# THÈSE

présentée pour obtenir le grade de Docteur en Sciences de l'Université d'Avignon

**SPÉCIALITÉ : Informatique**

École Doctorale ED 536 « Sciences and Agrosciences »
Laboratoire d'Informatique d'Avignon (UPRES No 4128)

# *Cell Design and Resource Allocation for Small Cell Networks*

par

**Sreenath Ramanath**

**Soutenue publiquement le 6 Octobre 2011 devant un jury composé de :**

| | | | |
|---|---|---|---|
| M. | CHAHED Tijani | Professeur, Telecom SudParis, France | Rapporteur |
| M. | BONALD Thomas | Maître de conférences, Telecom ParisTech, France | Rapporteur |
| M. | ALTMAN Eitan | Directeur de recherche, INRIA, France | Directeur |
| M. | DEBBAH Mérouane | Professeur, Supelec, France | Co-directeur |
| M. | GESBERT David | Professeur, Eurecom, France | Examinateur |
| M. | KUMAR Vinod | Directeur, Global Research Partnerships, Alcatel Lucent, France | Examinateur |


This dissertation is carried out in the framework of the INRIA and Alcatel-Lucent Bell Labs Joint Research Lab on Self Organized Networks.

To my parents, family, friends and teachers.

# Part I

# Prologue

# Preface

It was a warm February evening (2008) in Bangalore that I met Dr. Eitan Altman the first time. He had just offered my wife, Kavitha, a post doc position at INRIA and wanted to know what were my plans. I was intending to take a sabbatical from my job to support my wife's research interests, when, he presented an irresistible offer: to be his student!. A dormant ambition, a desire to learn and the challenges of an untrodden path beckoned me to a world of research. It is with great pleasure that i thank Eitan for giving me this opportunity. His ideas, encouragement and support have gone a long way in understanding new topics and trends in wireless and networking research.

I would like to thank Dr. Philippe Nain for accepting me into the Maestro Group at INRIA, Sophia Antipolis and encouraging me to actively participate in group meetings, conferences and workshops.

I would like to acknowledge Laurant Thomas for accepting me to be a part of the INRIA-Bell Labs initiative on Self Organizing Networks and proposing a very interesting and happening research topic. I appreciate Dr. Vinod Kumar's encouragement and initiative to make me feel comfortable and settle down with research as well as life in France.

A fortnight into my arrival in France, i found a surprise mail. This opened up a lot of new opportunities and dimensions in my research. Thank you Prof. Merouane Debbah for being my thesis co-advisor and for your support and encouragement.

Over the course of my thesis, i was introduced to Laurant Roullet. I would like to acknowledge his interest and patience in understanding and encouraging me in my research.

Innovation is key in any research and i do hope to find some of the ideas presented in this thesis to find use in practice. I would like to acknowledge Dr. Veronique Capdeville in showing interest in my research and encouraging me to participate in realizing these in practice.

I would like to thank the administrative staff at INRIA, especially, Ephie Deriche and Laurie Vermeersch for being patient with our many queries and helping us to plan visits to conferences and workshops. I would like to acknowledge the administrative staff at LIA, specifically, Simone Mouzac and Jocelyne Gourret for taking care of our day to day requirements and activities.

I want to express my sincere gratitude to Prof. Tijani Chahed and Prof. Thomas Bonald, who accepted to review the manuscript and Prof. David Gesbert and Dr. Vinod Kumar, who accepted to be my thesis examiners. Thankyou for your patient reading, valuable comments to improvise the thesis and kind words of encouragement.

A review comment sparked a thought process which resulted in a flash survey to find the top five innovations and advancements in the field of wireless communications in recent years. I would like to express my sincere thanks to many Professors, colleagues, colloborators and acquaintances who took part and shared their views.

I would like to thank my colleagues in INRIA, LIA and Supelec for many interesting discussions and words of encouragement. Special thanks go to Kostia, Tania, Rachid, Yezekael, Sabir, Tembine, Julio, Sulan, Yuedong, Amar, Manjesh, Baslam, Habib, Wissam, Romain, Alonso, Veronica, Leonardo, Antonia, Jakob, Salam and Subhash for their technical as well as friendly interactions.

A special mention of *Merci Beaucoup* goes to Richard Combes for his help in writing the thesis abstract in French (*Résumé*).

A very special thanks to my daughter, Prajna, who learnt and thought me French. Her school, friends and activities made my everyday life more colorful, joyous and fun, not to mention the feeling of revisiting my childhood.

Coming back to Kavitha, she has been an inspiration which i couldn't catch up!. Her willingness and dedication to understand and learn is exemplary. She has been a constant source of encouragement and help in understanding and solving many a tricky twists and turns that came in the course of my research. I owe her a lot.

I wish to thank my parents, sisters and their families, my in-laws, for their unconditioned love, encouragement and support.

And thanks to all my friends and relatives for their love and encouragement during the tenure of my studies.

Sep 2011, Avignon, France Sreenath Ramanath

# Résumé

Récemment, il y a eu une hausse massive du trafic dans les réseaux mobiles à cause de nouveaux services et applications. Les architectures actuelles des réseaux cellulaires ne sont plus capables de gérer de façon satisfaisante ce trafic. Les Réseaux de Petites Cellules (RPC), basées sur un déploiement dense de stations de bases portables, auto-organisantes et efficaces en termes d'énergie apparait comme une solution prometteuse à ce problème. Les RPC augmentent la capacité du réseau, réduisent sa consommation énergétique et améliorent sa couverture. Par contre, elles posent des défis importants en termes de design optimal.

Dans cette thèse, des aspects liés au design cellulaire et à l'allocation de ressources dans les RPC sont traités. La thèse se compose de deux parties.

Dans la première partie, le design cellulaire est étudié: une population statique d'utilisateurs est considérée, et la taille optimale de cellule maximisant le débit spatial est donnée en fonction du modèle de récepteur, des conditions radio et des partitions indoor/outdoor. En considérant des utilisateurs mobiles, la taille de cellule optimale est étudiée afin de minimiser le temps de service, et minimiser le blocage et la déconnexion en cours de communication, en fonction de la vitesse des utilisateurs et du type de trafic. Le problème de placement des stations de base optimal est traité en fonction de différents critères de qualité (maximisation de débit total, équité proportionnelle, minimisation de délai, équité max-min) pour différentes distributions d'utilisateurs et partitions de cellules. Le problème de scaling de capacité dans un RPC limité par l'interférence avec pré-codage est étudié, et la quantité optimale d'antennes par utilisateurs en fonction de l'interférence inter-cellules est dérivée. Dans le cadre d'un réseau "green", pour une charge du réseau donnée, on étudie les politiques optimales en boucle ouverte, afin de maximiser soit une fonction coût du système (contrôle centralisé) soit des fonctions de coût de chacune des stations de base (contrôle distribué).

Dans la seconde partie, nous étudions l'allocation de ressources, nous introduisons les concepts de d'équité T-échelle et équité multi-échelle. Ces concepts permettent de distribuer les ressources équitablement pour les différentes classes de trafic. Ces concepts sont illustrés par des applications au partage de spectre et à l'allocation de ressources dans les femto-cellules indoor/outdoor. L'allocation de puissance pour satisfaire les demandes de trafic des utilisateurs avec un grand nombre d'interféreurs est une tâche difficile. Ce problème est abordé, et nous proposons un algorithme universel

qui converge vers une configuration de puissance optimale qui satisfait les demandes des utilisateurs dans toutes les stations de base. Les performances de l'algorithme sont illustrées pour différentes configurations du système et différents niveaux de coopération entre les stations de base.

# Abstract

An ever increasing demand for mobile broadband applications and services is leading to a massive network densification. The current cellular system architectures are both economically and ecologically limited to handle this. The concept of small-cell networks (SCNs) based on the idea of dense deployment of self-organizing, low-cost, low-power base station (BSs) is a promising alternative.

Although SCNs have the potential to significantly increase the capacity and coverage of cellular networks while reducing their energy consumption, they pose many new challenges to the optimal system design.

Due to small cell sizes, the mobile users cross over many cells during the course of their service resulting in Frequent handovers. Also, due to proximity of base stations (BS), users (especially those at cell edges) experience a higher degree of interference from neighboring base stations. If one has to derive advantages from small networks, these alleviated effects have to be taken care either by compromising on some aspects of optimality (like dedicating extra resources) or by innovating smarter algorithms or by a combination of the two.

The concept of umbrella cells is introduced to take care of frequent handovers. Here extra resources are dedicated to ensure that the calls are not dropped within an umbrella cell. To manage interference, one might have to ensure that the neighboring cells always operate in independent channels or design algorithms which work well in interference dominant scenarios or use the backhaul to incorporate base station cooperation techniques. Further, small cell BS are most often battery operated, which calls for efficient power utilization and energy conservation techniques. Also, when deployed in urban areas, some of the small cells can have larger concentration of users throughout the cell, for example, hot-spots, which calls in for design of small cell networks with dense users. Also, with portable base stations, one has the choice to install them on street infrastructure or within residential complexes. In such cases, cell design and resource allocation has to consider aspects like user density, distribution (indoor-outdoor), mobility, attenuation, etc.

We present the thesis in two parts. In the first part we study the cell design aspects while the second part deals with the resource allocation. While the focus is on small cells, some of the results derived and the tools and techniques used are also applicable to conventional cellular systems.

In the first part, we study various aspects of cell design like cell dimensioning, base station placement, optimal fraction of users per transmit antenna, base station(s) activation policies etc.

For cell dimensioning, we consider two scenarios based on the mobility pattern of the users. For systems supporting only static users (**Chapter 2**), under fluid limits (i.e., when the number of users is large), we compute the cell size that maximizes the spatial throughput density. For systems that support mobile users (**Chapter 3**), we use the concept of umbrella cells. We model the small cell network using queueing theoretic models and obtain cell size that optimizes various performance measures like expected waiting times, service times, call block and drop probabilities, etc. We next forego this umbrella assumption and obtain the cell size that optimizes call block and drop probabilities. We make some interesting observations like that the optimal cell size is independent of the traffic type and that for a given power there is a limit velocity beyond which useful communication ceases.

Once the cell dimensioning is done, further design rules tend to assume that the BS is centrally located. But, does the placement of BS matter? Does it change depending on some criteria? We analyze this in the context of locating base stations (**Chapter 4**) according to some fairness criterion. We show that the location of the base station converges to the center of the cell as the fairness parameter tends to infinity; i.e, the max-min fair BS placement is the center of the cell. Further, this is true independent of the underlying user density or characteristics of the cell (eg. indoor-outdoor partitions, hot-spots, etc.).

While, it is well known that dividing a large cell into number of small cells enhances the system capacity, the spatial dimension can be exploited to enhance the capacity further. In this aspect, we consider a MIMO broadcast channel (**Chapter 5**) and investigate the effect of multi-cell interference in precoded small cell networks. We show that their exists an optimal ratio of number of antennas at the BS to the number of users for a given interference level. The problem is solved in the asymptotic limit using random matrix theory and we show via numerical simulations that the asymptotic expressions are reliable even in the finite case.

Cell design considering dense deployment of BS as in small cell networks need to be energy efficient. We come up with optimal open loop BS activation policies (**Chapter 6**), which depends on the system load. We use tools from multimodularity to derive the structure of the optimal policies.

In the second part of the thesis, we address resource allocation. Resources are to be allocated so as to fair share the average utilities that corresponds to the assignments. But the exact definition of average share depends on the application! Different applications require averaging over different time periods or time scales (eg., real time voice, file downloads, gaming, multimedia streaming). Hence fairness need to be defined over mixed timescales. In this context, we introduce T-scale and multiscale fairness (**Chapter 7**). This new concept allows one to distribute the network resources fairly among different classes of traffic. We illustrate this concept via some example applications in spectrum allocation and in indoor-outdoor femtocells.

Futher, given that we have users with varying QoS demands, how do we allocate power to satisfy their demands? What if we have different system architectures, supporting various standards and interfering with each other?. Is there a simple and effective self-organizing power allocation mechanism, which can work for any or a combination of these systems? We address this problem of power allocation to satisfy user demand rates (**Chapter 8**) in a multicell network. We propose a simple and universal power allocation algorithm which guarantees convergence to user demands, whenever the demands lie within the system specific rate region. This algorithm can work from a completely centralized to a fully distributed setting. Further, with macrocells and small cells co-existing, we propose an extension of the algorithm to multi-tier networks.

We have used a variety of tools to model and analyze the problems that arise in dimensioning and resource allocation. Specifically, queueing theory, random matrix theory, stochastic approximation, etc, have been used to understand, characterize and derive asymptotic results and practically implementable algorithms for self organizing networks.

Our hope is that this thesis will form a initial framework to explore and exploit the many dimensions that arise in designing optimal small cell systems, which will satisfy the next generation mobile broadband user.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 General introduction

Emergence of a variety of standards for Wireless Communication Networks in culmination with advances in Radio Access Technologies offer increased reach, higher capacity, improved quality of service and many more things, while reducing energy consumption and deployment costs, paving the way for new applications and services in mobile broadband access.

A pioneer of a computer networking systems is ALOHAnet [2], popularly known as ALOHA, developed at the University of Hawaii, which became operational in 1971, providing the first demonstration of a wireless data network.

ALOHA used experimental UHF frequencies to begin with; as frequency assignments for commercial applications were not available in the 1970s. Further, ALOHA was used in cable (Ethernet based) and satellite (Immarsat) applications. In the early 1980s frequencies for mobile networks became available, and in 1985 frequencies suitable for Wi-Fi were allocated in the US. These regulatory developments made it possible to use ALOHA in both Wi-Fi and in mobile telephone networks. Since then ALOHA has found applications across a multitude of wireline and wireless technologies.

While ALOHA has been a pioneer networking system, which spanned across wireline and wireless networks, the wireless technology itself has evolved over the past few years from using analog FM transmission for voice telephony to OFDM / OFDMA for mobile intenet and video streaming applications in the recent years. In table 1.1, we summarize the evolution of wireless generations over the past few decades [158].

| Generation | Period | Transmission | Services | Examples |
|---|---|---|---|---|
| 1G | 1980's | Analog FM<br>FDMA / FDD | Voice | AMPS |
| 2G | 1990's | Digital modulation<br>TDMA/CDMA | Voice, SMS | GSM, IS-95<br>GPRS, EDGE |
| 3G | 2000+ | Wideband<br>modulation | Internet, email,<br>Multi-media<br>streaming, etc. | WCDMA<br>HSDPA, HSUPA |
| 4G | 2005+ | OFDMA | Mobile internet<br>Mobile video, etc. | LTE, WiMAX |

***Table 1.1:*** *Evolution of wireless generations*

## 1.2 Recent innovations and advances in wireless communication

In this section, we shall discuss a number of recent innovations and advances in information theory and signal processing that have enabled reliable and fastet communication. While the initial advancement was to push channel coding to achieve near Shannon limits, the recent ones focus on exploiting the inherent properties of the wireless channel: fading and interference, which form the basic ingredients of Opportunistic communication.

The first big thing that comes to our mind is Turbo codes [23], followed by Gallager's *forgotten* LDPC codes [61]. Both of these came within 0.5 dB of the Shannon capacity limit. Though these codes came to light in the late 90's and early 2000, the advancement of VLSI, made its realisability in practice and they have become an integral part of todays and future wireless standards. Of late, there has been a lot of exitement about fountain codes [38] and polar codes [16]. They not only pack the technological breakthrough of turbo and LDPC codes, but also promise simpler implementation complexity. In fact, they are shown to come even closer to Shannon limits.

The spatial component of the wireless channels has seen tremendous advancement in recent, starting from Telatar's landmark paper [144] and became popular with the famed Alamouti [3] code. The Alamouti code has the unique property of being the simplest space time code [142, 109] which offers full diversity and multiplexing gain. Thus started the exploration and exploitation of spatial diversity, brought about by multiple antennas.

Multiple Input Multiple Output (MIMO), based on Multiple antennas at the transmitter and / or receiver is a technique which achieves diversity and / or multiplexing gains. This enables higher data rates between transmitters and receivers. The tradeoff between diversity and multiplexing has been well captured by the landmark paper of Tse [148]. Of late, MIMO has become a standard component of research problems and simpler versions of it (upto 4 X 4 tx/rx configurations) including Alamouti codes, Rate-2, Rate-3, Rate-4 space time codes (STC) have become part of WiMAX and LTE

standards. Multiple antenna transmissions also makes it possible to offer better QoS to cell edge users by cleverly beamforming or precoding the transmissions towards them.

Coming to multiplexing techniques, CDMA (Code Division Multiple Access) was the popular choice of 90's, while OFDM (Orthogonal Frequency Disision Multiplexing [40]) has become the defacto in the 2000s. The advancement in signal processing in culmination with the gains of narrow band signalling has made OFDM a popular choice in todays modems. The biggest disadvantage comes from fluctuations in the signal levels when the frequency domain information is converted to time domain via FFT. The extent of these fluctuations are measured by the metric peak to average power ratio (PAPR), which can be as high as 15-20 dB and large PAPR leads to problems with the design of power amplifiers (larger dynamic range). With the advancement in PAPR reduction techniques, using both signal processing and RF techniques and with improved power amplifier technology, this issue seems to be well taken care off and OFDM is here to stay.

OFDMA [169] is a multiple access technique based on OFDM. Here, the available sub carriers spreading across the spectrum of interest is split amongst multiple users. With this it is possible to choose a cluster of sub-carriers and users in an optimal way to combat fading and achieve diversity as well.

Ultra wide band (UWB) [163] was talked about a lot in recent years. This is a modulation scheme which uses narrow (time domain) pulses spreading over multiple GHz of frequency spectrum. They were proposed for multiple applications including short range communications (Personal Area Networks: communication over few tens of meters). Though there was a huge surge of interest initially, in recent years, it is still waiting to take off owing to spectral and economic viability issues.

We come back to our initial discussion about the two fundamental properties of wireless channel and discuss how in recent years, these are further exploited to improve spectral efficiency and resource utilization.

The fading nature of wireless channels is exploited to advantage via Opportunistic communication [72]. This involves the transmitters being aware of the channel towards its users (via feedback) and favoring those users with better channel conditions. Since fading is a time/frequency varying phenomenon, everyone stands to benefit over some averaging duration. Further, fair schedulers [91] can be employed to guarentee certain QoS, even to disadvantaged users (cell edge or non LOS).

The other interesting idea of opportunistic communication is the principle of Cognitive radio [83]. This exploits unused spectrum and transmission opportunities of primary subscribers to schedule secondary users.

Relaying [80], yet another aspect of Opportunistic communication, manages interfernce and improve end to end QoS. Signals from transmitters to receivers are routed through relays which offer the best channel conditions. Also, due to the reduced transmission ranges, power budgets are reduced and hence interfernce. The disadvantage is the extra resouces needed to setup and maintain relays.

Of late, co-operative as well as distributive strategies are gaining popularity. Co-operative strategies like multi cell co-opertion [67, 80] aim at managing interfernce actively. At any given time the channel state of the entire system is known and a central controller can now decide the most optimal way the communciation happens between individual transmitters and receivers. Recently, using these strategies, it has been shown that the multicell capacity is same as the single cell capacity multiplied by the number of cells [67]. However, in practice, due to multiple limitations involving backhaul, latency, processing power, etc., sub-optimal schemes like clustering [67, 80], where few neighboring base stations (e.g, clustering) share the channel state or schemes that use a very minimal form of channel states like channel statistics have been proposed to make it possible for practical realization. Some of these are already underway for standardization.

Decentralized and distributed processing is becoming popular to manage the complexity of central control. Here, each agent (base station), simultaneously updates its parameters (runs an algorithm) to achieve a certain purpose. For example, each base station could run a power allocation algorithm to meet its users demands with power budget contraints, while exchanging minimal information (channel statistics, rates allocated, etc.,) with its neighbours. Thus distributed processing becomes an essential part of a self organizing system, guaranteeing a certain minimum QoS to each user. Each agent (base station / mobile) learns and adapts to the environment and thus manages to do the best. This is where the concepts of learning and adaptive algorithms come in. Reinforcement learning [165, 138], stochastic approximation [93], etc are becoming an integral part of new generation base stations. Game theory [11] and ODE (Ordinary Differntial Equations) tools are extensively used to formulate, analyze and understand the behaviour of these algorithms.

Another point to mention is the recent surge in usage of tools like stochastic geometry to analyze cellular networks [17]. The traditional methods of network analysis assumes linear or hexagonal networks. Stochastic geometry takes into account the random location (distribution) of transmit and receive nodes, which is a more realisitc assumption and many a times, it is possible to obtain explicit expressions for important system metrics. For example, one can compute explicit expressions for the total interference at a base station with simultaneous transmissions from randomly located nodes, distributed according to a poisson process [17]. This can be plugged into the Signal to Interference plus Noise (SINR) equations to get a more realistic value as compared to a idealized value with regular placement of nodes. Thus the upper bounds and lower bounds of system metrics become more tighter. Though this is an emerging idea, it is still very difficult to analyze systems with spatial randomness and one often is comfortable using the popular Wyner-type cellular modelling [166] to get first cut understanding of new advancements like multi cell co-operation strategies [67].

Current view: With all these discussions, a point to ponder is that some of the innovations and advancements are a promise for the future, what about today? We need to find a quicker solution to meet todays smart phones, gaming devices and tablets needs. This is where topology comes to aid technology. Dividing a large cell into number of small cells [107, 49, 44, 82, 52] fits the old adage 'divide and conquer'. The reduced

cell sizes improve capacity and coverage. There are drawbacks related to infrastructural costs, backbone, mobility induced frequent handovers, etc. Some of these, we shall address in due course. But, small cells appear promising to overcome coverage and capacity bottlenecks of todays applications. Also, it is a matter of time before the technical advancements we discussed become integral part of the small cell technology.

A recent survey: With so many new and exiting possibilites as discussed in previous paragraphs, we wanted to know what were the most landmark innovations and advancements in recent years as seen by experts in the research and engineering community. We prepared a simple survey to identify the top few innovations and advancements in wirless communications and predict the next big thing in Information theory. We got some very interesting view points (See **Appendix 9**). Some said the survey itself was very thought provoking and needs careful thinking. Others felt that the guessing game is difficult, citing examples of LDPC. Few others opined that it is all related to simplicity of the idea, ease of implementation, standardization and economics of deployment. Citing from law of large numbers (of opinions), the top five without any particular order were Turbo/LDPC codes, OFDMA, MIMO, Opportunistic communication and Multicell co-operative networks. The answer for the next big thing: Network information theory.

Topology vs. Technology: Let us now illustrate via a few examples as to how reducing the cell size turns out to be more beneficial when combined with recente technical advancements.

For example, opportunistic communication aims at scheduling and allocating users which have good channel conditions. Users close to the base station enjoy the benefits of such schemes owing to better channel conditions (better signal strength, LOS components, etc.), while users at cell edge are disadvantaged. To circumvent this, a base station can employ a proportional-fair scheduler to improve the QoS of the cell edge user, but, this will be at the cost of decreasing the QoS for users with better channel conditions. Thus decreasing the cell size benefits opportunistic communication schemes and also increases the fairness in the system.

Another example, OFDMA, aims at dividing the available sub-carriers amongst the users to avail the benefits of frequency non-selective transmissions. But, again the limitations are apparent as one has to allocate more number of sub-carriers to cell edge users to maintain the QoS of the link. Shrinking the cell size translates to meeting the same QoS with lesser number of sub-carriers. The left over sub-carriers can be used to support additional users for example.

Multiple antennas at the base station and users can dramatically improve the achievable capacity on a given link. Beam forming or precoding can effectively focus a beam to create better SINR conditions for cell edge users, while reducing interference towards other users. But, this comes at the cost of using more resources at the transmitter. For example, a significant portion of the available power from the total power budget at the base station is used towards the cell edge users. Bringing the cell edges closer, power allocation towards the cell edge users dramatically decreases.

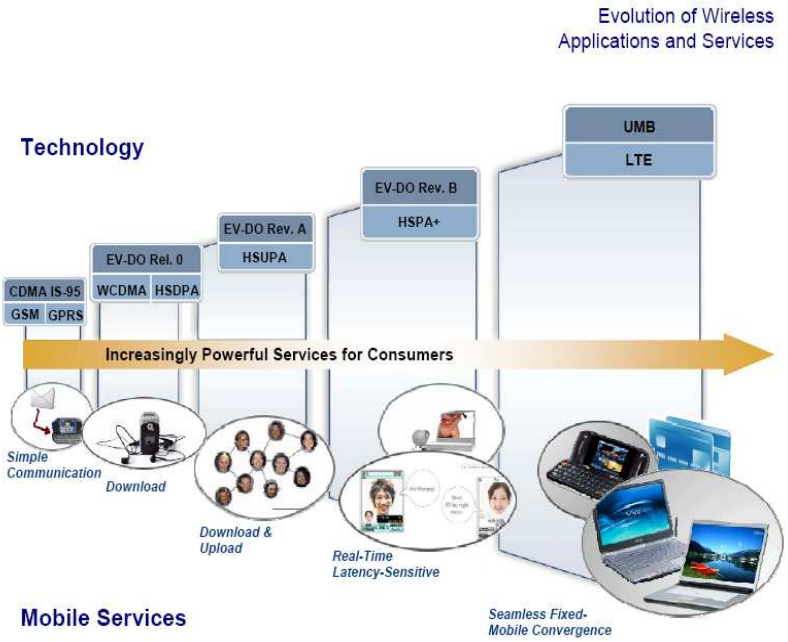


***Figure 1.1:*** *Evolution of wireless applications and services (Qualcomm [112])*

.

Takeaway point: From our discussions so far, the aim behind these successive generations and technology advancement is to simplify the mode of communication, equip the user with versatile features and finally satisfy him or her on the move. With the rapid growth in the number of wireless applications, services and devices, using a single wireless technology such as 2G or 3G will be inadequate to meet the data rate and QoS constraints in a seamless manner (see [79]). To provide seamless broadband connectivity to mobile users, the next generation wireless systems (4G and beyond) are being devised with the vision of heterogeneity (see figure 1.2) in which a mobile user/device will be able to connect to multiple wireless networks. This ensures that the user is always connected to the best network. Further, cell edge users can constrain the system to a very great extent and the benefits of technical advances still limit the system performance, especially to meet the traffic demands of recent mobile broadband applications. System designers have to exploit new dimensions to manage this.

Thus topology aiding technology advancement is seen as the next key step in meeting the capacity requirements of next generation wireless networks.

## 1.3 Small cell networks

While, advances in technology in culmination with heterogeneous networks adds new capabilities and features, the innovation in new services and applications always leave some users uncovered, under served and dissatisfied. Here is where the Small Cell

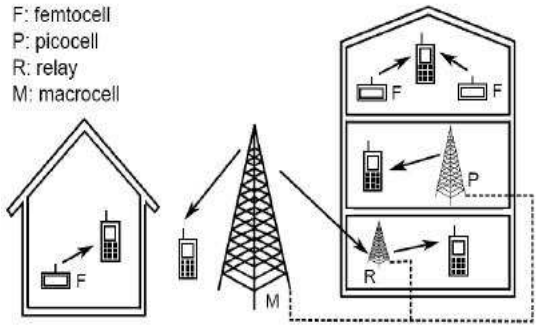


***Figure 1.2:*** *Example of heterogeneous network deployment with macrocells complemented by relays and pico/femto cells (Guillaume de la Roche, et. al., [50])*

.

Networks (SCNs) fit in. They fill these gaps [107, 49, 44, 82]. At the end of the day (or the week), the operators know how and where the network is loaded, the user behavior and demands. The best way to handle this is to put in some smart, portable base stations, which are self-configuring, self-organizing and self-healing. They are sufficiently intelligent to sense, learn and respond to the environment. Be it urban hotspots, social events, sports, malls, etc, small cells are the smart answer to an adapting wireless network.

The trend in traffic demand for wireless data and video services have seen an exponential growth in the recent years (e.g. figure 1.3 [44]). Market forecasts from industry leaders like CISCO (e.g. figure 1.4 [154]), predict the trend to continue in the coming years. In addition, users tend to expect a LAN-like experience on the move. The existing network with its backhaul capabilities has reached a bottleneck and the emerging networks, apart from being able to deliver high data throughput are expected to be energy efficient and environment friendly. Thus the design and deployment of next generation networks need to address contrasting requirements like increasing capacity and coverage, while keeping the energy consumption and emissions within reasonable limits.

Current 3G and emerging 4G wireless networks (like WiMAX and LTE) are unable to meet the demand. One of the promising solutions to this is to divide a large cell into a number of small cells ([107, 49, 44, 82]), thereby increasing the capacity and coverage in the existing network. To be cost effective and efficient, such a network needs to employ self-organizing, low-cost, low-power base stations (BSs). But, such a deployment poses many a new challenges to the optimal system design. In subsequent paragraphs, we will address some of the problems and challenges that arise in the deployment of small cell networks. We specifically address problems in cell design and resource allocation in this thesis.

Managing a dense network is a difficult task. Thus self configuring, organizing and healing mechanisms need to be built into small cell base stations. These networks should be able to configure radio, system parameters, optimize resource allocation and handover procedures, have energy conservation mechanisms and be capable of recovering from node failures. Also, since SCNs are connected via backhaul, challenges

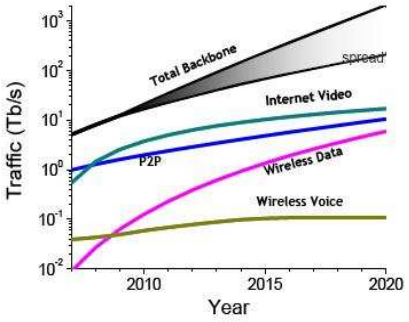


***Figure 1.3:*** *Traffic demand (Claussen [44])*

.

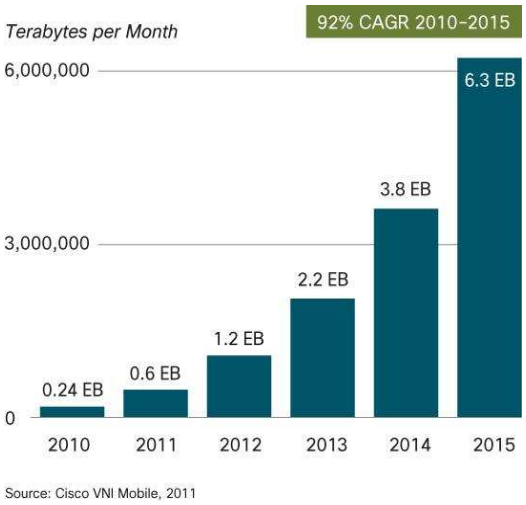


***Figure 1.4:*** *Mobile data forecast (Cisco [154])*

.

posed by delay sensitive traffic needs to be addressed as well.

One of the major challenges is managing frequent handovers when users are mobile. With cell sizes from 10's of meters to a couple of 100 meters, every few seconds is a potential handover. Thus new mechanisms need to be built to manage frequent handovers and HO signalling overhead. Grouping slow speed users to small cells and high speed to macro cells [44] is one of the solutions. Another alternative is the formation of virtual cells, i.e., a cluster of cooperating small cells that appears to the user as a single distributed BS [41][4]. In this setting, handovers would occur only at virtual cell boundaries. Further fast bast-station switching (FBSS) is also a possible solution. Cell dimensioning involving static and mobile users is a key issue to address.

Some of the other key challenges include interference management, resource allocation, energy conservation, etc. (refer [107, 49, 44, 82] for more details).

***Figure 1.5:*** *Example of Pico cell deployment serving static and moving users*

.

Typically, small cells comprise micro, pico and femto cells. Micro cells are miniature counterparts of large macrocells. The idea of micro cells is to ensure enhanced data rate and better coverage services in dense urban areas. Pico cells (see figure (1.5)) are smaller when compared to micro cells. Pico base stations are designed to be portable and easy to mount on existing street and outdoor infrastucture. Further, they can effectively manage mobile users. Thus, they are intended to provide capacity and coverage gains in dense urban areas with both static and mobile users. Such pico cells, while serving the outdoor users better, fall short to achieve higher throughputs and wider coverage while serving indoor users due to signal attenuation. Thus, indoor users can be better served with an even smaller form factor indoor base stations, the femtos. Thus Femto cells cater to indoor coverage serving homes and small offices. The future small cell base stations design can be configured to function either as a micro, pico or a femto BS. One of the key challenges to address is managing mobile users among these heterogeneous networks.

Thus small cell networks are a paradigm shift from the conventional network design and pose many a new challenges in optimal system design. In this context, we address cell design and resource allocation in small cell networks. In the rest of the thesis, we model, analyze and study key performance measures in SCNs (with static and mobile) related to cell deisgn and resource allocation. We use how tools from queuing theory, random matrix theory, multimodularity, stochastic approximation and new concepts in fairness can be effectively used to study cell design and resource allocation in Small Cell Networks.

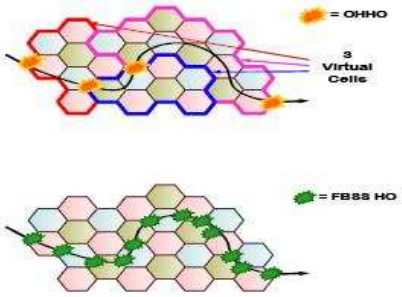


***Figure 1.6:*** *Possible alternatives to support mobility in small cells (Alcatel Lucent - Bell Labs [4])*

.

## 1.4 Thesis Overview

In this thesis, we address cell design and resource allocation for small cell networks.

One of the key issues addressed in this thesis is cell dimensioning. This aspect is studied with respect to achievable capacity using fluid models [90] with static users with different receiver structures, frequency reuse, path-loss effects and various type of cell partitions. Cell dimensioning with mobile users is studied in the context of different classes of traffic and various pedestrian and vehicular velocity profiles. Assuming random but fixed velocity of mobiles, we use tools from spatial queueing theory [164] to derive cell sizes which minimizes key performance metrics like expected waiting times, call block and drop probabilities, etc.

Next, we address the base station location problem. Where do we locate base stations to be throughput optimal? which location minimizes the delay? Is the center of the cell proportionally fair or max-min fair? We address this interesting base station location problem using the popular alpha-fair fairness criterion [105]. Using large population limits, we compute the base station locations that are optimal for a given degree of fairness. We indeed show that the center of the cell is a max-min fair BS location, while considering varying density of users and cell partitions.

How does the capacity scale in an interference limited multi-antenna precoded small cell network? Is there an optimal user density per antenna for a given degree of interference? We answer these questios via asymptotic analysis via random matrix theory [46]. Simulations establish the results to be true in the finite regime. Further, the asymptotic expressions are used to study power allocation to satisfy user rate demands.

Different applications need their utilities to be averaged over different time scales. Thus the amount of resource that a user gets depends on his application and each application has a different timescale. In this context, we introduce the concept of multiscale fairness, which encapsulates short-term, long-term and other notions of fairness. We demonstrate the application of this concept in some example applications in spectrum

allocation and indoor-outdoor femto cells.

How do we allocate power to satisfy user demands in a multi-cell, heterogeneous network?. We address this problem and propose a stochastic approximation [93] based universal power allocation algorithm. We demonstrate the working of this algorithm for systems with various degree of co-operation. This self-organizing algorithm fits the paradigm of self organizing small cell networks.

Dense deployment of small cell's address capacity and coverage holes. But, the system load varies over time. Hence, from a greener perspective, one can switch off a fraction of base stations via a central control or put a base station in an idle mode via decentralized control, depending on the load. We derive the structure of this control using tools from multimodularity [12].

High speed mobiles are subjected to frequent handovers due to the dimension of small cells during the course of their service. We address this problem of managing high speed mobility. We propose novel ways of power and resource allocation to manage high speed users in small cells.

## 1.5 Tools used in the thesis

In this section, we present a brief overview of the tools used in the thesis. We have used fluid limits [90] to address cell dimensioning with static users in **chapter 2**. Fluid limits many a times yield explicit expressions for relevant performance measures, which are tractable and are a good starting point to study complex problems (e.g, cell dimensioning). Further they reduce the simulation overhead. Similar tools are also used in **chapter 4** to address the base station placement problem.

Fluid limits are asymptotic limits and hence are valid when some underlying quantities tend to infinity. For example when the user density or base station density increases to infinity. Alternatively, queueing tools are useful when asymptotics are not valid. In **chapter 3**, we use queueing tools to derive cell dimensions that optimize service times, waiting times, call block and drop probabilities, while considering mobile users. Explicit expressions for various performance metrics can be easily found in most of the books on queueing theory, e. g., [164] and can be straightaway used if a system can be modeled as a certain type of a queue.

To analyze capacity scaling and per antenna user density that can be supported in interference limited multicell precoded systems in **chapter 5**, we have used random matrix theory [149, 46]. Most often, the limiting spectral distribution of these large random matrices, representative of the system under consideration can be expressed explicitly by the popular Stieltjes or other transforms. The asymptotic results obtained via such an approach has been shown to be quite effective even in the finite regime.

For the problem of deriving the structure of the control policy for base station activation in **chapter 6**, we have used tools from Multimodularity. Multimodularity addresses convex functions over integer spaces and if applied to systems whose cost func-

tion evolves in a max-plus algebra, the axioms of Multimodularity can be proven and the result is a simple well structured policy. This tool, introduced in [75] has been used to address many problems in admission control, routing and other applications [12].

Fairness comes at the cost of efficiency. But, users can demand the same QoS irrespective of their location w.r.t the serving base station. So, the resource allocation policies at the base stations cannot always be selfish to maximize their own revenue. To keep the user satisfied, they have to sometimes maximize the utility of the weakest user. So, it is necessary to be fair on many counts (global, proportional, delay minimizing, max-min, etc., ). These different notions of fairness are encapsulated in the concept of alpha fairness [105]. Further, different applications need the utilties to be averaged over different time scales. With this view, the concept of alpha fairness in conjunction with traffic type and averaging durations has been used to come up with new fairness concepts; the T-scale and multiscale fairness in **chapter 7**. The concept of alpha fairness is also used in fair location of base stations **chapter 4**.

Stochastic approximation (e.g., [93]) analysis are powerful and are used extensively in a variety of applications with iterative algorithms. They are handful in obtaining the transient as well as steady state behaviour of the iterative algorithms. Ordinary differential equation (ODE) approach is a popular one while studying the stochastic approximation based algorithms. In this approach, either the iterative algorithm is approximated by the trajectory of an appropriate ODE or the time asymptotic limits of an algorithm are obtained via the attractors of a ODE. In **chapter 8**, we proposed a universal (one which works in variety of systems) power allocation algorithm, which while running independently and simultaneosly at all the base stations of the network, assymptotically satisfies the demands of all the users of the network. We obtained the analysis of the proposed algorithm via the ODE analysis. The same chapter also uses Game theoretic tools to obtain the demand satisfying power profile as a nash equilibrium of an appropriate game.

These analysis are used for obtaining the For our power allocation problem to satisfy user demands with base station power constraints i we have proposed a stochastic approximation based universal algorithm which can work in a variety of systems. This algorithm converges to a power profile, which is the zero of the function being addressed in this case. Also, specific to this problem, we have used a simple game theoretic framework [11] to formulate the problem and as has been a popular approach, we use an ordinary differential equation (ODE) [93] approach to analyze the algorithm.

For the purpose of simulations, we have used MATLAB and MAPLE extensively. Especially MAPLE is a favorite tool to check if explicit expressions are possible for seemingly difficult integrals and other mathematical functions. We further propose to use the LTE system level simulator [161] to validate some of the ideas and algorithms developed during the course of our thesis.

## 1.6 Organization and Contributions

This dissertation focuses on cell design and resource allocation for small cell networks. The chapters are organized in two parts. The first part deals with cell design, while the second part discusses resource allocation. Our main contributions and the outline of the chapters content are the following

Part A: Cell design and dimensioning

1. Cell dimensioning with static users: A fluid perspective [115].
2. Cell dimensioning with moving users: A Spatial queuing approach [89], [119], [118].
3. Where to locate the base station?: A large population perspective [116].
4. Capacity scaling and per-antenna user density in multi-antenna precoded networks: A random matrix approach [117].
5. BS activation control for green networking: A multimodularity approach [120].

Part B: Resource allocation

1. Multiscale Fairness and its Application to Resource Allocation in Wireless Networks [8], [10], [9], [114].
2. Satisfying Demands in a Multicellular Network: A Universal Power Allocation Algorithm [121].

In **Introduction 1**, we provided a brief overview of the generation of wireless networks, the need for small cells, an overview of the thesis, tools used to address problems in the thesis and chapter highlights.

In **Chapter 2**, we present a systematic study of the uplink capacity and coverage of pico-cell wireless networks. Both the one dimensional as well as the two dimensional cases are investigated. Our goal is to compute the size of pico-cells that maximizes the spatial throughput density. To achieve this goal, we consider fluid models that allow us to obtain explicit expressions for the interference and the total received power at a base station. We study the impact of various parameters on the performance: the path loss factor, the spatial reuse factor and the receiver structure (matched filter or multiuser detector).

In **Chapter 3**, we characterize the performance of Picocell networks in presence of moving users. We model various traffic types between base-stations and mobiles as different types of queues. We derive explicit expressions for expected waiting time, service time and drop/block probabilities for both fixed as well as random velocity of mobiles. We obtain (approximate) closed form expressions for optimal cell size when the velocity variations of the mobiles is small for both non-elastic as well as elastic traffic. We conclude from the study that, if the expected call duration is long enough, the optimal cell size depends mainly on the velocity profile of the mobiles, its mean and

variance. It is independent of the traffic type or duration of the calls. Further, for any fixed power of transmission, there exists a maximum velocity beyond which successful communication is not possible. This maximum possible velocity increases with the power of transmission. Also, for any given power, the optimal cell size increases when either the mean or the variance of the mobile velocity increases.

In **Chapter 4**, we address the problem of fair assignment of base station locations in a cellular network. We use the generalized $\alpha$-fairness criterion, which encompasses the different notions of fairness: that of global, proportional, harmonic or max-min fairness in our study. We derive explicit expression for $\alpha$-fair BS locations under 'large population' limits in the case of simple 1D models. We show analytically that as $\alpha$ increases asymptotically, the optimal location for a single BS converges to the center of the cell. We validate our analysis via numerical examples. We further study throughput achievable as a function of $\alpha$-fair BS placement, path-loss factor $\beta$ and noise variance $\sigma^2$ via numerical examples. We also briefly address the problem of optimal placement of two base stations and obtain similar conclusions.

In **Chapter 5**, we study precoded MIMO based small cell networks. We derive the theoretical sum-rate capacity, when multi-antenna base stations transmit precoded information to its multiple single-antenna users in the presence of inter-cell interference from neighboring cells. Due to an interference limited scenario, increasing the number of antennas at the base stations does not yield necessarily a linear increase of the capacity. We assess exactly the effect of multi-cell interference on the capacity gain for a given interference level. We use recent tools from random matrix theory to obtain the ergodic sum-rate capacity, as the number of antennas at the base station, number of users grow large. Simulations confirm the theoretical claims and also indicate that in most scenarios the asymptotic derivations applied to a finite number of users give good approximations of the actual ergodic sum-rate capacity.

In recent years there has been an increasing awareness that the deployment as well as utilization of new information technology may have some negative ecological impact. This includes awareness to energy consumption which could have negative consequences on the environment. In recent years, it was suggested to increase energy saving by deactivating base stations during periods in which the traffic is expected to be low. In **Chapter 6**, we study the optimal deactivation policies, using recent tools from Multimodularity (which is the analog concept of convexity in optimization over integers). We consider two scenarios: In the first case, a central control derives the optimal open loop policies so as to maximize the expected throughput of the system given that at least a certain percentage of Base stations are deactivated (switched OFF). In the second case, we derive optimal open loop polices, which each base station can employ in a decentralized manner to minimize the average buffer occupancy cost when the fraction of time for which the BS station is deactivated (idle mode) is lower bounded. In both the cases, we show that the cost structure is Multimodular and characterize the structure of optimal policies.

Fair resource allocation is usually studied in a static context, in which a fixed amount of resources is to be shared. In dynamic resource allocation one usually tries to assign

resources instantaneously so that the average share of each user is split fairly. The exact definition of the average share may depend on the application, as different applications may require averaging over different time periods or time scales. In **Chapter 7**, we study dynamic resource allocation in wireless networks. Our main contribution is to introduce new refined definitions of fairness that take into account the time over which one averages the performance measures. We examine how the constraints on the averaging durations impact the amount of resources that each user gets.

Power allocation to satisfy user demands in the presence of large number of interferers in a multicellular network is a challenging task. Further, the power to be allocated depends upon the system architecture, for example upon components like coding, modulation, transmit precoder, rate allocation algorithms, available knowledge of the interfering channels, etc. This calls for an algorithm via which each base station in the network can simultaneously allocate power to their respective users so as to meet their demands (when they are within the achievable limits), using whatever information is available of the other users. In **Chapter 8**, we propose one such algorithm which in fact is universal: the proposed algorithm works from a fully co-operative setting to almost no co-operation and or for any configuration of modulation, rate allocation, etc. schemes. The algorithm asymptotically satisfies the user demands, running simultaneously and independently within a given total power budget at each base station. Further, it requires minimal information to achieve this: every base station needs to know its own users demands, its total power constraint and the transmission rates allocated to its users in every time slot. We formulate the power allocation problem in a system specific game theoretic setting, define system specific capacity region and analyze the proposed algorithm using ordinary differential equation (ODE) framework. Simulations confirm the effectiveness of the proposed algorithm.

Summary and future research directions are discussed in **Conclusion 9**. A list of publications during the course of the thesis is available in **Publications 9**

# Part II

# Cell Design

# Chapter 2

# Cell Dimensioning with Static Users, a fluid perspective

**Contents**



## 2.1 Introduction

In a Small cell, the shorter transmission distance coupled with lower transmit power, enhances both capacity as well as the Signal to Interference Noise Ratio (SINR) achievable within the cell. But, a designer or a system architect would like to answer questions such as: What is the optimum number of cells that one would want to divide the macrocell?, What is the optimum cell size which maximizes the throughput achievable at a Small cell?. Does the receiver configuration matter? How is throughput affected when one moves from a deployment of Small cells on a street (1D) to a deployment in a office space or shopping mall (2D)?. What if the entire cell is located indoors? What happens if the Small cell BS is within the building or located outside? What is the implication of frequency reuse on the throughput achievable?

In this chapter, we try to address several of these questions. In particular, we derive explicit expressions for the up-link (UL) SINR and throughput for simple 1D and 2D

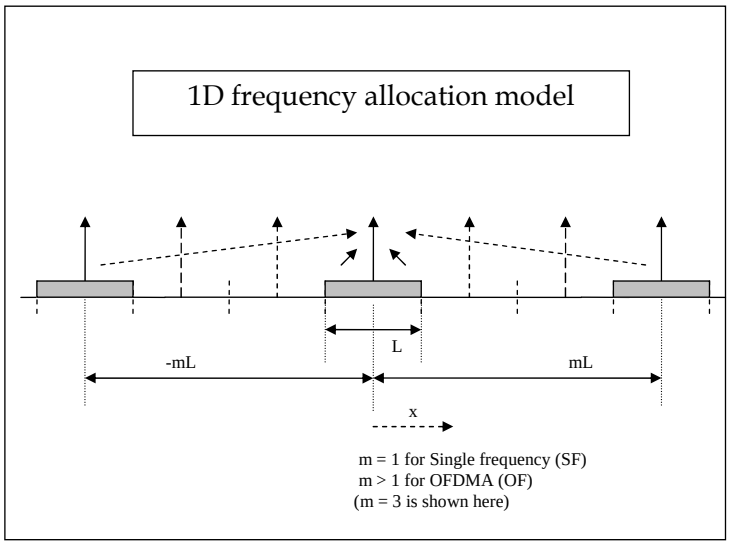


***Figure 2.1:*** *Frequency allocation in 1D.*

models and analyze the achievable throughput as a function of the cell size, coverage, etc. The goal of this work is to develop an analytical framework that can be used for preliminary dimensioning purposes in planning the Small cell network and thus provide an insight for answering the above questions. We derive closed form expressions of useful performance metrics considering free space path loss. We believe that the insight brought from our approach can be used as inputs for a more detailed model to include detailed propagation effects, fading, shadowing, mobility, etc.

We begin our study with the computation of the received power in single frequency and frequency reuse in Section 2.2. Next, in Section 2.3, we derive expressions for the throughput for both the modes with base-station (BS) receivers using the matched filter and multi-user detector. Using expressions derived in Section 2.3, we study the impact of cell size on the achievable throughput at the BS in Section 7.2.1 and propose a simple optimization criteria for both receiver configurations. In Section 2.5, we analyze the impact of cell size on the throughput in indoor scenarios. We derive expressions for the received power and throughput for a 2D model in Section 2.6 and use it to study the throughput as a function of cell size. Finally in Section 2.7, we look at the trade-off between capacity and coverage for the models presented in previous sections. We conclude our observations in Section 2.8.

### 2.1.1 Related work

We use a fluid model approach similar to the one used in [90], [69]. The simplicity of the fluid model approach eases laborious and time consuming simulations. Typically the cell structures are uniformly placed base-stations on a *line segment* (1D) or *hexagons* (2D). This model is simplistic. One can also use a Poisson-Voronoi model. Both models are discussed for example in [87]. The principal behind the hexagon approach is

to construct a disc with an equivalent area similar to the hexagon (OR hexagons of interest). While in the Poisson-Voronoi model, the base station locations and cell sizes are random. These two are extreme and complementary architectures: The hexagonal model represents perfectly structured networks, whereas the Poisson-Voronoi model takes into account irregularities of real networks in a statistical way. We shall treat in detail the hexagonal model in this work.

An alternative *fluid* model has been introduced in [31, 30] to study the performance and optimal cell size for CDMA networks. The model includes fading, and the fluid limits are those obtained when the density of mobiles become large. The main tool there has been random matrix theory. The approach in [31, 30] requires complete homogeneity within each cell: the gain from a mobile to a base station does not depend on the location of the mobile. In contrast, in our work, we take into account the detailed impact of the distance on the channel gains. In [135] the authors address a related problem of optimizing the spectral efficiency of cellular indoor wireless networks by adjusting the location and power of the base-stations. They apply both continuous and combinatorial approaches to find a solution to the optimization problem.

## 2.2 Received Power computation

We compute the total power received at the base station for single frequency and frequency reuse modes as a function of the cell size. The single frequency or frequency reuse modes can employ single carrier or OFDM modulation for transmission.

### 2.2.1 Single frequency (SF)

Assume a single frequency deployment. There is a uniform density of mobiles on the line $\mathcal{L} = ((-\infty, 0), (\infty, 0)$, each transmitting at a unit power (similar to the fluid model approach used in [90], [31]). There are base stations (BS) located at $(nL, 1)$, $n = ..., -1, 0, 1, ....$ Since we consider Small cells, we explicitly include in the geometric model the vertical distance (normalized to one) of the base station (which could be negligible in large cells). Let $\alpha$ represent the attenuation factor or path loss factor in the given wireless environment (for practical values of alpha one can refer for ex. to [123]). Thus the power received at the BS from a mobile at a distance of $x$ is equal to $(1 + x^2)^{-\alpha/2}$.

Denote $\beta_1 = (\alpha - 1)/2$ and $\beta_2 = (\alpha + 1)/2$. The total power received at the BS is

$$P_{bs}^{tot} = \int_{-\infty}^{\infty} (1 + x^2)^{-\alpha/2} \mathrm{d}x = \frac{\sqrt{\pi}\Gamma(\beta_1)}{\Gamma(\alpha/2)} \tag{2.1}$$

In particular, we have in Table 4.1 explicit expressions for some integer valued $\alpha$'s:

***Table 2.1:*** *total received power as a function of* $\alpha$

| $\alpha$ | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Total received power | $\pi$ | 2 | $\frac{\pi}{2}$ | $\frac{4}{3}$ | $\frac{3}{8}\pi$ |

**Power received at the BS in** $C_0$

Define the cell $C_0$ to be the segment $[-L/2, L/2]$ on the line $\mathcal{L}$. The total power received at the BS in $C_0$ from mobiles in $C_0$ is

$$\begin{aligned} P_{bs}^{C_0}(L) &= \int_{-L/2}^{L/2}(1+x^2)^{-\alpha/2}\mathrm{d}x \\ &= \frac{1}{\beta_1}\left(\frac{L}{2}\right)^{1-\alpha} hypergeom\left(\left[\frac{\alpha}{2},\beta_1\right],\beta_2,-\frac{4}{L^2}\right) - \frac{\pi^{3/2}sec\left(\frac{\pi\alpha}{2}\right)}{\Gamma(\frac{\alpha}{2})\Gamma\left(\frac{3-\alpha}{2}\right)} \end{aligned} \tag{2.2}$$

For the special case of $\alpha = 2$ this simplifies to

$$P_{bs}^{C_0}(L) = \pi + i\log\left(\frac{L+2i}{L-2i}\right) \tag{2.3}$$

Where $L$ is the cell size and $i = \sqrt{-1}$. In figure (2.2) we depict this case for $L$ taking the values from 0.01 to 10 for $\alpha = 2, 4$. We can see that $P_{bs}^{C_0}(L) \to P_{bs}^{tot}$ as L increases.

### 2.2.2 Frequency reuse (FR)

We consider some time slot and a given frequency and assume that this same frequency is not used at the same time slot in all cells: it is separated by $m-1$ cells. In figure (2.1), we present a typical frequency allocation for a one dimensional (1D) case. $m = 1$ is the single frequency case, whereas, any $m > 1$ represents frequency reuse. We show in the figure a typical reuse-3 case. i.e, every third cell uses the same frequency.

The total power received at the BS is

$$P_{bs}^{tot} = P_{bs}^{C_0}(L) + \sum_{i=-\infty, i\neq 0}^{\infty} P_{bs}^{C(im)}(L) \tag{2.4}$$

where $P_{bs}^{C(j)}(L)$ is the power received at the BS from cell $j$ that has size $L$. We note that

$$\begin{aligned} P_{bs}^{C(im)}(L) + P_{bs}^{C(-im)}(L) &= P_{bs}^{C_0}((2m+1)L) \\ &\quad - P_{bs}^{C_0}((2m-1)L) \end{aligned} \tag{2.5}$$

## 2.3 Throughput

We compute the total achievable throughput at a cell in the case where the BS receiver uses a matched filter or employs a multi-user detection scheme like successive interference cancellation.

### 2.3.1 Matched filter (MF)

We model the power received at BS in cell $C_0$ from a mobile at $x$ as the total power received from $[x, x+dx]$, i.e. $dP(x) = (1+x^2)^{-\alpha/2}\mathrm{d}x$. We use the Shannon capacity to compute the throughput while treating the interferences from all other mobiles at the same frequency and time as noise. Using a detection scheme based on the matched filter, the achievable throughput from the mobile at $x$ is

$$d\theta(x) = \log\left(1 + \frac{dP(x)}{\sigma^2 + P_{bs}^{tot}}\right)$$

where, $\sigma^2$ is the noise power.

Since the quantity $\frac{dP(x)}{P_{bs}^{tot}} << 1$, we use $\log(1+x) \approx x$ and rewrite

$$d\theta(x) = \frac{dP(x)}{\sigma^2 + P_{bs}^{tot}} \tag{2.6}$$

Hence the total throughput at the cell is given by

$$\Theta(L) = \frac{\int_{-L/2}^{L/2} dP(x)dx}{\sigma^2 + P_{bs}^{tot}} = \frac{P_{bs}^{C_0}(L)}{\sigma^2 + P_{bs}^{tot}}. \tag{2.7}$$

and the throughput density is given by

$$\Psi_{MF}(L) = \frac{\Theta_C}{L} = \frac{P_{bs}^{C_0}(L)/L}{\sigma^2 + P_{bs}^{tot}}. \tag{2.8}$$

Let $L^*$ denote the cell size which maximizes the throughput density.

**Lemma 2.3.2.** *In the case of the matched filter, the throughput density is maximized by taking base stations as dense as possible. i.e,* $L^* \to 0$.

**Proof** : From equation (2.8),

$$\begin{aligned} L^* &= \arg\max_L \frac{P_{bs}^{C_0}(L)/L}{(\sigma^2 + P_{bs}^{tot})} \\ &= \arg\max_L P_{bs}^{C_0}(L)/L \\ &= \arg\max_L \left(\frac{1}{L}\int_0^L (1+x^2)^{\frac{-\alpha}{2}} dx\right) \end{aligned}$$

as $\sigma^2$ and $P_{bs}^{tot}$ are independent of $L$. Clearly,

$$\int_0^L 1^{\frac{-\alpha}{2}} dx > \int_0^L (1+x^2)^{\frac{-\alpha}{2}} dx > \int_0^L (1+L^2)^{\frac{-\alpha}{2}} dx$$

and hence,

$$1 > \frac{P_{bs}^{C_0}(L)}{L} > (1+L^2)^{\frac{-\alpha}{2}}.$$

From the above, it is easy to see that,

$$\begin{aligned} \frac{P_{bs}^{C_0}(L)}{L} &< 1 \text{ if } L > 0, \\ \frac{P_{bs}^{C_0}(L)}{L} &\to 1 \text{ as } L \to 0. \end{aligned}$$

And thus, $L^* = 0$. ■

### 2.3.3 Multi-user detection (MD)

We assume that all the signal received at BS in cell $C_0$ from mobiles out of $C_0$ are considered noise; however within $C_0$, some multi-user detection scheme that maximizes the cell throughput is used. For example, successive interference cancellation is assumed. Then the SINR is given by

$$SINR = \frac{P_{bs}^{C_0}(L)}{\sigma^2 + P_{bs}^{tot} - P_{bs}^{C_0}(L)}. \tag{2.9}$$

and the throughput achievable using the Shannon capacity limit with the multi-user detection constraint [153]

$$\Theta(L) = \log(1 + SINR) \tag{2.10}$$

and the throughput density is

$$\Psi_{MD}(L) = \frac{\Theta}{L} = \frac{\log(1 + SINR_C)}{L} \tag{2.11}$$

**Lemma 2.3.4.** *In the case of multi-user detection, the cell size which maximizes throughput density is such that* $L^* > 0$

**Proof** : From equation (2.11),

$$L^* = \arg\max_L \frac{1}{L} \log\left(1 + \frac{P_{bs}^{C_0}(L)}{\sigma^2 + P_{bs}^{tot} - P_{bs}^{C_0}(L)}\right) \tag{2.12}$$

Let $\sigma^2 + P_{bs}^{tot} = K$. By monotonicity of log function,

$$L^* = \arg\max_L \left(\frac{K}{K - P_{bs}^{C_0}(L)}\right)^{\frac{1}{L}}.$$

Clearly,

$$L > P_{bs}^{C_0}(L) > L(1+L^2)^{\frac{-\alpha}{2}}$$

and hence,

$$\left(\frac{K}{K-L(1+L^2)^{\frac{-\alpha}{2}}}\right)^{\frac{1}{L}} > \left(\frac{K}{K-P_{BS}^{C_0}(L)}\right)^{\frac{1}{L}} > \left(\frac{K}{K-L}\right)^{\frac{1}{L}}$$

From the above, it is easy to see that

$$\begin{aligned}\left(\frac{K}{K-P_{BS}^{C_0}(L)}\right)^{\frac{1}{L}} &> 1 \text{ if } L>0,\\ \left(\frac{K}{K-P_{BS}^{C_0}(L)}\right)^{\frac{1}{L}} &\to 1 \text{ as } L\to 0.\end{aligned}$$

Thus there is an $L^* > 0$, which maximizes the throughput density. ■

**Computation of $L^*$:** By virtue of Lemma 2.3.4, $L^* > 0$, i.e., the maximizer is not at the boundary point and the objective function is clearly a differentiable function. Hence, the cell size which maximizes the throughput density is a zero of the derivative $d\Psi_{MD}/dL$ and hence is a zero of

$$\Omega(L^*;\alpha,\sigma^2) = 0 \tag{2.13}$$

where

$$\Omega(L;\alpha,\sigma^2) := L\left(1+\frac{L^2}{4}\right)^{\frac{-\alpha}{2}} - \left(\sigma^2 + P_{bs}^{tot} - P_{bs}^{C0}\right)\log\left(\frac{\sigma^2+P_{bs}^{tot}}{\sigma^2+P_{bs}^{tot}-P_{bs}^{C0}}\right)$$

Finding an explicit expression for $L^*$ is not easy. We outline a simple procedure to compute $L^*$ iteratively.

$$L_{k+1} = L_k + \epsilon\,\Omega(L_k;\alpha,\sigma^2) \tag{2.14}$$

$L_0 > 0$ is the initial value for the first iteration and $L^*$ is the converged value. One can confirm the above iteration has converged whenever the error $|L_{k+1} - L_k| < \mu$ for some small enough positive constant $\mu$.

For example, with $\alpha = 2$, $\sigma^2 = 1$, $L_0 = 0.1$ and $\epsilon = 0.1$, we converge to $L^* = 0.85$ in 150 iterations. This matches with the optimal (max) value for the curve labeled $SF\_MD$ in figure (2.5).

**Asymptotic Approximations for $L^*$ in Small cells:** We derive approximations for $L^*$ as the path-loss factor $\alpha$ converges to 0 or to $\infty$. We recall from equation (2.11),

$$L^* = \arg\max_{L\in[0,L_{max}]} \frac{1}{L}\log\left(1+\frac{P_{bs}^{C_0}(L)}{\sigma^2+P_{bs}^{tot}-P_{bs}^{C_0}(L)}\right).$$

In the above, $L_{max}$ represents the maximum cell size that we can design. Since we are dealing with Small cells $L_{max}$ itself is a small number and hence $P_{bs}^{C_0}(L) \ll P_{bs}^{tot}$. So we can use the approximation $log(1+x) \approx x$ and re-write

$$L^* = \arg \max_{L \in [0, L_{max}]} \frac{1}{L} \left( \frac{P_{bs}^{C_0}(L)}{K_\alpha - P_{bs}^{C_0}(L)} \right),$$

where $K_\alpha := \sigma^2 + P_{bs}^{tot}$.

**For large values of $\alpha$ :** By bounded convergence theorem as $\alpha \to \infty$, $P_{bs}^{C_0}(L) - L(1+L^2)^{-\alpha/2} \to 0$. Therefore for large values of $\alpha$,

$$L^* \approx \arg \max_L \frac{1}{L} \left( \frac{L(1+L^2)^{-\alpha/2}}{K_\alpha - L(1+L^2)^{-\alpha/2}} \right)$$

or $L^*$ approximately solves

$$\left. \frac{d\left( \frac{(1+L^2)^{-\alpha/2}}{K_\alpha - L(1+L^2)^{-\alpha/2}} \right)}{dL} \right|_{L=L^*} = 0.$$

Solving this yields

$$L^* \approx \frac{1}{\alpha K_\alpha} \text{ for large values of } \alpha. \tag{2.15}$$

**For small values of $\alpha$ :** By bounded convergence theorem again as $\alpha \to 0$, $P_{bs}^{C_0}(L) \to L$. Therefore, $L^* \approx \arg \max_L \left( \frac{1}{K_\alpha - L} \right)$ for small values of $\alpha$. Hence

$$L^* \approx L_{max} \text{ for small values of } \alpha. \tag{2.16}$$

In figure (2.9), we plot $L^*$ as a function of $\alpha$ computed using equations (2.15), (2.16) and compare it with the value of $L^*$ computed via numerical simulations. We see that the numerically evaluated values lie within the two bounds. Also, for smaller values of $\alpha$, the simulation results are closer to the $\alpha$ small bound and for larger values, the simulation results are closer to the $\alpha$ large bound.

### 2.3.5 Comments on the fluid approach

So far we have been working with a fluid model which had the advantage (over a more detailed discrete stochastic model) of being sufficiently simple to allow us to obtain explicit expressions for performance measures related to the system capacity. We now address the question of the validity of the fluid approximation: can we estimate what these expressions say on the original discrete system? To address the question we first define our stochastic model. We assume a Poisson arrival process with constant intensity.

Assume that mobiles are located according to a uniform Poisson point process $X = \{X_i\}$ with intensity $\lambda$ which we normalize to one unit. $N_x(A)$ denotes the number of points of the process $X$ in a set $A$. Consider first the single frequency setting of Section 2.2.1. Then the total power received at the BS is given by $\sum_{i=-\infty}^{\infty} (1 + X_i^2)^{-\alpha/2}$ and its expectation equals precisely to $\int_{-\infty}^{\infty} (1 + x^2)^{-\alpha/2} \mathrm{d}x$, for which equation (2.1) provides an explicit expression (for integer $\alpha$'s).

The power and the interferences received at the BS according to the fluid approximation are thus precisely the expectation of those corresponding to the discrete model. In particular, the remaining equations of Section 2.2.1 also hold for the expectations of the corresponding objects in the discrete model.

The total throughput of the cell is then given by equation (2.7) and its expectation satisfies

$$E[\Theta(L)] = E[P_{bs}^{C_0}(L)]E\left[\frac{1}{\sigma^2 + P_{bs}^{tot}}\right] \geq \frac{E[P_{bs}^{C_0}(L)]}{\sigma^2 + E[P_{bs}^{tot}]}$$

where the first equality follows from the independence properties due to the Poisson assumption on the location of mobiles, and the last step follows by Jensen's inequality. We conclude that the throughput results of the fluid model provide lower bounds to those of the discrete model.

## 2.4 Impact of cell size on throughput

In this section, we use the expressions obtained in the previous sections to study the impact of the cell size on throughput. This analysis will allow us to optimize the cell size for the different modes and receiver schemes considered. We perform the throughput analysis via some numerical examples.

### 2.4.1 Numerical results

In figure (2.4), we plot the total achievable throughput in $C_0$ as a function of the cell size $L$ which is obtained using equation (2.7), (2.10). We consider two cases, the first case when all the cells deployed use the same frequency and the second case where we use a reuse factor of 3. i.e, every third cell uses the same frequency. We compute the throughput for the matched filter as well as the multi-user detection scheme in both single frequency and frequency reuse. We plot the throughput as a function of $L$ in figure (2.4) for $\alpha = 2$. As expected, the throughput increases with $L$ and multi-user detection performs better than the matched filter.

Next, we look at the total throughput density achievable in $C_0$ as a function of $L$. For the single frequency case, when the BS receiver uses a matched filter (SF_MF), we note that the numerator in the throughput density (equation (2.8)) depends on $L$; it

represents the useful power density received. By Lemma 1, we know that the optimal cell size that maximizes the throughput density, $L^* \to 0$. Figure (2.3) plots the power density as a function of $L$. We observe that the power density decreases with $L$ and is maximum when $L^* \to 0$. Thus we can conclude that the *throughput density is maximized by taking base stations as dense as possible* for a BS receiver which uses a matched filter. With frequency reuse, the throughput density also depends upon $L$ via the total received power (see equations (2.4), (2.8)). Hence, we directly compute the throughput density in this case.

As before, we consider the two cases described previously. We plot the throughput density as a function of $L$ in figure (2.5), for $\alpha = 2$. We also plot the throughput density for a single frequency with a matched filter (SF_MF), frequency reuse with a matched filter (FR_MF), single frequency with multi-user detection (SF_MD) and frequency reuse with multi-user detection (FR_MD) in the same figure.

By Lemma 2, we would expect that the cell size which maximizes the throughput density, $L^* > 0$, for the BS receiver which employs multi-user detection. From the numerical example we conducted (figure (2.5) ), we indeed see that $L^* > 0$ and in fact, there is a unique optimal $L$ for a certain frequency reuse which maximizes the throughput density. The effect is more pronounced in the frequency reuse case. Also, we achieve a higher throughput density for smaller cell sizes for the more practical case of $\alpha = 4$ in contrast to $\alpha = 2$.

### 2.4.2 Optimizing the cell size

**Matched filter:** We see from figure (2.5) that in the matched filter case, the throughput density is maximized for smaller cell sizes. However there is a cost for deploying base stations. So it is more natural to pose the problem of maximizing

$$J(L) \quad = \quad \frac{1}{L}(c_1\Theta(L) - c_2) \tag{2.17}$$

where $c_1/L$ is the revenue per throughput density and $c_2$ is the cost of each BS.

We note from figure (2.10)

that the throughput density is not sensitive at all to the reuse parameter $m$ in the matched filter case. (*the cell size, $L^*$, which maximizes the throughput density is always equal to the smallest $L$ considered in the numerical analysis*). Also, note that the optimization is over the integers and $L$ is typically not larger than a few hundreds. This allows one to solve the optimization problem in a very short time even if exhaustive search is used.

**Multi-user detection:** For a BS which employs multi-user detection, we observe from numerical analysis (see figure (2.10)) that $L^*$ decreases as the reuse factor $m$ increases. So, as the reuse factor increases, one would prefer to reduce the cell size in order to maximize throughput density (densify base stations). However, this increases deployment costs. Hence, our problem is to find an optimal cell size which balances the throughput and deployment costs over the reuse factors of interest. Assuming that

the deployment cost is proportional to the frequency reuse factor $m$, we can formulate this problem as:

$$J(L,m) \quad = \quad \frac{1}{L}(c_1\Theta(L) - c_2 - mc_3) \tag{2.18}$$

as before, $c_1/L$ is the revenue per throughput density and $c_2$ is the cost of each BS. The additional cost $mc_3$ is incurred due to reuse co-ordination where $m$ is the reuse factor and $c_3$ is the cost per reuse.

## 2.5 Indoor analysis

### 2.5.1 BS located inside the building

Next, we consider the cell of interest to be enclosed within a wall. All the interferers are located outside of the wall. We use our earlier received power and throughput computation to analyze the performance of the desired cell for both the matched filter as well as the multi-user detection schemes. Here again, we consider single frequency as well as frequency reuse, first in the 1D case. We add to the gain, a constant multiplication term to take into account the attenuation due to penetration through walls (refer [131], [1], [168] for some example values).

Thus the throughput density equations (2.8) and (2.11) change as

$$\Psi_{MF}(L) = \frac{P_{bs}^{C_0}(L)/L}{\sigma^2 + \eta P_{bs}^{tot} + (1-\eta)P_{bs}^{C_0}(L)}.$$

and

$$\Psi_{MD}(L) = \frac{1}{L}\log\left(1 + \frac{P_{bs}^{C_0}(L)}{\sigma^2 + \eta(P_{bs}^{tot} - P_{bs}^{C_0}(L))}\right)$$

, where $\eta$ is the wall attenuation factor .

We plot the numerical results in figure (2.7) for attenuation $\eta = 12dB$

**Observations:**

- When the interference is attenuated, there does not exist an optimal cell size which maximizes the throughput density for both the matched filter as well as multi-user detection schemes in single frequency and frequency reuse respectively.
- Throughput density increases with path-loss $\alpha$.
- Throughput density increases with attenuation $\eta$.

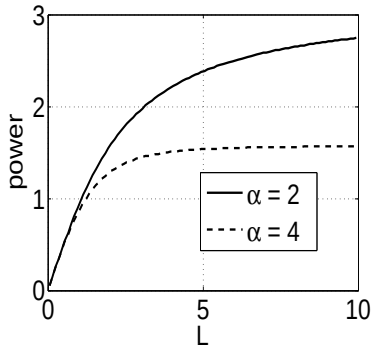


***Figure 2.2:*** *Total power received at the BS in $C_0$ vs L ($\alpha = 2, 4$; single frequency).*

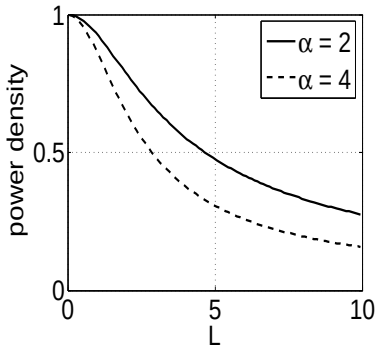


***Figure 2.3:*** *Total power density from $C_0$ vs L ($\alpha = 2, 4$; single frequency, matched filter).*

### 2.5.2 BS located outside the building

Another interesting case is the one where we assume that the serving BS is located outside of the wall, but geometrically within the cell (example, BS mounted on top of the roof, served mobiles inside the building (cell)). The throughput density equations (2.8) and (2.11) change as $\Psi_{MF}(L) = \dfrac{\eta P_{bs}^{C_0}(L)/L}{\sigma^2 + P_{bs}^{tot} - (1-\eta)P_{bs}^{C_0}(L)}$. and $\Psi_{MD}(L) = \dfrac{1}{L}\log\left(1 + \dfrac{\eta P_{bs}^{C_0}(L)}{\sigma^2 + P_{bs}^{tot} - P_{bs}^{C_0}(L)}\right)$.

We plot the throughput density as a function of the cell size in figure (2.8) for 12 dB attenuation.

**Observations:**

- For both the matched filter and multi-user detection, the maximum throughput is achieved for a certain optimal cell size.
- As expected, the achievable throughput density increases with path-loss $\alpha$.
- The optimal cell size decreases as path-loss $\alpha$ increases.
- The advantage of multi-user detection over the matched filter diminishes as the attenuation $\eta$ increases.
- The optimal cell size increases as the attenuation $\eta$ increases.

**Remarks:** From the above two cases, it is interesting to note that the placement of the BS matters, when the cell of interest is located indoors.

## 2.6 Dimension 2

In this section, we want to compute the optimal cell sizes in two dimension (2D). We begin our study with the power and throughput computation similar to the 1D case.

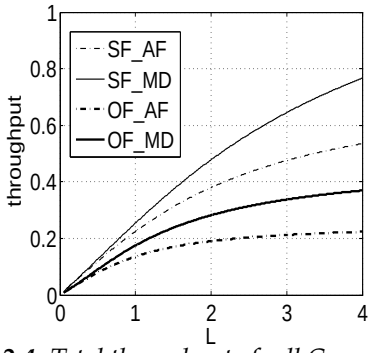


***Figure 2.4:*** *Total throughput of cell $C_0$ vs $L$ ($\alpha = 2$).*

- The total power received for $\alpha > 2$,

$$P_{bs}^{tot} = \int_0^{\infty} 2\pi x (1+x^2)^{-\alpha/2} \mathrm{d}x = \frac{2\pi}{\alpha - 2}$$

. Whereas, for $\alpha \leq 2$ the total received power is infinite.

- Consider a cell $C_0$ centered at the origin with radius $L$. The power received at the origin from mobiles within the cell is

$$P_{bs}^{C0} = \int_0^{L} 2\pi x (1+x^2)^{-\alpha/2} \mathrm{d}x = \frac{2\pi}{\alpha - 2}\left(1 - (1+L^2)^{1-\alpha/2}\right)$$

- The total throughput achievable at a cell when decoding each mobile using the matched filter (considering all the rest as noise) is

$$Thp_C^{adapt-filter} = \frac{1 - (1+L^2)^{1-\alpha/2}}{\sigma^2 \frac{2-\alpha}{2\pi} + 1}$$

- The total throughput achievable at a cell when multi-user detection is used (considering all the mobiles out of the cell as noise) is

$$Thp_C^{Mult-Acc} = \log\left(1 + \frac{1 - (1+L^2)^{1-\alpha/2})}{\sigma^2 \frac{2-\alpha}{2\pi} + (1+L^2)^{1-\alpha/2}}\right).$$

For the two dimensional case, we consider frequency reuse only.

### 2.6.1 A simple approximation to the hexagonal grid

In this case, we approximate the hexagon cells with virtual circles (similar to the examples used in [31], [100] or [69]). Given that we consider a circle with radius $L$, we can construct a hexagon with side $L'$, such that the area of the hexagon and the circle are the

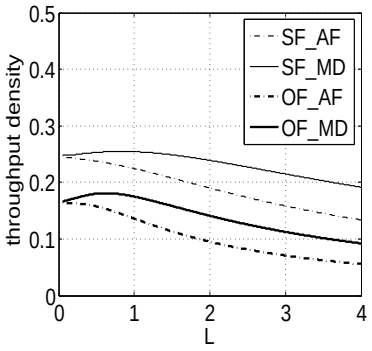


***Figure 2.5:*** *Throughput density vs L ($\alpha = 2$).*

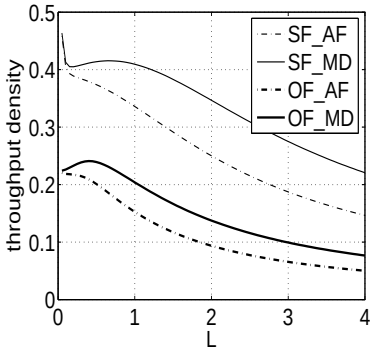


***Figure 2.6:*** *Throughput density vs L ($\alpha = 4$).*

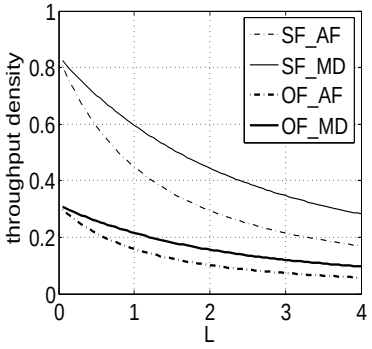


***Figure 2.7:*** *Throughput density vs L ($\alpha = 2$, wall attenuation $12dB$, BS indoors).*

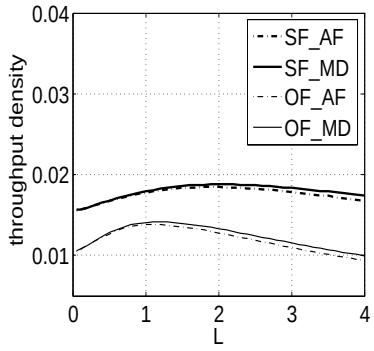


***Figure 2.8:*** *Throughput density vs L ($\alpha = 2$, wall attenuation $12dB$, BS outdoors).*

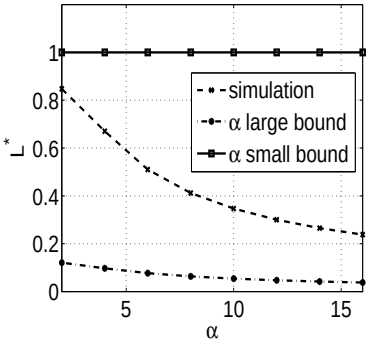


***Figure 2.9:*** *Optimal cell size $L^*$ vs path-loss factor $\alpha$ (reuse factor $m = 1$)*

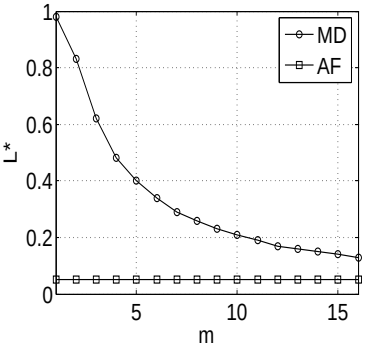


***Figure 2.10:*** *Optimal cell size $L^*$ vs reuse factor $m$ (path-loss factor $\alpha = 2$).*

same [87]. One can easily see that $L' = \left(\sqrt{\frac{2\pi}{3\sqrt{3}}}\right)L$. *For example*, if $L = 1$, $L' = 1.0996L$. Also, the number of hexagons grow by 6 for each tier.

The infinite hexagonal grid representation is shown in figure (2.11). The inner most hexagon is the cell of interest, where we desire to compute the throughput density. We consider a reuse 4 scenario. The subsequent hexagonal cells surrounding the cell of interest has the frequency allocation as shown in figure (2.11). As we see, every alternate tier repeats the frequency of the inner-most cell, thus contributing to interference to the cell of interest. We can also observe that the amount of interference is 1/2 of the total contribution from this tier of hexagons.

The interference power from this annular ring is given by

$$P_{bs}^{C(j)} = \frac{1}{m}\frac{1}{2}\frac{2\pi}{\alpha - 2}\Big\{\Big((1 + (a_j L)^2)^{1-\alpha/2}\Big) - \Big((1 + (b_j L)^2)^{1-\alpha/2}\Big)\Big\} \tag{2.19}$$

where $a_j = \sqrt{(2j)6 + 1}$, $b_j = \sqrt{(2j-1)6 + 1}$ and $j$ takes the value 1, 2, 3 ... representing the power from the $j^{th}$ odd hexagonal tier. Note that $m$ is the reuse factor.

### 2.6.2 A more precise approximation for the 2-D hexagonal grid

In this section, we propose an alternate approach to derive a more precise approximation for the 2-D hexagonal grid and compare it with the previously used approximation. In figure (2.11), every tier of hexagons around the cell of interest is enclosed by an hexagon (shown dotted), whose side $L'' = (\frac{2}{\sqrt{3}} + N\sqrt{3})L'$. Where, N is the index of the hexagonal tier. As before, we replace these hexagons with annular rings, where, the radius of the annular rings is related to the enclosed hexagons by $L'' = \left(\sqrt{\frac{2\pi}{3\sqrt{3}}}\right)L$. Using these relationships, we can see that the radii of the equivalent annular ring's for this model grows as $\frac{1}{\sqrt{3}}(2r, 5r, 8r, ...$ $(MethodB)$ as compared to $r, \sqrt{7}r, \sqrt{13}r, \sqrt{19}r...$ $(MethodA)$ in the previous case.

The numerical results are shown in figure (2.12). We conclude from this new approximation for the choice of radii in the earlier case *Method A* was conservative. Thus using *Method B* would result in a lower throughput density.

### 2.6.3 Throughput density with reuse

In this section, we would like to compare how the throughput density changes as as function of the reuse factor. We use the 2-D hexagonal model proposed (*Method B*) in

***Table 2.2:*** *Interference contribution for different reuse factors.*

| *tier no.* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| reuse 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| reuse 2 | 1/3 | 2/3 | 1/3 | 2/3 | 1/3 | 2/3 | 1/3 | 2/3 | 1/3 | 2/3 |
| reuse 3 | 0 | 1/2 | 1/3 | 1/4 | 2/5 | 2/6 | 2/7 | 3/8 | 3/9 | 3/10 |
| reuse 4 | 0 | 1/2 | 0 | 1/2 | 0 | 1/2 | 0 | 1/2 | 0 | 1/2 |
| reuse 7 | 0 | 0 | 1/3 | 0 | 1/5 | 1/6 | 1/7 | 1/8 | 1/9 | 2/10 |

our numerical simulations. For this study, we construct hexagonal grid with 10-tiers around the cell of interest. This would encompass 331 cells in total. So we would be considering interference from a possible 330 cells surrounding our cell of interest. For the numerical analysis, we use reuse 1, 2, 3, 4 and 7.

The amount of interference contributed from each tier for different reuse factors are listed in the table 4.4.

Now equation (2.19) is modified to accommodate the reuse factor, $m$ and re-written as shown

$$P_{bs}^{C(j)} = \frac{1}{m} c_j \frac{2\pi}{\alpha - 2} \left\{ \left( (1 + (a_j L)^2)^{1-\alpha/2} \right) - \left( (1 + (b_j L)^2)^{1-\alpha/2} \right) \right\} \tag{2.20}$$

now, $a_j = \sqrt{(3j) + 2}$, $b_j = \sqrt{(3(j-1) + 2}$, to make it generic to accommodate all possible reuse factors and $j$, the tier number goes from $1, 2, .., 10$. $c_j$ is the interference contribution from the $j^{th}$ tier (ex. from table 4.4, for reuse 4, the interference contribution is $1/2$ for every odd tier). We use equation (2.20), to compute the throughput density.

We see from figure (2.13) that the throughput density increases with the reuse factor. The other interesting point to note is that the throughput density for higher reuse does not fall rapidly as compared to the lower reuse factors.

### 2.6.4 Indoor analysis

For our next analysis, we assume that the cell of interest is located indoors and that the walls offer an attenuation of $\eta = 12dB$ to the interferer's. figure (2.14) captures this case for all reuse factors. We would also like to know what is the benefit or gain in throughput density when the interferer's are attenuated. This can be captured by including the case of reuse factor 4 from the previous numerical analysis. We see that for smaller cell sizes, this is almost a factor of 2 (see the region ellipse in the figure). But, one loses this advantage as the cell sizes tend to increase. Finally, we look at the case of the cell of interest located indoors, but the BS located outside, say, mounted on top of the roof. Now the signals from mobiles inside the building are attenuated by $\eta = 12dB$. figure (2.15) captures this analysis. The reduction in throughput density in such a case is very drastic as one can observe for the case of reuse factor 4. The benefit of reuse hardly seems to help in such situations (see the ellipse in the figure). In conclusion, we

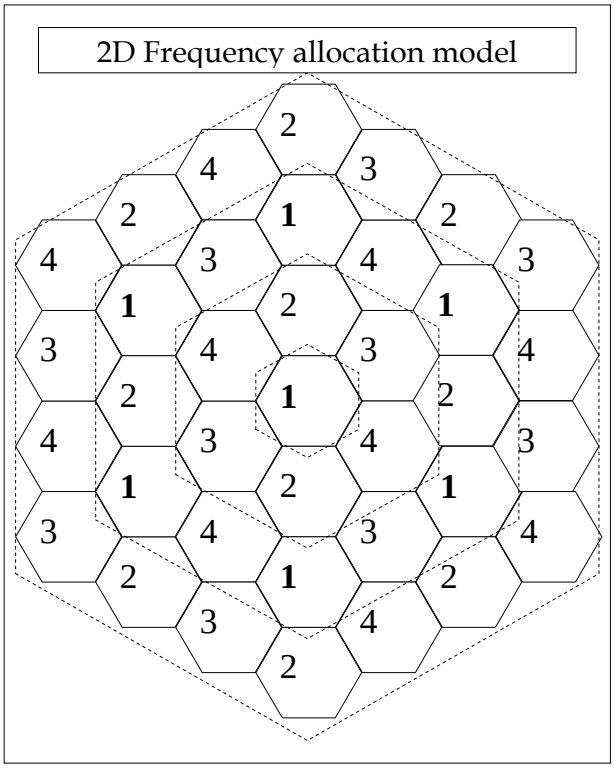


***Figure 2.11:*** *Frequency allocation in 2D.*

see that the gain in throughput density when the BS is located indoors is more than compensated when one moves it outside. This would mean that one would need to plan appropriate BS placement, depending on the nature of user's demand.

## 2.7 Coverage and capacity

Next, we want to study the trade-off between capacity and coverage. We assume that some portion towards the periphery of the cell is not covered by the BS and hence these mobiles are switched off (in power saving mode). We want to look at the *throughput per mobile* as a function of *coverage* and the *total capacity* achievable at the cell, again as a function of *coverage*.

### 2.7.1 Coverage and capacity in a single cell

We assume power control. i.e, mobiles at the boundary of the coverage area (distance $x$ from BS) transmit with power $P$ and the received power at the BS is

$$p = P(1 + x^2)^{-\alpha/2}. \tag{2.21}$$

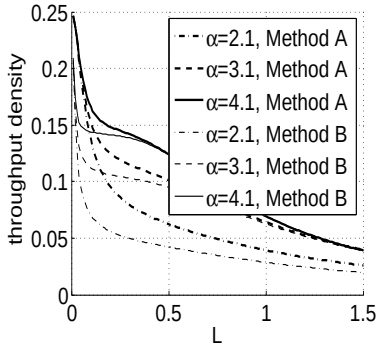


***Figure 2.12:*** *Throughput density vs L for different α and decoding schemes comparing MethodA with MethodB.*

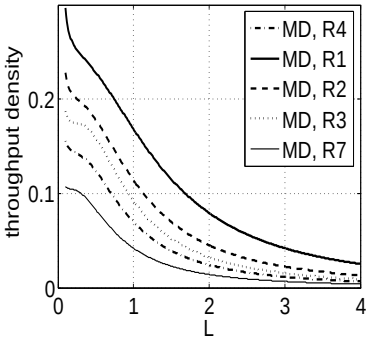


***Figure 2.13:*** *Throughput density vs L for $\alpha = 4.1$ and different reuse factors.*

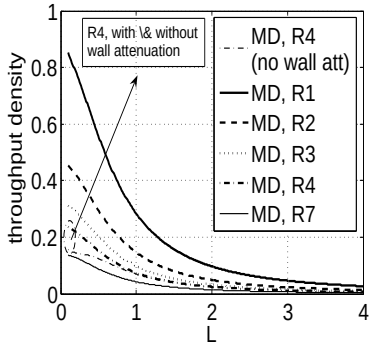


***Figure 2.14:*** *Throughput density vs L for $\alpha = 4.1$ and different reuse factors (BS indoors).*

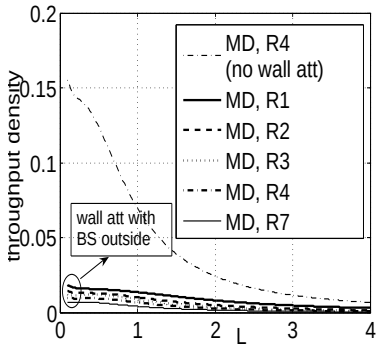


***Figure 2.15:*** *Throughput density vs L for $\alpha = 4.1$ and different reuse factors (indoor cell, BS outdoors).*

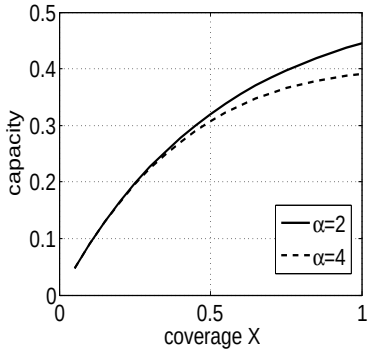


***Figure 2.16:*** *Capacity vs cell coverage x for $\alpha = 2, 4$.*

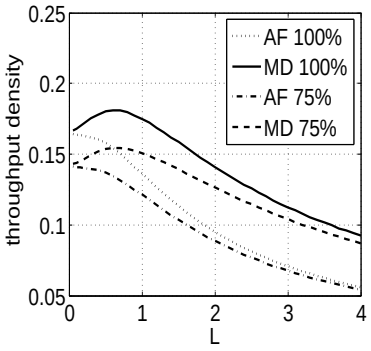


***Figure 2.17:*** *Throughput density vs L for 100 % and 75 % cell coverage ($\alpha = 2$, 1D).*

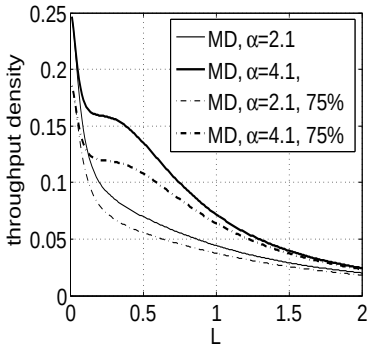


***Figure 2.18:*** *Throughput density vs L for 100 % and 75 % cell coverage ($\alpha = 4.1$, multi-user detection, 2D).*

Since we assume power control, any mobile within this coverage area will transmit with a power $P'$ lesser than $P$, such that the received power at the BS will always satisfy equation (2.21). The throughput per mobile can be computed as

$$d\theta(x) = \frac{p}{1 + \int_0^x p dx} = \frac{p}{1 + px} \tag{2.22}$$

while, the capacity of the cell as a function of coverage $x$ can be computed as

$$C(x) = x d\theta(x) = \frac{px}{1 + px} \tag{2.23}$$

We observe that (figure (2.16)) the capacity of the cell increases with increase in coverage.

### 2.7.2 Coverage and capacity on a line segment (1D)

Next, we want to extend the argument for the entire line segment. For the ease of analysis, we assume no power control and as before, the mobiles, which are not in the coverage area are in a power-down state. In this context, we want to analyze how the throughput density changes as a function of coverage for both multi-user detection (MD) and the matched filter (MF) case. We consider frequency reuse of 3.

We use equations (2.2), (2.3), (2.4) and (2.5) and account for the coverage by replacing $L$ with $L + \delta L$ or $L - \delta L$ appropriately. In our numerical examples, we compare full coverage with 75 % coverage.

Figure (2.17) shows the throughput density. We observe that for the multi-user detection scheme, the full as well as the 75 % coverage exhibits certain cell size for maximum throughput density, while in the matched filter case, the throughput density is maximized by densifying the cells. Hence, we conclude that coverage does not alter

the behavior of achievable maximum throughput, albeit at a different cell size for the multi-user detection case.

We observe that for small cell sizes, the throughput density (refer to figure (2.17)) is proportional to the coverage. But, as cell size increases, the throughput density for different coverage tends to converge.

### 2.7.3 Coverage and capacity in two dimension

Next, we look at coverage and capacity in 2-dimension. As before, we have an infinite hexagonal grid and our analysis assumes interference coming from hexagonal rings surrounding the cell of interest. Here again, we consider full and 75 % cell coverage and compute throughput density for different values of $\alpha$

The following are note-worthy observations (see figure (2.18)).

1. The achievable throughput density falls very sharply in the vicinity of a small cell radii, irrespective of the coverage.

2. We achieve a maximum throughput density proportional to the coverage. i.e for example, the MD scheme with 75 % coverage achieves about 75 % of maximum throughput density of the full coverage case at small cell radii. But, as the cell size increases, the achieve throughput density with 75 % coverage starts moving closer to the full coverage case.

## 2.8 Conclusions and future perspectives

We study aspects of cell dimensioning with static users under fluid limits. We characterize the throughput achievable at the BS as a function of the cell size. The aim of this study is to investigate if there exists a cell size which maximizes the achievable throughput. We derived explicit expressions for power and throughput for both single frequency and frequency reuse with different receiver configurations and numerically analyze the impact of the cell size on throughput for various 1D, 2D models in both indoor and outdoor scenarios. Our first cut analysis used free space path loss and we did not consider the impact of shadowing, fading, etc. The analysis shows that the throughput achievable is not always maximized by densifying the base-stations (BS), but rather depends on a case to case basis on factors like deployment (1D, 2D, indoor, outdoor), frequency reuse, etc. Our initial analysis did not incorporate radio fading, shadowing effects, mobility of users, etc. It would be interesting to study the behavior of throughput achieved as a function of cell size with these effects. Further, we assumed continuum of users, which would enable us to use fluid limits and thus obtain simple explicit expressions and an initial insight into cell design. A more realistic situation is to assume users at discrete points or users distributed randomly according to some distribution. Properties of such distributions can be used to analyze and obtain

dimensioning rules. It would be interesting to compare these results in the limiting regime.

## 2.9 Publications

1. Sreenath Ramanath, Eitan Altman, Vinod Kumar, Merouane Debbah, "Optimizing cell size in pico-cell networks", proceedings of the Workshop on Resource Allocation in Wireless Networks (RAWNET'09), Jun 27, Seoul, South Korea.

# Chapter 3

# Spatial Queueing Analysis for Design and Dimensioning of Small Cell Networks with Mobile Users

**Contents**



## 3.1 Introduction

In this chapter, we study cell design and dimensioning with mobile users.

To prevent a large number of handovers that would result from the small size of the cells ([127, 4]), it has been proposed to group together a number of Small cells in one virtual Macrocell and to restrict the effort of preventing losses due to the handover only to those handovers that occur between Small cells of the same virtual cell. In between the Small cells some fast switching mechanisms are proposed such as frequency following mechanism where the frequency used by a mobile follows it from one Small cell to the next. This requires reserving the same channel for a user in the entire Macrocell.

In this chapter, we consider a large Macrocell divided into a number of Small cells and study the impact of mobility on such systems, especially the effect of frequent handovers. We assume that the ongoing call is never dropped at the Small cell boundary, however base station switching (BSS) at any Small cell boundary requires some fixed amount of information (in terms of bytes) to be exchanged. There is however a possibility of calls being dropped at Macrocell boundaries. We further assume that the active users cross Macrocell boundaries at maximum once, i.e., the calls always end before reaching the second Macro boundary. The handovers at the Macrocell boundaries are modeled as independent Poisson process (this is a commonly made assumption, for example see [108, 53]).

We have several goals. First, to model the system so as to predict its performance measures. We are thus interested in developing tools using spatial queuing , that take into account not only the instantaneous geometry but also the way it varies in time. It should thus account for the impact of the speed of the users. We model the Macrocells by various types of queues and well known results from queuing theory (for example [164]) are used to obtain performance measures like expected waiting time , expected service time , drop or blocking probabilities , etc. We shall use these results for preliminary dimensioning purposes in planning the Small cell network catering to pedestrian and vehicular mobility, typical of urban and sub-urban areas. We derive closed form expressions of useful performance metrics considering free space path loss, handover constraints, traffic type etc. We also obtain closed form expressions for optimal cell size, optimal for various performance metrics, when all the users move at the same fixed velocity. To derive these performance measures, we would require the *moments* of the time taken by the system to serve[1] the customers, which in our case will equivalently be the time the Macrocell spends on a call. We derive the expressions for these service times, during which the information is exchanged between the moving user and the set of appropriate base stations (which it encounters during its journey), using variable rate of transmission. We further simplify the expressions for service time under the following assumption: in a typical Small cell network, a moving user would have traversed across a number of cells before the completion of call. In this system, arrivals occur in space and the service time depend on the position, movement of the user and the serving base station(s). The queues modeling these systems are referred to as spatial queues and have been used in interesting applications ([170, 88, 125, 18]). We make the following theoretical and or simulation based observations:

1) Maximum possible velocity: For any fixed power of transmission $P$, there exists a maximum user velocity $V_{lim}(P)$ and the users moving at speeds greater than $V_{lim}$ can not successfully communicate with the BS;

2) Larger cells for larger velocities: Given $P$, the optimal cell size increases with an increase in the highest velocity that the system has to support. This is true as long as the highest velocity is less than $V_{lim}(P)$. However the system cannot cater for velocities above the limit $V_{lim}(P)$, even if one increases the cell size indefinitely;

3) Insensitivity to application: The optimal cell size remains the same for non-elastic

[1] Throughout we use the terms from queuing theory like arrivals, service time etc.

(NES) as well as elastic (ES) calls for large file sizes, when the rest of the parameters remain same;

4) Two dimensional Manhattan Grid: The one dimensional results can be applied directly to a two dimensional regular street grid (for example [103, 14]).

We extended the analysis to systems, where (negligible number of ) call drops can also occur at Small cell boundaries, and obtained some initial results. To completely avoid call drops at Small cell boundaries, one needs a centralized call admission control (a control based on the total number of calls in the entire Macrocell). This requirement can be relaxed if one can design a system delivering required QoS, in spite of (negligible) call drops at Small cell boundaries. We consider one such system (with decentralized call admission control) and obtain closed form expressions for the optimal cell size catering to non elastic traffic.

In [118], we extended the analysis to a hexagonal 2D cellular structure in which the users can move in any direction across the cell. The 2D system also considers possible drops at Small cell boundaries as well as works with a more distributed call management system.

Several authors have examined the impact of mobility on the performance of wireless networks. The authors in [72], have shown that mobility in fact increases the capacity in ad hoc networks. In [21], the authors discuss the trade off between delay and achievable throughput in the presence of mobility in wireless ad-hoc networks. Further, In [28, 29, 36, 35], the authors discuss the impact of inter and intra-cell mobility on capacity, flow level performance, trade-off between throughput and fairness, etc. They show that mobility tends to increase the capacity with globally optimal as well as fair sharing policies, when the base stations interact. In conclusion, we see that the mobility (via multiuser diversity, opportunistic scheduling etc) can in fact improve the overall performance of the system. However, most of the work assume that the handovers occurs without extra cost. But in reality, each handover requires exchange of some information between the user and the new BS and has a risk of not finding free resources in the new cell. Our work mainly focuses on these issues which become significant whenever the frequency of handovers increases, as in the case of Small cell networks.

We describe our system in Section 3.2 while the service time is discussed in Subsection 3.3.1 and is optimized in 3.3.6. The NES, ES calls are modeled by appropriate queuing models and performance measures are derived in Subsections 3.3.8, 3.3.9 respectively. The two dimensional regular street grid is studied in Section 3.4 while the numerical examples are provided in Section 3.5. Some initial results on a system with possible call drops at Small cell boundaries and with a decentralized call admission control are in Section 3.6. Some lengthy derivations are provided in two appendices (Appendix M and P) placed at the end of the chapter.

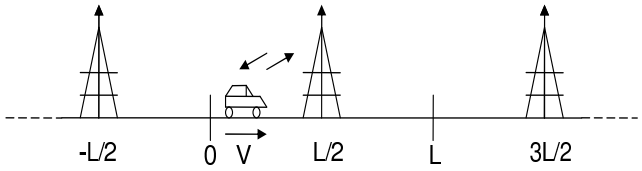


***Figure 3.1:*** *User moving with velocity V along a line*

.

## 3.2 System Model

We consider a Macrocell, $[-D, D]$, divided into a number of Small cells of length $L$. Each Small cell has a base station (BS) located at the center and all these BS communicate to a central unit (CU), which controls the entire system. *We assume that there is no interference between any two transmissions.*

Traffic Types: Define the waiting time as the duration between the arrival of the call and the instance its service starts. We consider two types of traffic: elastic (ES) and Non-elastic (NES). The two types of calls result from two different types of applications. The non-elastic traffic (multimedia streaming, voice calls etc) is very sensitive to the waiting time. These callers can tolerate some errors in transmission (for example in voice calls the ears can't distinguish the errors when they are up to a certain perseverance limit) but are very impatient. They infact drop the call if not picked up within a small waiting time. An elastic traffic (e.g., a data call) on the other hand can wait for some time before it is picked up, but is very sensitive to errors in transmission. In such calls, the natural performance metric is either the expected waiting time or the expected sojourn time, while the probability of a call being blocked, $P_B$ and the drop probability, i.e., the probability that an ongoing call is dropped before completion, $P_D$ are important performance measures for NES calls. Systems are designed with more stringent requirements on $P_D$ than $P_B$. The two types of users are assumed to be served using independent set of resources.

*Arrivals:* The two (ES, NES) arrivals are modeled by two independent Poisson arrivals with rates $\lambda$. Every arrival is associated with Marks $(X, V, S)$: $X \in [-D, D]$ the location of arrival, $S$ the file size requirement and $V$ the user velocity, distributed respectively according to $P_{n,X}$, $P_{n,V}$ and $P_{n,S}$ with respective densities$f_{n,X}$, $f_{n,V}$ and $f_{n,S}$. Let $P_n := (P_{n,X}, P_{n,V}, P_{n,S})$. We *assume symmetry in both directions, i.e., that* $P_n((X, V, S) \in A) = P_n((X, V, S) \in -A)$ *for all Borel sets* $A$. We thus *calculate and analyze without loss of generality (w.l.g.) for* $V \geq 0$.

Handovers: In major parts of this research (except for Section 3.6), we assume that handovers are completely successful at Small cell boundaries. However, we do not assume the same for Macrocell boundaries, i.e., a crossover into a new Macrocell results in a successful handover only if the new Macrocell has free servers. We model each handover into a Macrocell as a Poisson arrival, stochastically independent of the new call arrivals (as done for example in [108, 53]). We further assume that there can be at max-

imum one handover, i.e., the calls get finished before reaching the second Macrocell boundary. This simplifies the analysis to a good extent and is quite a good assumption as the Macrocells are typically large in dimension. We consider generalization of this assumption in our future work. Lastly in Section 3.6 we obtain some initial results by relaxing the 'handovers at Small cell boundaries are completely successful' assumption.

*Radio Conditions:* The BS communicates with the mobiles using a wireless link, at a rate that depends upon the distance between the two. Since our primary focus is on mobility we implicitly consider Small cells deployed outdoors, for example urban, suburban scenarios. Hence we can assume significant line of sight signal. Further, Small cells being small in size, it will be sufficient to consider only the direct path for communication. A user located at $x$ communicates with BS of cell $m$ using unit transmit power (when receiver noise variance is one) at rate[2] given by,

$$\bar{R}(x;m) := 1_{\{|x-(mL-\frac{L}{2})|\leq d_0\}} + 1_{\{|x-(mL-\frac{L}{2})|>d_0\}} {d_0}^{\beta} \left| x - \left( mL - \frac{L}{2} \right) \right|^{-\beta}, \qquad (3.1)$$

where $\beta \geq 1$ represents path loss factor and $d_0 > 0$ is a small distance up to which there is no propagation loss. The above model is valid for systems with low signal to noise ratios, where in the rates are directly given by the SNRs.

## 3.3 System Analysis

The users are moving continuously with a fixed but random velocity. The Macrocell can handle at maximum $K$ parallel calls. Transmission always occurs at fixed power $P$. Since Small cells are small in size, the movement of the users results in frequent handovers. The number of handovers will be quite large that it would be complicated to design a reliable system without redundancy: We assume that every BS can also handle $K$ parallel calls[3]. This ensures that, once a call is picked up it is not dropped at any Small cell boundary: when a user crosses over to a new BS, the new BS would at maximum be handling $K-1$ calls and hence will have at least one free server. However it is important to note that the maximum power used at any time in the system equals $KP$. We further assume that :

1) Every BSS (base station switching at a Small cell boundary) requires fixed $B_h$ bytes of information to be communicated (independent of the user's velocity), after which the

[2]The analysis will go through for any other rate functions, for example like $R(x;1) = (1 + (x - L/2)^2)^{-\beta/2}$ ([115]), $R(x;1) = log\left[1 + (1 + (x - L/2)^2)^{-\beta/2}\right]$ ([115]), $R(x;1) = log\left[1 + \left(d_0^{\beta}|x - L/2|^{-\beta}\ 1_{\{|x-L/2|>d_0\}} + 1_{\{|x-L/2|\leq d_0\}}\right)\right]$ etc. Some of the simplifications that we obtain in subsequent sections, may not be possible with these rate functions. However one can always conduct Monte Carlo simulations to obtain the required inferences.

[3]In practical systems, each BS will have $M$ backup servers to manage handovers. This means each BS can handle $M$ parallel calls. In general $M$ need not be equal to $K$, however $M$ has to be chosen large enough to ensure negligible call drops at Small cell boundaries, taking into consideration the large number of handover associated with Small cells. With this large enough $M$ the system's behavior will be close to the system considered in this work (the case with $M = K$).

user's service is resumed by the BS of the cell it just entered;

2) The user is served by the BS of the Small cell in which it is moving, as it is physically nearest to this BS.

In the following sections up to and excluding Section 3.6, all the results are derived under the assumption that, no call drops ever happen at Small cell boundaries.

### 3.3.1 Time required for communicating $S$ bytes ($B_c$)

Define by $B_c(S, X, V)$ the time required to communicate a packet of length $S$ bytes to a user located at $X$ (when the service starts) and moving with velocity $V$. If the user can communicate at a fixed rate $r$ bytes/sec then the communication time would have been $S/r$. The maximum rate at which a user can communicate with the BS in cell $m$ is given by (3.1). This position dependent rate varies: minimum when the user is at the cell edges and increases as the user moves towards the cell center. This poses a need to calculate the communication time considering the variable rates. The location of the user (under service) will change according to $X(t) = X + Vt$ (Figure 3.1). At time $t$, if the user is in cell $m$, i.e. if $X(t) \in [(m-1)L, mL]$, it communicates with the BS of $m^{\text{th}}$ cell. Hence the user gets service at time varying rate given by

$$R(t; X, V) := P\bar{R}(X(t); m) \text{ if } t \in \left[\frac{(m-1)L - X}{V}, \frac{mL - X}{V}\right].$$

Without loss of generality we consider the users, whose communication started in the first Small cell, i.e., with $X \in [0, L]$. The communication time $B_c$ required by the user, i.e., the time required to communicate $S$ bytes satisfies :

$$S = \int_0^{B_c} R(t; X, V)dt. \tag{3.2}$$

Let, $$g(l) := \int_0^{l/V} P\bar{R}(Vt; 1)dt = P\int_0^l \bar{R}(l'; 1)\frac{dl'}{V},$$

represent the number of bytes communicated while the mobile traverses interval $[0, l]$. Note that (throughout it is assumed that $L > 2d_0$: one can easily show that the optimal cell size has to be greater than $2d_0$),

$$\begin{aligned} g(L) &= \frac{Pd_0^\beta}{V}\int_0^{L/2-d_0}\left(\frac{L}{2} - l\right)^{-\beta} dl + \frac{Pd_0^\beta}{V}\int_{L/2+d_0}^{L}\left(l - \frac{L}{2}\right)^{-\beta} dl + \frac{2Pd_0}{V} \\ &= \begin{cases} \dfrac{2Pd_0}{V(\beta-1)}\left(\beta - \left(\dfrac{L}{2d_0}\right)^{1-\beta}\right) & \text{when} \quad \beta > 1 \\ \dfrac{2Pd_0}{V}\left(log(L/2) - log(d_0)\right) & \text{when} \quad \beta = 1. \end{cases} \end{aligned} \tag{3.3}$$

For any $m$, the number of bytes communicated as the user traverses through $m^{\text{th}}$ Small cell (by change of variable $l = X + Vt - (m-1)L$),

$$\int_{\frac{(m-1)L-X}{V}}^{\frac{mL-X}{V}} R(t; X, V)dt = \int_{\frac{(m-1)L-X}{V}}^{\frac{mL-X}{V}} P\bar{R}(X + Vt; m)dt = \frac{P}{V}\int_0^L \bar{R}(l; 1)dl = g(L)$$

and thus is independent of $m$. Out of this number, $B_h$ number of bytes are dedicated for BSS. Hence, irrespective of the cell which the user traverses, $g(L) - B_h$ number of bytes are transmitted during the user's journey via one Small cell. Thus the communication time can have three components : 1) Time taken in the originated cell: $(L - X)/V$, 2) Time taken to travel $N$ full cells, where (with $\lfloor t \rfloor$ representing the largest integer in $t$)

$$N = N(S, X, V) := \left\lfloor \frac{(S - (g(L) - g(X)))}{g(L) - B_h} \right\rfloor$$

represents the number of cells traveled during the communication of $S$ bytes and 3) Time taken in the cell in which the call terminates, i.e., time taken to communicate leftover bytes

$$S_l := S - (g(L) - g(X)) - N(g(L) - B_h).$$

Further a call can be handled only if the bytes that can be communicated while traversing through a cell $g(L)$, is greater than the number of bytes required for BSS $B_h$. From (3.2), the communication time $B_c(S, X, V)$ can be calculated as:

$$B_c(S, X, V) = \begin{cases} \frac{1}{V} \arg \inf_{l \in (X,L]} \{(g(l) - g(X)) \geq S\} & \text{if } S < (g(L) - g(X)) \\ \infty & \text{if } B_h > g(L) \\ \frac{L - X}{V} + N\frac{L}{V} + \frac{1}{V} \arg \inf_{l \in (0,L]} \{g(l) - B_h \geq S_l\} & \text{else.} \end{cases}$$

From (3.3), $g$ is continuous and monotonically increasing function, so $g^{-1}$ exists and thus:

**Theorem 3.3.1.1.** *Time to communicate $S$ bytes with a user initially located at $X$ and moving with velocity $V$ is,*

$$B_c(S, X, V) = \begin{cases} \frac{g^{-1}(S + g(X); V) - X}{V} & \text{if } S < (g(L) - g(X)) \\ \infty & \text{if } B_h > g(L) \\ \frac{(L - X) + NL + g^{-1}(S_l + B_h; V)}{V} & \text{else, where} \end{cases}$$

$$g^{-1}(s; v) = \begin{cases} \frac{L}{2}(1 - e^{-\frac{sv}{Pd_0}}) & \text{if } \beta = 1 \quad \frac{sv}{Pd_0} < log\left(\frac{L}{2d_0}\right) \\ \frac{L}{2} + \frac{2d_0^2 e^{-2}}{L} e^{\frac{sv}{Pd_0}} & \text{if } \beta = 1 \quad \frac{sv}{Pd_0} > log\left(\frac{L}{2d_0}\right) + 2 \\ \frac{L}{2} + \frac{sv}{P} - d_0 log\left(\frac{L}{2d_0}\right) - d_0 & \text{if } \beta = 1 \quad \text{else} \\ \frac{L}{2} - \left(\frac{sv(\beta - 1)}{Pd_0^\beta} + \left(\frac{L}{2}\right)^{-\beta+1}\right)^{\beta-1} & \text{if } \beta > 1 \quad \frac{sv(\beta-1)}{Pd_0^\beta} < d_0^{-\beta+1} - \left(\frac{L}{2}\right)^{-\beta+1} \\ \frac{L}{2} + \left(\frac{2\beta d_0^{-\beta+1}}{\beta - 1} - \frac{sv(\beta - 1)}{Pd_0^\beta} - \left(\frac{L}{2}\right)^{-\beta+1}\right)^{\beta-1} & \text{if } \beta > 1 \quad \frac{sv(\beta-1)}{Pd_0^\beta} > \frac{(\beta+1)d_0^{-\beta+1}}{\beta-1} - \left(\frac{L}{2}\right)^{-\beta+1} \\ \frac{L}{2} + \frac{d_0^\beta}{\beta - 1}\left(\frac{sv(\beta - 1)}{Pd_0^\beta} + \left(\frac{L}{2}\right)^{-\beta+1} - d_0^{-\beta+1}\beta\right) & \text{if } \beta > 1 \quad \text{else.} \end{cases}$$

□

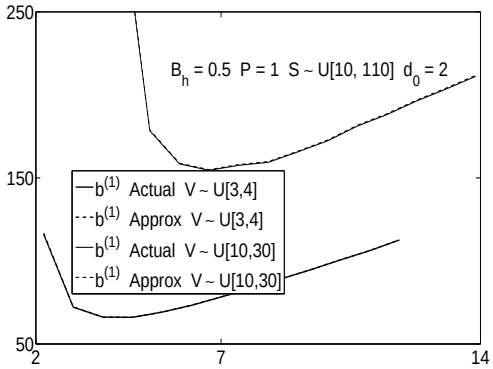


***Figure 3.2:*** *Approximation of Communication time, $B_c$*

Approximation : In Small cell based systems, a user traverses a large number of Small cells while receiving service. Hence the communication time can be approximated by the product of number of cells, $S/(g(L) - B_h)$, and the time taken for traversing each cell $L/V$:

$$B_c(S, X, V) \approx \frac{S}{g(L) - B_h} \frac{L}{V} \quad \text{when } g(L) > B_h. \tag{3.4}$$

In Figure 3.2 we show that this approximation is very good. We plot the expected value of actual communication time (given in Theorem 3.3.1.1) and the expected value of the approximation (3.4), for two different velocity profiles. As expected the approximation is very good, in fact for all velocity profiles (one can hardly distinguish the two lines in the figure).

**Remark 3.3.2.** *If a call is originated at position $X$ and moves with velocity $V \neq 0$ then, in case of ES applications, the call will be picked up at a position $X_s$ ($= X + VW$; $W$ the waiting time) different from $X$. It is difficult to estimate $X_s$ (as $W$ is unknown) and the time taken to communicate $S$ bytes, $B_c$, actually depends upon $X_s$ but not on $X$. However with the above approximation, $B_c(S, X_s, V) = B_c(S, V)$, i.e., $B_c$ does not depend upon the location of the user when its communication started.*

### 3.3.3 Maximum velocity handled by the system

Communication time, $B_c$, is finite only if the number of bytes transfered $g(L)$ per cell is strictly greater than the bytes required for BSS, $B_h$. The system can not handle velocities for which the communication times are infinite and hence we obtain (proof in Appendix **M.1**):

**Theorem 3.3.3.1.** *When $\beta = 1$, for any transmit power $P$, system can handle all velocity profiles. When $\beta > 1$, there exists a bound $V_{lim}(P) < \infty$ (increasing linearly with $P$) on the maximum velocity that can be handled by system, where*

$$V_{lim}(P) := \frac{2Pd_0\beta}{\beta - 1}. \qquad \square \tag{3.5}$$

Henceforth we consider only the velocity profiles that satisfy:

$$P_{n,V}(V < V_{max}) = 1, \text{ where } V_{max} < V_{lim}(P). \tag{3.6}$$

### 3.3.4 Service time : The time of the Macrocell spent for user's service

The user reaches the boundary of the Macrocell starting from a point $X$ in time:

$$B_\partial(X,V) := \frac{D-X}{V}. \tag{3.7}$$

The Macrocell has to serve the user either until all its $S$ bytes are transmitted (which takes time $B_c$) or till the user reaches the boundary. Thus, the overall service time requirement of the user in a Macrocell is $B_D := \min\{(D-X_s)/V, B_c(V,S)\}$, where $X_s$ (defined in Remark 3.3.2) is the user position when its service starts. By Remark 3.3.2, $B_c$ does not depend upon $X_s$. For NES applications $X_s = X$ the position of arrival. For ES applications, it is difficult to estimate $X_s$, instead we approximate $X_s$ with $X$, i.e., $B_D(X,V,S) \approx \min\{B_\partial(X,V), B_c(V,S)\}$. The error in this approximation is given (with $W$ representing the waiting time) by:

$$E_{rr} = W1_{\left\{B_c > \frac{D-X}{V}\right\}} + \left(B_c - \frac{D-X_s}{V}\right) 1_{\left\{\frac{D-X_s}{V} < B_c < \frac{D-X}{V}\right\}}$$

and so for any k,

$$E[E_{rr}^k] \leq E\left[W^k 1_{\left\{B_c > \frac{D-X}{V} - W\right\}}\right].$$

The *error is small either whenever the number of servers is large (so waiting times are small) or when the Macrocell is large in size (which usually is the case).*

### 3.3.5 Macro Handovers

Handover are modeled as Poisson arrivals (as done in [108, 53]). In this subsection we derive the other characteristics of the handover calls.

Distribution of handover call marks $(X,S)$: In general handover densities will be different from the new call densities $f_{n,X}$, $f_{n,S}$ and $f_{n,V}$. As the users move in either direction with equal probability, $P_{h,X}$ the position of handover arrival occurs either at $-D$ or at $D$ with half probability. If handover occurs at $-D$ the corresponding velocity will be positive, which is the case we consider w.l.g. We assume $f_{n,S}$ *is exponential*, i.e., $f_{n,S}(s) = \mu e^{-\mu s}$ for some $\mu > 0$, in which case by memoryless property, $f_{h,S} = f_{n,S}$.

Rate of handovers: Let

$$\nu(v,L) := \frac{\eta(L) - vB_h}{L} = v\frac{g(L) - B_h}{L}. \tag{3.8}$$

Then from (3.4) and (3.7),

$$P_{ho} \;:=\; Prob(B_\partial < B_c) \;=\; E_{n,X,V}\left[e^{-\mu\nu(V,L)(D-X)/V}\right] \tag{3.9}$$

gives the probability that a call is not completed in one Macrocell. This precisely represents that fraction of new arrivals which get converted into handover calls. So

$$\lambda_{ho} := \lambda P_{ho}$$

is the rate at which handovers occur.

Speed of handover arrival : A handover call arrives at $X = -D$ with velocity $v > 0$ only if a new call with velocity $v$ is not completed before reaching the boundary. Here we use the assumption that handover occurs at maximum at one Macrocell boundary, i.e., an handover call is not converted to another handover. Let

$$P_{ho,v} := Prob(B_\partial < B_c|v) = E_{n,X}\left[e^{-\mu_V(v,L)(D-X)/v}\right] \tag{3.10}$$

represent the conditional probability given $V = v$. Then the handover speed distribution,

$$f_{h,V}(v) = \frac{f_{n,V}(v)P_{ho,v}}{P_{ho}}.$$

### 3.3.6 Moments of Service time

Under assumption (3.6), the $k^{\text{th}}$ moment of $B_D$ exists (whenever the corresponding for $S$ and $V^{-1}$ exist) and equals, $b^{(k)} := E_{X,V,S}[(B_D(X,V,S))^k]$, where $E_{X,V,S}$ is the expectation w.r.t. the (new call and handover call) joint distribution,

$$P_{X,V,S} := \frac{\lambda(P_{n,X},P_{n,V},P_{n,S}) + \lambda_{ho}(P_{h,X,V},P_{h,S})}{\lambda + \lambda_{ho}}.$$

In Appendix **M.2** we obtain (recall $E_n$ is expectation w.r.t. $P_n$, the new call distribution):

$$b^{(k)} = E_n\left[\frac{B_D(x,v,s)^k + B_D(-D,v,s)^k P_{ho,v}}{1+P_{ho}}\right] \text{ and} \tag{3.11}$$

$$\frac{db^{(k)}}{dL} = E_n\left[\frac{d}{dL}\left(\frac{B_D(x,v,s)^k + B_D(-D,v,s)^k P_{ho,v}}{1+P_{ho}}\right)\right]. \tag{3.12}$$

### 3.3.7 Cell size optimizing the moments of the service time

The number of bytes that can be communicated in a cell increases with the increase in cell size: from (3.3) $g$ is continuous and monotonically increasing w.r.t. $L$. For any given velocity there exists a minimum cell size (the smallest cell size at which one can transmit more than $B_h$ bytes per cell), below which successful communication is not possible. When cell size is closer to this smallest one, the useful bytes transmitted per cell ($g(L) - B_h$) are very small and hence it takes more time to transmit $S$ bytes, i.e., the communication time $B_c$ will be large. As the cell size increases from this smallest size, the communication time $B_c$ starts reducing. However after some point, due to path loss, the number of bytes per cell starts saturating and hence the gain in terms of useful bytes transmitted per cell will be small in comparison with the extra time taken to traverse each cell, resulting in increasing the communication time again. *Thus*

*there exists an optimal cell size for every fixed velocity*. One can extrapolate similar things even for random velocity. We derive the optimal cell size $L^*_{b^{(1)}}$ and relate the same to the optimizer of more interesting performance measures for ES and NES calls in the subsequent sub-sections.

Define $L^*_\nu(v) := \arg\max_L \nu(v, L)$, the maximizer of the function $\nu$ given by (3.8). In Appendix **M.3** we show that, there exists an unique $L^*_\nu(v) > 0$ for every velocity $v$. In the Appendix **M.4** we further show that, for fixed velocities (i.e., when $V \equiv \bar{v}$), the derivatives of all the (existing) $k^{\text{th}}$ moments of the service time vanish only at $L^*_\nu(\bar{v})$. Thus *for fixed velocities, all the (existing) moments of the service time have unique and common minima*:

$$L^*_{b^{(k)}} := \arg\min_L b^{(k)} = \arg\max_L \nu(\bar{v}, L) = L^*_\nu(\bar{v}) \quad \text{for all } k.$$

The common optimizer $L^*$ (for example for $\beta > 1$) satisfies

$$\left.\frac{\partial \nu(\bar{v}, L)}{\partial L}\right|_{L=L^*} = 0 \quad \text{or} \quad 2P\left(\frac{L^*}{d_0}\right)^{-\beta} L^* - \eta(L^*) + \bar{v}B_h = 0 \text{ and therefore}$$

**Theorem 3.3.7.1.** *For fixed velocity profile, i.e., $P_{n,V}(V = \bar{v}) = 1$ the optimal cell size for the expected service time is,*

$$L^*_{b^{(1)}} \quad = \quad L^*_\nu \quad = \quad \begin{cases} 2\left(\dfrac{2Pd_0^\beta \frac{\beta}{\beta-1}}{2Pd_0\frac{\beta}{\beta-1} - \bar{v}B_h}\right)^{\frac{1}{\beta-1}} & \text{when} \quad \beta > 1 \\ 2d_0 e^{\frac{\bar{v}B_h}{2Pd_0}} & \text{when} \quad \beta = 1. \end{cases}$$

*Further, if the $k^{th}$ moment of the service time exists then $L^*_{b^{(k)}} = L^*_{b^{(1)}}$.* □

*For velocity profiles with small variance, the optimizers of all the moments of the service time will be equal approximately.* Hence when $P_{n,V}$ has small variance with mean $\bar{v}$ then $L^*_{b^{(1)}}$ is close to $L^*_\nu(\bar{v})$ . *From the above it is clear that $L^*_{b^{(1)}}$ increases when the mean $\bar{v}$ increases.*

Having obtained the service time, we now turn our attention to model various configurations of the Macrocell with appropriate queues to further obtain their performance measures.

### 3.3.8 ES Calls : Average Waiting time

Each Macrocell can handle at maximum $K$ parallel calls. The CU of the Macrocell keeps a record of the users entered into the system and serves them in FIFO (first in first out) order via the BSs of the Small cells. When a new user initiates a call, it is immediately picked up if there are less than $K$ active calls in the system. If not the user will have to wait. Its service will start at the time : 1) when one of the active $K$ users finish their service and exit 2) if there are no other waiting users arrived before it. The BS nearest to the user, at the time of its service start, will initiate the call. Hence after, its call is served (by the Macrocell under consideration) as discussed in subsection 3.3.1 either till its

service is over or till it reaches the Macrocell boundary. When it reaches the boundary the call will be transfered to the next Macrocell as a handover call and the handover call is treated by the new cell similar to that of a new call. Thus each Macrocell can be modeled by a $M/G/K$ queue with service time $B_D$ and Poisson arrivals at rate $\lambda + \lambda_{ho}$. This queue has been analyzed to a good extent and the system is stable only if ([140])

$$\rho := \frac{(\lambda + \lambda_{ho})b^{(1)}}{K} < 1.$$

For stable queues, the expected waiting time of a randomly arrived customer can be approximated by ([140]):

$$E[W]_K = \left(\frac{b^{(2)}}{2(b^{(1)})^2}\right)\left(\frac{b^{(1)}}{K(1-\rho)}\right)\left(\frac{(K\rho)^K}{K!}\right)\pi_0; \;\; \pi_0^{-1} = \frac{(K\rho)^K}{K!} + (1-\rho)\sum_{i=0}^{K-1}\frac{(K\rho)^i}{i!} \quad (3.13)$$

where $b^{(1)}, b^{(2)}$ are given by (3.11). If the system is unstable the number of waiting customers grows towards infinity and thus one should consider only the cell sizes $L$ with $\rho < 1$. Hence, the optimal size, which minimizes (3.13) is

$$L^*_{ES} := \arg \min_{\{L:\rho<1\}} E[W]_K.$$

We saw in the previous section that the optimizer of $b^{(2)}$ is same as that of $b^{(1)}$ for fixed velocities and will be close to each other for smaller velocity variances. In a similar way the same thing is true for $\rho$, i.e., for fixed velocities,

$$L^*_{\rho} := \arg \min_{\{L:\rho<1\}} \rho = L^*_{b^{(1)}}.$$

The expected waiting time (3.13) is continuously differentiable in $b^{(1)}$, $b^{(2)}$ and $\rho$. Thus (minimizer of (3.13) is a zero of its derivative and $E[W]_K$ depends upon $L$ only via $b^{(1)}$, $b^{(2)}$, $\rho$),

**Theorem 3.3.8.1.** *Optimal cell size for a system with elastic traffic, with $P_{n,V}(V = \bar{v}) = 1$ is,*

$$L^*_{ES} = \arg \min_{\{L:\rho<1\}} E[W]_k = L^*_{b^{(1)}}.$$

*So, $L^*_{b^{(1)}}$ minimizes both expected waiting time and expected service time. Also, it minimizes the expected sojourn time, as it is the sum of expected waiting and service times.* □

Also from (3.13), when $L^*_{b^{(1)}}$ and $L^*_{b^{(2)}}$ are close, it is easy to see that the optimizer of $E[W]_K$ will be close to that of the expected service time, $b^{(1)}$. Thus even for low velocity variances,

$$L^*_{ES} \approx \arg \min_{\{L:\rho<1\}} b^{(1)} = L^*_{b(1)}.$$

We see that this is true even for many general velocity profiles in examples section 3.5.

### 3.3.9 NES Calls : Block and Drop Probabilities

As before the system can handle at maximum $K$ parallel calls. The call is picked up immediately (by the BS of the Small cell in which the call is originated) only if the Macrocell is serving lesser than $K$ users at the time of its arrival. If all the $K$ servers are busy it is dropped. When an active customer reaches the boundary of a Macrocell, its call is continued in the next Macrocell only if the new Macrocell has free servers. Each Macrocell can thus be modeled by an $M/G/K/K$ queue. And its call block probability is given by the Erlang Loss formula ($\rho$ was defined in previous section),

$$P_B(L) \quad = \quad \frac{\rho(L)^K/K!}{\sum_{k=0}^{K}\rho(L)^k/k!}.$$

It is interesting to note that $P_B(L)$ and $\rho$ are both differentiable w.r.t. $L$ and further that if the derivative $d\rho/dL$ is zero at a $L^*$ so is the derivative $dP_B/dL$. By taking the second derivative, we can in fact show that their minimizers are the same. Hence,

$$L^*_{P_B} \quad = \quad \arg\min_{\{L:\rho(L)<1\}} P_B(L) \quad = \quad L^*_\rho.$$

Further at fixed velocities, $L^*_\rho = L^*_{b^{(1)}}$ and so we have,
**Theorem 3.3.9.1.** *The minimizer, $L^*_{b^{(1)}}$ also minimizes the block probability, $P_B$, for fixed velocities. For any velocity profile, $L^*_\rho$ also minimizes the block probability.* □

Drop Probability : Under the assumptions stated earlier, only a new call can reach the boundary and not a call which was already handed over once. Further, an active call is dropped only when it reaches the Macrocell boundary and the new Macrocell is busy. By independence of the two events (status of the new Macrocell prior to handover is independent of the call that is handed over), the drop probability is

$$P_D(L) = P_{ho}P_B(L).$$

One can design an optimal system, catering to NES calls, either by jointly minimizing both the block and drop probabilities or by minimizing one of the probabilities while placing a constraint on the other. Usually systems are designed with more stringent requirements on $P_D$ than on $P_B$. We note from the above calculations that $P_D$ is directly proportional to $P_B$ and will be smaller than $P_B$ by a factor $P_{ho}$. We make in the rest of this subsection, a commonly made assumption that, the location of the call arrivals is uniformly distributed, i.e. that $X \sim \mathcal{U}[-D, D]$. Under this assumption:

$$\begin{aligned}
P_{ho} &= P_n\left(B_\partial(X,V) < B_c(V,S)\right) = E_{n,S,V}\left[P_X(D-X < VB_c(V,S))\right] \\
&= \frac{E_{n,S,V}\left[\min\left\{2D, \frac{VSL}{\eta(L)-VB_h}\right\}\right]}{2D} \\
&= E_{n,V}\left[e^{-\mu 2D(\eta(L)-VB_h)/VL}\right] + \frac{E\left[1-e^{-\mu 2D(\eta(L)-VB_h)/VL}\left(1+\frac{2D\mu(\eta(L)-VB_h)}{VL}\right)\right]}{2D\mu} \\
&< E_{n,V}\left[e^{-\mu 2D(\eta(L)-VB_h)/VL}\right] + \frac{1}{2D\mu}.
\end{aligned}$$

Thus $P_{ho}$ decreases with $2D$, the Macrocell size. Macrocells are large in dimension and hence $P_D$ can be ensured to be within the prescribed limits (the limit is a design parameter) by directly minimizing $P_B$ itself. Thus we propose to choose cell size $L$ to minimize $P_B$ and hence equivalently $\rho$:

$$L^*_{NES} := L^*_\rho = \arg \min_{\{L:\rho(L)<1\}} \rho(L) \approx L^*_{b^{(1)}}.$$

**Remark 3.3.10.** *Thus for both ES and NES applications one needs to minimize the first moment of the service time, $b^{(1)}$, to obtain the optimal cell size. This optimal cell size has been discussed in the previous section for fixed velocities and for velocity profiles with small variances. The general situation is studied in section 3.5 via numerical examples.*

## 3.4 Mobility on a street grid

We assume that users are moving in a rectangular grid overlaying a large area $[-D, D] \times [-D, D]$ with grid size $b, d$ as shown in Figure 3.3. This example is typical of urban areas where the streets are in a criss-cross manner (this is a well known model, see for example [103, 14]) and hence is an interesting example for study.

In this case, we assume that the location of arrival $X$ is uniformly distributed on the lines, i.e, $X \sim \mathcal{U}[\mathcal{G}]$, where the grid

$$\mathcal{G} := \cup_{i=1}^{D/d}[-D, D] \times \left\{id + \frac{d}{2}\right\} \cup \cup_{i=1}^{D/b}\left\{ib + \frac{b}{2}\right\} \times [-D, D].$$

A one dimensional vector $V$ represents the speed of the vehicle, which is uniformly distributed, i.e., $V \sim \mathcal{U}[0, V_{max}]$ . It's direction depends upon the position of arrival $X$: it is horizontal if $X$ is on horizontal line and is vertical if on vertical line. One can easily extend the analysis to include zig-zag paths. In either case we assume it be equiprobable in the two possible directions; towards left or right in case of horizontal line and towards up or down in case of vertical line.

Any Small cell is a line segment of a street and a base station is placed at the center of this cell (if we neglect the small number of Small cells that might possibly span across two intersecting streets). The mobiles may change directions as they take a turn, but the rate they obtain with their BS once again follows periodic pattern as explained in section 3.3.1. Hence the time to reach the boundary $B_\partial$, the time to serve $S$ bytes $B_S$ and hence the service time $B_D$ are just the same as those derived in the previous sections. Thus *the analysis and the results of all the previous sections is applicable to the grid structure. So in the grid structure, the 2 dimensional analysis actually boils down to one dimensional analysis itself.*

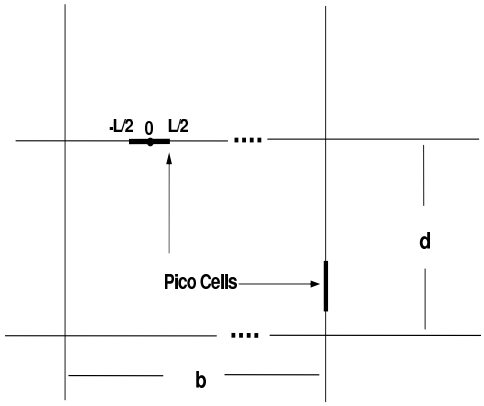


***Figure 3.3:*** *2D network for rectangular-grid small cell networks*

.

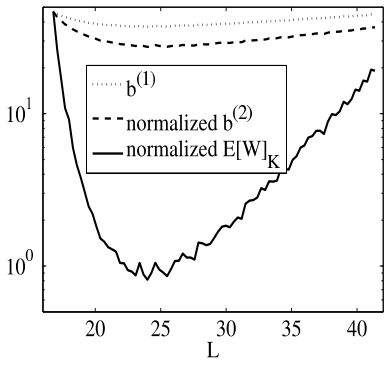


***Figure 3.4:*** *Moments of the service time and the expected waiting time versus L.*

## 3.5 Mobility Examples

In the numerical examples of this section, we consider uniformly distributed velocity profiles. The position of arrival $X$ is also uniformly distributed in the Macrocell $[-D, D]$. We use Monte Carlo simulations to estimate $b^{(1)}$, $b^{(2)}$ and use exhaustive search to find the optimizers. In figure 3.4 we plot normalized values of $b^{(1)}$, $b^{(2)}$ and $E[W]_K$ versus $L$. As discussed earlier we notice that the various performance measures decrease with cell size initially, reach an optimal value and increase again from then on. In fact, all the performance measures have unique minimum at the same $L$. We study more details of these minimizers in the following.

In figure 3.5 we plot the optimal cell size (optimal with respect to the moments of the service time $b^{(1)}$, $b^{(2)}$, block and drop probabilities $P_B$, $P_D$ of NES calls and the expected waiting time $E[W]_K$ of ES calls) versus mean velocity for two different values of variance. We set $d_0 = 5$, $\lambda = 0.1$, $B_h = 2$, $P = 1$, $\mu = 5$, $K = 20$ and consider a Macrocell of size $D = 1000$. We also plot $L^*_\nu$, which is the maximizer of $\nu(E_n[V], L)$. Note that $L^*_\nu$ is a single curve in the figure while the remaining 5 optimizers ($L^*_{b^{(1)}}$, $L^*_{b^{(2)}}$, $L^*_{P_B}$, $L^*_{P_D}$ and $L^*_{E[W]}$) are plotted for two values of variance and hence there are two curves for each of these 5 optimizers. In fact, $L^*_\nu$ is not visible separately as it completely

coincides with the minimizer $L^*_{E[W]}$, plotted for variance equal to 1. For small velocity variance (variance equal to 1), all the minimizers are close to $L^*_\nu$. For large velocity variance, we notice that all the minimizers (together) are away from $L^*_\nu$, but however are close to each other for most cases. That is, the minimizers of expected waiting time are the same as that of block as well as drop probabilities and all of them equal $L^*_{b^{(1)}}$. This suggests that *even for velocity profiles with high variances, it is sufficient to optimize the average service time $b^{(1)}$ for both ES as well as NES calls and hence the optimal cell size again remains independent of the application. However for high variance it is not sufficient to minimize $\nu(E_n[V], L)$, rather one needs to minimize $b^{(1)}$ directly.*

In Figure 3.6 we further illustrate the same, by plotting the various optimizers now as a function of velocity variance. We set mean, $E[V] = 10$, $d_0 = 5$, $\lambda = 0.1$, $B_h = 2$, $P = 1$, $\mu = 5$, $K = 20$. We once again note that all the minimizers are close to each other for many cases. We also note that all the minimizers are close to $L^*_\nu$ for low velocity variances. We further observe that the optimal cell size increases with increase in the variance also. Thus *larger the velocities the system has to support, the larger are the optimal cell sizes.*

We notice that, in both the figures (3.5 and 3.6) only the optimizer of the second moment $L^*_{b^{(2)}}$, is some times different from the rest of the minimizers. However, even when $L^*_{b^{(2)}}$ is different from $L^*_{b^{(1)}}$, the minimizer $L^*_{EW}$ (minimizing the expected waiting time) is equal to $L^*_{b^{(1)}}$ and thus for both the types of traffic $L^*_{b^{(1)}}$ (minimizer of expected service time) gives the optimal cell size.

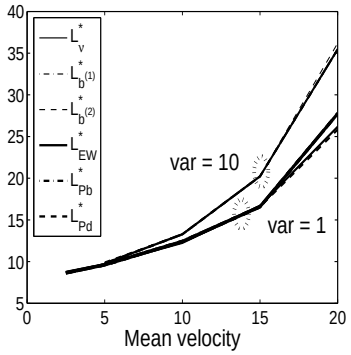


***Figure 3.5:*** *Optimal cell size versus mean velocity for different variances.*

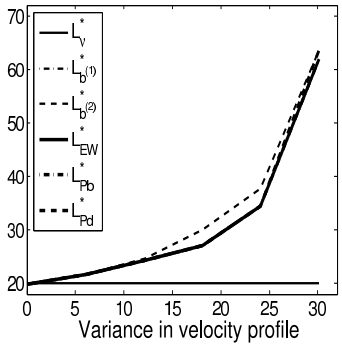


***Figure 3.6:*** *Optimal cell size versus variance of the velocity.*

## 3.6 Call drops at Small cell boundaries (NES calls)

So far in our analysis, we have assumed that there are no call drops during handover when the mobile traverses across Small cells. We hence called it as just BSS (Base station switching). In practice even the handovers at Small cell boundaries can fail, even

though with an extremely small probability. In this section we obtain some initial results considering possible Small cell drops for the special case of uniform arrivals, uniform velocity profile and the networks catering to NES calls. We further assume that the velocity profile does not include 0 so that $E[1/V] < \infty$.

Every Small cell handover, as before, needs $B_h$ bytes to perform the BS switching and in addition has a (small) possibility of not being successful. We now call the BSS near a Small cell boundary also as a handover. The BS of the Small cell admits a call (handover or a new one), as before $K$ calls at maximum, however some of the new calls are not picked up to reserve resources for the handover calls, now to keep the overall Drop probability under the required limit. Further in this case the call admission control can also be distributed: 1) in a centralized scheme the CU directs the Small cell BS as in the previous section, either to admit or not admit the new calls (base stations possibly are using fewer number of servers); 2) in a decentralized scheme, the Small cell BS always reserves $K_1$ servers exclusively for handover calls. That is, in a decentralized scheme, the Small cell BS (independent of the other Small cells) admits a new call if more than $K_1$ servers are free while a handover call is admitted whenever any server is free.

In this work we consider a call admission control scheme that is decentralized and is different from the above two schemes. A new call is *considered for picking up* with probability $p$ independent of everything else. And a call *considering for picking up* will be picked up when any one of the servers is free. This scheme is close to the decentralized scheme described above when $p$ is close to one, i.e., when $K_1$ is close to 0. The centralized scheme (with Small cell drops) on the other hand has to be handled using totally different queuing models and is not considered in this research. Note that smaller $p$ implies larger priority to handover calls. In this case *each Small cell itself is modeled as a separate $M/G/K/K$ queue*.

**Remark 3.6.1.** *We considered a simple, analytically tractable and a decentralized call admission control in this section. One can try to analyze a system which reserves $K_1$ servers for the handover calls using priority queues. On the other hand it will be more difficult to analyze a system with centralized control (with possible Small cell drops), in which the CU directs to admit new calls or not depending upon the total number of calls in the system. One will need the theory of interacting queues for such a study.*

### 3.6.2 Service Time : Time of the Small cell spent for user's service

The service time $B_L$, now represents the overall time of a Small cell used by the mobile. Thus the service time will be much smaller in comparison with the previous section and its analysis will be significantly different. There is no more periodicity while obtaining the service time, however the analysis in this case also simplifies, now due to small values of $L$. The service time components $B_c$ and $B_\partial$ are:

$$B_\partial = \frac{L - X}{V} \text{ and } B_c = \frac{g^{-1}\left(S + g(X); V\right) - X}{V} \text{ when } B_c \leq B_\partial$$

with the overall service time as before equal to $B_L = \min\{B_\partial, B_c\}$. We use notation $B_L$ instead of $B_D$ for service time, to emphasize that this is the service time of a Small cell

of length $L$. Further,

$$P_{ho} = Prob(B_\partial < B_c) = Prob(g(L) - g(X) < S) = E[e^{-\mu(g(L)-g(X))}] = E[e^{-\mu\frac{\eta(L)-\eta(X)}{V}}]$$

and $P_{ho,v}$ also changes similarly. However it is easy to see that as the cell size decreases towards zero the function inside the expectation converges to 1 for all the realizations and hence by BCT $P_{ho}$, $P_{ho,v}$ (for all $v$) converge to 1 as $L \to 0$. Thus for Small cells *we approximate* $P_{ho}$, $P_{ho,v}$ *by 1.*

The service time in a Small cell with high probability (the probability increasing as $L \to 0$) will be equal to the time taken to traverse the cell itself, as with high probability the user has to pass through many Small cells before completing his request and thus:

$$B_L(S, X, V) \approx \frac{L - X}{V}.$$

### 3.6.3 Small cell Handovers

Handover arrival rate : There is a major difference in the analysis while calculating the handover rate $\lambda_{ho}$: now we should consider that the (Small cell) handovers can happen many times (and not at most once as in case of a Macro handover) during a call duration. The calculations though different can easily be carried out. We introduce some new notations for this purpose. Let $P_{ho,\partial}$, $P_{ho,\partial,v}$ respectively represent terms equivalent to $P_{ho}$ and $P_{ho,v}$ for a call that is already handed over (at least once). It is easy to see that (when $P(V < V_{lim}(P)) = 1$) by BCT again (using equation (3.8)):

$$P_{ho,\partial} = Prob(g(L) < S + B_h) = E\left[e^{-\mu(g(L)-B_h)}\right] = E\left[e^{-\mu\frac{L\nu(L)}{V}}\right] \to 1 \text{ as } L \to 0$$

and similarly $P_{ho,v,\partial} \to 1$ as $L \to 0$ (for all $v$). By Taylor series, $e^{-x} \approx 1 - x$ for small values of $x$ and this is used for obtaining another approximation for $P_{ho,\partial}$ again using BCT:

$$(1 - P_{ho,\partial}) - \mu L E\left[\frac{\nu(L, V)}{V}\right] \to 0 \text{ as } L \to 0. \tag{3.14}$$

This approximation is more appropriate when we consider terms like $1/(1 - P_{ho,\partial})$. From (3.8) for $L$ small using (3.14),

$$1 - P_{ho,\partial} \approx \frac{\mu L \nu(\bar{v}_{inv}, L)}{\bar{v}_{inv}} \text{ where } \bar{v}_{inv} := \frac{1}{E\left[\frac{1}{V}\right]}. \tag{3.15}$$

Let, $\lambda_L := pL\lambda/D$ be the effective rate at which the users arrive into the Small cell of interest $[0, L]$. Note in the above that $L\lambda/D$ represents the actual rate at which the new calls arrive in $[0, L]$ while $\lambda_L$ represent the arrival rate of those sampled users who are considered for picking up. The handover rate $\lambda_{ho}$ in this case can be calculated as: 1) due to symmetry, the handovers from Small cell 0 ($[0, L]$) to cell 1 ($[L, 2L]$) are stochastically same as those from cell -1 ($[-L, 0]$) to cell 0; 2) the same is true for handovers when

a mobile travels from right to left; 3) so, the handovers into a cell of interest (cell 0) are stochastically same as those that go out of the cell 0; 4) thus under the assumption that handovers are Poisson in nature, the handover rate should satisfy the following fixed point equation:

$$\lambda_{hL} P_{ho,\partial} + \lambda_L P_{ho} = \lambda_{hL}.$$

Using the approximation in equation (3.15),

$$\lambda_{hL} = \frac{\lambda_L P_{ho}}{1 - P_{ho,\partial}} \approx \lambda \frac{p\bar{v}_{inv}}{D\mu\nu(\bar{v}_{inv}, L)}. \tag{3.16}$$

**Remark 3.6.4.** *Note that* $L^*_{\lambda_{hL}} := \arg\min_L \lambda_{hL} = L^*_\nu(\bar{v}_{inv})$ *and that for fixed velocities (when* $V \equiv \bar{v}$*)* $\bar{v}_{inv} = \bar{v}$*. Thus for fixed velocities, interestingly for any value of* $p$*, the cell size optimizing the the handover rate* $\lambda_{hL}$ *(considering the Small cell drops) is the same one that is optimal for NES as well as ES calls, obtained by neglecting the Small cell drops.*

Handover Speed distribution: Note in equation (3.16) the expectation $E[1/V]$ in the last term is w.r.t. to the distribution corresponding to the handover calls and hence one needs to calculate these distributions. The handover arrivals are either at 0 or at $L$ with half probability and since we *are considering the positive velocities without loss of generality the position is always at* 0. The (Small cell) handover speed distribution, $f_{h,V}(v)$, (after considering the drops at Small cells) once again satisfies another fixed point equation due to the statistical similarity between the arrivals into and out of the cell 0. In Appendix P, this fixed point equation is derived and the Small cell handover speed distribution is shown to converge to uniform speed distribution as the cell size $L$ tends to zero. Thus *the (Small cell) handover arrivals have approximately uniform speed distribution.*

### 3.6.5 Stability Factor

The stability factor $\rho$ in a Small cell queue is calculated in Appendix P using the approximations for small cell sizes and we obtain:

$$\rho_{pico}(L) \quad = \quad \frac{(\lambda_L + \lambda_{hL})b^{(1)}}{K} \quad \approx \quad \frac{\lambda pL}{KD\mu\nu(\bar{v}_{inv}, L)}. \tag{3.17}$$

Define

$$L^*_{\nu/L}(v) := \arg\max_L \frac{\nu(v, L)}{L}.$$

Thus the optimal cell size for stability factor is,

$$L^*_{\rho,pico} := \arg\min_L \rho_{pico}(L) = L^*_{\nu/L}(\bar{v}_{inv}).$$

### 3.6.6 New Call Block Probability

The Small cell can be modeled by an $M/G/K/K$ queue as in the previous section. Let $P_{Busy,pico}$ represent the probability that an arrival (a new or an handover call) finds all

the $K$ servers busy in the Small cell. This probability can be calculated using the results from queuing theory as done in the previous section using Erlang Loss formula,

$$P_{Busy,pico}(L) := \frac{\rho_{pico}(L)^K/K!}{\sum_{k=0}^{K}\rho_{pico}(L)^k/k!}.$$

A new arrival is not picked up either with probability $1-p$ when the BS intentionally does not consider it for picking up or with probability $pP_{Busy,pico}$ when all the servers are busy. Thus, the new call block probability for system when considering Small cell drops will be given by,

$$P_{B,pico}(L) = (1-p) + pP_{Busy,pico}(L).$$

From the above the cell size optimizing the block probability $P_{B,pico}$ will be same as that optimizing $P_{Busy,pico}$, which further is same as $L^*_{\rho,pico}$ (using the logic as in section 3.3.9). Thus it is clear that, $L^*_{P_B,pico} = L^*_{v/L}(\bar{v}_{inv})$ and from (3.17) this optimizer is a zero of

$$\frac{d\rho}{dL} = c_1\frac{2L(\eta(L)-\bar{v}_{inv}B_h) - L^2\eta'(L)}{(\eta(L)-\bar{v}_{inv}B_h)^2},$$

where $c_1$ is an appropriate positive constant. Hence we have,

**Theorem 3.6.6.1.** *The cell size optimizing the block probability of the NES calls considering the possible drops at Small cell boundaries, when resources are reserved for handover calls by intensionally dropping some of the new arrivals, is given by,*

$$L^*_{P_{B,pico}} = L^*_{\rho,pico} = L^*_{v/L} = \begin{cases} 2\left(\dfrac{Pd_0^{\beta}\frac{\beta+1}{\beta-1}}{2Pd_0\frac{\beta}{\beta-1} - \bar{v}_{inv}B_h}\right)^{\frac{1}{\beta-1}} & when \quad \beta > 1 \\ 2d_0 e^{\frac{\bar{v}_{inv}B_h - Pd_0}{2Pd_0}} & when \quad \beta = 1. \end{cases} \qquad \square$$

When this optimal cell size is compared at fixed velocities with that obtained after neglecting the Small cell drops (that obtained in Theorem 3.3.7.1) we find that: 1) when $\beta > 1$, the two cell sizes matches except for a factor of $((\beta+1)/(2\beta))^{1/(\beta-1)}$; 2) when $\beta = 1$, the difference is in the power of the exponent, an extra $-Pd_0$ factor for the cell size with Small cell drops. This differences are negligible only when $\beta$ is close to 1 (but not equal to 1). The two systems use different call admission control mechanisms and hence the difference. Nevertheless *the optimal cell size of the Theorem 3.6.6.1 is valid for a distributed call management system and further it gives the cell size even for velocity profiles with non zero variance.*

### 3.6.7 Drop Probability

It was straight forward to calculate the drop probability for the previous case (i.e., without Small cell drops) as the drop could have occurred at maximum at one Macro boundary. It is more tedious to calculate the same when drops are also possible at Small cell

boundary. This tedious job is carried out in Appendix P, wherein the drop probability is obtained by conditioning on two events. We obtain (see Appendix P):

$$P_{D,pico} = \frac{P_{Busy,pico}}{\frac{\mu(\eta(L)-\bar{v}_{inv}B_h)}{\bar{v}_{inv}} + P_{Busy,pico}}$$

It is interesting to see that both the call drop and the new call block probabilities depend upon the busy probability $P_{Busy,pico}$ and one can thus design a optimal system (considering Small cell drops) by minimizing the busy probability or equivalently the stability factor $\rho_{pico}$. We thus propose to choose the optimal cell size:

$$L^*_{NES,pico} := L^*_{\rho,pico} = L^*_{P_B,pico},$$

which is obtained in Theorem 3.6.6.1.

## 3.7 Conclusions and Future work

We looked at the problem of characterizing the performance of Small cell networks in the presence of mobility. We modeled various traffic types between base-stations and mobiles as different types of queues. We derived explicit expressions for expected waiting time, service time and drop/block probabilities for the various queuing models considered for both fixed as well as random velocity of mobiles. We showed that there exists an optimal cell size for a given velocity profile, which minimizes the service time for elastic applications as well as the drop and block probabilities of non-elastic applications. We obtained (approximate) closed form expressions for this optimal cell size when the velocity variations of the mobiles is very small. We find that if the call is long enough, the optimal cell size depends mainly on the velocity profile of the mobiles, its mean and variance.; It is independent of the traffic type or duration of the calls. We show that for any fixed power of transmission, there exists a maximum velocity beyond which successful communication between the mobile and the system is not possible. This maximum possible velocity increases with the power of transmission. Further, for any given power, the optimal cell size increases when either the mean or the variance of the mobiles velocity increases.

The mobility models considered in this work are suitable for modeling users traveling continuously with considerable speeds (example users traveling in a car). The movement of slower users (e.g., pedestrians) can be better captured by either Random Walk or Random Way-point model. Still better would be to consider systems with heterogeneous users (slow moving, fast moving and users that are at rest). We considered two dimensional Small cell networks in [118] catering to non elastic users, in which the users can move in any direction across the cell. In that work, we brought out some issues while designing the Small cell networks that are specific to two dimensional cellular networks. It would be interesting to extend those results for elastic users and further for heterogeneous users.

## 3.8 Appendix M: Calculations related to Macro queue

### 3.8.1 M.1 Proof of Theorem 3.3.3.1

The communication time is finite with probability one if and only if

$$Prob(B_h > g(L)) = P_{n,V}(VB_h > \eta(L)) = 0, \quad \text{where } \eta(L) := Vg(L).$$

Note from (3.3) that $\eta$ is only a function of $P$, $L$ and hence that this probability depends only upon the velocity profile $V$, cell size $L$ and the transmit power $P$. Because of the path loss, for any fixed $P$, $\eta(L)$ increases as $L$ increases and finally saturates (when $\beta > 1$). Thus for all $L$ (when $\beta > 1$),

$$\eta(L) \leq \eta_\infty := \lim_{L\to\infty} \eta(L) < \infty.$$

When $\beta = 1$, $\eta_\infty = \infty$ and this proves the first statement of the Theorem. As $Prob(B_h > g(L)) = Prob\,(VB_h > \eta(L))$, there exists a cell size $L$ with $P(B_h > g(L)) = 0$ if and only if

$$Prob\left(V > \frac{\eta_\infty}{B_h}\right) = 0.$$

Thus with a given power $P$, the system can handle all velocities that are strictly less than

$$V_{lim} := \frac{\eta_\infty}{B_h} = \frac{2Pd_0\beta}{\beta - 1}. \quad \square$$

### 3.8.2 M.2 Moments of service time and its derivatives

The moments can be rewritten as,

$$\begin{aligned} b^{(k)} &= \frac{1}{(1+P_{ho})} \int_0^\infty \int_{-D}^{D} \int_0^{V_{max}} \left[B_D(x,v,s)^k + B_D(-D,v,s)^k P_{ho,v}\right] \\ &\qquad\qquad f_{n,V}(v)dv \;\; f_{n,X}(x)dx \;\; \mu e^{-\mu s} ds. \\ &= E_n\left[\frac{B_D(x,v,s)^k + B_D(-D,v,s)^k P_{ho,v}}{1+P_{ho}}\right]. \end{aligned}$$

By Bounded Convergence Theorem (BCT), all $b^{(k)}$ are continuously differentiable (c.d.) in $L$ and the derivative is

$$\frac{db^{(k)}}{dL} = E_n\left[\frac{d}{dL}\left(\frac{B_D(x,v,s)^k + B_D(-D,v,s)^k P_{ho,v}}{1+P_{ho}}\right)\right]$$

because: 1) $B_D$ is almost surely c.d.; 2) $P_{ho,v}$ is c.d. everywhere in $L$; 3) $P_{ho}$ is c.d. and 4) all the derivatives involved are uniformly bounded almost surely; 4) hence by virtue of mean value theorem, the terms like

$$\frac{|B_D(X,V,S;L+\delta)-B_D(X,V,S;L)|}{\delta}, \frac{|P_{ho,v}(L+\delta)-P_{ho,v}(L)}{\delta} \text{ etc}$$

can be bounded uniformly by a constant.

### 3.8.3 M.3 $\nu$ has an unique maximizer:

From equation (3.3), $g$ and hence $\eta = vg$ are both concave in $L$ on $(0,\infty)$ for every $v$. Thus from (3.8), for any fixed velocity $v$, $\nu$ has a unique maxima,

$$L^*_\nu(v) := \arg\max_L \nu(v,L),$$

which satisfies $\partial\nu/\partial L = 0$ (as clearly $L^*_\nu(v) > 0$ for all $v$).

### 3.8.4 M.4 Derivatives $db^{(k)}/dL$ vanish only at $L^*_\nu(\bar{v})$ when $V \equiv \bar{v}$

Define $\Psi(L) := E_n\left[B_D(X,V,S)^k + B_D(-D,V,S)^k P_{ho,V}\right]$. Then from (3.11)

$$b^{(k)} = \frac{\Psi(L)}{1+P_{ho}}.$$

From (3.4), $B_c$ depends upon $L$ only via the function $\nu$ given by (3.8) and hence so is the service time $B_D(x,v,s) = \min\{B_c(v,s), B_\partial(x,v)\}$ for all $x,v,s$. Similarly from (3.10) and (3.9), $P_{ho,v}$ and $P_{ho}$ depend upon $L$ only via the function $\nu$. Hence with

$$\Theta(v) := -\frac{\partial P_{ho,v}}{\partial \nu} = \frac{\mu}{v} E_{n,X}\left[(D-X)e^{-\mu\nu(v,L)(D-X)/v}\right]$$

$$\begin{aligned}
\frac{db^{(k)}}{dL} &= \frac{1}{1+P_{ho}} E_n\left[\frac{\partial\nu(V,L)}{\partial L}\left(\frac{\partial B_D(X,V,S)^k}{\partial\nu} + P_{ho,V}\frac{\partial B_D(-D,V,S)^k}{\partial\nu} + \frac{\partial P_{ho,V}}{\partial\nu} B_D(-D,V,S)^k\right)\right] \\
&\quad - \frac{1}{(1+P_{ho})^2}\Psi(L)\frac{dP_{ho}}{dL} \\
&= \frac{1}{1+P_{ho}} E_n\left[\frac{\partial\nu(V,L)}{\partial L}\left(-kB_D(X,V,S)^{k-1}\frac{S1_{\{SV<(D-X)\nu(V,L)\}}}{\nu(V,L)^2}\right.\right. \\
&\qquad \left.\left. - kP_{ho,V}B_D(-D,V,S)^{k-1}\frac{S1_{\{SV<2D\nu(V,L)\}}}{\nu(V,L)^2} - B_D(-D,V,S)^k\Theta(V)\right)\right] \\
&\quad + \frac{1}{(1+P_{ho})^2}\Psi(L)E_n\left[\frac{\partial\nu(V,L)}{\partial L}\Theta(V)\right] \\
&= E_n\left[\frac{\partial\nu(V,L)}{\partial L}\left(\frac{-k\frac{S^k1_{\{SV<(D-X)\nu(V,L)\}}}{\nu(V,L)^{k+1}} - kP_{ho,V}\frac{S^k1_{\{SV<2D\nu(V,L)\}}}{\nu(V,L)^{k+1}} - B_D(-D,V,S)^k\Theta(V)}{1+P_{ho}}\right.\right. \\
&\qquad \left.\left. + \frac{\Psi(L)\Theta(V)}{(1+P_{ho})^2}\right)\right].
\end{aligned} \tag{3.18}$$

Thus the derivatives will have the form

$$\frac{db^{(k)}}{dL} = E_{n,V}\left[\frac{\partial \nu(V,L)}{\partial L} E_{n,X,S}[\Gamma^{(k)}(X,V,S,\nu(V,L))]\right] \tag{3.19}$$

for some functions $\Gamma^{(k)}$. Thus for fixed velocities, i.e., when $V \equiv \bar{v}$

$$\frac{db^{(k)}}{dL} = \frac{\partial \nu(\bar{v},L)}{\partial L} E_{n,X,S}[\Gamma^{(k)}(X,\bar{v},S,\nu(\bar{v},L))]. \tag{3.20}$$

**Claim :** The term $E_{n,X,S}[\Gamma^{(k)}(X,\bar{v},S,\nu(\bar{v},L))]$ is strictly negative.

**Proof of Claim:** For any velocity $v$, $B_D(X,v,S) \leq B_D(-D,v,S)$ for all $(X,S)$. Thus, for fixed velocities,

$$\Psi(L) \leq (1+P_{ho,\bar{v}}) E_{n,S}[B_D(-D,\bar{v},S)^k].$$

Further $P_{ho} = E_{n,V}[P_{ho,V}] = P_{ho,\bar{v}}$. Hence the sum of the last two inner terms of the equation (3.18),

$$E_{n,X,V}\left[-\frac{B_D(-D,V,S)^k \Theta(V)}{1+P_{ho}} + \frac{\Psi(L)\Theta(V)}{(1+P_{ho})^2}\right] \leq 0.$$

The remaining two inner terms of (3.18) are always negative and this proves the Claim. □

From (3.20), by the virtue of the Claim, derivative $db^{(k)}/dL$ is zero only at a zero of $\partial\nu/\partial L$. But $\partial\nu/\partial L$ has a unique zero at the maximizer $L^*_\nu(\bar{v})$. Thus $L^*_\nu(\bar{v})$ is the only zero of all the service time moments.

## 3.9 Appendix P: Calculations for Small cell queue

### 3.9.1 P.1 Small cell Handover Speed Distribution

In the following the event $\{\partial \ \ arrival\} = \{\partial\}$ means that the arrival in cell 0 was at the boundary (i.e., it was an handover arrival from cell -1), the event $\{int \ \ arrival\} = \{int\}$ meant a new arrival in cell 0 and the event $\{ho\}$ implies the cell 0 active user has reached the next boundary before finishing his service (i.e, has to be handed over to cell 1). Note that an handover from cell 0 occurs either due to an already handed over call or new call, whose service could not be completed before reaching the (other) boundary of cell 0 and thus:

$$Prob(ho) = \frac{\lambda_L P_{ho} + \lambda_{hL} P_{ho,\partial}}{\lambda_L + \lambda_{hL}}.$$

Note that $Prob(ho) \rightarrow 1$ as $L \rightarrow 0$ (as both $P_{ho}$, $P_{ho,\partial}$ converge to 1). With the above notations, the fixed point equation for the handover speed density can be obtained as:

$$\begin{aligned} f_{h,V}(v) &= Prob(V \in vdv | ho \ \ from \ \ cell \ \ 0) \\ &= Prob(V \in vdv \ \ int \ \ arrival | ho) + Prob(V \in vdv \ \ ho \ \ arrival | ho) \\ &= \frac{Prob(ho \ \ V \in vdv \ \ int) + Prob(V \in vdv \ \ \partial \ \ ho)}{Prob(ho)} \\ &= \frac{Prob(ho | int \ \ v) Prob(V \in vdv | int) Prob(int)}{Prob(ho)} \\ &\qquad + \frac{Prob(ho | \partial \ \ v) Prob(V \in vdv | \partial) Prob(\partial)}{Prob(ho)} \\ &= \frac{P_{ho,v} f_{n,V}(v)}{Prob(ho)} \frac{\lambda_L}{\lambda_L + \lambda_{hL}} + \frac{P_{ho,v,\partial} f_{h,V}(v)}{Prob(ho)} \frac{\lambda_{hL}}{\lambda_L + \lambda_{hL}} \end{aligned}$$

Solving (when $f_{n,V}(v) = 1/V_{max}$ for all $v \leq V_{max}$) the handover speed density is,

$$f_{h,V}(v) = \frac{\frac{\lambda_L}{Prob(ho)(\lambda_L + \lambda_{hL}) V_{max}} P_{ho,v}}{1 - \lambda_{hL} \frac{P_{ho,v,\partial}}{Prob(ho)(\lambda_L + \lambda_{hL})}}$$

Thus as $L$ tends to zero ($P_{ho,v}, P_{ho,v,\partial}, Prob(ho), P_{ho} \rightarrow 1$), the speed of an handover arrival will tend to uniform distribution. This effect is seen faster if the packet sizes $S$ are larger.

### 3.9.2 P.2 Small cell Stability Factor

The stability factor $\rho$ in a Small cell queue is given by:

$$\begin{aligned} \rho_{pico}(L) &= \frac{(\lambda_L + \lambda_{hL}) b^{(1)}}{K} = \frac{1}{K} E_V \left[ \lambda_L E_X[E_S[B_L(S, X, V)]] + \lambda_{hL} E_S[B_L(S + B_h, 0, V)] \right] \\ &\approx \frac{1}{K} \left( \lambda_L (L - E[X]) + L \lambda_{hL} \right) E\left[\frac{1}{V}\right] \quad = \quad \frac{1}{K} \left( \lambda_L \frac{L}{2} + L \lambda_{hL} \right) E\left[\frac{1}{V}\right] \\ &= c_0 L (\lambda_L + 2\lambda_{hL}) \overset{\text{as } P_{ho} \approx 1}{\approx} c_0 L \lambda_L \frac{1 - P_{ho,\partial} + 2}{1 - P_{ho,\partial}} \approx \frac{2 c_0 \lambda_L L}{1 - P_{ho,\partial}} \end{aligned}$$

where $c_0 = 1/(2K)E[1/V]$ is a constant independent of $L$. Using the approximation in (3.15),

$$\rho_{pico}(L) = \frac{\lambda p L}{K D \mu \nu(\bar{v}_{inv}, L)}. \tag{3.21}$$

### 3.9.3 P.3 Drop probability

The drop probability $P_{D,pico}$ considering possible Small cell drops can be calculated as below:

$$\begin{aligned} P_{D,pico} &= Prob(\text{ Call ever Dropped before completion} | \text{ Call is picked up}) \\ &= P_{ho} \left( P_{ho,D}(1 - P_{Busy,pico}) + P_{Busy,pico} \right) + (1 - P_{ho})(0) \end{aligned}$$

where $P_{ho,D}$ is defined as the probability of call drop at any of the future instances of handovers, given that the current handover (first handover in the context of the above equation) is successful. Because of the memoryless nature of $S$, this probability does not depend upon the number of the handover. Probability $P_{ho,D}$ can be calculated by first conditioning on the event that the call is completed in the current cell (call it as $\mathcal{C}$) and then on the event that the call is not picked up in the next cell (call it as S). Note that $P_{ho}(\mathcal{C}^c) = P_{ho,\partial}$ and $P_{ho}(\mathcal{S}|\mathcal{C}^c) = P_{ho,fail} = 1 - P_{Busy,pico}$. Thus, by conditioning

$$\begin{aligned}
P_{ho,D} &= P_{ho}(\text{ Call dropped } \cap \mathcal{C}) + P_{ho}(\text{ Call dropped } \cap \mathcal{C}^c) \\
&= 0 + P_{ho,\partial} P_{ho}(\text{ Call dropped } |\mathcal{C}^c) \\
&= P_{ho,\partial}\left(P_{ho}(\text{ Call dropped } \cap \mathcal{S}^c|\mathcal{C}^c) + P_{ho}(\text{ Call dropped } \cap \mathcal{S}|\mathcal{C}^c)\right) \\
&= P_{ho,\partial}\left(P_{ho,D}(1 - P_{Busy,pico}) + 1 P_{Busy,pico}\right) \\
&\overset{Solving}{=} \frac{P_{Busy,pico} P_{ho,\partial}}{1 - P_{ho,\partial}(1 - P_{Busy,pico})} \text{ and hence,} \\
P_{D,pico} &= \frac{P_{ho} P_{Busy,pico}}{1 - P_{ho,\partial}(1 - P_{Busy,pico})} \approx \frac{P_{Busy,pico}}{1 - P_{ho,\partial} + P_{Busy,pico}}. \qquad (3.22) \\
&= \frac{P_{Busy,pico}}{\frac{\mu L \nu(\bar{v}_{inv}, L)}{\bar{v}_{inv}} + P_{Busy,pico}} = \frac{P_{Busy,pico}}{\frac{\mu(\eta(L) - \bar{v}_{inv} B_h)}{\bar{v}_{inv}} + P_{Busy,pico}}
\end{aligned}$$

## 3.10 Publications

1. Veeraruna Kavitha, Sreenath Ramanath and Eitan Altman, "Spatial queuing analysis for design and dimensioning of Picocell networks with mobile users", Elseviers Performance Evaluation, Volume 68, Issue 8.
2. Sreenath Ramanath, Veeraruna Kavitha, Eitan Altman, "Spatial queuing analysis for mobility in pico cell networks", proceedings of WiOpt 2010, May 31-Jun 04, Avignon, France.
3. Sreenath Ramanath, Veeraruna Kavitha, Eitan Altman, "Impact of mobility on call block, call drops and optimal cell size in small cell networks", proceedings of the Workshop on Indoor and Outdoor Femto Cells (IOFC'10), Sep 26, Istanbul, Turkey.

# Chapter 4

# Fair Assignment of Base Station Locations

**Contents**



## 4.1 Introduction

In a cellular network, models used to derive analytic expressions for capacity, coverage, etc, often assume base station (BS) locations to be at the center of the cell. Such a model brings in a regular geometry to the problem being addressed and many a times results in closed-form analytic expressions for metrics of interest.

While, this indeed facilitates analysis, the actual throughput achievable at the BS, tends to vary significantly, depending on the BS placement and cell geometry. The regular geometric model with a centrally located BS is a good model, when one assumes uniform density of users. But, today's cellular networks have concentration of users, for example hot-spots or indoor-outdoor partitions that offer various levels of attenuation to radio signals, not to mention the ever present channel fading and shadowing effects above this.

The goal of our research is to place the BS in a manner which is optimal for any general fairness criterion; that of $\alpha$-fairness [105, 147], which addresses popular fairness criterion like global, proportional, harmonic and max-min fairness. We show that the $\alpha$-fair BS location varies continuously with fairness parameter $\alpha$ and moves close to the center of the cell as $\alpha$ increases asymptotically. This implies that, the regular geometric models (which place BS at the center) have max-min fair BS placement.

We further observe (via some numerical examples) that $\alpha$-fair BS location varies significantly based on system parameters like path-loss, noise variance. However we show that the max-min fair BS placement is close to the center of the cell irrespective of the system parameters.

To bring in the importance of BS placement, we consider cells which are completely outdoors or which have indoor-outdoor partitions (Split cells). We consider cases where user density can be uniform or tend to increase along the cell (a simplistic model for a hot spot) [99]. We consider cases where adjacent cells can use the same frequency or different frequencies. For the later case, we look at the problem of fair assignment of two base stations (BS), where the cell gets divided into sub-cells from the users' perspective based on SINR association criteria. We limit our study to free space path-loss and the analysis with fading and shadowing would be our subsequent focus.

We derive simplified expressions for $\alpha$-fair objective functions using large population limits , i.e., as the number of users become large. We use Strong Law of Large Numbers (SLLN) to replace summation of large number of terms in the objective function with appropriate expected value almost surely (AS). The expected value is expressed as integrals. Using these large population limits, we obtain both theoretical and numerical results. The $\alpha$-fair BS locations obtained are optimal for almost all realizations of the users locations.

We begin our study by introducing our model and review the generalized $\alpha$-fair fairness criterion in Section 4.2. In Section 4.3, we derive the $\alpha$-fair placement criterion under large population limits. In Section 4.4, we analyze $\alpha$-fair placement of base stations as $\alpha$ increases asymptotically and come across some interesting insights. In Section 4.5, we study the $\alpha$-fair BS locations for the case of a) an outdoor cell and b) a mixed partition cell (split-cell) via some numerical examples. Next, in Section 4.6, we derive the $\alpha$-fair BS locations for an outdoor cell which has two BS. We conclude our study in Section 4.7.

## 4.2 Our model and assumptions

Our focus is on communication in the uplink (UL) direction. Large number, $N$, of users are located on i.i.d. locations on the line segment $[-D, D]$. The line segment is divided into cells of length $L$ and one or more base stations, each of unit height, are placed in every cell.

The placement of the BS(s) is the issue that we address in this work. One is usually

interested in maximizing global throughput (the sum throughput due to all users) at each BS, i.e., place the BS(s) such that the global throughput is maximized. However, maximizing the global throughput can result in starving the users at a far away location, which in turn can reduce the network efficiency. Hence, several fairness criterion have been suggested and implemented in various network architectures ([105, 147]).

In [105] it is shown that all these fairness criterion are special cases of a generalized fairness concept: the $\alpha$-fairness . Given a positive constant $\alpha \neq 1$, consider for example the problem of determining $z$ so as to maximize

$$\max_{z\in[0,L]} \sum_{x_i\in[0,L]} \frac{\theta(x_i,z)^{1-\alpha}}{1-\alpha} \tag{4.1}$$

where, $\theta(x_i, z)$ is the throughput at the BS located at $z$ from a user located at $x_i$. Note that the above objective function is defined over the convex set $[0, L]$. Further, when the objective function is concave (we will show in later sections that this is the case most of the times) and the constraints are linear, this defines a unique allocation which we call the $\alpha$-fair allocation. It turns out that $\alpha$-fairness gives global optimum for $\alpha \to 0$, proportional fairness when $\alpha \to 1$, harmonic (delay minimization) fairness indexHarmonic fairnessfor $\alpha = 2$ and max-min fairness when $\alpha \to \infty$.

We begin our study by first deriving explicit expressions for power, throughput and $\alpha$-fair placement criterion under large population limits (as the number of users become large). Throughout the work, we use large population limits for analytical purposes. The idea is similar to fluid limits (see [90]), where summation of large number of terms is approximated by appropriate integrals.

## 4.3 Large population limits and problem statement

Large number, $N$, of users are located at $\{X_i\}_{i\leq N}$, where the locations $X_i$ of the users are i.i.d., according to some probability measure $P(dx) = \lambda(x)dx$. *We assume that each user uses the same power for transmission.* Without loss of generality, the total power in the system equals 1 and hence the power used by each user is $1/N$.

We first consider the case of a single BS in the cell and compute the total power received, throughput achievable and the $\alpha$-fair placement for the BS under large population limits. The case of two base stations follows in a similar way and is addressed in Appendix.

### 4.3.1 Power computation :

The power received at a BS located at $z$ from a user at $X_i$ is given by,

$$P(X_i, z) = \frac{1}{N}(1 + (z - X_i)^2)^{\frac{-\beta}{2}}.$$

Thus the total power received at the BS is

$$P_{tot}(z) = \sum_{i=1}^{N} P(X_i, z) = \frac{1}{N} \sum_{i=1}^{N} (1 + (z - X_i)^2)^{\frac{-\beta}{2}}.$$

This is a random power. By the Strong Law of Large Number (SLLN) this converges P-a.s. to a constant limit

$$\begin{aligned} \lim_{N \to \infty} P_{tot}(z) &= E\left[(1 + (z - X_i)^2)^{\frac{-\beta}{2}}\right] \\ &= \int_{-D}^{D} (1 + (z - x)^2)^{\frac{-\beta}{2}} \lambda(x) dx. \end{aligned} \tag{4.2}$$

Hence for large values of $N$ one can approximate $P_{tot}(z)$ almost surely with the above integral.

### 4.3.2 Throughput computation:

The signal to interference noise ratio (SINR) at the BS located at $z$ from a user at $X_i$ is

$$SINR(X_i, z) = \frac{P(X_i, z)}{\sigma^2 + P_{tot}(z) - P(X_i, z)},$$

where $\sigma^2$ is the noise variance. In the above, $P_{tot}(z)$ is approximated $P-$almost surely by a constant value, i.e., by the integral of (4.2). However $SINR(X_i, z)$ is still random because of the term $P(X_i, z)$. The Shannon capacity or throughput achievable at the BS located at $z$ from a user at $X_i$ is

$$\theta(X_i, z) = \log(1 + SINR(X_i, z))$$

Considering a receiver with an adapted filter and using the approximation $\log(1 + x) \approx x$ (for smaller values of $x$), the throughput achievable is

$$\theta(X_i, z) = \frac{P(X_i, z)}{\sigma^2 + P_{tot}(z) - P(X_i, z)}$$

The total (global) throughput achievable at the BS from all the users in the cell of interest is:

$$\begin{aligned} f(z) &= \sum_{i=1}^{N} 1_{\{X_i \in [0,L]\}} \theta(X_i, z) \\ &= \frac{1}{N} \sum_{i=1}^{N} 1_{\{X_i \in [0,L]\}} \frac{(1 + (z - X_i)^2)^{\frac{-\beta}{2}}}{\sigma^2 + P_{tot}(z) - P(X_i, z)} \\ &\approx \frac{1}{N} \sum_{i=1}^{N} 1_{\{X_i \in [0,L]\}} \frac{(1 + (z - X_i)^2)^{\frac{-\beta}{2}}}{\sigma^2 + P_{tot}(z)}, \end{aligned}$$

as for large values of $N$, $P(X_i, z)$ is negligible in comparison with $P_{tot}(z)$. Again, the above Random sum can be approximated using Strong Law of Large Numbers whenever the number of users inside the cell is large P-almost surely, giving rise to the following large population approximation:

$$\begin{aligned} f(z) &\approx E\left[I_{\{X_1 \in [0,L]\}} \psi(X_1, z)\right] \\ &= \int_0^L \psi(x,z)\lambda(x)dx \text{ with} \\ \psi(x,z) &:= \frac{(1+(z-x)^2)^{\frac{-\beta}{2}}}{\sigma^2 + P_{tot}(z)}. \end{aligned} \tag{4.3}$$

### 4.3.3 $\alpha$-fair placement criterion :

The $\alpha$-fair objective function of (4.1) in a similar way can be approximated almost surely under large population limits by:

$$\tilde{f}_\alpha(z) := N^\alpha \frac{1}{1-\alpha} \int_0^L \psi(x,z)^{(1-\alpha)} \lambda(x)dx$$

Thus $\alpha$-fair placement of the BS is given by,

$$\begin{aligned} z^*(\alpha) &= \arg\max_z \tilde{f}_\alpha(z) \\ &= \arg\max_z f_\alpha(z) \text{ where} \\ f_\alpha(z) &:= \frac{1}{1-\alpha} \int_0^L \psi(x,z)^{(1-\alpha)} \lambda(x)dx. \end{aligned} \tag{4.4}$$

Important point to note here is that, *for almost all realizations of the locations of the users the objective function is approximated by the constant integral and hence $z^*(\alpha)$ is optimal $\alpha$-fair location for almost all users locations.*

### 4.3.4 Problem statement

Now with this background, we pose the following problems:

1. Find BS location $z$ so as to maximize global throughput $f(z)$. See large population limit (4.3).

2. Find the $\alpha$-fair BS location $z^*$ which maximizes $f_\alpha(z)$ for various fairness criterion. See large population limit (4.4)

In subsequent sections, we analyze and apply the $\alpha$-fair placement criterion to obtain BS locations which are both optimal and fair in various cellular environments considered.

## 4.4 Analysis : Single BS Placement

We notice that both the global throughput (large population limit (4.3)) and the $\alpha$-fair placement (large population limit (4.4)) of the BS is dependent on the total power received $P_{tot}(z)$ at the BS, which in-turn depends on its location $z$. In many cases, the total-power received can be assumed independent of the location of the BS, whenever the cell size is small (which is typical of pico cells). This for example is true for cells with user density $\lambda(x)$ being symmetric about $\frac{L}{2}$ (uniform being the trivial case) and completely located outdoors.

The above assumption simplifies analysis to a good extent and is considered in the first subsection, while an approximate analysis is given in the following subsection without this assumption.

We consider asymptotic analysis in this section and hence consider only the cases with $\alpha > 1$. For notational simplicities, we redefine $f_\alpha(z)$ of equation (4.4) after dropping the division by $(1-\alpha)$ factor and now,

$$z^*(\alpha) := \arg\max_{z\in[0,L]} \left(-f_\alpha(z)\right).$$

### 4.4.1 $P_{tot}(z)$ is independent of BS location $z$ :

As $P_{tot}(z)$ is independent of $z$, the $\alpha$-fair location is obtained by minimizing the function,

$$\bar{f}_\alpha(z) := \int_0^L \left(1+(z-x)^2\right)^{-\frac{\beta}{2}(1-\alpha)} \lambda(x)dx$$

We can easily show that $\bar{f}_\alpha(z)$ is concave in $z$. We also have joint continuity in $(\alpha, z)$ by Bounded Convergence theorem. Hence, by maximum theorem [137] under convexity, we get

**Lemma 4.4.2.** *The function $z^*(\alpha)$ is continuous in $\alpha$.* ■

By differentiability of $\bar{f}_\alpha(.)$,

$$g(\alpha, z^*(\alpha)) = 0.$$

where with $\gamma := \frac{\beta}{2}(\alpha-1)-1$ (for some appropriate $c \neq 0$),

$$\begin{aligned} g(\alpha,z) &:= c\frac{\partial \bar{f}_\alpha(z)}{\partial z} \\ &= \int_0^L (z-x)\left(1+(z-x)^2\right)^\gamma \lambda(x)dx. \end{aligned}$$

If $z < \frac{L}{2}$ then,

$$\left(1+\left(\frac{L}{2}\right)^2\right)^{-\gamma} g(\alpha,z) = \int_0^{\frac{L}{2}+z}(z-x)\left(\frac{1+(z-x)^2}{1+(\frac{L}{2})^2}\right)^{\gamma}\lambda(x)dx + \int_{\frac{L}{2}+z}^{L}(z-x)\left(\frac{1+(z-x)^2}{1+(\frac{L}{2})^2}\right)^{\gamma}\lambda(x)dx$$

tends to $-\infty$ as $\alpha \uparrow \infty$, because the first term tends to zero while the later tends to $-\infty$ (by bounded convergence theorem). Therefore there exists $\alpha_0 > 0$ such that, $g(\alpha, z) < 0$ and hence such that,

$$g(\alpha,z) \neq 0 \text{ for all } \alpha > \alpha_0.$$

Similarly if $z > \frac{L}{2}$ then, $\left(1+\left(\frac{L}{2}\right)^2\right)^{-\gamma} g(\alpha,z)$ tends to $\infty$ as $\alpha \uparrow \infty$ and hence we have,

$$g(\alpha,z) \neq 0 \text{ for all } \alpha > \alpha_0(z) \text{ whenever } z \neq \frac{L}{2}.$$

However $\left(1+\left(\frac{L}{2}\right)^2\right)^{-\gamma} g\left(\alpha, \frac{L}{2}\right) \to 0$ as $\alpha \uparrow \infty$.

In fact, we have (by monotonicity arguments) for all $z_0 < \frac{L}{2}$:

$$g(\alpha,z) \neq 0 \text{ for all } \alpha > \alpha_0(z_0), z \in [0,z_0] \cup [L-z_0, L],$$

and hence the optimizer lies in a smaller interval around $\frac{L}{2}$ for all larger values of $\alpha$ and thus we get the following:

**Lemma 4.4.3.** *For every $\epsilon < \frac{L}{2}$ there exists an $\alpha_0(\epsilon)$ (depending upon $\epsilon$), such that for all $\alpha > \alpha_0(\epsilon)$*

$$z^*(\alpha) \in \left[\frac{L}{2}-\epsilon, \frac{L}{2}+\epsilon\right].$$

*i.e, the optimizer lies in a smaller interval around $\frac{L}{2}$ for all larger values of $\alpha$. That is, $z^*_\alpha(z) \to \frac{L}{2}$ as $\alpha \to \infty$.* ■

Whenever the density $\lambda(x)$ is symmetric about $\frac{L}{2}$ within the cell $[0, L]$, using similar derivative arguments one can get,

**Lemma 4.4.4.** *The partial derivatives under symmetric conditions, for all $\alpha$*

$$\left.\frac{\partial \bar{f}_\alpha(z)}{\partial \alpha}\right|_{z=z_0} = 0$$

*and hence optimal locations for all $\alpha$ are at $\frac{L}{2}$.* ■

**Summary of the results :**

1. $\alpha$-fair location is continuous in $\alpha$ (by Lemma 4.4.2).

2. When the density function $\lambda$ is symmetric about $\frac{L}{2}$ then by Lemma 4.4.4 all the $\alpha$-fair locations are at the center of the cell.

3. If density is not symmetric about $\frac{L}{2}$ then, by Lemma 4.4.3, the $\alpha$-fair locations tend to $\frac{L}{2}$ as $\alpha$ tends to infinity.

4. Lemma 4.4.3, 4.4.4 are correct as long as the support of measure $\lambda$ contains both the end points, i.e., $\{0, L\} \subset supp(\lambda)$. If not, the same results are true with $\frac{L}{2}$ replaced with $length(supp(\lambda))/2$.

### 4.4.5 $P_{tot}(z)$ is dependent on BS location $z$ :

Next, we consider cases when the total power $P_{tot}(z)$ is dependent on base station location $z$. This is true for cases with non-symmetric user densities, cells with partitions, etc.

Let $h(.;z)$ represent the following parametrized function :

$$h(x;z) := \left(\sigma^2 + P_{tot}(z)\right)\left(1 + (x-z)^2\right)^{\frac{\beta}{2}}.$$

and let $||h(.;z)||_p$ represent its $L_p$ norm with respect to the probability measure $\frac{\lambda(x)dx}{\int_0^L \lambda(x)dx}$.

With the above definitions, for $\alpha > 1$, we can equivalently write the optimal $\alpha$-fair location as,

$$z^*(\alpha) \quad = \quad \arg\min_{z\in[0,L]} ||h(.;z)||_{\alpha-1}.$$

As $\alpha \to \infty$, $||h(.;z)||_{\alpha-1} \to ||h(.;z)||_\infty$ and one can show that,

$$\lim_{\alpha\to\infty} z^*(\alpha) \approx \arg\min_{z\in[0,L]} ||h(.;z)||_\infty.$$

Since,

$$\begin{aligned} ||h(.;z)||_\infty &= \left(\sigma^2 + P_{tot}(z)\right) \sup_{x\in[0,L]} \left(1 + (x-z)^2\right)^{\frac{\beta}{2}} \\ &= \left(\sigma^2 + P_{tot}(z)\right)\left(1 + (\max\{z, L-z\})^2\right)^{\frac{\beta}{2}} \\ &= \left(\sigma^2 + P_{tot}(z)\right)\max\{(1+z^2)^{\frac{\beta}{2}}, (1+(L-z)^2)^{\frac{\beta}{2}}\}, \end{aligned}$$

the asymptotic $\alpha$-fair location approximately equals:

$$\lim_{\alpha\to\infty} z^*(\alpha) \quad \approx \quad \arg\min_{z\in[0,L]} \left(\sigma^2 + P_{tot}(z)\right) \max\{z, L-z\}\,.$$

and note that,

$$\max\{z, L-z\} = 1_{\left\{z\geq \frac{L}{2}\right\}}(z) + 1_{\left\{z< \frac{L}{2}\right\}}(L-z)\,.$$

Clearly if $P_{tot}(z)$ was independent of $z$, asymptotic $\alpha$-fair location would be at $\frac{L}{2}$.

In the subsequent sections, we consider some interesting examples and show the validity of the results of this Section. We also derive many more interesting conclusions using the large population limits of Section 4.3 for those examples.

## 4.5 Optimal and fair placement of a single BS

In this section, we consider two cases. In the first, we consider an outdoor cell, while in the second, we consider a cell which spans over both indoor and outdoor environment (split-cell).

### 4.5.1 Outdoor cell

An outdoor cell is typically characterized by a cell placed in open environment/free space, i.e., the signals from the users are attenuated only due to path-loss. We assume a cellular deployment which uses the same frequency throughout. i.e., the power received from the entire line segment $[-D, D]$ will interfere with the power received from the user under consideration.

By Lemma 4.4.4, the $\alpha$-fair solution for uniform user density, $\lambda(x) \equiv 1/2D$, is trivial (all the $\alpha$-fair locations are at the center of the cell $\frac{L}{2}$).

Next, we consider another interesting case where user density $\lambda(x) = x$; to mimic a simplistic hot-spot (i.e, the user density proportionally increases towards the hot-spot, which is located around $L$). Figure 4.1 depicts the scenario. We want to place the BS such that the locations are optimal and fair.

By Lemma 4.4.2, the $\alpha$-fair location varies continuously w.r.t. $\alpha$. Also, by Lemma 4.4.3 and discussions in Section 4.4.5, the $\alpha$-fair locations should tend to $\frac{L}{2}$ as $\alpha$ increases to infinity. We will indeed show that this is the case in the following numerical example. We further make some more interesting observations.

**Numerical example:** We evaluate equation (4.4) for some typical cases: for $\alpha = 0$ (global), $\alpha = 0.99$ (proportional), $\alpha = 2$ (harmonic) and $\alpha = 128$ (max-min). The example considers cell length $L = 10$, noise variance $\sigma^2 = 1$ and path-loss exponent $\beta =$

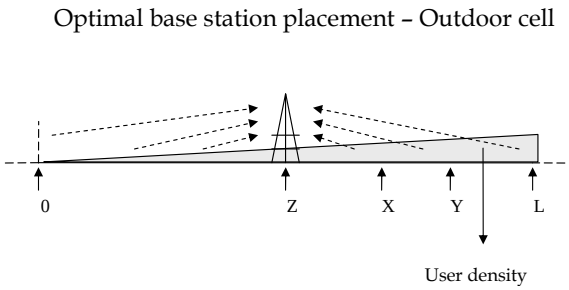


***Figure 4.1:*** *Open-cell: BS located at z, user density* $\lambda(x) = x$

2, 4. Fig 4.2 - 4.4 shows example plots for the $\alpha$-fair objective functions $f(z)$ (equation (4.3) corresponding to $\alpha = 0$), $f_\alpha(z)$ (equation (4.4)) for $\alpha = 0.99, 2$. Note that the case of global fairness ($\alpha = 0$) is also the case which maximizes sum throughput. Also, note that Fig 4.2 gives the global throughput as function of BS location $z$.

We compute the $\alpha$-fair BS placement for increasing values of $\alpha$. In Figure 4.5, we plot the $\alpha$-fair BS location as a function of $\alpha$. As given by Lemma 4.4.2 the $\alpha$-fair location is continuous in $\alpha$. We further, observe that the BS location shifts rapidly going from optimally fair to proportionally fair and finally tends to $\frac{L}{2}$ for being max-min fair.

We tabulate normalized throughput (ratio of the global throughput with BS at $\alpha$-fair location, $z^*(\alpha)$ to the maximum achievable global throughput, i.e., the total throughput achieved when BS is placed at $z^*(0)$) achievable for these $\alpha$-fair BS locations in Table 4.1.

We show the impact of path-loss factor $\beta$ and noise variance $\sigma^2$ on the optimally-fair BS placement in Table 4.2 and Table 4.3, respectively. In those tables $f(z; \beta, \sigma^2)$ represents the global throughput when BS is placed at $z$ and with path-loss factor $\beta$ and noise variance $\sigma^2$.

***Table 4.1:*** *Outdoor cell: The $\alpha$-fair BS locations and normalized throughput. User density* $\lambda(x) = x$, $L = 10$ *and path-loss* $\beta = 2$

| $\alpha$-fairness | BS lox | Normalized throughput |
|---|---|---|
| global ($\alpha = 0$) | 7.4 | 1.000 |
| proportional ($\alpha = 0.99$) | 6.8 | 0.998 |
| harmonic ($\alpha = 2$) | 6.3 | 0.995 |
| max-min ($\alpha = 128$) | 5.0 | 0.981 |

**Observations:**

a. We observe that the placement of BS affects the throughput achievable in case of an outdoor cell, modeling a hot-spot.

b. The BS location shifts rapidly going from globally fair to proportionally fair and finally settles close to $\frac{L}{2}$ for being max-min fair (Refer Figure 4.5). This is an interest-

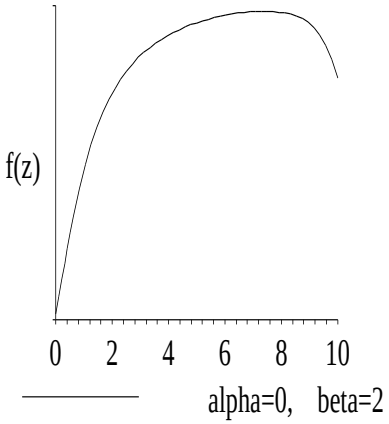


***Figure 4.2:*** *Open cell: Global throughput (4.3) as a function of the BS location. User density $\lambda(x) = x$ and $\beta = 2$)*

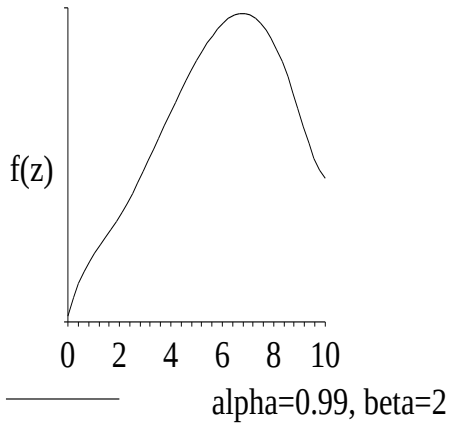


***Figure 4.3:*** *Proportional fair objective function $f_\alpha$ given by (4.4) with $\alpha \approx 1$, as function of BS location z. User density $\lambda(x) = x$, $\beta = 2$)*

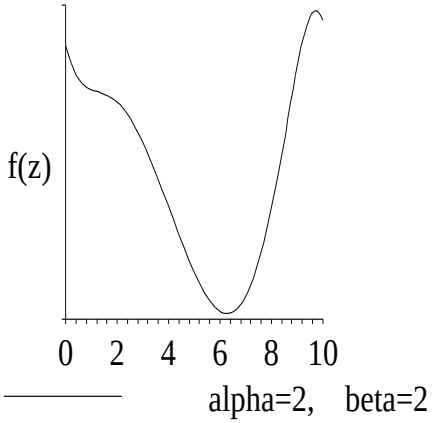


***Figure 4.4:*** *Harmonic fair objective function $f_\alpha$ given by (4.4) with $\alpha = 2$, as function of BS location z. User density $\lambda(x) = x$, $\beta = 2$)*

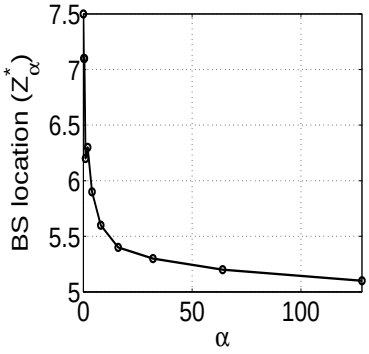


***Figure 4.5:*** *α-fair BS location, $z^*(\alpha)$ as a function of $\alpha$.*

***Table 4.2:*** *Outdoor cell: BS placement for globally-fair throughput for various path-loss $\beta$. User density $\lambda(x) = x$ and $L = 10$*

| Path-loss $\beta$ | BS lox | Throughput ratio $f(z^*(0);\beta,1)/f(z^*(0);2,1)$ (w.r.t $\beta = 2$) |
|---|---|---|
| 2 | 7.4 | 1.00 |
| 4 | 8.2 | 0.99 |
| 6 | 8.8 | 0.98 |

***Table 4.3:*** *Outdoor cell: BS placement for globally-fair throughput for various noise-variance $\sigma^2$. User density $\lambda(x) = x$, path-loss $\beta = 2$ and $L = 10$*

| Noise variance $\sigma^2$ | BS lox | Throughput ratio $f(z^*(0);2,\sigma^2)/f(z^*(0);2,1)$ (w.r.t $\sigma^2 = 1$) |
|---|---|---|
| $\frac{1}{4}$ | 6.9 | 1.05 |
| 1 | 7.4 | 1.00 |
| 4 | 7.9 | 0.58 |

ing observation which implies that the regular geometric models, which assume centrally placed BS are actually positioned to be max-min fair. But, such assumptions does not seem to impact the throughput achievable as seen in this case. The max-min fair throughput is just about 2% below the maximum achievable global throughput. For the other fair locations, the reduction in throughput is quite negligible (Refer Table 4.1).

c. The optimal throughput does not seem to be sensitive to path-loss (Refer Table 4.2)

d. The achievable optimal throughput is very sensitive to noise variance $\sigma^2$. A four fold increase in noise variance degrades the throughput by 40%.

### 4.5.2 Indoor-outdoor cell (Split-cell)

In this section, we consider a cell which covers both indoor and outdoor environments, partitioned by solid structures like walls etc. We consider a cell which has a single partition or wall, located at $y$ within the cell $[0, L]$ and offers an attenuation of $\eta$ dB. Here again, a single BS of unit height is located at $z$. The scenario is depicted in figure 4.6. We want to find BS locations which optimize various fairness criterion.

The total power, throughput and $\alpha-$ fair objective function for this case can be derived exactly in the same way as before to obtain the following large population limits :

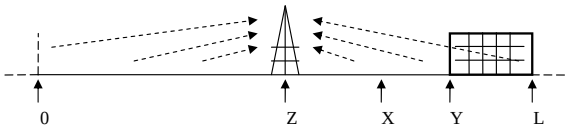


***Figure 4.6:*** *Split-cell: BS located at z, wall located at y*

$$
\begin{aligned}
P_{tot}(z) &= \int_{[-D,y)\cup[L,D)} (1+x^2)^{\frac{-\beta}{2}} \lambda(x)dx \\
&\quad +\eta \int_y^L (1+x^2)^{\frac{-\beta}{2}} \lambda(x)dx \qquad (4.5)\\
f(z) &= \int_0^y \psi(x,z)\lambda(x)dx + \eta \int_y^L \psi(x,z)\lambda(x)dx \qquad (4.6)\\
f_\alpha(z) &= \frac{1}{1-\alpha}\left[\int_0^y \psi(x,z)^{1-\alpha}\lambda dx \right.\\
&\quad \left. +\eta \int_y^L \psi(x,z)^{1-\alpha}\lambda(x)dx\right] \qquad (4.7)
\end{aligned}
$$

The equations (4.5), (4.6) and (4.7) are similar respectively to (4.2), (4.3) and (4.4) if $\lambda(x)$ is replaced by (an appropriate constant multiple of) $\lambda(x)(1_{\{x\notin[y,L]} + \eta 1_{\{x\in[y,L]\}})$. Hence the results of Section 4.4 hold good here also. From Lemma 4.4.2, Lemma 4.4.3 and discussions in Section 4.4.5, we would expect the $\alpha$-fair location to vary continuously w.r.t. $\alpha$ and tend close to $\frac{L}{2}$ as $\alpha$ increases asymptotically. We shall validate this via the following numerical example for uniform user density, i.e., for $\lambda(x) \equiv 1/2D$.

**Numerical example:** We evaluate equation (4.7) for some typical cases: for global ($\alpha = 0$), proportional ($\alpha = 0.99$) and harmonic ($\alpha = 2$) fairness with path-loss exponent $\beta = 2, 4$, noise variance $\sigma^2 = 1$ ,wall attenuation $\eta = 12dB$ and wall located at $y = 0.75L$. The results are presented in Fig 4.7 - 4.12. Note that the case of global fairness ($\alpha = 0$) is also the case which maximizes global throughput (Refer Figure 4.7).

In Fig 4.10, we plot the $\alpha$-fair BS location as a function of $\alpha$. We observe that the BS location shifts rapidly going from globally fair to proportionally fair and finally converges close to $\frac{L}{2}$ for max-min fair.

Next, we plot global throughput as a function of BS location $z$ and wall location $y$ in Figure 4.11. In Fig 4.12, we show global throughput as a function of BS location $z$ and attenuation $\eta$.

Table 4.4 tabulates the normalized throughput achievable for various $\alpha$-fairness criterion along with the $\alpha$-fair BS locations, While, Tables 4.5 and 4.6 tabulate the BS placement for globally-fair throughput for various path-loss factors $\beta$ and noise variance $\sigma^2$, respectively.

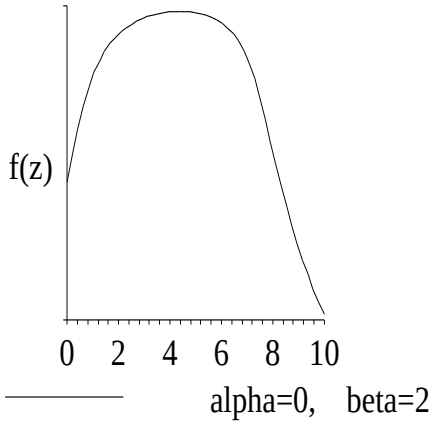


***Figure 4.7:*** *Split-cell: Global throughput (Objective function $f_\alpha(z)$ (4.6) with $\alpha = 0$) as a function of BS location $z$.*

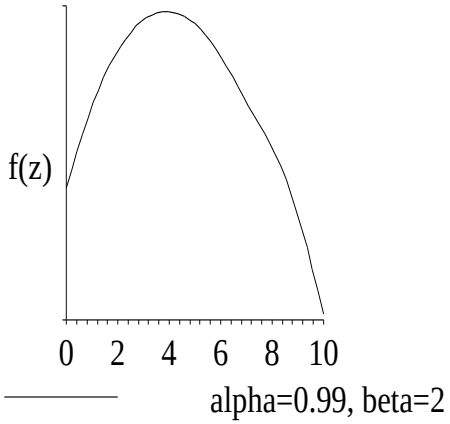


***Figure 4.8:*** *Split-cell: Objective function $f_\alpha(z)$ (4.7) for proportional fairness ($\alpha = 0.99$) as a function of BS location $z$.*

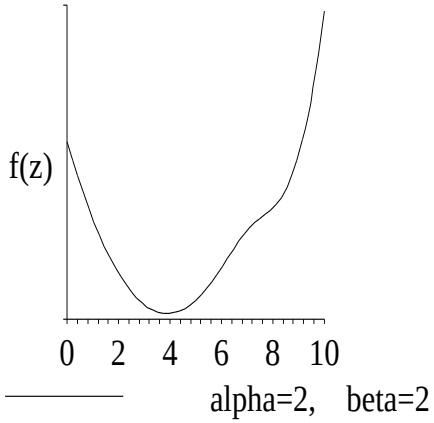


***Figure 4.9:*** *Split-cell: Objective function $f_\alpha(z)$ (4.7) for harmonic fairness ($\alpha = 2$) as a function of BS location $z$.*

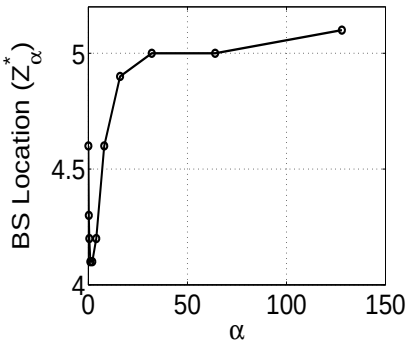


***Figure 4.10:*** *Split-cell: $\alpha$-fair BS location $z^*(\alpha)$ as a function of $\alpha$.*

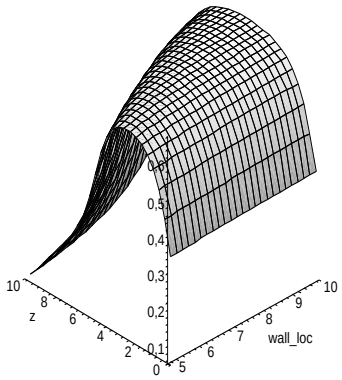


***Figure 4.11:*** *Split-cell: Global throughput (4.6) as a function of BS location z and wall location y.*

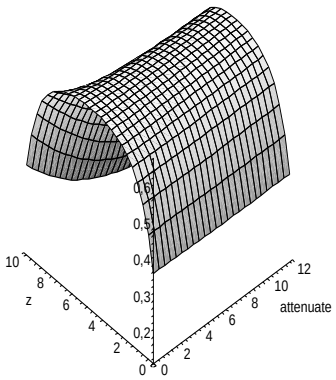


***Figure 4.12:*** *Split-cell: Global throughput (4.6) as a function of BS location z and attenuation $\eta$.*

In the tables, Normalized throughput, $f(z;\beta,\sigma^2)$ represent similar terms as in the previous section.

***Table 4.4:*** *Split-cell: The $\alpha$-fair BS location and normalized throughput. User density $\lambda(x) \equiv 1/2D$, $L = 10$, $y = 0.75L$, path-loss $\beta = 2$ and wall attenuation $\eta = 12dB$*

| $\alpha$-fairness | BS lox | Normalized throughput |
|---|---|---|
| global ($\alpha = 0$) | 4.35 | 1.0000 |
| proportional ($\alpha = 0.99$) | 3.90 | 0.9983 |
| harmonic ($\alpha = 2$) | 3.88 | 0.9983 |
| max-min ($\alpha = 128$) | 5.00 | 0.9981 |

**Observations:**

a. We observe that the BS location shifts rapidly going from globally fair to proportionally fair and finally settles at $\frac{L}{2}$ for being max-min fair (Refer Figure 4.10).

b. Further, we observe that the placement of BS does not seem to affect the throughput achievable in case of an indoor-outdoor cell.

c. The price in throughput is negligible and the deployment can satisfy various fairness criterion (Refer Table 4.4).

d. The reduction in globally-fair throughput is quite significant (as much as 20%) with an increase in path-loss factor $\beta$ (Refer Table 4.5)

e. As the indoor portion of the split-cell reduces, the globally-fair throughput response tends to become flat. (Refer Figure 4.11)

***Table 4.5:*** *Split cell: BS placement for globally-fair throughput for various path-loss $\beta$. User density $\lambda(x) \equiv 1/2D$ and $L = 10$*

| Path-loss $\beta$ | BS lox | Throughput ratio $f(z^*(0);\beta,1)/f(z^*(0);2,1)$ (w.r.t $\beta = 2$) |
|---|---|---|
| 2 | 4.35 | 1.00 |
| 4 | 4.15 | 0.93 |
| 6 | 4.00 | 0.83 |

***Table 4.6:*** *Split cell: BS placement for globally-fair throughput for various noise-variance $\sigma^2$. User density $\lambda(x) \equiv 1/2D$, path-loss $\beta = 2$ and $L = 10$*

| Noise variance $\sigma^2$ | BS lox | Throughput ratio $f(z^*(0);2,\sigma^2)/f(z^*(0);2,1)$ (w.r.t $\sigma^2 = 1$) |
|---|---|---|
| $\frac{1}{4}$ | 4.65 | 1.30 |
| 1 | 4.35 | 1.00 |
| 4 | 3.90 | 0.21 |

f. Wall attenuation does not seem to alter the globally-fair BS placement much, though one can observe a significant reduction in throughput initially (Refer Figure 4.12)

g. The reduction in globally-fair throughput is quite drastic (as much as 80%) with an increase in noise variance $\sigma^2$ (Refer Table 4.6)

## 4.6 Optimal and fair placement of two BS in an outdoor cell

In this section we consider optimal placement of two BS in a single cell for various $\alpha$-fair criterion. We consider a new scenario in this section, that of a single isolated cell (i.e., no interference from the other cells). One can easily study a single BS problem with this new scenario and vice versa using the tools of this research. This new scenario is considered for covering all varieties of the settings/scenarios.

Users are located on this segment with density $\lambda(x), x \in [-L, L]$. Assume $BS_1$ and $BS_2$ are located at $z_1$ and $z_2$, respectively and uses the same frequency and cooperate with each other. Further, we assume that the neighboring cells do not use the same frequency. The users associate themselves with one of the two base stations which maximize their SINR.

Under these assumptions, we first calculate the global (sum) throughput from all the users associated with a particular BS. Under cooperative setting, the sum of these two global throughputs would be the appropriate criteria for optimization. In Appendix A, we derived simplified expressions $f(z_1, z_2)$, for this sum of global throughputs, under

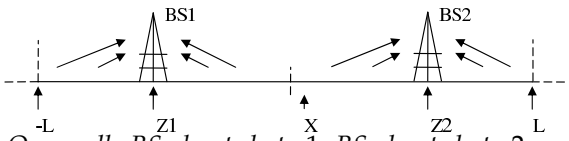


***Figure 4.13:*** *Open-cell: $BS_1$ located at z1, $BS_2$ located at z2, user density $\lambda(x) \equiv 1/2D$*

large population limits. In a similar way, a general simplified $\alpha$-fair objective function $f_\alpha(z_1, z_2)$ is also derived in the same Appendix. We now re-state the problem of Section 4.3 for the two BS case :

1. Find location $(z_1, z_2)$ so as to maximize global throughput $f(z_1, z_2)$. See large population limit (4.10) (same as $f_\alpha(z_1, z_2)$ with $\alpha = 0$) of Appendix A.

2. Find the $\alpha$-fair location $(z_1^*, z_2^*)$ which maximizes $f_\alpha(z_1, z_2)$ for various fairness criterion. See large population limit (4.13) of Appendix A.

We reproduce from Appendix A, the $\alpha$-fair location as given by

$$(z_{1\alpha}^*, z_{2\alpha}^*) = \arg\max_{z_1, z_2} f_\alpha(z_1, z_2).$$

where,

$$\begin{aligned} f_\alpha(z_1, z_2) &= \frac{\int_{C(z_1,z_2)} \left(1 + (x - z_1)^2\right)^{1-\alpha} \lambda(x) dx}{\left(\sigma^2 + P_{tot}(z_1)\right)^{1-\alpha}} \\ &\quad + \frac{\int_{C(z_1,z_2)^c} \left(1 + (x - z_2)^2\right)^{1-\alpha} \lambda(x) dx}{\left(\sigma^2 + P_{tot}(z_2)\right)^{1-\alpha}} \end{aligned}$$

**Numerical example:** We evaluate equation (4.13) for some typical cases: for global ($\alpha = 0$), proportional ($\alpha = 0.99$) and harmonic ($\alpha = 2$) fairness with path-loss exponent $\beta = 2$ and noise variance $\sigma^2 = 1$.

For the numerical analysis we have assumed that the BS are located symmetrically about the origin to ease the SINR based user association criteria (See Appendix A).

The results are presented in Fig 4.14 - 4.17. Note that the case of global fairness ($\alpha = 0$) is also the case of sum global throughput (Refer Figure 4.14). From the plots, we observe that the BS locations for global fairness is $(-6.5, 6.5)$.

In Figure 4.17, we plot the $\alpha$-fair $BS_2$ location as a function of $\alpha$. We observe that the BS location shifts rapidly going from globally fair to proportionally fair and finally settles at $L/2$ for being max-min fair. In a similar way, the $BS_1$ tends to $-L/2$ as $\alpha$ increases to infinity. In fact, we observe that the BS location exhibits max-min fair placement for values of $\alpha = 8$ and beyond.

Table 4.7 tabulates the normalized throughput achievable for various $\alpha$-fairness criterion along with the $\alpha$-fair BS locations, while, Table 4.8 tabulates the BS placement for

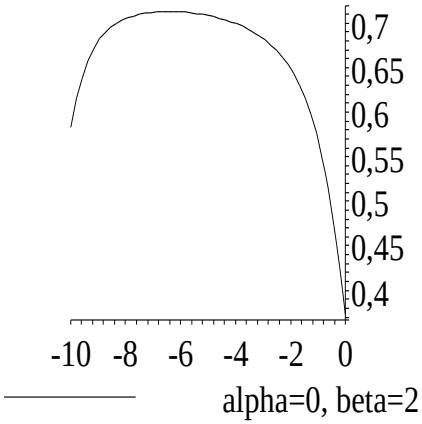


***Figure 4.14:*** *Outdoor cell, two BS: Global throughput (objective function $f_\alpha(z_1, z_2)$ with $\alpha = 0$) as a function of $BS_1$ location ($z_2 = -z_1$)).*

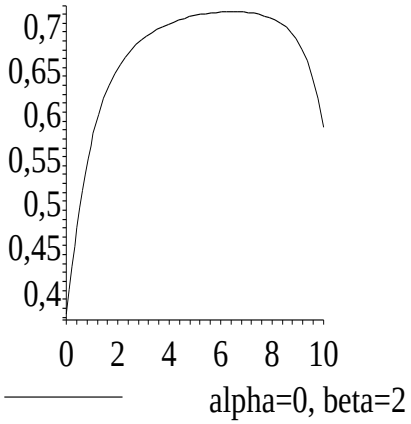


***Figure 4.15:*** *Outdoor cell, two BS: Global throughput (objective function $f_\alpha(z_1, z_2)$ with $\alpha = 0$) as a function of $BS_2$ location($z_1 = -z_2$)).*

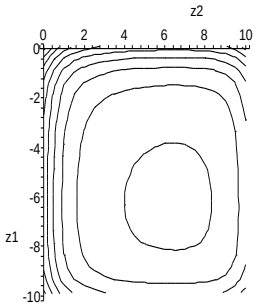


***Figure 4.16:*** *Outdoor cell, two BS: 3-D contour plot of global throughput (objective function $f_\alpha(z_1, z_2)$ with $\alpha = 0$) as a function of BS locations $(z_1, z_2)$*

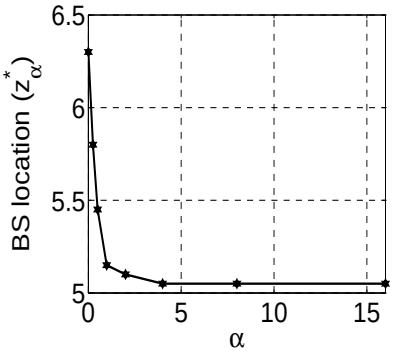


***Figure 4.17:*** *Outdoor cell, two BS: α-fair BS location $z_2^*(\alpha)$ as a function of α (Placement of $BS_2$ shown here).*

globally-fair throughput for various path-loss factors $\beta$.

***Table 4.7:*** *The $\alpha$-fair BS location(s) and normalized throughput for outdoor cell with two BS, user density $\lambda(x) = 1/2L$, $L = 10$, path-loss $\beta = 2$*

| $\alpha$-fairness | $BS_1$ lox | $BS_2$ lox | Normalized |
|---|---|---|---|
| global ($\alpha = 0$) | -6.45 | 6.45 | 1.0000 |
| proportional ($\alpha = 0.99$) | -5.15 | 5.15 | 0.9970 |
| harmonic ($\alpha = 2$) | -5.10 | 5.10 | 0.9950 |
| max-min ($\alpha = 128$) | -5.05 | 5.05 | 0.9940 |

***Table 4.8:*** *Outdoor cell with two BS: BS placement for globally-fair throughput for various path-loss $\beta$. User density $\lambda(x) = 1/2L$ and $L = 10$*

| Path-loss $\beta$ | $BS_1$ lox | Throughput ratio $f(z^*(0);\beta,1)/f(z^*(0);2,1)$ (w.r.t $\beta = 2$) |
|---|---|---|
| 2 | -6.45 | 1.00 |
| 4 | -5.55 | 0.85 |
| 6 | -5.35 | 0.76 |

**Observations:**

a. We observe that the BS locations shift rapidly going from globally fair to proportionally fair and finally settles at $(-L/2, L/2)$ for being max-min fair. In fact, the BS location exhibits max-min fair placement for values of $\alpha = 8$ onwards (Refer Figure 4.17).

b. Further, we observe that the placement of BS does not seem to affect the throughput achievable in case of an outdoor cell with two BS.

c. The price in throughput is negligible and the deployment can satisfy various fairness criterion.

d. The globally-fair throughput reduces by 25% with an increase in path-loss exponent $\beta$ from 2 to 6 (Refer Table 4.8).

e. We observe that for a uniform distribution of users, when placing fairly two base stations on the segment $[-L, L]$, the distance between the stations decrease as $\alpha$ increases (Refer Table 4.7). In particular, we note that a model similar to this has been already studied in [13] where the equilibrium location was computed in a non-cooperative context (each base station tries to maximize its own throughput) instead of the fair location. As in the fair placement case that we study here, it was shown there that the equilibrium distance is also closer than the distance corresponding to the globally fair location. As an example, the equilibrium location of the BS that corresponds to the data of Fig 4.17 here is 5.5 in Table 1 of [13] (the globally fair being around 6.4). This means that the non-cooperative equilibrium location is fairer than the globally fair one - it corresponds to the $\alpha$ fair placement where $\alpha$ is seen from Fig 4.17 to be around 0.5.

Further, our work can be extended to find the $\alpha$-fair BS locations when multiple BS are to be located on a line segment or on a 2D grid. This is a step towards optimal BS placement to satisfy various fairness criteria when a macro-cell is divided into a number of small cells. For example, the optimal placement of BS in pico-cell networks.

## 4.7 Conclusions and future perspectives

We studied the problem of optimal BS placement, optimal for various $\alpha$-fair criterion in cellular networks. We considered simple 1D models which characterize both indoor and outdoor cellular environments with mixed partitions. We derived explicit expressions for $\alpha$-fair criterion under large population limits. These limits were used to obtain the theoretical asymptotic analysis of the $\alpha$-fair locations. We show that the $\alpha$-fair locations converge close to center of the cell as $\alpha$ increases to infinity (which basically represents the max-min fair location).

The large population limits were also used to numerically compute BS locations which satisfy global, proportional, harmonic and max-min fairness. For the models considered, we presented results via plots and tables to show the variations in achievable throughput for the different fairness criterion. We also confirmed, via numerical examples, that the $\alpha$-fair locations converge to the center of the cell as $\alpha$ tends to infinity.

We next considered a two base station optimal placement problem again for various $\alpha$-fair criterion. We obtained large population limits under cooperative setting and using this we showed, via numerical examples, that the $\alpha$-fair BS locations converge to a pair of locations which divide the cell once again into equal regions.

We used large population limits to get some initial insight into this problem. Another interesting way would be to consider users distributed accord to a stochastic process according to some distribution. One can try to use tools like stochastic geometry or the likes to analyze the problem. It would be interesting to compare the results obtained via different approaches in the limiting regime.

## 4.8 Appendix A : Large population limits - power, throughput and $\alpha$-fair placement of two base stations:

In Section 4.3, we derived power, throughput and $\alpha$-fair placement expressions for a single BS located in the cell. In this appendix section, we derive the same for two BS. For simplicity, we consider the cell of interest to span $[-L, L]$. Also, in this case, we assume that neighboring cells use different frequencies (i.e, there is no frequency reuse)

As before, the power from a user located at $X_i$ received at $BS_1$ located at $z_1$ is

$$P(X_i, z_1) = \frac{1}{N}(1 + (z_1 - X_i)^2)^{\frac{-\beta}{2}}.$$

The total power received at $BS_1$ under large population limits is

$$P_{tot}(z_1) = \int_{-L}^{L} (1 + (z_1 - x)^2)^{\frac{-\beta}{2}} \lambda(x) dx,$$

assuming no frequency re-use.

The throughput (which is approximately equal to the SINR in case of an adaptive filter) at $BS_1$ is

$$\theta(X_i, z_1) \approx SINR(X_i, z_1) = \frac{P(X_i, z_1)}{\sigma^2 + P_{tot}(z_1)}.$$

Similarly throughput at $BS_2$ is,

$$\theta(X_i, z_2) \approx SINR(X_i, z_2) = \frac{P(X_i, z_2)}{\sigma^2 + P_{tot}(z_2)}.$$

The user at $X_i$ will associate itself with $BS_1$ if

$$SINR(X_i, z_1) > SINR(X_i, z_2). \tag{4.8}$$

Let

$$C(z_1, z_2) := \{x : SINR(x, z_1) \geq SINR(x, z_2)\} \tag{4.9}$$

represent the set of users which associate themselves with $BS_1$.

Under cooperative setting, the total sum throughput received at both the base stations is,

$$\begin{aligned} f(z_1, z_2) &:= \frac{1}{(1-\alpha)} \sum_{i=1}^{N} \Big[ \theta(X_i, z_1) 1_{\{X_i \in C(z_1, z_2)\}} \\ &\qquad + \theta(X_i, z_2) 1_{\{X_i \in C(z_1, z_2)^c\}} \Big] \end{aligned} \tag{4.10}$$

The $\alpha$-fair solution in this case is given by the BS location pair $(z_1^*, z_2^*)$ which maximizes $f_\alpha$ where,

$$\begin{aligned} \tilde{f_\alpha}(z_1, z_2) &:= \frac{1}{(1-\alpha)} \sum_{i=1}^{N} \Big[ \theta(X_i, z_1) 1_{\{\in C(z_1, z_2)\}} \\ &\qquad + \theta(X_i, z_2) 1_{\{X_i \in C(z_1, z_2)^c\}} \Big]^{1-\alpha} \\ &= \frac{1}{(1-\alpha)} \sum_{i=1}^{N} \Big[ \theta(X_i, z_1)^{1-\alpha} 1_{\{X_i \in C(z_1, z_2)\}} \\ &\qquad + \theta(X_i, z_2)^{1-\alpha} 1_{\{X_i \in C(z_1, z_2)^c\}} \Big] \end{aligned}$$

which under large population limits is approximated by,

$$\tilde{f}_\alpha(z_1, z_2) \tag{4.11}$$

$$\approx \frac{N^\alpha}{(1-\alpha)} \left[ \frac{\int_{C(z_1,z_2)} \left(1 + (x - z_1)^2\right)^{1-\alpha} \lambda(x) dx}{\left(\sigma^2 + P_{tot}(z_1)\right)^{1-\alpha}} + \frac{\int_{C(z_1,z_2)^c} \left(1 + (x - z_2)^2\right)^{1-\alpha} \lambda(x) dx}{\left(\sigma^2 + P_{tot}(z_2)\right)^{1-\alpha}} \right] \tag{4.12}$$

Thus $\alpha$ fair placement of the two BS is given by,

$$\begin{aligned} (z_{1\alpha}^*, z_{2\alpha}^*) &= \arg\max_{z_1, z_2} f_\alpha(z_1, z_2) \text{ where} \\ f_\alpha(z_1, z_2) &= \\ (-1)^{1_{\{\alpha>1\}}} &\left[ \frac{\int_{C(z_1,z_2)} \left(1 + (x - z_1)^2\right)^{1-\alpha} \lambda(x) dx}{\left(\sigma^2 + P_{tot}(z_1)\right)^{1-\alpha}} \right. \\ &\left. + \frac{\int_{C(z_1,z_2)^c} \left(1 + (x - z_2)^2\right)^{1-\alpha} \lambda(x) dx}{\left(\sigma^2 + P_{tot}(z_2)\right)^{1-\alpha}} \right] \end{aligned} \tag{4.13}$$

## 4.9 Publications

1. Sreenath Ramanath, Eitan Altman, Vinod Kumar, Veeraruna Kavitha, Laurent Thomas, "Fair assignment of base stations in cellular networks", proceedings of the 22nd World Wireless Research Forum (WWRF'09), May 5-7, Paris, France.

# Chapter 5

# Asymptotic Analysis of Precoded Small Cell Networks

**Contents**



## 5.1 Introduction

While, dividing a macro-cell into multiple small cells enhances the capacity, the spatial dimension has been exploited in the recent past to enhance the capacity further. It is now well established that Multiple antenna at the transmitter ($N_t$) and the receiver ($N_r$) achieve capacity gains which grow linearly as $\min(N_t, N_r)$.

Recently, the MIMO broadcast channel [155, 160, 77], where, a multi-antenna base station, transmitting on M antennas to K single antenna users is shown to achieve capacity gains which grow linearly as $min(M, K)$, provided the transmitter and receivers all know the channel [85]. To achieve this, several methods have been proposed among which linear precoders offer a good compromise between complexity and performance trade-off [156],[110].

Further, MIMO based systems have been studied in the framework of multi-cell networks. In a multi-cell scenario, the achievable sum-rate in the downlink, diminishes due to interference from neighboring base stations. Thus increasing the number of antennas at the base-stations does not necessarily yield a linear increase in capacity.

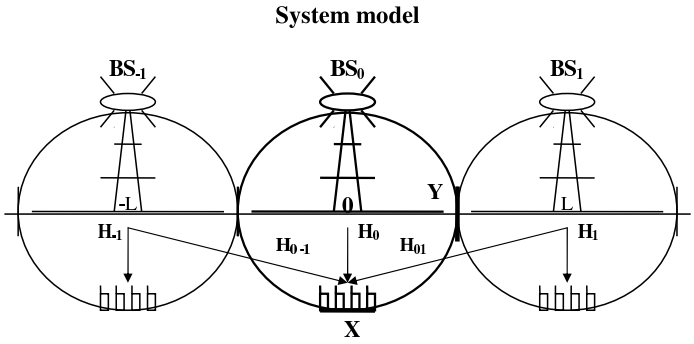


***Figure 5.1:*** *System model: multi-cell network. BS with M antennas, serving K users. Users at X experience nominal interference and users at Y experience high interference*

Frequency reuse and various forms of interference co-ordination [132], [86] have been proposed to achieve linear growth in capacity.

In our contribution, we want to asses exactly the effect of multi-cell interference in MIMO based small cell networks. Small cells being in close proximity experience higher levels of interference, which would bring down the capacity gains significantly. We want to study the impact of multi-cell interference when base-stations employ linear precoding techniques, such as channel inversion (CI) at the base station.

As mentioned before, linear precoding techniques such as channel inversion(CI) and regularized channel inversion (RCI) offer a convenient trade-off between complexity and achievable sum-rate performance [77, 110]. The behavior of CI in uncorrelated MIMO broadcast channels (MIMO-BC) has already been studied in [77, 110] for i.i.d. Gaussian channels. In particular, the authors in [77] showed that CI achieves linear growth in multiplexing-gain. Further, authors in [48], extended the case to include antenna correlations due to dense packing of the antennas at the transmitter. The analysis carried out considers single cell systems and they show that for the case of CI, the sum-rate is maximized when the number of antennas $M$ on the BS is equal to the number of users $K$.

For the multi-cell case, the problem of interference co-ordination in uplink has been discussed at length in [152]. In [86], authors address downlink macro-diversity in cellular systems. They study the potential benefit of base-station (BS) cooperation for downlink transmission in a modified Wyner-type [166] multicell model. They compare various precoders and obtain analytical sum rate expressions for both the fading and the non-fading case. They demonstrate via monte-carlo simulations the effectiveness of linear precoding. Authors in [155] suggests that asymptotically, equal power allocation is optimal when the channel is i.i.d. Gaussian.

In our work, we are interested in studying the impact of interference from adjacent base stations, which is more pronounced in MIMO based small cell networks on

the achievable sum-rate capacity. We consider multiple-input multiple-output (MIMO) multi-cell systems, each cell composed of a transmitter equipped with $M$ antennas and $K$ single-antenna receivers. We consider Wyner-type cellular models in our study. We neglect the effects of channel correlation due to densely packed antennas at the base-station transmitter, with a view to keep the analysis tractable.

The analytic expressions of the sum-rates for CI are derived by applying recent tools from random matrix theory (RMT . These expressions are independent of the specific channel realizations.

In our study, we find that

- The achievable sum-rate is significantly diminished by the effect of multi-cell interference in MIMO based small cell networks.
- The sum-rate capacity tends to grow sub-linearly with respect to the number of base-station antennas as long as the interference is non-zero.
- Also, there is an optimal number of users for a given number of antennas at the transmitter, which maximizes the sum-capacity. This depends on the interference level and the transmit power at the base-station.

The remainder of this chapter is organized as follows: Section 5.2 briefly reviews various tools of random matrix theory which will be used in later derivations. Section 5.3 introduces the multi-cell system model. In Section 5.4 we study channel inversion precoding. Section 5.5 provides simulation results which are shown to corroborate the theoretical derivations. Finally in Section 5.6 we provide our conclusions.

*Notations:* In the following, boldface lower-case symbols represent vectors, capital boldface characters denote matrices ($\mathbf{I}_N$ is the $N \times N$ identity matrix). The Hermitian transpose is denoted $(\cdot)^{\mathsf{H}}$. The operator $\mathrm{tr}[\mathbf{X}]$ represents the trace of matrix $\mathbf{X}$. The eigenvalue distribution of an Hermitian random matrix $\mathbf{X}$ is $\mu_{\mathbf{X}}(x)$. The symbol $\mathrm{E}[\cdot]$ denotes expectation. The derivative of a function $f(x)$ of a single variable $x$ is denoted $f'(x)$. All logarithms are base-2 logarithms.

## 5.2 Random Matrix Theory Tools

In this work, we are interested in the behavior of large random Hermitian matrices, and particularly in the asymptotic distribution of their eigenvalues. Specifically, the eigenvalue distribution of large Hermitian matrices converges, in many practical cases, to a definite probability distribution, hereafter called the *empirical distribution* of the random matrix, when the matrix dimensions grow to infinity.

A tool of particular interest in this work is the *Stieltjes transform* $\mathcal{S}_{\mathbf{X}}$ of a large Hermitian non-negative definite matrix $\mathbf{X}$, defined on the half the space $\mathbb{C}^+ = \mathbb{C} \setminus R^+ = \{z \in \mathbb{C}, \mathrm{Re}(z) < 0\}$, as

$$\mathcal{S}_{\mathbf{X}}(z) = \int_0^{+\infty} \frac{1}{\lambda - z} \mu_{\mathbf{X}}(\lambda) d\lambda \tag{5.1}$$

where $\mu_{\mathbf{X}}$ is the empirical distribution of $\mathbf{X}$.

Couillet et al. [47] derived a fixed-point expression of the Stieltjes transform for Gaussian matrices with correlations in the following theorem,

**Theorem 5.2.0.1.** *Let the entries of the $K \times M$ matrix $\mathbf{W}$ be i.i.d. Gaussian with zero mean and variance $1/M$. Let $\mathbf{X}$ and $\mathbf{Q}$ be respectively $K \times K$ and $M \times M$ Hermitian non-negative definite matrices with eigenvalue distributions $\mu_{\mathbf{X}}$ and $\mu_{\mathbf{Q}}$. We impose further that the largest eigenvalues of $\mathbf{X}$ and $\mathbf{Q}$ are bounded independently of $K$, $M$. Let $\mathbf{Y}$ be an $K \times K$ Hermitian matrix with the same eigenvectors as $\mathbf{X}$ and let $f$ be some function mapping the eigenvalues of $\mathbf{X}$ to those of $\mathbf{Y}$. Let $z \in \mathbb{C}^+ = \mathbb{C} \setminus \mathbb{R}^+$. Then, for $M$, $K$ large with $K/M = 1/\beta$, the Stieltjes transform $\mathcal{S}_{\mathbf{H}}(z)$ of $\mathbf{H} = \mathbf{X}^{1/2}\mathbf{W}\mathbf{Q}\mathbf{W}^{\mathsf{H}}\mathbf{X}^{1/2} + \mathbf{Y}$ is given*

$$\mathcal{S}_{\mathbf{H}}(z) = \int \left( f(x) + x \int \frac{q \cdot \mu_{\mathbf{Q}}(q) \dot{q}}{1 + \frac{1}{\beta} q \mathcal{T}_{\mathbf{H}}(z)} - z \right)^{-1} \mu_{\mathbf{X}}(x) \dot{x} \tag{5.2}$$

*where $\mathcal{T}_{\mathbf{H}}$ is a solution of the fixed-point equation*

$$\mathcal{T}_{\mathbf{H}}(z) = \int x \left( f(x) + x \int \frac{q \cdot \mu_{\mathbf{Q}}(q) \dot{q}}{1 + \frac{1}{\beta} q \mathcal{T}_{\mathbf{H}}(z)} - z \right)^{-1} \mu_{\mathbf{X}}(x) \dot{x} \tag{5.3}$$

An immediate corollary, when only right-correlation is considered, unfolds naturally as follows,

**Corollaire 5.2.1.** *[130] Let the entries of the $K \times M$ matrix $\mathbf{W}$ be i.i.d. Gaussian with zero mean and variance $1/M$. Let $\mathbf{Y}$ be an $K \times K$ Hermitian non-negative matrix with eigenvalue distribution $\mu_{\mathbf{Y}}(x)$. Moreover, let $\mathbf{Q}$ be a $M \times M$ non-negative definite matrix with eigenvalue distribution $\mu_{\mathbf{Q}}(x)$, such that the eigenvalues of $\mathbf{Q}$ are bounded irrespectively of $M$. Then, for large $K$, $M$, such that $K/M = \alpha$, the Stieltjes transform on $\mathbb{C}^+$ of the matrix*

$$\mathbf{H} = \mathbf{W}\mathbf{Q}\mathbf{W}^{\mathsf{H}} + \mathbf{Y} \tag{5.4}$$

*verifies*

$$\mathcal{S}_{\mathbf{H}}(z) = \mathcal{S}_{\mathbf{Y}} \left( z - \int \frac{q}{1 + \alpha q \mathcal{S}_{\mathbf{H}}(z)} \mu_{\mathbf{Q}}(q) dq \right) \tag{5.5}$$

## 5.3 System model and assumptions

We discuss the system model in this section. We consider a multi-cell Wyner-type model, for example as shown in figure (5.1). For simplicity and to be able to keep the analysis tractable, we consider a three-cell network. The cell at the center is our reference. The users in this cell experience interference from the neighboring base stations as shown. Each cell serves $K$ users from a base-station with $M$ antennas. We assume that the base station antennas are uncorrelated. The information from the base-station to its user set is precoded assuming perfect channel state information at the transmitter (CSIT). i.e, each base station knows perfectly the channel towards the users in its

cell, but not the interfering channels. Users receive desired signal plus interference signals from adjacent base stations. We assume channel inversion (CI) precoding at the transmitter. The transmitted signals from the base stations undergo Rayleigh fading and path-loss. Further, we assume that the channel is constant for some interval long enough for the transmitter to learn and use it until it changes to a new value. We are interested in the behavior of the system and its sum-rate capacity. Many of our results are obtained for large limits, because the limiting results are often tractable. Nevertheless, we often consider $M, K$ small in our simulation examples. Further, all users are assumed to have the same average (but not instantaneous) received signal power, so our model assumes that the users are similar distances from the base station and are not in deep shadow fades.

## 5.4 Channel inversion precoding

Channel inversion precoding, also referred to as zero-forcing (ZF) precoding, annihilates all the inter-user interference by performing an inversion of the channel matrix $\mathbf{H}$ at the transmitter. We begin our analysis with the single cell case, which is discussed in detail in [48], [77], and further we shall consider the multi-cell case.

### 5.4.1 Single cell

Without loss of generality, we consider cell 0. The signal received by users in this cell is

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}. \tag{5.6}$$

where, $\mathbf{H}$ is the $K \times M$ channel matrix with zero-mean unit-variance i.i.d complex Gaussian entries, $\mathbf{x} = \mathbf{G}\mathbf{s}$ is the transmit vector obtained by linear precoding of the symbol vector $\mathbf{s}$ with the precoding matrix $\mathbf{G}$. Symbol $s_k \in \mathbf{s}$ for any user $k$ is complex Gaussian with zero mean and unit variance. The $M \times K$ linear precoding matrix is defined as

$$\mathbf{G} = \alpha \mathbf{H}^{\mathsf{H}} \left( \mathbf{H}\mathbf{H}^{\mathsf{H}} \right)^{-1}. \tag{5.7}$$

where $\alpha$ is chosen appropriately to satisfy the total transmit power constraint $\mathrm{tr}(E[\mathbf{x}\mathbf{x}^H]) \leq \mathrm{tr}(\mathbf{G}\mathbf{G}^H) \leq P$.

Now the received vector in Cell 0

$$\mathbf{y} = \alpha \mathbf{s} + \mathbf{n}. \tag{5.8}$$

The parameter $\alpha$ which satisfies the transmit power constraint and depends only on the channel realization $\mathbf{H}$ is given by

$$\alpha^2 = \frac{P}{\mathrm{tr}\left( (\mathbf{H}\mathbf{H}^{\mathsf{H}})^{-1} \right)} \tag{5.9}$$

The SNR (signal to noise ratio) for any user $k$ is defined as

$$\eta_k = \frac{E_s\left[|\alpha s_k|^2\right]}{E|n_k|^2} = \frac{\alpha^2}{\sigma^2}. \tag{5.10}$$

is independent of the selected user. $\sigma^2$ is the noise variance.

The ergodic capacity for user $k$ is

$$C_k = \log(1+\eta_k). \tag{5.11}$$

and the sum-rate is

$$R_{ci} = \sum_{k=1}^{K} \log(1+\eta_k). \tag{5.12}$$

### 5.4.2 Asymptotic analysis for a single-cell

$\alpha$ is a function of $\mathbf{H}$ and as $M, K \to \infty$, $\alpha$ tends to a constant. Thus the sum-rate can be written as

$$\mathcal{R}_{\text{ci}} = K \log\left(1+\eta_k\right) \tag{5.13}$$

Let us denote $\mathbf{H}' = \frac{1}{\sqrt{M}}\mathbf{H}$. It follows from (5.9) that When $M$ is large with $M/K = \beta$,

$$\begin{aligned}
\frac{1}{M}\text{tr}\left(\mathbf{H}'\mathbf{H}'^{\mathsf{H}}\right)^{-1} &= \frac{1}{M}\sum_{i=1}^{K}\frac{1}{\lambda_i} \\
&= \frac{K}{M}\left(\frac{1}{K}\sum_{i=1}^{K}\frac{1}{\lambda_i}\right) \\
&= \frac{K}{M}\int\frac{1}{\lambda}\left(\frac{1}{K}\sum_{i=1}^{K}\delta(\lambda-\lambda_i)\right)d\lambda \\
&= \frac{1}{\beta}\int\frac{1}{\lambda}\mu^{K}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(\lambda)d\lambda \\
&= \frac{1}{\beta}\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0)
\end{aligned}$$

As a consequence, for large $(K, M)$

$$\frac{\alpha^2}{\sigma^2} \to \frac{\rho\beta}{\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0)}, \text{ where } \rho = P/\sigma^2 \tag{5.14}$$

and the sum-rate is

$$\mathcal{R}_{\text{ci}} = K\log\left(1+\frac{\rho\beta}{\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0)}\right) \tag{5.15}$$

According to Corollary 5.2.1, $\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0)$ is the solution of[1]

$$\begin{aligned}\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0) &= \left(\int \frac{\lambda}{1+\frac{\lambda}{\beta}\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0)}\mu(\lambda)d\lambda\right)^{-1} \\ &= \left(\int \frac{\lambda\ \ \delta(\lambda-1)}{1+\frac{\lambda}{\beta}\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0)}\right)^{-1} \\ &= \left(1+\frac{\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0)}{\beta}\right) \end{aligned} \tag{5.16}$$

Solving for $\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0)$ yields,

$$\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0) = \frac{\beta}{(\beta-1)} \tag{5.17}$$

and the sum-rate is re-written as

$$\mathcal{R}_{\text{ci}} = K\log\left(1+\rho(\beta-1)\right) \text{ for } \beta \geq 1 \tag{5.18}$$

The rate-per-antenna is

$$\frac{\mathcal{R}_{\text{ci}}}{M} = \frac{1}{\beta}\log\left(1+\rho(\beta-1)\right). \tag{5.19}$$

As $\beta \to 1$, $\mathcal{R}_{\text{ci}}/M \to 0$, which implies that the sum rate of channel inversion does not increase linearly with $M$ (or $K$)

### 5.4.3 Optimizer $\beta^*$ for the single cell

Following [77] we now look for a value $\beta^\star$ of the ratio $M/K$ such that, for a fixed number of transmit antennas $M$, the sum-rate $\mathcal{R}_{\text{ci}}(\beta)$ is maximized. By differentiating eqn (5.19) with respect to $\beta$ and setting the derivative to zero, $\beta^*$ is the solution of the implicit equation

$$\rho\beta^* = (1+\rho(\beta^*-1))\,log\,(1+\rho(\beta^*-1)) \tag{5.20}$$

[1] it is important to note here that we slightly misapply Corollary 5.2.1 since the result is only proven valid outside for any $z > 0$.

### 5.4.4 Multi-cell

In this section, we study the effect of multi-cell interference. Without loss of generality, we consider users in Cell 0 affected by interference from adjacent base-stations. We consider a 3-cell Wyner-type model as shown in figure 5.1. Cell $C_0$ is at the center. Adjacent cells are designated Cell $C_1$ and Cell $C_{-1}$.

Following our analysis of the single cell case, the received vector for users of cell $C_0$, is

$$\mathbf{y} = \mathbf{H}_0\mathbf{G}_0\mathbf{s}_0 + \sqrt{\gamma}\mathbf{H}_{01}\mathbf{G}_1\mathbf{s}_1 + \sqrt{\gamma}\mathbf{H}_{0-1}\mathbf{G}_{-1}\mathbf{s}_{-1} + \mathbf{n}. \tag{5.21}$$

As before, $\mathbf{H}_0$ is the channel matrix from base station in cell $\mathbb{C}_0$ to its users. $\mathbf{H}_{01}$ and $\mathbf{H}_{0-1}$ are interfering channels from cell $\mathbb{C}_1$ and $\mathbb{C}_{-1}$, respectively. $\mathbf{G}_1$ and $\mathbf{G}_{-1}$ are precoding matrices for users in cell $\mathbb{C}_1$ and $\mathbb{C}_{-1}$, respectively. $\gamma$ is the signal (interference) attenuation.

As stated earlier, all users in cell $\mathbb{C}_0$ are assumed to have the same average received signal power, so our model assumes that the users are similar distances from the base station and are not in deep shadow fades.

The precoding matrices in cell $i$ can be written as

$$\mathbf{G}_i = \alpha_i\mathbf{H}_i^{\mathsf{H}}(\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}})^{-1} \tag{5.22}$$

The ergodic capacity for user $k$ in cell $\mathbb{C}_0$ is expressed as

$$C_k = \log\left(1 + \frac{\alpha_0^2}{\mathrm{E}[|n_k|^2]}\right) \tag{5.23}$$

Where, $n_k$ is the $k^{th}$ element of the covariance matrix $\mathbf{n}$. The expectation of this matrix can be written as

$$\begin{aligned}\mathrm{E}[\mathbf{n}\mathbf{n}^{\mathsf{H}}] &= \gamma\mathbf{H}_{01}\mathbf{G}_1\mathbf{G}_1^{\mathsf{H}}\mathbf{H}_{01}^{\mathsf{H}} \\ &+ \gamma\mathbf{H}_{0-1}\mathbf{G}_{-1}\mathbf{G}_{-1}^{\mathsf{H}}\mathbf{H}_{0-1}^{\mathsf{H}} + \sigma^2\mathbf{I}\end{aligned} \tag{5.24}$$

Expanding and simplifying,

$$\begin{aligned}\mathrm{E}[\mathbf{n}\mathbf{n}^{\mathsf{H}}] &= \gamma\alpha_1^2\mathbf{H}_{01}\mathbf{H}_1^{\mathsf{H}}(\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}})^{-2}\mathbf{H}_1\mathbf{H}_{01}^{\mathsf{H}} \\ &+ \gamma\alpha_{-1}^2\mathbf{H}_{0-1}\mathbf{H}_{-1}^{\mathsf{H}}(\mathbf{H}_{-1}\mathbf{H}_{-1}^{\mathsf{H}})^{-2}\mathbf{H}_{-1}\mathbf{H}_{0-1}^{\mathsf{H}} \\ &+ \sigma^2\mathbf{I}\end{aligned} \tag{5.25}$$

Since,

$$\mathrm{E}[|n_1|^2] = \mathrm{E}[|n_2|^2] \ldots = \mathrm{E}[|n_k|^2] \tag{5.26}$$

We can write,

$$\begin{aligned}\mathrm{E}[|n_k|^2] &\rightarrow \frac{1}{K}\sum_{k=1}^{K}\mathrm{E}[|n_i|^2] \\ &= \frac{1}{K}\mathrm{tr}\left(E[\mathbf{n}\mathbf{n}^{\mathsf{H}}]\right)\end{aligned} \tag{5.27}$$

$$\mathrm{E}[|n_k|^2] = \frac{1}{K}\mathrm{tr}\left(\gamma\alpha_1^2\mathbf{H}_{01}\mathbf{H}_1^{\mathsf{H}}(\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}})^{-2}\mathbf{H}_1\mathbf{H}_{01}^{\mathsf{H}} + \gamma\alpha_{-1}^2\mathbf{H}_{0-1}\mathbf{H}_{-1}^{\mathsf{H}}(\mathbf{H}_{-1}\mathbf{H}_{-1}^{\mathsf{H}})^{-2}\mathbf{H}_{-1}\mathbf{H}_{0-1}^{\mathsf{H}}\right) + \sigma^2\mathbf{I} \tag{5.28}$$

### 5.4.5 Asymptotic analysis for the multi-cell

**Lemma 5.4.6.** *As* $K, M \to \infty$

$$\frac{1}{K}\mathrm{tr}\left(\mathbf{H}_{01}\mathbf{H}_1^{\mathsf{H}}(\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}})^{-2}\mathbf{H}_1\mathbf{H}_{01}^{\mathsf{H}}\right) \to \frac{1}{\beta-1}$$

**Proof:** Denote

$$\mathbf{A} = \mathbf{H}_1^{\mathsf{H}}(\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}})^{-2}\mathbf{H}_1$$

Now,

$$\begin{aligned}
\frac{1}{K}\mathrm{tr}\left(\mathbf{H}_{01}\mathbf{A}\mathbf{H}_{01}^{\mathsf{H}}\right) &= \frac{1}{K}\mathrm{E}\left[\mathrm{tr}\left(\mathbf{H}_{01}\mathbf{A}\mathbf{H}_{01}^{\mathsf{H}}\right)\right] \\
&= \frac{1}{K}\mathrm{tr}\left(\mathrm{E}\left[\mathbf{H}_{01}\mathbf{A}\mathbf{H}_{01}^{\mathsf{H}}\right]\right) \\
&= \frac{1}{K}\mathrm{tr}\left(\mathrm{E}\left[\mathrm{tr}(\mathbf{A})\right]\mathbf{I}_{K\times K}\right) \\
&= \mathrm{E}\left[\mathrm{tr}(\mathbf{A})\right] \\
&= \mathrm{E}\left[\mathrm{tr}\left(\mathbf{H}_1^{\mathsf{H}}(\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}})^{-2}\mathbf{H}_1\right)\right] \\
&= \mathrm{E}\left[\mathrm{tr}\left(\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}})^{-1}\right)\right]
\end{aligned}$$

If $K \times M$ matrix $\mathbf{H}_1$ is zero-mean, i.i.d. Gaussian, then $\mathbf{W} = \mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}$ is a *Wishart* matrix. For a Wishart matrix [2],

$$\mathrm{E}\left[\mathrm{tr}\left(\mathbf{W}\right)^{-1}\right] = \frac{K}{M-K}$$

$$\mathrm{E}\left[\mathrm{tr}\left(\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}})^{-1}\right)\right] = \frac{K}{M-K} = \frac{1}{\beta-1}$$

and hence,

$$\frac{1}{K}\mathrm{tr}\left(\mathbf{H}_{01}\mathbf{H}_1^{\mathsf{H}}(\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}})^{-2}\mathbf{H}_1\mathbf{H}_{01}^{\mathsf{H}}\right) \to \frac{1}{\beta-1} \qquad \square$$

[2] Refer section 2.1.6, equation (2.9) of [149] and the references there-in ([71, 97])

Thus the expectation in eq. (5.28) reduces to,

$$\mathrm{E}[|n_k|^2] \to \alpha_1^2\gamma\frac{1}{\beta-1} + \alpha_{-1}^2\gamma\frac{1}{\beta-1} + \sigma^2 \tag{5.29}$$

And hence, the sum-rate is

$$\mathcal{R}_{\mathrm{ci}} = K\log\left(1 + \frac{\alpha_0^2(\beta-1)}{\alpha_1^2\gamma + \alpha_{-1}^2\gamma + \sigma^2(\beta-1)}\right) \tag{5.30}$$

Following (5.14), for large $(K, M)$,

$$\frac{\alpha_0^2}{\sigma^2} = \frac{\alpha_1^2}{\sigma^2} = \frac{\alpha_{-1}^2}{\sigma^2} \to \frac{\rho\beta}{\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0)}, \text{ where } \rho = P/\sigma^2 \tag{5.31}$$

Thus the above sum-rate expression can be simplified as

$$\mathcal{R}_{\mathrm{ci}} = K\log\left(1 + \frac{\rho\beta(\beta-1)}{(\beta-1)\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0) + 2\gamma\rho\beta}\right) \tag{5.32}$$

Substituting for $\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0)$,

$$\mathcal{R}_{\mathrm{ci}} = K\log\left(1 + \frac{\rho(\beta-1)}{1+2\gamma\rho}\right) \tag{5.33}$$

Re-writing,

$$\frac{\mathcal{R}_{\mathrm{ci}}}{M} = \frac{1}{\beta}\log\left(1 + \frac{\rho(\beta-1)}{1+2\gamma\rho}\right) \tag{5.34}$$

We observe that when $\gamma = 0$, that is when there is no interference, the capacity formula is that of the single-cell case.

As $\beta \to 1$, $\mathcal{R}_{\mathrm{ci}}/M \to 0$, which implies that the sum rate of channel inversion does not increase linearly with $M$ (or $K$)

### 5.4.7 Optimizer $\beta^*$ for the multi-cell

Following on similar lines of the single-cell case, we now look for a value $\beta^\star$ of the ratio $M/K$ such that, for a fixed number of transmit antennas $M$, the sum-rate $\mathcal{R}_{\mathrm{ci}}(\beta)$ is maximized. By differentiating eqn (5.34) with respect to $\beta$ and setting the derivative to zero, $\beta^*$ is the solution of the implicit equation

$$\rho\beta^* = [\rho(\beta^*-1) + (1+2\gamma\rho)]\log\left[1 + \frac{\rho(\beta^*-1)}{1+2\gamma\rho}\right] \tag{5.35}$$

One can observe that by setting $\gamma = 0$, we fall back to the implicit equation (5.20) of the single cell case.

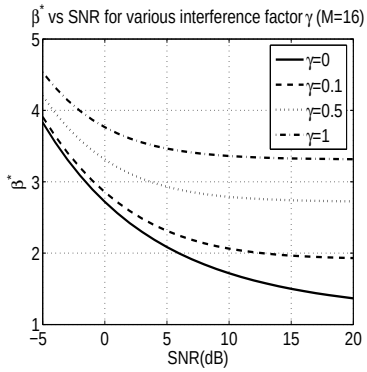


***Figure 5.2:*** *$\beta^*$ vs SNR for various interference factors ($M = 16$)*

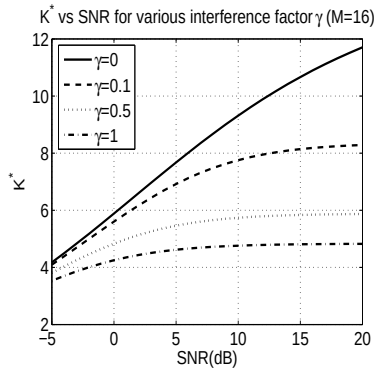


***Figure 5.3:*** *$K^*$ vs SNR for various interference factors ($M = 16$)*

### 5.4.8 Some observations:

Following our single cell and multi-cell analysis, we plot in figure 5.2, the optimal $\beta$, i.e, $\beta^*$ (refer equation 5.35), which maximizes the sum rate and in figure 5.3 the corresponding optimal number of users $K^* = M/\beta^*$ for M = 16 and different SNR. We observe that,

1) With increasing SNR more and more users should be served to maximize the sum rate.

2) Also, the number of users required to maximize the sum rate tends to saturate with an increase in the interference factor $\gamma$.

Next, we plot the optimal sum rate (refer equation 5.34), i.e, the sum rate achieved when $\beta = \beta^*$ in figure 5.4. We compare this for example with $\beta = 2$, shown in figure 5.5. We obtain the sum-rate by computing the rate per user in the asymptotic regime and then multiplying this with a finite number of antennas $M$ at the BS. For this example we have used $M = 16$.

There are some interesting observations here:

1) The sum-rate tends to increase at a constant rate when $\beta = \beta^*$, when there is no interference ($\gamma = 0$).

2) The sum-rate tends to saturate with interference and the saturation occurs sooner when the interference is higher.

3) The sum-rate with interference for any other $\beta$, for example $\beta = 2$ (fig 5.5), is not much different from $\beta = \beta^*$ (fig 5.4) in the presence of interference. The rate per transmit antenna tends to saturate with interference.

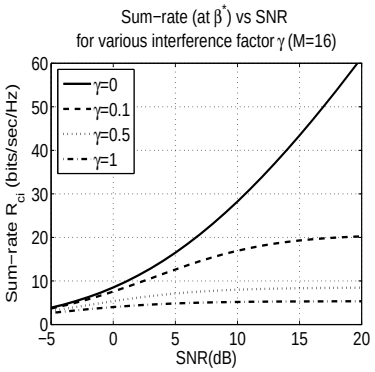


***Figure 5.4:*** *Sum rate at $\beta^*$ for various interference factors ($M = 16$)*

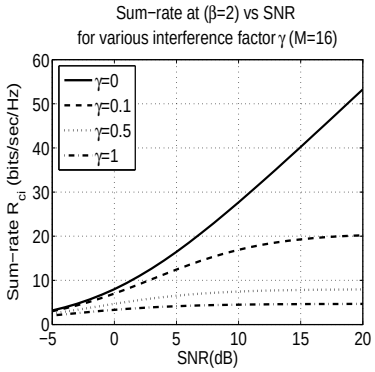


***Figure 5.5:*** *Sum rate at $\beta = 2$ for various interference factors ($M = 16$)*

### 5.4.9 Single cell and multi-cell with unequal power

We re-define the power-constraint as

$$\text{tr}[\mathbf{x}\mathbf{P}\mathbf{x}^{\mathsf{H}}] \leq \text{tr}[\mathbf{G}\mathbf{P}\mathbf{G}^{\mathsf{H}}] \leq P \tag{5.36}$$

such that the $k^{th}$ diagonal element of $\mathbf{P}$ represents power $p_k$ for user $k$ with $\sum_{k=1}^{K} p_k = \text{tr}(\mathbf{P})$.

Expanding $\text{tr}[\mathbf{G}\mathbf{P}\mathbf{G}^{\mathsf{H}}]$,

$$\begin{aligned}
\text{tr}\left(\mathbf{G}\mathbf{P}\mathbf{G}^{\mathsf{H}}\right) &= \text{tr}\left(\mathbf{H}^{\mathsf{H}}(\mathbf{H}\mathbf{H}^{\mathsf{H}})^{-1}\mathbf{P}(\mathbf{H}\mathbf{H}^{\mathsf{H}})^{-1}\mathbf{H}\right) \\
&= \text{tr}\left((\mathbf{H}\mathbf{H}^{\mathsf{H}})^{-1}\mathbf{P}\right) \\
&= \text{tr}(\mathbf{H}\mathbf{H}^{\mathsf{H}})^{-1}\ \frac{1}{K}\text{tr}(\mathbf{P}) \\
&= \frac{1}{M}\text{tr}(\mathbf{H}'\mathbf{H}'^{\mathsf{H}})^{-1}\ \frac{1}{K}\text{tr}(\mathbf{P}) \\
&= \frac{1}{\beta}\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0)\ \frac{1}{K}\text{tr}(\mathbf{P})
\end{aligned} \tag{5.37}$$

From eqn (5.36) and (5.37), we see that

$$\frac{1}{K}\text{tr}(\mathbf{P}) \leq \frac{P\beta}{\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0)}$$

With $\rho = P/\sigma^2$,

$$\frac{1}{\sigma^2} \leq \frac{\rho\beta}{\frac{1}{K}\text{tr}(\mathbf{P})\ \mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0)} \tag{5.38}$$

Substituting $\mathcal{S}_{\mathbf{H}'\mathbf{H}'^{\mathsf{H}}}(0) = \beta/(\beta-1)$, the ergodic capacity for user $k$ in the single-cell case is

$$C_k = \log\left(1 + \frac{p_k}{\sigma^2}\right) = \log\left(1 + \frac{p_k\rho(\beta-1)}{\frac{1}{K}\mathrm{tr}(\mathbf{P})}\right). \tag{5.39}$$

and the sum-rate is

$$R_{ci} = \sum_{k=1}^{K} \log\left(1 + \frac{p_k\rho(\beta-1)}{\frac{1}{K}\mathrm{tr}(\mathbf{P})}\right). \tag{5.40}$$

We can easily see that with equal power for all users, $\frac{1}{K}\mathrm{tr}(\mathbf{P}) = p = p_k$ and the above expression will reduce to the expressions derived for the single-cell case with equal power constraint (eqn 5.18).

The rate per antenna is

$$\frac{R_{ci}}{M} = \frac{1}{\beta}\frac{1}{K}\sum_{k=1}^{K} \log\left(1 + \frac{p_k\rho(\beta-1)}{\frac{1}{K}\mathrm{tr}(\mathbf{P})}\right) \tag{5.41}$$

For the multi-cell case, the ergodic capacity eqn (5.23) for user $k$ is

$$C_k = \log\left(1 + \frac{p_k}{\mathrm{E}[|n_k|^2]}\right) \tag{5.42}$$

Where,

$$\begin{aligned}\mathrm{E}[\mathbf{n}\mathbf{n}^{\mathsf{H}}] &= \gamma\mathbf{H}_{01}\mathbf{G}_1\mathbf{P}_1\mathbf{G}_1^{\mathsf{H}}\mathbf{H}_{01}^{\mathsf{H}} \\ &+ \gamma\mathbf{H}_{0-1}\mathbf{G}_{-1}\mathbf{P}_{-1}\mathbf{G}_{-1}^{\mathsf{H}}\mathbf{H}_{0-1}^{\mathsf{H}} \\ &+ \sigma^2\mathbf{I}\end{aligned} \tag{5.43}$$

After suitable simplification similar to the multi-cell analysis in the previous section, we can re-write the above expression as

$$\begin{aligned}\mathrm{E}[|n_k|^2] = \frac{1}{K}\mathrm{tr}\Big(&\mathbf{H}_{01}\mathbf{H}_1^{\mathsf{H}}(\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}})^{-1}\mathbf{P}_1(\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}})^{-1}\mathbf{H}_1\mathbf{H}_{01}^{\mathsf{H}} \\ &+\mathbf{H}_{0-1}\mathbf{H}_{-1}^{\mathsf{H}}(\mathbf{H}_{-1}\mathbf{H}_{-1}^{\mathsf{H}})^{-1}\mathbf{P}_{-1}(\mathbf{H}_{-1}\mathbf{H}_{-1}^{\mathsf{H}})^{-1}\mathbf{H}_{-1}\mathbf{H}_{0-1}^{\mathsf{H}}\Big) \\ &+\sigma^2\mathbf{I}\end{aligned} \tag{5.44}$$

As $M, K \to \infty$,

$$\begin{aligned}\mathrm{tr}\left(\mathbf{H}_{01}\mathbf{H}_1^{\mathsf{H}}(\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}})^{-1}\mathbf{P}_i(\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}})^{-1}\mathbf{H}_1\mathbf{H}_{01}^{\mathsf{H}}\right) \\ \to \frac{1}{\beta-1}\mathrm{tr}(\mathbf{P}_i)\end{aligned} \tag{5.45}$$

Therefore, the expectation can be written as,

$$\mathrm{E}[|n_k|^2] = \gamma \frac{1}{\beta - 1} \frac{\mathrm{tr}(\mathbf{P}_1)}{K} + \gamma \frac{1}{\beta - 1} \frac{\mathrm{tr}(\mathbf{P}_{-1})}{K} + \sigma^2 \mathbf{I}. \tag{5.46}$$

The capacity of user $k$ is

$$C_k = \log \left( 1 + \frac{(\beta - 1)\ p_k}{\gamma \frac{\mathrm{tr}(\mathbf{P}_1)}{K} + \gamma \frac{\mathrm{tr}(\mathbf{P}_{-1})}{K} + \sigma^2 (\beta - 1)} \right) \tag{5.47}$$

Substituting eqn (5.38) for $1/\sigma^2$,

$$C_k = \log \left( 1 + \frac{\rho(\beta - 1)\ p_k}{\gamma\rho\ \frac{\mathrm{tr}(\mathbf{P}_1)}{K} + \gamma\rho\ \frac{\mathrm{tr}(\mathbf{P}_{-1})}{K} + \frac{\mathrm{tr}(\mathbf{P})}{K}} \right) \tag{5.48}$$

and the sum-rate is expressed as

$$\mathcal{R}_{\mathrm{ci}} = \sum_{k=1}^{K} \log \left( 1 + \frac{\rho(\beta - 1)\ p_k}{\gamma\rho\ \frac{\mathrm{tr}(\mathbf{P}_1)}{K} + \gamma\rho\ \frac{\mathrm{tr}(\mathbf{P}_{-1})}{K} + \frac{\mathrm{tr}(\mathbf{P})}{K}} \right) \tag{5.49}$$

Notice that if $\frac{1}{K}\mathrm{tr}(\mathbf{P}_i) = p = p_k$, the above expression will reduce to the expressions derived for the multi-cell case with equal power constraint (eqn 5.33).

The sum-rate per antenna is

$$\frac{\mathcal{R}_{\mathrm{ci}}}{M} = \frac{1}{\beta} \frac{1}{K} \sum_{k=1}^{K} \log \left( 1 + \frac{\rho(\beta - 1)\ p_k}{\gamma\rho\ \frac{\mathrm{tr}(\mathbf{P}_1)}{K} + \gamma\rho\ \frac{\mathrm{tr}(\mathbf{P}_{-1})}{K} + \frac{\mathrm{tr}(\mathbf{P})}{K}} \right) \tag{5.50}$$

We observe two things here. 1) One can come up with an optimal power allocation policy (for ex. based on the channel characteristics) which maximizes the sum-capacity in the unequal power allocation scheme. 2) If some of the users in the adjacent base stations are not being serviced, i.e, their respective antenna at the transmitter is switched off, the interference comes down (for ex. if one or more user links are inactive in cell 1, then $\mathrm{tr}(\mathbf{P}_1) < P_1$) and hence the sum-capacity scales up.

## 5.5 Simulation results

In this section we evaluate by simulation how interference from neighboring base stations impacts the behavior of the sum-rate of linearly precoded MIMO small cell networks when the antenna array at the transmitter are large. We compare numerical

results obtained by Monte-Carlo simulations with our previously derived asymptotic expressions for finite $(K, M)$. In particular, we have the following cases.

1) We fix the SNR ($\rho = 20$ dB) and calculate rate achieved per antenna as we vary $\beta = M/K$ (refer equation (5.34)). We plot this in figure (5.6) for various interference factors $\gamma$. We observe that the rate per antenna is maximized for a certain $\beta = \beta^*$. This matches with the $\beta^*$ computed by solving the implicit eqn (5.35). It is also interesting to observe that $\beta^*$ increases with increasing interference. Also, beyond $\beta^*$, the capacity growth is not in proportion to the growth in number of antennas $M$ at the base station.

2) We fix the SNR ($\rho = 20$ dB) and the ratio $M/K = \beta = 2$. We compute the rate achieved per antenna as we vary the interference factor $\gamma$. We compare asymptotic results via monte-carlo simulations. We plot this in figure 5.7. We observe that the achievable rate is very sensitive to interference. The drop in rate is very steep in the beginning and tends to saturate for higher interference. Thus, the rate per antenna saturates with $\gamma$. This seems to indicate that the high amount of interference envisaged in small cells might not be as harmful. Many of the proposed interference management and co-ordination schemes might work well even in the case of small cells.

3) Next we show how the sum-rate increases with increasing number of base-station antennas $M$ at SNR ($\rho = 0, 20dB$) for various interference factors $\gamma$, when $\beta = 2$. We compute the rate per antenna from equation (5.34) for the asymptotic part to compare it with monte-carlo simulations. The observations are plotted in figures (5.8), (5.9). We observe that the increase in sum-rate is linear when interference is nil. The increase is sub-linear for other interference factors. Since the number of antennas at the base station and number of users are increasing simultaneously, the capacity is expected to grow in proportion to $\min(M, K)$, scaled by a factor, that depends on the interference factor $\gamma$ and the SNR $\rho$.

In all the cases, we observe that in all simulations the asymptotic results closely match the numerical results even for small values of $(K, M)$.

## 5.6 Conclusions

We looked at the problem of inter-cell interference in MIMO based small cell networks. We started our analysis with a single cell, where multi-antenna base station employ channel inversion precoding to communicate with multiple single-antenna users. We extended the case to multi-cell scenario, using a simple wyner-type model. We derived the sum-rate capacity in the asymptotic regime, i.e, when the number of antennas at the base station and number of user grow large, but, with a fixed ratio. We used recent tools from random matrix theory, which have proven to give reliable results even when the quantities involved are practical and finite. We further derived $\beta^*$, the ratio of number of transmit antennas to users, which maximizes the achievable sum-rate. This ratio provides the user density per antenna at the BS. The asymptotic analysis was validated with monte-carlo simulations in the finite regime.

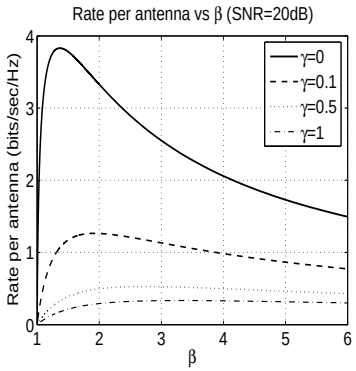


***Figure 5.6:*** *Rate per antenna vs $\beta$ at SNR of 20 dB*

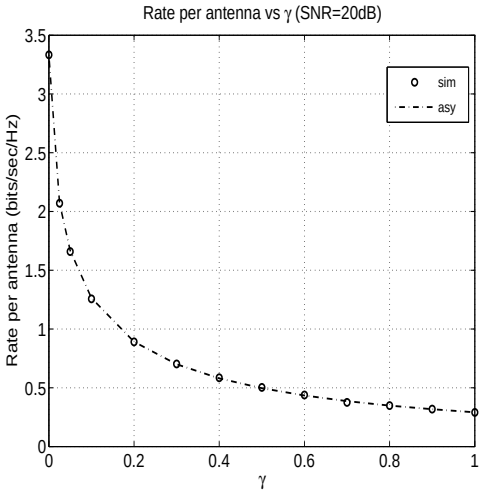


***Figure 5.7:*** *Rate per antenna vs $\gamma$, when, $\beta = 2$, Sl*

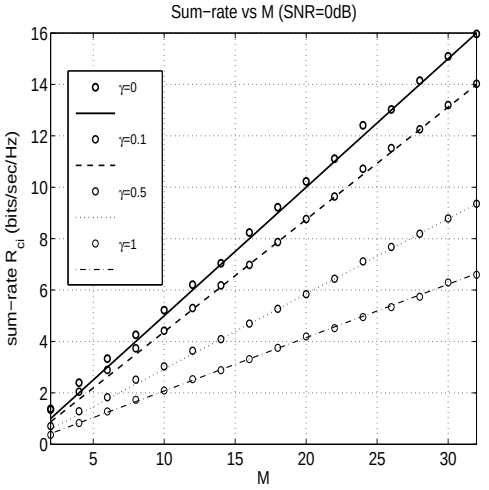


***Figure 5.8:*** *Sum rate per antenna as a function of M for $\beta = 2$ at SNR of 0 dB for various interference factors*

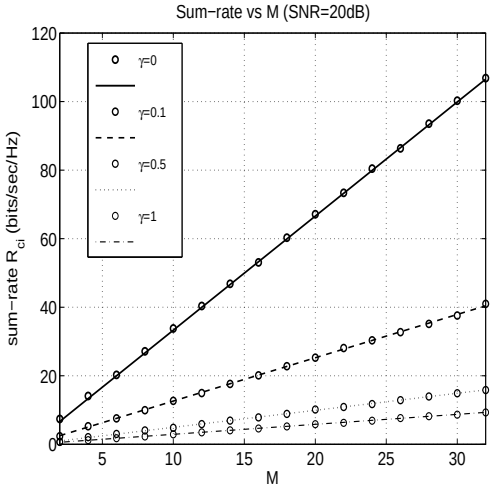


***Figure 5.9:*** *Sum rate per antenna as a function of M for $\beta = 2$ at SNR of 20 dB for various interference factors*

We conclude that the achievable sum-rate is significantly diminished by the effect of multi-cell interference in MIMO based small cell networks. The sum-rate capacity tends to grow sub-linearly with increasing interference. Also, there is an optimal number of users for a given number of antennas at the transmitter, which maximizes the sum-capacity. This depends on the interference level and the transmit power at the base-station. For a given number of transmit antenna, moving away from the optimal, $\beta^*$, tends to saturate the capacity growth at high SNR. The saturation occurs sooner with higher interference.

## 5.7 Publications

1. Sreenath Ramanath, Merouane Debbah, Eitan Altman, Vinod Kumar, "Asymptotic analysis of precoded small cell networks", proceedings of InfoCom 2010, Mar 15-19, San Diego, USA.

# Chapter 6

# Open Loop Control of BS Deactivation

**Contents**



## 6.1 Introduction

In recent years, there is a paradigm shift towards greener and denser networks [39, 81]. While green networks aim at reducing overall power and energy consumption in networks, denser networks increase the capacity and coverage of networks. Typically, dense networks, popularly referred to as small cell networks comprise Femto and pico cells serving indoor, hot-spot and urban mobility regions. The base stations used in such cells are compact and small portable devices, which can be easily installed on existing infrastructure and are often battery operated. They recharge periodically and it would be beneficial from a greener perspective if this can be as less frequent as possible. Thus these base stations need to judiciously use the available battery power. Further, if one assumes that all these base stations are accessible to a central control unit via a back haul link, depending on the load, some of these devices can be switched OFF to conserve their battery life.

In our current work, we derive optimal policies to conserve energy in two scenarios. 1) A central control unit, which, depending on the load in the system can switch OFF

certain base stations to reduce overall energy conservation in the system, but, with an objective to maximize the expected throughput. 2) Each base station can derive its policy independent of its neighbors to minimize an average cost metric (for example, its buffer occupancy), while keeping a lower bound on the fraction of time the it is switched OFF. To address these problems, we use tools from Multimodularity.

The natural counterpart of convex functions over integer sets turn out to be the so called *Multimodular functions* ([75]) and for such functions we have indeed the property that local minima are global minima. The property of Multimodularity can handle the control of discrete events and we utilize the same to obtain optimal activation policies for base stations. These tools were earlier used in the context of stochastic control of Queuing systems ([12] and reference therein). In a recent work [34], while addressing the problem of energy limited wireless handsets, the authors use tools from Multimodularity to address open loop control and establish optimality of bracket sequences based control.

We consider a regular network, wherein the base stations form a linear network and here the tools of Multimodularity fit in. Tools like stochastic geometry can be used to address networks which have base stations and mobiles distributed stochastically according to some given process (for eg. Poisson) and this study is not addressed here. Our objective is to obtain open loop policies. Later, we plan to derive closed loop policies for the same. The main results from this chapter are:

1. We show that the cost structure for the both centralized and decentralized scenarios is Multimodular.
2. For the central control, among all feasible policies with at least an asymptotic fraction $\rho$ of the base stations being switched OFF, the bracket policy with rate $\rho$ is optimal.
3. For the decentralized control, among all feasible policies with the BS switched OFF for at least an asymptotic fraction $\rho$ of the time, the bracket policy with rate $\rho$ is optimal.
4. In both the cases, the optimal policy is given by ($\lfloor x \rfloor$ denotes the largest integer smaller or equal to $x$)
$$a_n = \lfloor n\rho \rfloor - \lfloor (n-1)\rho \rfloor.$$
Here, $a_n = 0(1)$, if the $n$-th BS (centralized) or BS in the $n$-th time slot (decentralized) is switched OFF (ON).
5. The optimal policies depend only upon the conservation factor $\rho$ and are independent of all other system parameters, for example path loss coefficient, power per transmission, etc.

In both the problems, we call $\rho$, the switch OFF fraction (the total fraction of base stations to be switched OFF in central control or the fraction of time a base station is switched OFF in decentralized control), as the *conservation factor*. In the first case, this is decided by the central control based on system statistics (like traffic type, load distribution, QoS, time of hour, day of week, power saving, etc.), while in the decentralized

case, it is derived based on resources available and the QoS setting (memory, power saving, waiting time, sojourn time, etc.). Further, in our analysis, we assume that this factor is known a priori.

**Organization of the chapter:** In Section 6.2, we introduce Multimodularity concepts and the tools relevant in this work. In Section 6.3, we address the central control problem, while Section 6.4 studies the decentralized control. In Section 6.5, we discuss some possible future directions and we conclude our work in Section 6.6.

**Notations:** Bold letters, for example $\mathbf{a}$, represent an infinite sequence while $\mathbf{a}_j^k$ represents a part of this sequence defined by $(a_j, \cdots, a_k)$. Let $\mathbf{1}_1^n$ represent a $n$ vector of all ones while $\mathbf{1}$ represents the infinite sequence of all ones. In the sections, where the infinite sequences are not used and where all the vectors used are of the same length, then bold letters (for example $\mathbf{a}$) themselves are used to represent the finite length vectors. $\underline{\lim}$, $\overline{\lim}$ is limit infimum and supremum, respectively.

## 6.2 Multimodularity

Multimodularity can be used to address a wide class of control problems over sequences of integer numbers. *Multimodular* functions turn out to be the natural counter part of Convex functions, in the case of integer valued functions, for which the existence of a local minima guarantees a global minima. Also, they induce a particular form of optimal policies, which turn out to be very *regular*, and are described by the well known *bracket* sequences. We reproduce the related definitions (see [12] for general definitions) specific to the spaces considered in this work.

In this section, we use notation $\mathbf{a}$ to represent the $N-$length vector $\mathbf{a}_1^N$ as here we do not need the infinite length sequences.

**Definition 6.2.0.1.** *A function $f : \{0,1\}^N \to R$ is Multimodular if*

$$f(\mathbf{a}+\mathbf{v}) + f(\mathbf{a}+\mathbf{u}) \geq f(\mathbf{a}) + f(\mathbf{a}+\mathbf{u}+\mathbf{v}) \tag{6.1}$$

*for all $\mathbf{a} \in \{0,1\}^N$ and for all $\mathbf{u}, \mathbf{v} \in \mathcal{F}$ with $\mathbf{u} \neq \mathbf{v}$ and such that $\mathbf{a}+\mathbf{u}, \mathbf{a}+\mathbf{v}, \mathbf{a}+\mathbf{u}+\mathbf{v} \in \{0,1\}^N$.*

*The Multimodular base $\mathcal{F}$ contains the vectors $\{-\mathbf{e}_1, \mathbf{s}_2, \mathbf{s}_3, \cdots, \mathbf{s}_N, \mathbf{e}_N\}$, where,*

$$\begin{array}{lcl}
-\mathbf{e}_1 & = & (-1\ 0\ 0\ 0\ 0\ \ldots\ 0\ 0), \\
\mathbf{s}_2 & = & (1\ -1\ 0\ 0\ 0\ \ldots\ 0\ 0), \\
\mathbf{s}_3 & = & (0\ 1\ -1\ 0\ 0\ \ldots\ 0\ 0), \\
\cdots, & & \\
\mathbf{s}_N & = & (0\ 0\ 0\ 0\ 0\ \ldots\ 1\ -1) \text{ and} \\
\mathbf{e}_N & = & (0\ 0\ 0\ 0\ 0\ \ldots\ 0\ 1)
\end{array}$$

**Definition 6.2.0.2.** *The bracket sequence $\mathbf{a}(\rho,\theta) := \{a_n(\rho,\theta)\}$ with rate $\rho \in [0,1)$ and initial phase $\theta \in [0,1)$ is defined as*

$$a_n(\rho,\theta) = \lfloor n\rho + \theta \rfloor - \lfloor (n-1)\rho + \theta \rfloor \tag{6.2}$$

In this work, we mainly use Theorem 6, pp. 25, [12] which establishes the optimality of bracket sequences and the same is reproduced here.

**Theorem 6.2.0.3.** *A bracket sequence* $\mathbf{a}(\rho, \theta)$ *for any* $\theta \in [0, 1)$ *minimizes the cost*

$$\overline{\lim_{N\to\infty}} \frac{1}{N} \sum_{n=1}^{N} f_n(a_1, \cdots, a_n)$$

*over all the sequences that satisfy*

$$\underline{\lim_{N\to\infty}} \frac{1}{N} \sum_{n=1}^{N} a_n \geq \rho,$$

*when* $\rho \in [0, 1)$, *under the following assumptions:*

**A**.1 $f_n$ *is Multimodular* $\forall n$.

**A**.2 $f_n(a_1, \ldots, a_n) \geq f_{n-1}(a_2, \ldots, a_n)$, $\forall n > 1$ *and*

**A**.3 *for any sequence* $\{a_n\}$, $\exists$ *a sequence* $\{b_n\}$ *such that* $\forall n, m$ *with* $n > m$,

$$f_n(b_1, \ldots, b_{n-m}, a_1, \ldots, a_m) = f_m(a_1, \ldots, a_m)$$

**A**.4 *all functions* $f_n(a_1, \cdots, a_n)$ *are increasing in all* $a_i$.

## 6.3 Centralized optimal control

We have uniformly placed points on a line, which are separated by a distance $d$. Each point can potentially contain a BS and / or a mobile. The mobile at any point is active (i.e., has a communication request) with probability $q$ and this process is identical and independent across the space. Some of the base stations are switched OFF to optimize the battery performance. Every BS that is ON is associated to some of the mobiles based on the nearest distance criterion. Further the BS uses directional antennae and can only communicate with the users to its right. The throughput at the base station from a mobile located at distance $r$ is given by

$$\theta = q \log\left(1 + \frac{pr^{-\beta}}{\sigma^2}\right),$$

where $p, r, \beta, \sigma^2$ respectively represent the transmit power from mobile, distance between BS and mobile, path loss factor and the noise variance. Further we assume that there is no intra or inter-cell interference. Note that the throughput is achieved only when the mobile has a request, which occurs with probability $q$.

Our goal is to find an *optimal switch OFF pattern of the base stations so as to maximize the sum of the expected throughputs of all the mobiles when one has to at least switch OFF a fraction* $\rho$ *of the base stations.* The conservation factor $\rho$ at any time period will be decided by the

network based on the load and for a given $\rho$ the network prefers to adopt an optimal ON-OFF pattern.

We obtain this answer in the asymptotic as $N$, the number of points in the space tend to $\infty$. Consider a sequence $\mathbf{a} \in \{0,1\}^{\infty}$ to represent the control sequence in the following sense: $a_i$, the action at $i^{th}$ point is 1 if BS is switched OFF and 0 if BS is ON. The goal is to find an optimal sequence $\mathbf{a}^*$ which maximizes the expected throughput (defined via Cesaro limit) such that the total fraction of the base stations, that are switched OFF, is lower bounded by $\rho$. We assume that the system starts at point 0 where a BS is always switched ON and we control the ON-OFF status of the remaining base stations, i.e., the ones starting from point 1 onwards. We neglect the throughput due to the mobile at 0 as it does not contribute to optimization. Thus, we maximize

$$\max_{\mathbf{a}} \overline{\lim_{N\to\infty}} \frac{1}{N} \sum_{n=1}^{N} \theta_n(\mathbf{a}) \tag{6.3}$$

$$\text{subject to } \underline{\lim_{N\to\infty}} \frac{1}{N} \sum_{n=1}^{N} a_n \geq \rho \tag{6.4}$$

where $\theta_n(\mathbf{a})$ is the throughput due to the mobile at position $n$ (which depends upon the position of the nearest base station). This depends upon the sequence $\mathbf{a}$ and equals (note the base station at point 0 is always ON):

$$\theta_n(\mathbf{a}) = q \log\left(1 + \frac{p\ \left(1+r_n^2\right)^{-\beta/2}}{\sigma^2}\right) \text{ where} \tag{6.5}$$

$$r_n := \begin{cases} nd & \text{if } \mathbf{a}_1^n = \mathbf{1}_1^n \\ \inf\limits_{1\leq j\leq n} \{|n-j|d : a_j = 0\} & \text{else.} \end{cases} \tag{6.6}$$

We use tools from Multimodularity [12] to address this problem. The related definitions are summarized in section 6.2. We use Theorem 6.2.0.3 and obtain (proof is in Appendix A)

**Theorem 6.3.0.4.** *The function $f_n(\mathbf{a}_1^n) := -\theta_n(\mathbf{a})$ is Multimodular for every $n$. Further, the centralized problem (6.3) is optimized by a bracket sequence (6.2),*

$$\mathbf{a}^* = \mathbf{a}(\rho,\theta) \text{ for some } \theta \in [0,1). \diamond$$

From the above theorem it is clear that the optimal sequence depends only upon the conservation factor $\rho$ and nothing else. We now give some examples of bracket sequences. **Example 1:** The bracket sequence 100100100 maximizes the expected throughput for $(\rho,\theta) = (0.33, 0.9)$ . **Example 2:** The bracket sequence 1001001000 maximizes the expected throughput for $(\rho,\theta) = (0.3, 0.9)$ . If the factor $\rho$ is rational, then the sequence is periodic ([75]). In this case the optimal policy is to switch OFF the base stations in a periodic fashion, for example with $\rho = 0.33$ one needs to switch OFF every third BS. The optimality of bracket sequence is established in the limit $N$, the number of points, tending to infinity. This would imply the bracket sequence would be nearly optimal for

| Sequence **a** | **100100100** | 101000100 | 010010010 | 101000010 | 100010001 | 001001001 |
|---|---|---|---|---|---|---|
| Throughput $\Theta$ | 5.1 | 5.0 | 4.9 | 4.86 | 4.6 | 4.4 |

**Table 1:** Expected system throughput $\Theta = \frac{1}{N}\sum_{1}^{N}\theta_N(\mathbf{a})$ for different sequences.
The underlined sequence is the bracket sequence which optimizes the throughput.

systems with large $N$. It would also be optimal for not so large values of $N$ and this is established using a numerical example. We consider a system with $N = 9$ points in the space, $\rho = 1/3$ and obtain the optimal control sequence by exhaustive search. The results are tabulated in Table **1**. We observe that the system throughput is maximized again for the bracket sequence 100100100.

## 6.4 Decentralized optimal control

We have base stations deployed in a network. Each BS is powered by a battery and can transmit up to a maximum of $B$ bytes during a transmission opportunity, which are slotted over time. The BS can either be in an active (ON) state where it transmits packets or in an idle state where it shuts OFF its activity to conserve the battery. Note that whenever a BS enters idle (OFF) mode, more number of packets get stored in the buffer and the buffer occupancy cost increases. *What is an optimal policy to minimize the average buffer occupancy such that the BS is switched OFF at least for a fraction $\rho$ of the time ?*

Let $\mathbf{a} = \{a_t\}_{t\geq 1}, a_t \in \{0,1\}$, be a sequence of controls such that $a_t = 1$ indicates BS is OFF at the $t^{\overline{th}}$ time slot and $a_t = 0$ indicates that the BS can serve at maximum $B$ packets. With any general ON-OFF policy, **a**, the buffer occupancy, $x_t$ evolves as

$$\begin{aligned} x_t(\mathbf{a}) &= (x_{t-1}(\mathbf{a}) - (1-a_t)B)^+ + w_t \\ \text{and it begins with } x_0 &= w_0. \end{aligned} \tag{6.7}$$

In the above, $w_t$ represents the new arrivals in $t$-th time slot. We assume $\{w_t\}_{n\geq 0}$ is an Identically and Independently distributed (IID) sequence and that it is bounded by $B$, i.e., $w_t \leq B$ with probability one. We now have the following problem of minimizing

$$\begin{aligned} &\min_{\mathbf{a}} \overline{\lim_{T\to\infty}} \frac{1}{T}\sum_{t=1}^{T} E[x_t(\mathbf{a})] \\ &\text{subject to} \quad \underline{\lim_{T\to\infty}} \frac{1}{T}\sum_{t=1}^{T} a_t \geq \rho. \end{aligned} \tag{6.8}$$

One needs to choose the conservation factor, $\rho$ such that $B(1-\rho) > E[W]$. This has to be done to ensure that the system can be stable at least for some of the control sequences **a**. With the above condition, the system for example is stable for all those sequences whose switch OFF fraction exactly equals $\rho$. We again use the Multimodularity Theorem 6.2.0.3 and obtain (proof is in Appendix B)

**Theorem 6.4.0.5.** *The function* $f_t(\mathbf{a}_1^t) := E[x_t(\mathbf{a})] = E[x_t(\mathbf{a}_1^t)]$ *is Multimodular for every* $t$. *And, the decentralized problem (6.8) is optimized by a bracket sequence.*

$$\mathbf{a}^* = \mathbf{a}(\rho, \theta) \text{ for some } \theta \in [0, 1). \diamond$$

Thus the optimal sequence is again a bracket sequence which depends only on $\rho$, the conversation factor. Hence, in both centralized and decentralized problems, we have similar optimal control pattern which in both the cases depends only upon the conservation factor $\rho$ and none of the other design parameters.

## 6.5 Future Directions

We saw in previous examples how individual BS can decide switch OFF (in time) or a central control can switch OFF base stations (in space) to conserve the energy. In future, one can consider some more interesting problems which can be solved using Multimodularity tools.

*Optimization in space and time:* We have uniformly placed $M$ points on a line, which are separated by a distance $d$. Each point can potentially contain a BS and a queue of waiting mobiles. We assume that the BS service is directional and they can potentially serve up to a maximum of $B$ mobiles to the right in every time slot. The service order of the mobiles depend on their proximity to the BS (to its left). Further, to conserve power, a central control can decide to switch OFF certain number of BS every time slot. This scenario results from a network equipped with cooperative base stations and in this case one needs to design an ON-OFF control that not only spans across the space but also across time. We expect to obtain again a bracket sequence as the optimal sequence and now we further expect this sequence to have 'periodicity' along space (when confined to the same time slot), have a small phase change in the next time slot and the periodicity in space continues.

*Coverage problems:* A mobile is covered, i.e., serviced when it is within a distance $nd$ from a BS that is ON. One can study the optimal ON-OFF pattern which optimizes the fraction of the mobiles covered.

*Closed Loop Policies:* For the decentralized problem we would like to obtain the closed loop policies. That is, we obtain the optimal policies, which at any time $t$ depend upon the history up to time $t-1$, $(\mathbf{a}_1^{*t-1}, \mathbf{x}_1^{t-1})$. We here consider minimizing the finite time horizon problem

$$\min_{\mathbf{a}_1^T \in \mathcal{A}^T} \sum_{t=1}^{T} E\left[x_t(\mathbf{a}_1^T) - \lambda a_t\right]$$

where,

$$\mathcal{A}^T := \{\mathbf{a}_1^T; \text{ for every } t \leq T, \;\; a_t : (\mathbf{a}_1^{t-1}, \mathbf{w}_1^{t-1}) \mapsto \{0, 1\}\},$$

that is for every $t$ the action at time $t$, $a_t$ is a function of the history $(\mathbf{a}_1^{t-1}, \mathbf{w}_1^{t-1})$.

One can again use the tools of Multimodularity to obtain the structure of close loop optimal policy. We propose to use Theorem 64, page 214, [12] to show that the optimal policy is a threshold policy in the following sense: for any given $t$ and history $\mathbf{a}_1^{*t-1}, \mathbf{x}_1^{t-1}$ (optimal actions and optimal state trajectory till time $t-1$), the optimal action at time $t$, $a_t^*$ is 0 until the state $x_t$ is above a threshold and is 1 otherwise.

## 6.6 Conclusions

In this chapter, we derived energy conserving policies for Base stations in regular green networks using tools from Multimodularity. We considered two example scenarios and show how the cost functions are Multimodular. In the first case, for the case of central control, we derived the optimal open loop policies so as to maximize the expected throughput of the system given that at least a certain percentage of Base stations are switched OFF. In the second case, we derived optimal open loop polices, which each base station can employ in a decentralized manner to minimize buffer occupancy costs, while keeping the long term average fraction of the BS switch OFF time at least above a given threshold. We established the optimality of bracket policies for both the cases. We established that these regular sequences optimize the (respective) performance(s) and these sequences depend only upon the conservation factor and nothing else. We conclude that Multimodularity can be applied in several interesting example scenarios to derive optimal control in Green Networks and provide some examples for future direction of research in this area.

# Appendix A: Proofs related to centralized control

**Proof of Theorem 6.3.0.4:** The proof is obtained using Theorem 6.2.0.3 of section 6.2. By Theorem 6.6.0.6, the function $f_n(a_1, \cdots, a_n) = -\theta_n(\mathbf{a})$ is Multimodular and hence assumption **A**.1 of Theorem 6.2.0.3 is satisfied.

Given $\mathbf{a}_1^n$, define $a_0 = 0$ and then define $n_b := \arg \inf_{0 \le j \le n} \{|n-j| : a_j = 0\}$, to denote the index of the nearest base station in the left for the mobile at point $n$. When $n_b > 1$, assumption **A**.2 is satisfied as then $f_n(\mathbf{a}_1^n) = f_{n-1}(\mathbf{a}_2^n) = f(\mathbf{a}_{n_b}^n)$. When $n_b = 1$ (i.e., when $\mathbf{a}_1^n = (0, 1, \cdots, 1)$) the equality still holds in **A**.2 by the definition (6.5) of $\theta_n(\mathbf{a})$. When $\mathbf{a}_1^n = \mathbf{1}_1^n$, then the assumption **A**.2 is satisfied with inequality as $\theta_n(\mathbf{1}_1^n) < \theta_{n-1}(\mathbf{1}_1^{n-1})$.

Assumption **A**.3 is satisfied by taking $\{b_n\}$ to be all zeros and clearly $f_n$ satisfies **A**.4. Thus all the hypothesis of Theorem 6.2.0.3 are satisfied and hence the theorem follows. ◇

**Theorem 6.6.0.6.** *For every $n$, $f_n$ is Multimodular.*

**Proof:** All the sequences in this proof are $n$ length vectors and hence we use the short notation $\mathbf{a}$ in place of $\mathbf{a}_1^n$. Consider any sequence $\mathbf{a}$. We need to show for $\mathbf{u} \neq \mathbf{v} \in \mathcal{F}$

(see section 6.2 for definitions) $f_n$ satisfies

$$f_n(\mathbf{a}+\mathbf{u}) + f_n(\mathbf{a}+\mathbf{v}) \geq f_n(\mathbf{a}) + f_n(\mathbf{a}+\mathbf{u}+\mathbf{v})$$

$$\text{whenever, } \mathbf{a}+\mathbf{u},\ \mathbf{a}+\mathbf{v},\ \mathbf{a}+\mathbf{u}+\mathbf{v} \text{ are all in } \{0,1\}^n.$$

Without loss of generality let $\mathbf{v} = \mathbf{s}_j$ (see section 6.2) which result in shifting an ON position of the base station from $j-1$ location to location $j$. Since we can only consider such $\mathbf{s}_j$ for which $\mathbf{a}+\mathbf{s}_j$ is in $\{0,1\}^n$ the sequence $\mathbf{a}$ should have $a_{j-1} = 0$ and $a_j = 1$. Further, $(a+v)_{j-1} = 1$ and $(a+v)_j = 0$ while for every $i \neq j$ or $i \neq j-1$, $a_i = (a+v)_i$. Thus, addition of vector $\mathbf{v}$ to $\mathbf{a}$ results in changes in the base station associations and hence the throughputs, only for the the mobiles located in $(j-1, \cdots, \eta_j^r - 1)$, where $\eta_j^r$ is defined as the nearest base station to the right of the location $j$ that is switched ON in $\mathbf{a}$:

$$\eta_j^r := \begin{cases} \arg\inf_{n>j}\{|n-k| : a_k = 0\} & \text{if set is non empty} \\ n+1 & \text{else.} \end{cases} \tag{6.9}$$

Hence,

$$\begin{aligned} f_n(\mathbf{a}) - f_n(\mathbf{a}+\mathbf{v}) &= \sum_{k=1}^{n} \theta_k(\mathbf{a}) - \sum_{k=1}^{n} \theta_k(\mathbf{a}+\mathbf{v}) \\ &= \sum_{k=j-1}^{\eta_j^r - 1} \theta_k(\mathbf{a}) - \theta_k(\mathbf{a}+\mathbf{v}). \end{aligned}$$

Let $\mathbf{u} = \mathbf{s}_l$ with $l > j+1$ ($l \neq j+1$ as then it is not possible that both $\mathbf{a}+\mathbf{u}$ and $\mathbf{a}+\mathbf{v}$ are in $\{0,1\}^n$). The addition of $\mathbf{u}$ to $\mathbf{a}$ will introduce changes in mobile throughputs only at locations $(l-1, \cdots, \eta_l^r)$, which do not overlap with locations changed by $\mathbf{v}$, $(j-1, \cdots, \eta_j^r)$. Further, the addition of $\mathbf{v}$ to $\mathbf{a}+\mathbf{u}$ also changes the mobile throughputs only in locations $(j-1, \cdots, \eta_j^r)$ (w.r.t. the mobile throughputs under $\mathbf{a}+\mathbf{u}$). Because of the independence of the locations of the changes due to $\mathbf{v}$ and $\mathbf{u}$ and because the mobile throughputs only depend upon the distance w.r.t. the serving BS, the mobile throughput changes from $\mathbf{a}+\mathbf{u}$ to $\mathbf{a}+\mathbf{u}+\mathbf{v}$ will be same as that when $\mathbf{a}$ is changed to $\mathbf{a}+\mathbf{v}$. Thus,

$$\begin{aligned} f_n(\mathbf{a}+\mathbf{u}) \quad - \quad & f_n(\mathbf{a}+\mathbf{u}+v) = \\ & \sum_{k=j-1}^{\eta_j^r} \theta_k(\mathbf{a}+\mathbf{u}) - \theta_k(\mathbf{a}+\mathbf{u}+\mathbf{v}) = \\ & \sum_{k=j-1}^{\eta_j^r} \theta_k(\mathbf{a}) - \theta_k(\mathbf{a}+\mathbf{v}) = \\ & f_n(\mathbf{a}) - f_n(\mathbf{a}+\mathbf{v}). \end{aligned} \tag{6.10}$$

The second vector $\mathbf{u}$ can either be $\mathbf{s}_l$ with $l > j+1$ or $l < j-1$ (as both $\mathbf{a}+\mathbf{u}$ and $\mathbf{a}+\mathbf{v}$ have to be in $\{0,1\}^n$) or $\mathbf{u}$ can be $\mathbf{e}_n$ when $j < n$ or it can $\mathbf{e}_1$ when $j > 1$. In all the combinations, addition of vectors $\mathbf{u}$ and $\mathbf{v}$ results in changes to the base station association at independent locations as above. Thus for any $\mathbf{u} \neq \mathbf{v}$, as in (6.10), one can show that $f_n(\mathbf{a}+\mathbf{u}) + f_n(\mathbf{a}+\mathbf{v}) = f_n(\mathbf{a}) + f_n(\mathbf{a}+\mathbf{u}+\mathbf{v})$. $\diamond$

## Appendix B: Proofs related to decentralized control

**Proof of Theorem 6.4.0.5:** The proof is obtained using Theorem 6.2.0.3 of section 6.2. By Theorem 6.6.0.7, the function $g_t(a_1, \cdots, a_t) = x_t(\mathbf{a})$ is Multimodular for every sample path of $\mathbf{w}_0^t$. The sample path wise Multimodularity implies the Multimodularity of the average function $f_t(\mathbf{a}_1^t) = E[x_t(\mathbf{a}_1^t)]$ and thus the first part of the theorem is established as well as the Assumption **A.**1 is satisfied.

The initial buffer size is $w_0$, i.e., $x_0(\mathbf{a}) = w_0$ for all sequences $\mathbf{a}$ and all samples. The function value $f_t(\mathbf{a}_1^t)$ is the average buffer size at $t-1$ time slot, obtained by progressing (Lindley's recursion) $t-1$ time slots using the control sequence $\mathbf{a}_2^t$ and when initial buffer size is given by $x_1(a_1) \geq w_1$ while the function value $f_{t-1}(\mathbf{a}_2^t)$ is the average buffer size at the same time slot obtained again by progressing $t-1$ time slots using the same control sequence $\mathbf{a}_2^k$ but now with initial buffer size equal to $w_0$. Note that $w_1$ is distributed same as $w_0$ and hence $f_t(\mathbf{a}_1^t) \geq f_{t-1}(\mathbf{a}_2^t)$. Thus, the assumption **A**.2 is satisfied.

For assumption **A**.3 take $\{b_t\}$ to be an all zero sequence. When the control sequence is all zeros, i.e., the BS serves in all the time slots, since the maximum number of arrivals in a slot is $B$, the buffer size at the end of every slot $t$, will exactly be $w_t$, the new arrivals. Thus the function value $f_t(\mathbf{b}_1^{t-\tau}, \mathbf{a}_1^\tau)$ represents the average buffer size after $\tau$ time slots when the control sequence is $\mathbf{a}_1^\tau$ and when the initial buffer size is $w_{t-\tau}$ while $f_\tau(\mathbf{a}_1^\tau)$ represents the same after $\tau$ time slots and with the same control sequence $\mathbf{a}_1^\tau$ but with initial buffer size $w_0$ and hence the two average values are equal. Thus, assumption **A**.3 is satisfied.

Clearly, assumption **A**.4 is also satisfied and hence the theorem follows by Theorem 6.2.0.3. ◇

**Theorem 6.6.0.7.** *For every $t$ and for every sample path $\mathbf{w}_0^t$, the function $g_t$ is Multimodular.*

**Proof:** In [34], while addressing the problem of energy limited wireless handsets, the authors show that the function $x_t(\mathbf{a}_1^t)$ is Multimodular for every sample path of the arrival sequence $\{w_t\}$. The functions used in describing their Cesaro limit (see [34]) are exactly the same as the functions $\{x_t\}$ of the decentralized problem. The sample path wise Multimodularity is proved as Theorem 17, page 6 [34] (details of this proof are in their technical report, Theorem 20, [33]). The proof there is little difficult to read and hence we provide a brief overview of the same below:

If $x_{t-1} \leq B$ then $a_t = 0$ results in an empty queue and $x_t = w_t$. On the other hand, if $x_{t-1} > B$, some part $\Delta x_{t-1} = x_{t-1} - B$, remains in the queue and thus $x_t = \Delta x_{t-1} + w_t$. Using this one can show that

$$\begin{aligned} x_t(\mathbf{a}) &= x_t(\mathbf{a} + \mathbf{s}_j) \text{ if } x_{j-2} > B \\ \text{and } x_t(\mathbf{a}) &= x_t(\mathbf{a} + \mathbf{s}_j) \text{ if } x_{j-2} \leq B \end{aligned} \tag{6.11}$$

Without loss of generality, let $\mathbf{v} = \mathbf{s}_j$ with $j = 2, \cdots, t$. Then, there are three possible cases: i) $\mathbf{u} = -\mathbf{e}_1$, ii) $\mathbf{u} = \mathbf{s}_k$, $k > j$ and iii) $\mathbf{u} = \mathbf{e}_t$ and we need to show for every

combination that the equation (6.1) is satisfied, by the function $g_t = x_t$, to complete the Multimodularity proof. In the following we present the proof for the case (ii) and the remaining cases are much simpler and follow similar logic.

Note that $x_l(\mathbf{a}+\mathbf{u}) = x_l(\mathbf{a})$ for all $l < k-1$. Thus, $x_{j-2}(\mathbf{a}) \geq B$ if and only if $x_{j-2}(a+v) \geq B$. Note further that this case is possible only if $a_{j-1} = 0$, $a_j = 1$ and $a_{k-1} = 0$, $a_k = 1$.

a) If $x_{j-2}(\mathbf{a}) \geq B$ then $x_t(\mathbf{a}+\mathbf{v}) = x_t(\mathbf{a})$ and $x_t(\mathbf{a}+\mathbf{u}+\mathbf{v}) = x_t(\mathbf{a}+\mathbf{u})$. Thus (6.1) is satisfied with equality.

b) If $x_{j-2}(\mathbf{a}) < B$ and further if $x_{k-2}(\mathbf{a}) \geq B$ and $x_{k-2}(\mathbf{a}+\mathbf{v}) \geq B$ then again $x_t(\mathbf{a}+\mathbf{u}) = x_t(\mathbf{a})$ and $x_t(\mathbf{a}+\mathbf{v}+\mathbf{u}) = x_t(\mathbf{a}+\mathbf{v})$ and so again (6.1) is satisfied with equality. On the other hand, if $x_{k-2}(\mathbf{a}) \geq B$ and $x_{k-2}(\mathbf{a}+\mathbf{v}) < B$ we have $x_t(\mathbf{a}+\mathbf{u}) = x_t(\mathbf{a})$ and $x_t(\mathbf{a}+\mathbf{v}+\mathbf{u}) \leq x_t(\mathbf{a}+\mathbf{v})$ and then (6.1) is satisfied, but need not be with equality. Now if $x_{k-2}(\mathbf{a}) < B$ and $x_{k-2}(\mathbf{a}+\mathbf{v}) < B$ then because $a_{k-1} = 0$ so is $(a+v)_{k-1} = 0$ (the controls in $a$, $a+v$ are same after $j+1$), $x_{k-1}(\mathbf{a}) = x_{k-1}(\mathbf{a}+\mathbf{v})$ as they both result only because of new arrivals at $k-1$ and older ones (which might be different) were flushed out completely. Since there is no difference in both the controls $\mathbf{a}$, $\mathbf{a}+\mathbf{v}$ after the time point $j$, $x_t(\mathbf{a}) = x_t(\mathbf{a}+\mathbf{v})$. From (6.11) we do have $x_t(\mathbf{a}+\mathbf{u}) \geq x_t(\mathbf{a}+\mathbf{u}+\mathbf{v})$ and so again (6.1) is satisfied. ◇

## 6.7 Publications

1. Sreenath Ramanath, Veeraruna Kavitha, Eitan Altman, "Open Loop Optimal Control of Base Station Activation for Green Networks", proceedings of WiOpt 2011, May 9-13, Princeton, USA.

# Part III

# Resource Allocation

# Chapter 7

# Multiscale Fairness and its Application in Wireless Networks

**Contents**



## 7.1 Introduction

Fair resource allocation is usually studied in a static context, in which a fixed amount of resources is to be shared. In dynamic resource allocation one usually tries to assign resources instantaneously so that the average share of each user is split fairly. The exact definition of the average share may depend on the application, as different applications may require averaging over different time periods or time scales. We study dynamic resource allocation and examine how the constraints on the averaging durations impact the amount of resources that each user gets.

Let us consider some set $S$ of resource that we wish to distribute among $I$ users by assigning user $i$ a subset $S_i$ of it. We shall be interested in allocating subsets of the resource fairly among the users. The set $S$ may actually correspond to one or to several resources. We shall consider standard fairness criteria for sharing the resources among users. We shall see, however, that the definition of a resource will have a major impact on the fair assignment.

We associate with each user $i$ a measurable function $x_i$ that maps each point in $S$ to some real number. Then, we associate with each $i$ a utility $u_i$ which maps all measurable subsets $S_i$ to the set of real numbers. We shall say that $S$ is a resource if $u_i(S_i)$ can be written as

$$u_i(S_i) = f\left(\int_{S_i} x_i(s)ds\right)$$

for each $S_i \subset S$.

As an example, consider $I$ mobiles that wish to connect to a base station between 9h00 and 9h10 using a common channel. The time interval is divided into discrete time slots whose number is $N$. Assume that the utility for each mobile $s$ of receiving a subsets $\mathcal{N}_i$ of slots depend only on the number of slots $N_i$ it receives. Then the set of $N$ slots is considered to be a resource.

Next assume that if mobile $i$ receives the channel at time slot $t$ then it can transmit at a throughput of $X_t^i$. Assume that the utility of user $i$ is a function of the total throughput it has during this fraction of an hour. Then again the $N$ slots are considered as a resource.

We adopt the idea that fair allocation should not be defined in terms of the object that is split but in terms of the utility that corresponds to the assignments. This is in line with the axiomatic approach for defining the Nash bargaining solution for example. With this in mind, we may discover that the set of $N$ slots cannot always be considered as a resource to be assigned fairly. Indeed, a real time application may consider the $N$ slots as a set of $n$ resources, each containing $B = N/n$ consecutive slots. A resource may correspond to the number of time slots during a period of 100 msec. The utility of the application is defined as a function of the instantaneous rate, i.e. the number of slots it receives during each period of 100 msec. (With a playout buffer that can store 100 msec of voice packets, the utility of the mobile depends only on how many slots are assigned to it during 100 msec and not which slots are actually assigned to it.)

What is the impact on data transfer applications of splitting the resource of $N$ slots into $B$ smaller resources? We shall show that allocating fairly each of these $B$ resources results in performance degradation for the data transfer applications. This raises the question of how to define fair assignment when the very notion of a resource varies from one user to another.

Another example where this question arises is frequency allocation. Assume that frequency bandwidth needs to be split between users, who bid for $N$ carriers, each of bandwidth $b$. There may be users who need carriers of bandwidth $mb$. They can make use of a carrier only if they receive a set of $m$ consecutive carriers. For these users, a resource may correspond to the set of $N/m$ group of carriers, each of which containing $m$ consecutive carriers.

### Related work

Our work is based on the $\alpha$-fairness notion introduced by Mo and Walrand [105]. This notion provides a continuum of fairness definitions through the real parameter $\alpha$ and it includes various known fairness concepts that are obtained for some specific values of the real parameter $\alpha$ (the max-min fairness, the proportional fairness and the harmonic fairness). This, as well as other fairness notions can be defined through a set of axioms, see [94]. This work is inspired by several papers which already observed or derived fairness at different time-scales [7, 5, 6, 37, 113, 91]. However, we would like to mention that the $T$-scale fairness (a unifying generalization of long- and short- term fairness) and multiscale fairness are new concepts introduced in the present work.

### Structure of the chapter

The chapter is organized as follows: In the next Section 7.2 we introduce a resource sharing model which is particularly suitable for wireless network applications. In Section 7.2 we also define several fairness criteria, illustrate them by examples and prove theoretical properties of the introduced fairness criteria. In Section 7.3 we derive explicit formulae for instantaneous $\alpha$-fairness in the case of linear resources. The case of linear resources corresponds to the frequency as a resource in wireless networks. In Sections 7.4 and 7.5 we apply different fairness criteria to spectrum allocation in fading channels and to indoor-outdoor scenario, respectively. Section 7.6 concludes the chapter and provides avenues for future research.

## 7.2 Resource Sharing model and different fairness definitions

Consider $n$ mobiles located at points $x_1, x_2, ..., x_n$, respectively. We assume that the utility $U_i$ of mobile $i$ depends on its location $x_i$ and on the amount of resources $s_i$ it gets.

Let $\mathbf{S}$ be the set of assignments; an assignment $s \in \mathbf{S}$ is a function from the vector $x$ to a point in the $n$-dimensional simplex. Its $i$th component, $s_i(x)$ is the fraction of resource assigned to mobile $i$.

**Definition 7.2.0.8.** *($\alpha$-fair assignment) An assignment $s$ is $\alpha$-fair if it is a solution of:*

$$\begin{aligned} Z(x,\alpha) &:= \max_s \sum_i Z_i(s_i, x_i, \alpha) \\ s.t. \quad \sum_i s_i &= 1,\ s_i \geq 0 \forall i = 1, ..., n \end{aligned} \tag{7.1}$$

*where,*

$$Z_i(x_i, s_i, \alpha) := \frac{(U_i(x_i, s_i))^{1-\alpha}}{1-\alpha}$$

*for $\alpha \neq 1$. For $\alpha = 1$ we define*

$$Z_i(x_i, s_i, \alpha) := \log(U_i(x_i, s_i))$$

We shall assume throughout that $U_i$ is non-negative, strictly increasing and is concave in $s_i$. Then for any $\alpha > 0$, $Z_i(x_i, s_i, \alpha)$ is strictly concave in $s_i$. We conclude that $Z(x_i, s_i, \alpha)$ is strictly concave in $s$ for any $\alpha > 0$ and therefore there is a unique solution $s^*(\alpha)$ to (7.1).

**Definition 7.2.0.9.** *(Mo and Walrand [105]) We call $Z_i(s_i, \cdot, \alpha)$ the fairness utility of mobile $i$ under $s_i$, and we call $Z(s, \cdot, \alpha)$ the instantaneous degree of $\alpha$-fairness under $s$.*

In applications, the state $X$ will be random, so that the instantaneous amount of resource assigned by an $\alpha$-fair allocation will also be a random variable. Thus, in addition to instantaneous fairness we shall be interested in the expected amount assigned by being fair at each instant.

**Definition 7.2.0.10.** *We call $E[Z(s, X, \alpha)]$ the expected instantaneous degree of $\alpha$-fairness under $s$.*

In Section 7.2.1 we introduce the expected long-term fairness in which the expected amount of resource is assigned fairly.

**Definition 7.2.0.11.** *We say that a utility is linear in the resource if it has the form:*

$$U_i(x_i, s_i) := s_i q_i(x_i).$$

For example, consider transmission between a mobile source and a base station, and assume
(i) that the base station is in the origin ($x = 0$) but at a height of one unit, whereas all mobiles are on the ground and have height 0. Thus, the distance between the base station and a mobile located on the ground at point $x$ is $\sqrt{1 + ||x||^2}$.
(ii) that the Shannon capacity can be used to describe the utility. If the resource that is shared is the frequency then the utility has the linear form:

$$\begin{aligned} U(C, x) &:= Cq(x) \\ \text{where } q(x) &= \log\left(1 + \frac{P(x^2+1)^{-\beta/2}}{\sigma^2}\right) \end{aligned}$$

Note: if the power and not the frequency, were taken to be the resource then we would not obtain the linear form of the resource.

In the linear case, we write $Z_i$ as:

$$Z_i(s, x, \alpha) := s_i^{1-\alpha} v_i(x), \qquad v_i := \frac{(q_i(x_i))^{1-\alpha}}{1-\alpha}$$

### 7.2.1 Fairness over time: Instantaneous Versus Long term $\alpha$-fairness

Next we consider the case where $x_i(t)$, $i = 1, ..., n$, may change in time.

**Definition 7.2.1.1.** *We define an assignment to be instantaneous α-fair if at each time t each mobile is assigned a resource so as to be α-fair at that instant.*

Consider the instantaneous $\alpha$-fair allocation and assume that time is discrete. We thus compute the instantaneous $\alpha$-fair fair assignment over a period of $T$ slot as the assignment that maximizes (for $\alpha \neq 1$)

$$\sum_{i=1}^{n} \frac{(U_i(x_i(t), s_i(t)))^{1-\alpha}}{1-\alpha}.$$

for every $t = 1, ..., T$. This is equivalent to maximizing

$$\sum_{t=1}^{T} \sum_{i=1}^{n} \frac{(U_i(x_i(t), s_i(t)))^{1-\alpha}}{1-\alpha}. \tag{7.2}$$

For $\alpha = 1$ the same is true but where we replace

$$\frac{(U_i(x_i(t), s_i(t)))^{1-\alpha}}{1-\alpha}$$

by

$$\log[U_i(x_i(t), s_i(t))]$$

We make the following surprising observation: The optimization problem (7.2) corresponds to the $\alpha$-fair assignment problem in which there are $nT$ players instead of $n$ players, where the utility of player $i = kn + j$ ($k = 0, ..., T-1, j = 1, ..., n$) is defined as

$$\mathcal{U}_i(x_i, s_i) = U_j(x_j(k+1), s_j(k+1))$$

. Thus the expected instantaneous fairness criterion in the stationary and ergodic case regards assignments at different time slots of the same player as if it were a different player at each time slot!

Note that when considering the proportional fair assignment, then the resulting assignment is the one that maximizes

$$\prod_{i=1}^{n} \prod_{t=1}^{T} U_i(x_i(t), s_i(t))$$

**Definition 7.2.1.2.** *Assume that the state process $X(t)$ is stationary ergodic. Let $\lambda_i$ be the stationary probability measure of $X(0)$. The long term α-fairness index of an assignment $s \in \mathbf{S}$ of a stationary process $X(t)$ is defined as*

$$\begin{aligned} \overline{Z}_\lambda(s) &:= \sum_{i=1}^{n} \overline{Z}_\lambda^i(s) \\ \text{where } Z_\lambda^i(s) &= \frac{\left(E_\lambda\left[U_i(X_i(0), s_i(X(0)))\right]\right)^{1-\alpha}}{1-\alpha} \end{aligned}$$

*An assignment s is long-term α-fair if it maximizes $Z_\lambda(s)$ over $s \in \mathbf{S}$.*

As we see, instead of attempting to have a fair assignment of the resources at every $t$, it is the expected utility in the stationary regime that one assigns fairly according to the the long-term fairness. Under stationarity and ergodicity conditions on the process $X(t)$ this amounts in an instantaneous assignment of the resources in a way that the time average amount allocated to the users are $\alpha$-fair.

### 7.2.2 Fairness over time: $T$-scale $\alpha$-fairness

Next we define fairness concepts that are in between the instantaneous and the expected fairness. They are related to fairness over a time interval $T$. Either continuous time is considered or discrete time where time is slotted and each slot is considered to be of one time unit. Below, we shall understand the integral to mean summation when ever time is discrete.

**Definition 7.2.2.1.** *The T-scale $\alpha$-fairness index of an assignment $s \in \mathbf{S}$ is defined as*

$$\begin{aligned} Z_T(s) &:= \sum_{i=1}^{n} Z_T^i(s) \\ \text{where } Z_T^i &= \frac{\left[\frac{1}{T}\int_0^T U_i(X_i(t), s_i(X(t)))dt\right]^{1-\alpha}}{1-\alpha} \end{aligned}$$

*The expected T-scale $\alpha$-fairness index is its expectation. An assignment $s$ is T-scale $\alpha$-fair if it maximizes $Z_T(s)$ over $s \in \mathbf{S}$.*

**Definition 7.2.2.2.** *The T-scale expected $\alpha$-fairness index of an assignment $s \in \mathbf{S}$ is defined as*

$$\begin{aligned} Z_T(s) &:= \sum_{i=1}^{n} Z_T^i(s) \\ \text{where } Z_T^i &= \frac{\left[\frac{1}{T}\int_0^T E[U_i(X_i(t), s_i(X(t)))]dt\right]^{1-\alpha}}{1-\alpha} \end{aligned}$$

We shall consider the following simple example of 2-scale fairness

**Example 7.2.2.3.** *Consider two time slots and two mobile stations. To whoever the first time slot will be allocated, that mobile would send or receive 25 units. At the second slot, a rate of 5 (resp. 10) units will be used if the slot is assigned to mobile 1 (resp. 2). We make the following observations. By [i,j] we shall denote the allocation that assigns slot 1 to mobile i and slot 2 to mobile j. The allocation [1,2] maximizes the global utility and moreover, the $\alpha$-fair 2-scale utility for any $\alpha$.*

Thus, we observe that the $\alpha$-assignment is not monotone: The player with larger utilities received less at the $\alpha$-fair utility, for all values of $\alpha$!

**Example 7.2.2.4.** *(Example 7.2.2.3 continued) We now change a single utility in the last example: assume that if mobile 2 receives the first slot then it earns $10^2$ units.*

*(i) Now the global optimal solution is the assignment [2,2].*

*(ii) The proportional fair solution ($\alpha = 1$) is [2,1].*

*(iii) The maxmin fair assignment is [1,2].*

*We depict in Figure 7.1 the performance index of the assignments [1,2] and [2,1]. We see that the max-min fair assignment [2,1] is 2-scale α-fair for all α larger than* 1.36*, whereas the assignment [1,2] is α fair for* $\alpha \in [1, 1.36]$.

*For* $\alpha < 1$ *the two best assignments are [2,1] and [2,2]. The former is optimal over* $\alpha \in [0.17, 1]$ *and the latter over* $\alpha \in [0, 0.17]$. *This is seen from Figure 7.2.*

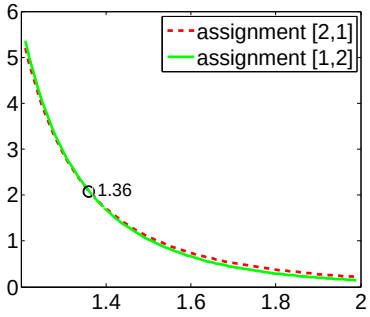


***Figure 7.1:*** *Performance index of [2,1] (dashed line) and [1,2] (solid line) assignments as a function of α (horizontal axis)*

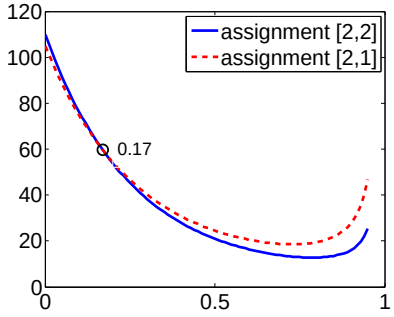


***Figure 7.2:*** *Performance index of [2,1] (dashed line) and [2,2] (solid line) assignments as a function of α (horizontal axis)*

Assume that the state processes is stationary ergodic. Then for any assignment $s \in \mathbf{S}$ we would have by the Strong Law of Large Numbers:

$$\begin{aligned}\lim_{T\to\infty} \frac{1}{T}\int_0^T U_i(X_i(t), s_i(X(t)))dt \\ = E_\lambda\left[U_i(X_i(0), s_i(X(0)))\right]\end{aligned}$$

P-a.s. Hence, for every $i$ and $s$, we have P-a.s.

$$\begin{aligned}\lim_{T\to\infty} Z_T^i(s) &= \lim_{T\to\infty} \frac{\left[\frac{1}{T}\int_0^T U_i(X_i(t), s_i(X(t)))dt\right]^{1-\alpha}}{1-\alpha} \\ &= \frac{\left(E_\lambda\left[U_i(X_i(0), s_i(X(0)))\right]\right)^{1-\alpha}}{1-\alpha} \\ &= \overline{Z}_T^i(s).\end{aligned}$$

Assume that $U_i$ is bounded. Then $Z_T^i$ is bounded uniformly in $T$. The bounded convergence then implies that

$$\lim_{T\to\infty} E[Z_T^i(s)] = \overline{Z}_T^i(s). \tag{7.3}$$

**Theorem 7.2.2.5.** *Assume that the convergence in (7.3) is uniform in s. Let* $s^*(T)$ *be the T-scale α fair assignment and let* $s^*$ *be the long term α-fair assignment. Then the following holds:*

- $s^* = \lim_{T\to\infty} S^*(T)$
- *For any $\epsilon > 0$, $s^*$ is an $\epsilon$-optimal assignment for the $T$-scale criterion for all $T$ large enough.*
- *For any $\epsilon > 0$, $s^*(T)$ is an $\epsilon$-optimal assignment for the long term fairness for all $T$ large enough.*

**Proof.** According to [145], any accumulation point of $s^*(T)$ as $T \to \infty$ is an optimal solution to the problem of maximizing $\overline{Z}_T$ over $S$. Due to the strict concavity of $\overline{Z}_T$ in $s$ it has a unique solution and it is coincides with any accumulation point of $s^*(T)$. This implies the first statement of the theorem. The other statements follow from Appendices A and B in [145]. ◇

### 7.2.3 Fairness over different time scales: Multiscale fairness

We consider real time (RT) and non-real time (NRT) traffic. Resource allocation policy for RT traffic is *instantaneous-fair*, while for the NRT traffic, it is *expected-fair*. The available resources are divided amongst the RT and NRT traffic so as to guarantee a minimum quality of service (QoS) requirement for the RT traffic and to keep service time as short as possible for the NRT traffic.

The real time traffic would like the allocation to be instantaneously $\alpha$-fair. For $\alpha > 0$, this guarantees that at any time it receives a strictly positive allocation.

The non-real time traffic does not need to receive at each instant a positive amount of allocation. It may prefer the resources to be assigned according to the $T$-scale $\alpha$-fair assignment where $T$ may be of the order of the duration of the connection. Moreover, different non real time applications may have different fairness requirements. For instance, bulk FTP transfer can prefer fairness over time scale longer than a time scale for some streaming application.

In order to be fair, we may assign part (say half) of the resource according to the instantaneous $\alpha$-fairness and the rest of the resources according to the $T$-scale $\alpha$-fairness. We thus combine fairness over different time scales.

We may now ask how to choose what part of the resource would be split according to the instantaneous assignment and what part according to the $T$-scale assignment. We propose to determine this part using the same $\alpha$-fair criterion.

Specifically we define the multiscale fairness as follows:
**Definition 7.2.3.1.** *The multiscale $\alpha$-fairness index of an assignment $s \in \mathbf{S}$ is defined as*

$$\begin{aligned} Z_{T_1,\ldots,T_n}(s) &:= \sum_{i=1}^{n} Z^i_{T_i}(s) \\ \text{where } Z^i_{T_i} &= \frac{\left[\frac{1}{T_i}\int_0^{T_i} U_i(X_i(t), s_i(X(t)))dt\right]^{1-\alpha}}{1-\alpha} \end{aligned}$$

*The expected multiscale α-fairness index is its expectation. An assignment s is multiscale α-fair if it maximizes $Z_{T_1,...,T_n}(s)$ over $s \in \mathbf{S}$. We also say that multiscale α-fair assignment is $(T_1, ..., T_n)$-scale fair assignment.*

**Example 7.2.3.2.** *Let us consider an example of multiscale fairness. Say, we have in total N time slots. The allocation happens in a bundle of 6 slots, such that, either we allocate all of it to an outdoor user located at $x_1$ or fair share them amongst 3 indoor users located at $(y_1, y_2, y_3)$ with $y_i \in (0, L)$, with any user getting two consecutive slots. Now the question is "Given that we fair-share among the indoor users, how do we fair share between the outdoor and the indoor users?".*

*In this example, we assume any user gets a throughput $q \in [0, 1]$. Let $\{U_1, U_2\}$, represent the utility of user 1 and sum utility of users $2-4$ and let $\{T_1, 1-T_1\}$ represent their respective assignment of resources. Now, utility of user 1,*

$$U_1(T_1) = 6T_1 q_1(x_1).$$

*Let $\bar{s} = \{s_1, s_2, 1-s_1-s_2\}$ represent the assignment of resources for the indoor users. Then, utility of users $2-4$ is*

$$u_2(T_1, \bar{s}) = 6s_1(1-T_1)q_2(y_2),$$

$$u_3(T_1, \bar{s}) = 6s_2(1-T_1)q_3(y_3)$$

*and*

$$u_4(T_1, \bar{s}) = 6(1-s_1-s_2)(1-T_1)q_4(y_4)$$

*Now the α-fair share $\bar{s}^* = \{s_1^*, s_2^*, 1-(s_1^*-s_2^*)\}$ is given by,*

$$\bar{s}^* = \arg\max_{\bar{s}} \sum_{i=2}^{4} \frac{E[u_i(T_1, \bar{s})]^{1-\alpha_1}}{1-\alpha_1}$$

*The sum utility of users $2-4$ is,*

$$U_2(T_1) = \sum_{i=2}^{4} 6\bar{s_i}^*(1-T_1)q_i(y_i)$$

*The α-fair share between the outdoor and indoor users is,*

$$T_1^* = \arg\max_{T_1} \frac{E\left[U_1(T_1)\right]^{1-\alpha} + E\left[U_2(1-T_1)\right]^{1-\alpha}}{1-\alpha}$$

## 7.3 Instantaneous $\alpha$-fairness for linear resources

In the case of linear resources the instantaneous $\alpha$-fairness has a nice explicit expression.

**Theorem 7.3.0.3.** *(i) The α-fair share is given by*

$$s_i^* = \frac{v_i(x_i)^{1/\alpha}}{\sum_j v_j(x_j)^{1/\alpha}} = \frac{q_i(x_i)^{1/\alpha-1}}{\sum_j q_j(x_j)^{1/\alpha-1}}$$

*(ii) The utility for mobile i under the fair assignment is then*

$$U_i(s^*, x) = \frac{v_i(x_i)^{1/\alpha}}{\sum_j v_j(x_j)^{1/\alpha}} q_i(x_i) = \frac{q_i(x_i)^{1/\alpha}}{\sum_j q_j(x_j)^{1/\alpha-1}}$$

*(iii) The optimal value Z is given by*

$$Z = \frac{1}{1-\alpha} \sum_i \left( \frac{q_i(x_i)^{1+1/\alpha}}{\sum_j q_j(x_j)^{1/\alpha}} \right)^{1-\alpha}$$

**Proof.** We relax the constraint and use KKT condition. $s$ is optimal if and only if there is some $\lambda > 0$ such that $s$ maximize $L^\lambda$ s.t. $s_i \geq 0$ for all $i$, where

$$L^\lambda = \sum_i s_i^{1-\alpha} v_i + \lambda(1 - \sum_i s_i)$$

Equating the derivative w.r.t. $s_i$ to zero gives

$$\begin{aligned} s_i^{-\alpha} v_i(x_i) &= \frac{\lambda}{1-\alpha} \\ \text{so that } s_i &= \left( \frac{1-\alpha}{\lambda} v_i(x_i) \right)^{1/\alpha} \end{aligned}$$

Since the sum of $s_i$ is 1, we conclude that

$$\frac{\lambda}{1-\alpha} = \left( \sum_j v_j(x_j)^{1/\alpha} \right)^\alpha$$

Substituting in the previous equation yields (i), which then implies the rest. ◇

**Example 7.3.0.4.** *Consider as an example a path loss $\beta = 2$ and let the base station be located one unit above the mobiles. We assume that $q_i(x)$ is proportional to the attenuation between the mobile and the base station: $q_i(x) = c_i q(x)$ where $q(x) = (1 + x^2)^{-1/2}$. For $\beta = \alpha = 2$ we have*

$$s_i^*(x) = \frac{c_i^{-1/2}(1 + x_i^2)^{-1/2}}{\sum_j c_j^{-1/2}(1 + x_j^2)^{-1/2}}.$$

*Furthermore,*

$$U_i(s^*, x) = s_i^*(x) q_i(x_i) = \frac{c_i^{1/2}(1 + x_i^2)}{\sum_j c_j^{-1/2} \sqrt{1 + x_j^2}}.$$

## 7.4 Application to spectrum allocation in random fading channels

We consider two users: fast-changing user and slowly-changing user. The users' channels are modeled by the Gilbert model. The users can be either in a good or in a bad state. The dynamics of the fast-changing user is described by a Markov chain $\{Y_1(t)\}_{t=0,1,...}$ with the following transition matrix

$$P_1 = \begin{bmatrix} 1-\alpha_1 & \alpha_1 \\ \beta_1 & 1-\beta_1 \end{bmatrix}.$$

Its stationary distribution is given by

$$\pi_1 = \begin{bmatrix} \frac{\beta_1}{\alpha_1+\beta_1} & \frac{\alpha_1}{\alpha_1+\beta_1} \end{bmatrix}.$$

The slowly-changing user is described by a Markov chain $\{Y_2(t)\}_{t=0,1,...}$ with the following transition matrix

$$P_2 = \begin{bmatrix} 1-\epsilon\alpha_2 & \epsilon\alpha_2 \\ \epsilon\beta_2 & 1-\epsilon\beta_2 \end{bmatrix}.$$

Its stationary distribution is given by

$$\pi_2 = \begin{bmatrix} \frac{\beta_2}{\alpha_2+\beta_2} & \frac{\alpha_2}{\alpha_2+\beta_2} \end{bmatrix}.$$

Note that the parameter $\epsilon$ does not have an effect on the stationary distribution but it influences for how long the slowly-changing user stays in some state. The smaller $\epsilon$, the more seldom the user changes the states. If we choose $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$, then the fast-changing user and the slowly-changing user have the same stationary distribution.

We assume that state 1 is a bad state and state 2 is a good state. When the fast-changing user is in the bad state, its channel gain coefficient is $h_{11}$ and when the fast-changing user is in the good state, its channel gain coefficient is $h_{12}$. Of course, we have $h_{11} < h_{12}$. Thus, the achievable throughputs in different states are given by

$$\begin{aligned} U_{11} &= s_{11}q_{11} = s_{11}\log\left(1+\frac{h_{11}P_1}{\sigma^2}\right), \\ U_{12} &= s_{12}q_{12} = s_{12}\log\left(1+\frac{h_{12}P_1}{\sigma^2}\right) \end{aligned}$$

where $P_1$ is the power applied by the fast-changing user.

Similarly, for the slowly-changing user we associate with the bad state (state 1) the channel gain $h_{21}$ and with the good state (state 2) the channel gain $h_{22}$. Again we have $h_{21} < h_{22}$, and the achievable throughputs in different states are given by

$$\begin{aligned} U_{21} &= s_{21}q_{21} = s_{21}\log\left(1+\frac{h_{21}P_2}{\sigma^2}\right), \\ U_{12} &= s_{22}q_{22} = s_{22}\log\left(1+\frac{h_{22}P_2}{\sigma^2}\right) \end{aligned}$$

where $P_2$ is the power applied by the slowly-changing user.

First, we would like to analyze $T$-scale fairness and to see the effect of the time scale on the resource allocation. Specifically, we consider the following optimization criterion

$$\sum_{i=1}^{2} \frac{1}{1-\alpha} \left[ \frac{1}{T} \sum_{t=0}^{T} U_i(t) \right]^{1-\alpha} \quad \to \quad \max_{s_1, s_2} \tag{7.4}$$

with $U_i(t) = s_i(t) q_{i,Y_i(t)}$ and $s_1(t) + s_2(t) = 1$.

Let us consider several options for the time horizon $T$:

**Instantaneous fairness.** If we take $T = 1$ we obtain the instantaneous fairness. Namely, the criterion (7.4) takes the form

$$\frac{1}{1-\alpha} \left[ U_1^{1-\alpha}(0) + U_2^{1-\alpha}(0) \right] \quad \to \quad \max_{s_1, s_2}$$

The solution of the above optimization problem (follows from Theorem 1) is given by

$$s_i(0) = \frac{q_{i,Y_i(0)}^{(1-\alpha)/\alpha}}{q_{1,Y_1(0)}^{(1-\alpha)/\alpha} + q_{2,Y_2(0)}^{(1-\alpha)/\alpha}}$$

This allocation results in the following expected throughputs

$$\begin{aligned} \theta_1 &= \sum_{i,j} \frac{q_{1,i}^{1/\alpha}}{q_{1,i}^{(1-\alpha)/\alpha} + q_{2,j}^{(1-\alpha)/\alpha}} \pi_{1,i} \pi_{2,j}, \\ \theta_2 &= \sum_{i,j} \frac{q_{2,j}^{1/\alpha}}{q_{1,i}^{(1-\alpha)/\alpha} + q_{2,j}^{(1-\alpha)/\alpha}} \pi_{1,i} \pi_{2,j}. \end{aligned} \tag{7.5}$$

**Mid-term fairness.** Let us take the time horizon as a function of the underlying dynamics time parameter $\epsilon$, that is $T = T(\epsilon)$, satisfying the following conditions: (a) $T(\epsilon) \to \infty$ and (b) $T(\epsilon)\epsilon \to 0$. The condition (a) ensures that

$$\frac{1}{T(\epsilon)} \sum_{t=0}^{T(\epsilon)} 1\{Y_1(t) = i\} \to \pi_{1,i}, \quad \text{as} \quad \epsilon \to 0,$$

and the condition (b) ensures that

$$\frac{1}{T(\epsilon)} \sum_{t=0}^{T(\epsilon)} 1\{Y_2(t) = i\} \to \delta_{Y_2(0),i}, \quad \text{as} \quad \epsilon \to 0.$$

This follows from the theory of Markov chains with multiple time scales (see e.g., [58]). It turns out to be convenient to take the following notation for the resource allocation: We denote by $s(t)$ the allocation for the fast-changing user and by $1 - s(t)$ the resource

allocation for the slowly-changing user. Thus, we have $s_1(t) = s(t)$ and $s_2(t) = 1 - s(t)$. We denote by $\bar{s}_{i,j} = E[s(t)|Y_1(t) = i, Y_2(t) = j]$. We note that since the fast-changing user achieves stationarity when $T(\epsilon) \to \infty$ we are able to solve (7.4) in stationary strategies. Then, the criterion (7.4) takes the form

$$\begin{aligned}\frac{1}{1-\alpha} \quad & \Big[(\pi_{1,1}q_{1,1}\bar{s}_{1,Y_2(0)} + \pi_{1,2}q_{1,2}\bar{s}_{2,Y_2(0)})^{1-\alpha} \\ & + \ ((1 - \pi_{1,1}\bar{s}_{1,Y_2(0)} - \pi_{1,2}\bar{s}_{2,Y_2(0)})q_{2,Y_2(0)})^{1-\alpha}\Big] \\ & \to \quad \max_{\bar{s}_{1,Y_2(0)},\, \bar{s}_{2,Y_2(0)}}\end{aligned}$$

The above nonlinear optimization problem can be solved numerically. The expected throughputs in the mid-term fairness case are given by

$$\begin{aligned}\theta_1 &= (\pi_{1,1}q_{1,1}\bar{s}_{1,1} + \pi_{1,2}q_{1,2}\bar{s}_{2,1})\pi_{2,1} \\ &+ (\pi_{1,1}q_{1,1}\bar{s}_{1,2} + \pi_{1,2}q_{1,2}\bar{s}_{2,2})\pi_{2,2} \\ \theta_2 &= (1 - \pi_{1,1}\bar{s}_{1,1} - \pi_{1,2}\bar{s}_{2,1})q_{2,1}\pi_{2,1} \\ &+ (1 - \pi_{1,1}\bar{s}_{1,2} - \pi_{1,2}\bar{s}_{2,2})q_{2,2}\pi_{2,2}\end{aligned} \tag{7.6}$$

**Long-term fairness.** In the case of long-term fairness we set $T = \infty$ which results in the following criterion

$$\frac{1}{1-\alpha}\Big[E[U_1]^{1-\alpha} + E[U_2]^{1-\alpha}\Big] \quad \to \quad \max_{s_1,s_2}$$

Due to stationarity, we can solve the above optimization problem over sequences in stationary strategies. Namely, we have the following optimization problem

$$\begin{aligned}\frac{1}{1-\alpha} \quad & ((\pi_{1,1}\pi_{2,1}\bar{s}_{1,1} + \pi_{1,1}\pi_{2,2}\bar{s}_{1,2})q_{1,1} \\ & + \ (\pi_{1,2}\pi_{2,1}\bar{s}_{2,1} + \pi_{1,2}\pi_{2,2}\bar{s}_{2,2})q_{1,2})^{1-\alpha} \\ & + \ ((\pi_{2,1} - \pi_{1,1}\pi_{2,1}\bar{s}_{1,1} - \pi_{1,2}\pi_{2,1}\bar{s}_{2,1})q_{2,1} \\ & + \ (\pi_{2,2} - \pi_{1,1}\pi_{2,1}\bar{s}_{1,1} - \pi_{1,2}\pi_{2,2}\bar{s}_{2,2})q_{2,2})^{1-\alpha}] \\ & \to \quad \max_{\bar{s}_{1,1},\, \bar{s}_{1,2},\, \bar{s}_{2,1},\, \bar{s}_{2,2}}\end{aligned}$$

The expected throughputs in the long-term fairness case are given by

$$\begin{aligned}\theta_1 &= (\pi_{1,1}\pi_{2,1}\bar{s}_{1,1} + \pi_{1,1}\pi_{2,2}\bar{s}_{1,2})q_{1,1} \\ &+ (\pi_{1,2}\pi_{2,1}\bar{s}_{2,1} + \pi_{1,2}\pi_{2,2}\bar{s}_{2,2})q_{1,2} \\ \theta_2 &= (\pi_{2,1} - \pi_{1,1}\pi_{2,1}\bar{s}_{1,1} - \pi_{1,2}\pi_{2,1}\bar{s}_{2,1})q_{2,1} \\ &+ (\pi_{2,2} - \pi_{1,1}\pi_{2,1}\bar{s}_{1,1} - \pi_{1,2}\pi_{2,2}\bar{s}_{2,2})q_{2,2}\end{aligned} \tag{7.7}$$

Let us also consider the expected instantaneous fairness which is given by the criterion

$$\frac{1}{1-\alpha}\Big[E[U_1^{1-\alpha}(t)] + E[U_2^{1-\alpha}(t)]\Big] \quad \to \quad \max_{s_1,s_2}$$

which is equivalent to

$$\frac{1}{1-\alpha} \quad \left[ \sum_{ij} \pi_{1,i}\pi_{2,j} \int_0^1 (sq_{1,i})^{1-\alpha} dF_{ij}(s) \right.$$
$$\left. + \sum_{ij} \pi_{1,i}\pi_{2,j} \int_0^1 ((1-s)q_{2,j})^{1-\alpha} dF_{ij}(s) \right] \quad \rightarrow \quad \max_{F_{ij}}$$

where $F_{ij}(s)$ is the distribution for $s(t)$ conditioned on the event $\{Y_1(t) = i, Y_2(t) = j\}$. The above criterion is maximized by

$$F_{ij}(s) = \begin{cases} 0, & \text{if } s < q_{1,i}^{(1-\alpha)/\alpha}/(q_{1,i}^{(1-\alpha)/\alpha} + q_{2,j}^{(1-\alpha)/\alpha}), \\ 1, & \text{if } s \geq q_{1,i}^{(1-\alpha)/\alpha}/(q_{1,i}^{(1-\alpha)/\alpha} + q_{2,j}^{(1-\alpha)/\alpha}). \end{cases}$$

Thus, we can see that the expected instantaneous fairness criterion is equivalent to instantaneous fairness.

Let us consider a numerical example. The parameters are given in Table 1. We consider three typical cases. For these three cases, we plot the expected throughput of the mobiles for various fairness criteria (see Figures 7.3-7.5). The first case corresponds to the symmetric scenario. Since in this scenario the users have the same stationary distributions and the same conditional Shannon capacities, the expected throughputs are the same when we use either instantaneous fairness criterion or long-term fairness criterion. Interesting, in the long-term fairness case, both users experience degradation in throughput when $\alpha$ increases. In the case of mid-term fairness, the expected throughput of the fast-changing user is higher as in the mid-term fairness criterion the utility of the fast-changing user is the $\alpha$-fairness function of the expected throughput. In the second case, the fast-changing user has in general better channel conditions. In this case, different fairness criteria provide different resource allocation. We observe that instantaneous and mid-term fairness allocations are more sensitive with respect to the parameter $\alpha$ than the long-term fairness allocation. In the third scenario the slowly-changing user (user 2) is more often in the good channel state than the fast-changing user (user 1). Now, for all the criteria the slowly-changing user has better expected throughput.

Next, let us consider multiscale fairness over time. Specifically, $(T_1, T_2)$-scale fairness is defined by the following criterion

$$\frac{1}{1-\alpha} \left[ \left( \frac{1}{T_1} \sum_{t=0}^{T_1} U_1(t) \right)^{1-\alpha} + \left( \frac{1}{T_2} \sum_{t=0}^{T_2} U_2(t) \right)^{1-\alpha} \right] \quad \rightarrow \quad \max_{s_1,s_2}$$

In this particular example, there are 6 possible combinations of different time scales. It turns out that in this example only the $(1,\infty)$-scale fairness gives a new resource allocation. The other combinations of time scales reduce to some $T$-scale fairness. Thus, let us first consider the multiscale fairness when we apply instantaneous fairness to the

***Table 7.1:*** *Case 1,2 & 3: Shannon capacity (q)/probability($\pi$)*

| **Case-1** | state-1 (bad) | state-2 (good) |
|---|---|---|
| User-1 | 2/0.2 | 8/0.8 |
| User-2 | 2/0.2 | 8/0.8 |
| **Case-2** | state-1 (bad) | state-2 (good) |
| User-1 | 3/0.1 | 9/0.9 |
| User-2 | 1/0.3 | 7/0.7 |
| **Case-3** | state-1 (bad) | state-2 (good) |
| User-1 | 3/0.9 | 9/0.1 |
| User-2 | 1/0.3 | 7/0.7 |

fast-changing user and long-term fairness to the slowly-changing user. The $(1,\infty)$-scale fairness corresponds to the following optimization criterion

$$\frac{1}{1-\alpha}\left[U_1(0)^{1-\alpha}+E[U_2(t)]^{1-\alpha}\right] \quad \rightarrow \quad \max_{s_1,s_2}$$

which is equivalent to

$$\begin{aligned}\frac{1}{1-\alpha} \quad & \Big[(q_{1,Y_1(0)}(\bar{s}_{Y_1(0),1}\pi_{2,1}+\bar{s}_{Y_1(0),2}\pi_{2,2}))^{1-\alpha} \\ & +(q_{2,1}(1-\bar{s}_{Y_1(0),1})\pi_{2,1}+q_{22}(1-\bar{s}_{Y_1(0),2})\pi_{2,2})^{1-\alpha}\Big] \\ & \rightarrow \quad \max_{\bar{s}_{Y_1(0),1},\,\bar{s}_{Y_1(0),2}}\end{aligned}$$

The expected throughputs in the $(1,\infty)$-scale fairness case are given by

$$\begin{aligned}\theta_1 &= q_{1,1}(\bar{s}_{1,1}\pi_{1,1}\pi_{2,1}+\bar{s}_{1,2}\pi_{1,1}\pi_{2,2}) \\ &+ q_{1,2}(\bar{s}_{2,1}\pi_{1,2}\pi_{2,1}+\bar{s}_{2,2}\pi_{1,2}\pi_{2,2}) \\ \theta_2 &= (q_{2,1}(1-\bar{s}_{1,1})\pi_{2,1}+q_{22}(1-\bar{s}_{1,2})\pi_{22})\pi_{1,1} \\ &+ (q_{2,1}(1-\bar{s}_{2,1})\pi_{2,1}+q_{22}(1-\bar{s}_{2,2})\pi_{22})\pi_{1,2}.\end{aligned}$$

As we have mentioned above, the other combinations of time scales reduce to some $T$-scale fairness. In particular, $(1,T(\epsilon))$-fairness reduces to the instantaneous fairness, $(T(\epsilon),\infty)$-fairness reduces to long-term fairness, and $(T(\epsilon),1)$-, $(\infty,1)$- and $(\infty,T(\epsilon))$-fairness all reduce to mid-term fairness.

We also plot the expected throughputs for $(1,\infty)$-scale fair allocation for the numerical example with three cases (see Figures 7.3-7.5). We observe that in the symmetric case $(1,\infty)$-scale fairness criterion provides an allocation which is opposite to the allocation provided by the mid-term fairness criterion. This indicates that the $T$-scale fairness and multiscale fairness concepts provide a versatile framework for resource allocation which takes into account the dynamics of the users. From Figures 7.4 and 7.5

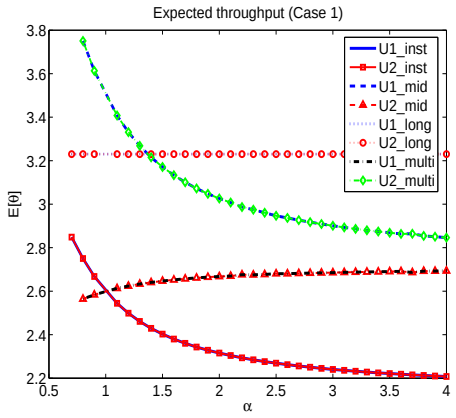


***Figure 7.3:*** *Throughput($\theta$) as a function of $\alpha$ for instantaneous, mid-term, long-term and (1,∞)-scale fairness criteria (Case 1).*

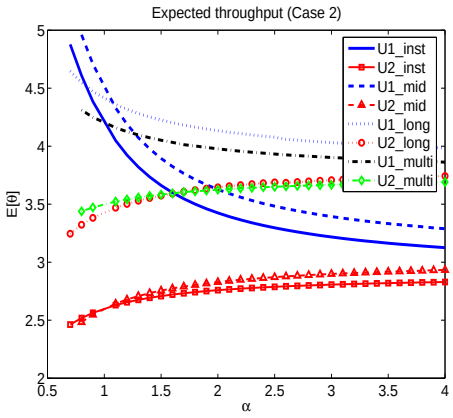


***Figure 7.4:*** *Throughput($\theta$) as a function of $\alpha$ for instantaneous, mid-term, long-term and (1,∞)-scale fairness criteria (Case 2).*

we conclude that multiscale fairness provide good sensitivity to the variation of the fairness parameter and good performance in expected throughput.

**Coefficient of variation:** We compute the coefficient of variation for short-term, mid-term, long-term and multiscale fairness. For this, we first compute the second moment of the throughput and then find the ratio of the standard deviation to its mean. For any user $i$, the coefficient of variation is

$$\Gamma_i = \frac{\sqrt{\mathbf{E}[\theta_i^2]}}{\mathbf{E}[\theta_i]}$$

In Figure 7.6, we plot the coefficient of variation in throughput for the considered above various fairness criteria. It is very interesting to observe that except the $(1,\infty)$-scale fairness criterion all the other fairness criteria behave similarly with respect to the coefficient of variation. Only in the case of $(1,\infty)$-scale fairness the coefficient of variation decreases for sort-term fairness oriented user. This is a very desirable property of the multiscale fairness as a short-term fairness oriented user is typically a user with a delay sensitive application.

## 7.5 Application to indoor-outdoor scenario

Let $\Omega$ be the line segment $[-L, L]$, and let there be a wall at $x = 0$. Assume that the base station is located just to the left of the wall. Mobile 1 is at some point $x \leq 0$ outdoor and user 2 remains always indoor and is located at some $Y_t$ which is uniformly distributed over $[0, L]$. We let $q_i(x) = c_i q(x)$ with $c_1 = 1$ and $c_2$ is equal to some large fixed number.

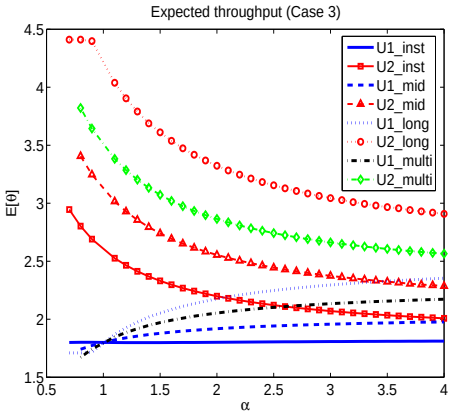


***Figure 7.5:*** *Throughput($\theta$) as a function of $\alpha$ for instantaneous, mid-term, long-term and (1,∞)-scale fairness criteria (Case 3).*

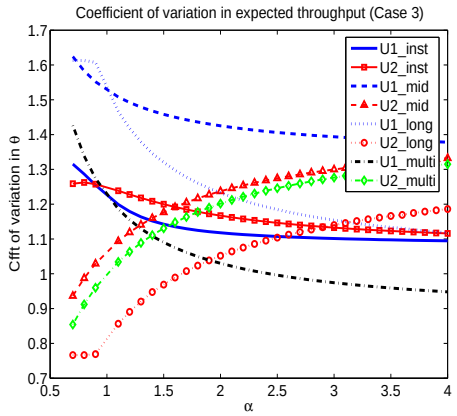


***Figure 7.6:*** *Coefficient of variation in expected throughput as a function of $\alpha$ for instantaneous, mid-term, long-term and (1,∞)-scale fairness criteria (Case 3).*

Thus the presence of the wall between the base station and mobile 2 is modeled by a multiplicative attenuation by some constant $c_2$. Assume that the mobility pattern of mobile 2 is uniform over the indoor part $[0, L]$.

We consider allocation of the fraction of time between the two mobiles.

### 7.5.1 Instantaneous Fairness

**Example 7.3.0.4 (continued).** We compute the expected utility for each user when assigning the channel so as to achieve instantaneous fairness. The expected utility for mobile 1 under the instantaneous optimal fairness **s** is given by

$$\begin{aligned} U_1(s^*, x) &= s^*(x) q_1(x), \text{ where } s^*(x) := \frac{a}{a+b} \\ a &:= c_1^{-1/2} \log\left(1 + \frac{1}{x^2}\right)^{-1/2} \\ b &:= c_2^{-1/2} \left[\log(1 + \frac{1}{L^2}) + \frac{2}{L}\tan^{-1}(L)\right]^{-1/2} \end{aligned}$$

Note that mobile 2 has a mobility pattern which is uniform over the indoor part [0,L] and hence its utility is given by $\frac{1}{L}\int_0^L \log(1+\frac{1}{x^2})dx = \log(1+\frac{1}{L^2}) + \frac{2}{L}\tan^{-1}(L)$, which is the second term in the denominator.

◇

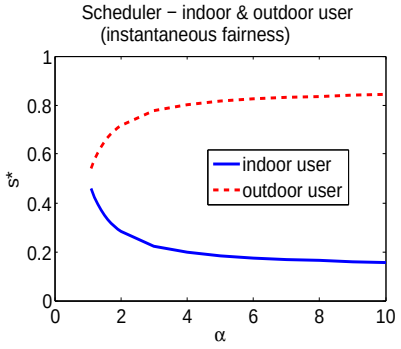


***Figure 7.7:*** *Scheduler $s^*$ for the indoor and outdoor user with instantaneous fairness as a function of $\alpha$ for $\alpha > 1$. Wall attenuation 6 dB, path-loss $\beta = 3$, position of outdoor user $x = -3$.*

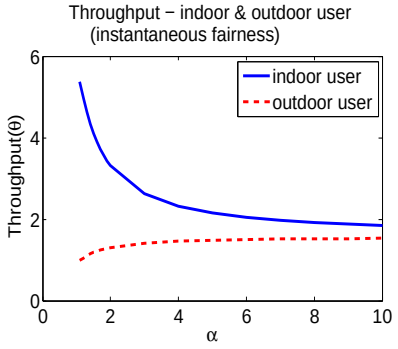


***Figure 7.8:*** *Throughput $\theta$ for the indoor and outdoor user with instantaneous fairness as a function of $\alpha$ for $\alpha > 1$. Wall attenuation 6 dB, path-loss $\beta = 3$, position of outdoor user $x = -3$.*

In figure (7.7) and (7.8), we plot the scheduler and the instantaneous throughput for the indoor and outdoor user, as a function of $\alpha$. We fix the location of the outdoor user at $x = -3$, path-loss $\beta = 3$. We set $L = 3$ for this example. The indoor user is located at some point which is uniformly distributed over $[0, L]$. When the fairness index $\alpha$ is small, we observe that the instantaneous throughput achieved is higher as the outdoor user is located at the boundary $(-L)$. But, as the fairness index increases, the throughput of the indoor and the outdoor user starts to converge. Notice that the scheduler starts to schedule the outdoor user more as $\alpha$ increases, which results in an increase in the outdoor users throughput.

### 7.5.2 Long term Fairness

Next we consider the long-term fairness. The long term allocation $s \in \mathbf{S}$ (which is a function of $x$ and $Y_t$) is given by maximizing

$$Z(s) := \frac{\left[\frac{1}{L}\int_0^L dy s_1(x,y)q(x)\right]^{1-\alpha} + \left[\frac{1}{L}\int_0^L dy c_2 s_2(x,y)q(y)\right]^{1-\alpha}}{1-\alpha}$$

**Theorem 7.5.2.1.** *The long term $\alpha$-fair policy is given by $s_2(x_2) = 1$ for $x_2 \leq l(\alpha)$ and is otherwise zero, where $l(\alpha)$ is the solution of the fixed point equation*

$$l(\alpha) = c_2^{1-\frac{1}{\alpha}} \left(\frac{q(l(\alpha))}{q(x)}\right)^{-\beta(1-\frac{1}{\alpha})}$$

*where $q(x)$ is a monotone decreasing function of the form $x^{-\beta}$*

**Proof.** It is easy to see that $\alpha$-fair policy has to have the form mentioned in the theorem statement. If not, for example say there exists an optimal policy which allocates mobile 2 in two disjoint intervals. Then, one can construct a better policy by shifting the right most interval to the end of the left interval and this contradicts the optimality.

Thus the optimization simplifies to one-dimensional optimization

$$\max_{s} Z(s) = \max_{l \in [0,L]} Z(s^l) \text{ where } s_2^l(x) = 1 \text{ for } \{x \leq l\}.$$

It is easy to see that

$$Z(s^l) = \frac{1}{1-\alpha}\left[\left(\frac{L-l}{L}\, q(x)\right)^{1-\alpha} + \left(\frac{1}{L}\int_0^l c_2\, q(y)dy\right)^{1-\alpha}\right].$$

The optimal $l(\alpha)$ is obtained by differentiating the above equation w.r.t $l$ and equating to zero, which results in the fixed point equation

$$\big((L-l(\alpha))q(x)\big)^{-\alpha} q(x) - \left(c_2 \int_0^{l(\alpha)} q(y)dy\right)^{-\alpha} q(l(\alpha)) = 0.$$

Specially when $q(x) = x^{-\beta}$ then the fixed point equation simplifies to

$$l(\alpha) = c_2^{1-\frac{1}{\alpha}} \left(\frac{q(l(\alpha))}{q(x)}\right)^{-\beta(1-\frac{1}{\alpha})}$$

◇

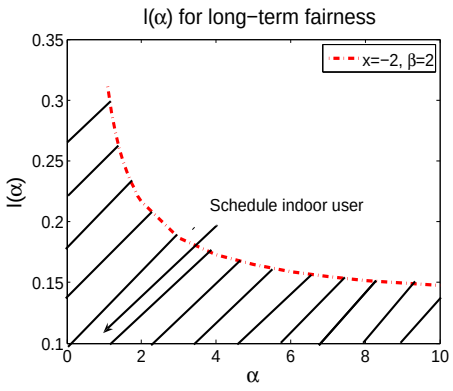


***Figure 7.9:*** *$l(\alpha)$ for long-term fairness as a function of $\alpha$ ($\alpha > 1$) and wall attenuation of 6 dB ,path-loss $\beta = 2$, position of outdoor user $x = -2$.*

We plot in figure (7.9) a numerical example to observe how $l(\alpha)$ varies with $\alpha$ for $\alpha > 1$. In this example, we consider path loss $\beta = 2$, location of outdoor user $x = -2$ and wall attenuation of 6 dB. We observe that as $\alpha$ increases, the value of $l(\alpha)$ monotonically decreases and starts to saturate. It is interesting to note that the indoor user is scheduled when its *mobility* and the *fairness* of resource allocation, $(l(\alpha), \alpha)$, lie within the dashed region below the curve. Also, when the user exhibits higher mobility, the range of fairness applicable reduces.

## 7.6 Conclusion and Future Research

We have introduced $T$-scale fairness and multiscale fairness. The notion of $T$-scale fairness allows one to address in a flexible manner requirements of emerging applications (like YouTube) which demand quality of service requirement between strict real time traffic and best effort traffic. The notion of multiscale fairness allows one to use a single optimization criterion for resource allocation when different applications are present in the network. We have compared the new fairness notions with previously known criteria of instantaneous and long-term fairness criteria. We have illustrated the new notions by their application in wireless networks. Specifically, we have considered spectrum allocation when users with different dynamics are present in the system. We have demonstrated that the multiscale fairness provides a versatile framework for resource allocation. We have also considered the resource allocation in indoor-outdoor scenario and have observed how the spacial component influences the resource scheduling under different fairness criteria. In the near future we plan to investigate in detail how multiscale fairness criterion allocates resources when a number of applications with different QoS requirements are present in the network. It is also interesting to investigate $T$-scale fairness in the non-stationary regime.

## 7.7 Publications

1. Eitan Altman, Konstantin Avrachenkov, Sreenath Ramanath, "Multiscale Fairness and its Application to Resource Allocation in Wireless Networks", proceedings of IFIP Networking 2011, LNCS 6631, May 9-13, Valencia, Spain.
2. Sreenath Ramanath, Eitan Altman, Konstantin Avrachenkov, "A heterogeneous approach to fair resource allocation and its application in femtocell networks", proceedings of the Workshop on Indoor and Outdoor Femto Cells (IOFC'11), May 13, Princeton USA.
3. Eitan Altman, Konstantin Avrachenkov, Sreenath Ramanath, "Multiscale Fairness and its Networking Application", Poster, Performance 2010, Nov 17-19, Namur, Belgium.

# Chapter 8

# Satisfying Demands in Multicell Networks: A Universal Power Allocation Algorithm

**Contents**



## 8.1 Introduction

Multi-input multi-output (MIMO) combined with network densification promise improved network coverage and capacity for mobile broadband access. But, due to an increased number of transmit antennas and or the proximity of base stations (BS), users at cell edges experience a higher degree of interference from neighboring base stations.

Network MIMO or other forms of BS co-operation enable sharing complete or statistical knowledge of channel states (CS) amongst neighbors via back-haul links to alleviate interference and offer better rates to users. When back-haul is not available, each BS may estimate the local channel state information and use the same for better performance. In some cases, a low rate feedback from the receiver indicating the QoS of the

current transmissions is utilized, while in the worst case the transceivers are designed with no CS information. Thus we have a variety of systems with varying degrees of the information about the interfering channels. However the goal in each is the same: satisfy the demands of all the users. We may require higher power profiles to satisfy the same demands when working with lesser information. Further diverse situations can arise because of the system configuration like modulation, precoding, channel coding, resource allocation etc.

For a given vector of power constraints at various base stations, Shannon capacity gives the maximum achievable rate, i.e., the capacity region. This is an upper bound. We define "system specific capacity region" (achievable rate region of a given system) which depend on coding (space-time, channel), modulation, channel state information availability, synchronization, feedback errors and many other things. Given a system architecture with a chosen set of parameters which define its rate allocation, modulation, etc, the achievable rates are usually inferior to the theoretical rates and the system specific capacity region is defined based on these rates. The system-specific capacity region for the same power constraint varies: for example it shrinks if the number of supported discrete rates reduce. Thus, the power allocated to any user to achieve the same demand rate varies with the set of system parameters.

The main contribution of this work is an universal algorithm which can work with many of the systems mentioned above. It satisfies asymptotically the demands of all the users irrespective of the system in which it is operating, albeit with different power profiles. *Each base station requires minimal information: its user's demands, its total power constraint and the current transmission rates to its users.* The current transmission rates are decided by the serving base stations either using complete CSIT (algorithm can also be used as a centralized scheme in this case) or has to be estimated completely blindly or using some partial information. The following are the contributions of this work:

1) A system specific game theoretic problem formulation using the system specific capacity region.

2) A *Stochastic Approximation* based universal power allocation algorithm in an interference limited multi-cell network.

3) Various properties (eg., convergence) of the proposed algorithm is analyzed using an ODE framework.

4) Simulation results demonstrate the effectiveness of the proposed algorithm for a variety of systems.

**Related Work:** For an excellent survey on power control in wireless networks, the reader is referred to [43] and the references there-in (eg. [60, 165, 78, 150, 55]). In recent years, several authors have addressed distributed power control strategies with various levels of co-operation for a given system configuration (eg. [171]). Typically, the design objective is to *maximize* the total sum rate of all the users subject to BS power constraints or to *minimize* the total transmit power satisfying some SINR constraints of the users.

Most of the existing algorithms aim at either optimizing the total power spent keep-

ing the QoS above a required level and or optimize the QoS while keeping the power utilized within a given budget. But our algorithm does not optimize, it only meets the demands (in the form of average transmission rates) on average asymptotically[1]. This relaxation helps us in proposing an algorithm that requires minimal information (hence has minimal complexity) at the transmitters: rates at which the information is correctly transmitted to the user in every slot. Data is pumped out from the transmitter and hence these rates are readily known to the transmitter. Hence this algorithm does not require any extra information and this can be exploited in many more ways. For example, one can probably use this algorithm in networks with heterogeneous cells, i.e., when each cell has a system configuration that can be different from the other cells.

**Organization:** We introduce the system model in section 8.2. In section 8.3, we describe the system specific problem formulation. The algorithm and its analysis is presented in section 8.4. Section 8.5 provides simulations. Appendix contains example systems and proofs.

**Notations:** Boldface lower-case symbols represent vectors, capital boldface symbols denote matrices ($\mathbf{I}_N$ is the $N \times N$ identity matrix). Hermitian transpose is denoted $(\cdot)^{\mathsf{H}}$ while $\mathbf{tr}[\mathbf{X}]$ represents the trace of matrix $\mathbf{X}$. All logarithms are base-2 logarithms. Small letters represent the scalars. Let $a_k$ represent the $k^{\text{th}}$ component of the vector $\mathbf{a}$. If the vector is already indexed like for example in $\mathbf{p}_j$, then $p_{k,j}$ represents its $k^{\text{th}}$ component. Let $(\mathbf{p}.\mathbf{s})$ represent the component-wise product, i.e., $(\mathbf{p}.\mathbf{s})_k = p_k s_k$ for all $k$ while $\sqrt{\mathbf{p}}$ represents component wise square root. $\mathbf{E}[\cdot]$ denotes expectation and $\mathbf{E}_{\mathbf{s}}$ is expectation w.r.t to $\mathbf{s}$ when conditioned (if any) on the other random variables.

## 8.2 System model

We consider a multi-cell MIMO system. Each base station has $M$ transmit antennas and is communicating with $K$ single-antenna users (see figure 8.1). Every user experiences both intra-cell (transmissions from parent BS) and inter-cell (transmissions from neighboring BS) interference. Each user in a cell demands a certain rate and all these rates have to be jointly satisfied by the BS (present in the cell) while operating within a total power constraint.

Let $\mathbf{H}_{j,l}$ represent the $K \times M$ channel matrix, when the users in cell $j$ receive signals from the BS of cell $l$ and let its elements be given by zero-mean unit-variance i.i.d. complex Gaussian entries. Let $\mathbf{n}_j$ represent the additive white Gaussian noise at the receivers of cell $j$, $\mathbf{x}_j$ be the $M$ length transmit vector in cell $j$ and $\gamma_l \in [0,1]$ be the interference factor, representative of the level of interference from cell $l$. For example, as base stations become denser, interference increases and hence $\gamma_l \to 1$. The signal

[1]We show that the demand meeting power profile to be a NE of a 'leaky' game. We call this game 'leaky', because the utility of the game is upper bounded by the demands (see definition (8.5), section 8.3.1). In summary our aim is to provide a channel, to each one of the users, whose (system specific) capacity is more than or equal to the user's demand.

vector (of length $K$) received by users in cell $j$ is given by,

$$\mathbf{y}_j = \mathbf{H}_{j,j}\mathbf{x}_j + \sum_{l=1,l\neq j}^{N} \gamma_l \mathbf{H}_{j,l}\mathbf{x}_l + \mathbf{n}_j \text{ for all } j \leq N. \tag{8.1}$$

In the above the first term represents the useful signal part as well as the intra-cell interference while the second term (summation) represents the inter-cell interference to the $j^{th}$ cell from its neighbors.

If $\bar{P}_j$ represents the total power constraint in cell $j$, then $\mathbf{tr}(\mathbf{E}[\mathbf{x}_j\mathbf{x}_j^H]) \leq \bar{P}_j$ to satisfy the power constraint. As an example if the BS in cell $j$ uses power levels specified by $\mathbf{p}_j$ and a *precoding matrix* $\mathbf{G}_j$ *(of size $M \times K$),* then the transmit vector is given by $\mathbf{x}_j = \mathbf{G}_j(\sqrt{\mathbf{p}_j}.\mathbf{s}_j)$ ($\mathbf{s_j}$ is a $K$ length independent symbol vector of zero mean and unit variance components). In this case the power constraint leads to,

$$\mathbf{tr}(\mathbf{E}[\mathbf{x}_j\mathbf{x}_j^H]) \leq \mathbf{tr}(\mathbf{E}[\mathbf{G}_j\sqrt{\mathbf{p}_j}(\mathbf{G}_j\sqrt{\mathbf{p}_j})^H]) \leq \bar{P}_j \text{ for any } j.$$

Given a precoding scheme, this constraint can equivalently be represented by (for a possibly different $\bar{P}_j$) $\sum_k p_{k,j} \leq \bar{P}_j$. The symbol, $y_{k,j}$, received by the user $k$ of cell $j$ is,

$$\begin{aligned} y_{k,j} &= \mathbf{h}_{k,j,j}^{\mathsf{H}}\mathbf{x}_j + \sum_{i=1,i\neq k}^{K} \mathbf{h}_{i,j,j}^{\mathsf{H}}\mathbf{x}_j + \sum_{l=1,l\neq j}^{N}\sum_{i=1}^{K} \gamma_l \mathbf{h}_{i,j,l}^{\mathsf{H}}\mathbf{x}_l + n_{k,j} \\ &= u_{k,j} + i_{k,j,j} + \sum_{l\neq j} i_{k,j,l} + n_{k,j} \end{aligned} \tag{8.2}$$

where $\mathbf{h}_{k,j,l}$, the $k^{th}$ row of matrix $\mathbf{H_{j,l}}$, represents the $M$ length channel vector for user $k$ of cell $j$ as received from the BS of cell $l$. In the above, $u_{k,j}$, $i_{j,j,k}$ and $i_{k,j,l}$ respectively represent the useful, intra-cell interference and inter-cell interference signal, respectively.

### System with No Precoding:

This work proposes an algorithm which works for any system in general. By *system, we mean a particular multi-cell network with a given configuration like, precoding scheme, channel coding, resource allocation etc.* We will derive the exact received signal characteristics for one such example system. The received signal characteristics of the others system can be derived in a similar way. We consider a system with no precoding (for example, systems which does not have access to channel state information). Further we consider a system with $M = K$ and with $\mathbf{x}_j = (\sqrt{\mathbf{p}_j}.\mathbf{s}_j)$. The average power in the useful, intra-cell, inter-cell interference signals of the received signal (after channel coding at the transmitter and channel decoding at the receiver) after averaging w.r.t. to the symbol statistics $\{\mathbf{s}_j\}$ for any given channel state:

$$\begin{aligned} \mathbf{E}_{\mathbf{s}_j, 1\leq j\leq N}[|u_{k,j}|^2] &= p_{k,j}|h_{k,j,j,k}|^2, \\ \mathbf{E}_{\mathbf{s}_j, 1\leq j\leq N}[|i_{k,j,j}|^2] &= \sum_{\bar{k}\neq k} p_{\bar{k},j}|h_{k,j,j,\bar{k}}|^2 \text{ and} \\ \mathbf{E}_{\mathbf{s}_j, 1\leq j\leq N}[|i_{k,j,l}|^2] &= \sum_{\bar{k}} \gamma_l p_{\bar{k},l}|h_{k,j,l,\bar{k}}|^2 \end{aligned} \tag{8.3}$$

where, $h_{k,j,l,\bar{k}}$ is the $(k,\bar{k})^{\text{th}}$ component of the matrix $\mathbf{H_{j,l}}$. In the above we used $E[s_{k,j}s^*_{k',j'}] = 1_{\{k=k',j=j'\}}$.

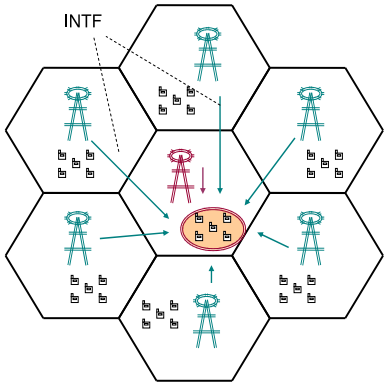


***Figure 8.1:*** *2D Wyner model*

| I CSIT | | II TX rate | | III Precoder | |
|---|---|---|---|---|---|
| **A** | Asymptotic | **I** | Ideal | **ZF** | Zero-forcing |
| **C** | Full CSIT | **D** | Discrete | **NO** | No precoder |
| **L** | Local CSIT | **RA** | Rate adaptation | | |
| **N** | No CSIT | **RAE** | RA with errors | | |

**Table 1:** System specification (I-II-III)

## 8.3 System specific problem formulation

Every BS has to meet its users demands, for example BS $j$ has to meet its users demand rates represented by $\mathbf{r}_j := \{r_{k,j}, k \leq K\}$. It has to tune its power levels $\mathbf{p}_j$ to achieve this. But the rates achieved will also depend upon the powers used by the other base stations. Our goal is to find a simple universal power allocation algorithm which runs independently and simultaneously at all the base stations and tunes the power levels to achieve the user demands using minimal information. The power levels depend upon the system configuration (for example channel precoding scheme, rate allocation scheme). We consider some interesting example systems briefed in Table **2** and described in Appendix A. These systems are referred using a three part code, I-II-III, as explained below:

1) The first part (I) represents the **availability of channel state information at transmitter**[2]: a) **A** represents an asymptotic (large number of antennas/users) system, where achievable rates for all most all CS are approximated by a constant (see [117] and references there in), b) **C** for systems with complete CSIT, c) **L**, systems with local CSIT, i.e., BS $j$ knows $\mathbf{H}_{j,j}$ part of the CS, d) **N**, systems with no CSIT.

[2]One can also consider systems which have an estimate of the CS.

| **System Description** (more details in Appendix A) | $R_{k,j}^{sys}(\mathcal{P},\mathcal{H})$ |
|---|---|
| **A-I-ZF**: Large no. of antennas and users with $M > K$. Asymptotic rates of [117] approximate the instantaneous rates for almost all CS and transmission at ideal rates. Zero forcing precoder. | $\log\left(1+\frac{p_{k,j}}{\frac{1}{\beta-1}\sum_{l=1,l\neq j}^{N}\gamma_l\,\frac{\mathrm{tr}(\mathbf{P}_l)}{K}+\sigma_{k,j}^2}\right)$ |
| **C-I-ZF** : Number of antennae/users not large enough. Asymptotic results not accurate. Every BS has CSIT, computes theoretical rates and transmits at ideal rates. ZF precoder. | $\log\left(1+\frac{p_{k,j}^t}{\sum_{l=1,l\neq j}^{N}\gamma_l\,\frac{\mathrm{tr}\left(\mathbf{H}_{j,l}^t Q_l \mathbf{H}_{j,l}^{t\ \mathrm{H}}\right)}{K}+\sigma_{k,j}^2}\right)$ |
| **C-D-ZF:** Similar to C-I-ZF, but TX rate allocation from discrete set $\mathbb{R}$. | $\inf_{r\in\mathbb{R}}\{r\leq R_{k,j}^{\mathbf{C-I-ZF}}(\mathcal{P},\mathcal{H})\}$ |
| **C-I-NO:** Similar to C-I-ZF, but without Precoder. | $\log\left(1+\frac{\mathbf{E}_{\mathbf{s}_j}[\lvert u_{k,j}\rvert^2]}{\sum_l \mathbf{E}_{\mathbf{s}_l}[\lvert i_{k,j,l}\rvert^2]+\sigma_{k,j}^2}\right)$ |
| **C-D-NO:** Similar to C-I-NO, but TX rate allocation from discrete set $\mathbb{R}$. | $\inf_{r\in\mathbb{R}}\{r\leq R_{k,j}^{\mathbf{C-I-NO}}(\mathcal{P},\mathcal{H})\}$ |
| **N-RA-NO:** Rate adaptation w/o CSIT. Uses blind methods to adapt to the correct rate as long as the underlying channel can support the same. No TX precoder. | $R_{k,j}^{\mathbf{C-D-NO}}(\mathcal{P},\mathcal{H})$ |
| **L-RA-ZF:** Rate adaptation with local CSIT. BS has local CS, Uses blind methods to assign rates (as in N-RA-NO) and local CS for precoding. | $R_{k,j}^{\mathbf{C-D-ZF}}(\mathcal{P},\mathcal{H})$ |
| **N-RAE-NO:** Similar to N-RA-NO, but with rate estimation errors, $E_{k,j}(r)$. | $R_{k,j}^{\mathbf{N-RA-NO}}(\mathcal{P},\mathcal{H})-E_{k,j}(R_{k,j}^{\mathbf{N-RA-NO}}(\mathcal{P},\mathcal{H}))$ |
| **L-RAE-ZF:** Similar to L-RA-ZF, but with rate estimation errors, $E_{k,j}(r)$. | $R_{k,j}^{\mathbf{L-RA-ZF}}(\mathcal{P},\mathcal{H})-E_{k,j}(R_{k,j}^{\mathbf{L-RA-ZF}}(\mathcal{P},\mathcal{H}))$ |

**Table 2:** Some example Systems.
Right column gives the rate at which data is transmitted when CS is $\mathcal{H}$ and system uses power profile $\mathcal{P}$

2) The second part (II) represents **the transmission rates used at the system**[3]: a) **I** for ideal systems which can channel code to achieve any feasible rate, b) **D** for the those systems which can only operate at one of the discrete rates in the set $\mathbb{R}=\{r_1,r_2,\cdots r_{N_R}\}$ (arranged in decreasing order), c) **RA** for systems which estimate the current rate (i.e., pick the current maximum possible rate from the set $\mathbb{R}$) using (blind) rate adaptation schemes (eg. [19]) without CSIT, d) **RAE** when there are estimation errors in the rate adaptation algorithm.

3) The third part (III) represents the **precoder**[4]: a) **ZF** for zero forcing precoding, b) **NO** for no channel precoding.

[3]We illustrate these concepts using simple rate allocation schemes. One can extend it to other rate allocations, for eg. schemes that incorporate fairness.

[4]One can also consider other types of precoders (eg. MMSE). Our analysis and proofs hold for these configurations as long as they satisfy the assumptions **A**.1 to 4 (refer Section 8.4).

### 8.3.1 Game theoretic formulation

As the base stations influence each other, the problem can best be captured using a game theoretic formulation. We begin by introducing the components of the game. The calligraphic letters (for example $\mathcal{P}$) represent the ensemble of either vectors, matrices or scalars for all the base stations.

**Power profile,** $\mathcal{P} := \{p_{k,j}\}_{k \leq K,\ j \leq N}$, represents the vector comprising of the powers used at all the base stations and for all the users. Recall, $p_{k,j}$ represents the power used by the BS of cell $j$ for user $k$ in cell $j$.

**Channel State (CS),** $\mathcal{H} := \{\mathbf{H}_{1,1}, \mathbf{H}_{1,2}, \cdots, \mathbf{H}_{N,N}\}$, arranged as a matrix of dimension $KN \times MN$, represents the channel state of the entire system.

**Rate for a given power profile and system,** $R^{sys}_{k,j}(\mathcal{P}, \mathcal{H})$, represents the transmission rates allocated, to the user $k$ by the base station $j$, in system represented by sys (eg. **N-RAE-NO** in Table **2**) when the base stations use powers $\mathcal{P}$ and when the CS is $\mathcal{H}$. These rates are given in the right column of the Table **2** for various example systems, whose detailed descriptions are provided in Appendix A.

**Average Rate for a given system and power profile,** is the rate that is achieved on average when a given system uses the power profile $\mathcal{P}$ : $R^{sys}_{avg,k,j}(\mathcal{P}) = \mathbf{E}_{\mathcal{H}}[R^{sys}_{k,j}(\mathcal{P}, \mathcal{H})]$. Let $R^{sys}_{avg} := \{R^{sys}_{avg,k,j}\}_{k,j}$.

**Power constraint ($\mathcal{P} \leq \bar{\mathcal{P}}$)** We use $\leq$ in a special manner to facilitate defining the power constraint. We say a power profile $\mathcal{P}$ is "less that or equal to" and hence satisfies the constraint defined in terms of another power profile $\bar{\mathcal{P}}$ if the two profiles satisfy the constraints for each base station as: $\sum_k p_{k,j} \leq \sum_k \bar{p}_{k,j}$ for all $j \leq N$.

**System Specific Capacity Region** for any given power profile constraint $\bar{\mathcal{P}}$ and a system, sys, is defined as the collection of all possible tuple of average rates while using powers that satisfy the constraints defined in terms of $\bar{\mathcal{P}}$, i.e.,

$$\mathbb{C}^{sys}(\bar{\mathcal{P}}) := \Big\{ \{R_{k,j}\} \in \mathcal{R}^{NK} : \text{ for all } k, j \\ R_{k,j} = R^{sys}_{avg,k,j}(\mathcal{P}) \text{ for some } \mathcal{P} \text{ with } \mathcal{P} \leq \bar{\mathcal{P}} \Big\}. \tag{8.4}$$

This region is different for different systems. For a system with ideal rates the capacity region coincides with the theoretical one. A system with discrete rates cannot always achieve the maximum possible rate and hence its capacity region shrinks. It further depends upon $\mathbb{R}$, the set of supported rates. If the system has estimation errors, the capacity region shrinks further.

**Utilities and Players :** Each BS $j$ is a player and its strategy is $K$-dimensional power vector, $\mathbf{p}_j := [p_{1,j}, \cdots, p_{K,j}]$. Note that $\mathcal{P} = [\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_N]$. Define the utility of player

$j$ as[5],

$$\begin{aligned} U_j^{sys}(\mathbf{p}_j, \mathcal{P}_{-j}) &:= \sum_k \min\left\{R_{avg,k,j}^{sys}(\mathbf{p}_j, \mathcal{P}_{-j}),\ r_{k,j}\right\} \\ \text{with } \mathcal{P}_{-j} &:= [\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_{j-1}, \mathbf{p}_{j+1}, \cdots, \mathbf{p}_N]. \end{aligned} \tag{8.5}$$

In the above, $\mathcal{P}_{-j}$ is the power vector profile excluding only the powers of BS of cell $j$ and $r_{k,j}$ is the demand of user $k$ of cell $j$. Every system with given power constraint $\bar{\mathcal{P}}$ and demand vectors $\{r_{k,j}\}$ defines an $N$-player non cooperative strategic form game: $([1, 2 \cdots N],\ \{U_j^{sys}\}_{j \le N})$. The Nash equilibrium (NE) of this game is a power profile $\mathcal{P}^*$ that satisfies,

$$\mathbf{p}_j^* \in \arg\max_{\mathcal{P} \le \bar{\mathcal{P}}} U_j^{sys}(\mathbf{p}_j, \mathcal{P}_{-j}^*) \text{ for all } j. \tag{8.6}$$

From the above definitions, it is evident that,

**Lemma 8.3.2.** *For any given system and power constraints $\bar{\mathcal{P}}$, if the vector of the demands $\{r_{k,j}\}_{k,j}$ is in the corresponding capacity region $\mathbb{C}^{sys}(\bar{\mathcal{P}})$, then there exists a $\mathcal{P}^* \le \bar{\mathcal{P}}$, which is a NE satisfying all the demands:*

$$R_{avg,k,j}^{sys}(\mathcal{P}^*) = r_{k,j} \text{ for all } k, j. \quad \diamond$$

Thus, when all the base stations use the NE power profile $\mathcal{P}^*$ of Lemma 8.3.2, all the users in each cell achieve an average rate which equals their demand, i.e., will be able to receive the information at the demand rate on average. The *main aim of this work is to obtain this NE (time) asymptotically (if required in a completely distributed way) for any given system when the demands satisfy Lemma 8.3.2.* This NE depends on the system considered (for example higher amount of power may be required if one uses discrete rates in the place of ideal rates) even if the power constraint and demands are same. *The proposed algorithm is a general iterative algorithm which works irrespective of the system considered,* i.e, the proposed algorithm converges to the system specific NE.

**Remark on hypothesis of Lemma 8.3.2:** It requires that the demands equal one of the average rates of the capacity region. Lemma 8.4.4 of the next section gives an easily verifiable assumption which ensures this hypothesis of Lemma 8.3.2.

**Set of demand meeting NE,** $\mathbb{L}^{sys} \subset \mathbb{C}^{sys}(\bar{\mathcal{P}})$, is the set of NE which meet the demands as in Lemma 8.3.2.

We now present the Universal Power Allocation algorithm for power constrained Multi Cell Networks (UPAMCN).

[5]The utility of an user is the average rate at which its data is transferred. The user $k$ of cell $j$ requires transmission at maximum at its demand rate $r_{k,j}$ and hence his utility is upper bounded by the same.

| | Type | Intf BS | Tx Ant | Users |
|---|---|---|---|---|
| **L1** | Linear | 2 | 16 | 8 |
| **L2** | Linear | 2 | 32 | 8 |
| **H1** | Hexagon | 6 | 16 | 8 |
| **H2** | Hexagon | 6 | 2 | 2 |

**Table 3:** Network configurations

| | |
|---|---|
| **S1** | Asymptotic Ideal with ZF Precoder (**A-I-ZF**) |
| **S2** | Rate adaptation with local CSIT and ZF precoder (**L-RA-ZF**) |
| **S3** | Rate adaptation with local CSIT, ZF Precoder and with estimation errors (**L-RAE-ZF**) |
| **S4** | Rate adaptation without CSIT (**N-RA-NO**) |
| **S5** | Rate adaptation without CSIT and estimation errors (**N-RAE-NO**). |

**Table 4:** System configurations

## 8.4 Universal Algorithm : UPAMCN

We consider a quasi-static channel and obtain the NE of Lemma 8.3.2 asymptotically by iteratively updating the power profile at the beginning of every slot, during which the CS is assumed constant.

**Basic idea**[6] **:** Each BS $j$ in every time slot knows the rates at which data is transmitted to its users, $\{R^{sys}_{k,j}(\mathcal{P},\mathcal{H})\}_k$. The characterization of these rates is provided for some examples in Table **2**. An iterative algorithm can find the average value of it. One can then update the power vectors to force this average towards the demands $\{r_{k,j}\}$.

Let $d^{t+1}_{k,j}$ represent the number of bytes of data transmitted successfully in time slot $t+1$ by the $j^{\text{th}}$ base station to its user $k$ divided by the duration of the time slot. This ratio depends upon the power profile of the entire system in the previous slot ($\mathcal{P}^t$) and the entire CS in the current slot ($\mathcal{H}^{t+1}$), but ($\mathcal{P}^t$, $\mathcal{H}^{t+1}$) are only partially known at the base stations. However $d^{t+1}_{k,j}$ is still known at base station $j$ as it is the source that pumps out the data. Infact, it will be precisely equal to $d^{t+1}_{k,j} = R^{sys}_{k,j}(\mathcal{P}^t,\mathcal{H}^{t+1})$ of Table **2** by definition. Let $\{\mu^t\}$ represent the step sizes.

### 8.4.1 UPAMCN algorithm

With $\Pi_{\mathbb{A}}$ representing the projection in to the set $\mathbb{A}$

$$p^{t+1}_{k,j} = \Pi_{\mathbb{A}_j}\left[p^t_{k,j} - \mu^t\left(d^t_{k,j} - r_{k,j}\right)\right] \text{ with}$$
$$\mathbb{A}_j := \left\{\mathbf{p}\in\mathcal{R}^K : \sum_k p_k \le \bar{P}_j\right\};\ \ \mathbb{A} := \mathbb{A}_1\times\mathbb{A}_2\cdots\times\mathbb{A}_N. \tag{8.7}$$

[6]Most of the cases stochastic approximation algorithms obtain optimum of a function as the zero of its derivative. In contrast, this algorithm obtains the profile that satisfies the demands, as the zero of the function given by *the average rate minus demand.*

### 8.4.2 Analysis

We obtain the asymptotic analysis of the algorithm using the ordinary differential equations (ODE) approach of [93]. We establish Theorem 8.4.2.1 given below, under:

**A**.1 There exists a sequence

$\alpha_t \to \infty$ with $\lim_t \sup_{0 \le i \le \alpha_t} \mu^{t+i}/\mu^t = 0$.

**A**.2 The channel state $\{\mathcal{H}^t\}$ is an independent and identically distributed (IID) sequence with finite mean and variance.

**A**.3 The instantaneous rates are bounded by the same constant, i.e., $|R^{sys}_{k,j}(\mathcal{P},\mathcal{H})| \le B$ for all $k, j, \mathcal{P}$ and $\mathcal{H}$.

**A**.4 The average rate $R^{sys}_{avg,k,j}$ is continuous in $\mathcal{P}$ for all $k, j$.

We will show that the UPAMCN trajectory (8.7) can be approximated by the solution $(\mathcal{P}(t))$ of the following ODE (to be precise a differential inclusion).

$$\dot{p}_{k,j} \quad = \quad r_{k,j} - R_{avg,k,j}(\mathcal{P}) + z_{k,j}(\mathcal{P}) \text{ for all } k, j \tag{8.8}$$

where $z_{k,j}(\mathcal{P})$ represents the projection term. Define the limit set of this ODE :

$$\mathbb{L}^{ODE} := \lim_{t\to\infty} \cup_{\mathcal{P}\in\mathcal{A}} \{\mathcal{P}(s) : s \ge t \text{ and } \mathcal{P}(0) = \mathcal{P}\}.$$

The $\delta$-neighborhood of this set is defined as:

$$\mathbb{B}_\delta(\mathbb{L}^{ODE}) := \left\{\mathcal{P} : |\mathcal{P} - \bar{\mathcal{P}}| \le \delta \text{ for some } \bar{\mathcal{P}} \in \mathbb{L}^{ODE}\right\}.$$

Theorem 8.4.2.1 establishes that the trajectory ultimately spends time in this limit set. We first establish the theorem and later study the systems of previous section using this Limit set.

**Theorem 8.4.2.1.** *Assume* **A**.*1-4. Then for every $\delta > 0$, the fraction of time the tail of the algorithm (for any initial power profile with $\tilde{\mathcal{P}} < \bar{\mathcal{P}}$)*

$$\{\mathcal{P}^\tau\}_{\tau \ge t} \text{ with initialization } \mathcal{P}^t = \tilde{\mathcal{P}}$$

*spends in the $\delta$-neighborhood of the limit set $\mathbb{B}_\delta(\mathbb{L}^{ODE})$ tends to one (in probability) as $t \to \infty$.*

**Proof:** Refer Appendix B.

### 8.4.3 Analysis of the specific systems

Most of the systems considered in this chapter (for example, **C-D-ZF**) transmit at one of rates from a discrete set $\mathbb{R}$ depending on the instantaneous CS and for these one need to explicitly prove the continuity of the average rates. This is achieved in the following (proof in Appendix B).

**Lemma 8.4.4.** *Assume* **A**.*1 and* **A**.*2. Then, for all the systems considered in Table* **2**, *assumptions* **A**.*3 and* **A**.*4 are satisfied, Theorem 8.4.2.1 applies and hence for the UPAMCN trajectory (8.7) asymptotically spends most of its time in the limit set,* $\mathbb{L}^{ODE}$.

*Further, the demand meeting NE set,* $\mathbb{L}^{sys}$, *is non empty and these form the stationary points of the ODE (8.8), whenever for all* $k, j$ *the demands satisfy*

$$r_{k,j} \leq \sup_{\mathcal{P} \leq \bar{\mathcal{P}}} R_{avg,k,j}(\mathcal{P}). \diamond$$

For further analysis, one needs to study the limit set of ODE (8.8). A limit set of a ODE usually contains limit cycles or attractors. The demand meeting NE of $\mathbb{L}^{sys}$ would be in the limit set if further we could show that they are attractors. In that case, the algorithm spends most of its time in these attractors or in other words the *UPAMCN algorithm asymptotically meets the demands of all the users.* Right now, we can only say that, every stationary point of ODE (8.8) is a demand meeting NE and any attractor of the ODE must be a stationary point. We will show via numerical examples in the next section that the algorithm indeed converges to a demand meeting NE for all the systems considered in this work.

### 8.4.5 Extensions to UPAMCN

The UPAMCN algorithm works under the basic assumption that the BS always has sufficient data to transmit. But in reality, data often arrives in real time and hence there can be situations when the BS can transmit at a higher rate but does not have sufficient data. In this case we propose the following extension to UPAMCN:

$$\begin{aligned} b_{k,j}^{t+1} &= b_{k,j}^{t} + B_{k,j}^{t+1} - \min\left\{d_{k,j}^{t}, b_{k,j}^{t}\right\} \text{ and} \\ p_{k,j}^{t+1} &= \Pi_{\mathbb{A}_j}\left[p_{k,j}^{t} + \mu_t\left(\min\{d_{k,j}^{t}, b_{k,j}^{t}\} - r_{k,j}\right)\right] \end{aligned} \tag{8.9}$$

where $b_{k,j}^{t}$ represents the remaining (accumulating) bytes of data to be transmitted by BS $j$ to the user $k$ at the begining of time slot $t$ and $B_{k,j}^{t}$ represents the fresh sample of data added to the corresponding buffer.

## 8.5 Simulation

We consider two types of cellular networks in our simulations (Table **3**). The first one is a Hexagonal network, where users in each cell experience interference from BS transmissions of surrounding cells (typically assumed to be from the $1^{st}$ tier of surrounding 6 cells (see for example figure (8.1)). The second one is a linear network, where users in each cell experience interference from BS transmissions of adjacent cells (typically two adjacent neighbors). The system configurations are summarized in Table **4**. Each BS equipped with $M$ transmit antennas is serving $K$ users in its cell.

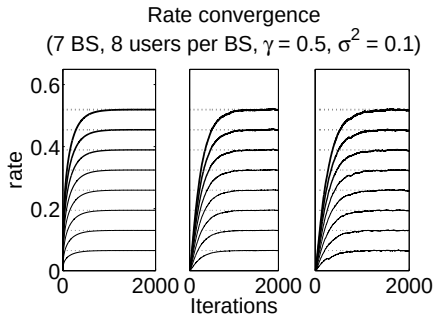


***Figure 8.2:*** *Rate convergence for Systems S1, S2 and S3 (H1 Network)*

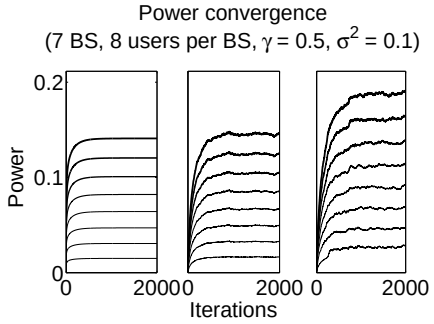


***Figure 8.3:*** *Power convergence for Systems S1, S2 and S3 (H1 Network)*

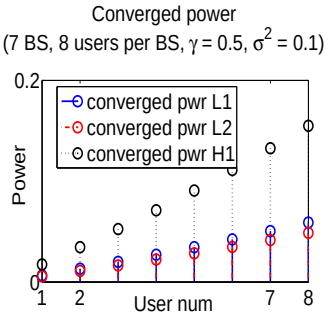


***Figure 8.4:*** *Demand satisfying NE. System S2. (L1, L2 & H1 networks)*

| System | Demand satisfying NE (converged power) |
|---|---|
| ID | .015 .032 .049 .067 .085 .105 .126 .147 |
| D1 | .017 .035 .053 .072 .092 .113 .135 .158 |
| D2 | .031 .049 .071 .097 .121 .146 .174 .203 |

**Table 5:** System S2
ID: Ideal, D1: Discrete - 100 levels, D2: Discrete - 20 levels

In all the simulations we also compute the average rates via the following iteration: $\phi_{k,j}^{t+1} = \phi_{k,j}^{t} + \mu^t \left( d_{k,j}^{t} - \phi_{k,j}^{t} \right)$ for all $k, j$. This iteration is only a measurement procedure that is used for the purpose of calculating average rates of the system for the numerical examples considered. That it represents the average rate can be understood by noticing that $\phi_{k,j}^{t}$ is actually a weighted average of all the instantaneous rates $\{d_{k,j}^{\tau}; \tau \leq t\}$ up to time $t$. These average rates are used to illustrate that systems considered in these examples, asymptotically (as time progresses) satisfy the user's demands on average.

The power limit on each BS is set to 1 unit. Interference factor $\gamma_l$ from each interfering BS, $l$, is set to 0.5. For the simulations considered here, we choose the demand rate vector (to lie within the capacity region and is common for all the base stations) as: $\mathbf{r} = [.065 \ .130 \ .195 \ .260 \ .325 \ .389 \ .454 \ .520]$.

In the first set of simulations, we consider the hexagonal network (H1). The rate and the power convergence behavior of the algorithm for systems S1, S2 and S3 is plotted in figure 8.2 and 8.3, respectively. We observe that: **(1)** The algorithm converges to the demand meeting NE: we see in Figure 8.2 that for all the systems, the average rate achieved asymptotically converges towards the demand rates. **(2)** As discussed in the previous sections, we notice from Figure 8.3, that the converged power profile (demand

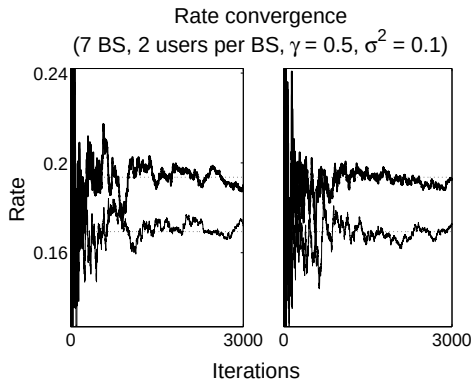


***Figure 8.5:*** *Rate convergence for System S4 and S5 (H2 Network)*

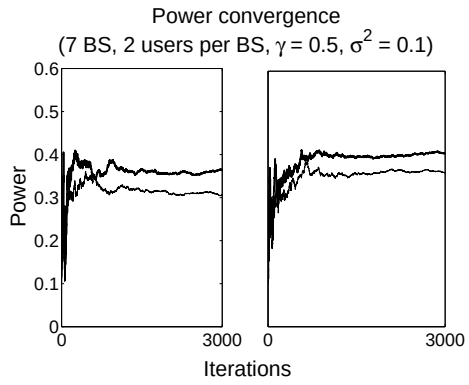


***Figure 8.6:*** *Power convergence for Systems S4 and S5 (H2 Network)*

meeting NE) is system specific. S3 is a system with errors, the proposed algorithm still satisfies the demands asymptotically, however, the converged power profile has higher power levels in comparison with the error free systems S2 and S1. **(3)** Note that S2 can also represents **C-D-ZF**, a complete CSIT system (see details on Table **2** and Appendix A). From figure 8.3, we observe that the converged power profile of **C-D-ZF** (S2) is close to that of **A-I-ZF** (S1) system. Thus the demand meeting power profile of systems with large number of transmit antennas and or users and large number of discrete levels in $\mathbb{R}$ is close to that of the asymptotic ideal rate system. Further convergence is faster with S1 system. Thus for such systems, UPAMCN algorithm can be used to estimate (approximately) the demand meeting power profile, much faster, using the asymptotic rate expressions in place of instantaneous transmit rates allocated, $\{d^t_{k,j}\}$. Note that this further avoids the need of complete CSIT, as we need only local CSIT for precoding. **(4)** As the discrete levels increase, the power profile decreases and finally converges to that of the ideal rate. This is tabulated in Table **5**.

In the second set of simulations, for given demand rates, we compare the algorithm behavior for different network configurations L1, L2 and H1 with system S2. We observe that: to satisfy the same demands, the base stations in L2 expend the least power, followed by L1 and then H1. L2 performs better than L1 due to improved transmit diversity. H1 is the worst (larger number of interfering base stations).

In the final set of simulations, for network configuration H2, we consider the least informed (No CSIT) systems, the rate adaptation systems S4 and S5. We choose the common demand rate vector as $[0.1694 \;\; 0.1936]$. Figures 8.5and 8.6 illustrate the average rate and power profile convergence. As CSIT (even local) is not available at the base stations, they cannot use any precoders. Thus, it is a totally interference dominated system and hence the achievable capacity region is small. But the UPAMCN algorithm works even for this least informed system: it asymptotically satisfies the demand rates, albeit using a higher power profile. Further, the rate estimation errors in system S5 demand higher power levels in comparison with the error free system S4 to achieve the same demands.

Further, we observe that the convergence to the demand meeting NE is quicker in those systems where base stations have more information (see for eg. figure 8.3).

## 8.6 Conclusions

Mobile broadband users demand certain rates depending on the end application and QoS requirements. The base station serving these users has to allocate power to satisfy user demands operating within its own total power budget. Intra-cell and inter-cell interference diminish the available rates in multicell networks. Neighboring base stations can co-operate to exchange some form of channel state information depending on backhaul capacity and processing power to alleviate interference and thus enhance achievable rates. Further, system specific components like modulation, coding, rate allocation, channel estimation and synchronization impacts the achievable rates and hence the power allocation. In our work, we propose an universal power allocation algorithm which works in this setting. The stochastic approximation based universal power allocation algorithm runs at each BS, independently and simultaneously to meet the user demands as long as the demands are achievable. The power allocation is formulated as a game problem. A system specific capacity region is defined and the proposed algorithm is analyzed with an ODE framework. The proposed algorithm works well in a multitude of system configurations as demonstrated via simulations and analysis.

Our algorithm assumes that the serving BS always has sufficient amount of data to transmit. However, in many applications, the data is available in real time. We mentioned a possible extension of the same in the work.

## 8.7 Appendix A: Example Systems

1) *Asymptotic Ideal Rate system :* In a multicellular system with large number of antennas at the BS and large number of users, the rate for a given CS can be obtained using random matrix theory. For example, in [117] the asymptotic rates are derived for a zero forcing (ZF) precoder. It is shown that for almost all realizations of CS, the rate can be approximated by the expression given below in equation (8.10). Further, we consider a system in which, the base stations use channel coding schemes to transmit very close to the theoretical rates. When this system (which we call as *asym-ideal-zeroforcing* or in short **A-I-ZF** according to our notations) uses power profile $\mathcal{P}$ and when the channel state (CS) is $\mathcal{H}$, the BS $j$ transmits to the user $k$ at rate ([117]) (when $M > K$):

$$R_{k,j}^{\mathbf{A-I-ZF}}(\mathcal{P},\mathcal{H}) \approx \log\left(1+\frac{p_{k,j}}{\frac{1}{\beta-1}\sum_{l=1,l\neq j}^{N}\gamma_l\,\frac{\mathrm{tr}(\mathbf{P}_l)}{K}+\sigma_{k,j}^2}\right) \tag{8.10}$$

where, $\beta = M/K$ is the ratio of number of transmit antennas on the BS to the number of users and $\gamma_l \in (0,1)$ represents the interference from cell $l$. This rate is same for

almost all CS $\mathcal{H}$ as it is an asymptotic rate. Similar expression is available for the case with $M = K$ in [117].

2) *Ideal rates using complete CSIT :* If the number of antennae/number of users is not large enough, the asymptotic results are not accurate. If BS has access to CSIT (and if each BS could channel code to obtain rates closer to the ideal rate) then with ZF precoder it transmits at rate:

$$R_{k,j}^{\mathbf{C-I-ZF}}(\mathcal{P},\mathcal{H}) = \log\left(1+\frac{p_{k,j}^t}{\sum_{l=1,l\neq j}^{N}\gamma_l\,\frac{\mathrm{tr}\left(\mathbf{H}_{j,l}^t Q_l \mathbf{H}_{j,l}^{t\,\mathsf{H}}\right)}{K}+\sigma_{k,j}^2}\right)$$
$$Q_l := \mathbf{H}_l^{t\mathsf{H}}(\mathbf{H}_l^t\mathbf{H}_l^{t\mathsf{H}})^{-1}\mathbf{P}_l^t(\mathbf{H}_l^t\mathbf{H}_l^{t\mathsf{H}})^{-1}\mathbf{H}_l^t$$

For the same configuration, but without transmitter precoding, the the instantaneous transmission rate (as obtained using Shannon's capacity expression), from equation (8.2) is:

$$R_{k,j}^{\mathbf{C-I-NO}}(\mathcal{P},\mathcal{H}) = \log(1+\eta_{k,j}) \text{ where} \tag{8.11}$$
$$\text{SINR, } \eta_{k,j} := \frac{\mathbf{E}[|u_{k,j}|^2]}{\rho_{k,j}} \text{ with noise + interference, } \rho_{k,j} := \sum_l \mathbf{E}[|i_{k,j,l}|^2]+\sigma_{k,j}^2.$$

3) *Finite number of Rates :* Ideal rate systems are not realistic, they can't be implemented in practice. We consider a system, in which the BS can transmit at one of the available discrete rates from the set $\mathbb{R}$. When transmitter has CSIT, it knows the exact theoretical rate and hence will pick the largest rate from set $\mathbb{R}$ that is smaller than the current theoretical rate:

$$R_{k,j}^{\mathbf{C-D-ZF}}(\mathcal{P},\mathcal{H}) = \inf_{r\in\mathbb{R}}\{r \leq R_{k,j}^{\mathbf{C-I-ZF}}(\mathcal{P},\mathcal{H})\}, \tag{8.12}$$
$$R_{k,j}^{\mathbf{C-D-NO}}(\mathcal{P},\mathcal{H}) = \inf_{r\in\mathbb{R}}\{r \leq R_{k,j}^{\mathbf{C-I-NO}}(\mathcal{P},\mathcal{H})\}. \tag{8.13}$$

4) *Rate adaptation Without CSIT :* It is once again not realistic to assume the knowledge of complete CSIT. There are many schemes that estimate the rate blindly or using some partial CSIT (eg. [19]). The UPAMCN algorithm is a very general algorithm and works with all those systems which satisfy assumptions **A.**1-4. These are quite simple assumptions and most of the systems can satisfy these and hence the algorithm works for majority of the blind/partial CSIT systems.

We explain one such blind system wherein, the BS estimates the transmission rates without knowledge of CSIT. Each time, the BS begins by attempting at the highest available rate $r_1$. If the data is not received correctly (information obtained via a feedback from the receiver), the BS sends some more information about the same data packet so that the overall rate now is the second highest $r_2$. This procedure repeats until the two agree upon the correct rate. We assume that this rate adaptation system is always successful, i.e, it can estimate the actual rates without errors. Such a system does not require CSIT, however the final rate at which the transmission takes place depends upon the current channel state in exactly the same way as in the case of **C-D** (or **A-D** for large antenna and user case and note there is no channel coding in this case as there is no CSIT) and hence,

$$R_{k,j}^{\mathbf{N-RA-NO}}(\mathcal{P},\mathcal{H}) = R_{k,j}^{\mathbf{C-D-NO}}(\mathcal{P},\mathcal{H}) \tag{8.14}$$

5) *Rate Adaptation with local CSIT:* All the base stations have local CSIT, i.e., BS $j$ knows the $\mathbf{H}_{j,j}$ part of CS. However they can't estimate the current rates just based on local CSIT. So, they once again use rate adaptation technique as in the system (4). They can however design a better system by using for example a zero forcing precoder. In this case, as in system (4) the rate will be adapted to the actual underlying rate and hence will be same as that in **C-D-ZF**:

$$R^{\mathbf{L-RA-ZF}}(\mathcal{P},\mathcal{H}) = R^{\mathbf{C-D-ZF}}_{k,j}(\mathcal{P},\mathcal{H}) \tag{8.15}$$

6) *Rate Adaptation with errors :* There can be some errors in rate adaptation algorithm of system (4) or (5). In this case

$$\begin{aligned} &R^{\mathbf{N-RAE-NO}}_{k,j}(\mathcal{P},\mathcal{H}) \\ &= R^{\mathbf{N-RA-NO}}_{k,j}(\mathcal{P},\mathcal{H}) - E_{k,j}(R^{\mathbf{N-RA-NO}}_{k,j}(\mathcal{P},\mathcal{H})) \qquad (8.16)\\ &R^{\mathbf{L-RAE-ZF}}_{k,j}(\mathcal{P},\mathcal{H}) \\ &= R^{\mathbf{L-RA-ZF}}_{k,j}(\mathcal{P},\mathcal{H}) - E_{k,j}(R^{\mathbf{L-RA-ZF}}_{k,j}(\mathcal{P},\mathcal{H})) \qquad (8.17) \end{aligned}$$

where (assuming independent errors) $E_{k,j}(\bar{r})$ can take values in the subset $\mathbb{R} \cap \{r \leq \bar{r}\}$ with a given probability distribution.

## 8.8 Appendix B: Proofs

**Proof of Theorem 1:** As a first step, we rewrite the algorithm as in [93]:

$$Y^t_{k,j} := r_{k,j} - d^t_{k,j}, \quad p^{t+1}_{k,j} = \Pi_{\mathbb{A}_j}\left[p^t_{k,j} + \mu^t Y^t_{k,j}\right].$$

Define, $\mathcal{F}_t := \sigma\left(\mathcal{P}^\tau, \{Y^{\tau-1}_{k,j}\}_{k,j}, \text{ for all } \tau \leq t\right)$ and let $\mathbb{E}_t$ represent the expectation w.r.t. $\mathcal{F}_t$, the filtration. Under the assumptions **A**.2 and **A**.3 clearly the condition expectation

$$\mathbb{E}_t[Y^t_{k,j}] = g^p_{k,j}(\mathcal{P}^t) := r_{k,j} - R_{avg,k,j}(\mathcal{P}^t) \text{for all } k, j \text{ and } t.$$

For every $j$, the constrain set $\mathbb{A}_j$ satisfies the assumption (A3.2), page 107 of [93]. By assumption **A**.3 $\{Y^t_{k,j}; t\}$ is uniformly integrable and hence satisfies assumption A.2.1, pp. 258 [93]. They also satisfy the assumption A.2.3 to A.2.7 of pages 258, 259 [93] with $g_t = \bar{g} = g^p$ and with $\beta_t = 0$ $\xi_t = 0$ for all time $t$. Assumption A.2.2, pp 258, [93] is satisfied because of our Assumption **A**.4. Let $\mathbf{z}_j$ represent the projection or constraint term, the minimum force needed to keep the vector $\mathbf{p}_j$ in $\mathbb{A}_j$. Then by Theorem 2.3, pp. 259, [93] the UPAMCN algorithm trajectory $\mathcal{P}^t$ converges weekly to the trajectory of the solution of the ODE (8.8) (in the sense as explained in [93]). Further by the same theorem of [93], for any $\delta > 0$, the fraction of time that the tail sequence $\{\mathcal{P}^\tau\}_{\tau \geq t}$, with initializations $p^t_{k,j} = \tilde{p}_{k,j}$ for every $(k,j)$, spends in the $\delta$-neighborhood of the limit of set of the above ODE (8.8), $\mathbb{B}_\delta(\mathbb{L}^{ODE})$, goes to one (in probability) as $t \to \infty$. $\diamond$

**Proof of Lemma 8.4.4:** The boundedness assumption **A**.3 is direct for discrete rate systems and is also true for ideal rate systems as seen from the formulas. The ideal rates are point wise continuous and are bounded and hence by bounded convergence theorem satisfy the continuous assumption **A**.4. The same for the discrete rates is given by Lemma 8.8.1.

The continuity assumption **A**.4 now also holds for the rate adaptation system with errors, **L-RAE-NO**, whenever the statistics of the errors $\{E_{k,j}\}$ are independent of the power profile or when they are continuous in $\mathcal{P}$. *Thus for all the systems considered in this work Theorem 8.4.2.1 applies.*

**Conditions for existence of demand meeting NE :** For all the systems considered so far, the hypothesis of Lemma 8.3.2 is satisfied, i.e., $\mathbb{L}^{NE}$ is non empty whenever the power constraints are sufficient to cater to the demand rates. This fact is established by the continuity of the average rates w.r.t. the power profile, i.e., the establishment of the assumption **A**.4. *To be precise Lemma 8.3.2 is satisfied, i.e., $\mathbb{L}^{NE}$ is non empty whenever for all $k, j$ $r_{k,j} \leq \sup_{\mathcal{P} \leq \bar{\mathcal{P}}} R_{avg,k,j}(\mathcal{P})$.* ◇

**Lemma 8.8.1.** *The average rates $R^{sys}_{avg,k,j}$ for systems* **C-D-ZF** *and* **C-D-NO** *are continuous w.r.t. power profile $\mathcal{P}$ for all $k, j$.*

**Proof :** Let $R_{k,j}(\mathcal{P}, \mathcal{H})$ represent the corresponding ideal rate (the rate before discretization) for the given CS $\mathcal{H}$. From all the rate formulas in this work, we can see that these rates bounded and are continuous in $\mathcal{P}$, for all $\mathcal{H}$. For the discretized systems, the average rates can be written as,

$$\begin{aligned} R_{avg,k,j}(\mathcal{P}) &= \sum_{i \leq N} q(i,\mathcal{P}) r_i \text{ where} \\ q(i,\mathcal{P}) &:= \int 1_{\{r_{i-1} \leq R_{k,j}(\mathcal{P},\mathcal{H}) \leq r_i\}} d\Gamma(\mathcal{H}) \end{aligned} \tag{8.18}$$

with $d\Gamma$ representing the Gaussian measure. For a given $\mathcal{P}$, the probability of the sets of the type (the boundaries of the sets used while defining the indicators in (8.18))

$$\Gamma\left(\left\{\mathcal{H} : R_{k,j}(\mathcal{P},\mathcal{H}) = r_i\right\}\right) = 0,$$

because of the continuity of the Gaussian measure. Hence, the point wise functions of integral (8.18) are continuous w.r.t. to $\mathcal{P}$ for almost all $\mathcal{H}$. Thus the lemma follows by bounded convergence theorem. ◇

## 8.9 Publications

1. Sreenath Ramanath, Veeraruna Kavitha, Merouane Debbah, "Satisfying Demands in a Multicellular Network: An Universal Power Allocation Algorithm", proceedings of WiOpt 2011, May 9-13, Princeton, USA.

# Part IV

# Epilogue

# Chapter 9

# Conclusions

## Summary and general discussion

Small cell networks are a promising alternative to increasing data demands arising due to new applications and services in mobile broadband. Optimal design of small cells is a challenging task. In this thesis we addressed two important aspects of small cells: cell dimensioning and resource allocation.

We began our discussion with dimensioning of static and moving users. For the case of static users, in chapter 2, we used fluid limits to derive explicit expressions for own cell and interference power and computed throughput achievable as a function of cell size. Our analysis considered frequency reuse, path-loss, different receiver structures, indoor-outdoor partitions, etc. Our results showed that optimal cell sizes exist in some configurations, while in the others, the way to maximize achievable throughput was to increase the base station density. A scope for future directions is to consider more sophisticated radio propagation models, fading and shadowing effects, etc. Further, fluid models were used to get first-cut results for cell dimensioning, but one can consider a more closer to realistic situations, like the users are distributed according to a stochastic process and thus meaningful results can be obtained using tools like stochastic geometry [17].

For the case of mobile users, in chapter 3, we used queuing theoretic tools to derive optimal cell sizes which optimize important system metrics like expected waiting/service times, call drop/block probabilities, etc. Our analysis was predominantly for users traversing in single dimensions and further once a velocity is chosen by a certain user, it is assumed that the user continues with the same velocity until his service is completed or until the time he is blocked or dropped. Many interesting possibilities exist as scope towards future research. Some of these have been mentioned in our work [89]. Extensions would include analysis in 2-dimensions and further assuming that users travel with a velocity with a certain mean and variance. We have carried out partial analysis of some of these in our work [118]. Further, it would be interesting to consider more sophisticated and realistic user movements which can be modeled by

random-walk, random way point or Brownian motion.

Similar to our observations in the previous discussion on chapter 2 and chapter 3, for the location of base station(s), it would be interesting to study with discrete/random distribution of users, instead of the continuum approach and extend it to the case of 2-dimensions.

In chapter 5, we derived the per antenna user density via asymptotic analysis of an interference limited precoded small cell network. Our analysis considered the simplest case of zero force precoders (ZF). Zero force precoders, can churn up unreliable results when there are channel realizations with deep fade, due to the matrix inversion in computing an estimate of the channel matrix. A more reliable precoder is the Minimum Mean Square Error (MMSE) precoder, which has an additional scaling term to take care of such situations. A single cell analysis for such cases has already been explored in [48]. It would be interesting to extend this work for the multicell case. Further, in our analysis, many aspects like the transmit correlation, path-loss, etc, have been avoided to keep the analysis simple and tractable. A more eloborate model to take these things into account would throw up some interesting observations. Further, we used Wyner-type cellular layouts, where we assume that users are experiencing same interference on an average. It would be worthwhile to see for possible approaches that can relax these assumptions.

For the case of energy conservation towards green cellular initiative, our work considered simple models to derive open loop control of base station activation in chapter 6. In the case of central control, we made an assumption that each BS has a unidirectional antenna. This assumption simplified the problem and gave rise to a very nice analytic form of the control policy in bracket sequences. It would be worthwhile to explore if the assumptions to be satisfied for the cost functions to be multimodular still hold with omni-direction antenna. Further, one can explore closed loop control policies by including state information.

In chapter 7, we defined T-scale and multiscale fairness. We believe these new fairness concepts allows one to share resources fairly, depending on the traffic type. We demonstrated some applications, which included spectrum sharing and resource allocation in indoor-outdoor femtocells. These concepts can be extended to frequency or time-frequency systems (OFDMA). Further our analysis considered utilities are linear in resources. It would be interesting to deal with the case when this is not true. For example, if the resource to be shared is not the throughput, but the power. We also discussed an application of this new concept to an indoor-outdoor femtocell [114]. We can apply these new concepts to other variants of cell partitions.

We propose stochastic approximation based power allocation algorithm in chapter 8. These algorithms are shown to satisfy user demand rates assymptotically, while operating within power constraints at each base station. We could obtain partial analysis (via ordinary differential equation (ODE) approximation approach) and relied more on simulations to show that the algorithm indeed works. It would be interesting to attempt and show that the converged power profiles are indeed stationary points or elements in the limit set of the approximating ODE, which completes the analysis. Further, the ex-

tension to multi-tier and heterogeneous networks, considering this algorithm is again demonstrated via simulations. It would be challenging to analyze this problem via some sort of two-time scale stochastic approximation based analysis (e.g., see [32]).

Finally, some of the ideas developed during the process of study should find use in practical implementation of small cell networks. It would be of great value to demonstrate few key ideas used in cell dimensioning and resource allocation on a more comprehensive evaluation platform, like an LTE simulator [161].

# Appendix

## Recent Innovations and Advances in Wireless Communications

Inspired by a review comment, we conducted a survey inviting leading experts in the research community to share their views on the recent innovations and advances in Wireless Communications. The questions we posed were:

1. What according to you are the top 5 innovations or advances in Information theory / Signal processing in recent years that has pushed up the spectral efficiency?
2. What according to you is the next 'big bang' in Information theory which can push it further?

In the table (following page), we collate the responses.

Perspectives: We got some very interesting view points from the Survey. Some said the survey itself was very thought provoking and needs careful thinking. Others felt that the guessing game is difficult, citing examples of LDPC. Few others opined that it is all related to simplicity of the idea, ease of implementation, standardization and economics of deployment. Citing from law of large numbers (of opinions), the top five without any particular order were Turbo/LDPC codes, OFDMA, MIMO, Opportunistic communication and Multicell co-operative networks. The answer for the next big thing: Network information theory.

We wish to thank the experts for their opinions and participating in the Survey.

List of Participants: Shlomo Shamai (Technion ISREAL), David Gesbert (Eurecom FRANCE), Arogyaswamy Paulraj (Stanford University USA), Thomas Bonald (Telecom ParisTech FRANCE), Bert Hochwald (University of Notre Dame USA), Piyush Gupta (Alcatel Lucent Bell Labs USA), Merouane Debbah (Supelec, FRANCE), Rajesh Sundaresan (IISc INDIA), Mung Chiang (Princetion University USA), Chandra Murthy (IISc INDIA), Vincent Poor (Princeton University USA), Rahul Vaze (TIFR INDIA), Tijani Chahed (Telecom SudParis FRANCE), Pramod Vishwanath (UIUC USA), Angel Lozano (UPF SPAIN), Eitan Altman (INRIA FRANCE), Jeffery Andrews (UT Austin USA), Gerhard Kramer (TUMunich GERMANY), Nilesh Mehta (IISc INDIA), Sivakumaran Kalyanaraman (IBM Research INDIA), Sergio Verdu (Princeton University USA), Andrea Goldsmith (Stanford University USA), Emre Teletar (EPFL Switzerland), Robert Heath (UT Austin USA)

| | Researcher | What according to you are the top innovations or advances in Information theory / Signal processing in recent years that has pushed up the spectral efficiency? | What according to you is the next 'big bang' in Information theory which can push it further? |
|---|---|---|---|
| 1 | David Gesbert<br><br>Eurecom<br>Sophia Antipolis, France | • MIMO<br>• Relays<br>• Adaptive coding/modulation<br>• Multi-user diversity scheduling<br>• OFDMA | I wish I knew.<br><br>Cooperative MIMO is already known. See<br><br>http://www.eurecom.fr/~gesbert/papers/multicellMIMO_jsactutorial.pdf |
| 2 | Piyush Gupta<br><br>Alcatel Lucent Bell Labs<br>NJ, USA | You are asking tough questions :-) and I'll only attempt to provide partial and quick answers:<br>• MIMO<br>• LDPC,<br>• Cognitive radio<br>• Interference alignment | Network information theory |
| 3 | Merouane Debbah<br><br>SUPELEC<br>Paris, France | • Network MIMO<br>• Relaying<br>• cooperative communications<br><br>http://ieeexplore.ieee.org/Xplore/login.jsp?url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel4%2F18%2F15554%2F00720543.pdf%3Farnumber%3D720543&authDecision=-203 | Network information theory<br><br>Green information theory related to physics concepts...<br><br>Also see:<br>http://www.flexible-radio.com/publication/theoretical-foundations-mobile-flexible-networks |
| 4 | Rajesh Sundaresan<br><br>IISc<br>Bangalore, India | • Adaptive modulation and coding<br>• Turbo codes and LDPC codes with iterative decoding<br>• Opportunistic scheduling for delay tolerant traffic<br>• Signal processing for MIMO, particularly multiple receivers<br>• Incremental redundancy and hybrid ARQ<br>• interference planning across base stations<br><br>See:<br>P.Subrahmanya, R.Sundaresan, and D.Shiu, An overview of high-speed packet data transport in CDMA systems, IETE Technical Review, vol. 21, no. 5, September-October 2004 (invited).<br><br>Of course there have been concomitant | An understanding of optimal layering architectures for distributed decision making in multiterminal communication systems |

| | | | |
|---|---|---|---|
| | | advances in devices and VLSI that were equally important to support these complex algorithms. | |
| 5 | Eitan Altman<br><br>INRIA<br>Sophia Antipolis, France | • Tse's work on relay channel<br>• Capacity and Couverage of networks whose BS and/or mobiles are located according to a Poisson process on the plane<br>• non-cooperative information theory using non-cooperative games | |
| 6 | Shlomo Shamai<br><br>Technion<br>Haifa, Isreal | • (Turbo-LDPC) were a serious factor in enhancing communications capabilities on one hand.<br>• On the other hand, information theoretic results and paradigms that emerge from the theoretical considerations of multi-terminal information theory (network coding included!), have not only revolutionized over view of what is possible (see for example 5th generation of cellular systems), but also will, I believe, have a serious future impact. | As for the future, with all respect, I am reluctant even to address these questions.<br><br>Note that neither Shannon and nor the greatest figures in information theory, could predict that Information Theory will turn to be so useful.<br><br>Note further that all predictions of the greatest experts in the sixties (and that is after the introduction of LDPC codes in Gallager's thesis) that coding theory is good just as a theory, and at best will be used in outer space communications, did not exactly come true. |
| 7 | Jeff Andrews<br><br>UT Austin<br>USA | Thanks for the nice email but these are extremely open-ended questions.<br>To be honest I don't think the spectral efficiency has increased much recently, and I don't think it will in the future either. I believe most gains will come from small cells.<br><br>See the "Cellular 1000x" talk on my homepage for my vision. | |
| 8 | Angel Lozano<br><br>UPF<br>Barcelona, Spain | Is the PHY layer dead<br>http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5741160 | |
| 9 | Thomas Bonald<br><br>Telecom ParisTech<br>France | • Turbo/LDPC codes<br>• Opportunistic scheduling<br>• Adapative modulation & coding<br>• MIMO<br>• OFDMA | |
| 10 | Pramod Viswanath | There are no clear cut answers to your questions. | |

| | | |
|---|---|---|
| | UIUC<br>Urbana, IL, USA | Wireless is too complicated (intersection of "rocket-science" technology, economics, circuits/memory technology, public policy (spectrum allocation)) and information theory too specific to link the two in any conclusive manner.<br>I would answer differently to the questions based on what aspect I think the listener is looking for |
| 11 | Rahul Vaze<br><br>TIFR<br>Mumbai, INDIA | Hmm You got me here.<br>1) Top 5. a) Relaying Strategies for network coverage b) Cooperative MIMO c) Interference Alignment (still to be used in practice) d) There is something called DIDO (refer Rearden Inc. white paper, though I don't believe it) e) MediaFlow (Broadcast channel stuff)<br>2) Next Bang? We are still waiting for it :) I can't really pinpoint, but Network Information Theory is still not developed enough. |
| 12 | Bert Hochwald<br><br>University of Notre Dame, USA | I am not sure ppl would agree on the same subjects, but significant theoretical and practical improvements in point-to-point spectral efficiencies are due to MIMO.<br>There are also recent improvements in coding such as due to turbo codes and LDPC codes, but the gains relative to Convolutional codes are 1-2 dB, as opposed to the MIMO gains that promise more.<br><br>On the network side, the adoption of OFDMA as the multi-access protocol (over CDMA and TDMA) is significant, such as being used in WiMax and LTE.<br><br>TDD multiplexing is getting more attention because of the way spectrum is allocated, and the gains from channel knowledge that comes along with it. |
| 13 | Arogyaswami. Paulraj<br><br>Stanford University, USA | http://www.ew2011.org/index.php?id=62 |
| 14 | Vincent Poor<br><br>Princeton University, USA | They would take a good bit of thought, actually … I don't think I can rattle of five things, but I can certainly give you one: namely the recognition that fading can be exploiting rather than simply mitigated. This is behind MIMO, physical layer security, etc. |
| 15 | Mung Chiang<br><br>Princeton University, USA | I haven't been working on information theory or signal processing, so hard for me to tell. As to networking, see www.network20q.com |
| 16 | Chandramurthy<br><br>IISc<br>Bangalore, INDIA | These are interesting questions! I will need to think about these - I don't know the answers off the top of my head! |
| 17 | Gerhard Kramer<br><br>TU Munich, GERMANY | I think turbo codes was one innovation, I have no idea what's next. |

# Index

# Publications

## Journals

1. Veeraruna Kavitha, Sreenath Ramanath and Eitan Altman, "Spatial queuing analysis for design and dimensioning of Picocell networks with mobile users", Elseviers Performance Evaluation, Volume 68, Issue 8.

## Conferences and workshops

1. Sreenath Ramanath, Veeraruna Kavitha, Eitan Altman, "Open Loop Optimal Control of Base Station Activation for Green Networks", proceedings of WiOpt 2011, May 9-13, Princeton, USA.
2. Sreenath Ramanath, Veeraruna Kavitha, Merouane Debbah, "Satisfying Demands in a Multicellular Network: An Universal Power Allocation Algorithm", proceedings of WiOpt 2011, May 9-13, Princeton, USA.
3. Sreenath Ramanath, Eitan Altman, Konstantin Avrachenkov, "A heterogeneous approach to fair resource allocation and its application in femtocell networks", proceedings of the Workshop on Indoor and Outdoor Femto Cells (IOFC'11), May 13, Princeton USA.
4. Eitan Altman, Konstantin Avrachenkov, Sreenath Ramanath, "Multiscale Fairness and its Application to Resource Allocation in Wireless Networks", proceedings of IFIP Networking 2011, LNCS 6631, May 9-13, Valencia, Spain.
5. Eitan Altman, Konstantin Avrachenkov, Sreenath Ramanath, "Multiscale Fairness and its Networking Application", Poster, Performance 2010, Nov 17-19, Namur, Belgium.
6. Sreenath Ramanath, Veeraruna Kavitha, Eitan Altman, "Impact of mobility on call block, call drops and optimal cell size in small cell networks", proceedings of the Workshop on Indoor and Outdoor Femto Cells (IOFC'10), Sep 26, Istanbul, Turkey.
7. Sreenath Ramanath, Veeraruna Kavitha, Eitan Altman, "Spatial queuing analysis for mobility in pico cell networks", proceedings of WiOpt 2010, May 31-Jun 04, Avignon, France.

8. Sreenath Ramanath, Merouane Debbah, Eitan Altman, Vinod Kumar, "Asymptotic analysis of precoded small cell networks", proceedings of InfoCom 2010, Mar 15-19, San Diego, USA.
9. Sreenath Ramanath, Eitan Altman, Vinod Kumar, Merouane Debbah, "Optimizing cell size in pico-cell networks", Invited paper, Workshop on Resource Allocation in Wireless Networks (RAWNET'09), Jun 27, Seoul, South Korea.
10. Sreenath Ramanath, Eitan Altman, Vinod Kumar, Veeraruna Kavitha, Laurent Thomas, "Fair assignment of base stations in cellular networks", proceedings of the 22nd World Wireless Research Forum (WWRF'09), May 5-7, Paris, France.

### TECHNICAL REPORTS

1. Eitan Altman, Konstantin Avrachenkov and Sreenath Ramanath, "Multiscale Fairness and its Application to Dynamic Resource Allocation in Wireless Networks", INRIA research report number RR-7382 available at http://hal.inria.fr/inria-00515430/en/

# Bibliography


[1] Robert. A, Dinesh. T, and Xinrong. Li. Indoor propogation modeling at 2.4 ghz for 1eee 802.11 networks. In *Proceedings of the International conference on wireless and optical communication*, Jul 2006.

[2] N. Abramson. The aloha system - another alternative for computer communications. In *Proceedings of the Fall Joint Computer Conference*, 1970.

[3] S. M. Alamouti. A simple transmit diversity technique for wireless communications. *IEEE JSAC*, 16:1451–1458, 1998.

[4] Alcatel-Lucent. Beyond the base station router. technical note.

[5] E. Altman, K. Avrachenkov, and A. Garnaev. Generalized alpha-fair resource allocation in wireless networks. In *Proceedings of CDC*, Cancun, Mexico, 2008.

[6] E. Altman, K. Avrachenkov, and A. Garnaev. Alpha-fair resource allocation under incomplete information and presence of a jammer. In *Proceedings of NETCOOP, LNCS 5894*, 2009.

[7] E. Altman, K. Avrachenkov, and B. J. Prabhu. Fairness in mimd congestion control algorithms. In *Proceedings of the IEEE INFOCOM*, Miami, USA, 2005.

[8] E. Altman, K. Avrachenkov, and S. Ramanath. Multiscale fairness and its application to dynamic resource allocation in wireless networks, September 2010.

[9] E. Altman, K. Avrachenkov, and S. Ramanath. Multiscale fairness and its networking application. In *Poster session, IFIP Performance*, Namur, Belgium, 2010.

[10] E. Altman, K. Avrachenkov, and S. Ramanath. Multiscale fairness and its application to resource allocation in wireless networks. In *Proceedings of IFIP Networking (LNCS 6631)*, Valencia, Spain, 2011.

[11] E. Altman, T. Boulogne, R. El Azouzi, T. Jiménez, and L. Wynter. A survey on networking games in telecommunications. *Computers and Operations Research*, 33(2):286–311, 2006.

[12] E. Altman, B. Goujal, and A. Hordijk. *Discrete-Event Control of Stochastic Networks: Multimodularity and Regularity*. Springer, Lecture Notes in Mathematics, No. 1829, 2001.

[13] Eitan Altman, Anurag Kumar, Chandramani Singh, and Rajesh Sundaresan. Spatial sinr games combining base station placement and mobile association. In *Proceedings of INFOCOM*, 2009.

[14] J.B Andersen, T.S Rappaport, and S Yoshida. Propagation measurements and models for wireless communications channels. *IEEE Communications Magazine*, 33:42–49, Aug 1996.

[15] J. Andrews. Interference cancellation for cellular systems: A contemporary overview. *IEEE Wireless Communications Magazine*, 12, Apr 2005.

[16] E Arikan. Channel polarization: A method for constructing capacity-achieving codes. In *Proceedings of IEEE ISIT*, pages 1173–1177, 2008.

[17] F. Baccelli and B. Blaszczyszyn. *Stochastic Geometry and Wireless Networks: Vol I, II*. Now, 2009.

[18] F. Baccelli, B. Blaszczyszyn, and M. Karray. A spatial markov queuing process and its applications to wireless loss systems, 2007.

[19] K. Balachandran, S. R. Kadaba, and S. Nanda. Channel quality estimation and rate adaptation for cellular mobile radio. *IEEE JSAC*, 17, Jul 1999.

[20] N. Bansal and Z. Liu. Capacity, delay and mobility in wireless ad-hoc networks. In *Proceedings of the IEEE Infocom*, pages 1553–1563, 2003.

[21] N. Bansal and Z. Liu. Capacity, delay and mobility in wireless ad-hoc networks. In *Proceedings of IEEE INFOCOM*, 2003.

[22] B. Baynat, G. Nogueira, M. Maqbool, and M. Coupechoux. An efficient analytical model for the dimensioning of wimax networks. In *Proceedings of the IFIP Networking*, pages 521–534, 2009.

[23] C. Berrou, A. Glavieux, and P. Thitimajshima. Shannon limit error-correcting coding and decoding: Turbo-codes. In *Proceedings of the IEEE ICC*, pages 1064–1070, 1993.

[24] D. Bertsekas and R. Gallager. *Data Networks*. PRENTICE HALL, Englewood Cliffs, New Jersey, 1987.

[25] N. Bisnik and A.A. Abouzeid. Queuing network models for delay analysis of multihop wireless ad hoc networks. *Ad Hoc Networks*, 7:79–97, 2009.

[26] M. Bloem, T. Alpcan, and T. Basar. A stackelberg game for power control and channel allocation in cognitive radio networks. In *Proceedings of the ACM/ICST GameComm workshop*, Nantes, France, October 2007.

[27] O. Blume and et al. Energy savings in mobile networks based on adaptation to traffic statistics. *Bell Labs Tech. J.*, 15, Sep 2010.

[28] T. Bonald, S. Borst, N. Hegde, M. Jonckheere, and A. Proutiere. Flow-level performance and capacity of wireless networks with user mobility. *QUESTA*, 63, Dec 2009.

[29] T. Bonald, S. Borst, and A. Proutiere. How mobility impacts the flow-level performance of wireless data systems. In *Proceedings of IEEE INFOCOM*, 2004.

[30] N. Bonneau, M. Debbah, and E. Altman. Spectral efficiency of cdma downlink cellular networks with matched filter. *EURASIP Journal on Wireless Communications and Networking*, 2005.

[31] N. Bonneau, M. Debbah, E. Altman, and G. Caire. Spectral efficiency of cdma uplink cellular networks. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Philadelphia, USA, Mar 2004.

[32] V. S. Borkar. Stochastic approximation with two time scales. *Systems and Control Letters*, 29, 1997.

[33] V. S. Borkar, A. A. Kherani, and B. J. Prabhu. Control of buffer and energy of a wireless device: Closed and open loop approaches, Dec 2004.

[34] V. S. Borkar, A. A. Kherani, and B. J. Prabhu. Closed and open loop optimal control of buffer and energy of a wireless device. In *Proceedings of 3rd Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt'05)*, Trentino, Italy, 2005.

[35] S. Borst, N. Hegde, and A. Proutiere. Mobility-driven scheduling in wireless networks. In *Proceedings of IEEE INFOCOM*, 2009.

[36] S. Borst, A. Proutiere, and N. Hegde. Capacity of wireless data networks with intra- and inter-cell mobility. In *Proceedings of IEEE INFOCOM*, 2006.

[37] M. Bredel and M. Fidler. Understanding fairness and its impact on quality of service in ieee 802.11. In *arXiv*, 2008.

[38] John W. Byers, Michael Luby, Michael Mitzenmacher, and Ashutosh Rege. A digital fountain approach to reliable distribution of bulk data. In *Proceedings of ACM SIGCOMM*, 1998.

[39] Han. C., Harrold. T., Krikidis. I., Ku. I., Le. T. A., Videv. S., Zhang. J., Armour. S., Grant. P. M., Haas. H., Hanzo. L., Nakhai. M. R., Thompson. J. S., and C. X. Wang. Green radio: Radio techniques to enable energy efficient wireless networks. *IEEE Comm. Mag., Special Issue: Green Communications*.

[40] R. W Chang. Synthesis of band-limited orthogonal signals for multi-channel data transmission. In *Bell System Technical Journal*, pages 1775–1796, 1966.

[41] P. Charriere, J. Brouet, and V. Kumar. Optimum channel selection strategies for mobility management in high traffic tdma-based networks with distributed coverage. In *Proceedings of the IEEE Int. Conf. Personal Wireless Communications*, Bombay, India, Dec 1997.

[42] Kwang-Cheng Chen and J. Roberto B. de Marca. *Mobile WiMAX*. John Wiley & Sons, Ltd, Publication. ISBN 978-0-470-51941-7, 2008.

[43] M. Chiang, P. Hande, T. Lan, and C. W. Tan. *Power Control in Wireless Cellular Networks*. now Publishers, 2008.

[44] H. Claussen. The future of small cell networks. *IEEE COMMSOC MMTC E-Lett.*, Sep 2010.

[45] R. B. Cooper. *Introduction to Queueing Theory*. North Holland, 2nd edition, 1981.

[46] R. Couillet and M. Debbah. *Random matrix theory methods for wireless communications*. Cambridge University Press, 2011.

[47] R. Couillet, M. Debbah, and J. Silverstain. A deterministic equivalent approach for the capacity analysis of multi-user mimo channels. *IEEE Trans. on Info. Theory, Submitted*, 2009.

[48] Romain Couillet, Sebastian Wagner, and Merouane Debbah. Asymptotic analysis of linear precoding techniques in correlated multi-antenna broadcast channels. *IEEE Trans. on Info. Theory, Submitted*, 2009.

[49] G. Davi. Using picocells to build high-throughput 802.11 networks, Jul 2004.

[50] Guillaume de la Roche and Jie Zhang. Femtocell networks: Perspectives before wide deployments, Sep 2010.

[51] M. Debbah and R. Muller. Mimo channel modeling and the principle of maximum entropy. *IEEE Trans. on Info. Theory*, 51, 2005.

[52] Merouane Debbah. Theoretical foundations of mobile flexible networks. *Rev Journal on Electronics and Communications*, 1, Jan 2011.

[53] S. Dharmaraja, Kishor S. Trivedi, and Dimitris Logothetis. Performance analysis of cellular networks with generally distributed handoff interarrival times. *Elsevier Computer Communications*, 10, 2003.

[54] D. Dutta, A. Goel, and J. Heidemann. Oblivious aqm and nash equilibria. In *Proceedings of the IEEE Infocom*, April 2003.

[55] T. ElBatt and A. Ephremides. Joint scheduling and power control for wireless ad hoc networks. *IEEE Trans. on Wireless Comm.*, 3, Jan 2004.

[56] Kumar et al. Fountain broadcast for wireless networks. In *Second International Workshop on Networked Sensing Systems (INSS)*, 2005.

[57] R. Fantacci and D. Tarchi. Bridging solutions for a heterogeneous wimax-wifi scenario. *Journal of Communications and Networks*, 8:1–9, 2006.

[58] J. Filar, H.A. Krieger, and Z. Syed. Cesaro limits of analytically perturbed stochastic matrices. *Linear Algebra Applications*, 353, 2002.

[59] G. Fodor, A. Eriksson, and A. Tuoriniemi. Providing quality of service in always best connected networks. *IEEE Communications Magazine*, 41:154–163, June 2003.

[60] G. J. Foschini and Z. Miljanic. A simple distributed autonomous power control algorithm and its convergence. *IEEE Trans. on Veh. Tech.*, 42, Nov 1993.

[61] R. G. Gallager. *Low Density Parity Check Codes*. Cambridge, MA, USA: MIT Press, 1963.

[62] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah. Optimal throughput-delay scaling in wireless networks: part i: the fluid model. *IEEE/ACM Transactions Networking*, 14:2568–2592, 2006.

[63] A.E. Gamal, J. Mammen, B. Prabhakar, and D. Shah. Throughput-delay trade-off in wireless networks. In *Proceedings of the IEEE Infocom*, March 2004.

[64] Stephane Gaubert and et. al. Methods and applications of (max,+) linear algebra, Jan 1997.

[65] V. Gazis, N. Alonistioti, and L. Merakos. Toward a generic "always best connected" capability in integrated wlan/umts cellular mobile networks (and beyond). *IEEE Wireless Communications*, 12:20–29, 2005.

[66] L. Georgiadis, M. Neely, and L. Tassiulas. Resource allocation and cross layer control in wireless networks. *Foundations and Trends in Networking*, 1, 2006.

[67] D. Gesbert and et al. Multi-cell mimo cooperative networks: A new look at interference. *IEEE J. Select. Areas Commun.*, 28, Dec 2010.

[68] P. Glasserman and D. D. Yao. *Monotone structures in discrete event systems*. Wiley, New York, 1994.

[69] P. Godlewski, M. Maqbool, M. Coupechoux, and J-M. Kelif. Analytical evaluation of various frequency reuse schemes in cellular ofdma networks. In *Proceedings of ValueTools*, 2008.

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

[71] P. Graczyk, G. Letac, and H. Massam. The complex wishart distribution and the symmetric group. *Annals of Statistics*, 31, 2003.

[72] M. Grossglauser and D. N. C. Tse. Mobility increases the capacity of ad hoc wireless networks. *IEEE/ACM Transactions on Networking*, 10:477–486, Aug 2002.

[73] The Climate Group and Global e Sustainability Initiative (GeSI). Smart 2020: Enabling the low carbon economy in the information age, 2008.

[74] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46:388–404, 2000.

[75] B. Hajek. External splittings of point processes. *Mathematics of operation research*, 10, 1985.

[76] Y. Hayel, S. Lasaulce, R. El-Azouzi, and M. Debbah. Introducing hierarchy in energy efficient power control games. In *Proceedings of the ACM/ICST International Conference on Performance Evaluation Methodologies and Tools (valuetools'08, GameComm)*, Athens, Greece, October 2008.

[77] B. Hochwald and S. Vishwanath. pace-time multiple access: Linear growth in the sum rate. In *Proceedings of ALERTON*, 1998.

[78] T. Holliday, A. Goldsmith, P. Glynn, and N. Bambos. Distributed power and admission control for time varying wireless networks. In *Proceedings of the IEEE Globecom*, Nov 2004.

[79] Ekram Hossain. *Heterogeneous Wireless Access Networks Architectures and Protocols*. Springer, ISBN: 978-0-387-09776-3, 2009.

[80] Ekram Hossain, Dong In Kim, and Vijay K. Bhargava (Editors). *Cooperative Cellular Wireless Networks*. Cambridge University Press, 2011.

[81] J. Hoydis and M. Debbah. Green, cost-effective, flexible, small cell networks. *IEEE Comsoc MMTC E-Letter*, 5, Sep 2010.

[82] J. Hoydis and et al. Green small-cell networks. *IEEE Vehicular Technology Magazine*, Mar 2011.

[83] J. Mitola III and Jr. G. Q. Maguire. Cognitive radio: making software radios more personal. *IEEE Personal Communications Magazine*, 6:13–18, August 1999.

[84] R. Jain, D. W. Chiu, and W. R. Hawe. A quantitative measure of fairness and discrimination for resource allocation in shared computer system. dEC research report tr-301., September 1984.

[85] N. Jindal and A. Goldsmith. Dirty-paper coding versus tdma for mimo broadcast channels. *IEEE Trans. on Info. Theory*, 51, 2005.

[86] Zheng Jing, David N. C. Tse, Joseph B. Soriaga I, Jilei Hou, John E. Smeet, and Roberto Padovani. Downlink macro-diversity in cellular networks. In *Proceedings of ISIT*, 2007.

[87] M. K. Karray. Analytic evaluation of wireless cellular networks performance by a spatial markov process accounting for their geometry, dynamics and control schemes, doctoral thesis, Sep 2007.

[88] V. Kavitha and E. Altman. Queueing in space: design of message ferry routes in sensor networks. In *Proceedings of the International Teletraffic Congress (ITC)*, 2009.

[89] V. Kavitha, S. Ramanath, and E. Altman. Spatial queuing analysis for design and dimensioning of picocell networks with mobile users. *Elsevier Performance Evaluation*, 68, 2011.

[90] J. M. Kelif and E. Altman. Downlink fluid model of cdma networks. In *Proceedings of IEEE VTC Spring*, May 2010.

[91] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan. Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of Operations Research*, 1998.

[92] S.R. Kulkarni and P. Viswanath. A deterministic approach to throughput scaling in wireless networks. *IEEE Transaction on Information Theory*, 50:1041–1049, 2004.

[93] H. J. Kushner and G. Yin. *Stochastic Approximation and Recursive Algorithms and Applications*. Springer, Second edition, 2003.

[94] T. Lan, D. Kao, M. Chiang, and A. Sabharwal. An axiomatic theory of fairness for resource allocation. In *Proceedings of the IEEE INFOCOM*, San Diego, USA, 2010.

[95] S. Lasaulce, M. Debbah, and E. Altman. Methodologies for analyzing equilibria in wireless games. *IEEE Signal Processing Magazine, Special issue on Game Theory*, 26:51–52, 2009.

[96] J. D. C. Little. A proof for the queueing formula: L $= \lambda$W. *Operations Research*, 9:383–387, May-June 1961.

[97] A. Lozano, A. M. Tulino, and S. Verdu. Multiple-antenna capacity in the low-power regime. *IEEE Trans. on Info. Theory*, 49, Oct 2003.

[98] M. Luby. Lt codes. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 271–282, 2002.

[99] Jingxiang Luo and Carey Williamson. Managing hotspot regions in wireless cellular networks with partial coverage picocells. In *Proceedings of the 6th ACM international symposium on Mobility management and wireless access*, 2008.

[100] Zorzi. M. On the analytical computation of the interference statistics with applications to the performance evaluation of mobile radio systems. *IEEE Transactions on Communications*, 45:103–109, Jan 1997.

[101] D. J. C. Mackay. Fountain codes. *IEE Proceedings - Communications*, 152:1062–1068, December 2005.

[102] P. Marbach and R. Pang. Transmission costs, selfish nodes, and protocol design. In *Proceedings of the Third International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt'05)*, pages 31–40, 2005.

[103] J. G. Markoulidakis, G. L. Lyberopoulos, D. F. Tsirkas, and E. D. Sykas. Mobility modeling in third-generation mobile telecommunications systems. *IEEE Personal Communications*, 4:41–56, Aug 1997.

[104] F. Meshkati, H.V. Poor, S.C. Schwartz, and N.B. Mandayam. An energy efficient approach to power control and receiver design in wireless data networks. *IEEE Transactions on Communications*, 53:1885–1894, November 2005.

[105] J. Mo and J. Walrand. Fair end-to-end window-based congestion control. *IEEE ACM Transactions on Networking*, 5, Oct 2000.

[106] Femtocell Networks. Special issue. *EURASIP Journal on Wireless Communication and Networking*, 2010.

[107] Tropos Networks. Picocell mesh: Bringing low-cost coverage, capacity and symmetry to mobile wimax, Mar 2007.

[108] Philip V. Orlik and Stephen S. Rappaport. On the handoff arrival process in cellular communications. *Wireless Networks*, 7, 2001.

[109] Arogyaswami Paulraj, Rohit Nabar, and Dhananjay Gore. *Introduction to Space-Time Wireless Communications*. Cambridge University Press, 2003.

[110] C. Peel, B. Hochwald, and A. Swindlehurst. A vector-perturbation technique for near-capacity multiantenna multiuser communication-part i: Channel inversion and regularization. *IEEE Trans. on Communications*, 53, 2005.

[111] J. Pérez-Romero, O. Sallent, and R. Agustí. Policy-based initial rat selection algorithms in heterogeneous networks. In *Proceedings of the 7th IFIP International Conference on Mobile and Wireless Communication Networks (MWCN'05)*, Marrakech, Morocco, September 2005.

[112] Qualcomm. Evolution of wireless applications and services, Dec 2007.

[113] V. Ramaiyan, A. Kumar, and E. Altman. Fixed point analysis of single cell ieee 802.11e wlans: uniqueness, multistability and throughput differentiation. In *Proceedings of the ACM SIGMETRICS, Performance Evaluation Rev.*, volume 33, pages 109–120, Barcelona, Spain, April 2005.

[114] S. Ramanath, E. Altman, and K. Avrachenkov. A heterogeneous approach to fair resource allocation and its application in femtocell networks. In *Proceedings of the Workshop on Indoor and Outdoor Femtocells (IOFC'11)*, Princeton, USA, 2011.

[115] S. Ramanath, E. Altman, V. Kumar, and M. Debbah. Optimizing cell size in picocell networks. In *Proceedings of the Workshop on Resource Allocation in Wireless Networks (RAWNET'09)*, Seoul, South Korea, 2009.

[116] S. Ramanath, E. Altman, V. Kumar, V. Kavitha, and L. Thomas. Fair assignment of base stations in cellular networks. In *Proceedings of 22nd World Wireless Research Forum (WWRF'09)*, Paris, France, 2009.

[117] S. Ramanath, M. Debbah, E. Altman, and V. Kumar. Asymptotic analysis of precoded small cell networks. In *Proceedings of IEEE Infocom*, San Diego, USA, 2010.

[118] S. Ramanath, V. Kavitha, and E. Altman. Impact of mobility on call block, call drops and optimal cell size in small cell networks. In *Proceedings of 2rd Intl. Workshop on Indoor and Outdoor Femtocells (IOFC'10)*, Istanbul, Turkey, 2010.

[119] S. Ramanath, V. Kavitha, and E. Altman. Spatial queuing analysis for mobility in pico cell networks. In *Proceedings of 8th Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt'10)*, Avignon, France, 2010.

[120] S. Ramanath, V. Kavitha, and E. Altman. Open loop optimal control of base station activation for green networks. In *Proceedings of 9th Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt'11)*, Princeton, USA, 2011.

[121] S. Ramanath, V. Kavitha, and M. Debbah. Satisfying demands in a multicellular network: An universal power allocation algorithm. In *Proceedings of 9th Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt'11)*, Princeton, USA, 2011.

[122] R. Rao and A. Ephremides. On the stability of interacting queues in a multiple-access system. *IEEE Transactions on Information Theory*, 34:918–930, September 1988.

[123] T. Rappaport. *Wireless Communications: Principles and Practice (2nd Edition)*. Prentice Hall, 2001.

[124] S. Ray, D. Starobinski, and J.B. Carruthers. Performance of wireless networks with hidden nodes: A queuing-theoretic analysis. *Computer Communications*, 28:1179–1192, 2005.

[125] W. Saad, Z. Han, T. Basar, M. Debbah, and A. Hjørungnes. A selfish approach to coalition formation among unmanned air vehicles in wireless networks. In *Proceedings of the International Conference on Game Theory for Networks (Gamenets)*, 2009.

[126] M. Schwartz. *Information, Transmission, Modulation and Noise*. Third edition, McGraw-Hill, New York, 1980.

[127] S. Sen, A. Arunachalam, K. Basu, and M. Wernik. A qos management framework for 3g wireless mobile network. In *Proceedings of the Wireless Conference on Networking and Communications (WCNC)*, 1999.

[128] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423 and 623–656, April 1948.

[129] A. Shokrollahi. Raptor codes. *IEEE Transactions on Information Theory*, 52:2551–2567, June 2006.

[130] J. Silverstein and Z. Bai. On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *Journal of Multivariate Analysis*, 54, 1995.

[131] G. L. Siqueira, E. V. Vasquez, R. A. Gomes, C. B. Sampaio, V. C. F. Costa, and M. A. Socorro. Propagation measurements for indoor mobile picocell coverage. In *Proceedings of the Microwave and Optoelectronics Conference*, Aug 1997.

[132] O. Somekh, B. M. Zaidel, and S. Shamai. Asymptotic analysis of linear precoding techniques in correlated multi-antenna broadcast channels. *IEEE Trans. on Info. Theory*, 53, Dec 2007.

[133] G. Song and Y. Li. Cross-layer optimization for ofdm wirless networks-part i: theoritical framework. *IEEE transactions on wireless communications*, 4:614–624, March 2005.

[134] V. Srinivasan, P. Nuggehalli, C.F. Chiasserini, and R.R. Rao. Cooperation in wireless ad hoc networks. In *Proceedings of the IEEE Infocom*, volume 2, pages 808–817, San Francisco, March 2003.

[135] D. Stamatelos and A. Ephremides. Spectral efficiency and optimal base placement for indoor wireless networks. *IEEE JSAC*, 14:651–661, May 1996.

[136] R. Stemm and R. Katz. Vertical handoffs in wireless overlay networks. *Journal on Mobile Networks and applications*, 3:335–350, 1998.

[137] R. K. Sundaram. *A first course in optimization theory*. Cambridge university press, 2007.

[138] Richard S Sutton and Barto Andrew G. *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[139] W. Szpankowski. Stability conditions for some distributed systems. *Buffered random access systems, Advances in Applied Probability*, 26:498–515, 1994.

[140] Y. Takahashi. An approximation formula for the mean waiting time of an m/g/c queue. *Journal on Operations Research*, pages 150–163, 1977.

[141] R. Tandra, A. Sahai, and S. M. Mishra. What is a spectrum hole and what does it take to recognize one? *Proc. IEEE*, 97, May 2009.

[142] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: Performance criteria and code construction. *IEEE Trans on Information Theory*, March 1998.

[143] Emre Telatar. Capacity of multi-antenna gaussian channels. *European Transactions on Telecommunications*, 10:585–595, 1999.

[144] Emre Telatar. Capacity of multi-antenna gaussian channels. *European Transactions on Telecommunications*, 10:585–595, 1999.

[145] M. Tidball, A. Lombardi, O. Pourtallier, and E. Altman. Continuity of optimal values and solutions for control of markov chains with constraints. *SIAM Journal on Control and Optimization*, 38, 2000.

[146] C. K. Toh, W. K. Tsai, V. O. L. Li, and G. Guichai. Transporting audio over wireless ad hoc networks: Experiments and new insights. In *Proceedings of the 14th Personal, Indoor and Mobile Radio Communications Symposium (PIMRC)*, volume 1, pages 772–777, China, 2003.

[147] Corrine Touati, Eitan Altman, and Jerome Galtier. Fair power and transmission rate control in wireless networks. In *Proceedings of GLOBECOM*, 2002.

[148] D.N.C. Tse, P. Viswanath, and Lizhong Zheng. Diversity-multiplexing tradeoff in multiple-access channels. *IEEE Transactions on Information Theory*, 50:1859–1874, 2004.

[149] Antonia M. Tulino and Sergio Verdu. *Random Matrix Theory and Wireless Communications*. Now, 2004.

[150] S. Ulukus and R. Yates. Stochastic power control for cellular radio systems. *IEEE Trans. on Comm.*, 46, Jun 1998.

[151] A. Urpi, M. A. Bonuccelli, and S. Giordano. Modeling cooperation in mobile ad hoc networks: a formal description of selfishness. In *Proceedings of the International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, Sophia Antipolis, France, 2003.

[152] Sivarama Venkatesan, Angel Lozano, and Reinaldo Valenzuela. Network mimo: Overcoming intercell interference in indoor wireless systems. In *Proceedings of ASILOMAR*, 2007.

[153] S. Verdu. *Multiuser Detection*. Cambridge press, 2003.

[154] Cisco visual networking index:. Global mobile data traffic forecast update,2009-2014.

[155] H. Viswanathan and S. Venkatesan. Asymptotics of sum rate for dirty paper coding and beamforming in multiple-antenna broadcast channels. In *Proceedings of ALERTON*, 1998.

[156] Mai Vu and Arogyaswami Paulraj. Mimo wireless linear precoding. *IEEE Signal Processing Mag.*, Sep 2007.

[157] J.G. Wardrop. Some theoretical aspects of road traffic research. In *Proceedings of the Institution of Civil Engineers*, volume 2, pages 325–378, 1952.

[158] W. Webb. *Wireless Communications: The Future*. New York: Wiley, 2007.

[159] R. R. Weber and S. Stidham. Optimal control of service rates in a network of queues. *Advances in applied probability*, 19, 1987.

[160] H. Weingarten, Y. Steinberg, and S. Shamai. The capacity region of the gaussian multiple-input multiple-output broadcast channel. *IEEE Trans. on Info. Theory*, 52, 2006.

[161] T. U. Wien. Lte system level simulator. http://www.nt.tuwien.ac.at/about-us/staff/josep-colom-ikuno/lte-system-level-simulator.

[162] Wikipedia. Femtocell. available at http://en.wikipedia.org/wiki/femtocell.

[163] Wikipedia. Ultra-wideband. available at http://en.wikipedia.org/wiki/ultra-wideband.

[164] R. W. Wolff. *Stochastic Modeling and the Theory of Queues*. Prentice Hall, 1989.

[165] C. Wu and D. P. Bertsekas. Distributed power control algorithms for wireless networks. In *Proceedings of the IEEE VTC*, Mar 2001.

[166] Aaron D. Wyner. Shannon-theoretic approach to a gaussian cellular multiple-access channel. *IEEE Trans. on Info. Theory*, 40, Nov 1994.

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

[168] Ding Xu, Jianhua Zhang, Xinying Gao, Ping Zhang, and Yufei Wu. Indoor office propagation measurements and path loss models at 5.25 ghz. In *Proceedings of the VTC*, Sep 2007.

[169] Hujun Yin and Siavash Alamouti. Ofdma: A broadband wireless access technology. In *Proceedings of IEEE Sarnoff Symposium*, pages 1–4, 2006.

[170] W. J. Yuan, Y. Smeers, X. Huamg, and R. F. Serfozo. Spatial queueing processes. *Mathematics of Operations Research*, 24:865–886, Nov 1999.

[171] R. Zakhour and D. Gesbert. Coordination on the miso interference channel using the virtual sinr framework. In *Proceedings of the International Workshop on Smart Antennas*, 2009.

[172] M. Zorzi and R.R. Rao. Comments on "capture and retransmission control in mobile radio". *IEEE Journal on Selected Area in Communications*, 24:2341–2342, December 2006.