# Distributed Multiple Access for OFDMA Femtocells

Abdulmohsen Almutairi

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2014

Reading Committee:

Sumit Roy, Chair

Archis Vijay Ghate

Thomas R Henderson

Program Authorized to Offer Degree:
Electrical Engineering

University of Washington

**Abstract**

Distributed Multiple Access for OFDMA Femtocells

Abdulmohsen Almutairi

Chair of the Supervisory Committee:
Professor Sumit Roy
Electrical Engineering

Demand for data services in cellular networks is growing exponentially, due to the proliferation of high-end, multimedia-enabled mobile devices. To meet this challenge, cellular operators are moving toward a heterogeneous network architecture consisting of macrocells for wide-area coverage and smaller cells such as femtocells for capacity boost in local hotspots. The primary problem in such heterogeneous networks is mitigating inter-cell interference especially in dense deployments of residential femtocells. Traditionally, interference in wide-area cellular networks has been studied from a multi-cell resource allocation perspective, where radio resources, e.g. power, bandwidth, are allocated to different cells to reduce interference. This approach generally assumes a fully loaded network (i.e. many simultaneous active users in a cell) where the system performance is insensitive to the activity of a single user. This assumption is not suitable for femtocells which are designed to serve very few users and thus lacks the presumed traffic aggregation.

In this thesis, we first discuss the infeasibility of applying the classical multi-cell resource allocation framework to the femtocell case. Then, we motivate the case of using distributed, random access protocols as apposed to centralized interference mitigation techniques in the context of LTE femtocells. For this, we employ queuing analysis to compare the performance of centralized resource allocation schemes represented by an ideal power control to the performance of a simple random access protocols. The results show that, with a high probability, a simple un-optimized random access protocol such as Aloha would perform

much better than an optimized power controlled network in most of the unsaturated traffic scenarios. This result serves as a strong motivation for the design of OFDMA-aware random access protocols that utilize the inherent frequency diversity of OFDMA to improve the MAC performance.

Next, we analyze OFDMA-Aloha which is the simplest protocol that utilize the frequency-dimension in OFDMA to improve random access performance. The protocol attempts to reduce the packet retransmission time using collision resolution over the frequency domain by switching subchannels randomly after each collision. However, this comes at the expense of expanded time scale, or larger slot size due to lower channel rates. We showed that when the network is lightly loaded, the reduction in the collision rate outweighs the effect of expanded slot size thus in these situations OFDMA-Aloha enjoys smaller packet delays than the single channel Aloha.

Then, we propose the Exponential Backoff in Frequency (EBF) algorithm to address the case when multiple packets need to be transmitted over multiple subchannels simultaneously. Instead of spreading the packets uniformly over $K$ subchannels as would be done in a multichannel variant of Aloha (K-Aloha), the EBF algorithm keeps packet transmissions clustered at few frequency branches using a synchronized binary tree branching process over the $K$ subchannels. Collisions are handled by reducing the accessed bandwidth exponentially and using an access probability that is inversely proportional to the accessed bandwidth. Analysis of a lightly loaded network, shows that EBF enjoys considerably less packet delay compared to the basic K-Aloha protocol.

Finally, we propose the OFDM-based Reservation Random access (OFDM-RR) protocol that utilizes the frequency-dimension in a different way. In this protocol, reservation requests are transmitted on randomly selected subchannels and the indices of these subchannels are then used to create an implicit ordering among competing nodes for conflict resolution. Our analysis shows that beyond some critical value of the network load, OFDM-RR significantly improves the system throughput compared to the canonical Reservation-Aloha (R-Aloha) protocol.

## TABLE OF CONTENTS

Page

Glossary . . . . . iii

Chapter 1: Introduction . . . . . 1
1.1 Practical Solutions to The Capacity Challenge . . . . . 2
1.2 The Femtocell Concept . . . . . 5
1.3 Femtocell Interference . . . . . 9
1.4 Thesis Statement . . . . . 10

Chapter 2: Femtocell Interference Mitigation . . . . . 13
2.1 The Centralized Multi-cell Resource Allocation Problem . . . . . 13
2.2 The Case for Distributed Interference Mitigation in Femtocells . . . . . 18
2.3 Proposed Approach . . . . . 20
2.4 Integrating OFDMA-based Random Access in LTE Femtocells . . . . . 24
2.5 Simplified MAC Model for OFDMA Femtocells . . . . . 33
2.6 Classical Tools for Random MAC Analysis . . . . . 37

Chapter 3: Optimal Power Control Versus Random Access . . . . . 42
3.1 Practical Self-Optimization Approaches in Femtocells . . . . . 42
3.2 System Model and Assumptions . . . . . 45
3.3 Optimal Power Allocation . . . . . 48
3.4 Queuing Analysis of the Packet Delay . . . . . 54
3.5 Numerical Evaluation . . . . . 58
3.6 Conclusions . . . . . 61

Chapter 4: Delay Analysis of OFDMA-Aloha . . . . . 63
4.1 System Model and Assumptions . . . . . 64
4.2 Saturated Case Analysis . . . . . 66
4.3 Unsaturated Case Analysis . . . . . 74
4.4 Conclusion . . . . . 86

4.5 Appendix . . . . . 87

Chapter 5: Exponential Backoff in Frequency (EBF) . . . . . 90
5.1 System Model and Baseline Scenario . . . . . 90
5.2 Proposed MAC Algorithm . . . . . 92
5.3 Performance Analysis . . . . . 93
5.4 Numerical and Simulation Results . . . . . 99
5.5 Conclusion . . . . . 101

Chapter 6: OFDM-based Reservation Random Access . . . . . 103
6.1 Proposed Random Access Protocol . . . . . 104
6.2 Throughput Analysis . . . . . 108
6.3 Throughput in a Fixed-slot System . . . . . 114
6.4 Discussion and Numerical Results . . . . . 116
6.5 Conclusions . . . . . 118
6.6 Appendix A: Proof of proposition 3 . . . . . 120
6.7 Appendix B: Approximation of $f_n(K) = Kd_n(K)$ . . . . . 121
6.8 Appendix C: Proof of proposition 4 . . . . . 124

Chapter 7: Summary and Future Work . . . . . 125

Bibliography . . . . . 127

## GLOSSARY

3GPP: Third Generation Partnership Project.

4G: Fourth Generation.

AMC: Adaptive Modulation and Coding.

ARQ: Automatic Repeat Request.

AWGN: Additive White Gaussian Noise.

CDMA: Code Division Multiplexing.

DCH: Dedicated Channel.

DSCH: Downlink Shared Channel.

FAP: Femtocell Access Point.

FCC: Federal Communications Commission.

FDD: Frequency Division Duplex.

FDMA: Frequency Division Multiplexing.

GPS: Global Positioning System.

GSM: Global System for Mobile communications.

H-ARQ: Hybrid Automatic Repeat Request.

HSPA: High Speed Packet Access.

IP: Internet Protocol.

ISM: Industrial, Science, Medical.

LTE: Long Term Evolution.

LTE-A: Long Term Evolution - Advanced.

MAC: Medium Access Control.

MCS: Modulation and Coding Scheme.

MIMO: Multi Input Multi Output.

OFDM: Orthogonal Frequency Division Multiplexing.

OFDMA: Orthogonal Frequency Division Multiple Access.

PC: Power Control

PDU: Protocol Data Unit.

PHY: Physical layer.

RB: Resource Block.

RLC: Radio Link Control.

RNC: Radio Network Controller.

RTS/CTS: Ready To Send/Clear To Send.

SINR: Signal to Interference and Noise Ratio.

SIR: Signal to Interference Ratio.

SNR: Signal to Noise Ratio.

TB: Transport Block.

TDD: Time Division Duplex.

TDMA: Time Division Multiplexing.

TTI: Transmission Time Interval.

UE: User Equipment (Mobile phone).

UMTS: Universal Mobile Telecommunications System.

WIFI: Wireless Fidelity.

WLAN: Wireless Local Area Network.

## ACKNOWLEDGMENTS

I would like to thank my committee for their thoughtful commentary and helpful revision suggestions. Specifically, I would like to thank my advisor Sumit Roy for his exceptional guidance and mentorship. His advise and effort in reviewing my work over the years has helped improve my research skills and I am deeply grateful to have had the opportunity to work with him.

I would like to thank the Ministry of Higher Education in Saudi Arabia for funding my PhD study. Without this generous scholarship, I would not have the opportunity to pursue a graduate career.

Also, I would like to thank my colleagues in the Fundamental of Networking Lab and in the EE department for their support and helpful discussions.

Finally, I would like to thank my family for their constant encouragement and support during my graduate school career.

## DEDICATION

to my dear wife, Mariam

# Chapter 1

# INTRODUCTION

In the first and second generations of cellular networks, voice was the driving factor for the design and optimization of the wireless network. A voice call has a fixed transmission rate that requires a minimum received signal strength measured in terms of Signal-to-Interference Ratio (SIR). The key performance metric was the coverage probability distribution $\Pr\{\, \mathrm{SIR} \geq \mathrm{SIR}_{th} \,\}$ assuming random users' distribution over a given geographic area. The network was optimized in terms of the placement of base stations (topology) and their transmission power, which defines the cell area, to achieve this target SIR throughout the network. Upon arrival of a new call, the network controller reserves a radio resource unit, e.g. a frequency channel in an FDMA system or a time slot in a TDMA system (e.g. GSM), for the sole use of this call. Active calls within a cell and among adjacent cells are allocated different "orthogonal" resources (frequency channels or time slots) to reduce interference thus increasing the SIR. The number of frequency channels (or time slots) allocated for a given cell is limited, hence some voice calls are blocked and not carried by the network. The allocation of radio resources among cells is optimized to achieve a low blocking probability assuming common load patterns in the network.

This "call-centric" theme remained largely the same in the third generation (3G) of cellular networks which are based on CDMA. The main feature of this generation compared to prior ones was the support of video calling and multimedia streaming and some data services albeit limited. Video/streaming calls were handled pretty much the same as traditional voice calls: radio resources were reserved for the session during the call setup phase. For example, a video call or a TV streaming session in the 3GPP UMTS cellular standard was created by reserving multiple CDMA spreading codes on the Dedicated Channel (DCH) between the base station and the mobile device. The capacity of the network was often quantified in terms of the maximum number of simultaneous video and voice calls

that can be admitted with acceptable quality of service.

However, the theme is changing drastically; data (not voice) is now the key driving force behind the cellular evolution. The proliferation of smart phones and tablets has brought the full potential of the Internet into the cellular world. Traffic classified as "Mobile Internet" traffic has been more than doubling since the past few years [1]. It is not unreasonable to see this mobile Internet traffic surpassing the desktop Internet traffic in the near future as some have already projected. The impact of this growth on the cellular network is significant. Unlike fixed-rate voice communication, data applications are intrinsically capacity-hungry; they consume whatever capacity available in the network. Capacity is now the key target and coverage is becoming a second-priority. Also, our notion of network capacity has become the aggregate transmission rate of all active users, rather than just the number of active fixed-rate users.

In response, there has been a paradigm shift in the architecture and design of cellular networks starting from later revisions of 3G standards such as 3GPP HSPA (sometimes called 3.5G) and newer 4G standards such as 3GPP LTE/LTE-A. The old "call-centric" theme has been dropped in favor of a "packet-centric" theme where network resources are only allocated on a packet-by-packet basis rather than being reserved for the whole session. Furthermore, variable-rate communication, where the instantaneous transmission rate is "adapted" to the conditions of the wireless channel, is becoming a standard technique. With such "adaptive" communication, mobile users close to the base station (high SIR) enjoy better transmission rates compared to cell-edge users (low SIR). This disparity on its own has brought up new problems with regard to fairness and priority scheduling of users within a single cell. The bigger question however, is how the cellular network is going to overcome the formidable challenge of the exploding mobile traffic.

### *1.1 Practical Solutions to The Capacity Challenge*

Cellular operators around the world are attacking the capacity problem on multiple fronts: via acquiring new radio spectrum as well as via squeezing more bits/second/Hz of current spectrum allocations. The latter can be achieved by either using new wireless technologies with better spectral efficiency (measured in bits/s/Hz) or evolving the architecture of the

network to allow smaller cell sizes that provide better spatial spectral efficiency (measured in bits/s/Hz/area) [2].

### *1.1.1 More Spectrum*

The most straightforward way to increase capacity is to add more bandwidth. The maximum channel bandwidth in 4G cellular standards has been increased from 20 MHz in LTE up to 100 MHz in LTE-A by means of channel aggregation [3]. However, the availability of more licensed spectrum is a challenging task in many parts of the world. Beside licensed spectrum, the wireless industry is looking for new ways to capitalize on existing unlicensed bands such as the ISM bands commonly used by WiFi networks. Also, regulatory bodies such as the FCC are adopting new licensing policies that allow shared use of under-utilized spectrum. A secondary service provider would be allowed to utilize a licensed spectrum band provided that it does not interfere with the primary license holder. The so-called TV white space is a good candidate of this policy because TV towers have well-known temporal and spatial transmission patterns.

### *1.1.2 Better Spectral Efficiency*

The theoretical maximum spectral efficiency in an ideal AWGN channel is $\log_2(1 + \mathrm{SNR})$ bits/s/Hz. In reality however, the wireless channel is far from being ideal. Wireless signals suffer from path loss due to attenuation, fading due to multipath and interference due to nearby co-channel transmissions. To counter path loss, the transmission range is reduced using smaller cell sizes. To counter multi-path fading, OFDM modulation coupled with advanced coding techniques over time and over space using multiple antennas can be used to achieve significant gains. To counter the interference problem, we need coordinated scheduling between adjacent cells in addition to advanced receiver structures supporting interference cancellation. Inter-cell interference is a subject of great importance to cellular networks especially with the addition of many small cells as we will discuss in Section 1.3.

Multiplexing gains can also be achieved using MIMO where multiple data streams are transmitted toward the same user in parallel using multiple antennas. 3GPP LTE-A targets

a spectral efficiency of 30 bits/sec/Hz using 8x8 MIMO transmission [3]. On the cell level, channel-aware scheduling can be used to serve mobile users based on their instantaneous channel quality. By serving users with better channel conditions first and deferring users with worse channel conditions to a later time, higher cell capacity can be achieved from this opportunistic multi-user diversity. Furthermore, the highly flexible Orthogonal Frequency Division Multiple Access (OFDMA) scheme allows for more granular bandwidth allocations to serve a wide range of broadband data and multimedia applications. OFDMA divides a wide-band channel into many orthogonal narrow-band channels called sub-carriers which can be independently assigned to different users. Bandwidth flexibility is realized by allocating fewer sub-carriers to low-rate communication such as voice calls and more sub-carriers to capacity-demanding data applications. OFDMA is also a corner stone in the inter-cell interference mitigation techniques proposed in this research and will be described in depth in Chapter 2.

### *1.1.3 Smaller Cells*

Since capacity is proportional to the received signal strength, the surest way to increase the network capacity for a fixed bandwidth is to bring the receiver closer to the transmitter. In cellular networks, this means shrinking the cell size by using lower transmission power and deploying many small cells. However, this approach faces two problems. First, the infrastructure cost can become very prohibitive. Each cell requires its own power facility and communication facility to connect to the rest of the network via a reliable backhaul link. Also, the operator needs to maintain a high-speed (usually fiber-optic) backbone network to transport huge volumes of traffic from these small cells. The second problem is related to mobility. The small cell size implies that fast moving users have to switch from cell to cell more frequently. Handoff between cells requires control signaling in the network and can generate unwanted load. Also, frequent handoffs increase the probability of dropped calls and degraded quality of service.

To overcome these problems, the cellular industry is adopting a heterogeneous architecture with an overlay of different cell sizes and various backhaul options. Traditional

high-power, long-range *macrocells* are kept to handle wide-area mobility and coverage for truly mobile users while smaller cells are being introduced to provide localized capacity boost inside residential/office buildings and around public hot spots. The term *small cells* is an umbrella term for low-powered radio access base stations such as picocells, microcells and femtocells that operate in the licensed spectrum and have a range of 10 meters to several hundreds meters [4]. Picocells and microcells are compact base stations deployed by the operator to enhance capacity in densely populated public areas and enterprises. They are installed in strategic locations such as rooftops and lamp posts and utilize low-cost backhaul options such as ADSL/Cable connections. Femtocells are another class of small cells that follow a significantly different architectural model which we will explain the following section.

### *1.2 The Femtocell Concept*

Femtocells, also called Home Base Stations, are low cost, short range indoor wireless access points installed by the consumer to enhance the cellular network at his premises. Similar to a WiFi access point, a Femtocell Access Point (FAP) uses low power up to 21 dBm, compared to 43 dBm in macrocells, and thus covers a very small area of radius less than 30 m. The FAP utilizes the same licensed spectrum used by the outdoor macrocell base station and therefore can only be associated with one cellular operator. The functionality of the FAP is identical to the conventional macrocell base station; both provide voice and data services to registered users. The FAP, however, requires a modification to the network architecture as shown in Figure 1.1. In a standard cellular network, the mobile User Equipment (UE) connects to a macrocell Base Station (BS) which is linked to the operator's core network via a dedicated backhaul link; usually a fiber-optic or a microwave connection. The network operator is responsible for all radio network planning, cell site ownership, operation and maintenance of the macrocell tower. With femtocells, the operator only needs to provide the device and load some basic configuration into the FAP to allow easy plug-and-play installation by the customer. The customer connects the FAP to a broadband Internet connection such as ADSL or cable modems. Once powered on, the FAP establishes a secure tunnel over the Internet to a femtocell gateway that is connected to the core network. The

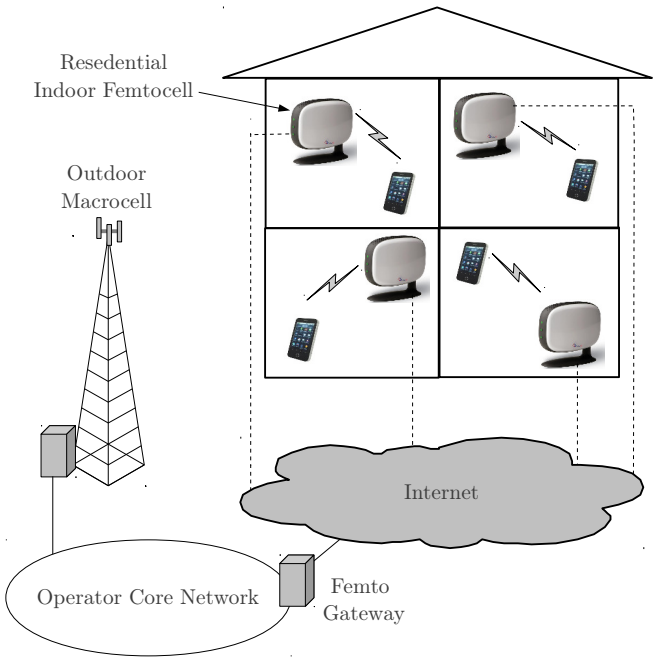


Figure 1.1: Femtocell network architecture

femtocell gateway is responsible for the configuration and authorization of the FAP. Once the users phone number is registered with the FAP, his UE will be handed-off from the macrocell to the femtocell and all his traffic is routed via the FAP over the Internet. When the user moves out of range of the femtocell, his UE will be handed back to the macrocell.

Access to the femtocell can be open, restricted or a hybrid of the two modes. In open-access femtocells, all nearby users can connect to the FAP. This turns the femtocell into a public hot spot in places like airports, shopping malls and large office buildings. This improves the indoor capacity, but increases the hand-off frequency which could strain the control plane of the network. In closed-access femtocells, only few authorized users can connect to the FAP. This is a typical residential scenario where the customer is generally not welling to let other users utilize his Internet connection. In this case, the FAP is pre-configured with the phone numbers associated with the customer, or may be equipped with a management interface that allow the user to add or remove registered phone numbers. Although this model is more commercially sound, it can create serious interference issues

to nearby users who are not authorized to use the femtocell, as will be discussed soon. The hybrid access model allows non-authorized users to connect to the FAP, but limits the radio and Internet bandwidth that they can utilize. For example, if a user is suffering from a strong interference from a nearby femtocell, then he could be handed-off to that femtocell to maintain connectivity with limited capacity.

### *1.2.1 Benefits*

From the customer's perspective, femtocells guarantee excellent indoor coverage (or 5-bar signal) which provides better quality of service and much higher speeds for data applications. Also, the low-power indoor transmission helps prolong the battery life of the mobile device, which is a notable issue in broadband-enabled smart phones. Simplicity or convenience is another advantage for the user. With femtocell, the user only needs one device with a single radio which can work seamlessly in both indoor femtocells and outdoor macrocells. In addition, the user can get personalized services from his femtocell, for example reduced tariffs, single family number service and many other home zone services.

Despite these benefits, the incentive for the customer may not be so strong especially with the ubiquity of WiFi-enabled smart phones and intelligent battery saving apps. It is the network operators who are pushing for the idea for two main reasons. First, femtocells provide huge capacity gains at very low cost if interference is handled properly. Because the femtocell is deployed and managed by the customer and utilizes his own power and Internet bandwidth, the operator is relieved from all associated operational cost. Second and more importantly, the traffic generated by the femtocell is carried over the Internet connection which is often provided by a third party. Since the majority of mobile Internet traffic originates from indoor users [5], femtocells provide an attractive approach to offload large volumes of Internet traffic away from the macrocellular network and save on otherwise necessary infrastructure upgrades. Overall, femtocells help cellular operators offload traffic and operational cost away from the macro network and hence reduce the cost per bits/sec [4].

### *1.2.2 Main Technical Challenges*

From its clear advantages, the femtocell market is expected to grow rapidly with several vendors specializing in building self-organizing femtocells and operators already offering the service. However, there are major technical challenges that need to be addressed before any large scale deployment of femtocells. The first and most important issue is the interference caused by femtocells. Because the FAP is installed by the customer and may be moved to another location at any time, it is difficult to apply traditional frequency planning mechanisms to reduce inter-cell interference. Therefore, with large number of femtocells in the network, the interference between femtocells and macrocells and between femtocells themselves becomes a serious issue. The problem will be discussed in more details in the next section.

The second challenge is the time synchronization between the femtocell and its neighboring macrocells. This is required for successful hand-off between femtocells and macrocells. Also, synchronization is essential for TDD systems which is typically achieved in 3G systems using high accuracy oscillators that are calibrated periodically over the dedicated low-latency backhaul links. This is not possible with the Internet-based connectivity of femtocells. GPS-based solutions have been suggested, but they may not be practical for indoor femtocells. Network-based protocols such as the IEEE 1588 Precision Timing Protocol (PTP) which can provide sub-microsecond accuracy may be more suitable for femtocells [5].

The third issue is how to identify the location of the femtocell. In order to configure the femtocell with the proper licensed spectrum, its location needs to be identified automatically. The location is also required by some regulations for emergency calls purposes. In some countries, the spectrum management authority requires the location of every base station using a licensed band to be identified. In some commercial femtocells, GPS with external antennas is used to lock the femtocell to a specific area. The user needs to arrange with operator before he can relocate the femtocell to a different area. Another solution is for the FAP to read the cell IDs of neighboring macro base stations to identify its location. This does not provide the precision required and does not address the cases where no macrocell coverage is available.

## *1.3 Femtocell Interference*

In traditional cellular networks, cell association which is the procedure for selecting the serving cell for each user is merely based the strongest signal received from all surrounding cells. With femtocells, cell association is based on ownership relationship rather than signal strength. Therefore, a user may receive a stronger signal from a nearby femtocell to which he is not allowed to connect. The femtocell then becomes an interference source to the user and could interrupt his service. There are many interference scenarios that can be caused by femtocells depending on who is the aggressor and who is the victim [6]. We summarize these scenarios into two main categories: macro-femto interference and femto-femto interference as illustrated in Figure 1.2.

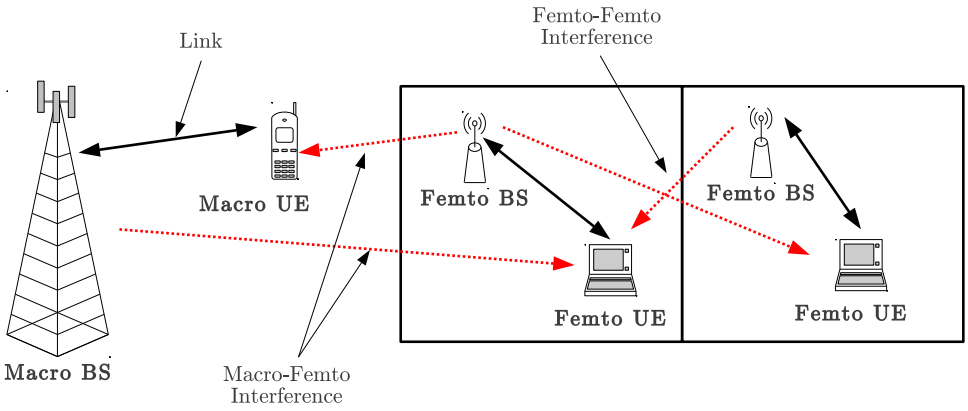


Figure 1.2: Femtocell interference scenarios

### *1.3.1 Macro-femto Interference*

A femtocell can create severe interference to a nearby macro user who is not allowed to join the femtocell. For example, a visiting macro user can be located in the same apartment where a closed-access femtocell is installed. The macro UE will receive the stronger signal and will continuously try to camp on it unsuccessfully. This effectively creates a dead zone around the femtocell where cellular service cannot be received. The opposite scenario in the uplink direction is also possible whereby a strong uplink signal from a macro UE can

degrade the service in a nearby femtocell UE. Although the probability of this event in a large network is small, it cannot be ignored in practice by network operators. Practical solutions have been proposed by the industry, e.g. switching the victim macro UE to a different carrier or capping the femtocell power below some threshold or even allowing the macro UE to join the femtocell temporarily [6].

### *1.3.2 Femto-femto Interference*

A cluster of closely located femtocells (e.g. in the same building) can interfere with each other and degrade the overall service. This scenario is commonly encountered in WiFi networks especially in dense apartment buildings. Because the current level of deployment of femtocells did not reach that of WiFi, this problem is currently not seen as serious issue by the industry. However, the problem will become more pronounced in the future with mass adoption of this technology. In contrast to WiFi networks, the fact that femtocells occupy the expensive licensed spectrum makes the mitigation of femto-femto interference extremely important for cellular operators.

Most of the current research on femtocells focuses on the macro-femto interference issue because of its timeliness and relevance to current market needs [7]. The majority of the papers assume a shared spectrum between femtocells and macrocells and propose some interference mitigation techniques either based on power control or orthogonal time-frequency splitting between macrocells and femtocells. However, very few papers addressed the issue of the competition (interference) among femtocells over whatever time-frequency resources allocated to them. This by itself is a challenging problem even in the absence of macrocells altogether, i.e. even if femtocells were allocated a dedicated spectrum band. The challenge arises from the ad-hoc nature of femtocell locations and the randomness of the user activity in each femtocell. This problem will be the main focus in this research.

## ***1.4 Thesis Statement***

The objective of this research is to study and develop simple distributed techniques for interference avoidance between adjacent femtocells, taking full account of the randomness of the user activity. We will target a specific class of femtocells that are based on OFDMA

since it is the technology of choice in most 4G cellular networks. Specifically, we will study a class of distributed random access protocols adapted carefully to exploit the inherent flexibility of OFDMA.

### *1.4.1 Motivation*

The motivation behind this research is the strong similarity between femtocells and ad-hoc IEEE 802.11 wireless networks (WiFi) where distributed random access protocols play a central role. WiFi runs on the unlicensed ISM band and is designed to serve few users at a time. Now as almost every household connected to the Internet has a WiFi access points, the ISM spectrum is becoming very crowded. WiFi uses a random access protocol with built-in radio-sensing capability to handle interference within the same access point while interference from adjacent access points is handled by manually switching to a different, hopefully less crowded channel. This can be a challenging task in a densely populated building and hence the wireless link quality often becomes very sensitive to the activity of other users in adjacent WiFi networks. In peak hours, this is commonly experienced as poor wireless performance due to frequent packet collisions caused by interference. The same situation can happen to femtocells if interference is not handled properly. Fortunately, femtocells are not completely ad-hoc. They belong to the same cellular operator and are tightly synchronized with the network. These properties, in addition to OFDMA, provide a clear opportunity to design more efficient random access protocols than what is currently used in WiFi.

### *1.4.2 Contributions*

The main philosophy of this thesis is to look at the femtocell interference problem from a multiple access (MAC) perspective rather than from a classical cellular resource allocation perspective. For this, we first discuss the infeasibility of applying the classical multi-cell resource allocation framework to the femtocell case. Then, we motivate the case of using distributed, random access protocols as apposed to centralized interference mitigation techniques in the context of LTE femtocells. Armed with appropriate MAC models that we

developed for this purpose, we propose and analyze different OFDMA-aware random access protocols that utilize the inherent frequency diversity of OFDMA to improve the MAC performance. The specific contributions of this work are the following:

1. We present a strong case for using distributed random access protocol to handle inter-cell interference in dense femtocell networks. For this, we employ queuing analysis to compare the packet-level performance of centralized resource allocation schemes represented by a highly idealized optimum power control with the performance of a simple interference mitigation using random access protocols. Because the randomness of the user activities cannot be accounted for in practical resource allocation schemes, channel resources are often wasted, thus the packet-level performance is generally worse than random access protocols.

2. Next, we analyze OFDMA-Aloha, the simplest way to utilize the frequency-dimension in OFDMA to improve random access performance. In this protocol, the throughput is improved by switching the channel randomly after each packet collision. We derive its mean packet delay and compare it to the basic Aloha in unsaturated load scenarios.

3. Then, we propose the Exponential Backoff in Frequency (EBF) algorithm which allows transmitting multiple packets on multiple subchannels simultaneously and reduces the number of subchannels (bandwidth) in a structured way after each packet collision. The delay performance of the protocol is derived and compared to a simple Aloha-based protocol that access multiple subchannel randomly.

4. Finally, we propose the OFDM-based Reservation Random access (OFDM-RR) protocol that utilizes the frequency-dimension in a different way. In this protocol, reservation requests are transmitted on randomly selected subchannels and the indices of these subchannels are used to create an implicit ordering among competing nodes which resolve conflicts in a distributed way. The throughput of this protocol is derived and compared to the canonical Reservation-Aloha protocol.

# Chapter 2

# FEMTOCELL INTERFERENCE MITIGATION

A typical femtocell covers a small area and serves very few users at a time. Therefore, it already operates at maximum spectral efficiency whereby additional gains from multiuser diversity and opportunistic channel-aware scheduling are marginal at best. In such environments, the main performance limiter is the interference from nearby co-channel cells. The downlink interference from macrocells is predictable because they are well planned and their location and power settings are fairly static. The real challenge comes from nearby "unplanned" femtocells that appear and disappear frequently.

Traditionally, inter-cell interference has been handled via careful frequency planning and power control. Frequency planning refers to the process of assigning different channels to adjacent cells and reusing the same channels in far-away cells. The amount of spectrum available at a given cell is limited to a fraction (called frequency reuse factor) of the total spectrum. This type of static channel assignment is no longer suitable for today's cellular networks where maximum utilization of the expensive spectrum is needed in every part of the network. In 4G cellular networks, full frequency reuse (reuse factor of 1) where all cells operate on the full spectrum is adopted and interference is handled in a dynamic way. Dynamic interference mitigation is often described as multi-cell resource allocation problem [8].

## *2.1 The Centralized Multi-cell Resource Allocation Problem*

The general downlink multi-cell resource allocation in cellular networks is illustrated in Figure 2.1. We will follow the centralized resource allocation model described in [8]. The network is composed of $N$ cells each serving $U_n$ users, $n = 1, \cdots, N$. Full frequency reuse is employed throughout the network so that all $N$ cells are operating on the same frequency channel (reuse factor 1). At the beginning of every time slot, a scheduling vector $\mathbf{U} =$

$(u_1, u_2, \cdots, u_N)$, where $u_n$ is the active user in cell $n$, is selected from the feasible set of all scheduling vectors $\Upsilon = \{\mathbf{U} : 1 \leq u_n \leq U_n \ \forall \ n = 1, \cdots, N\}$. Therefore, at any considered instant of time we have $N$ active transmit-receive pairs in the network while other users remain silent. Then, a vector of transmit powers $\mathbf{P} = (P_{u_1}, P_{u_2}, \cdots, P_{u_N})$ is selected from the set of all feasible power vectors $\Omega = \{\mathbf{P} : 0 \leq P_{u_n} \leq P_{\max} \ \forall \ u_n = 1, \cdots, U_N\}$. For notational clarity we use $P_{u_n}$ to denote the *downlink* transmit power of the active user $u_n$ in cell $n$.

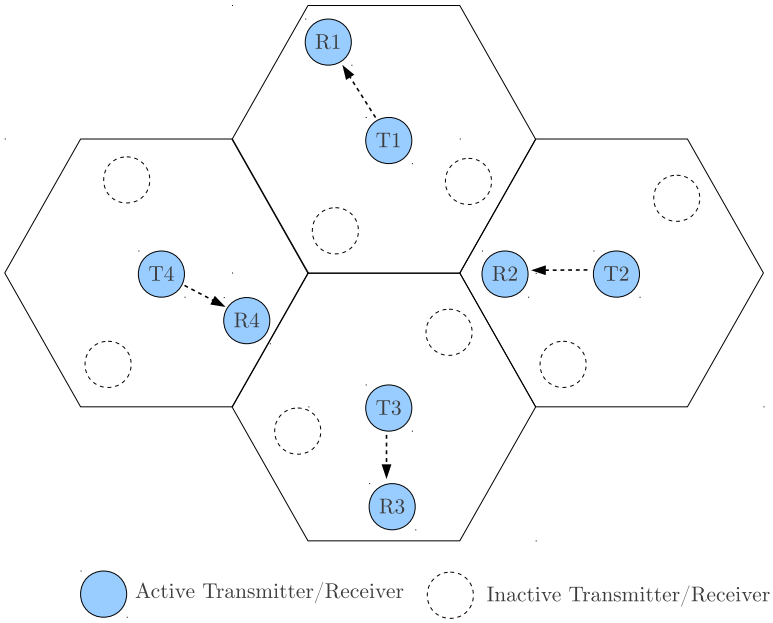


Figure 2.1: Network model for multi-cell resource allocation (adapted from [8])

In an interference-limited system, the merit associated with a particular choice of $(\mathbf{U}, \mathbf{P})$ is usually a utility function of $\Gamma_{u_n}$ the Signal to Interference and Noise Ratio (SINR) at user $u_n$ defined as:

$$\Gamma_{u_n}(\mathbf{U}, \mathbf{P}) = \frac{G_{n,u_n} P_{u_n}}{\sigma^2 + \sum_{i \neq n} G_{i,u_n} P_{u_i}}, \tag{2.1}$$

where $G_{i,u_n}$ is the channel gain from cell $i$ to user $u_n$ and $\sigma^2$ is the AWGN power. In this formulation, it is usually assumed that the channel gains of all transmit-receive pairs are known at the central controller. In a typical cellular network, each cell transmits a unique

reference signal (pilot symbols) used for synchronization and cell search. The users can detect these pilot symbols and estimate the channel gains from all surrounding cells.

Most wireless systems employ Adaptive Modulation and Coding (AMC) where the modulation scheme and the coding rate are adapted to the varying channel condition based on feedback from the user. For example, 3GPP LTE supports 15 AMC levels starting from QPSK modulation with rate 1/3 channel coding up to 64-QAM modulation with rate 4/5 channel coding [9]. The feedback from the user is usually a function of the channel gain and the SINR. Therefore, a reasonable choice of utility is a monotonically piecewise increasing function of the SINR, reflecting the various AMC levels in the system. Assuming an ideal, continuous-rate AMC, most researchers use the theoretical spectral efficiency $\log_2(1+\text{SINR})$ as a measure of the instantaneous user data rate, treating interference as noise. The total utility of the network is then taken to be the *sum rate* capacity defined as follows [8]:

$$\mathcal{C}(\mathbf{U},\mathbf{P}) \equiv \sum_{n=1}^{N} \log\left[1+\Gamma_{u_n}(\mathbf{U},\mathbf{P})\right]. \tag{2.2}$$

With this notion of network utility as our optimization objective, the optimal multi-cell resource allocation can be formalized as follow:

$$(\mathbf{U}^*,\mathbf{P}^*) \;=\; \underset{\substack{\mathbf{U}\in\Upsilon\\ \mathbf{P}\in\Omega}}{\arg\max} \quad \mathcal{C}(\mathbf{U},\mathbf{P}) \tag{2.3}$$

This utility maximization problem extends easily to multi channel systems such as OFDMA. Since OFDMA divides the channel into a fixed number, $K$, of orthogonal sub-channels (as will be discussed later), resource allocation in one sub-channel is independent from the allocation in other sub-channels. Therefore, the overall multi-cell resource allocation problem can be decomposed into $K$ parallel optimization problems of the form given by (2.3).

The optimization problem (2.3) is a very general formulation encompassing a wide range of resource allocation problems. If we ignore the scheduling of users within a cell and fix the set of active users, the problem reduces to the classical power control problem. On the other hand, if we restrict the power to $\Omega = \{0, P_{\max}\}$, i.e. binary "ON-OFF" power control, we end up with a combinatorial optimization problem for selecting an optimum schedule

of transmit-receive pairs. Despite the appeal of formality, solving (2.3) is non-trivial for theoretical and practical reasons as we will highlight in the following subsections.

### *2.1.1 Theoretical Aspects*

From a theoretical standpoint, the key difficulty of multi-cell resource allocation lies in the non-convexity of the problem. Because the SINR $\Gamma_{u_n}(\mathbf{U}, \mathbf{P})$ is not a convex function of $\mathbf{P}$, standard convex optimization techniques do not apply directly. Also, exhaustive search for $(\mathbf{U}^*, \mathbf{P}^*)$ is prohibitive. Fairness is also another issue. The network-wide optimal solution may favor users with good channel conditions and starve other users especially cell-edge users. Adding fairness measures into the problem serves only to complicate it further. There are however, different formulations that try to simplify the problem by targeting less ambitious objectives. For example, instead of maximizing the sum capacity, we can seek the minimization of the transmit power subject to a given quality of service constraint such as minimum data rate. Alternatively, we can achieve fairness by maximizing the minimum SINR in the network subject to a minimum SINR threshold.

### *2.1.2 Practical Aspects*

The major issue with the above approach to multi-cell resource allocation is the implementation complexity. Even if all computational issues were to be resolved, the optimal solution still requires global information about the network state in terms of channel gains in every time slot. Every base station (BS) is required to collect the channel gains seen by every user and to report the information to a central radio network controller (RNC). There are two issues with this requirement. First, the interferer's channel gain $G_{i,u_n}, i \neq n$ are not readily available at the user equipment (UE). In practical wireless systems, the user equipment (UE) does not estimate the channel gain from every surrounding cell other than its serving cell. Instead, the feedback from the UE to the BS is based on the aggregate interference from all sources of interference in the network. Requiring each UE to decode the pilot symbols of all surrounding cells is not practical due to the signal processing overhead.

Second, the radio network controller needs to process a large amount of information in

order to compute the optimal solution and distribute it back to the concerned cells in real time. This near real-time signaling may not be feasible in large cellular networks where the controller covers a wide area with hundreds of base stations. The channel conditions may even change before the optimal solution is computed thus rendering the optimization useless. Actually, the mere existence of a central controller implies a complex hierarchical network architecture which the industry strives to avoid. For this reason, the RNC entity has been totally eliminated in the 4G cellular standard 3GPP LTE/LTE-A and a flat, distributed architecture has been adopted. Practical interference mitigation solutions have to be distributed with minimal information exchange between base stations.

### *2.1.3 Related Work on Inter-cell Interference Coordination in LTE Networks*

Despite the complexity of the multi-cell resource allocation problem, it has been investigated extensively in recent years with a variety of mathematical formulations for various settings. While the utility maximizing power control problem in (2.3) remains of a particular theoretical interest, it is often used as a starting point for simpler, reduced problems that may yield sub-optimal, but practical solutions. Most of the current research activities are focused on simpler, distributed Inter-cell Interference Coordination (ICIC) techniques in the context of 4G LTE networks [10–12]. In LTE, distributed coordination can be realized using the new X2 interface between base stations. Also, the advent of OFDMA in 4G cellular systems has brought new degrees of freedom to the system designer. As we will discuss in the next section, the slotted, channelized nature of OFDMA allows dividing the channel into a time-frequency resource grid where each block or "tile" represents a time slot in one subchannel. Adjacent cells can be assigned different *Resource Blocks* (RBs) at different times, a process essentially equivalent to the classical channel assignment but on a "finer" scale.

Most of the proposed algorithms are based on partitioning of the resource blocks among cells in some way or another. A widely studied technique is called Fractional Frequency Reuse (FFR) which has been adopted by the industry. In FFR, the cell is divided into a center which is allowed to use the full spectrum and an edge which is restricted to a

fraction of it. The edges of two adjacent cells are allocated different parts of the spectrum to reduce the interference suffered by cell-edge users. FFR can be combined with static power control to improve the performance further. Also, inter-cell coordination based on coarse-time measurements has been proposed to reduce interference at the cell edge. A survey of different resource allocation algorithms for OFDMA networks is presented in [13].

### *2.2 The Case for Distributed Interference Mitigation in Femtocells*

Since femtocells evolved from cellular networks, it seems logical to extend the above centralized multi-cell resource allocation framework to handle femtocell interference as well. As femtocells represent a new "tier" overlaid over the conventional macro network, a multi-tier view to the multi-cell resource allocation is needed to handle the co-tier and cross-tier interference. This is indeed the approach followed by most researchers who studied femtocell interference. Numerous heuristic algorithms have been proposed for mitigating inter-cell interference either by power control only, channel allocation only or joint channel and power allocation, see for example [7, 14–17].

However, almost all the proposed ideas assume a full-buffer traffic model in every femtocell and ignore the user activity factor. This saturated buffer assumption, while justified in macro cellular networks, is simply not valid in femtocells. By definition, a femtocell is designed to serve very few users with at most one or two users being active at any instant. Even in a cluster of co-located femtocells, e.g. in apartment buildings, the chance that all users are active in the same time is very small. This is in contrast to macrocells where the high traffic aggregation from many users makes the system insensitive to the activity of a single user. *Therefore, the randomness of the user activity must be accounted for through appropriate traffic models in any proper study of the femtocell interference problem.*

Unfortunately, accounting for the traffic dynamics in centralized interference mitigation strategies is practically infeasible. In the network architecture of femtocells, there is no central node suitably positioned to play a centralized role other than the Femto gateway (Figure. 1.1). However, the Internet-based backhaul (VPN tunnel) between the femtocells and the gateway cannot be used for fast tracking and controlling the highly dynamic (due to user activity) interference situation in femtocells. The latency of this link is much larger

than the resource allocation time slot which is in the order of 1 ms in 3GPP LTE. This prohibits any form of centralized real-time resource allocation, i.e. on a slot-by-slot basis or even on a packet-by-packet basis.

Moreover, distributed interference coordination mechanisms that rely on explicit message passing between adjacent cells become very challenging for two reasons. First, there is no practical way to setup and maintain the required control channels between neighboring femtocells because these devices can be turned off or relocated by the customer at any time. Second, any distributed interference coordination algorithm certainly requires multiple rounds of communication exchange in order to converge to an optimal solution. The slow convergence time makes it infeasible to coordinate interference on a packet-by-packet basis.

Therefore, the current industry practice and research trend are focused on the concept of Self-Organizing Networks (SON) to handle femtocell interference [18]. In SON-enabled framework, each femtocell runs a self-optimization function with the help of a central Femtocell Management Server (FMS) which stores a database of location-specific network configuration such as the list of available channels and the list of macrocells and other femtocells in a specific geographic area. The self-optimization function utilizes local measurements of the interference received from surrounding cells and database information to optimize the femtocell channel selection and/or transmit power to provide a target coverage while minimizing interference to the rest of the network. This function is run when the femtocell is first activated and repeated afterward either periodically or in response to network-triggered events such as the addition of a new femtocell in the neighborhood. Most self-optimizing femtocells in current commercial deployments are based on this scheme [19]. Some SON functionality permits explicit coordination between adjacent cells using advanced inter-cell links such as the X2 interface in 3GPP LTE.

Because the self-optimization process involves RF scanning for interference measurements and requires control-plane signaling and back and forth communication between multiple network elements, the re-configuration cycle is slow and runs on a large time scale, in the order of hours and days. It is not expected to be run on the time scale of a user session, let alone a packet-by-packet time scale. This means the femtocell configuration,

e.g. channel selection, bandwidth and/or transmitter power, and hence the average transmission rate, would remain static throughout a single user activity session even if no other users are active. Therefore, radio resources may be wasted especially in dense networks of femtocells with bursty packet arrival models. In Chapter 3, we investigate the use of an idealized power control to mitigate interference in femtocells and compare its packet-level performance with simple random access protocols under regular and bursty packet arrival processes.

### *2.3 Proposed Approach*

In this research, we will take a different approach to the femtocell interference problem. We note that at the packet level, handling interference between femtocells is essentially a local area multiple access (MAC) design problem similar to MAC design in IEEE 802.11 WLANs (WiFi) but with the added flexibility of OFDMA. By definition, the role of any MAC algorithm is to provide non-interfering communication channels for each active transmit-receive pair in the network. Therefore, we will look at the femto-femto interference problem from a multiple access perspective that allow us to factor-in the randomness of the user activity using packet traffic models. Since femtocells are unsaturated, the appropriate figure of merit is the delay induced by the MAC algorithm at the packet-level instead of the throughput of the algorithm. This is in contrast to the multi-cell resource allocation framework which only looks at the capacity from a PHY-layer perspective and ignores the packet-level delay performance.

Because a femtocell is designed to serve very few users with at most one or two users being active at any instant, the generated traffic load at the cell level is often bursty with low duty cycles. We can view a cluster of co-located single-user femtocells as an ad-hoc network of interfering transmit-receive pairs competing for channel access without a central arbitrator. In such situations, using random multiple access protocols to mitigate interference in a distributed way is a viable approach if the MAC algorithm is designed properly. In a random access MAC, users contend for the channel, rather than being allocated a resource. Conflicts are then resolved using a collision resolution protocol. Random access MAC protocols have been studied for many years and have been successfully implemented in

many wireless standards such as the popular IEEE 802.11 (WiFi). In the simplest random access protocol, which is known as Aloha, the user transmits a packet at will without regard to other users. If the packet is not acknowledged by the receiver, the user backs off and tries again after a random backoff time. In Carrier Sense Multiple Access (CSMA) protocols, the user "listens" to the channel before transmitting a packet in order to avoid collisions whenever possible. Collisions can still occur and are then resolved using some backoff algorithms. Effectively, the random access MAC schedules the users into non-interfering time slots in a distributed way without explicit coordination.

Our point of departure from the vast amount of research on wireless random access MAC protocols such as WiFi, lies in OFDMA. The fine-grained time-frequency channelization of OFDMA enables the design of flexible multiple-access strategies to accommodate many users with varying applications, data rates and QoS requirements. In fact, we can exploit the flexibility of OFDMA and time synchronism to design more efficient random access MAC protocols than classical protocols. For example, nodes in an OFDMA network can in principle contend for subchannels and multiple "winning" nodes can transmit simultaneously on different subchannels without collisions. This provides an additional degree of freedom in the random access scheme by allowing *contention resolution over the frequency as well as over the time dimension.*

In the following sections, the concept of OFDMA-based Random Access MAC is defined and contrasted with the concept of multi-channel MAC from prior wireless MAC literature. Then, we discuss the use of these MAC protocols for interference avoidance in OFDMA femtocells and describe an appropriate MAC models for future investigation.

### *2.3.1 OFDMA-based Random Multiple Access*

OFDMA combines the desirable features of both FDMA and TDMA by allowing multiple users to transmit in the same time slot over different sets of subcarriers. High bandwidth flexibility is achieved by allocating more subcarriers to users demanding higher data rates on a slot-by-slot basis. Typically, subcarriers are grouped into subchannels in order to reduce the complexity of the system. For example, the smallest unit of allocation in LTE is called

a *Resource Block* (RB) comprising of 12 subcarriers (180 KHz) for the duration of 7 OFDM symbols (0.5 ms) as depicted in Figure 2.2.

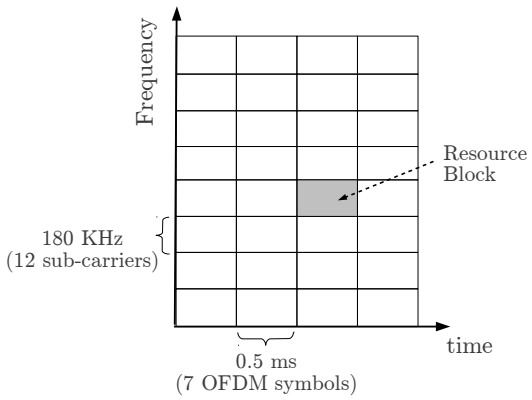


Figure 2.2: LTE Resource Grid (OFDMA channelization)

The time-frequency resource grid in OFDMA (Figure 2.2) can be thought of as a slotted multi-channel system. The grouping of sub-carriers into Resource Blocks (RBs) divides the channel bandwidth into equally-sized subchannels which can be accessed on a slot-by-slot basis. We will use the terms "subchannel" and "channel" interchangeably to denote one RB in the frequency domain. Because these subchannels are orthogonal, the channel access is a Multi-Channel MAC problem. Numerous contention-based algorithms have been studied for Multi-channel MAC mostly in the context of ad-hoc and wireless mesh networks [20]. In multi-channel ad-hoc networks, the user usually contends for a single channel and then transmits one packet over that channel. The advantage is that when the number of users is much greater than the number of channels, the multi-channel MAC can distribute the contention load and hence reduce collision rate.

Two unique features distinguish OFDMA from traditional multi-channel systems proposed in the context of wireless ad-hoc and mesh networks. First, channel switching is instantaneous because by definition OFDMA allows control over which subchannels are accessible and which are not on a slot-by-slot basis. This type of channelization being entirely realized in the digital processing domain marks a big leap over traditional multi-channel wireless systems where the transmitter has to switch its radio from one frequency carrier

to another, thus incurring a significant time penalty. The second feature is the ability to transmit and receive over a variable number of subchannels simultaneously with a single radio. For example, a packet in LTE is segmented or concatenated with other packets to fit into a variable-size block that is mapped to a variable number of allocated subchannels. This enables the random MAC protocol to adapt dynamically to various network loads and traffic scenarios.

### *2.3.2 Related Work on OFDMA-based Random Access*

Several random access algorithms have been proposed to exploit the multi-channel nature of OFDMA. Opportunistic multi-channel Aloha (discussed in [21] among others) attempts to exploit the multi-user diversity by adapting the transmission probability in each subchannel to its corresponding channel quality. However, the first work that studied contention resolution over the frequency domain in OFDMA appeared in [22]. It proposed what we term OFDMA-Aloha - a fast retrial algorithm for frequency-domain backoff. In OFDMA-Aloha, instead of waiting for a random backoff period after a collision, the user tries another (randomly selected) subchannel immediately subject to a maximum retry limit. The basic idea of OFDMA-Aloha was extended to OFDMA-based CSMA/CA systems in [23]. In OFDMA-CSMA/CA, the user first senses all subchannels and randomly selects one subchannel from the set of idle subchannels for transmission. Then, it backs off for a random Collision Avoidance (CA) interval whose length is proportional to the number of idle subchannels sensed in every slot. There are other follow-up papers on [22, 23], but none of the published work had investigated packet delay which we believe necessary in bursty low-duty cycle scenarios such as residential femtocell networks. We will provide a detailed analysis of the packet delay of OFDMA-Aloha in Chapter 4.

Other works tried to extend this frequency diversity to reservation-based random access protocols such as Reservation-Aloha (R-Aloha). In LTE, the Random Access Channel (RACH) is used by the UE to synchronize with the network and for initial registration with the base station. Once the UE is successfully admitted, the BS schedules all subsequent transmission toward the UE. LTE utilizes an enhanced form of R-Aloha in which each

UE transmits a random preamble sequence on one of few resource block (RBs) dedicated for RACH in each radio frame and conflicts are resolved by the BS in multiple rounds. A straightforward extension of R-Aloha to OFDMA uses frequency-domain backoff, e.g. OFDMA-Aloha, to expedite contention resolution of the reservation request. This has been proposed for distributed scheduling in OFDMA and WiFi networks in several papers [24–28]. Other reservation-based protocols employ special reservation requests (physical layer preambles) that are mapped to different OFDMA sub-channels to send the required reservation information [29, 30]. However, most of these reservation-based protocols are designed for uplink channel access within a single cell where a base station can assist in resolving conflicts between contending nodes. None of the published work so far has considered the use of OFDMA-based random access whether conventional or reservation-based for handling inter-cell interference in dense femtocell networks.

## *2.4 Integrating OFDMA-based Random Access in LTE Femtocells*

Random access protocols are not used today in cellular networks such as LTE except for initial channel acquisition and synchronization. However, the scheduling and interference coordination entity at each femtocell's BS can be configured to implement any channel access policy including random access protocols to handle inter-cell interference. To see how a random access protocol fits in the standard architecture, we briefly describe the downlink protocol stack in LTE and the relevant functionalities.

### *2.4.1 Downlink Protocol Stack*

In LTE, IP packets travel through four layers before they are transmitted over the air as shown in Figure 2.3. At the Base Station side, the Radio Link Control (RLC) layer receives packets (or PDUs) from the Packet Data Convergence Protocol (PDCP) which is responsible for IP header compression and ciphering. The RLC layer is responsible for the segmentation and concatenation of packets into RLC PDUs, whose sizes vary from one slot to another depending on the resource allocation offered by the MAC layer. The main function of the MAC layer is resource allocation and scheduling in addition to implementing the Hybrid-ARQ (H-ARQ) process. The PHY layer performs coding, modulation and resource mapping

as assigned by the MAC scheduler [31]. The reverse processing occurs at the User Equipment (UE) side.

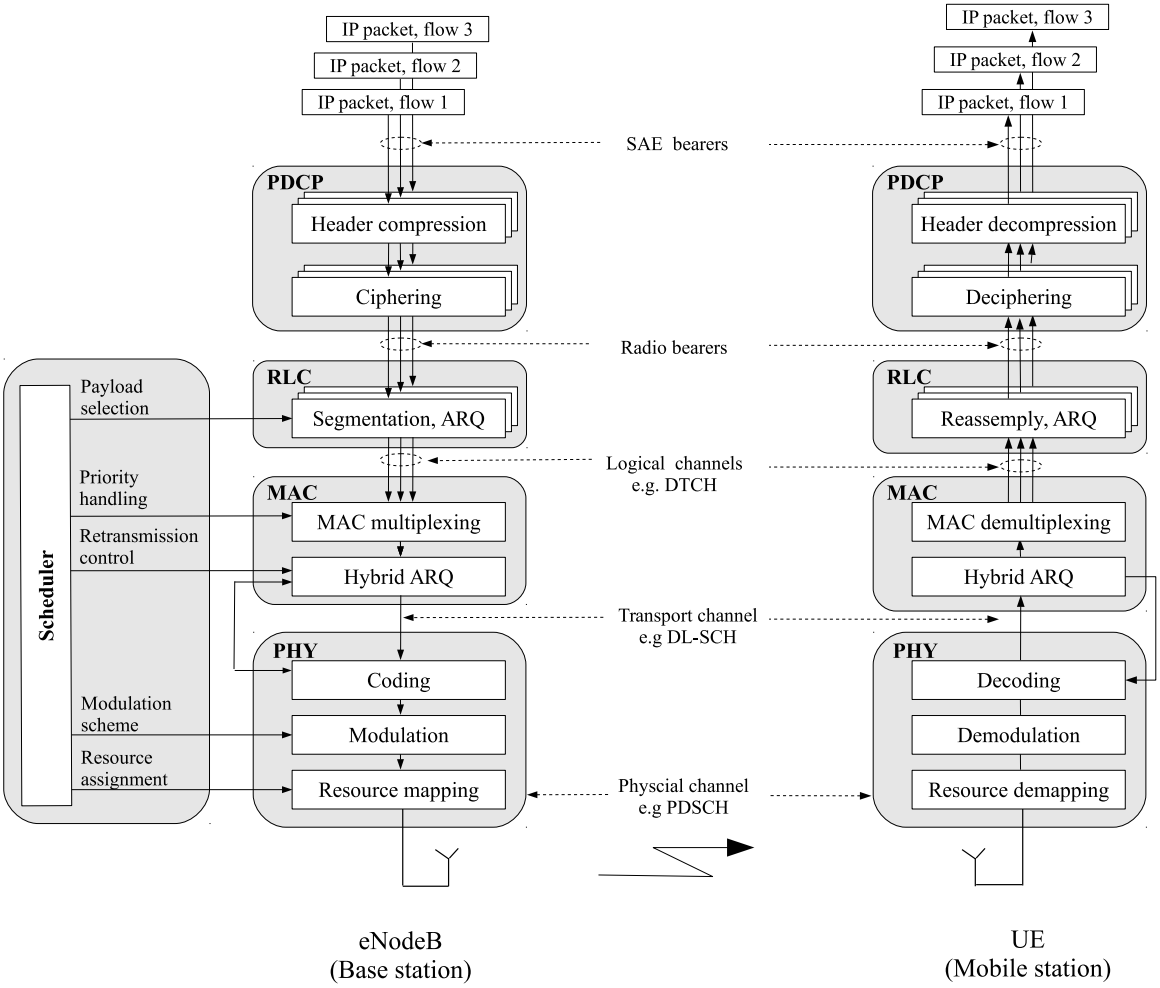


Figure 2.3: LTE Downlink Protocol Stack, edited from [31]

### *2.4.2 Logical Channels Mapping*

The MAC layer provides services to the RLC layer in the form of *Logical channels* which are defined by the type of information they carry. *Control* channels are used to carry control and configuration information and *Traffic* channels are used for user data. Table 2.1 lists the logical channels defined in LTE [31]. The main logical channel that is used for transmission

of all unicast downlink user data is called the Dedicated Traffic Channel (DTCH).

Table 2.1: Logical Channels in 3GPP LTE [31]

| Channel | Name | Purpose |
|---|---|---|
| BCCH | Broadcast Control Channel | broadcast of system information |
| PCCH | Paging Control Channel | paging of terminals whose cell-level location is unknown |
| CCCH | Common Control Channel | transmission of common control information |
| DCCH | Dedicated Control Channel | transmission of control information to/from a terminal |
| DTCH | Dedicated Traffic Channel | transmission of user data to/from a terminal |
| MCCH | Multicast Control Channel | transmission of multicast control information |
| MTCH | Multicast Traffic Channel | downlink transmission of multicast services |

Each UE in the cell is associated with one DTCH and multiple DTCH channels are multiplexed by the MAC layer into one *Transport channel*. Transport channels are defined by how the information is transmitted over the radio interface. In LTE, data transmission on a transport channel is organized into Transport Blocks (TB) which are transmitted every Transmission Time Interval (TTI). Associated with each transport block is a Transport Format (TF) specifying the transport-block size, the modulation and coding scheme (MCS), and the resource blocks and antenna mapping. Therefore, a transport channel is defined by a set of resource blocks (RBs) and an associated Transport Format (TF). Table 2.2 lists the transport channels defined in LTE and Figures 2.4 and 2.5 illustrate the mapping of logical channels to transport channels [31].

A *Physical channel* corresponds to the set of time-frequency resources assigned for the transmission of a particular transport channel and each transport channel is mapped to a corresponding physical channel as shown in Figures 2.4 and 2.5. In addition, there are physical channels that are not mapped to transport channels but are used to carry Downlink Control Information (DCI) needed for proper reception of the data transmission. Figure 2.6 illustrates the physical channels that are involved in the downlink data transmission [31].

1. The Physical Downlink Shared Channel (PDSCH) is the main physical channel used

Table 2.2: Transport Channels in 3GPP LTE [31]

| Channel | Name |
|---|---|
| BCH | Broadcast Channel |
| PCH | Paging Channel |
| DL-SCH | Downlink Shared Channel |
| UL-SCH | Uplink Shared Channel |
| RACH | Random-Access Channel |
| MCH | Multicast Channel |

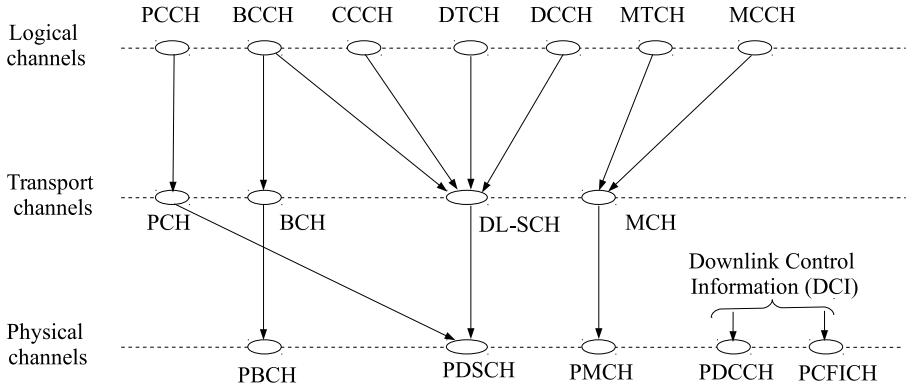


Figure 2.4: LTE Downlink channels mapping, edited from [31]

for unicast data transmission.

2. The Physical Downlink Control Channel (PDCCH) is used for downlink control information (DCI), mainly scheduling decisions (resource assignment), required for reception of PDSCH, and for scheduling grants for the uplink channel.
3. The Physical Control Format Indicator Channel (PCFICH) is a channel providing the UE with information about the format of the above downlink control channel (PDCCH).
4. The Physical Uplink Control Channel (PUCCH) is used by the UE to send uplink control information (UCI) such as hybrid-ARQ acknowledgements, channel-state reports

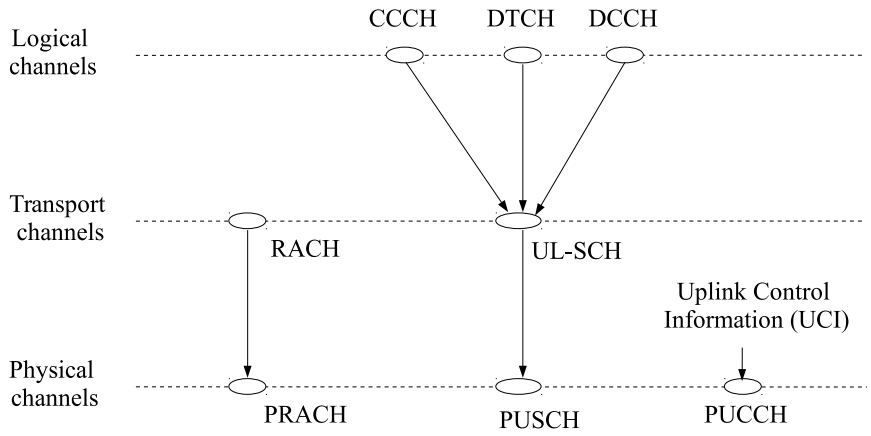


Figure 2.5: LTE Uplink channels mapping, edited from [31]

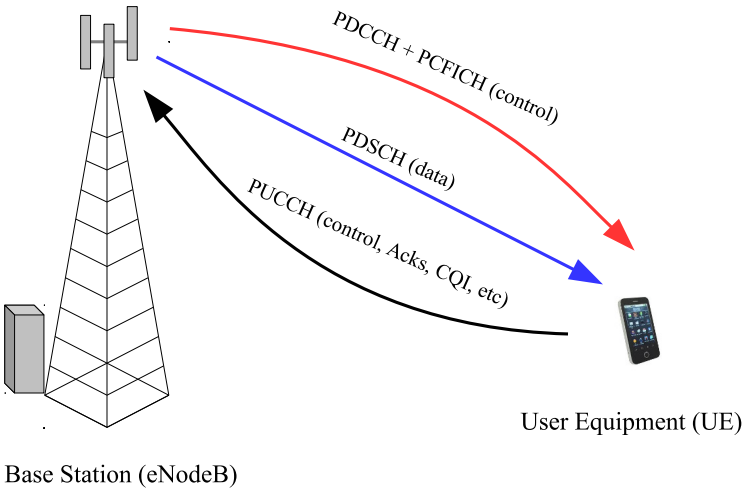


Figure 2.6: LTE Physical channels involved in downlink data transmission

(CQI), and uplink scheduling requests.

The mapping of these channels to physical resource elements is controlled by the base station and is done on a 1 ms sub-frame basis. Figure 2.7 illustrates the split of the subframe into (variable-sized) control (PDCCH and PCFICH) and data (PDSCH) regions. The uplink control information (PUCCH) is mapped to two or more (depending on the type of UCI) resource blocks (RBs) at the edge of the system bandwidth in every 1 ms subframe as

illustrated in Figure 2.8.

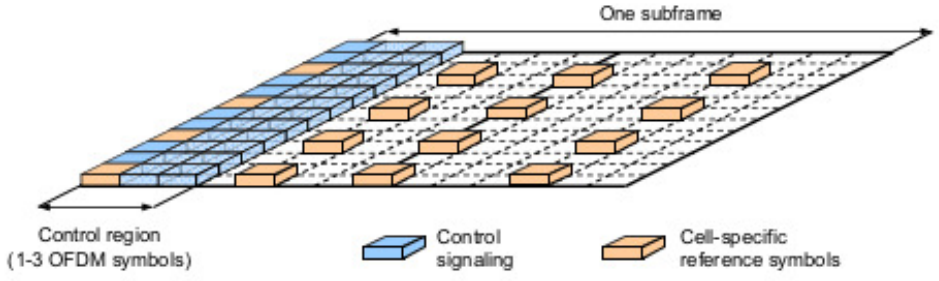


Figure 2.7: LTE Downlink time/frequency grid illustrating the split of the subframe (1 ms) into (variable-sized) control and data regions, from [31]

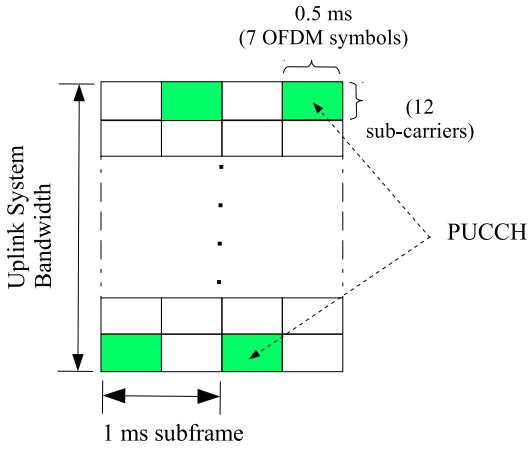


Figure 2.8: The PUCCH is mapped to two Resource Block (RB) at the edge of the uplink system bandwidth in each subframe

### *2.4.3 Reference Signals and Channel Quality Feedback*

In standard 3GPP LTE systems, 4 OFDM symbols in every Resource Block (RB) are reserved for cell-specific reference signals (CS-RS) as shown in Figure 2.7. These reference signals are pseudo-noise random sequences that are derived from the cell identity (CID) and transmitted with a pre-configured, constant power [31, chapter 10]. The location (in

frequency) of the CS-RS depends also on the cell ID and can be orthogonalized between up to six adjacent cells. The UE obtains the configuration of the CS-RS (location and power) from the system information which is transmitted on the broadcast (BCH) channel. Reference signals enable several important functions. First, they are used by the UE to estimate the channel for coherent demodulation at the receiver. Second, they are used to compute the average received power from a particular cell to facilitate handoff decisions. The third key function is the computation of the channel quality indicator (CQI) which is fed into the Adaptive Modulation and Coding (AMC) and channel dependent scheduling algorithms.

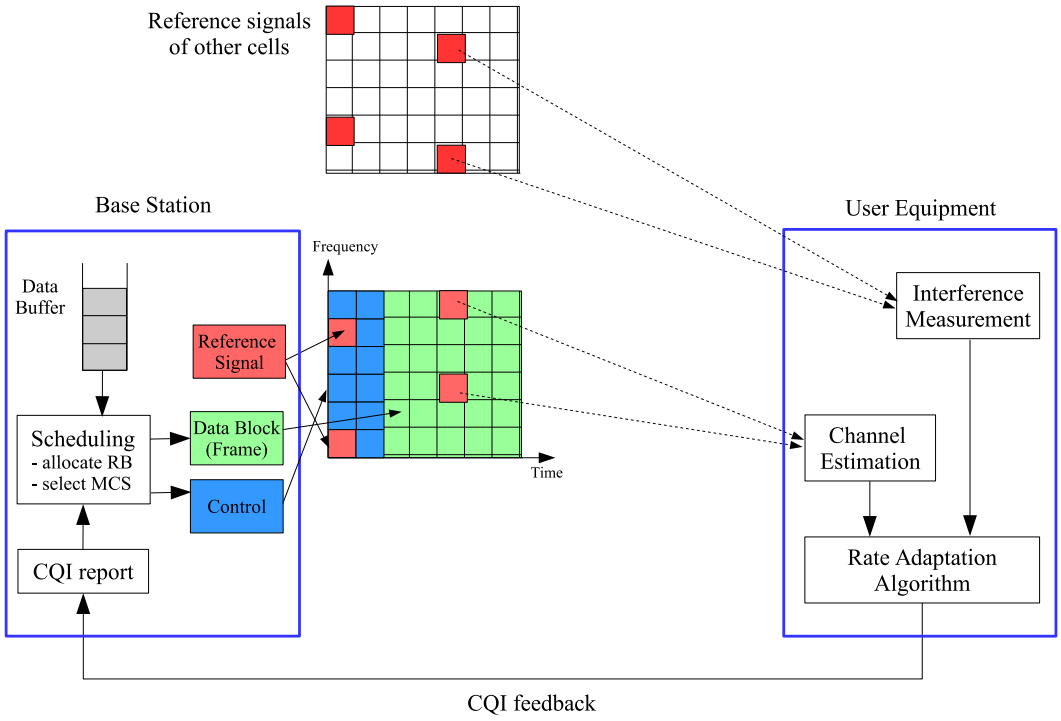


Figure 2.9: Link adaptation and CQI feedback loop in LTE

Adaptive Modulation and Coding (link adaptation) is used on all data channels in LTE to improve the utilization of the radio resources. The link adaptation process is illustrated in Figure 2.9. The UE uses reference signals to measure the quality of the channel which depends on the random channel fading and interference from other cells. Then, the mea-

surements are used by a rate adaptation algorithm that compute a channel quality index (CQI). The CQI basically indicates the maximum rate (modulation order and coding rate) that it can receive with a Block Error Rate (BLER) less than 10% [31, Chapter 13]. CQIs are reported periodically and upon request from the BS. Also, some CQI reports cover the entire downlink bandwidth (wide-band CQI) and some are for specific sub-bands. The BS scheduler then uses the CQI report to determine the Modulation and Coding Scheme (MCS) for each Resource Block assigned to the UE. Note that the rate adaptation algorithm is an internal UE function that depends on the receiver complexity, hence it is left for vendor implementation and not specified by 3GPP standards [31].

### *2.4.4 Downlink Packet Flow and ARQ*

A sample data flow is shown in Figure 2.10. In each slot, the BS scheduler assigns a group of resource blocks (RBs) to be used for each user in the next Downlink Shared Channel (DSCH) transmission. The size of the Transport Block (TB), which is a coded MAC frame, is computed from the number of assigned RBs and the modulation and coding scheme (MCS) selected for each RB. Then, the RLC layer segments/concatenates IP packets into multiple RLC PDUs whose combined size is equal to the TB size plus the MAC header. The MAC layer combines the RLC PDUs, adds a header and creates a TB for the H-ARQ process. After coding and modulating the TB with the selected MCS, the PHY layer maps the modulation symbols over the subcarrier groups (RBs) assigned by the MAC scheduler. Information about the assigned RBs and their MCS is conveyed to the receiver over the downlink control channel (PDCCH). If the TB is decoded correctly at the receiver, the RLC PDUs are extracted from the frame, reassembled and IP packets are delivered up the stack.

To handle frame errors, there are two layers of ARQ in LTE. At the lower PHY/MAC layer, a fast Hybrid ARQ (HARQ) process runs between the transmitter and the receiver and uses Incremental Redundancy or Chase Combining to correct transmission errors. After three unsuccessful attempts (the original transmission plus two retransmissions), the TB is discarded and the HARQ process frees itself for a new TB [31]. The HARQ protocol

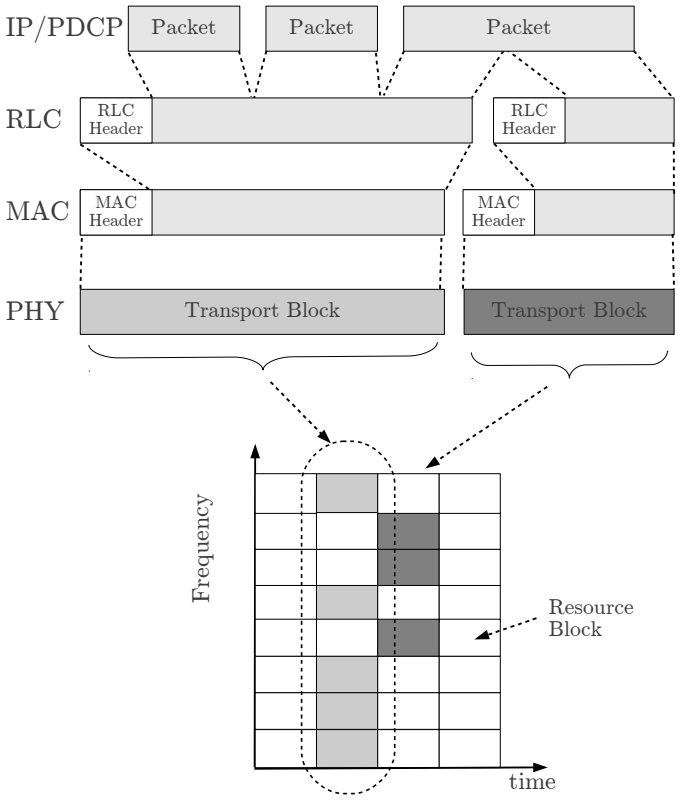


Figure 2.10: LTE data flow and resource mapping

is designed to correct most transmission errors caused by noise and unpredictable channel variations, but cannot handle persistent errors caused by collisions or excessive interference. To provide a reliable error-free radio link, another ARQ at the Radio Link Control (RLC) layer is used in LTE. The RLC Acknowledged Mode (AM) is a window-based selective repeat ARQ designed to handle all transmission errors due to channel, interference or imperfect feedback loop. Retransmissions at the RLC layer may be based on timeouts, explicit status reports from the receiver or triggered immediately after a HARQ failure. The LTE standard specifies the signaling and the format of the RLC protocol, but leaves the number and the timing of retransmissions to vendor implementation [31].

### *2.4.5 How a Random Access can be realized?*

In LTE, access to the channel is fully controlled by the BS scheduler. The scheduling algorithm is an internal function of the BS and hence it is entirely a vendor implementation and not specified by 3GPP standards [31]. Generally, the main function of the scheduler is fair allocation of resource blocks among different users within the cell, i.e. UE's associated with the BS. In addition, the scheduler can also perform interference coordination with adjacent cells via the X2 interface. For example, the scheduler may leave some RB's unassigned in order to be utilized by adjacent cells, a technique called Almost Blank Subframe (ABS) scheduling [31, Chapter 15]. In an interference-limited dense femtocell network, the scheduling algorithm along with the feedback loop of both the HARQ and RLC protocols can be adapted to implement a random access algorithm that coordinates channel access between adjacent femtocells. In this scenario, the BS could, for example, transmit at maximum rate on the downlink channel (PDSCH) and signals the highest MCS index to the UE on the associated downlink control channel (PDCCH), see Figure 2.6. Next, the HARQ entity awaits for the ACK/NACK from the UE on the associated uplink control channel (PUCCH). A HARQ error would then be taken as an indication of a collision caused by interference from a nearby cell. The scheduler can then delay the RLC retransmission for a random number of slots according to the employed random access protocol.

## *2.5 Simplified MAC Model for OFDMA Femtocells*

In the OFDMA-based random MAC algorithms proposed in the literature, the user is not allowed to access more than one subchannel at a time because the target is to distribute the offered load (attempt rate) over the available subchannels. The implicit assumption here is that the number of users is greater than the number of subchannels. The packet-level MAC model in this case is the traditional multi-channel MAC with *Single Access* where exactly one packet is transmitted over a single subchannel. This MAC model is not suitable for OFDMA femtocells where the number of subchannels is significantly greater than the number of active users. Hence, an important design criteria in this case is to allow the user access to as many subchannels as possible to maximize bandwidth utilization. The MAC

frame in OFDMA is a single coded block of bits (a.k.a. Transport Blocks in LTE) mapped to multiple subchannels. Since the user can initiate transmissions over multiple parallel subchannels, we need to define the notion of collision for such parallel transmissions. This motivates us to propose the *Bulk Access* model to facilitate the design and analysis of random multiple access in OFDMA femtocells.

### *2.5.1 Bulk Access Model*

The term "Bulk Access" is derived from *bulk service* queuing models where a group of customers are served together in bulks or "batches" by a single server [32]. The distinction from multi-server model is that the service time of each customer in the bulk is not independent, i.e. they are served together as one unit. This is precisely how OFDMA handles packet transmission over multiple subchannels. The information bits which are transmitted over multiple subchannels are actually treated as a single code block in the physical processing chain.

From Figure 2.10, we see that the Transport Block (TB) is physically mapped into multiple small resource blocks (RBs), one per subchannel. These RBs however, are only a physical layer realization and cannot be processed independently. If one RB is corrupted due to strong interference in the corresponding subchannel, the entire TB is deemed lost and must be retransmitted. If the H-ARQ fails to transmit the TB successfully, it deletes it from its buffer and notifies the upper layers. In RLC Acknowledged Mode (AM), RLC PDUs must be delivered reliably using RLC-level ARQ. In the next available transmission opportunity, lost RLC PDUs go through the same segmentation/concatenation process, possibly with a different TB size due to the varying allocation size provided by the MAC scheduler. Such sophisticated cross-layer interaction does not lend itself to easy analysis of the multiple access protocol with existing tools. This is where the Bulk Access model offers a suitable theoretical framework.

Looking at Figure 2.10 in the bottom-up direction, each Resource Block (RB) may be conceived as a packet because essentially it contains coded information bits. Then assuming a link with maximum spectral efficiency (as often the case in femtocells), consider a fixed

RB size and ignore any MCS link adaptation. We can also match the packet size to that of the traffic model to abstract out the RLC segmentation/concatenation functionality. Therefore, we have a multi-channel MAC with Bulk Access where packets are served in bulks. Any collision in one packet renders the whole bulk corrupted and all packets must be retransmitted, possibly with other packets depending on the MAC decision. This model matches the bulk-service queuing system where a single server takes a number of customers, depending on available capacity, and serves them as a bulk [32]. The simplified Bulk Access model is illustrated in Figure 2.11.

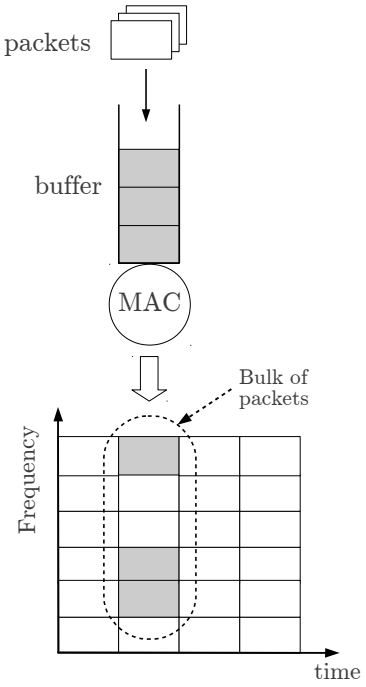


Figure 2.11: Multi-channel Bulk Access model for OFDMA femtocells

Note that this model does not impose any restriction on the choice of the actual MAC algorithm itself. It is the function of the MAC algorithm to specify the rule for how many and what subchannels are accessed in each time slot. The purpose of the proposed model is to allow the researcher to study different random MAC options under unsaturated traffic scenarios in which a stochastic packet arrival process is a necessary assumption. The model hides most of the complexity of the LTE protocol stack but captures the important MAC

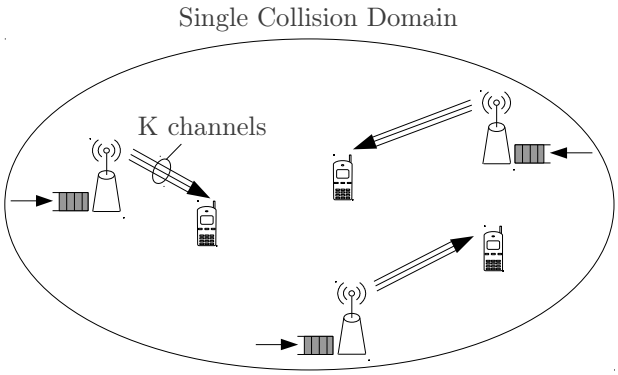


Figure 2.12: Multi-channel MAC model for OFDMA femtocells

interactions at the packet level. To this end, we present a case study for applying this model in evaluating two simple OFDMA-based random MAC algorithms.

### *2.5.2 Case Study: K-Aloha vs. Aloha*

For illustration purposes, we apply the Bulk Access model to compare two simple random access schemes in a small network of OFDMA femtocells with low traffic load. Consider a slotted OFDMA system with $K$ subchannels and $N$ femtocells, each serving exactly one user as depicted in Figure 2.12. Assume also that all femtocells are within a single collision domain so that only one packet can be successfully transmitted in each subchannel. Packets arrive to each user's queue according to a Bernoulli arrival process with probability $\alpha$ in every slot. According to our bulk access model, a group of packets is transmitted as a bulk where each packet occupy one subchannel for exactly one time slot. The bulk size $M$ is a random variable that depends on the current queue length $Q$ and is upper-bounded by $K$.

In low traffic scenarios, it is likely that the user does not have enough packets to occupy all $K$ subchannels in every slot, i.e. $M \ll K$. Looking at Figure 2.11, we see that we have some freedom in selecting $M$ out of $K$ subchannels to transmit the bulk of packets. In a random access network, the most natural choice is to select the $M$ subchannels randomly with the hope for reducing collisions. For example, suppose we have two users each having exactly one packet in his queue ($Q = 1$) and employing slotted Aloha MAC with transmission probability $p$. In this case, both users will create a bulk of size $M = 1$. If each user maps his

bulk starting at subchannel 1 (i.e. sequentially in order of channel index), then the collision probability is simply $p$. However, if the bulk is randomized over the $K$ subchannels, the collision probability is reduced to $p \times 1/K$ (for $M = 1$). Clearly, as the bulk size $M$ increases the collision probability increases until the gain from subchannel randomization diminishes at $M = K$.

We simulated the first scheme (Simple-Aloha) and the second randomized version (K-Aloha) in a network of $N = 12$ single-user femtocells with $K = 32$ subchannels and Aloha transmission probability $p = 1/N$. In both schemes, the user transmits a bulk of size $M = \min\{Q, K\}$ with probability $p$ in every slot. In Simple-Aloha, the bulk is always mapped to subchannels indexed 1 through $M$. In K-Aloha, $M$ subchannels are selected randomly out of the $K$ subchannels. The input traffic load is increased by increasing $\alpha$ of the Bernoulli arrival process. Figure 2.13 shows the packet service time which is the time needed to successfully transmit the Head of the Line packet (also known as access delay). Clearly, in high traffic load with high queue levels ($Q \geq K$) the difference between the two schemes diminishes as indicated by the results. However, significant reduction in the service time can be achieved in the low load region if we randomize the selection of channels, i.e. spreading the bulk over the frequency domain as in K-Aloha. Clearly, this kind of randomization is not the best option to utilize the frequency dimension, but it helps illustrate the bulk access model and opens up new ideas for random multiple access in OFDMA femtocells. We will use the Bulk Access model in the analysis of a new OFDMA-based random MAC algorithm in Chapter 5.

## *2.6 Classical Tools for Random MAC Analysis*

Before we conclude this chapter, it is necessary to briefly describe classical analytical approaches for evaluating the performance of MAC algorithms. Usually, the starting point of any new MAC algorithm is its throughput performance. In a slotted MAC system, the throughput (output rate) is the average number of packets transmitted successfully in a slot. The maximum throughput is reached when the network is saturated, i.e. every node has at least one packet in its buffer. On the other hand when the network is not saturated, some nodes may be idle and thus will not contend for channel access. In such cases, the

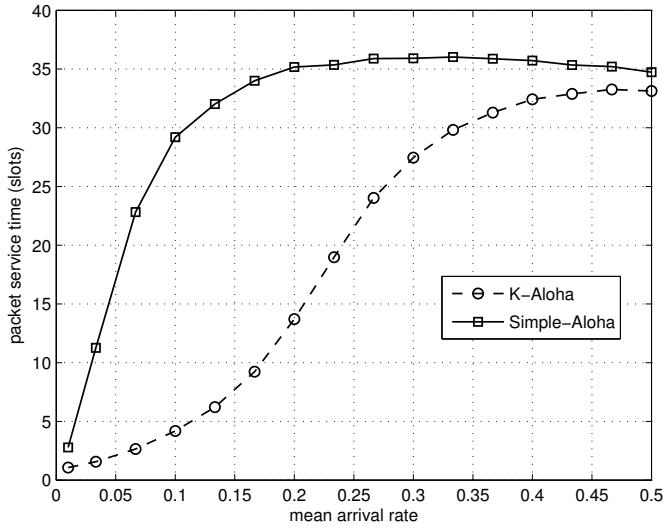


Figure 2.13: Packet service time (access delay) of K-Aloha vs. Simple-Aloha, $K = 32$, $N = 12$, $p = 1/N$

steady-state output rate (throughput) equals the input rate (offered traffic load) in stable operation. Therefore, the useful performance measure in unsaturated networks is the packet delay.

In the literature, there are two models for MAC analysis: the *Queue-free* model and *Queue-based* model. In the queue-free model, the system is modeled as an infinite population of buffer-less nodes which collectively generate traffic that can be considered equivalent to an independent Poisson source. Because there is no buffering, it is assumed that each newly arriving packet arrives at a new node. This model is the most-tractable and can approximate the throughput of finite networks with relatively large population. However, it fails to capture important performance measures such as the delay and stability of the MAC protocol.

In the queue-based model, the network is modeled as a finite number of buffered nodes each attached to a queue where packets arrive according to a random arrival process. A group of $N$ wireless nodes competing for channel access can be viewed as a multi-queue system as shown in Figure 2.14. However, the service of each queue is not independent from

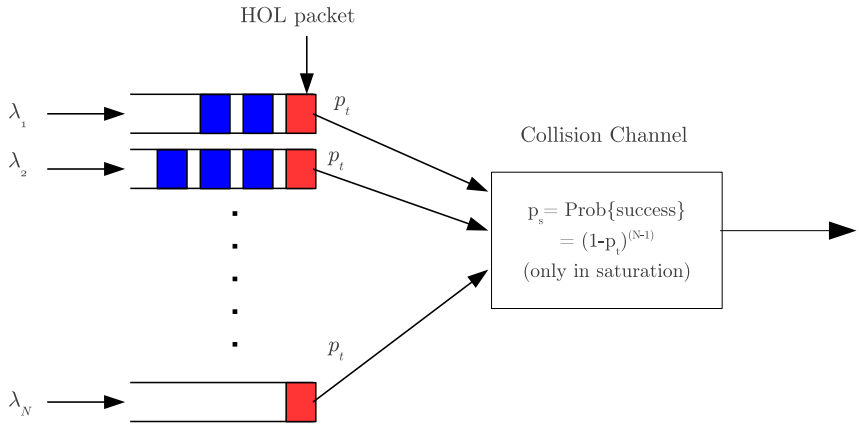


Figure 2.14: Multi-queue model: $N$ queues competing for a single shared channel

the other queues because of channel contention. The service time of the Head of Line (HOL) packet in one queue depends on the MAC algorithm and on the state of all other queues. For example, consider a simple slotted Aloha system where every active node attempts to transmit with probability $p_t$ in every slot. The success probability (and hence the service time) seen by the HOL packet depends on the number of active (non-empty) queues in the network which is by itself an unknown random quantity. The minimum success probability (worst case) is $p_s = (1 - p_t)^{N-1}$ which is only attained when all queues are not empty, i.e. in saturation.

Because of the interaction between these queues, classical queuing theory cannot be applied directly to solve for the service time and packet delay. Exact queuing analysis for coupled queues has not been achieved except for the case of two queues in Aloha systems [33] and in CSMA/CD systems [34]. In principle, Markov analysis can be applied to any multiple access system because the state transition probabilities can be derived from the packet generation probabilities specified by the arrival process and the packet transmission probabilities specified by the MAC algorithm. However, the resultant state space is very large and extremely difficult to handle. For example, the state of the system in Figure 2.14 is an $N$-dimensional vector $\mathbf{q} = (q_1, q_2, \cdots, q_N)$, where $q_n$ is the length of the $n$th queue. Therefore, a direct brute-force Markovian analysis of this system is intractable and

approximate analysis is necessary. Broadly speaking, there are four approximation strategies in the literature of queue-based MAC analysis, briefly described in the following.

### *2.6.1 Reduction of the State Space*

The first approximation approach is to summarize the state space into a manageable set using some assumptions [35–40]. One method considers the state of the system to be $(q, n)$ where $q$ is the number of packets queued at a "typical" tagged user and $n$ is the number of other busy users (with non-empty queues). Then using the symmetry and fairness properties of the MAC algorithm, it estimates the probability that a busy user has exactly one packet in queue which is needed for the transition probability. Another similar method considers two separate states one for the tagged user and one for the system summarizing the activity of all other users. A different approach that does not consider a specific tagged user uses the state $(n, Q)$ where $n$ is the number of busy users and $Q$ is the total number of packets in all queues. Then by symmetry and fairness of the MAC protocol, the $Q$ packets are assumed equiprobably distributed over the $n$ busy users to compute the transition probabilities. All these methods produce a system of non-linear (or linear in some cases) equations that can be solved numerically for the queue length and then the mean packet delay is computed from Little's Theorem.

### *2.6.2 Decomposition*

The second approximation approach decomposes the coupled queuing system into $N$ independent queuing systems and captures the interaction between the queues in the service time distribution [41–44]. Since all queues are assumed statistically identical in a homogeneous system, it suffices to derive the packet delay at one tagged queue. There are different ways of approximating the service distribution of the tagged queue. The simplest method is to assume that all other queues are saturated and focus on the dynamics of the tagged queue, which is clearly a poor approximation. A better approximation is to summarize the activity of the other users by $p_b$ the probability of being busy (non-empty queue). Because all queues are statistically identical, the busy probability of the tagged queue is also $p_b$ which

can be related to the busy period of the queue. This can be used to produce two coupled equations of the success probability and busy probability which can be solved numerically. We will explain this approach and use it in our analysis of OFDMA-Aloha in Chapter 4.

### *2.6.3 Equilibrium Point Approximation*

The equilibrium point analysis is a fluid-type approximation of the system [45]. It assumes the system is always at equilibrium where the input rate equals the output rate. The state of the system is taken to be $\mathbf{n} = (n_0, n_1, \cdots, n_L)$ where $n_i$ is the number of users in mode $i$, i.e. having $i$ packets in their queues, assuming maximum buffer size $L$. Then assuming equilibrium, the average number of users moving out of mode $j$ equals the average number of users moving into mode $j$. A set of non-linear equations can be obtained and numerically solved for the equilibrium point. If one solution is found, the system is stable and we can get the total system throughput and mean packet delay from Little's Theorem.

### *2.6.4 Idle-Busy Model*

In the previous three techniques, the goal is to provide an approximate analysis of an exact system model. A different approach is to provide an exact analysis of an equivalent system that approximates the unsaturated multiple access network. In this approach, the system is modeled as $N$ users alternating between two states: idle (or thinking) and backlogged (or busy). In each slot, an idle user generates a packet with some probability. Once a packet is generated, the user becomes backlogged (busy) and does not generate any new packets until the pending packet is successfully transmitted. With this model, it is enough to know the number of backlogged (busy) users to derive all necessary state transition probabilities of the Markov chain. Although there is no actual queuing of packets, this model approximate a lightly loaded network and can be used to study the impact of the different MAC parameters on the packet delay. This approach was first used by Kleinrock and Lam in the analysis of slotted Aloha in finite networks in [46] and since then has been followed by many researchers. We will adopt this approach in our analysis of the Exponential Back-off in Frequency in Chapter 5.

# Chapter 3

# OPTIMAL POWER CONTROL VERSUS RANDOM ACCESS

In Chapter 2, we formulated the general multi-cell resource allocation problem and discussed why it is not feasible for mitigating interference in femtocells. Then, we briefly described the Self-Organizing Network (SON) concept which is the current industry practice and research focus regarding femtocells [18, 47]. We argued in Section 2.2 that the slow re-configuration cycle in SON-enabled femtocells cannot cope with the highly dynamic interference situation caused by the random user activity. As a result, radio resources may be wasted especially in dense networks of femtocells with bursty packet arrival models. In this chapter, we investigate the loss in channel capacity by analyzing the average packet delay of an *ideal* power control scheme in comparison to the average packet delay that would result if the interference is mitigated solely by a simple random access protocol.

We will first describe how power control is used to mitigate femtocell interference in the SON framework . Then, we present a hypothetical, highly idealized power control scheme that represents a best-case situation for comparison with worst-case random access scheme. Next, we formulate the power allocation problem as a convex program and solve it for the optimal solution. Given the optimal power allocation and a stochastic packet arrival process, we carry out a queuing analysis to derive the mean packet delay and compare it to the corresponding delay of Aloha-based random access protocol.

## *3.1 Practical Self-Optimization Approaches in Femtocells*

The SON framework includes many functions such as Self Configuration, Self-Optimization and Self-Healing [47]. The relevant interference mitigation aspects are part of the self-optimization function. In OFDMA-based femtocells, the optimization mainly concerns the selection of the frequency subchannels (Resource Blocks) and the allocation of the transmitter power. Here we describe a typical self-optimization process for the downlink channel

in a dense network of 3GPP LTE femtocells. The process consists of three phases described in the following.

1. **Interference measurement phase:** In this phase, the femtocell BS with the help of database information from a central configuration server, tries to estimate the interference from all surrounding femtocells and macrocells either via the UE or directly using a Network Monitoring Module (NMM) to be explained below. In standard 3GPP LTE systems, 4 OFDM symbols at fixed locations in every Resource Block (RB) are reserved for cell-specific reference signals which are required for channel state information and other functions as explained in Section 2.4. The BS can request the UE to measure the reference signal power of the surrounding cells to estimate the interference from each cell. Because the self-optimization process is intended to be run in the background and the UE may not be always present in the cell to perform the required measurement, some vendors suggests the use of a Network Monitoring Module which runs as a User Equipment (UE) inside the femtocell BS [18, 19, 48]. The NMM listens to the downlink reference signals of the surrounding cells and, with the knowledge of the configuration (location and power ) of the reference signal of each cell from a central server, it can measure the corresponding interference. Note that because the NMM is located at the BS, this approach can only provide coarse, approximate measurements of the interference in a small area around femtocell BS.

2. **Optimization algorithm:** The femtocell BS runs a local algorithm based on the data obtained in the measurement phase to determine the best subset of subchannels and the transmit power for each subchannel. If a long-term optimization is sought, then the measurement reports can be sent to the centralized femtocell management server (FMS) for further optimization as suggested in [18]. An ideal optimization algorithm will be described later.

3. **Re-configuration phase:** The outcome of the optimization algorithm is a set of subchannels and their corresponding transmit power. Subchannel selection is enforced by the femtocell BS scheduler which can restrict the pool of available resource blocks

(RBs) accordingly. This does not require additional control signaling because the UE always reads the list of assigned RBs from the associated control channel at the beginning of each allocation interval as described in Section 2.4. The transmit power allocation is enforced by setting the cell-specific reference power accordingly. This step requires Radio Resource Control (RRC) signaling to inform the user equipment (UE) of the new reference signal power [49, chapter 6]. The transmit power of the data signal is derived from the reference signal power using an RRC-configured offset. Once the transmit power is configured, the UE estimates the channel quality in every time slot using the new reference signal power and reports the CQI value to the BS to be used to determine next slot's transmission rate.

The above self-optimization is driven by measurements derived from the downlink reference signals which are always transmitted in every subframe (always ON) regardless of the traffic load in each cell. This approach over-estimates the interference in lightly loaded scenarios as is the case in femtocells [31, Chapter 16]. As explained in the link adaptation process (Section 2.9), the power assigned to reference signals determines the CQI value reported by the UE. As a result, the transmission rate of the downlink data channel is effectively limited by the perceived interference from neighboring cells as seen from reference signals, even if no users are active in those cells.

This problem is not noticeable in macrocells as their data channels are expected to be always busy and their SINR can be predicted from reference signals. However, if data channels from neighboring cells have low duty cycles, then the SINR of the data signal may be much larger (due to less interference) than the SINR perceived from reference signals. The possibility of this situation is significant in dense networks of femtocells, especially with bursty traffic applications. In this chapter, we will try to quantify the impact of this mismatch between the traffic on data channels and the power control loop on the packet-level performance.

### *3.1.1 Ideal Power Control*

In order to simplify the analysis, we will assume an ideal power control protocol and ignore any measurement and control signaling overhead. Suppose there exists a central controller that measures the channel gains of all transmit-receive pairs in a cluster of femtocell and computes the optimal transmit power for each user sessions. Also, suppose the controller can re-compute the optimal power whenever the network topology changes due to for example user relocation or the addition/removal of femtocells. This hypothetical scheme can be considered *ideal* or a best-case scenario compared to the above self-optimization scheme for two reasons. First, it assumes availability of the channel gains as observed by the UE whereas in the measurement process described earlier, only the average interference observed at the femtocell BS is measured and is taken as an estimate of the interference experienced at the UE. Clearly, this estimate does not hold for indoor deployments where both the transmitter and the interferer can be located randomly anywhere around the receiver. Second, it assumes global channel information is available for the optimization algorithm. In the above SON-based self-optimization, only local channel information is collected and the optimization is run with a localized view of the interference situation. Therefore, we can consider the result obtained from this ideal scheme as the best that can be obtained from any practical self-optimization algorithm.

## *3.2 System Model and Assumptions*

Consider a cluster of residential OFDMA femtocells in a dense apartment building as illustrated in Figure 3.1. Because we are interested in mitigating the inter-cell interference, we assume each cell serves only one user and we implicitly account for the multiple user case by increasing the traffic load in each cell. There are $N$ active users corresponding to $N$ interfering transmit-receive pairs. The UE and BS are deployed randomly in each cell without coordination among cells. Denote by $g_{ij}$ the average channel gain between BS $i$ and UE $j$ and by $P_i$ the transmit power allocated for each BS $i$ where $i, j = 1, \cdots, N$. We make the following assumptions that characterize a typical femtocell scenario.

1. The femtocell network is assigned a dedicated channel which implies no interference

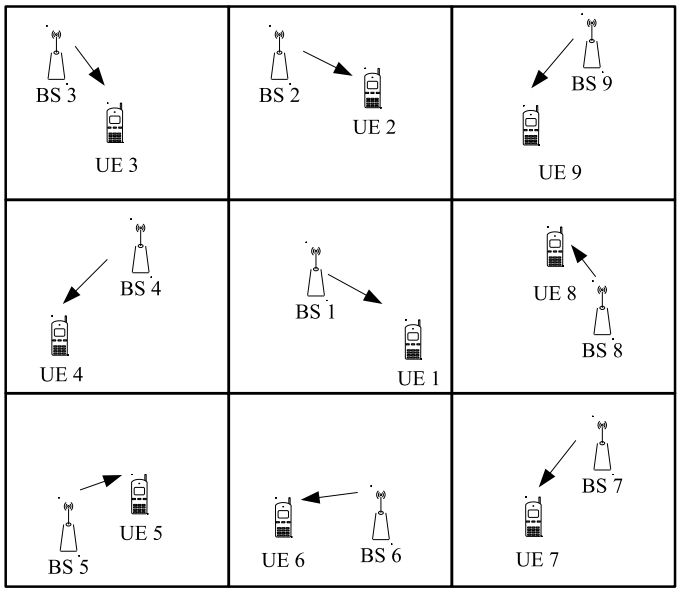


Figure 3.1: Dense residential femtocell network

from macrocells is considered. This is justified because the current 3GPP standard interference mitigation approaches are mostly based on orthogonal resource partitioning between macrocells and femtocells either in frequency by assigning dedicated channels or in time by using Almost Blank Subframe (ABS) scheduling [50], [31, Chapter 16]. The focus of this research is on inter-cell femto-to-femto interference on whatever orthogonal resource blocks assigned to them.

2. The number of active session $N$ (transmit-receive pairs) remains fixed in one optimization cycle. However, during each session packets arrive randomly according to an independent arrival process for each session (to be described later).

3. The total channel bandwidth $W$ is fixed and the maximum spectral efficiency is $C_{\max}$. The achievable spectral efficiency of each link $i$ with $\text{SINR}_i$ is given by the following

continuous rate adaptation function:

$$C_i = \begin{cases} 0\ , & \mathrm{SINR}_i \leq \Gamma_{\min} \\ \beta \log_2(1+\mathrm{SINR}_i)\ , & \Gamma_{\min} < \mathrm{SINR}_i \leq \Gamma_{\max} \\ C_{\max}\ , & \mathrm{SINR}_i > \Gamma_{\max} \end{cases} \tag{3.1}$$

where $\beta$ is some factor to capture the implementation gap to capacity. The thresholds $\Gamma_{\min}$ and $\Gamma_{\max}$ are practical bounds on the rate adaptation function. This model treats interference as additive Gaussian noise and hence can have too optimistic results, thus we consider it as an ideal case. Nevertheless, this model is used by 3GPP for system-level evaluation [6].

4. The BS and UE positions, thus the topology, are stationary during each activity session which is a reasonable assumption in residential scenarios. Therefore, the average channel gains $g_{ij}$ (ignoring fast fading) are assumed constant in one optimization cycle.

5. The femtocell network is dense and highly interference-limited which means the probability of the following condition is very high:

$$g_{ij}P_i \gg P_{\text{noise}}, \quad \forall \quad i,j = 1,\cdots,N. \tag{3.2}$$

The SNR in the absence of interference is very high ($\mathrm{SNR} \gg \Gamma_{\max}$) and would result in the maximum possible transmission rate ($C_{\max}$) from the AMC process. With the short distances in a dense residential femtocell deployment, this condition holds with high probability.

6. The path loss (in dB) between BS $i$ and UE $j$ is given by the following simplified model (see [51, chapter 2] and [6]):

$$L_{ij} = 37 + 10\log_{10}(d_{ij}^{\eta}) + S_{ij} \tag{3.3}$$

where $d_{ij}$ is the distance, $\eta$ is the path loss exponent and $S_{ij}$ is a zero-mean Normal random variable that captures the random shadowing due to walls and other indoor

objects. We use two shadowing standard deviations $\sigma_S$ and $\sigma_I$ to differentiate between the shadowing gain within a cell (i.e. for the desired signal) and the shadowing gain across cells (i.e. for interference). This simplified model is justified because the more sophisticated path loss models commonly used in femtocells' evaluations consist mainly of a distance-dependent term, a shadow fading term and some constants related to the carrier frequency and antenna parameters [6]. As the system is interference limited and all femtocells are assumed identical, all constants in the path loss equation cancel out in the SINR expression except the distance and the shadowing dependent terms. Therefore, the model in (3.3) suffices for our purpose.

### *3.3 Optimal Power Allocation*

With the ideal power control assumptions and system model described above, we are now ready to formulate and solve the optimization problem. Recall from Chapter 2 that the capacity maximizing power allocation problem in (2.3) does not provide fairness among interfering links which is an important practical requirement. It turns out that if we make fairness as our optimization objective, we can reformulate the power allocation problem into a convex problem. Instead of maximizing the sum rate, we would like to mitigate interference by maximizing the minimum SINR in the network subject to maximum transmit power constraints. In an interference-limited situation, this would result in the maximum possible equal SINR in all cells. Because $\log_2(1 + \text{SINR})$ is a monotonically increasing function of the SINR, all cells would get equal rate which achieves our fairness objective. Define the SINR of link $i$ ($\text{BS}_i \rightarrow \text{UE}_i$) as follows:

$$\text{SINR}_i = \frac{g_{ii}P_i}{\sum_{j\neq i} g_{ji}P_j \ + P_{\text{noise}}}, \quad i = 1, \cdots, N \tag{3.4}$$

where the channel gain between $\text{BS}_i$ and $\text{UE}_i$ is determined from the path loss $L_{ij}$ as:

$$g_{ij} = 10^{-L_{ij}/10}, \quad i = 1, \cdots, N \tag{3.5}$$

With the necessary constraints on minimum SINR and maximum transmit power from our system model, the optimal fair power allocation problem can be written as:

$$\begin{aligned} &\max_{P_i} \quad \min_{1\le i\le N} \{\mathrm{SINR}_i\} \\ &\text{subject to:} \\ &\quad \mathrm{SINR}_i \ge \Gamma_{\min}, \quad i = 1, \cdots, N \\ &\quad 0 \le P_i \le P_{\max}, \quad i = 1, \cdots, N \end{aligned} \tag{3.6}$$

Using a slack variable $t = \min\{\mathrm{SINR}_i\}$, and defining $a_{ji} = g_{ji}/g_{ii}$ and $b_i = P_{\mathrm{noise}}/g_{ii}$, we can rewrite the problem as the following Geometric Program (GP) in standard form [52]:

$$\begin{aligned} &\text{minimize} \quad t^{-1} \\ &\text{subject to:} \\ &\sum_{j\ne i} a_{ji} P_j \, P_i^{-1} t \; + b_i \, P_i^{-1} t \; \le 1, \qquad i = 1, \cdots, N \\ &\Gamma_{\min} \, t^{-1} \; \le 1, \\ &0 \le P_i \le P_{\max}, \qquad i = 1, \cdots, N \\ &\text{variables:} \quad t \, , P_i \, , \qquad i = 1, \cdots, N \end{aligned} \tag{3.7}$$

Given the pair-wise channel gains $\{g_{ij}\}$ and the system parameters $P_{\max}$ and $\Gamma_{\min}$, we can solve the problem using standard software packages. To solve problem (3.7) we used CVX, a package for specifying and solving convex programs [53].

If the problem is feasible, then the optimal power vector $(P_1, \cdots, P_N)$ guarantees an equal SINR level ($\mathrm{SINR}_{\mathrm{opt}}$) in every cell. This optimal solution is a function of the underlying channel state. Therefore, for each realization of the random topology (pair-wise distances) and shadowing, we get a different solution. In general, there is no closed-form expression for $\mathrm{SINR}_{\mathrm{opt}}$ in terms of the channel matrix $G = \{g_{ij}\}$, thus no probability distribution can be derived for $\mathrm{SINR}_{\mathrm{opt}}$ even if the joint probability distribution of $\{g_{ij}\}$ is known. Next, we derive a closed-form expression for $N = 2$ and then present an empirical probability distribution of $\mathrm{SINR}_{\mathrm{opt}}$ for $N = 9$ assuming two widely used topology models.

### *3.3.1 Solution for the two-cell case*

Let $N = 2$ and define the following channel gain ratios: $a = g_{21}/g_{11}$ and $b = g_{12}/g_{22}$. Denote the normalized power for $\text{BS}_1$ and $\text{BS}_2$ by $x_1, x_2 \in (0, 1]$. Ignoring the noise power (by assumption 5), our problem becomes:

$$\begin{aligned} &\text{maximize} \quad \min\{\frac{x_1}{ax_2}, \frac{x_2}{bx_1}\} \\ &\text{subject to:} \\ &\quad \frac{x_1}{ax_2} \geq \Gamma_{\min}, \quad \frac{x_2}{bx_1} \geq \Gamma_{\min} \\ &\quad 0 \leq x_1, x_2 \leq 1 \end{aligned} \tag{3.8}$$

Since the optimal solution is determined by $a$ and $b$, we have three cases:

1. If $a = b$, then the optimal SINR is achieved by any $x_1 = x_2 \in (0, 1]$. Taking $x_1 = x_2 = 1$, i.e. full power, gives us $\text{SINR}_{\text{opt}}=1/a$.

2. If $a < b$, then the full power allocation $x_1 = x_2 = 1$ would results in $\text{SINR}_1 > \text{SINR}_2$. Therefore, we set $x_2 = 1$ and solve for $x_1$ in the following equality (from the fairness condition):
$$\frac{x_1}{a} = \frac{1}{bx_1} \tag{3.9}$$
which gives us $x_1 = \sqrt{a/b}$. The optimal SINR is given by $\text{SINR}_{\text{opt}}=1/\sqrt{ab}$ if the condition $ab \leq 1/\sqrt{\Gamma_{\min}}$ holds. Otherwise, the problem is infeasible.

3. If $a > b$, then a similar argument gives us $\text{SINR}_{\text{opt}}=1/\sqrt{ab}$ if the condition $ab \leq 1/\sqrt{\Gamma_{\min}}$ holds. Otherwise, the problem is infeasible.

Using the assumed path loss model (3.3), the optimal solution if the problem is feasible is given by:

$$\text{SINR}_{\text{opt}} = \left(\frac{d_{21}\, d_{12}}{d_{11}\, d_{22}}\right)^{\eta/2} 10^{S/10} \tag{3.10}$$

where $d_{ij}$ is the distance, $\eta$ is the path loss exponent and $S \sim \mathcal{N}(0,\, 2\sigma_S^2 + 2\sigma_I^2)$. Clearly, the optimal SINR is a function of the ratio between the interferer distance and the transmitter

distance. If we simplify our path loss model by ignoring the shadowing component and assume a symmetric network, i.e. equal distance ratios for both cells $\alpha = d_{21}/d_{11} = d_{12}/d_{22}$, then the optimal SINR reduces to:

$$\mathrm{SINR}_{\mathrm{opt}} = \alpha^{\eta} \tag{3.11}$$

We will use this result later to characterize the range of $\alpha$ in which a simple Aloha-based random access protocol would perform better than this optimal power allocation under a given traffic load condition.

### *3.3.2 Solution in the general case*

Due to the algorithmic nature of the convex program solver, it is difficult to provide a closed-form expression of the optimal solution for the general case $N > 2$. Therefore, it is not possible to obtain an analytic probability distribution of $\mathrm{SINR}_{\mathrm{opt}}$ even for the simplest input distribution. Here, we provide an empirical probability distribution of the solution in two widely used topology models.

1. **Grid model:** In this model, a rectangular network structure is assumed whereby square apartments are arranged on a grid (Figure 3.1). A femtocell BS and a UE are randomly dropped in each apartment with a minimum separation distance of 1m. A widely used model is the 5-by-5 grid of 10m×10m square apartments representing dense residential femtocells [6]. For this work, we used a 3-by-3 grid of 10m×10m square apartments as illustrated by the sample realization in Figure 3.2. The reason for using the 3-by-3 model is because in the 5-by-5 model, the interference at the center cell would be much higher than the edge cells, hence our max-min fairness criteria would force the optimal solution toward the low SINR region. Since we are looking for the best-case scenario to compare against random access, the 3-by-3 is a good choice.

2. **Circle model:** In this model, the network is a circle of radius $R_1$ and the cell is a circle of radius $R_0$. The first BS is located at the origin and all other $N-1$ BSs are dropped uniformly randomly over the network circle. Then, each UE is dropped

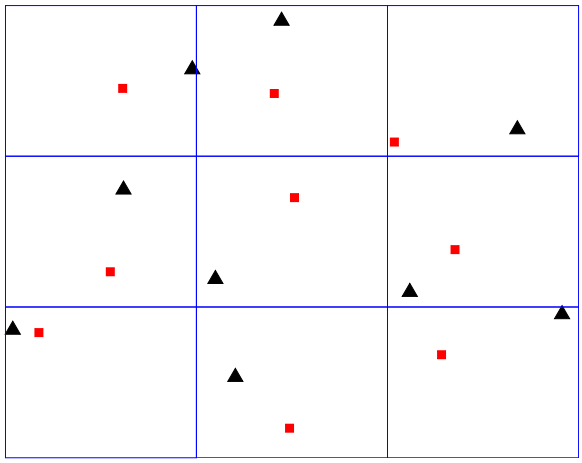

Figure 3.2: Sample realization of the 3-by-3 grid of 10m×10m square model

uniformly randomly in a circle of radius $R_0$ around its BS. Similar to the grid model, a minimum separation distance of 1m is enforced. The values of $R_0$ and $R_1$ are selected such that the average interference distance and the average transmitter distance are approximately equal to those of the grid model. A sample realization of this model is illustrated in Figure 3.3

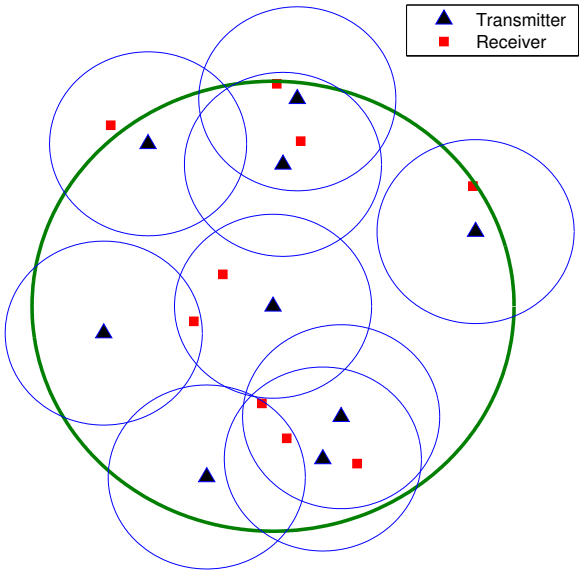


Figure 3.3: Sample realization of the circle model. $R_0$=7.85 m, $R_1$=19.2 m

Table 3.1: System parameters (based on [6])

| Parameter | Value |
|---|---|
| Max power: $P_{\max}$ | 20 dBm, 10 dBm |
| Minimum SINR: $\Gamma_{\min}$ | -10 dB |
| Intra-cell shadowing std: $\sigma_S$ | 4 dB |
| Inter-cell shadowing std: $\sigma_I$ | 8 dB |
| Path loss exponent: $\eta$ | 3 |
| Bandwidth: $W$ | 10 MHz |
| Slot time: $T_{\text{slot}}$ | 1 ms |
| Max spectral efficiency: $C_{\max}$ | 4.4 bits/sec/Hz |
| Implementation gap to capacity: $\beta$ | 0.6 |

With the system parameters in Table 3.1, we generated 40K random samples of each topology and computed the resultant channel matrix $G$ for each sample. Next, for each sample, we try to solve the Geometric Program (3.7) using CVX and if the problem is feasible, the optimal value $\text{SINR}_{\text{opt}}$ is recorded. Then for the same sample of $G$, we compute the SINR of each cell at maximum power ($P_i = P_{\max}, i = 1, \cdots, N$) which represents the uncontrolled-power case, i.e no interference-mitigation. The network is said to be in outage for a particular sample realization, if either the optimization problem is infeasible or, in the uncontrolled-power case, at least one cell is disconnected ($\text{SINR} < \Gamma_{\min}$).

We first compare the probability of network outage in the optimized power-controlled case and the uncontrolled-power case. From Table 3.2, we can see that probability of network outage due to interference is very large compared to the probability of outage due to optimization failure. This asserts our dense network assumption and indicates that this kind of network operate in a *single collision domain* most of the time.

Next, we plot the empirical probability distribution of the optimal solution ($\text{SINR}_{\text{opt}}$) for the two topology models in Figure 3.4. Clearly, the structured grid model results in less interference than the random circle model and hence the optimal solution is about 3

Table 3.2: Probability of network outage

| Topology | Pr{network outage at $P_{\max}$} | Pr{optimization infeasible} |
|---|---|---|
| Grid model | 86% | 18.6% |
| Circle model | 93% | 46% |

dB better in most of the cases. The optimal solution is insensitive to the value $P_{\max}$ which asserts our assumption about the network being interference-limited.

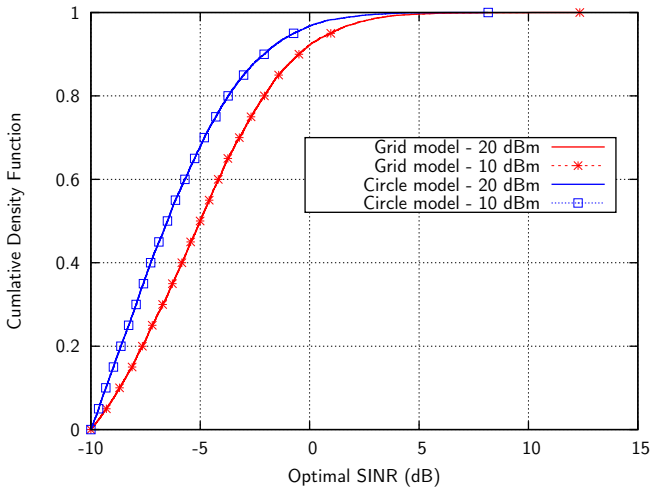


Figure 3.4: Empirical probability distribution of $\text{SINR}_{\text{opt}}$

### *3.4 Queuing Analysis of the Packet Delay*

From the optimal max-min fair power allocation in (3.6), we obtain the same SINR level ($\text{SINR}_{\text{opt}}$) in every cell. Since the optimal solution for a particular random topology realization remains fixed throughout the user session, the transmission rate can be assumed constant for each packet that arrive during that session. Clearly, the downlink buffer in each femtocell behaves like an M/D/1 queuing system where the service time of each packet is some constant D. Note that each random topology realization gives rise to a different value of D.

Consider the femtocell network in Figure 3.1 with $N = 9$ operating in a channel of bandwidth $W$. Time is slotted into intervals of $T_{\text{slot}}$ seconds and all transmissions are synchronized to the slot boundary. In each slot, packets arrive independently at each femtocell BS buffer according to a common arrival process (to be described later). We assume a constant packet size $L_{\max} = C_{\max} W T_{\text{slot}}$ bits which is the maximum packet size that can be carried by the channel in a single slot. We will apply discrete-time queuing analysis to derive the mean packet delay of the following interference-mitigation protocols:

1. Ideal Power Control (PC): In this protocol, the optimal transmit power is pre-computed using the optimization problem (3.6) and is kept fixed during the user session. Whenever a packet arrives, the femtocell BS transmits the packet in the full channel bandwidth $W$ with a fixed rate $C_{\text{opt}}$ as determined by the optimal transmit power.

2. Random Access (Aloha): In this protocol, each femtocell BS transmits at full power ($P_{\max}$) with full rate ($C_{\max}$) in the full channel bandwidth $W$ with a fixed access probability $q$ in every slot. If two or more cells transmit at the same time, a collision occurs and all packets are lost. If there is no collision, one packet is successfully transmitted in exactly one slot. We assume all cells are in a single collision domain which is a worst-case assumption.

3. Orthogonal channel partitioning (FDMA): In this protocol, the channel bandwidth is divided equally among all $N$ femtocells and each femtocell BS transmits at full power ($P_{\max}$) with full rate ($C_{\max}$) in a subchannel of bandwidth $W/N$. This represents the second approach to self-optimization in which the transmit power is fixed at $P_{\max}$ and the optimal sub-set of sub-channels is computed. Since the coherence bandwidth of indoor wireless channels is typically large, there is not much frequency selectivity among the various OFDMA subchannels in one cell. If there is strong interference in one subchannel, it is most likely that other subchannels are experiencing the same level of interference. Hence, any fair subchannel allocation would result in equal subchannel partitioning among the $N$ femtocells as in Frequency Division Multiple Access (FDMA) systems.

### *3.4.1 Queueing system model*

A discrete-time queuing model is used to model each of the three protocols: PC, Aloha and FDMA. The service discipline is First Come First Serve (FCFS) with late arrivals which means packets arrive late in the slot and must wait for the beginning of the next slot to get service. Two discrete-time arrival processes are considered. The first is a Bernoulli arrival process where a packet arrives in every slot with probability $\lambda$. This case can be modeled as a **Geo/G/1** queuing system because the packet inter-arrival time is geometrically distributed with mean $1/\lambda$. The second is a Batch Bernoulli arrival process where a batch of packets (of random size) arrives in every slot with probability $\lambda$. In this case, the inter-arrival time is still geometrically distributed and we have a **batch arrival** Geo/G/1 system which is denoted by $\mathbf{Geo^X/G/1}$, see [54]. The batch size is drawn from a geometric distribution with mean $b = \mathbb{E}[\mathbf{B}]$. The Batch Bernoulli process models bursty traffic applications while the simple Bernoulli process models more regular sources.

Next, we define the following random variables with subscripts $i = 1, 2, 3$ denoting the three protocols: PC, Aloha and FDMA respectively. Denote by $\mathbf{X}_i$ the service time of the Head of Line (HOL) packet in the queue; from the instant it reaches the HOL until it is successfully transmitted and its first two moments by $x_i = \mathbb{E}[\mathbf{X}_i]$ and $x_i^{(2)} = \mathbb{E}[\mathbf{X}_i^2]$. Also, denote by $\mathbf{\Lambda}_i$ the number of packets that arrive in a slot and its first two **factorial moments**[1] by $\lambda_i = \mathbb{E}[\mathbf{\Lambda}_i]$ and $\lambda_i^{(2)} = \mathbb{E}[\mathbf{\Lambda}_i(\mathbf{\Lambda}_i - 1)]$. Finally, the packet delay which is the total time spent in the system is denoted by $\mathbf{T}_i$. All these random variables are expressed in terms of the slot size $T_{\text{slot}}$.

### *3.4.2 Mean Packet Delay*

In order to completely solve the above queuing system, we need to derive the service time distribution. However, since we are only interested in the mean packet delay $\mathbb{E}[\mathbf{T}_i]$, we only need the first two moments of the service time distribution and the inter-arrival time distribution. We apply the following result from discrete-time queuing theory for $\text{Geo}^{\text{X}}/\text{G}/1$

[1]The first factorial moment of a random variable $\mathbf{Y}$ with a PGF $Y(z)$ is defined as $y = \mathbb{E}[\mathbf{Y}] = Y^{'}(1)$ and the second factorial moment is defined as $y^{(2)} = \mathbb{E}[\mathbf{Y}(\mathbf{Y}-1)] = Y^{''}(1)$.

systems to compute the mean packet delay, see [54]:

$$\mathbb{E}[\mathbf{T}_i] = \frac{\lambda^2 x_i^{(2)} - \lambda\rho_i + \lambda^{(2)} x_i}{2\lambda(1-\rho_i)} + x_i \tag{3.12}$$

where $\rho_i = \lambda_i x_i$ is the server utilization.

In the Power Control (PC) protocol, the service (transmission) time of each packet is constant in a given user session and is determined by the optimal solution $\text{SINR}_{\text{opt}}$. From (3.1), we compute the spectral efficiency $C_{\text{opt}}$ at $\text{SINR}_{\text{opt}}$ and obtain the packet service time normalized by the slot time as follows:

$$\begin{aligned} x_1 &= \left\lfloor \frac{L_{\max}}{W C_{\text{opt}} T_{\text{slot}}} \right\rfloor \\ &= \left\lfloor \frac{C_{\max}}{C_{\text{opt}}} \right\rfloor \end{aligned} \tag{3.13}$$

Rounding is needed for our discrete-time model and it does not affect our results since we are looking at the ideal case (i.e. with the smallest possible service time). The second moment is given by:

$$x_1^{(2)} = x_1^2 \tag{3.14}$$

The service time distribution in the Aloha random access protocol is more challenging as explained in Section 2.6 due to the interaction (channel contention) among the interfering nodes. The service of each queue is not independent from the other queues and the system behaves as $N$ coupled queues which is difficult to analyze. Exact queuing analysis for coupled queues has not been achieved except for the case of two queues ($N = 2$) in Aloha systems [33] and in CSMA/CD systems [34]. Several approximation methods exist for deriving the service time distribution of random access protocols, see for example [41–44,55]. For a slotted Aloha protocol, the first two moments of the service time are given by [55]:

$$x_2 = 1 + \frac{1-p}{pq} \tag{3.15}$$

$$x_2^{(2)} = 1 + \frac{(1-p)(2+pq)}{(pq)^2} \tag{3.16}$$

where $q$ is the channel access probability and $p$ is the solution of the following fixed-point equation:

$$p = (1 - \lambda/p)^{N-1} \tag{3.17}$$

In the FDMA protocol, each femtocell BS transmits at full power ($P_{\max}$) with full rate ($C_{\max}$) in a subchannel of bandwidth $W/N$. Similar to the PC case, the service (transmission) time of each packet is constant during the user session. Its first moment is given by:

$$\begin{aligned} x_3 &= \frac{L_{\max}}{\frac{W}{N} C_{\max} T_{\text{slot}}} \\ &= N \end{aligned} \tag{3.18}$$

and the second moment is given by:

$$x_3^{(2)} = N^2 \tag{3.19}$$

Substituting $x_i$ and $x_i^{(2)}$ for $i = 1, 2, 3$ in (3.12), we obtain the mean packet delay of each respective protocol for any given input load provided that the queue stability condition $\rho_i < 1$ holds.

### *3.5 Numerical Evaluation*

We evaluated the mean packet delay of the three interference mitigation protocols in the 3-by-3 grid network ($N = 9$) under bursty and regular input traffic load. To see the effect of burstiness under the same average load, the parameters of the Batch Bernoulli arrival process are set such that the average arrival rate matches that of the simple Bernoulli process. The mean batch size is fixed at $b = 10$ and the arrival probability of a batch is set to $\hat{\lambda} = \lambda/b$ where $\lambda$ is the arrival probability of a packet in the simple Bernoulli process[2].

With the system parameters listed in Table 3.1, we generated 40K samples of the random grid topology and solved the optimization problem (3.6) for $\text{SINR}_{\text{opt}}$ using CVX. If the problem is infeasible, the sample is dropped, otherwise the service time moments $x_1$ and $x_1^{(2)}$ are computed using (3.13) and (3.14). Next, for each feasible case corresponding to one sample of $x_1$, the mean packet delay $\mathbb{E}[\mathbf{T}_1]$ is evaluated for all arrival rates in the stable region $\lambda < 1/x_1$. Then, we computed the $5^{\text{th}}$, $25^{\text{th}}$ and $50^{\text{th}}$ percentiles[3] of the mean

[2]In this case, the second factorial moment of the Batch Bernoulli arrival process is $\lambda^{(2)} = 2\hat{\lambda}(b^2 - b) = 2\lambda(b-1)$ while $\lambda^{(2)} = 0$ for the simple Bernoulli process.

[3]The $\text{p}^{\text{th}}$ percentile of $\mathbb{E}[\mathbf{T}_1]$ is the value of $t$ such that $\Pr\{\mathbb{E}[\mathbf{T}_1] < t\} = \frac{\text{p}}{100}$

packet delay $\mathbb{E}[\mathbf{T}_1]$. The mean packet delay of Aloha is evaluated with the channel access probability fixed at $q = 1/N$ for all arrival rates in the stable region $\lambda < q(1-q)^{N-1}$.

In Figure 3.5, we compare the mean packet delay of Aloha and FDMA with the 5-percentile (PC-5%), 25-percentile (PC-25%) and 50-percentile (PC-50%) of the Power Control (PC) protocol in the simple Bernoulli arrival process. We only show the load region in which 50% of the 40K topology samples are stable, i.e. the condition $\lambda < 1/x_1$ holds for at least 50% of the cases. Clearly, the mean delay in Aloha is less than the 5-percentile of the mean delay in the PC protocol for most of the stable region. This means that in 95% of the topology samples, the optimum power allocation results in larger mean packet delay than Aloha. The results also show that FDMA performs better than optimum power allocation in at least 75% of the cases.

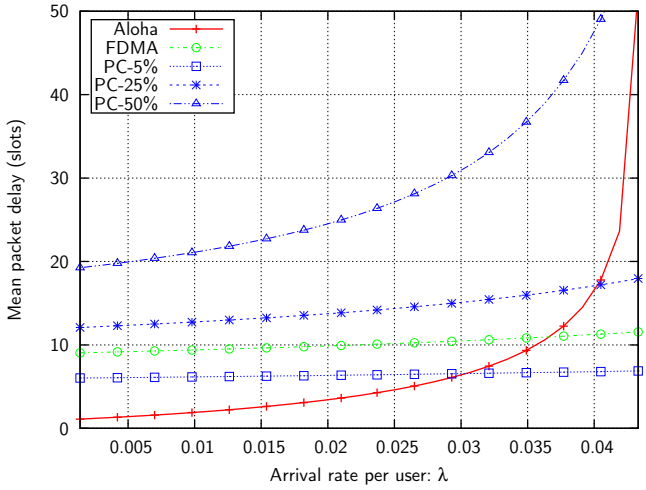


Figure 3.5: Mean packet delay with Bernoulli arrivals. PC-p% is the p-percentile of the mean delay in the optimized PC protocol

The mean packet delay for the Batch Bernoulli arrival process is shown in Figure 3.6. Here too the delay performance of PC is much worse than Aloha with bursty traffic because of the larger second moment $\lambda^{(2)}$ of the arrival process. It is well-known that random access protocols such as Aloha work best with bursty traffic because the large gaps between arrivals reduce the chance of collision.

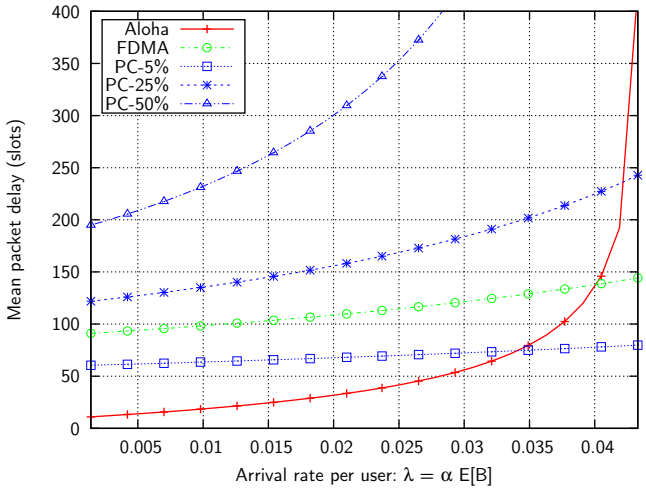


Figure 3.6: Mean packet delay with Batch Bernoulli arrivals. PC-p% is the p-percentile of the mean delay in the optimized PC protocol

In both bursty and regular arrivals, the gap between the delay in Aloha and the $50^{\text{th}}$ percentile of the delay in PC is significant in all the unsaturated load region. This result indicates clearly that an optimized power-controlled access to the channel would perform worse than a simple random access protocol such as Aloha with a high probability. The reason is that the optimum solution tends to protect a few weak links (e.g. center cells) at the expense of everyone else and regardless of the offered load. Considering the ideal, best-case assumptions that we made about this optimum power control and the worst-case assumptions that we made about Aloha, we can safely conclude that any practical power control approach is not expected to perform better than efficiently-designed random access in most of the unsaturated load region. If the network is very saturated, then the best strategy for mitigating interference in a dense femtocell network is a fair subchannel allocation (FDMA) which is agnostic to the network topology and the underlying channel conditions.

### *3.5.1 Two-cell case*

Before we conclude this chapter, we return to the two cell case and examine the conditions at which Aloha performs better than PC. An analytic solution of $\text{SINR}_{\text{opt}}$ for $N = 2$ is given in (3.11) in terms of $\alpha$ (the ratio of the interferer distance to the transmitter distance). We would like to characterize a region for $\alpha$ and $\lambda$ in which the mean delay of PC is guranteed to be larger than the mean delay of Aloha. The exact solution for the 2-node Aloha system with Bernoulli arrivals was derived by [33]. The mean packet delay of Aloha with $N = 2$ and $q = 1/2$ is given by:

$$\mathbb{E}[\mathbf{T}_2] = \frac{2 - 3\lambda}{1 - 4\lambda} \tag{3.20}$$

The mean packet delay of PC in the symmetric 2-cell case with no shadowing can be written in terms of $\alpha$ as follows:

$$\mathbb{E}[\mathbf{T}_1] = \frac{2x_1 - \lambda x_1 - \lambda x_1^2}{2(1 - \lambda x_1)} \tag{3.21}$$

where $x_1 = C_{\max}\left[\beta \log_2(1 + \alpha^{\eta})\right]^{-1}$. We set $\mathbb{E}[\mathbf{T}_1] = \mathbb{E}[\mathbf{T}_2]$ and solve for $\alpha^*$ at which the PC delay equals Aloha delay. Since $x_1$ increases with decreasing $\alpha$ and $\mathbb{E}[\mathbf{T}_1]$ is an increasing function of $x_1$, we can say $\mathbb{E}[\mathbf{T}_1] > \mathbb{E}[\mathbf{T}_2]\ \forall\ \alpha < \alpha^*$. Figure 3.7 shows the region described by the pair $(\lambda, \alpha)$ for which the PC delay is larger than Aloha delay in the two user case. These values of $\alpha$ are not uncommon in dense network of femtocells. The region is even larger if Aloha is optimized by selecting $q$ that minimizes the delay as described in [33].

## *3.6 Conclusions*

In this chapter, we reviewed self-optimization (SON) approaches for interference mitigation in femtocells and described how a SON-based power control protocol works in the current standard framework. Because of the slow re-configuration cycle in such SON-based approaches, the interference coordination algorithm cannot cope with the highly dynamic interference situation caused by the random user activity. As a result, radio resources may be wasted especially in dense networks of femtocells with bursty traffic patterns. We investigated the loss in channel capacity by analyzing the average packet delay of an ideal power control scheme in comparison to the average packet delay that would result if the interference is mitigated solely by a simple random access protocol. It turned out that with

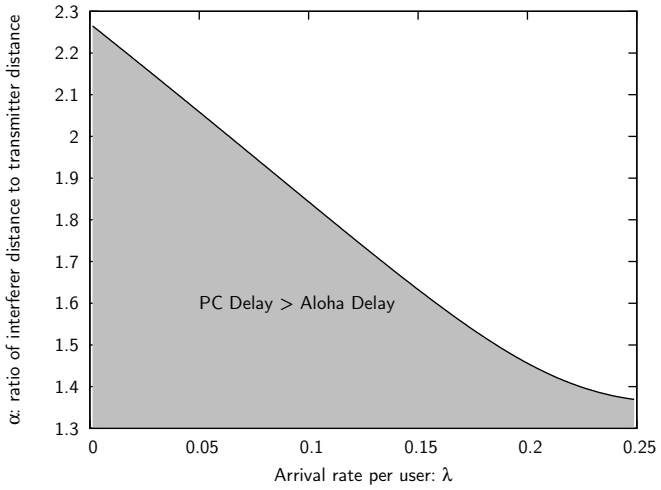


Figure 3.7: Ratio of interferer distance to transmitter distance versus input load. The shaded area describes the region $(\lambda, \alpha)$ in which PC delay is larger than Aloha delay.

a high probability a simple un-optimized random access protocol such as Aloha would perform much better than an optimized power-controlled network in most of the unsaturated load region. If the network is very saturated, then a fixed orthogonal partitioning of the channel would be the best strategy to handle the interference. This result serves as a strong motivation for revisiting the design of random access protocols especially with the flexibility of OFDMA which will be target of the remaining chapters of this thesis.

# Chapter 4

# DELAY ANALYSIS OF OFDMA-ALOHA

In this chapter, we focus on an OFDMA-Aloha system which attempts to exploit the flexibility of OFDMA by extending the basic time-domain backoff procedure employed in slotted Aloha using a single channel, to the frequency-domain in the multi-channel scenario. The idea of OFDMA-Aloha was originally proposed in [22] as a fast-retrial algorithm which is explained in detail in Section 4.1. The authors assumed a finite network with $N$ users and approximated all new arrivals plus fast-retransmissions from frequency- and time-domain retransmissions by an aggregate Poisson model. They derived the distribution of the access delay for a fixed number of channels $K$. However, their results do not allow a fair comparison between the single channel Aloha versus OFDMA-Aloha for different values of $K$. It is not clear from their analysis whether or not the total system bandwidth is fixed for both Aloha and OFDMA-Aloha, as would be necessary for a fair comparison.

The primary contributions of this chapter are as follows. First, we provide a delay analysis of OFDMA-Aloha in the saturated case which is helpful to study the limits of the protocol under various system settings. Here, the *exact distribution of the packet access delay* is derived in terms of the number of channels $K$, number of users $N$ and other system parameters. Our results allow us to study the scalability of the OFDMA-Aloha protocol with varying number of channels and compare it to the single channel Aloha under fixed system bandwidth and load conditions. Second, we derive the mean queue length and mean packet delay in the unsaturated case using an approximate queuing analysis. For this, we utilized the symmetry in the network to decompose the system of interfering queues into multiple independent queues which can be studied using standard queuing theory. The analysis reveals an interesting phase transition in the performance of OFDMA-Aloha (compared to Aloha) when going from the unsaturated to the saturated region. The major results are organized into two main sections- the analysis of the saturated case is presented in Section

4.2 and the queuing model and analysis of the unsaturated case are presented in Section 4.3.

### *4.1 System Model and Assumptions*

The basic operation of OFDMA-Aloha is depicted in Figure 4.1. The total system bandwidth is divided equally into $K$ channels (or sub-channels in OFDMA terms). We assume a *collision channel* model where transmission errors can only occur because of collisions, i.e. we ignore noise and other channel imperfections. A node with a packet to transmit, selects a channel randomly from the $K$ channels and transmits at the beginning of the next slot. If a collision occurs, the node tries to retransmit the packet on a different channel immediately in the next slot. If the packet collides again, the node persists in retransmission until a maximum number of retrials $M$ is reached, at which time the node backs off for a random amount of time and resets its fast-retry counter. This fast retrial feature is what differentiates OFDMA-Aloha from standard multi-channel Aloha.

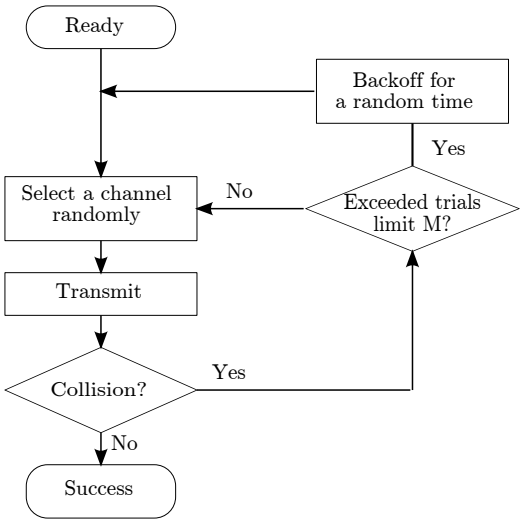


Figure 4.1: OFDMA-Aloha MAC Algorithm

For a consistent and fair comparison between OFDMA-Aloha with $K$ channels and the single channel Aloha, we analyze the two systems under the same total channel bandwidth and net arrival rates. We use the subscript 1 to denote Aloha and subscript 2 to

denote OFDMA-Aloha. The system model in both cases is depicted in Figure 4.2. We assume a (usual) noise-free collision channel with total bandwidth $B$ (total channel rate $R$ bits/second) and fixed packet size $L$. The network consists of $N$ users, each with an infinite buffer. Both MAC protocols are slotted and packet transmission is only allowed at the slot boundary. Time is discretized into *mini-slots* of duration $\tau$ corresponding to the packet transmission time over the full bandwidth, i.e., $\tau = L/R$ seconds. All time related quantities are defined in multiples of the mini-slot $\tau$. With this, the slot duration becomes $\tau_1 = \tau$ seconds in Aloha and $\tau_2 = K\tau$ seconds in OFDMA-Aloha because the rate of each sub-channel is $R/K$. The queuing model is described in more detail in Section 4.3.

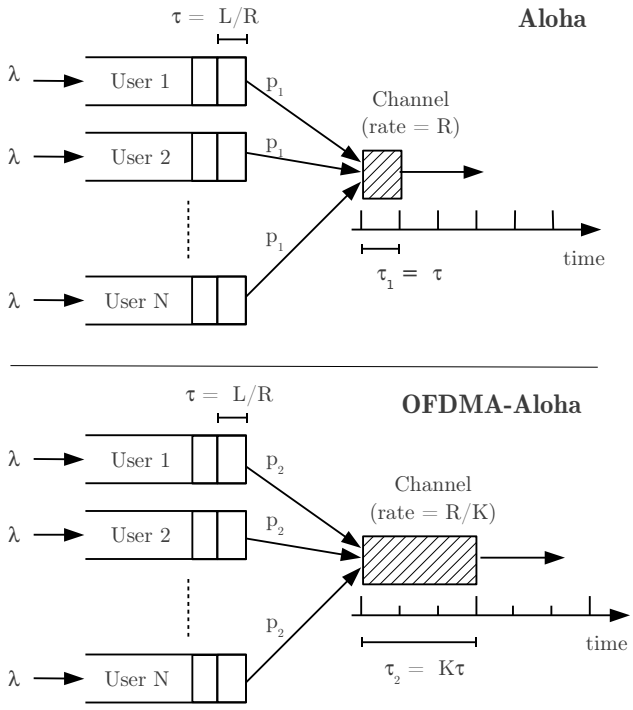


Figure 4.2: System Model

For tractability, we employ the Delayed First Transmission (DFT) model which is widely used in the literature [56]. Under this model, the same transmission probability is used for all packets, whether new or old. That is, a newly arriving packet is forced to wait for the same random backoff period as any other retransmitted packet. Figure 4.3 and Figure 4.4

show the two MAC protocols under the DFT model. The consequence of this model is that when all users are busy, the transmission probability of each user is fixed in all time slots. Using this DFT assumption, upon a new arrival, the node moves immediately into the backoff mode. In the backoff mode, the node transmits (or retransmits) the Head of Line (HOL) packet with probability $p$ or delays it to the next slot with probability $1-p$. The transmission probabilities are $p_1$ and $p_2$ for Aloha and OFDMA-Aloha respectively. Finally, we assume the number of users $N$ is strictly greater than the number of channels $K$ in OFDMA-Aloha; otherwise, the bandwidth becomes underutilized.

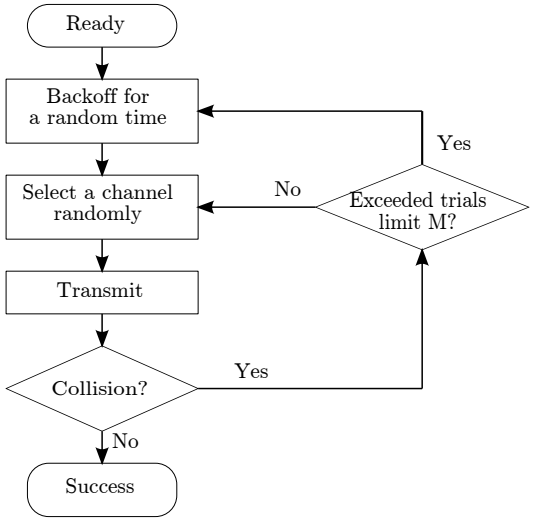


Figure 4.3: OFDMA-Aloha MAC flow chart under the DFT model

## *4.2 Saturated Case Analysis*

In the saturated case, we assume all buffers are always full, such that every successful packet is replaced immediately with a new packet. In this case, the quantity of interest is the throughput or equivalently its reciprocal: the packet access delay $d$. One approach to the analysis of such system is to construct a Markov chain representing the state of the system in any slot and finding the state transition probability matrix. While the system state for Aloha in the saturated case is trivial, we need an $M+1$ dimensional Markov chain to represent the state of the system in OFDMA-Aloha. If $m_i$ is the fast-retry counter at

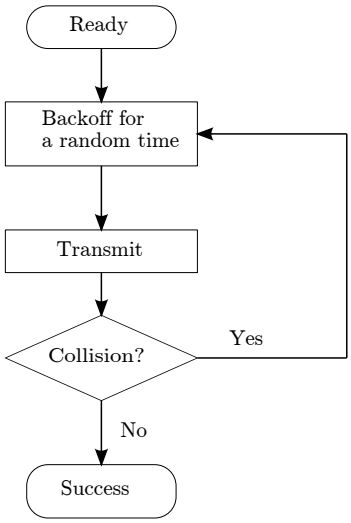


Figure 4.4: Aloha MAC flow chart under the DFT model

user $i$, then the system state is given by $\mathbf{n} = (n_1, n_2, \cdots, n_M, n_{bk})$ where $n_j,\ j = 1, \cdots, M$ represents the number of users with fast-retry counter $m = j$ and $n_{bk}$ represents the number of users in the backoff state, i.e. their fast-retry counters have exceeded $M$. This Markov chain is difficult to analyze even for the simplest case of $N = 2$.

By utilizing the symmetry in the system, we carry out a much simpler analysis using state flow graph techniques [57]. Since all users are statistically identical, we derive the quantity of interest by analyzing the state flow graph of the Markov process representing packet transmission in a typical, *tagged* user. The steady state distribution and moments of the process are derived using transform analysis of the state flow graph described in [57]. This approach has been used previously, for example for the analysis of Tone Sense Multiple Access in [58] and for an approximate queuing analysis of Aloha in [42] and CSMA/CD in [59]. We start by illustrating this technique for Aloha in order to facilitate the exposition of the more complicated OFDMA-Aloha system.

### *4.2.1 Aloha Saturation Analysis*

The state transition diagram of the packet transmission process in a typical user in Aloha is shown in Figure 4.5. The state flow graph is constructed from the original state transition

diagram of the process by multiplying each branch between any two states by the Probability Generating Function (PGF) of the time (or number of steps) required for the transition. Under the DFT assumption, a ready packet moves to the backoff state immediately. In every subsequent slot, the user transmits with probability $p_1$ and defers transmission with probability $1 - p_1$. If we denote the probability of successful transmission by $q_1$, then the packet moves into the success state with probability $p_1 q_1$ and stays in the backoff state with probability $(1 - p_1) + p_1(1 - q_1)$. Each of these two transitions require one time slot, and hence the PGF is just $z$ times these probabilities. Upon successful packet transmission, a new packet moves to the ready state immediately and the process repeats again.

From the state flow graph, the access delay $d_1$ seen by each packet in Aloha is the total time required to move from the "Ready" state to the "Success" state. The PGF of this delay $D_1(z)$ is the transfer function from state $S_\mathrm{r}$ to state $S_\mathrm{s}$ in the flow graph. $D_1(z)$ can be obtained using flow graph reduction methods or directly using Mason's rule, see [60]:

$$\begin{aligned} D_1(z) &= \frac{p_1 q_1 z}{1 - [(1 - p_1)z + p_1(1 - q_1)z]} \\ &= \frac{p_1 q_1 z}{1 - (1 - p_1 q_1)z} \end{aligned} \tag{4.1}$$

where the average success probability $q_1$ of the tagged user in Aloha is given by:

$$q_1 = (1 - p_1)^{N-1}$$

By differentiating $D_1(z)$ and evaluating at $z = 1$ we get the mean access delay $d_1$ (in mini-slots):

$$\begin{aligned} d_1 &= D_1'(1) \\ &= \frac{1}{p_1 q_1} = \frac{1}{p_1(1 - p_1)^{N-1}} \end{aligned} \tag{4.2}$$

### *4.2.2 OFDMA-Aloha Saturation Analysis*

Similarly, we show the state flow graph of the packet transmission process in OFDMA-Aloha in Figure 4.6, where $p_2$ and $q_2$ denote the transmission probability and the success

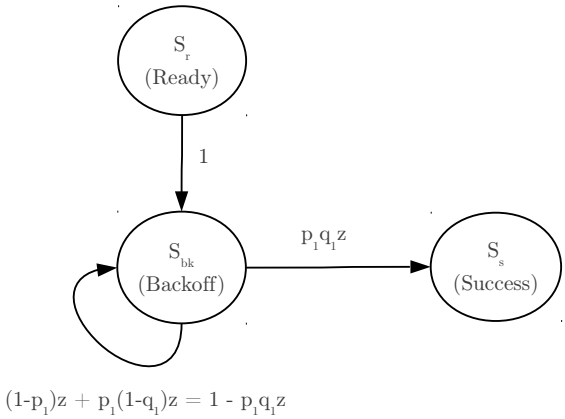


Figure 4.5: State Flow Graph of Packet Transmission Process in Aloha

probability respectively. In this case we have $M$ additional states representing each stage of the fast retry mode. As in Aloha, a ready packet moves to the backoff state $S_{\rm bk}$ immediately. In $S_{\rm bk}$, the user transmits with probability $p_2$ and defers transmission to the following slot with probability $1 - p_2$. When a collision occurs for the first time, the retry counter is incremented to $m = 1$, and the packet moves to state $S_1$. When a second collision occurs in the following slot, the packet moves to state $S_2$, continuing this way until the maximum retry limit $m = M$ is reached in state $S_{\rm M}$. If a collision occurs in state $S_{\rm M}$, the user "gives up" retrying and falls back into the backoff state $S_{\rm bk}$ and resets $m = 0$. Note that in all the fast retry states: $S_1, S_2, \cdots, S_{\rm M}$, the user transmits with probability 1, whereas in the backoff state he transmits with probability $p_2$.

As before, we find the PGF of the access delay $D_2(z)$ in OFDMA-Aloha by writing down the transfer function from state $S_{\rm r}$ to state $S_{\rm s}$ in the state flow graph using Mason's rule:

$$D_2(z) = \frac{p_2 q_2 z \left[1 - (1-q_2)^{M+1} z^{M+1}\right]}{\left[1 - (1-p_2)z - p_2(1-q_2)^{M+1} z^{M+1}\right]\left[1 - (1-q_2)z\right]} \tag{4.3}$$

We could differentiate $D_2(z)$ and evaluate at $z = 1$ to get the mean access delay $d_2$ in OFDMA-Aloha. However, we follow an easier approach which will prove useful later in finding the success probability $q_2$. For this, we apply the transient Markov process analysis described in [57, ch. 4]. Note that from the perspective of a newly arrived packet, the states: $S_{\rm bk}, S_1, \cdots, S_{\rm M}$ are transient states and state $S_{\rm s}$ is a trapping, or an absorbing

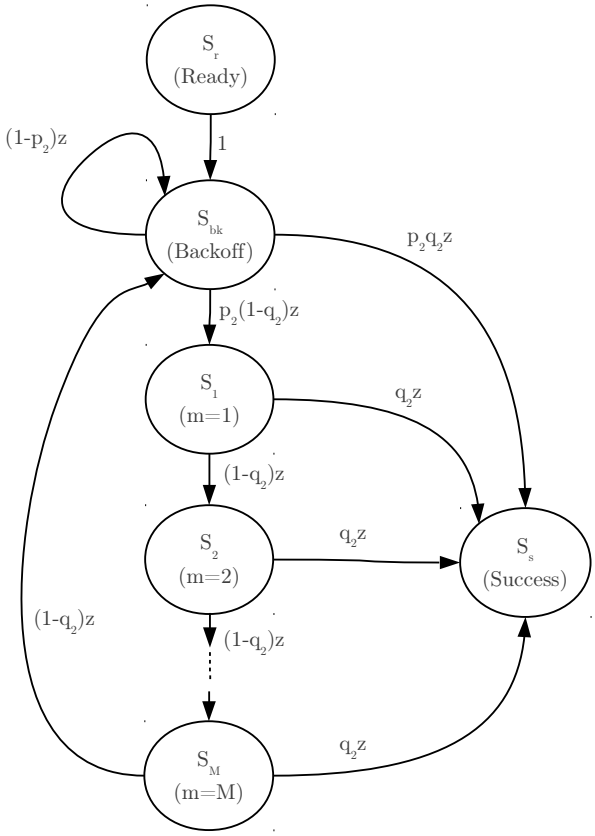


Figure 4.6: State Flow Graph of OFDMA-Aloha

state. Therefore, the delay seen by the packet is the total time spent in the transient process which is the sum of the time spent in the backoff state and all fast retry states.

To find the average time spent in the transient backoff state, $T_{\rm bk}$, we could find the transfer function from the input state $S_{\rm r}$ to state $S_{\rm bk}$ to obtain the PGF of $T_{\rm bk}$ and then differentiate to get $\mathbb{E}[T_{\rm bk}]$. Alternatively, since we are interested in the first moment, we could find the path transmission gain from $S_{\rm r}$ to $S_{\rm bk}$ using a modified state flow graph with $z = 1$. Applying Mason's rule between state $S_{\rm r}$ and state $S_{\rm bk}$ in the modified flow graph, we get the average time spent in the backoff state:

$$T_{\rm bk} = \frac{1}{p_2 - p_2(1-q_2)^{M+1}} = \frac{1}{\Delta}$$

where $\Delta \equiv p_2 - p_2(1-q_2)^{M+1}$. Similarly, we use the modified flow graph to find the average time spent in the fast retry states:

$$\begin{aligned} T_1 &= \frac{p_2(1-q_2)}{\Delta} \\ T_2 &= \frac{p_2(1-q_2)^2}{\Delta} \\ &\vdots \\ T_{\rm M} &= \frac{p_2(1-q_2)^M}{\Delta} \end{aligned}$$

The mean access delay is the total time spent in the transient process (in mini-slots):

$$\begin{aligned} d_2 &= (T_{\rm bk} + T_1 + T_2 + \cdots + T_{\rm M}) \times K \\ &= \left( \frac{1}{\Delta} + \sum_{m=1}^{M} \frac{p_2(1-q_2)^m}{\Delta} \right) \times K \\ &= \frac{q_2 - p_2 q_2 + p_2[1-(1-q_2)^{M+1}]}{p_2 q_2 [1-(1-q_2)^{M+1}]} \times K \end{aligned} \tag{4.4}$$

where the multiplier $K$ signifies that the slot size in OFDMA-Aloha is $K$ mini-slots.

Next, we find the probability of success $q_2$ for the tagged user in OFDMA-Aloha. Denote by $p_t$ the average transmission probability of any user in any slot and the set of all states by $\mathcal{S} = \{S_{\rm bk}, S_1, \cdots, S_M\}$:

$$\begin{aligned} p_t &= \sum_{i \in \mathcal{S}} \Pr[\text{user transmits} \,|\, \text{state } i] \Pr[\text{state } i] \\ &= p_2 \times p_{bk} \; + \; 1 \times p_f \end{aligned} \tag{4.5}$$

where $p_{bk}$ and $p_f$ are the probabilities of being in the backoff mode (state $S_{\rm bk}$) and the fast retry mode (states $S_1, \cdots, S_{\rm M}$) respectively. To find $p_{bk}$ and $p_f$, we utilize the above transient process analysis. $p_{bk}$ is the proportion of time spent in the backoff mode and is given by:

$$\begin{aligned} p_{bk} &= \frac{\mathbf{E}\{\text{time spent in the backoff state}\}}{\mathbf{E}\{\text{total time spent in the transient process}\}} \\ &= \frac{T_{\rm bk}}{d_2} \\ &= \frac{q_2}{q_2 - p_2 q_2 + p_2[1-(1-q_2)^{M+1}]} \end{aligned} \tag{4.6}$$

Similarly, $p_f$ is the proportion of time spent in the fast retry mode and is given by:

$$\begin{aligned} p_f &= \frac{\mathbf{E}\{\text{time spent in all fast retry states}\}}{\mathbf{E}\{\text{total time spent in the transient process}\}} \\ &= \frac{\sum_{m=1}^{M} T_m}{d_2} \\ &= \frac{p_2[1-(1-q_2)^{M+1}] - p_2 q_2}{q_2 - p_2 q_2 + p_2[1-(1-q_2)^{M+1}]} \end{aligned} \tag{4.7}$$

Now, suppose that the tagged user selects channel $c$ for transmission. Since all remaining $N$-1 users are statistically identical and channel selection is uniform with probability $1/K$, the probability of success of the tagged user on this channel is given by:

$$q_2 = \left(1 - p_t \times \frac{1}{K}\right)^{N-1} \tag{4.8}$$

Substituting $p_{bk}$ and $p_f$ in (4.5) and (4.8), we get the following non-linear relation between the packet success probability and the system parameters in OFDMA-Aloha:

$$q_2 = \left[1 - \frac{1}{K}\left(\frac{p_2[1-(1-q_2)^{M+1}]}{q_2 - p_2 q_2 + p_2[1-(1-q_2)^{M+1}]}\right)\right]^{N-1} \tag{4.9}$$

which can be solved numerically for $q_2$ to derive the mean access delay $d_2$ in (4.4).

In the analysis above, we assumed the DFT model in Figure 4.3 and Figure 4.4. For Aloha in the saturated case, this assumption does not pose any significant difference, and hence we can use the result for $d_1$ in (4.2) without this assumption. From this we can state the following result for the original OFDMA-Aloha in Figure 4.1. If we allow unlimited number of fast retrials ($M = \infty$), the access delay in OFDMA-Aloha is always larger than

the access delay in Aloha regardless of the number of channels, as quantified in the following proposition:

**Proposition 1.** *Let $d_2^\infty$ be the mean access delay in OFDMA-Aloha with $K$ channels and with $M = \infty$, i.e. unlimited number of fast retrials. Let $d_1$ be the mean access delay in Aloha with the same total bandwidth and with transmission probability $p_1 = 1/N$ where $N$ is the number of users. If all users are saturated, then:*

$$d_2^\infty \geq d_1 \qquad \forall \quad K > 1 \tag{4.10}$$

*with the equality satisfied only when $K = N$.*

*Proof.* Note that when $M = \infty$ in OFDMA-Aloha without the DFT assumption (Figure 4.1), every user transmits with probability $1/K$ in every slot until the packet is successfully transmitted. This means no time-domain backoff will be encountered. In this case, the success rate of the tagged user in each slot is $q_2 = (1 - 1/K)^{(N-1)}$ and the mean access delay is $1/q_2$ slots. In terms of mini-slots, the mean access delay is given by:

$$\begin{aligned} d_2^\infty &= \frac{1}{q_2} \times K \\ &= \frac{K}{(1 - \frac{1}{K})^{(N-1)}} \end{aligned} \tag{4.11}$$

From the first and second derivatives of (4.11) with respect to $K$, we see that $d_2^\infty$ is strictly convex on the interval $(1, \infty)$ and attains its minimum value at $K = N$. At this value of $K$, $d_2^\infty = d_1$, the mean access delay of Aloha with $p_1 = 1/N$ given by (4.2):

$$d_1 = \frac{N}{(1 - \frac{1}{N})^{(N-1)}}$$

□

### *4.2.3 Numerical and Simulation Results*

In order to verify our analytical results, we simulated the two MAC algorithms using the flow charts in Figure 4.3 and Figure 4.4. The simulation environment represents the abstract system model described in 4.1. We present numerical and simulation results for OFDMA-Aloha with $K = 5,\ 10\,, 15$ and $M = 10,\ 20$. In all cases, the transmission probabilities of all

users were fixed and the load line is varied by changing the number of users. Since we are studying a "symmetric" system of $N$ homogeneous users, we set the transmission probability (attempt rate) of Aloha to $p_1 = 1/N$. For a fair comparison, we set the transmission probability in OFDMA-Aloha to $p_2 = K/N$ which is $K$ times the attempt rate of Aloha because the slot size is $K$ times larger. Also, $p_2 = K/N$ gives identical performance to Aloha during the initial startup phase when all users are in the backoff state. For the numerical evaluation, we used the fzero routine in MATLAB to compute the solution of $d_2$ in (4.4). In all cases of interest ($K < N$), the numerical results agree very well with the simulation results.

The mean access delay in the saturated cased is shown in Figure 4.7 for $M = 10$. The results show that the single channel Aloha always performs better than OFDMA-Aloha when all users are saturated. This can be interpreted as follows; the key idea behind OFDMA-Aloha is to reduce the retransmission time by reducing the collision probability. This is achieved at the expense of an expanded slot duration due to the lower channel rate. The reduction in the collision rate has to be significant to overcome this expense and this is only achieved with larger number of channels, i.e $K$ close to $N$. For example at $N = 20$, the access delay of OFDMA-Aloha with $K = 15$ is very close to that of Aloha. However, as the gap between $K$ and $N$ increases, the reduction in the collision rate decreases and the impact of the expanded slot duration becomes more dominant. This suggests that OFDMA-Aloha might be helpful for lightly loaded network with low collision rates. We will study this case in the next section.

To investigate the effect of the fast retry limit $M$, we compare the access delay of OFDMA-Aloha with 3 different values of $M$ in Figure 4.8 for $K = 15$. Clearly, higher values of $M$ perform worse because they permit more fast retrials which result in higher number of collisions in the saturated case.

### *4.3 Unsaturated Case Analysis*

In the unsaturated case, some user buffers may be empty, and hence will not contend for channel access. This necessitates a queuing analysis of the system. As depicted in Figure 4.2, our system model consists of $N$ queues served by one server (channel). Due to channel

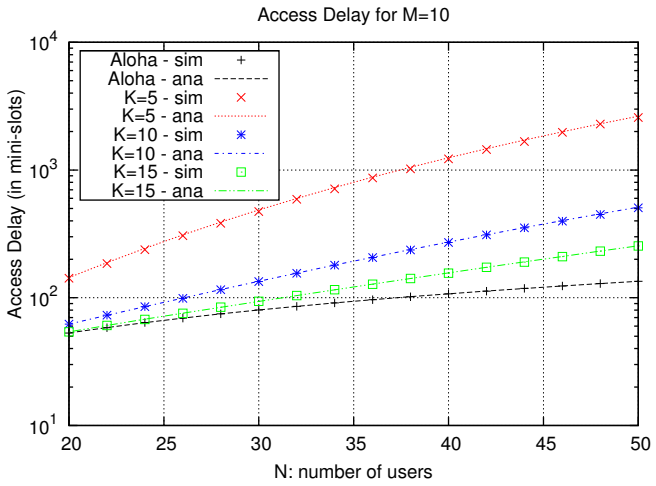


Figure 4.7: Mean access delay in the saturated case for $M = 10$

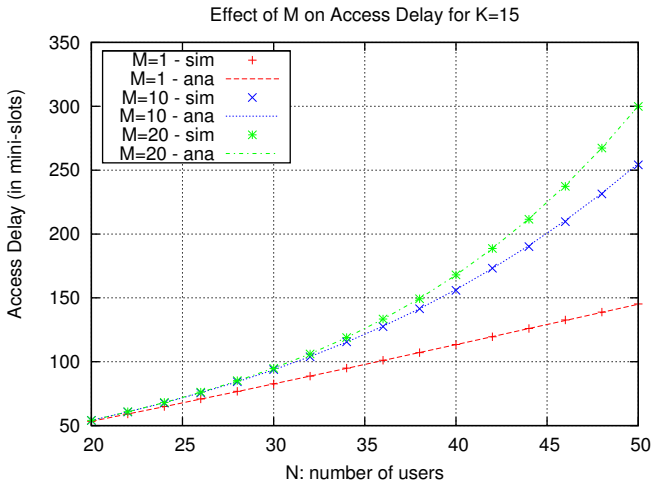


Figure 4.8: Effect of $M$ (Max retry limit) on access delay in the saturated case for $K = 15$

contention, the service time of each queue depends on the state of all other queues. The state of the system is thus an N-dimensional vector $\mathbf{Q} = (Q_1, Q_2, \ldots, Q_N)$, where $q_n$ is the queue size of the $n$th queue. A direct brute-force analysis of this system is intractable due to the large state space, and approximate analysis is necessary. This problem of interfering or interacting queues in multiple access networks has been studied extensively over the past 30 years with various approximation techniques proposed in the literature, see for example [43] and the references therein.

A notable approximation technique is based on decomposing the multi-queue system into $N$ independent queuing systems and capturing the interaction between the interfering queues in the service time distribution using some simplifying assumption. One approach is to assume that each user operates at its steady state independently of all other users and is busy with probability $\rho < 1$. In this case, the distribution of the service time of a typical, "tagged" user can be expressed as a function of the busy probabilities of all other users. Because of the symmetry and fairness of the protocol, the busy probabilities of all other users are identical to that of the tagged user. Therefore, the service time distribution of the tagged user can be expressed as a function of its own busy probability. Then, by applying classical queuing models on the tagged user's queue, we can obtain a relation between the busy probability and the service time distribution. These two relations can be used to solve the system for the quantities of interest. This approach has been used for a queuing analysis of Aloha in [42] and CSMA/CD in [59].

Utilizing the general idea above, we proceed to compare the two systems: Aloha and OFDMA-Aloha with unsaturated load. Particularly, we derive the mean packet delay which is the sum of the queuing delay and the access delay. For clarity of exposition, we split the analysis into two parts: a contention analysis and a queuing analysis and then we combine the results to solve for the quantity of interest. First, we state our queuing model and assumptions.

### *4.3.1 Queuing Model and Assumptions*

In addition to the general system assumptions stated earlier, we further invoke the following in our queuing analysis:

1. A discrete-time queuing model is used to model each of the two systems: Aloha and OFDMA-Aloha with respective slot sizes $\tau_1 = \tau$ second and $\tau_2 = K\tau$ seconds as illustrated in Figure 4.2.

2. The common arrival process is Bernoulli with parameter $\lambda$ with respect to the mini-slot $\tau$, i.e. in every mini-slot a packet arrives with probability $\lambda$ and does not arrive with probability $1 - \lambda$.

3. The service discipline is First Come First Serve (FCFS) with late arrival mode. This means a packet arriving to an empty queue must wait for the next slot to get service.

Next, we define several quantities needed in the analysis. All discrete-time related quantities are expressed in slots where the slot size is defined for each system accordingly. We use subscript $i = 1$ for Aloha and $i = 2$ for OFDMA-Aloha.

1. $\mathbf{T}_i$: the total time spent by a packet in the system.

2. $\mathbf{X}_i$: the service time of the Head of Line (HOL) packet in the queue; from the instant it reaches the HOL until it is successfully transmitted. The probability generating function (PGF) of $\mathbf{X}_i$ and its first two moments are defined as follows:

$$X_i(z) \equiv \sum_{k=0}^{\infty} \Pr[\mathbf{X}_i = k] z^k$$

$$x_i = \mathbb{E}[\mathbf{X}_i] = X_i^{'}(1)$$

$$x_i^{(2)} = \mathbb{E}[\mathbf{X}_i^2] = X_i^{''}(1) + X_i^{'}(1)$$

3. $\mathbf{Q}_i$: the number of packets in the queuing system (including the HOL packet) at the boundary of an arbitrary slot and its PGF is defined as follows:

$$Q_i(z) \equiv \sum_{k=0}^{\infty} \Pr[\mathbf{Q}_i = k] z^k$$

4. $\mathbf{\Lambda}_i$: the number of packets that arrive in a slot and its PGF and first two **factorial moments** are defined as follows::

$$\Lambda_i(z) \equiv \sum_{k=0}^{\infty} \Pr[\mathbf{\Lambda}_i = k] z^k$$

$$\lambda_i = \mathbb{E}[\mathbf{\Lambda}_i] = \Lambda_i^{'}(1)$$

$$\lambda_i^{(2)} = \mathbb{E}[\mathbf{\Lambda}_i(\mathbf{\Lambda}_i - 1)] = \Lambda_i^{''}(1)$$

With the above assumptions, the queuing models for the two systems are described below.

- **Queuing Model for Aloha:** Because the slot duration in Aloha matches with the mini-slot ($\tau_1 = \tau$) in the common arrival process, no more than one packet can arrive in a slot, hence the arrival process in this system is also Bernoulli with parameter $\lambda_1 = \lambda$. This system can be modeled as **Geo/G/1** queuing system because the packet inter-arrival time is geometrically distributed with mean $1/\lambda_1$.

- **Queuing Model for OFDMA-Aloha:** The slot size in OFDMA-Aloha is ($\tau_2 = K\tau$). The number of packets that can arrive in one slot is Binomially distributed with parameters ($K$,$\lambda$). To simplify the analysis, we assume that all packets arriving in a slot arrive in a batch at the end of the last mini-slot of this slot (late arrival model), see Figure 4.9. This approximation is needed because we will derive the mean packet delay of this discrete queuing system in terms of the slot time $\tau_2$ and later convert it to mini-slots by multiplying by $K\tau$. With the batch arrival model, the inter-arrival time is still geometrically distributed and we have a **batch arrival** Geo/G/1 system which is denoted by $\mathbf{Geo^X/G/1}$, see [54]. The PGF and the first two factorial moments[1] of this arrival process in this case are given by:

$$\Lambda_2(z) = (1 - \lambda + \lambda z)^K$$

$$\lambda_2 = K\lambda$$

[1]The first factorial moment of a random variable $\mathbf{X}$ with a PGF $X(z)$ is defined as $x = \mathbb{E}[\mathbf{X}] = X^{'}(1)$ and the second factorial moment is defined as $x^{(2)} = \mathbb{E}[\mathbf{X}(\mathbf{X} - 1)] = X^{''}(1)$.

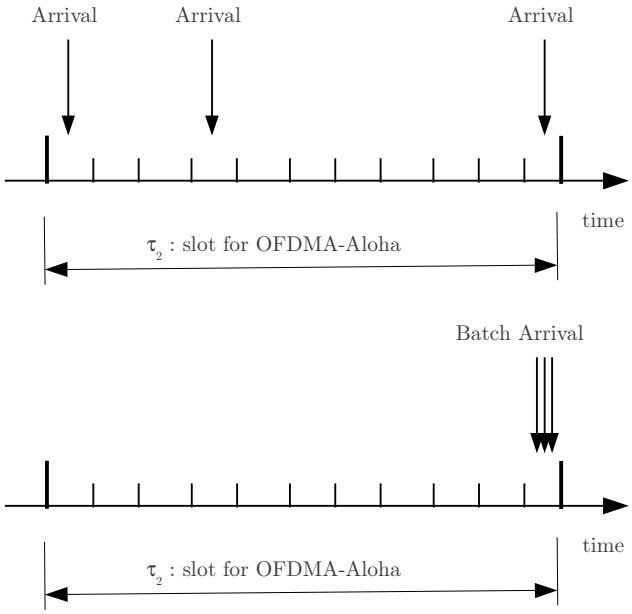


Figure 4.9: Bernoulli arrivals vs. Batch arrivals

$$\lambda_2^{(2)} = (1 - \frac{1}{K})\lambda_2^2$$

### *4.3.2 Contention Analysis*

We first start with the contention analysis of Aloha to obtain $X_1(z)$, the service time distribution, of the tagged user in terms of the busy probability: $p_b$, or equivalently the idle probability $p_0 = 1 - p_b$ . Using the fact that only busy users contend for channel access, we repeat the state flow graph analysis given in the previous section, but this time multiplying the transmission probability of each user by its busy probability $(1 - p_0)$. Following the same approach we used to derive the access delay in (4.1), we get

$$X_1(z) = \frac{p_1 q_1 z}{1 - (1 - p_1 q_1)z} \tag{4.12}$$

and by differentiating and setting $z = 1$ we get the mean service time in Aloha:

$$x_1 = \frac{1}{p_1 q_1} \tag{4.13}$$

The success probability $q_1$ in this case is given by:

$$q_1 = [1 - (1 - p_0)p_1]^{N-1} \tag{4.14}$$

where $p_1$ is the transmission probability in the Aloha system.

Similarly, we use the state flow graph in Figure 4.6 to derive the service time distribution $X_2(z)$ of OFDMA-Aloha:

$$\begin{aligned} X_2(z) = & \frac{p_2 q_2 z \left[1 - (1 - q_2)^{M+1} z^{M+1}\right]}{[1 - (1 - p_2)z - p_2(1 - q_2)^{M+1} z^{M+1}]} \\ & \times \frac{1}{[1 - (1 - q_2)z]} \end{aligned} \tag{4.15}$$

and its mean service time:

$$x_2 = \frac{q_2 - p_2 q_2 + p_2[1 - (1 - q_2)^{M+1}]}{p_2 q_2 [1 - (1 - q_2)^{M+1}]} \tag{4.16}$$

Note that this is the same as the mean access delay $d_2$ given in (4.4) in the saturated case. However, the probability of success here is given by:

$$\begin{aligned} q_2 = & \left[1 - (1 - p_0)\frac{1}{K} \right. \\ & \left. \times \left( \frac{p_2[1 - (1 - q_2)^{M+1}]}{q_2 - p_2 q_2 + p_2[1 - (1 - q_2)^{M+1}]} \right) \right]^{N-1} \end{aligned} \tag{4.17}$$

From equations (4.13-4.16), we see that the mean service time of each system can be expressed as a function of the idle probability $p_0$ of the tagged user in addition to the system parameters: $M$, $K$, $N$, $p_1$ and $p_2$.

### *4.3.3 Queuing Analysis*

Second, we utilize known results from discrete-time queuing theory to obtain another relation between the service time and the idle probability $p_0$ in each system. In both Geo/G/1 and $\text{Geo}^{\text{X}}$/G/1 systems, the PGF of the queue size at the boundary of an arbitrary slot is given by, see [54]:

$$Q_i(z) = \frac{(1 - \rho_i)(1 - z)X_i[\Lambda_i(z)]}{X_i[\Lambda_i(z)] - z} \tag{4.18}$$

where $\rho_i = \lambda_i x_i$ is the load or the utilization of the server. This relation is only valid for $\rho < 1$. From this, we get:

$$\begin{aligned} p_0 &= \Pr[\mathbf{Q}_i = 0] \\ &= 1 - \rho_i \\ &= 1 - \lambda_i x_i \end{aligned} \tag{4.19}$$

Therefore, we have two independent relations which can be used to solve for the success probability $q_i$ in each system in terms of the system parameters and the arrival rate. We substitute (4.13) and (4.19) into (4.14) to get the success probability in Aloha:

$$q_1 = \left[1 - \frac{\lambda_1}{q_1}\right]^{N-1} \tag{4.20}$$

Similarly, we substitute (4.16) and (4.19) into (4.17) to get the success probability in OFDMA-Aloha:

$$q_2 = \left[1 - \frac{\lambda_2}{K q_2}\right]^{N-1} \tag{4.21}$$

Once the success probability $q_i$ is obtained for each respective system, the first and second moments of the service time, can be derived from their PGFs in (4.12) and (4.15) for Aloha and OFDMA-Aloha respectively. Once the service time distribution is determined, the queuing system is actually solved. From (4.18), we get the mean queue size:

$$\mathbb{E}[\mathbf{Q}_i] = \frac{\lambda_i^2 x_i^{(2)} - \lambda_i \rho_i + \lambda_i^{(2)} x_i}{2(1 - \rho_i)} + \rho_i \tag{4.22}$$

Applying the discrete-time version of Little's theorem, we get the mean packet delay:

$$\mathbb{E}[\mathbf{T}_i] = \frac{\mathbb{E}[\mathbf{Q}_i]}{\lambda_i} \tag{4.23}$$

Finally, we stress that this approximate queuing analysis is based on the assumption that every user operates at steady state independently of all other users. This assumption is violated in two cases. The first is when the number of users is small because in this case the interaction among the users is strong and the above approximation is poor. The second is when the mean number of arrivals in a slot exceeds the mean service rate, i.e $\rho > 1$. In

this case, the steady state probability distribution does not exist as all queues saturate and all users will be busy with probability 1. When this happens, no solution can be found for the probability of success in (4.20) and (4.21) and the above analysis cannot be used. The maximum admissible arrival rate before the system saturates defines the stability of the system as discussed in the following subsection.

### *4.3.4 Stability Analysis*

A queuing system is said to be *stable* if $\rho = \lambda x < 1$, where $\lambda$ is the mean arrival rate and $x$ is the mean service time. In buffered multiple access systems, the mean service time $x$ depends on the interaction among all competing queues which makes their stability very difficult to characterize. The general stability region of even the simplest multiple access systems like Aloha has not been fully understood except for few simple cases and several approximate bound. However, for Aloha with homogeneous users, i.e. all users having the same arrival rate $\lambda_1$ and the same transmission probability $p_1 = 1/N$, then the necessary and sufficient condition for system stability was derived by Tsybakov and Mikailov [61] as follows:

$$\lambda_1 < \lambda_{\text{aloha}}^{\max} = \frac{1}{N}(1 - \frac{1}{N})^{N-1} \tag{4.24}$$

This stability result was generalized in [62] for any arrival vector $\boldsymbol{\lambda}$ and any transmission probability vector $\boldsymbol{p}$ as

$$\lambda_i < p_i \prod_{\substack{j \neq i \\ j=1}}^{N} (1 - p_j), \quad \forall \quad i = 1, \cdots, N$$

The stability bound of OFDMA-Aloha is difficult to express algebraically because it involves the unknown zeros of (4.21) which must be found numerically. However, for the settings assumed in this analysis, it can be shown that the stability region of OFDMA-Aloha is strictly smaller than the corresponding region for Aloha. This indicates that OFDMA-Aloha saturates faster than Aloha under the same arrival rate as highlighted in the following proposition.

**Proposition 2.** *In OFDMA-Aloha with $N$ homogeneous users, $K$ channels, a finite number of fast retrials $M$ and a transmission probability $p_2 = \frac{K}{N}$, the maximum admissible arrival*

*rate* $\lambda_{ofdma}^{max}$ *for which the system does* not *saturate, satisfies:*

$$\lambda_{ofdma}^{max} < \lambda_{aloha}^{max} = \frac{1}{N}(1-\frac{1}{N})^{N-1} \tag{4.25}$$

*for all* $1 < K < N$ *and* $M \geq 1$.

*Proof.* The proof hinges on using the dominant system approach introduced in [63] and extended in [62]. Denote by S1 our original $N$-queue OFDMA-Aloha system with the queuing model described before. Define another OFDMA-Aloha system S2 which is identical to S1, with the same arrival rate $\lambda_2$ and the same transmission probability $p_2 = K/N$. The only difference is that when a queue becomes empty in S2, it continues to transmit "dummy" packets with the same transmission probability $p_2$. Dummy packets can result in collisions, but the successful transmission of a dummy packet does not reduce the queue size. Therefore, S2 always has larger queue size than S1 if both start from the same initial conditions, i.e. S2 dominates S1. Clearly, if S2 is stable then our original system S1 is also stable.

The mean service time of the dominant system is given by (4.16), but its success probability $q_2$ is given by (4.9) because it is indistinguishable from OFDMA-Aloha under saturation, i.e. $p_0 = 0$. Substituting (4.16) in (4.9), we get:

$$q_2 = \left[1 - \frac{1}{Kq_2x_2}\right]^{N-1}$$

or equivalently,

$$x_2 = \frac{1}{Kq_2(1-q_2^{1/(N-1)})}$$

For a stable queuing system, we must have $\lambda_2 x_2 < 1$, giving the following stability bound for $\lambda$ under OFDMA-Aloha:

$$\begin{aligned} \lambda_2 &< Kq_2(1-q_2^{\frac{1}{N-1}}) \\ K\lambda &< Kq_2(1-q_2^{\frac{1}{N-1}}) \\ \lambda &< q_2(1-q_2^{\frac{1}{N-1}}) \end{aligned} \tag{4.26}$$

where $q_2$ is the solution of (4.9). Define

$$h(x) = x(1-x^{\frac{1}{N-1}}), \quad 0 < x < 1$$

It can be shown that $h(x)$ is maximized at $x_0 = (1 - \frac{1}{N})^{N-1}$. In addition, $h'(x) > 0$ for $0 < x < x_0$ and $h'(x) < 0$ for $x_0 < x < 1$ and hence:

$$\begin{aligned} \max_{0<x<1} h(x) &= h(x_0) \\ &= \frac{1}{N}(1 - \frac{1}{N})^{N-1} \end{aligned}$$

It is sufficient for our purpose to prove that $x_0 = (1 - \frac{1}{N})^{N-1}$ is not a solution of (4.9). In the Appendix, we show that if $p_2 = K/N$ then there exist a solution $q_2^*$ of (4.9) and it is strictly less than $x_0$, hence

$$\lambda_{\text{ofdma}}^{\max} = h(q_2^*) < h(x_0) = \lambda_{\text{aloha}}^{\max}$$

□

### *4.3.5 Numerical and Simulation Results*

For the unsaturated case, we compare Aloha versus OFDMA-Aloha with $K = 10\,, 20$ and $M = 10$. We fix the number of users at $N = 30$ and vary the load line by increasing $\lambda$ which gives $\lambda_1 = \lambda$ for Aloha and $\lambda_2 = K\lambda$ for OFDMA-Aloha. We use the same transmission probability as in the saturated case: $p_1 = 1/N$ for Aloha $p_2 = K/N$ for OFDMA-Aloha.

The mean packet delay is plotted in Figure 4.10. When the load is very low, there is no queuing delay and very few collisions occur in both systems. In this case, the mean packet delay is just the initial backoff time under the DFT assumption. The expected backoff period is the same in both systems: $1/p_1 = N$ mini-slots in Aloha and $1/p_2 = N/K \times K = N$ mini-slots in OFDMA-Aloha. For relatively small $\lambda$, OFDMA-Aloha enjoys smaller packet delays as compared to Aloha. Also, within OFDMA-Aloha, smaller values of $K$ perform better than larger values. This is because the collision rate is relatively low and fewer channels are sufficient to absorb the load and reduce the number of collisions. The fast retrial feature of OFDMA-Aloha works best in this region and achieves faster retransmission time. However, increasing the number of channels increases the slot duration without any further reduction in the number of collisions, and hence results in slightly larger delay.

When the load increases beyond Aloha's saturation line at $\lambda_{\text{aloha}}^{\max}$, both systems become saturated and the delay grows without bound. However, we note that for smaller values of

$K$, OFDMA-Aloha saturates faster than larger values and also faster than Aloha, a result expected from proposition 2. When the system saturates, we cannot find a numerical solution and the simulation returns very large numbers. We omit these large numbers from the plot in order not to obscure the other results. This behavior near the saturation line is somehow expected because we have seen from the analysis in section 4.1 that OFDMA-Aloha with smaller values of $K$ performs the worst in saturation. To verify this, we plot the mean service time of the two systems in Figure 4.11 and the average load $\rho$ in Figure 4.12. Note that in these two figures, no analytical results are shown for values of $\lambda$ near Aloha's saturation line because no solutions could be found here as explained before. This is particularly evident in Figure 4.12 where the analytical results are missing when $\rho$ approaches 1.

In all cases, the mean service time beyond Aloha's saturation line converges to a constant which equals the mean access delay that was derived in the saturated case, see (4.2) for Aloha and (4.4) for OFDMA-Aloha as shown by the solid horizontal lines in Fig 4.11. Therefore, the behavior of the two MAC protocols in the unsaturated case is exactly opposite to that of the saturated case. This confirms our earlier reasoning that the reduction in the collision rate in OFDMA-Aloha has to be significant enough to overcome the effect of the expanded time scale. When the collision rate is relatively low, a small number of channels is sufficient to absorb the load and improve the retransmission time. However, when the collision rate is very high, e.g. in saturation, channelization only helps to expand the time scale and hence performs worse than the single channel.

From the above, we see that the performance of the two MAC protocols undergoes a phase transition when going from the unsaturated region to the saturated region. This phenomena was also reported in the past in the analysis of multi-channel Aloha by Yue and Matsumoto in [64, chapter 2]. Multi-channel Aloha allows a collided node to switch to another random channel, but only after a random back-off period, so it does not have the fast retrial feature of OFDMA-Aloha. We note though, that their analysis is for nodes with single-packet buffer size and is substantially different from our approach.

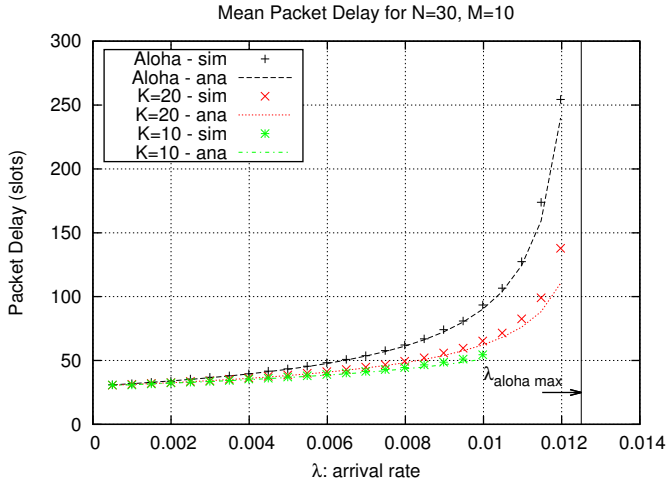


Figure 4.10: Mean packet delay in the unsaturated case for $N = 30$ and $M = 10$

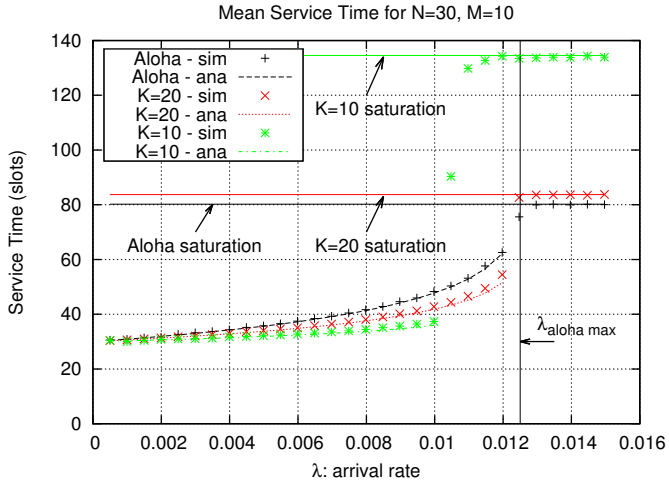


Figure 4.11: Mean service time in the unsaturated case for $N = 30$ and $M = 10$

## *4.4 Conclusion*

OFDMA-Aloha is a new MAC protocol that promises to exploit the channel switching flexibility of OFDMA. By allowing collision resolution over the frequency as well as time domains, the protocol attempts to reduce the packet retransmission time. However, this comes at the great expense of expanded time scale, or larger slot size due to lower channel

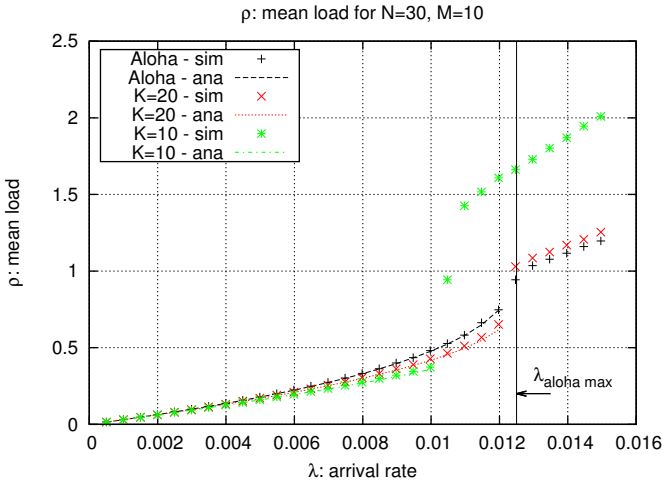


Figure 4.12: Mean load $\rho$ in the unsaturated case for $N = 30$ and $M = 10$

rates. We showed that when the network is already saturated, channelization does not bring substantial reduction in the collision rate to the point where it outweighs the effect of expanded slot size; single channel Aloha performs better than OFDMA-Aloha especially when the gap between the number of channels and the number users is large. On other hand, when the network is lightly loaded, OFDMA-Aloha enjoys smaller packet delays, but not for long as it saturates faster than the single channel Aloha. This suggests the need for further study on the stability region of OFDMA-Aloha as it may help develop practical adaptive algorithms for the future.

## *4.5 Appendix*

Here we show that all solutions of eq (4.9) in the interval $(0, 1)$ are strictly less than $x_0 = (1 - 1/N)^{N-1}$ when $p_2 = K/N$.

*Proof.* The proof is tedious, but straightforward. Substitute $p_2 = K/N$ and rewrite (4.9) as follows:

$$\begin{aligned} 0 = &(N-K)q_2(1-q_2^{\frac{1}{N-1}}) \\ &+ \left[1-(1-q_2)^{M+1}\right](K - Kq_2^{\frac{1}{N-1}} - 1) \end{aligned}$$

Define $z = q_2^{\frac{1}{N-1}}$, and for convenience define $n = N-1$, $m = M+1$, and $a = N-K$. Then, we have the following polynomial:

$$\begin{aligned} g(z) = & az^n(1-z) + K - Kz - 1 \\ & - (1-z^n)^m(K - Kz - 1) \end{aligned}$$

Note that $n > 1$, $m > 1$, $1 < K \leq n$ and $1 \leq a \leq n$ from the assumptions in this chapter. For our purpose, it is sufficient to show that $g(z)$ has at least one real root in the interval $(0, z_0)$ and no real roots in $(z_0, 1)$ where $z_0 = 1 - 1/N$.

First, we show the existence of the solution and in particular, we show that there is a real root in $(0, z_0)$. Let $z_1 = 1/N$.

$$\begin{aligned} g(z_1) = & a(\frac{1}{N})^n(1 - \frac{1}{N}) \\ & + (K - \frac{K}{N} - 1)\left[1 - (1 - (\frac{1}{N})^n)^m\right] \end{aligned}$$

Since $\frac{K}{N} < 1$ and $K \geq 2$, all above terms are positive and $g(z_1) > 0$.

$$\begin{aligned} g(z_0) = & (1 - \frac{K}{N})(1 - \frac{1}{N})^n \\ & - (1 - \frac{K}{N})\left[1 - (1 - (1 - \frac{1}{N})^n)^m\right] \\ = & -(1 - \frac{K}{N})(1 - w) \end{aligned}$$

where $w$ is the term:

$$w = (1 - \frac{1}{N})^n + \left[1 - (1 - \frac{1}{N})^n\right]^m$$

For $m \geq 2$, $w < 1$ and $g(z_0) < 0$. Hence, by the Intermediate Value Theorem, $g(z)$ must have a zero in $(z_1, z_0)$ or equivalently $g(z)$ has at least one real root in $(0, z_0)$.

Next, we show that there are no real roots of $g(z)$ in $(z_0, 1)$. Taking the first derivative of $g(z)$:

$$\begin{aligned} g'(z) = & anz^{n-1} - a(n+1)z^n \\ & + mn(K - Kz - 1)z^{n-1}(1 - z^n)^{m-1} \\ & + K(1 - z^n)^m - K \end{aligned}$$

Defining $w_1 = (1 - z^n)^{m-1}$ and $w_2 = (1 - z^n)^m$ and after some algebra, we can write $g'(z)$ as:

$$\begin{aligned} g'(z) = & -z^{n-1}(ah_1 + mnw_1h_2) \\ & -K(1 - w_2) \end{aligned}$$

where:

$$h_1 = (n + 1)z - n$$

and

$$h_2 = Kz + 1 - K$$

It can be shown that $h_1$ and $h_2$ are always positive for $z \in [z0, 1]$ and $K \in (1, n]$. Since $0 < w_1, w_2 < 1$ and $a$, $n$ $m$ are all positive, $g'(z) < 0$ and hence $g(z)$ is decreasing on $[z0, 1]$. Since $g(z_0) < 0$ (from above analysis) and $g(1) = -1$, $g(z)$ has no real roots in $(z_0, 1)$.

□

# Chapter 5

# EXPONENTIAL BACKOFF IN FREQUENCY (EBF)

In Chapter 2, we discussed OFDMA-based Random Access protocols and presented a case for random access in femtocells. Two packet-level MAC models were described in Section 2.5: a *Single-Access* model which allows only a single packet per transmission attempt and a *Bulk-Access* model which allows multiple packets per transmission attempt. For the Single-Access model, a detailed analysis of the OFDMA-Aloha MAC algorithm was presented in Chapter 4. In this chapter, we propose and analyze a new Random Access MAC algorithm using the Bulk-Access model. The proposed Exponential Backoff in Frequency (EBF) algorithm allows femtocells to transmit packets in bulk and resolves collisions in the frequency-domain using a synchronized binary tree branching process over the channels that reduces the accessed bandwidth after every failed attempt. The main goal of this MAC algorithm is to reduce packet delays in dense OFDMA femtocells under typical "unsaturated" traffic scenarios.

The novel MAC algorithm offers new insights into the design and analysis of multiple access techniques for LTE femtocells outside the usual framework of cellular resource allocation. Aided by the packet-level MAC model proposed in Section 2.5, we are able to hide most of the LTE protocol stack details and focus on the impact of the multiple access algorithm over the 2-dimensional OFDMA time-frequency grid (Figure 2.2). The proposed MAC algorithm is described in Section 5.2. The analysis of packet delay along with simulation results are presented in Section 5.3 followed by concluding remarks.

## *5.1 System Model and Baseline Scenario*

We consider a cluster of residential LTE femtocells in a dense apartment building. There are $N$ femtocells each serving exactly one user and competing for $K$ OFDMA channels (or subchannels in OFDMA context), see Figure 5.1. We assume a fully connected network

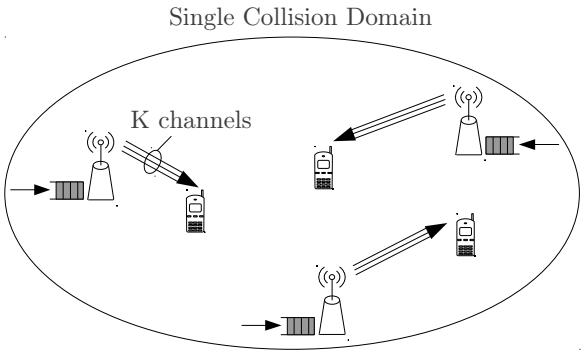


Figure 5.1: Basic scenario and system model: $N$ femtocells, $K$ channels

where all transmissions are heard by all nodes, i.e. the femtocell cluster belongs to a single collision domain. This assumption is commonly used to simplify the analysis of basic packet-level performance of any MAC algorithm [65]. Also, we assume collisions to be the only source of transmission errors and ignore errors due to channel noise and link-level imperfections. Only one packet can be successfully transmitted in any given channel, i.e. no packet capture is modeled. The time access is slotted and all nodes are assumed to be synchronized to the slot boundary. Furthermore, since co-channel femtocells typically belong to the same service provider, we assume that each femtocell knows the network size $N$ either through configuration or through initial environment sensing.

In an OFDMA femtocell environment, the number of channels (subchannels) is usually larger than the number of active users. Hence, an important design criterion for any MAC algorithm is to allow the user access to as many channels as possible to maximize utilization. Since the user can initiate transmissions over multiple parallel channels, we need to define the notion of collision for such parallel transmissions. In Section 2.5, we proposed an abstract access model called the *Bulk Access* model to facilitate the design and analysis of random multiple access in OFDMA femtocells. In this model, the size of each packet is fixed and equals one Resource Block (RB) - one time-frequency slot in an LTE network. Each node can transmit a group of packets as a bulk over multiple channels simultaneously. The key assumption in this model is that a collision in any channel renders the whole bulk corrupted and all packets in the bulk must be retransmitted. This is a result of abstracting the slot-by-

slot link adaptation and the segmentation/reassembly of data packets in OFDMA systems, see Section 2.5 for more discussions.

As a baseline, we use two simple MAC schemes for comparison with our proposed MAC algorithm. The first is the simplest random access MAC algorithm (**Aloha**) where the user transmits with probability $p$ in every slot. In this scheme, a user with $Q \leq K$ packets transmits his bulk of packets over fixed channels 1 to $Q$. An improved scheme (**K-Aloha**) is to spread the bulk uniformly over the available $K$ channels. In K-Aloha, the first packet in the bulk is mapped randomly to one of $K$ channels; the second is then mapped randomly to one of the remaining $K-1$ channels, and so on for the remaining packets in the batch. Therefore, K-Aloha attempts to utilize the frequency dimension to reduce collisions especially in low traffic scenarios, i.e. $Q \ll K$.

### *5.2 Proposed MAC Algorithm*

The proposed *Exponential Backoff in Frequency* (**EBF**) MAC employs a binary tree branching process with respect to the number of accessible channels $C$. The basic idea of the algorithm is as follows. Initially, a user tries to access the full bandwidth ($C = K$ channels) and transmits $C$ packets in parallel. If there is a collision, the user backs off and halves the number of accessible channels and attempts to transmit $C = K/2$ packets on one out of two branches chosen at random. In case of another collision, it halves again and attempts to transmit $C = K/4$ packets on one out of four branches. This continues until it reaches $C = 1$ after $\log_2(K)$ steps at which the user starts over at $C = K$ and repeats the procedure again. In this way, packets' transmissions are always clustered into branches chosen randomly, instead of spreading all packets uniformly over the $K$ channels as done in K-Aloha. To minimize collisions, the backoff branching process is synchronized among all users in the system. Every user maintains a backoff stage $m$ that counts down over $\{M-1, M-2, \cdots, 1, 0\}$ in every slot regardless of any transmission activity, see Figure 5.2. The $(M-1)^{\text{th}}$ backoff stage, where $(M = \log_2(K)+1)$, allows the user to access all $K$ channels while the $0^{\text{th}}$ stage limits the accessible channels to only 1 channel. In addition, we attach an access probability $p_m$ which is proportional to the number of accessible channels as illustrated in Table 5.1. For example, at stage $m$, a user can only access (with probability

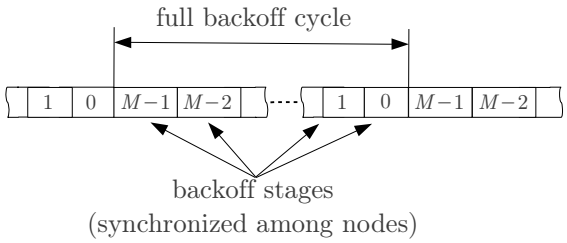


Figure 5.2: Example backoff cycle of EBF MAC, backoff stage $m$ rotates between $M$ phases

Table 5.1: Relationship between the backoff stage $m$ and different EBF variables

| $m$: backoff stage | $C$: channels limit | $p_m$: access probability |
|---|---|---|
| $m = 0$ | 1 | 1 (for $K > N$) |
| $m = i$ | $2^i$ | $p_m = \min\{1, K/(N2^i)\}$ |
| $m = M - 1$ | $2^{M-1} = K$ | $p_m = 1/N$ |

$p_m$) a branch of size $2^m$ channels selected randomly from $K/2^m$ branches. The detailed flow of the EBF algorithm is described in Figure 5.3. Note that only colliding users are restricted by the backoff stage whereas successful users always start with full bandwidth, i.e. $C = K$.

### *5.3 Performance Analysis*

In a typical femtocell environment, packet traffic is often bursty with low duty cycle. To model this traffic, we use a Batch Bernoulli arrival process in which an idle user generates a batch of packets with probability $\alpha$ and does not generate any traffic with probability $1 - \alpha$. Once a batch is generated, the user becomes busy and does not generate any more traffic until all packets are successfully transmitted. This idle/busy model is widely used in the analysis of MAC protocols in non-saturated scenarios, see for example [65]. The size of each packet is fixed so that its transmission occupies exactly one time-frequency slot as described in the Bulk Access model (Section 2.5). The batch size $B$ is a random variable drawn from a *strictly positive* Binomial distribution with maximum size $K$ and mean $\mathbf{E}[B]$;

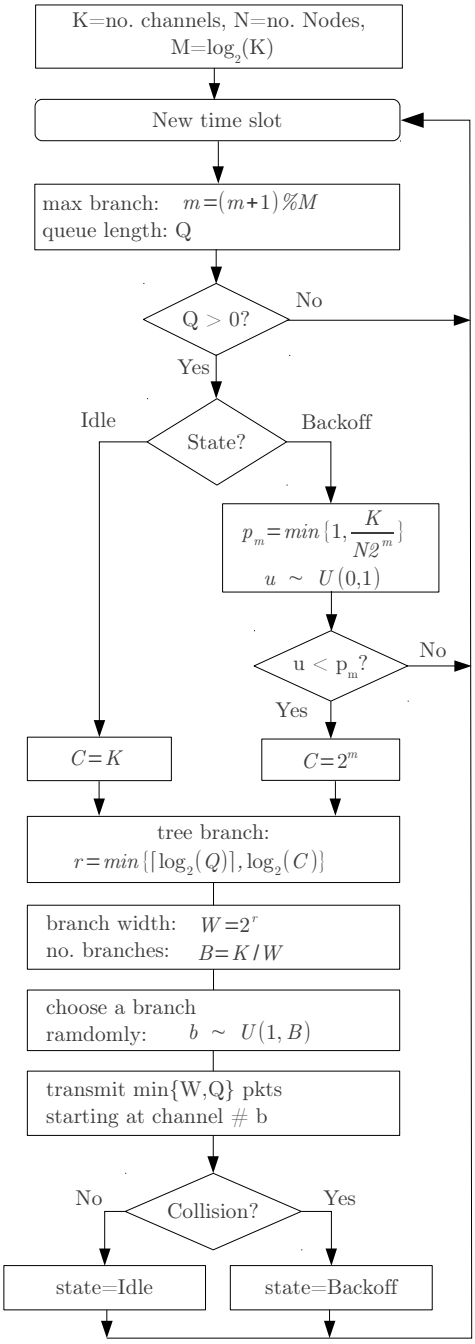


Figure 5.3: Flow chart of the EBF Algorithm. Note: all nodes are synchronized with the same value of $m$ in every slot

i.e., $B = B_0 + 1$ where $B_0 \sim$ Bino(K-1, $\beta$) which gives

$$\begin{aligned} \mathbf{B}_n &\equiv \Pr[B = n] \\ &= \frac{n}{K\beta}\binom{K}{n}\beta^n(1-\beta)^{K-n}. \end{aligned} \tag{5.1}$$

The coin toss probability $\beta$ is chosen according to the desired mean batch size $\mathbf{E}[B]$, i.e. $\beta = (\mathrm{E}[B]-1)/(K-1)$. Limiting the maximum batch size to $K$ simplifies the analysis considerably because otherwise we have to deal with the residual packets after every successful batch transmission if $B > K$.

In unsaturated scenarios, some users may be idle, thus will not contend for channel access. In steady state, a MAC protocol's output rate (throughput) equals the offered input traffic rate for stable operation. Therefore, we use the packet delay as our performance measure for the proposed MAC protocol. In single-channel systems, the state of the system is simply the number of busy users which determines channel contention. However, because we have multiple channels in our system we need to keep track of how many packets at each busy user. The state of this system is a $K$-dimensional vector $\mathbf{A} = (a_1, a_2, \cdots, a_K)$, where $a_n$ is the number of users having $n$ packets in their buffer. A direct brute-force analysis of this system is intractable due to the large state space, hence approximate analysis is necessary. We use such an analysis based on renewal theory to derive the relationship between the mean packet delay and arrival rate, batch size and other system parameters. Also, we use the Delayed First Transmission (DFT) model in which the user upon generating new packets moves immediately to the backoff state so that the first transmission attempt has the same access probability as all subsequent retransmissions. We start with the analysis of the K-Aloha protocol to facilitates the derivation of the more complicated EBF MAC algorithms.

### *5.3.1 K-Aloha Delay Analysis*

In K-Aloha, the user alternates between the *idle* state and the *busy* state. Consider user $i$ and denote by $Q_i$ the number of packets in his queue. Because K-Aloha does not allow the user to divide the transmission of his batch of packets over multiple slots (in contrast to

EBF) and the batch size is always bounded by $K$, the queue length $Q_i$ will always be equal to the batch size $B$. From (5.1), the distribution of $Q_i$ is

$$\begin{aligned} A_i(n) &\equiv \Pr[Q_i = n] \\ &= \frac{n}{K\beta}\binom{n}{K}\beta^n(1-\beta)^{K-n}. \end{aligned} \tag{5.2}$$

We first derive the average pair-wise collision probability $p_c$ between any two active (transmitting) users:

$$\begin{aligned} p_c &\equiv \Pr\text{ [collision between any two active users]} \\ &= \sum_{n=1}^{K}\sum_{m=1}^{K} b_{i,j}(n,m)\; A_{i,j}(n,m) \end{aligned} \tag{5.3}$$

where $b_{i,j}(n,m)$ is the conditional pair-wise collision probability between user $i$ having $n$ packets and user $j$ having $m$ packets and the joint queue distribution is simply $A_{i,j}(n,m) = A_i(n)A_j(m)$ because users are independent. To find $b_{i,j}(n,m)$, consider an urn with $K$ balls and $n$ of them are marked. We draw $m$ balls at random and we consider the outcome to be "success" if none of the $m$ balls is marked, and a failure or "collision" otherwise. This is because in our Bulk Access model, a collision in any channel used in the transmission of the bulk, renders the whole bulk corrupted and all packets must be retransmitted. It can be easily shown that

$$b_{i,j}(n,m) = 1 - \prod_{l=0}^{n-1}\left(1 - \frac{m}{K-l}\right). \tag{5.4}$$

From (5.2-5.4), we can obtain the average pair-wise collision probability $p_c$ between any two active users. The average success probability of an arbitrary "tagged" user in K-Aloha is then given by

$$q = p\,(1 - p_b\,p\,p_c)^{N-1}, \tag{5.5}$$

where $p_b$ is the probability of the user being busy. We utilize the Alternating Renewal Theorem (ART) [66] to compute $p_b$. Note that the idle time is geometrically distributed with mean $1/\alpha$ (where $\alpha$ is the batch generation probability) while the busy time is geometrically distributed with mean $1/q$. This gives us

$$p_b = \frac{1/q}{1/q \;+\; 1/\alpha} \tag{5.6}$$

which can be used with (5.5) to solve for $p_b$ numerically. Since all packets in the batch are transmitted as a bulk and stay together until the whole batch is successfully transmitted, the delay seen by one packet is the same as the delay seen by the whole batch which is simply the duration of the busy period. Therefore the mean packet delay in K-Aloha $T_{\text{K-Aloha}}$ is given by:

$$T_{\text{K-Aloha}} = \frac{1}{q} = \frac{1}{p\,(1 - p_b\,p\,p_c)^{N-1}}.$$

### *5.3.2 EBF Delay Analysis*

In the EBF algorithm, the user also alternates between idle and busy periods. However, during the busy period the user rotates between $M = \log_2(K)+1$ backoff stages continuously until it becomes idle when the whole batch is transmitted successfully. Therefore, the busy period is divided into $M$ sub-periods which have different characteristics. At stage $m$, a busy user transmits with probability $p_m$ on a branch of width $2^m$ at most. To simplify the analysis we assume that each user always selects a branch of width $2^m$ even if his queue has less than $2^m$ packets. In this case, the number of possible branches that the user could choose from is $K/2^m$. Therefore, the pair-wise collision probability between any two active (transmitting) users at stage $m$ is always fixed and equal to $2^m/K$. The success probability of the tagged user at stage $m$ is given by

$$q_m = p_m \left(1 - p_b\,p_m\,\frac{2^m}{K}\right)^{N-1}, \tag{5.7}$$

where $p_m$ is the transmission probability at stage $m$ (see Table 5.1). The average success probability during the entire busy period is not simply $\sum q_m/M$ because not all packets get the same chance of transmission in every stage. While the first packet of the batch can be transmitted in all stages, the second packet cannot be transmitted at stage $m = 0$ because the branch width at the $0^{\text{th}}$ stage can only accommodate one packet position ($2^0 = 1$) which goes to the first packet. Similarly, the $n^{\text{th}}$ packet in the batch can only be transmitted in stages higher than $\log_2(n)$. Therefore the average success probability seen by the $n^{\text{th}}$ packet

in the batch is given by

$$w(n) = \frac{1}{M} \sum_{m=v(n)}^{M-1} q_m$$
$$= \frac{1}{M} \sum_{m=v(n)}^{M-1} p_m \left(1 - p_b\, p_m \frac{2^m}{K}\right)^{N-1}, \tag{5.8}$$

where $v(n) = \lceil \log_2(n) \rceil$ and $q_m$ is given by (5.7). From Table 5.1, we see that $p_m = \frac{K}{N2^m}$ for $m \geq \lceil \log_2(\frac{K}{N}) \rceil$ and $p_m = 1$ otherwise. However, since $v(n) \geq \lceil \log_2(\frac{K}{N}) \rceil$ for most of the sample space of our batch size distribution (e.g. for $n = 3, \cdots, K$), we can approximate the summation as follows:

$$w(n) \approx \frac{1}{M} \sum_{m=v(n)}^{M-1} \frac{K}{N2^m} \left(1 - p_b \frac{K}{N2^m} \frac{2^m}{K}\right)^{N-1}$$
$$= \frac{K}{MN} \left(1 - \frac{p_b}{N}\right)^{N-1} \sum_{m=v(n)}^{M-1} \frac{1}{2^m}$$
$$= \frac{2K}{MN} \left(1 - \frac{p_b}{N}\right)^{N-1} \left[(1/2)^{v(n)} - (1/2)^M\right]. \tag{5.9}$$

Note the dependency on the packet position within the batch through $v(n)$ and the busy probability $p_b$. The delay seen by the $n^{\text{th}}$ packet is geometrically distributed with mean

$$D_n = \frac{1}{w(n)}. \tag{5.10}$$

To evaluate the unknown $p_b$, we use ART as we did for the K-Aloha case. In the EBF algorithm, the user stays busy until the last packet of the batch is successfully transmitted. Therefore, the total busy duration is just the delay seen by the last packet in the batch. We utilize this and the distribution of the batch size in (5.1) to derive the average busy period duration as follows:

$$\text{E[busy time]} = \sum_{n=1}^{K} D_n \mathbf{B}_n = \sum_{n=1}^{K} \frac{\mathbf{B}_n}{w_n}$$
$$\approx \frac{MN}{2K\,(1 - p_b/N)^{N-1}} \sum_{n=1}^{K} \frac{\mathbf{B}_n}{\left[(1/2)^{v(n)} - (1/2)^M\right]}$$
$$= \frac{c}{(1 - p_b/N)^{N-1}}, \tag{5.11}$$

where the constant $c$ can be evaluated from the batch size distribution $\mathbf{B}_n$ defined in (5.1) and system parameters. From ART, we get

$$p_b = \frac{\frac{c}{(1-p_b/N)^{N-1}}}{\frac{c}{(1-p_b/N)^{N-1}} + \frac{1}{\alpha}}. \tag{5.12}$$

which can be solved numerically for the busy probability $p_b$. We can then compute $w(n)$ from (5.9) for all $n = 1, \cdots, K$. Next, we compute the average success probability of a packet in a batch of size $i$ as follows:

$$\bar{s}_i = \frac{1}{i} \sum_{n=1}^{i} w(n). \tag{5.13}$$

Conditioning on the batch size $i$, we get the average success probability $\bar{s}$ of an arbitrary packet:

$$\bar{s} = \sum_{i=1}^{K} \bar{s}_i B_i. \tag{5.14}$$

Finally, since the delay is geometrically distributed with the success probability, the mean packet delay of an arbitrary packet is given by

$$T_{\text{EBF}} = 1/\bar{s}. \tag{5.15}$$

### *5.4 Numerical and Simulation Results*

To verify the approximate analysis above, we developed a simulation model in NS-3 that implements the details of the EBF and K-Aloha protocols. The code provides a custom NS-3 stack implementing a simplified PHY layer on top of $K$ parallel collision channels based on the Bulk Access model described in Section 5.1 and a full MAC layer for three protocols: EBF, K-Aloha and Aloha. Also, we developed custom NS-3 traffic generators for various traffic models including the Binomial Batch Arrival idle/busy model described at the beginning of this section. In the simulation, only the downlink (Femto BS $\rightarrow$ user) traffic is considered. A perfect feedback channel is also used to inform the BS of the result of the transmission at the end of the slot. The overall simulation model follows the abstract system model depicted in Figure 2.12.

We simulated a network of $N = 15$ femtocells competing for $K = 32$ channels and mean batch size $\mathbf{E}[B] = 10$ packets. We varied the input traffic load by changing the

arrival probability $\alpha$. In Figure 5.4, we plot the mean packet delay of EBF, K-Aloha and Aloha against the arrival rate $\lambda = \alpha \mathbf{E}[B]$ per user. The symbols denotes the analytical results while the lines denotes the simulation results. As expected, there is a gap between the analytical and simulation results of the EBF protocol due to the approximation used in (5.9). Our proposed EBF MAC algorithm achieves considerably smaller packet delay compared to K-Aloha and Aloha. The clustering of packets' transmissions at few discrete branches performs better than simply spreading them uniformly over the $K$ channels. Also, the synchronized nature of EBF ensures that all contending nodes have the same branch width $2^m$ at each stage of the backoff period, hence the pair-wise collision probability is $p_c = 2^m/K$ at most. The worst case is the simple Aloha protocol because it does not spread the bulk over the $K$ channels regardless of the bulk size and hence it has the largest pair-wise collision probability ($p_c = 1$). K-Aloha enhances the packet delay slightly over Aloha because its average pair-wise collision probability is $p_c < 1$.

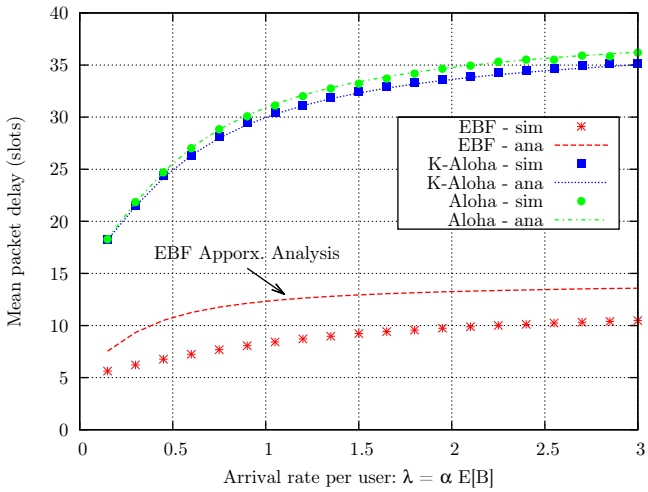


Figure 5.4: Mean packet delay; $N = 15$, $K = 32$, $\mathbf{E}[B] = 10$

In addition to the simulation above, we implemented the more general buffered model in which the input traffic is not restricted by the idle/busy state of the user. In this model, packets that arrive while the user is busy transmitting/retransmitting the Head of Line (HOL) packet are queued in a buffer. The analysis of this buffered model is challenging

because of the interaction between multiple competing queues, hence it is not considered here. However, we present simulation results of this model in order to check the performance of the EBF in the general case. In Figure 5.5, we plot the mean packet delay of the three MAC protocols in a network of $N = 15$ femtocells competing for $K = 32$ channels. The traffic model here is a Batch Bernoulli arrival process with an arrival probability $\alpha$ in every slot and a batch size geometrically distributed with mean $\mathbf{E}[B] = 10$. The input load was varied from 0 up until all the three MAC protocols start to saturate their queues. Clearly, the EBF protocol outperforms K-Aloha and Aloha for most of the unsaturated region. Notice that the packet delay grows unbounded as we approach the saturation region in contrast to Figure 5.4. This is because in this buffered-model, there is no blocking of packet generation when the user becomes busy as opposed to the idle/busy model analyzed above.

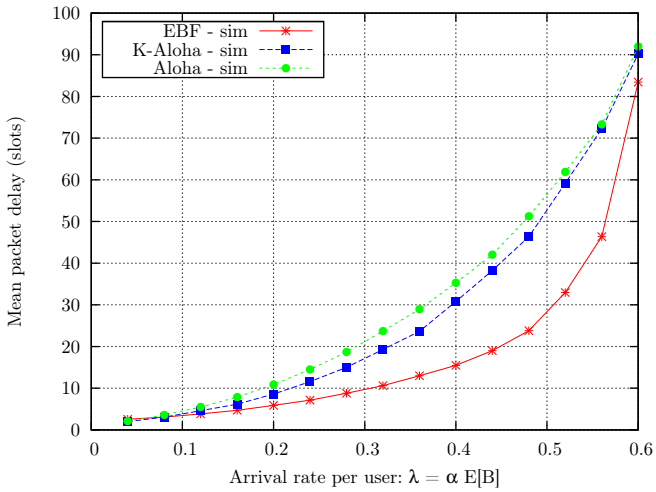


Figure 5.5: Mean packet delay in buffered model; $N = 15$, $K = 32$, $\mathbf{E}[B] = 10$

## *5.5 Conclusion*

OFDMA-based random access MAC can utilize the frequency dimension for collision resolution more effectively. Multiple packets from multiple sources can be transmitted over multiple subchannels with a low collision rate if the MAC is designed properly. This chap-

ter proposed a new Exponential Backoff in Frequency MAC algorithm for random access in OFDMA femtocells. The algorithm tries to keep packets transmission clustered at few discrete frequency branches in order to reduce collisions that usually result from random channel selection as in K-Aloha. Approximate analysis of the mean packet delay and simulation results show that the EBF algorithm reduces packet delay considerably compared to basic Aloha and its multichannel variant (K-Aloha) in a dense network of OFDMA femtocells.

# Chapter 6

# OFDM-BASED RESERVATION RANDOM ACCESS

There are multiple ways in which an OFDM-aware random access protocol can exploit the time-frequency grid to improve performance. The simplest method is by varying the location (index) of the occupied OFDM sub-channel randomly after each collision as done in the OFDMA-Aloha protocol studied in Chapter 4. A second way is by varying the number of occupied sub-channels (bandwidth) after each collision as done in the EBF protocol proposed in Chapter 5. These can be regarded as conventional MAC protocols in which the right to channel access is gained via random contention among competing data packets. In a time-slotted MAC model, the slot size is usually taken to be the transmission time of a single packet. The channel contention time, i.e. the time between successful slots, is an integral multiple of the slot size thus can be quite significant if the packet size is large.

To handle this issue, a different class of MAC protocols called reservation protocols are often used. Reservation protocols such as Reservation-Aloha (R-Aloha) and the RTS/CTS mode of the WiFi MAC try to reduce the contention time by first reserving the channel using very short packets called Reservation Requests in a random contention phase and then sending large data packets in the transmission phase without collisions. Since, the contention time is a multiple of the short reservation slot, the MAC efficiency is improved significantly especially with large data packets.

In this chapter, we propose and analyze an OFDM-aware Reservation-based random access protocol for handling inter-cell interference in OFDMA femtocells. In this protocol, each femtocell transmits a short reservation request on a randomly chosen subchannel and then utilizes receiver-feedback information about the indices of the subchannels occupied by other femtocells to determine the status of its reservation attempt. An implicit ordering of femtocells based on subchannels' indices is used to resolve conflicts in almost one slot which reduces the reservation time significantly. Although, this idea uses simple ordering

in the frequency dimension, the concept can be extended to more sophisticated reservation schemes that exploit both time and frequency dimensions, i.e. reservation over the 2-Dimensional time-frequency grid which can potentially improve the performance of the reservation system.

Since this kind of reservation in the frequency dimension has not been analyzed before, it is important to analyze its maximum throughput when all transmitters are active all the times. This first-order analysis helps understand the relationship between the performance of the protocol and the network load and system parameters. It can be considered as a first step toward a more thorough analysis that accounts for the activity in each femtocell. In this chapter, we will derive the throughput of this protocol and present numerical and simulation results to compare it to the throughput of the canonical R-Aloha protocol.

## *6.1 Proposed Random Access Protocol*

Several works tried to extend the concept of Reservation-Aloha to OFDMA networks as discussed in Section 2.3.2. However, a novel way of utilizing frequency domain backoff in a distributed protocol was proposed in [67] in the context of WiFi networks. In standard WiFi (IEEE 802.11) systems, each user draws a random count-down timer between 1 and CW, the contention window size, and starts counting down after the channel is sensed free. The countdown continues at each slot where the channel is idle (and is suspended when it is not); the user whose count-down timer expires first starts transmission and all other users freeze their count-down timers. This effectively creates a random ordering among users; the user with the minimum slot number wins the contention. We can in principle substitute OFDM sub-carrier index for the time slot index and achieve the same random ordering among users in one time slot.

The protocol in [67] works as follows. A ready user listens to the channel and once it is determined to be free, selects an OFDM sub-carrier index randomly out of the 48 data sub-carriers in WiFi and transmits a reservation request (essentially a busy tone) on the selected sub-carrier. Using a second antenna for listening, the user detects any signal presence on all 48 sub-carriers and then compares the index of its sub-carrier to the indices of all other users' sub-carriers. The user with the highest index wins the contention and

starts transmitting the data packet on the full channel (all 48 sub-carriers). Therefore, the backoff procedure (or conflict resolution) is performed using frequency slots (sub-carrier index) rather than time slots, hence the notion of frequency-domain backoff.

If two or more users select the same "highest-index" sub-carrier, a collision would occur on the subsequent packet transmission. To resolve this conflict, a second round of contention follows the first. Only the users who selected the "highest-index" sub-carrier in the first round are admitted into the second round and the same procedure is repeated. Collisions could still occur since the same conflict can happen again in the second round, but the experience in [67] showed that two rounds of contention are sufficient for conflict resolution in most practical scenarios. An associated problem is the lack of a reliable way of determining the status of the reservation request. The suggested method in [67] of using a second "listening" antenna is susceptible to strong self-interference from the transmitting antenna. No performance analysis was presented in [67] to assess these issues.

We generalize this idea into a collision-free reservation-based random access called OFDM-based Reservation Random Access (OFDM-RR) protocol. The key extension consists of allowing a variable number of contention rounds for reservation which guarantee subsequent collision-free data transmission. Although it is applicable to any slotted OFDM wireless system with a feedback mechanism, we propose the OFDM-RR protocol for resolving inter-cell interference in a dense network of OFDMA femtocells.

### *6.1.1 Protocol Description*

We will first describe the proposed random access algorithm in a generic OFDMA-based femtocell network and then we will describe how it can be used in LTE-based femtocells. Consider a cluster of co-located femtocells each serving a single user as depicted in Figure 6.1. All femtocells belong to the same network provider and thus operate on the same licensed spectrum. It is assumed that the nearby macro-cells operate on different spectrum so that there is no macro-to-femto interference. In this paper, we only consider the downlink direction, but the same protocol can be applied with some modifications to the uplink direction.

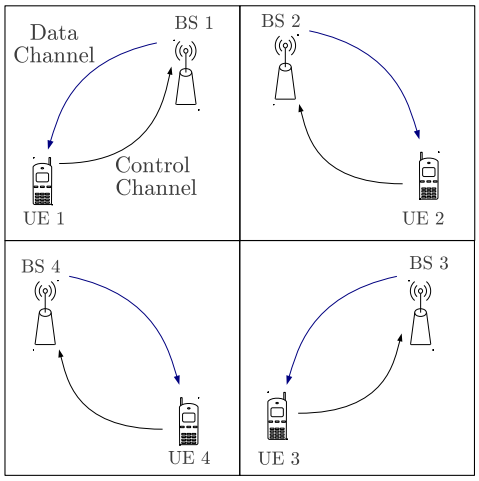


Figure 6.1: A cluster of 4 femtocells each serving one user. The downlink data channel carries reservation requests and data packets while the uplink control channel carries reservation status reports.

The OFDM-RR protocol works as follow. The channel is divided into $K$ sub-channels and time is divided into equal slots. Whenever the channel becomes available for reservation (i.e. at the end of a packet transmission), each femtocell BS picks a number $k = 1, \ldots, K$ uniformly randomly and transmits a short reservation request on the $k$th sub-channel. The reservation packet contains a unique femtocell ID and a checksum. The corresponding receiver decodes each sub-channel and compares any received request with the locally stored ID. At the end of the slot, the receiver sends back a report indicating the status of any reservation attempt in each sub-channel. Next, the femtocell compares its chosen value of $k$ with that of other femtocells (which is inferred from the status report as will be explained shortly). The femtocell who picked the highest sub-channel index $k_{\max}$ wins the reservation. If the highest occupied sub-channel's index $k_{\max}$ was picked by two or more femtocells, a collision occurs and a second round of conflict resolution is needed. Only those femtocells who picked $k_{\max}$ in this round are admitted into the next round. Other femtocells stop their reservation attempts but continue monitoring the channel (via uplink status reports) to determine when it becomes available for reservation again. The conflict resolution process is repeated until the highest occupied sub-channel $k_{\max}$ is picked by exactly one femtocell.

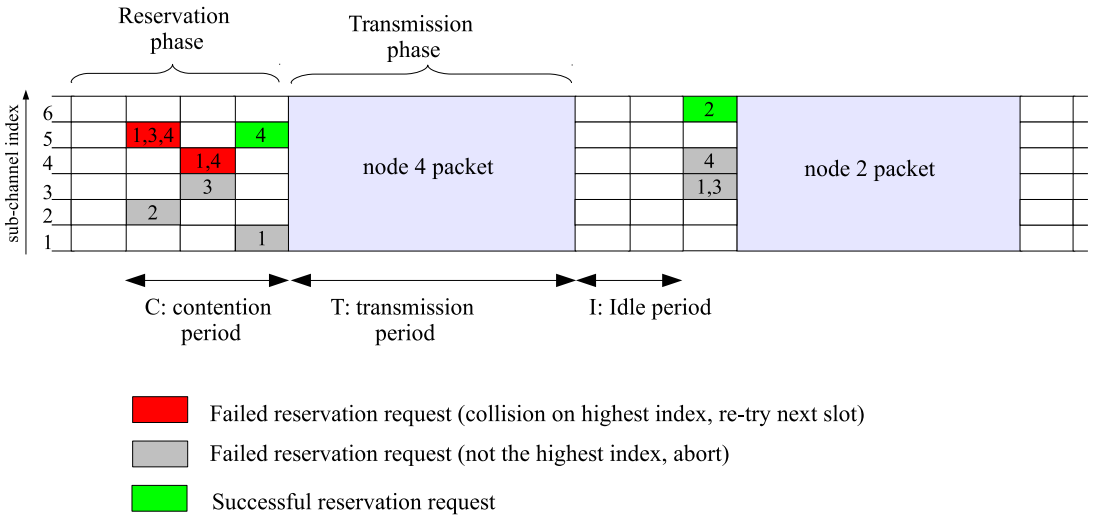


Figure 6.2: OFDM-RR protocol example with $N = 4$ nodes and $K = 6$ sub-channels

Once the reservation is established, the femtocell can occupy the whole channel for a fixed interval to transmit its data packets, possibly to more than one user.

Figure 6.2 illustrates an example of OFDM-RR with $N = 4$ femtocells and $K = 6$ sub-channels. In the first attempt after an idle period, the random values of the sub-channel index $k$ picked by femtocells 1 to 4 is $(5, 2, 5, 5)$ respectively. The highest occupied sub-channel index is $k_{\max} = 5$ which is picked by femtocells 1, 3 and 4, hence a collision occur and the three femtocells continue to the next round while femtocell 2 quits. In the second round, femtocells 1 and 4 collide on the highest occupied sub-channel index so they proceed to a third round. In the third round, the highest occupied sub-channel index $k_{\max} = 5$ is picked by exactly one femtocell, thus the reservation is successfully established for femtocell 4. After the transmission of femtocell 4's packet, the channel becomes idle and the reservation process restarts again. It can be seen here that femtocell 2 was able to reserve the channel from the first attempt. Clearly, the number of rounds needed to resolve the conflict is a random variable that depends on the network load and the degree of channelization $K$.

### *6.1.2 Using OFDM-RR in LTE Femtocells*

In LTE, the channel bandwidth is divided into resource blocks (RB) of 12 OFDM subcarriers each as explained in Section 2.4. For example, a 10 MHz LTE downlink channel contains about 50 RBs. Resource blocks are grouped into Resource Block Groups (RBG) (subchannels) which can be allocated to UEs every 1 ms Transmission Time Interval (TTI). The group size is a configuration parameter and therefore a variable number $K$ of subchannels can be realized [31]. The random access policy can be implemented in the scheduling algorithm as explained in Section 2.4. The feedback mechanism required for the proposed OFDM-RR can be implemented using the Physical Uplink Control Channel (PUCCH), see Figure 2.6. With OFDM-RR, the BS scheduler sends the reservation request which contains a unique cell ID and a checksum on one of the $K$ subchannels (RBG) and awaits for the feedback from the UE on the PUCCH. The UE tries to decode any transmission on all $K$ downlink subchannels. If a subchannel contains a valid reservation request (i.e. with a correct checksum), the UE compares the cell ID in the request with its locally stored cell ID. If the request matches its cell ID, then the status of this subchannel is "success". Otherwise, if the checksum fails or the request is valid but with another Cell ID, then the subchannel status is "error". If nothing is detected on the subchannel, the status is reported as "idle". The UE then transmits the full status report of all $K$ subchannels to its BS using the PUCCH. The BS scheduler then uses the status report to determine the result of the reservation attempt as explained above.

## *6.2 Throughput Analysis*

We now analyze the performance of the proposed OFDM-RR protocol in terms of the maximum MAC throughput. We will first compare the throughput of this protocol to that of Reservation-Aloha (R-Aloha) with the same total bandwidth. Here, we will assume that R-Aloha has access to the full channel bandwidth, and hence can transmit the reservation request $K$ times faster than an OFDM-RR system with $K$ subchannels. This means the reservation slot in R-Aloha is $1/K$ shorter than the reservation slot in OFDM-RR. This provides a fair comparison in terms of the normalized throughput (packet transmission

time over total reservation and transmission time). Then, we will look at the throughput when both protocols are used in a system with fixed slot size such as LTE in the following section.

### *6.2.1 System Model*

To evaluate the baseline performance of OFDM-RR, we use a simplified MAC model that captures the key features of OFDM-RR in a generic setting. We model the femtocell cluster in Figure 3.1 as a group of $N$ wireless links belonging to a single collision domain, which means all transmissions in the group interfere with each other. This assumption greatly simplifies the analysis and is commonly used in studying random access protocols [65]. Each link consists of a transmit-receive pair, i.e. a BS serving a single user [1]. A single wireless channel of bandwidth $W$ is shared among all cells, where channel access is determined using the reservation-based random access protocol. Each transmitting node first attempts to reserve the channel by sending a short reservation request of size $m$ bits; the successful node then transmits the data packet in the subsequent slot without any collision. The data packet size $M$ is fixed and assumed to be an integer multiple of the reservation request size.

We are interested in comparing the performance of OFDM-RR to Reservation-Aloha (R-Aloha) which was originally proposed to improve the efficiency of Slotted Aloha [68]. There are many variant systems that employ the basic idea of R-Aloha but differ in how the reservation process is implemented. In this paper, we adopt the "canonical" reservation system described in [69]. In this system, it is assumed that each node receives an instantaneous ternary feedback {0, 1, e} report indicating the idle, success or collision status of any reservation attempt in the previous slot. If there was a collision or the slot was idle, nodes continue to contend for reservation. If the reservation request was successful, the winning node occupies the channel to transmit a data packet.

In extending this "canonical" scheme to the case of OFDM-RR where the channel is divided into many sub-channels, we require that each node receives a ternary feedback {0,

[1]Since we are looking at the inter-cell contention (interference), the number of users in each cell simply scales the offered load with no other impact, and the single-user assumption can be invoked without loss of generality.

1, e} report per sub-channel at the end of the slot. This is necessary for the node to determine its "rank" among contending nodes. If the sub-channel chosen by the node has a "success" (1) status and all higher order sub-channels have an "idle" (0) status, then the reservation is successful. If the chosen sub-channel has a "collision" (e) status and all higher-order sub-channels are idle, then the node can retry the reservation request in the next slot. In all other cases, the node stops the reservation attempt, but continues to monitor the channel until all sub-channels become idle, indicating the end of a packet transmission and the start of a new reservation cycle.

For comparing the throughput of R-Aloha and OFDM-RR, our analysis assumes all nodes are saturated. Whenever the channel becomes available (i.e. at the end of a packet transmission), a node transmits a reservation request with probability $p$. In the R-Aloha protocol, if there is a collision, the process is repeated in the next slot until the channel is successfully reserved. In OFDM-RR, if more than one node transmit their reservation request in the same slot, those nodes enter the conflict resolution phase while all other nodes stop contending until the reservation is completed and a packet is transmitted.

Time is slotted and we assume all transmissions are synchronized to the beginning of each slot. The slot size is taken to be the the transmission time of the reservation packet in each respective protocol. Therefore, if we let $\tau$ be the slot size in R-Aloha, then the corresponding slot size in OFDM-RR is $K\tau$ because here the reservation packet is transmitted in a single sub-channel whose bandwidth is $1/K$ of the whole channel bandwidth. The key parameter is the ratio of the transmission time of the reservation request to the transmission time of the data packet $a = \tau/T$. In saturation, the total offered load (attempt rate) is $G = Np$. The throughput $S$ is defined as the proportion of time in which the channel is occupied by useful transmissions (i.e. data packets). Labeling R-Aloha as protocol 1 and OFDM-RR as protocol 2, we proceed to derive $S_1$ and $S_2$ in terms of $G$, $a$ and $K$.

### *6.2.2 Throughput of R-Aloha*

The throughput of R-Aloha has been derived assuming an infinite Poisson network in [70], but for consistency we re-derive it here for our finite-network system model. On the channel,

we observe a sequence of reservation and transmission cycles. The length of the reservation cycle is a geometric random variable with mean $1/p_s$ slots where $p_s = Np(1-p)^{N-1}$ is the probability of a successful reservation in any given slot. Hence the throughput is

$$\begin{aligned} S_1 &= \frac{T}{T + \frac{\tau}{Np(1-p)^{N-1}}} \\ &= \frac{Np(1-p)^{N-1}}{Np(1-p)^{N-1} + a} \\ &\approx \frac{Ge^{-G}}{Ge^{-G} + a} \end{aligned} \tag{6.1}$$

where the last approximation holds for large $N$ [2]. Note that this is slightly different than the result in [70], because it was assumed that each successful packet transmission was followed by a full idle slot before nodes can resume contention. We do not require this idle slot after each successful packet transmission because it would make the comparison unfair as the slot size in OFDM-RR is $K$ times larger than in R-Aloha. Instead, we assume nodes instantly know when the channel becomes available from the instantaneous feedback as described in the system model.

### *6.2.3 Throughput of OFDM-RR*

In OFDM-RR, the channel time can also be viewed as a sequence of reservation and transmission cycles, see Fig 6.2. Since nodes transmit their reservation requests with probability $p$, some slots may remain idle if no requests are transmitted. Once a request or more is transmitted, the channel becomes busy until contention is resolved and a data packet is transmitted. Therefore, the reservation cycle consists of two periods: an Idle period followed by a Contention period. Let $I = 0, 1, \ldots$ and $C = 1, 2, \ldots$ be the number of slots in the Idle period and the Contention period respectively. The random variable $I$ is geometrically distributed with mean $(1-q)/q$ where $q$ is the probability of at least one reservation

[2] From $\lim_{N\to\infty}(1-p)^N = \lim_{N\to\infty}(1-G/N)^N = e^{-G}$

attempt in a slot. In terms of the slot size $K\tau$, we have:

$$\begin{aligned} \mathrm{E}[I] &= \frac{1-q}{q} \times K\tau \\ &= \frac{(1-p)^N}{1-(1-p)^N} \times K\tau \end{aligned} \tag{6.2}$$

The derivation of the mean contention time is more challenging. Note that we need at least one slot to resolve the contention. To derive $\mathrm{E}[C]$, we cast the contention resolution process as the following game. We have $n$ labeled balls and $K$ labeled urns. The urns are arranged in a row from left to right starting with urn 1 and ending with urn $K$. The game starts by placing each ball randomly in one of the urns and observing the rightmost *occupied* urn which we label as $k_{\max}$. If the number of balls in urn $k_{\max}$ is exactly 1, the game ends. Otherwise, we repeat the game with only the balls in urn $k_{\max}$ and discard all other balls. We are interested in the expected number of rounds required to end the game.

Let $X_n(t)$ be the number of balls in urn $k_{\max}$ at the end of round $t$ given that we start with $n$ balls. Then, $X_n(t)$ is an absorbing Markov chain with an initial state $X_n(0) = n$ and an absorbing state at $X_n(t^*) = 1$. Let $A(n,k)$ be the number of different ways to place $n$ labeled balls in $k$ labeled urns. Because balls are placed in urns $1,\ldots,k$ independently, we have $A(n,k) = k^n$. To derive the transition probability $P_{i,j}$, we need to count the number of arrangements in which $j$ balls are placed in urn $k_{\max}$ while $i-j$ balls are placed in the remaining urns $1,\ldots,k_{\max}-1$. Conditioning on $k_{\max} = r$, we have exactly $\binom{i}{j}A(i-j,r-1)$ ways in which $j$ balls are placed in the $r^{\text{th}}$ urn and all urns $r+1,\ldots,K$ are empty. This gives us

$$\begin{aligned} P_{i,j} &= \sum_{r=1}^{K} \frac{\binom{i}{j}A(i-j,r-1)}{A(i,K)} \\ &= \sum_{r=1}^{K} \frac{\binom{i}{j}(r-1)^{i-j}}{K^i} \\ &= \binom{i}{j}\frac{1}{K^i}\sum_{m=0}^{K-1} m^{i-j}. \end{aligned} \tag{6.3}$$

Therefore, the absorbing chain is fully described by the following lower triangular transition

matrix $\mathbf{P}$:

$$P_{i,j} = \begin{cases} 1 & \text{if } i = j = 1 \\ 0 & \text{if } j > i \\ \binom{i}{j} \frac{1}{K^i} \sum_{m=0}^{K-1} m^{i-j} & \text{if } j \leq i, \quad i = 2, \ldots, n \end{cases} \tag{6.4}$$

Note that $P_{i,j}$ is a function of $K$ and so is the expected time to absorption. Denote by $d_n(K)$ the expected time to absorption of the Markov chain $X_n(t)$ starting at state $n$. Before we proceed to derive the exact expression for $d_n(K)$, we prove the following bound which will be useful later.

**Proposition 3.** *For the absorbing Markov chain $X_n(t)$ described above, the expected absorption delay $d_n(K)$ satisfies*

$$1 < d_n(K) < n + 1 \tag{6.5}$$

*for all $K \geq 2$ and $n \geq 2$*

*Proof.* See Appendix 6.6 for the complete proof. Here we note that the lower bound is straightforward because we need at least one round to go from any state to the absorbing state. For the upper bound, we observe that the triangular structure of $\mathbf{P}$ implies that once we are in a state we can only go to lower states. The worst case path from state $n$ to the absorbing state occurs when all states $n-1, n-2, \ldots, 2$ are visited in sequence. Therefore, the sum of the expected time spent in each state gives us an upper bound to the absorbing delay of state $n$. □

A closed form expression for $d_n(K)$ can be derived using the theory of absorbing Markov Chains [71]. Writing $\mathbf{P}$ in the canonical form

$$\mathbf{P} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{R} & \mathbf{Q} \end{pmatrix}$$

where $\mathbf{Q}$ is the $n-1 \times n-1$ sub-matrix of transition probabilities among the transient states $i = 2, \ldots, n$. Using the Fundamental matrix of the chain $(\mathbf{I} - \mathbf{Q})^{-1}$ and defining the vector $\mathbf{e}^T = [0\ 0 \ldots 0\ 1]$, we have from [71, Ch.3]:

$$d_n(K) = \mathbf{e}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1}\,, \qquad n \geq 2. \tag{6.6}$$

By taking the expectation over all $n = 1, \ldots, N$ and defining $d_1(K) = 1$, we get the expected length of the Contention period $\mathrm{E}[C] = \mathrm{E}[d_n(K)]$. Let $w_n$ be the probability that we start the Contention period with $n$ nodes. Since each user transmits the reservation request with probability $p$, we have $w_n = \binom{N}{n} p^n (1-p)^{N-n}$ and $\mathrm{E}[C]$ is given by:

$$\mathrm{E}[C] = \sum_{n=1}^{N} \frac{w_n}{1-(1-p)^N} \, d_n(K) \times K\tau \tag{6.7}$$

In the above, the division by $1-(1-p)^N$ is necessary since we are considering the probability of $n$ active users given that we already started the contention period. The throughput $S_2$ of OFDM-RR follows from (6.2) and (6.7):

$$\begin{aligned} S_2 &= \frac{T}{T + \mathrm{E}[I] + \mathrm{E}[C]} \\ &= \frac{T}{T + \frac{(1-p)^N}{1-(1-p)^N} \times K\tau + \sum_{n=1}^{N} \frac{w_n}{1-(1-p)^N} \, d_n(K) \times K\tau} \\ &= \frac{1-(1-p)^N}{1-(1-p)^N + Ka(1-p)^N + Ka \sum_{n=1}^{N} w_n \, d_n(K)} \end{aligned} \tag{6.8}$$

Although (6.8) allows us to compute $S_2$ exactly for any values of $K$ and network load $Np$, the dependence on $K$ is obscured by the term $d_n(K)$ in the summand which is a function of the Fundamental matrix of the chain. We desire a simpler equation of $S_2$ as a function of $G$, $a$ and $K$ for easier comparison with $S_1$. In Appendix 6.7, we utilize an alternative approach that allows us to approximate the factor $K\, d_n(K)$ as a linear function of both $n$ and $K$ which allows an approximate, but highly accurate matrix-free equation for the throughput of OFDM-RR, as below:

$$S_2 \approx \frac{1-e^{-G}}{(1+\frac{1}{9}a)(1-e^{-G}) - \frac{11}{18}aGe^{-G} + \frac{1}{2}aG + Ka}. \tag{6.9}$$

Eq. (6.9) shows explicitly the effect of the channelization factor $K$ on the throughput of OFDM-RR. From Figure 6.3, we can see that (6.9) is a good approximation especially around the region of the maximum throughput.

### *6.3 Throughput in a Fixed-slot System*

OFDMA networks such as LTE operate with a fixed, minimum slots size where all resource allocation and scheduling actions are synchronized to the slot boundary. In this case, the

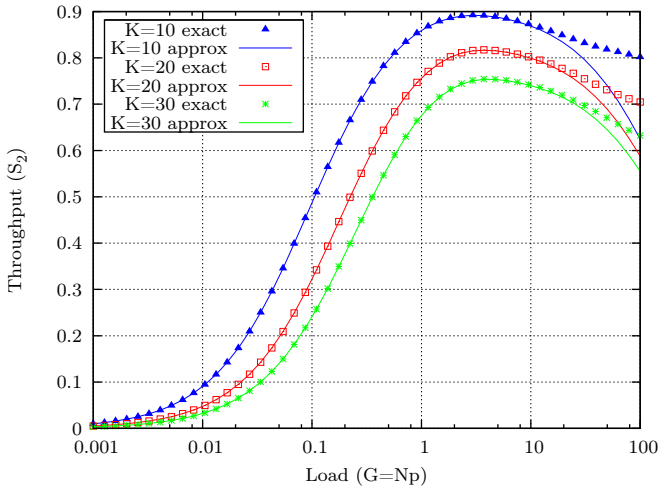


Figure 6.3: Exact vs. approximate values of $S_2$ for $a = 0.01$

reservation slot cannot be smaller than this system slot even if the reservation request can be transmitted on the full bandwidth. Also, the reservation request is a small packet containing an ID and a CRC and it can be carried by a single resource block (RB) which can fit tens of bits in one slot (e.g. 1 ms TTI). The reservation slot does not need to be greater than the system slot size even if no grouping of RBs is used to create higher capacity subchannels. Therefore, it is necessary to compare the throughput of both protocols under this system constraint.

Taking the system slot size as our reservation slot size, the throughput of R-Aloha is still governed by equation (6.1):

$$\begin{aligned} S_1 &= \frac{Np(1-p)^{N-1}}{Np(1-p)^{N-1}+a} \\ &\approx \frac{Ge^{-G}}{Ge^{-G}+a} \end{aligned} \tag{6.10}$$

However, the average idle period E[$I$] and the average contention period E[$C$] in OFDM-RR are expressed in multiples of the reservation slot which in this system model equals the slot size of R-Aloha, i.e. we don't need to multiply by $K$ in (6.2) and (6.7). Therefore, the

throughput of OFDM-RR is given by:

$$S_2 = \frac{1-(1-p)^N}{1-(1-p)^N + a(1-p)^N + a\sum_{n=1}^{N} w_n\, d_n(K)} \tag{6.11}$$

### *6.4 Discussion and Numerical Results*

In this section, we compare the throughput of OFDM-RR to that of R-Aloha using numerical and simulation results. Our results highlight the following important fact that has been under-emphasized in previous works; although partitioning the channel reduces the collision probability sharply and hence reduces the reservation time, the reservation slot itself is increased by a factor of $K$. Clearly, the reduction in collision probability must be significant enough to outweigh the impact of the bigger slot size and this can only be attained at a sufficiently high load which we quantify in the following proposition.

**Proposition 4.** *With fixed total channel bandwidth and network size $N$ and the same probability $p$ of transmission in a free slot, we have the following inequalities between the throughput of R-Aloha $S_1$ and the throughput of OFDM-RR $S_2$:*

1. *For a fixed offered load $G = Np$, we have*

$$S_1 \;>\; S_2, \quad \forall \quad K \geq K^* = \frac{e^G - 1}{G} \tag{6.12}$$

2. *For a fixed $K$, we have*

$$\begin{aligned} &S_1 \;>\; S_2, \\ &\forall \quad G \leq G^* = -\frac{1}{K} - W_{-1}\left(-\frac{1}{K}e^{-1/K}\right) \end{aligned} \tag{6.13}$$

*where $W_{-1}(x)$ is the -1 branch of the Lambert W function* [3].

*Proof.* We state the key step in the proof here and leave the details to Appendix 6.8. From Proposition 3, we have $d_n(K) > 1$ which gives us the following upper bound on $S_2$:

$$S_2 \;<\; S_2^{\max} = \frac{1-e^{-G}}{1-e^{-G}+Ka}, \tag{6.14}$$

[3] The Lambert $W$ is defined by the equation $z = W(z)e^{W(z)}$ for any complex number $z$, See full discussion in [72]

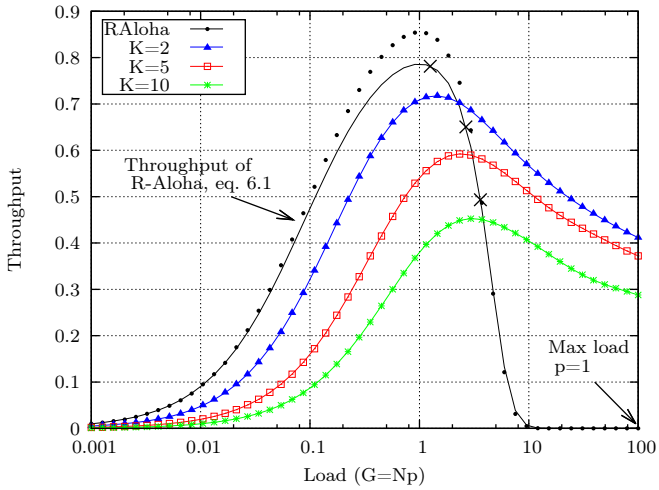


Figure 6.4: Throughput comparison between R-Aloha and OFDM-RR for $K$=2, 5 and 10 for $a = 0.1$. The value of $G^*$ is marked with $\times$ at 1.26, 2.66, 3.62 for $K$=2, 5, 10 respectively. Symbols are for simulation results and lines are for equations (6.1) and (6.8).

where $S_2^{\max}$ is derived by substituting $d_n(K) = 1$ in (6.8) and using $\lim_{N\to\infty}(1-p)^N = \lim_{N\to\infty}(1-G/N)^N = e^{-G}$. In the Appendix, we provide the complete proof by comparing $S_1$ to $S_2^{\max}$ and obtaining the critical points $K^*$ and $G^*$ defined above. □

Note that this result does not imply $S_1 < S_2$ for loads beyond $G^*$ since for $G > G^*$ we can only prove that $S_1 < S_2^{\max}$ which is the upper bound on $S_2$. However, this threshold $G^*$ is very close to the true intersection point $S_1 = S_2$ as evident from the numerical results. Also, note that this threshold is independent of the parameter $a$ which significantly determines the performance of any reservation protocol.

To verify the above equations, we simulated the two MAC protocols in a network of $N$=100 nodes with $K$=2, 5 and 10 at different load points $G$=$Np$. The simulator is a MATLAB program that implements the two protocols exactly as described in the system model in Section 6.2.1. We plot the throughput of OFDM-RR along with that of R-Aloha for $a = 0.1$ in Figure 6.4. In all cases, the simulation results closely match the numerical results obtained from (6.1) and (6.8).

Clearly, the throughput of R-Aloha is higher than OFDM-RR for most of the low load

region, i.e. below R-Aloha's saturation load at $G$=1 (which is also the saturation load in Slotted Aloha). In the high load region, OFDM-RR significantly improves the throughput of the channel even at the maximum load of $p = 1$, i.e. when every node always transmits in a free slot. The reason for this is that the collision rate increases with the load and, as we have seen, OFDM-RR reduces the collision rate sharply which amounts to higher throughput. In the low load region, the gain from this "reduction" in collision rate is not significant enough to outweigh the negative effect of increasing the slot size by a factor of $K$. We can also see from Figure 6.4 that smaller value of $K$ achieve better throughput than larger value which suggests that a minimal number of sub-channels is sufficient to alleviate contention load considerably.

Finally, it is well-known that the throughput of reservation MAC protocols improves with decreasing $a$, the ratio of the reservation request size to packet size. As shown in Figure 6.5 for $a = 0.01$, the throughput of both protocols improves compared to Figure 6.4. Although R-Aloha remains superior in this case too, the intersection point (near $G^*$) is now close to the peak of R-Aloha throughput. We also compared the two protocols for the extreme case of $a = 0.001$ and the results show further reduction in the gap between the two curves.

### *6.4.1 Results in the Fixed Slot Systems*

Here we evaluate the two protocols in a fixed-slot system with $K$ subchannels as explained in Section 6.3. In Figure 6.6, we plot the throughput of R-Aloha and OFDM-RR with $K$=2, 5, 10 for $a = 0.1$. Clearly, OFDM-RR enjoys higher throughput than R-Aloha especially at higher load (as $p \to 1$). The larger the number of subchannels $K$, the higher the throughput because the contention load reduces sharply with increasing $K$. Also, the gain with larger packet sizes (smaller $a$) is significant as depicted in Figure 6.7 where the throughput of OFDM-RR approaches 1 for $a = 0.01$.

## *6.5 Conclusions*

Frequency-domain backoff is a novel idea for the dispersion of the contention load in random access protocols at high loads. We integrated frequency-domain backoff into a more general

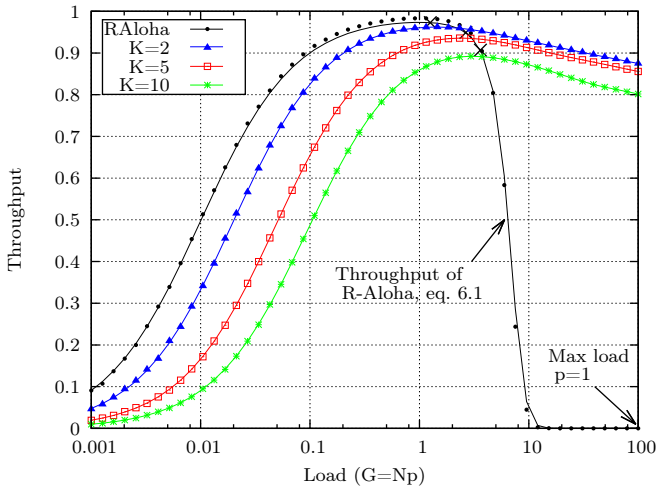


Figure 6.5: Throughput comparison between R-Aloha and OFDM-RR for $K$=2, 5 and 10 for $a = 0.01$. The value of $G^*$ is marked with $\times$ at 1.26, 2.66, 3.62 for $K$=2, 5, 10 respectively. Symbols are for simulation results and lines are for equations (6.1) and (6.8).

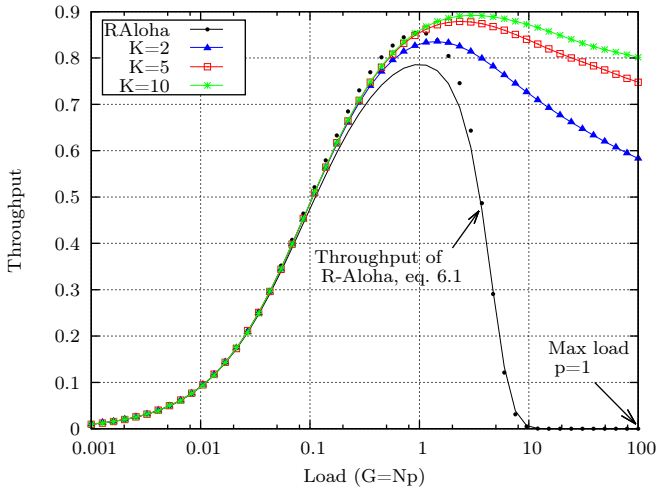


Figure 6.6: Throughput comparison between R-Aloha and OFDM-RR for $K$=2, 5 and 10 for $a = 0.1$ in the fixed slot system. Symbols are for simulation results and lines are for equations (6.1) and (6.11).

OFDM-based Reservation Random multiple access (OFDM-RR) protocol and described how it can be used to mitigate inter-cell interference in dense networks of OFDMA femtocells such as LTE Home eNBs. We then developed a unified system model that facilitates a fair

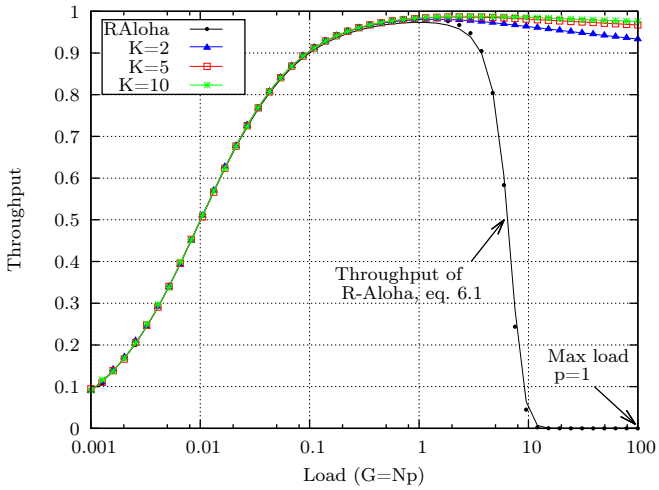


Figure 6.7: Throughput comparison between R-Aloha and OFDM-RR for $K$=2, 5 and 10 for $a = 0.01$ in the fixed slot system. Symbols are for simulation results and lines are for equations (6.1) and (6.11).

comparison between this protocol and the baseline reservation-based random access protocol (R-Aloha) and derived the throughput in terms of the network load, packet size and the number of sub-carriers. We proved that OFDM-RR throughput is superior to R-Aloha in the high load region and derived the critical value of the network load beyond which the proposed protocol significantly improves the system throughput.

## *6.6 Appendix A: Proof of proposition 3*

To prove the strict inequality $1 < d_n(K)$, we note that we need at least one round to go from state $n$ to the absorbing state. However, since we have a positive probability of staying in this state $P_{n,n} > 0$, we must have $d_n(K) > 1$. For the upper bound, we observe that the triangular structure of $\mathbf{P}$ implies that once we are in a state we can only go to lower states. The worst case path from state $n$ to the absorbing state occurs when all states $n-1, n-2, \ldots, 2$ are visited in sequence. Let $t_i(K)$ be the expected time spent in state $i$. Since the probability of exiting state $i$ is $1 - P_{i,i}$, we have $t_i(K) = 1/(1 - P_{i,i})$. From (6.4),

$P_{i,i} = 1/K^{i-1}$ and we have the following inequality:

$$d_n(K) \leq \sum_{i=2}^{n} t_i(K) \tag{6.15a}$$

$$= \sum_{i=2}^{n} \frac{1}{1 - 1/K^{i-1}} \tag{6.15b}$$

$$= \sum_{l=1}^{n-1} \left(1 + \frac{(1/K)^l}{1 - (1/K)^l}\right) \tag{6.15c}$$

$$= n - 1 + \sum_{l=1}^{n-1} \frac{(1/K)^l}{1 - (1/K)^l} \tag{6.15d}$$

For $K > 1$ and $l \geq 1$, we have the following inequality:

$$\frac{(1/K)^l}{1 - (1/K)^l} \ \leq \ \frac{(1/K)^l}{1 - 1/K} \tag{6.16}$$

The last summation in (6.15d) can be bounded as

$$\sum_{l=1}^{n-1} \frac{(1/K)^l}{1 - (1/K)^l} \ \leq \ \sum_{l=1}^{n-1} \frac{(1/K)^l}{1 - (1/K)} \tag{6.17a}$$

$$= \frac{1}{1 - (1/K)} \left[\frac{1/K[1 - (1/K)^{n-1}]}{1 - 1/K}\right] \tag{6.17b}$$

$$= \frac{1/K}{(1 - 1/K)^2} - \frac{(1/K)^n}{(1 - 1/K)^2} \tag{6.17c}$$

$$< \frac{1/K}{(1 - 1/K)^2} \tag{6.17d}$$

$$< 2 \tag{6.17e}$$

where the last inequality holds for all $K \geq 2$. Substituting this in (6.15d), we get

$$d_n(K) \ < n + 1. \tag{6.18}$$

## *6.7 Appendix B: Approximation of $f_n(K) = Kd_n(K)$*

Instead of computing the absorption delay $d_n(K)$ using the Fundamental matrix as in (6.6), we can use the following recursive relation which follows from the fact that the Markov

chain $X_n(t)$ is monotonically non-increasing:

$$\begin{aligned} d_n(K) &= 1 + \sum_{j=2}^{n} P_{n,j}\, d_j(K) \\ &= \frac{1}{1 - P_{n,n}} \left[ 1 + \sum_{j=2}^{n-1} P_{n,j}\, d_j(K) \right]. \end{aligned} \tag{6.19}$$

Now, define a new function as $f_n(K) = K\, d_n(K)$. From (6.19), we derive the first few instants of $f_n(K)$ for $n = 2, 3, 4, 5, 6$:

$$f_2(K) = \frac{K^2}{K-1} \tag{6.20}$$

$$f_3(K) = \frac{2K^3 + 3K^2}{2K^2 - 2} \tag{6.21}$$

$$f_4(K) = \frac{K^3 + 2K^2}{K^2 - 1} \tag{6.22}$$

$$f_5(K) = \frac{6K^5 + 15K^4 + 5K^3 + 15K^2}{6K^4 - 6} \tag{6.23}$$

$$f_6(K) = \frac{2K^5 + 6K^4 + K^3 + 6K^2}{2K^4 - 2} \tag{6.24}$$

Observe that the power of the numerator polynomial is always one plus the power the denominator polynomial which suggests that $f_n(K)$ can be approximated as $K + o(\frac{1}{K})$ plus some constant that depends on $n$. Using the first-order approximation around $K = 10$ and noting that $f_n'(10) \approx 1$ and the difference $f_n(10) - f_{n-1}(10) \approx 0.5$ for $n \geq 2$, we get the following linear relationship between $f_n(K)$ and both $K$ and $n$:

$$\begin{aligned} f_n(K) &\approx f_n(10) + f_n'(10)(K - 10) \\ &\approx f_2(10) + \frac{n-2}{2} + (K - 10) \\ &= K + \frac{1}{9} + \frac{1}{2}n. \end{aligned} \tag{6.25}$$

This linear approximation is highly accurate as shown in Figure 6.8 especially for larger values of $K$. Finally, we use (6.25) to simplify the sum in (6.8) as follows:

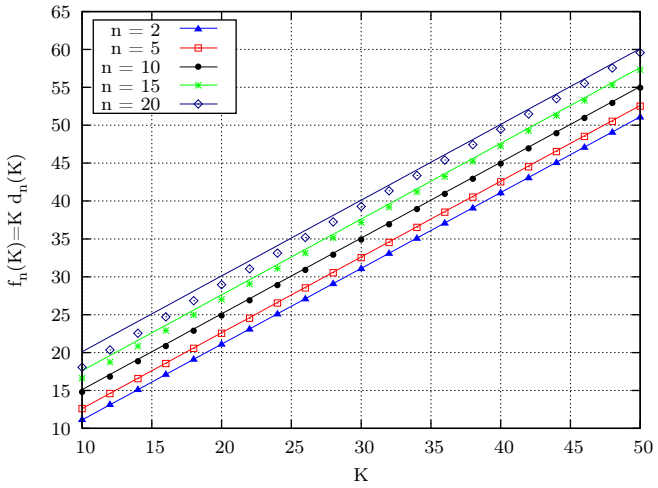


Figure 6.8: Comparing the approximation for $f_n(K)$. Solid lines for $f_n(K) \approx K + \frac{1}{9} + \frac{n}{2}$ and symbols for the exact values of $f_n(K) = Kd_n(K)$

$$
\begin{aligned}
& Ka \sum_{n=1}^{N} \binom{N}{n} p^n (1-p)^{N-n}\, d_n(K) \\
=& KaNp(1-p)^{N-1} \\
& + Ka \sum_{n=2}^{N} \binom{N}{n} p^n (1-p)^{N-n} d_n(K) \\
\approx& KaNp(1-p)^{N-1} \\
& + a \sum_{n=2}^{N} \binom{N}{n} p^n (1-p)^{N-n} (K + \frac{1}{9} + \frac{1}{2}n) \\
=& KaNp(1-p)^{N-1} \\
& + a(K + \frac{1}{9}) \left[1 - (1-p)^N - Np(1-p)^{N-1}\right] \\
& + \frac{1}{2} a \sum_{n=2}^{N} \frac{N}{n} \binom{N-1}{n-1} p^n (1-p)^{N-n} n \\
=& KaNp(1-p)^{N-1} \\
& + a(K + \frac{1}{9}) \left[1 - (1-p)^N - Np(1-p)^{N-1}\right] \\
& + \frac{1}{2} aNp \left[1 - (1-p)^{N-1}\right].
\end{aligned}
\tag{6.26}
$$

With $G = Np$, equation (6.9) follows from $\lim_{N\to\infty}(1-p)^N = \lim_{N\to\infty}(1-G/N)^N = e^{-G}$.

### *6.8 Appendix C: Proof of proposition 4*

First, we prove that $S_1 > S_2$, by showing that $S_1 > S_2^{\max}$ for all $K > K^*$. From

$$\begin{aligned} S_1 - S_2^{\max} &= \frac{Ge^{-G}}{Ge^{-G}+a} - \frac{1-e^{-G}}{1-e^{-G}+Ka} \\ &= \frac{KaGe^{-G} + ae^{-G} - a}{(Ge^{-G}+a)(1-e^{-G}+Ka)} \end{aligned} \tag{6.27}$$

Since $a > 0$ and the denominator is positive, we have

$$\begin{aligned} S_1 \;>\; S_2^{\max} \quad &\text{if} \quad KGe^{-G} + e^{-G} - 1 \geq 0 \\ \implies \quad S_1 \;>\; S_2^{\max} \quad &\forall \quad K \geq K^* = \frac{e^G - 1}{G} \end{aligned} \tag{6.28}$$

and since $S_2^{\max} > S_2$, the proof of the first part of the proposition is complete.

To prove the second part, we solve the inequality for $G$

$$\begin{aligned} S_1 \;>\; S_2^{\max} \quad &\text{if} \quad KGe^{-G} + e^{-G} - 1 \geq 0 \\ \implies \; S_1 \;>\; S_2^{\max} \quad &\text{if} \quad e^G \leq KG + 1. \end{aligned} \tag{6.29}$$

Note that $e^G$ is convex and intersects the line $KG+1$ at exactly two points, $G = 0$ and $G = G^*$. To find $G^*$, we utilize the solution for Lambert's $W$ function [72]:

$$e^{G^*} = KG^* + 1 \tag{6.30}$$

$$\implies G^* = -\frac{1}{K} - W_{-1}\left(-\frac{1}{K}e^{-1/K}\right). \tag{6.31}$$

Since $e^1 < K + 1$ for $K \geq 2$, we must have $G^* > 1$ and hence $W(.) < -1$ which implies the -1 branch of the $W$ function. Due to convexity, the curve $e^G$ lies below the line $KG+1$ for $G < G^*$ and above it for $G > G^*$. We conclude from (6.29) that $S_1 > S_2$ for all $G < G^*$ which is the required proof.

# Chapter 7

# SUMMARY AND FUTURE WORK

In this thesis, we investigated the problem of inter-cell interference in dense networks of OFDMA femtocells. We showed that the classical multi-cell resource allocation approach which handles interference by controlling either the transmit power or the bandwidth is not feasible in the femtocell case. The reason is that this approach fails to track the highly dynamic interference situation caused by random user activities. This motivated us to consider simple, and distributed random access protocols that are designed to exploit the frequency diversity in OFDMA to improve MAC efficiency. An OFDM-aware random access protocol can exploit the frequency domain in multiple ways. The simplest method is by varying the location (index) of the occupied OFDM sub-channel randomly after each collision as done in OFDMA-Aloha. A second way is by varying the number of occupied sub-channels (bandwidth) after each collision as done in the EBF protocol. A different reservation-based approach utilizes the sub-channel index in deciding who gets the channel next in a distributed way.

The overall objective of this work was to show the potential of the OFDM-aware random access concept taking OFDMA femtocells as the best usage case due to their time synchronism and light traffic loading. The concept, however, generally applies to any OFDM-based local area wireless network such as WiFi. A possible extension of this work would look at using carrier-sensing to improve the frequency-domain backoff idea. The carrier sensing result can be used as the feedback mechanism instead of relying on packet collisions as done in the above OFDM-aware random access protocols. For example, a CSMA-based OFDM-aware protocol, may start by sensing the full channel and if it is busy, the channel bandwidth is reduced and the sensing is repeated until a minimum channel bandwidth is reached.

Further work is necessary to optimize the operation of these OFDM-aware random ac-

cess protocols in practical settings. For example, the channel access probability $p$ can be optimized based on the traffic load and need not be fixed as assumed in this thesis. Another issue that was not investigated here is the hidden node problem which is common in local wireless networks. Furthermore, the fact that the receiver (UE) in OFDMA systems such as LTE has a dedicated control channel for feedback to the transmitter provides further opportunities for optimizing the random access protocol.

# BIBLIOGRAPHY

[1] (2012, Feb.) Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2011-2016. Cisco. [Online]. Available: http://www.cisco.com

[2] (2012, Oct.) The 1000x Data Challenge. Qualcomm. [Online]. Available: http://www.qualcomm.com/solutions/wireless-networks/technologies/1000x-data

[3] J. Wannstrom. (2012, Dec.) LTE-Advanced Overview. 3G Partnership project (3GPP). [Online]. Available: http://www.3gpp.org/lte-advanced

[4] (2012, Feb.) Small cells - what's the big idea? Small Cell Forum Ltd. [Online]. Available: http://www.smallcellforum.org

[5] J. Zhang and G. de la Roche and, *Femtocells: Technologies and Deployment.* John Wiley & Sons, Feb. 2010.

[6] (2010, Mar.) Interference Management in OFDMA Femtocells. Small Cell Forum Ltd. [Online]. Available: http://www.smallcellforum.org

[7] T. Zahir, K. Arshad, A. Nakata, and K. Moessner, "Interference Management in Femtocells," *Communications Surveys Tutorials, IEEE*, vol. PP, no. 99, pp. 1–19, 2012.

[8] D. Gesbert, S. G. Kiani, A. Gjendemsjo, and G. E. Oien, "Adaptation, Coordination, and Distributed Resource Allocation in Interference-Limited Wireless Networks," *Proceedings of the IEEE*, vol. 95, no. 12, pp. 2393–2409, Dec. 2007.

[9] M. Nohrborg. (2012, Aug.) LTE Overview. 3G Partnership project (3GPP). [Online]. Available: http://www.3gpp.org/LTE

[10] G. Fodor, C. Koutsimanis, A. Rcz, N. Reider, A. Simonsson, and W. Mller, "Intercell Interference Coordination in OFDMA Networks and in the 3GPP Long Term Evolution System," *Journal of Communications*, vol. 4, no. 7, Aug. 2009.

[11] G. Boudreau, J. Panicker, N. Guo, R. Chang, N. Wang, and S. Vrzic, "Interference coordination and cancellation for 4G networks," *Communications Magazine, IEEE*, vol. 47, no. 4, pp. 74–81, Apr. 2009.

[12] K. Yang, "Interference management in LTE wireless networks [Industry Perspectives]," *Wireless Communications, IEEE*, vol. 19, no. 3, pp. 8–9, jun 2012.

[13] E. Yaacoub and Z. Dawy, "A Survey on Uplink Resource Allocation in OFDMA Wireless Networks," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 2, pp. 322–337, Second Quarter 2012.

[14] D. Lopez-Perez, I. Guvenc, G. de la Roche, M. Kountouris, T. Q. S. Quek, and J. Zhang, "Enhanced intercell interference coordination challenges in heterogeneous networks," *Wireless Communications, IEEE*, vol. 18, no. 3, pp. 22–30, jun 2011.

[15] R. Madan, J. Borran, A. Sampath, N. Bhushan, A. Khandekar, and T. Ji, "Cell Association and Interference Coordination in Heterogeneous LTE-A Cellular Networks," *Selected Areas in Communications, IEEE Journal on*, vol. 28, no. 9, pp. 1479–1489, dec 2010.

[16] N. Saquib, E. Hossain, L. B. Le, and D. I. Kim, "Interference Management in OFDMA Femtocell Networks: Issues and Approaches," *Wireless Communications, IEEE*, vol. 19, no. 3, pp. 86–95, Jun. 2012.

[17] J. Kim and D.-H. Cho, "A Joint Power and Subchannel Allocation Scheme Maximizing System Capacity in Indoor Dense Mobile Communication Systems," *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 9, pp. 4340–4353, Nov 2010.

[18] J. Weitzen, M. Li, E. Anderland, and V. Eyuboglu, "Large-Scale Deployment of Residential Small Cells," *Proceedings of the IEEE*, vol. 101, no. 11, pp. 2367–2380, 2013.

[19] (2012, Aug.) Residential femtocells. [Online]. Available: http://www.ubiquisys.com/small-cells-femtocells-residential

[20] J. Mo, H.-S. W. So, and J. Walrand, "Comparison of Multichannel MAC Protocols," *Mobile Computing, IEEE Transactions on*, vol. 7, no. 1, pp. 50–65, Jan. 2008.

[21] K. Bai and J. Zhang, "Opportunistic Multichannel Aloha: Distributed Multiaccess Control Scheme for OFDMA Wireless Networks," *Vehicular Technology, IEEE Transactions on*, vol. 55, no. 3, pp. 848–855, May 2006.

[22] Y.-J. Choi, S. Park, and S. Bahk, "Multichannel random access in OFDMA wireless networks," *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 3, pp. 603–613, Mar. 2006.

[23] H. Kwon, H. Seo, S. Kim, and B. G. Lee, "Generalized CSMA/CA for OFDMA Systems: Protocol Design, Throughput Analysis, and Implementation Issues," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 8, pp. 4176–4187, Aug. 2009.

[24] Y.-J. Chang, F.-T. Chien, and C.-C. J. Kuo, "Opportunistic Access with Random Subchannel Backoff (OARSB) for OFDMA Uplink," in *Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE*, nov 2007, pp. 3240–3244.

[25] S. Valentin, T. Freitag, , and H. Karl, "Integrating Multiuser Dynamic OFDMA into IEEE 802.11 WLANs - LLC/MAC Extensions and System Performance," in *Communications, 2008. ICC '08. IEEE International Conference on*, may 2008, pp. 3328–3334.

[26] D. Wang, H. Minn, and N. Al-Dhahir, "A distributed opportunistic access scheme and its application to OFDMA systems," *Communications, IEEE Transactions on*, vol. 57, no. 3, pp. 738–746, mar 2009.

[27] K. Tan, J. Fang, Y. Zhang, S. Chen, L. Shi, J. Zhang, and Y. Zhang, "Fine-grained channel access in wireless LAN," in *Proceedings of the ACM SIGCOMM 2010 conference*, ser. SIGCOMM '10, 2010, pp. 147–158.

[28] E. Yaacoub, M. A. Al-Alaoui, and Z. Dawy, "Novel time-frequency reservation Aloha scheme for OFDMA systems," in *Wireless Communications and Networking Conference (WCNC), 2011 IEEE*, mar 2011, pp. 7–12.

[29] J. Fang, K. Tan, Y. Zhang, S. Chen, L. Shi, J. Zhang, Y. Zhang, and Z. Tan, "Fine-grained channel access in wireless lan," *Networking, IEEE/ACM Transactions on*, vol. 21, no. 3, pp. 772–787, 2013.

[30] P. Huang, X. Yang, and L. Xiao, "WiFi-BA: Choosing arbitration over backoff in high speed multicarrier wireless networks," in *INFOCOM, 2013 Proceedings IEEE*, 2013, pp. 1375–1383.

[31] E. Dahlman, S. Parkval, J. Skold, and P. Beming, *4G: LTE/LTE-Advanced for Mobile Broadband*, 2nd ed. London: Academic Press, 2014.

[32] M. L. Chaudhry and J. Templeton, *A First Course in Bulk Queues*. New York: Wiley, 1983.

[33] M. Sidi and A. Segall, "Two Interfering Queues in Packet-Radio Networks," *Communications, IEEE Transactions on*, vol. 31, no. 1, pp. 123–129, jan 1983.

[34] H. Takagi and L. Kleinrock, "Mean Packet Queueing Delay in a Buffered Two-User CSMA/CD System," *Communications, IEEE Transactions on*, vol. 33, no. 10, pp. 1136–1139, oct 1985.

[35] S. Kamal and S. Mahmoud, "A Study of Users' Buffer Variations in Random Access Satellite Channels," *Communications, IEEE Transactions on*, vol. 27, no. 6, pp. 857–868, jun 1979.

[36] T. N. Saadawi and A. Ephremides, "Analysis, stability, and optimization of slotted ALOHA with a finite number of buffered users," *Automatic Control, IEEE Transactions on*, vol. 26, no. 3, pp. 680–689, jun 1981.

[37] E. D. Sykas, D. E. Karvelas, and E. N. Protonotarios, "Queueing Analysis of Some Buffered Random Multiple Access Schemes," *Communications, IEEE Transactions on*, vol. 34, no. 8, pp. 790–798, aug 1986.

[38] T. K. Apostolopoulos and E. N. Protonotarios, "Queueing Analysis of Buffered CSMA/CD Protocols," *Communications, IEEE Transactions on*, vol. 34, no. 9, pp. 898–905, sep 1986.

[39] A. Ephremides and R.-Z. Zhu, "Delay Analysis of Interacting Queues with an Approximate Model," *Communications, IEEE Transactions on*, vol. 35, no. 2, pp. 194–201, feb 1987.

[40] A. Ganz and I. Chlamtac, "A linear solution to queueing analysis of synchronous finite buffer networks," *Communications, IEEE Transactions on*, vol. 38, no. 4, pp. 440–446, apr 1990.

[41] T. Takine, Y. Takahashi, and T. Hasegawa, "An approximate analysis of a buffered CSMA/CD," *Communications, IEEE Transactions on*, vol. 36, no. 8, pp. 932–941, aug 1988.

[42] T. Wan and A. U. Sheikh, "Performance and Stability Analysis of Buffered Slotted ALOHA Protocols Using Tagged User Approach," *Vehicular Technology, IEEE Transactions on*, vol. 49, no. 2, pp. 582–593, Mar. 2000.

[43] A. Sheikh, T. Wan, and Z. Alakhdhar, "A Unified Approach to Analyze Multiple Access Protocols for Buffered Finite Users," *Journal of Network and Computer Applications*, vol. 27, no. 1, pp. 49–76, 2004.

[44] S. C. Liew, Y. J. Zhang, and D. R. Chen, "Bounded-mean-delay throughput and non-starvation conditions in Aloha network," *Networking, IEEE/ACM Transactions on*, vol. 17, no. 5, pp. 1606–1618, oct 2009.

[45] S. Tasaka, "Dynamic Behavior of a CSMA-CD System with a Finite Population of Buffered Users," *Communications, IEEE Transactions on*, vol. 34, no. 6, pp. 576–586, jun 1986.

[46] L. Kleinrock and S. S. Lam, "Packet Switching in a Multiaccess Broadcast Channel: Performance Evaluation," *Communications, IEEE Transactions on*, vol. 23, no. 4, pp. 410–423, apr 1975.

[47] M. Peng, D. Liang, Y. Wei, J. Li, and H.-H. Chen, "Self-configuration and self-optimization in LTE-advanced heterogeneous networks," *Communications Magazine, IEEE*, vol. 51, no. 5, pp. 36–45, May 2013.

[48] (2010, Dec.) LTE Network Monitor Mode Specification v1.01. Femto Forum Ltd. [Online]. Available: http://www.smallcellforum.org/resources-technical-papers

[49] H. Holma and A. Toskala, *LTE for UMTS - OFDMA and SC-FDMA Based Radio Access*. West Sussex, United Kingdom: John Wiley & Sons, 2009.

[50] J. Wannstrom and K. Mallinson. (2014, Jan.) Heterogeneous Networks in LTE. 3G Partnership project (3GPP). [Online]. Available: http://www.3gpp.org/technologies/1576-hetnet

[51] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[52] M. Chiang, *Geometric Programming for Communication Systems*, ser. Foundations and Trends in Communications and Information Theory Series. Boston: Now, 2005.

[53] M. Grant and S. Boyd, "CVX: Matlab Software for Disciplined Convex Programming, version 2.0 beta," http://cvxr.com/cvx, Sep. 2013.

[54] H. Takagi, *Queueing Analysis: Discrete-time systems*, ser. Queueing Analysis: A Foundation of Performance Evaluation. New York: North-Holland, 1993.

[55] L. Dai, "Stability and Delay Analysis of Buffered Aloha Networks," *Wireless Communications, IEEE Transactions on*, vol. 11, no. 8, pp. 2707–2719, Aug. 2012.

[56] F. A. Tobagi, "Analysis of a Two-Hop Centralized Packet Radio Network-Part I: Slotted ALOHA," *Communications, IEEE Transactions on*, vol. 28, no. 2, pp. 196–207, Feb. 1980.

[57] R. A. Howard, *Dynamic Probabilistic Systems*. New York: Wiley, 1971.

[58] M.-K. Lo, , and T.-S. P. Yum, "The Tone Sense Multiaccess Protocols with Partial Collision Detections (TSMA/PCD) for Packet Satellite Communications," *Communications, IEEE Transactions on*, vol. 41, no. 6, pp. 820–824, Jun. 1993.

[59] T. Wan and A. Sheikh, "Performance Analysis of Buffered CSMA/CD Systems," *Wireless Personal Communications*, vol. 18, pp. 45–65, Jul. 2001.

[60] S. J. Mason and H. J. Zimmermann, *Electronic Circuits, Signals, and Systems*. New York: Wiley, 1960.

[61] B. S. Tsybakov and M. V. Mikhailov, "Ergodicity of a Slotted ALOHA System," *Problemy Peredachi Informatsii*, vol. 15, no. 4, pp. 73–87, 1979.

[62] W. Luo and A. Ephremides, "Stability of N Interacting Queues in Random-Access Systems," *Information Theory, IEEE Transactions on*, vol. 45, no. 5, pp. 1579–1587, Jul. 1999.

[63] R. R. Rao and A. Ephremides, "On the Stability of Interacting Queues in a Multiple-Access System," *Information Theory, IEEE Transactions on*, vol. 34, no. 5, pp. 918–930, Sep. 1988.

[64] W. Yue and Y. Matsumoto, *Performance Analysis of Multi-Channel and Multi-Traffic on Wireless Communication Networks*. New York: Kluwer Academic, 2002.

[65] Raphael Rom and Moshe Sidi, *Multiple Access Protocols: Performance and Analysis*. New York: Springer-Verlag, 1990.

[66] S. M. Ross, *Stochastic Processes*, 2nd ed. Boston: John Wiley & Sons, 1996.

[67] S. Sen, R. R. Choudhury, and S. Nelakuditi, "No time to countdown: migrating backoff to the frequency domain," in *17th annual international conference on Mobile computing and networking*, ser. MobiCom '11, 2011, pp. 241–252.

[68] W. Crowther, R. Rettberg, D. Walden, S. Ornstein, and F. Heart, "A system for broadcast communication: Reservation-ALOHA," in *6th Hawaii International Conference on System Sciences*, jan 1973, pp. 596–603.

[69] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Upper Saddle River, NJ 07458: Prentice-Hall, 1992.

[70] K. J. Crisler and M. L. Needham, "Throughput analysis of reservation ALOHA multiple access," *Electronics Letters*, vol. 31, no. 2, pp. 87–89, jan 1995.

[71] J. G. Kemeny and J. L. Snell, *Finite Markov Chains: With a New Appendix "Generalization of a Fundamental Matrix"*. New York: Springer, 1976.

[72] R. M. Corless, G. H. Gonnet, D. E. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W function," *Advances in Computational Mathematics*, vol. 5, no. 1, pp. 329–359, 1996.