Shashika Manosha Kapuruhamy Badalge

CONVEX OPTIMIZATION BASED RESOURCE ALLOCATION IN MULTI-ANTENNA SYSTEMS
SHASHIKA MANOSHA KAPURUHAMY BADALGE

CONVEX OPTIMIZATION BASED RESOURCE ALLOCATION IN MULTI-ANTENNA SYSTEMS

Academic dissertation to be presented with the assent of the Doctoral Training Committee of Technology and Natural Sciences of the University of Oulu for public defence in Kuusamonsali (YB210), Linnanmaa, on 9 January 2018, at 12 noon

UNIVERSITY OF OULU, OULU 2018
Abstract

The use of multiple antennas is a fundamental requirement in future wireless networks as it helps to increase the reliability and spectral efficiency of mobile radio links. In this thesis, we study convex optimization based radio resource allocation methods for the downlink of multi-antenna systems.

First, the problem of admission control in the downlink of a multicell multiple-input single-output (MISO) system has been considered. The objective is to maximize the number of admitted users subject to a signal-to-interference-plus-noise ratio (SINR) constraint at each admitted user and a transmit power constraint at each base station (BS). We have cast the admission control problem as an $\ell_0$ minimization problem; it is known to be combinatorial, NP-hard. Centralized and distributed algorithms to solve this problem have been proposed. To develop the centralized algorithm, we have used sequential convex programming (SCP). The distributed algorithm has been derived by using the consensus-based alternating direction method of multipliers in conjunction with SCP. We have shown numerically that the proposed admission control algorithms achieve a near-to-optimal performance. Next, we have extended the admission control problem to provide fairness, where long-term fairness among the users has been guaranteed. We have focused on proportional and max-min fairness, and proposed dynamic control algorithms via Lyapunov optimization. Results show that these proposed algorithms guarantee fairness.

Then, the problem of admission control for the downlink of a MISO heterogeneous networks (hetnet) has been considered, and the proposed centralized and distributed algorithms have been adapted to find a solution. Numerically, we have illustrated that the centralized algorithm achieves a near-to-optimal performance, and the distributed algorithm’s performance is closer to the optimal value.

Finally, an algorithm to obtain the set of all achievable power-rate tuples for a multiple-input multiple-output hetnet has been provided. The setup consists of a single macrocell and a set of femtocells. The interference power to the macro users from the femto BSs has been kept below a threshold. To find the set of all achievable power-rate tuples, a two-dimensional vector optimization problem is formulated, where we have considered maximizing the sum-rate while minimizing the sum-power, subject to maximum power and interference threshold constraints. This problem is known to be NP-hard. A solution method is provided by using the relationship between the weighted sum-rate maximization and weighted-sum-mean-squared-error minimization problems. The proposed algorithm was used to evaluate the impact of imposing interference threshold constraints and the co-channel deployments in a hetnet.

Keywords: convex approximation techniques, distributed optimization, dynamic control, heterogeneous networks, multi-antenna systems, radio resource allocation
Tiivistelmä

Monen antennin käyttö on perusvaatimus tulevissa langattomissa verkoissa, koska se auttaa lisäämään matkaviestinyhteyksien luotettavuutta ja spektritehokkuutta. Tässä väitöskirjassa tutkitaan konveksiin optimointiin perustuvia radioresurssien allokointimenetelmiä moniantennijärjestelmissä alalinkin suunnassa.


Tämän jälkeen käsitellään heterogeenisen langattoman MISO-verkon pääsynvalvonnan ongelmia. Edellä ehdotettuja keskitettyjä ja hajautettuja algoritmeja on muokattu tämän ongelman ratkaisemiseksi. Työssä numerisesti osoitetaan, että sekä keskitetyllä että hajautetulla algoritmillä saavutetaan lähes optimaalinen suorituskyky.


Asiasanat: dynaaminen hallinta, hajautettu optimointi, heterogeeniset verkot, Konveksit approksimaatiomenetelmät, moniantennijärjestelmät, radioresurssien allokointi
To my ever-loving wife Dilani, wonderful daughter Vinudi, and parents
Preface

This doctoral thesis contains the results of the research that has been conducted at the Centre for Wireless Communications (CWC), University of Oulu, Finland. My stay at CWC has been quite educational and challenging, and hence, I was able to progress towards my personal and career development while working in this place. Here, I would like to take this opportunity to express my sincere gratitude to all the people who have supported me towards the successful completion of my thesis.

First of all, I would like to thank my supervisor, Professor Matti Latva-aho, for offering me an opportunity to pursue my PhD in a top-notch research unit. I am deeply grateful to him for his immense support, guidance, invaluable encouragement, and the confidence he has shown in me during my doctoral research. He has always directed me to work on high quality research projects, which has helped me to gain a great deal of experiences in the past years. I am grateful to my advisor Adjunct Professor Marian Codreanu for the valuable guidance provided to me during my doctoral studies. The comments and the criticisms he has provided to improve the scientific quality of the research work is appreciable. Then, I wish to express my deep gratitude to my advisor Professor Premanandana Rajatheva for the guidance, support, inspiration, encouragement he has provided and the confidence shown in me throughout my research career. Without all your support and encouragement I would not have come this far.

I would further wish to thank the pre-examiners of my thesis, Professor Dr.-Ing. Eduard A. Jorswieck, from Dresden University of Technology (TUD) in Germany and Associate Professor Marina Petrova, School of Information and Communication Technology, from the KTH Royal Institute of Technology in Sweden. Their insightful comments and constructive feedback helped me to further improve the quality of my thesis. My gratitude also goes to Professor Markku Juntti, Dean of the University of Oulu Graduate School, Dr. Harri Posti and Dr. Ari Pouttu, the director of the CWC during my doctoral studies, for providing a dynamic and inspiring working environment. I also would like to extend my gratitude to the members of my follow-up group Adjunct Professor Mehdi Bennis, Dr. Pedro Nardelli and Dr. Keigo Hasegawa.
During the years with the CWC, I have had the privilege to contribute to multiple projects at CWC: EARTH, LOCON, CRUCIAL, 5Gto10G, CORE++, and uO5G. I would like to thank the managers of these projects, Dr. Pekka Pirinen, Dr. Jouko Leinonen, Dr. Anna Pantelidou, Tuomo Hänninen, and Dr. Marja Matinmikko. These projects were funded by the European Commission, the Finnish Funding Agency for Technology and Innovation, the Finnish Communications Regulatory Authority, Ministry of Transport and Communications Finland, Nokia Networks, Elektrobit, and many other industrial partners. I would like to acknowledge them all. I was also fortunate to receive personal research grants for my doctoral studies from the following Finnish foundations: the Tutkissäätiö Foundation and the Riiatta and Jorma J. Takanen Foundation.

Within the period of my doctoral study I had the pleasure of being surrounded by many researchers in CWC who are from different parts of the world. I would like to thank the former colleagues of CWC Dr. Animesh, Bret, Dr. Carlos, Prasanth, Dr. Ratheesh, Tamur, Simon, Dr. Namal and the present colleagues Ayotunde, Ayswarya, Bidushi, Cheng, ElBamby, Dr. Pennanen, Hamidreza, Heshani, Dr. Hirley, Hossein, Inosha, Jarkko, Jari, Kalle, Kien-Giang, Dr. Madhusanka, Markus, Markku, Majidzadeh, Pawan, Dr. Petri, Praneeth, Dr. Qiang Xue, Dr. Tölli, Tachporn, Kien, Upul, amongst all the others for the nice and friendly environment you have made and the unforgettable memories I had with you all. Special thanks go for Dr. Chathuranga, Ganesh, Dr. Keeth, Satya, Sumudu, and Uditha for the nice times we had spent while having discussions on both technical and philosophical topics. I also would like to thank the collaborators from Nokia Dr. Seppo Yrjölä and Kari Horneman. I extend my gratitude also to the administrative staff from the CWC, more specifically Antero Kangas, Ann Niskanen, Elina Komminaho, Hanna Saarela, Jari Sillanpää, Kirsu Ojutkangas, Eija Pajunen and many others.

My stay in Oulu was made remarkable by the nice little community: Professor Rajatheva & family, Buddhika-Pawani, Chathuranga-Kanchana, Ganesh-Ayswarya, Madhusanka-Ruwanthi, Saliya-Bhagya, Sandun-Chamari, Satya-Sujeetha, Sumudu-Inosha, Tharanga-Dimuthu, Bidushi, Namal, Uditha, Upul, and the bunch of joyful kids Aditya, Ashini, Ayod, Diliru, Kesa, Sahas, Savith, Senehas, and Somnas. I also wish to extend my thanks to all my friends (specially Indika, Isuru, Udesha, Hasitha, Ranga, Asanka, Samitha, Viraj, Gayantha, Ranil, Haritha, Amara, Gayan, Madushan, Kasun, Darshana, and Madhura) and
all the relatives (specially nenda, loku amma, Maldini akka and their families) in Sri Lanka for their immense support. Special thanks also go to Dr. K. Gunawickrama and Dr. K. Pirapaharan at the Faculty of Engineering, University of Ruhuna, and Prof. K. Liyanage at Faculty of Engineering, University of Peradeniya.

I want to express my unreserved gratitude to my mother Manel and father Sarath (who recently passed away) for their love, kindness, support and encouragement provided throughout my life. The sacrifices you have made during my life are priceless. I am also grateful to my parents-in-law for their blessings. I extend my gratitude to my brother (Chathu), two sisters (Hashani and Ashini), and to my brothers and sisters in-laws and their families for their love, joy and help during my life. Then, to my sweet little angel Vinudi, thank you for bringing happiness to my life and that crucial motivation required to get things done. Finally, I would like to express deepest gratitude to my loving wife Dilani. Without your love, concern, understanding, encouragement, and support none of this would be possible.
Abbreviations

Roman-letter notations

\( a_{ki}(t) \) Admission status of \( i \)th user of base station \( k \) at time \( t \)

\( \mathbf{c}_{ki}^l \) Small scale fading coefficient vector from \( l \)th base station to \( i \)th user of base station \( k \)

\( \mathbf{c}_{ki} \) Small scale fading coefficient matrix from \( l \)th base station to \( i \)th user of base station \( k \)

\( \mathcal{C} \) Set of small scale fading coefficients

\( d_0 \) Far field reference distance

\( d_{ki}(t) \) Information symbol associated with \( i \)th user of base station \( k \) during time slot \( t \)

\( D_{	ext{BS}} \) Distance between two adjacent base stations

\( D \) Set of distances

\( \mathbf{E}_{ki} \) Mean-squared-error of \( i \)th user of base station \( k \)

\( \mathbf{E}_{	ext{MMSE}}^{ki} \) Minimum-mean-squared-error matrix of \( i \)th user of base station \( k \)

\( G_{ki}(t) \) Virtual queue of \( i \)th user of base station \( k \) at time slot \( t \) for proportional fair admission control

\( \mathbf{G}(t) \) Vector of virtual queues for proportional fair admission control

\( \mathbf{h}_{ki}^l(t) \) Channel vector from base station \( l \) to \( i \)th user of base station \( k \) during time slot \( t \)

\( \mathbf{H}_{ki}^l \) Channel matrix from base station \( l \) to \( i \)th user of base station \( k \)

\( \mathbf{J}_{ki} \) Received signal covariance matrix of \( i \)th user of base station \( k \)

\( \mathcal{K} \) Set of all base stations

\( L(\cdot) \) Lyapunov function

\( L_p(\cdot) \) Function used to represent the augmented Lagrangian

\( \mathbf{m}_{ki}(t) \) Transmit beamformer associated with \( i \)th user of base station \( k \) during time slot \( t \)

\( \mathbf{m} \) Vector obtained by stacking \( \mathbf{m}_{ki} \) for all \( i \in \mathcal{U}_k, k \in \mathcal{K} \)

\( M_k \) Number of antennas at base station \( k \)

\( n_{ki}(t) \) Circularly symmetric complex Gaussian noise with variance \( \sigma^2_{ki} \)

\( \mathbf{n}_{ki}(t) \) Circularly symmetric complex Gaussian noise vector

\( N_{ki} \) Number of antennas at \( i \)th user of base station \( k \)
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{O}$</td>
<td>The set of directly achievable power-rate tuples</td>
</tr>
<tr>
<td>$P_{\text{sum}}(\cdot)$</td>
<td>Sum-power function</td>
</tr>
<tr>
<td>$P_k^{\text{max}}$</td>
<td>Maximum transmit power of $k$th base station</td>
</tr>
<tr>
<td>$Q_k$</td>
<td>Feasible set of the local optimization problem associated with $k$th base station in admission control problem</td>
</tr>
<tr>
<td>$r$</td>
<td>Femtocell radius</td>
</tr>
<tr>
<td>$R$</td>
<td>Macrocell radius</td>
</tr>
<tr>
<td>$R_{ki}$</td>
<td>Achievable rate of $i$th user of base station $k$</td>
</tr>
<tr>
<td>$R_{\text{BS}}$</td>
<td>Radius of a cell</td>
</tr>
<tr>
<td>$R_{\text{INS}}(\cdot)$</td>
<td>Sum-rate function</td>
</tr>
<tr>
<td>$R_{\text{AVG}}$</td>
<td>Instantaneous power-rate region</td>
</tr>
<tr>
<td>$R_{\text{AVG}}(\cdot)$</td>
<td>Average power-rate region</td>
</tr>
<tr>
<td>$s_{ki}(t)$</td>
<td>Nonnegative auxiliary variable of $i$th user of base station $k$ at time slot $t$</td>
</tr>
<tr>
<td>$s_k$</td>
<td>Vector obtained by stacking $s_{ki}$ for all $i \in U_k$</td>
</tr>
<tr>
<td>$s$</td>
<td>Vector obtained by stacking $s_{ki}$ for all $i \in U_k$, $k \in \mathcal{K}$</td>
</tr>
<tr>
<td>$s_{ki}$</td>
<td>Data stream transmitted by base station $k$ to $i$th user of base station $k$</td>
</tr>
<tr>
<td>$\hat{s}_{ki}$</td>
<td>Data vector estimated by $i$th user of base station $k$</td>
</tr>
<tr>
<td>$t$</td>
<td>Time slot</td>
</tr>
<tr>
<td>$T$</td>
<td>Transmit antennas</td>
</tr>
<tr>
<td>$T_{\text{max}}$</td>
<td>Maximum number of time slots</td>
</tr>
<tr>
<td>$U_{ki}$</td>
<td>Receive beamformer of $i$th user of base station $k$</td>
</tr>
<tr>
<td>$U$</td>
<td>Collection of all the receive beamformers in the network</td>
</tr>
<tr>
<td>$U_{\text{MMSE}}_{ki}$</td>
<td>Linear minimum-mean-squared-error receiver of $i$th user of base station $k$</td>
</tr>
<tr>
<td>$U_k$</td>
<td>Set of all users associated with $k$th base station</td>
</tr>
<tr>
<td>$\tilde{U}_k$</td>
<td>Set of admissible users at $k$th base station</td>
</tr>
<tr>
<td>$v_k$</td>
<td>Scaled dual variable for $k$th base station</td>
</tr>
<tr>
<td>$V$</td>
<td>Trade-off parameter between the size of the virtual queue backlogs and the objective function in fair admission control problem</td>
</tr>
<tr>
<td>$V_{ki}$</td>
<td>Transmit beamforming matrix associated to $i$th user of base station $k$</td>
</tr>
<tr>
<td>$V$</td>
<td>Collection of all the transmit beamformers in the network</td>
</tr>
<tr>
<td>$w_{ki}, \tilde{w}_{ki}$</td>
<td>Weight associated with $i$th user of base station $k$</td>
</tr>
</tbody>
</table>
\( \mathbf{W}_{ki} \) Weight matrix associated with \( i \)th user of base station \( k \)

\( \mathbf{W} \) Collection of all the weight matrices in the network

\( x_{ki,k}^l \) Local copy of \( z_{ki}^l \) saved at \( k \)th base station

\( x_{ki,l}^l \) Local copy of \( z_{ki}^l \) saved at \( l \)th base station

\( \mathbf{x}_k(t) \) Antenna signal vector transmitted by \( k \)th base station during time slot \( t \)

\( \mathbf{x}_k \) Local copies of the intercell interference terms associated with \( k \)th base station

\( y_{ki}(t) \) Received signal by \( i \)th user of base station \( k \) during time slot \( t \)

\( \mathbf{y}_{ki}(t) \) Received signal vector by \( i \)th user of base station \( k \) during time slot \( t \)

\( z_{ki}^l \) Interference generated by \( l \)th base station to \( i \)th user of base station \( k \)

\( \mathbf{z}_k \) Global variables associated with \( k \)th base station

\( Z_{ki}(t) \) Virtual queue of \( i \)th user of base station \( k \) at time slot \( t \) for max-min fair admission control

\( \mathbf{Z}(t) \) Vector of virtual queues for max-min fair admission control

Greek-letter notations

\( \boldsymbol{\alpha} \) Vector of weights associated with sum-rate function

\( \beta_{ki} \) Interference-plus-noise experienced by \( i \)th user of base station \( k \)

\( \boldsymbol{\beta} \) Vector of weights associated with sum-power function

\( \gamma_{ki} \) SINR threshold of \( i \)th user of base station \( k \)

\( \Gamma_{ki} \) Received SINR of \( i \)th user of base station \( k \)

\( \Delta(\cdot) \) Drift in a Lyapunov function from one slot to the next

\( \eta \) Path loss exponent

\( \theta_{ij}^k \) Maximum interference that can be generated by \( k \)th base station towards \( j \)th user of base station 1

\( \mu_{ki}(t) \) Auxiliary variable associated with \( i \)th user of base station \( k \) at time slot \( t \) in fair admission control problem

\( \nu_{ki}(t) \) Auxiliary variable associated at time slot \( t \)

\( \sigma_{ki}^2 \) Variance of the white Gaussian noise at \( i \)th user of base station \( k \)

\( \Upsilon_{ki} \) Interference-plus-noise covariance matrix of \( i \)th user of base station \( k \)
Mathematical Operator notations and symbols

**conv(\mathcal{X})**  
Convex hull of the set \mathcal{X}

**E\{\cdot\}**  
Expectation

**I**  
Identity matrix; size of the matrix is implicit

**\log \det(X)**  
Log-determinant of matrix \(X\)

**\max(\cdot)**  
Maximum

**\min(\cdot)**  
Minimum

**\text{rank}(X)**  
Rank of matrix \(X\)

**\text{Tr}(\cdot)**  
Trace of a matrix

**\text{vec}(\cdot)**  
Vec-operator; if \(X = [x_1, \ldots, x_n]\), then \text{vec}(X) = [x_1^T, \ldots, x_n^T]^T

**|x|**  
Absolute value of the complex number \(x\)

**\langle x \rangle^+**  
Positive part of scalar \(x\), i.e., \max(0, x)

**||x||_0**  
Cardinality of vector \(x\) (number of nonzero elements)

**||x||_1**  
\(\ell_1\)-norm of vector \(x\)

**||x||_2**  
\(\ell_2\)-norm of vector \(x\)

**[x]_n**  
nth component of vector \(x\)

**X^H**  
Conjugate transpose (Hermitian) of matrix \(X\)

**X^T**  
Transpose of matrix \(X\)

**X^{-1}**  
Inverse of matrix \(X\)

**[X]_{i,j}**  
Element at the \(i\)th row and the \(j\)th column of matrix \(X\)

**|\mathcal{X}|**  
Cardinality of set \(\mathcal{X}\)

**\nabla f(x)**  
Gradient of function \(f\) at \(x\)

**\mathbb{C}**  
Set of complex numbers

**\mathbb{C}^n**  
Set of complex \(n\)-vectors

**\mathbb{C}^{m \times n}**  
Set of complex \(m \times n\) matrices

**\mathcal{C}N(m, K)**  
Circularly symmetric complex Gaussian distribution with mean \(m\) and variance \(K/2\) per dimension

**\mathcal{C}N(m, K)**  
Circularly symmetric complex Gaussian vector distribution with mean \(m\) and covariance matrix \(K\)

**R(x)**  
Real part of scalar \(x\)

**(\cdot)^*\)**  
Solution of an optimization problem

**\sim\)**  
Distributed according to

**\geq\)**  
Greater than or equal operator; between real matrices, it represents componentwise inequality
\(\triangleq\) Defined to be equal to

<table>
<thead>
<tr>
<th>Acronyms</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>5G</td>
<td>Fifth Generation</td>
</tr>
<tr>
<td>ASA</td>
<td>Authorized Shared Access</td>
</tr>
<tr>
<td>ADMM</td>
<td>Alternating Direction Method of Multipliers</td>
</tr>
<tr>
<td>BS</td>
<td>Base station</td>
</tr>
<tr>
<td>CBRS</td>
<td>Citizens Broadband Radio Service</td>
</tr>
<tr>
<td>CC</td>
<td>Central Controller</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Distribution Function</td>
</tr>
<tr>
<td>CSI</td>
<td>Channel State Information</td>
</tr>
<tr>
<td>DC</td>
<td>Difference of Convex</td>
</tr>
<tr>
<td>FBS</td>
<td>Femto Base Station</td>
</tr>
<tr>
<td>FU</td>
<td>Femto User</td>
</tr>
<tr>
<td>hetnet</td>
<td>Heterogeneous Network</td>
</tr>
<tr>
<td>KKT</td>
<td>Karush-Kuhn-Tucker</td>
</tr>
<tr>
<td>LMMSE</td>
<td>Linear Minimum-Mean-Squared-Error</td>
</tr>
<tr>
<td>LSA</td>
<td>Licensed Shared Access</td>
</tr>
<tr>
<td>MBS</td>
<td>Macro Base Station</td>
</tr>
<tr>
<td>MU</td>
<td>Macro User</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple-Input Multiple-Output</td>
</tr>
<tr>
<td>MISO</td>
<td>Multiple-Input Single-Output</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean-Squared-Error</td>
</tr>
<tr>
<td>MMSE</td>
<td>Minimum Mean-Squared-Error</td>
</tr>
<tr>
<td>OFDM</td>
<td>Orthogonal Frequency Division Multiplexing</td>
</tr>
<tr>
<td>OFDMA</td>
<td>Orthogonal frequency-division multiple access</td>
</tr>
<tr>
<td>QCQP</td>
<td>Quadratically Constrained Quadratic Program</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>SAS</td>
<td>Spectrum Access System</td>
</tr>
<tr>
<td>SDR</td>
<td>Semidefinite-Relaxation</td>
</tr>
<tr>
<td>SINR</td>
<td>Signal-To-Interference-Plus-Noise Ratio</td>
</tr>
<tr>
<td>SISO</td>
<td>Single-Input Single-Output</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-To-Noise Ratio</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>SOC</td>
<td>Second Order Cone</td>
</tr>
<tr>
<td>SOC</td>
<td>Second Order Cone Programming</td>
</tr>
<tr>
<td>WSR</td>
<td>Weighted Sum-Rate</td>
</tr>
<tr>
<td>WSRMax</td>
<td>Weighted Sum-Rate Maximization</td>
</tr>
<tr>
<td>WMMSE</td>
<td>Weighted-Sum-Mean-Squared-Error</td>
</tr>
</tbody>
</table>
Contents

Abstract 9
Tiivistelmä
Preface 13
Abbreviations 19
Contents 21
1 Introduction

1.1 Motivation ......................................................... 23
1.2 Review of earlier and parallel work ................................. 24
1.2.1 Weighted sum-rate maximization .................................. 24
1.2.2 Energy efficient resource allocation .................................. 26
1.2.3 Admission control ............................................... 31
1.3 Aims and the outline of the thesis .................................. 33
1.4 The author’s contribution to the publications 35

2 Admission control in MISO downlink wireless networks 37

2.1 Introduction ......................................................... 37
2.2 System model and problem formulation ............................. 39
2.3 Algorithm derivation ................................................ 41
2.3.1 Centralized algorithm ........................................... 42
2.3.2 Distributed algorithm ........................................... 45
2.3.3 Distributed algorithm via ADMM ................................. 48
2.3.4 The signaling overhead of Algorithm 2.1 and Algorithm 2.2 ................................. 53
2.3.5 Complexity of Algorithm 2.1 and Algorithm 2.2 .................. 53
2.3.6 Convergence of Algorithm 2.1 and Algorithm 2.2 .................. 54
2.4 Fair admission control .............................................. 55
2.4.1 Transformed problem via auxiliary variables ................................. 56
2.4.2 Solving the transformed problem ................................... 58
2.4.3 Solving the resource allocation problem of the dynamic control algorithm ................................. 61
2.4.4 Max-Min fair admission control ................................... 62
2.5 Admission control in heterogeneous networks ....................... 66
1 Introduction

The usage of wireless communication services has become an essential part of our everyday life, and hence, wireless connectivity has been considered as an important commodity like electricity and water. Originally, the wireless networks were designed to provide reliable voice only services. Then, with the evolution of digital communication techniques, it was possible to provide low data rate services, and thus facilitated mobile internet access. Since then, this technology has undergone a tremendous progress, and has significantly contributed to enhancing the life style of people in countless ways. The technology will keep on improving further, and currently industry along with academia have already set the path towards the next generation [1-6]. Surprisingly, in the future, we will be living in an era where we can anticipate a zero latency-gigabit experience in wireless communication networks [1].

It is envisioned that the networks will deliver 1GB of personalized data per user per day in the near future. Furthermore, it is predicted that by the year 2030 the expected amount of traffic will be 10,000 times greater than that was in 2010 [7]. From a different view point, in fifth generation (5G) systems the vision is to provide a peak data rate of 10Gbit/s and to increase the data rate up to 100Mbit/s at least for 95% of the cell-edge users [1]. This increase in the volume of traffic will be mainly driven by the growth in the number of applications which require higher data rates and the increase in the size of the contents. For example, entertainment and information retrieval applications that support augmented reality, interactive video connectivity (with high resolution cameras), sharing 3D videos, and other audio video streaming services consume a significant amount of data [1, 2, 8]. Interestingly, these predictions are quite evident by looking at the current situation where there are many people who own a phone, laptop, tablet, and few low-range communication devices. Undoubtedly, this trend will continue, and we will see people with a few tens of wireless devices in the next decade.

Wireless communication will also play a key role in the sectors like automotive, health, logistics and freight tracking sectors, and in enabling the concepts such as smart society and smart grids [1, 2]. Specifically, in the automotive industry there
are a wide range of applications where the wireless connectivity is mandatory. For example, to provide entertainment for the passengers, to enable augmented reality dashboards in vehicles, and for safety related applications like driver assistance systems the vehicles need to be connected to a network. In the case of health sector, areas such as telemedicine have benefited with this technology. Similarly to these use cases, there are many other scenarios where wireless communication will be imperative. Hence, it is clear that there are an ever-increasing number of devices connecting to wireless networks expecting a better mobile broadband experience. The main challenge for the research community is to investigate how to support such diverse use cases with the required data rates and reliability.

To increase the data rate of a wireless link, several significant technical developments such as multiple-antenna techniques [9, pp. 290-483], [10], novel spectrum sharing methods [11, 12], and network modifications (multi-tier heterogeneous wireless networks) [13–15] have been introduced in the past. Multiple antenna techniques provide diversity, array, and spatial multiplexing gains; thus helping to increase the spectral efficiency and the reliability of wireless links [9, pp. 71-101], [16, pp. 15-16]. Novel spectrum sharing techniques such as authorized shared access-licensed shared access (ASA-LSA) [17], co-primary spectrum sharing [17], and citizens broadband radio service-spectrum access system (CBRS-SAS) [12] have been able to allocate more spectrum for commercial usage, which also helps to increase the data rate. Heterogeneous wireless networks with multiple-tiers enable the transmitter to be brought closer to the receiver. Hence, networks with multiple-tiers allow transmitting with low power, and thus contribute to boosting data rates by reducing the interference in the network [13–15].

While designing the future wireless networks designers require a thorough understanding of the system design objectives and resource control methods to achieve those objectives. The reason is due to the scarcity of wireless resources, the performance improvements of a network that can be gained by solely using the above techniques (multiple antennas, spectrum sharing techniques, etc.) is limited. Hence, it is evident that, efficient resource allocation plays a central role in wireless networks [18], [19, pp. 115-170]. The goal of a resource allocation algorithm is to improve the system and individual performance by optimally using the scarce available resources.
In the introductory chapter, we present the motivation of the thesis in Section 1.1. The relevant literature that is associated with the scope of this thesis is presented in Section 1.2. The aims and the outlines are discussed in Section 1.3, and the author’s contribution to the publication is provided in Section 1.4.

1.1 Motivation

Resource allocation in wireless communication systems is a complex task due to the interference that occurs in these systems. Although, orthogonal allocation of resources can reduce interference, it is not an efficient method for utilizing scarce resources. Fortunately, in the case of nonorthogonal allocation the use of multi-antenna techniques helps to exploit the spatial domain; thus it enables efficient resource allocation in contrast to the single antenna systems. However, resource allocation in multi-antenna systems is complex compared to single antenna systems, because the decision variable space is larger and obtaining the channel state information is challenging. Furthermore, in these systems it requires the joint optimization of transmit beamformers and transmit power, and this is a difficult task. In addition, since commonly used resource allocation problems such as weighted sum-rate maximization and admission control are NP-hard, they add extra complications to the resource allocation problems. Thus, developing resource allocation algorithms for multi-antenna systems is a challenging research topic.

There are systematic methods such as branch and bound \cite{20} to find the globally optimal solutions for the above mentioned NP-hard resource allocation problems. However, such global methods are often proven to be slow in practice, and hence they are of little practical use. Thus, it is advantageous to recognize or reformulate the problem as a convex optimization problem, because such problems can be solved, very reliably and efficiently, using interior-point methods or other special methods in convex optimization \cite{21}. To this end, there are a variety of local optimization techniques, such as difference of convex (DC) programming \cite{22}, sequential convex programming \cite{23}, alternating optimization \cite[pp. 272-276]{24}, dual relaxation, etc., available in the literature.

The centralized implementation of a resource allocation algorithm requires the collection of the channel state information (CSI) of all the users, solving
the optimization problem, and providing information about the decision back to each base station. To carry out these tasks, a lot of backhaul signaling and high computational power are needed; thus centralized algorithms may limit the performance of a network in terms of scalability, delay, and backhaul efficiency etc. Therefore, it is interesting to investigate the distributed implementation of resource allocation problems.

1.2 Review of earlier and parallel work

A variety of resource allocation problems of interest within the research community include weighted sum-rate maximization, sum-power minimization, admission control, maximization of minimum signal-to-interference-plus-noise ratio (SINR), energy efficiency maximization, etc. All these problems are interesting research topics due to their unique challenges. However, in this thesis, we only provide a literature review of the relevant resource allocation problems, which are necessary for the detailed analysis presented in Chapters 2 and 3. That is, we specifically focus on the weighted sum-rate maximization (WSRmax), energy efficiency maximization, and the admission control problems. We further narrow down our literature review, and consider the following aspects of these problems: (1) the nature of the problem (i.e., whether it is convex or not), (2) the methods that have been used to solve the problems, and (3) the implementation method (centralized or distributed).

1.2.1 Weighted sum-rate maximization

The WSRmax problem is a highly focused resource allocation problem within the research community [25, pp. 4-10]. It is central to many other problems that we encounter in wireless network resource allocation. There are a wide variety of applications where WSRmax arises, for example, network utility maximization, link scheduling, cross-layer control, energy efficiency, and spectrum sharing, etc [25, pp. 4-6]. This problem is convex in orthogonal multiple-access channels [26]; however, in the context of interference channels it is nonconvex, in fact it is an NP-hard problem [27]. To solve this problem, for different system setups, both globally and locally (suboptimal) optimal methods have been proposed in the literature [25, pp. 7-9].
For single-input single-output (SISO) systems, globally optimal methods to solve the WSRmax problem have been considered in [28–30]. Specifically, in [28] the proposed solution is based on the branch and bound technique [20] in conjunction with DC programming [31]. However, this method is applicable only in problems where the objective function is convertible to DC form. In contrast to [28], the solution method provided in [29] is also based on the branch and bound technique, but it does not rely on the convertibility of the objective function into a DC form. Apart from [28, 29], the solution method proposed in [30] has been developed by transforming the problem into a linear fractional program [32], and then by exploiting the monotonically increasing property of the objective function with SINR.

In the case of multiple-input single-output (MISO) systems, the research works in [33–35] have proposed global optimization methods to solve the WSRmax problem. In [33], the authors use the outer polyblock approximation method [36] to provide a solution to this problem. The proposed solution is for a two-user case, and its extension to the general case (i.e., for more than two-users) is nontrivial. The work in [34] is an extension of [29] to a MISO system, using the branch and bound method. In [35], an approach to obtain the global optimal solution for this problem is proposed by jointly utilizing the outer polyblock approximation algorithm and rate profile techniques [37].

Due to the nonconvex nature of the WSRmax problem, the convergence speed of the algorithms proposed above can be slow for large networks [38]. Hence, these algorithms are not suitable for practical implementation in large networks. Therefore, fast-converging suboptimal algorithms are desirable in practice.

For SISO systems, suboptimal algorithms to solve the WSRmax problem have been proposed in [39, 40]. Specifically, in [39] an efficient algorithm that finds an approximate solution for this problem has been proposed by exploiting a relationship between two related problems: an approximated WSRMax problem (using high SINR approximation) and the max-min weighted SINR optimization problem. In [40] the authors have provided an algorithm that performs well in a high SINR regime, using geometric programming [41], [42, pp. 47-107].

For multi-antenna networks, i.e., for MISO and multiple-input multiple-output (MIMO) systems, suboptimal methods for the WSRmax problem have been investigated in [43–47]. For MIMO broadcast channels, an iterative algorithm...
has been proposed to find an approximate solution for the WSRMax problem in [43]. There, the WSRMax problem has been transformed to an equivalent weighted sum mean-squared-error (MSE) minimization problem. The weights for the MSE of each user are chosen in such a way that they depend on the optimal beamforming matrix. The proposed algorithm iterates between a weighted minimum mean-squared-error (WMMSE) transmit filter design, a minimum MSE receive filter design, and a weighting matrix update. The algorithm in [43] has been extended to a MIMO interference broadcast channel in [45]. In this study also, the authors rely on the equivalence between the WSRMax problem and the WMMSE minimization problem to develop the proposed algorithm. In addition, the algorithm in [45] can handle general utility functions (which includes WSR as a special case), and it can also be extended to solve WSRMax problems when there are channel estimation errors. The algorithm proposed in [44] to solve this problem is also developed by exploiting the relationship between MSE and the sum-rate. This algorithm iterates between the downlink system and a virtual uplink system in order to update the receive filters and allocate power (using geometric programming [40, 41]).

For MISO systems, a method to find an approximate solution for the WSRmax problem has been proposed in [46]. In this work, the problem is converted to a convenient form by approximating the objective function with a convex function and using the uplink-downlink duality [9, pp. 449-453]. Then, the block coordinate descent method [24, pp. 272-276] is applied to find a suboptimal solution. Following a different approach, in [47] the WSRmax problem has been formulated as a second order cone program (SOCP), and an iterative fast converging algorithm was proposed to find an approximate solution. In this work, the authors have numerically compared the convergence speed of the proposed algorithm with the one in [45], and have shown that their algorithm outperforms the one in [45].

It is worth noting that most of the above listed schemes are centralized methods, i.e., they need a central controller for collecting the channel state information and computing the precoders and transmit powers. However, due to backhaul constraints in practical (large) networks, the central processing is not an efficient approach. Thus, distributed algorithms are practically useful for solving resource allocation problems that appear in wireless networks.
For multi-antenna networks, the distributed implementation of the WSRmax problem has been considered in [45, 48–55]. In [48] a distributed algorithm for a multicell MISO system has been proposed, which utilizes only the local CSI. That is, a particular base station knows the channel state information of all the users in a network from that base station. To decouple the problem across each base station so that each subproblem involves only the local beamformers and local CSI, the authors have used a high SINR approximation method [56]. The distributed algorithm proposed in [49] is obtained by solving the Karush-Kuhn-Tucker conditions. In this method, the base stations need to coordinate with each other in each iteration; however, it shows a better performance compared to the uncoordinated case. To derive the algorithm proposed in [49] some parameters need to be found heuristically, and for that they have not provided a systematic method. The proposed method in [50] is developed by using primal decomposition method [57, 58]. The goal is to decompose the problem into a master problem and to a set of subproblems (specific to each base station). The master problem is solved by using sequential convex approximation technique and the subproblems are iteratively solved by using SOCP [21, pp. 156-160] in conjunction with geometric programming.

The authors in [45, 55] have proposed distributed algorithms for multiuser MIMO downlink systems, and have used the equivalence between the WSRMax problem and WMMSE problem to derive their algorithms. To carry out the distributed implementation, the authors in [45] assume that local CSI is available for each user. Here, each user estimates the received signal covariance matrix, and updates the receive beamformers and weight matrix and informs the base stations. It is assumed that there is a separate channel available for each user to transmit this information to each base station. Then, using the received information each base station updates the transmit beamformers of its own users, without communicating with other base stations. The algorithms proposed in [55] are based on primal decomposition and Lagrangian relaxation. The main focus in [55] is to provide a low complexity distributed algorithm with minimum amount of over the air signalling.
1.2.2 Energy efficient resource allocation

Recently, the research community has steered their research towards investigating resource allocation schemes for improving the energy efficiency of wireless networks [59–72]. In the literature, the objective of an energy efficient resource allocation problem is defined in one of the following ways: maximizing the sum-rate per unit of expended energy in the network or minimizing the energy consumption per a transmitted bit in the network. In other words, this can be expressed as optimizing the ratio between the sum-rate and sum-power or vice versa. Hence, this problem inherits all the challenges encountered while solving the WSRMax problem.

Based on the existing studies, the approaches that have been used to improve the energy efficiency of a network can be categorized as follows [73]:

- **Resource allocation**: This method introduces novel resource allocation algorithms that maximize the energy efficiency of a network. Specifically, the investigations have been carried out for single and multiple antenna systems, orthogonal frequency-division multiple access (OFDMA) systems, heterogeneous networks, and cooperative networks.

- **Network planning and deployment**: This approach proposes infrastructure management methods, such as base station switch-on/switch-off algorithms and antenna subset selection techniques.

- **Energy harvesting and transfer**: This method proposes energy harvesting methods to operate communication systems. This technique does not reduce the energy consumption of a network, instead, it enables the networks to be powered by energy harvesting methods.

- **Hardware solutions**: The focus here is on developing new hardware that reduces the energy consumption of base stations. Special attention has been given to designing low power consuming RF chains and simplified transceiver structures.

In this thesis, we mainly focus on the radio resource allocation techniques that have been used to improve energy efficiency, and we also pay attention to the literature that considers resource allocation schemes which incorporate antenna selection techniques.
To quantify the energy efficiency of a wireless network, it is essential to consider the power consumption of the entire base station [71, 72]. This is because in addition to the transmit power a large amount of power is consumed by base station components. Here, the base station components include an antenna interface, a power amplifier, a radio frequency small-signal transceiver section, a baseband interface including a receiver (uplink) and transmitter (downlink) section, a DC-DC power supply, an active cooling system, and an AC-DC unit (mains supply) [72]. Thus, unlike the other problems, a distinct feature in energy efficient resource allocation is that the total power consumption is modeled by considering the power consumption of the base station components, in addition to the total transmit power [72].

It is possible to obtain a linear model, particularized to the type of base station, that captures the power consumption of the base station components [72]. However, a large number of publications in the literature consider a simplified linear power consumption model. In this model, the total power consumption of the base station components is approximated as the summation of the power consumed by the circuits and the power amplifier [60–67]. Note that in these research works the power consumption of the base station components is a constant value. Hence, from the convex optimization point of view, this is not a decision variable for the optimization problem. Thus, the resource allocation decision does not depend on the power consumption of the base station components. For example, the impact of circuit power is evaluated in [63] by changing it to some predefined values, via numerical simulations. Similarly, in [64] the effect of circuit power on the energy efficiency is measured by changing the number of antennas in the simulations. Interestingly, in recent studies [70, 74], the authors have also considered the number of selected antennas as a decision variable. More specifically, in [70] an energy efficient resource allocation scheme has been considered for the downlink of a MISO system, in which the number of active antennas is a decision variable. With this approach, it is possible to find the optimal number of antennas that should be activated in the system so that the energy efficiency is maximized. Furthermore, the research work in [74] considers a hybrid beamforming scheme that maximizes the energy efficiency of a massive MIMO network while dynamically selecting the radio frequency (RF) chains.
The main focus in [68, 69] is different compared to the other related literature to energy efficiency. Specifically, in [68] the goal is to find the theoretically achievable energy efficiency in a point-to-point MIMO channel for static and fast-fading cases. In [69] the main contribution is to obtain the convex-hull of the power-rate region for a MIMO heterogeneous networks. The advantage of this study is, by observing the power-rate region one can easily calculate the energy efficiency of a heterogeneous network subject to the given constraints (e.g., maximum transmit power). It is worth mentioning that the authors in studies [68, 69] have not considered the total power consumption of a base station. Nevertheless, it is easy to incorporate a static power consumption model [72] with these two works. However, incorporating an active antenna set selection for these studies is not a straightforward extension, and hence such investigations are interesting future works of the problems considered in [68, 69].

In the following, we provide a summary of investigations that consider energy efficient resource allocation in multi-antenna systems, and we pay attention to the method of implementation. The centralized implementation of the energy efficiency maximization problem for the downlink of MISO systems has been considered in [63, 64, 75]. Specifically, in [63] the energy efficiency is maximized subject to individual SINR constraints. A suboptimal solution for this problem is obtained via a zero-gradient based iterative algorithm. In [64], a heterogeneous network is considered, and the objective is to minimize the energy consumption per bit of the femtocells subject to individual rate constraints of the femto users and interference constraints of the macro users. A suboptimal solution to this problem is obtained by using the alternating optimization technique in conjunction with SOCP and geometric programming. Following a different approach, that is by using the uplink-downlink duality method, an energy efficient resource allocation algorithm is proposed in [75]. Although, the aforementioned schemes provide suboptimal solutions, they outperform each other in terms of their performances such as the convergence speed, complexity and the obtained objective value. For a MIMO heterogeneous network, an approximate solution for energy efficiency maximization problem is provided via Dinkelbach’s method [60, 76] in [77].

The authors in [78] have considered the problem of maximizing the weighted sum of per-cell energy efficiencies in a heterogeneous network. Instead of defining the energy efficiency as the ratio between sum-rate and sum-power, the purpose
of defining the objective function as in [78] is to address a variety of energy efficiency requirements in different cells (macro and femto) in a heterogeneous network. By utilizing the relationship between the WSRmax problem and WSMME problem [45], an iterative algorithm is proposed to find a suboptimal solution to this problem via block coordinate ascent method.

For MISO systems, a distributed implementation of the energy efficiency maximization problem has been considered in [79, 80]. The proposed solution in [79] is derived by exploiting the equivalence between the WSRMax problem and a WMMSE problem [45]. Although the proposed solution is suboptimal, the problem can be solved in parallel at each base station. However, this method needs to perfectly estimate and feedback the received signal covariance during each iteration, which can be challenging in the presence of user terminal imperfections, such as estimation feedback errors [50]. In [80], the authors have proposed a pricing-based distributed algorithm for a MISO interference channel for weighted sum energy efficiency maximization problem. The pricing metric represents the marginal decrease in the energy efficiency of a user due to a marginal increase in the interference that the user experiences. This algorithm updates the beamformers in a sequential order, and hence, it is not suitable for a network with large number of base stations. In a different context, i.e., for a heterogeneous network, the authors in [77] have proposed a distributed algorithm based on the alternating direction method of multipliers (ADMM) to solve the energy efficiency maximization problem for the downlink of a MIMO system. In [81], an asynchronous distributed algorithm is proposed for a MIMO interference broadcast channel using a noncooperative game theoretic approach. Here, each link updates its transmit covariance matrix without following any specific order, i.e., some users may update the precoders more frequently than other users.

### 1.2.3 Admission control

Another important objective in wireless networks is to maximize the number of admitted users to the network subject to guaranteed QoS constraints of them and the maximum transmit power of the base stations [82, pp. 33-46]. By solving it one can find the maximum number of users that can be admitted to the
network, with a guaranteed QoS. Unfortunately, this is a difficult combinatorial optimization problem, and it is known to be NP-hard [83, 84].

The exhaustive search method is one approach to finding the global optimal solution to the admission control problem. However, the computational complexity of an exhaustive search method increases exponentially with the number of users. Systematic approaches like the branch and bound method have been proposed to optimally solve this problem [84]. Although the solution in [84] is optimal, it is not suitable for practical scenarios due to the complexity of the branch and bound method [20]. Hence, fast suboptimal algorithms are desirable in practice [83, 85, 86].

In the case of SISO systems, suboptimal methods for the problem of admission control have been proposed in [87–93]. Specifically, in [87] this problem has been cast as an $\ell_0$ minimization problem and an $\ell_1$-norm relaxation technique has been used to find an approximate solution. In [88] different priority user groups are considered and, by exploiting a relationship between the signal-to-interference-plus-noise ratios (SINR) and the transmit power of the users, an iterative algorithm is proposed. Both works in [87] and [88] have proposed centralized algorithms. By using the dual decomposition technique [94] a distributed algorithm is proposed in [89]. Since the dual decomposition method relies on the subgradient method [95], the convergence speed of [89] can be slow and highly sensitive on the choice of a step size. By solving the Karush-Kuhn-Tucker (KKT) optimality conditions, a distributed algorithm is proposed in [90]. Note that a solution obtained by solving the KKT conditions is only a stationary point [21, pp. 215-272], and in general its optimality is not guaranteed for a nonconvex (admission control) problem. By using the Foschini and Miljanic power control technique [96], distributed methods for admission control have been studied in [91–93]. However, these solutions cannot be directly applied to multi antenna networks.

For MISO systems, the centralized implementation of the admission control problem has been studied in [83]. Specifically, the authors in [83] have formulated this problem as an integer nonlinear optimization problem [97, pp. 1-40] for a single cell. Then, two approximate solutions are proposed via semidefinite-relaxation (SDR) method [98, 99] and SOCP [21, pp. 156-160]. The work in [83] can be directly applied to a multiecell scenario, however, its distributed implementation is not straightforward, and there is no reported work on that.
In the context of a multicell system, this problem has been cast as an $\ell_0$ minimization problem in [85]. Then, the $\ell_1$-norm relaxation technique [100, 101] in conjunction with the SDR method is used to provide both centralized and distributed algorithms. In particular, the distributed algorithm is based on a block coordinate descent method [24, pp. 272-276], where the subproblems associated with the base stations are solved in a cyclic order. Both algorithms in [83] and [85] are derived by using the deflation based approach, which relies on dropping users at each iteration of the algorithm. In a different context (i.e., for a heterogeneous network) the problem of admission control has been addressed in [86]. By imposing cross-tier interference constraints to protect macro users, the authors of [86] proposed several suboptimal centralized and distributed algorithms.

In cellular networks, users who are closer to the base station have a better average SINR compared to the cell edge users. Additionally, some users can be located in a rich scattering environment and others with no scatterers around them [9, pp. 228-278]. These differences in the fading statistics of the users make it challenging to provide fairness for the users while they are being dynamically admitted to the system [16, pp. 279-292]. In [83] and [87], it is mentioned that the fairness can be provided by scheduling the non-admitted users with orthogonal resources. However, the authors in [83] and [87] have not provided any methods to achieve fairness while admitting users. Furthermore, since the resources are scarce, orthogonal allocation may not be spectrally efficient even though it can provide fairness [9, pp. 228-275],[102]. Hence, providing fairness in the admission control problem (i.e., admitting users with a guaranteed SINR level) is an important, and yet unaddressed point in the literature.

1.3 Aims and the outline of the thesis

The aim of this thesis is to apply convex optimization techniques to develop radio resource allocation algorithms for multi-antenna (i.e., MISO and MIMO) systems, to solve important and challenging problems that appear in the field of wireless communications. Specifically, we focus on designing linear transmit beamforming techniques to optimize the performance of wireless networks (multicell cellular and heterogeneous wireless networks). For some of the considered problems we provide only a centralized solution, and for others we propose both centralized and
distributed algorithms. In the following we briefly outline the main contributions of this thesis, which are presented in two different chapters.

In Chapter 2, we consider the problem of admission control in the downlink of a multicell MISO system. The objective is to maximize the number of admitted users subject to a SINR constraint for each admitted user and a transmit power constraint at each base station. We cast the admission control problem as an $\ell_0$ minimization problem. This problem is known to be combinatorial, and NP-hard. Hence, we have to rely on suboptimal algorithms to solve it. We first approximate the $\ell_0$ minimization problem via a non-combinatorial one. Then, we propose centralized and distributed algorithms to solve the non-combinatorial problem. To develop the centralized algorithm, we use sequential convex programming method. The distributed algorithm is derived by using alternating direction method of multipliers in conjunction with sequential convex programming. We numerically show that the proposed admission control algorithms achieve a near-to-optimal performance. Next, we extend the admission control problem to provide fairness, where the long term fairness between users is guaranteed. We focus on proportional and max-min fairness and propose dynamic control algorithms via Lyapunov optimization. It is shown numerically that the proposed fair admission control algorithms guarantee fairness between the users. The results are presented in [103, 104].

Finally, in this chapter, the problem of admission control for the downlink of a MISO heterogeneous wireless network has been considered. Compared to the multicell cellular networks, in this system all the macro users have to be served with their required SINR targets. Hence, the original admission control problem has been modified so that it addresses this requirement. For this problem, we have found a suboptimal solution by adapting the centralized and distributed algorithms that have been proposed for a multicell system. Numerically, we have illustrated that the centralized algorithm achieves a near-to-optimal performance, and the distributed algorithm’s performance is close to the optimal value.

In Chapter 3, we consider a single-macrocell heterogeneous MIMO network, where the macrocell shares the same frequency band with a femto network. The interference power for the macro users from the femto base stations is kept below a threshold to guarantee that the performance of the macro users is not degraded due to the femto network. We consider the problem of finding the set of all achievable power-rate tuples for this setting. To do this, we first formulate
a two-dimensional vector optimization problem in which we consider maximizing
the sum-rate and minimizing the sum-power, subject to maximum transmit
power and interference threshold constraints. The considered problem is NP-hard.
We provide a method to solve this problem by using the relationship between
the weighted sum-rate maximization and weighted-sum-mean-squared-error
minimization problems. Furthermore, using the proposed algorithm, we evaluate
the impact of imposing interference threshold constraints and the impact of
coop-channel deployment in heterogeneous networks. The proposed algorithm can
be used to evaluate the performance of realistic heterogeneous networks via
off-line numerical simulations. The results have been presented in [69, 105].

Finally, in Chapter 4 we have provided the conclusions of the research works
presented in each chapter. Then, possible future extensions of the problems
considered in this thesis are briefly described.

1.4 The author’s contribution to the publications

The thesis is based on two journal papers [69, 103], and two related conference
papers [104, 105]. The papers [69, 105] have already been published, paper [103]
is in minor revision round, and paper [104] has been accepted for publication.

The author of this thesis had the main responsibility in carrying out the
analysis, developing MATLAB simulation codes, generating the numerical results,
and writing papers [69, 103–105]. Other authors provided comments, criticism,
and support during the process.

In addition to the papers [69, 103–105], the author has published four
conference papers [64–66, 106] which are not included in this thesis. The author
has also contributed to the journal paper [107].
2 Admission control in MISO downlink wireless networks

2.1 Introduction

In this chapter, we consider the problem of admission control for the downlink of a multicell MISO system. The objective is to maximize the number of admitted users subject to an SINR constraint for each (admitted) user and a maximum transmit power constraint at each base station. Inspired by [87] and [90], we cast this problem as an $\ell_0$ minimization problem, where we minimize the number of non-admitted users (instead of maximizing the number of admitted users). This $\ell_0$ minimization problem is known to be NP-hard [100, 108]. Compared to the $\ell_0$ based problem formulations in [87] and [90], which are restricted to the SISO case, we have generalized the formulation for a MISO system. This extension is highly nontrivial, because the beamformers are optimization variables in our problem; hence, the methods proposed in [87] and [90] cannot be applied to solve the admission control problem considered in this chapter.

Next, we extend the admission control problem to take into account the fairness among the users. To the best of our knowledge all the existing admission control algorithms consider a static case (i.e., the admission control problem for a given instance). Thus, when these algorithms are applied to a dynamic network over a period of time, they may not provide fairness while admitting users. In this chapter, we consider the problem of fair admission control in a dynamic network. Specifically, we focus on providing long term fairness [109] based on proportional and max-min fairness criteria. To do this, we combine the Lyapunov drift-plus-penalty framework [110] with the recent results in sparse optimization [101], and cast the fair admission control problem as a stochastic optimization problem.

The main contributions of this chapter are two-fold. The first is to propose both centralized and distributed algorithms to solve the problem of admission control for the static case. Compared to the deflation based methods proposed in the literature, our centralized algorithm is developed based on sequential convex programming [22]. The distributed algorithm is derived by using the consensus-
based alternating direction method of multipliers (ADMM) in conjunction with sequential convex programming. In contrast to the existing distributed algorithm for multicell systems \cite{85} (that solves the subproblems in a cyclic order), our proposed algorithm solves the subproblems independently in parallel at all base stations. Thus, compared to the distributed algorithm in \cite{85} our algorithm is desirable for a network with a large number of BSs. Simulation results show that the proposed admission control algorithms achieve near-to-optimal performance. Furthermore, the results show that for a network with large number of users, the execution time of the proposed centralized algorithm is much less than that of the algorithm in \cite{83}.

The second contribution of this chapter is to provide dynamic control algorithms that guarantee proportional and max-min fairness while admitting users to the network. We derive these algorithms by using the drift-plus-penalty framework in \cite{110}. Specifically, we introduce a virtual queue for each user; this grows if the user is unserved in the current time slot and reduces otherwise. The proposed dynamic control algorithms use these virtual queues as the scheduling priorities for the users. Numerical results show that the proposed algorithms guarantee proportional and max-min fairness among the users.

Furthermore, we also investigate the problem of admission control for heterogeneous networks where there are users with different priorities. Specifically, we consider a two-tier macro femto network, where the macrocell shares the same frequency band with a set of femtocells. Since the macro users are high priority users (compared to the femto users), all of them have to be admitted with the required SINR targets. The femto users are admitted by guaranteeing their SINR targets, only if they can coexist with the macro users while satisfying the macro users requirements. We show that by modifying the original admission control problem we can accomplish this task. To find a suboptimal solution to this problem, we adapt the proposed centralized and distributed admission control algorithms. Numerically, we show that the centralized algorithm achieves near-to-optimal performance, and the distributed algorithm performance is close to the optimal value.
2.2 System model and problem formulation

We consider the downlink of a multicell MISO system with \( K \) base stations as shown in Fig. 2.1. The set of all base stations is denoted by \( \mathcal{K} \) and we label them with the integer values \( k = 1, \ldots, K \). We assume that each base station is equipped with \( T \) transmit antennas and each user with a single receive antenna, that is a MISO system is considered. We denote the set of all users associated with \( k \)th base station by \( \mathcal{U}_k \), and we label them with the integer values \( i = 1, \ldots, I_k \). The network is assumed to be operating in slotted time with normalized slots \( t \in \{1, 2, \ldots\} \).

![Fig. 2.1. System model for the downlink multicell MISO network.](image)

The antenna signal vector transmitted by \( k \)th base station during time slot \( t \) is given by

\[
\mathbf{x}_k(t) = \sum_{i \in \mathcal{U}_k} \mathbf{m}_{ki}(t)d_{ki}(t), \tag{2.1}
\]
where \( d_{ki}(t) \in \mathbb{C} \) and \( \mathbf{m}_{ki}(t) \in \mathbb{C}^{T} \) represent the information symbol and the transmit beamformer associated with \( i \)th user of base station \( k \) during time slot \( t \), respectively. We assume that the information symbol \( d_{ki}(t) \) is normalized such that \( E|d_{ki}(t)|^2 = 1 \). Furthermore, we assume that the data streams are independent, i.e., \( E\{d_{ki}(t)d_{kj}(t)^*\} = 0 \) for all \( i \neq j \), where \( i, j \in \mathcal{U}_k, \ k \in \mathcal{K} \).

The signal received by \( i \)th user of base station \( k \) in time slot \( t \) can be written as

\[
y_{ki}(t) = (h_{ki}^H(t))\mathbf{m}_{ki}(t)d_{ki}(t) + \sum_{j \in \mathcal{U}_k, j \neq i} (h_{ki}^H(t))\mathbf{m}_{kj}(t)d_{kj}(t) + \sum_{l \in \mathcal{K} \setminus \{k\}} \sum_{j \in \mathcal{U}_k} (h_{ki}^H(t))\mathbf{m}_{lj}(t)d_{lj}(t) + n_{ki}(t),
\]

where \( h_{ki}^H(t) \in \mathbb{C}^{T} \) is the channel vector from base station \( l \) to \( i \)th user of base station \( k \) and \( n_{ki} \) is circularly symmetric complex Gaussian noise with variance \( \sigma_{ki}^2 \). The received SINR of \( i \)th user of base station \( k \) is given by

\[
\Gamma_{ki}(\mathbf{m}(t)) = \frac{|(h_{ki}^H(t))\mathbf{m}_{ki}(t)|^2}{\sum_{j \in \mathcal{U}_k, j \neq i} |(h_{ki}^H(t))\mathbf{m}_{kj}(t)|^2 + \sum_{l \in \mathcal{K} \setminus \{k\}} \sum_{j \in \mathcal{U}_k} |(h_{ki}^H(t))\mathbf{m}_{lj}(t)|^2 + \sigma_{ki}^2},
\]

where we use the notation \( \mathbf{m}(t) \) to denote a vector obtained by stacking \( \mathbf{m}_{ki}(t) \) for all \( i \in \mathcal{U}_k, \ k \in \mathcal{K} \), on top of each other, i.e., \( \mathbf{m}(t) = [\mathbf{m}_{11}(t)^T, \mathbf{m}_{12}(t)^T, \ldots, \mathbf{m}_{K\ell_k}(t)^T]^T \).

We assume that the power allocation is subject to a maximum transmit power constraint at each base station, i.e.,

\[
\sum_{i \in \mathcal{U}_k} \| \mathbf{m}_{ki}(t) \|^2 \leq P_{k}^{\text{max}}, \quad k \in \mathcal{K},
\]

where \( P_{k}^{\text{max}} \) is the maximum transmit power of \( k \)th base station. Furthermore, we assume that the QoS of \( i \)th user of base station \( k \) is assured if its SINR is greater than a threshold \( \gamma_{ki} \), i.e.,

\[
\Gamma_{ki}(\mathbf{m}(t)) \geq \gamma_{ki}.
\]

Let \( \tilde{\mathcal{U}}_k(\mathbf{m}(t)) \) denote a generic set of admissible users at \( k \)th base station. Specifically, \( \tilde{\mathcal{U}}_k(\mathbf{m}(t)) \) denotes a set of users who satisfy their SINR thresholds under the power constraint, i.e.,

\[
\tilde{\mathcal{U}}_k(\mathbf{m}(t)) = \{ ki \mid \Gamma_{ki}(\mathbf{m}(t)) \geq \gamma_{ki}, \sum_{i \in \mathcal{U}_k} \| \mathbf{m}_{ki}(t) \|^2 \leq P_{k}^{\text{max}}, \quad i \in \mathcal{U}_k \}, \quad k \in \mathcal{K}.
\]
Our goal is to maximize the number of admitted users to the system, i.e., to maximize the sum of the cardinalities of $\tilde{U}_k(m(t))$ for all $k \in K$. We now formulate this design problem as a mathematical optimization problem. To do this, let us introduce the nonnegative auxiliary variables $s_{ki}(t)$ for all $k \in K$, $i \in U_k$, and consider a set of relaxed SINR constraints as follows:

$$\Gamma_{ki}(m(t)) \geq \gamma_{ki} - s_{ki}(t), \quad i \in U_k, k \in K. \quad (2.7)$$

In (2.7), when $s_{ki}(t) = 0$ we recover constraint (2.5), i.e., the SINR constraint of $i$th user of $k$th base station is satisfied. Furthermore, by making $s_{ki}(t)$ large enough the set of relaxed SINR constraints in (2.7) can be always made feasible.

Note that maximizing the number of admitted users who satisfy the SINR constraints (2.5) is equivalent to minimizing the number of users who require a strictly positive value of the auxiliary variable $s_{ki}$ that satisfies constraint (2.7). Hence, by using expressions (2.4) and (2.7) the problem of admission control can be expressed as

$$\begin{align*}
\text{minimize} & \quad \sum_{k \in K} \| s_k(t) \|_0 \\
\text{subject to} & \quad \Gamma_{ki}(m(t)) \geq \gamma_{ki} - s_{ki}(t), \quad i \in U_k, k \in K \quad (2.8a) \\
& \quad \sum_{i \in U_k} \| m_{ki}(t) \|_2^2 \leq P_{k}^{\text{max}}, \quad k \in K \quad (2.8b) \\
& \quad s_{ki}(t) \geq 0, \quad i \in U_k, k \in K, \quad (2.8c)
\end{align*}$$

where $s_k(t) = [s_{k1}(t), \ldots, s_{k|U_k|}(t)]^T$ for all $k \in K$ and variables are $s_{ki}(t)$ and $m_{ki}(t)$ for all $k \in K, i \in U_k$. In problem (2.8), constraint (2.8c) is redundant\(^1\). However, this constraint is useful while deriving an algorithm to solve problem (2.8), as will be clarified in the next section.

### 2.3 Algorithm derivation

Problem (2.8) is a difficult combinatorial optimization problem [100] due to the $\ell_0$ objective function. In fact, this problem is known to be NP-hard [100, 108], and it requires exponential complexity to find a global solution. Therefore, we have to rely on suboptimal algorithms to solve it. In this section, we provide suboptimal centralized and distributed algorithms to solve problem (2.8). The

\(^1\)We can show by contradiction that negative values of variables $s_{ki}(t), i \in U_k, k \in K$ are not optimal solutions of problem (2.8).
proposed centralized algorithm is based on the $\ell_0$ approximation method [111] and sequential convex programming [23]. The distributed algorithm is derived by using consensus-based ADMM [112] in conjunction with sequential convex programming. In this section, we solve problem (2.8) for a single time slot, hence we drop the time index $t$ for notational simplicity.

### 2.3.1 Centralized algorithm

We start by approximating the objective function of problem (2.8) with a concave function

$$
\sum_{k \in K} \sum_{i \in U_k} \log (s_{ki} + \epsilon),
$$

where $\epsilon$ is a small positive constant\(^2\) and $s_{ki} \geq 0$ for all $i \in U_k$, $k \in K$ [111]. Let us now define a new variable $\beta_{ki}$ to denote the interference-plus-noise experienced by the $i$th user of the $k$th base station, i.e.,

$$
\beta_{ki} = \sum_{j \in U_k, j \neq i} |(h^k_{ki})^H m_j|^2 + \sum_{l \in K \setminus \{k\}} \sum_{j \in U_l} |(h^l_{ki})^H m_j|^2 + \sigma^2_{ki},
$$

for all $i \in U_k$, $k \in K$. Then, a solution of problem (2.8) can be approximated by solving the following optimization problem:

\[
\text{minimize} \quad \sum_{k \in K} \sum_{i \in U_k} \log (s_{ki} + \epsilon) \\
\text{subject to} \quad \gamma_{ki} - s_{ki} - \frac{|(h^k_{ki})^H m_{ki}|^2}{\beta_{ki}} \leq 0, \quad i \in U_k, k \in K \\
\sum_{j \in U_k, j \neq i} |(h^k_{ki})^H m_j|^2 + \sum_{l \in K \setminus \{k\}} \sum_{j \in U_l} |(h^l_{ki})^H m_j|^2 + \sigma^2_{ki} \leq \beta_{ki}, \quad i \in U_k, k \in K \\
\sum_{i \in U_k} \|m_{ki}(t)\|^2 \leq P_{k}^{\text{max}}, \quad k \in K \\
s_{ki}(t) \geq 0, \quad i \in U_k, k \in K
\]

with variables $\{s_{ki}, m_{ki}, \beta_{ki}\} \in \mathbb{R}$. In constraint (2.9a), we have replaced the interference-plus-noise term of $\Gamma_{ki}(m)$ with $\beta_{ki}$. Note that the objective function of problem (2.9) increases in $s_{ki}$, hence, it can be shown (e.g., by contradiction) that constraint (2.9b) holds with equality at the optimal point. The slope of the function $\log (s_{ki} + \epsilon)$ at the origin increases as $1/\epsilon$ when $\epsilon$ tends to zero. Hence, similarly to $\ell_0$, the usage of $\log(\cdot)$ function allows large penalties to be enforced on smaller nonzero coefficients, and thus, encourages the terms with smaller

\(^2\)We refer the reader to [111] for a better understanding (with illustrations) of the advantages of approximating an $\ell_0$ function with a log function, instead of using $\ell_1$-norm to approximate it.
nonzero coefficients to be set to zero [111]. Hence, this approach provides a
sparser solution compared to the $\ell_1$-norm approach.

Problem (2.9) is a non-combinatorial optimization problem. However, it is still
nonconvex because the objective function is concave and constraint function (2.9a)
is not convex. In the sequel, we apply sequential convex programming [22] to
to solve problem (2.9). Here, we approximate the objective function and constraint
function (2.9a) with their best convex approximations. Then, we iteratively
solve the approximated convex problem to find a solution for problem (2.9).

Let the objective function of problem (2.9) is denoted by

$$f(s) = \sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{U}_k} \log(s_{ki} + \epsilon),$$

where $s = [s_{11}, \ldots, s_{K_I K}].$ Note that $f(s)$ is a concave function [21, pp.
67-112]. Hence, the best convex approximation of function $f(s)$ can be
obtained by replacing it with its first order approximation [22]. The first order
approximation of function $f(s)$ near an arbitrary positive point $\hat{s} = [\hat{s}_{11}, \ldots, \hat{s}_{K_I K}]$
can be expressed as

$$\hat{f}(s) = f(\hat{s}) + \sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{U}_k} (s_{ki} - \hat{s}_{ki})/(\hat{s}_{ki} + \epsilon). \quad (2.10)$$

We now focus on constraint (2.9a). Let $g_{ki}(m_{ki}, \beta_{ki})$ be a function defined
as $g_{ki}(m_{ki}, \beta_{ki}) = |(h_{ki}^H m_{ki})^H / \beta_{ki}$ for all $i \in \mathcal{U}_k, k \in \mathcal{K}.$ It is easy to see that
constraint function (2.9a) is a difference of the affine function $\gamma_{ki} - s_{ki}$ and the
convex function $g_{ki}(m_{ki}, \beta_{ki})$ [21, pp. 67-112]. The best convex approximation
of constraint function (2.9a) near an arbitrary point $(\hat{m}_{ki}, \hat{\beta}_{ki})$ can be obtained
by replacing $g_{ki}(m_{ki}, \beta_{ki})$ with its first order approximation [22], and it can be
expressed as:

$$\hat{g}_{ki}(m_{ki}, \beta_{ki}) = g_{ki}(\hat{m}_{ki}, \hat{\beta}_{ki}) + \nabla g_{ki}(\hat{m}_{ki}, \hat{\beta}_{ki})^T ((m_{ki}, \beta_{ki}) - (\hat{m}_{ki}, \hat{\beta}_{ki})), \quad (2.11)$$

where $\nabla g_{ki}(\hat{m}_{ki}, \hat{\beta}_{ki})$ is the gradient of function $g_{ki}(m_{ki}, \beta_{ki})$ evaluated at point
$(\hat{m}_{ki}, \hat{\beta}_{ki}),$ defined as:

$$\nabla g_{ki}(\hat{m}_{ki}, \hat{\beta}_{ki}) = \left(\frac{2h_{ki}^H h_{ki}^H m_{ki}}{\beta_{ki}}, -\frac{h_{ki}^H h_{ki}^H h_{ki}^H m_{ki}}{\beta_{ki}^2} \right). \quad (2.12)$$

Now by using expressions (2.10) and (2.11), we approximate problem (2.9)
near an arbitrary positive point $(\hat{s}_{ki}, \hat{m}_{ki}, \hat{\beta}_{ki})$ for all $i \in \mathcal{U}_k, k \in \mathcal{K},$ as the

---

3Here, we use parentheses to construct a column vector from a comma separated list, e.g.,
$(a, b, c) = [a^T, b^T, c^T]^T$
following convex optimization problem:

\[
\begin{align*}
    \text{minimize} & \quad \sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{U}_k} w_{ki} s_{ki} \\
\text{subject to} & \quad \gamma_{ki} - s_{ki} - \hat{g}_{ki}(m_{ki}, \beta_{ki}) \leq 0, \quad i \in \mathcal{U}_k, k \in \mathcal{K} \\
& \quad \sum_{j \in \mathcal{U}_k, j \neq i} |(h_{ki}^j)^H m_{kj}|^2 + \sum_{i \in \mathcal{K} \setminus \{k\}} \sum_{j \in \mathcal{U}_i} |(h_{ki}^j)^H m_{kj}|^2 + \sigma^2_{ki} \leq \beta_{ki}, \quad i \in \mathcal{U}_k, k \in \mathcal{K} \\
& \quad \sum_{i \in \mathcal{U}_k} \|m_{ki}(t)\|_2^2 \leq P^\text{max}_k, \quad k \in \mathcal{K} \\
& \quad s_{ki}(t) \geq 0, \quad i \in \mathcal{U}_k, k \in \mathcal{K},
\end{align*}
\]

where \( w_{ki} = 1/(\hat{s}_{ki} + \epsilon) \), variables are \( \{s_{ki}, m_{ki}, \beta_{ki}\}_{k \in \mathcal{K}, i \in \mathcal{U}_k} \), and \( w_{ki} \) acts as a weight associated with user \( i \) of \( k \)-th base station for all \( i \in \mathcal{U}_k, k \in \mathcal{K} \). Note that in the objective function of problem (2.13) we dropped the constant term \( f(\hat{s}) - \hat{s}_{ki}/(\hat{s}_{ki} + \epsilon) \), since it does not affect the solution of problem (2.13).

Finally, we summarize the proposed centralized algorithm for problem (2.9) in Algorithm 2.1.

**Algorithm 2.1.** Centralized algorithm for solving problem (2.9)

1. Initialization: given maximum transmit power \( \{P^\text{max}_k\}_{k \in \mathcal{K}} \), SINR thresholds \( \{\gamma_{ki}\}_{i \in \mathcal{U}_k, k \in \mathcal{K}} \), initial starting points \( \{s^0_{ki}, m^0_{ki}, \beta^0_{ki}\}_{i \in \mathcal{U}_k, k \in \mathcal{K}} \), \( \{w_{ki} = 1\}_{i \in \mathcal{U}_k, k \in \mathcal{K}} \), and the parameter \( \epsilon \). Set the iteration index \( p = 0 \).
2. Set \( \hat{m}_{ki} = m^p_{ki} \) and \( \hat{\beta}_{ki} = \beta^p_{ki} \) for all \( i \in \mathcal{U}_k, k \in \mathcal{K} \). Form \( \hat{g}_{ki}(m_{ki}, \beta_{ki}) \) by using expressions (2.11) and (2.12), for all \( i \in \mathcal{U}_k, k \in \mathcal{K} \).
3. Solve problem (2.13). Denote the solution by \( s^*_i, m^*_i, \) and \( \beta^*_i \) for all \( i \in \mathcal{U}_k, k \in \mathcal{K} \).
4. Stopping criterion: if the stopping criterion is satisfied, STOP by returning the solution \( \{s^*_i, m^*_i, \beta^*_i\}_{i \in \mathcal{U}_k, k \in \mathcal{K}} \). Otherwise go to step 5.
5. Update \( s^{p+1}_{ki} = s^*_i, \beta^{p+1}_{ki} = \beta^*_i, m^{p+1} = m^*_i, \) and \( w_{ki} = 1/(s^*_i + \epsilon) \) for all \( i \in \mathcal{U}_k, k \in \mathcal{K} \). Set \( p = p + 1 \), and go to step 2.

The first step initializes the algorithm. At step (2), problem (2.9) is approximated as a convex problem (2.13) near an arbitrary point \( \{\hat{s}_{ki}, \hat{m}_{ki}, \hat{\beta}_{ki}\}_{i \in \mathcal{U}_k, k \in \mathcal{K}} \). At step (3) problem (2.13) finds an approximate solution for problem (2.9).
near an arbitrary point \( \{\hat{s}_{ki}, \hat{m}_{ki}, \hat{\beta}_{ki}\}_{i \in U_k, k \in K} \). Hence, to obtain the best local solution for problem (2.9), Algorithm 2.1 solves problem (2.13) repeatedly for different values of \( \{\hat{s}_{ki}, \hat{m}_{ki}, \hat{\beta}_{ki}\}_{i \in U_k, k \in K} \) until the stopping criteria is satisfied at step 4. The algorithm can be stopped when the difference between the achieved objective values of problem (2.9) in two successive iterations is less than a given threshold. It is worth mentioning that relying on the analytical justification provided by [111], it is easy to see that the solution method proposed in this chapter to solve problem (2.13) is equivalent to using the \( \text{reweighed } \ell_1 \)-norm minimization algorithm.

### 2.3.2 Distributed algorithm

In this subsection we extend Algorithm 2.1 to derive a distributed algorithm for admission control problem (2.8). The distributed algorithm is derived by solving step 3 of Algorithm 2.1 (i.e., problem (2.13)) distributively over the base stations using ADMM [112].

We start by introducing an auxiliary variable \( z_{ki}^l \) to denote the interference generated by \( l \)-th base station to \( i \)-th user of base station \( k \), i.e.,

\[
 z_{ki}^l = \sum_{j \in U_l, j \neq i} |(h_{ki}^l)^H m_j|^2 \text{ for all } i \in U_k, k \in K \text{ and } l \in K \setminus \{k\}.
\]

Then problem (2.13) can be equivalently written as

\[
\begin{align*}
\text{minimize} & \quad \sum_{k \in K} \sum_{i \in U_k} w_{ki} s_{ki} \\
\text{subject to} & \quad \gamma_{ki} - s_{ki} - \hat{g}_{ki}(m_{ki}, \beta_{ki}) \leq 0, \quad i \in U_k, k \in K \\
& \quad \sum_{j \in U_k, j \neq i} |(h_{ki}^l)^H m_j|^2 + \sum_{l \in K \setminus \{k\}} z_{ki}^l + \sigma_{ki}^2 \leq \beta_{ki}, \quad i \in U_k, k \in K \\
& \quad \sum_{j \in U_k} |(h_{ki}^l)^H m_j|^2 \leq z_{ki}^l, \quad i \in U_k, k \in K, \quad l \in K \setminus \{k\} \\
& \quad \sum_{i \in U_k} \|m_{ki}(t)\|^2 \leq P_{\max}^k, \quad k \in K \\
& \quad s_{ki}(t) \geq 0, \quad i \in U_k, k \in K,
\end{align*}
\]

(2.14)

with variables \( \{s_{ki}, m_{ki}, \beta_{ki}\}_{k \in K, i \in U_k} \) and \( \{z_{ki}^l\}_{i \in U_k, k \in K, l \in K \setminus \{k\}} \). Note that problems (2.13) and (2.14) are equivalent as it can be easily shown (e.g., by contradiction) that the third inequality holds with equality at the optimal point.

In problem (2.14), variable \( z_{ki}^l \) represents the power of the intercell interference caused by \( l \)-th base station to \( i \)-th user of base station \( k \), and thus variable \( z_{ki}^l \) couples the two base stations \( l \) and \( k \). We use consensus technique [112, pp. 48-56] to distribute problem (2.14) over the base stations. The method consists...
of introducing local copies of the coupling variables $z_{ki}^l$ for all $i \in U_k$, $k \in \mathcal{K}$, and $l \in \mathcal{K}\{k\}$, at each base station (see Figure 2.2).

![Diagram](diagram.png)

**Fig. 2.2.** Illustration of base station coupling, and introducing local copies to decouple a problem. BS2 and BS3 are coupled with BS1 by coupling variables $z_{ki}^2$ and $z_{ki}^3$, respectively. To distribute the problem, local copy $x_{ki,1}^2$ of $z_{ki}^2$ at BS1 and local copy $x_{ki,2}^2$ of $z_{ki}^2$ at BS2 are introduced. Similarly, local copy $x_{ki,1}^3$ of $z_{ki}^3$ at BS1 and local copy $x_{ki,3}^3$ of $z_{ki}^3$ at BS3 are introduced.

Let us define $x_{ki,k}^l$ as the local copy of $z_{ki}^l$ saved at $k$th base station, and $x_{ki,l}^l$ as the local copy of $z_{ki}^l$ saved at $l$th base station. Thus for each $z_{ki}^l$, we make two local copies (see Figure 2.2). Then, problem (2.14) can be written equivalently in *global consensus* form as

\[
\begin{align*}
\text{minimize} & \quad \sum_{k \in \mathcal{K}} \sum_{i \in U_k} w_{ki} s_{ki} \\
\text{subject to} & \quad \gamma_{ki} - s_{ki} - \hat{g}_{ki}(m_{ki}, \beta_{ki}) \leq 0, \quad i \in U_k, \quad k \in \mathcal{K} \quad (2.15a) \\
& \quad \sum_{j \in U_k, j \neq i} |(h_{ki})^\text{T} m_{kj}|^2 + \sum_{l \in \mathcal{K}\{k\}} x_{ki,k}^l + \sigma_{ki}^2 \leq \beta_{ki}, \quad i \in U_k, \quad k \in \mathcal{K} \quad (2.15b) \\
& \quad \sum_{j \in U_k} |(h_{ki})^\text{T} m_{kj}|^2 + x_{ki,k}^l \leq x_{ki,l}^j, \quad i \in U_k, \quad k \in \mathcal{K}, \quad l \in \mathcal{K}\{k\} \quad (2.15c) \\
& \quad x_{ki,k}^l = z_{ki}^l, \quad i \in U_k, \quad k \in \mathcal{K}, \quad l \in \mathcal{K}\{k\} \quad (2.15d) \\
& \quad x_{ki,l}^l = z_{ki}^l, \quad i \in U_k, \quad k \in \mathcal{K}, \quad l \in \mathcal{K}\{k\} \quad (2.15e) \\
& \quad \sum_{i \in U_k} \| m_{ki}(t) \|_2^2 \leq P_{k}^{\text{max}}, \quad k \in \mathcal{K} \quad (2.15f) \\
& \quad s_{ki}(t) \geq 0, \quad i \in U_k, \quad k \in \mathcal{K}, \quad (2.15g)
\end{align*}
\]
with variables \( \{s_{ki}, m_{ki}, \beta_{ki}\}_{k \in K, i \in U_k}, \) and \( \{x_{ki,k}^l, z_{ki}^l\}_{i \in U_k, k \in K, l \in K \setminus \{k\}} \). Note that in constraint (2.15b) we have replaced \( z_{ki}^l \) with the local copy \( x_{ki,k}^l \). In constraints (2.15c), we have replaced \( z_{ki}^l \) with the local copy \( x_{ki,l}^l \). Constraints (2.15d) and (2.15e) are called consistency constraints and they enforce local copies \( x_{ki,k}^l \) and \( x_{ki,l}^l \) to be equal to the corresponding global variable \( z_{ki}^l \).

In problem (2.15), the objective function and all the inequality constraints are separable in \( k \in K \) (one for each base station). Thus, without constraints (2.15d) and (2.15e), problem (2.15) can now be easily decoupled into \( K \) subproblems, one for each base station.

We now express problem (2.15) more compactly. To do this, we first express the consistency constraints of problem (2.15) more compactly by using vector notation. The set of local variables associated with \( k \)th base station includes two components: 1) copies of the interference experienced by all the users associated with \( k \)th base station from all the other base stations, i.e., \( \{x_{ki,k}^l\}_{l \in K \setminus \{k\}, i \in U_k} \), and 2) copies of the interference generated by \( k \)th base station to all the other users, i.e., \( \{x_{kli,k}^l\}_{l \in K \setminus \{k\}, i \in U_l} \). Thus, we can compactly write the local copies of the intercell interference terms associated with \( k \)th base station as

\[
x_k = (\{x_{ki,k}^l\}_{l \in K \setminus \{k\}, i \in U_k}, \{x_{kli,k}^l\}_{l \in K \setminus \{k\}, i \in U_l}) .
\] (2.16)

Similarly, we can compactly write the global variables associated with \( k \)th base station as

\[
z_k = (\{z_{ki}^l\}_{l \in K \setminus \{k\}, i \in U_k}, \{z_{kli}^l\}_{l \in K \setminus \{k\}, i \in U_l}) .
\] (2.17)

It is worth noting that some components of \( z_k \) are also in the components of \( z_l \) for all \( l \in K \setminus \{k\} \). For example, the components \( \{z_{ki}^l\}_{i \in U_k} \) and \( \{z_{kli}^l\}_{j \in U_l} \) are common to both \( z_k \) and \( z_l \). With the compact notations introduced above we can equivalently write the equality constraints (2.15d) and (2.15e) as

\[
x_k = z_k, \quad k \in K .
\] (2.18)
Now, for the sake of brevity, let us define the following set

\[
Q_k = \left\{ \mathbf{s}_k, \mathbf{m}_k, \beta_k, \mathbf{x}_k \right\}
\]

where

\[
\mathbf{Q}_k = \left\{ \mathbf{s}_k, \mathbf{m}_k, \beta_k, \mathbf{x}_k \right\}
\]

\[
\begin{align*}
&\gamma_{ki} - s_{ki} - \hat{g}_{ki}(\mathbf{m}_{ki}, \beta_k) \leq 0, \quad i \in \mathcal{U}_k, \ k \in \mathcal{K} \\
&\sum_{j \in \mathcal{U}_k, j \neq i} |(\mathbf{h}_{ki}^H \mathbf{m}_{kj})|^2 + \sum_{l \in \mathcal{K} \setminus \{k\}} \sum_{i \in \mathcal{U}_k} x_{ki,l}^i + \sigma_{ki}^2 \leq \beta_k, \quad i \in \mathcal{U}_k \\
&\sum_{i \in \mathcal{U}_k} |(\mathbf{h}_{ij}^k)^H \mathbf{m}_{ki}|^2 \leq x_{ij,k}^l, \quad j \in \mathcal{U}_l, \ l \in \mathcal{K} \setminus \{k\} \\
&s_{ki} \geq 0, \quad i \in \mathcal{U}_k, \ k \in \mathcal{K}
\end{align*}
\]

(2.19)

where \( \beta_k = [\beta_{k1}, \ldots, \beta_{kl_k}] \), and the following function

\[
f_k(\mathbf{s}_k, \mathbf{m}_k, \beta_k, \mathbf{x}_k) = \begin{cases} \sum_{i \in \mathcal{U}_k} w_{ki}s_{ki} & \{\mathbf{s}_k, \mathbf{m}_k, \beta_k, \mathbf{x}_k\} \in \mathbf{Q}_k \\ \infty & \text{otherwise} \end{cases}
\]

(2.20)

Then by using notations in (2.16), (2.17), (2.19), and (2.20), consensus problem (2.15) can be compactly written as

\[
\begin{align*}
\text{minimize} & \quad \sum_{k \in \mathcal{K}} f_k(\mathbf{s}_k, \mathbf{m}_k, \beta_k, \mathbf{x}_k) \\
\text{subject to} & \quad \mathbf{x}_k = \mathbf{z}_k, \quad k \in \mathcal{K},
\end{align*}
\]

(2.21)

with variables \( \mathbf{s}_k, \mathbf{m}_k, \beta_k, \mathbf{x}_k, \) and \( \mathbf{z}_k \) for all \( k \in \mathcal{K} \).

\section*{2.3.3 Distributed algorithm via ADMM}

In this subsection, we derive a distributed algorithm for problem (2.21). The proposed algorithm is based on ADMM \cite{112}. Let \( \mathbf{y}_{ki,k}^l \) and \( \mathbf{y}_{ki,l}^k \) for all \( i \in \mathcal{U}_k, \ k \in \mathcal{K}, \) and \( l \in \mathcal{K} \setminus \{k\} \), be the dual variables associated with constraints (2.15d) and (2.15e) of problem (2.15). Specifically, the dual variable associated with \( k \)th base station in constraints (2.15d) and (2.15e) can be compactly written as

\[
\mathbf{y}_k = (\{y_{ki,k}^l\}_{i \in \mathcal{K} \setminus \{k\}, i \in \mathcal{U}_k}, \{y_{ki,l}^k\}_{i \in \mathcal{K} \setminus \{k\}, i \in \mathcal{U}_l}).
\]

(2.22)

Now we can write the augmented Lagrangian for problem (2.21) as

\[
L_\rho(\{\mathbf{s}_k, \beta_k, \mathbf{x}_k, \mathbf{m}_k, \mathbf{z}_k\}_{k \in \mathcal{K}}) = \sum_{k \in \mathcal{K}} \left( f_k(\mathbf{s}_k, \beta_k, \mathbf{x}_k, \mathbf{m}_k) + \mathbf{y}_k^T (\mathbf{x}_k - \mathbf{z}_k) + \frac{\rho}{2} \| \mathbf{x}_k - \mathbf{z}_k \|^2 \right),
\]

(2.23)

where \( \{\mathbf{y}_k\}_{k \in \mathcal{K}} \) are the dual variables associated with the equality constraint of (2.21), and \( \rho > 0 \) is the penalty parameter that adds the quadratic penalty term
to the standard Lagrangian $L_0$ when the equality constraints of problem (2.21) are violated.

Each iteration of the ADMM algorithm consists of the following three steps [112, pp. 13-23]:

\[\begin{align*}
&\mathbf{s}_{q+1}^k, \mathbf{\beta}_{q+1}^k, \mathbf{x}_{q+1}^k, \mathbf{m}_{q+1}^k = \arg\min_{\mathbf{s}_k, \mathbf{\beta}_k, \mathbf{x}_k, \mathbf{m}_k} L_\rho(\mathbf{s}_k, \mathbf{\beta}_k, \mathbf{x}_k, \mathbf{m}_k, \mathbf{z}_q^k, \mathbf{y}_q^k), \quad k \in K \\
&\{\mathbf{z}_{q+1}^k\}_{k \in K} = \arg\min_{\mathbf{z}_k \in K} L_\rho(\{\mathbf{s}_{q+1}^k, \mathbf{\beta}_{q+1}^k, \mathbf{x}_{q+1}^k, \mathbf{m}_{q+1}^k\}_{k \in K}, \{\mathbf{z}_k\}_{k \in K}, \{\mathbf{y}_q^k\}_{k \in K}) \\
&\mathbf{y}_{q+1}^k = \mathbf{y}_q^k + \rho(\mathbf{x}_{q+1}^k - \mathbf{z}_{q+1}^k), \quad k \in K,
\end{align*}\]

where the superscript $q$ is the ADMM iteration counter. Note that steps (2.24a) and (2.24c) are completely decentralized, and hence, they can be carried out independently in parallel at each base station. Recall that the component of $\mathbf{z}_k$ couples with two local variables that are associated with the interfering base stations (see the consistency constraints of problem (2.15) and the definition of $\mathbf{z}_k$ in (2.17)). Thus, step (2.24b) requires gathering the updated local variables and the dual variables from all $K$ base stations. In the sequel, we first explain in detail how to solve the ADMM steps (2.24a) and (2.24b). Then, we present the proposed ADMM based distributed algorithm.

We start by providing a method to compute step (2.24a). The local variables update $(\mathbf{s}_{q+1}^k, \mathbf{\beta}_{q+1}^k, \mathbf{x}_{q+1}^k, \mathbf{m}_{q+1}^k)$ in (2.24a) is a solution of the following optimization problem:

\[
\begin{align*}
\text{minimize} \quad & f_k(\mathbf{s}_k, \mathbf{\beta}_k, \mathbf{x}_k, \mathbf{m}_k) + \mathbf{y}_q^T(\mathbf{x}_k - \mathbf{z}_q^k) + \frac{\rho}{2} \| \mathbf{x}_k - \mathbf{z}_q^k \|^2, \\
\text{with variables} \quad & \mathbf{s}_k, \mathbf{\beta}_k, \mathbf{x}_k, \mathbf{m}_k. \quad \text{Here we use } \mathbf{y}_q^T \text{ instead of } (\mathbf{y}_q^T)^T \text{ to lighten the notation. Let } \mathbf{v}_k \text{ be a scaled dual variable, i.e., } \mathbf{v}_k = (1/\rho)\mathbf{y}_k. \text{ By using the} \\
\end{align*}
\]
notations (2.19) and (2.20), problem (2.25) can be equivalently written as

$$\begin{align*}
\text{minimize} & \quad \sum_{i \in \mathcal{U}_k} w_{ki} s_{ki} + \frac{\rho}{2} \| x_k - z_k^q + v_k^q \|_2^2 \\
\text{subject to} & \quad \gamma_{ki} - s_{ki} - \tilde{g}_{ki}(m_{ki}, \beta_{ki}) \leq 0, \quad i \in \mathcal{U}_k \\
& \quad \sum_{j \in \mathcal{U}_k, j \neq i} |(h^k)_{ij}^H m_{kj}|^2 + \sum_{l \in \mathcal{K} \setminus \{i\}} x_{kl,k} + \sigma_{ki}^2 \leq \beta_{ki}, \quad i \in \mathcal{U}_k \\
& \quad \sum_{s \in \mathcal{U}_k} \| m_{ks} \|_2^2 \leq \rho_{k}^\max, \\
& \quad s_{ki} \geq 0, \quad i \in \mathcal{U}_k,
\end{align*}$$

(2.26)

with variables $s_k$, $\beta_k$, $x_k$, and $m_k$ (see variable definitions given in (2.16) and (2.17)). Note that the second term of the objective function of problem (2.26) is obtained by: 1) combining the linear and quadratic terms of the objective function of problem (2.25) as $y_k^q (x_k - z_k^q) + \frac{\rho}{2} \| x_k - z_k^q \|_2^2 = \frac{\rho}{2} \| x_k - z_k^q + v_k^q \|_2^2 - \frac{\rho}{2} \| v_k^q \|_2^2$, and then 2) dropping the constant term $\frac{\rho}{2} \| v_k^q \|_2^2$. This constant term is dropped since it does not affect the solution of the problem. Let us denote $s_k^*, \beta_k^*, z_k^*$, and $m_k^*$ as a solution of problem (2.26) and we denote $s_k^{q+1} = s_k^*$, $\beta_k^{q+1} = \beta_k^*$, $x_k^{q+1} = x_k^*$, and $m_k^{q+1} = m_k^*$ as the updates.

We now consider the second step of the ADMM algorithm, i.e., (2.24b), and provide a solution for the global variable update. The update $\{z_k^{q+1}\}_{k \in \mathcal{K}}$ is a solution of the following optimization problem:

$$\begin{align*}
\text{minimize} & \quad \sum_{k \in \mathcal{K}} y_k^q (x_k^{q+1} - z_k) + \frac{\rho}{2} \| x_k^{q+1} - z_k \|_2^2, \\
& \quad \text{subject to} \quad \sum_{i \in \mathcal{U}_k} \sum_{i \in \mathcal{U}_k} \left( y_{ki,k}^q (x_{ki,k}^{q+1} - z_{ki,k}^q) + \frac{\rho}{2} (x_{ki,k}^{q+1} - z_{ki,k}^q)^2 \right) + \\
& \quad \sum_{i \in \mathcal{U}_k} \sum_{i \in \mathcal{U}_k} \left( b_{ki,k}^q (x_{ki,k}^{q+1} - z_{ki,k}^q) + \frac{\rho}{2} (x_{ki,k}^{q+1} - z_{ki,k}^q)^2 \right),
\end{align*}$$

(2.27)

with variables $\{z_k^i\}_{i \in \mathcal{U}_k, k \in \mathcal{K}, i \in \mathcal{K} \setminus \{k\}}$. Note that problem (2.28) is an expanded version of problem (2.27), in which we have substituted elements for each vector using (2.16), (2.17), and (2.22). Here, the square brackets in superscript represent the ADMM iteration index.

Problem (2.28) decouples across $z_{ki}^q$, since the objective function is separable in $\{z_{ki}^q\}_{i \in \mathcal{U}_k, k \in \mathcal{K}, i \in \mathcal{K} \setminus \{k\}}$. Moreover, the objective function of problem (2.28) is
Algorithm 2.2. Distributed algorithm for admission control problem (2.9)

1. Initialization: given maximum transmit power \( \{P^\text{max}_k\}_{k \in \mathcal{K}} \), SINR thresholds \( \{\gamma_k\}_{i \in \mathcal{U}_k, k \in \mathcal{K}} \), initial starting points \( \{x^0_{ki}, \mathbf{m}^0_k, \beta^0_{k}\}_{i \in \mathcal{U}_k, k \in \mathcal{K}} \), \( \{w_{ki} = 1\}_{i \in \mathcal{U}_k, k \in \mathcal{K}} \), and the parameters \( \epsilon \) and \( \rho \). Set the iteration indexes \( q = 0 \).

2. By setting \( \mathbf{m}^q_{ki} = \mathbf{m}^0_{ki} \) and \( \beta^q_{ki} = \beta^0_{k} \), form \( \hat{g}_{ki}(\mathbf{m}^q_{ki}, \beta^q_{ki}) \) for all \( i \in \mathcal{U}_k, k \in \mathcal{K} \) and by using expressions (2.11) and (2.12).

3. ADMM algorithm:
   a) each base station \( k \in \mathcal{K} \) updates the local variables \( \left( s^{q+1}_k, \beta^{q+1}_k, x^{q+1}_k, \mathbf{m}^{q+1}_k \right) \) by solving (2.26).
b) base station $k$ and $l$ exchange their local copies $x_{ki,l}^l$ and $x_{ki,k}^l$ for all $i \in U_k$, $k \in K$, and $l \in K \setminus \{k\}$.

c) each BS $k$ updates global variable $z_{ki}^{q+1}$ using (2.31).

d) each base station $k$ updates dual variable $y_{ki}^{q+1}$ by solving (2.24c).

e) ADMM stopping criterion: if the stopping criterion is satisfied, go to step 4. Otherwise, set $q = q + 1$, and go to step 3.a.

4. Stopping criterion: if the stopping criterion is satisfied STOP. Otherwise, set update \(\{w_{ki} = 1/(s_{ki}^{q+1} + \epsilon)\}\) for all $i \in U_k$, set $q = q + 1$, and go to step 2.

We refer to each execution of steps 2-4 as an outer iteration and to each execution of ADMM algorithm (i.e., step 3) as an inner iteration. In each outer iteration, at step 2, base station $k$ evaluates the expression $\hat{g}_{ki}(m_{ki}, \beta_{ki})$ for all $k \in K$ individually, thus decoupling step 2 over the base stations. In inner iteration, steps 3.a, 3.c, and 3.d are decentralized over the base stations. Step 3.b requires a coordination between the base stations to exchange the updated values of the local variables and this step can be done without a central controller. Step 3.e checks the ADMM stopping criteria. In ADMM algorithm, the standard stopping criteria is to check the primal and dual residuals [112, pp. 18-20]. But, as ADMM can produce acceptable results of practical use within a few iterations, a finite number of iterations is used to stop the ADMM algorithm in step 3.e [112, pp. 18-20]. Step 4 checks the algorithm stopping criterion.

In many practical applications, we have to stop the distributed algorithm after a finite number of iterations before converging the algorithm. On the other hand, an intermediate solution provided by Algorithm 2.2 at each inner iteration does not necessarily result in a feasible solution for problem (2.15). That is, constraints (2.15d) and (2.15e) of problem (2.15) may be violated by an intermediate solution provided by Algorithm 2.2. However, at the cost of solving one additional subproblem by each base station, we can find a feasible solution for problem (2.15). To do this, we make the local copies $x_{ki,k}^l$ and $x_{ki,l}^l$ to be equal, by setting them to be equal to the consensus value $z_{ki}^l$, i.e., $x_{ki,k}^l = z_{ki}^l$ and $x_{ki,l}^l = z_{ki}^l$ for all $i \in U_k$, $l \in K \setminus \{k\}$, $k \in K$. Then, each base station can solve problem (2.26) independently for variables $s_k$, $\beta_k$, and $m_k$ to find a feasible solution for problem (2.15).
2.3.4 The signaling overhead of Algorithm 2.1 and Algorithm 2.2

The proposed Algorithm 2.1 is a centralized algorithm. Hence, to solve problem (2.13) in step 3 of the algorithm, it is required to collect the CSI of all the users in the network at a central controller (CC). That is, \( k \)th base station needs to send \( \{ h_{ki}^k \}_{i \in U_k} \) and \( \{ h_{lj}^k \}_{j \in U_l \setminus \{ k \}} \) to the CC. Let \( U \) denote the total number of users in the network, i.e., \( U = \sum_{k \in K} |U_k| \). Then, the number of real scalars that has to be sent from \( K \) base stations to the CC is \( 2TKU \), where the factor 2 is considered to take into account the real and complex parts of a channel coefficient. Next, the CC needs to send the computed beamformers of the admitted users to their respective base stations. This requires \( 2T \sum_{k \in K} |\tilde{U}_k| \) scalars to be sent to \( K \) base stations. Thus, the total number of real scalars required to be exchanged between the base stations and the CC in Algorithm 2.1 is \( 2T(KU + \sum_{k \in K} |U_k|) \). Following a similar approach, we can show that the algorithm proposed in [83] requires the same amount of signaling overhead if it is applied in a multicell network. As the exhaustive search algorithm is also run in a CC, its signaling overhead is the same as that of centralized Algorithm 2.1.

In the case of distributed Algorithm 2.2, coordination between the base stations is required at step 3.b, only to exchange auxiliary variables that are used to achieve the consensus on the out-of-cell interference values. Specifically, \( k \)th base station sends \( \{ x_{ki,k}^l \}_{i \in U_k \setminus \{ k \}, i \in U_k} \) and \( \{ x_{li,k}^l \}_{l \in K \setminus \{ k \}, i \in U_l} \) to BS \( l \) for all \( l \in K \setminus \{ k \} \). Thus, the number of scalars that \( k \)th base station has to send to all the other base stations per iteration is \( \omega_k = (K - 1) |U_k| + \sum_{l \in K \setminus \{ k \}} |U_l| \). Hence, the total number of scalars that \( K \) base stations exchange per iteration is \( \sum_{k \in K} \omega_k = 2(K - 1)U \). The signaling overhead of the algorithm in [85] per iteration is \( U \). The additional signaling overhead of the proposed Algorithm 2.2 is worth, as our algorithm provides a better objective value compared to the algorithm in [85], which is illustrated in the Section 2.6.1.

2.3.5 Complexity of Algorithm 2.1 and Algorithm 2.2

Algorithm 2.1 and Algorithm 2.2 are iterative algorithms, and both of them solve convex problems at each iteration (i.e., problem (2.13) is solved in Algorithm 2.1 and problem (2.26) is solved in Algorithm 2.2). Thus, these problems can be efficiently solved by using the interior-point method that relies on Newton’s
method [21, pp. 561-615]. It is easy to see that problem (2.13) consists of $U(T + 2)$ variables and $3U + K$ inequality constraints. Hence, the complexity of Algorithm 2.1 per iteration is $O\left(\sqrt{3U + K}U^3(T + 2)^3 \log(1/\epsilon)\right)$, where $\epsilon$ is the required relative accuracy of the duality gap at termination [21, pp. 561-615]. Furthermore, when $T$ and $K$ are much smaller than $U$, the complexity of Algorithm 2.1 per iteration can be approximated as $O\left(U^{3.5} \log(1/\epsilon)\right)$. For a multicell scenario as considered in this chapter, the complexity of the algorithm in [83] per iteration is the same as Algorithm 2.1. To get an insight into the overall complexity of Algorithm 2.1 and the one in [83], we compare the average execution times of these two algorithms in the numerical results in Section 2.6.1.

In Algorithm 2.2, each base station solves problem (2.26), which consists of $|U_k|(K + T) + U$ variables and $2|U_k| + U + 1$ inequality constraints. Hence, the complexity of solving problem (2.26) at $k$th BS is $O\left(\sqrt{2|U_k| + U + 1}\left(|U_k|(K + T) + U\right)^3 \log(1/\epsilon)\right)$. By following a similar approach (i.e., by counting the total number of variables and the inequality constraints), the complexity of solving the BS specific problem in [85] can be expressed as $O\left(\sqrt{|U_k| + 3U + 1}\left(2U + T^2|U_k|\right)^3 \log(1/\epsilon)\right)$. To compare the complexities of these two algorithms, assume that the number of users associated with each BS is equal, i.e., let $|U_k| = \tilde{U}$ for all $k \in K$. Then, the complexity of solving a base station specific problem in Algorithm 2.2 can be expressed as $O\left(\sqrt{(K+2)\tilde{U} + 1}\left(2K + T\right)^3 \tilde{U}^3 \log(1/\epsilon)\right)$ and that in [85] is $O\left(\sqrt{(3K+1)\tilde{U} + 1}\left(2K + T^2\right)^3 \tilde{U}^3 \log(1/\epsilon)\right)$. Thus, it is easy to see that Algorithm 2.2 has a lower level of complexity than the one in [85] for solving a base station specific problem.

### 2.3.6 Convergence of Algorithm 2.1 and Algorithm 2.2

Recall that problem (2.8) is an NP-hard problem. Hence, the optimal solution for problem (2.8) can only be obtained by using some sort of exhaustive search methods such as branch and bound [20]. In practice, the quality of the solution achieved within the first few iterations is a more important performance criterion than the asymptotic results, because there is only time to perform a finite
(usually small) number of iterations. The goal of this chapter is to provide polynomial-time algorithms that find suboptimal solutions for problem (2.8). Therefore, the convergence of our proposed polynomial-time algorithms to the optimal value of problem (2.8) cannot be established. Nonetheless, a weaker form of convergence of Algorithm 2.1 and Algorithm 2.2 can be established.

Note that Algorithm 2.1 obtains an approximate solution of problem (2.8) by solving problem (2.9), which is a difference of convex problem. Hence, it can be shown that the sequence of objective values generated by Algorithm 1 is nonincreasing [22]. Moreover, the optimal objective value of problem (2.9) is bounded. Thus, the monotonic convergence of Algorithm 2.1 can be guaranteed.

To establish the convergence of Algorithm 2.2, we first note that the steps of Algorithm 2.2 are similar to those of centralized Algorithm 2.1, except step 3 of both algorithms. At step 3 of Algorithm 2.1, it solves problem (2.13) in a central controller. However, in Algorithm 2.2 the same problem is solved distributively using ADMM algorithm. Note that problem (2.13) is convex, and hence, by following the approach in [113, Prop. 4.2], [112], we can show that the ADMM algorithm converges. Since, Algorithm 2.2 inherits the properties of Algorithm 2.1, a sequence of the objective values of problem (2.13) which are produced upon ADMM convergence are guaranteed to be convergent.

2.4 Fair admission control

In Section 2.3, we proposed a centralized and a distributed algorithm to find a set of admissible users during a single time slot. We know that the provided solution cannot always admit all the users within a single time slot due to limited resources. Hence some users remain unserved and thus fairness among the users has to be considered as well.

In this section, we first consider the problem of proportional fair admission control. We model this problem as a stochastic optimization problem, and propose a dynamic control algorithm via the Lyapunov optimization method [110] to solve it. Finally, we modify this algorithm to obtain max-min fairness.

Recall from section 2.3 that \( i \)th user of base station \( k \) is admitted to the network if \( s_{ki} = 0 \), i.e., when \( \Gamma_{ki}(m(t)) \geq \gamma_{ki} \) (see problem (2.8)). Based on this observation, we define an indicator function \( a_{ki}(t) \) for all \( i \in \mathcal{U}_k, \ k \in \mathcal{K}, \) to
model the admission status of $i$th user of base station $k$ in time slot $t$ as

$$a_{ki}(t) = \begin{cases} 1 & s_{ki}(t) = 0 \\ 0 & s_{ki}(t) > 0. \end{cases}$$

(2.32)

It is easy to see from (2.32) that admission status $a_{ki}(t) = 1$ means $\Gamma_{ki}(\mathbf{m}(t)) \geq \gamma_{ki}$.

We follow the approach in [110, pp. 29-129] to formulate the problem of proportional fair admission control. The idea is to proportionally maximize the time average admissions for each user, subject to the SINR targets of each user and the power constraints of each base station. To formulate this problem, let us define the time average expectation of $a_{ki}(t)$ over $t$ time slots as

$$\bar{a}_{ki}(t) \triangleq \frac{1}{t} \sum_{\tau=0}^{t-1} E\{a_{ki}(\tau)\},$$

(2.33)

where the expectation depends on the control policy and is with respect to the random channel states and the control actions made in reaction to these channel states [110, 114] $^5$. Then, by following the approach in [110, pp. 97-129],[115] the problem of proportional fair admission control can be expressed as the following optimization problem:

$$\text{maximize} \quad \lim_{t \to \infty} \inf \sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{U}_k} \log(\bar{a}_{ki}(t))$$

subject to

$$\Gamma_{ki}(\mathbf{m}(t)) \geq \gamma_{ki} - s_{ki}(t), \quad i \in \mathcal{U}_k, \quad k \in \mathcal{K}, \quad \forall t \quad (2.34a)$$

$$\sum_{i \in \mathcal{U}_k} \|\mathbf{m}_{ki}(t)\|_2^2 \leq P_k^{\max}, \quad k \in \mathcal{K}, \quad \forall t \quad (2.34b)$$

$$s_{ki}(t) \geq 0, \quad i \in \mathcal{U}_k, \quad k \in \mathcal{K}, \quad \forall t, \quad (2.34c)$$

with variables $\{s_{ki}(t), \mathbf{m}_{ki}(t)\}_{k \in \mathcal{K}, i \in \mathcal{U}_k}$ for all $t \in \{0, 1, \ldots \}$.

### 2.4.1 Transformed problem via auxiliary variables

In this section we use the Lyapunov optimization method [110, 114] to find a solution of problem (2.34). To do this we transform problem (2.34) so that it conforms to the structure required by the drift-plus-penalty minimization method in [110, 114].

$^5$Note that we use a commonly used procedure and express the long term average as the time average of expectation, which leads to a tractable algorithm [110, 114-117]
We start by equivalently reformulating problem (2.34) by introducing an auxiliary variable $\bar{\mu}_{ki}$ for all $i \in U_k$, $k \in K$, as follows:

\[
\begin{align*}
\text{maximize} & \quad \sum_{k \in K} \sum_{i \in U_k} \log(\bar{\mu}_{ki}) \\
\text{subject to} & \quad \bar{\mu}_{ki} \leq \liminf_{t \to \infty} \bar{a}_{ki}(t), \quad i \in U_k, \ k \in K \\
& \quad \Gamma_{ki}(\mathbf{m}(t)) \geq \gamma_{ki} - s_{ki}(t), \quad i \in U_k, \ k \in K, \ \forall t \\
& \quad \sum_{i \in U_k} \| \mathbf{m}_{ki}(t) \|_2^2 \leq P_k^{\text{max}}, \quad k \in K, \ \forall t \\
& \quad s_{ki}(t) \geq 0, \quad i \in U_k, \ k \in K, \ \forall t,
\end{align*}
\]

with variables $\{\bar{\mu}_{ki}\}_{k \in K, i \in U_k}$ and $\{s_{ki}(t), \mathbf{m}_{ki}(t)\}_{k \in K, i \in U_k}$ for all $t \in \{0, 1, \ldots\}$.

Note that the first set of inequality constraints of problem (2.35) holds with equality at the optimal solution due to the monotonic increasing property of the objective function.

Note that problem (2.35) involves the statistics of random channel states, and that we do not even know these. Our goal is to provide a dynamic control algorithm that makes a decision in each time slot and solves problem (2.35). This we achieve by using the drift-plus-penalty minimization method [110, 114], which converts a long term objective of problem (2.35) into a series of myopic optimizations. To do this we assume that the auxiliary variable $\bar{\mu}_{ki}$ is a time average of auxiliary variable $\mu_{ki}(t)$ for all $t = \{0, 1, \cdots\}$, i.e., $\bar{\mu}_{ki} = \lim_{t \to \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} E\{\mu_{ki}(\tau)\}$. Then, by following the approach in [110, pp. 99-100], we modify problem (2.35) as the following optimization problem:

\[
\begin{align*}
\text{maximize} & \quad \sum_{k \in K} \sum_{i \in U_k} \log(\mu_{ki}) \\
\text{subject to} & \quad \mu_{ki} \leq \liminf_{t \to \infty} a_{ki}(t), \quad i \in U_k, \ k \in K \\
& \quad \Gamma_{ki}(\mathbf{m}(t)) \geq \gamma_{ki} - s_{ki}(t), \quad i \in U_k, \ k \in K, \ \forall t \\
& \quad \sum_{i \in U_k} \| \mathbf{m}_{ki}(t) \|_2^2 \leq P_k^{\text{max}}, \quad k \in K, \ \forall t \\
& \quad s_{ki}(t) \geq 0, \quad i \in U_k, \ k \in K, \ \forall t,
\end{align*}
\]

with variables $\{s_{ki}(t), \mathbf{m}_{ki}(t), \mu_{ki}(t)\}_{k \in K, i \in U_k}$ for all $t \in \{0, 1, \ldots\}$, where $\log(\mu_{ki})$ is defined as

\[
\log(\mu_{ki}) = \lim_{t \to \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} E\{\log(\mu_{ki}(\tau))\}.
\]

Note that by using Jensens inequality we can easily verify that $\log(\mu_{ki})$ is a lower bound on $\log(\bar{\mu}_{ki})$. Thus, a solution of problem (2.36) is also feasible for the original problem (2.34), and hence problem (2.36) provides a reasonable lower bound for the original problem (2.34).
2.4.2 Solving the transformed problem

In this subsection we apply the drift-plus-penalty minimization method introduced in [110] to solve problem (2.36). In particular to this method, the time average inequality constraint (2.36a) is enforced by transforming problem (2.36) into a queue stability problem. In other words, for each time average inequality constraint in (2.36a), a virtual queue is associated with each constraint; the stability of these virtual queues implies the feasibility of constraint (2.36a).

Let \( \{\bar{G}_{ki}(t)\}_{k \in K, i \in U_k} \) be the virtual queues associated with constraint (2.36a).

We update virtual queue \( \bar{G}_{ki}(t + 1) \) for all \( k \in K, i \in U_k \), in every time slot as

\[
\bar{G}_{ki}(t + 1) = \max\{\bar{G}_{ki}(t) + \mu_{ki}(t) - a_{ki}(t), 0\}, \tag{2.38}
\]

Note that \( \bar{G}_{ki}(t) \) can be viewed as a backlog in a virtual queue with an input process \( \mu_{ki}(t) \) and a service process \( a_{ki}(t) \). Here, we adopt the notion of strong stability \(^6\), and we state that the queues are strongly stable if [110, pp. 15-25]

\[
\bar{G}_{ki} \triangleq \limsup_{t \to \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} E\{\bar{G}_{ki}(t)\} < \infty, \quad k \in K, \; i \in U_k. \tag{2.39}
\]

Intuitively, (2.39) means that a queue is strongly stable if its time average backlog is finite.

We next define the Lyapunov function and its drift, which will be used to define the queue stability problem. Let \( \bar{G}(t) \) be a vector of the virtual queues, i.e., \( \bar{G}(t) = [\bar{G}_{11}(t), \ldots, \bar{G}_{KI_k}(t)]^T \), then a quadratic Lyapunov function \( L(\bar{G}(t)) \) can be defined as [110, pp. 29-43]

\[
L(\bar{G}(t)) = \frac{1}{2} \sum_{k \in K} \sum_{i \in U_k} \bar{G}_{ki}(t)^2. \tag{2.40}
\]

Note that if \( L(\bar{G}(t)) \) is small then all queues are small, and if \( L(\bar{G}(t)) \) is large then at least one queue is large. Hence, by minimizing the drift in the Lyapunov function (i.e., by minimizing a difference in the Lyapunov function from one slot to the next), queues \( \{\bar{G}_{ki}(t)\}_{k \in K, i \in U_k} \) can be stabilized [110, pp. 45-77]. Furthermore, by using expression (2.40), the drift in the Lyapunov function from one slot to the next can be defined as

\[
\Delta(\bar{G}(t)) = E\{L(\bar{G}(t + 1)) - L(\bar{G}(t))|\bar{G}(t)\}. \tag{2.41}
\]

\(^6\)A definition of strong stability is general, and it also implies other forms of stabilities [110, Th. 2.8]
Now we use the drift-plus-penalty minimization method \[110\] to solve problem (2.36). In this method, a control policy that solves problem (2.36) is obtained by minimizing an upper bound on the following drift-plus penalty expression \[110\]:

\[
\Delta(G(t)) - V \sum_{k \in K} \sum_{i \in U_k} E\{\log(\mu_{ki}(t)) \mid G(t)\},
\]

where \( V \geq 0 \), subject to the constraints (2.36b)-(2.36d) in each time slot, i.e.,

\[
\Gamma_{ki}(m(t)) \geq \gamma_{ki} - s_{ki}(t), \quad i \in U_k, \ k \in K 
\]

\[
\sum_{i \in U_k} \| m_{ki}(t) \|^2 \leq P_k^{\text{max}}, \quad k \in K 
\]

\[
s_{ki}(t) \geq 0, \quad i \in U_k, \ k \in K, 
\]

(2.43a) (2.43b) (2.43c)

Note that the expression (2.42) has two terms. The first term \( \Delta(G(t)) \) is the drift; by minimizing this we ensure the inequality constraint (2.36a) \[110\], pp. 45-77. Furthermore, by minimizing the second term \( -\sum_{k \in K} \sum_{i \in U_k} E\{\log(\mu_{ki}(t)) \mid G(t)\} \), the objective function of problem (2.36) is maximized.

In the rest of this section, to simplify algorithm development, we first determine an upper bound of expression (2.42). Then, we present a dynamic control algorithm to solve problem (2.36) that, at each time slot, minimizes an upper bound of expression (2.42) subject to the constraints (2.43a)-(2.43c).

To obtain an upper bound on the objective function of problem (2.42), using expression (2.38), we first note that \( G_{ki}(t+1)^2 \) can be bounded as

\[
G_{ki}(t+1)^2 \leq G_{ki}(t)^2 + a_{ki}(t)^2 + \mu_{ki}(t)^2 + 2G_{ki}(t)[\mu_{ki}(t) - a_{ki}(t)].
\]

(2.44)

In order to find expression (2.44), we have used the fact that \((\max[G + \mu - a, 0])^2 \leq G^2 + \mu^2 + a^2 + 2G(\mu - a)\) for any \( G \geq 0, \mu \geq 0, \) and \( a \geq 0 \). Then, by using expression (2.41) and inequality (2.44), an upper bound on the expression (2.42) can be expressed as

\[
\Delta(G(t)) - V \sum_{k \in K} \sum_{i \in U_k} E\{\log(\mu_{ki}(t)) \mid G(t)\} \leq D - V \sum_{k \in K} \sum_{i \in U_k} E\{\log(\mu_{ki}(t)) \mid G(t)\}
\]

\[
+ \sum_{k \in K} \sum_{i \in U_k} G_{ki}(t)E\{\mu_{ki}(t) \mid G(t)\} - \sum_{k \in K} \sum_{i \in U_k} G_{ki}(t)E\{a_{ki}(t) \mid G(t)\},
\]

(2.45)

where \( D \) is a finite positive constant that satisfies the following condition for all \( t \):

\[
D \geq \frac{1}{2} \left[ \sum_{k \in K} \sum_{i \in U_k} E\{\mu_{ki}(t)^2 \mid G(t)\} + \sum_{k \in K} \sum_{i \in U_k} E\{a_{ki}(t)^2 \mid G(t)\} \right].
\]

(2.46)
Note that the right hand side of (2.45) consists of cross terms between \( G_{ki}(t) \) and \( \mu_{ki}(t) \) as well as \( G_{ki}(t) \) and \( a_{ki}(t) \) for all \( i \in U_k \) and \( k \in K \). Furthermore, the value of \( G_{ki}(t) \) can be interpreted as the scheduling priority that should be given to \( i \)th user of base station \( k \). Thus, by minimizing the right hand side of (2.45) more emphasis is put on the users that have been served least.

We now summarize the proposed proportional fair admission control algorithm to solve problem (2.36) in Algorithm 2.3. The proposed algorithm observes the virtual queues \( \{G_{ki}(t)\}_{k \in K, i \in U_k} \) and the channel states \( \{h'_{ki}(t)\}_{k \in K, i \in U_k} \) in every time slot \( t \), and it takes a control action to minimize the right hand side of expression (2.45) subject to constraints (2.43a)-(2.43c). The minimization of the right hand side of expression (2.45) can be decoupled across variables \( \{\mu_{ki}(t)\}_{k \in K, i \in U_k} \) and \( \{s_{ki}(t), m_{ki}(t)\}_{k \in K, i \in U_k} \) for all \( t \in \{0, 1, \ldots \} \), resulting in subproblems as shown in Algorithm 2.3. Note that the drift-plus-penalty minimization method uses the concept of opportunistically minimizing an expectation [110, Ch 1.8] to solve the subproblems.

---

**Algorithm 2.3. Proportional fair dynamic admission control algorithm for problem (2.36)**

1. Auxiliary variable: the auxiliary variables \( \{\mu_{ki}(t)\}_{k \in K, i \in U_k} \) are obtained by solving the following optimization problem:

   \[
   \begin{align*}
   \text{maximize} & \quad \sum_{k \in K} \sum_{i \in U_k} \left( V \log(\mu_{ki}(t)) - G_{ki}(t)\mu_{ki}(t) \right) \\
   \text{subject to} & \quad 0 \leq \mu_{ki}(t) \leq \mu_{\text{max}}, \quad i \in U_k, \quad k \in K,
   \end{align*}
   \]  

   (2.47)

   with variables \( \{\mu_{ki}(t)\}_{k \in K, i \in U_k} \) and \( \mu_{\text{max}} > 0 \) is an algorithm parameter as described in [110, Ch. 5].

2. Resource allocation: the transmit beamformer \( m_{ki}(t) \) and the variable \( s_{ki}(t) \) for all \( k \in K, i \in U_k \) are obtained by solving the following optimization problem:

   \[
   \begin{align*}
   \text{maximize} & \quad \sum_{k \in K} \sum_{i \in U_k} G_{ki}(t)a_{ki}(t) \\
   \text{subject to} & \quad \Gamma_{ki}(m(t)) \geq \gamma_{ki} - s_{ki}(t), \quad i \in U_k, \quad k \in K, \\
   & \quad \sum_{i \in U_k} \| m_{ki}(t) \|_2^2 \leq P_{ki}^{\text{max}}, \quad k \in K, \\
   & \quad s_{ki}(t) \geq 0, \quad i \in U_k, \quad k \in K,
   \end{align*}
   \]  

   (2.48a)-(2.48c)

   with variables \( \{s_{ki}(t), m_{ki}(t)\}_{k \in K, i \in U_k} \).
3. Queue update: update \( \{G_{ki}(t)\}_{k \in K, i \in U_k} \) using (2.38). Set \( t = t + 1 \), and go to step (1).

First step updates the auxiliary variables. Specifically, in problem (2.47) the objective function is maximized when the derivative with respect to \( \mu_{ki}(t) \) is zero, i.e., \( \mu^*_{ki}(t) = \min(V/G_{ki}(t), \mu_{\text{max}}) \) for all \( k \in K, i \in U_k \). The parameter \( \mu_{\text{max}} \) has to be set large enough such that \( \mu_{\text{max}} \) does not affect the optimal solution of problem (2.34), i.e., \( \mu_{\text{max}} \geq \liminf_{t \to \infty} \bar{a}_{ki}(t) \). Step 2 solves the resource allocation problem (2.48). A solution of this is found by adopting Algorithm 2.1, and is explained in subsection 2.4.3. Step 3 updates the virtual queues.

### 2.4.3 Solving the resource allocation problem of the dynamic control algorithm

In this subsection, we show how Algorithm 2.1 can be adopted to find a solution of problem (2.48). To do this, we first compactly write expression (2.32) as [100]:

\[
a_{ki}(t) = 1 - \| s_{ki}(t) \|_0. \tag{2.49}
\]

By substituting expression (2.49) into the objective function of problem (2.48), we simplify it as:

\[
\sum_{k \in K} \sum_{i \in U_k} G_{ki}(t)a_{ki}(t) = \sum_{k \in K} \sum_{i \in U_k} G_{ki}(t) - \sum_{k \in K} \sum_{i \in U_k} G_{ki}(t) \| s_{ki}(t) \|_0. \tag{2.50}
\]

Then, by using expression (2.50) as the objective function, we can equivalently reformulate problem (2.48) as follows:

\[
\begin{align*}
\text{minimize} & \quad \sum_{k \in K} \sum_{i \in U_k} G_{ki}(t) \| s_{ki}(t) \|_0 \\
\text{subject to} & \quad \Gamma_{ki}(m(t)) \geq \gamma_{ki} - s_{ki}(t), \quad i \in U_k, \ k \in K \tag{2.51a} \\
& \quad \sum_{i \in U_k} \| m_{ki}(t) \|^2_2 \leq P_{\text{max}}^k, \quad k \in K \tag{2.51b} \\
& \quad s_{ki}(t) \geq 0, \quad i \in U_k, \ k \in K. \tag{2.51c}
\end{align*}
\]

with variables \( \{s_{ki}(t), m_{ki}(t)\}_{k \in K, i \in U_k} \). Note that in the objective function of problem (2.51) we have removed the constant term \( \sum_{k \in K} \sum_{i \in U_k} G_{ki}(t) \), since it does not affect the optimal solution. Furthermore, by changing the sign of the objective function we have converted it to a minimization problem. The
objective function of problem (2.51) is combinatorial, and constraint (2.51a) is nonconvex; therefore, this problem is a difficult optimization problem.

To solve problem (2.51), we follow a similar approach that we have carried out in subsection 2.3.1. That is, we approximate the objective function of problem (2.51) and the constraint function (2.51a) with their best convex approximations. Using the aforementioned technique, a solution of problem (2.51) can be approximated near an arbitrary positive point \((\hat{s}_{ki}, \hat{m}_{ki}, \hat{\beta}_{ki})\) for all \(i \in U_k, k \in K\), by solving the following optimization problem:

\[
\begin{align*}
\text{minimize} & \quad \sum_{k \in K} \sum_{i \in U_k} \tilde{w}_{ki}(t) s_{ki}(t) \\
\text{subject to} & \quad \gamma_{ki} - s_{ki}(t) - \hat{g}_{ki}(m_{ki}(t), \beta_{ki}(t)) \leq 0, \quad i \in U_k, k \in K \\
& \quad \sum_{j \in U_k, j \neq i} |(h^t_k(t))^H m_{kj}(t)|^2 + \sum_{l \in K \setminus \{k\}} \sum_{j \in U_l} |(h^t_l(t))^H m_{kj}(t)|^2 + \sigma^2_{ki} \leq \beta_{ki}(t), \quad i \in U_k, k \in K \\
& \quad s_{ki}(t) \geq 0, \quad i \in U_k, k \in K
\end{align*}
\]

where \(\tilde{w}_{ki}(t) = \frac{G_{ki}(t)}{(\hat{s}_{ki}(t) + \epsilon)}\), the variables are \(\{s_{ki}(t), m_{ki}(t), \beta_{ki}(t)\}_{k \in K, i \in U_k}\), and \(\hat{g}_{ki}(m_{ki}, \beta_{ki})\) is as defined in expression (2.11). Here, the weight \(\tilde{w}_{ki}(t)\) of each user is associated with the virtual queue of that user (compare with weight \(w_{ki}(t)\) in problem (2.13)). We can see that problem (2.52) is similar to problem (2.13), and hence we use the proposed Algorithm 2.1 to find a solution of problem (2.48).

### 2.4.4 Max-Min fair admission control

In this subsection, we briefly discuss the method for providing max-min fairness while users are being admitted to the network. This notion of fairness can be achieved by following similar steps that we have presented in subsection 2.4.2.

The problem of max-min fair admission control can be expressed, by modifying the objective function of problem (2.34), as the following optimization problem [110, pp. 97-129]:

\[
\begin{align*}
\text{maximize} & \quad \min_{k \in K, i \in U_k} \lim_{t \to \infty} \bar{a}_{ki}(t) \\
\text{subject to} & \quad \Gamma_{ki}(m(t)) \geq \gamma_{ki} - s_{ki}(t), \quad i \in U_k, k \in K, \forall t
\end{align*}
\]

(2.53a)
\[ \sum_{i \in U_k} \| m_{ki}(t) \|_2^2 \leq P_k^{\max}, \quad k \in \mathcal{K}, \forall t \quad (2.53b) \]

\[ s_{ki}(t) \geq 0, \quad i \in U_k, k \in \mathcal{K}, \forall t, \quad (2.53c) \]

where \( \bar{a}_{ki}(t) \) is as defined in (2.33), and variables are \( \{s_{ki}(t), m_{ki}(t)\}_{k \in \mathcal{K}, i \in U_k} \) for all \( t \in \{0, 1, \ldots\} \).

We now present a method to find a solution of problem (2.53). Let \( \bar{\nu} \) be an auxiliary variable. Then, problem (2.53) can be equivalently reformulated as the following optimization problem:

maximize \( \bar{\nu} \)

subject to

\[ \bar{\nu} \leq \min_{k \in \mathcal{K}, i \in U_k} \liminf_{t \to \infty} \bar{a}_{ki}(t) \quad (2.54a) \]

\[ \Gamma_{ki}(m(t)) \geq \gamma_{ki} - s_{ki}(t), \quad i \in U_k, k \in \mathcal{K}, \forall t \quad (2.54b) \]

\[ \sum_{i \in U_k} \| m_{ki}(t) \|_2^2 \leq P_k^{\max}, \quad k \in \mathcal{K}, \forall t \quad (2.54c) \]

\[ s_{ki}(t) \geq 0, \quad i \in U_k, k \in \mathcal{K}, \forall t, \quad (2.54d) \]

with variables \( \bar{\nu} \) and \( \{s_{ki}(t), m_{ki}(t)\}_{k \in \mathcal{K}, i \in U_k} \) for all \( t \in \{0, 1, \ldots\} \).

Next, we use the drift-plus-penalty minimization method \([110, 114]\) to find a solution of problem (2.54). To do this we assume that the auxiliary variable \( \bar{\nu} \) is a time average of auxiliary variable \( \nu(t) \) for all \( t = \{0, 1, \ldots\} \), i.e., \( \bar{\nu} \triangleq \lim_{t \to \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \mathbb{E}\{\nu(\tau)\} \). Then, by expressing the inequality constraint (2.54a) as a set of separate inequalities, we modify problem (2.54) as

maximize \( \bar{\nu} \)

subject to

\[ \bar{\nu} \leq \liminf_{t \to \infty} \bar{a}_{ki}(t), \quad i \in U_k, k \in \mathcal{K} \quad (2.55a) \]

\[ \Gamma_{ki}(m(t)) \geq \gamma_{ki} - s_{ki}(t), \quad i \in U_k, k \in \mathcal{K}, \forall t \quad (2.55b) \]

\[ \sum_{i \in U_k} \| m_{ki}(t) \|_2^2 \leq P_k^{\max}, \quad k \in \mathcal{K}, \forall t \quad (2.55c) \]

\[ s_{ki}(t) \geq 0, \quad i \in U_k, k \in \mathcal{K}, \forall t, \quad (2.55d) \]

with variables \( \nu(t) \) and \( \{s_{ki}(t), m_{ki}(t)\}_{k \in \mathcal{K}, i \in U_k} \) for all \( t \in \{0, 1, \ldots\} \). Hence, we follow the steps that we have used in subsection 2.4.2 to find a solution of problem (2.55).

Let \( \{Z_{ki}(t)\}_{k \in \mathcal{K}, i \in U_k} \) be the virtual queues associated with constraint (2.55a). We update virtual queues \( Z_{ki}(t) \) for all \( k \in \mathcal{K}, i \in U_k \), in every time slot as

\[ Z_{ki}(t + 1) = \max[Z_{ki}(t) + \nu(t) - a_{ki}(t), 0]. \quad (2.56) \]
Note that the queue update equation (2.56) is different from the one given in expression (2.38). The stability of virtual queues \( \{Z_{ki}(t)\}_{k \in K, i \in U_k} \) implies the feasibility of constraint (2.55a). Then, a quadratic Lyapunov function \( L(Z(t)) \) can be defined as [110, pp. 29-43]

\[
L(Z(t)) = \frac{1}{2} \sum_{k \in K} \sum_{i \in U_k} Z_{ki}(t)^2,
\]

where \( Z(t) = [Z_{11}(t), \ldots, Z_{KI}(t)]^T \). Furthermore, the drift in the Lyapunov function from one slot to the next can be defined, using expression (2.57), as follows:

\[
\Delta(Z(t)) = \mathbb{E}\{L(Z(t+1)) - L(Z(t))|Z(t)\},
\]

(2.58)

Similar to the discussion in section 2.4.2, we design a control policy that solves problem (2.55), by minimizing an upper bound of the following drift-plus penalty expression [110]:

\[
\Delta(Z(t)) - V\mathbb{E}\{\nu(t)|Z(t)\},
\]

(2.59)

where \( V \geq 0 \), subject to constraints (2.55b)-(2.55d) in each time slot, i.e.,

\[
\Gamma_{ki}(m(t)) \geq \gamma_{ki} - s_{ki}(t), \quad i \in U_k, k \in K
\]

(2.60a)

\[
\sum_{i \in U_k} \|m_{ki}(t)\|_2^2 \leq P_k^{\text{max}}, \quad k \in K
\]

(2.60b)

\[
s_{ki}(t) \geq 0, \quad i \in U_k, k \in K.
\]

(2.60c)

An upper bound of expression (2.59) can be expressed as

\[
\Delta(Z(t)) - V\mathbb{E}\{\nu(t)|Z(t)\} \leq \tilde{D} - V\mathbb{E}\{\nu(t)|Z(t)\} + \mathbb{E}\{\nu(t)|Z(t)\} \sum_{k \in K} \sum_{i \in U_k} Z_{ki}(t) - \sum_{k \in K} \sum_{i \in U_k} Z_{ki}(t)\mathbb{E}\{a_{ki}(t)|Z(t)\},
\]

(2.61)

where \( \tilde{D} \) is a finite positive constant that satisfies the following condition for all \( t \):

\[
\tilde{D} \geq \frac{1}{2} \left[ \sum_{k \in K} \sum_{i \in U_k} \mathbb{E}\{\nu(t)^2|Z(t)\} + \sum_{k \in K} \sum_{i \in U_k} \mathbb{E}\{a_{ki}(t)^2|Z(t)\} \right].
\]

(2.62)

We now summarize the proposed dynamic control algorithm to solve problem (2.55) in Algorithm 2.4. In every time slot \( t \), the proposed algorithm observes the virtual queues \( \{Z_{ki}(t)\}_{k \in K, i \in U_k} \) and the channel states \( \{h_{ki}(t)\}_{k \in K, i \in U_k} \).
Then, a control action is taken to minimize the righthand side of expression (2.61) subject to the constraints (2.60a)-(2.60c). The minimization of the righthand side of expression (2.61) can be decoupled across variables \( \nu(t) \) and \( \{s_{ki}(t), m_{ki}(t)\}_{k \in \mathcal{K}, i \in \mathcal{U}_k} \) for all \( t \in \{0, 1, \ldots\} \), resulting in subproblems as shown in Algorithm 2.4. To solve the subproblems, the concept of opportunistically minimizing an expectation has been used.

Algorithm 2.4. Max-Min fair dynamic admission control algorithm for problem (2.55)

1. Auxiliary variable: the auxiliary variable \( \nu(t) \) is obtained by solving the following optimization problem

\[
\text{maximize } V\nu(t) - \nu(t) \sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{U}_k} Z_{ki}(t) \\
\text{subject to } 0 \leq \nu(t) \leq \nu^{\text{max}},
\]

(2.63)

with variable \( \nu(t) \), and \( \nu^{\text{max}} > 0 \) is an algorithm parameter.

2. Resource allocation: the transmit beamformer \( m_{ki}(t) \) and the variable \( s_{ki}(t) \) for all \( k \in \mathcal{K}, i \in \mathcal{U}_k \) are obtained by solving the following optimization problem:

\[
\text{maximize } \sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{U}_k} Z_{ki}(t)a_{ki}(t) \\
\text{subject to } \Gamma_{ki}(m(t)) \geq \gamma_{ki} - s_{ki}(t), \quad i \in \mathcal{U}_k, \ k \in \mathcal{K} \quad (2.64a) \\
\sum_{i \in \mathcal{U}_k} \|m_{ki}(t)\|_2^2 \leq P_{k_{\text{max}}}, \quad k \in \mathcal{K} \quad (2.64b) \\
s_{ki}(t) \geq 0, \quad i \in \mathcal{U}_k, \ k \in \mathcal{K}, \quad (2.64c)
\]

with variables \( \{s_{ki}(t), m_{ki}(t)\}_{k \in \mathcal{K}, i \in \mathcal{U}_k} \).

3. Queue update: update \( \{Z_{ki}(t)\}_{k \in \mathcal{K}, i \in \mathcal{U}_k} \) using expression (2.56). Set \( t = t + 1 \), and go to step 1.

The first step updates the auxiliary variables, where we obtain the solution as \( \nu^*(t) = \nu^{\text{max}} \) if \( V \geq \sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{U}_k} Z_{ki}(t) \), and \( \nu^*(t) = 0 \) otherwise. Note that problem (2.64) is identical to problem (2.48), and hence, we use the method described in Section 2.4.3 to update step 2 of Algorithm 2.4. The virtual queues are updated in step 3.
The dynamic control Algorithm 2.3 and Algorithm 2.4 use Algorithm 2.1 to solve the optimization problem in step 2 of these algorithms. Compared to the computational complexity of step 2, the complexity of solving the problem in step 1 of both these algorithms is negligible. Hence, the computational complexity of Algorithm 2.3 and Algorithm 2.4 per time slot is the same as that of Algorithm 2.1. Furthermore, as signaling is required only at step 2 of the dynamic control algorithms, and the signaling overhead of these algorithms per time slot is the same as that of Algorithm 2.1.

2.5 Admission control in heterogeneous networks

In section 2.2, we considered the problem of admission control in the downlink of a multicell MISO system. An interesting extension of this problem is to find the maximum number of users that can be admitted to the network when they have different priorities. For example, in a two-tier macro-femto network (i.e., a heterogeneous network) macro users have a higher priority compared to the femto users [118]. Hence, more priority should be given to the macro users while they are being admitted to the network. In this section, we focus the problem of admission control in a heterogeneous network (hetnet).

We consider the downlink of a single-macrocell MISO hetnet. We assume that the hetnet consists of a macrocell and a set of femtocells overlaid within the coverage area of the macrocell. The macrocell uses the space division multiple access method to share its frequency band with the femto network. We consider the index one is used to represent the macrocell (i.e., \( k = 1 \) represents macrocell), and the rest (i.e., \( k = 2, ..., K \) represent femtocells.

We assume that all the macro users must be served with their required SINR targets. Furthermore, we assume that there are not enough resources to admit all the femto users at their specified SINR targets. Hence, it is necessary to find the maximum number of femto users that can be admitted to the network while satisfying the SINR requirements of all the macro users and the admitted femto users. Thus, the problem of admission control for the downlink of a MISO hetnet can be expressed, by modifying problem (2.8), as the following optimization problem:

\[ \text{maximize} \quad \sum_{k=2}^{K} \theta_k \]
\[ \text{subject to} \quad \sum_{k=2}^{K} \theta_k \leq \frac{1}{N} \sum_{n=1}^{N} y(n) \]
\[ \theta_k \geq 0 \quad \forall k \]

66
minimize \[ \sum_{k \in K \backslash \{1\}} \| s_k(t) \|_0 \]
subject to \[
\Gamma_{1i}(m(t)) \geq \gamma_{1i}, \quad i \in U_1 \tag{2.65a}
\]
\[
\Gamma_{ki}(m(t)) \geq \gamma_{ki} - s_{ki}(t), \quad i \in U_k, k \in K \backslash \{1\} \tag{2.65b}
\]
\[
\sum_{i \in U_k} \| m_{ki}(t) \|_2^2 \leq P_{k}^{\max}, \quad k \in K \tag{2.65c}
\]
\[
s_{ki}(t) \geq 0, \quad i \in U_k, k \in K \backslash \{1\}, \tag{2.65d}
\]

where \( s_k(t) = [s_{k1}(t), \ldots, s_{kK}(t)]^T \) for all \( k \in K \backslash \{1\} \) and variables are \( s_{ki}(t) \) for all \( k \in K \backslash \{1\}, i \in U_k \) and \( m_{ki}(t) \) for all \( k \in K, i \in U_k \). The notation \( \Gamma_{ki}(m(t)) \) is as defined in (2.3), \( P_{k}^{\max} \) is the maximum transmit power of \( k \)th base station, and \( \gamma_{ki} \) is the SINR target of \( i \)th user of base station \( k \).

Note that due to the additional constraint (2.65a), problem (2.65) is different compared to problem (2.8). Thus, we cannot directly use the proposed centralized Algorithm 2.1 or the distributed algorithm Algorithm 2.2 to find a solution of this problem. However, these two algorithms can be adapted to approximately solve problem (2.65) (in centralized and distributed manner), by expressing constraint (2.65a) as a convex constraint.

It is easy to show that, by following the approach in [119], constraint (2.65a) can be expressed as the following second-order cone constraint

\[
\begin{bmatrix}
\sqrt{1 + \frac{1}{\gamma_{1i}}(h_{1i}^H m_{ki})}
\end{bmatrix}^\mathsf{T} \succeq 0, \quad i \in U_1, \tag{2.66}
\]

and this a convex constraint [21, pp. 127-174]. Then, by using constraint (2.66), problem (2.65) can be equivalently reformulated as follows:

minimize \[ \sum_{k \in K \backslash \{1\}} \| s_k(t) \|_0 \]
subject to \[
\begin{bmatrix}
\sqrt{1 + \frac{1}{\gamma_{1i}}(h_{1i}^H m_{ki})}
\end{bmatrix}^\mathsf{T} \succeq 0, \quad i \in U_1, \tag{2.67}
\]
\[
\Gamma_{ki}(m(t)) \geq \gamma_{ki} - s_{ki}(t), \quad i \in U_k, k \in K \backslash \{1\} \tag{2.65b}
\]
\[
\sum_{i \in U_k} \| m_{ki}(t) \|_2^2 \leq P_{k}^{\max}, \quad k \in K \tag{2.65c}
\]
\[
s_{ki}(t) \geq 0, \quad i \in U_k, k \in K \backslash \{1\}, \tag{2.65d}
\]
with variables \( s_{ki}(t) \) for all \( k \in \mathcal{K}\setminus\{1\}, i \in \mathcal{U}_k \) and \( \mathbf{m}_{ki}(t) \) for all \( k \in \mathcal{K}, i \in \mathcal{U}_k \). Note that in problem (2.67), the objective function and the second constraint are not convex.

We now follow the approach (i.e., sequential convex programming) that is presented in subsection 2.3.1 to solve problem (2.67). Let \((\hat{s}_{ki}, \hat{\mathbf{m}}_{ki}, \hat{\beta}_{ki})\) for all \( i \in \mathcal{U}_k, k \in \mathcal{K}\) be an arbitrary positive point, then problem (2.67) can be approximated near this point as the following convex optimization problem:

\[
\begin{align*}
\text{minimize} & \quad \sum_{k \in \mathcal{K}\setminus\{1\}} \sum_{i \in \mathcal{U}_k} w_{ki} s_{ki} \\
\text{subject to} & \quad \sqrt{1 + \frac{\gamma_{ki}}{\hat{s}_{ki}}} (\mathbf{h}_{ki}^H \mathbf{m}_{ki} + \mathbf{H}_{\mathcal{K}}^H \mathbf{m}_{\mathcal{K} \setminus \{1\}})^T \succeq \text{SOC}_0, \quad i \in \mathcal{U}_k, \\
& \quad \gamma_{ki} - s_{ki} - \hat{g}_{ki}(\mathbf{m}_{ki}, \beta_{ki}) \leq 0, \quad i \in \mathcal{U}_k, k \in \mathcal{K}\setminus\{1\} \\
& \quad \sum_{j \in \mathcal{U}_k \setminus \{i\}} |(\mathbf{h}_{ki}^H \mathbf{m}_{kj})|^2 + \sum_{l \in \mathcal{K}\setminus\{k\}} \sum_{j \in \mathcal{U}_l \setminus \{i\}} |(\mathbf{h}_{ki}^H \mathbf{m}_{lj})|^2 + \sigma_{ki}^2 \leq \beta_{ki}, \quad i \in \mathcal{U}_k, k \in \mathcal{K}\setminus\{1\} \\
& \quad \sum_{i \in \mathcal{U}_k} \| \mathbf{m}_{ki}(t) \|_2^2 \leq P_{k_{\max}}, \quad k \in \mathcal{K} \\
& \quad s_{ki}(t) \geq 0, \quad i \in \mathcal{U}_k, k \in \mathcal{K}\setminus\{1\},
\end{align*}
\]

(2.68)

where \( w_{ki} = 1/(\hat{s}_{ki} + \epsilon) \), variables are \{\mathbf{m}_{ki}\}_{k \in \mathcal{K}, i \in \mathcal{U}_k} \) and \{\hat{s}_{ki}, \hat{\beta}_{ki}\}_{k \in \mathcal{K}\setminus\{1\}, i \in \mathcal{U}_k}, \) and \( \hat{g}_{ki}(\mathbf{m}_{ki}, \beta_{ki}) \) is as defined in expression (2.11). Note that compared to problem (2.13), in problem (2.68) the objective function is modified (i.e., by removing the terms for \( k = 1 \)) and an additional SOC constraint is introduced.

We can easily adapt the proposed Algorithm 2.1 and distributed Algorithm 2.2 to find centralized and distributed solutions of problem (2.68). Specifically, by solving problem (2.68) at step 3 of Algorithm 2.1 we can obtain a centralized algorithm. To obtain a distributed algorithm to solve this problem we can follow the approach given in sections 2.3.2 and 2.3.3.

2.6 Numerical results

In this section, we evaluate the performance of the proposed algorithms for multicell cellular networks and heterogeneous wireless networks.

68
2.6.1 Numerical simulations for multicell cellular networks

Here, we evaluate the performance of the proposed Algorithm 2.1, Algorithm 2.2, Algorithm 2.3, and Algorithm 2.4. In our simulations, we consider two multicell network setups, with the number of base stations, users, and antennas given in Table 2.1.

Table 2.1. Network configurations.

| Multicell network | Number of BS ($K$) | Total number of users ($\sum_{k \in K} |U_k|$) | Number of antennas ($T$) |
|-------------------|-------------------|---------------------------------|---------------------|
| Network 1         | 3                 | 12                              | 4                   |
| Network 2         | 7                 | 28                              | 4                   |

In each network, the base stations are placed in such a way that they form an equilateral triangle as shown in Fig. 2.3, and the distance between two base stations is denoted by $D_{BS}$. We assume circular cells, where the radius of each is denoted by $R_{BS}$. Furthermore, we consider that each base station serves four users.

We use an exponential path loss model, where the channel gains between base station $l$ and $i$th user of base station $k$ is modeled as

$$h_{lki} = \left(\frac{d_{lki}}{d_0}\right)^{-\eta/2} c_{ki}^l,$$  \hspace{1cm} (2.69)

where $d_{lki}$ is the distance from $l$th base station to $i$th user of base station $k$, $d_0$ is the far field reference distance [120], $\eta$ is the path loss exponent, and $c_{ki}^l \in \mathbb{C}^T$ is arbitrarily chosen from the distribution $CN(0, I)$ (i.e., frequency-flat fading with uncorrelated antennas). The first term in (2.69) represents the path loss component, and the second term is the Rayleigh small-scale fading. We refer to an arbitrarily generated set of fading coefficients $\mathcal{C} = \{c_{ki}^l | l, k \in K, i \in U_k\}$ as a single fading realization, and an arbitrarily generated set of distances $\mathcal{D} = \{d_{lki}^l | l, k \in K, i \in U_k\}$ as a single topology realization.

We assume that $\{P_{k}^{\text{max}} = P_0^{\text{max}} \}_{k \in K}$, $\{\sigma_{ki} = \sigma, \gamma_{ki} = \gamma \}_{i \in U_k, k \in K}$. We define the signal-to-noise ratio (SNR) operating point at a distance $d$ as

$$\text{SNR}(d) = \left(\frac{P_0^{\text{max}}}{\sigma^2}\right) (d/d_0)^{-\eta}.$$  \hspace{1cm} (2.70)
Fig. 2.3. Base station layout: (a) Network 1 with 3-base stations, (b) Network 2 with 7-base stations.
In our simulations, we set $\text{SNR}(R_{\text{BS}}) = 5\, \text{dB}$, $d_0 = 1$, $\eta = 4$, $P_{0}^{\text{max}}/\sigma^2 = 45\, \text{dB}$, and $D_{\text{BS}} = 1.6R_{\text{BS}}$.

We now evaluate the performance of proposed Algorithm 2.1 and Algorithm 2.2 for Network 1 and Network 2. In all our simulations, Algorithm 2.1 is stopped if the improvement in the objective value is less than some $\kappa$, where we set $\kappa = 0.01$ [22]. In Algorithm 2.2, the ADMM algorithm (i.e. the inner iterations) is stopped after running a fixed number of $Q$ iterations [112, pp. 15-18]. The outer iteration is stopped when the improvement in the objective function of problem (2.9) is less than 0.01 or when the algorithm runs for a maximum 20 iterations. Furthermore, parameter $\rho$ is set to 0.1. To illustrate the convergence behavior of Algorithm 2.1 and Algorithm 2.2, we consider an arbitrarily chosen single fading realization and single topology realization.

Fig. 2.4 shows the convergence behavior of the proposed Algorithm 2.1 in the case of Network 1. Note that we have also used the same figure to plot the number of admitted users in each iteration. We denote by $F^p_C = \sum_{k \in K} \sum_{i \in U_k} \log(s^p_{ki} + \epsilon)$ the objective value problem (2.9) calculated at $p$th iteration of Algorithm 2.1, where $s^p_{ki}$ is the auxiliary variable of $i$th user of base station $k$ in $p$th iteration.
We evaluate $F_C^p$ after step 2 of Algorithm 2.1. In each iteration, the number of admitted users are counted by checking how many users have satisfied condition (2.5). The results show that the proposed Algorithm 2.1 converges within the first few iterations.

Fig. 2.5 shows the convergence behavior of the distributed Algorithm 2.2, along with the results obtained by the centralized Algorithm 2.1 for Network 1. It also shows the number of admitted users in each iteration of the distributed Algorithm 2.2 compared to that of Algorithm 2.1. Here we set $Q = 10$ in the ADMM algorithm. We denote by $F_D^q = \sum_{k \in K} \sum_{i \in U_k} \log(\hat{s}_{ki}^q + \epsilon)$ the objective value problem (2.9) calculated at iteration $q$ of Algorithm 2.2, where $s_{ki}^q$ is the auxiliary variable of $i$th user of base station $k$ in $q$th iteration. The “square” and “circle” markers in Fig. 2.5 are to indicate $F_D^q$ and the number of admitted users, respectively. Furthermore, these markers represent the start of ADMM algorithm for a new point $\{\hat{m}_{ki}, \hat{\beta}_{ki}\}_{i \in U_k, k \in K}$ which is set at step 2 of the algorithm. The results show that the proposed distributed algorithm converges to the objective value obtained via the centralized algorithm, and it also admits the same number of users as the centralized Algorithm 2.1.
Fig. 2.6. Centralized algorithm comparison for the average admitted users versus the SINR target: (a) Network 1; (b) Network 2.
Fig. 2.7. Distributed algorithm comparison for the average admitted users versus the SINR target: (a) Network 1; (b) Network 2.
We now discuss the average behavior of Algorithm 2.1 and Algorithm 2.2. We run both algorithms on Network 1 and Network 2 over 500 fading and topology realizations. In Fig. 2.6, we compare the average number of admitted users versus the SINR target $\gamma$ of the proposed Algorithm 2.1 with an exhaustive search algorithm, the D-SOC algorithm in [83] (named Alg. [83] in this chapter\(^7\)), and the centralized algorithm in [85] (named Alg. [85]-centralized). The results in Fig. 2.6(a) are for Network 1, and they show that the average number of admitted users of Algorithm 2.1 is close to that of the exhaustive search algorithm. In addition, the results show that when $\gamma$ is low the proposed Algorithm 2.1 slightly outperforms Alg. [83], and when $\gamma$ is high Alg. [83] marginally performs better than Algorithm 2.1. Furthermore, it is shown that proposed Algorithm 2.1 is better than Alg. [85]-centralized for both high and low SINR targets. In the case of Network 2 (see Fig. 2.6(b)), the results show that the performance obtained by the proposed Algorithm 2.1 and D-SOC algorithm display similar behavior as they did in Network 1. However, our Algorithm 2.1 outperforms

\(^7\)Note that we compare our results with one of the best algorithms, i.e., Algorithm 2: Deflation based on SOC programming (D-SOC), proposed in [83]. Furthermore, the algorithm parameter $M$ in the D-SOC algorithm is set to zero.
Fig. 2.9. Centralized algorithm performance comparison: (a) Average execution time (b) Average number of admitted users.
Alg. [85]-centralized for all considered SINR targets. In the case of NW2, the number of combinations required to be checked using an exhaustive search algorithm is very large. Therefore, we have not use the exhaustive search method to find the optimal solution for NW2. It is worth noting that, it may not be practical to use this method when the number of users are greater than fifteen [83, 100].

After that, we evaluate the average number of admitted users versus the SINR target $\gamma$ of the proposed distributed Algorithm 2.2 in Fig.2.7. As benchmarks, we consider an exhaustive search algorithm, the proposed centralized Algorithm 2.1, and the distributed algorithm in [85] (named Alg. [85]-distributed). For Network 1 (see Fig. 2.7(a)), the average number of admitted users obtained by distributed Algorithm 2.2 is plotted by running the ADMM algorithm for $Q = 1, 5, 10$, and 50 iterations. The results show that for $Q = 5, 10$, and 50 proposed Algorithm 2.2 outperforms Alg. [85]-distributed for all the considered SINR targets. When $Q = 1$, our Algorithm 2.2 performs slightly better than Alg. [85]-distributed at low SINR targets. However, for large SINR targets, proposed Algorithm 2.2 outperforms Alg. [85]-distributed when $Q = 1$. Furthermore, the results show that when $\gamma$ is low, proposed centralized Algorithm 2.1 admits more users on average than that of the distributed Algorithm 2.2. But
for all the simulated cases, when $\gamma$ is high the number of average admitted users obtained by using Algorithm 2.2 is closer to that of the centralized Algorithm 2.1, for the considered values of $Q$. For Network 2 (see Fig. 2.7(b)), the average number of admitted users obtained by Algorithm 2.2 is plotted by varying the maximum number of inner iterations as $Q = 1$ and 10. The results show that our Algorithm 2.2 outperforms Alg. [85]-distributed, when $Q = 1$ and 10, for all considered SINR targets. Furthermore, the results show that when the number of ADMM iterations are increased, the performance of Algorithm 2.2 is closer to that of Algorithm 2.1; but it needs more ADMM iterations to reach the centralized solution. Fig. 2.8 shows the average number of admitted users per BS versus iteration $q$ for Algorithm 2.2. We have set $Q = 10$ for the ADMM algorithm, i.e., step 3 of Algorithm 2.2. In each graph a marker is drawn to show the start of ADMM algorithm for a new point $\{\hat{m}_k, \hat{\beta}_k\}_{i \in I, k \in K}$ that is set at step 2 of Algorithm 2.2. The results show that once Algorithm 2.2 approximates problem (2.9) as a convex problem at step 2, (i.e., after each 10th iteration) the average number of admitted users per BS increases.

We now evaluate the average execution times and the average number of admitted users for the proposed centralized Algorithm 2.1, for $\gamma = 9$dB and 15dB. We consider a network with 7 base stations. Then, the simulations are carried
out by varying the total number of users in the network as $U = 28, 35, \ldots, 56$, such that each base station has an equal number of users (i.e., each base station serves $U/7$ users). Fig. 2.9(a) shows the average execution time (in seconds) versus the total number of users in the network. The results show that for a small number of users (i.e., $U \leq 35$), the average execution time of the proposed Algorithm 2.1 is slightly higher than that of Alg. [83]. However, for a large number of users (i.e., $U > 35$) the average execution time of Alg. [83] is higher compared to the proposed algorithm. In order to show the gain in the objective value (i.e., the total admitted users), we also plot the average number of admitted users versus the total number of users in the network in Fig. 2.9(b). The results show that for all simulated cases both algorithms display a similar performance. However, it is worth noting that the proposed algorithm has a low execution time for larger networks.

To provide a statistical description for the speed of convergence of the proposed centralized Algorithm 2.1 and Alg. [83], we have also plotted the cumulative distribution function (CDF) of the total number of iterations that each algorithm requires to terminate (i.e., find an approximate solution for the admission control problem). We ran this simulation for Network 1 and a network
with 19 base stations (named Network 3). In Network 3 also each base station serves four users and $T = 4$.

Fig. 2.10 shows the empirical CDF plots for the total number of iterations required to terminate Algorithm 2.1 and Alg. [83] for Network 1. The results show that Alg. [83] terminates in 10 iterations for almost 100% of the simulated cases. However, in our proposed Algorithm 2.1 only 60% of the simulated cases are terminating in 10 iterations. Note that although Alg. [83] terminates in a fewer number of iterations, it achieves the same average performance as Algorithm 2.1 (see Fig. 2.6(a)). The results further show that Algorithm 2.1 requires around 20 iterations to terminate 100% of the simulated cases.

Fig. 2.11 shows the empirical CDF plots of the total number of iterations required to terminate Algorithm 2.1 and Alg. [83] for Network 3. The results show that the proposed Algorithm 2.1 terminates in less than 35 iterations for almost 100% of the simulated cases. However, in Alg. [83] only 10% of the simulated cases terminate in fewer than 35 iterations. Note that although our Algorithm 2.1 terminates in a fewer number of iterations it provides the same average performance as Alg. [83] (see Fig. 2.12).

Next, we evaluate the average behavior of the proposed proportional fair admission control Algorithm 2.3 and the max-min fair admission control Alg-
Algorithm 2.1 and Algorithm 2.3 for Network 1 and Network 2. In our simulations, we consider that two of the users in each cell are located within a circular region with $2R_{BS}/3$ radius around the base station (we call them in-cell users). The other two users are located at the cell boundary, and we name them as cell-edge users. We label the in-cell users in each cell with the integer values 1 and 2 and the cell-edge users with the integer values 3 and 4. We set the SINR target $\gamma = 9$ dB, and the obtained results are averaged over $T_{\text{max}} = 15000$ time slots. Because we want to check the performance of the cell-edge users, the simulations are run by changing only the fading realizations while keeping the topology realization fixed. We compare both Algorithm 2.3 and Algorithm 2.4 with the average performance obtained by running the proposed Algorithm 2.1 (which does not account for fairness) for $T_{\text{max}}$ time slots.

Fig. 2.13 shows the objective value of problem (2.35) versus parameter $V$ for Network 1. The results show that the objective value increases as $V$ is increased, and it quickly converges to its maximum value. Fig. 2.14 shows the average number of admissions of users in each base station for Network 1. It can be seen that the average admissions of the cell-edge users (i.e., users 3 and 4 of each base station) have been proportionally increased in proportional fair admission control Algorithm 2.3, compared to that of Algorithm 2.1. However, it can be
seen that this performance gain for the cell-edge users is obtained at the expense of the performance of the in-cell users (i.e., users 1 and 2 of each base station).

In Fig. 2.15 and Fig. 2.16 we see the average behavior of Algorithm 2.4 in the case of Network 1. The results in Fig. 2.15 show that the objective value improves as $V$ is increased. Furthermore, Fig. 2.16 illustrates the average admissions of each user, and it shows Algorithm 2.4 has admitted all the users nearly equally within $T_{\text{max}}$, compared to Algorithm 2.1. Note that the gains of the cell-edge users are obtained at the cost of a performance reduction for in-cell users.

For Network 2, the average performance of Algorithm 2.3 is shown in Fig. 2.17(a) and that of Algorithm 2.4 is shown in Fig. 2.17(b). The results show that the proposed fair algorithms show a similar performance as we have seen in Network 1 (compare with Fig. 2.14 and Fig. 2.16). Hence, it can be seen that the proposed Algorithm 2.3 and Algorithm 2.4 provide the required fairness for a network with a higher number of users.

Table 2.2 shows the fairness index calculated for the results obtained by the proposed algorithms using the Jain’s fairness index equation [121]. The results show that for the proportional fairness case the Jain’s fairness index obtained by
Table 2.2. Jain’s fairness index comparison for Network 1 and Network 2 using proposed algorithms.

<table>
<thead>
<tr>
<th>Network</th>
<th>Fairness index (proportional fairness)</th>
<th>Fairness index (maxmin fairness)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Algorithm 2.1</td>
<td>Algorithm 2.3</td>
</tr>
<tr>
<td>Network 1</td>
<td>0.5405</td>
<td>0.7459</td>
</tr>
<tr>
<td>Network 2</td>
<td>0.5185</td>
<td>0.7174</td>
</tr>
</tbody>
</table>

proposed Algorithm 2.3 is higher than that of Algorithm 2.1. Furthermore, the results show that for the maxmin fairness case the Jain’s fairness index obtained by Algorithm 2.4 is also higher than that of Algorithm 2.1, and the fairness index is equal to 1.

### 2.6.2 Numerical simulations for heterogeneous networks

Here, we evaluate the performance of the proposed Algorithm 2.1 and Algorithm 2.2 by applying them for a single-macro cell hetnet.
Fig. 2.17. Average number of admissions of users for Network 2 when $V = 25$: (a) Algorithm 2.3, proportional fairness; (b) Algorithm 2.4, max-min fairness.

(a)

(b)
We assume that the macro base station and the femto base stations have a circular coverage region, where the radius of the macrocell transmission region is $R_{BS}$, and the radius of the femtocell transmission region is $r_{BS}$. We consider that the macrocell is overlaid by three femtocells. These femtocells are uniformly distributed within the macrocell such that they are $R_{BS}/3$ distance away from the macro base station. By deploying the femto base station $R_{BS}/3$ distance away, we reduce the high interference generated by the macro base station towards the femto users. Furthermore, we assume that there are two macro users in the macrocell, and four femto users in each femtocell. In our simulations, the macro and femto users are uniformly distributed, and the macro users are not located inside the femtocells.

The channel gains between the users are as defined in (2.69) and the SINR operating point at a distance $d$ is as given in (2.70). In our simulations, we set $\text{SNR}(R_{BS}) = 5\text{dB}$, $\text{SNR}(r_{BS}) = 10\text{dB}$, $d_0 = 3$ for the femtocell and $d_0 = 10$ for the macrocell, $\eta = 4$, $P_1^{\max}/\sigma^2 = 45\text{dB}$, $P_k^{\max}/\sigma^2 = 35\text{dB}$ for all $k \in \mathcal{K}\setminus\{1\}$, and $T = 4$. The obtained results are averaged over 500 fading and topology realizations.

Fig. 2.18 shows the average number of admitted users versus the SINR target $\gamma$ for the case of the hetnet obtained by solving problem (2.68) in a centralized and distributed manner. Results for both centralized and distributed cases are obtained by adapting the proposed algorithms Algorithm 2.1 and Algorithm 2.2, respectively. In the distributed algorithm, we have set $Q = 10$ and 50, i.e., the number of iterations in the ADMM algorithm. The results show that the performance obtained in the centralized implementation is closer to that of the exhaustive search algorithm. Furthermore, the results show that the performance obtained in the distributed case is slightly low compared to that obtained with the centralized implementation. However, the results further show that when the number of iterations in the ADMM algorithm are increased, the average performance reaches the performance of the centralized algorithm.

Fig. 2.19 shows the average number of admissions for each user in the hetnet obtained by solving problem (2.68) in a centralized and distributed manner for $\gamma = 9\text{dB}$ and $\gamma = 15\text{dB}$. The results shows that in both scenarios the two macro users were always admitted to the network for all the simulated cases. The results further show that, when the target SINR $\gamma = 15\text{dB}$ the average admissions of femto users have been reduced compared to the $\gamma = 9\text{dB}$ case.
(compare the average admission of femto users in Fig. 2.19(a) and Fig. 2.19(b)). The reason for this is that to admit all the macro users with required high SINR targets more femto users have to remain unserved. Thus, the interference towards the macro users can be reduced and the required SINR target of macro users can be achieved.

![Graph](image)

**Fig. 2.18.** Average number of admitted users versus SINR target $\gamma$ of the hetnet.

### 2.7 Summary and discussion

In this chapter, the problem of admission control for the downlink of a multicell multiple-input single-output system has been considered. This problem has been cast as an $\ell_0$ minimization problem, which is combinatorial, and NP-hard. Then, the formulated $\ell_0$ minimization problem has been approximated as a non-combinatorial problem. To solve this problem, suboptimal centralized and distributed algorithms based on sequential convex programming and alternating direction method of multipliers have been proposed. Numerically, we have shown that the proposed centralized and distributed admission control algorithms achieve near-to-optimal performance. Next, the admission control problem has
Fig. 2.19. Average number of admissions of users of the hetnet: (a) $\gamma = 9\,\text{dB}$ (b) $\gamma = 15\,\text{dB}$. 
been extended to provide fairness, where a long term fairness among the users has been guaranteed. Specifically, the proportional and max-min fairness have been considered, and dynamic control algorithms via Lyapunov optimization have been proposed to solve them. Using numerical simulations, it has been shown that the proposed fair admission control algorithms guarantee fairness among the users. Finally, the problem of admission control for a heterogeneous network has been considered. For this problem, a suboptimal solution has been found by adapting the proposed centralized and distributed algorithms. The study of fairness guaranteed distributed dynamic admission control algorithms is left as a future interesting work.
3 Power-Throughput tradeoff in MIMO downlink heterogeneous networks

In this chapter, we consider a single-macrocell MIMO heterogeneous network, and provide an algorithm to find the achievable power-rate region for this setup. We assume that the macrocell shares the same frequency band with the femto network. The interference power towards the macro users from the femto base stations is kept below a threshold to guarantee that the performance of the macro users is not degraded due to the femto network. We first formulate a two-dimensional vector optimization problem in which we consider maximizing the sum-rate and minimizing the sum-power, subject to the maximum transmit power and interference threshold constraints. We then generalize the two-dimensional problem as a multi-dimensional one which can be used to address different requirements of a heterogeneous network. The considered problem is NP-hard. We provide a solution method based on the relationship between the weighted sum-rate (WSR) maximization and weighted-sum-mean-squared-error (WMMSE) minimization problems [45].

We use the proposed algorithm to numerically find the power-rate region of the heterogeneous network. Furthermore, we evaluate the impact on the power-rate region when the interference generated by femtocells towards the macro users is limited to different threshold levels and when the number of transmit antennas is varied. Next, by using the proposed algorithm we study the power-rate regions when the users are served by both macrocell and femtocells and only by macrocell. The proposed algorithm can be used to evaluate the performance of real heterogeneous networks via off-line numerical simulations.

Note that the method proposed in this chapter is based on an extension of the algorithm proposed in [45]. However, this extension is nontrivial because the optimization problem considered in this chapter has a different objective function and extra interference constraints compared to [45]. Hence, the precoder updating step of our algorithm is completely different from that of [45].
3.1 System model

We consider the downlink of a single-macrocell multiple-input multiple-output (MIMO) heterogeneous network, as illustrated in Figure 3.1. We assume that the heterogeneous network consists of a macrocell and a set of femtocells overlaid within the coverage area of the macrocell. The macrocell uses the space division multiple access method to share its frequency band with the femto network.

We assume that there are $K$ cells in the network (including the macrocell), where the base station of cell $k = 1, \ldots, K$ has $M_k$ transmit antennas. The index one is used to represent the macrocell (i.e., $k = 1$ represents macrocell), and the rest (i.e., $k = 2, \ldots, K$) represent femtocells. Base station $k$ serves a predefined set of $I_k$ users, $i = 1, \ldots, I_k$. We assume that $i$th user of base station $k$ has $N_{ki}$ receive antennas. We denote the number of data streams transmitted by $k$th base station to $i$th user as $d_{ki}$.

Assuming linear transmit precoding, the antenna signal vector transmitted by $k$th base station is given by

$$x_k = \sum_{i=1}^{I_k} V_{ki} s_{ki},$$

(3.1)

where $s_{ki} \in \mathbb{C}^{d_{ki}}$ and $V_{ki} \in \mathbb{C}^{M_k \times d_{ki}}$ represent the data stream and the transmit beamformers associated with $i$th user of base station $k$. We assume that the data streams are independent, zero-mean, and normalized to unit variance, i.e., $\mathbb{E}\{s_{ki}s_{ki}^H\} = I$. Moreover, we assume that the total number of data streams transmitted by $k$th base station (i.e., $\sum_{i=1}^{I_k} d_{ki}$) is smaller than $M_k$, i.e., the system is not spatially overloaded.

The signal received at $i$th user of base station $k$ can be expressed as

$$y_{ki} = H_{ki}^k V_{ki} s_{ki} + \sum_{j=1, j \neq i}^{I_k} H_{kj}^k V_{kj} s_{kj} + \sum_{i=1, i \neq k}^{K} \sum_{j=1}^{I_i} H_{ki}^l V_{lj} s_{lj} + n_{ki},$$

(3.2)

where $H_{ki}^l \in \mathbb{C}^{N_{ki} \times M_l}$ represents the channel matrix from $l$th base station to $i$th user of base station $k$, and $n_{ki}$ is circular symmetric complex Gaussian noise with $n_{ki} \sim \mathcal{CN}(0, \sigma^2 I)$. We assume that the channel is quasi-static.
Fig. 3.1. System model for the downlink MIMO heterogeneous network.

We assume a linear receiver \(^8\) where \(i\)th user of base station \(k\) estimates data vector \(s_{ki}\) by linearly combining the received signal vector \(y_{ki}\) with the receive beamformer \(U_{ki}\), i.e., \(s_{ki} = U_{ki}^H y_{ki}\). Thus, the data vector estimated by \(i\)th user of base station \(k\) can be written as

\[
\hat{s}_{ki} = U_{ki}^H y_{ki} = U_{ki}^H H_{ki} V_{ki} s_{ki} + U_{ki}^H \sum_{j=1,j \neq i}^{K} H_{ki}^j V_{kj} s_{kj} + U_{ki}^H \sum_{l=1,l \neq k}^{K} \sum_{j=1}^{L_l} H_{ki}^l V_{lj} s_{lj} + U_{ki}^H n_{ki}.
\]

\(^8\)In general, the linear receiver is sum-rate optimal only when the minimum mean-squared-error (MMSE) matrix is diagonal \([43]\). However, for any precoder matrix, there is a method to obtain a diagonal MMSE matrix. That is, there exists a unitary transformation matrix, which can be used to rotate any precoder matrix at the transmitter side so that the MMSE matrix is diagonal \([122, pp. 2207]\). The rate obtained by this unitary transformed precoder is the same as the rate that can be achieved when the MMSE matrix is diagonal. This transformation does not affect the SDMA precoding, i.e., the rate of other users will not be changed, because the interference-plus-noise covariance is not changed due to this transformation \([43, pp. 4796]\).
The mean-squared-error (MSE) matrix of \(i\)th user of base station \(k\) can be written as
\[
E_{ki} = E\{(\hat{s}_{ki} - s_{ki})(\hat{s}_{ki} - s_{ki})^H\} = I - U_{ki}^H H_{ki}^k V_{ki} - (U_{ki}^H H_{ki}^k V_{ki})^H + U_{ki}^H J_{ki} U_{ki}, \tag{3.4}
\]
where \(J_{ki}\) is the received signal covariance matrix of \(i\)th user of base station \(k\), and is given by
\[
J_{ki} = E\{y_{ki} y_{ki}^H\} = \sum_{l,j} \sum_{i=1}^K I_l \sum_{j=1}^H H_{li}^k V_{lj}^k (H_{li}^k)^H + \sigma^2_{ki} I. \tag{3.5}
\]

For notational simplicity, we also define the interference-plus-noise covariance matrix of \(i\)th user of base station \(k\) as \(\Upsilon_{ki} = \sum_{(l,j) \neq (k,i)} H_{li}^k V_{lj}^k (H_{li}^k)^H + \sigma^2_{ki} I\). Moreover, we assume that all users are capable of estimating \(\Upsilon_{ki}\). Note that the MSE corresponding to \(n\)th data symbol of \(i\)th user of base station \(k\) is the \(n\)th diagonal element of \(E_{ki}\), i.e., \(e_{ki,n} = E\{|\hat{s}_{ki,n} - s_{ki,n}|^2\} = [E_{ki}]_{nn}\). As a result of minimizing the MSE for all the data streams of \(i\)th user of base station \(k\), for a fixed \(V_{ki}\), the linear minimum-mean-squared-error (LMMSE) receiver of \(i\)th user of base station \(k\) is given by [10, Ch. 2]
\[
U_{ki}^{\text{LMMSE}} = J_{ki}^{-1} H_{ki}^k V_{ki}. \tag{3.6}
\]

When \(i\)th user of base station \(k\) employs the LMMSE receiver, the MSE matrix \(E_{ki}\) reduces to [123, Th. 12.1]
\[
E_{ki}^{\text{LMMSE}} = I - (H_{ki}^k V_{ki})^H J_{ki}^{-1} H_{ki}^k V_{ki} = I - (H_{ki}^k V_{ki})^H U_{ki}^{\text{LMMSE}}, \tag{3.7}
\]
where in the second equality we have used the definition of \(U_{ki}^{\text{LMMSE}}\) in (3.6).

We assume that Gaussian signaling is used, and the interference from all other users is treated as noise. Thus, the rate of \(i\)th user of base station \(k\) can be written as [45]
\[
R_{ki} = \log \det \left( I + H_{ki}^k V_{ki} (H_{ki}^k V_{ki})^H \Upsilon_{ki}^{-1} \right). \tag{3.8}
\]

The transmit power constraint at \(k\)th base station is given by
\[
\sum_{i=1}^{I_k} \mathbb{E} \left\{ \|V_{ki} s_{ki}\|^2 \right\} = \sum_{i=1}^{I_k} \text{Tr}(V_{ki} V_{ki}^H) \leq P_{\text{max}}^{\text{max}}, \quad k = 1, \ldots, K, \tag{3.9}
\]
where $P_{\text{max}}^k$ is the maximum transmit power at $k$th base station. The interference power experienced by macro user $j$, due to the communication between $k$th femto base station and $i$th femto user of femto base station $k$ is given by $\text{Tr} (H_{k}^{j} V_{k_i} (H_{k}^{j} V_{k_i})^H)$. Thus, the total interference power generated by $k$th femto base station at macro user $j$ can be expressed as $\sum_{i=1}^{I_k} \text{Tr} (H_{k}^{j} V_{k_i} (H_{k}^{j} V_{k_i})^H)$. We assume that the total interference power generated by $k$th femto base station, over all the receiving antennas at macro user $j$ is limited by a predefined threshold $\theta_{kj}^k$. Thus we have

$$\sum_{i=1}^{I_k} \text{Tr} (H_{k}^{j} V_{k_i} (H_{k}^{j} V_{k_i})^H) \leq \theta_{kj}^k, \quad k = 2, \ldots, K, \ j = 1, \ldots, I_1. \quad (3.10)$$

### 3.2 Problem formulation

The sum-power minimization and the sum-rate maximization problems are commonly used resource allocation criteria in wireless network designs. However, they are competing objectives, because one cannot maximize sum-rate and minimize sum-power simultaneously. Hence, there is a trade-off between these two objectives. In this section, we formulate the problem of finding the set of all achievable power-rate tuples in a single-macrocell MIMO heterogeneous network.

We start by defining the sum-power and the sum-rate functions as

$$P_{\text{sum}}(V) = \sum_{k=1}^{K} I_k \sum_{i=1}^{I_k} \text{Tr}(V_{k_i} V_{k_i}^H), \quad (3.11)$$

and

$$R_{\text{sum}}(V) = \sum_{k=1}^{K} I_k \sum_{i=1}^{I_k} R_{ki}(V), \quad (3.12)$$

respectively, where $V = \{V_{k_i}\}_{k=1, \ldots, K, \ i=1, \ldots, I_k}$ is the collection of all the transmit beamformers in the network. By considering the maximum transmit power constraint in (3.9) and the interference threshold constraint in (3.10), the set of
directly achievable power-rate tuples can be expressed as

\[
\mathcal{O} = \left\{ (P_{\text{sum}}(\mathbf{V}), R_{\text{sum}}(\mathbf{V})) \right\} \quad \begin{aligned}
\sum_{i=1}^{I_k} \text{Tr}(\mathbf{V}_{ki} \mathbf{V}_{ki}^H) &\leq P_{\text{max}}^k, \quad k = 1, \ldots, K \\
\sum_{i=1}^{I_k} \text{Tr}(\mathbf{H}_{ij} \mathbf{V}_{ki} (\mathbf{H}_{ij} \mathbf{V}_{ki})^H) &\leq \theta_{ij}^k \\
& \quad k = 2, \ldots, K, \quad j = 1, \ldots, I_1
\end{aligned}
\]

(3.13)

Here we use the term “directly achievable” to mean that for any point in set \( \mathcal{O} \) there exists a set of transmit precoders \( \mathbf{V} = \{\mathbf{V}_{ki}\}_{k=1}^{K} \) that satisfy the power and interference constraints. In other words, any point in set \( \mathcal{O} \) can be achieved by operating continuously with a single set of precoders. However, the set of achievable power-rate tuples can be further extended to the convex hull of set \( \mathcal{O} \) (i.e., \( \text{conv} \mathcal{O} \)) by considering a transmission strategy which allows time sharing between directly achievable points. Specifically, any power-rate tuple in \( \text{conv} \mathcal{O} \) can be achieved by letting the system to use a set of feasible precoders for a fraction \( f \) of time and another set of feasible precoders for the remaining \((1 - f)\) fraction of time. This procedure is illustrated in Figure 3.2, where points \( A = \{P_i, R_i\} \) and \( B = \{P_j, R_j\} \) are directly achievable (i.e., \( A, B \in \mathcal{O} \)), and point \( E = \{fP_i + (1 - f)P_j, fR_i + (1 - f)R_j\} \) can be achieved by operating on point \( A \) for a fraction \( f \) of time and on point \( B \) for the remaining \((1 - f)\) fraction of time. Clearly, by varying \( f \) between 0 and 1 we can achieve any point on the line segment \( AB \). Thus, we refer to \( \text{conv} \mathcal{O} \) as the set of all achievable power-rate tuples. Note that Figure 2 illustrates the case where points \( A \) and \( B \) are located on the boundary of set \( \mathcal{O} \), but the time sharing argument holds for any location of the points \( A, B \in \mathcal{O} \).

Next, a method to find \( \text{conv} \mathcal{O} \) is provided, where we capitalize on a standard technique called scalarization [21, Sec. 4.7]. The scalarization method is used to solve vector (multi-objective) optimization problems. In this method, we associate some scalars, called weights, with each component of the vector objective, and then add all the weighted objective components. Thus, we transform the vector objective to a scalar objective. It is worth noting that the weight associated with an objective component can be used to control its impact on the other components. For a detailed discussion of the scalarization method we refer the reader to [21, Sec. 4.7].
\[
\lambda P_{\text{sum}} - (1 - \lambda)R_{\text{sum}} = r(< s)
\]

Fig. 3.2. The shaded area shows the set \( \mathcal{O} \) of directly achievable power-rate tuples. The region \( \mathcal{O} \) is the convex hull of set \( \mathcal{O} \). Any point located on the boundary segments \( OA \) and \( BG \) of set \( \mathcal{O} \) can be found by the scalarization method. None of the points located on the boundary segment \( AB \) of set \( \mathcal{O} \) can be found by the scalarization method. Point \( E \) can be obtained by applying the time-sharing method between points \( A \) and \( B \). By moving the hyperplane normal to \( \lambda \) (\( r \) and \( s \) are scalars where \( r < s \)) in \( -\lambda \) direction, we can find \( (P_{\text{sum}}^\ast, R_{\text{sum}}^\ast) \).

As explained above, to find \( \text{conv} \mathcal{O} \) we associate nonnegative weights with the sum-power \( P_{\text{sum}}(V) \) and with the sum-rate \( R_{\text{sum}}(V) \). We denote by \( \lambda \in [0, 1] \) the weight associated with \( P_{\text{sum}}(V) \) and by \( (1 - \lambda) \) the weight associated with \( R_{\text{sum}}(V) \). Next, let us consider the following related problem parameterized by \( \lambda \):

\[
\begin{align*}
\text{minimize} & \quad \lambda \sum_{k=1}^{K} \sum_{i=1}^{I_k} \text{Tr}(V_{ki} V_{ki}^H) - (1 - \lambda) \sum_{k=1}^{K} \sum_{i=1}^{I_k} R_{ki}(V) \\
\text{subject to} & \quad \sum_{i=1}^{I_k} \text{Tr}(V_{ki} V_{ki}^H) \leq P_{\text{max}}^k, \quad k = 1, \ldots, K \\
& \quad \sum_{i=1}^{I_k} \text{Tr} \left( H_{kj}^k V_{ki} (H_{kj}^k V_{ki})^H \right) \leq \theta_{kj}^k, \quad k = 2, \ldots, K, \quad j = 1, \ldots, I_1,
\end{align*}
\]

(3.14)
with variable $V$. Problem (3.14) can be interpreted geometrically as an optimization problem in the graph space $(P_{\text{sum}}; R_{\text{sum}})$, where we minimize $\lambda P_{\text{sum}} - (1 - \lambda) R_{\text{sum}}$ over set $O$, see Figure 3.2. Note that $\lambda = (\lambda, -(1 - \lambda))^T$ represents the gradient of the objective function of problem (3.14) with respect to $(P_{\text{sum}}, R_{\text{sum}})$. In other words, the hyperplane normal to $\lambda$ is moved in the direction of $-\lambda$ until it becomes tangent to $O$ at point $(P_{\text{sum}}^*, R_{\text{sum}}^*)$ (i.e., it becomes a supporting hyperplane of $O$ at $(P_{\text{sum}}^*, R_{\text{sum}}^*)$). Thus, by varying $\lambda$ from 0 to 1, and by solving problem (3.14) for every $\lambda$, one can find sufficient (nontrivial) boundary points for approximating $\text{conv } O$. Note that only the points on the convex boundary of set $O$ (e.g., on the segments OA and BG) can be found via the scalarization method while others (e.g., points C and D) cannot. However, as illustrated in Figure 3.2, these points are not needed for computing $\text{conv } O$. Finally, it is worth pointing out that problem (3.14) parameterized by $\lambda = 0$ is equivalent to the sum-rate maximization.

In heterogeneous networks, there are applications where we might want to differently emphasize the sum-powers of different femtocells and the rates of different users. In such cases we can generalize the two-dimensional scalarization method described above as follows. We denote by $\alpha_{ki}$ the weight associated with the rate of each user and by $\beta_k$ the weight associated with the sum-power of each base station in the network. Then the generalized problem formulation is given by

\[
\begin{align*}
\text{minimize} & \quad \sum_{k=1}^{K} \beta_k \sum_{i=1}^{I_k} \text{Tr}(V_{ki} V_{ki}^H) - \sum_{k=1}^{K} \sum_{i=1}^{I_k} \alpha_{ki} R_{ki} (V) \\
\text{subject to} & \quad \sum_{i=1}^{I_k} \text{Tr}(V_{ki} V_{ki}^H) \leq P_{\text{max}}^k, \quad k = 1, \ldots, K \\
& \quad \sum_{i=1}^{I_k} \text{Tr} \left( H_{kj}^i V_{ki} (H_{kj}^i V_{ki})^H \right) \leq \theta_k^i, \quad k = 2, \ldots, K, \quad j = 1, \ldots, I_1,
\end{align*}
\] (3.15)

with variable $V$. For the sake of completeness, in the following section we provide a solution for problem (3.15) (which includes problem (3.14) as a particular case).

Some optimization problems that exist in many engineering fields naturally appear with multiple objective functions [124, pp. 5-36]. Hence, to see the trade-off between the competing objective functions, it is common to formulate such problems as multi-objective optimization problems. To solve these problems,
apart from the scalarization method used in this chapter, one can use methods such as no-preference methods, a posteriori methods, a priori methods, and interactive methods as described in [124, pp. 61-211], [19, pp. 161-168].

3.3 Algorithm derivation

In this section we derive a linear transceiver design algorithm for problem (3.15). First, we equivalently reformulate problem (3.15) in a more convenient form. To do this we revisit the equivalence between the weighted sum-rate (WSR) maximization and weighted-sum-mean-squared-error (WMMSE) minimization problems [45]. Then we derive the proposed linear transceiver design algorithm.

3.3.1 A useful reformulation

We start by defining the following cost function

\[
c(V, U, W) = \sum_{k=1}^{K} \beta_{k} \sum_{i=1}^{I_{k}} \text{Tr}(V_{ki} V_{ki}^{H}) + \sum_{k=1}^{K} \sum_{i=1}^{I_{k}} \alpha_{ki} \left( \text{Tr}(W_{ki} E_{ki}) - \log \det(W_{ki}) \right),
\]

with variables \( V, \{ U_{ki} \}_{i=1}^{I_{k}}, \) and \( \{ W_{ki} \}_{i=1}^{I_{k}}, \) where \( W_{ki} \in \mathbb{C}^{d_{ki} \times d_{ki}} \) is a nonzero positive semidefinite weight matrix associated with \( i \)th user of base station \( k, \) \( W_{ki} \succeq 0. \) For the sake of brevity, let us denote \( U = \{ U_{ki} \}_{i=1}^{I_{k}}, \) and \( W = \{ W_{ki} \}_{i=1}^{I_{k}}. \)

Now we consider the following optimization problem.

\[
\begin{align*}
\text{minimize} & \quad \sum_{k=1}^{K} \beta_{k} \sum_{i=1}^{I_{k}} \text{Tr}(V_{ki} V_{ki}^{H}) + \sum_{k=1}^{K} \sum_{i=1}^{I_{k}} \alpha_{ki} \left( \text{Tr}(W_{ki} E_{ki}) - \log \det(W_{ki}) \right) \\
\text{subject to} & \quad \sum_{i=1}^{I_{k}} \text{Tr}(V_{ki} V_{ki}^{H}) \leq P_{k}^{\text{max}}, \quad k = 1, \ldots, K \\
& \quad \sum_{i=1}^{I_{k}} \text{Tr} \left( H_{kj}^{\text{H}} V_{ki} (H_{kj}^{H} V_{ki})^{\text{H}} \right) \leq \theta_{kj}, \quad k = 2, \ldots, K, \quad j = 1, \ldots, I_{1},
\end{align*}
\]

with variables \( V, U, \) and \( W. \) Note that problem (3.17) minimizes the cost function (3.16), subject to the maximum transmit power constraint in (3.9) and the interference threshold constraint in (3.10).
Following a methodology similar to the one used in \cite[Theorem 1]{45}, in the sequel, we show that problem (3.17) is equivalent to problem (3.15), in the sense that they achieve local optimal solutions for identical precoders $V$. To do this, let us first consider the partial minimization of $c(V, U, W)$ over $U, W$. We denote by $f(V)$ the corresponding optimal value of problem (3.17) for fixed $V$ and this is given by equation (3.18)

$$f(V) = \inf_{W,U} c(V, U, W) = \inf_{W} \inf_{U} c(V, U, W)$$

(3.18a)

$$= \inf_{W} \inf_{U} \left( \sum_{k=1}^{K} \beta_k \sum_{i=1}^{I_k} \text{Tr}(V_{ki}V_{ki}^H) + \sum_{k=1}^{K} \sum_{i=1}^{I_k} \alpha_{ki} \left( \text{Tr}(W_{ki}E_{ki}) - \log \det(W_{ki}) \right) \right)$$

(3.18b)

$$= \inf_{W} \left( \sum_{k=1}^{K} \beta_k \sum_{i=1}^{I_k} \text{Tr}(V_{ki}V_{ki}^H) + \sum_{k=1}^{K} \sum_{i=1}^{I_k} \alpha_{ki} \left( \text{Tr}(W_{ki}E_{MMSE}^{ki}) - \log \det(W_{ki}) \right) \right)$$

(3.18c)

$$= \sum_{k=1}^{K} \beta_k \sum_{i=1}^{I_k} \text{Tr}(V_{ki}V_{ki}^H) + \sum_{k=1}^{K} \sum_{i=1}^{I_k} \alpha_{ki} \left( \text{Tr} \left( \frac{(E_{MMSE}^{ki})^{-1}E_{MMSE}^{ki}}{1} \right) \right) - \log \det((E_{MMSE}^{ki})^{-1})$$

(3.18d)

$$= \sum_{k=1}^{K} \beta_k \sum_{i=1}^{I_k} \text{Tr}(V_{ki}V_{ki}^H) + \zeta - \sum_{k=1}^{K} \sum_{i=1}^{I_k} \alpha_{ki} \log \det \left( I + (H_{ki}^k V_{ki})^H Y^{-1} H_{ki}^k V_{ki} \right)$$

(3.18e)

$$= \sum_{k=1}^{K} \beta_k \sum_{i=1}^{I_k} \text{Tr}(V_{ki}V_{ki}^H) + \zeta - \sum_{k=1}^{K} \sum_{i=1}^{I_k} \alpha_{ki} \log \det \left( I + H_{ki}^k V_{ki} (H_{ki}^k V_{ki})^H Y^{-1} \right)$$

(3.18f)

$$= \sum_{k=1}^{K} \beta_k \sum_{i=1}^{I_k} \text{Tr}(V_{ki}V_{ki}^H) + \zeta - \sum_{k=1}^{K} \sum_{i=1}^{I_k} \alpha_{ki} R_{ki}.$$  

(3.18g)

Here (3.18a) and (3.18b) follow from straightforward substitutions. The equality (3.18c) follows from the fact that given $V$ and $W_{ki} \succeq 0$, $\text{Tr}(W_{ki}E_{ki})$ is minimized when $E_{ki} = E_{MMSE}^{ki}$, where $E_{MMSE}^{ki}$ is given in (3.7) and is achieved with the receive beamformer $U_{MMSE}^{ki}$, see (3.6). The equality (3.18d) follows by setting the gradient of $\text{Tr}(W_{ki}E_{MMSE}^{ki}) - \log \det(W_{ki})$ to 0, which implies that $W_{ki} = (E_{MMSE}^{ki})^{-1}$. Finally, (3.18e) follows from (3.7), where the constant $\zeta = \sum_{k=1}^{K} \sum_{i=1}^{I_k} \alpha_{ki} \text{Tr}(I_{ki})$.  

98
(3.18f) follows from $\det(I + AB) = \det(I + BA)$, and (3.18g) follows from (3.8). Note that $f(V)$ differs to the objective function of problem (3.15) only by a constant. Thus, by noting that the objective function of problem (3.17) is equivalent to that of problem (3.15) (according to the above proof) and combining the fact that the problems (3.17) and (3.15) have identical constraints, we can see that problem (3.15) and problem (3.17) are equivalent.

From here onwards, we consider the problem (3.17) instead of (3.15). It is worth noting that unlike problem (3.15), problem (3.17) has desirable properties, which can be used to design efficient algorithms based on alternating optimization methods, as we will see next.

3.3.2 Solution via alternating optimization method

Note that problem (3.17) is not jointly convex in variables $V$, $U$, and $W$. Moreover, it is an NP-hard problem. This can be easily verified by noting that weighted sum-rate maximization is NP-hard [27, 45], which is a special case of problem (3.17), i.e., $\beta_k = 0$ for all $k = 1, \ldots, K$. As a result, to find the solution, one has to rely on exponentially complex global optimization approaches, which are usually intractable for large scale problems. Therefore, in the following we present a suboptimal method to solve problem (3.17).

First note that problem (3.17) is convex in each optimization variable $V$, $U$, and $W$ [21, pp. 67-113], [45]. Hence, we can apply the alternating optimization method [24, pp. 272-276]. The key idea is to perform the optimization first with respect to $U$ while keeping $V$ and $W$ fixed, then optimize with respect to $W$ while keeping $U$ and $V$ fixed, and next optimize with respect to $V$ while keeping $U$ and $W$ fixed. The process is repeated until an approximate solution can be found for problem (3.17). Figure 3.3 illustrates the idea indicated above. Note that the solution of (3.17) obtained by this method is a locally optimal (suboptimal) one, and the global optimality can not be guaranteed as the problem is NP-hard.

Note from (3.18c) that for fixed $V$ and $W$, the optimal $U$ of problem (3.17) is given by $\{U^*_k = U^\text{NMBE}_{k}\}_{k=1,\ldots,K}$. Moreover, for fixed $V$ and $U$, the optimal $W$ of problem (3.17) is given by $\{W^*_k = (E_k^{\text{NMBE}})^{-1}\}_{k=1,\ldots,K}$, compare with (3.18d). Furthermore, for fixed $U$ and $W$, problem (3.17) can be decoupled into $K$. 99
Fig. 3.3. Illustration of applying the alternating optimization method to solve problem (3.17).

subproblems (one for each base station). Then, kth subproblem can be written as follows:

\[
\begin{aligned}
\text{minimize} & \quad \beta_k \sum_{i=1}^{I_k} \text{Tr}(V_{ki}V_{ki}^H) + \\
& \quad \sum_{i=1}^{I_k} \text{Tr} \left( \alpha_{ki} W_{ki}(I - U_{ki}^H H_{ki}^k V_{ki})(I - U_{ki}^H H_{ki}^k V_{ki})^H \right) + \\
& \quad \sum_{i=1}^{I_k} \sum_{(l,j) \neq (k,i)} \text{Tr} \left( \alpha_{lj} W_{lj}(U_{lj}^H H_{lj}^k V_{ki})(U_{lj}^H H_{lj}^k V_{ki})^H \right)
\end{aligned}
\] (3.19)

subject to

\[
\begin{aligned}
\sum_{i=1}^{I_k} \text{Tr}(V_{ki}V_{ki}^H) & \leq P_{max}^k \\
\sum_{i=1}^{I_k} \text{Tr}(H_{lj}^k V_{ki}(H_{lj}^k V_{ki})^H) & \leq \theta_{lj}^k, \quad j = 1, \ldots, I_1,
\end{aligned}
\]

with variables \( \{V_{ki}\}_{i=1,\ldots,I_k} \). Unlike \( U_{ki}^* \) and \( W_{ki}^* \), the solution \( \{V_{ki}^*\}_{i=1,\ldots,I_k} \) of problem (3.19) does not have a closed-form expression. Nevertheless, we can equivalently reformulate problem (3.19) as a quadratically constrained quadratic program (QCQP) \cite[Sec 4.4]{21} as we will show next.

The objective function of problem (3.19) consists of quadratic, linear, and constant terms, and by separating those terms we can rewrite the objective function of problem (3.19) as follows:
$$c_k = \sum_{i=1}^{I_k} \left( \beta_k \text{Tr}(V_{ki} V_{ki}^H) + \text{Tr}(\alpha_{ki} W_{ki} U_{ki}^H H_{ki}^H V_{ki} U_{ki}^H H_{ki}^H V_{ki}) \right)$$

$$+ \sum_{(l,j) \neq (i,k)} \text{Tr}(\alpha_{lj} W_{lj} U_{lj}^H H_{lj}^H V_{ki})$$

$$- \text{Tr} \sum_{i=1}^{I_k} (\alpha_{ki} W_{ki} U_{ki}^H H_{ki}^H V_{ki}) + \text{Tr}(\alpha_{ki} W_{ki} (U_{ki}^H H_{ki}^H V_{ki})$$

$$+ \text{Tr} \sum_{i=1}^{I_k} (\alpha_{ki} W_{ki}) \right).$$

(3.20)

Now consider the quadratic terms associated with $i$th user of base station $k$ in (3.20), and we can rewrite it compactly as

$$\beta_k \text{Tr}(V_{ki} V_{ki}^H) + \text{Tr}(\alpha_{ki} W_{ki} U_{ki}^H H_{ki}^H V_{ki} U_{ki}^H H_{ki}^H V_{ki})$$

$$+ \sum_{(l,j) \neq (i,k)} \text{Tr}(\alpha_{lj} W_{lj} U_{lj}^H H_{lj}^H V_{ki}(U_{lj}^H H_{lj}^H V_{ki})$$

$$- \text{Tr} \sum_{i=1}^{I_k} (\alpha_{ki} W_{ki} U_{ki}^H H_{ki}^H V_{ki} + \alpha_{ki} W_{ki} (U_{ki}^H H_{ki}^H V_{ki})$$

$$+ \text{Tr} \sum_{i=1}^{I_k} (\alpha_{ki} W_{ki}).$$

(3.21)

Thus, we can equivalently rewrite the quadratic part of (3.20) as

$$\text{Tr}(V_{ki}^H Z_{ki} V_{ki} + V_{ki}^H Z_{ki} V_{ki} + \cdots + V_{ki}^H Z_{ki} V_{ki} = \text{Tr}(V_{ki}^H Z_{ki} V_{ki}),$$

(3.23)

where the matrices $V_k \in \mathbb{C}^{(l_k \times M_k) \times \sum_{i=1}^{I_k} d_{ki}}$ and $Z_k \in \mathbb{C}^{(l_k \times M_k) \times (l_k \times M_k)}$ are defined as follows:

$$V_k = \begin{bmatrix}
V_{k1} & O_{M_k \times d_{k2}} & \cdots & O_{M_k \times d_{kI_k}} \\
O_{M_k \times d_{k1}} & V_{k2} & \cdots & O_{M_k \times d_{kI_k}} \\
& \ddots & \ddots & \vdots \\
& & O_{M_k \times d_{k1}} & O_{M_k \times d_{k2}} & \cdots & V_{kI_k}
\end{bmatrix}.$$ 

(3.24)
and
\[
Z_k = \begin{bmatrix}
Z_{k1} & O_{M_k \times M_k} & \cdots & O_{M_k \times M_k} \\
O_{M_k \times M_k} & Z_{k2} & O_{M_k \times M_k} & \cdots & O_{M_k \times M_k} \\
& \vdots & \ddots & \vdots \\
O_{M_k \times M_k} & O_{M_k \times M_k} & \cdots & Z_{kI_k}
\end{bmatrix}.
\] (3.25)

Next, we equivalently rewrite the linear term in equation (3.20) as
\[
\text{Tr} \left( \sum_{i=1}^{I_k} (\alpha_{ki} W_{ki} U_{ki} H_{ki}^H V_{ki} + (\alpha_{ki} W_{ki} U_{ki} H_{ki}^H V_{ki})^H) \right)
= \text{Tr} \left( \sum_{i=1}^{I_k} (V_{ki}^H B_{ki} + V_{ki} B_{ki}^H) \right)
= \text{Tr} \left( (B_k^H V_k) + B_k V_k \right),
\] (3.26)

where \( B_{ki} = \alpha_{ki} W_{ki} U_{ki} H_{ki}^H \), and \( B_k \in \mathbb{C}^{(\sum_{i=1}^{I_k} d_k) \times (I_k \times M_k)} \) is given by
\[
B_k = \begin{bmatrix}
B_{k1} & O_{d_k \times M_k} & \cdots & O_{d_k \times M_k} \\
O_{d_k \times M_k} & B_{k2} & O_{d_k \times M_k} & \cdots & O_{d_k \times M_k} \\
& \vdots & \ddots & \vdots \\
O_{d_k \times M_k} & O_{d_k \times M_k} & \cdots & B_{kI_k}
\end{bmatrix}.
\] (3.27)

Now let \( c \) be the constant term in equation (3.20), and it can be expressed as
\[
c = \text{Tr} \left( \sum_{i=1}^{I_k} (\alpha_{ki} W_{ki}) \right).
\] (3.28)

We can remove the constant \( c \) from the objective function, as it does not affect solution of the problem. Hence, the reformulated objective function of problem (3.19) is given by
\[
c_k = \text{Tr}(V_k^H Z_k V_k) - 2\Re(\text{Tr}(B_k V_k)).
\] (3.29)

Finally, we reformulate the set of constraints in problem (3.19) in quadratic matrix function form, as described below. Note that the constraints in problem (3.19) have only quadratic terms. Therefore, the equivalent quadratic matrix function of each constraint contains only a quadratic term. Let us first consider the power constraint, i.e., the first constraint in problem (3.19). It can be reformulated as
\[
\text{Tr}(V_k^H V_k) \leq P_{k_{\text{max}}},
\] (3.30)
where $\mathbf{V}_k$ is defined as in (3.24). Now consider the set of interference threshold constraints, i.e., the second set of constraints in problem (3.19). They can be expressed as

$$\text{Tr}(\mathbf{V}_k^H \mathbf{H}_{1j}^k \mathbf{V}_k) \leq \theta_{1j}^k, \quad \forall j = 1, \ldots, I_1,$$

(3.31)

where $\mathbf{V}_k$ is defined as in (3.24), and $\mathbf{H}_{1j}^k \in \mathbb{C}^{(I_k \times M_k) \times (I_k \times M_k)}$ is given by

$$\mathbf{H}_{1j}^k = \begin{bmatrix} \mathbf{H}_{1j}^k \mathbf{H}_{1j}^k & \mathbf{O}_{M_k \times M_k} & \cdots & \mathbf{O}_{M_k \times M_k} \\ \mathbf{O}_{M_k \times M_k} & (\mathbf{H}_{1j}^k)^H \mathbf{H}_{1j}^k & \cdots & \mathbf{O}_{M_k \times M_k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O}_{M_k \times M_k} & \mathbf{O}_{M_k \times M_k} & \cdots & (\mathbf{H}_{1j}^k)^H \mathbf{H}_{1j}^k \end{bmatrix}.$$  

(3.32)

Thus, by using the equivalent reformulations of the objective function (i.e., (3.29)) and the constraint functions (i.e., (3.30) and (3.31)), we equivalently reformulate problem (3.19) as

$$\begin{align*}
\text{minimize} & \quad \text{Tr}(\mathbf{V}_k^H \mathbf{Z}_k \mathbf{V}_k) - 2\Re\left(\text{Tr}(\mathbf{B}_k \mathbf{V}_k)\right) \\
\text{subject to} & \quad \text{Tr}(\mathbf{V}_k^H \mathbf{V}_k) \leq P_{\text{max}}^k \\
& \quad \text{Tr}(\mathbf{V}_k^H \mathbf{H}_{1j}^k \mathbf{V}_k) \leq \theta_{1j}^k, \quad j = 1, \ldots, I_1,
\end{align*}$$

(3.33)

where the variable is $\mathbf{V}_k$.

Note that problem (3.33) contains a transmit power constraint and a set of interference threshold constraints. Recall that we impose the interference threshold constraints to limit the interference generated by femto base stations towards the macro users. Hence, the interference threshold constraints should be considered only in the subproblems specified for femtocells (i.e., for $k \neq 1$). Thus, to find the optimal transmit beamformers associated with macro users (i.e., $\mathbf{V}_{1i}$ for all $i = 1, \ldots, I_1$) we must solve subproblem (3.33) without the interference threshold constraints in problem (3.33). In the sequel we express the subproblems associated with the macrocell (i.e., for $k = 1$) and with $k$th femtocell separately.

1. The subproblem of finding the set of precoders at macrocell

$$\begin{align*}
\text{minimize} & \quad \text{Tr}(\mathbf{V}_1^H \mathbf{Z}_1 \mathbf{V}_1) - 2\Re\left(\text{Tr}(\mathbf{B}_1 \mathbf{V}_1)\right) \\
\text{subject to} & \quad \text{Tr}(\mathbf{V}_1^H \mathbf{V}_1) \leq P_{\text{max}}^1,
\end{align*}$$

(3.34)

where the variable is $\mathbf{V}_1$. 

103
2. The subproblem of finding the set of precoders at $k$th femtocell

$$\begin{align*}
\text{minimize} & \quad \text{Tr}(V_k^H Z_k V_k) - 2\Re\{\text{Tr}(B_k V_k)\} \\
\text{subject to} & \quad \text{Tr}(V_k^H V_k) \leq P_{\text{max}}^k \\
& \quad \text{Tr}(V_k^H H_{kj}^k V_k) \leq \theta_{kj}, \quad j = 1, \ldots, I_k,
\end{align*}$$

(3.35)

where the variable is $V_k$, and $k = 2, \ldots, K$.

### 3.3.3 Proposed transceiver design algorithm

In this section, we summarize the proposed transceiver design algorithm for determining the convex hull of the set of directly achievable power-rate tuples. The proposed algorithm returns a point on the boundary of $\text{conv } C$ for a given set of weights. To lighten the notations, let us denote

$$\begin{align*}
\alpha &= \{\alpha_{ki}\}_{k=1,\ldots,K}^{i=1,\ldots,I_k} \\
\beta &= \{\beta_k\}_{k=1,\ldots,K}.
\end{align*}$$

(3.36)

**Algorithm 3.1.** Transceiver design algorithm to obtain the convex hull of the set of directly achievable power-rate tuples

1. Initialization: given maximum transmit power $\{P_{\text{max}}^k\}_{k=1,\ldots,K}$, interference threshold $\{\theta_{kj}^k\}_{k=1,\ldots,K}^{j=1,\ldots,I_k}$, weights $\alpha$ and $\beta$, and initial transmit beamformers $\{V_{ki}^{(0)}\}_{k=1,\ldots,K}^{i=1,\ldots,I_k}$. Set the iteration index $t = 0$.

2. Compute $U_{ki}^{(t+1)}$ for all $k = 1, \ldots, K$ and $i = 1, \ldots, I_k$ as follows:

$${U}_{ki}^{(t+1)} = (J_{ki}^{(t)})^{-1} \text{H}_{ki}^k V_{ki}^{(t)}.$$ 

3. Compute $W_{ki}^{(t+1)}$ for all $k = 1, \ldots, K$ and $i = 1, \ldots, I_k$ as follows:

$${W}_{ki}^{(t+1)} = (1 - (H_{ki}^k V_{ki}^{(t)})^H U_{ki}^{(t+1)})^{-1}.$$ 

4. For $k = 1$ to $K$

   a) If $k = 1$: Compute $V_{1i}^{(t+1)}$ for $i = 1, \ldots, I_1$ using problem (3.34).

   b) Else: Compute $V_{ki}^{(t+1)}$ for all $k = 2, \ldots, K$ and $i = 1, \ldots, I_k$ using problem (3.35).
5. Stopping criterion: if the stopping criterion is satisfied, STOP. Otherwise set $t = t + 1$, and go to step (2).

The first step initializes the algorithm. In step (2) we update $U_{ki}$ for all $k = 1, \ldots, K$ and $i = 1, \ldots, I_k$. In step (3) we update $W_{ki}$ for all $k = 1, \ldots, K$ and $i = 1, \ldots, I_k$. Next, in step (4) we update $V_{ki}$ for all $k = 1, \ldots, K$ and $i = 1, \ldots, I_k$. To update the transmit beamformers of macro users (i.e., $V_{ki}$ for all $i = 1, \ldots, I_k$) we use problem (3.34), and for updating the transmit beamformers of femto users (i.e., $V_{ki}$ for all $k = 2, \ldots, K$ and $i = 1, \ldots, I_k$) we use problem (3.35). Finally, step (5) checks the stopping criterion. The algorithm terminates if the stopping criterion is satisfied. Otherwise, the algorithm continues iterating from steps (2)-(5) until the stopping criterion is satisfied. We use the stopping criterion $\left| c^{(t+1)}(V, U, W) - c^{(t)}(V, U, W) \right| \leq \epsilon$, where $c^{(t)}(V, U, W)$ is the objective function of problem (3.17) at $t$th iteration, and $\epsilon$ is a small value. The proposed Algorithm 3.1 is a block coordinate descent algorithm, and it solves a convex problem at each block of the algorithm (i.e., in steps (2)-(4)). Hence, Algorithm 3.1 converges to a stationary point of problem 3.17 [24, pp. 272-276].

Here we discuss a few important facts about the proposed Algorithm 3.1. In the system model we assumed that the number of independent data streams transmitted by $k$th base station should be smaller than the number of antennas at the base station, i.e., $\sum_{i=1}^{I_k} d_{ki} \leq M_k$. We can relax this assumption while implementing the proposed algorithm. This is because in the transmit beamformer optimization step (i.e., in step (4)), the transmit beamformers associated with some data streams are set to zero, so that the system will not be spatially overloaded. To initialize the algorithm, we consider the following procedure. We initialize $V_{ki}$ for all $i = 1, \ldots, I_k$ by equally dividing the maximum transmit power (i.e., $P_{k}^{\text{max}}$) between all the users associated with $k$th base station. Note that the proposed algorithm is intended to be used for finding power-rate regions of heterogeneous networks, and it can be used to evaluate the impact of co-channel deployment and the impact of imposing an interference threshold on practical systems.
3.4 Numerical results

In this section we run the proposed Algorithm 3.1 in a single-macrocell MIMO heterogeneous network environment, and numerically evaluate the convex hull of the set of directly achievable power-rate tuples for different scenarios. First, we draw a plot illustrating the convergence behavior of the proposed algorithm. Then we obtain the convex hull of the set of directly achievable power-rate tuples for given power and interference threshold constraints. Next, we evaluate the impact of imposing an interference threshold constraint on a heterogeneous network. Finally, we investigate the impact (on the power-rate region) of co-channel deployment in a heterogeneous network.

In our simulation model, we consider that a macrocell is overlaid by ten femtocells, as shown in Figure 3.4. We assume that the base station of each cell has a circular coverage region, where the radius of the macrocell transmission region is $R$, and the radius of the femtocell transmission region is $r$. For simplicity, we consider there is only a single user per cell. Furthermore, we assume that all base stations are equipped with $M_k = 3$ antennas for all $k = 1, \ldots, K$ and all mobile users are equipped with $N_{ki} = 2$ for all $k = 1, \ldots, K$, $i = 1, \ldots, I_k$ antennas.

The femtocells are uniformly distributed within a region, which is defined by a circle with radius $R$ (i.e., the outer boundary is defined by the macrocell transmission region) and a circle with radius $R/3$ (i.e., the inner boundary is defined by a circle with radius $R/3$) centered at the macro base station, as illustrated in Figure 3.4. The reason for defining the inner exclusion region (i.e., the inner boundary with radius $R/3$) is to avoid deploying femtocells very close to the macro base station, and therefore femto users will not experience a very high interference from the macro base station. Note that in a typical indoor femtocell deployment, the high interference generated by macro base station towards the femto users is reduced due to wall attenuation. However, since we are not modelling the wall attenuation loss in our path loss model (this will be discussed next in this section), it is useful to define an exclusion zone to avoid deploying femtocells overly close to the macro base station. In our simulations, the femtocells are deployed so that their coverage areas do not overlap. Furthermore, we assume that the macro and femto users are uniformly distributed, and we consider that the macro users are not inside the femtocells.
We also assume that the macro users are not located inside the femtocells, because when a macro user is inside a femtocell, then that macro user has a better signal strength from the femto base station than the macro base station, and therefore that macro user can be connected to the femto base station. This scenario can be considered as a base station association problem, which is beyond the scope of the study considered in this chapter.

![Diagram showing the deployment of femtocells within the coverage area of the macrocell and a femtocell.](image)

**Fig. 3.4.** (a) Ten femtocells are deployed within the coverage area of the macrocell. The femtocells can be deployed only outside of an exclusion zone with radius $R/3$. (b) A femtocell.

We use an exponential path loss model, where the channel gains between base station $l$ and $i$th user of base station $k$ are given by

$$H_{ki}^l = \left(\frac{d_{ki}^l}{d_0}\right)^{-\eta/2} C_{ki}^l,$$  

(3.37)

where $d_{ki}^l$ is the distance from $l$th base station to $i$th user of base station $k$, $d_0$ is the far field reference distance [120], $\eta$ is the path loss exponent, and $C_{ki}^l \in \mathbb{C}^{N_i \times M_l}$ is arbitrarily chosen from the distribution $\mathcal{CN}(0, 1)$ (i.e., frequency-flat fading with uncorrelated antennas). The first term in (3.37) represents the path loss component and the second term denotes the Rayleigh small-scale fading. Here, we refer to an arbitrarily generated set of fading coefficients $\mathcal{C} = \{C_{ki}^l | l, k = 1, \ldots, K, i = 1, \ldots, I_k\}$ as a single fading realization. Furthermore, we refer to
an arbitrarily generated set of distances \( D = \{d_{ki}^l | l, k = 1, \ldots, K, i = 1, \ldots, I_k\} \) as a single topology realization.

We define the signal-to-noise-ratio (SNR) operating point at a distance \( d \) as

\[
\text{SNR}(d) = \frac{P_{\text{max}}^k}{\sigma^2} \left( \frac{d}{d_0} \right)^{-\eta}.
\]  

(3.38)

In our simulations, we set \( \frac{P_{\text{max}}^k}{\sigma^2} = 45 \text{dB} \), \( \frac{P_{\text{max}}^k}{\sigma^2} = 35 \text{dB} \) for all \( k = 2, \ldots, 11 \), \( \eta = 4 \), and \( d_0 = 3 \) for femtocells, \( d_0 = 10 \) for the macrocell, \( \text{SNR}(R) = 5 \text{dB} \) for the macrocell and \( \text{SNR}(r) = 10 \text{dB} \) for the femtocells. Furthermore, we assume that the interference threshold for \( i \)th macro user is \( \theta_{ki}^1 = 5\sigma^2 \) for all \( k = 2, \ldots, 11 \).

Recall that the proposed Algorithm 3.1 is designed to solve the general problem (3.15). There, we associate different weights with the sum-power of each base station and the rate of each user. However, when the number of weights is greater than three, the individual effect of varying each weight on the achievable region cannot be graphically depicted. Therefore, in our simulations we use the proposed algorithm to numerically evaluate \( \text{conv } O \) for a special case of problem (3.15), i.e., problem (3.14). To do this we set the weights as \([\alpha]_i = 1 - \lambda \) and \([\beta]_j = \lambda \) in our simulations.

Note that in problem (3.14) the weight associated with \( P_{\text{sum}}(V) \), i.e., \( \lambda \), can be considered as the emphasis on \( P_{\text{sum}}(V) \), and the weight associated with \( R_{\text{sum}}(V) \), i.e., \( 1 - \lambda \), can be considered as the emphasis on \( R_{\text{sum}}(V) \). Hence, to obtain the minimum \( P_{\text{sum}}(V) \) we associate a large weight with \( P_{\text{sum}}(V) \) (i.e., we set \( \lambda = 1 \)), and if we want \( P_{\text{sum}}(V) \) to be maximum we assign a small weight to \( P_{\text{sum}}(V) \) (i.e., we set \( \lambda = 0 \)). Here, we achieve maximum \( R_{\text{sum}}(V) \) for the maximum \( P_{\text{sum}}(V) \), and the minimum \( R_{\text{sum}}(V) \) for the minimum \( P_{\text{sum}}(V) \). Thus, for \( \lambda = 0 \) and \( \lambda = 1 \) we achieve extreme values for \( P_{\text{sum}}(V) \) and \( R_{\text{sum}}(V) \). All the other values for \( P_{\text{sum}}(V) \) and \( R_{\text{sum}}(V) \) can be achieved by selecting \( \lambda \in (0, 1) \).

In order to select \( \lambda \) in the interval (0, 1), we first choose a set of uniformly distributed values for \( \lambda \) in the interval (0, 1) with a chosen step size. Then by solving optimization problem (14) for the selected \( \lambda \in (0, 1) \), we evaluate \( P_{\text{sum}}(V) \) using (11) and \( R_{\text{sum}}(V) \) using (12). Next, we plot the obtained values of \( P_{\text{sum}}(V) \) and \( R_{\text{sum}}(V) \) in the graph space \((P_{\text{sum}}, R_{\text{sum}})\). By observing this graph, we identify the regions where \( R_{\text{sum}}(V) \) shows a rapid change for a small change of \( P_{\text{sum}}(V) \), and within these regions we select more values for \( \lambda \) with smaller step sizes.
Figure 3.6 shows the convergence behavior of the proposed Algorithm 3.1, obtained by changing the weights $\alpha$ and $\beta$ in the algorithm. All the plots in Figure 3.6 are drawn for the topology configuration given in Figure 3.5.

Figure 3.5. A random deployment of ten femtocells within the macrocell. Each femtocell serves a single femto user and the macrocell serves a single macro user.

Figure 3.6(a) shows the convergence behavior of Algorithm 3.1 for $\lambda = 0$ (i.e., $[\alpha]_i = 1$ and $[\beta]_j = 0$). In fact, for this weight configuration, problem (3.15) now reduces to a sum-rate maximization problem. Figure 3.6(b) shows the convergence behavior of the algorithm for $\lambda = 0.5$ (i.e., $[\alpha]_i = 0.5$ and $[\beta]_j = 0.5$). Note that the objective value in Figure 3.6(b) is less than that in Figure 3.6(a). Hence, the results show that by changing the weights associated with $P_{\text{sum}}(V)$ and $R_{\text{sum}}(V)$, the proposed algorithm can be configured to minimize the sum-power, maximize the sum-rate or to optimize both the sum-power and sum-rate simultaneously, based on the selected value of $\lambda$.

Next, we numerically evaluate $\mathbf{conv} \mathcal{O}$ for a heterogeneous network defined by the simulation parameters described earlier in this section. In fact, as explained in Section 3.2 $\mathbf{conv} \mathcal{O}$ represents the achievable power-rate region. Inspired by the discussion on defining the rate regions in [25, pp. 52-53], we follow a similar means to define the directly achievable power-rate region, the instantaneous power-rate region, and the average power-rate region. For a given
Fig. 3.6. $R_{\text{sum}}$ versus iterations for different weights: (a) $\lambda = 0$; (b) $\lambda = 0.5$ ([69] ©2013 IEEE).

fading realization $C^t$, topology realization $D^t$, and maximum transmit power $P_{\text{max}}^k$ for all $i = 1, \ldots, K$, the directly achievable power-rate region is given by

$$R_{\text{DIR}}(C^t, D^t, P_{\text{max}}^1, \ldots, P_{\text{max}}^k) = \left\{ (P_{\text{sum}}(V), R_{\text{sum}}(V)) \left| \begin{array}{l}
\sum_{i=1}^{I_k} \text{Tr}(V_{ki} V_{ki}^H) \leq P_{\text{max}}^k, \ k = 1, \ldots, K \\
\sum_{i=1}^{I_k} \text{Tr}(H_{1j}^{k_i} V_{ki} (H_{1j}^{k_i} V_{ki})^H) \leq \theta_{1j}^k, \\
k = 2, \ldots, K, \ j = 1, \ldots, I_1
\end{array} \right. \right\}. \quad (3.39)$$

Here, we write $H_{1j}^{k_i}$ instead of $(H_{1j}^{k_i})^T$ to lighten the notation. The instantaneous power-rate region $R_{\text{INS}}(C^t, D^t, P_{\text{max}}^1, \ldots, P_{\text{max}}^k)$ can be obtained by applying the time sharing argument to $R_{\text{DIR}}(C^t, D^t, P_{\text{max}}^1, \ldots, P_{\text{max}}^k)$. That is,

$$R_{\text{INS}}(C^t, D^t, P_{\text{max}}^1, \ldots, P_{\text{max}}^k) = \text{conv} \{ R_{\text{DIR}}(C^t, D^t, P_{\text{max}}^1, \ldots, P_{\text{max}}^k) \}, \quad (3.40)$$

Note that we can find a boundary point in the instantaneous power-rate region, by solving the optimization problem (3.15) for the given $[\alpha]_i = \lambda$, and $[\beta]_j = 1 - \lambda$, for some $\lambda \in [0, 1]$. 

110
Fig. 3.7. (a) Instantaneous power-rate region. (b) Average power-rate region: shaded area shows the set of all achievable power-rate tuples of a single-macrocell MIMO heterogeneous network, i.e., $\text{conv } \mathcal{O}$. The dark line outlines the boundary of $\text{conv } \mathcal{O}$ ([69] ©2013 IEEE).
We define the average power-rate region $R_{\text{INS}}^i(P_1^{\max}, \ldots, P_k^{\max})$ for given maximum transmit power $P_k^{\max}$ for all $i = 1, \ldots, K$ as

$$R_{\text{AVG}}(P_1^{\max}, \ldots, P_k^{\max}) = \frac{1}{T} \sum_{t=1}^{T} R_{\text{INS}}(C^{t}, D^{t}, P_1^{\max}, \ldots, P_k^{\max}).$$

where addition and scalar multiplication of sets is used [21, pg. 66], and $T$ denotes the number of fading and topology realizations we used in averaging. Let $[P_{\text{sum}}^{b}(V), R_{\text{sum}}^{b}(V)]$ denote any power-rate tuple located on the boundary of $R_{\text{AVG}}(P_1^{\max}, \ldots, P_k^{\max})$. Note that any boundary point $[P_{\text{sum}}^{b}(V), R_{\text{sum}}^{b}(V)]$ of $R_{\text{AVG}}(P_1^{\max}, \ldots, P_k^{\max})$ can be obtained by using the following steps: (a) solve problem (3.15) for $[\alpha]_i = \lambda$, and $[\beta]_j = 1 - \lambda$, where $\lambda \in [0, 1]$, (b) for any fading realization and topology realization $t \in \{1, \ldots, T\}$, evaluate the sum-power $P_{\text{sum}}^{t}(V)$ using (3.11) and sum-rate $R_{\text{sum}}^{t}(V)$ using (3.12), (c) average $P_{\text{sum}}^{t}(V)$ and $R_{\text{sum}}^{t}(V)$ over all $T$ fading and topology realizations to obtain $P_{\text{sum}}^{b}(V) = \frac{1}{T} \sum_{t=1}^{T} P_{\text{sum}}^{t}(V)$ and $R_{\text{sum}}^{b}(V) = \frac{1}{T} \sum_{t=1}^{T} R_{\text{sum}}^{t}(V)$.

Figure 3.7(a) shows the instantaneous power-rate region $R_{\text{INS}}(C^{t}, D^{t}, P_1^{\max}, \ldots, P_k^{\max})$ obtained by varying $\lambda$ between 0 and 1 for the topology given in Figure 3.5. Figure 3.7(b) shows the average power-rate region $R_{\text{AVG}}(P_1^{\max}, \ldots, P_k^{\max})$ achieved for $T = 500$. Figure 3.7(b) can also be interpreted as a trade-off curve between $P_{\text{sum}}(V)$ and $R_{\text{sum}}(V)$. The endpoint at the left in Figure 3.7(b) shows the smallest possible value of $R_{\text{sum}}(V)$, for the smallest possible value of $P_{\text{sum}}(V)$. The endpoint on the right shows the highest possible value of $R_{\text{sum}}(V)$, for the highest possible value of $P_{\text{sum}}(V)$. By finding the intersection of the curve with a horizontal line $R_{\text{sum}}(V) = \rho$, we can see how large $P_{\text{sum}}(V)$ should be to achieve $R_{\text{sum}}(V) \geq \rho$. The results show that there exists a strong trade-off between $P_{\text{sum}}(V)$ and $R_{\text{sum}}(V)$ near the endpoint on the left and the endpoint on the right on the boundary of $R_{\text{AVG}}(P_1^{\max}, \ldots, P_k^{\max})$.

Figure 3.8 shows the average power-rate regions obtained for different interference thresholds and different number of antennas at the base station. We set $M_k = 3$ for all $k = 1, \ldots, 11$, and run Algorithm 3.1 for $\theta_{ij}^k = 5\sigma^2$ and $\theta_{ij}^k = 0.001\sigma^2$ for all $k = 2, \ldots, K, j = 1$. Then we change $M_k = 2$ for all $k = 1, \ldots, K$, and run the algorithm again for $\theta_{ij}^k = 5\sigma^2$ and $\theta_{ij}^k = 0.001\sigma^2$ for all $k = 2, \ldots, K, j = 1$. In this set of simulations we used $T = 500$. For simplicity, we have listed the above combinations of simulations as:
- Case A - $M_k = 3$ for all $k = 1, \ldots, 11$ and $\theta_{1j}^k = 5\sigma^2$ for all $k = 2, \ldots, 11$ and $j = 1$
- Case B - $M_k = 3$ for all $k = 1, \ldots, 11$ and $\theta_{1j}^k = 0.001\sigma^2$ for all $k = 2, \ldots, 11$ and $j = 1$
- Case C - $M_k = 2$ for all $k = 1, \ldots, 11$ and $\theta_{1j}^k = 5\sigma^2$ for all $k = 2, \ldots, 11$ and $j = 1$
- Case D - $M_k = 2$ for all $k = 1, \ldots, 11$ and $\theta_{1j}^k = 0.001\sigma^2$ for all $k = 2, \ldots, 11$ and $j = 1$

Let us first focus on the curves drawn for Case A and Case B in Figure 3.8. The results show that the sum-rate maximization points (i.e. the end points on the right), the maximum transmit power that can be used in Case A is $4.33 \times 10^4$, and for Case B it is $4.21 \times 10^4$. The reason for this power difference is the relaxation of the interference threshold in Case A compared to Case B, and therefore, the femto base stations can transmit at high power. Hence, at the sum-rate maximization point, the gain in Case A compared to Case B is $1.7$ bit/s/Hz. Now let us focus on the curves drawn for Case C and Case D in Figure 3.8. The results show that at the sum-rate maximization points, the maximum transmit power that can be used in Case C is $4.12 \times 10^4$, and that of Case D is $3.64 \times 10^4$. Hence, at the sum-rate maximization point, the gain in Case C compared to Case D is $3.6$ bit/s/Hz. Thus, the results show that at the sum-rate maximization points, the difference in $R_{\text{sum}}(V)$ for Case A and Case B is lower than the difference in $R_{\text{sum}}(V)$ for Case C and Case D. Note that the sum-rate gain (i.e., the gain in $R_{\text{sum}}(V)$) obtained when the femto base stations are transmitting at high power (i.e., when $\theta_{1j}^k = 5\sigma^2$) is higher than the sum-rate gain obtained when the femto base stations are transmitting at low power (i.e., when $\theta_{1j}^k = 0.001\sigma^2$). Furthermore, when the number of transmit antennas are reduced at the base stations, we can observe a significant difference in the sum-rate of the network. More specifically, when $M_k = 2$ (i.e., Case C and D) the sum-rate of the network is lower than the sum-rate obtained when $M_k = 3$ (i.e., Case A and B).

We next compare the gain in $R_{\text{sum}}(V)$ for a unit increment of $P_{\text{sum}}(V)$ for all the cases in Figure 3.8. The results show that at the sum-rate maximization point, the gain in Case A compared to Case B is $0.014$ bit/s/Hz per unit of $P_{\text{sum}}(V)$. Similarly, the gain in Case C compared to Case D is $7.5 \times 10^{-4}$ bit/s/Hz per
unit of $P_{\text{sum}}(V)$. To increase the readability, a summary of the above discussion is presented in Table 3.1.

Table 3.1. Comparison of gains in $R_{\text{sum}}(V)$ for different interference thresholds and different numbers of antennas at the base station ([69] © 2013 IEEE).

<table>
<thead>
<tr>
<th>Comparison of</th>
<th>Case A with</th>
<th>Comparison of</th>
<th>Case C with</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case B</td>
<td>1.7</td>
<td>Case D</td>
<td>3.6</td>
</tr>
<tr>
<td>Gain in $R_{\text{sum}}(V)$</td>
<td>bit/s/Hz</td>
<td>Gain in $R_{\text{sum}}(V)$ per unit of $P_{\text{sum}}(V)$</td>
<td>0.014</td>
</tr>
</tbody>
</table>

![Fig. 3.8](image)

Fig. 3.8. Set of all achievable objective values for different interference thresholds and different antenna configurations (at the base station). Curves are drawn for $\theta_{ij} = 5\sigma^2$ and $\theta_{ij} = 0.001\sigma^2$ for all $k = 2, \ldots, 11$, $j = 1$ by changing $M_k = 2$ to $M_k = 3$ for all $k = 1, \ldots, 11$ ([69] © 2013 IEEE).

Figure 3.9 shows the empirical cumulative distribution function (CDF) plots of the rate obtained by the macro user and the average rate obtained by a femto user in the femto network, for $\lambda = 0$ and $\lambda = 0.001$. Each curve is
averaged over $T = 500$ fading and topology realizations. In the simulations, we set $\theta_{ij}^k = 5\sigma^2$ for all $k = 2, \ldots, 11$, $j = 1$. The results show that the rate obtained by the macro user and the rate of a femto user is reduced when the algorithm reduces the weights associated with $R_{\text{sum}}(V)$ and increases the weights associated with $P_{\text{sum}}(V)$.

Next, we investigate the impact of co-channel deployment in heterogeneous networks on the average power-rate region $R^{\text{AVG}}(P_{1}^{\max}, \ldots, P_{k}^{\max})$, by using the proposed algorithm. For this simulation, we assume that the macrocell shares the frequency band with two femtocells. Here, we compare four co-channel deployment cases described as follows. In Case 1, the macro base station (i.e., the MBS) serves a macro user (MU), and two femto base stations serve two femto users, one per each femto base station. In Case 2, the first femto base station (i.e., FBS$_1$) is switched-off, so that the femto user which was served by FBS$_1$ (i.e., FU$_1$) is now served by the MBS. Thus, we have one femto user (i.e., FU$_2$) served by the second femto base station (i.e., by FBS$_2$) and two macro users served by the MBS. In Case 3, the second femto base station (i.e., FBS$_2$) is switched-off, so that the associated femto user (i.e., FU$_2$) is now served by the MBS. At the same time, we assume that FBS$_1$ is now switched-on, and FU$_1$ is served by FBS$_1$. Hence, there are two macro users and one femto user in the system. Finally, in Case 4, both femto base stations (i.e., FAP$_1$ and FAP$_2$) are switched-off, so that both femto users are now served by the MBS. That is, in Case 4 we have only the MBS and three users in the network. Table 3.2 summarizes the above introduced co-channel deployment cases.

<table>
<thead>
<tr>
<th>Case</th>
<th>served by the MBS</th>
<th>served by FBS$_1$</th>
<th>served by FBS$_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>MU</td>
<td>FU$_1$</td>
<td>FU$_2$</td>
</tr>
<tr>
<td>Case 2</td>
<td>MU + FU$_1$</td>
<td>-</td>
<td>FU$_2$</td>
</tr>
<tr>
<td>Case 3</td>
<td>MU + FU$_2$</td>
<td>FU$_1$</td>
<td>-</td>
</tr>
<tr>
<td>Case 4</td>
<td>MU + FU$_1$ + FU$_2$</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3.2. Different co-channel deployment cases ([69] ©2013 IEEE).

Figure 3.10 shows the average power-rate region $R^{\text{AVG}}(P_{1}^{\max}, \ldots, P_{k}^{\max})$ for the cases listed in Table 3.2. In this simulation, we consider the SNR operating
Fig. 3.9. Empirical CDF plot of sum rate for different $\lambda$ when $\theta_{1j}^k = 5$ for all $k = 2, \ldots, 11$, $j = 1$: (a) macro user (b) femto user ([69] ©2013 IEEE).
point is given by (3.38), and we set $\theta_{ij}^k = 5\sigma^2$ for all $k = 2, \ldots, 11$, $j = 1$ and $T = 1000$.

The results show that for the same amount of $P_{\text{sum}}(V)$ (i.e., $P_{\text{sum}}(V) = P$ in the plot), the achievable $R_{\text{sum}}(V)$ in Case 1 (i.e., when the macrocell shares the spectrum with two femtocells) outperforms the achievable $R_{\text{sum}}(V)$ in all other cases. Furthermore, the results show that the set of all achievable objective values for Case 2 and Case 3 are identical. This is because there is no difference in Case 2 and Case 3 when we average the results over a sufficient number of fading and topology realizations. However, both Case 2 and Case 3 (i.e., when the macrocell shares the spectrum with a femtocell) outperform the achievable $R_{\text{sum}}(V)$ in Case 4. Hence, the results show that by sharing the frequency band with femtocells the sum-rate of the network can be increased.

![Fig. 3.10. Average power-rate regions for different co-channel deployment scenario when $\theta_{ij}^k = 5\sigma^2$ for all $k = 2, \ldots, 11$, $j = 1$ ([69] ©2013 IEEE).](image)

### 3.5 Summary and discussion

In this chapter, a single-macrocell heterogeneous multiple-input multiple-output network has been considered, where the macrocell shares the same frequency band with the femto network. To guarantee the performance of macro users, the
interference generated by the femto users is kept below a predefined threshold. For this setting, the problem of finding the set of all achievable power-rate tuples has been considered. In fact, this problem is NP-hard. A solution method has been provided based on the relationship between weighted sum-rate maximization and weighted-sum-mean-squared-error minimization problems. The proposed algorithm has been used to numerically evaluate the performance of the considered heterogeneous network. Specifically, the power-rate regions have been obtained when: the interference generated by femtocells towards the macro users is limited by different interference threshold levels and the users are served by both the macrocell and femtocells and only by the macrocell. Numerical results have shown that for a given amount of sum-power the sum-rate obtained by the network is higher when the interference threshold is larger. Furthermore, the results have shown that the sum-rate obtained by the network is increased when the macrocell shares the frequency with the femtocells, and the sum-rate further increases when the number of femtocells is increased. We believe that the proposed algorithm can be used to determine whether the performance of a heterogeneous network can be increased by: handing over macro users to femtocells, increasing the number of transmit antennas at the base stations or by relaxing the interference threshold constraint.
4 Conclusions and future work

In this chapter we first summarize the conclusions of this thesis, highlighting our contributions and the main results. Then, we present some future research directions.

4.1 Conclusions

The main focus of this thesis was to develop radio resource management algorithms for multi-antenna wireless communication networks by using known convex optimization techniques. Specifically, the admission control problem and the problem of finding the set of all achievable power-rate tuples have been considered.

In Chapter 1, the motivation, a precise literature survey that is relevant to the scope of this thesis, and the aims of this thesis have been presented. In Chapter 2, the admission control problem for the downlink of a multicell multiple-input single-output system has been considered. In other words, the objective was to find the maximum number of users that can be admitted to the network at their specified SINR targets. This problem has been cast as an $\ell_0$ minimization problem, where we minimized the number of non-admitted users (instead of maximizing the number of admitted users). The $\ell_0$ minimization problem is combinatorial, and NP-hard. Hence, we have to rely on suboptimal algorithms to solve it. To do this, the $\ell_0$ minimization problem has been approximated as a non-combinatorial one. The approximated non-combinatorial problem is a nonconvex one, and hence a convex approximation technique has been used to find an approximate solution for this problem. Then, suboptimal centralized and distributed algorithms have been proposed to find a solution to this problem. To develop the centralized algorithm, the sequential convex programming method has been used. The distributed algorithm has been derived by using the consensus-based alternating direction method of multipliers in conjunction with sequential convex programming. In contrast to the existing distributed algorithm (for admission control in multicell systems) that solves the subproblems in a cyclic order, the algorithm proposed in this thesis solves the subproblems independently in parallel at all BSs. Using numerical simulations, it
has been shown that the proposed centralized and distributed admission control algorithms achieve near-to-optimal performance.

Next, the admission control problem has been extended to provide fairness, where long-term fairness among the users has been guaranteed. Two commonly used fairness criteria, i.e., proportional fairness and max-min fairness have been used, and dynamic control algorithms via Lyapunov optimization have been proposed to solve them. Numerical results have shown that the proposed algorithms guarantee proportional and max-min fairness among the users.

Finally, in this chapter, the proposed centralized and distributed admission control algorithms have been adapted to control admissions in a heterogeneous network where there are users with different priorities. That is, in two-tier macro femto networks, macro users have a higher priority compared to the femto users. In such networks, all the macro users have to be admitted with the required SINR targets; the femto users are admitted by guaranteeing their targets, only if they can coexist with the macro users while satisfying the macro users' requirements. We have shown that by a simple modification to the problem we can accomplish this task. With numerical results, we have illustrated that the centralized algorithm achieves near-to-optimal performance, and the distributed algorithm performance is close to the optimal value.

In Chapter 3, we have looked at the problem of finding the set of all achievable power-rate tuples in a MIMO heterogeneous network. We have considered a single macrocell is overlaid with a set of femtocells, where the macrocell shares the same frequency band with the femtocells. The interference power affecting the macro users from the femto base stations has been kept below a certain threshold to guarantee that the performance of the macro users does not degrade due to the femto network. To find the set of all achievable power-rate tuples, a two-dimensional vector optimization problem has been formulated in which we have considered maximizing the sum-rate while minimizing the sum-power, subject to maximum power and interference threshold constraints. Then, the two-dimensional problem has been generalized to a multi-dimensional one, which can be used to address different requirements of a heterogeneous network. This problem is known to be NP-hard, because it involves the sum-rate maximization as one component of the objective function. A method to solve this problem has been provided by using the relationship between the weighted sum-rate maximization and weighted-sum-mean-squared-error minimization problems.
The proposed algorithm has been used to numerically find the power-rate regions of the heterogeneous network. Specifically, power-rate regions have been obtained when: 1) the interference generated by femtocells towards macro users is limited by different interference threshold levels, 2) the users are served by both macrocell and femtocells and only by the macrocell. Numerically it has been shown that the sum-rate obtained by the network is increased when the macrocell shares the frequency with the femtocells, and the sum-rate further increases when the number of femtocells has been increased. The proposed algorithm can be used to evaluate the performance of real heterogeneous networks via off-line numerical simulations. Furthermore, we believe this algorithm is useful to determine whether the performance of a heterogeneous network can be increased by handing over macro users to femtocells, increasing the number of transmit antennas at the base stations or by relaxing the interference threshold constraints.

### 4.2 Future directions

In Chapter 2, we have considered the problem of admission control for the downlink of a multicell MISO system and have proposed a suboptimal method to identify the number of users to be admitted to a network. Although the suboptimal methods are practically appealing, it is useful to investigate optimal methods for this problem. Such optimal methods can be used to provide performance benchmarks for other suboptimal algorithms, and are also essential in dimensioning wireless networks. In [84], a global optimal method to solve this problem is proposed, but it is for a SISO system. To the best of our knowledge, there are no global optimal methods proposed for solving the admission control problem for MISO systems.

It is envisioned that wireless networks will consist of densely deployed base stations with large number of user devices requesting heterogeneous services [7, 125–127]. Hence, for dense networks, the admission control algorithms proposed in Chapter 2 may converge slowly. In [87] a method, based on Perron-Frobenius theory [82, 128], has been introduced for SISO systems to accelerate the admission control process. Hence, it would be interesting to investigate such a mechanism for dense networks with multiple antennas, so that it can be used
in conjunction with the proposed admission control algorithms to accelerate the admission control process.

Another straightforward extension of the work considered in Chapter 2 is to provide fairness guaranteed distributed dynamic admission control algorithms. Specifically, for a MIMO system the methods provided in Chapter 2 may not be directly applicable to provide a distributed solution that guarantees fairness in the admission control for a dynamic network. Thus, we may rely on the relationship between the WMMSE and WSRmax method to provide a solution, as that technique can be used to find a distributed solution.

In addition, directly applying the proposed distributed algorithm may degrade the performance of a large network, because, even though the subproblems are solved in parallel, all the base stations have to wait until each base station has finished its local variable updates. This can create unnecessary delays in finding a solution in the case of dense networks. Hence, it will be interesting to investigate asynchronous algorithms where the base stations do not have to wait for the latest local variable updates from other base stations, and can even work with outdated information. An asynchronous ADMM \[129\] method could be one suitable method that we can rely on to modify the proposed distributed admission control algorithm to accomplish the above mentioned challenge.

In Chapter 3, we have not considered the total power consumption of a base station while computing the achievable power-rate tuples. Including static power consumption for that problem is a trivial extension, because it is not a decision variable, and does not affect the solution of the problem. However, incorporating an active antenna set selection for this study is not a straightforward extension, and hence such investigations would be interesting future extensions of the problem considered in Chapter 3. Furthermore, in this analysis we have assumed that the channel is perfectly known at the base stations. However, in practice it is quite challenging to know the channels without encountering errors. Hence, it would be interesting to compute the achievable power-rate regions in the presence of channel estimation errors.

Millimeter wave communications have been identified as a key technology in 5G wireless networks \[127, 130\]. Hence, investigating energy efficient resource allocation in millimeter communications is also important. In the recent study \[74\], an energy efficient resource allocation algorithm has been proposed assuming
perfect CSI. Thus, another interesting research direction could be looking at this problem with CSI errors.

Energy efficient resource allocation is also important in dense wireless networks. Hence, by using the relation between the WSRmax and WMMSE minimization problems (i.e., the method used in Chapter 3) it would be interesting to investigate centralized and distributed optimization algorithms. If antenna selection is also incorporated in this problem, then using the above approach may not be directly applicable, and hence, it may require investigating novel approaches to solve this problem.
References


131


Li, Xiaobai (2017) Reading subtle information from human faces
Luoto, Markus (2017) Managing control information in autonomic wireless networking
Mustonen, Miia (2017) Analysis of recent spectrum sharing concepts in policy making
Visuri, Ville-Valteri (2017) Mathematical modelling of chemical kinetics and rate phenomena in the AOD Process
Lappi, Tuomas (2017) Formation and governance of a healthy business ecosystem
Samarakoon, Sumudu (2017) Learning-based methods for resource allocation and interference management in energy-efficient small cell networks
Pargar, Farzad (2017) Resource optimization techniques in scheduling: Applications to production and maintenance systems
Uusitalo, Pauliina (2017) The bound states in the quantum waveguides of shape Y, Z, and C
Palosaari, Jarkko (2017) Energy harvesting from walking using piezoelectric cymbal and diaphragm type structures
Mononen, Petri (2017) Socio-economic impacts of a public agency—enhancing decision support for performance management
Karikkainen, Marja-Liisa (2017) Deactivation of oxidation catalysts by sulphur and phosphorus in diesel and gas driven vehicles
Viitlala, Harri (2017) Selected methods for WBAN communications — FM-UWB and SmartBAN PHY
Akram, Saad Ullah (2017) Cell segmentation and tracking via proposal generation and selection
Ylimäki, Markus (2017) Methods for image-based 3-D modeling using color and depth cameras
Bagheri, Hamidreza (2017) Mobile clouds: a flexible resource sharing platform towards energy, spectrum and cost efficient 5G networks
Heikkinen, Kari-Pekka (2018) Exploring studio-based higher education for T-shaped knowledge workers, case LAB studio model
Shashika Manosha Kapuruhamy Badalge

CONVEX OPTIMIZATION BASED RESOURCE ALLOCATION IN MULTI-ANTENNA SYSTEMS