CLUSTER BASED USER SCHEDULING SCHEMES TO EXPLOIT MULTIUSER DIVERSITY IN WIRELESS BROADCAST CHANNELS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

YUSUF SOYDAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

SEPTEMBER 2008

Approval of the thesis:

CLUSTER BASED USER SCHEDULING SCHEMES TO EXPLOIT MULTIUSER DIVERSITY IN WIRELESS BROADCAST CHANNELS

submitted by **YUSUF SOYDAN** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen ____________________
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. İsmet Erkmen ____________________
Head of Department, **Electrical and Electronics Engineering**

Assist. Prof. Dr. Çağatay Candan ____________________
Supervisor, **Electrical and Electronics Engineering Dept., METU**

Examining Committee Members:

Prof. Dr. Mete Severcan ____________________
Electrical and Electronics Engineering Dept., METU

Assist. Prof. Dr. Çağatay Candan ____________________
Electrical and Electronics Engineering Dept., METU

Assoc. Prof. Dr. Melek Yücel ____________________
Electrical and Electronics Engineering Dept., METU

Assist. Prof. Dr. A. Özgür Yılmaz ____________________
Electrical and Electronics Engineering Dept., METU

Assist. Prof. Dr. Emre Aktaş ____________________
Electrical and Electronics Engineering Dept., Hacettepe University

Date: September 2, 2008

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Yusuf Soydan

Signature :

ABSTRACT

CLUSTER BASED USER SCHEDULING SCHEMES TO EXPLOIT MULTIUSER DIVERSITY IN WIRELESS BROADCAST CHANNELS

Soydan, Yusuf
M.Sc., Department of Electrical and Electronics Engineering
Supervisor: Assist. Prof. Dr. Çağatay Candan

September 2008, 99 pages

Diversity methods are used to improve the reliability of the communication between transmitter and receiver. These methods use *redundancy* to reduce the errors in the communication link. Apart from the conventional diversity methods, *multiuser diversity* has an aim of maximizing the *sum capacity* of a multi-user system. To benefit from multiuser diversity, the *opportunistic scheduling method* grants the channel access to the user which has the best channel quality among all users. Therefore, the cumulative sum of the information sent to all users, which is the sum capacity, is maximized in the long term.

Although the opportunistic scheduling maximizes the sum capacity, it has some drawbacks such as the feedback load growing with the number of users and the problem of fairness for the users which may have lower channel quality than some other users. In this thesis, these two issues are investigated for the broadcast channels.

Feedback quantization, which gives partial information on the channel state, is studied to mitigate the feedback load with a goal of little loss in the sum capacity. The *thresholds* for the finite feedback quantization are determined to provide fairness and to reduce the feedback load at the same time. To provide fairness, users are grouped into *clusters* and thresholds are optimized for each cluster. A method is proposed by extending the one given by Floren *et. al*. to solve the mentioned problems and the proposed method is compared with some other scheduling methods in the literature.

Keywords: Opportunistic Communication, Feedback Quantization, Scheduling, Broadcast Channel, Clustering

ÖZ

KABLOSUZ YAYIN KANALLARI İÇİN ÇOKLU KULLANICI ÇEŞİTLİLİĞİ KULLANAN KÜMELEME TABANLI ZAMANDA PAYLAŞIM TEKNİKLERİ

Soydan, Yusuf
Yüksek Lisans, Elektrik ve Elektronik Mühendisliği Bölümü
Tez Yöneticisi: Yard. Doç. Dr. Çağatay Candan

Eylül 2008, 99 sayfa

Çeşitleme yöntemleri, alıcı ile verici arasındaki haberleşmenin iyileştirilmesi için kullanılmaktadır. Bu yöntemler haberleşme hattında oluşan hataların azaltılması için *artıklık* içermektedir. Geleneksel çeşitleme yöntemlerinden farklı olarak *çoklu kullanıcı çeşitliliği*, sistemin toplam kapasite değerini enbüyütme amacıyla kullanılır. Çoklu kullanıcı çeşitliliğinin avantajından yararlanmak için *kazançsal çizelgeleme yöntemi*, kanal kalitesi en iyi olan kullanıcıya kanal erişimini verir. Bu sayede, tüm kullanıcılara gönderilen kümülatif bilgi olan toplam kapasite değeri uzun dönemde enbüyütülmüş olur.

Kazançsal çizelgeleme toplam kapasiteyi enbüyütürken, artan kullanıcı sayısı ile doğru orantılı olarak artan geri besleme yükü ve kanal kalitesi daha düşük olan kullanıcıların kanal erişiminde elde ettikleri kapasite değerinin diğer kullanıcılarınkine göre daha az olması (adaletsizlik) gibi bazı dezavantajları barındırmaktadır. Bu tezde, bahsi geçen iki durum yayın kanalları için araştırılmıştır.

Geribesleme yükünü toplam kapasite değerinde çok az bir düşüş ile azaltmak için kullanılan ve kanal bilgisini kısmi olarak içeren geribesleme nicemlemesi üzerine çalışılmıştır. Kanal erişiminde elde edilen kapasite değerinde adaletin sağlanması ve geribesleme yükünün azaltılmasında kullanılmak üzere sonlu geribesleme nicemlemesi için *eşik değerleri* belirlenmiştir. Adaletin sağlanmasında kullanıcıların *kümelere* gruplanması ve eşik değerlerinin her küme için eniyilenmesi gerekmektedir. Bu doğrultuda, Floren vd. tarafından önerilen yöntem genişletilerek bahsi geçen problemlere çözüm olarak yeni bir yöntem önerilmiş olup bu yöntem literatürdeki diğer yöntemlerle de karşılaştırılmıştır.

Anahtar kelimeler: Kazançsal Haberleşme, Geribesleme Nicemlemesi, Çizelgeleme, Yayın Kanalı, Kümeleme

ACKNOWLEDGEMENT

I would like to express my gratitude to Assist. Prof. Dr. Çağatay Candan. His infinite support, constructive criticism and enlightening advice on my study are well appreciated. I am very grateful that my technical interests were always welcome. His confidence on my enthusiasm is also important for me. His supervision, guidance, and intuitions on the study enhanced my insight and academical development.

I would like to record my appreciation to my friends with whom I discuss about the study and general issues in the field. Those conversations were so helpful and contributed to my study and vision very much.

Last but not the least, I gratefully thank my family. It would not be possible to complete this study without their encouragement, the only motivation of mine to keep going after a long period of silence in my M.S. study.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CQMS	Clustered Quantized Maximum SNR
CQMS-max-agg	Clustered Quantized Maximum SNR with Aggregate Capacity Maximization
CQMS-P_o	Clustered Quantized Maximum SNR with Scheduling Outage
CQWS	Clustered Quantized Weighted SNR
CSI	Channel State Information
FDMA	Frequency Division Multiple Access
iid	Independent and Identically Distributed
LOS	Line of Sight
MIMO	Multi Input Multi Output
MISO	Multi Input Single Output
MOO	Multi-Objective Optimization
MS	Maximum SNR
OB	Opportunistic Beamforming
OFDMA	Orthogonal Frequency Division Multiple Access
PDF	Probability Density Function
QMS	Quantized Maximum SNR
RR	Round Robin
SER	Symbol Error Rate
SIMO	Single Input Multi Output
SNR	Signal to Noise Ratio
SOO	Single-Objective Optimization
TDMA	Time Division Multiple Access

CHAPTER 1

INTRODUCTION

Wireless systems have some additional difficulties on establishing reliable communication in comparison with wireline systems. One of the main problems is *fading*. In a wireless medium, transmitter and receiver have to combat low channel gains due to fading process so that a reliable communication link can be set up.

A general system model for communication is given in Figure 1.1.

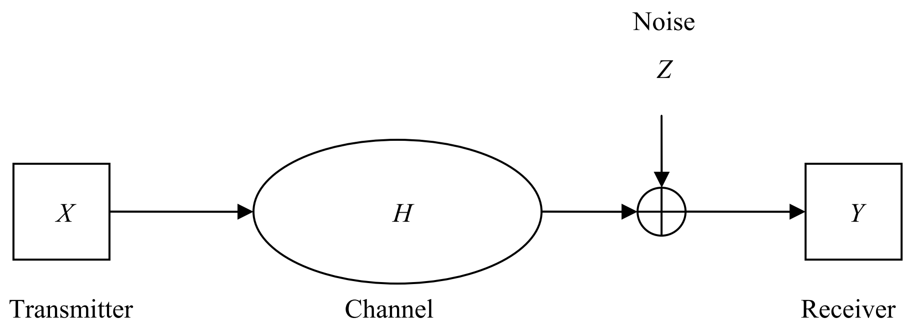

Figure 1.1 The system model for point-to-point communication

Here, H represents channel state and can be formed according to fading type of the channel. The distribution of H is related to the channel models. *Rayleigh*, *Rice* and *Nakagami-m* are some common examples of the statistical channel distribution that are used to model communications over wireless channels in different environments [1].

Establishing a *reliable communication link* between transmitter and receiver is the central issue of telecommunications. The performance criteria can be *bit* / *symbol* *error rate (BER / SER)* or *outage probability*. In an additive white Gaussian noise (AWGN) channel, BER is a function of the received signal to noise ratio (SNR) and its value is constant; whereas in a fading channel, the received signal power varies randomly. For fading channels, the outage probability is defined as the probability of received SNR being below a certain level. Since the received SNR randomly changes, the average BER is considered for fading channels.

A performance comparison of binary phase shift keying (BPSK) in AWGN and Rayleigh fading channels is shown in Figure 1.2. As it can be seen from the figure, the BER curve decreases exponentially for AWGN channel and linearly for Rayleigh fading channel with increasing received SNR in dB. Namely, received SNR must be 5 dB in AWGN channel; whereas it should be 17 dB in Rayleigh fading channel for a BER value of 10^{-2}. To approach to the AWGN performance, receivers make use of *diversity*.

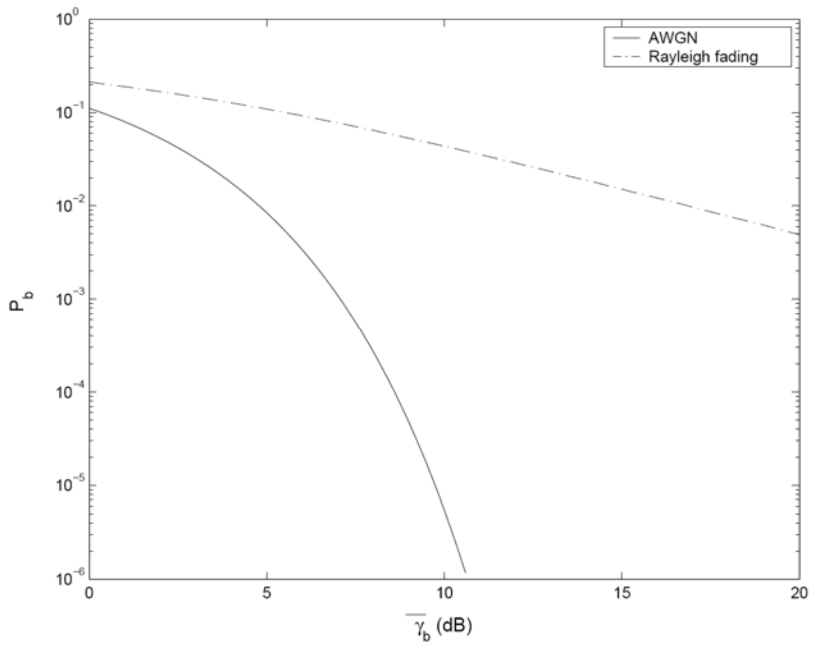

Figure 1.2 The average BER for BPSK in the Rayleigh fading and AWGN channels [2]

The capacity is defined as the mutual information between transmitted and received signals in a discrete memoryless channel shown in Figure 1.1 [3].

$$C = \max_{p(x)} I(X;Y) \qquad (1.1)$$

where *p(x)* is the distribution of input. When eq. (1.1) is maximized at constant average transmit power, we get the following, [4]

$$C = W \log_2 \left(1 + \frac{P}{N}\right) \quad \text{bits/sec} \qquad (1.2)$$

where P is the average signal power, N is the average channel noise power and W is the bandwidth of the channel. The capacity theorem, defined by Shannon, states that if the rate is less than the capacity of the channel, the data can be received with an arbitrarily small error probability [3].

The capacity of fading channels changes depending on the level of knowledge on channel state information (CSI). CSI can be known at

1. the receiver with the help of a pilot sequence
2. both the transmitter and receiver which can be established by sending CSI to the transmitter with a feedback mechanism

Different from classical diversity techniques which are used to set up communications with low error probability; *multiuser diversity* is used to set up higher capacity achievement for a multiuser system. Rather than combating fading nature of the channel, utilizing variations of channel states provides higher system capacity or sum (aggregate) capacity.

In [5] multiuser diversity is discussed for the uplink channel in the cellular communication model and the optimum scheduling is found as selecting the user which has the largest SNR. The downlink dual of this work is studied in [6] and it is found that giving the service solely to the user with the largest SNR is still optimum.

Since multiuser diversity maximizes the sum capacity, the fairness issues become more important for users which have different channel statistics. In [7], proportional fair scheduling algorithm is introduced, which balances the ratio of requested rate over past throughput. In addition to a fair scheduling algorithm, a kind of beamforming, called *opportunistic beamforming*, is proposed for the base station to increase the randomness on the dynamic range of the channel, [7].

Due to the necessity of CSI estimation at the user side and feeding it back to the base station, sending CSI fully can be a bottleneck in the cell with a large number of users. It also causes delay in the interval of scheduling decision at the base station. Hence, the reduced feedback load is discussed in many papers. The s*elective multiuser diversity* is introduced in [8] that any user in good condition considering its channel sends feedback to the base station. Therefore, users with lower CSI value do not waste the bandwidth. A threshold on the instantaneous received SNR is put to reduce the size of feedback sent by the users. Reducing feedback size means reducing the accuracy of the CSI known by the transmitter. When the base station schedules the users according to the partial information of the channel, the achieved capacity approaches the value of full feedback situation with increasing number of users, [9].

The effect of independent and identically distributed (iid) and non-iid users on sending the quantized CSI feedback is discussed in [10]. The quantized feedback levels according to the pre-defined BER are sent to the base station by iid users without a loss in spectral efficiency in comparison with the full feedback scheme. However, in non-iid case the quantized feedback levels, which are defined as normalizing the previous ones with the average SNR, are sent to the base station to avoid the user with the highest SNR monopolizing the channel.

In [11] the optimal threshold scheme is compared with the uniform threshold scheme having constraints of the probability minimization of incorrectly identification of the best user and the throughput maximization. The effect of feedback quantization is discussed in [12] that the throughput with increasing quantization levels approaches the throughput achieved with the full feedback case. However increasing number of users is necessary with increasing quantization levels to get the same throughput. In [13], full feedback with maximum SNR scheme, full feedback with the ratio of instantaneous SNR over the mean SNR (weighted SNR) scheme and a random select scheduling scheme by using thresholds are compared with round robin (RR) scheme. Although full feedback with maximum SNR (MS) and full feedback with weighted SNR is better than random select scheduling scheme, it reduces the feedback load of the system with little gains over RR.

In this thesis, we focus on both the feedback issue on exploiting the multiuser diversity and the fairness issue which exists in case of non-iid users. Feedback reduction, which can be done by quantizing the feedback value using pre-determined thresholds, plays a key role on decreasing feedback load and delay especially in a cell with a large number of users. Floren et al. [12] studied on this topic for a homogeneous cell model in which each user has the same channel statistics. Here we extend Floren's results to the non-homogeneous (heterogeneous) case. The *heterogeneous cell model* is introduced, in which every user has its own channel statistics that may or may not be identical with the other users in the cell. Although heterogeneous cell model is more realistic, its computational complexity on finding the optimum thresholds, which maximizes the sum capacity, grows exponentially with increasing number of users. Therefore, *clustering* users with the same channel statistics is an approach closer to the realistic case and more suitable for the optimization process. Clustering is also a partial solution to the issue of fairness. In this study, the fairness issue is also addressed during the threshold optimization process.

The organization of the thesis as follows: In the next chapter uplink and downlink channels are introduced. Some common multiple access methods are explained with an emphasis on TDMA. The TDMA capacity regions are discussed for AWGN and fading channels with or without CSI known at transmitter. The usage of channel variations towards our advantage, which is the main idea of *multiuser diversity*, is explained in Chapter 3. The implementation of multiuser diversity brings out some issues. These issues and their solutions with their drawbacks are introduced in the same chapter. Chapter 4 focuses on feedback mechanism for homogeneous and heterogeneous users. The feedback reduction and the structure of feedback quantization for homogeneous and heterogeneous cell models are introduced. The effects of multiuser diversity and the feedback quantization methods on the sum capacity are illustrated in Chapter 5. Finally, the study ends with the conclusions and further research directions.

CHAPTER 2

MULTIPLE USER COMMUNICATIONS

We have briefly pointed out the point-to-point communication over a channel between one transmitter and one receiver in Chapter 1. Here, we focus on multiuser communication. Transmission from one node to multiple terminals is called the *broadcast channel* (one-to-many connection). Commercial television and radio are common examples for the broadcast channel. In cellular communication, broadcast channel is also called as the *downlink channel* as illustrated in Figure 2.1.

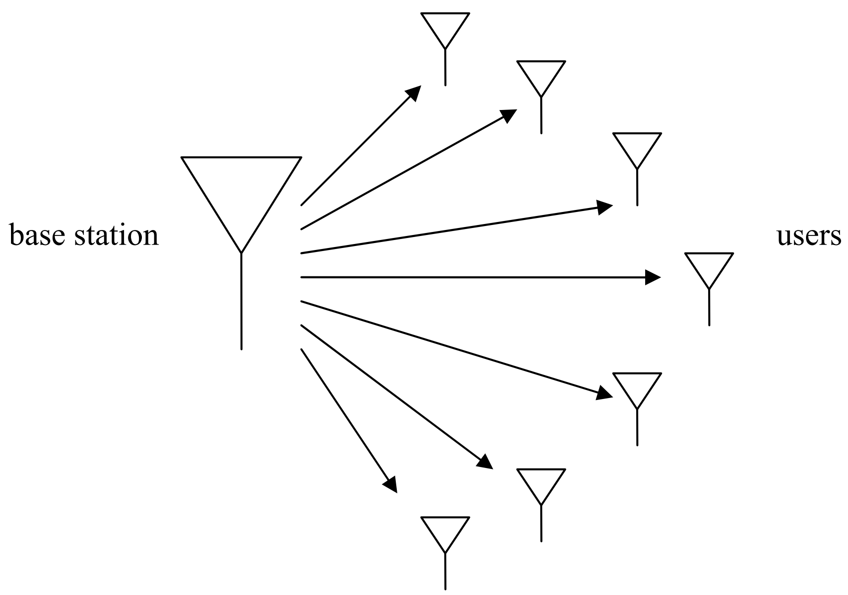

Figure 2.1 The downlink channel

The channel over which many transmitters (users) send signal to one receiver (base station) is called the *multiple access channel* (many-to-one connection). This channel, as seen on Figure 2.2, is also called the *uplink channel*.

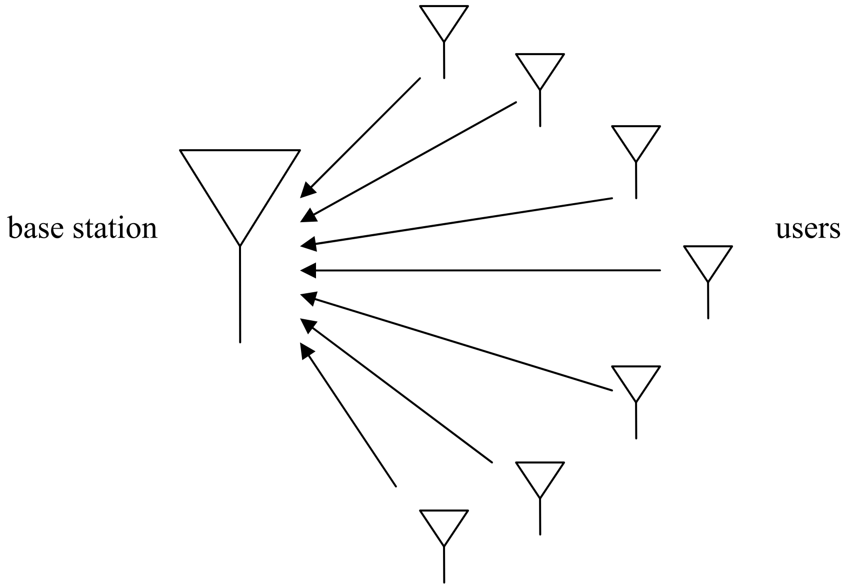

Figure 2.2 The uplink channel

Since there are many users to be accommodated, some issues such as resource allocation, interference mitigation, quality of service, continuity of communication (hand-off problem), coverage area, reusability of frequencies, etc. have to be addressed. These issues are discussed in [2] and [14] in details. Further information on wireless communications, especially cellular communications, can be found in [15]. The resource allocation issue is also discussed briefly in the subsequent chapters of the thesis.

2.1 MULTIPLE ACCESS SCHEMES

In a multiuser cell each user demands access to the channel. Therefore, the base station has to give each user access privilege according to some defined rules. The role of the base station can be interpreted as sharing the resources such as time, frequency, space etc. among the users. Resource sharing is applicable in both uplink and downlink channels. Next, we briefly explain some multiple access schemes.

2.1.1 Code Division Multiple Access

Code Division Multiple Access (CDMA) is a multiple access scheme which uses whole frequency band and time for all users in the cell as seen in Figure 2.3. CDMA uses orthogonal codes to spread the signal over the bandwidth. At the receiver side, the received signal is multiplied with the same code sequence used in the transmitter offer synchronization to extract any user’s data. Further information on CDMA principles is given in [1].

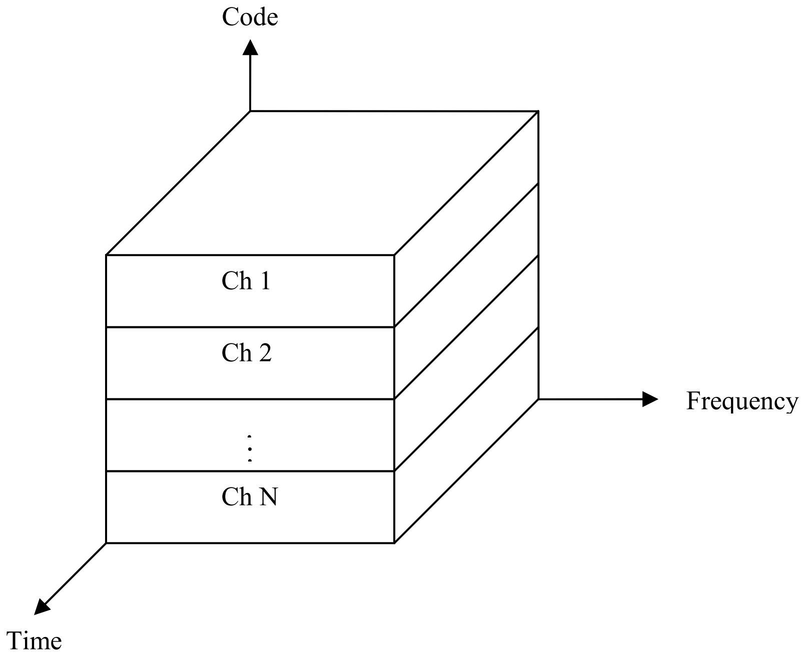

Figure 2.3 Code division multiple access, [2]

2.1.2 Frequency Division Multiple Access

Frequency Division Multiple Access (FDMA) is a multiple access method which divides the whole spectrum into frequency bands. Since these bands do not overlap with each other, they can be thought as orthogonal channels. Therefore, the interference between the channels can be eliminated. FDMA scheme can be drawn in time, frequency and code dimensions as in Figure 2.4.

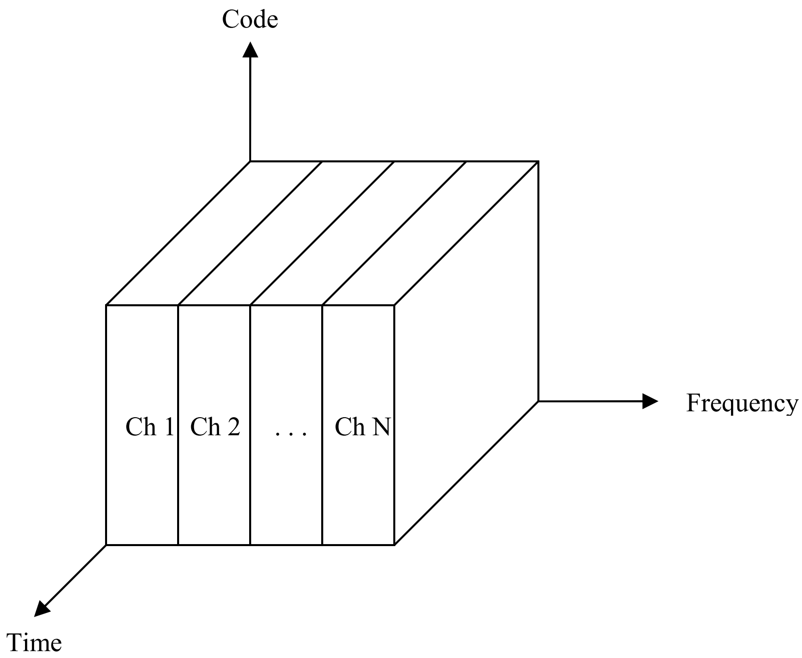

Figure 2.4 Frequency division multiple access, [2]

2.1.3 Time Division Multiple Access

Time division multiple access (TDMA) is another access scheme which gives users access to the channel according to time-sharing system. Each user accesses the channel for a limited period of time called time slot. Since time slots are orthogonal to each other, the interference between users is eliminated. Figure 2.5 shows the shares over the dimensions.

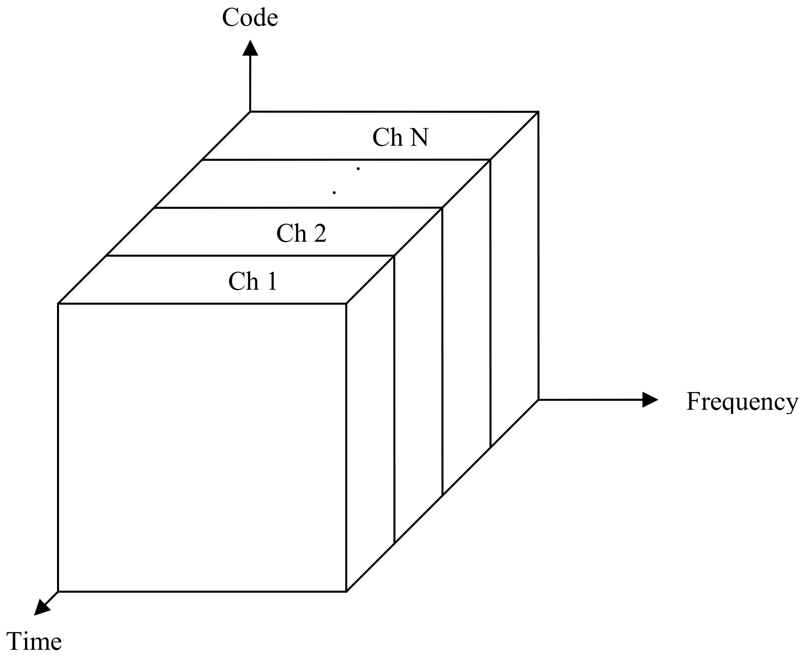

Figure 2.5 Time division multiple access, [2]

In this study the main concern is the orthogonalized systems such as TDMA scheme. Next we examine the TDMA scheme in more details.

2.2 DOWNLINK CHANNEL CAPACITY FOR TDMA

Unlike in the point-to-point communication, the user channel capacities can not be reduced to a single number as in single user communication [2]. A *capacity region* is defined for all users. Any point (vector) in this region is a set of achievable rates with arbitrarily small error probability for the users. The capacity region does not tell us a method to achieve a point in the capacity region. Instead, it shows the set of rates that can be achieved by all users with arbitrarily small communication error probability.

2.2.1 Additive White Gaussian Noise Channel Capacity for TDMA

As the capacity equation (1.2), no matter how many receivers are there, a single user (user-*m*) can achieve

$$C_m = W \log_2 (1 + \frac{|h_m|^2 P}{N_o W}), \ \mathrm{m} \in \{1,2,...,\mathrm{M}\} \quad (2.1)$$

where $N_o/2$ is power spectral density of AWGN channel, P is the total average power and h_m is the channel gain for user-m. It is assumed that the channel for each user is a degraded broadcast channel in which the channel gains stay constant, [3]. Eq. (2.1) gives an upper bound for the capacity region which is user-m has C_m and the other M-1 users have zero capacity.

The achievable rates for each user having equal power constraint,

$$C = \bigcup_{\sum_{m=1}^{M} \tau_m = 1} \left\{ (R_1, R_2 ..., R_M) : R_m = \tau_m W \log_2 \left(1 + \frac{|h_m|^2 P}{N_o W} \right) \right\} \quad (2.2)$$

where τ_m is the fraction of time in which only user-m can access the channel. Figure 2.6 shows capacity region for two users.

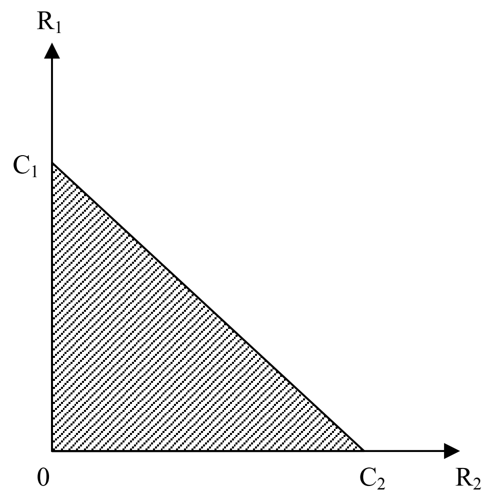

Figure 2.6 The capacity region (shaded area) for 2 users

Instead of fixing average power to P, the user dependent power allocation can be considered in time division access as

$$C = \bigcup_{\sum_{m=1}^{M} \tau_m = 1, \sum_{m=1}^{M} \tau_m P_m = \bar{P}} \left\{ (R_1, R_2 ..., R_M) : R_m = \tau_m W \log_2 \left(1 + \frac{|h_m|^2 P_m}{N_o W} \right) \right\} \quad (2.3)$$

Eq. (2.3) is also valid for FDMA, [16].

The maximum value for the sum of the rates in the region is called the sum capacity and formulated as

$$C_{sum} = \max_{(R_1,...,R_M) \in C} \sum_{m=1}^{M} R_m \quad (2.4)$$

The sum capacity is a single number that gives the limits of the capacity region.

2.2.2 Fading Channel Capacity for TDMA

In the fading channel two types of capacity are concerned: ergodic capacity and outage capacity. Outage capacity occurs when the signal to noise ratio (SNR) is below the minimum SNR for arbitrarily small probability of error. Further study on outage capacity for time division, frequency division and code division methods can be found in [18]. Ergodic capacity for time division can be calculated when channel state information (CSI) is available at the receiver or both at the receiver and transmitter. An example of fading channel model is shown in Figure 2.7. h_m is the CSI that belongs to user-m and z_m is the noise, $z_m \sim C(0,N_o)$. It is also assumed that channel is slowly fading so that transmitter can adapt its transmission before the channel conditions change.

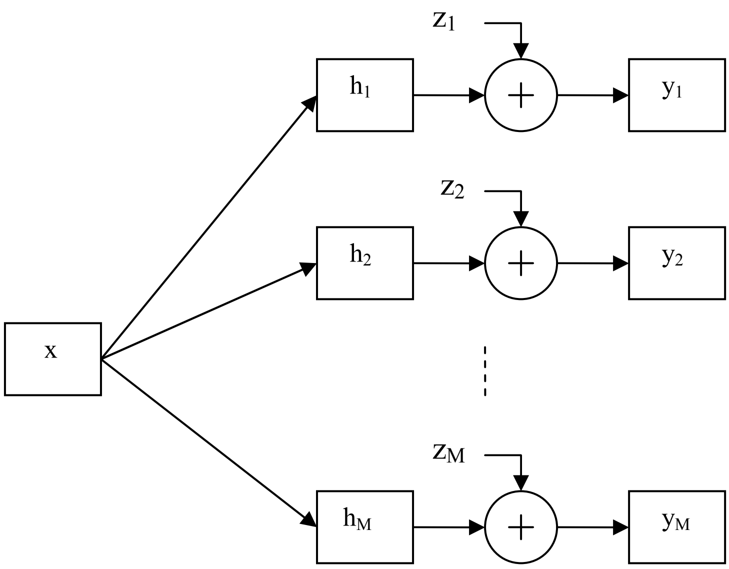

Figure 2.7 The broadcast channel model

CSI known at receiver

If the CSI is only known at the receiver, no dynamic power allocation can be done at the transmitter. Therefore, the best way is to transmit equal power for each transmission sessions, [14]. The capacity of this case,

$$C = \bigcup \left\{ \left(R_1, R_2 ..., R_M\right) : R_m = E_{h_m} \left\{ \frac{W}{M} \log_2 \left(1 + \frac{|h_m|^2 \overline{P}}{N_o W} \right) \right\} \right\} \tag{2.5}$$

where E_{h_m} takes the expectation of the function over h_m, the CSI for user-*m* for a certain time; $1/M$ is fraction of time in which user-*m* is transmitted; $\overline{P}$ is total average transmit power. Here it is assumed that any user can estimate its own CSI perfectly.

CSI known at both transmitter and receiver

This case states that there is a feedback channel between transmitter (base station) and each receiver (user) is assumed to estimate the CSI perfectly and send it back to the transmitter over a feedback channel without any error. Therefore, the transmitter has the freedom of varying the instantaneous transmitted power based on the CSI value. The system model is shown in Figure 2.8.

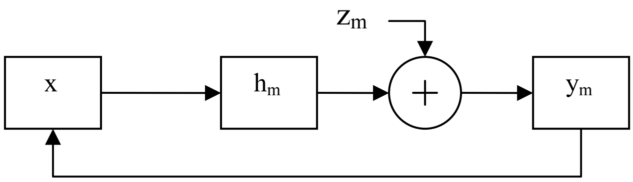

Figure 2.8 The broadcast channel belonging to user-*m* with a feedback channel to transmitter

Ergodic capacity derived in [17] as follows: $\hbar = (h_1, h_2, ..., h_M)$ and $\hbar \in H$ where H is the set of all possible joint fading states. Thus,

$$E_{\hbar}\left\{\sum_{m=1}^{M} \tau_m(\hbar) P_m(\hbar)\right\} \leq \overline{P}$$
$$\sum_{m=1}^{M} \tau_m(\hbar) = 1, \quad \forall \hbar \in H \tag{2.6}$$

where $\overline{P}$ is total average transmit power, $P_m(\hbar)$ is the transmit power for user-*m* at $\hbar$ according to a certain power allocation policy $\widetilde{P}$ and $\tau_m(\hbar)$ $(0 \leq \tau_m(\hbar) \leq 1)$ is the fraction of transmission time allocated to the same user. If $\widetilde{F}$ is defined as the set of all possible power allocation policies, the achievable rate region for the variable power and variable transmission time can be

$$C(\overline{P}) = \bigcup_{\widetilde{P} \in \widetilde{F}} C_{TD}(\widetilde{P}) \tag{2.7}$$

where

$$C_{TD}(\widetilde{P}) = \left\{ R : R_m \leq E_{\hbar}\left\{ \tau_m(\hbar) W \log_2\left(1 + \frac{|h_m|^2 P_m(\hbar)}{N_o W}\right)\right\}, 1 \leq m \leq M \right\} \tag{2.8}$$

Since CSI is known both at the receivers and transmitter, water-filling power allocation is the optimal power allocation scheme to achieve the maximum capacity [17]. The sum capacity can be achieved choosing the best user and giving it the allocated power found by water-filling [14]. Details of water-filling can be found in [21].

Eq. (2.8) shows the rate user-*m* can achieve using a practical scheme, called round robin (RR) scheme. It is discussed in the following chapters. RR scheduling is allocating the channel to any user for a certain time slot in the frame, which is formed by time slots, regardless of channel states of the users. Figure 2.9 shows an example of RR scheduling.

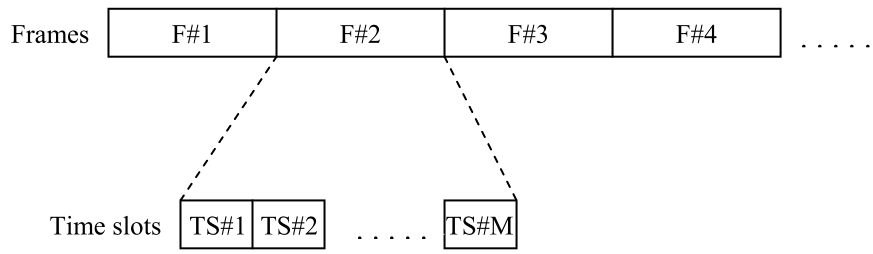

Figure 2.9 The round robin scheduling frame structure

As shown in Figure 2.9 each frame includes *M* (number of users) time slots and each user has the access in the same time slot of each frame. Although RR scheduling does not consider the channel state, it has a simple algorithm with almost no computational complexity. The capacity equation of RR scheme C_{TD_RR} [19] is a modified version of (2.8),

$$C_{TD\_RR} = \left\{ R : R_m \leq E_{h_m} \left\{ \frac{W}{M} \log_2 \left(1 + \frac{|h_m|^2 P}{N_o W} \right) \right\},\ 1 \leq m \leq M \right\} \tag{2.9}$$

where $\tau_m(\hbar) = 1/M$ for all *m* and it is assumed that there is no feedback channel from users to base station. Therefore, the equal power allocation is used for all $m, P_m(\hbar) = P$. (2.9) can be simplified as

$$C_{TD\_RR} = \left\{ R : R_m \leq \frac{W}{M} \int_0^\infty \log_2(1+\gamma) f_{\gamma^{(m)}}(\gamma) d\gamma,\ 1 \leq m \leq M \right\} \tag{2.10}$$

$$= \left\{ R : R_m \leq \frac{We^{1/\bar{\gamma}^{(m)}}}{M \ln 2} E_1\left(1/\bar{\gamma}^{(m)}\right), \quad 1 \leq m \leq M \right\} \tag{2.11}$$

where $E_1(x) = \int_x^{\infty} t^{-1} e^{-t} dt$ is the exponential integral function [20]; $\gamma^{(m)} = {|h_m|^2 P} / {N_o W}$ and $\bar{\gamma}^{(m)}$ are the instantaneous and mean SNR of user-*m* respectively; $f_{\gamma^{(m)}}$ is the probability density function (PDF) of $\gamma^{(m)}$ which is exponentially distributed.

CHAPTER 3

MULTIUSER DIVERSITY

In the previous chapter, some multiple access schemes are reviewed. Especially time division multiple access (TDMA) is discussed in more details and the capacity regions are given for additive white Gaussian noise (AWGN) and fading channels. In Chapter 1, the channel model is given for point-to-point communication between two nodes. For a fading channel, the channel fluctuations may cause difficulty in establishing reliable communication. Some techniques are used to provide robustness against deep fade situations. These techniques are called *diversity*. In mean sense, diversity means sending information from a transmitter to a receiver in multiple copies over independent paths. These independent paths can be in temporal, frequency, spatial dimensions.

Temporal diversity uses multiple intervals in time to transmit the same signal. Generally, error correction coding techniques use diversity in time. A very simple example is *repetition coding*. Repetition coding is repeating the same symbol. If the *coherence time* of the channel is large in comparison with the symbol duration, *interleaving* can be used to extract time diversity. For a block fading channel, the benefit of coding and interleaving can be seen in Figure 3.1.

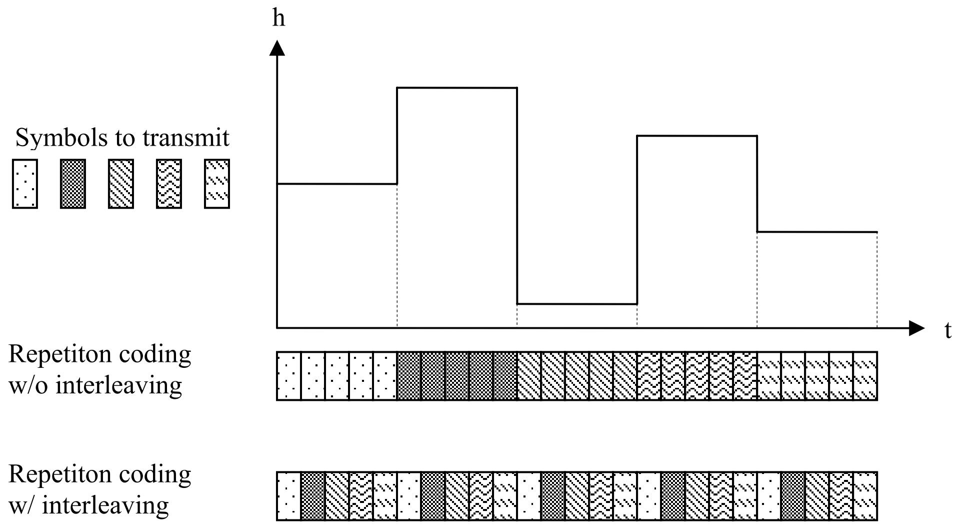

Figure 3.1 A repetition coding with interleaving

Frequency diversity is available for frequency selective channels. If the bandwidth of the signal sent is greater than the *coherence bandwidth* of the channel, the channel is *frequency selective*. Thus, it can be said that frequency response of the channel has multiple taps.

Spatial diversity is using multiple antennas separated far enough from each other to send or receive independently faded signals. Multiple antenna usage can be established in both transmitter side and receiver side. Such a system is called as a multiple input multiple output (MIMO) scheme. Antenna diversity is also used by single transmitter multiple receiver (single input multiple output - SIMO) or vice versa (multiple input single output - MISO) scenarios. An illustration of spatial diversity is shown in Figure 3.2.

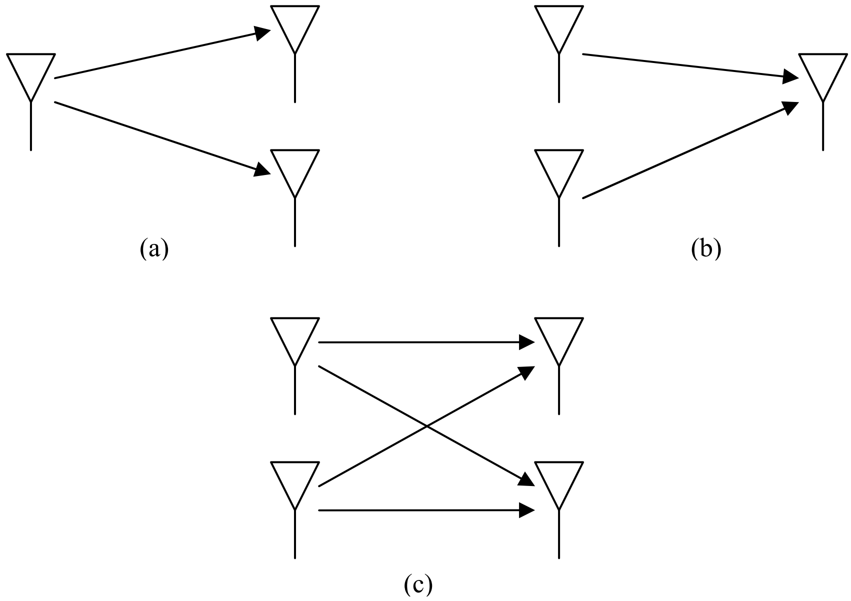

Figure 3.2 Spatial diversity: (a) SIMO, (b) MISO, (c) MIMO channels

Further details about diversity techniques discussed above can be in [1], [2], [14].

3.1 MULTIUSER DIVERSITY

Multiuser diversity is a different type of diversity in comparison with the common diversity techniques discussed before. The conventional diversity techniques are used to provide a more reliable communication, namely reducing bit error probability (BER), between the transmitter and receiver by using independent paths on dimensions such as time, frequency, etc. Improvement in BER can be achieved using diversity techniques. However, the multiuser diversity uses fading states of the channel to increase the sum capacity. Hence, multiuser diversity is used to maximize the sum capacity of the whole system. It does not promise a high capacity achievement for a small interval of time. Instead, a single user can achieve higher capacity due to the maximization of the sum capacity in the long term.

Multiuser diversity is discussed in [5] for a multiple access channel in a single cell. It is assumed in [5] that each receiver can track its channel with the help of a pilot signal and feeds its channel state back to the transmitter. The optimum scheduling method to maximize the sum capacity is the one giving the channel access to the user whose channel condition is the best among all users. In [6], similar results are given for the broadcast channels.

3.2 MULTIUSER DIVERSITY GAIN

The multiuser diversity gain benefits from channel fluctuations. The base station collects the user CSI from each user and transmits to the user with the best channel condition. At this point, the distribution of the channel gains has an important role on accessing the channel. The more random the channel changes for any user, the more capacity the system can achieve. Figure 3.3 shows an example of this case [14].

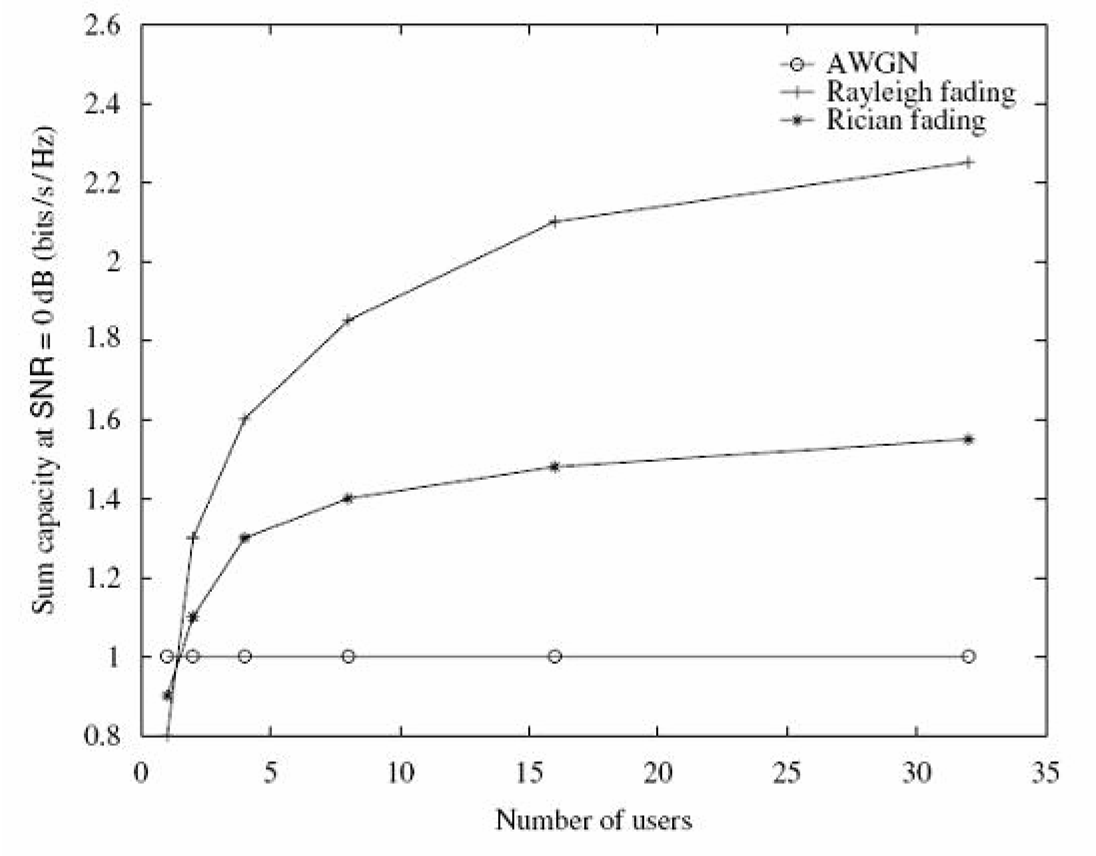

Figure 3.3 The multiuser diversity gain, [14]

The diversity gain on the sum capacity of the Rayleigh, Rician and AWGN channels are plotted for 0 dB SNR. As a conclusion, the Rayleigh fading channel which does not have a line of sight (LOS) path has the most random structure. The Rician fading channel has a LOS path which makes it less random as compared with the Rayleigh fading channel and provides less gain than the Rayleigh channel does. The AWGN channel at the constant SNR can not benefit from diversity.

3.3 SYSTEM ASPECTS

There are some issues about exploiting multiuser diversity detailed in [14]:

- Fairness and delay: For a symmetric scenario (every user has the same distribution of channel gains with the same average received power), the method of choosing a user with the best channel gain is fair in the long term due to the fact that maximizing the sum capacity also maximizes a single user capacity. However, in practice every user does not have the same channel statistics. Some users may have better channel conditions than others. When the prime goal is to raise the sum capacity, the system does not care about the capacity values achieved by an individual user. To establish fairness, some modifications have to be applied. For the real time applications the delay requirement is another issue to be handled. The fairness problem is addressed in the following chapters.

- Channel gains and feedback: Since multiuser diversity benefits when the CSI is known at the transmitter and receiver, a method of channel estimation has to be provided. The base station sends a common pilot to every user and each user estimates its channel perfectly according to the pilot sequence. After that every user sends its channel state over a feedback channel to the transmitter. Under the ideal conditions, feedback arrives without any error. However, sending the full feedback of the channel state may not be practical. Growing number of users causes a large amount of the feedback. Instead of full feedback, partial feedback may be used to avoid bottleneck with a sacrifice on the sum capacity. A partial feedback method is discussed in this thesis.

- Channel fluctuations: Channel fluctuations are another point to benefit multiuser diversity. Fast channel variations are preferred due to sharing the resources fairly and having low delay. That is, the slow channel variations make the base station to transmit to the best user for a long time. Hence any user having poor, channel would be punished with no transmission. In [7], a scheduling method is proposed to cope with slow channel variations. We also discuss this method for the sake of completeness.

3.4 PROPORTIONAL FAIR SCHEDULING

In [14], [7] and [22] a scheduling algorithm, called proportional fair scheduling, is discussed to provide fairness in the whole cell. The algorithm is adopted in TDMA for time slots of transmission. The scheduler works as follows:

1. The base station keeps track of the average throughput *T(k)* belonging to each user for a period of t_c at time *k*.
2. Each user request a rate *R(k)* according to its channel condition which is calculated by the pilot signal sent by base station.
3. The scheduler chooses

$$\max \frac{R_m(k)}{T_m(k)} \tag{3.1}$$

   over $m=1,2,\ldots M$ and transmits to the user m^* which satisfies (3.1).

4. Average throughputs are updated according to

$$T_m(k+1) = \begin{cases} \left(1-\frac{1}{t_c}\right)T_m(k)+\frac{1}{t_c}T_m(k), & m = m^* \\ \left(1-\frac{1}{t_c}\right)T_m(k) & , \quad m \neq m^* \end{cases} \tag{3.2}$$

Since the algorithm chooses any user according to (3.1) which means every user's rate request is normalized before the decision and updates (3.2), it is fair in the long term by serving users considering the past of each one. Details of proportional scheduling can be found in [14], [7], [22].

CHAPTER 4

CLUSTER BASED USER SCHEDULING WITH FEEDBACK LOAD REDUCTION

In Chapter 3, it is discussed that the sum capacity can be maximized using multiuser diversity. This can be achieved only when the user SNR values are known by the base station and the total power of the base station can be allocated only to the best user on the dimension of concern (time, frequency etc.). Feedback from each user is therefore to be sent to the base station. The full feedback (the unquantized SNR value) load for a single user can be very large in symbols (such as *bits*).

In this chapter, a system model is given first. After that, a resource allocation method for the quantized feedback system is explained and the sum capacity calculations are given according to this system model for both homogeneous and heterogeneous cells. Scheduling outage and fairness issues are also mentioned. The chapter ends with the discussion of threshold optimization to maximize the sum capacity.

4.1 SYSTEM MODEL

Time division multiple access (TDMA) scheme is used to allocate every time slot to an individual user which means that only one user can be served at any time.

It has been discussed that CSI known by transmitter (base station) is the key point of taking the advantage of multi user diversity, a pilot sequence for the estimation of the SNR belonging to each user. The channel is assumed to fade slowly enough such that the channel status stays same for the time slot during the SNR estimation. It is assumed that the SNR values of the users are estimated perfectly and the base station can get those estimates without any errors. It is also assumed that the feedback mechanism does not affect the throughput of the system.

Each user's channel is exposed to Rayleigh fading and the channel fades independently in each time slot. The channel bandwidth is assumed to be narrow enough to have a frequency flat transmission.

Since it is assumed Rayleigh fading channel, the received SNR belonging to the i^{th} user $\Gamma^{(i)}$ is a stochastic variable. The cumulative distribution function (CDF) and the probability density function (PDF) of $\Gamma^{(i)}$ is denoted by $F_{\Gamma^{(i)}}$ and, $f_{\Gamma^{(i)}}$ respectively. $\Gamma^{(i)}$ is exponentially distributed with mean SNR $\bar{\gamma}^{(i)}$ [1],

$$F_{\Gamma^{(i)}}(\gamma) = 1 - e^{-\gamma/\bar{\gamma}^{(i)}{}_i}, \quad \gamma \geq 0 \tag{4.1}$$

and PDF of $\Gamma^{(i)}$ is,

$$f_{\Gamma^{(i)}}(\gamma) = \frac{1}{\bar{\gamma}^{(i)}} e^{-\gamma/\bar{\gamma}^{(i)}}, \quad \gamma \geq 0 \tag{4.2}$$

This system model is used for both homogeneous and heterogeneous cell models.

4.2 QUANTIZED FEEDBACK IN THE HOMOGENEOUS CELL MODEL

As in [12], the homogeneous case states that each user in the cell has the same mean SNR. That is, all users' SNR values are distributed with $F_{\Gamma}(\gamma)$.

This case states that there are statistically identical M users in a single cell. All users feed their SNR values back to the base station and the base station selects the user with the best SNR for service.

Since the feedback load grows rapidly as the user number increases, a reduction method, *quantization*, is applied to the feedbacks. In order to quantize user feedbacks, *quantization levels* determined by *thresholds* must be found. Let the whole interval of instantaneous SNR be divided into K levels. Level k can be denoted as Q_k and defined between the limits,

$$Q_k = [q_k, q_{k+1}), \quad 0 \le k \le K-1$$
$$q_0 = 0$$
$$q_K = \infty$$

q_k's $k \ne 0, K$ are called *thresholds* of the quantization levels. Γ lies in the k^{th} level, $\Gamma \in Q_k$, and Q_k has limits as $q_k \le \Gamma < q_{k+1}$. Figure 4.1 shows the quantization levels and the limits that define these levels.

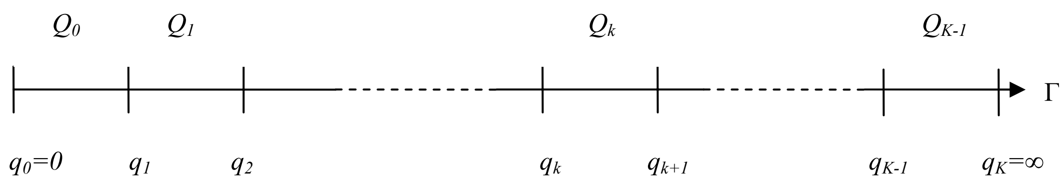

Figure 4.1 Quantization levels

The base station collects the quantized information for each user and makes a decision to transmit in the given time slot. It sorts all the feedbacks beginning from the highest level received and chooses a random user among the users in the highest level. The users in the same quantization level are chosen equally likely.

According to the mentioned scheduling method, explained in [12], each user has a probability of channel access in a given time slot which is given as $\Pr(A/\Gamma \in Q_k)$. A denotes the event of a reference user has a SNR in the highest level, Q_k, and the reference user is chosen to be granted the access to the channel. That is, there is no other user whose SNR falls in a higher level. Hence, the average gain per user can be found as,

$$\sum_{k=0}^{K-1} \int_{q_k}^{q_{k+1}} g(\gamma) \Pr(A/\Gamma \in Q_k) f_\Gamma(\gamma) d\gamma \qquad (4.3)$$

Here $g(\gamma)$ is the gain function and can be defined as in [3],

$$g(\gamma) = \frac{1}{2}\log_2(1+\gamma) \quad \text{bits/channel use} \tag{4.4}$$

If the transmitted signal is represented with the Nyquist rate and the average signal power is written as the function of consumed energy per bit ε_b; eq. (1.2), which is the capacity of a continuous bandlimited channel, becomes equivalent to eq. (4.4) which is the capacity of a discrete parallel channel, with $\gamma = {}^{2\varepsilon_b}\!/\!_{N_0}$. More information on the capacity relations are available at [1].

$\Pr(A/\Gamma \in Q_k)$ can be calculated as [12],

$$\Pr(A/\Gamma \in Q_k) = \sum_{m=0}^{M-1} \frac{1}{m+1}\binom{M-1}{m}\left[\Pr\left(\Gamma \in Q_k\right)\right]^m \left[\Pr\left(\Gamma \in \bigcup_{l<k} Q_l\right)\right]^{M-m-1} \tag{4.5}$$

In eq. (4.5), it is assumed that there are *m* users lie in the quantization level Q_k in addition to the reference user and *M*-*m*-1 users lie in the lower quantization levels (denoted by *l*<*k*). Since all users which have SNR values in Q_k can be chosen equally likely, the reference user has the selection probability of $\frac{1}{m+1}$. The product of probabilities with the number of events of users in Q_k, $\binom{M-1}{m}$, is added with respect to *m*. According to (4.5),

$$\Pr\left(\Gamma \in Q_k\right) = F_\Gamma\left(q_{k+1}\right) - F_\Gamma\left(q_k\right) \tag{4.6}$$

and

$$\Pr\left(\Gamma \in \bigcup_{l<k} Q_l\right) = F_\Gamma\left(q_k\right) \tag{4.7}$$

Substituting (4.6) and (4.7) into (4.5) yields,

$$\begin{aligned}\Pr(A/\Gamma \in Q_k) &= \sum_{m=0}^{M-1} \frac{1}{m+1}\binom{M-1}{m}\left[F_\Gamma\left(q_{k+1}\right) - F_\Gamma\left(q_k\right)\right]^m \left[F_\Gamma\left(q_k\right)\right]^{M-m-1} \\ &= \frac{\left[F_\Gamma\left(q_{k+1}\right)\right]^M - \left[F_\Gamma\left(q_k\right)\right]^M}{M\left[F_\Gamma\left(q_{k+1}\right) - F_\Gamma\left(q_k\right)\right]}\end{aligned} \tag{4.8}$$

Hence, the capacity of quantized maximum SNR (QMS) method (C_{QMS}), [12]

$$C_{QMS\_homogeneous} = \sum_{k=0}^{K-1} \frac{\left[F_\Gamma(q_{k+1})\right]^M - \left[F_\Gamma(q_k)\right]^M}{M\left[F_\Gamma(q_{k+1}) - F_\Gamma(q_k)\right]} \int_{Q_k} \frac{1}{2}\log_2(1+\gamma) f_\Gamma(\gamma) d\gamma \qquad (4.9)$$

The unit of eq. (4.9) is bits/channel use/user. The *aggregate capacity* for the cell becomes

$$C_{agg\_QMS\_homogeneous} = M \times C_{QMS\_homogeneous} \qquad (4.10)$$

4.3 QUANTIZED FEEDBACK IN THE HETEROGENEOUS CELL MODELS

It is assumed in the previous model that in a single cell every user has the same average SNR. However, in practice that is not possible due to mobility, the cell environment which is full of obstacles, scatterers and additive or destructive effects of the environment to the signal sent by the transmitter. Thus, it is possible to have users with average SNR values different from each other.

If the cell has *M* users, the set of user numbers is $R = \{1,2,3,...,M\}$ and the set of user numbers except the i^{th} user is $R_i = R - \{i\} = \{1,2,3,...,i-1,i+1,...,M\}$. The n^{th} *m-element* subset of R_i is denoted by $S_m^{(n)}$ and there are $\binom{M-1}{m}$ subsets, $S_m^{(1)}$, $S_m^{(2)}$ up to $S_m^{\left(\binom{M-1}{m}\right)}$. The complement of $S_m^{(n)}$ is denoted by $S_m^{(n)'}$ and $S_m^{(n)'} = R_i - S_m^{(n)}$. Like the previous cell model, every user has its own quantization level with the limits and the number of quantization levels is equal to each other for all users. The number of quantization levels of the i^{th} and j^{th} users is the same but the threshold values may not be so. The SNR range is divided into *K* quantization levels and k^{th} of them, $Q_k^{(i)}$ has the limits, $q_k^{(i)}$ and $q_{k+1}^{(i)}$, can be defined for the i^{th} user as follows,

$$Q_k^{(i)} = \left[q_k^{(i)}, q_{k+1}^{(i)}\right), \quad 0 \le k \le K-1$$
$$\left.\begin{aligned} q_0^{(i)} &= 0 \\ q_K^{(i)} &= \infty \end{aligned}\right\} \quad \text{for } \forall i \tag{4.11}$$

The probability of scheduling the i^{th} user, which has SNR in the k^{th} quantization level, in the given time slot is,

$$\Pr\left(A \mid \Gamma^{(i)} \in Q_k^{(i)}\right) = \sum_{m=0}^{M-1} \Pr\left(\text{the } i^{th} \text{ user} \mid m\right) \times \left\{ \sum_{n=1}^{\binom{M-1}{m}} \left[\prod_{u \in S_m^n} \Pr\left(\Gamma^{(u)} \in Q_k^{(u)}\right) \prod_{v \in S_m^{n'}} \Pr\left(\Gamma^{(v)} \in \bigcup_{l<k} Q_l^{(v)} \right) \right] \right\} \tag{4.12}$$

There are m users, whose SNR is in the k^{th} quantization level in addition to the i^{th} user's, denoted by u, and $|u| = m$; M-m-1 users, denoted by v, have SNR values lie in a quantization level lower than k and $|v| = (M-1) - m$. Thus, $|R_i| = |u| + |v|$. The probability of users in the k^{th} quantization level, $\Pr\left(\Gamma^{(u)} \in Q_k^{(u)}\right)$, can be defined as,

$$\Pr\left(\Gamma^{(u)} \in Q_k^{(u)}\right) = F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \tag{4.13}$$

The probability of the event that the other users, represented by v, have SNR values lie in a quantization level lower than k can be defined as,

$$\Pr\left(\Gamma^{(v)} \in \bigcup_{l<k} Q_l^{(v)} \right) = F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \tag{4.14}$$

If all users in k^{th} quantization level are chosen equally likely, $\Pr\left(\text{the } i^{th} \text{ user} \mid m\right)$ is the probability of selection of the i^{th} user over the users whose SNR values lie in the same quantization level can be defined as

$$\Pr\left(\text{the } i^{th} \text{ user} \mid m\right) = \frac{1}{m+1} \tag{4.15}$$

Substituting (4.13), (4.14) and (4.15) into (4.12),

$$\Pr\left(A \mid \Gamma^{(i)} \in Q_k^{(i)}\right) = \sum_{m=0}^{M-1} \frac{1}{m+1} \left\{ \sum_{n=1}^{\binom{M-1}{m}} \left[\prod_{u \in S_m^{(n)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right)\right] \prod_{v \in S_m^{(n)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] \right\} \tag{4.16}$$

Hence, the capacity equation for the i^{th} user becomes

$$C^{(i)}_{QMS\_heterogeneous} = \sum_{k=0}^{K-1} \Pr\left(A \mid \Gamma^{(i)} \in Q_k^{(i)}\right) \int_{Q_k^{(i)}} \frac{1}{2}\log_2(1+\gamma) f_{\Gamma^{(i)}}(\gamma)\,d\gamma \tag{4.17}$$

The aggregate capacity can be calculated as

$$C_{agg\_QMS\_heterogeneous} = \sum_{m=1}^{M} C^{(m)}_{QMS\_heterogeneous} \tag{4.18}$$

The examples A.1 and A.2 in the Appendix make the discussion above more clear. As it is seen in the given examples, the complexity of the probability relation grows rapidly as M increases. This brings significant difficulty on finding the optimal quantization levels. Every user must be considered separately due to difference in SNR values. Therefore, the new model, which has lower computational complexity than the discussed model above called *clustered heterogeneous cell model,* is proposed.

In the homogeneous cell model, which represents a set of users having the same average SNR, all users are subject to the same scheduling rule. Unlike the homogeneous model, the clustered heterogeneous cell model states that there are more than one set whose members have the same average SNR. Each set, called *cluster* from now on, has the certain property that all users in it have the same average SNR values. In other words, this method groups users into clusters according to their average SNR values. Figure 4.2 shows a *symbolical drawing* of both the homogeneous and heterogeneous models (Note that, each color represents an average SNR value and these drawings aim to show the mathematical equivalent of the models).

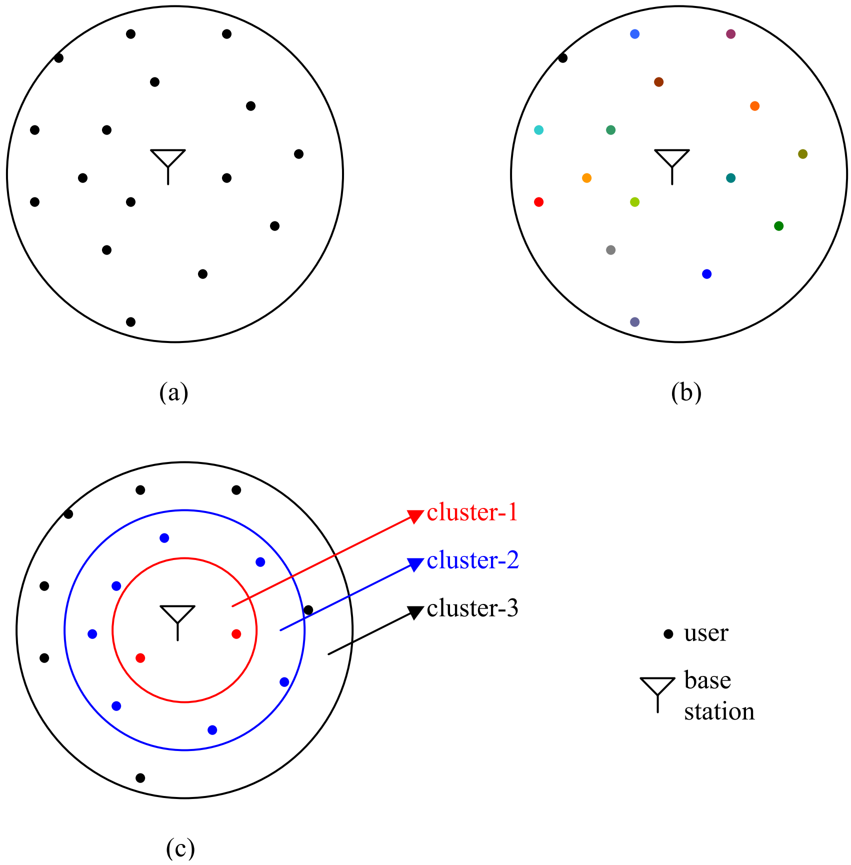

Figure 4.2 (a) Homogeneous, (b) heterogeneous and (c) clustered heterogeneous cell models

We proposed two scheduling rules for the clustered heterogeneous cell model, are explained in the following sections.

4.3.1 Random User Selection Scheduling Rule

The procedure is as follows,

I. sort the users into clusters with respect to their mean SNR values

II. collect the feedbacks (instantaneous SNR values) sent by each user according to the pre-determined threshold(s)

III. select the users which are in the highest quantization level in each cluster

IV. randomly choose one of users selected in III for service

There are T clusters and the reference user's cluster is *cluster-r*, $r \in \{1,2,...T\}$. Any cluster other than the reference user's is *cluster-i*, Every user in *cluster-i* has $\Gamma^{(i)}$. The probability of users being in the k^{th} quantization level for the *cluster-i* is,

$$\Pr\left(\Gamma^{(i)} \mid Q_k^{(i)}\right) = \sum_{m^{(i)}=1}^{M^{(i)}} \binom{M^{(i)}}{m^{(i)}} \left[\Pr\left(\Gamma^{(i)} \in Q_k^{(i)}\right)\right]^{m^{(i)}} \left[\Pr\left(\Gamma^{(i)} \in \bigcup_{l<k} Q_l^{(i)}\right)\right]^{M^{(i)}-m^{(i)}} \tag{4.19}$$

The probability of the reference user chosen in the k^{th} quantization level given the other users of the clusters,

$$\Pr\left(\Gamma^{(r)} \mid Q_k^{(r)}\right) = \sum_{m^{(r)}=0}^{M^{(r)}-1} \frac{1}{1+\sum_{z=1}^{T} m^{(z)}} \binom{M^{(r)}-1}{m^{(r)}} \left[\Pr\left(\Gamma^{(r)} \in Q_k^{(r)}\right)\right]^{m^{(r)}} \left[\Pr\left(\Gamma^{(r)} \in \bigcup_{l<k} Q_l^{(r)}\right)\right]^{M^{(r)}-m^{(r)}-1} \tag{4.20}$$

where

$$\Pr\left(\Gamma^{(r)} \in Q_k\right) = F_{\Gamma^{(r)}}\left(q_{k+1}^{(r)}\right) - F_{\Gamma^{(r)}}\left(q_k^{(r)}\right) \tag{4.21}$$

and

$$\Pr\left(\Gamma^{(r)} \in \bigcup_{l<k} Q_l^{(r)}\right) = F_{\Gamma^{(r)}}\left(q_k^{(r)}\right) \tag{4.22}$$

Thus, the probability of scheduling the reference user,

$$\begin{aligned}
\Pr\left(A \mid \Gamma^{(r)} \in Q_k^{(r)}\right) &= \prod_{t=1}^{T} \Pr\left(\Gamma^{(t)} \mid Q_k^{(t)}\right) \\
&= \prod_{\substack{t=1 \\ t\neq r}}^{T} \left\{ \sum_{m^{(t)}=1}^{M^{(t)}} \binom{M^{(t)}}{m^{(t)}} \left[\Pr\left(\Gamma^{(t)} \in Q_k^{(t)}\right)\right]^{m^{(t)}} \left[\Pr\left(\Gamma^{(t)} \in \bigcup_{l<k} Q_l^{(t)}\right)\right]^{M^{(t)}-m^{(t)}} \right\} \times \\
&\quad \sum_{m^{(r)}=0}^{M^{(r)}-1} \frac{1}{1+\sum_{z=1}^{T} m^{(z)}} \binom{M^{(r)}-1}{m^{(r)}} \left[\Pr\left(\Gamma^{(r)} \in Q_k^{(r)}\right)\right]^{m^{(r)}} \left[\Pr\left(\Gamma^{(r)} \in \bigcup_{l<k} Q_l^{(r)}\right)\right]^{M^{(r)}-m^{(r)}-1} \\
&= \prod_{\substack{t=1 \\ t\neq r}}^{T} \left\{ \sum_{m^{(t)}=1}^{M^{(t)}} \binom{M^{(t)}}{m^{(t)}} \left[F_{\Gamma^{(t)}}\left(q_{k+1}^{(t)}\right) - F_{\Gamma^{(t)}}\left(q_k^{(t)}\right)\right]^{m^{(t)}} \left[F_{\Gamma^{(t)}}\left(q_k^{(t)}\right)\right]^{M^{(t)}-m^{(t)}} \right\} \times \\
&\quad \sum_{m^{(r)}=0}^{M^{(r)}-1} \frac{1}{1+\sum_{z=1}^{T} m^{(z)}} \binom{M^{(r)}-1}{m^{(r)}} \left[F_{\Gamma^{(r)}}\left(q_{k+1}^{(r)}\right) - F_{\Gamma^{(r)}}\left(q_k^{(r)}\right)\right]^{m^{(r)}} \left[F_{\Gamma^{(r)}}\left(q_k^{(r)}\right)\right]^{M^{(r)}-m^{(r)}-1}
\end{aligned} \tag{4.23}$$

If T=1, there is only reference cluster exists and (4.23) becomes

$$\Pr\left(A \mid \Gamma^{(r)} \in Q_k^{(r)}\right) = \sum_{m^{(r)}=0}^{M^{(r)}-1} \frac{1}{1+m^{(r)}} \binom{M^{(r)}-1}{m^{(r)}} \left[F_{\Gamma^{(r)}}\left(q_{k+1}^{(r)}\right) - F_{\Gamma^{(r)}}\left(q_k^{(r)}\right)\right]^{m^{(r)}} \left[F_{\Gamma^{(r)}}\left(q_k^{(r)}\right)\right]^{M^{(r)}-m^{(r)}-1} \tag{4.24}$$

Thus, heterogeneous cell model becomes homogeneous cell model with T=1, Eq. (4.24) is equal to eq. (4.8).

Since there are K quantization level, capacity of the reference user is

$$C_{CQMS\_randuser}^{(r)} = \sum_{k=0}^{K-1} \Pr\left(A \mid \Gamma^{(r)} \in Q_k^{(r)}\right) \int_{Q_k^{(r)}} \frac{1}{2} \log_2(1+\gamma) f_{\Gamma^{(r)}}(\gamma) d\gamma \tag{4.25}$$

The unit of eq. (4.25) is bits/channel use/user in *cluster-r*. The aggregate capacity for the cell becomes the addition of the all cluster total capacities.

$$C_{agg\_CQMS\_randuser} = \sum_{t=1}^{T} M^{(t)} C_{CQMS\_randuser}^{(t)} \tag{4.26}$$

Example A.3 in Appendix shows the calculation of aggregate capacity for T=3.

4.3.2 Random Cluster Selection Scheduling Rule

The procedure is as follows,

I. sort users into clusters with respect to their mean SNR values
II. collect feedbacks (instantaneous SNR values) sent by each user according to the pre-determined threshold(s)
III. select the users which are in the highest quantization level in the clusters
IV. choose one of the clusters randomly
V. choose one user randomly (from the cluster chosen in IV) among users selected in III give the service

There are T clusters and the reference user's cluster is *cluster-r*, $r \in \{1,2,...T\}$. Any cluster other than the reference user's is *cluster-i*. Every user in *cluster-i* has $\Gamma^{(i)}$. The probability of the reference user chosen in the k^{th} quantization level given the other users of the clusters,

$$\Pr\left(\Gamma^{(r)} \mid Q_k^{(r)}\right)=\sum_{m^{(r)}=0}^{M^{(r)}-1}\frac{1}{1+m^{(r)}}\binom{M^{(r)}-1}{m^{(r)}}\left[\Pr\left(\Gamma^{(r)}\in Q_k^{(r)}\right)\right]^{m^{(r)}}\left[\Pr\left(\Gamma^{(r)}\in\bigcup_{l<k}Q_l^{(r)}\right)\right]^{M^{(r)}-m^{(r)}-1}$$

(4.27)

where

$$\Pr\left(\Gamma^{(r)}\in Q_k\right)=F_{\Gamma^{(r)}}\left(q_{k+1}^{(r)}\right)-F_{\Gamma^{(r)}}\left(q_k^{(r)}\right) \tag{4.28}$$

and

$$\Pr\left(\Gamma^{(r)}\in\bigcup_{l<k}Q_l^{(r)}\right)=F_{\Gamma^{(r)}}\left(q_k^{(r)}\right) \tag{4.29}$$

The event that the number of users in *cluster-i*, whose instantaneous SNR values lie in the k^{th} quantization level, is not equal to zero, is denoted as $s_{k,1}^{(i)}$ which means at least one user's instantaneous SNR is over $q_k^{(i)}$. The probability of this event is

$$\Pr\left(\Gamma^{(i)} \mid Q_k^{(i)}, s_{k,1}^{(i)}\right)=\sum_{m^{(i)}=1}^{M^{(i)}}\binom{M^{(i)}}{m^{(i)}}\left[\Pr\left(\Gamma^{(i)}\in Q_k^{(i)}\right)\right]^{m^{(i)}}\left[\Pr\left(\Gamma^{(i)}\in\bigcup_{l<k}Q_l^{(i)}\right)\right]^{M^{(i)}-m^{(i)}} \tag{4.30}$$

The event, denoted by $s_{k,0}^{(i)}$, describes none of the users in *cluster-i* can exceed $q_k^{(i)}$ of k^{th} quantization level then

$$\Pr\left(\Gamma^{(i)} \mid Q_k^{(i)}, s_{k,0}^{(i)}\right)=\left[\Pr\left(\Gamma^{(i)}\in\bigcup_{l<k}Q_l^{(i)}\right)\right]^{M^{(i)}} \tag{4.31}$$

The probability of selecting *cluster-r* is equal to

$$\Pr\{\text{selection of cluster - r}\}=\frac{1}{1+\sum_{\substack{z=1\\ z\neq r}}^{T} i^{(z)}} \tag{4.32}$$

where,

$$i^{(t)}=\begin{cases}0, & s_{k,0}^{(t)}\\ 1, & s_{k,1}^{(t)}\end{cases} \tag{4.33}$$

Thus, the probability of scheduling to the reference user in the k^{th} quantization level is

$$\Pr\left(A \mid \Gamma^{(r)} \in Q_k^{(r)}\right) = \Pr\left(\Gamma^{(r)} \mid Q_k^{(r)}\right) \sum_{i^{(1)}=0}^{1} \Pr\left(\Gamma^{(1)} \mid Q_k^{(1)}, s_{k,i^{(1)}}^{(1)}\right) \times \ldots \times$$
$$\sum_{i^{(r-1)}=0}^{1} \Pr\left(\Gamma^{(r-1)} \mid Q_k^{(r-1)}, s_{k,i^{(r-1)}}^{(r-1)}\right) \times \sum_{i^{(r+1)}=0}^{1} \Pr\left(\Gamma^{(r+1)} \mid Q_k^{(r+1)}, s_{k,i^{(r+1)}}^{(r+1)}\right) \times \ldots \times$$
$$\sum_{i^{(T)}=0}^{1} \Pr\left(\Gamma^{(T)} \mid Q_k^{(T)}, s_{k,i^{(T)}}^{(T)}\right) \frac{1}{1 + \sum_{\substack{z=1 \\ z \neq r}}^{T} i^{(z)}} \tag{4.34}$$

Eq. (4.34) becomes eq (4.24) with T=1. Since there are K quantization levels, the capacity belonging to the reference user is

$$C_{CQMS\_randcluster}^{(r)} = \sum_{k=0}^{K-1} \Pr\left(A \mid \Gamma^{(r)} \in Q_k^{(r)}\right) \int_{Q_k^{(r)}} \frac{1}{2} \log_2(1+\gamma) f_{\Gamma^{(r)}}(\gamma) d\gamma \tag{4.35}$$

The unit of (4.35) is bits/channel use/user in *cluster-r*. The total capacity of the cell, *aggregate capacity* for the cell becomes the addition of the all clusters total capacities. That is,

$$C_{agg\_CQMS\_randcluster} = \sum_{t=1}^{T} M^{(t)} C_{CQMS\_randcluster}^{(t)} \tag{4.36}$$

Example A.4 in Appendix shows the calculation of aggregate capacity for T=3.

The probability of scheduling to the reference user in the k^{th} quantization level for both methods is calculated using the following parameters,

i) number of users in every cluster
ii) mean SNR of every cluster (mean SNR values of identical users)
iii) thresholds for the quantization levels

The capacity belonging to each user is calculated with

i) number of quantization levels
ii) the probability of scheduling the reference user for every quantization level

The aggregate capacity is calculated with the followings,

i) number of clusters

ii) the user capacities

If all the number of clusters, the number of users in every cluster, mean SNR of the clusters and the number of quantization levels are fixed, the capacity achieved by each user will become a function of the thresholds. Hence, an optimization problem comes up: finding the thresholds which maximizes the capacity.

4.4 THRESHOLD OPTIMIZATION

A limited feedback rule is used for both homogeneous and heterogeneous cell models. Therefore, the base station has a role of defining a scheduling mechanism by determining thresholds which quantize SNR values. Hence, finding thresholds is critical for capacity maximization. In Appendix A.5, an example of optimum threshold calculation for homogeneous cell with two users is given. It is found as

$$q^* = e^{E_1(1/\bar{\gamma})e^{1/\bar{\gamma}}} - 1 \tag{4.37}$$

The optimum threshold depends on mean SNR, which can be seen in eq. (4.37), and also on the number of users in the cell, which is not obvious in eq. (4.37) but the following discussions illustrate the issue. The example in the Appendix A.5 is very simple with one threshold and two users in a homogeneous cell, it is easy to calculate the optimum threshold analytically. However, if there are more than two users, the problem becomes more complex. The optimum threshold is found for the same cell model as in Appendix A.5 but ten users. Due to the complexity of the capacity equation (4.9) for ten users, calculation of the optimum threshold has to be done numerically. The mean SNR is taken 5 dB and eq. (4.9) is plotted within the threshold interval of [-10, 15] dB in Figure 4.3.

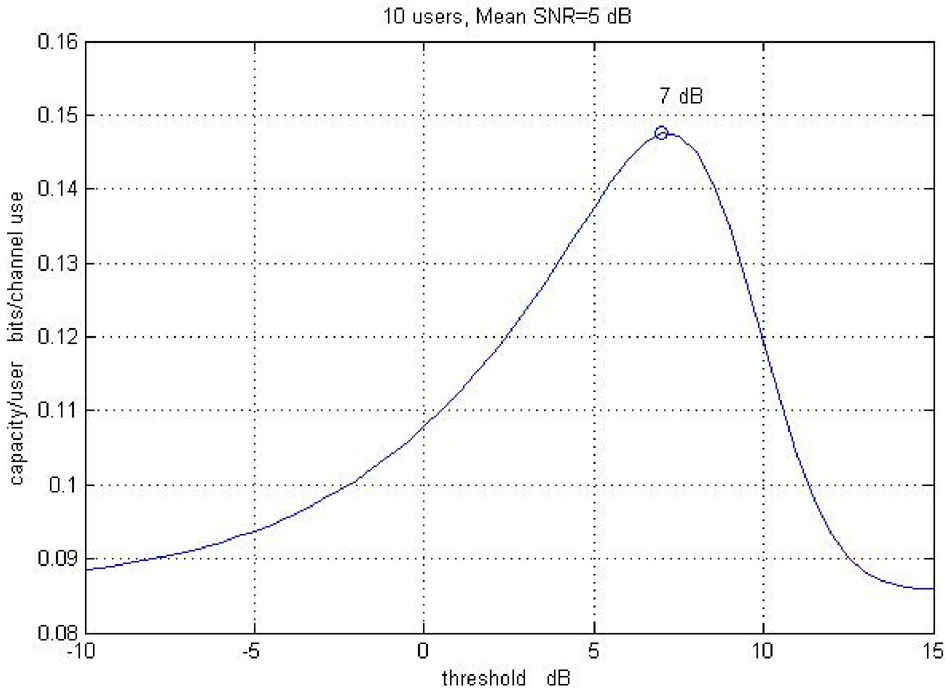

Figure 4.3 Capacity vs threshold value for 10 users

Considering the capacity curve plotted in Figure 4.3, if the threshold is defined so low, there will be many users above the threshold and they send feedback to the base station. Among those users, the transmitter randomly selects one of them for service. The difference between the best user and the worst user among the users which has sent feedback can be large. If the threshold is defined so high, it is highly probable that none of the users are above the threshold. In this situation the multiuser diversity can not be used effectively as previously.

The following graph is helpful to show that the threshold value also depends on number of users. In Figure 4.4 eq. (4.9) is plotted for 2, 4, 6, 8 and 10 users.

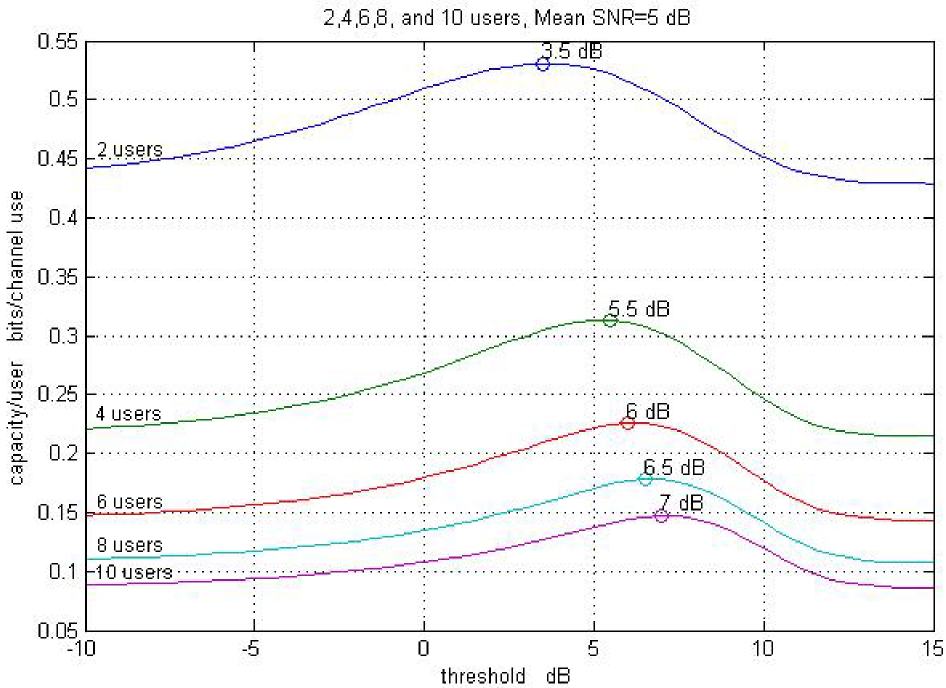

Figure 4.4 Capacity vs threshold value for 2, 4, 6, 8, and 10 users

As it can be seen from Figure 4.4, threshold value increases as the number of users in the cell increases. Addition of a user increases multiuser diversity. Therefore, the optimum threshold value to achieve the maximum capacity increases.

As a conclusion, capacity equations (4.9), (4.17), (4.25) and (4.35) have peaks for some threshold values. It is easy for a single threshold and a set of homogeneous users to calculate the optimum threshold. However, the optimization methods can be used for more complex situations.

In the previous sections, we have seen that the aggregate capacity of the cell and also the individual user capacity is a function of mean SNR belonging to the users and number of users. When users have identical mean SNR value, performance is only a function of user number and a single SNR value (homogeneous cell). However, if users have different mean SNR values, the problem becomes more difficult to solve. Here we propose to group users into *clusters* (clustered heterogeneous cell). Each cluster has a mean SNR level and assigned users. Any cluster is like a single homogeneous cell because of having identical users.

Figure 4.5 shows the process of optimization. The inputs are

→ cluster number

→ number of users in each cluster

→ mean SNR values of clusters

The outputs are

← thresholds

← capacity values calculated by thresholds found

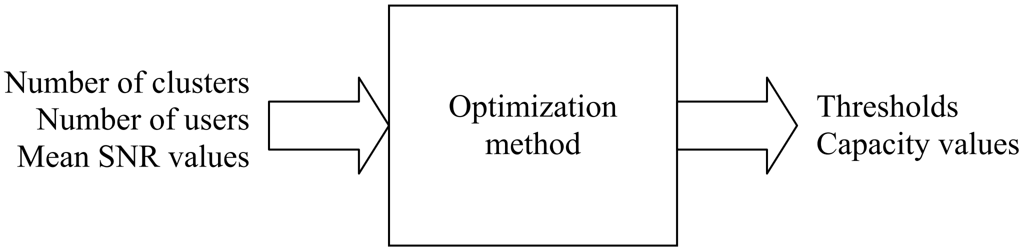

Figure 4.5 The inputs and outputs of the optimization method

Two optimization methods used in this study,

- *minimax optimization*
- *multivariable optimization*

a) minimax optimization

Minimax is a kind of multi-objective optimization (MOO). When there are multiple objective functions, this approach minimizes the maximum value of the objective functions. In other words, free variables control the risk value of user-1, user-2, etc. The minimax approach tries to minimize the maximum risk among all users.

b) multivariable optimization

Multivariable optimization, which is a single objective optimization (SOO), is done over a single objective function with multiple parameters. This can be interpreted as summing all the risks of users and then minimizing it instead of doing it individually.

4.5 THE METHODS TO DETERMINE THE THRESHOLDS FOR CLUSTERED USERS

4.5.1 Clustered Quantized Maximum SNR (CQMS) Method

In CQMS method, multi-objective optimization (MOO) is used to obtain the optimum thresholds. This method maximizes the capacity achieved by the worst cluster among all clusters. Therefore, it is a maxmin optimization.

The capacity equation C can be written as a function of thresholds q, mean SNR values $\bar{\gamma}$, number of quantization levels K, number of users vector M and number of clusters T. Thresholds can be put into q matrix,

$$q = \begin{bmatrix} q_1^{(1)} & q_2^{(1)} & \cdots & q_{K-1}^{(1)} \\ q_1^{(2)} & q_2^{(2)} & \cdots & q_{K-1}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ q_1^{(T)} & q_2^{(T)} & \cdots & q_{K-1}^{(T)} \end{bmatrix} \tag{4.38}$$

Mean SNR values can also be put into $\bar{\gamma}$ vector,

$$\bar{\gamma} = \begin{bmatrix} \bar{\gamma}^{(1)} & \bar{\gamma}^{(2)} & \cdots & \bar{\gamma}^{(T)} \end{bmatrix}^T \tag{4.39}$$

The optimization problem becomes,

$$\max_q \min_i C_{CQMS}^{(i)}(q, \bar{\gamma}, K, M, T)$$

subject to

$$q_0^{(t)} = 0 < q_1^{(t)} < \cdots < q_{K-1}^{(t)} < q_K^{(t)} = \infty \quad \text{for all } t \tag{4.40}$$

At the end of the optimization, user capacities in each cluster are obtained equalized by the optimum thresholds. However, in some cases, the optimization process can not find the optimum thresholds, especially when there is a large mean SNR gap between the clusters.

4.5.2 Clustered Quantized Maximum SNR Method With Normalization

CQMS with normalization method is produced by modifying the previous CQMS method. The normalization is introduced for two reasons:

1) having user capacities as a ratio of maximum SNR (MS) capacities
2) finding optimum thresholds with large mean SNR differences between the clusters exists

The normalization step, which is dividing the user CQMS capacity to MS capacity, can be inserted into optimization process of a multiple cluster cell model. The optimization problem is the modified form of the one in (4.40),

$$\max_{q} \min_{i} \frac{C^{(i)}(q,\bar{\gamma},K,M,T)}{C_{MS}^{(i)}}$$

subject to

$$q_0^{(t)} = 0 < q_1^{(t)} < \cdots < q_{K-1}^{(t)} < q_K^{(t)} = \infty \quad \text{for all } t \tag{4.41}$$

$C^{(i)}(q,\bar{\gamma},K,M,T)$ represents one of the equations (4.25) or (4.35). $C_{MS}^{(i)}$ is MS capacity for the i^{th} cluster. $C_{MS}^{(i)}$ is calculated as if the only cluster in the cell was the i^{th} cluster and all users in this cluster send their feedback fully to the base station which transmits to the best user. In [13], the maximum SNR is defined in the i^{th} cluster with

$$\Gamma_m^{(i)} = \max_{n \neq m} \Gamma_n^{(i)} \tag{4.42}$$

Since all users have the same mean SNR,

$$f_{\Gamma_n^{(i)}} = f_{\Gamma^{(i)}} \quad \text{for all n} \tag{4.43}$$

The probability of the m^{th} user's instantaneous SNR being the largest with the i^{th} cluster size $M^{(i)}$

$$\Pr(\Gamma_m^{(i)} \le \gamma) = \left[F_{\Gamma^{(i)}}(\gamma) \right]^{M^{(i)}-1} \tag{4.44}$$

The average capacity per slot per user is

$$C_{MS}^{(i)} = \int_0^\infty \frac{1}{2} \log_2 (1+\gamma) \left[F_{\Gamma^{(i)}}(\gamma) \right]^{M^{(i)}-1} f_{\Gamma^{(i)}}(\gamma) d\gamma \tag{4.45}$$

Since the MS capacity is the function of mean SNR and the user number in the cluster, the capacity ratio in the optimization problem in (4.41), called *loss factor*. The loss factor is equalized by the optimum thresholds at the end of the optimization process. As a result, the optimum thresholds equalize the loss in all clusters. This is the basic idea behind the normalization and as it is explained in Chapter 5 that some fairness is also gained by the described normalization.

Although the optimization problem (4.41) includes a ratio of capacities, the main interest is $C^{(i)}(q, \bar{\gamma}, K, M, T)$ and at the end of the process, optimum thresholds are found according to the ratio but $C^{(i)}(q, \bar{\gamma}, K, M, T)$ is calculated by multiplying the ratio with $C_{MS}^{(i)}$ which only depends on cluster size and mean SNR.

4.5.3 Clustered Quantized Weighted SNR (CQWS) Method

This method which is a quantized variant of scheme discussed in [7] is used when there are multiple clusters and single threshold for each one. The thresholds are defined like the mean SNR values of the corresponding clusters. That is,

- if the instantaneous SNR of a user is above the mean SNR of itself ($\hat{\Gamma}^{(i)} = \Gamma^{(i)} / \bar{\gamma}^{(i)} \ge 1$), it responds to the base station with a bit represents this case
- otherwise ($\hat{\Gamma}^{(i)} = \Gamma^{(i)} / \bar{\gamma}^{(i)} < 1$) the user stays in silence

The base station collects the feedbacks, which are all identically distributed due to weighting, from all users and chooses one of them to transmit randomly. From the above decision rule, this method is used for only one threshold.

According to the method, the PDF of $\hat{\Gamma}^{(i)}$ $f_{\hat{\Gamma}^{(i)}}$, is modified by using eq. (4.2) [23],

$$f_{\hat{\Gamma}^{(i)}}(\hat{\gamma}) = e^{-\hat{\gamma}} = e^{-\gamma/\bar{\gamma}^{(i)}} \tag{4.46}$$

Eq. (4.46) shows that the PDF of instantaneous SNR is same for all users. This means all users can be considered equally by base station no matter what their mean SNR values are. However, a user with a good channel state can achieve more capacity when it is given the channel access.

4.5.4 Clustered Quantized Maximum SNR With Scheduling Outage (CQMS-P_o) Method

The scheduling outage criteria can provide us only a single threshold due to its definition which is stated in [8]. The *scheduling outage* for the cluster is the event that the instantaneous SNR for all users falling below the lowest threshold. The probability of scheduling outage is

$$\begin{aligned} P_o &= \Pr\left(\Gamma_k < q \text{ for all } k\right) \\ &= \left(\mathrm{F}_{\Gamma_k}(q)\right)^M = \left(1 - e^{-q/\bar{\gamma}}\right)^M \end{aligned} \tag{4.47}$$

Taking out q in eq. (4.47),

$$q = -\bar{\gamma}\ln\left(1 - P_o^{\,1/M}\right) \tag{4.48}$$

The threshold, eq. (4.48), can be calculated for each cluster. Hence, it is possible to calculate a single threshold for every cluster given a probability of scheduling outage.

It is easy to find thresholds with this method since they can be directly written as a function of mean SNR, scheduling outage probability and number of users unlike in random user and random cluster selection rules. The thresholds found from outage requirement are put into eq. (4.25) or (4.35) to calculate user capacity. Since the dependency of user capacity on thresholds is still valid, now the optimization is handled over the scheduling outage probability interval, (0,1] instead of threshold interval, [0,∞). The optimum thresholds are found again as maximizing the user capacities.

4.5.5 Clustered Quantized Maximum SNR with Aggregate Capacity Maximization (CQMS-max-agg) Method

Different from the other methods which give the optimum thresholds by maximizing the user capacities in each cluster, CQMS-max-agg maximizes the aggregate capacity. Therefore this method is the maximum SNR (MS) method with quantized feedbacks. Since there is a single objective function, a single object optimization (SOO) is used to find the thresholds.

The capacity equation C can be written as a function of thresholds q, mean SNR values $\bar{\gamma}$, number of quantization levels K, number of users vector M and number of clusters T. Thresholds can be put into q matrix,

$$q = \begin{bmatrix} q_1^{(1)} & q_2^{(1)} & \cdots & q_{K-1}^{(1)} \\ q_1^{(2)} & q_2^{(2)} & \cdots & q_{K-1}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ q_1^{(T)} & q_2^{(T)} & \cdots & q_{K-1}^{(T)} \end{bmatrix} \tag{4.49}$$

Mean SNR values can be also put into $\bar{\gamma}$ vector,

$$\bar{\gamma} = \begin{bmatrix} \bar{\gamma}^{(1)} & \bar{\gamma}^{(2)} & \cdots & \bar{\gamma}^{(T)} \end{bmatrix}^T \tag{4.50}$$

The optimization problem as follows,

$$\max_q C_{CQMS-\max-agg}(q, \bar{\gamma}, K, M, T)$$

subject to

$$q_0^{(t)} = 0 < q_1^{(t)} < \cdots < q_{K-1}^{(t)} < q_K^{(t)} = \infty \quad \text{for all } t \tag{4.51}$$

It is obvious that high mean SNR cluster always dominates the scheduling job, since CQMS-max-agg is similar to MS method.

CHAPTER 5

NUMERICAL RESULTS AND COMPARISONS

In the previous chapter, the homogeneous and heterogeneous cell models are discussed. Partial feedback schemes are used to schedule users. The schemes can be optimized for different criterions. A base station handles scheduling according to the received feedback. In this chapter these methods are compared.

5.1 THE EFFECTS OF MULTIUSER DIVERSITY ON CAPACITY

In Figure 5.1 the capacity for ten users are plotted for round robin (RR) scheduling and multiuser diversity (maximum SNR) scheduling methods at different mean SNR levels.

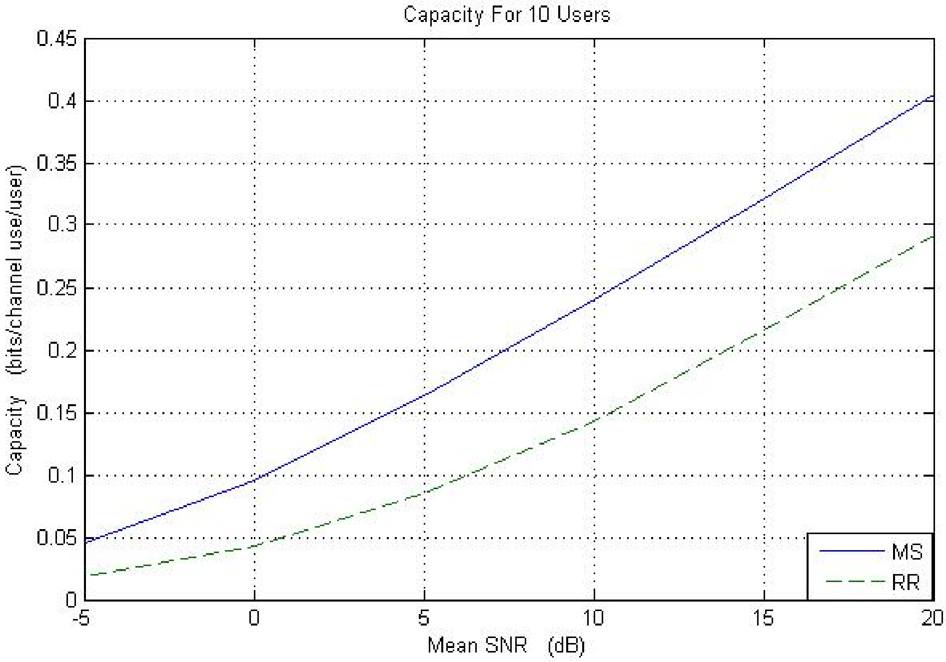

Figure 5.1 Capacity of RR and MS methods for 10 users [12]

As it is examined in Chapter 3, multiuser diversity benefits from increasing number of users which also increases the probability of selecting the best user which has better instantaneous SNR. The deviation from the RR capacity curve is the multiuser diversity gain.

Multiuser diversity takes the advantage of the number of users. The more users the cell includes, the more aggregate capacity can be achieved. However, RR method has fixed aggregate capacity for any number of users. That is because RR does not use the diversity and picks users regardless of their channel quality. Figure 5.2 shows aggregate capacities obtained by MS and RR methods separately.

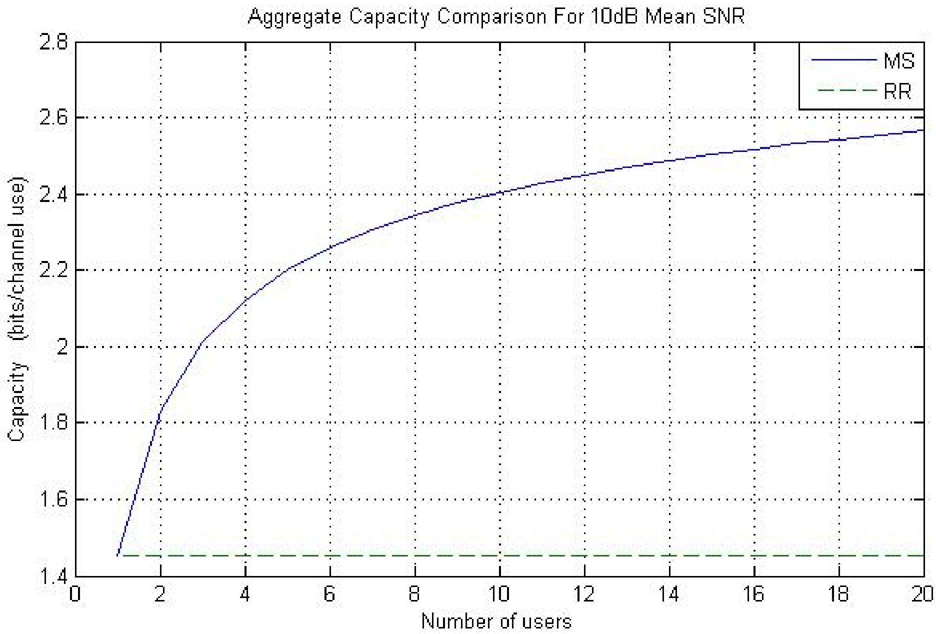

Figure 5.2 Aggregate capacity comparison of MS and RR against number of users

In Figure 5.3, capacity gain of multiuser diversity over RR scheduling is plotted by dividing the capacity achieved by multiuser diversity with capacity achieved by RR method.

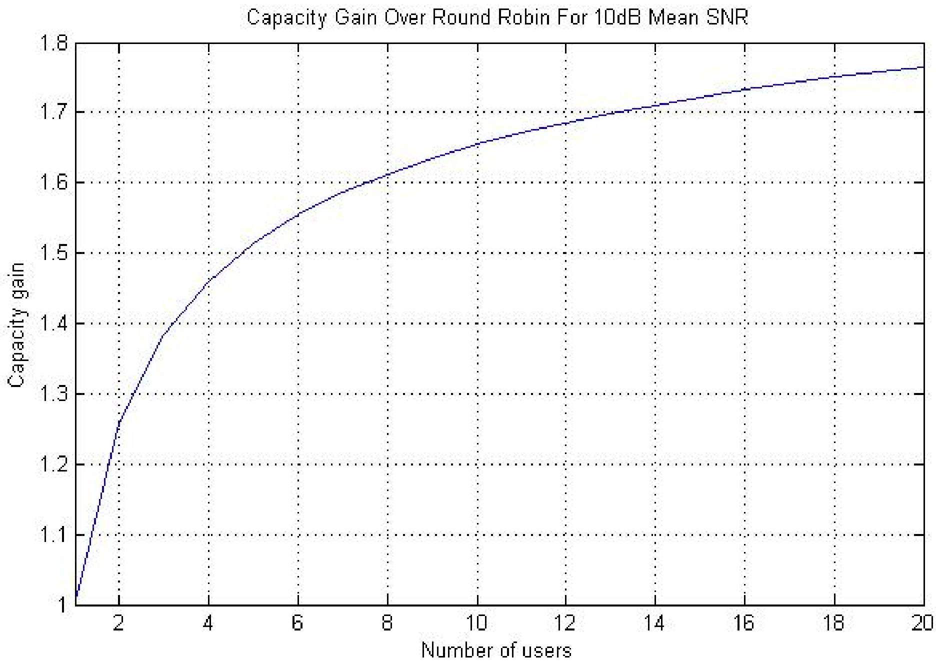

Figure 5.3 The capacity gain by using multiuser diversity against increasing number of users

RR chooses users for the same time slot uniformly. Therefore, the probability of user channel access is still equal to each other. MS method also preserves the user channel access probabilities for homogeneous users. While the best user is chosen, the probability of user selection is the same for every user in a sufficiently long time interval of service. Thus, the probability of user channel access for MS method is equal to one found for RR, showed in Figure 5.4.

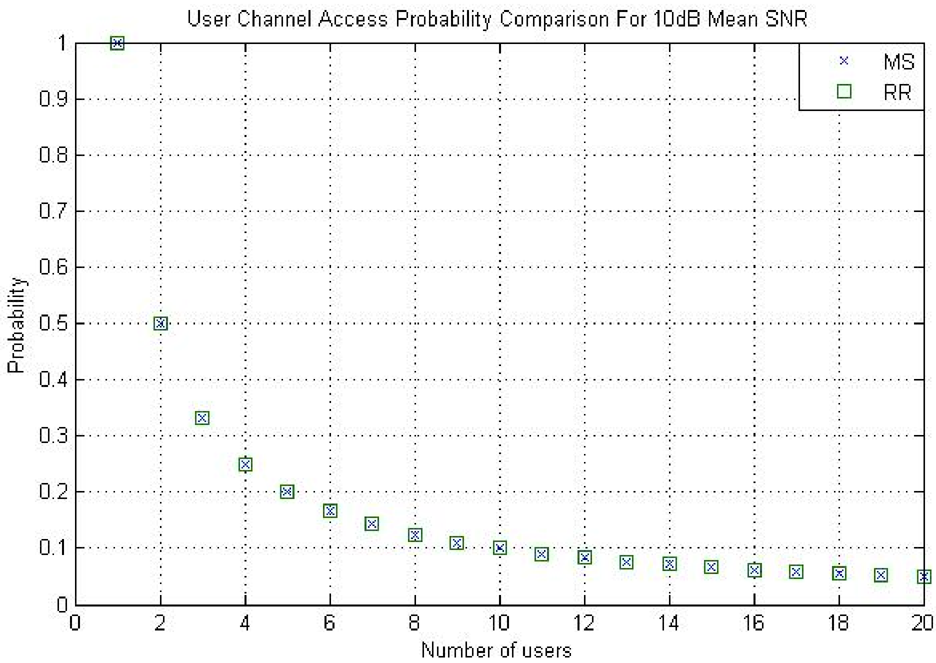

Figure 5.4 Probability of user channel access against increasing number of users

5.2 THE EFFECT OF FEEDBACK QUANTIZATION ON CAPACITY IN THE HOMOGENEOUS CELL

Quantized feedback in the homogeneous cell model is mentioned in the previous chapter. The formulation of the capacity is derived in Chapter 4. The optimum thresholds are found by maximizing eq. (4.9) with respect to thresholds. After finding the optimum thresholds, they are used in simulation of the method. The comparison of capacities acquired by theoretical calculation and simulation is showed in Figure 5.5 below.

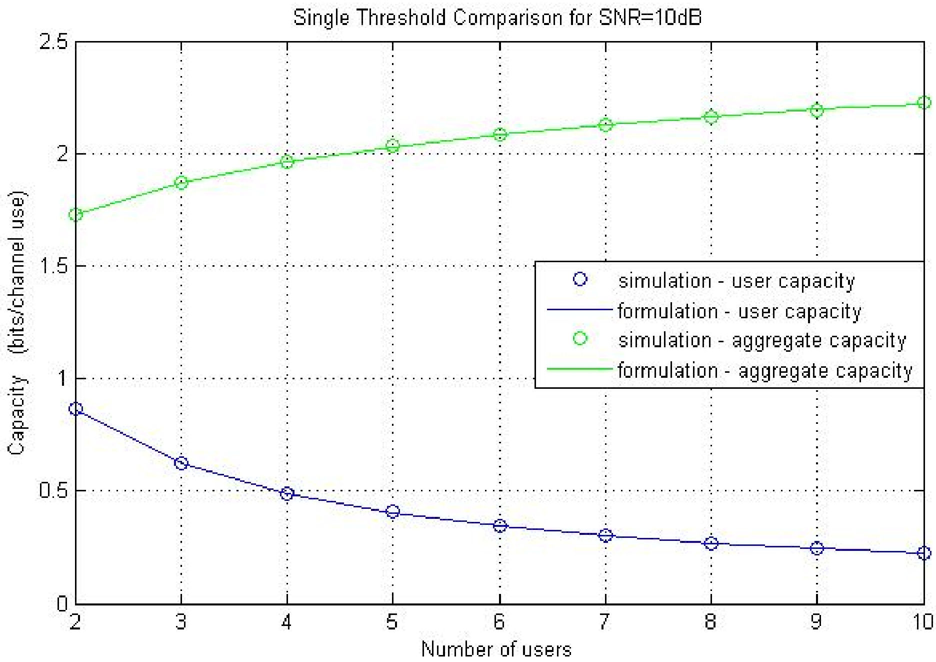

Figure 5.5 Comparison of theoretical and simulation results

Mean SNR of each user is equal to 10 dB and the acquired capacity is plotted against the number of users with a single threshold. As it can be seen from the figure, both analytical and simulation results give the same values. Since the resource is limited for all users, as the number of users increases the capacity achieved by each user decreases. On the other hand, the aggregate capacity increases as the number of users increases due to enhancement in diversity.

Having multiple thresholds for feedback quantization approaches the full feedback case. This can be seen on Figure 5.6. The ratio of the capacity gained by quantized feedback over the capacity gained by full feedback is drawn against number of thresholds.

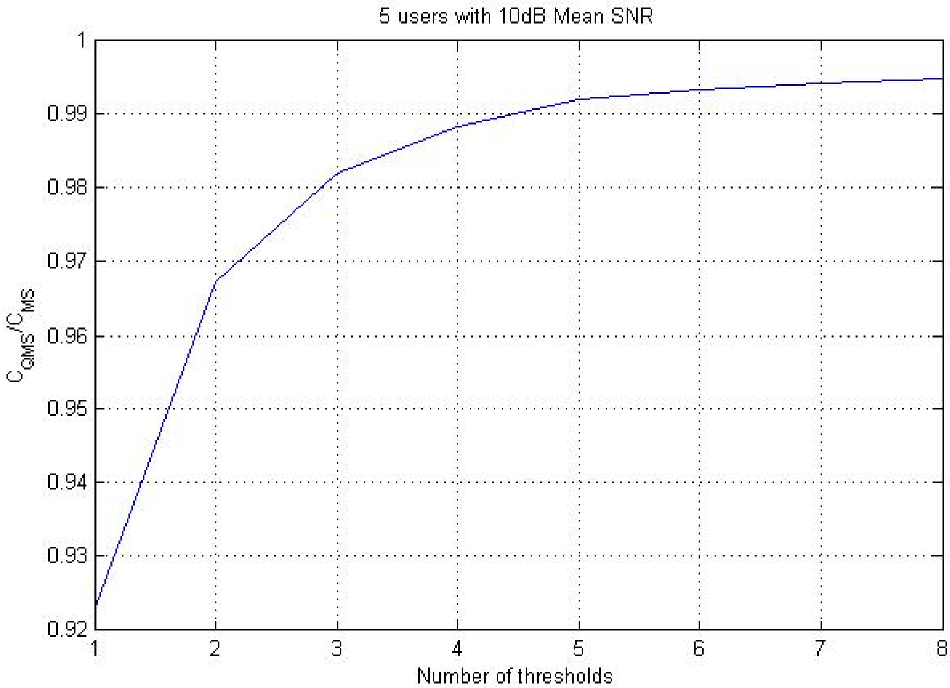

Figure 5.6 The capacity achieved by quantized feedback normalized with full feedback capacity

As it is seen from Figure 5.6 above, the capacity achieved with a single threshold quantization (C_{QMS}), which means one bit feedback is sufficient for the SNR interval, can exceed 92% the capacity achieved by full feedback (C_{MS}). Since the quantization levels increase, the capacity approaches the full feedback case at higher number of thresholds.

It has been mentioned in the previous chapters that the number of users is the key parameter for multiuser diversity gain. The effect of the number of users on the quantization levels are shown in Figure 5.7.

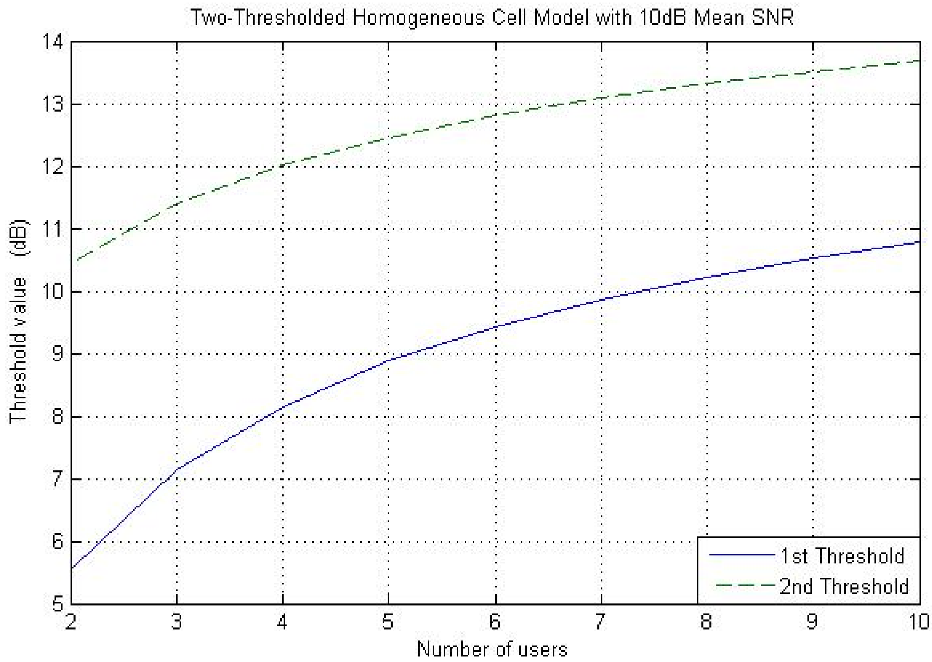

Figure 5.7 Threshold values found by optimization for 10 dB mean SNR

As in the full feedback scheme, the capacity increases as the number of user increases also in the quantized feedback scheduling. The threshold value becomes larger for higher number of users.

An investigation of the effect of single optimum threshold, determined for a number of users, on the interval of mean SNR is studied in [12]. The results are shown in Figure 5.8. Namely, the optimum threshold is first found for all users having 5 dB mean SNR. Then, the threshold for 5 dB mean SNR is used for the mean SNR range which is between -5 and 20 dB. The capacity found by MS and single threshold case are plotted relative to ones found by RR scheme. The capacity gains are calculated for 2, 10 and 20 users. Even the capacity achieved by single threshold can be equal or greater than in 25 dB SNR interval. The gain can reach the top and also approaches MS gain at 5 dB. Increasing the number of users also increases the capacity gain due to getting more diversity. On the other hand, increasing the number of users also causes a rise in the threshold value. This makes sense because there will be more users whose instantaneous SNR value above the threshold with growing number of users and fixed threshold. This causes the probability of choosing a user which has lower SNR will be higher. Hence, the threshold value gets more in the optimization to maximize the capacity. On the other hand, a very high threshold is not also suitable for maximizing the capacity because the probability of any user being above the threshold will be very low. Because of high and low mean SNR users are below the threshold, the achieved capacity will be lower.

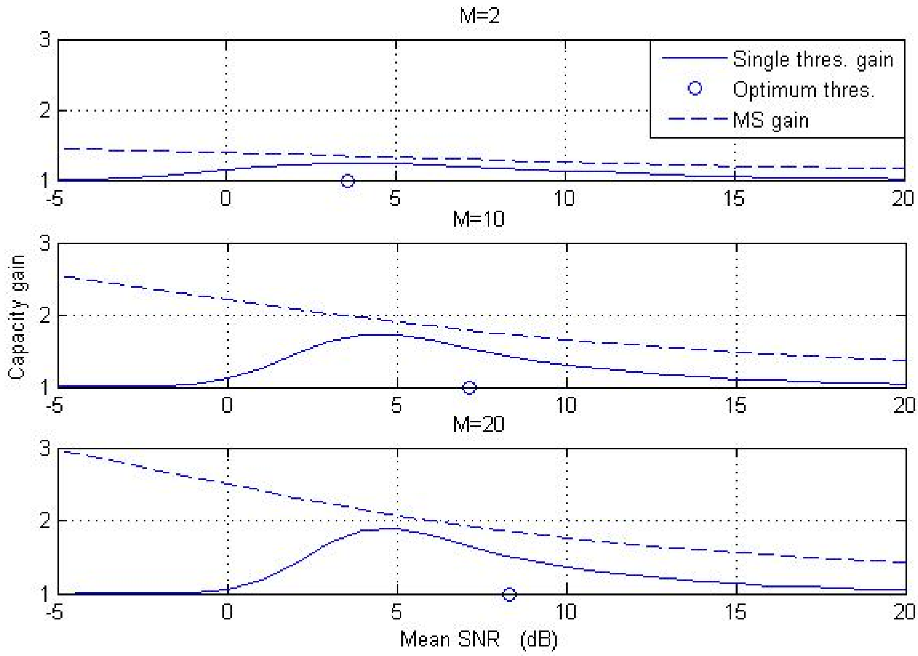

Figure 5.8 Capacity gains of a user relative to RR for 2–,10–, and 20–user cells [12]

The plots in Figure 5.8 also show the effect of multiuser diversity on the capacity gain. As the number of user increases, the capacity gain also increases. Namely, capacity gain is much higher in the case of 20 users than that of 2 users.

5.3 THE EFFECT OF CLUSTERED HETEROGENEOUS CELL MODEL ON CAPACITY

In Chapter 4, two rules are mentioned to schedule users in clustered heterogeneous cell model: *random user selection* and *random cluster selection* rules. A comparison of these methods with respect to the aggregate capacity against increasing number of users for two clusters is showed in Figure 5.9 and 5.10. In Figure 5.9, the capacity curves are plotted against the number of users in high SNR cluster. The same curves for increasing number of users in low SNR cluster are plotted in Figure 5.10. *Multi-objective optimization* is used to find the optimum thresholds for both random user selection and random cluster selection rules. In each figure, cluster capacities, which are equal to the summation of capacity achieved by each user in the cluster, and aggregate capacity, which is the summation of all cluster capacities, are shown.

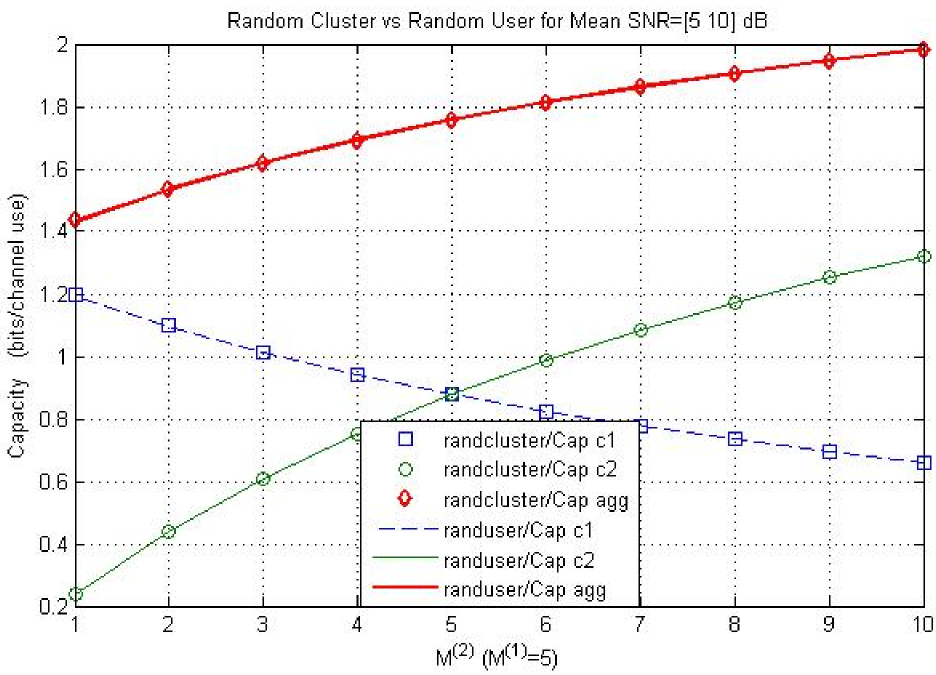

Figure 5.9 Random cluster selection rule vs random user selection rule with increasing number of users in high SNR cluster

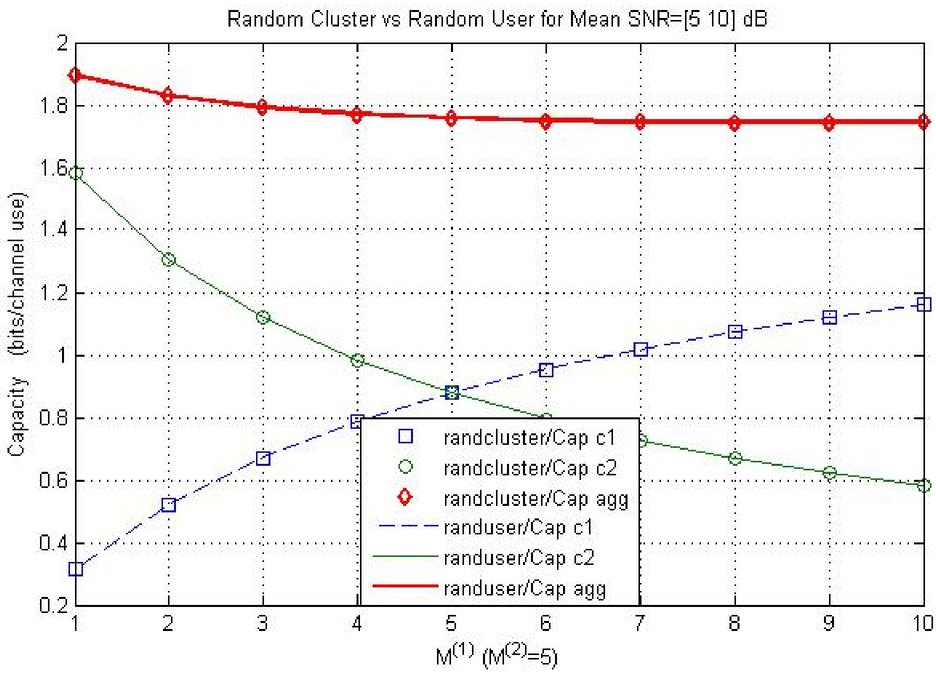

Figure 5.10 Random cluster selection rule vs random user rule with increasing number of users in low SNR cluster

As seen on both Figure 5.9 and 5.10, the achieved capacities are too close to each other for two methods. Using random cluster selection rule is a better choice due to its lower computational complexity.

For Figure 5.9, increasing the number of users also increases the cluster capacity to which the users belong because the number of users in low SNR cluster is fixed and each user coming from the high SNR cluster enhance its cluster capacity as well as the aggregate capacity. Unlike Figure 5.9, if the number of users belonging to the low SNR cluster increases, the cluster capacity also increases while the aggregate capacity decreases because the addition of each user to the low SNR cluster can not contribute as much as in the former case.

A common point for two figures is that capacity achieved by increasing number of users decreases the capacity achieved by the users of the other cluster. In order to alleviate this issue, the capacity normalization, is introduced in the optimization process.

5.4 THE EFFECTS OF MS NORMALIZATION ON CAPACITY IN THE OPTIMIZATION PROCESS

In Figure 5.11 and 5.12, capacity curves are plotted in the same way as the previous two figures except the optimum thresholds found by one user's MS capacity normalization as mentioned in Chapter 4 (loss factor equalization).

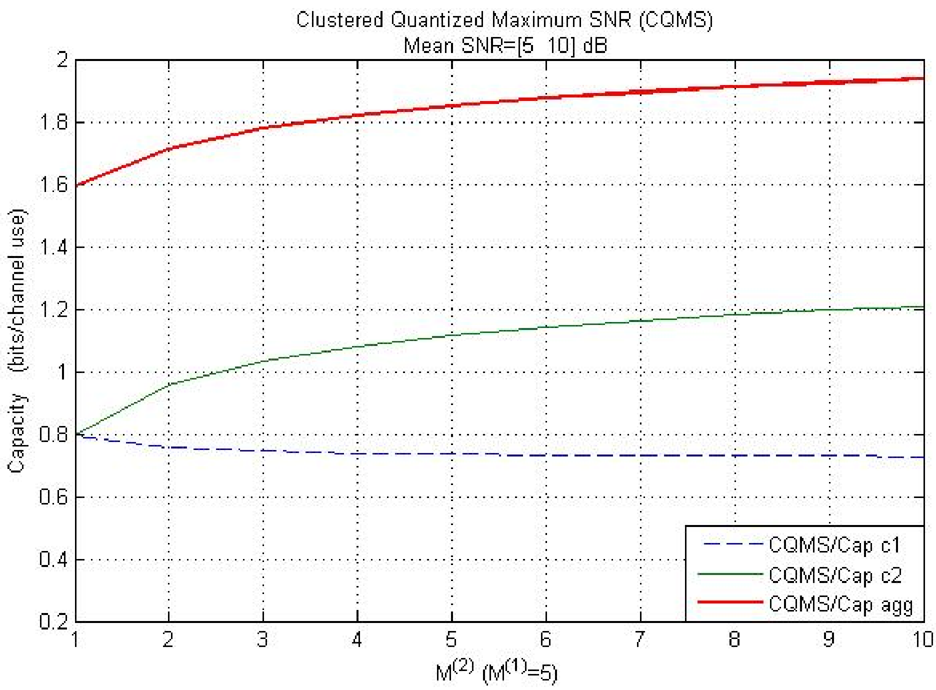

Figure 5.11 The random cluster selection rule with one user's MS capacity normalization is plotted against the number of users in high SNR cluster

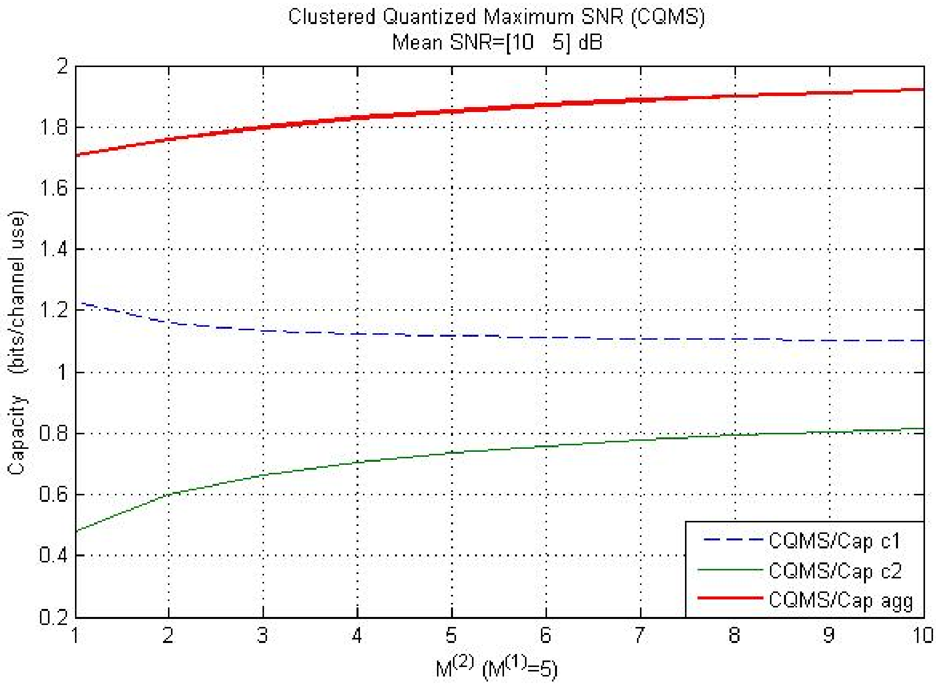

Figure 5.12 The random cluster selection rule with one user's MS capacity normalization is plotted against the number of users in low SNR cluster

Both Figure 5.11 and 5.12 are plotted with the modified optimization problem, discussed in Chapter 4 under the title of *Clustered Quantized Maximum SNR Method with Normalized Optimization*. The normalization step in the optimization problem makes the whole system to be fair for all users in any cluster. That is, adding a user to any cluster contributes to his cluster capacity and the aggregate capacity. On the other hand, the increase in capacity affects the other cluster very little. As a conclusion, increasing number of users does not affect very much the channel access of the other cluster.

5.5 THE COMPARISON OF CLUSTERED HETEROGENEOUS CELL MODEL WITH THE HOMOGENEOUS CELL

In this situation, the effects of scheduling the clustered users as if they are in a homogeneous cell are compared. These are two specific comparisons:

1) base station schedules using clustered users
2) base station schedules using homogeneous model

The flowchart in Figure 5.13 shows the steps of the analysis.

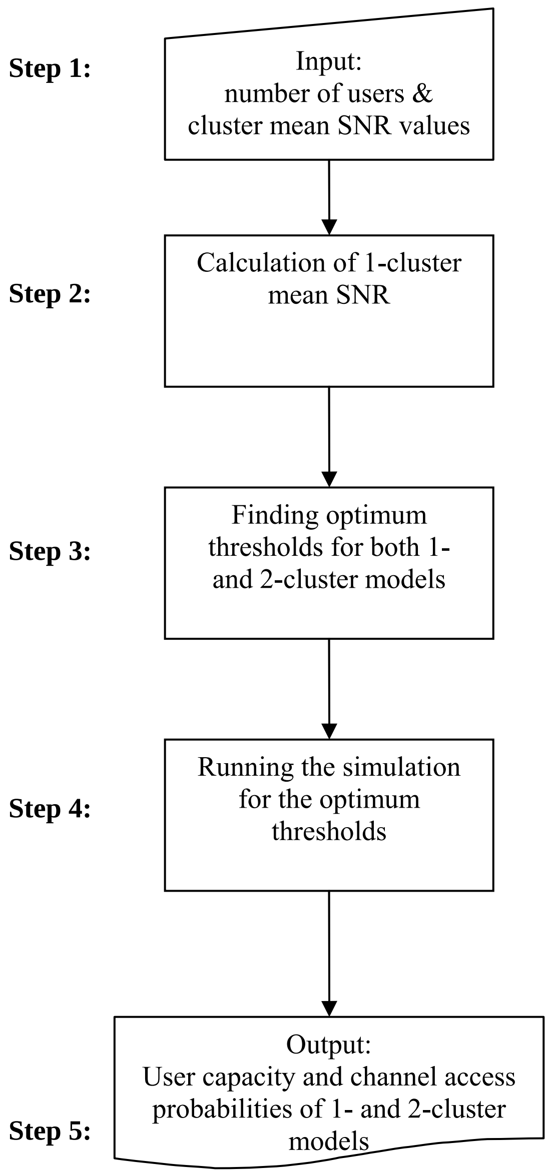

Figure 5.13 The flowchart of the 1- and 2-cluster model comparison

Step 1: User numbers in the clusters are determined. Each user in any cluster has the same channel statistics, that is, the same mean SNR.

Step 2: 1-cluster cell is created with the existing users. This cluster has all users from each cluster and the mean SNR of this single cluster is found by averaging the user mean SNR values.

Step 3: Each cluster has only one threshold in 2-cluster model and it is found by the optimization. The mean SNR calculated in Step 2 represents mean SNR of the 1-cluster model and is used to find the optimum threshold.

Step 4: Once the optimum thresholds for 1- and 2-cluster cells are obtained, Monte Carlo simulations are done.

Step 5: As outputs, user capacities and channel access probabilities are found and showed in Figure 5.14, 5.15, 5.16 and 5.17.

Figure 5.14 and 5.15 compare group/cluster and aggregate capacities.

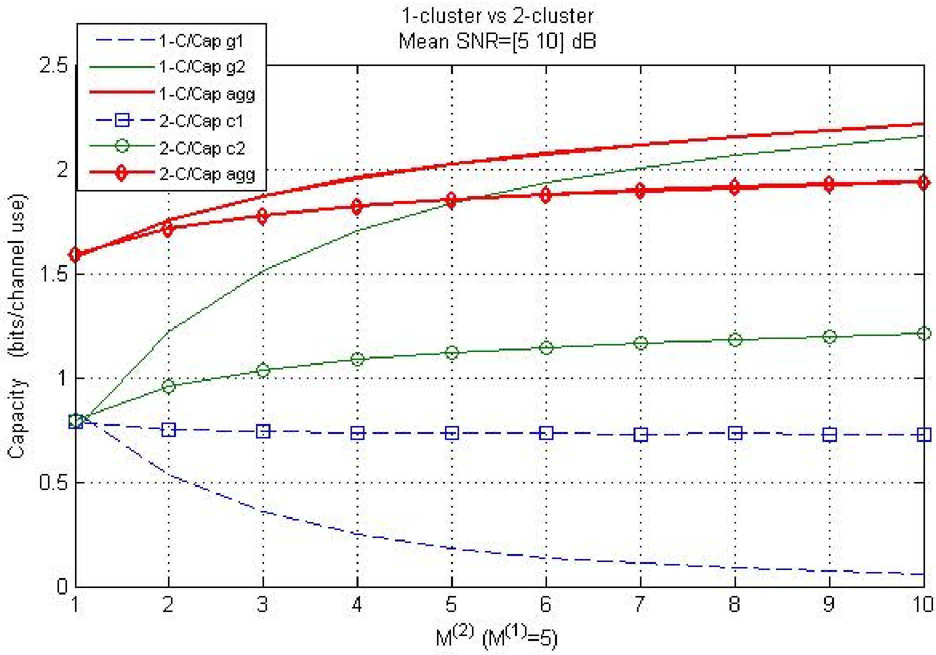

Figure 5.14 Comparison of 1- and 2-cluster cell models with increasing number of users in high SNR

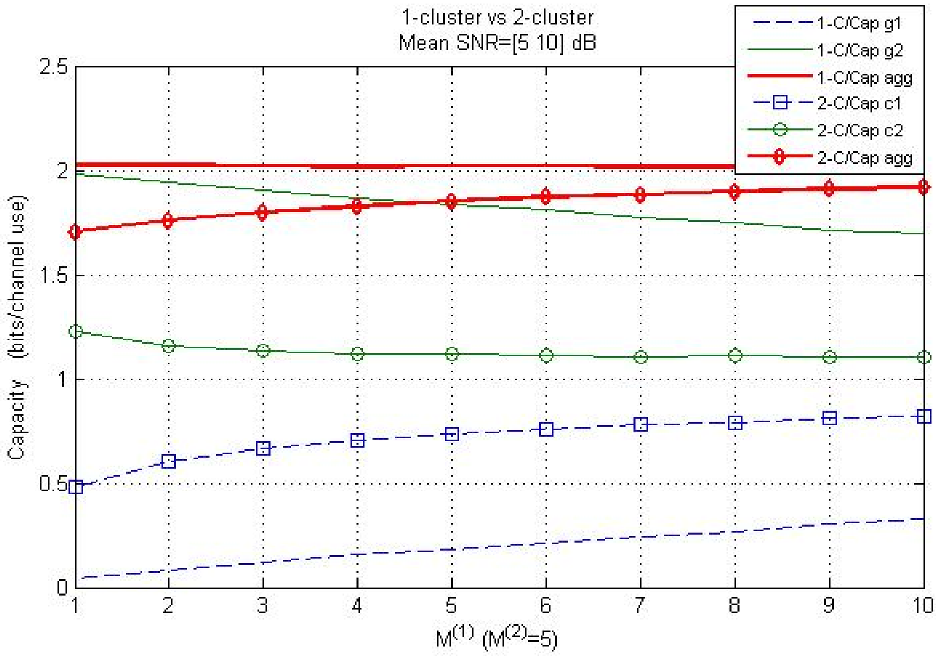

Figure 5.15 Comparison of 1- and 2-cluster cell models with increasing number of users in low SNR cluster

In Figure 5.14/5.15, there are two clusters and capacity curves are plotted against the number of users in the high/low SNR clusters is increasing. The curves with symbols are the same as Figure 5.11 and 5.12. The others represent the capacity curves of 1-cluster cell (homogeneous).

Groups, which include users with same mean SNR, exist in the 1-cluster cell. The capacity values achieved by all users which have same SNR values in the 1-cluster cell are also plotted for the sake of analogy.

If two clusters are united and become one cluster, the base station selects the best user in it. On the average, the users with high mean SNR values are chosen much often than the other ones which is also showed in the following two figures, Figure 5.16 and 5.17. Therefore, the high SNR group achieves better capacity than the other group for both cases. The capacity curves belonging to 1-cluster cell are very similar to MS method in which full feedback structure is used to schedule users are plotted in Figure 5.24 and 5.26. Another point is the addition of user to any group contributes its group capacity while decreases the other group.

Beside the capacity curves, channel access probabilities of groups/clusters are also plotted in Figure 5.16 and 5.17.

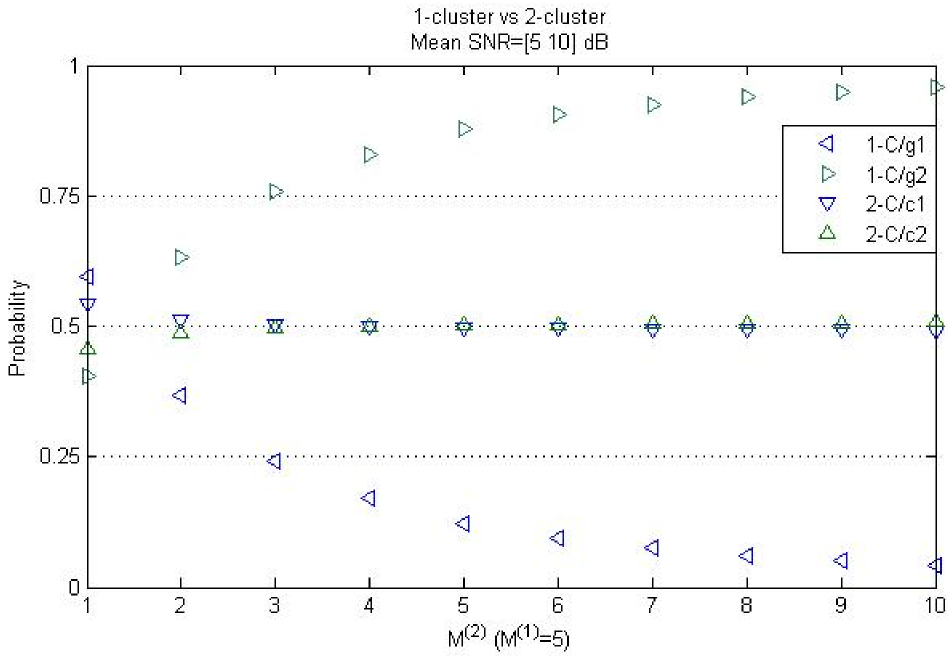

Figure 5.16 Comparison of channel access probabilities of 1- and 2-cluster cell models with increasing number of users in high SNR cluster

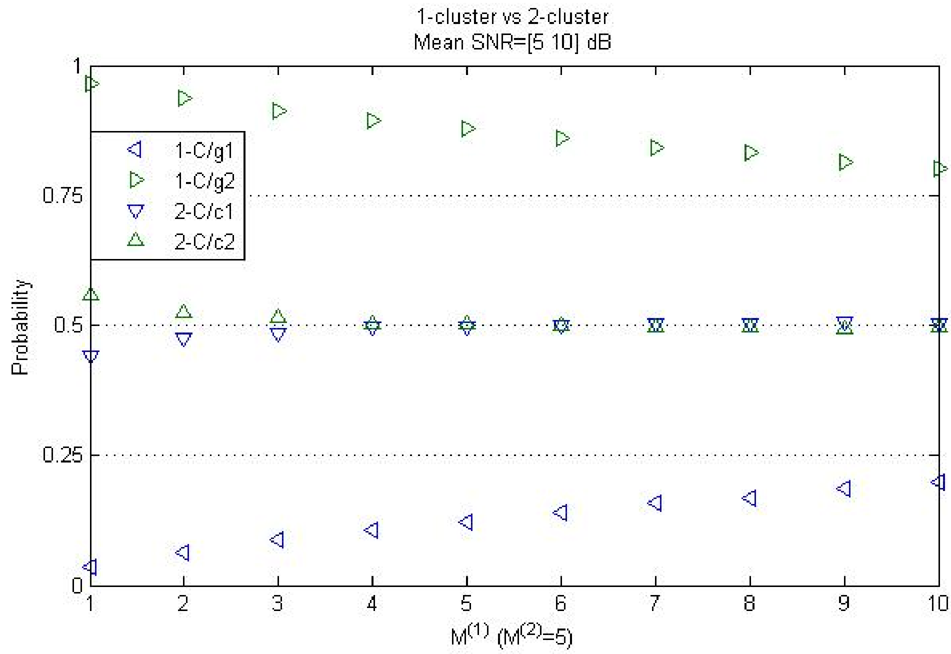

Figure 5.17 Comparison of channel access probabilities of 1- and 2-cluster cell models with increasing number of users in low SNR cluster

As it is seen from Figure 5.16 and 5.17, one user's MS capacity normalization in the optimization process brings out fairness which can be observed as equalized channel access probabilities of each cluster, found by summing all users channel access probabilities in the same cluster. On the other hand, the group probabilities are observed as higher for the high SNR group since choosing the best user frequently comes from this group.

5.6 THE CASE OF CLUSTERED HETEROGENEOUS CELL MODEL WITH NON-IDENTICAL USERS

A general case in which all users have different mean SNR values is presented and adapted to the clustered structure of heterogeneous cell model. The cell includes 12 users and the i^{th} user has i dB mean SNR. All users are ordered with respect to their mean SNR values. The cell is divided into 1 (homogeneous cell), 2, 3, 4 and 6 clusters. The number of users in every cluster is equal to each other. A simulation is handled for every clustered model. The procedure is the same as the previous analysis. According to the flowchart in Figure 5.13, the thresholds are found by using CQMS method with MS normalization.

The users ordered with respect to their mean SNR values in the clusters are showed in Table 5.1 and 5.2. The aim is putting users whose mean SNR values are close to each other into the same cluster. Hence, this combination of users is the best suitable one considering origin of the idea, clustering identical users.

Table 5.1 The user ranks in 1-, 2-, and 3-cluster cells

Rank	**1-cluster**	**2-cluster**		**3-cluster**		
1	user-12	user-6	user-12	user-4	user-8	user-12
2	user-11	user-5	user-11	user-3	user-7	user-11
3	user-10	user-4	user-10	user-2	user-6	user-10
4	user-9	user-3	user-9	user-1	user-5	user-9
5	user-8	user-2	user-8			
6	user-7	user-1	user-7			
7	user-6					
8	user-5					
9	user-4					
10	user-3					
11	user-2					
12	user-1					

Table 5.2 The user ranks in 4- and 6-cluster cells

Rank	**4-cluster**				**6-cluster**					
1	user-3	user-6	user-9	user-12	user-2	user-4	user-6	user-8	user-10	user-12
2	user-2	user-5	user-8	user-11	user-1	user-3	user-5	user-7	user-9	user-11
3	user-1	user-4	user-7	user-10						

The aggregate capacities for all clustered model and the user channel access probabilities are showed on Figure 5.18 and 5.19 respectively.

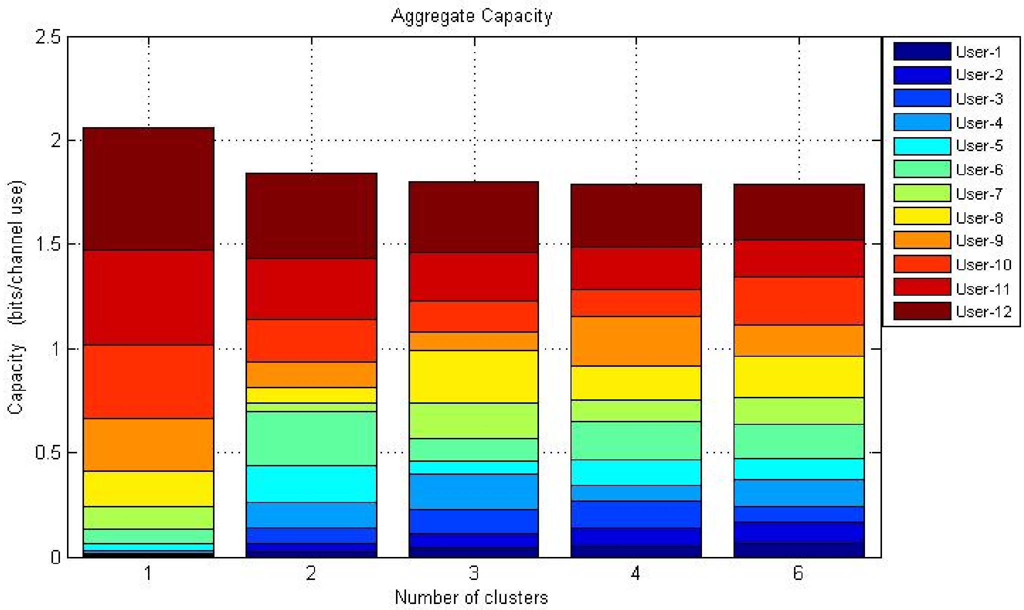

Figure 5.18 Aggregate capacities

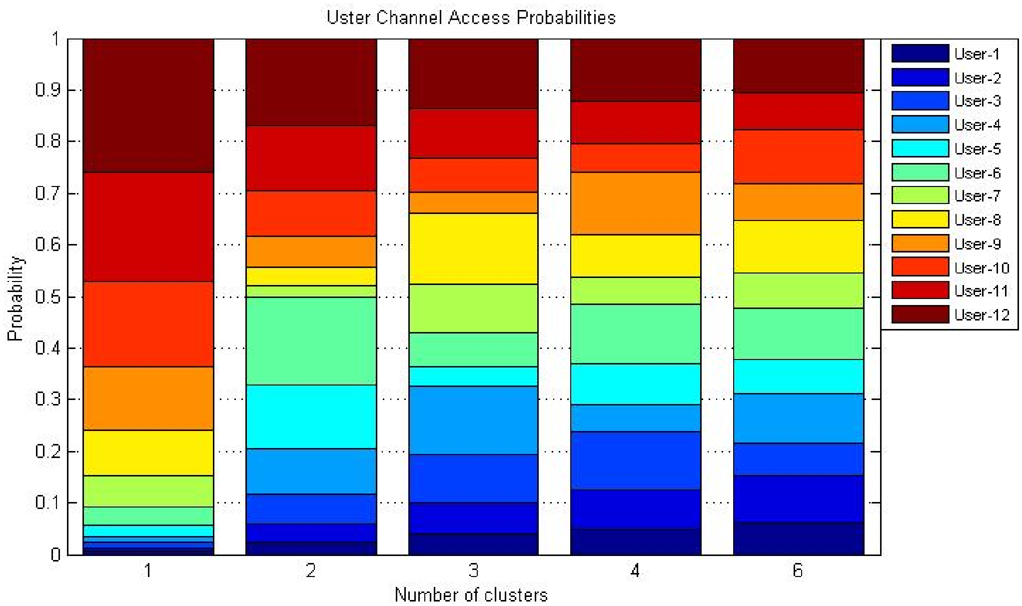

Figure 5.19 User channel access probabilities

It is obvious that the fewer clusters that the cell is divided, the more aggregate capacity it can achieve. This is because dividing the cell into the clusters affects the fairness. However, fairness is inversely proportional with aggregate capacity which can be maximized by giving the opportunity of channel access all the time for the best user in the cell. Hence, the other users get the channel very small time intervals. Another point in Figure 5.18 is that any user which has lower rank in any cluster can achieve less capacity when compared with a user having a better rank in another cluster number model. Besides rank, the cluster size is also a matter of concern that being in the same rank in a larger cluster is a better case. For example, user-3 which has 3 dB mean SNR achieves too little in the 1-cluster model and it is the 10^{th} in 12; it gets more in 2-cluster model and it is the 4^{th} in 6; it still gets more in 3-cluster model and it is the 2^{nd} in 4; it gets its best in 4-cluster model and it is the best one in the cluster; it decreases in 6-cluster model because it is the 2^{nd} in the two-user cluster but it is less than it achieves in 3-cluster cell despite it is the 2^{nd} too. The reason is the cluster size in 3-cluster cell is larger than the size in 6-cluster cell. Thus any user in the same rank in a larger cluster size benefits more. For another example, the user-12 is all times best user but its capacity always reduces despite it is always the best user. The reduction can be explained that same rank with the decreasing cluster size provides less capacity.

Figure 5.19 shows the basic reason of the capacity changes in the examples given. Because the channel access probabilities change, the capacities achieved by users also change. Any two users with the same rank in the same cluster size have identical channel access probabilities. However, this identical situation is not projected identically to the capacity plots (Figure 5.18) due to having different mean SNR values. Thus, the user with higher mean SNR has larger capacity than the one with a lower mean SNR for this identical channel access probabilities situation.

Like Figure 5.18 and 5.19, the share of the total capacities among the clusters and cluster channel access probabilities, which are the sum of all user probabilities in the same clusters, are showed on Figure 5.20 and 5.21 respectively.

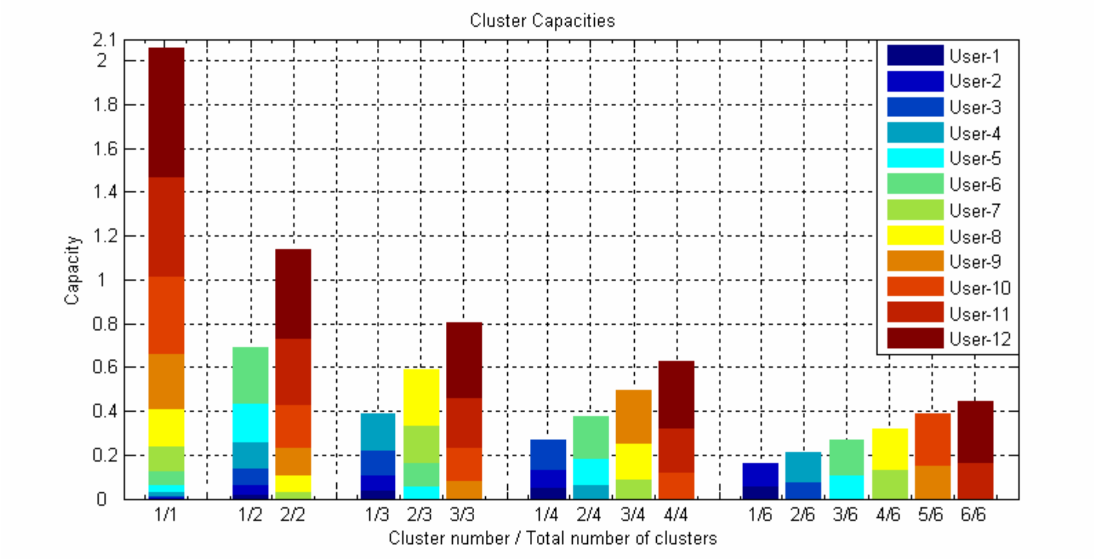

Figure 5.20 Cluster capacities

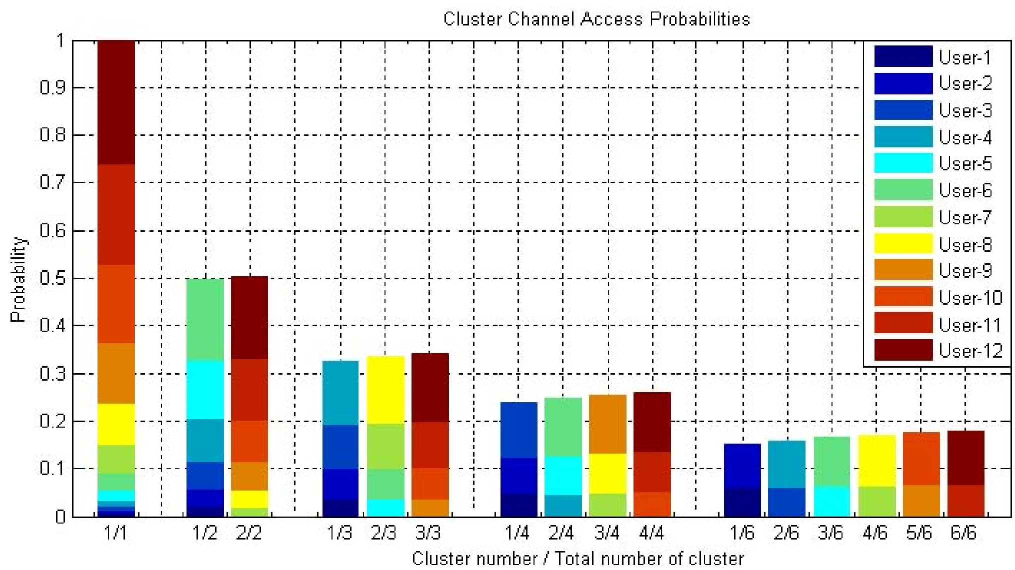

Figure 5.21 Cluster channel access probabilities

In Figure 5.20, the total cluster capacities are displayed cumulatively. In this figure, the cluster, which includes users which have large mean SNR values, has much more total capacity. However, the difference between the total cluster capacities in the same cell decreases with increasing number of clusters in the cell. That is, the difference between the total cluster capacities in 2-cluster cell is larger than the difference in 3-cluster cell.

It is stated that the user channel access probability affects the user capacity. The capacity differences exist due to the difference between mean SNR values. Thus, the user channel access probabilities are nearly identical for the same rank in the same cell model in Figure 5.21 but the capacities are different in Figure 5.20. Figure 5.21 also shows that the total cluster capacities are nearly same in the same cluster cell model. For example for 2-cluster cell, the cluster channel access probabilities are equal to 0.5, ≈ 0.33 for 3-cluster cell, ≈ 0.25 for 4-cluster cell. This is because the one user maximum SNR capacity is normalized with the one user MS capacity in the optimization problem. The found thresholds at the end of the optimization bring us the fairness of the equalized cluster channel access probabilities.

5.7 THE COMPARISON OF CLUSTERED USER SCHEDULING SCHEMES

So far, we have compared the heterogeneous cell model with different rules of choosing a user (random cluster selection vs. random user selection), fairness issues and the effects of increasing number of users to aggregate and cluster capacities, capacity and channel access probabilities of homogeneous and heterogeneous cell models with identical and non-identical users. Now the methods, maximum SNR (MS), clustered quantized weighted SNR (CQWS), clustered quantized maximum SNR with maximizing aggregate capacity (CQMS-max-agg), clustered quantized maximum SNR with scheduling outage probability (CQMS-P_o), round robin (RR) for heterogeneous cell model compared with clustered quantized maximum SNR (CQMS) method. Figure 5.22 and 5.23 show that the aggregate capacity curves for increasing number of users in high and low mean SNR clusters respectively. Single threshold is used for quantized methods. All users in the same cluster have the same mean SNR.

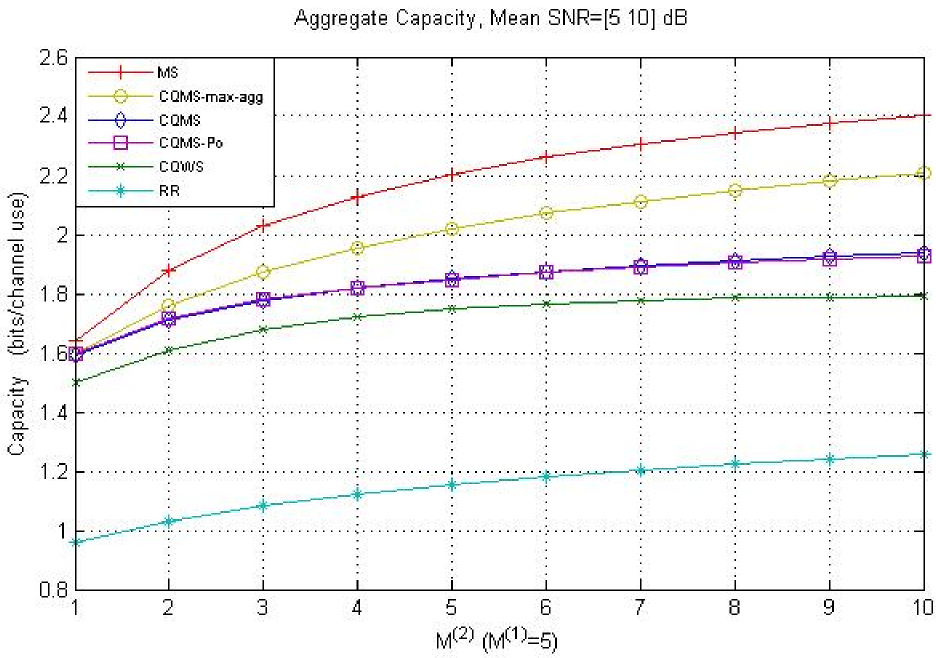

Figure 5.22 A comparison of aggregate capacities of different methods against number of users in high SNR cluster.

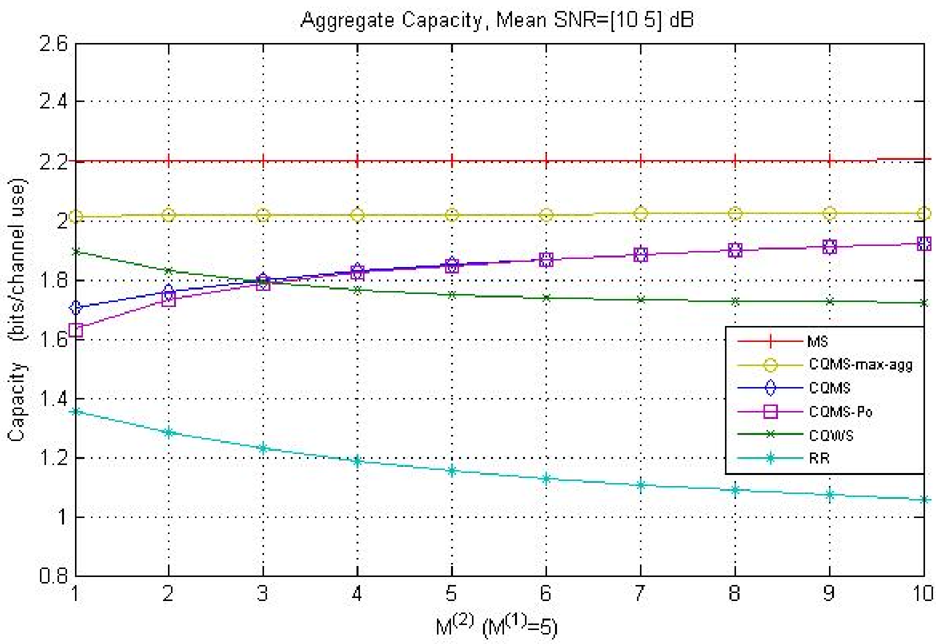

Figure 5.23 A comparison of aggregate capacities of different methods against number of users in low SNR cluster

MS method gives the highest aggregate capacity for both cases intuitively due to the selection of the best user all times by the help of the full feedbacks coming from the users. The rise on the aggregate capacity curve is much more significant in Figure 5.22 as compared with Figure 5.23 since the contribution of increasing number of high mean SNR users is more than that of increasing number of low mean SNR users.

CQMS-max-agg method gives the second maximum aggregate capacity curves for both figures. The optimization is handled over the aggregate capacity in CQMS-max-agg method. Thus it is the version of MS method with quantization feedback. In other words, it is possible to say that CQMS-max-agg is the quantized equivalent of MS. Therefore, CQMS-max-agg gives the aggregate capacity defining the upper bound for the quantized feedback scheme using multiuser diversity as the capacity achieved by MS which is the upper limit of full feedback schemes using multiuser diversity.

A remarkable point is the CQMS and CQMS-P_o aggregate capacity curves are close to each other. Here the best scheduling outage probability giving the maximum capacity is investigated and the optimum scheduling outage probability (P_o) is chosen to plot the capacity curves.

Aggregate capacity curves calculated by CQWS method do not have the same characteristics for both figures. The thresholds are equal to mean SNR values of the users. Since there is no optimization, thresholds do not have any optimality.

The worst aggregate capacity curves are obtained by RR method which has the simplest computational complexity. Addition of a high mean SNR user contributes the aggregate capacity while addition of a low mean SNR user does not.

Figures 5.24 - 5.43 show the aggregate and cluster capacities together in capacity graphs and user channel access probabilities in probability graphs.

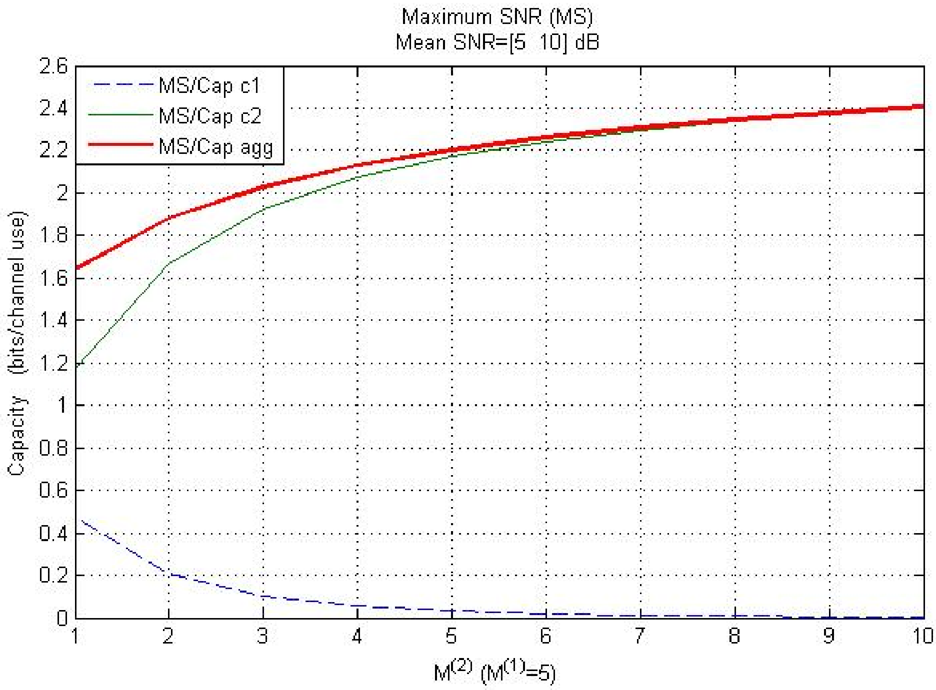

Figure 5.24 The aggregate and cluster capacities for MS method against the number of users in high SNR cluster

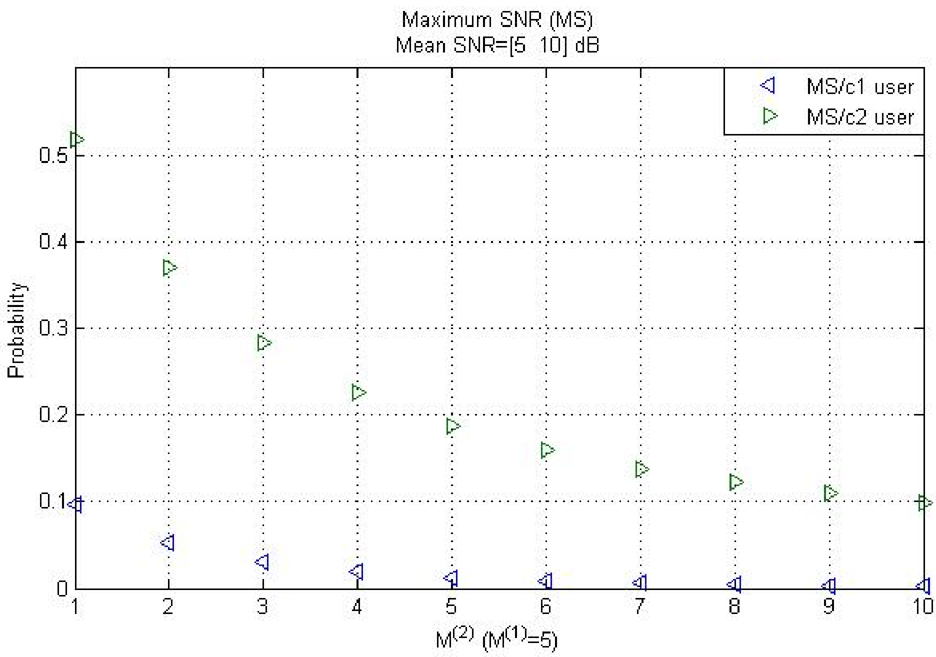

Figure 5.25 The user channel access probabilities for MS method against the number of users in high SNR cluster

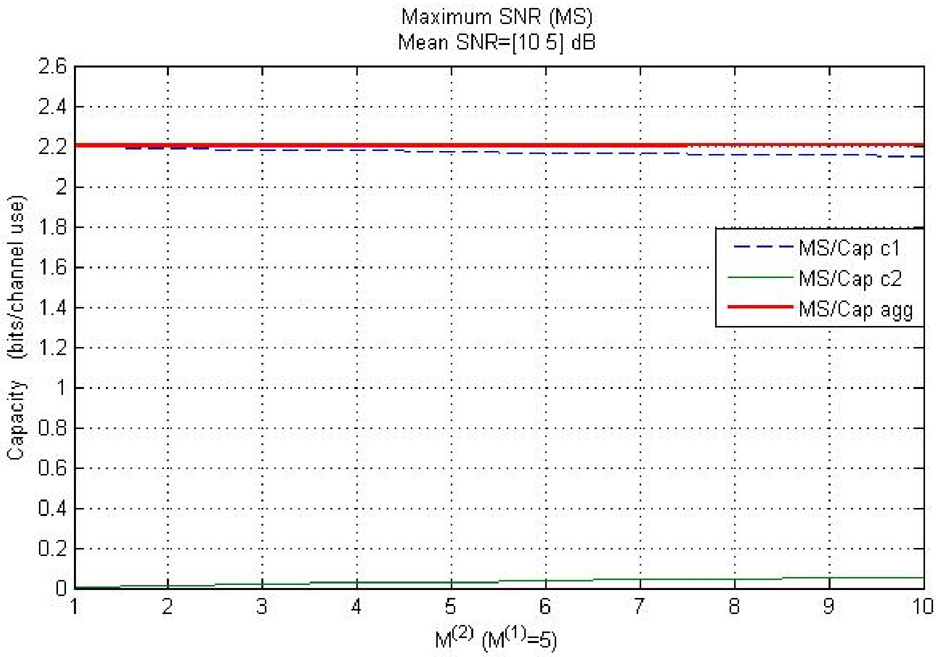

Figure 5.26 The aggregate and cluster capacities for MS method against the number of users in low SNR cluster

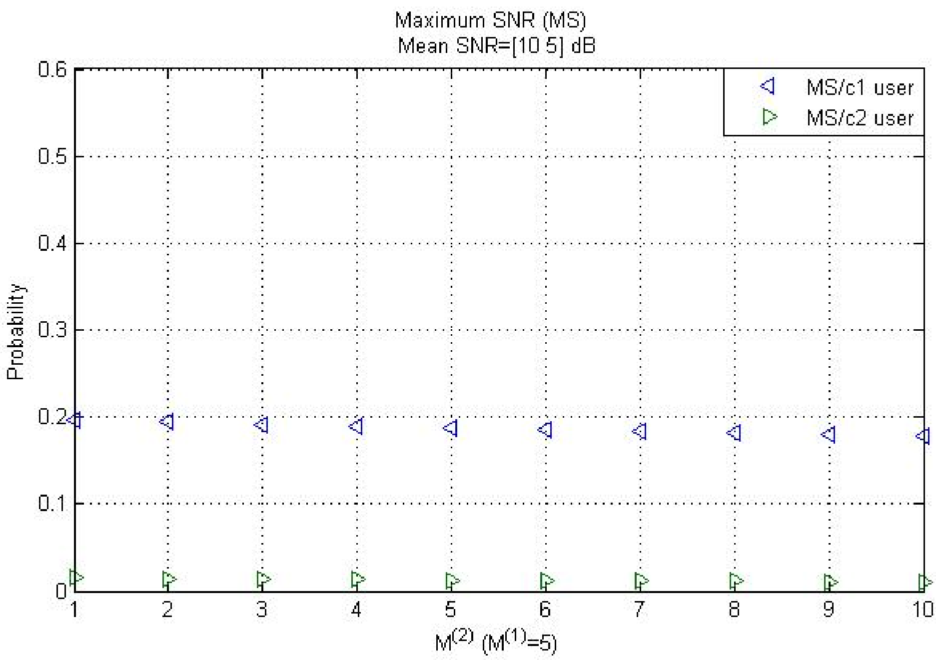

Figure 5.27 The user channel access probabilities for MS method against the number of users in low SNR cluster

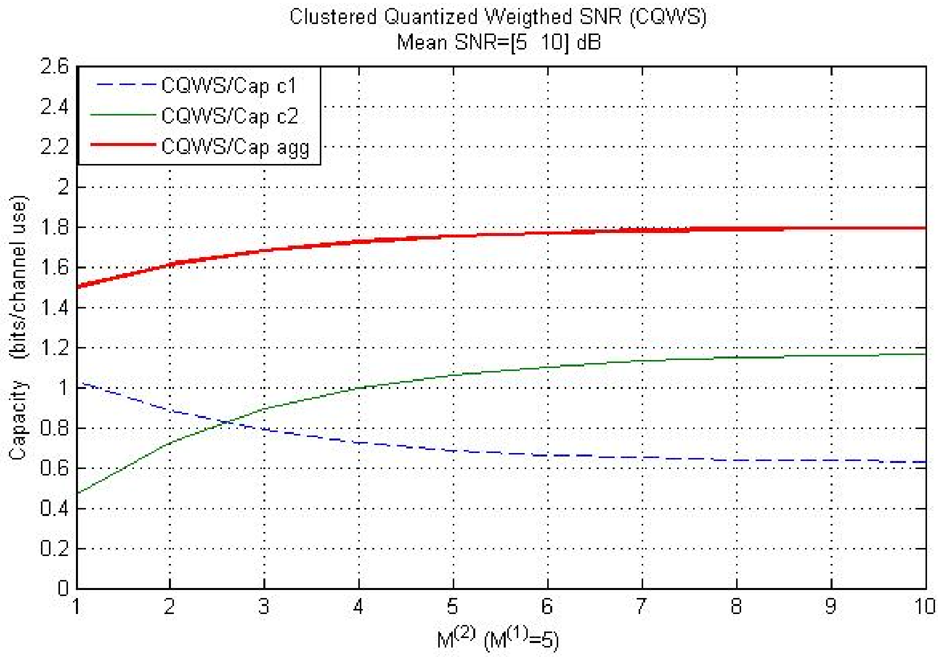

Figure 5.28 The aggregate and cluster capacities for CQWS method against the number of users in high SNR cluster

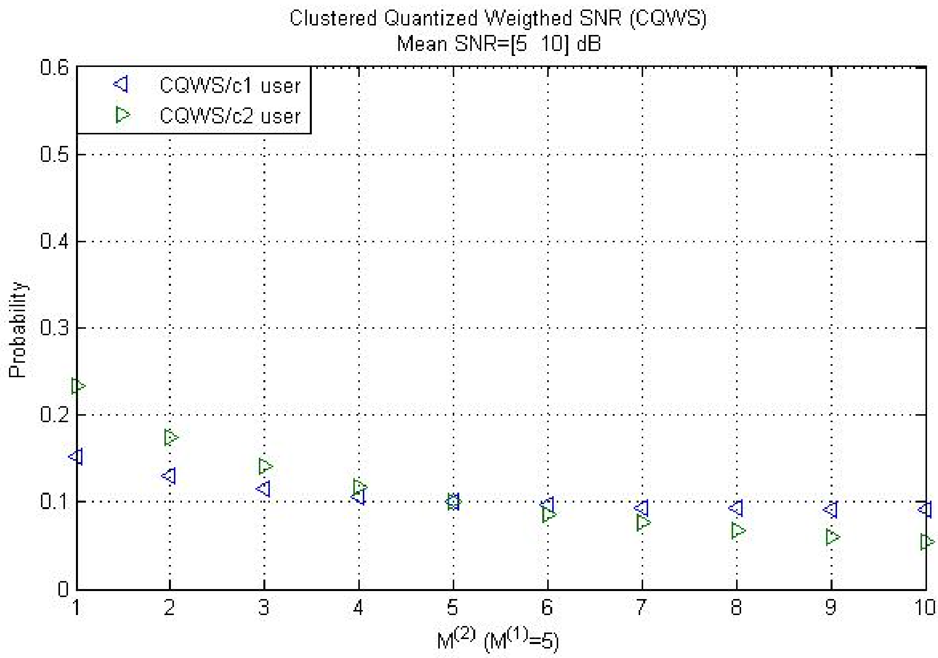

Figure 5.29 The user channel access probabilities for CQWS method against the number of users in high SNR cluster

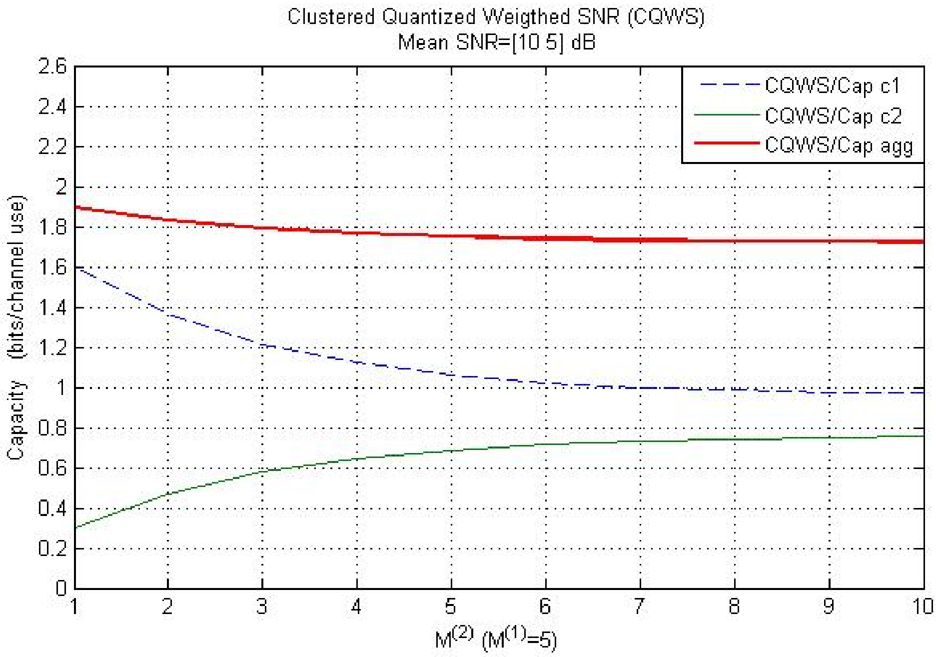

Figure 5.30 The aggregate and cluster capacities for CQWS method against the number of users in low SNR cluster

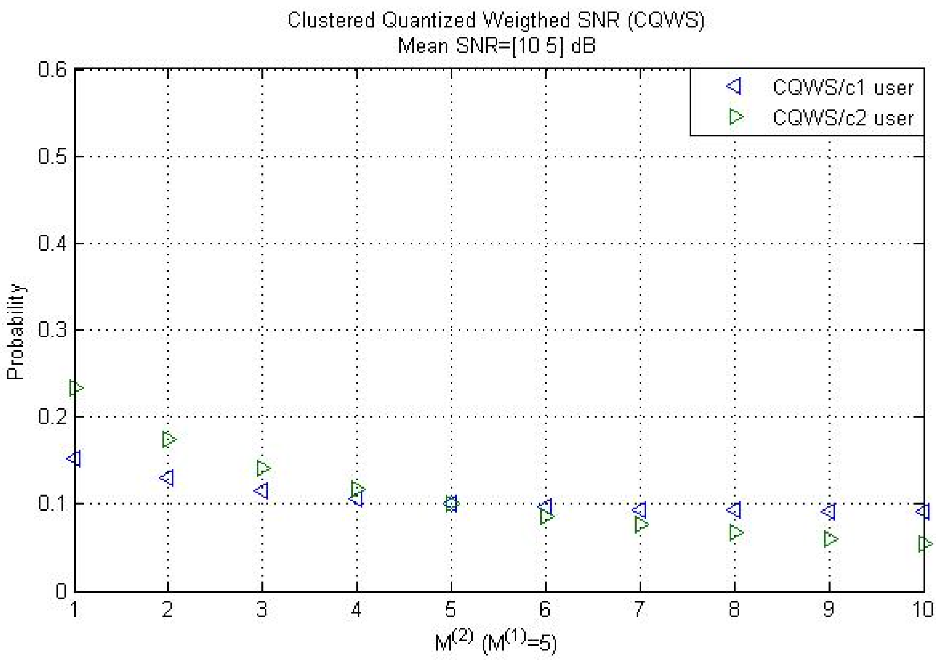

Figure 5.31 The user channel access probabilities for CQWS method against the number of users in low SNR cluster

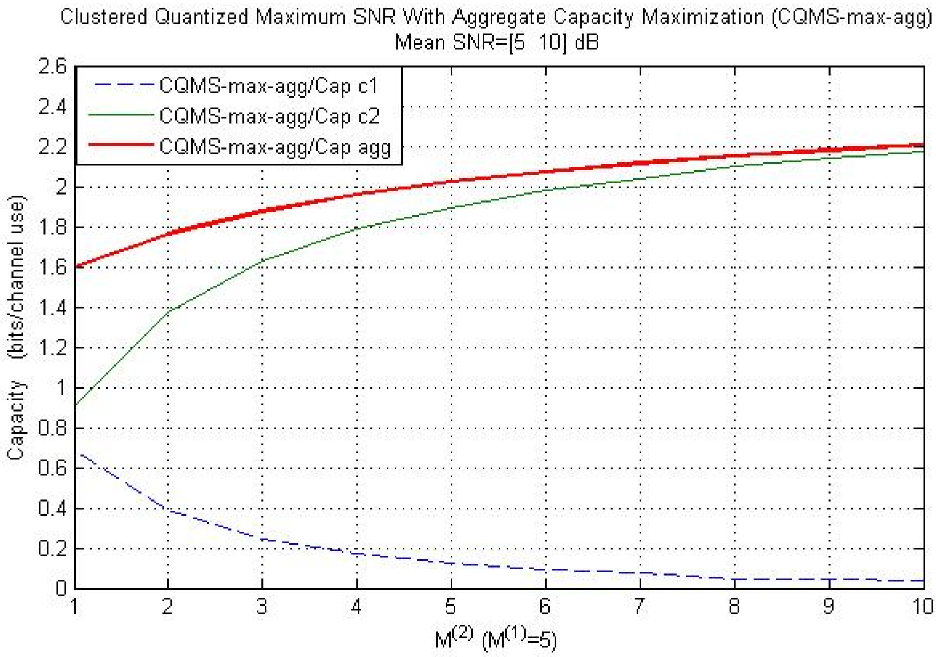

Figure 5.32 The aggregate and cluster capacities for CQMS-max-agg method against the number of users in high SNR cluster

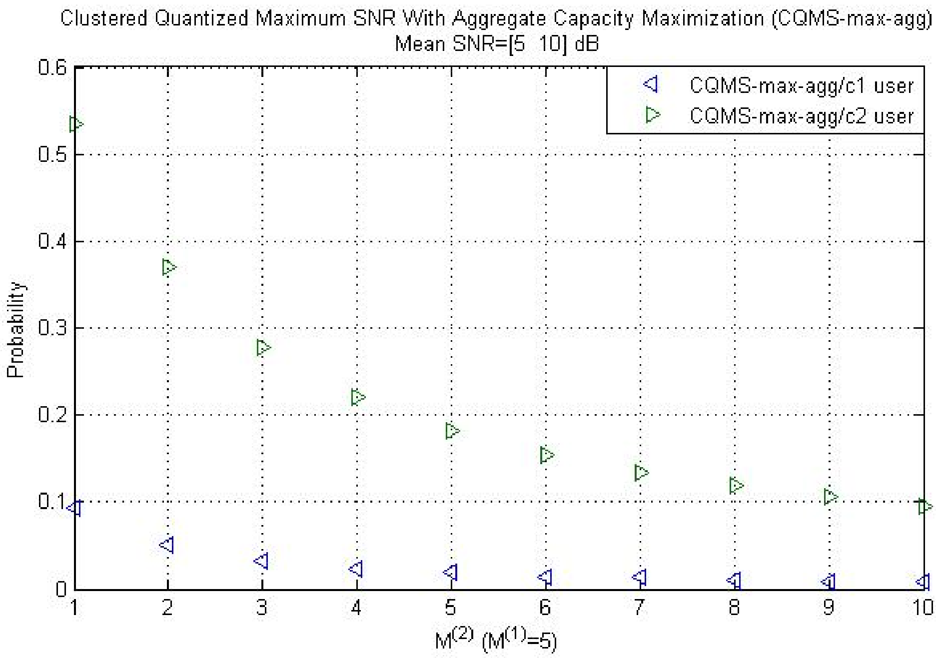

Figure 5.33 The user channel access probabilities for CQMS-max-agg method against the number of users in high SNR cluster

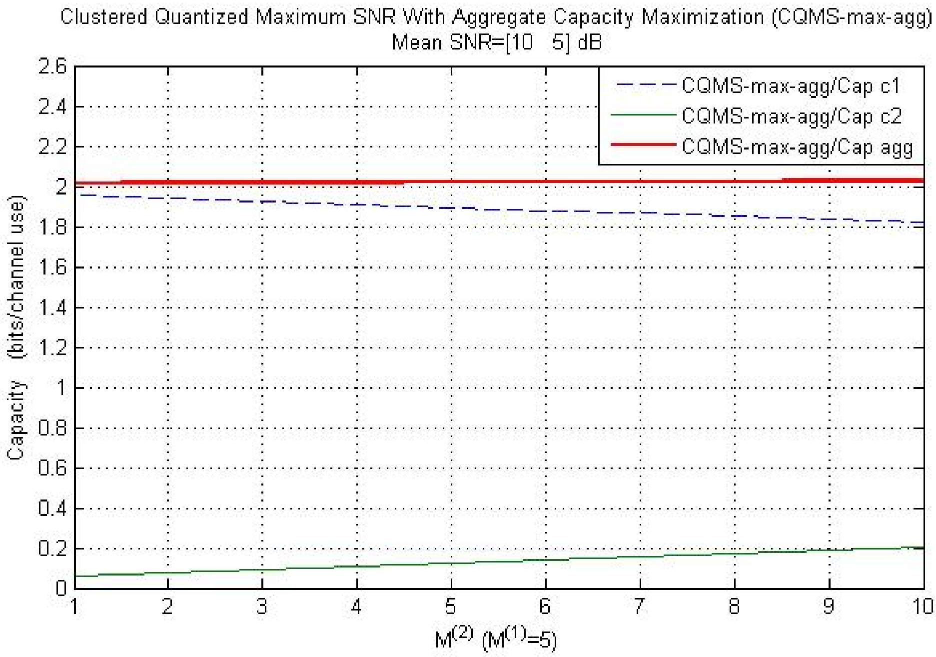

Figure 5.34 The aggregate and cluster capacities for CQMS-max-agg method against the number of users in low SNR cluster

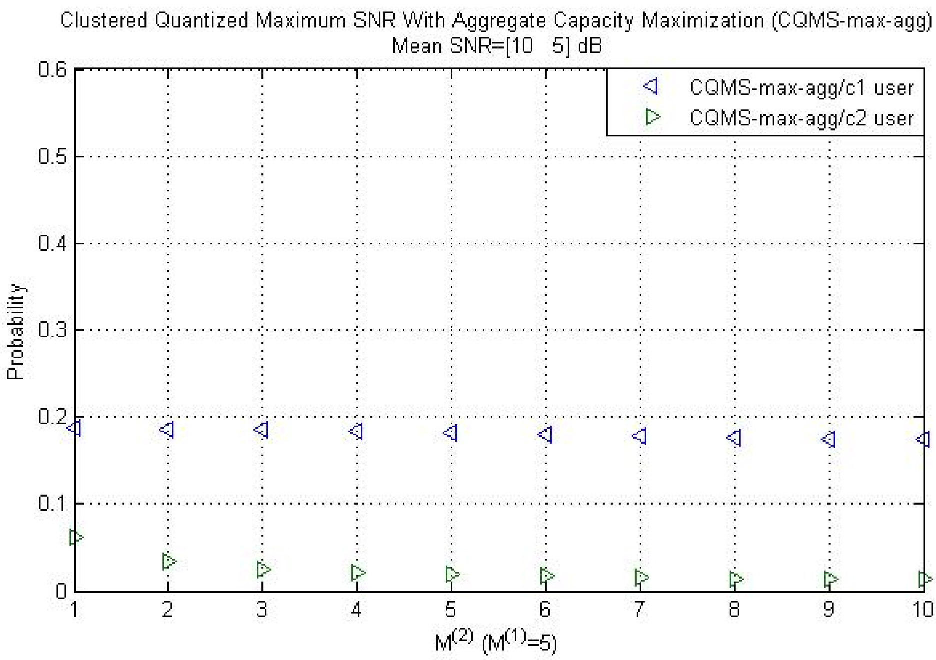

Figure 5.35 The user channel access probabilities for CQMS-max-agg method against the number of users in low SNR cluster

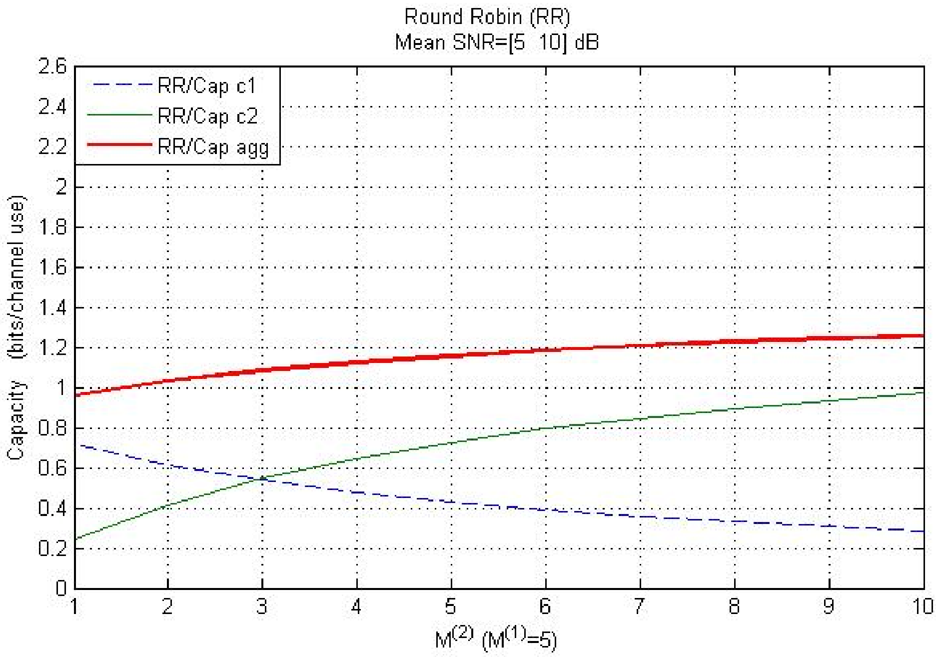

Figure 5.36 The aggregate and cluster capacities for RR method against the number of users in high SNR cluster

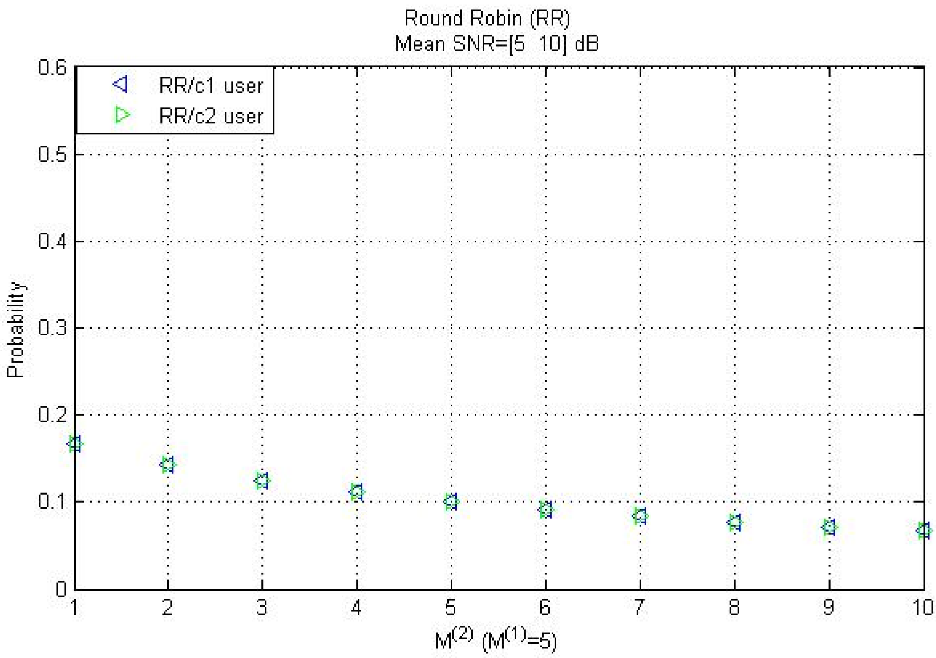

Figure 5.37 The user channel access probabilities for RR method against the number of users in high SNR cluster

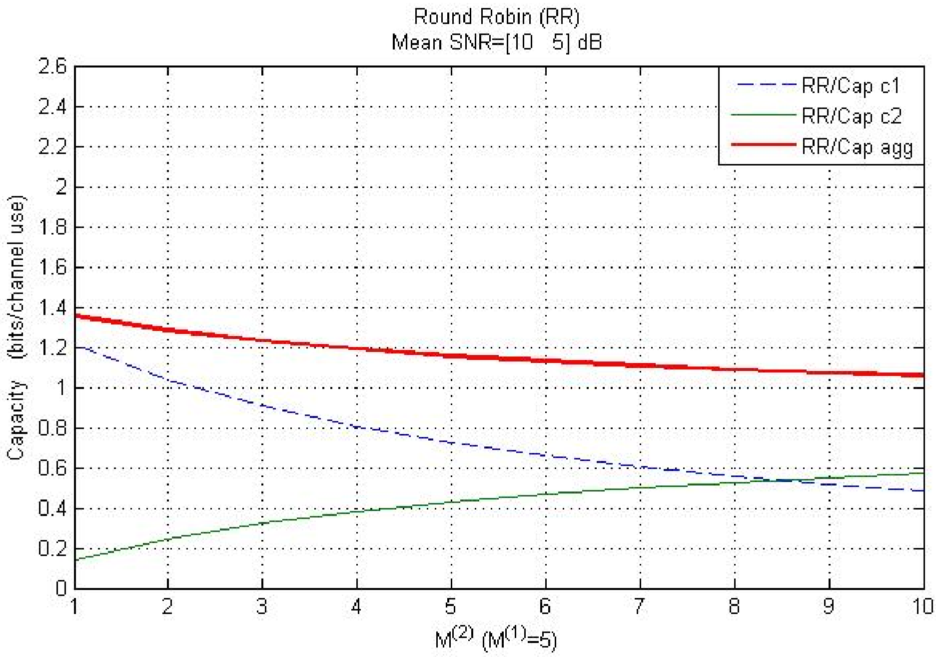

Figure 5.38 The aggregate and cluster capacities for RR method against the number of users in low SNR cluster

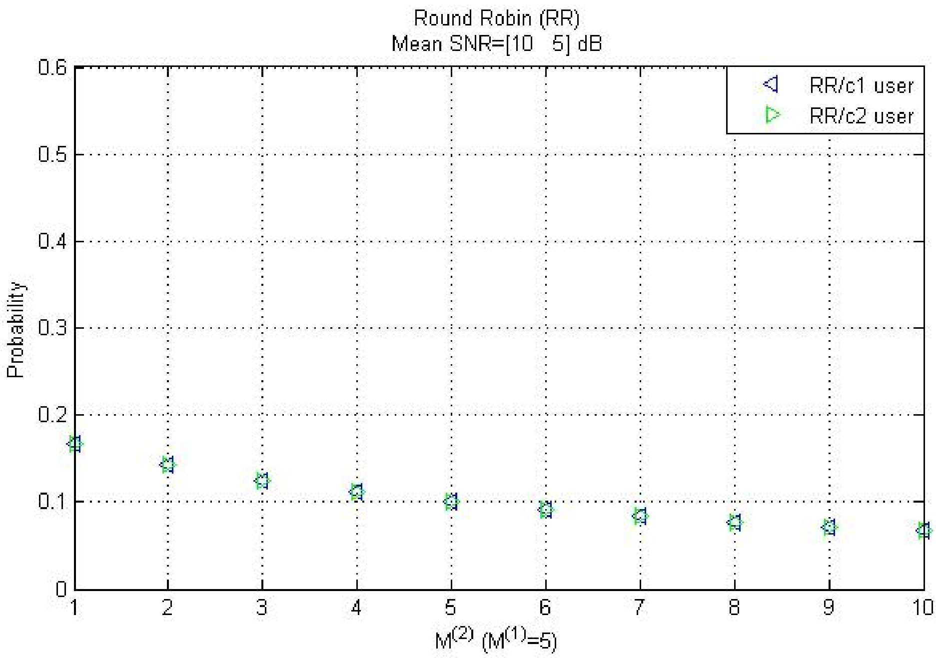

Figure 5.39 The user channel access probabilities for RR method against the number of users in low SNR cluster

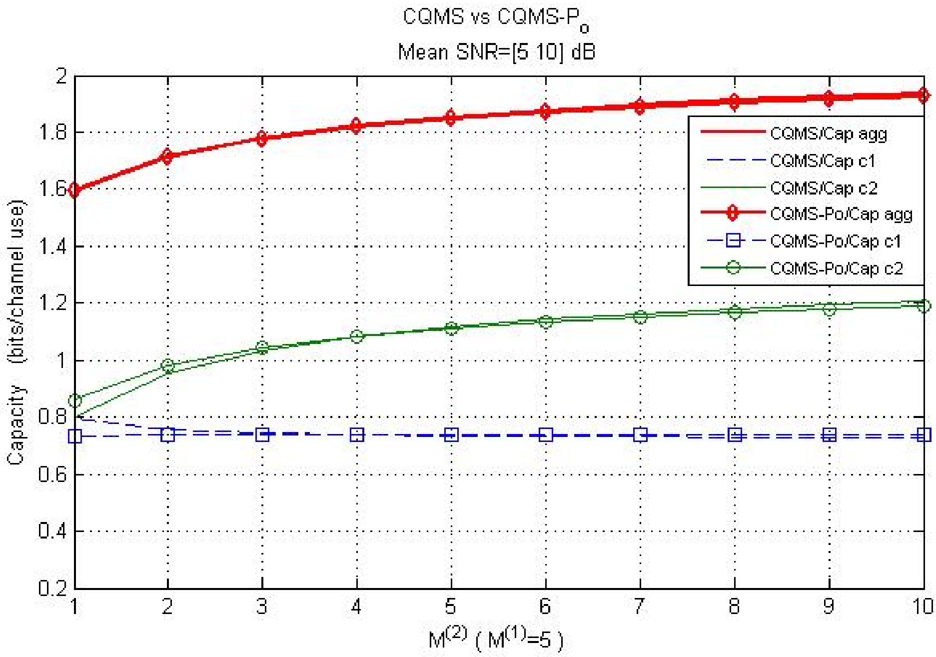

Figure 5.40 The aggregate and cluster capacities for CQMS and CQMS-P_o methods against the number of users in high SNR cluster

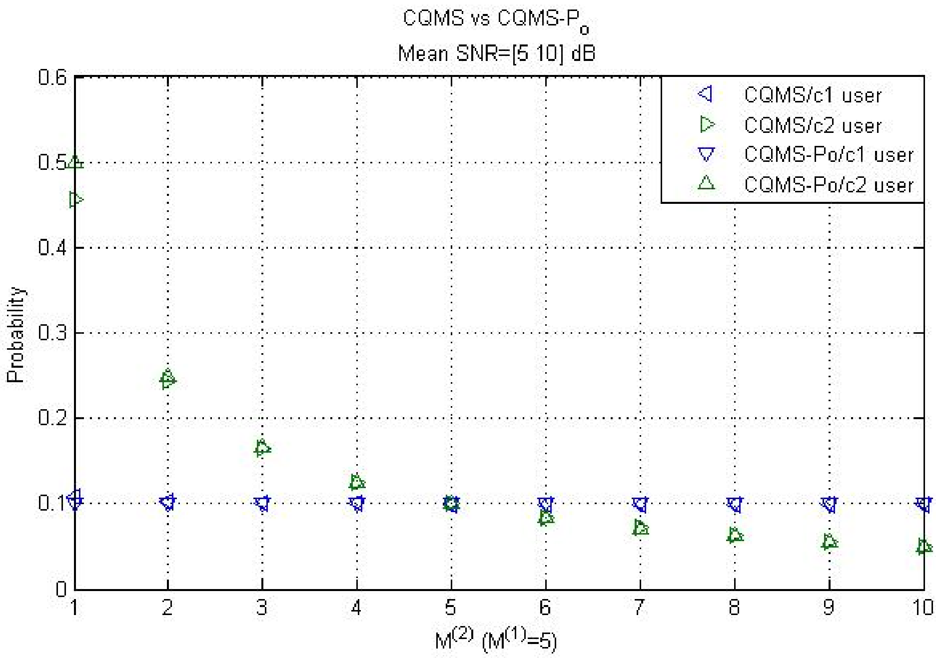

Figure 5.41 The user channel access probabilities for CQMS and CQMS-P_o methods against the number of users in high SNR cluster

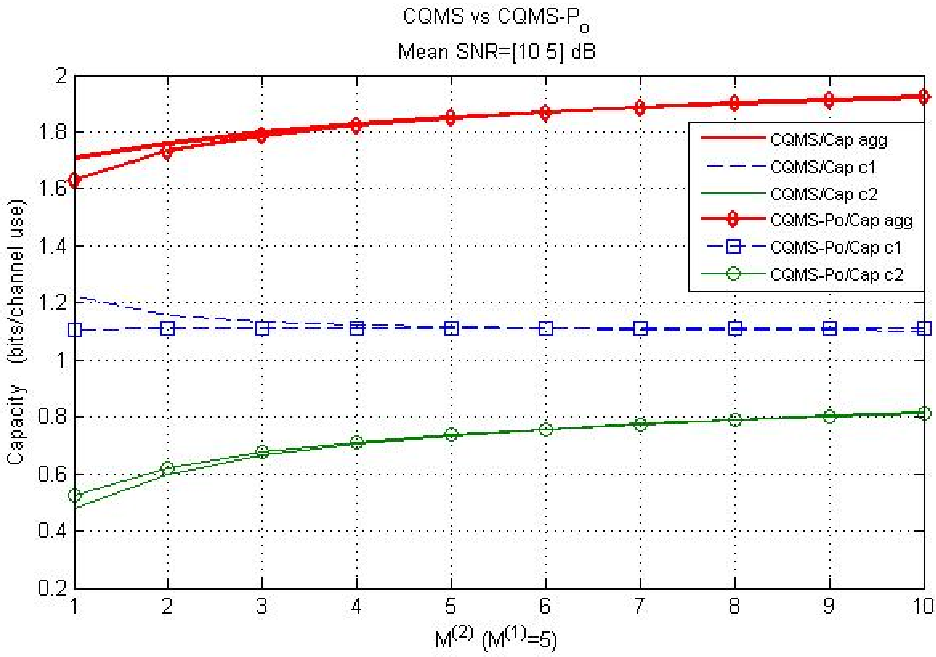

Figure 5.42 The aggregate and cluster capacities for CQMS and CQMS-P_o methods against the number of users in low SNR cluster

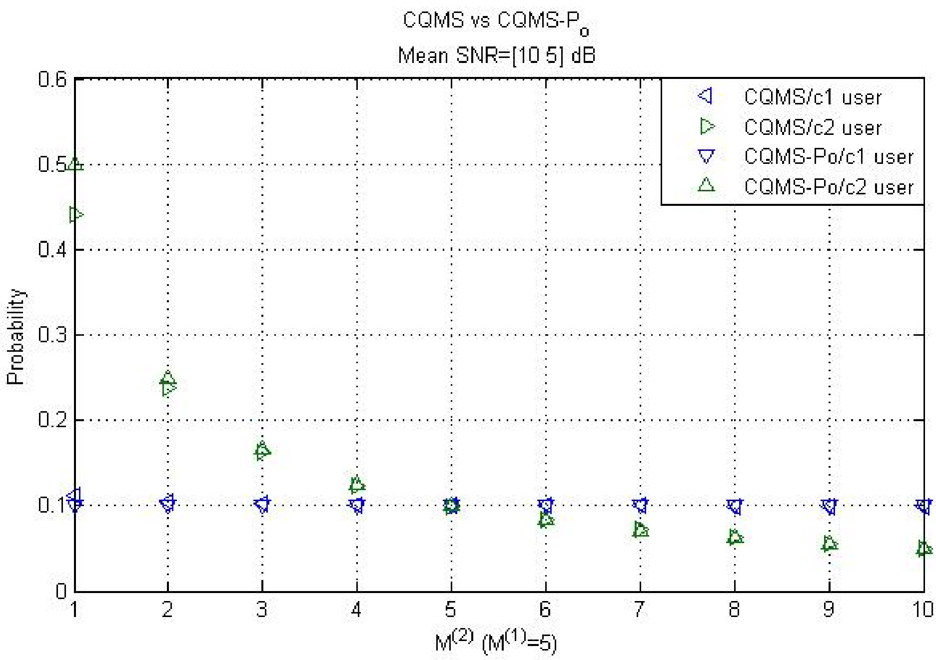

Figure 5.43 The user channel access probabilities for CQMS and CQMS-P_o methods against the number of users in low SNR cluster

Figure 5.31 (a) and 5.32 (a) show not only CQMS and CQMS-P_o aggregate capacities are nearly same but also their cluster capacities are close to each other. Channel access probabilities are showed on Figure 5.31 (b) and 5.32 (b). It is observed that no matter size of which cluster is increasing, the user channel access probabilities stay same. As a conclusion, the user channel access probabilities are not affected by the mean SNR of the cluster.

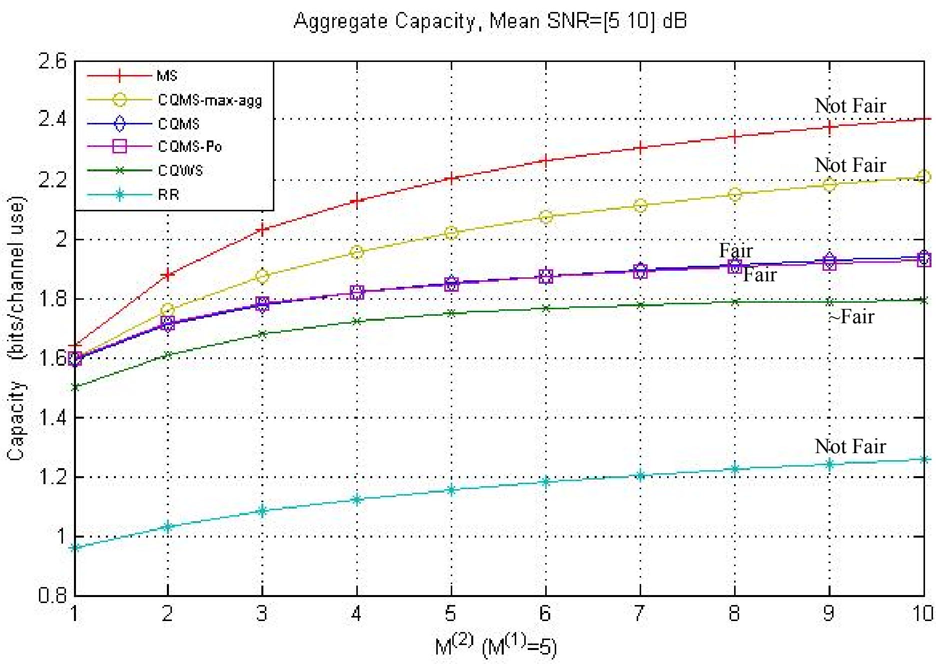

Figure 5.44 A comparison of aggregate capacities of different methods against number of users in high SNR cluster.

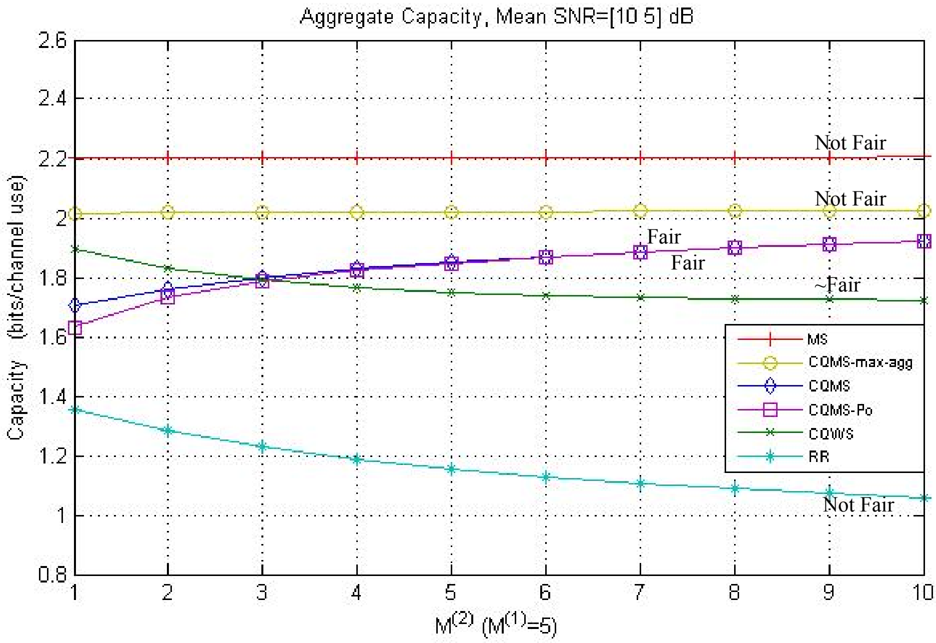

Figure 5.45 A comparison of aggregate capacities of different methods against number of users in low SNR cluster

Figure 5.44 and 5.45 have fairness tags in addition to the same plots in Figure 5.22 and 5.23. Here, the fairness criterion is that the addition of a user to any cluster does not affect the other clusters' capacity. CQMS and CQMS-P_o fully, CQWS slightly provide fairness for users with non-identically distributed channels. CQMS and CQMS-P_o perform equivalently showed in Figure 5.40, 5.41, 5.42 and 5.43. We were not able to analytically justify the close relation between CQMS and CQMS-P_o. On the other hand, in CQWS, enlarging cluster decreases the capacity achieved by the other cluster but at some user value it gets stable (in the observation of given users numbers). Besides, MS, CQMS-max-agg and RR are all unfair methods.

CHAPTER 6

CONCLUSION

Opportunistic user scheduling is reserving the channel access right to the user which has the best channel quality among all users. It works well for the dynamic channels where the more randomness channel has, the better the scheduling performs. The randomness of user channels makes the instantaneous SNR values fluctuate in time. In the long term average, the sum (aggregate) capacity which is defined as the summation of capacity achieved by each user is maximized. The system is fair for identically distributed users. Any user is chosen when its instantaneous SNR is at its peak. If channels vary slowly in time, the opportunistic user scheduling does not work fairly for all users. Therefore, the base station must have some modifications as discussed in [7] to improve the channel fluctuations.

Although opportunistic scheduling has many advantageous (maximizing capacity with fairly sharing time intervals to each user), there are some technical problems to be resolved. One of them is the feedback load which would be a problem with increasing number of users. In this thesis, methods for feedback quantization are described to reduce the feedback load. Since the base station can not fully know about the user channels with quantized feedbacks, some capacity degradation occurs. Another issue is that the opportunistic scheduling does not equally share the resources for non-identical set of users. Therefore, clustering users according to their channel quality along with the feedback quantization can provide more fairness and feedback load reduction.

In this thesis, first the impact of utilizing multiuser diversity is investigated. Maximum SNR (MS) method, which uses multiuser diversity, is compared with round robin (RR) scheduling which has low computational complexity. MS is more beneficial in comparison with RR scheduling for the capacity aspect. The capacity increase is much higher with the increasing mean SNR; since the largest instantaneous SNR value gets higher. It has been showed that the multiuser diversity is fair when the users remain identically distributed for a sufficiently long period of time.

The effects of feedback quantization for identical users (homogeneous cell) are studied. Feedback quantization is done by maximizing the capacity with an optimum selection of thresholds. The capacity equation is derived for finite feedback case and it is used in the optimization process. At the end of optimization, it gives optimum thresholds which achieve the maximum value for capacity. Besides evaluating capacity by analytically, some Monte Carlo simulations are also done to verify the analytical results.

The capacity achieved by multi level feedback quantization approaches the capacity achieved by the full feedback system. Even a single threshold can provide more than 92% of MS capacity, the quantized feedback capacity approached MS capacity with increasing number of quantization levels.

The effect of the number of users on the optimal threshold values are investigated for the homogenous cell model. When the number of users increases, the threshold values also increases to maximize the capacity. Since the addition of a user reduces the probability of any user being above the threshold, the threshold value is increased to have fewer users above it enough to get maximized capacity.

Opportunistic scheduling is fair for a set of identical users. However, it is not so fair for non-identical users due to selection of the best user. User with low mean SNR is probably not chosen as frequently as high mean SNR users. Hence, clustered structure is proposed to provide fairness to the disadvantaged users. In Chapter 4, two rules are introduced: random user selection and random cluster selection. However, it has been shown by simulations that both of them perform similarly.

Since random cluster selection rule has less computational complexity, it is preferred for further study in this thesis.

Clustered scheduling can achieve lower aggregate capacity than achieved by the MS method. However, users with low mean SNR can access channel more frequently. Some comparisons are done for a 2-cluster cell. The effect of increasing number of users in the high and low mean SNR clusters is an increase in the capacity belonging to the growing cluster. Addition of any user to a cluster punishes the users in the other cluster. To prevent this situation, an optimization problem is formulated. Instead of finding the thresholds which maximizes the quantized mean SNR capacity, finding the thresholds to minimize the maximum of the loss factor is proposed. The loss factor is the ratio of quantized capacity over the MS capacity. With this change, it is observed that adding a user to a cluster does affect very little the capacity of the other cluster.

A comparison of the homogeneous and clustered heterogeneous cell models is done. The cell includes two groups of users with high and low SNR values. Every user in a group has the same mean SNR. The users in these two groups are scheduled according to the homogeneous and clustered heterogeneous cell models. Homogeneous cell model treats all users as if they have the same mean SNR whereas the clustered heterogeneous cell model separates the different users and schedules them. The homogeneous cell model produces the capacity graphs for increasing number of users similar to ones produced by MS method. That means scheduling users according to the homogeneous cell model always seeks the user with the best SNR. Therefore the users belonging to high SNR group access the channel more. This result is also observed on the channel access probabilities. The users with high SNR access the channel much more frequently than the other users with low SNR as when users are treated homogeneously. However, the channel access probabilities belonging to groups are close to each other (≈ 0.5 for two groups).

In Chapter 4, a scheduling scheme is introduced for non-identical users. However, its computational complexity grows exponentially with a rise in the number of users. Instead of this complex scheme, non-identical users are scheduled according to the clustered heterogeneous cell which has reasonably lower complexity. In a computer simulation, a cell with users having different mean SNR values are taken and divided into clusters with the same sizes. With this grouping, users whose mean SNR values are close to each other are put into the same cluster. The user with largest mean SNR has the 1^{st} rank in the cluster. It is observed that the aggregate capacity reduces when more clusters are created. The highest ranked users achieve more capacity in each cluster. Therefore dividing users into more clusters make the lower ranked users achieve more capacity because of increasing user channel access probabilities. The users in different clusters having equal channel access probabilities (equal ranked users) may have different capacities. The sole reason of this is the mean SNR differences between them. Another fairness point is dividing users into clusters make the channel access probabilities of clusters, the summation of user channel access probabilities, close to each other. That is, the probability of choosing any cluster by the base station depends on the number of clusters and its density is almost uniform. The user channel access probabilities become almost equal as the number of clusters gets higher. Consequently, clustering of users offers fairness on the user/cluster channel access probabilities.

Some other user scheduling methods are compared with the proposed clustered user scheduling (CQMS) scheme. These methods: clustered quantized weighted SNR (CQWS), full feedback (MS), round robin (RR), clustered quantized mean SNR with scheduling outage probability (CQMS-P_o), clustered quantized mean SNR with aggregate capacity maximization (CQMS-max-agg). As expected, MS gives the highest aggregate capacity for increasing user number because of full feedback. The second highest aggregate capacity is achieved by CQMS-max-agg method because it finds the optimum thresholds which maximize the aggregate capacity. This is the quantized version of MS method. CQWS method is a method with low computational complexity. This method uses mean SNR values of clusters as thresholds. However, these thresholds are not optimized in any sense. Namely, aggregate capacity decreases with increasing number of users in low SNR cluster by using CQWS method. The decrease in capacity also occurs for RR method as in CQWS method. However, RR is the simplest way to schedule users. An important result of this comparison is that CQMS and CQMS-P_o methods perform almost identically. CQMS finds thresholds by optimizing them over the SNR range to maximize the capacity; whereas CQMS-P_o finds the thresholds by optimizing them over scheduling outage probability.

Future Work

In this study, it is assumed that cells are isolated against interference coming from the neighboring cells. However, in practice, there may be some interference and the capacity equations should be modified for this case.

It is also assumed that the fading channel has the characteristics of Rayleigh fading channel which has no line of sight (LOS) path and many scatterers, absorbers, obstacles etc. in the medium. The capacity derivations can be given for a channel with a LOS path. It is obvious that having a LOS path reduces the channels randomness. Although the key parameter is the dynamical range of channels attained by randomness in opportunistic communications, the way of increasing the dynamic range of channel should be investigated and its performance can be compared.

In this study, the feedback collection time is neglected. It is assumed that that delay in feedback does not affect the system capacity. However, there is a time interval in which the base station polls the users for their channel information. For a more precise capacity calculation, this time interval (polling time) can be included in the calculations. Improving the polling time is a better way to get higher capacity values.

Scheduling is done over the time domain (TDMA). However, opportunism in communication could be broadened to the frequency domain. Feedback reduction is applicable on a system which uses orthogonal frequency division multiple access (OFDMA). The adaptation of the study presented here to multiple carriers is possible research direction.

An immediate future work of this thesis would be an analytical study on the relationship between CQMS and CQMS-P_o methods which perform almost equivalently. In this thesis, it is observed that CQMS-P_o can also find the optimum thresholds of the CQMS. An explanation of this behavior is not evident to us.

REFERENCES

[1] J. Proakis, *Digital Communications*, 4th ed. Boston: McGraw-Hill, 2001.

[2] A. Goldsmith, *Wireless Communications*, 1st ed. New York: Cambridge University Press, 2005.

[3] T. M. Cover, J. A. Thomas, *Elements of Information Theory*, 1st ed. New York: Wiley, 1991.

[4] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal,* vol. 27, pp. 379-423, 623-656, July, October, 1948.

[5] R. Knopp, P. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. IEEE Int. Computer Conf. (ICC'95),* Seattle, WA, vol. 1, pp. 331-335, June 1995.

[6] D. N. C. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," in *Proc. IEEE Int. Symp. Information Theory,* Ulm. Germany, pp. 27, 1997.

[7] P. Viswanath, D. N. C. Tse, R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Information Theory,* pp. 1277-1294, June 2002.

[8] D. Gesbert, M. Alouini, "How much feedback is multiuser diversity really worth," in Proc. Int. Conference on Communications, Paris, France, June 2004.

[9] M. Johansson, "Benefits of multiuser diversity with limited feedback," in *Proc. of IEEE Workshop on Signal Processing Advances in Wireless Comm.*, pp. 155-159, June 2003.

[10] Y.Al-Harti, A. Tewfik, M. S. Alouini, "Multiuser diversity with quantized feedback," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 330-337 January 2007.

[11] A. P. T. Lau, F. R. Kchinschang, "Feedback quantization strategies for multiuser diversity systems," *IEEE Transactions on Information Theory*, vol. 53, pp. 1386-1400, April 2007.

[12] F. Floren, O. Edfors, B. A. Molin, "The effect of feedback quantization on the throughput of a multiuser diversity scheme," in *Proc. IEEE GLOBECOM*, vol. 1, pp. 497-501, December 2003.

[13] F. Floren, O. Edfors, B. A. Molin, "Throughput analysis of three multiuser diversity scheme," in *Proc. IEEE Vechicular Technology Conf.*, VTC 2003 Orlando, FL, USA, pp.2686-2690, Spring 2003.

[14] D. Tse, P. Viswanath, *Fundamentals of Wireless Communication,* 1st ed. New York: Cambridge University Press, 2005.

[15] T. S. Rappaport, *Wireless Communications: Principles and Practice,* 2nd ed. N.J.: Prentice Hall, 2002.

[16] P. P. Bergmans, T. M. Cover, "Cooperative broadcasting", *IEEE Transactions on Information Theory*, vol. IT-20, pp. 317-324, May 1974.

[17] L. Li, A. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels: Part I, ergodic capacity," *IEEE Transactions on Information Theory*, vol. 47, pp. 1083-1102, March 2001.

[18] L. Li, A. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels: Part II, outage capacity," *IEEE Transactions on Information Theory*, vol. 47, pp. 1103-1127, March 2001.

[19] L. H. Ozarow, S. Shamai, A. D. Wyner, "Information Theoretic Considerations for Cellular Mobile Radio," *IEEE Transactions on Vechicular Technology*, vol. 43, pp. 359-378, May 1994.

[20] L. Rade and B. Westergren, *Mathematics Handbook for Science and Engineering,* 4th ed. Springer 2001.

[21] R. G. Gallager, *Information Theory and Reliable Communicatio,* New York: Wiley, 1968.

[22] H. Bölcskei, D. Gespert, C. B. Papadis, A. van der Veen, *Space-Time Wireless Systems: from array processing to MIMO communications,* Cambridge: Cambridge University Press, 2006.

[23] A. Papoulis, *Probability, Random Variables, and Stochastic Processes,* 2nd ed., New York: McGraw-Hill, 1984.

[24] H. K. Daas, *Advanced Engineering Mathematics*, S. India: Chand, 2007.

APPENDIX A

THE EXAMPLES

A.1 THE HETEROGENEOUS CELL MODEL EXAMPLE #1

Calculating the probability of selecting the reference user in the k^{th} quantization level for the heterogeneous cell model consisting of 3 users (M=3):

Let the reference user be the first one, then

$R = \{1,2,3\}$

$R_1 = R - \{1\} = \{2,3\}$

Since $|R_1| = 2$, there exist 0, 1 and 2-element subsets of R_1.

0) Number of 0-element subset of R_1 is $\binom{|R_1|}{0} = \binom{2}{0} = 1$.

$S_0^{(1)} = \varnothing\,;\, S_0^{(1)'} = R_1 - S_0^{(1)} = R_1 - \varnothing = R_1 = \{2,3\}$

1) Number of 1-element subset of R_1 is $\binom{|R_1|}{1} = \binom{2}{1} = 2$.

$S_1^{(1)} = \{2\}\,;\, S_1^{(1)'} = R_1 - S_1^{(1)} = \{3\}$

$S_1^{(2)} = \{3\}\,;\, S_1^{(2)'} = R_1 - S_1^{(2)} = \{2\}$

2) Number of 2-element subset of R_1 is $\binom{|R_1|}{2} = \binom{2}{2} = 1$.

$S_2^{(1)} = \{2,3\}\,;\, S_2^{(1)'} = R_1 - S_2^{(1)} = \varnothing$

Hence, the probability of scheduling the 1st user is

$$\Pr\left(A \mid \Gamma^{(1)} \in Q_k^{(1)}\right) = \sum_{m=0}^{2} \frac{1}{m+1} \left\{ \sum_{n=1}^{\binom{2}{m}} \left[\prod_{u \in S_m^{(n)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_m^{(n)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] \right\}$$

$$= \left\{ \sum_{n=1}^{\binom{2}{0}} \left[\prod_{u \in S_0^{(n)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_0^{(n)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] \right\} +$$

$$\frac{1}{2} \left\{ \sum_{n=1}^{\binom{2}{1}} \left[\prod_{u \in S_1^{(n)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_1^{(n)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] \right\} +$$

$$\frac{1}{3} \left\{ \sum_{n=1}^{\binom{2}{2}} \left[\prod_{u \in S_2^{(n)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_2^{(n)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] \right\}$$

$$= \left\{ \prod_{u \in S_0^{(1)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_0^{(1)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right\} +$$

$$\frac{1}{2} \left\{ \left[\prod_{u \in S_1^{(1)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_1^{(1)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] + \right.$$

$$\left. \left[\prod_{u \in S_1^{(2)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_1^{(2)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] \right\} +$$

$$\frac{1}{3} \left\{ \left[\prod_{u \in S_2^{(1)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_2^{(1)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] \right\}$$

$$= \left\{ F_{\Gamma^{(2)}}\left(q_k^{(2)}\right) F_{\Gamma^{(3)}}\left(q_k^{(3)}\right) \right\} +$$

$$\frac{1}{2} \left\{ \left[F_{\Gamma^{(2)}}\left(q_{k+1}^{(2)}\right) - F_{\Gamma^{(2)}}\left(q_k^{(2)}\right) \right] F_{\Gamma^{(3)}}\left(q_k^{(3)}\right) + \right.$$

$$\left. \left[F_{\Gamma^{(3)}}\left(q_{k+1}^{(3)}\right) - F_{\Gamma^{(3)}}\left(q_k^{(3)}\right) \right] F_{\Gamma^{(2)}}\left(q_k^{(2)}\right) \right\} +$$

$$\frac{1}{3} \left\{ \left[F_{\Gamma^{(2)}}\left(q_{k+1}^{(2)}\right) - F_{\Gamma^{(2)}}\left(q_k^{(2)}\right) \right] \left[F_{\Gamma^{(3)}}\left(q_{k+1}^{(3)}\right) - F_{\Gamma^{(3)}}\left(q_k^{(3)}\right) \right] \right\}$$ ■

A.2 THE HETEROGENEOUS CELL MODEL EXAMPLE #2

Calculating the probability of selecting the reference user in the k^{th} quantization level for the heterogeneous cell model consisting of 4 users (M=4):

Let the reference user be the first one, then

$R = \{1,2,3,4\}$

$R_1 = R - \{1\} = \{2,3,4\}$

Since $|R_1| = 3$, there exist 0, 1, 2 and 3-element subsets of R_1.

0) Number of 0-element subset of R_1 is $\binom{|R_1|}{0} = \binom{3}{0} = 1$.

   $S_0^{(1)} = \varnothing$; $S_0^{(1)'} = R_1 - S_0^{(1)} = R_1 - \varnothing = R_1 = \{2,3,4\}$

1) Number of 1-element subset of R_1 is $\binom{|R_1|}{1} = \binom{3}{1} = 3$.

   $S_1^{(1)} = \{2\}$; $S_1^{(1)'} = R_1 - S_1^{(1)} = \{3,4\}$

   $S_1^{(2)} = \{3\}$; $S_1^{(2)'} = R_1 - S_1^{(2)} = \{2,4\}$

   $S_1^{(3)} = \{4\}$; $S_1^{(3)'} = R_1 - S_1^{(3)} = \{2,3\}$

2) Number of 2-element subset of R_1 is $\binom{|R_1|}{2} = \binom{3}{2} = 3$.

   $S_2^{(1)} = \{2,3\}$; $S_2^{(1)'} = R_1 - S_2^{(1)} = \{4\}$

   $S_2^{(2)} = \{2,4\}$; $S_2^{(2)'} = R_1 - S_2^{(2)} = \{3\}$

   $S_2^{(3)} = \{3,4\}$; $S_2^{(3)'} = R_1 - S_2^{(3)} = \{2\}$

3) Number of 3-element subset of R_1 is $\binom{|R_1|}{3} = \binom{3}{3} = 1$.

   $S_3^{(1)} = \{2,3,4\}$; $S_3^{(1)'} = R_1 - S_3^{(1)} = \varnothing$

Hence, the probability of scheduling the 1^{st} user is

$$\Pr\left(A \mid \Gamma^{(1)} \in Q_k^{(1)}\right) = \sum_{m=0}^{3} \frac{1}{m+1} \left\{ \sum_{n=1}^{\binom{3}{m}} \left[\prod_{u \in S_m^{(n)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_m^{(n)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] \right\}$$

$$= \left\{ \sum_{n=1}^{\binom{3}{0}} \left[\prod_{u \in S_0^{(n)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_0^{(n)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] \right\} +$$

$$\frac{1}{2} \left\{ \sum_{n=1}^{\binom{3}{1}} \left[\prod_{u \in S_1^{(n)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_1^{(n)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] \right\} +$$

$$\frac{1}{3} \left\{ \sum_{n=1}^{\binom{3}{2}} \left[\prod_{u \in S_2^{(n)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_2^{(n)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] \right\} +$$

$$\frac{1}{4} \left\{ \sum_{n=1}^{\binom{3}{3}} \left[\prod_{u \in S_3^{(n)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_3^{(n)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] \right\}$$

$$= \left\{ \prod_{u \in S_0^{(1)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_0^{(1)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right\} +$$

$$\frac{1}{2} \left\{ \left[\prod_{u \in S_1^{(1)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_1^{(1)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] + \right.$$

$$\left[\prod_{u \in S_1^{(2)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_1^{(2)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] +$$

$$\left. \left[\prod_{u \in S_1^{(3)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_1^{(3)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] \right\} +$$

$$\frac{1}{3} \left\{ \left[\prod_{u \in S_2^{(1)}} \left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right) - F_{\Gamma^{(u)}}\left(q_k^{(u)}\right) \right] \prod_{v \in S_2^{(1)'}} F_{\Gamma^{(v)}}\left(q_k^{(v)}\right) \right] + \right.$$

$$\left[\prod_{u\in S_2^{(2)}}\left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right)-F_{\Gamma^{(u)}}\left(q_k^{(u)}\right)\right]\prod_{v\in S_2^{(2)'}}F_{\Gamma^{(v)}}\left(q_k^{(v)}\right)\right]+$$

$$\left.\left[\prod_{u\in S_2^{(3)}}\left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right)-F_{\Gamma^{(u)}}\left(q_k^{(u)}\right)\right]\prod_{v\in S_2^{(3)'}}F_{\Gamma^{(v)}}\left(q_k^{(v)}\right)\right]\right\}+$$

$$\frac{1}{4}\left\{\prod_{u\in S_3^{(1)}}\left[F_{\Gamma^{(u)}}\left(q_{k+1}^{(u)}\right)-F_{\Gamma^{(u)}}\left(q_k^{(u)}\right)\right]\prod_{v\in S_3^{(1)'}}F_{\Gamma^{(v)}}\left(q_k^{(v)}\right)\right\}$$

$$=\left\{F_{\Gamma^{(2)}}\left(q_k^{(2)}\right)F_{\Gamma^{(3)}}\left(q_k^{(3)}\right)F_{\Gamma^{(4)}}\left(q_k^{(4)}\right)\right\}+$$

$$\frac{1}{2}\left\{\left[F_{\Gamma^{(2)}}\left(q_{k+1}^{(2)}\right)-F_{\Gamma^{(2)}}\left(q_k^{(2)}\right)\right]F_{\Gamma^{(3)}}\left(q_k^{(3)}\right)F_{\Gamma^{(4)}}\left(q_k^{(4)}\right)+\right.$$

$$\left[F_{\Gamma^{(3)}}\left(q_{k+1}^{(3)}\right)-F_{\Gamma^{(3)}}\left(q_k^{(3)}\right)\right]F_{\Gamma^{(2)}}\left(q_k^{(2)}\right)F_{\Gamma^{(4)}}\left(q_k^{(4)}\right)+$$

$$\left.\left[F_{\Gamma^{(4)}}\left(q_{k+1}^{(4)}\right)-F_{\Gamma^{(4)}}\left(q_k^{(4)}\right)\right]F_{\Gamma^{(2)}}\left(q_k^{(2)}\right)F_{\Gamma^{(3)}}\left(q_k^{(3)}\right)\right\}+$$

$$\frac{1}{3}\left\{\left[F_{\Gamma^{(2)}}\left(q_{k+1}^{(2)}\right)-F_{\Gamma^{(2)}}\left(q_k^{(2)}\right)\right]\left[F_{\Gamma^{(3)}}\left(q_{k+1}^{(3)}\right)-F_{\Gamma^{(3)}}\left(q_k^{(3)}\right)\right]F_{\Gamma^{(4)}}\left(q_k^{(4)}\right)+\right.$$

$$\left[F_{\Gamma^{(2)}}\left(q_{k+1}^{(2)}\right)-F_{\Gamma^{(2)}}\left(q_k^{(2)}\right)\right]\left[F_{\Gamma^{(4)}}\left(q_{k+1}^{(4)}\right)-F_{\Gamma^{(4)}}\left(q_k^{(4)}\right)\right]F_{\Gamma^{(3)}}\left(q_k^{(3)}\right)\Big\}$$

$$\left.\left[F_{\Gamma^{(3)}}\left(q_{k+1}^{(3)}\right)-F_{\Gamma^{(3)}}\left(q_k^{(3)}\right)\right]\left[F_{\Gamma^{(4)}}\left(q_{k+1}^{(4)}\right)-F_{\Gamma^{(4)}}\left(q_k^{(4)}\right)\right]F_{\Gamma^{(2)}}\left(q_k^{(2)}\right)\right\}+$$

$$\frac{1}{4}\left\{\left[F_{\Gamma^{(2)}}\left(q_{k+1}^{(2)}\right)-F_{\Gamma^{(2)}}\left(q_k^{(2)}\right)\right]\left[F_{\Gamma^{(3)}}\left(q_{k+1}^{(3)}\right)-F_{\Gamma^{(3)}}\left(q_k^{(3)}\right)\right]\times\right.$$

$$\left.\left[F_{\Gamma^{(4)}}\left(q_{k+1}^{(4)}\right)-F_{\Gamma^{(4)}}\left(q_k^{(4)}\right)\right]\right\} \qquad \blacksquare$$

A.3 THE CLUSTERED HETEROGENEOUS CELL MODEL WITH THE RANDOM USER SELECTION RULE

Calculating the aggregate capacity belonging to the reference user in the k^{th} quantization level, using the random user selection rule in the clustered heterogeneous cell model with 3 clusters (T=3):

Let the reference user be in the first cluster.

$$\Pr\left(A \mid \Gamma^{(1)} \in Q_k^{(1)}\right) = \prod_{t=2}^{3} \left\{ \sum_{m^{(t)}=0}^{M^{(t)}} \binom{M^{(t)}}{m^{(t)}} \left[\Pr\left(\Gamma^{(t)} \in Q_k^{(t)}\right)\right]^{m^{(t)}} \left[\Pr\left(\Gamma^{(t)} \in \bigcup_{l<k} Q_l^{(t)} \right)\right]^{M^{(t)}-m^{(t)}} \right\} \times$$

$$\sum_{m^{(1)}}^{M^{(1)}-1} \frac{1}{1+\sum_{z=1}^{3} m^{(z)}} \binom{M^{(1)}-1}{m^{(1)}} \left[\Pr\left(\Gamma^{(1)} \in Q_k^{(1)}\right)\right]^{m^{(1)}} \left[\Pr\left(\Gamma^{(1)} \in \bigcup_{l<k} Q_l^{(1)} \right)\right]^{M^{(1)}-m^{(1)}}$$

$$= \sum_{m^{(2)}=0}^{M^{(2)}} \binom{M^{(2)}}{m^{(2)}} \left[\Pr\left(\Gamma^{(2)} \in Q_k^{(2)}\right)\right]^{m^{(2)}} \left[\Pr\left(\Gamma^{(2)} \in \bigcup_{l<k} Q_l^{(2)} \right)\right]^{M^{(2)}-m^{(2)}} \times$$

$$\sum_{m^{(3)}=0}^{M^{(3)}} \binom{M^{(3)}}{m^{(3)}} \left[\Pr\left(\Gamma^{(3)} \in Q_k^{(3)}\right)\right]^{m^{(3)}} \left[\Pr\left(\Gamma^{(3)} \in \bigcup_{l<k} Q_l^{(3)} \right)\right]^{M^{(3)}-m^{(3)}} \times$$

$$\sum_{m^{(1)}}^{M^{(1)}-1} \frac{1}{1+\sum_{z=1}^{3} m^{(z)}} \binom{M^{(1)}-1}{m^{(1)}} \left[\Pr\left(\Gamma^{(1)} \in Q_k^{(1)}\right)\right]^{m^{(1)}} \left[\Pr\left(\Gamma^{(1)} \in \bigcup_{l<k} Q_l^{(1)} \right)\right]^{M^{(1)}-m^{(1)}}$$

The capacity for any user in the 1st cluster is,

$$C_{CQMS\_randuser}^{(1)} = \sum_{k=0}^{K-1} \Pr\left(A \mid \Gamma^{(1)} \in Q_k^{(1)}\right) \int_{Q_k^{(1)}} \frac{1}{2} \log_2(1+\gamma) f_{\Gamma^{(1)}}(\gamma) d\gamma$$

Hence, the aggregate capacity can be calculated as

$$C_{agg\_CQMS\_randuser} = M^{(1)} C_{CQMS\_randuser}^{(1)} + M^{(2)} C_{CQMS\_randuser}^{(2)} + M^{(3)} C_{CQMS\_randuser}^{(3)} \qquad \blacksquare$$

A.4 THE CLUSTERED HETEROGENEOUS CELL MODEL WITH THE RANDOM CLUSTER SELECTION RULE

Calculating the aggregate capacity belonging to the reference user in the k^{th} quantization level, using the random cluster selection rule in the clustered heterogeneous cell model with 3 clusters (T=3):

Let the reference user be in the first cluster.

$$\Pr\left(A \mid \Gamma^{(1)} \in Q_k^{(1)}\right) = \Pr\left(\Gamma^{(1)} \mid Q_k^{(1)}\right) \sum_{i^{(2)}=0}^{1} \Pr\left(\Gamma^{(2)} \mid Q_k^{(2)}, s_{k,i^{(2)}}^{(2)}\right) \sum_{i^{(3)}=0}^{1} \Pr\left(\Gamma^{(3)} \mid Q_k^{(3)}, s_{k,i^{(3)}}^{(3)}\right) \frac{1}{1+i^{(2)}i^{(3)}}$$

$$= \Pr\left(\Gamma^{(1)} \mid Q_k^{(1)}\right) \Pr\left(\Gamma^{(2)} \mid Q_k^{(2)}, s_{k,0}^{(2)}\right) \Pr\left(\Gamma^{(3)} \mid Q_k^{(3)}, s_{k,0}^{(3)}\right) \times$$

$$\frac{1}{2} \Pr\left(\Gamma^{(2)} \mid Q_k^{(2)}, s_{k,0}^{(2)}\right) \Pr\left(\Gamma^{(3)} \mid Q_k^{(3)}, s_{k,1}^{(3)}\right) \times$$

$$\frac{1}{2} \Pr\left(\Gamma^{(2)} \mid Q_k^{(2)}, s_{k,1}^{(2)}\right) \Pr\left(\Gamma^{(2)} \mid Q_k^{(2)}, s_{k,0}^{(2)}\right) \times$$

$$\frac{1}{3} \Pr\left(\Gamma^{(3)} \mid Q_k^{(3)}, s_{k,1}^{(3)}\right) \Pr\left(\Gamma^{(3)} \mid Q_k^{(3)}, s_{k,1}^{(3)}\right)$$

$$= \sum_{m^{(1)}=0}^{M^{(1)}-1} \frac{1}{m^{(1)}+1} \binom{M^{(1)}-1}{m^{(1)}} \left[\Pr\left(\Gamma^{(1)} \in Q_k^{(1)}\right)\right]^{m^{(1)}} \left[\Pr\left(\Gamma^{(1)} \in \bigcup_{l<k} Q_l^{(1)}\right)\right]^{M^{(1)}-m^{(1)}-1} \times$$

$$\left[\Pr\left(\Gamma^{(2)} \in \bigcup_{l<k} Q_l^{(2)}\right)\right]^{M^{(2)}} \left[\Pr\left(\Gamma^{(3)} \in \bigcup_{l<k} Q_l^{(3)}\right)\right]^{M^{(3)}} \times$$

$$\frac{1}{2} \left[\Pr\left(\Gamma^{(2)} \in \bigcup_{l<k} Q_l^{(2)}\right)\right]^{M^{(2)}} \times$$

$$\sum_{m^{(3)}=1}^{M^{(3)}} \binom{M^{(3)}}{m^{(3)}} \left[\Pr\left(\Gamma^{(3)} \in Q_k^{(3)}\right)\right]^{m^{(3)}} \left[\Pr\left(\Gamma^{(3)} \in \bigcup_{l<k} Q_l^{(3)}\right)\right]^{M^{(3)}-m^{(3)}} \times$$

$$\frac{1}{2} \sum_{m^{(2)}=1}^{M^{(2)}} \binom{M^{(2)}}{m^{(2)}} \left[\Pr\left(\Gamma^{(2)} \in Q_k^{(2)}\right)\right]^{m^{(2)}} \left[\Pr\left(\Gamma^{(2)} \in \bigcup_{l<k} Q_l^{(2)}\right)\right]^{M^{(2)}-m^{(2)}} \times$$

$$\left[\Pr\left(\Gamma^{(3)} \in \bigcup_{l<k} Q_l^{(3)}\right)\right]^{M^{(3)}} \times$$

$$\frac{1}{3} \sum_{m^{(2)}=1}^{M^{(2)}} \binom{M^{(2)}}{m^{(2)}} \left[\Pr\left(\Gamma^{(2)} \in Q_k^{(2)}\right)\right]^{m^{(2)}} \left[\Pr\left(\Gamma^{(2)} \in \bigcup_{l<k} Q_l^{(2)}\right)\right]^{M^{(2)}-m^{(2)}} \times$$

$$\sum_{m^{(3)}=1}^{M^{(3)}} \binom{M^{(3)}}{m^{(3)}} \left[\Pr\left(\Gamma^{(3)} \in Q_k^{(3)}\right)\right]^{m^{(3)}} \left[\Pr\left(\Gamma^{(3)} \in \bigcup_{l<k} Q_l^{(3)}\right)\right]^{M^{(3)}-m^{(3)}}$$

The capacity for any user in the 1st cluster is,

$$C_{CQMS\_randcluster}^{(1)} = \sum_{k=0}^{K-1} \Pr\left(A \mid \Gamma^{(1)} \in Q_k^{(1)}\right) \int_{Q_k^{(1)}} \frac{1}{2} \log_2(1+\gamma) f_{\Gamma^{(1)}}(\gamma)\, d\gamma$$

Hence, the aggregate capacity can be calculated as

$$C_{agg\_CQMS\_randcluster} = M^{(1)} C_{CQMS\_randcluster}^{(1)} + M^{(2)} C_{CQMS\_randcluster}^{(2)} + M^{(3)} C_{CQMS\_randcluster}^{(3)} \quad \blacksquare$$

A.5 THE OPTIMUM THRESHOLD FOR A HOMOGENEOUS CELL WITH 2 USERS

Considering a homogeneous cell model with a single threshold, number of users is 2 and each user has $\bar{\gamma}$ mean SNR. There are two quantization levels, Q_0 and Q_1,

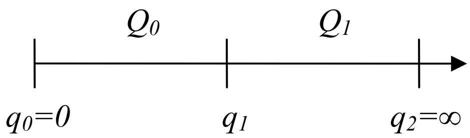

Figure A.1 The quantization levels of the example

Eq. (4.8) can be calculated for k=0 and k=1 as

$$\Pr(A/\Gamma \in Q_0) = \frac{F(q_1)}{2} \text{ and } \Pr(A/\Gamma \in Q_1) = \frac{F(q_1)+1}{2}$$

$$\begin{aligned} C_{CQMS\_homogeneous} &= \Pr(A/\Gamma \in Q_0)\int_{Q_0} \frac{1}{2}\log_2(1+\gamma) f_\Gamma(\gamma) d\gamma + \\ &\quad \Pr(A/\Gamma \in Q_1)\int_{Q_1} \frac{1}{2}\log_2(1+\gamma) f_\Gamma(\gamma) d\gamma \\ &= \frac{F(q_1)}{2}\int_0^{q_1} \frac{1}{2}\log_2(1+\gamma) f_\Gamma(\gamma) d\gamma + \frac{F(q_1)+1}{2}\int_{q_1}^{\infty} \frac{1}{2}\log_2(1+\gamma) f_\Gamma(\gamma) d\gamma \\ &= \frac{F(q_1)}{4}\int_0^{\infty} \log_2(1+\gamma) f_\Gamma(\gamma) d\gamma + \frac{1}{4}\int_0^{q_1} \log_2(1+\gamma) f_\Gamma(\gamma) d\gamma \end{aligned} \tag{A.1}$$

Since the optimum threshold q^* maximizes $C_{CQMS\_homogeneous}$,

$$\left.\frac{\partial C_{CQMS\_homogeneous}(q_1)}{\partial q_1}\right|_{q_1 = q^*} = 0 \tag{A.2}$$

According to [20]

$$\int_0^{\infty} \log_2(1+\gamma) f_\Gamma(\gamma) d\gamma = \frac{E_1(1/\bar{\gamma}) e^{1/\bar{\gamma}}}{\ln 2} \tag{A.3}$$

where

$E_1(x)=\int_x^{\infty}\frac{1}{t}e^{1/t}\,dt$ is the exponential integral function.

Putting eq. (A.1) into eq. (A.2),

$$\frac{\partial C_{CQMS\_homogeneous}(q_1)}{\partial q_1}=\frac{E_1(1/\bar{\gamma})e^{1/\bar{\gamma}}}{4\ln 2}\frac{\partial}{\partial q_1}F(q_1)+\frac{1}{4}\frac{\partial}{\partial q_1}\left(\int_0^{q_1}\log_2(1+\gamma)f_{\Gamma}(\gamma)d\gamma\right) \qquad \text{(A.4)}$$

using rule of differentiation under the integral sign when the limits of integral are the functions of the parameter [24],

$$\frac{\partial C_{CQMS\_homogeneous}(q_1)}{\partial q_1}=\frac{E_1(1/\bar{\gamma})e^{1/\bar{\gamma}}}{4\ln 2}\frac{1}{\bar{\gamma}}e^{-q/\bar{\gamma}}+\frac{1}{4}\left(-\log_2(1+q_1)\frac{1}{\bar{\gamma}}e^{-q/\bar{\gamma}}\right) \qquad \text{(A.5)}$$

Equalizing eq. (A.5) to 0 with substituting q^*, the optimum threshold can be found as

$$q^*=e^{E_1(1/\bar{\gamma})e^{1/\bar{\gamma}}}-1$$ ■