and its allocation, according to the channel states and the circuit power consumed. We demonstrate the existence of a unique globally optimal link adaptation solution and develop iterative algorithms to obtain it. We further consider the special case of flat-fading channels to develop an upper bound on energy efficiency and to characterize its variation with bandwidth, channel gain, and circuit power. Our results for OFDM systems demonstrate improved energy savings with energy optimal link adaptation as well as illustrate the fundamental tradeoff between energy-efficient and spectrum-efficient transmission.

The rest of the chapter is organized as follows. In Section 5.2, we investigate optimal conditions for energy-efficient transmission and develop algorithms to obtain the globally optimal solution. In Section 5.2.2, we consider a special case when the channel is with flat fading. We also consider energy-efficient link adaptation when the user has either data rate requirement or peak power limit in Section 5.3. As an example of energy-efficient link adaptation, we apply the energy-efficient scheme in OFDM systems and provide simulation results to demonstrate energy efficiency improvement in Section 5.5. Finally, we summarize in Section 5.6.

5.1 Problem Formulation

In this section, we formulate the problem of energy-efficient link adaptation.

Assume that K subchannels are used for transmission, each with a different channel gain. An example of this scenario is OFDM transmission over frequency-selective channels. Assume block fading [88, 98], that is, the channel state remains constant during each data frame and is independent from one to another. Denote the data rate on Subchannel i as r_i and the data rate vector on all subchannels as

$$\mathbf{R} = [r_1, r_2, \cdots, r_K]^T, \tag{5.1}$$

where $[]^T$ is the transpose of a vector. The data rate vector, $\mathbf{R}$, depends on the channel state, coding, and power allocation. Correspondingly, the overall data rate is

$$R = \sum_{i=1}^{K} r_i. \tag{5.2}$$

For a given channel state, the transmit power on each subchannel is determined by the requirement of reliable data transmission. If we denote W as the subchannel bandwidth, N_o the power spectral density, g_i the power gain, and P_{T_i} the allocated transmit power on Subchannel i, the channel output SNR will be

$$\eta_i = \frac{P_{T_i} g_i}{N_o W} \tag{5.3}$$

and the achievable data transmission rate r_i is determined by [99]

$$r_i = W \log(1 + \frac{\eta_i}{\Gamma}), \tag{5.4}$$

where Γ is the SNR gap that defines the gap between the channel capacity and a practical coding and modulation scheme. The SNR gap depends on the coding and modulation scheme used and on the target probability of error. For a coded *quadrature amplitude modulation* (QAM) system, the gap is given by [99]

$$\Gamma = 9.8 + \gamma_m - \gamma_c \text{ (dB)}, \tag{5.5}$$

where γ_m is the system design margin and γ_c is the coding gain. For Shannon capacity [100], $\Gamma = 0$ dB. Denote the overall transmit power as $P_T(\mathbf{R})$ and

$$P_T(\mathbf{R}) = \frac{\sum_{i=1}^{K} P_{T_i}}{\zeta} = \sum_{i=1}^{K} (e^{\frac{r_i}{W}} - 1)\frac{N_o W \Gamma}{g_i \zeta}, \tag{5.6}$$

where $\zeta \in [0, 1]$ is the power amplifier efficiency and depends on the design and implementation of the transmitter. $P_T(\mathbf{R})$ is strictly convex and monotonically increasing in $\mathbf{R}$. In fact, the developed theory and approaches can be used for any $P_T(\mathbf{R})$ that is strictly convex and monotonically increasing in $\mathbf{R}$ with $P_T(\mathbf{0}) = 0$, where $\mathbf{0} = [0, 0, \cdots, 0]^T$.

In addition to transmit power, mobile devices also incur additional circuit power during transmissions which is relatively independent of the transmission rate [58, 64]. While the transmit power models all the power used for reliable data transmission, we let the circuit power represents the average energy consumption of device electronics, such as mixers, filters, and digital-to-analog converters, and this portion of energy consumption excludes that of the power amplifier and is independent of the transmission state. If we denote the circuit power as P_C, the overall power consumption given a data rate vector will be

$$P(\mathbf{R}) = P_C + P_T(\mathbf{R}). \tag{5.7}$$

For energy-efficient communications, it is desirable to maximize the amount of data sent with a given amount of energy. Hence, given any amount of energy $\triangle e$ consumed in a duration, $\triangle t$, i.e. $\triangle e = \triangle t(P_C + P_T(\mathbf{R}))$, the mobile wants to send a maximum amount of data by choosing the data rate vector to maximize

$$\frac{R \triangle t}{\triangle e}, \tag{5.8}$$

which is equivalent to maximizing

$$U(\mathbf{R}) = \frac{R}{\triangle e / \triangle t} = \frac{R}{P_C + P_T(\mathbf{R})}. \tag{5.9}$$

$U(\mathbf{R})$ is called energy efficiency. The unit of the energy efficiency is bits per Joule, which has been frequently used in literature for energy-efficient communications [54, 56, 68, 101,

57]. The optimal energy-efficient link adaptation achieves maximum energy efficiency, i.e.

$$\mathbf{R}^* = \arg\max_{\mathbf{R}} U(\mathbf{R}) = \arg\max_{\mathbf{R}} \frac{R}{P_C + P_T(\mathbf{R})}. \tag{5.10}$$

Note that if we fix the overall transmit power, the objective of Equation (5.10) is equivalent to maximizing the overall throughput and the existing water-filling power allocation approach [50] gives the solution. However, besides adapting the power distributions on all subchannels, the overall transmit power can also be adapted according to the states of all subchannels to maximize the energy efficiency. Hence, the solution to Equation (5.10) is in general different from existing power allocation schemes that maximize throughput with power constraints.

5.2 Principles of Energy-Efficient Link Adaptation

In the following, we demonstrate that a unique globally optimal data rate vector always exists and give the necessary and sufficient conditions for a data rate vector to be globally optimal.

5.2.1 Conditions of Optimality

The concept of quasiconcavity will be used in our discussion and is defined as [102].

Definition 5.2.1 *A function f, which maps from a convex set of real n-dimensional vectors, $\mathcal{D}$, to a real number, is called strictly quasiconcave if for any $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{D}$ and $\mathbf{x}_1 \neq \mathbf{x}_2$,*

$$f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) > \min\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}, \tag{5.11}$$

for any $0 < \lambda < 1$.

Any strictly monotonic function is quasiconcave. Besides, any strictly concave function is also strictly quasiconcave but the reverse is not generally true. An example is the Gaussian function, which is strictly quasiconcave but not concave.

It is proved in Appendix D.3 that $U(\mathbf{R})$ has the following properties.

Lemma 5.2.2 *If $P_T(\mathbf{R})$ is strictly convex in $\mathbf{R}$, $U(\mathbf{R})$ is strictly quasiconcave. Furthermore, $U(\mathbf{R})$ is either strictly decreasing or first strictly increasing and then strictly decreasing in any r_i of $\mathbf{R}$, i.e. the local maximum of $U(\mathbf{R})$ for each r_i exists at either 0 or a positive finite value.*

For strictly quasiconcave functions, if a local maximum exists, it is also globally optimal [102]. Hence, a unique globally optimal transmission rate vector always exists and its characteristics are summarized in Theorem 5.2.3 according to the proofs in Appendix D.3.

Theorem 5.2.3 *If $P_T(\mathbf{R})$ is strictly convex, there exists a unique globally optimal transmission data rate vector $\mathbf{R}^* = [r_1^*, r_2^*, \cdots, r_K^*]^T$ for (5.10), where r_i^* is given by*

1. *when $\frac{P_C+P_T(\mathbf{R}_i^{(0)})}{R_i^{(0)}} \geq \left.\frac{\partial P_T(\mathbf{R})}{\partial r_i}\right|_{\mathbf{R}=\mathbf{R}_i^{(0)}}$, $\left.\frac{\partial U(\mathbf{R})}{\partial r_i}\right|_{\mathbf{R}=\mathbf{R}^*} = 0$, i.e. $\frac{1}{\frac{\partial P_T(\mathbf{R}^*)}{\partial r_i^*}} = \frac{R^*}{P_C+P_T(\mathbf{R}^*)} = U(\mathbf{R}^*)$;*
2. *when $\frac{P_C+P_T(\mathbf{R}_i^{(0)})}{R_i^{(0)}} < \left.\frac{\partial P_T(\mathbf{R})}{\partial r_i}\right|_{\mathbf{R}=\mathbf{R}_i^{(0)}}$, $r_i^* = 0$,*

where $\mathbf{R}_i^{(0)} = [r_1^*, r_2^*, \cdots, r_{i-1}^*, 0, r_{i+1}^*, \cdots, r_K^*]$ and $R_i^{(0)} = \sum_{j\neq i} r_j^*$, i.e. the overall data rate on all other subchannels except i.

Theorem 5.2.3 has clear physical insights. $P_C + P_T(\mathbf{R}_i^{(0)})$ is the power consumption of both circuit and all other subchannels when Subchannel i is not used. $\frac{P_C+P_T(\mathbf{R}_i^{(0)})}{R_i^{(0)}}$ is the per-bit energy consumption when Subchannel i is not used and the overall per-bit energy consumption needs to be minimized for energy-efficient communications. $\left.\frac{\partial P_T(\mathbf{R})}{\partial r_i}\right|_{\mathbf{R}=\mathbf{R}_i^{(0)}}$ is the per-bit energy consumption transmitting infinitely small data rate on Subchannel i conditioned on the optimal status of all other subchannels. Hence, Subchannel i should not transmit anything when $\frac{P_C+P_T(\mathbf{R}_i^{(0)})}{R_i^{(0)}} < \left.\frac{\partial P_T(\mathbf{R})}{\partial r_i}\right|_{\mathbf{R}=\mathbf{R}_i^{(0)}}$. Otherwise, there should be a tradeoff between the desired data rate on Subchannel i and the incurred power consumption. The tradeoff closely depends on the power consumption of both circuits and transmission on all other subchannels and can be found through the unique zero derivative of $U(\mathbf{R})$ with respect to r_i.

To further understand Theorem 5.2.3, we consider an example when each subchannel achieves the Shannon capacity and the transmit power on each subchannel is given in (5.6) with $\Gamma = 0$ dB and $\zeta = 1$. The overall transmit power is

$$P_T(\mathbf{R}) = \sum_{k=1}^{K} (e^{\frac{r_k}{W}} - 1)\frac{N_o W}{g_k}. \tag{5.12}$$

According to Condition (i) of Theorem 5.2.3, when $r_k > 0$, we have

$$\frac{1}{\frac{\partial P_T(\mathbf{R})}{\partial r_k}} = \frac{1}{e^{\frac{r_k}{W}}\frac{N_o}{g_k}} = U(\mathbf{R}^*). \tag{5.13}$$

Hence, the transmit power on Subchannel k is

$$P_{T_n} = (e^{\frac{r_k}{W}} - 1)\frac{N_o W}{g_k} = \frac{W}{U(\mathbf{R}^*)} - \frac{N_o W}{g_k}, \tag{5.14}$$

which is a water-filling to level $\frac{W}{U(\mathbf{R}^*)}$. Since the water level is determined by the optimal energy efficiency, we refer to our scheme as *dynamic energy-efficient water-filling*. Note that while the absolute value of power allocation is determined by the maximum energy efficiency $U(\mathbf{R}^*)$, which relies on both the circuit power and channel state, the relative differences of power allocation on different subchannels depend only on the channel gains on those subchannels.

5.2.2 A Special Case: When the Channel is Flat Fading

To facilitate the understanding of the fundamental dependence of energy efficiency on the channel gain, circuit power, and bandwidth, we consider a special case that the channel is experiencing flat fading in this section. Hence, all subchannels are with the same channel gain and the same link adaptation is applied on all subchannels. The overall data rate is

$$R = Kr. \tag{5.15}$$

According to Theorem 5.2.3, the optimal transmission data rate follows immediately and is summarized by Theorem 5.2.4, where the upper bound is proved in Appendix C.2.

Theorem 5.2.4 *If $P_T(R)$ is monotonically increasing and strictly convex in R, there exists a unique globally optimal transmission data rate to maximize energy efficiency and is given by*

$$R^* = \frac{P_C + P_T(R^*)}{P_T^{'}(R^*)}, \tag{5.16}$$

where $P_T^{'}(\cdot)$ is the first order derivative of function $P_T(\cdot)$. Besides, energy efficiency is upper bounded by $\frac{1}{P_T^{'}(0)}$.

When Shannon capacity is achieved in AWGN channels, the upper bound is $\frac{g}{N_o}$.

In the following, we investigate some basic properties of energy-efficient link adaptation. Propositions 5.2.5, 5.2.6, and 5.2.7 summarize the impact of channel gain, circuit power, and the number of subchannels on the optimal energy-efficient transmission, and are proved in Appendix C.3.

Proposition 5.2.5 *Both the data rate and energy efficiency increase with channel gain.*

Proposition 5.2.6 *The data rate increases with circuit power, while the energy efficiency decreases with it. With zero circuit power, the highest energy efficiency, $\frac{1}{P_T^{'}(0)}$, is obtained by transmitting with infinite small data rate.*

From Proposition 5.2.6, when circuit power dominates power consumption, which is usually true with short-range communication, the highest data rate should be used to finish transmission as soon as possible, which has been commonly assumed by most MAC layer energy-efficient optimization schemes as describe in the introduction of this chapter. However, when the circuit power is negligible, which is usually true with long-range communication like satellite communications, the lowest data rate should be used, which coincides with the results in [57] and [103].

Proposition 5.2.7 *The data rate on each subchannel decreases with increasing number of subchannels while the energy efficiency increases with it. With infinite number of subchannels, the highest energy efficiency, $\frac{1}{P_T^{'}(0)}$, is obtained by transmitting with infinite small data rate.*

Propositions 5.2.5, 5.2.6, and 5.2.7 discover three ways to improve energy efficiency: increasing channel power gain, reducing circuit power, and allocating more subchannels. The energy-efficiency upper bound is achieved by transmitting with infinite small data rate when either circuit power is zero or infinite number of subchannels is assigned.

5.3 Constrained Energy-Efficient Link Adaptation

In this section, we study energy-efficient link adaptation when user has either a data rate requirement or a peak power limit.

With a data rate requirement Γ, the energy-efficient link adaptation is given by

$$\widehat{\mathbf{R}}^* = \arg\max_{\mathbf{R}} \frac{R}{P_C + P_T(\mathbf{R})}, \tag{5.17a}$$

subject to

$$R \geq \Gamma. \tag{5.17b}$$

If the optimal data rate vector without constraint in (5.10) satisfies $R^* \geq \Gamma$, it is also the solution to Problem (5.17), i.e. $\widehat{\mathbf{R}}^* = \mathbf{R}^*$. Otherwise, Problem (5.17) is equivalent to

$$\widehat{\mathbf{R}}^* = \arg\max_{\Gamma} \frac{\Gamma}{P_C + P_T(\mathbf{R})} = \arg\min_{\mathbf{R}} P_T(\mathbf{R}), \tag{5.18a}$$

subject to

$$R = \Gamma. \tag{5.18b}$$

Since $P_T(\mathbf{R})$ is strictly convex, a unique globally optimal $\widehat{\mathbf{R}}^*$ exists. Denote

$$f_k(r_k) = \frac{\partial P_T(\mathbf{R})}{\partial r_k} \tag{5.19}$$

and its inverse function to be $f_k^{-1}()$. Then $\widehat{\mathbf{R}}^*$ can be easily obtained via the Lagrangian technique [104] and is

$$\widehat{r_k^*} = \max\left\{f_k^{-1}(\lambda), 0\right\} \tag{5.20}$$

for $k = 1, \cdots, K$, where λ is determined by

$$\sum_{k=1}^{K} \widehat{r_k^*} = \Gamma. \tag{5.21}$$

When the channel capacity is achieved on each subchannel, the corresponding optimal power allocation is a water-filling allocation, which achieves the sum channel capacity Γ.

Similarly, with a maximum transmit power constraint, the problem is to find

$$\widetilde{\mathbf{R}}^* = \arg\max_{\mathbf{R}} \frac{R}{P_C + P_T(\mathbf{R})}, \tag{5.22a}$$

subject to

$$P_T(\mathbf{R}) \leq P_m. \tag{5.22b}$$

If the optimal data rate vector without constraint in (5.10) satisfies $P_T(\mathbf{R}^*) \leq P_m$, it is also the solution to Problem (5.22), i.e. $\widetilde{\mathbf{R}}^* = \mathbf{R}^*$. Otherwise, via the the Lagrangian technique again, we have the unique optimal solution as follows

$$\widetilde{r}_k^* = \max\left\{f_k^{-1}(\lambda), 0\right\}, k = 1, \cdots, K, \tag{5.23}$$

where λ is determined by

$$P_T(\widetilde{\mathbf{R}}^*) = P_m. \tag{5.24}$$

When channel capacity is achieved on each subchannel, the power allocation is the classical water-filling where the water level is determined by P_m [50].

5.4 Algorithm Design

Theorem 5.2.3 provides the necessary and sufficient conditions for a rate vector to be the unique and globally optimum one. However, it is usually difficult to directly solve the joint nonlinear equations according to Theorem 5.2.3 to obtain the optimal vector $\mathbf{R}^*$. Therefore, we develop iterative methods to search the optimal $\mathbf{R}$ for maximizing $U(\mathbf{R})$. The global optimality of the proposed methods is guaranteed by the strict quasiconcavity of $U(\mathbf{R})$. In the following, we describe our low-complexity iterative algorithms.

5.4.1 Gradient Assisted Binary Search

When there is only one subchannel, Lemma 5.2.2 shows that function $U(\mathrm{r})$ has a unique r^* such that for any $r < r^*$, $\frac{dU(\mathrm{r})}{dr} > 0$, and for any $r > r^*$, $\frac{dU(\mathrm{r})}{dr} < 0$. Hence, we have the following lemma to seek two points r_1 and r_2 such that $r_1 \leq r^* \leq r_2$.

Proposition 5.4.1 *Let the initial setting* $r^{[0]} > 0$ *and set* $\alpha > 1$. *For any* $i \geq 0$, *let*

$$r^{[i+1]} = \begin{cases} \frac{r^{[i]}}{\alpha} & \left.\frac{dU(r)}{dr}\right|_{r^{[0]}} < 0 \\ \alpha r^{[i]} & otherwise \end{cases} . \tag{5.25}$$

Repeat (5.25) until $r^{[I]}$ *such that* $\left.\frac{dU(r)}{dr}\right|_{r^{[I]}}$ *has a different sign from* $\left.\frac{dU(r)}{dr}\right|_{r^{[0]}}$. *Then* r^* *must be between* $r^{[I]}$ *and* $r^{[I-1]}$.

To locate r^* between r_1 and r_2, let $\widehat{r} = \frac{r_1+r_2}{2}$. If $\left.\frac{dU(r)}{dr}\right|_{\widehat{r}} = 0$, r^* is found. If $\left.\frac{dU(r)}{dr}\right|_{\widehat{r}} < 0$, $r_1 < r^* < \widehat{r}$ and replace r_2 with $\widehat{r}$; otherwise, replace r_1 with $\widehat{r}$. This leads to the *gradient assisted binary search* (GABS) for maximizing $U(r)$, which is summarized in Table 5.1.

Table 5.1. Gradient assisted binary search

Algorithm *GABS*(r_o)
($*$ algorithm for single-subchannel transmission. $*$)
Input: initial guess: $r_o > 0$
Output: optimal transmission rate: r^*

1. $r_1 = r_o$, $h_1 \leftarrow \left.\frac{dU(r)}{dr}\right|_{r_1}$, initialize $\alpha > 1$ (e.g.10)
2. **if** $h_1 < 0$
   ($*$ seek r_1 and r_2 such that $r_1 < r^* < r_2$ $*$)
3. **then** $r_2 \leftarrow r_1$, $r_1 \leftarrow \frac{r_1}{\alpha}$, and $h_1 \leftarrow \left.\frac{dU(r)}{dr}\right|_{r_1}$
4. **while** $h_1 < 0$
5. **do** $r_2 \leftarrow r_1$, $r_1 \leftarrow \frac{r_1}{\alpha}$, and $h_1 \leftarrow \left.\frac{dU(r)}{dr}\right|_{r_1}$
6. **else** $r_2 \leftarrow r_1 * \alpha$ and $h_2 \leftarrow \left.\frac{dU(r)}{dr}\right|_{r_2}$
7. **while** $h_2 > 0$
8. **do** $r_1 \leftarrow r_2$, $r_2 \leftarrow r_2 * \alpha$, and $h_2 \leftarrow \left.\frac{dU(r)}{dr}\right|_{r_2}$
9. **while** no convergence
   ($*$ seek r^* between r_1 and r_2 $*$)
10. **do** $\widehat{r} \leftarrow \frac{r_2+r_1}{2}$; $\widehat{h} \leftarrow \left.\frac{dU(r)}{dr}\right|_{\widehat{r}}$
11. **if** $\widehat{h} > 0$
12. **then** $r_1 = \widehat{r}$;
13. **else** $r_2 = \widehat{r}$
14. **return** $\widehat{r}$

5.4.2 Binary Search Assisted Ascent

To find the optimal data rate vector for the multiple subchannel case, we design a gradient ascent method to produce a maximizing sequence $\mathbf{R}^{[i]}$, $n = 0, 1, \cdots$, and

$$\mathbf{R}^{[i+1]} = \left[\mathbf{R}^{[i]} + \mu \nabla U(\mathbf{R}^{[i]})\right]^+, \tag{5.26}$$

where $[\mathbf{R}]^+$ sets the negative part of the vector $\mathbf{R}$ to be zero, $\mu > 0$ is the search step size, and $\nabla U(\mathbf{R}^{[i]})$ is the gradient at iteration i. With sufficiently small step size, $U(\mathbf{R}^{[i+1]})$ will be always bigger than $U(\mathbf{R}^{[i]})$ except when $\nabla U(\mathbf{R}^{[i]}) = 0$ that indicates the optimality of $\mathbf{R}^{[i]}$ [104]. However, small step size leads to slow convergence. Besides, each element of the gradient depends on the corresponding subchannel power gain, which potentially differs from each other by orders of magnitude. Hence, a line search of the optimal step size needs to cover a large range to assure global convergence on all subchannels, which is computationally expensive. Therefore, at each $\mathbf{R}^{[i]}$, an efficient algorithm is needed to find the optimal step size. Denote

$$f_i(\mu) = U(\left[\mathbf{R}^{[i]} + \mu \nabla U(\mathbf{R}^{[i]})\right]^+). \tag{5.27}$$

Similar to the proof of Lemma 5.2.2, it is easy to show that $g_i(\mu)$ is also strictly quasiconcave in μ and has a unique globally maximum μ^* such that for any $\mu < \mu^*$, $\frac{df_i(\mu)}{d\mu} > 0$, and for any $\mu > \mu^*$, $\frac{df_i(\mu)}{d\mu} < 0$. Let $\nabla U(\mathbf{R}^{[i]}) = [\widehat{g}_1, \widehat{g}_2, \cdots, \widehat{g}_K]$. Replace $\frac{dU(r)}{dr}$ in GABS to be

$$\frac{df_i(\mu)}{d\mu} = [\nabla U(\mathbf{R}^{[i+1]})]^T \widetilde{\mathbf{G}}[i], \tag{5.28}$$

where $\widetilde{\mathbf{G}}[i] = \frac{d\left[\mathbf{R}^{[i]} + \mu \nabla U(\mathbf{R}^{[i]})\right]^+}{d\mu} = [\widetilde{g}_1, \widetilde{g}_2, \cdots, \widetilde{g}_K]$, in which $\widetilde{g}_k = \widehat{g}_k$ if the kth component of $\mathbf{R}^{[i]} + \mu \nabla U(\mathbf{R}^{[i]})$ is positive and $\widetilde{g}_k = 0$ otherwise. Then GABS can be used for quick location of the optimal step size. This leads to the *binary search assisted ascent* (BSAA) algorithm in Table 5.2.

5.4.3 The Rate of Convergence

While the global convergence of both GABS and BSAA is guaranteed by the strict quasiconcavity of $U(\mathbf{R})$ [105], we further study the convergence rate in this section.

Table 5.2. Binary search assisted ascent

Algorithm *BSAA*($\mathbf{R}_o$)
($*$ algorithm for multi-subchannel transmission. $*$)
Input: initial guess: $\mathbf{R}_o$(default transmission rate can be used)
Output: optimal transmission rate vector: $\mathbf{R}^*$
1. $\mathbf{R} = \mathbf{R}_o$,
2. **while** no convergence
3. **do** use GABS to find the optimal step size μ^*;
4. $\mathbf{R} = [\mathbf{R} + \mu^* \nabla U(\mathbf{R})]^+$
5. **return** $\mathbf{R}$

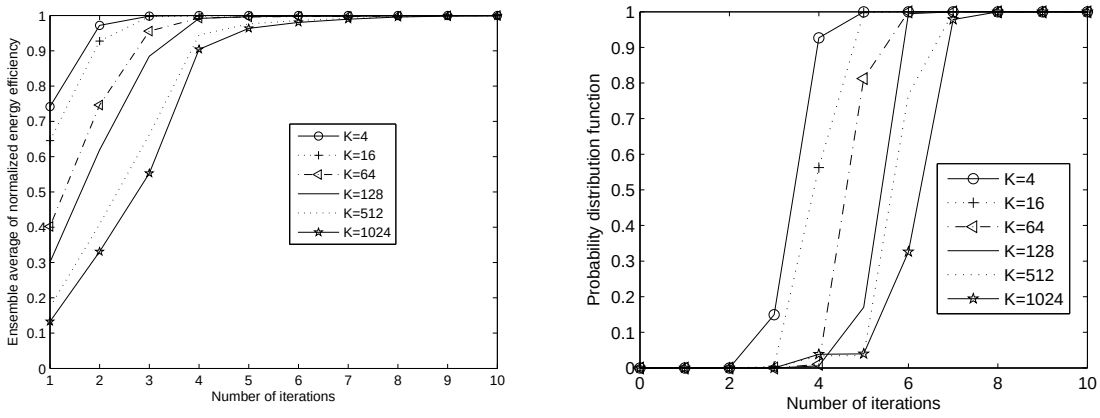
Theorem 5.4.2 characterizes the convergence of GABS and is proved in Appendix C.4.

Theorem 5.4.2 *GABS converges to the globally optimal transmission data rate r^*. A rate r, which satisfies $|r - r^*| \leq \epsilon$, can be found within at most M iterations, where M is the minimum integer such that $M \geq \log_2(\frac{(\alpha-1)r^*}{\epsilon} - 1)$.*

It is difficult to theoretically analyze the global convergence rate of BSAA because of the nonconcavity of $U(\mathbf{R})$. Instead, we run numerical simulations and observe the convergence. Figure 5.1(a) illustrates the improvement of energy efficiency with iterations. Here we assume the channel gain of each subchannel has Rayleigh distribution with a unit average. The circuit power is 5. The noise power on each subchannel is 0.01. The transmit power is given by Equation (5.6) with $\Gamma = 0$ dB. The energy efficiency is normalized by the optimal value and the curves are the ensemble averages of 5000 channel instances. Figure 5.1(b) shows the corresponding probability distribution functions of the numbers of iterations necessary for convergence. In both figures, we vary the number of subchannels to verify its impact on the convergence rate. We can see that BSAA converges very fast to the global optimum, even with 1024 subchannels.

5.5 Simulation Results for OFDM

The proposed energy-efficient link adaptation is general and can be applied to different kinds of OFDM, MIMO, and MIMO-OFDM systems. To apply it, we only need to find the



(a) Converging process: relationship between ensemble average and iterations

(b) Probability distribution function of the number of iterations for convergence

Figure 5.1. Convergence rate of BSAA.

transmit power relationship $P_T(\mathbf{R})$ of those systems. In this section, we discuss the optimal energy-efficient link adaptation for OFDM with subchannelization as an example.

5.5.1 Modeling of OFDM with Subchannelization

In OFDM systems with subchannelization, subcarriers are grouped into subchannels and the subcarriers forming one subchannel may, but not necessarily be adjacent, such as the contiguous and distributed subchannelization schemes in 802.16e [12]. Each subchannel is treated to be flat fading and the *effective channel power gain*, $\overline{g}$, rather than physical channel power gain of each subcarrier, is used as a metric. For simplicity, $\overline{g}$ is the average of channel power gains of all subcarriers within the subchannel. Note that classical OFDM is a special case when each subchannel has one subcarrier. The frame structure is shown in Figure 5.2. Each transmission slot consists of a data interval, T_s, and a signalling interval, τ. In each data interval, l symbols are transmitted. We use uncoded M-QAM. The transmit power on each subchannel needs to be determined.

Consider Subchannel i that consists of c_i subcarriers. The number of bits transmitted per symbols on each subcarrier is $b_i = \frac{\frac{r_i}{c}(T_s+\tau)}{l}$. Hence, the modulation order M is given by $M_i = 2^{b_i} = 2^{B_i r_i}$, where $B_i = \frac{(T_s+\tau)}{c_i l}$. The BER for coherently detected M-QAM with Gray

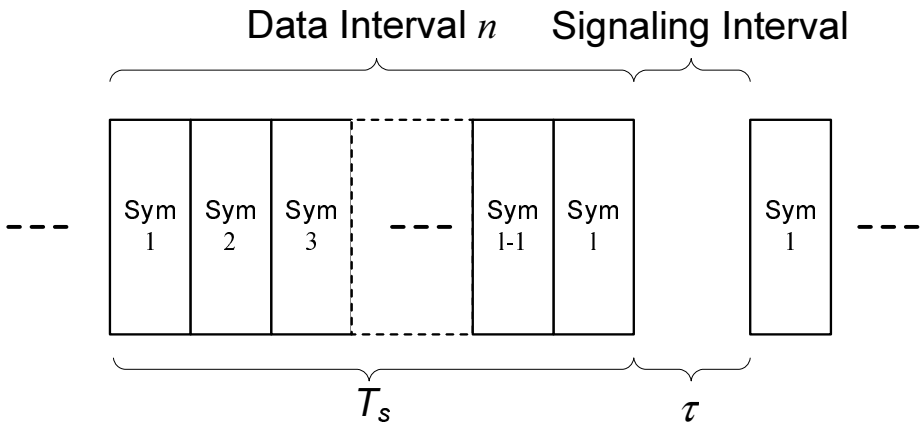


Figure 5.2. Frame structure

mapping over an AWGN channel is approximated by [74]

$$P_e(\gamma) \approx 0.2 \exp\left(-\frac{1.5\gamma}{M-1}\right), \tag{5.29}$$

where γ is the SNR. For a BER target, P_e, the required SNR on Subchannel i is

$$\gamma_i = \frac{2}{3}(1 - M_i)\ln(5P_e) = \frac{2}{3}(1 - 2^{B_i r_i})\ln(5P_e). \tag{5.30}$$

Hence, the overall transmit power on Subchannel i is

$$P_{T_i}(r_i) = \frac{\gamma_i c_i N_o W}{\overline{g}_i} = A_i(1 - 2^{B_i r_i}), \tag{5.31}$$

where W is the signal bandwidth of each subcarrier and

$$A_i = \frac{2c_i \ln(5P_e)N_o W}{3\overline{g}_i}. \tag{5.32}$$

Assuming no coupling between transmit powers among subchannels, the overall transmit power will be the cumulative of the transmit powers of all subchannels, that is,

$$P_T(\mathbf{R}) = \sum_{i=1}^{K} P_{Ti}(r_i), \tag{5.33}$$

which is monotonically increasing and strictly convex in $\mathbf{R}$. The energy-efficient link adaptation immediately follows from Section 5.2.

Table 5.3. System parameters

Carrier frequency	1.5 GHz
Subcarrier number	256
Subcarrier bandwidth	10 kHz
BER requirement	10^{-3}
Symbol number of data interval, l	100
Time duration of data interval, T_s	0.01s
Time duration of signalling interval, τ	0.001s
Thermal noise power, N_o	-141 dBW/MHz
User antenna height	1.6 m
BS antenna height	40 m
Environment	Macro cell in urban area
Circuit power, P_C	100 mW
Modulation	Uncoded M-QAM
Subchannelization	Fixed-interval and contiguous
Propagation Model	Okumura-Hata model
Shadowing	Log-normal with standard deviation of 10 dB
Frequency-selective fading	ITU pedestrian channel B
User speed	3 km/h

5.5.2 Performance Comparison

In this section, we compare the performance of energy-efficient OFDM transmission with that of traditional transmission schemes. The system parameters are listed in Table 7.1. The *International Telecommunication Union* (ITU) pedestrian channel model B [106] is used to implement the multipath frequency-selective fading. We implement two subchannelization schemes as in Figure 5.3, fixed-interval and contiguous, both of which group 10 subcarriers into a subchannel. In the fixed-interval subchannelization, one draws subcarriers out of all subcarriers with a fixed interval to form a subchannel, while in the contiguous one, each subchannel consists of a block of contiguous subcarriers.

Figures 5.4(a) and 5.4(b) and Figures 5.5(a) and 5.5(b) compare energy efficiency and throughput of different transmission schemes with contiguous subchannelization and with fixed-interval subchannelization respectively. Two energy-efficient OFDM transmission schemes are implemented: FS EE, that is the optimal energy-efficient transmission developed in this chapter, and flat EE, that treats the channel as flat fading. Transmissions with

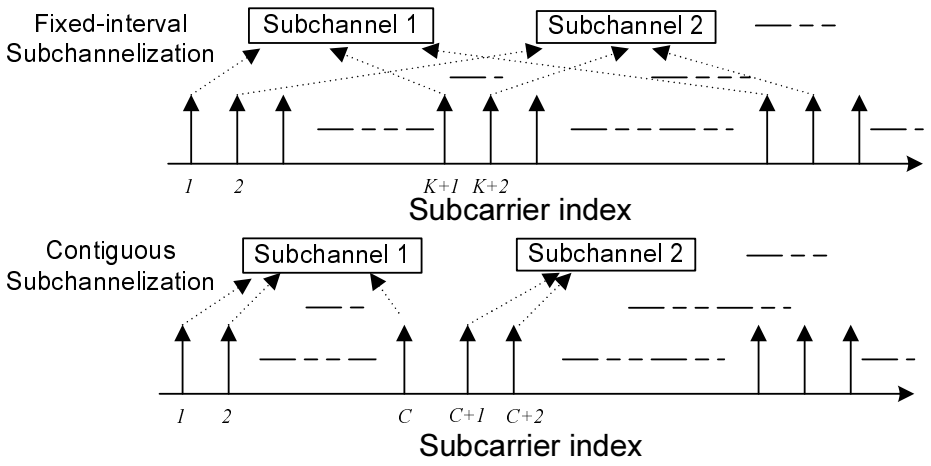
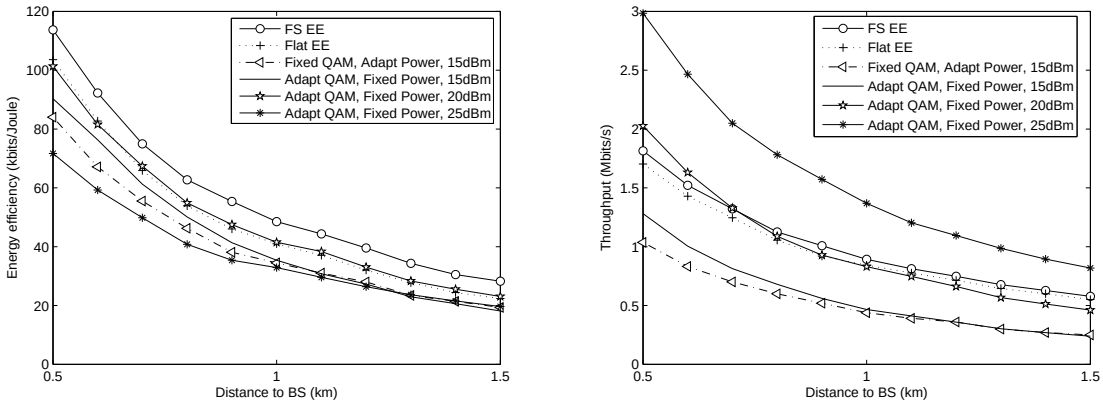


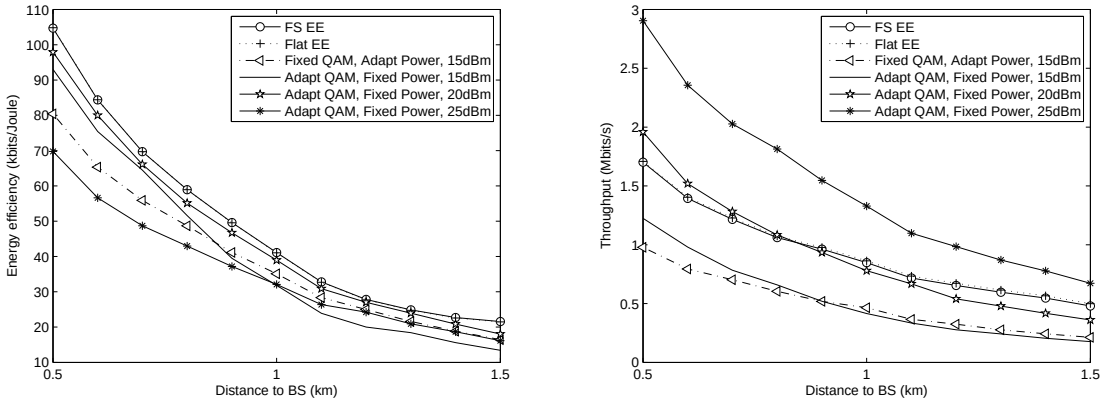
Figure 5.3. OFDM subchannelization (K subchannels, each with c subcarriers)

both fixed and adaptive QAM modulations are also implemented for comparison. For fixed modulation, the transmit power is adapted to meet BER requirement while not exceeding 15 dBm maximum power constraint. For adaptive modulation, transmit power is equally distributed over all subchannels and the modulation is adapted to meet BER requirement. From Figures 5.4(a) and 5.4(b), fixed and adaptive modulations perform closely to each other, especially when far away from BS, for both energy efficiency and throughput, when the maximum transmit power is 15 dBm. By increasing the transmit power from 15 dBm to 25 dBm, the throughput of adaptive modulation increases, however, the energy efficiency first increased and then decreases. Due to the global optimality, the proposed energy-efficient transmission for frequency-selective channels always achieves the highest energy efficiency, and outperforms the others by at least 15%. However, the throughput is not necessarily maximum; the other schemes, especially the adaptive QAM modulation with 25 dBm transmit power, sacrifice power to obtain higher throughput. Similar results can also be observed in Figures 5.5(a) and 5.5(b). Furthermore, we note that when fixed-interval subchannelization is used, different subchannels have trivial differences in average channel gain and the energy-efficient transmission treating the channel to be flat fading performs the same as the one considering the difference of different subchannels. This indicates energy-efficient link adaptation treating channels to be flat fading is sufficient for performance



(a) Energy efficiency (b) Throughput

Figure 5.4. Performance comparison for contiguous subchannelization.



(a) Energy efficiency (b) Throughput

Figure 5.5. Performance comparison for fixed-interval subchannelization.

optimization.

5.6 Conclusion

In this chapter, we have investigated the energy-efficient link adaptation. While the usefulness of the proposed technique is illustrated using frequency-selective OFDM as an example, the solution developed is applicable to more general transmission scenarios where transmission occurs over resources experiencing different channel conditions. Joint circuit and transmit power consumptions are taken into account to maximize energy efficiency

rather than throughput. We demonstrate the existence of a unique globally optimal link adaptation solution and provide iterative algorithms to obtain this optimum. The optimal power allocation is shown to be a dynamic water-filling where the water level is determined by the maximum energy efficiency. We further consider a special case when the channel is experiencing flat fading and show the upper bound of energy efficiency as well as two ways to achieve this bound. We explicitly demonstrate that energy efficiency is improved by increasing channel power gain, bandwidth, and by reducing circuit power consumption. From the simulation results, we observed at least 15% improvement in energy utilization when frequency selectivity is exploited and the improvement depends on how much frequency diversity exists within the channels.

CHAPTER 6

LOW-COMPLEXITY ENERGY-EFFICIENT SCHEDULING FOR UPLINK OFDMA

The work in Chapter 5 uses iterative approaches to maximize the instantaneous bits-per-Joule energy efficiency. Iterative approaches may be complex to be implemented in practical systems and running iterative algorithms consumes additional power. In this chapter, we use a time-averaged bits-per-Joule metric to develop low-complexity schemes. Specifically, we obtain closed-form solutions for energy-efficient link adaptation in frequency-selective channels. On the other hand, in an OFDMA cellular network, the BS determines subchannel assignment to optimize the overall network performance. As indicated previously, while increasing subchannels always improves energy efficiency, the entire system bandwidth can not be allocated exclusively to one user in a multi-user system since this may hurt the energy efficiency of other users as well as that of the overall network. Hence, resource management in OFDMA is critical in determining the overall network energy efficiency and we also derive closed-form approaches for the maximum arithmetic and geometric mean energy-efficient schedulers.

In the following, we describe the system model and design objectives in Section 6.1. Then we develop energy-efficient link adaptation in Section 6.2 and resource allocation schemes in Section 6.3. Simulation results are provided in Section 7.4. Finally, we conclude the chapter in Section 6.5.

6.1 System Description

We focus on the uplink OFDMA system shown in Figure 6.1, as the *radio frequency* (RF) transmit power for a user dominates the limited power budget of a battery-constrained mobile device. The BS assigns subchannels for each user to optimize the overall network energy efficiency. Channels are assumed to be frequency-selective and with block fading,

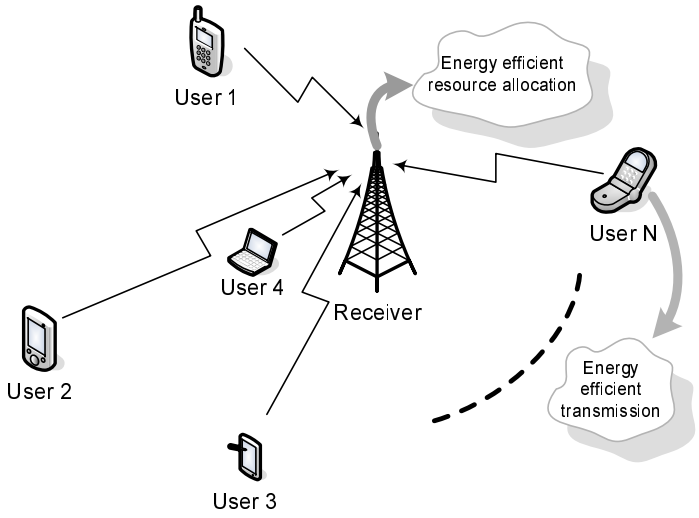


Figure 6.1. Network architecture

i.e. the channel state is constant within each frame [88]. Accurate channel state information is available to both BS and mobile users to optimize energy-efficient communications. The link adaptation and resource allocation settings are allowed to vary from one frame to another according to the channel state information.

Consider a network with N users and K subchannels. Denote the index set of all subchannels as $\mathcal{K} = \{1, 2, \cdots, K\}$. Denote the index set of subchannels assigned to User n at Frame t to be $C_n[t]$. Each subchannel is only assigned to one user during a frame. Then,

$$\begin{aligned} C_i[t] \bigcap C_j[t] = \emptyset, \forall i \neq j \\ \bigcup_i C_i[t] \subseteq \mathcal{K}, \end{aligned} \tag{6.1}$$

where $\emptyset$ is an empty set. The data rate of User n is

$$r_n[t] = \sum_{k \in C_n[t]} r_{nk}[t], \tag{6.2}$$

where $r_{nk}[t]$ is the data rate of User n at subchannel k. The average throughput of User n at Frame t, $T_n[t]$, is obtained using an exponentially weighted low-pass filter,

$$T_n[t] = (1 - \frac{1}{w})T_n[t-1] + \frac{1}{w} r_n[t], \tag{6.3}$$

where $w \gg 1$.

Denote the SNR for reliable reception of $r_{nk}[t]$ to be

$$\eta_{nk} = S(r_{nk}[t]). \tag{6.4}$$

Similar to the argument in [60], function $S(r)$ is assumed to be strictly convex in r and $S(0) = 0$. Denote the signal power attenuation of User n at Subchannel k at Frame t to be $g_{nk}[t]$, then required power to transmit at a rate of $r_{nk}[t]$ will be

$$p_{nk}[t] = \frac{\eta_{nk}\sigma^2}{g_{nk}[t]} = \frac{S(r_{nk}[t])\sigma^2}{g_{nk}[t]}, \tag{6.5}$$

where σ^2 is the noise power on each subchannel.

The overall transmit power of User n is

$$p_n[t] = \sum_{k \in C_n[t]} p_{nk}[t]. \tag{6.6}$$

As indicated in [60, 107], circuit power, p_c, in addition to the transmit power, also plays an important role in energy-efficient communications. While transmit power is used for reliable data transmission, circuit power represents energy consumption of device electronics. The overall average power consumption, $P_n[t]$, is also obtained using an exponentially weighted moving average low-pass filter, that is,

$$P_n[t] = (1 - \frac{1}{w})P_n[t-1] + \frac{1}{w}(p_n[t] + p_{cn}[t]). \tag{6.7}$$

The circuit power, $p_{cn}[t]$, is user and time dependent. $p_{cn}[t]$ is measured at Frame t by each mobile.

For energy-efficient communications, users want to send as much data as possible with a given amount of energy. Hence, with energy $\triangle e$ consumed in a duration $\triangle t$, User n wants to send a maximum amount of data by choosing $r_{nk}[t], k \in C_n[t]$, to maximize

$$\frac{T_n[t] \triangle t}{\triangle e}, \tag{6.8}$$

which is equivalent to maximize

$$u_n[t] = \frac{T_n[t]}{\triangle e / \triangle t} = \frac{T_n[t]}{P_n[t]}. \tag{6.9}$$

u_n is called average energy efficiency of User n. Adapting transmission rate and power to optimize equation (7.6) is referred to as energy-efficient link adaptation.

If we fix the overall transmit power, the objective of Equation (7.6) is equivalent to maximizing the overall throughput and existing water-filling power allocation approaches [50,51] can be used. However, besides adapting the power distributions on all subchannels, the overall transmit power can also be adapted according to the states of all subchannels and the history of data transmission and power consumption to maximize the average energy efficiency. Hence, the solution to Equation (7.6) is different from existing power allocation schemes that maximize throughput with power constraints.

The BS determines subchannel assignment to optimize the overall network performance. We consider two multi-user performance metrics, arithmetic and geometric means. The resource management is optimized to maximize the arithmetic or geometric average of the performances of all users with arithmetic- or geometric-mean metric. Considering these performance metrics in the context of spectral efficiency, we note that the arithmetic-mean metric leads to power allocation for sum throughput maximization, which assures no fairness since some users may have zero throughput. However, the geometric-mean metric leads to a solution for throughput product maximization and assures proportional fairness among all users [76, 16]. Analogously, we call energy efficiency optimization schemes using geometric- or arithmetic-mean metrics to be energy-efficient schedulers with or without fairness.

When using the arithmetic-mean metric, the subchannels are allocated to maximize the arithmetic average of the energy efficiency of all users, i.e. to maximize

$$U[t] = \sum_{n=1}^{N} u_n[t]. \tag{6.10}$$

When using the geometric-mean metric, the subchannels are allocated to maximize the geometric average of the energy efficiency of all users, i.e. to maximize

$$V[t] = \sum_{n=1}^{N} \log(u_n[t]). \tag{6.11}$$

For the above optimization, consider a special case that the circuit power dominates the power consumption, i.e.

$$p_{cn}[t] \gg p_n[t], \qquad \forall t. \tag{6.12}$$

This is usually true for short-range communications as low transmit power is needed to compensate for path loss. In this case, maximizing energy efficiency (7.6) is equivalent to maximizing throughput $T_n[t]$ as $P_n[t]$ is almost independent of power allocation and rate adaptation. Correspondingly, (6.10) is equivalent to maximizing the sum of throughput weighted by the inverse of circuit power and (6.11) equals maximizing the product of throughput. The dependence of the optimization on circuit power will be further demonstrated later.

In the following, we develop link adaptation and resource allocation strategies in closed-forms, based on optimizing the energy efficient metrics discussed in this section.

6.2 Energy-Efficient Link Adaptation

In this section, we investigate energy-efficient link adaptation for a user with a given channel assignment. Therefore, user index n is dropped in the subsequent discussion in this section.

We need to determine the data rates at all subchannels to maximize

$$\begin{aligned} u[t] =& \frac{T[t]}{P[t]} \\ =& \frac{(1-\frac{1}{w})T[t-1] + \frac{1}{w}\sum_k r_k[t]}{(1-\frac{1}{w})P[t-1] + \frac{1}{w}(\sum_k p_k[t] + p_c[t])}, \end{aligned} \tag{6.13}$$

where $p_k[t+1]$ is given by (6.5).

Denote $c = |C[t]|$, which is the number of elements in $C[t]$ and $C[t] = \{k_i | k_1 < k_2 < \cdots < k_c\} \subseteq \mathcal{K}$. Denote the data rate vector to be $\mathbf{r}[t] = [r_{k_1}[t], r_{k_2}[t], \cdots, r_{k_c}[t]]$. Then $u[t]$ is a function of $\mathbf{r}[t]$. It is easy to see that the sublevel sets

$$S_\alpha = \{\mathbf{r}[t] \big| u[t] \geq \alpha\} \qquad \text{for any real } \alpha, \tag{6.14}$$

are strictly convex. Hence, $u[t]$ is a strictly quasi-concave function optimized on a convex set $\mathbf{r}[t]$ [102] and a unique globally optimal rate vector, $\mathbf{r}^*[t]$, exists and every element in $\mathbf{r}^*[t]$ satisfies

$$\frac{\partial u[t]}{\partial r_k[t]} = 0 \tag{6.15}$$

if $r_k[t] > 0$. Solving (6.15) yields the following optimal rate condition

$$\frac{\partial p_k[t]}{\partial r_k[t]} = \frac{P[t]}{T[t]} = \frac{1}{u[t]}, \forall k. \tag{6.16}$$

Since $w \gg 1$, $P[t] \approx P[t-1]$, and $T[t] \approx T[t-1]$,

$$\frac{\partial p_k[t]}{\partial r_k[t]} = \frac{P[t-1]}{T[t-1]} = \frac{1}{u[t-1]}, \forall k. \tag{6.17}$$

Together with (6.5), we have

$$S^{'}(r_k[t]) = \frac{1}{u[t-1]} \frac{g_k[t]}{\sigma^2}, \forall k. \tag{6.18}$$

where $S^{'}(\cdot)$ is the first-order derivative of the function $S(\cdot)$. Consequently, the optimal data rate follows immediately,

$$r_k^*[t] = \max\left(S^{'-1}(\frac{1}{u[t-1]} \frac{g_k[t]}{\sigma^2}), 0\right) \forall k \in C[t]. \tag{6.19}$$

where $S^{'-1}()$ is the inverse function of $S^{'}$. The corresponding optimal power allocation is

$$p_k^*[t] = \frac{S(r_k^*[t])\sigma^2}{g_k[t]}, \forall k \in C[t]. \tag{6.20}$$

If the Shannon capacity [50] is achieved on each subchannel, $S(r) = 2^{\frac{r}{B}} - 1$, where B is the subchannel bandwidth. The optimal data rate is

$$r_k^*[t] = \max\left(B \log_2 \left(\frac{B g_k[t]}{u[t-1]\sigma^2 \log 2}\right), 0\right) \forall k \in C[t]. \tag{6.21}$$

The corresponding optimal power allocation is

$$p_k^*[t] = \max\left(\frac{B}{u[t-1] \log 2} - \frac{\sigma^2}{g_k[t]}, 0\right) \forall k \in C[t], \tag{6.22}$$

which is a water-filling power allocation with a water level of $\frac{B}{u[t-1]\log 2}$. We can see that the energy-efficient link adaptation in (6.19), (6.20), (6.21), and (6.22) is determined by $u[t-1]$ and $g_k[t]$, and is expressed in closed form. This significantly reduces the complexity associated with the iterative solutions developed earlier in [60]. The low-complexity energy-efficient water-filling power allocation in (6.22) can be illustrated in Figure 6.2, in which every shadowed part corresponds to the power allocated on each subchannel.

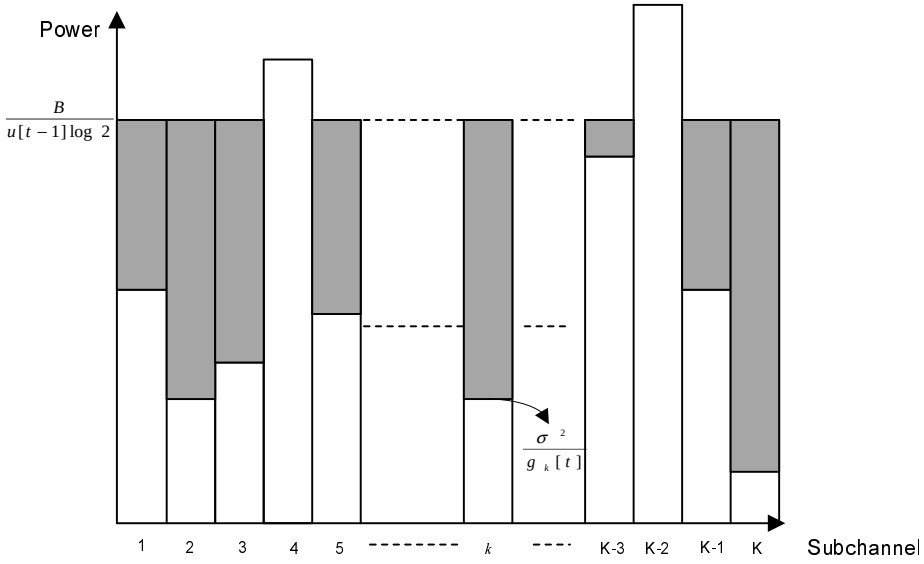


Figure 6.2. Low-complexity energy-efficient water-filling power allocation

6.3 Energy-Efficient Resource Allocation

In this section we will consider low-complexity energy-efficient resource allocation. We will be using the index n to refer to a particular user in this multi-user case. Schedulers based on both the arithmetic and the geometric mean will be considered.

6.3.1 Max Arithmetic Mean Energy-Efficient Scheduler

In this section, the subchannels are assigned such that the sum energy efficiency $U[t]$ is maximized. Since $U[t-1]$ is fixed, it is equivalent to maximize

$$\begin{aligned} \triangle U &= U[t] - U[t-1] \\ &= \sum_{n=1}^{N} u_n[t] - \sum_{n=1}^{N} u_n[t-1] \\ &= \sum_{n=1}^{N} (u_n[t] - u_n[t-1]). \end{aligned} \tag{6.23}$$

We can see that

$$\begin{aligned} u_n[t] - u_n[t-1] &= \frac{T_n[t]}{P_n[t]} - \frac{T_n[t-1]}{P_n[t-1]} \\ &= \frac{T_n[t]P_n[t-1] - P_n[t]T_n[t-1]}{P_n[t]P_n[t-1]}. \end{aligned} \tag{6.24}$$

Substituting Equations (6.3) and (6.7) into (6.24), we have

$$\begin{aligned} & u_n[t] - u_n[t-1] \\ =& \left(P_n[t-1] \sum_{k \in C_n[t]} r_{nk}[t] - T_n[t-1] (\sum_{k \in C_n[t]} p_{nk}[t] + p_{cn}[t]) \right) / (w P_n[t] P_n[t-1]) \\ =& \sum_{k \in C_n[t]} \frac{P_n[t-1] r_{nk}[t] - T_n[t-1] p_{nk}[t]}{w P_n[t] P_n[t-1]} - \frac{T_n[t-1] p_{cn}[t]}{w P_n[t] P_n[t-1]} \\ =& \sum_{k=1}^{K} I_k(C_n[t]) \frac{P_n[t] r_{nk}[t] - T_n[t-1] p_{nk}[t]}{w P_n[t] P_n[t-1]} - \frac{T_n[t-1] p_{cn}[t]}{w P_n[t] P_n[t-1]}, \end{aligned}$$

where indicator function $I_k(C_n)$ is defined as

$$I_k(C_n) = \begin{cases} 1 & k \in C_n, \\ 0 & \text{otherwise.} \end{cases} \tag{6.25}$$

Hence, the subchannel assignment is to maximize

$$\begin{aligned} \triangle U =& \sum_{n=1}^{N} (u_n[t] - u_n[t-1]) \\ =& \sum_{n=1}^{N} \sum_{k=1}^{K} I_k(C_n[t]) \frac{P_n[t-1] r_{nk}[t] - T_n[t-1] p_{nk}[t]}{w P_n[t] P_n[t-1]} \end{aligned}$$

$$
\begin{aligned}
&- \sum_{n=1}^{N} \frac{T_n[t-1]p_{cn}[t]}{wP_n[t]P_n[t-1]} \\
&= \sum_{k=1}^{K}\sum_{n=1}^{N} I_k(C_n[t]) \frac{P_n[t-1]r_{nk}[t] - T_n[t-1]p_{nk}[t]}{wP_n[t]P_n[t-1]} \\
&- \sum_{n=1}^{N} \frac{T_n[t-1]p_{cn}[t]}{wP_n[t]P_n[t-1]}.
\end{aligned}
$$

Denote the allocation metric to be

$$
\begin{aligned}
J(n,k) &= \frac{P_n[t-1]r_{nk}[t] - T_n[t-1]p_{nk}[t]}{P_n[t]P_n[t-1]} \\
&\approx \frac{P_n[t-1]r_{nk}[t] - T_n[t-1]p_{nk}[t]}{P_n^2[t-1]} \\
&= \frac{r_{nk}[t]}{P_n[t-1]} - u_n[t-1]\frac{p_{nk}[t]}{P_n[t-1]},
\end{aligned}
\tag{6.26}
$$

where $r_{nk}[t]$ is given by (6.19) and $p_{nk}[t]$ (6.20).

It is easy to see that $\triangle U$ is maximized by assigning subchannel k to the user with the highest allocation metric $J(n,k)$ on that subchannel, that is, the optimal subchannel assignment is

$$
C_n^* = \{k | J(n,k) > J(m,k), \forall m \neq n\}, \forall n. \tag{6.27}
$$

When the circuit power dominates the power consumption, the allocation metric is

$$
J_t(n,k) \approx \frac{r_{nk}[t]}{P_n[t-1]}. \tag{6.28}
$$

Assume all users consume the same circuit power and $P_n[t-1]$ is the same for all users. Since the user with the maximum $r_{nk}[t]$ is the same as the one with the maximum SINR on that subchannel, the energy-efficient scheduler is equivalent to applying the traditional max-SINR scheduler on each subchannel to achieve the highest spectral efficiency [14], which is,

$$
C_n^* = \{k | r_{n,k} > r_{m,k}, \forall m \neq n\}, \forall n. \tag{6.29}
$$

6.3.2 Max Geometric Mean Energy-Efficient Scheduler

In order to maximize the geometric mean of the energy efficiency of all users, the subchannels are assigned to maximize

$$V[t] = \sum_{n=1}^{N} \log(u_n[t]), \tag{6.30}$$

which is equivalent to maximize

$$\begin{aligned} \triangle V &= V[t] - V[t-1] \\ &= \sum_{n=1}^{N} \log(u_n[t]) - \sum_{n=1}^{N} \log(u_n[t-1]) \\ &= \sum_{n=1}^{N} \left(\log\left(\frac{T_n[t]}{T_n[t-1]}\right) - \log\left(\frac{P_n[t]}{P_n[t-1]}\right)\right). \end{aligned} \tag{6.31}$$

Using the Taylor series expansion and the fact that $w \gg 1$, we can express

$$\begin{aligned} \log\left(\frac{T_n[t]}{T_n[t-1]}\right) &= \log\left(1 - \frac{1}{w} + \frac{\frac{1}{w}\sum_{k\in C_n} r_{nk}[t]}{T_n[t-1]}\right) \\ &\approx \log(1 - \frac{1}{w}) + \frac{\sum_{k\in C_n} r_{nk}[t]}{T_n[t-1](w-1)}. \end{aligned} \tag{6.32}$$

Similarly, we have

$$\begin{aligned} &\log\left(\frac{P_n[t]}{P_n[t-1]}\right) \\ &\approx \log(1 - \frac{1}{w}) + \frac{\sum_{k\in C_n} p_{nk}[t] + p_{cn}[t]}{P_n[t-1](w-1)}. \end{aligned} \tag{6.33}$$

Hence, $\triangle V$ can be expressed by

$$\begin{aligned} \triangle V &= \sum_{n=1}^{N} \left(\frac{\sum_{k\in C_n[t]} r_{nk}[t]}{T_n[t-1](w-1)} \right. \\ &\left. - \frac{\sum_{k\in C_n[t]} p_{nk}[t] + p_{cn}[t]}{P_n[t-1](w-1)} \right) \\ &= \sum_{n=1}^{N} \sum_{k=1}^{K} \left(I_k(C_n[t])\left(\frac{r_{nk}[t]}{T_n[t-1](w-1)}\right.\right. \\ &\left.\left. - \frac{p_{nk}[t]}{P_n[t-1](w-1)}\right)\right) - \sum_{n=1}^{N} \frac{p_{cn}[t]}{P_n[t-1](w-1)} \\ &= \sum_{k=1}^{K} \sum_{n=1}^{N} \left(I_k(C_n[t])\left(\frac{r_{nk}[t]}{T_n[t-1]}\right.\right. \end{aligned}$$

$$-\frac{p_{nk}[t]}{P_n[t-1]})/(w-1)\Big) - \sum_{n=1}^{N} \frac{p_{cn}[t]}{P_n[t-1](w-1)}.$$

Denote the allocation metric to be

$$J_f(n,k) = \frac{r_{nk}[t]}{T_n[t-1]} - \frac{p_{nk}[t]}{P_n[t-1]}, \tag{6.34}$$

where $r_{nk}[t]$ is given by (6.19) and $p_{nk}[t]$ (6.20).

$\triangle V$ is maximized by assigning subchannel k to the user with the highest allocation metric $J_f(n,k)$ on that subchannel, that is, the optimal subchannel assignment achieving proportional fairness is

$$C_n^* = \{k | J_f(n,k) > J_f(m,k), \forall m \neq n\}, \forall n. \tag{6.35}$$

When the circuit power dominates the power consumption, the allocation metric is

$$J_{tf}(n,k) \approx \frac{r_{nk}[t]}{T_n[t-1]}, \tag{6.36}$$

and the energy-efficient scheduler is equivalent to applying the traditional proportional-fair scheduler [76, 16] on each subchannel, that is,

$$C_n^* = \{k | J_{tf}(n,k) > J_{tf}(m,k), \forall m \neq n\}, \forall n. \tag{6.37}$$

6.4 Simulation Results

In the previous sections, we have obtained closed-form and approximate expressions for energy efficient link adaptation and resource allocation, using the average energy efficiency metric. In this section, we compare the proposed schemes with the global optima to evaluate the suboptimality gap. The global optima are obtained by exhaustive search. Since the weight of the exponentially weighted low-pass filter determines approximation accuracy, we focus on its impact on the system performance.

We consider a system with 8 subchannels to reduce complexity of exhaustive search. The subchannels are experiencing independent and identically-distributed Rayleigh fading

with unit average power gain. Capacity approaching coding is assumed. Figure 6.3 shows the suboptimality gap of energy-efficient link adaptation. The energy efficiency of the proposed link adaptation is normalized by the energy efficiency of the global optimal solution. We show the normalized energy efficiency when different weights, ws, are used. We also change the average transmit power to circuit power ratio, ϵ, by varying the circuit power. We can observe that the proposed link adaptation performs closely to the global optimum, with a performance loss of less than 2% when $w > 10$. Similarly, we show the normalized energy efficiency of different schedulers in Figure 6.4 when there are three users in the system. The performance loss is within 5% when $w > 20$ for the proposed schedulers.

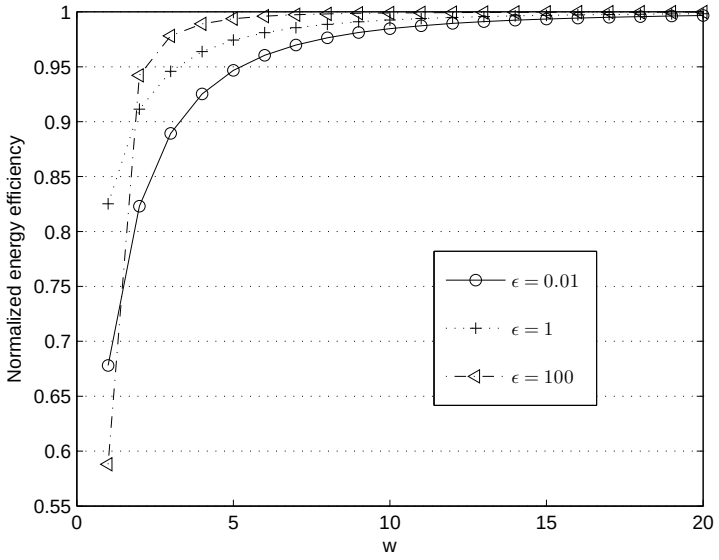


Figure 6.3. Normalized energy efficiency of one link

6.5 Conclusion

We have considered uplink energy-efficient communications in OFDMA systems since mobile stations are battery powered. Time-varying circuit power is accounted for system design. Based on optimizing a time averaged energy efficiency metric, we first obtain a closed-form link adaptation scheme for frequency-selective channels. Furthermore, as a

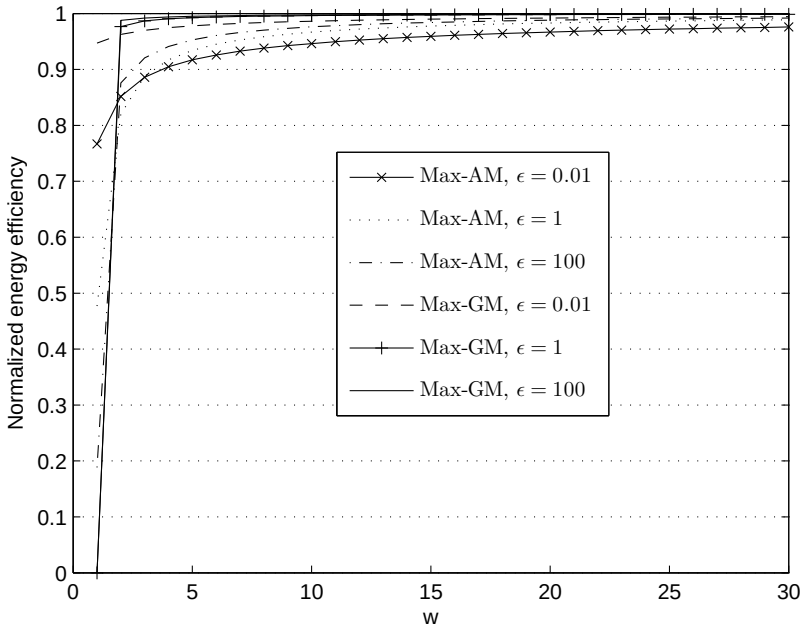


Figure 6.4. Normalized average energy efficiency of a three-user network

system approach is critical in determining the overall network performance, we also design maximum-energy-efficiency and proportional fair energy-efficient schedulers, both in closed forms. Our simulation results show that the proposed low-complexity schemes perform close to the optimum that is obtained by exhaustive search.

CHAPTER 7

INTERFERENCE-AWARE ENERGY-EFFICIENT POWER OPTIMIZATION

In Chapter 5, we have observed that in an interference free environment, a tradeoff between *energy efficiency* (EE) and *spectral efficiency* (SE) exists, as increasing transmit power always improves throughput but not necessarily EE. In this chapter, we consider multi-cell interference-limited scenarios and develop power optimization and resource allocation schemes to improve EE. We note that the general power optimization problem in the presence of interference is intractable even when ideal user cooperation is assumed. We first study this problem for a simple two-user network with ideal user cooperation. Then we develop a noncooperative game for energy-efficient power optimization. We show that the equilibrium always exists. Furthermore, when there is only one subchannel or the channel experiences flat fading, there will be a unique equilibrium. However, in frequency-selective channels, this is not true in general as demonstrated by a counter example. We reveal a sufficient condition that assures the uniqueness. Then we investigate the tradeoff between EE and SE. We show that in interference-limited scenarios, since increased transmit power also brings higher interference to the network, SE is not necessarily increased. Energy-efficient power optimization not only improves system EE but also improves the tradeoff between EE and SE due to the conservative nature of power allocation, which effectively controls interference from other cells to improve network throughput. Later, we also design an implementation of the noncooperative power optimization game.

The rest of the chapter is organized as follows. We first formulate the interference aware power control problem in Section 7.1. In Section 7.2, a two-user network with ideal user cooperation is discussed to gain insights into energy-efficient power control. Noncooperative energy-efficient power optimization is discussed in Section 7.3 and the performance

improvement is demonstrated by simulations in Section 7.4. Finally, we conclude the chapter in Section 7.5.

7.1 Problem Description

We introduce interference-aware energy-efficient power optimization in this section.

Consider a system with K subchannels. Each of them experiences independent and flat fading and AWGN. There are N users, each consisting of a pair of transmitter and receiver and operating on these subchannels. All users interfere with each other. Accurate channel state information is available to any pair of transmitter and receiver. Denote the signal power attenuation of User i at Subchannel k to be $g_{ii}^{(k)}$ and the interference power gain from the transmitter of User i to the receiver of User j at Subchannel k to be $g_{ij}^{(k)}$. The noise power on each subchannel is σ^2. The power allocation of User n on all subchannels is denoted by vector $\mathbf{p}_n = [p_n^{(1)} p_n^{(2)} \cdots p_n^{(K)}]^T$, where T is the transpose. The interference on all subchannels of User n is denoted by vector $\mathbf{I}_n = [I_n^{(1)} I_n^{(2)} \cdots I_n^{(K)}]^T$, where

$$I_n^{(k)} = \sum_{i=1, i\neq n}^{N} p_i^{(k)} g_{in}^{(k)}. \tag{7.1}$$

Consequently, the *signal-to-interference-plus-noise ratio* (SINR), $\eta_n^{(k)}$, of User n at Subchannel k can be expressed as

$$\eta_n^{(k)} = \frac{p_n^{(k)} g_{nn}^{(k)}}{\sum_{i=1, i\neq n}^{N} p_i^{(k)} g_{in}^{(k)} + \sigma^2}. \tag{7.2}$$

The data rate at Subchannel k of User n, $r_n^{(k)}$, is assumed to be a function of η_n and can be expressed as

$$r_n^{(k)} = R(\eta_n^{(k)}), \tag{7.3}$$

where $R()$ depends on the modulation and coding and is assumed to be strictly concave and increasing with $R(0) = 0$. For capacity approaching coding [50], $r_n^{(k)} = w \log(1 + \eta_n^{(k)})$, where w is the bandwidth of each subchannel, and the above assumptions are obviously satisfied.

Let the data rate vector of User n across the K subchannels be $\mathbf{r}_n = [r_n^{(1)}, r_n^{(2)}, \cdots, r_n^{(K)}]^T$, then the overall data rate is

$$r_n = \sum_{k=1}^{K} r_n^{(k)}. \tag{7.4}$$

The total transmit power is

$$p_n = \sum_{k=1}^{K} p_n^{(k)}. \tag{7.5}$$

Note that as in [60, 107], both transmit power and circuit power, p_c, are important for energy-efficient communications. While transmit power is used for reliable data transmission, circuit power represents average energy consumption of device electronics. As in [107], we optimize the energy efficiency, which can be expressed as

$$u_n = \frac{r_n}{\triangle e / \triangle t} = \frac{r_n}{p_n + p_c}, \tag{7.6}$$

where r_n is given by (7.4) and p_n by (7.5). u_n is called EE of User n.

Note that if we fix the overall transmit power, the objective of Equation (7.6) is equivalent to maximizing the overall throughput of all subchannels and existing water-filling power allocation approach [50] gives the solution. However, besides power distributions on all subchannels, the overall transmit power needs to be adapted according to the states of all subchannels to maximize energy efficiency. Hence, the solution to Equation (7.6) is in general different from existing power allocation schemes that maximize throughput with power constraints. The power control in a multi-cell setting to optimize the overall network energy efficiency is also different from traditional power control schemes that emphasize throughput improvement.

We define EE of the overall network to be

$$u = \sum_{n=1}^{N} u_n, \tag{7.7}$$

which is a function of $p_n^{(k)}$ for all n and k. This definition is based on summation of EE of all users rather than the ratio of sum network throughput to sum network power consumption because powers of different users can not be shared and so are their throughput and EE.

We now need to determine power allocation of all users to optimize overall network EE subject to the interference scenario.

7.2 Cooperative Two-User Case

Note that the solution maximizing sum network EE is difficult to obtain as the objective function, in general, is non-concave in $p_n^{(k)}$. To gain some insight, we investigate the case where two users transmit simultaneously on a single channel in this section. We assume both users have complete network knowledge and cooperate to maximize the sum energy efficiency,

$$u(p_1, p_2) = \frac{r_1}{p_1 + p_c} + \frac{r_2}{p_2 + p_c}, \tag{7.8}$$

where

$$r_1 = w \log(1 + \frac{p_1 g_1}{p_2 g_{21} + \sigma^2}) \qquad \text{and} \qquad r_2 = w \log(1 + \frac{p_2 g_2}{p_1 g_{12} + \sigma^2}). \tag{7.9}$$

As u is non-concave in p_1 and p_2, finding the global maximum is intractable. However, we can get some effective approaches by restricting our attention to some special regimes.

7.2.1 Circuit Power Dominated Regime

In this regime, circuit power dominates power consumption, i.e. $p_c \gg p_n$ for $n = 1, 2$. This is usually true for short-range communications as small transmit power is needed to compensate for path loss. In this case, we have

$$u(p_1, p_2) \approx \frac{w}{p_c} \left(\log(1 + \frac{p_1 g_1}{p_2 g_{21} + \sigma^2}) + \log(1 + \frac{p_2 g_2}{p_1 g_{12} + \sigma^2}) \right). \tag{7.10}$$

Hence, maximizing EE is equivalent to maximizing sum network capacity, which has been discussed in literature [70, 69]. The optimal solution takes on the form of binary power control where each user either shuts down or transmits with full power [70]. Whether two users transmit simultaneously or exclusively depends on interference strength.

7.2.2 Transmit Power Dominated Regime

When the circuit power is negligible, e.g. in extremely long distance communications where transmit power should be strong enough to compensate for large path loss,

$$u(p_1, p_2) \approx \frac{w \log(1+\frac{p_1 g_1}{p_2 g_{21}+\sigma^2})}{p_1} + \frac{w \log(1+\frac{p_2 g_2}{p_1 g_{12}+\sigma^2})}{p_2}. \tag{7.11}$$

By examining derivatives of $u(p_1, p_2)$ in Appendix D.1, we can see that $u(p_1, p_2)$ is strictly decreasing with both p_1 and p_2. Hence, the optimal solution is to allocate as low power as possible. However, the above conclusion holds only when the circuit power is negligible. When the transmit power is comparable to the circuit power, other approaches are needed to determine the optimal power.

7.2.3 Noise Dominated Regime

Now we look at the problem from a different perspective. When noise is much stronger than interference, we have

$$u(p_1, p_2) \approx \frac{w \log(1 + \frac{p_1 g_1}{\sigma^2})}{p_1 + p_c} + \frac{w \log(1 + \frac{p_2 g_2}{\sigma^2})}{p_2 + p_c}. \tag{7.12}$$

Hence, the problem is decoupled and the sum network EE is maximized when each user selects power to maximize their own EE, which has already been given in [107, 60].

7.2.4 Interference Dominated Regime

In the interference dominated regime, interference is much stronger than noise, i.e. $p_1 g_{12} \gg \sigma^2$ and $p_2 g_{21} \gg \sigma^2$ for any feasible p_1 and p_2 that support reliable transmission. To be specific, we require that $p_1 g_{12} \gg \sigma^2$ and $p_2 g_{21} \gg \sigma^2$ are significant enough that the *interference-to-noise ratio* (INR) and SINR of each user satisfies

$$INR > 1 + SINR. \tag{7.13}$$

Note that Equation (7.13) does hold when the interference is strong engough since INR increases with interference power while SINR decreases with it. The interference dominated

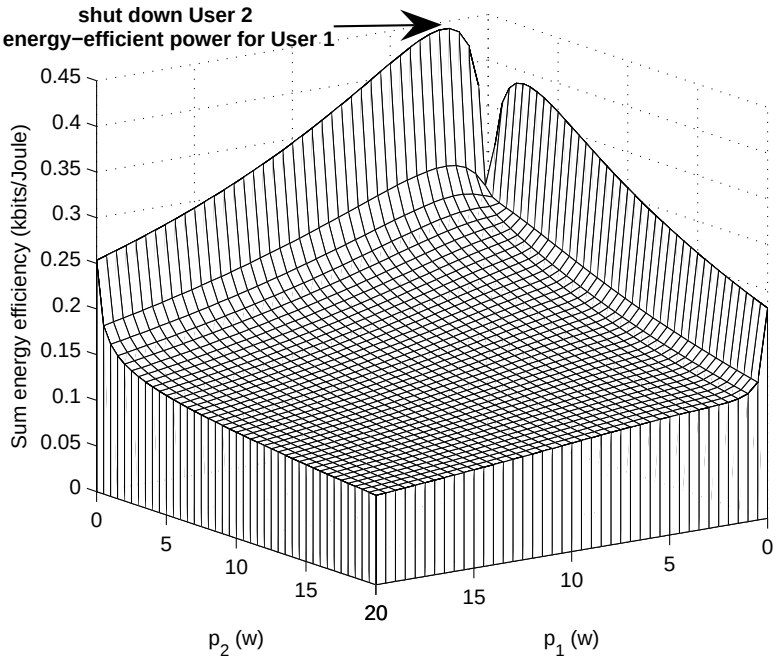
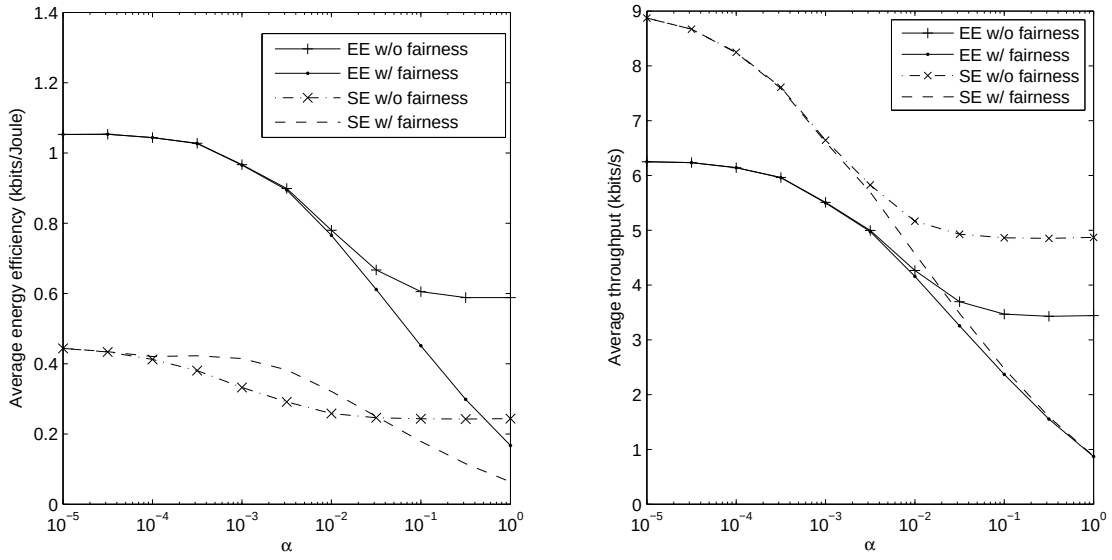


Figure 7.1. Sum energy efficiency and transmit powers in interference dominated regime.

regime exists when different transmissions are close to each other, e.g. closely coupled. Hence,

$$u(p_1, p_2) \approx \frac{w \log(1 + \frac{p_1 g_1}{p_2 g_{21}})}{p_1 + p_c} + \frac{w \log(1 + \frac{p_2 g_2}{p_1 g_{12}})}{p_2 + p_c}. \tag{7.14}$$

In Appendix D.2, we show that $u(p_1, p_2)$ is maximized by an ON-OFF approach, i.e. letting the user with higher channel gain to transmit with energy-efficient power selection and shutting down the other. Figure 7.1 illustrates an example when the average interference-to-noise ratio is 20 dB. In Figure 7.1, the sum energy-efficiency is maximized by shutting down User 2 and choosing power for User 1 to maximize its EE. In this regime, the design of time-division protocols determines network spectral efficiency. Our work on DOMRA, CIA-MAC, and CAD-MAC in Chapters 2, 3, and 4 focus on optimizing the time division in this regime and can be used to improve energy efficiency as well. With these schemes, energy-efficient link adaptation should be used when a user accesses the channel if energy efficiency is the primary concern.



(a) Energy efficiency.

(b) Spectral efficiency.

Figure 7.2. Comparison of cooperative EE and SE.

7.2.5 Spectral Efficiency and Energy Efficiency Tradeoff with Cooperation

Our previous research [60] has shown that maximizing EE and maximizing SE usually disagree. Therefore, tradeoff between them exists. To examine the impact of interference on this tradeoff when ideal cooperation exists, we consider a symmetric two-user network and compare energy-efficient schemes with spectral-efficient ones. Both users experience Rayleigh fading. Different power optimizations result in different interference scenarios. To characterize interference level, we need to use a metric that is independent of transmit powers. Hence, define network coupling factor α,

$$\alpha = \frac{\text{average interference channel gain}}{\text{average signal channel gain}}. \tag{7.15}$$

α characterizes what level different transmissions interfere with each other and higher α represents heavier interfering scenario.

Consider two network performance metrics, arithmetic and geometric means. The power is optimized to maximize the arithmetic- or geometric-mean metric. It has been indicated in [94] that optimization based on the arithmetic average of SE leads to power

allocation for sum throughput maximization that considers no fairness since some users may have zero throughput, and optimization based on the geometric average also assures proportional fairness among all users. Optimization based on the arithmetic and geometric averages of EE has similar characteristics and we call arithmetic- or geometric-mean metrics for EE optimization to be energy-efficient power optimization schemes without or with fairness.

In Figure 7.2, the EEs and throughputs of four schemes are compared when α has different values. The schemes are the energy-efficient and spectral-efficient power allocation either with or without proportional fairness. A peak power constraint is applied in the power allocation. The schemes without fairness allocate power to maximize the sum of either EE (energy-efficient) or throughput (spectral-efficient) while those with proportional fairness maximize the product. From Figure 7.2, we can see that the tradeoff between SE and EE depends on the network coupling.

7.3 Noncooperative Energy-Efficient Communications

The above section discusses energy-efficient power optimization with ideal cooperation in a two-user network. Extension to special regimes for a multi-user network is straightforward and omitted. However, in general, it is difficult to determine the globally optimal power allocation due to the nonconcavity of sum energy-efficiency functions. More users and subchannels will result in more local maximums and searching the globally optimal power allocation would be a daunting task. Even if the globally optimal solution can be found, it is still impractical since it requires complete network knowledge, including interference channel gains. In this section, we consider a more practical case and assume no cooperation among users. To assure fairness, all users apply the same policy using local information. In the following, we first model the noncooperative energy-efficient control from a game-theory perspective and then discuss the existence and uniqueness of its equilibrium. Then we investigate SE and EE tradeoff assuming symmetric channel condition

to obtain insights. After that, we further develop a noncooperative energy-efficient power control scheme to facilitate implementation.

7.3.1 Noncooperative Energy-Efficient Power Optimization Game

Since the network energy efficiency depends on the behaviors of two or more users, we model the power control to be a noncooperative game in game theory [108]. Rooted in economics, game theory has been broadly applied in wireless communications for random access and power control optimizations [68, 69].

In a noncooperative game, each user optimizes power allocation to maximize its energy efficiency. Consider the power allocation of User n and denote the power vectors of other users to be vector

$$\mathbf{p}_{-n} = (\mathbf{p}_1^T, \mathbf{p}_2^T, \cdots, \mathbf{p}_{n-1}^T, \mathbf{p}_{n+1}^T, \cdots, \mathbf{p}_N^T)^T. \tag{7.16}$$

Given $\mathbf{p}_{-n}$, the best response of the power allocation of User n is

$$\mathbf{p}_n^o = f_n(\mathbf{p}_{-n}) = \arg\max_{\mathbf{p}_n} u_n(\mathbf{p}_n, \mathbf{p}_{-n}), \tag{7.17}$$

where u_n is given by (7.6) and is a function of both $\mathbf{p}_n$ and $\mathbf{p}_{-n}$. $f_n(\mathbf{p}_{-n})$ is called the best response function of User n. The existence and uniqueness of $\mathbf{p}_n^o$, i.e. the best response, is assured by Theorem I in our previous work [107].

Note that noncooperative power control is not efficient in terms of SE optimization since users tend to act selfishly by increasing their transmit power beyond what is reasonable [68]. Hence, pricing mechanisms are introduced to regulate the aggressive power transmission by individuals to produce more socially beneficial outcome towards improving sum throughput of all users [69]. Different from SE optimal power control, energy-efficient power optimization desires a power setting that is greedy in EE but chary of power. Furthermore, Problem (7.17) is equivalent to

$$\begin{aligned} \mathbf{p}_n^o &= \arg\max_{\mathbf{p}_n} \log(u_n(\mathbf{p}_n, \mathbf{p}_{-n})) \\ &= \arg\max_{\mathbf{p}_n} \left(\log(r_n) - \log(p_n + p_c)\right), \end{aligned} \tag{7.18}$$

which implies that energy-efficient power control can be regarded as a variation of traditional spectral-efficient one with power pricing [69]. Since this power-conservative expression is socially favorable in interference-limited scenarios, energy-efficient power control is desirable to reduce interference and improve throughput in a noncooperative setting.

Each user optimizes their power independently. The variation of power allocation of one user impacts those of all others. Equilibrium is the condition of a network in which competing influences are balanced assuming invariant channel conditions. Its properties are important to network performance. Hence, we characterize the equilibrium of noncooperative energy-efficient power optimization in the following three sections.

7.3.2 Existence of Equilibrium

In a noncooperative game, a set of strategies is said to be at Nash equilibrium, referred as equilibrium in the following, if no user can gain individually by unilaterally altering its own strategy. Denote the equilibrium as

$$\mathbf{p}^* = (\mathbf{p}_1^*, \mathbf{p}_2^*, \cdots, \mathbf{p}_N^*). \tag{7.19}$$

Nash equilibrium can be described by the following definition.

Definition 7.3.1 *In an energy-efficient noncooperative game, an equilibrium is a set of power allocation that no user can unilaterally improve its energy efficiency by choosing a different set of power allocation, i.e.*

$$\mathbf{p}^* = f(\mathbf{p}^*) = (f_1(\mathbf{p}_{-1}^*), f_2(\mathbf{p}_{-2}^*), \cdots, f_N(\mathbf{p}_{-N}^*)), \tag{7.20}$$

where $f(\mathbf{p})$ *is the network response function.*

The network response relies on energy efficiency of all users. In the following, we first give the properties of energy efficiency function and then study the existence of equilibrium. To facilitate our discussion, we first introduce the concept of quasiconcavity.

Definition 7.3.2 *As defined in [102], a function z, which maps from a convex set of real n-dimensional vectors, $\mathcal{D}$, to a real number, is called strictly quasiconcave if for any $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{D}$ and $\mathbf{x}_1 \neq \mathbf{x}_2$,*

$$z(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) > \min\{z(\mathbf{x}_1), z(\mathbf{x}_2)\}, \tag{7.21}$$

for any $0 < \lambda < 1$.

Lemma 7.3.3 characterizes the energy efficiency function and is proved in Appendix D.3.

Lemma 7.3.3 *$u_n(\mathbf{p}_n, \mathbf{p}_{-n})$ is strictly quasiconcave in $\mathbf{p}_n$.*

Based on Lemma 7.3.3, the existence of equilibrium $\mathbf{p}^*$ is given by Theorem 7.3.4. A necessary and sufficient condition for a set of power allocation to be an equilibrium is also summarized in Theorem 7.3.4. The proof can be found in Appendix D.4.

Theorem 7.3.4 (Existence) *There exists at least one equilibrium $\mathbf{p}^*$ in the noncooperative energy-efficient power optimization game defined by (7.17). A set of power allocation of all users, $\mathbf{p}^* = (\mathbf{p}_1^*, \mathbf{p}_2^*, \cdots, \mathbf{p}_N^*)$, is an equilibrium if and only if it satisfies that, for any Subchannel i of any User n,*

1. *if $\frac{\sum_{j\neq i} r_n^{(j)*}}{p_c + \sum_{j\neq i} p_n^{(j)*}} \leq R'(0)\gamma_n^{(i)*}$, $\left.\frac{\partial u_n(\mathbf{p}_n, \mathbf{p}_{-n}^*)}{\partial p_n^{(i)}}\right|_{\mathbf{p}_n = \mathbf{p}_n^*} = 0$, i.e. $R'(\gamma_n^{(i)*} p_n^{(i)*})\gamma_n^{(i)*} = u(\mathbf{p}_n^*, \mathbf{p}_{-n}^*)$;*
2. *otherwise, $p_n^{(i)*} = 0$,*

where $\gamma_n^{(i)*} = \frac{g_{nn}^{(i)}}{\sum_{j=1, j\neq n}^{N} p_j^{(i)*} g_{jn}^{(i)} + \sigma^2}$.

7.3.3 Uniqueness of Equilibrium in Flat Fading Channels

In this section, we discuss the uniqueness of the equilibrium. First, we consider a special case when there is a single subchannel in a network and

$$p_n^o = f_n(\mathbf{p}_{-n}) = \arg\max_{p_n} u_n(p_n, \mathbf{p}_{-n}). \tag{7.22}$$

Proposition 7.3.5 shows the properties of the response functions and is proved in Appendix D.5.

Proposition 7.3.5 *When there is only one subchannel, the power allocation, i.e. the response functions, of all users satisfy*

- *Concavity:* $f_n(\mathbf{p}_{-n})$ *is strictly concave in* $\mathbf{p}_{-n}$*;*
- *Positivity:* $f_n(\mathbf{p}_{-n}) > 0$*;*
- *Monotonicity: If* $\mathbf{p}_{-n} > \mathbf{q}_{-n}$, $f_n(\mathbf{p}_{-n}) > f_n(\mathbf{q}_{-n})$*;*
- *Scalability: For all* $\alpha > 1$, $\alpha f_n(\mathbf{p}_{-n}) > f_n(\alpha \mathbf{p}_{-n})$,

where $>$ *denotes vector inequality and each element of the vector satisfies the inequality.*

Note that the monotonicity indicates that increasing interference results in increasing transmit power while the scalability indicates that variation of transmit power is always smaller than that of the interference power. These assure the convergence to a unique equilibrium.

The properties in Proposition 7.3.5 can be extended to networks with multiple subchannels where all subchannels experience the same channel gain, i.e. flat-fading channels. This can be done by defining $f_n(\mathbf{p}_{-n})$ to be the optimal total transmit power on all subchannels and the four properties can be easily verified by the approaches in Appendix D.5.

Theorem 7.3.6 (Uniqueness) *When the channel experiences flat fading, there exists one and only one equilibrium* $\mathbf{p}^*$ *in the noncooperative energy-efficient power optimization game defined by (7.17).*

Proof It has been shown in [109] that a noncooperative power control with positivity, monotonicity, and scalability has a unique fixed point $\mathbf{p} = f(\mathbf{p})$. Hence, we have the above theorem.

7.3.4 Uniqueness of Equilibrium in Frequency-Selective Channels

When there are multiple subchannels which experience frequency-selective fading, whether there is a unique equilibrium depends on channel conditions.

Consider a network with two users as an example. Let $p_c = 1, w = 1, \sigma^2 = 1, g_{11}^{(1)} = g_{11}^{(2)} = g_{22}^{(1)} = g_{22}^{(2)} = 1, g_{12}^{(1)} = g_{21}^{(1)} = 1e^{-10}, g_{12}^{(2)} = g_{21}^{(2)} = 1e^{10}$. We show in Appendix D.6 that one of the equilibrium has the form $\mathbf{p}_1^* = [p_a \ p_b]$ and $\mathbf{p}_2^* = [p_c \ 0]$, where p_a, p_b, and p_c are positive. Due to the symmetry of network conditions, $\mathbf{p}_1 = [p_c \ 0]$ and $\mathbf{p}_2 = [p_a \ p_b]$ also form an equilibrium. Hence, the network has at least two equilibria. When there are more users and subchannels, more equilibria will exist in general. However, in the following, we will show that when the interfering channels satisfy a certain condition, there will be a unique equilibrium.

We consider a general noncooperative power control over multiple subchannels where each user selfishly chooses power allocation to maximize its own utility in an interference-limited environment. The utility, denoted by $U_n(\mathbf{p}_n, \mathbf{I}_n(\mathbf{p}_{-n}))$, is assumed to be quasiconcave in $\mathbf{p}_n$ given $\mathbf{I}_n$, interferences on all subchannels. $\mathbf{I}_n$ is a function of $\mathbf{p}_{-n}$ and is determined by (7.1). The best response of power allocation of User n is denoted to be

$$\mathbf{p}_n^o = F_n(\mathbf{p}_{-n}) = \tilde{F}_n(\mathbf{I}_n(\mathbf{p}_{-n})) = \arg\max_{\mathbf{p}_n} U_n(\mathbf{p}_n, \mathbf{I}_n(\mathbf{p}_{-n})). \tag{7.23}$$

The noncooperative energy-efficient power optimization in (7.17) is an example of (7.23).

Denote the Jacobian matrix of $\tilde{F}_n$ at $\mathbf{I}_n$ to be $\frac{\partial \tilde{F}_n}{\partial \mathbf{I}_n}$ and the Jacobian matrix of $\mathbf{I}_n$ at $\mathbf{p}_{-n}$ to be $\frac{\partial \mathbf{I}_n}{\partial \mathbf{p}_{-n}}$. Then

$$\frac{\partial \tilde{F}_n}{\partial \mathbf{I}_n} = \begin{pmatrix} \frac{\partial p_n^{(1)o}}{\partial I_n^{(1)}} & \cdots & \frac{\partial p_n^{(K)o}}{\partial I_n^{(1)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial p_n^{(1)o}}{\partial I_n^{(K)}} & \cdots & \frac{\partial p_n^{(K)o}}{\partial I_n^{(K)}} \end{pmatrix} \tag{7.24}$$

and

$$\frac{\partial \mathbf{I}_n}{\partial \mathbf{p}_{-n}} = \begin{pmatrix} g_{1n}^{(1)} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & g_{1n}^{(K)} \\ & \vdots & \\ g_{(n-1)n}^{(1)} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & g_{(n-1)n}^{(K)} \\ g_{(n+1)n}^{(1)} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & g_{(n+1)n}^{(K)} \\ & \vdots & \\ g_{Nn}^{(1)} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & g_{Nn}^{(K)} \end{pmatrix}. \tag{7.25}$$

Denote $\|A\|$ to be the Frobenius norm of matrix $A = (a_{ij})$, i.e. $\|A\| = \sqrt{\sum_{i,j} a_{ij}^2}$. We know that when a contraction mapping has a fixed point, the fixed point is unique [110]. Readily, we have the following sufficient condition, which comes from [111], that assures a unique equilibrium.

Theorem 7.3.7 (Uniqueness) *In frequency selective channels, if for any User n,* $\|F_n(\mathbf{p}_{-n}) - F_n(\check{\mathbf{p}}_{-n})\| < \|\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n}\|$ *for any different* $\mathbf{p}_{-n}$ *and* $\check{\mathbf{p}}_{-n}$*, there exists one and only one equilibrium* $\mathbf{p}^*$ *in the noncooperative power control game defined by (7.23).*

Intuitively, Theorem 7.3.7 says that if other users change their transmit powers by some amount, the best power allocation of the user is altered by a lesser amount, then the equilibrium is unique. Note that the transmit powers of other users and the best response $F_n(\mathbf{p}_{-n})$ in (7.23) are related through interference channel gains, which therefore determines the variation of the best response and whether the sufficient condition can be guaranteed. Stronger

interference channel gains result in higher correlation and vice versa. The above two-user network illustrates an example where one subchannel has extremely strong interference channel gains. In this case, the sufficient condition is violated and there are multiple equilibria.

Based on Theorem 7.3.7, Theorem 7.3.8 explicitly shows the impact of interference channel gains on the number of equilibria and is proved in Appendix D.7.

Theorem 7.3.8 (Uniqueness) *In frequency selective channels, if for any User n,*

$$\left\| \frac{\partial \mathbf{I}_n}{\partial \mathbf{p}_{-n}} \right\| < \frac{1}{\sup_{\mathbf{I}_n} \left\| \frac{\partial \tilde{F}_n}{\partial \mathbf{I}_n} \right\|}, \tag{7.26}$$

where $\sup_{\mathbf{I}_n}$ *is the supremum on all feasible* $\mathbf{I}_n$*, there exists one and only one equilibrium* $\mathbf{p}^*$ *in the noncooperative power control game defined by (7.23).*

After examining the Jacobian matrices, we see that the left hand side of (7.26) depends on interference channel gains only, while the right hand side is independent of interference channel gains. Hence, interference channel gains directly impacts the number of equilibria. Consider an example where different users are sufficiently far away and all interference channel gains are close to zero. It is easy to see that transmit powers of other users have almost no effect on the best response of the user and there is a unique equilibrium.

Note that while a sufficient condition of a unique equilibrium for distributed power control over a single channel is given in [109], we provide sufficient conditions of a unique equilibrium for distributed multichannel power controls in Theorems 7.3.7 and 7.3.8, which can be applied to different kinds of distributed *multiple input multiple output* (MIMO) and *orthogonal frequency-division multiplexing* (OFDM) systems.

Given Theorems 7.3.7 and 7.3.8, a sufficient condition to assure a unique equilibrium of the noncooperative energy-efficient power optimization follows immediately.

Theorem 7.3.9 (Uniqueness) *In frequency selective channels, the noncooperative energy-efficient power optimization game defined by (7.17) has a unique equilibrium if for any User n,* $\|f_n(\mathbf{p}_{-n}) - f_n(\check{\mathbf{p}}_{-n})\| < \|\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n}\|$ *for any different* $\mathbf{p}_{-n}$ *and* $\check{\mathbf{p}}_{-n}$ *or* $\left\| \frac{\partial \mathbf{I}_n}{\partial \mathbf{p}_{-n}} \right\| < \frac{1}{\sup_{\mathbf{I}_n} \left\| \frac{\partial f_n}{\partial \mathbf{I}_n} \right\|}$.

Note that the above theorem only gives sufficient conditions of uniqueness that may not be necessary ones. For example for a single-channel network, due to the strict concavity of $f_n(\mathbf{p}_{-n})$, $\sup_{I_n} \left\| \frac{\partial f_n}{\partial I_n} \right\| = \left. \frac{\partial f_n}{\partial I_n} \right|_{I_n=0}$. However, for all interference channel gains, there is always a unique equilibrium, as shown in Theorem 7.3.6.

7.3.5 SE and EE Tradeoff without Cooperation

In this section, we investigate the tradeoff between noncooperative energy-efficient power optimization and noncooperative spectral-efficient power control schemes. Here, no peak power constraint is assumed to investigate performance limit.

Consider a symmetric single-channel network to simplify analysis and to get insights. There are N users, all experiencing the same channel power gain g. All interference channels have the same power gain $\widetilde{g}$. The network coupling factor is

$$\alpha = \frac{\widetilde{g}}{g}. \tag{7.27}$$

Consider the equilibrium, which is unique according to Theorem 7.3.6. Due to the assumption of network symmetry, all users transmit with the same power in the equilibrium. Denote the transmit power of all users to be p.

The overall network EE is

$$u(p) = \sum_{n=1}^{N} \frac{w \log\left(1 + \frac{pg}{\sum_{i,i\neq n} p\widetilde{g} + \sigma^2}\right)}{p + p_c} = \frac{Nw \log\left(1 + \frac{p}{(N-1)\alpha p + \frac{\sigma^2}{g}}\right)}{p + p_c}, \tag{7.28}$$

and the network SE is

$$r(p) = N \log\left(1 + \frac{p}{(N-1)\alpha p + \frac{\sigma^2}{g}}\right). \tag{7.29}$$

With noncooperative spectral-efficient power control, every user allocates power to selfishly maximize its SE. Without power limit, the transmit power tends to infinity in the equilibrium. Besides, we can see that $r(p)$ is strictly increasing in p. Hence, the maximum network SE is obtained in the equilibrium and the upperbound is

$$r_{SE} = \lim_{p\to\infty} r(p) = N \log(1 + \frac{1}{(N-1)\alpha}) \tag{7.30}$$

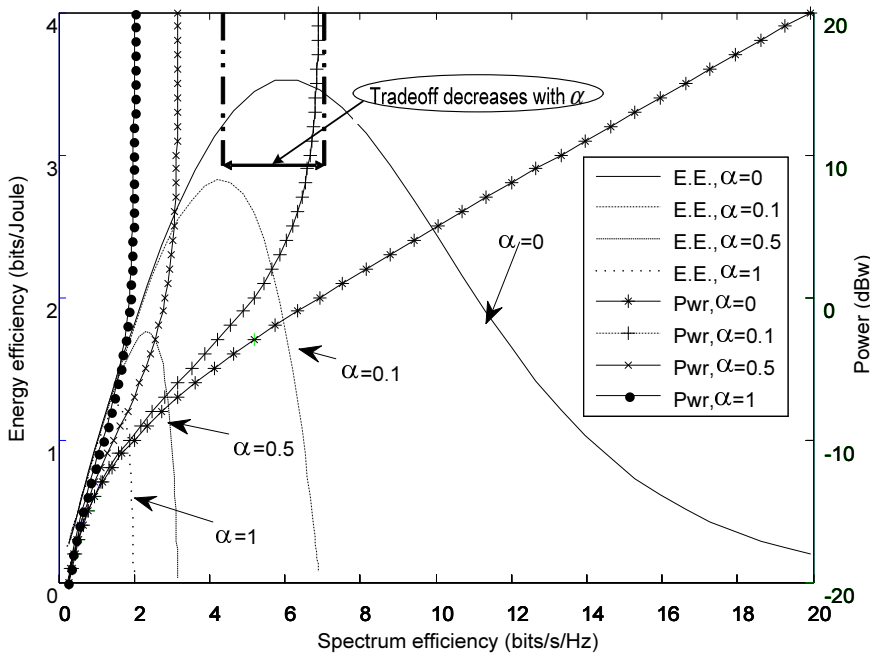


Figure 7.3. Tradeoff of EE and SE with different interfering scenarios ($p_c = 1, g = 1, \sigma^2 = 0.01, N = 2$)**.**

with the corresponding EE $u_{SE} = \lim_{p\to\infty} u(p) = 0$, which is completely energy inefficient and noncooperative SE optimal power control is not desired for energy efficiency.

With noncooperative energy-efficient power optimization, the network energy efficiency at the equilibrium is $u_{EE} = u(p^*)$ with the corresponding SE $r_{EE} = r(p^*)$. Hence, the SE penalty of energy-efficient power optimization is

$$r_{tr} = r_{SE} - r_{EE} = N\log(1 + \frac{1}{(N-1)\alpha}) - r(p^*). \tag{7.31}$$

In an interference-free scenario, i.e. $N = 1$ or $\alpha = 0$, the penalty is infinite. Otherwise, whenever interference exists, it is bounded.

To further understand the tradeoff, Figure 7.3 illustrates a case when two users transmit with the same power and interfere with each other. Curves with markers draw the relationship between transmit power and SE when the network has different couplings while those without markers draw the corresponding energy efficiency. When $\alpha = 0$, arbitrary SE can be obtained by choosing enough transmit power. When $\alpha > 0$, regions beyond the SE upperbound is not achievable. Furthermore, EE is much more sensitive to power selection than SE. In interference-limited scenarios, increasing transmit power beyond the optimal

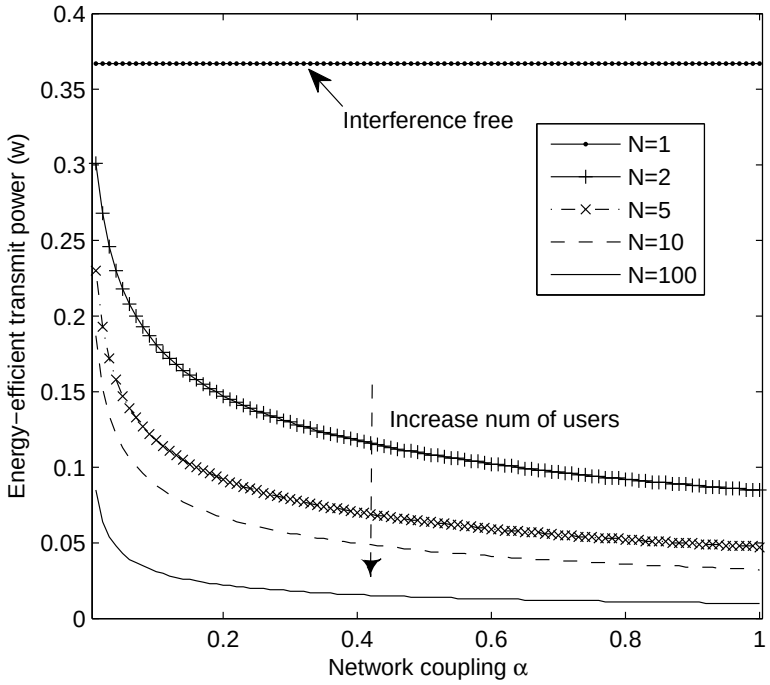


Figure 7.4. Noncooperative energy-efficient power optimization in the equilibrium($P_c = 1, g = 1, \sigma^2 = 0.01$).

power for EE has little SE improvement but significantly hurts EE. Furthermore, power optimization to achieve the highest energy efficiency will also have reduced SE tradeoff with the increase of α. Figure 7.4 shows the transmit power in the equilibrium when the network has different couplings and numbers of users. The equilibrium power decreases with either user number or α and automatically alleviates network interference.

7.3.6 Implementation of Noncooperative Energy-Efficient Power Optimization

In the previous section we know that energy-efficient power optimization is advantageous in interference-limited scenarios due to its conservative power allocation nature. In this section, we will develop practical approaches for noncooperative energy-efficient power optimization.

In (7.17), the best response of User n depends on the transmit power vectors of all other users, $\mathbf{p}_{-n}$, which can not be obtained in a noncooperative setting. Instead, we observe that $\mathbf{p}_{-n}$ affects the best response in the form of interference, which thus contains sufficient

information of $\mathbf{p}_{-n}$ to determine the best response and can be acquired locally. Hence, we let each user measure interferences on all subchannels to determine the power optimization.

At time $t-1$, the measured interference powers on all subchannels of User n are denoted by $\mathbf{I}_n[t-1] = [I_n^{(1)}[t-1], I_n^{(2)}[t-1], \cdots, I_n^{(K)}[t-1]]^T$. Denote the predicted SINR to be

$$\widehat{\eta}_n^{(k)}[t] = \frac{p_n^{(k)}[t] g_{nn}^{(k)}}{\widehat{I}_n^{(k)}[t-1] + \sigma^2} = \frac{p_n^{(k)}[t] g_{nn}^{(k)}}{I_n^{(k)}[t-1] + \sigma^2}. \tag{7.32}$$

Hence, the predicted EE is

$$\widehat{u}_{n[t]}(\mathbf{p}_n[t]) = \frac{\widehat{r}_n[t]}{p_n[t] + p_c} = \frac{\sum_k R(\widehat{\eta}_n^{(k)}[t])}{\sum_k p_n^{(k)}[t] + p_c}. \tag{7.33}$$

The best response at time t of User n is

$$\mathbf{p}_n^o[t] = \arg\max_{\mathbf{p}_n[t]} \widehat{u}_{n[t]}(\mathbf{p}_n[t]). \tag{7.34}$$

Due to the strictly quasi-concavity of $\widehat{u}_{n[t]}$, numerical methods like gradient ascent algorithms can be used to find the optimal power allocation at each time slot. A *Binary Search Assisted Ascent* algorithm has been developed in [107]. However, if we obtain the optimal power allocation at each time slot, it requires intensive computations. Instead, we introduce a *temporal iterative binary search* (TIBS) algorithm to track channel temporal variation and search for the optimal power allocation with reduced complexity.

The basic idea of TIBS is to search a better power allocation along the gradient at each time slot and enable iterative search along time. The power at t is updated by

$$\mathbf{p}_n[t] = \mathbf{p}_n[t-1] + \mu(\nabla \widehat{u}_{n[t]})_{\mathbf{p}_n[t-1]}, \tag{7.35}$$

where $(\nabla \widehat{u}_{n[t]})_{\mathbf{p}_n[t-1]}$ is the gradient of $\widehat{u}_{n[t]}$ at $\mathbf{p}_n[t-1]$ and μ is a small step size. Fixing channel states and transmit powers of all other users, the EE at t will always be bigger than that at $t-1$ with sufficiently small step size except when the gradient is zero, i.e. $\mathbf{p}_n[t-1]$ is already optimal [104]. However, small step size leads to slow convergence and channel tracking capability. Denote

$$g(\mu) = \widehat{u}_{n[t]}(\mathbf{p}_n[t-1] + \mu(\nabla \widehat{u}_{n[t]})_{\mathbf{p}_n[t-1]}). \tag{7.36}$$

It is easy to show that $g(\mu)$ is also strictly quasi-concave in μ and binary search can be used for rapid location of the optimal step size μ^* [107]. TIBS is summarized in the following algorithm.

Algorithm *Temporal Iterative Binary Search (TIBS)*

($*$ noncooperative energy-efficient power optimization $*$)

Input: $\mathbf{p}[t-1]$, $\mathbf{I}[t-1]$

Output: $\mathbf{p}[t]$

1. use *Gradient Assisted Binary Search*([107]) to find the optimal step size μ^*;
2. $\mathbf{p}[t] = \mathbf{p}[t-1] + \mu^*(\widehat{\nabla u}[t])_{\mathbf{p}[t-1]}$,
3. **return** $\mathbf{p}[t]$

Table 7.1. System parameters

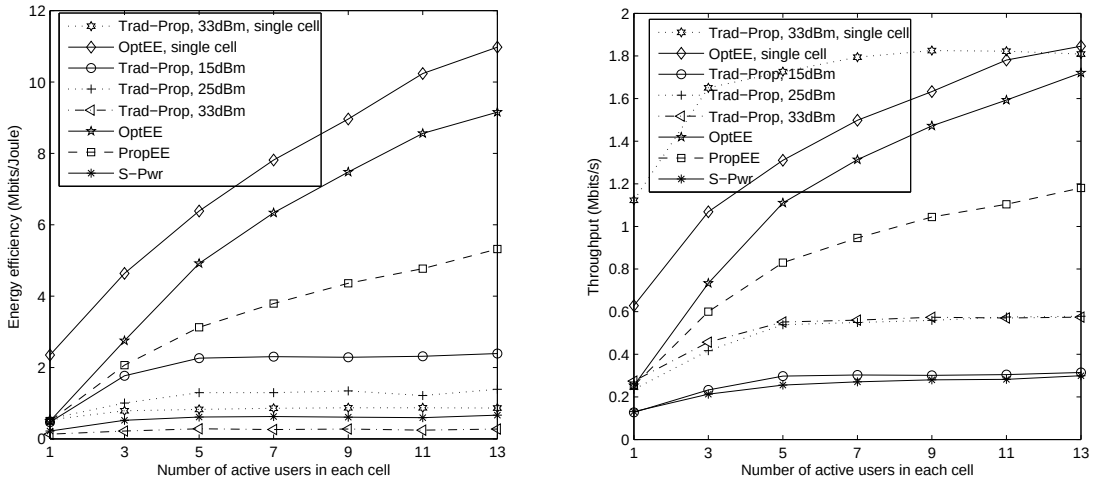
Carrier frequency	1.5 GHz
Number of subchannels	96
Subchannel bandwidth	10 kHz
Target BER	10^{-3}
Thermal noise power, N_o	-141 dBW/MHz
Circuit power, P_C	100 mW
Maximum transmit power	33 dBm
Propagation model	Okumura-Hata model
Shadowing	Log-normal
Modulation	Uncoded M-QAM

7.4 Simulation Results

Table 7.2. Scheduling and power control

Legend	Scheduler	Power control
OptEE	Energy-efficient scheduler w/o fairness	TIBS
PropEE	Energy-efficient scheduler w/ proportional fairness	TIBS
Trad-Prop	Traditional proportional fair	Fixed power
S-Pwr	Traditional proportional fair	Traditional power control

In this section we present simulation results for an interference-limited uplink OFDMA cellular network with reuse one. The network consists of seven hexagonal cells and the



(a) Energy efficiency.

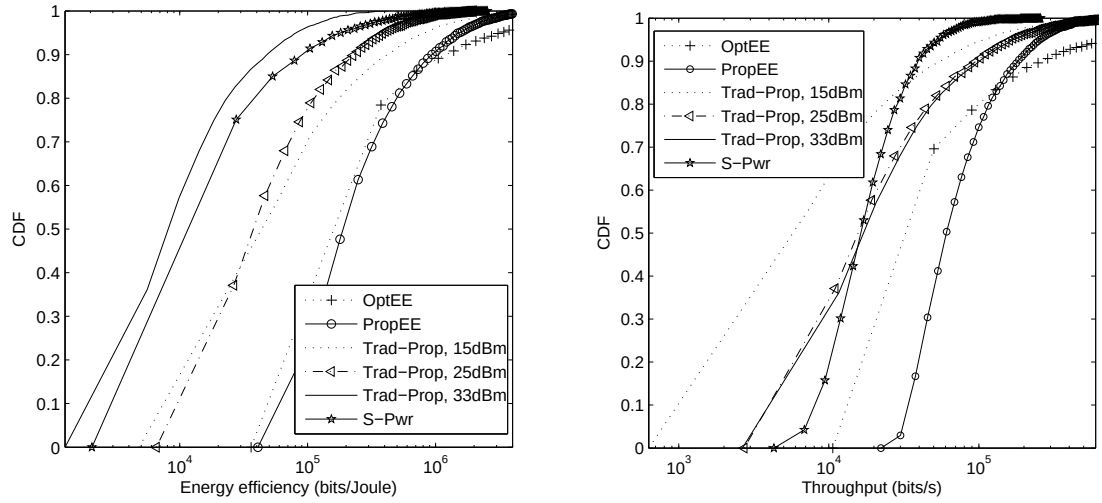
(b) Throughput.

Figure 7.5. Performance comparison of different schemes.

center cell is surrounded by the other six. Users are uniformly dropped into each cell at each simulation trial. The system parameters are listed in Table 7.1. The BS schedules subchannels to maximize different network performance metrics. All schedulers and corresponding power control schemes are listed in Table 7.2. The energy-efficient schedulers in [60] are used. The traditional proportional fair scheduler assigns subchannels to the user with the highest $\frac{r}{T}$, where r is the instantaneous data rate on that subchannel and T the average total throughput [16]. While energy-efficient schedulers assign subchannels to different users to maximize EE either with or without fairness, the traditional proportional fair scheduler assigns all subchannels to one user at each time slot due to flat fading. We also implement a traditional soft power control scheme [112]. In this scheme, parameters are selected to maximize the throughput of cell-edge users while not hurting the throughput of other users too much.

Figure 7.5 compares the average sum network EE and the corresponding throughput performance respectively. For fixed-power transmission, the transmit powers are shown in the legend. To see performance loss due to interference, the energy-efficient scheduler without fairness and the traditional proportional scheduler with the maximum transmit

power is also simulated in a single cell network. We can see that transmitting with the highest power brings the highest interference and causes significant throughput loss for the traditional scheduler. In contrast, energy-efficient power control effectively reduces network interference and has much less throughput loss. While our previous results in [60] show that EE and throughput efficiency do not necessarily agree for an interference-free single cell scenario, the situation is different for a multi-cell interference-limited network. Here energy-efficient schemes optimize both throughput and energy utilization and exhibit an improved SE tradeoff. Figure 7.6 further shows the *cumulative distribution functions* (CDFs) of energy efficiency and throughput when there are nine users in the network. Observe the throughput CDF of the soft power control scheme. Compared with other traditional schemes, it maximizes cell-edge throughput that is illustrated in low-throughput range. However, it performs much worse than other traditional schemes in high-throughput range. From the CDFs, we can see that the proposed EE schemes not only improve the sum energy efficiency and throughput, but also uniformly improve the performance of all users in the cell.



(a) Energy efficiency. (b) Throughput.

Figure 7.6. Comparison of different schemes.

7.5 Conclusion

Since power optimization is critical for both interference management and energy utilization, we investigate energy-efficient power optimization schemes for interference-limited communications. The optimal power allocation solution in an interference-limited setting is intractable due to the non-convexity of the objective function. To gain insight into this problem, we first study a two-user network with ideal user cooperation and get effective approaches for specific regimes. Then we develop a noncooperative energy-efficient power optimization game. We show that the equilibrium always exists. Furthermore, when there is only one subchannel or the channel experiences flat fading, there will be a unique equilibrium. However, in frequency-selective channels, this is not true in general. We give a sufficient condition that assures the uniqueness. We further show that the spectral efficiency tradeoff of energy-efficient power control is reduced in interference-limited scenarios. Then we develop a practical approach of the noncooperative power optimization game. Simulation results show that the proposed scheme improves not only energy efficiency but also spectral efficiency uniformly for all users due to the conservative nature of power allocation achieved, which reduces other-cell interference to improve the overall network throughput.

CHAPTER 8

CONCLUSION

In this dissertation, we exploit CSI to improve both spectral and energy efficiency of wireless communications. The main contribution is summarized as follows.

We have investigated a series of general treatments of exploiting CSI in a distributed way to control the medium access for networks with arbitrary topologies and traffic distributions. As the first step, we have designed DOMRA, which uses local CSI and two-hop static neighborhood information to improve slotted Aloha. DOMRA adapts to the inhomogeneous spatial traffic distribution and achieves performance comparable with the global optimum, which can only be obtained using complete network knowledge. The generic framework developed in DOMRA proved to be very useful in improving cellular networks as well. We have developed CIA-MAC to deal with the downlink severe cochannel interference that is usually experienced by cell-edge users. CIA-MAC requires only minor changes of existing protocols to obtain significant performance gain. Aloha-based schemes have low channel utilization efficiency because of the collision of entire data frames. To further improve performance, we have developed a scheme with signaling negotiation ahead of data transmission to avoid collision. We noticed that the backoff-after-collision approach in traditional multiple access schemes like CSMA/CA ignored channel variations and deferring transmission without considering the variations might result in data transmission in deep fades. Hence, in our design, each frame is divided into contention and transmission periods and conflicts are optimally resolved in the Aloha-based contention period using the methodology of DOMRA. CAD-MAC completely resolves the contention of networks with arbitrary topologies. Besides, it achieves throughput close to that using centralized schedulers and is robust to any channel uncertainty.

We have also addressed energy-efficient communications in this dissertation. We have

discovered the global optimal energy-efficient OFDM transmission using the strict quasi-concavity of energy efficiency functions. This transmission optimally balances the power consumption of electronic circuits and that of data transmission on each subchannel. The global optimal energy-efficient transmission can be obtained using iterative operations, which may be complex to be implemented in a practical system. Besides, running iterative algorithms consumes additional energy. Hence, using a locally linear approximation, we have developed a closed-form link adaptation scheme, which performs close to the global optimum. Besides, since subchannel allocation in OFDMA systems determines the energy efficiency of all users, we have further developed closed-form resource allocation approaches that also achieve near-optimal performance. Later we observed that in an interference-free environment, a tradeoff between energy efficiency and spectral efficiency exists, as increasing transmit power always improves spectral efficiency but not necessarily energy efficiency. Hence we have continued the investigation in interference-limited scenarios and found that since increased transmit power also brings higher interference to the network, spectral efficiency is not necessarily higher. When interference dominates, energy-efficient power optimization not only improves energy efficiency but also spectral efficiency because of the conservative nature of power allocation, which effectively controls interference to the network. Especially, in interference-dominated regimes, e.g., local area networks, both spectral- and energy-efficient communications desire optimized time-division protocols and the proposed DOMRA, CIA-MAC, and CAD-MAC can be used to improve both spectral and energy efficiency.

APPENDIX A

PROOFS FOR CHAPTER 2

A.1 Proof of Theorem 2.3.1

Since in each transmission time slot, User i sends packets to User j on subchannel k only when this subchannel has the best channel gain among all users in $\mathcal{T}_i$, and the subchannel power gain is above $\overline{H}_{(i,j)_k}$, we get the following:

$$\begin{aligned}
p_{(i,j)_k} &= \Pr\left\{h_{(i,j)_k} = \max_{a\in\mathcal{T}_i}(h_{(i,a)_k}), h_{(i,j)_k} \geq \overline{H}_{(i,j)_k}\right\} \\
&= \Pr\left\{h_{(i,j)_k} = \max_{a\in\mathcal{T}_i}(h_{(i,a)_k})\right\} \cdot \Pr\left\{h_{(i,j)_k} \geq \overline{H}_{(i,j)_k} \,\middle|\, h_{(i,j)_k} = \max_{a\in\mathcal{T}_i}(h_{(i,a)_k})\right\} \\
&= \frac{1}{|\mathcal{T}_i|} \Pr\left\{\max_{a\in\mathcal{T}_i}(h_{(i,a)_k}) \geq \overline{H}_{(i,j)_k}\right\} = \frac{1}{|\mathcal{T}_i|}\left(1 - \prod_{a\in\mathcal{T}_i} \Pr\left(h_{(i,a)_k} < \overline{H}_{(i,j)_k}\right)\right) \\
&= \frac{1}{|\mathcal{T}_i|}\left(1 - F^{|\mathcal{T}_i|}(\overline{H}_{(i,j)_k})\right)
\end{aligned}$$

A.2 Proof of Theorem 2.3.2

According to (2.16), we can see that $\overline{H}^*_{(i,j)_k}$ is independent of j and k. Hence,the first constraint of (2.22) is $\sum_{j\in\mathcal{T}_i,k=1,...,K} P_{r(i,j)_k} \leq P_a|\mathcal{T}_i|(\int_{\overline{H}^*_{(i,j)_k}}^{\infty} \frac{1}{h}\mathrm{d}F^{|\mathcal{T}_i|}(h))^{-1}$. Since data rate function $R()$ is assumed to be a strictly concave function, $\sum_{j\in\mathcal{T}_i,k} R(P_{r(i,j)_k}) \leq |\mathcal{T}_i|KR(\frac{\sum_{j\in\mathcal{T}_i,k} P_{r(i,j)_k}}{K|\mathcal{T}_i|})$. The equation holds if and only if $P_{r(i,j)_k}$ is the same value for all $j \in \mathcal{T}_i$ and k. Hence, for optimal solution, $\max_j \frac{P^*_{r(i,j)_k}}{\overline{H}^*_{(i,j)_k}}$ is the same for all $k = 1, ..., K$, and the second constraint of (2.22) is equivalent to $P^*_{r(i,j)_k} \leq \frac{P_m\overline{H}^*_{(i,j)_k}}{K}$. Then, it is easy to see that when the first constraint in (2.22) takes effect, the optimal solution is the first term in (3.4), while when the second constraint takes effect, the optimal solution is the second term in (3.4). Hence, (3.4) satisfies both constraints, and the objective value will be maximized when one constraint takes effect while satisfying the other constraint.

A.3 Proof of Theorem 2.3.3

According to the symmetry of all subchannels, we see that $\max_{h,j} P_{(i,j)_k}$ is the same for all subchannels. Hence, constraint (2.26c) equals $\max_{h,j} P_{(i,j)_k}(h) \leq \frac{P_m}{K}$, which is the same as $P_{(i,j)_k}(h) \leq \frac{P_m}{K}$. According to (2.16), $\overline{H}^*_{(i,j)_k}$ is independent of j and k. Problem (2.26) is equivalent to:

$$\mathcal{P}_i^* = \arg\min_{\mathcal{P}_i} -\sum_{j\in\mathcal{T}_i,k} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} R\left(\frac{hP_{(i,j)_k}(h)}{n_o W/K}\right) \mathrm{d}F^{\mathcal{T}_i}(h) \tag{A.1a}$$

subject to

$$\sum_{j\in\mathcal{T}_i,k} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} P_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) - |\mathcal{T}_i|P_a \leq 0, \tag{A.1b}$$

and

$$P_{(i,j)_k}(h) - \frac{P_m}{K} \leq 0. \tag{A.1c}$$

Introducing Lagrange multipliers $\lambda_{(i,j)_k}(h)$, $\gamma_{(i,j)_k}(h)$ and $\nu \geq 0$ for the three inequalities respectively, the Lagrange function associated with problem (A.1) is:

$$\begin{aligned}
&L(\mathcal{P}_i, \lambda_{(i,j)_k}, \gamma_{(i,j)_k}, \nu) \\
&= -\sum_{j\in\mathcal{T}_i,k} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} R\left(\frac{\widehat{hP_{(i,j)_k}(h)}}{n_o W/K}\right) \mathrm{d}F^{|\mathcal{T}_i|}(\widehat{h}) + \sum_{j\in\mathcal{T}_i,k} \lambda_{(i,j)_k}(h)(-P_{(i,j)_k}(h)) + \sum_{j\in\mathcal{T}_i,k} \gamma_{(i,j)_k}(h)(P_{(i,j)_k}(h) - \frac{P_m}{K}) + \\
&\nu\left(\sum_{j\in\mathcal{T}_i,k} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} P_{(i,j)_k}(\widehat{h})\mathrm{d}F^{|\mathcal{T}_i|}(\widehat{h}) - |\mathcal{T}_i|P_a\right) \\
&= \sum_{j\in\mathcal{T}_i,k} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} \overline{L}_{(i,j)_k}(\widehat{h})\mathrm{d}F^{|\mathcal{T}_i|}(\widehat{h}) - \sum_{j\in\mathcal{T}_i,k} \gamma_{(i,j)_k}(h)\frac{P_m}{K} - \nu|\mathcal{T}_i|P_a,
\end{aligned}$$

where $\overline{L}_{(i,j)_k}(\widehat{h}) = -R\left(\frac{\widehat{hP_{(i,j)_k}(h)}}{n_o W/K}\right) + \nu P_{(i,j)_k}(\widehat{h}) - \lambda_{(i,j)_k}(h)P_{(i,j)_k}(h)/C + \gamma_{(i,j)_k}(h)(P_{(i,j)_k}(h))/C$, in which $C = 1 - F^{|\mathcal{T}_i|}(\overline{H}^*_{(i,j)_k})$. According to [104], we obtain the following Karush-Kuhn-Tucker (KKT) conditions for optimal power allocation when $h \geq \overline{H}^*_{(i,j)_k}$:

$$\sum_{j\in\mathcal{T}_i,k} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} P^*_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) - |\mathcal{T}_i|P_a \leq 0, \tag{A.2}$$

$$P^*_{(i,j)_k}(h) - \frac{P_m}{K} \leq 0, \tag{A.3}$$

$$\lambda^*_{(i,j)_k}(h) \geq 0, \gamma^*_{(i,j)_k}(h) \geq 0, \text{ and } \nu^* \geq 0, \tag{A.4}$$

$$\lambda^*_{(i,j)_k}(h) P^*_{(i,j)_k}(h) = 0, \tag{A.5}$$

$$\gamma^*_{(i,j)_k}(h) \left(P^*_{(i,j)_k}(h) - \frac{P_m}{K} \right) = 0, \tag{A.6}$$

$$\nu^* \left(\sum_{j \in \mathcal{T}_i, k} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} P^*_{(i,j)_k}(h) \mathrm{d}F^{|\mathcal{T}_i|}(h) - |\mathcal{T}_i| P_a \right) = 0, \tag{A.7}$$

and

$$\left. \frac{\partial \overline{L}_{(i,j)_k}(h)}{\partial P_{(i,j)_k}(h)} \right|_{P^*_{(i,j)_k}(h)} = -R^{'}\left(\frac{hP_{(i,j)_k}(h)}{n_o W/K}\right)\frac{hK}{n_o W} + \nu^* - \lambda^*_{(i,j)_k}(h)/C + \gamma^*_{(i,j)_k}(h)/C = 0. \tag{A.8}$$

1°. When

$$\sum_{j \in \mathcal{T}_i, k} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} P^*_{(i,j)_k}(h) \mathrm{d}F^{|\mathcal{T}_i|}(h) < |\mathcal{T}_i| P_a, \tag{A.9}$$

according to (A.7), $\nu^* = 0$. From (A.8), $\gamma^*_{(i,j)_k}(h) > 0$. Hence, $P^*_{(i,j)_k}(h) = \frac{P_m}{K}$ from (A.6). (A.9) equals $P_m < \frac{P_a}{1-F^{|\mathcal{T}_i|}(\overline{H}^*_{(i,j)_k})}$.

2°. When $\sum_{j \in \mathcal{T}_i, k} \int_{\overline{H}^*_{(i,j)_k}}^{\infty} P^*_{(i,j)_k}(h) \mathrm{d}F^{|\mathcal{T}_i|}(h) = |\mathcal{T}_i| P_a$, i.e. $P_m \geq \frac{P_a}{1-F^{|\mathcal{T}_i|}(\overline{H}^*_{(i,j)_k})}$,

a). if $\gamma^*_{(i,j)_k}(h) > 0$, $P^*_{(i,j)_k}(h) = \frac{P_m}{K}$ from (A.6), and $\lambda^*_{(i,j)_k}(h) = 0$ from (A.5). According to (A.8), $\nu^* < R^{'}\left(\frac{hP_m}{n_o W}\right)\frac{hK}{n_o W}$

b). if $\gamma^*_{(i,j)_k}(h) = 0$, from (A.8), $\nu^* = R^{'}\left(\frac{hP^*_{(i,j)_k}(h)}{n_o W/K}\right)\frac{hK}{n_o W} + \lambda^*_{(i,j)_k}(h)/C \geq R^{'}\left(\frac{hP^*_{(i,j)_k}(h)}{n_o W/K}\right)\frac{hK}{n_o W}$. According to the strictly decreasing assumption of the first order derivative of $R(\eta)$, $R^{'}\left(\frac{hP^*_{(i,j)_k}(h)}{n_o W/K}\right) \leq R^{'}(0)$, i.e. $\nu^* \leq R^{'}(0)\frac{hK}{n_o W} + \lambda^*_{(i,j)_k}(h)/C$, and equality holds only when $P^*_{(i,j)_k}(h) = 0$. Hence, with (A.4) and (A.5), if $\nu^* \geq R^{'}(0)\frac{hK}{n_o W}$, it is easy to see that $P^*_{(i,j)_k}(h) = 0$. If $\nu^* < R^{'}(0)\frac{hK}{n_o W}$, $P^*_{(i,j)_k}(h) > 0$, $\lambda^*_{(i,j)_k}(h)=0$, and $R^{'}\left(\frac{hP^*_{(i,j)_k}(h)}{n_o W/K}\right)\frac{hK}{n_o W} = \nu^*$. Then $P^*_{(i,j)_k}(h) = R^{'-1}\left(\frac{\nu^* n_o W}{hK}\right)\frac{n_o W}{Kh}$. Hence, when $h \geq \overline{H}^*_{(i,j)_k}$, we have

$$P^*_{(i,j)_k}(h) = \begin{cases} \frac{P_m}{K} & \nu^* < R^{'}\left(\frac{hP_m}{n_o W}\right)\frac{hK}{n_o W}, \\ 0 & \nu^* \geq R^{'}(0)\frac{hK}{n_o W}, \\ R^{'-1}\left(\frac{\nu^* n_o W}{hK}\right)\frac{n_o W}{Kh} & \text{otherwise}, \end{cases} \tag{A.10}$$

Observing that both $\overline{H}^*_{(i,j)_k}$ and $P^*_{(i,j)_k}(h)$ are independent of j and k, substituting (A.10) into

the condition $\sum_{j\in\mathcal{T}_i,k}\int_{\overline{H}^*_{(i,j)_k}}^{\infty} P^*_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) = |\mathcal{T}_i|P_a$, we get

$$\int_{\overline{H}^*_{(i,j)_k}}^{\infty} P^*_{(i,j)_k}(h)\mathrm{d}F^{|\mathcal{T}_i|}(h) = \frac{P_a}{K}. \tag{A.11}$$

Since $P^*_{(i,j)_k}(h)$ is a piecewise-linear decreasing function of ν^* with breakpoints at $R^{'}(0)\frac{hK}{n_oW}$ and $R^{'}\left(\frac{hP_m}{n_oW}\right)\frac{hK}{n_oW}$, (A.11) has a unique solution of ν^*.

Theorem 2.3.3 is readily obtained from both 1° and 2°. The solution is globally optimal since for convex optimizations, KKT conditions are both necessary and sufficient for a local minimum to be a global minimum. If, in addition, the objective function is strictly convex, the globally optimal solution is unique. It is easy to see that in (A.1), given that the first order derivative $R^{'}(\eta)$ is positive and strictly decreasing, the constraints are convex, and the objective function is strictly convex of $P_{(i,j)_k}(h)$, $j \in \mathcal{T}_i, k = 1, \cdots, K$, and the unique global optimality follows.

APPENDIX B

PROOFS FOR CHAPTER 4

B.1 Proof of Theorem 4.4.2

Proof We prove that the two conditions of the definition hold for CAD-MAC.

1) Suppose two links, (i, j) and (k, l), that have won the contention have collision and the transmission of User i interferes with the reception of User l. First, (i, j) and (k, l) should not have won the contention at the same CRS since the REQUESTs of the two links collide at User l and User l will not acknowledge SUCCESS. If (i, j) receives SUCCESS first, the OCCUPIED signal of User i will prevent User l from acknowledging SUCCESS. If (k, l) wins first, the broadcasting of SUCCESS by User l will prevent User i from sending REQUEST. Hence, Condition 1 always holds.

2) To verify Condition 2, suppose that there exists a link (I, J) that has not won access and does not collide with any link that has won. Besides, within the interference range of Link (I, J), no other link could win as otherwise, after that link wins, Link (I, J) should not contend and the contention is completely resolved. There are two possibilities. (1) User I does not send any REQUEST all the time or (2) whenever User I sends a REQUEST, it collides with that of other links. We show in the following that both have zero possibility.

(1) User I does not send any REQUEST all the time. This indicates that $h_{IJ} < \widehat{H}_{IJ}[k]$ for all $k > K$, where $K > 0$. Obviously nobody that interferes with User J should send anything. Hence, User J will keep on sending IDLE signals to User I and $\widehat{H}_{IJ}[k]$ will be lowered successively. It is easy to see that in this case $\lim_{k\to\infty} \widehat{H}_{IJ}[k] = \widehat{H}^m_{IJ}$. Hence, the probability that $h_{IJ} < \widehat{H}_{IJ}[k]$ for all $k > K$ is zero and sooner or later User I will send a REQUEST and win.

(2) Whenever (I, J) sends a REQUEST, it collides with others. Denote the CRSs that (I, J) sends REQUESTs by $C = \{c_1, c_2, \cdots\}$, where $c_1 < c_2 < \cdots$. Suppose there are N links that collide with (I, J). According to (4.16), the contention probability of any link

using ADs I or II is $\frac{1}{2}$ after sending the first REQUEST. If using AD III, the contention probability is zero. We consider the CRSs after all the interfering links have sent the first REQUEST and denote $N_k \leq N$ to be the number of interfering links that contends with probability $\frac{1}{2}$ in CRS k. The probability that (I, J) keeps on contending and never succeeds is given by

$$\begin{aligned}
&\Pr\{(I,J) \text{ never wins}\} \\
&= \lim_{|C|\to\infty} \prod_{k\in C} \Pr\{\text{at least one interferer contends in CRS } k\} \\
&= \lim_{|C|\to\infty} \prod_{k\in C} \left(1 - (1-\frac{1}{2})^{N_k}\right) \qquad \text{(B.1)} \\
&\leq \lim_{|C|\to\infty} \prod_{k\in C} (1 - (\frac{1}{2})^{N}) = \lim_{|C|\to\infty} (1 - (\frac{1}{2})^{N})^{|C|} < \sigma \qquad \text{(B.2)}
\end{aligned}$$

for any $\sigma > 0$. Hence, the probability that (I, J) never resolves its contention is zero. That is, with probability one, (I, J) always wins the contention when none of its neighbors can win and the network contention within the interference range of (I, J) can always be resolved.

Theorem 4.4.2 follows immediately.

B.2 Proof of Theorem 4.4.3

Proof Suppose there are N links and in CRS 1, each has the contention probability $p_{i,j}[1] = \frac{1}{N}$. According to (4.16), the contention probability in CRS k is

$$p_{ij}[k] = \begin{cases} \frac{1}{N}, \text{IDLE in all the previous CRSs,} \\ \frac{1}{2}, \text{otherwise.} \end{cases} \qquad \text{(B.3)}$$

Denote by $\bar{k}_n$ the average number of CRSs necessary to resolve the collision involving n links. From (B.3), these links will contend with probability $\frac{1}{2}$ if they still contend in the following CRSs. Hence,

$$\bar{k}_n = \left(\frac{1}{2}\right)^n \left[\sum_{i=2}^{n} \binom{n}{i} (\bar{k}_i + 1) + \binom{n}{0} (\bar{k}_n + 1) + \binom{n}{1} 1\right], \qquad \text{(B.4)}$$

where $\left(\frac{1}{2}\right)^n \binom{n}{i}$ is the probability that i users have their gains above the thresholds and on average, $\overline{k}_i$ additional CRSs are needed if $i > 1$. If $i = 0$, all users have their gains below the thresholds and are involved in the following contention. If $i = 1$, the contention is resolved. It has been proved in [26] that $\overline{k}_n$ in (B.4) satisfies

$$\log_2(n) \leq \overline{k}_n \leq \log_2(n) + 1 \tag{B.5}$$

for all n. Before a collision happens, all users may have channel gains so low that several CRSs are necessary for them to lower their thresholds successively until some users are allowed to send REQUESTs. Hence, the average number of CRSs necessary for completely resolving the network contention is

$$\overline{K}_N = \sum_{i=0}^{\infty} \left((1-\frac{1}{N})^{Ni} \left(\sum_{n=1}^{N} \binom{N}{n} (\frac{1}{N})^n (1-\frac{1}{N})^{N-n} (\overline{k}_n + i + 1) \right) \right), \tag{B.6}$$

where $(1-\frac{1}{N})^{Ni}$ is the probability that all users have their gains below their thresholds in all the first i CRSs and $\binom{N}{n}(\frac{1}{N})^n(1-\frac{1}{N})^{N-n}$ is the probability that in the $i+1$st CRS, n users send REQUESTs and collide. Let $M_N = \sum_{n=1}^{N} \binom{N}{n}(\frac{1}{N})^n(1-\frac{1}{N})^{N-n}(\overline{k}_n + 1)$. Then,

$$\begin{aligned} M_N &\leq \sum_{n=1}^{N} \binom{N}{n} (\frac{1}{N})^n (1-\frac{1}{N})^{N-n} (\log_2(n) + 1) \\ &< \sum_{n=1}^{N} \binom{N}{n} (\frac{1}{N})^n (1-\frac{1}{N})^{N-n} (n+1) \\ &= 2 - (1-\frac{1}{N})^N. \end{aligned} \tag{B.7}$$

Hence, $\overline{K}_N$ equals

$$\begin{aligned} \overline{K}_N &= M_N \sum_{i=0}^{\infty} (1-\frac{1}{N})^{Ni} + \sum_{i=0}^{\infty} i(1-\frac{1}{N})^{Ni} \\ &= \frac{M_N}{1-(1-\frac{1}{N})^N} + \frac{(1-\frac{1}{N})^N}{(1-(1-\frac{1}{N})^N)^2} \\ &< \frac{2-(1-\frac{1}{N})^N}{1-(1-\frac{1}{N})^N} + \frac{(1-\frac{1}{N})^N}{(1-(1-\frac{1}{N})^N)^2} \\ &= 1 + \frac{1}{[1-(1-\frac{1}{N})^N]^2} \end{aligned}$$

$$< 1 + [1 - e^{-1}]^{-2}. \tag{B.8}$$

Hence, $\overline{K}_N$ is bounded for all N and the right hand side of (B.6) converges. A tighter bound is

$$\overline{K}_N \leq \frac{\widehat{M}_N}{1-(1-\frac{1}{N})^N} + \frac{(1-\frac{1}{N})^N}{(1-(1-\frac{1}{N})^N)^2},$$

where $\widehat{M}_N = \sum_{n=1}^{N} \binom{N}{n} (\frac{1}{N})^n (1-\frac{1}{N})^{N-n} (\log_2(n)+1)$. As N goes to infinity, using computer calculation, we have

$$\overline{K}_N < \overline{K}_\infty \leq 2.43. \tag{B.9}$$

B.3 Proof of Theorem 4.4.4

Proof Let K be the number of CRSs necessary to completely resolve network contention in a frame slot and $\mathcal{K} = \{1, 2, \cdots, K\}$ the corresponding set of CRSs. Let L be the number of links winning the contention and $\mathcal{K}_l, i = 1 \cdots, L$, the corresponding set of CRSs that the lth winning link is involved in the contention. Assume that $\mathcal{K}_l, l = 1, \cdots, L$ are independently and identically distributed and independent of L. Obviously,

$$\mathcal{K} = \bigcup_i \mathcal{K}_i \text{ and } K = |\mathcal{K}| \leq \sum_i |\mathcal{K}_i|, \tag{B.10}$$

where $|\mathcal{X}|$ is the cardinality of set $\mathcal{X}$. Define the contention coexistence factor as

$$\beta = \frac{\mathbf{E}(\sum_i |\mathcal{K}_i|)}{\mathbf{E}(|\mathcal{K}|)}. \tag{B.11}$$

It is easy to see that β is the average number of simultaneous resolutions in each CRS. For example, if all users interfere with all others, then $L = 1$ and $\beta = 1$, meaning only one resolution in each CRS. If a network consists of L groups of users and the communication of any group does not interfere with that of any other group, then these L groups can resolve the contention within each group to produce L winners. If we further assume $\mathcal{K}_1, \mathcal{K}_2, \cdots, \mathcal{K}_L$ are independently and identically distributed, then $\beta = L$, indicating L simultaneous resolutions in each CRS on average. Then we have

$$\overline{K} = \mathbf{E}(K) = \mathbf{E}(|\mathcal{K}|) = \frac{\mathbf{E}(\sum_i^L |\mathcal{K}_i|)}{\beta}. \tag{B.12}$$

Furthermore, L is a stopping time for K_i and according to Wald's equation [113], we have

$$\overline{K} = \frac{\mathbf{E}(|\mathcal{K}_i|)\mathbf{E}(L)}{\beta}. \tag{B.13}$$

Obviously from Theorem 4.4.3, $\mathbf{E}(|\mathcal{K}_i|) < \overline{K}_\infty$. Hence,

$$\overline{K} < \frac{2.425 \cdot \mathbf{E}(L)}{\beta}. \tag{B.14}$$

APPENDIX C

PROOFS FOR CHAPTER 5

C.1 Proof of Lemma 5.2.2

Proof Denote the upper contour sets of $U(\mathbf{R})$ as

$$S_\alpha = \{\mathbf{R} \geq \mathbf{0} | U(\mathbf{R}) \geq \alpha\}, \tag{C.1}$$

where symbol $\geq$ denotes vector inequality and $\mathbf{R} \geq \mathbf{0}$ means each element of $\mathbf{R}$ is non-negative. According to Proposition C.9 of [102], $U(\mathbf{R})$ is strictly quasiconcave if and only if S_α is strictly convex for any real number α. When $\alpha < 0$, no points exist on the contour $U(\mathbf{R}) = \alpha$. When $\alpha = 0$, only $\mathbf{0}$ is on the contour $U(\mathbf{0}) = \alpha$. Hence, S_α is strictly convex when $\alpha \leq 0$. Now we investigate the case when $\alpha > 0$. S_α is equivalent to $S_\alpha = \{\mathbf{R} \geq \mathbf{0} | \alpha P_C + \alpha P_T(\mathbf{R}) - R \leq 0\}$. Since $P_T(\mathbf{R})$ is strictly convex in $\mathbf{R}$, S_α is also strictly convex. Hence, we have the strict quasiconcavity of $U(\mathbf{R})$.

The partial derivative of $U(\mathbf{R})$ with r_i is

$$\frac{\partial U(\mathbf{R})}{\partial r_i} = \frac{P_C + P_T(\mathbf{R}) - RP_T^{'}(\mathbf{R})}{(P_C + P_T(\mathbf{R}))^2} \triangleq \frac{\beta(r_i)}{(P_C + P_T(\mathbf{R}))^2}, \tag{C.2}$$

where $P_T^{'}(\mathbf{R})$ is the first partial derivative of $P_T(\mathbf{R})$ with respect to r_i. According to Lemma 5.2.2, if r_i^* exists such that $\frac{\partial U(\mathbf{R})}{\partial r_i}\Big|_{r_i = r_i^*} = 0$, it is unique, i.e. if there is a r_i^* such that $\beta(r_i^*) = 0$, it is unique. In the following, we investigate the conditions when r_i^* exists.

The derivative of $\beta(r_i)$ is

$$\beta^{'}(r_i) = -RP_T^{''}(\mathbf{R}) < 0, \tag{C.3}$$

where $P_T^{''}(\mathbf{R})$ is the second partial derivative of $P_T(\mathbf{R})$ with respect to r_i. Hence, $\beta(r_i)$ is strictly decreasing. According to the L'Hopital's rule, it is easy to show that

$$\begin{aligned} \lim_{r_i->\infty} \beta(r_i) &= \lim_{r_i->\infty} (P_C + P_T(\mathbf{R}) - RP_T^{'}(\mathbf{R})) = \lim_{r_i->\infty} \left(\frac{P_C + P_T(\mathbf{R}) - RP_T^{'}(\mathbf{R})}{r_i} r_i \right) \\ &= \lim_{r_i->\infty} \left(\frac{P_T^{'}(\mathbf{R}) - P_T^{'}(\mathbf{R}) - RP_T^{''}(\mathbf{R})}{1} r_i \right) = \lim_{r_i->\infty} -P_T^{''}(\mathbf{R}) R r_i < 0. \end{aligned} \tag{C.4}$$

Besides,

$$\begin{aligned}\lim_{r_i->0}\beta(r_i) &= \lim_{r_i->0}(P_C + P_T(\mathbf{R}) - RP_T^{'}(\mathbf{R})) \\ &= P_C + P_T(\mathbf{R}_i^{(0)}) - R_i^{(0)}P_T^{'}(\mathbf{R}_i^{(0)}),\end{aligned} \tag{C.5}$$

where $\mathbf{R}_i^{(0)} = [r_1, r_2, \cdots, r_{i-1}, 0, r_{i+1}, \cdots, r_K]^T$ and $R_i^{(0)} = \sum_{j\neq i} r_j$.

(1^o)When $P_C + P_T(\mathbf{R}_i^{(0)}) - R_i^{(0)}P_T^{'}(\mathbf{R}_i^{(0)}) \geq 0$, $\lim_{r_i->0}\beta(r_i) \geq 0$. Together with (C.4), we see that t_i^* exists and $U(\mathbf{R})$ is first strictly increasing and then strictly decreasing in r_i.

(2^o)When $P_C + P_T(\mathbf{R}_i^{(0)}) - R_i^{(0)}P_T^{'}(\mathbf{R}_i^{(0)}) < 0$, $\lim_{r_i->0}\beta(r_i) < 0$. Together with (C.3) and (C.4), t_i^* does not exist. However, $U(\mathbf{R})$ is always strictly decreasing in r_i. Hence, $U(\mathbf{R})$ is maximized at $r_i = 0$.

Lemma 5.2.2 is readily obtained.

C.2 Proof of the Upperbound in Theorem 5.2.4

Proof $U(R) = \frac{R}{P_C+P_T(R)} \leq \frac{R}{P_T(R)}$. Denote $\widehat{U}(R) = \frac{R}{P_T(R)}$. $\widehat{U}^{'}(R) = \frac{d\widehat{U}(R)}{dR} = \frac{P_T(R)-RP_T^{'}(R)}{P_T^2(R)}$. According to the L'Hopital's rule, $\lim_{R\to 0}\widehat{U}^{'}(R) = \lim_{R\to 0}\frac{P_T(R)-RP_T^{'}(R)}{P_T^2(R)} = \lim_{R\to 0}\frac{-RP_T^{''}(R)}{2P_T(R)P_T^{'}(R)} = \lim_{R\to 0}\frac{-P_T^{''}(R)}{2(P_T^{'}(R))^2} \leq 0$. Besides $P_T(R) - RP_T^{'}(R)$ is 0 when $R = 0$ and has negative derivative when $R > 0$. Hence, $P_T(R) - RP_T^{'}(R) < 0$ when $R > 0$. Thus, $\widehat{U}^{'}(R)$ is negative when $R > 0$ and $\widehat{U}(R)$ is maximized when R approaches zero, i.e. $U(R) \leq \lim_{R\to 0}\frac{R}{P_T(R)} = \frac{1}{P_T^{'}(0)}$.

C.3 Proof of Propositions 5.2.5, 5.2.6, and 5.2.7

Proof Denote $P_R(r)$ to be the received power on a subchannel for reliable detection when the data rate on the subchannel is r. We have $P_T(R) = \frac{KP_R(r)}{g} = \frac{KP_R(\frac{R}{K})}{g}$, where g is the channel power gain. It is easy to see that $P_R(r)$ is monotonically increasing and strictly convex, and $P_T(0) = P_R(0) = 0$. According to Theorem 5.2.4, we have $R^*P_T^{'}(R^*) = P_C + P_T(R^*)$, which is equivalent to $R^*P_R^{'}(\frac{R^*}{K}) - KP_R(\frac{R^*}{K}) = P_C g$. By differentiating the left hand side with respect to R^*, $\frac{\partial\left(R^*P_R^{'}(\frac{R^*}{K})-cP_R(\frac{R^*}{K})\right)}{\partial R^*} = \frac{R^*}{K}P_R^{''}(\frac{R^*}{K}) > 0$. Hence, the left hand side is strictly increasing in R^*. Therefore, higher data rate should be used when the channel has higher power gain. Suppose $g_1 > g_2$, and the corresponding optimal modulation and

coding result in data rates R_1^* and R_2^* respectively. Hence, $U_1(R_1^*) > U_1(R_2^*)$. Besides, $U_1(R_2^*) = \frac{R_2^*}{P_C + \frac{KP_R(R_2^*/K)}{g_1}} > \frac{R_2^*}{P_C + \frac{KP_R(R_2^*/K)}{g_2}} = U_2(R_2^*)$. Hence, the energy efficiency increase with channel gain.

According to Theorem 5.2.4, $R^* P_T^{'}(R^*) - P_T(R^*) = P_C$. The derivative of the left hand side is $R^* P_T^{''}(R^*) > 0$. Hence, R^* increases with P_C. The proof that the energy efficiency decreases with circuit power is similar to the proof that energy efficiency increases with channel gain. When $P_C = 0$, according to proof in C.2, $U(R)$ is maximized when R approaches zero, i.e. $U_{max} = \lim_{R\to 0} \frac{R}{P_T(R)} = \frac{1}{P_T^{'}(0)}$.

$R = Kr$ and $P_T(R) = K\overline{P}_T(\frac{R}{K})$, where $\overline{P}_T(r)$ is the transmit power on each subchannel, and is monotonically increasing and strictly convex in r. According to Theorem 5.2.4, we have $R^* \overline{P}_T^{'}(\frac{R^*}{K}) = P_C + K\overline{P}_T(\frac{R^*}{K})$, which is equivalent to $r^* \overline{P}_T^{'}(r^*) - P_T(r^*) = \frac{P_C}{K}$. The left hand side is increasing in r^* while the right hand side is decreasing in K. Hence, the data rate on each subchannel should decrease with increasing number of subchannels assigned. The proof that the energy efficiency increases with the number of subchannels assigned is also similar to the proof in C.3 and is omitted. The highest energy efficiency is obtained with infinite number of subchannels, i.e. $U(R) = \lim_{K\to\infty} \frac{R}{P_C + P_T(R)} = \frac{r}{\overline{P}_T(r)}$. Similar to the proof in C.2, $U(R)$ is maximized when r approaches zero. We have $U_{max} = \lim_{r\to 0} \frac{r}{\overline{P}_T(r)} = \frac{1}{P_T^{'}(0)}$.

C.4 Proof of Theorem 5.4.2

Proof The global convergence is straightforward from Lemmas 5.2.2. Since $r_2^{[0]} = \alpha r_1^{[0]}$ and $r_1^{[i]} \le r^* \le r_2^{[i]}$, with induction, we have $r_2^{[i]} - r_1^{[i]} = \frac{r_2^{[0]} - r_1^{[0]}}{2^i} \le \frac{(\alpha-1)r^*}{2^i}$. Hence, $\widehat{r}^{[i]} = \frac{r_1^{[i]} + r_2^{[i]}}{2} \ge (2r_2^{[i]} - \frac{(\alpha-1)r^*}{2^i})/2 \ge r^* - \frac{(\alpha-1)r^*}{2^{i+1}}$ and $\widehat{r}^{[i]} \le r^* + \frac{(\alpha-1)r^*}{2^{i+1}}$. Then $|\widehat{r}^{[i]} - r^*| \le \frac{(\alpha-1)r^*}{2^{i+1}}$. Let $\frac{(\alpha-1)r^*}{2^{i+1}} \le \epsilon$. We have $i \ge \log_2(\frac{(\alpha-1)r^*}{\epsilon} - 1)$. Theorem 5.4.2 follows immediately.

APPENDIX D

PROOFS FOR CHAPTER 7

D.1 Proof for Transmit-Power-Dominated Regime

Proof It is obvious that the second term in $u(p_1, p_2)$ is strictly decreasing in p_1. To determine the first term, we need to verify that $F(p) = \frac{\log(1+ap)}{p}, \forall a > 0$, is strictly decreasing in p, i.e.

$$\frac{\partial F(p)}{\partial p} = \frac{ap - \log(1+ap) - ap\log(1+ap)}{p^2(1+ap)} < 0 \tag{D.1}$$

which is equivalent to $G(p) = ap - \log(1+ap) - ap\log(1+ap) < 0$. $G(0) = 0$. Besides $\frac{\partial G(p)}{\partial p} = -a\log(1+ap) < 0$. Hence $G(p) < 0, \forall p > 0$. Thus $F(p)$ is strictly decreasing in p.

D.2 Proof for Interference-Dominated Regime

Proof Since we are considering interference dominated regime, whenever Users 1 and 2 are sending data, $p_1 g_{12} \gg \sigma^2$ and $p_2 g_{21} \gg \sigma^2$ and $INR > 1 + SINR$ due to close coupling between these transmissions. In wireless communications, radio links exhibit a threshold effect where link quality is acceptable where signal-to-noise ratio must exceed certain thresholds [114]. This indicates that the power allocation should not be too small. We assume that feasible p_1 and p_2 satisfies $p_1 \geq \widehat{p_1}$ and $p_2 \geq \widehat{p_2}$; otherwise, the user is shut down. Besides, in the interference dominated regime, $\widehat{p_1} g_{12} \gg \sigma^2$ and $\widehat{p_2} g_{21} \gg \sigma^2$. We compare two schemes. The first is to let both users send data simultaneously and the other is to shut down one user. First, we will show that when both users transmit, for User 1,

$$\frac{2w\log(1+\frac{p_1 g_1}{p_2 g_{21}+\sigma^2})}{p_1 + p_c} < \frac{w\log(1+\frac{p_1 g_1}{\sigma^2})}{p_1 + p_c}, \tag{D.2}$$

which is equivalent to show that $(1+\frac{p_1 g_1}{p_2 g_{21}+\sigma^2})^2 < 1+\frac{p_1 g_1}{\sigma^2}$. This inequality equals to

$$1+\frac{p_1 g_1}{p_2 g_{21} + \sigma^2} < \frac{p_2 g_{21}}{\sigma^2}. \tag{D.3}$$

Since $INR > SINR + 1$, (D.3) holds and so is (D.2). Similarly we have the same result for User 2. When both users are sending, the maximum energy efficiency is

$$\max_{\substack{p_1 \geq \widehat{p}_1 \gg \frac{\sigma^2}{g_{12}} \\ p_2 \geq \widehat{p}_2 \gg \frac{\sigma^2}{g_{21}}}} \left(\frac{w \log(1 + \frac{p_1 g_1}{p_2 g_{21} + \sigma^2})}{p_1 + p_c} + \frac{w \log(1 + \frac{p_2 g_2}{p_1 g_{12} + \sigma^2})}{p_2 + p_c} \right). \tag{D.4}$$

Assume the above maximum energy efficiency is obtained by p_1° and p_2°. According to (D.2),

$$\frac{w \log(1 + \frac{p_1^\circ g_1}{p_2^\circ g_{21} + \sigma^2})}{p_1^\circ + p_c} < \frac{w \log(1 + \frac{p_1^\circ g_1}{\sigma^2})}{2(p_1^\circ + p_c)} \leq \frac{1}{2} \max_{p_1} \left(\frac{w \log(1 + \frac{p_1 g_1}{\sigma^2})}{p_1 + p_c} \right) \tag{D.5}$$

Similarly, $\frac{w \log(1 + \frac{p_2^\circ g_2}{p_1^\circ g_{12} + \sigma^2})}{p_2^\circ + p_c} < \frac{1}{2} \max_{p_2} \left(\frac{w \log(1 + \frac{p_2 g_2}{\sigma^2})}{p_2 + p_c} \right)$. Suppose $g_1 \geq g_2$. It is easy to see that

$$\max_{p_1} \left(\frac{w \log(1 + \frac{p_1 g_1}{\sigma^2})}{p_1 + p_c} \right) \geq \max_{p_2} \left(\frac{w \log(1 + \frac{p_2 g_2}{\sigma^2})}{p_2 + p_c} \right). \tag{D.6}$$

Comparing the above inequalities, we can see that

$$\max_{p_1} \left(\frac{w \log(1 + \frac{p_1 g_1}{\sigma^2})}{p_1 + p_c} \right) > \frac{w \log(1 + \frac{p_2^\circ g_2}{p_1^\circ g_{12} + \sigma^2})}{p_2^\circ + p_c} + \frac{w \log(1 + \frac{p_2^\circ g_2}{p_1^\circ g_{12} + \sigma^2})}{p_2^\circ + p_c}. \tag{D.7}$$

The conclusion follows immediately. Extension to multi-user case is straightforward.

D.3 Proof of Lemma 7.3.3

Proof Denote the upper contour sets of $u_n(\mathbf{p}_n, \mathbf{p}_{-n})$ as $S_\alpha = \{\mathbf{p}_n \succeq \mathbf{0} | u_n(\mathbf{p}_n, \mathbf{p}_{-n}) \geq \alpha\}$, where symbol $\succeq$ denotes vector inequality and $\mathbf{R} \succeq \mathbf{0}$ means each element of $\mathbf{R}$ is nonnegative. According to Proposition C.9 of [102], $u_n(\mathbf{p}_n, \mathbf{p}_{-n})$ is strictly quasiconcave in $\mathbf{p}_n$ if and only if S_α is strictly convex for any real number α. It is obvious that when $\alpha \leq 0$, S_α is strictly convex when $\alpha \leq 0$. Now we investigate the case when $\alpha > 0$. Since $u_n(\mathbf{p}_n, \mathbf{p}_{-n}) = \frac{\sum_{k=1}^K R(\frac{p_n^{(k)} g_{nn}^{(k)}}{\sum_{i=1, i \neq n}^N p_i^{(k)} g_{in}^{(k)} + \sigma^2})}{p_c + \sum_{k=1}^K p_n^{(k)}} \geq \alpha$, S_α is equivalent to $S_\alpha = \{\mathbf{p}_n \succeq \mathbf{0} \Big| \sum_{k=1}^K R(\frac{p_n^{(k)} g_{nn}^{(k)}}{\sum_{i=1, i \neq n}^N p_i^{(k)} g_{in}^{(k)} + \sigma^2}) - (p_c + \sum_{k=1}^K p_n^{(k)})\alpha \geq 0\}$. Since $R()$ is strictly concave, S_α is also strictly convex. Hence, we have Lemma 5.2.2.

D.4 Proof of Theorem 7.3.4

Proof In [111], it has been shown a Nash equilibrium exists in a noncooperative game if for any n, (1) $\mathbf{p}_n$ is a nonempty, convex, and compact subset of some Euclidean space $\Re^L$ and (2) $u_n(\mathbf{p}_n, \mathbf{p}_{-n})$ is continuous and quasi-concave in $\mathbf{p}_n$, both of which are satisfied in our noncooperative energy-efficient control game. Hence, the existence of the equilibrium immediately follows. According to our previous work in [107], in a point-to-point energy-efficient transmission, the necessary and sufficient condition for a data rate vector of User n, $\mathbf{r}_n^o = [r_n^{(1)o}, r_n^{(2)o}, \cdots, r_n^{(K)o}]^T$, to be globally optimal is given by, for any Subchannel i,

1. if $\frac{p_c+\sum_{j\neq i} p_n^{(j)}}{\sum_{j\neq i} r_n^{(j)}} \geq \left.\frac{\partial(\sum_j p_n^{(j)})}{\partial r_n^{(i)}}\right|_{\mathbf{r}_n=\mathbf{r}_n^{(i0)}}$, $\left.\frac{\partial u_n(\mathbf{p}_n,\mathbf{p}_{-n})}{\partial r_n^{(i)}}\right|_{\mathbf{r}_n=\mathbf{r}_n^o} = 0$, i.e. $\left.\frac{\partial(\sum_j p_n^{(j)})}{\partial r_n^{(i)}}\right|_{\mathbf{r}_n=\mathbf{r}_n^o} = \frac{1}{u(\mathbf{p}_n^o,\mathbf{p}_{-n})}$;

2. otherwise, $r_n^{(i)o} = 0$,

where $\mathbf{r}_n^{(i0)} = [r_n^{(1)o}, r_n^{(2)o}, \cdots, r_n^{(i-1)o}, 0, r_n^{(i+1)o}, \cdots, r_n^{(K)o}]$.

By transformation of parameters, $\frac{\partial f}{\partial r_n^{(i)}} = \frac{\partial f}{\partial p_n^{(i)}} \left/ \frac{\partial r_n^{(i)}}{\partial p_n^{(i)}} \right. = \frac{\partial f}{\partial p_n^{(i)}} \frac{1}{R'(\eta_n^{(i)})\gamma_n^{(i)}}$, where $R'()$ is the first order derivative of $R()$ and $\gamma_n^{(i)} = \frac{\eta_n^{(i)}}{p_n^{(i)}} = \frac{g_{nn}^{(i)}}{\sum_{j=1,j\neq n}^N p_j^{(i)} g_{jn}^{(i)} + \sigma^2}$. Hence, we have the following equivalent condition for each user. For any Subchannel i,

1. if $\frac{\sum_{j\neq i} r_n^{(j)}}{p_c+\sum_{j\neq i} p_n^{(j)}} \leq R'(0)\gamma_n^{(i)}$, $\left.\frac{\partial u_n(\mathbf{p}_n,\mathbf{p}_{-n})}{\partial p_n^{(i)}}\right|_{\mathbf{p}_n=\mathbf{p}_n^o} = 0$, i.e.

$$R'(\gamma_n^{(i)} p_n^{(i)o})\gamma_n^{(i)} = u(\mathbf{p}_n^o, \mathbf{p}_{-n}); \tag{D.8}$$

2. otherwise, $p_n^{(i)o} = 0$.

It is easy to see that the network achieves an equilibrium if and only if the power settings of all users satisfy the above conditions. Theorem 7.3.4 is readily obtained.

D.5 Proof of Proposition 7.3.5

Proof $p_n^o = f_n(\mathbf{p}_{-n}) = \arg\max_{p_n} u_n(p_n, \mathbf{p}_{-n})$. Since $u_n(0, \mathbf{p}_{-n}) = 0$ and $u_n(p_n, \mathbf{p}_{-n}) > 0$ for any $p_n > 0$, $f_n(\mathbf{p}_{-n}) > 0$ and we have the positivity. Denote $I_n = \sum_{j=1,j\neq n}^N p_j g_{jn}$ **and** $\gamma_n = \frac{g_{nn}}{I+\sigma^2}$. According to (D.8), p_n^o satisfies

$$R'(\gamma_n p_n^o)\gamma_n = u(p_n^o, \mathbf{p}_{-n}) = \frac{R(\gamma_n p_n^o)}{p_c + p_n^o}. \tag{D.9}$$

Substituting $R(\eta) = w\log(1+\eta)$ in to (D.9), we have the following equivalent condition,

$$w(p_n^o, I) = g_{nn}(p_c + p_n^o) - (p_n^o g_{nn} + I + \sigma^2)\log(1 + \frac{p_n^o g_{nn}}{I+\sigma^2}) = 0. \tag{D.10}$$

Hence, $\frac{\partial p_n^o}{\partial I} = -\frac{\partial w}{\partial I}\Big/\frac{\partial w}{\partial p_n^o} = \frac{p_n^o\gamma_n - \log(1+p_n^o\gamma_n)}{g_{nn}\log(1+p_n^o\gamma_n)}$. Since $x > \log(1+x)$ for all $x > 0$, we have $\frac{\partial p_n^o}{\partial I} > 0$. The monotonicity follows immediately. Furthermore,

$$\frac{\partial^2 p_n^o}{\partial I^2} = \frac{\partial \frac{\partial p_n^o}{\partial I}}{\partial I} = -\frac{p_n^o(-p_n^o\gamma_n + (1+p_n^o\gamma_n)\log(1+p_n^o\gamma_n))}{(I+\sigma^2)(I+\sigma^2+p_n^o g_{nn})\log(1+p_n^o\gamma_n)^2}. \tag{D.11}$$

We can easily show that $(1+x)\log(1+x) > x$ for all $x > 0$ since $(1+0)\log(1+0) = 0$ and $(1+x)\log(1+x) - x$ has positive first-order derivative when $x > 0$. Thus, $\frac{\partial^2 p_n^o}{\partial I^2} < 0$. Since I is a linear combination of $\mathbf{p}_{-n}$, $f_n(\mathbf{p}_{-n})$ is strictly concave in $\mathbf{p}_{-n}$. We get the scalability immediately by letting $F(\alpha) = \alpha f_n(\mathbf{p}_{-n}) - f_n(\alpha\mathbf{p}_{-n})$ and observing that $F(1) = 0$ and $\frac{\partial^2 F(\alpha)}{\partial\alpha^2} < 0$.

D.6 Proof of An Equilibrium Form

Proof We need to show that one of the equilibrium has the form $\mathbf{p}_1^* = [p_a p_b]$ and $\mathbf{p}_2^* = [p_c 0]$, where p_a, p_b, and p_c are positive. We only need to verify that there exist p_a, p_b, and p_c that satisfy Theorem 7.3.4. Suppose $\mathbf{p}_2^* = [p_c 0]$. After some calculation, it is easy to see that $\sigma^2 \gg p_c g_{21}^{(1)}$ and $\eta_1^{(1)} \approx \frac{p_a g_{11}^{(1)}}{\sigma^2}$. Hence, both subchannels of User 1 have approximately the same SINR condition. Thus in the equilibrium, the transmit powers on the two subchannels of User 1 are almost the same. Besides, they cannot be zero. Hence, both are positive and satisfy the first condition of Theorem 7.3.4. Assume $\mathbf{p}_1^* = [p_a p_b]$. Now we verify $\mathbf{p}_2^*$. Since User 2 does not transmit on the second subchannel, $\frac{\sum_{j\neq 1} r_n^{(j)*}}{p_c + \sum_{j\neq 1} p_n^{(j)*}} = 0$ and the first condition of Theorem 7.3.4 should be satisfied. Hence, a positive power is allocated on the first subchannel in the equilibrium of User 2. Regarding the second subchannel, $\gamma_n^{(2)*} = \frac{g_{nn}^{(2)}}{p_b g_{12}^{(2)} + \sigma^2} = \frac{1}{p_b 1e^{10} + 1} \to 0$. Hence, $\frac{\sum_{j\neq 2} r_2^{(j)*}}{p_c + \sum_{j\neq 2} p_n^{(j)*}} > R'(0)\gamma_n^{(2)*} \to 0$ and condition 2 of Theorem 7.3.4 is satisfied. Hence, $\mathbf{p}_2^* = [p_c 0]$. Numerical methods can be used to determine the exact values of p_a, p_b, and p_c.

D.7 Proof of Theorem 7.3.8

Proof For any two power vectors $\mathbf{p}_{-n}$ and $\check{\mathbf{p}}_{-n}$, define the function $\mathfrak{F}_n(\theta) = F_n(\check{\mathbf{p}}_{-n} + \theta(\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n}))$. It is clear that $\mathfrak{F}_n(0) = F_n(\check{\mathbf{p}}_{-n})$ and $\mathfrak{F}_n(1) = F_n(\mathbf{p}_{-n})$; By the chain rule, we know that $\frac{\partial \mathfrak{F}_n}{\partial \theta} = (\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n}) \frac{\partial F_n}{\partial (\check{\mathbf{p}}_{-n} + \theta(\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n}))}$. Hence, we have

$$F_n(\mathbf{p}_{-n}) - F_n(\check{\mathbf{p}}_{-n}) = \mathfrak{F}_n(1) - \mathfrak{F}_n(0) = \int_0^1 \mathfrak{F}_n'(\theta) d\theta = (\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n}) \int_0^1 \frac{\partial F_n}{\partial (\check{\mathbf{p}}_{-n} + \theta(\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n}))} d\theta.$$

$$\begin{aligned}
&\text{Thus, } \|F_n(\mathbf{p}_{-n}) - F_n(\check{\mathbf{p}}_{-n})\| = \left\| (\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n}) \int_0^1 \frac{\partial F_n}{\partial (\check{\mathbf{p}}_{-n} + \theta(\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n}))} d\theta \right\| \\
&\leq \|(\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n})\| \left\| \int_0^1 \frac{\partial F_n}{\partial (\check{\mathbf{p}}_{-n} + \theta(\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n}))} d\theta \right\| \leq \|(\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n})\| \int_0^1 \left\| \frac{\partial F_n}{\partial (\check{\mathbf{p}}_{-n} + \theta(\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n}))} \right\| d\theta \\
&\leq \|(\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n})\| \int_0^1 \left\| \sup_{\mathbf{p}_{-n}} \frac{\partial F_n}{\partial \mathbf{p}_{-n}} \right\| d\theta = \|(\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n})\| \left\| \sup_{\mathbf{p}_{-n}} \frac{\partial F_n}{\partial \mathbf{p}_{-n}} \right\|.
\end{aligned}$$

Besides, according to the chain rule, $\frac{\partial F_n}{\partial \mathbf{p}_{-n}} = \frac{\partial \mathbf{I}_n}{\partial \mathbf{p}_{-n}} \frac{\partial \tilde{F}_n}{\partial \mathbf{I}_{-n}}$; Hence, we have

$$\frac{\|F_n(\mathbf{p}_{-n}) - F_n(\check{\mathbf{p}}_{-n})\|}{\|\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n}\|} \leq \sup_{\mathbf{p}_{-n}} \left\| \frac{\partial F_n}{\partial \mathbf{p}_{-n}} \right\| = \sup_{\mathbf{p}_{-n}} \left\| \frac{\partial \mathbf{I}_n}{\partial \mathbf{p}_{-n}} \frac{\partial \tilde{F}_n}{\partial \mathbf{I}_{-n}} \right\| \leq \left\| \frac{\partial \mathbf{I}_n}{\partial \mathbf{p}_{-n}} \right\| \sup_{\mathbf{I}_{-n}} \left\| \frac{\partial \tilde{F}_n}{\partial \mathbf{I}_{-n}} \right\|;$$

When $\left\| \frac{\partial \mathbf{I}_n}{\partial \mathbf{p}_{-n}} \right\| < \frac{1}{\sup_{\mathbf{I}_n} \left\| \frac{\partial \tilde{F}_n}{\partial \mathbf{I}_n} \right\|}$, $\frac{\|F_n(\mathbf{p}_{-n}) - F_n(\check{\mathbf{p}}_{-n})\|}{\|\mathbf{p}_{-n} - \check{\mathbf{p}}_{-n}\|} < 1$. The uniqueness of equilibrium follows immediately from Theorem 7.3.7.

REFERENCES

[1] K. Lahiri, A. Raghunathan, S. Dey, and D. Panigrahi, "Battery-driven system design: A new frontier in low power design," in *Proc. Intl. Conf. on VLSI Design*, (Bangalore, India), pp. 261–267, Jan. 2002.

[2] U. C. Kozat, I. Koutsopoulos, and L. Tassiulas, "Cross-layer design for power efficiency and QoS provisioning in multi-hop wireless networks," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 3306–3315, Nov. 2006.

[3] P. P. Pham, S. Perreau, and A. Jayasuriya, "New cross-layer design approach to *ad hoc* networks under rayleigh fading," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 28–39, Jan. 2005.

[4] M. A. Haleem and R. Chandramouli, "Adaptive downlink scheduling and rate selection: a cross-layer design," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 1287–1297, Jan. 2005.

[5] M. Pedram, "Power optimization and management in embedded systems," in *Proc. ASP-DAC 2001*, (Yokohama, Japan), pp. 239–244, Feb. 2001.

[6] L. Benini, A. Bogliolo, and G. De Micheli, "A survey of design techniques for system-level dynamic power management," *IEEE Trans. VLSI Syst.*, vol. 8, pp. 299–316, June 2000.

[7] A. Raghunathan, N. Jha, and S. Dey, *High-level power analysis and optimization*. Norwell, MA: Kluwer Academic Publishers, 1998.

[8] C. Schurgers, *Energy-Aware Wireless Communications*. PhD thesis, University of California Los Angeles, 2002.

[9] IEEE, "IEEE 802.16e-2004, part 16: air interface for fixed and mobile broadband wireless access systems - amendment for physical and medium access control layers for combined fixed and mobile operation in licensed bands," Nov. 2004.

[10] Y. Xiao, "Energy saving mechanism in the IEEE 802.16e wireless MAN," *IEEE Commun. Let.*, vol. 9, pp. 595–597, July 2005.

[11] Atheros Communications, "White paper: Power consumption & energy efficiency," 2003.

[12] Draft IEEE std 802.16e/D9, "IEEE standard for local and metropolitan area networks part 16: air interface for fixed and mobile broadband wireless access systems," June 2005.

[13] 3GPP, TR 25.913, "Requirements for evolved UTRA and evolved UTRAN," www.3gpp.org.

[14] G. Song and Y. (G.) Li, "Cross-layer optimization for OFDM wireless networks-part i: theoretical framework," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 614–624, 2005.

[15] G. Song and Y. (G.) Li, "Cross-layer optimization for OFDM wireless networks-part ii: algorithm development," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 625–634, 2005.

[16] P. Viswanath, D. N. C. Tse, and R. L. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory.*, vol. 48, pp. 1277–1294, June 2002.

[17] R. Knopp and P. Humblet, "Information capacity and power controlin single-cell multiuser communications," in *In Proc., IEEE Int. Conf. on Commun.*, pp. 331–335, June 1995.

[18] G. W. Miao and N. Himayat, "Low complexity utility based resource allocation for 802.16 OFDMA systems," in *Proc. IEEE WCNC 2008*, pp. 1465–1470, Mar. 2008.

[19] D. Bertsekas and R. Gallager, *Data Networks (2nd Edition)*. Englewood Cliffs, NJ: Prentice Hall, 1992.

[20] V. Bhaghavan, A. Demers, S. Shenker, and L. Zhang, "MACAW: A media access protocol for wireless lan's," in *Proc. Sigcomm 1994*, Oct. 1994.

[21] X. Qin and R. Berry, "Exploiting multiuser diversity for medium access control in wireless networks," in *Proc. IEEE INFOCOM 2003*, pp. 1084–1094, Apr. 2003.

[22] K. Bai and J. Zhang, "Opportunistic multichannel Aloha: distributed multiaccess control scheme for OFDMA wireless networks," *IEEE Trans. Veh. Tech.*, vol. 55, pp. 848–855, Mar. 2006.

[23] G. Ganesan, G. Song, and Y. (G.) Li,, "Asymptotic throughput analysis of distributed multichannel random access schemes," in *Proc. IEEE ICC 2005*, pp. 3637–3641, 2005.

[24] S. Adireddy and L. Tong, "Exploiting decentralized channel state information for random access," *IEEE Trans. Inf. Theory.*, vol. 51, pp. 537–561, Feb. 2005.

[25] S. Adireddy and L. Tong, "Optimal channel-aware aloha protocol for random access in wlans with multipacket reception and decentralized channel state information," vol. 56, pp. 2575–2588, June 2008.

[26] X. Qin and R. Berry, "Opportunistic splitting algorithms for wireless networks," in *Proc. IEEE INFOCOM 2004*, pp. 1662–1672, Apr. 2004.

[27] Y. Xue, T. Kaiser, and A. B. Gershman, "Channel-aware Aloha-based OFDM subcarrier assignment in single-cell wireless communications," *IEEE Trans. Commun.*, vol. 55, pp. 953–962, May 2007.

[28] S. Adireddy, and L. Tong, "Exploiting decentralized channel state information for random access," *IEEE Trans. Inf. Theory.*, vol. 51, pp. 537–561, Feb. 2005.

[29] A. Maharshi, L. Tong, and A. Swami, "Cross-layer designs of multichannel reservation MAC under rayleigh fading," vol. 51, pp. 2054–2067, Aug. 2003.

[30] V. Naware, G. Mergen, and L. Tong, "Stability and delay of finite-user slotted Aloha with multipacket reception," *IEEE Trans. Inf. Theory.*, vol. 51, pp. 2636–2656, July 2005.

[31] I. F. Akyildiz, S. Weilian, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Magazine*, vol. 40, pp. 102–114, Aug. 2002.

[32] I. Chlamtac, M. Conti b, J. J.-N. Liu, "Mobile *ad hoc* networking: imperatives and challenges," *Ad Hoc Networks*, vol. 1, pp. 13–64, July 2003.

[33] I. F. Akyildiz and X. Wang, "A survey on wireless mesh networks," *IEEE Commun. Magazine*, vol. 43, pp. S23–S30, Sept. 2005.

[34] 3GPP TS 25.321, "3rd generation partnership project; technical specification group radio access network; medium access control (MAC) protocol specification," July 2007.

[35] R. Prakash, N.Shivaratri, and M. Singhal, "Distributed dynamic fault-tolerant channel allocation for mobile computing," *IEEE Trans. Veh. Tech.*, vol. 48, pp. 1874–1888, Nov. 1999.

[36] G. Cao and M. Singhal, "An adaptive distributed channel allocation strategy for mobile cellular networks," *J. Parallel and Dist. Comput.*, vol. 60, pp. 451–473, 2000.

[37] G. L. Stüber, *Principles of Mobile Communication*. Norwell, MA: Kluwer Academic Publishers, 2001.

[38] K. Begain, G. I. Rozsa, A. Pfening, and M. Telek, "Performance analysis of GSM networks with intelligent underlay-overlay," in *Proc. 7th Int. Symp. on Comp. and Commun. (ISCC 2002)*, pp. 135–141, 2002.

[39] M. C. Necker, "Coordinated fractional frequency reuse," in *Proc. 10th ACM Symp. on Mod., Anal., and Sim. of wireless and Mob. Syst.*, pp. 296–305, 2007.

[40] 3GPP TSG-RAN R1-050764, "Inter-cell interference handling for E-UTRA," Ericsson, Sept. 2005.

[41] 3GPP TSG RAN R1-050841, "Further analysis of soft frequency reuse scheme," Huawei, Aug. 2005.

[42] J. G. Andrews , "Interference cancellation for cellular systems: a contemporary overview," *IEEE Wireless Commun.*, vol. 12, pp. 19–29, Apr. 2005.

[43] H. Schoeneich and P. A. Hoeher, "Single antenna interference cancellation: iterative semi-blind algorithm and performance bound for joint maximum-likelihood interference cancellation," in *Proc. IEEE Global Commun. Conf. 2003*, pp. 1716–1720, 2003.

[44] J. Li, B. K. Letaief, and Z. Cao , "Co-channel interference cancellation for space-time coded OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 41–49, Jan. 2003.

[45] H. Dai, A. F. Molisch, and H. V. Poor, "Downlink capacity of interference-limited MIMO systems with joint detection," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 442–453, Mar. 2004.

[46] S. Shamai and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *Proc. Conf. Veh. Tech. Spring, 2005*, vol. 3, pp. 1745–1749, 2001.

[47] H. Zhang and H. Dai, "Cochannel interference mitigation and cooperative processing in downlink multicell multiuser MIMO networks," *EURASIP J. Wireless Commun. and Networking*, pp. 222–235, Feb. 2004.

[48] W. Choi and J. G. Andrews, "Base station cooperatively scheduled transmission in a cellular MIMO TDMA system," in *Proc. 40th Annual Conf. Inf. Sci. Sys.*, pp. 105–110, 2006.

[49] P. E. Omiyi and H. Haas, "Improving time-slot allocation in 4^{th} generation OFDM/TDMA TDD radio access networks with innovative channel-sensing," in *Proc. IEEE ICC 2004*, vol. 6, pp. 3133–3137, 2004.

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

[51] Y. (G.) Li and G. L. Stüber, *OFDM for Wireless Communications*. Springer, 2006.

[52] S. Sampei, S. Komaki, S., and N. Morinaga, "Adaptive modulation/TDMA scheme for personal multimedia communication systems," in *Proc. IEEE GLOBECOM 1994*, pp. 989–993, Nov. 1994.

[53] M. Ouchi, H. J. Lee, S. Komaki, and N. Morinaga, "Proposal for modulation level controlled radio system applied to ATM networks," in *Proc. of Fourth European Conference on Radio Relay Systems*, pp. 322–327, Oct. 1993.

[54] R. G. Gallager, "Power limited channels: Coding, multiaccess, and spread spectrum," in *Proc. Conf. Inform. Sci. and Syst.*, vol. 1, Mar. 1988.

[55] S. Verdu, "On channel capacity per unit cost," *IEEE Trans. Inf. Theory.*, vol. 36, pp. 1019–1030, Sept. 1990.

[56] S. Verdu, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory.*, vol. 48, pp. 1319–1343, June 2002.

[57] F. Meshkati, H. V. Poor, S. C. Schwartz, and N. B. Mandayam, "An energy-efficient approach to power control and receiver design in wireless networks," *IEEE Trans. Commun.*, vol. 5, pp. 3306–3315, Nov. 2006.

[58] A. Y. Wang, S. Cho, C. G. Sodini, and A. P. Chandrakasan, "Energy efficient modulation and MAC for asymmetric RF microsensor system," in *Int. Symp. Low Power Electronics and Design*, (Huntington Beach, CA), pp. 106–111, 2001.

[59] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2349–2360, Sept. 2005.

[60] G. Miao, N. Himayat, Y. Li, and D. Bormann, "Energy-efficient design in wireless OFDMA," in *Proc. IEEE ICC 2008*, pp. 3307–3312, May 2008.

[61] C. Schurgers and M. B. Srivastava, "Energy optimal scheduling under average throughput constraint communications," in *Proc. IEEE ICC 2003*, vol. 3, pp. 1648–1652, 2003.

[62] S. Cui, A. Goldsmith, and A. Bahai, "Joint modulation and multiple access optimization under energy constraints," in *Proc. IEEE GLOBECOM 2004*, vol. 1, pp. 151–155, Nov. 2004.

[63] R. Mangharam, R. Rajkumar, S. Pollin, F. Catthoor, B. Bougard, L. Van der Perre, and I. Moeman, "Optimal fixed and scalable energy management for wireless networks," in *Proc. IEEE INFOCOM 2005*, vol. 1, pp. 114–125, Mar. 2005.

[64] S. Cui, A. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," in *IEEE J. Sel. Areas Commun.*, vol. 22, pp. 1089–1098, Aug. 2004.

[65] S. K. Jayaweera, "An energy-efficient virtual MIMO architecture based on V-BLAST processing for distributed wireless sensor networks," in *Proc. IEEE SECON 2004*, pp. 299–308, Oct. 2004.

[66] J. Rabaey, J. Ammer, J. L. da Silva Jr., and D. Patel, "PicoRadio: *Ad-hoc* wireless networking of ubiquitous low-energy sensor/monitor nodes," in *IEEE VLSI*, pp. 9–12, 2000.

[67] M. Haenggi and D. Puccinelli, "Routing in *ad hoc* networks: a case for long hops," *IEEE Commun. Magazine*, vol. 43, pp. 112–119, Oct. 2005.

[68] D. Goodman and N. Mandayam, "Power control for wireless data," vol. 7, pp. 48–54, Apr. 2000.

[69] D. Gesbert, S. G. Kiani, A. GjendemsjØ, and G. E. Øien , "Adaptation, coordination, and distributed resource allocation in interference-limited wireless networks," vol. 95, pp. 2393–2409, Dec. 2007.

[70] A. GjendemsjØ, D. Gesbert, G. E. Øien, and S. G. Kiani, "Optimal power allocation and scheduling for two-cell capacity maximization," in *Proc. RAWNET (WiOpt)*, Apr. 2006.

[71] F. Meshkati, H. V. Poor, and S. C. Schwartz, "Energy-efficient resource allocation in wireless networks," *IEEE Commun. Magazine*, pp. 58–68, May 2007.

[72] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inf. Theory.*, vol. 45, pp. 1468–1489, July 1999.

[73] IEEE, P802.11D5, Draft Standard IEEE 802.11, "Wireless LAN medium access control (MAC) and physical layer (phy) spec," May 1996.

[74] A. J. Goldsmith and S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 1218–1230, Oct. 1997.

[75] E. L. Hahne, "Round-robin scheduling for max-min fairness in data networks," *IEEE J. Sel. Areas Commun.*, vol. 9, pp. 1024–1039, Sept. 1991.

[76] R. Mazumdar, L. G. Mason, and C. Douligieris, "Fairness in network optimal flow control: optimality of product forms," *IEEE Trans. Commun.*, vol. 39, pp. 775–782, May 1991.

[77] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability," *J. Operat. Res. Soc.*, vol. 49, pp. 237–252, 1998.

[78] P. D. Straffin, *Game theory and strategy*. Mathematical Association of America, 1993.

[79] K. Kar, S. Sarkar, and L. Tassiulas, "Achieving proportional fairness using local information in Aloha networks," *IEEE Trans. Autom. Control*, vol. 49, pp. 1858–1863, Oct. 2004.

[80] P. Gupta and A. L. Stolyar, "Optimal throughput allocation in general random-access networks," in *Proc. Conf. Information Science and Systems (CISS)*, (Princeton, NJ), pp. 1254–1259, 2006.

[81] A. Qayyum, L. Viennot, and A. Laouiti, "Multipoint relaying: an efficient technique for flooding in mobile wireless networks," in *Proc. of 35th Annual Hawaii Int. Conf. on Sys. Sci.*, pp. 3866– 3875, 2002.

[82] J. Lipman, P. Boustead, and J. Judge, "Neighbor aware adaptive power flooding in mobile ad hoc networks," in *Int. J. Foundations of Comp. Sci.*, vol. 14, pp. 237–252, Apr. 2003.

[83] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory.*, vol. 43, pp. 1986–1992, July 1997.

[84] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *Europ. Trans. Telec.*, vol. 10, pp. 585–595, Nov. 1999.

[85] J. Jang, and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 171–178, Feb. 2003.

[86] G. W. Miao and Z. Niu, "Practical feedback design based OFDM link adaptive communications over frequency selective channels," in *Proc. IEEE ICC 2006*, pp. 4624–4629, June 2006.

[87] T. Keller and L. Hanzo, "Adaptive multicarrier modulation: a convenient framework for timefrequency processing in wireless communications," in *Proc. of the IEEE*, vol. 88, pp. 611–640, May 2000.

[88] R. J. McEliece and W. E. Stark, "Channels with block interference," *IEEE Trans. Inf. Theory.*, vol. 30, pp. 44–53, Jan. 1984.

[89] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[90] M. Hellebrandt, R. Mathar, and M. Scheibenbogen, "Estimating position and velocity of mobiles in a cellular radio network," *IEEE Trans. Veh. Tech.*, vol. 46.

[91] C. Drane, M. Macnaughtan, and C. Scott, "Positioning GSM telephones," vol. 36.

[92] M. Hellebrandt and R. Mathar, "Location tracking of mobiles in cellular radio networks," *IEEE Trans. Veh. Tech.*, vol. 48.

[93] G. Sun, J. Chen, W. Guo, and K.J.R. Liu, "Signal processing techniques in network-aided positioning: a survey of state-of-the-art positioning designs," *IEEE Sig. Processing Mag.*, vol. 22.

[94] M. Pischella and J.-C. Belfiore,, "Power control in distributed cooperative OFDMA cellular networks," vol. 7, pp. 1900–1906, May 2008.

[95] K. K. Leung, B. McNair, L. J. Cimini, J. H. Jr. Winters, "Outdoor IEEE 802.11 cellular networks: MAC protocol design andperformance," in *Proc. IEEE ICC 2002*, vol. 1, pp. 595–599, 2002.

[96] A. Kumar, D. Manjunath, and J. Kuri, *Wireless Networking*. Morgan Kaufmann, 2008.

[97] G. W. Miao, Y. (G.) Li, and A. Swami, "Decentralized optimization for multichannel random access," to appear in *IEEE Trans. Commun.*

[98] D Tse and P Viswanath , *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[99] J. M. Cioffi, *A Multicarrier Primer*. ANSI T1E1, 1999.

[100] C. E. Shannon, "Communication in the presence of noise," in *Proc. IRE*, vol. 37, pp. 10–21, Jan. 1949.

[101] N. Feng, S. C. Mau, and N. B. Mandayam, "Pricing and power control for joint network-centric and user-centric radio resource management," *IEEE Trans. Commun.*, vol. 52, pp. 1547–1557, Sept. 2004.

[102] E. Wolfstetter, *Topics in Microeconomics: Industrial Organization, Auctions, and Incentives*. Cambridge University Press, 1999.

[103] B. Prabhakar, E. U. Biyikoglu, and A. E. Gamal, "Energy-efficient transmission over a wireless link via lazy packet scheduling," in *Proc. IEEE Infocom 2001*, vol. 1, pp. 386–394, Apr. 2001.

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

[105] K. C. Kiwiel and K. Murty, "Convergence of the steepest descent method for minimizing quasiconvex functions," in *J. of Optimization Theory and Applications*, vol. 89, pp. 221–226, Sept. 2005.

[106] ITU-R Recommendation M.1225, "Guidelines for evaluation of radio transmission technologies for IMT-2000," 1997.

[107] G. Miao, N. Himayat, and Y. Li, "Energy-efficient transmission in frequency-selective channels," in *Proc. IEEE Globecom 2008*, pp. 1–5, Dec. 2008.

[108] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA: MIT Press, 1991.

[109] R. D. Yates, "A framework for uplink power control in cellular radio systems," pp. 1341–1347, Sept. 1995.

[110] R. G. Bartle, *The elements of real analysis*. New York: John Wiley, 1964.

[111] J. W. Friedman, *Oligopoly and the theory of games*. North-Holland Pub. Co., 1977.

[112] A. T. Koc, S. Talwar, A. Papathanassiou, R. Yang, N. Himayat, and H. Yin, "IEEE c802.16m-08/666r2: Uplink power control recommendations for IEEE 802.16m.".

[113] S. M. Ross, *Stochastic Process*. John Wiley & Sons, Inc., 1996.

[114] K. Feher, *Advanced Digital Communications*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

VITA

Guowang Miao was born in Taizhou City, Jiangsu Province, China. He received a B.S. and a M.S. in Electronic Engineering from Tsinghua University, Beijing, China, in 2003 and 2006 respectively. In 2009, he received a second M.S. and a Ph.D., both in Electrical and Computer Engineering at Georgia Institute of Technology.