COOPERATIVE TRANSMISSION FOR THE DOWNLINK OF MULTIUSER MIMO CELLULAR NETWORKS

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL AND

ELECTRONICS ENGINEERING

AND THE INSTITUTE OF ENGINEERING AND SCIENCES

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

By

Yakup Kadri Yazarel

August, 2007

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Asst. Prof. Dr. Defne Aktaş (Supervisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof. Dr. Erdal Arıkan

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Asst. Prof. Dr. Çağatay Candan

Approved for the Institute of Engineering and Sciences:

Prof. Dr. Mehmet Baray
Director of Institute of Engineering and Sciences

ABSTRACT

COOPERATIVE TRANSMISSION FOR THE DOWNLINK OF MULTIUSER MIMO CELLULAR NETWORKS

Yakup Kadri Yazarel

M.S. in Electrical and Electronics Engineering

Supervisor: Asst. Prof. Dr. Defne Aktaş

August, 2007

In this thesis, we propose a distributed transmission scheme for the downlink of a multiuser system. The base-stations (BSs) cooperate with each other with limited, local message-passing to find the optimum beamforming vectors, where there are individual signal-to-interference-plus-noise-ratio (SINR) targets for each user. Majority of the previous work on this problem assumed a total power constraint on the BSs. However, since each transmit antenna is limited by the amount of power it can transmit due to the limited linear region of the power amplifiers, a more realistic constraint is to place a limit on the per-antenna power.

In a recent work, Yu and Lan proposed an iterative algorithm for computing the optimum beamforming vectors minimizing the power margin over all antennas under individual SINR and per-antenna power constraints. However, from a system designer point of view, it may be more desirable to minimize the total transmit power rather than minimizing the power margin, especially when the system is not symmetric. Reformulating the transmitter optimization problem to minimize the total transmit power subject to individual SINR constraints on the users and per-antenna power constraints on the base stations, the algorithm proposed by Yu and Lan is modified. Performance of the modified algorithm is compared with the existing methods for various cellular array scenarios.

The modified algorithm requires inversion of a matrix, which cannot be implemented fully distributively using limited information exchange between BSs. By approximating the matrix as tridiagonal, a suboptimal distributed algorithm for computing the beamforming vectors in a cooperative system is obtained. The proposed distributed algorithm is shown to achieve near optimal performance when the target SINRs and the size of the array are small.

Keywords: Downlink beamforming, distributed transmission, base-station cooperation, broadcast channel, per-antenna power constraints.

ÖZET

ÇOKLU KULLANICILI ÇOKLU ANTENLİ HÜCRESEL AĞLARDA BAZ İSTASYONU-YER BAĞI İÇİN İŞBİRLİKLİ İLETİM

Yakup Kadri Yazarel
Elektrik ve Elektronik Mühendisliği Bölümü Yüksek Lisans
Tez Yöneticisi: Yard. Doç. Dr. Defne Aktaş
Ağustos, 2007

Bu tezde çok kullanıcılı bir sistem için dağıtılmış bir baz istasyonu-yer bağı iletim algoritması önerilmektedir. Bu algoritmada baz istasyonlarının kendi aralarında sınırlı sayıda ve yerel mesaj alışverişi yaparak, her kullanıcı için varolan sinyal-girişim oranı hedeflerini sağlayacak şekilde gönderecekleri dalgaları ayarlaması amaçlanmaktadır. Bu konuda daha önce yapılan çalışmalar genel olarak baz-istasyonu antenlerinde toplam güç sınırı varolduğunu kabul etmişlerdir. Ancak, her antenin iletim gücü bağlı olduğu yükseltici devrelerin doğrusal bölgesi ile kısıtlanmıştır. Dolayısıyla antenler üzerinde güç kısıtlaması yerine anten başına bir güç kısıtlaması düşünülmesi daha gerçekçi bir varsayımdır.

Yakın zamandaki bir çalışmada, Yu ve Lan gönderdikleri dalgaları ayarlayarak anten başına bir güç sınırını aşmadan, her kullanıcı için varolan sinyal-girişim oranı hedeflerini sağlamaya çalışan ve anten başına düşen güç payını en aza indirgeyen bir algoritma önermişlerdir. Ancak, bir sistem tasarımcısı gözüyle, anten başına düşen güç payından ziyade toplam iletim gücünü en aza indirgemek, özellikle sistem asimetrik olduğunda daha fazla istenen bir durumdur. Bu yüzden, Yu ve Lan'ın önerdiği algoritma değiştirilerek, toplam gücü en aza indirgeyen ve baz istasyonu başına bir güç sınırını aşmadan, her kullanıcı için varolan sinyal-girişim oranı hedeflerini sağlamaya çalışan bir algoritma önerilmiştir. Önerilen algoritmanın performansı varolan metodlarla değişik hücresel sistem senaryoları için karşılaştırılmıştır.

Değiştirilen algoritma bir matrisin tersinin alınmasını gerektirmektedir. Ancak, bu baz istasyonları arasında sınırlı bilgi alışverişi kullanarak tamamen dağıtılmış bir şekilde yapılamamaktadır. Bu matris yaklaşık olarak 3 köşegenel olarak alınıp işbirlikli bir sistemde dalgalar ayarlanarak en iyiye yakın bir algoritma elde edilmiştir. Önerilen işbirlikli algoritmanın, sinyal-girişim oranı hedefleri ve hücre sayısı az olduğunda en iyiye yakın olduğu gösterilmiştir.

Anahtar Kelimeler: Baz istasyonu-yer bağı, hüzme oluşturma, dağıtılmış iletim teknikleri, baz istasyonu işbirliği, yayın kanalı, anten başına güç sınırlamaları.

ACKNOWLEDGMENTS

I would like to thank my supervisor Dr. Defne Aktaş for her invaluable help, encouragement and motivation during my MS studies. I would also like to thank Dr. Erdal Arıkan and Dr. Çağatay Candan for accepting to be in the thesis committee and commenting on the thesis. Finally, I appreciate my office colleagues.

Contents

1 INTRODUCTION **1**

1.1 Overview 1

1.2 Contributions 5

1.3 Outline of the Thesis 6

1.4 Notation 6

2 BACKGROUND **7**

2.1 System Model 7

2.2 MIMO BC Channel Capacity 9

2.3 Uplink-Downlink Duality 12

3 BEAMFORMING ALGORITHMS IN THE LITERATURE **14**

3.1 A Beamforming Technique Under Sum-Power Constraint on the Antennas 14

3.2 Beamforming Techniques Under Per-Antenna Power Constraints . 20

3.2.1 Equal-Rate Zero-Forcing Transmission 21

3.2.2 Power Margin Minimization in the Downlink 24

4 PROPOSED BEAMFORMING ALGORITHMS 30

4.1 Centralized Algorithm 31

4.2 Suboptimal Distributed Algorithm 35

5 NUMERICAL RESULTS 39

5.1 Centralized Algorithm 39

5.2 Distributed Algorithm 47

6 CONCLUSIONS 50

APPENDIX 52

A Test of SINR Feasibility 52

B Lagrangian Dual Formulation 53

C Subgradient Projection 56

List of Figures

1.1 General MIMO system 2

2.1 Multiple antenna BS and single-antenna mobile users 8

2.2 Wyner's circular cellular array model 10

2.3 Wyner's linear cellular array model 11

2.4 MIMO BC and dual MAC 13

3.1 The convexity of the area between logarithm function and a constant 24

4.1 Message-passing between neighbour BSs 37

5.1 Comparison of the proposed method with the method in [1] (denoted as Boche et.al.) in terms of total transmit power for Wyner's symmetric circular array with $N = 10$ and various α values 40

5.2 Proposed algorithm vs. Yu-Lan algorithm in [2] for symmetric circular array and linear array for $N = 10$ and $\alpha = 0.3$ 41

5.3 Proposed algorithm vs. Yu-Lan algorithm in [2] for asymmetric Wyner's circular array with $N = 10$ 43

5.4 Comparison of proposed and Yu-Lan algorithm in terms of maximum, average and minimum BS antenna powers for an asymmetric Wyner's circular array with $N = 10$ and random interference parameters 43

5.5 Time for convergence of proposed algorithm vs. Yu-Lan algorithm for symmetric circular and linear array with $N = 10$ and $\alpha = 0.3$. 44

5.6 Comparison of proposed algorithm and ZF algorithm for a symmetric Wyner's circular array with $N = 5$ and various α values . . 45

5.7 Comparison of beamforming only scheme and beamforming with DPC scheme for a symmetric Wyner's circular array with $N = 9$ and $\alpha = 0.25$ 46

5.8 The absolute value of the duality gap for distributed algorithm for Wyner's linear array with $N = 3$ and $\alpha = 0.25$ and $\alpha = 0.5$ 49

List of Tables

5.1 Number of iterations for convergence for various array scenarios for SINR target 5 for all users 42

List of Acronyms

BC : Broadcast Channel

BS : Base-Station

DPC : Dirty Paper Coding

LMMSE : Linear Minimum-Mean-Squared-Error

MAC : Multiple-Access Channel

MIMO : Multiple Input Multiple Output

LMMSE : Linear Minimum-Mean-Squared-Error

MMSE : Minimum-Mean-Squared-Error

QoS : Quality of Service

SIC : Successive Interference Cancelation

SINR : Signal-to-Interference-plus-Noise Ratio

SNR : Signal-to-Noise Ratio

SVD : Singular Value Decomposition

ZF : Zero Forcing

Dedicated to my parents

Chapter 1

INTRODUCTION

In this thesis, we consider downlink beamforming for a multiuser multiple input multiple output (MIMO) cellular network with base-station (BS) cooperation. We propose a centralized and a distributed algorithm that computes the optimal beamforming vectors under individual SINR and per-antenna power constraints. In this chapter, we will give an overview of existing literature on multiuser MIMO cellular networks and transmission schemes for the downlink. We summarize the contributions of the thesis and introduce the notation used in the sequel.

1.1 Overview

Today's communication systems have a need of very high data rates. On the other hand, some systems have a limit in terms of power and bandwidth. To satisfy these needs, multiple antennas both at the transmitter and receiver can be used. These systems are referred as MIMO systems. Recent advances show that MIMO systems promise high spectral efficiency and data rates over wireless links without increasing transmit power and requiring extra bandwidth. Providing resistivity to fading and increased coverage, MIMO systems require complex algorithms and design methods [3],[4]. A general MIMO system is shown in Fig. 1.1.

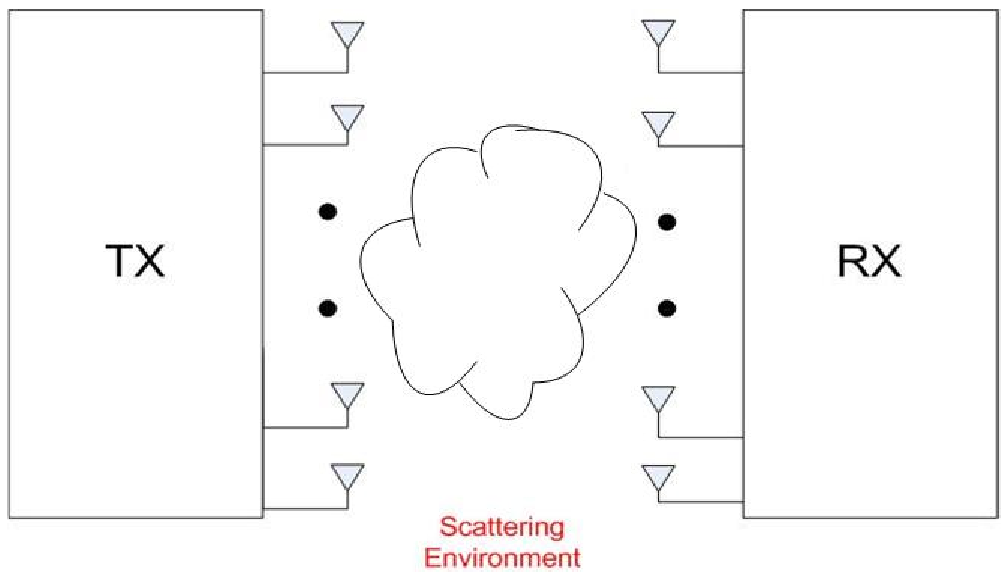

Figure 1.1: General MIMO system

Single user MIMO systems require less complexity than multiuser MIMO systems. In multiuser MIMO, because of the interference caused by other users' signals, performance of the users can be limited. Suppression of this interference often requires complex algorithms. To mitigate interference, a technique called beamforming which adjusts the beam-patterns of antenna arrays to minimize the effect of interference on the terminals, is used [5].

Recently many researchers have especially studied spectrally efficient multi-antenna BS processing (transmit beamforming) since downlink is typically the bottleneck in cellular systems. In the downlink, since receivers are mobile users with limited number of antennas, processing power and energy constraints; the task of mitigating the interference is typically shifted to the transmitter (BS) side. It is more effective to minimize the effect of interference at the transmitter side [5].

In a multiuser MIMO cellular system, the downlink is modeled as MIMO Gaussian broadcast channel (BC) whereas the uplink is modeled as Gaussian multiple-access channel (MAC). Recent work showed that Gaussian BC and Gaussian MAC are duals of each other. That means, the signal-to-interference plus noise (SINR) region of a downlink channel is equal to the SINR region of a dual uplink channel under a sum-power constraint [6], [7]. Exploiting this duality, Costa's "Dirty-Paper Coding" (DPC) strategy [8] together with downlink beamforming is found to be optimal in achieving the sum capacity for MIMO downlink channel [6], [9]. However, DPC is an information theoretic coding scheme, which is not practical to be implemented in a real system.

The reason why duality is used in the downlink transmission problem is the following: downlink beamforming is more complicated and analytically difficult problem to solve since beamformers need to be optimized jointly. That is, one user's beamformer may increase the interference of another user and degrade the quality of service for that user. Because of the crosstalk of the users which may affect each other's SINR values, downlink beamforming becomes a complex and difficult-to-solve problem [10]. Therefore, dual uplink model which is easy-to-compute, is used while computing the downlink beamformers.

The problem of computing the optimal beamforming vectors and adjusting transmit powers for antennas have been analyzed for various schemes in [1], [5], [10], [11]. Widely used performance metric is the rates (SINR values) of the users. Commonly used system resource is the transmit powers of the antennas. From a network designer point of view, performance (rates) must be hold above a certain threshold while minimum of system resources (transmit powers) are used. Another approach is to maximize the achievable SINR region under maximum power constraints [10].

Previous works mainly consider a single cell scenario with multiple-antenna BS with a total power constraint and single antenna mobile users. The transmission schemes are previously derived for achieving sum capacity given a sum power constraint or for achieving minimum transmit power given SINR constraints (corresponding to different quality of service (QoS) requirements) on the users. Based on these results, our primary goal is to develop distributed transmission schemes for a cellular network with cooperative single antenna BSs and multiple decentralized single antenna users. We focus on Wyner's cellular network model and study the performance of the proposed method for this simplified network model to gain further insight.

We aim to jointly optimize the transmit user power allocation and beamforming vectors minimizing the total transmit power subject to individual SINR constraints on the mobile terminals and transmit power constraints on the BS. Transmitted power constraints may either be on the total power or per-antenna power. While total power constraint is analytically easier to solve [10], per-antenna power constraint is more practical and realistic since all antennas have their own front-end amplifiers which are limited by their linear regions [2]. For macrodiversity systems where the BSs cooperate in transmission of the information to the users, per-antenna power constraints are a reasonable assumption since antennas in the BS cooperation case are geographically distributed.

In [12], zero-forcing (ZF) beamforming is implemented for downlink with per-antenna power constraints. ZF is used for interference suppression but it does not concern about optimizing the SINR [13]. It is suboptimal since it uses more power to null out the interference. In [2], Yu and Lan proposed a numerical algorithm for the downlink which computes the beamformers with minimum power margin under per-antenna power constraints. But this approach fails to give optimal results when the system is not symmetric as illustrated in Chapter 5. In order to minimize power margin, it tries to satisfy a power balance between the antennas and uses high transmit power in certain cases. This notion tells us that rather than minimizing the power margin, it is more reasonable to minimize the total transmit power.

In this thesis our aim is to build a distributed, iterative algorithm that computes beamforming vectors satisfying SINR and power constraints. We reformulate the optimization problem in [2] and investigate distributed implementation of the proposed transmission scheme which can be implemented by limited local information exchange between cooperative BSs. The BSs are assumed to be connected to each other by a high capacity backbone and cooperate in transmission of information to the users. As we will illustrate in Chapter 5, the reformulation of the problem yields better results in terms of performance.

1.2 Contributions

An optimal algorithm that computes beamforming vectors satisfying SINR and per-antenna power constraints is proposed. The proposed algorithm performs better in terms of performance and convergence time over Yu-Lan algorithm [2], which computes the minimum power margin satisfying the same constraints, and the ZF beamforming algorithm in [12] as illustrated in Chapter 5.

The distributed implementation of the proposed algorithm is investigated. An algorithm based on limited local information exchange between BSs is presented. Due to an approximation done in one of the steps of the centralized algorithm to limit the amount of information exchange between the BSs, the performance of the distributed algorithm is suboptimal.

1.3 Outline of the Thesis

The rest of the thesis is organized as follows. In Chapter 2, the system model and the MIMO channel capacity in the literature are given. In Chapter 3, we review previous algorithms that computes optimal beamformers under SINR constraints. The proposed algorithm and its distributed implementation are presented in Chapter 4. The proposed method is compared with existing methods in Chapter 5 through numerical results and conclusions are given in Chapter 6.

1.4 Notation

In the sequel, we use small boldface letters to denote vectors and capital boldface letters to denote matrices. For a given matrix $\mathbf{A}$; $\mathbf{A}^{-1}$,$\mathbf{A}^{T}$, $\mathbf{A}^{H}$,$\mathrm{Tr}(\mathbf{A})$ and $\mathbf{A}_{i,k}$ denote the inverse, the transpose, the conjugate transpose, the trace and the (i,k)th element of $\mathbf{A}$ respectively. $\mathbf{A}^{(n)}$ denotes the value of $\mathbf{A}$ at nth iteration of an iterative algorithm. $\mathbf{I}$ denotes the identity matrix with appropriate dimensions and $\mathrm{diag}(\mathbf{A})$ denotes the vector of diagonal elements of any square matrix $\mathbf{A}$. $[\mathbf{A}]^{+}$ operation takes the maximum with respect to the elements of all-zero matrix with the same size of $\mathbf{A}$. $\mathcal{E}[.]$ denotes the expectation operation. $\mathbb{R}$ and $\mathbb{C}$ denote the set of real and complex numbers, respectively. $\|.\|_2$ denotes the l_2 norm.

Chapter 2

BACKGROUND

In this chapter, a brief introduction to the downlink of multiuser MIMO cellular networks will be given. First, the system model under consideration will be presented and then the capacity region of MIMO BC, modeling the downlink, and uplink-downlink duality used in optimization problems involving MIMO BC are summarized.

2.1 System Model

We consider a cellular network with BSs and single-antenna mobile users. The base-stations are assumed to be connected to each other via a high-capacity backbone and cooperate with each other. This scenario is identical to the case where a single BS with geographically distributed antennas communicate with single antenna mobile users as depicted in Fig. 2.1.

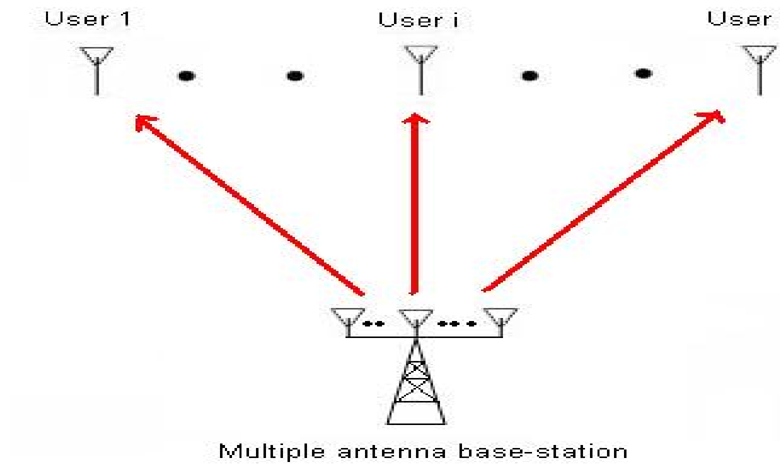

Figure 2.1: Multiple antenna BS and single-antenna mobile users

We assume that the BSs and mobiles have perfect channel knowledge and the channel is flat-fading. The scenario under investigation consists of N base-stations and K remote decentralized users all with a single antenna. The downlink channel is modeled as:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{2.1}$$

where $\mathbf{x} = [x_1 \ldots x_N]^T$ is an N x 1 vector representing the transmit signal, $\mathbf{H}$ is a K x N channel matrix and $\mathbf{n} = [n_1 \ldots n_N]^T$ is an N x 1 vector whose components are additive white Gaussian noise with variance σ^2. The rows of channel matrix $\mathbf{H}$ are denoted as $\mathbf{h}_i^H \in \mathbb{C}^{1 \text{ x } N}$, $i = 1, \ldots, K$ which represents the complex path gains from BS antennas to user i's antenna.

The form of the transmit signal is as follows :

$$\mathbf{x} = \sum_{i=1}^{K} d_i \mathbf{w}_i \tag{2.2}$$

where d_i is a scalar denoting the information to be transmitted to the ith user which is of unit energy, i.e. $\mathcal{E}\left[|d_i|^2\right] = 1$ and $\mathbf{w}_i$ is a N x 1 beamforming vector for user i. With this formulation, the power allocated to ith user is given as $p_i = \mathbf{w}_i^H \mathbf{w}_i$. The power transmitted by kth antenna is given as $\tilde{p}_k = \left(\sum_{i=1}^{K} \mathbf{w}_i \mathbf{w}_i^H\right)_{k,k}$.

To gain insight on gains from BS cooperation, in the distributed implementation and the numerical studies, we focus on a simplified cellular array model described by Wyner [14]. In Wyner's cellular model, each BS has one active user due to an orthogonal intra cell access scheme and each user is exposed to interference only from the two neighbouring cells. This interference is exploited to improve performance using BS cooperation. Mathematically, the cellular scenario can be formulated with the following channel matrix $\mathbf{H}$ (for $N = K$) with interference factors $0 < \alpha_i^+ < 1$ and $0 < \alpha_i^- < 1$:

$$\mathbf{H} = \begin{bmatrix} 1 & \alpha_1^+ & 0 & \cdots & 0 & \alpha_1^- \\ \alpha_2^- & 1 & \alpha_2^+ & & & 0 \\ 0 & & \ddots & & & \vdots \\ \vdots & & & \ddots & & 0 \\ 0 & & & \alpha_{N-1}^- & 1 & \alpha_{N-1}^+ \\ \alpha_N^+ & \cdots & & 0 & \alpha_N^- & 1 \end{bmatrix}$$

The above channel model represents the so called Wyner's circular cellular array model in which the cells are located on a circle with BSs at the center. The model is depicted in Fig. 2.2.

If all the interference factors are same, i.e. $\alpha_i^+ = \alpha_i^- = \alpha$, $\forall i$, the above cellular array is symmetric for all the base-stations and mobiles. If we set α_1^- and α_N^+ to 0, we obtain Wyner's linear cellular array model as shown in Fig. 2.3.

2.2 MIMO BC Channel Capacity

The capacity region for the general degraded BC has been known but the capacity region for the general non-degraded BC has not been derived yet. MIMO BC is in general a non-degraded channel whose capacity region has remained an open problem until recently.

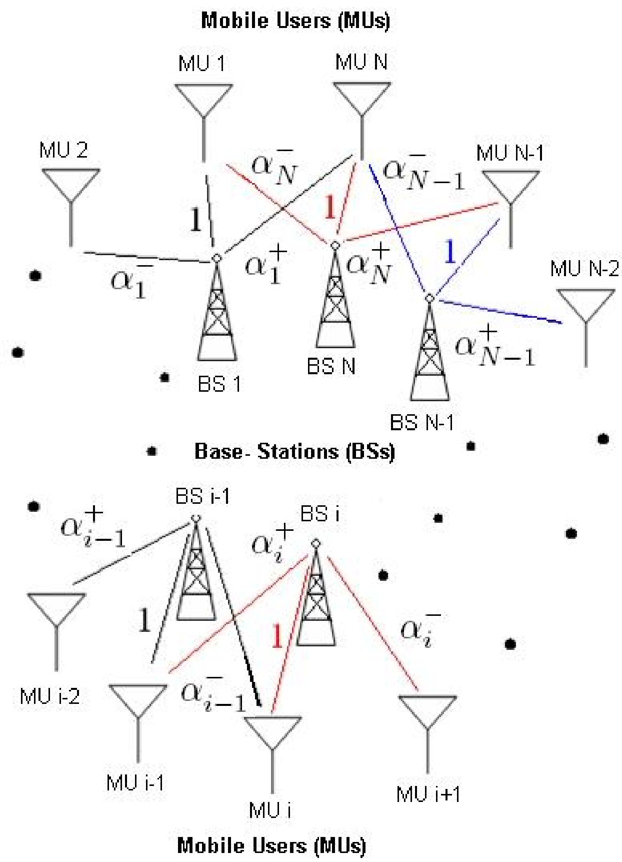

Figure 2.2: Wyner's circular cellular array model

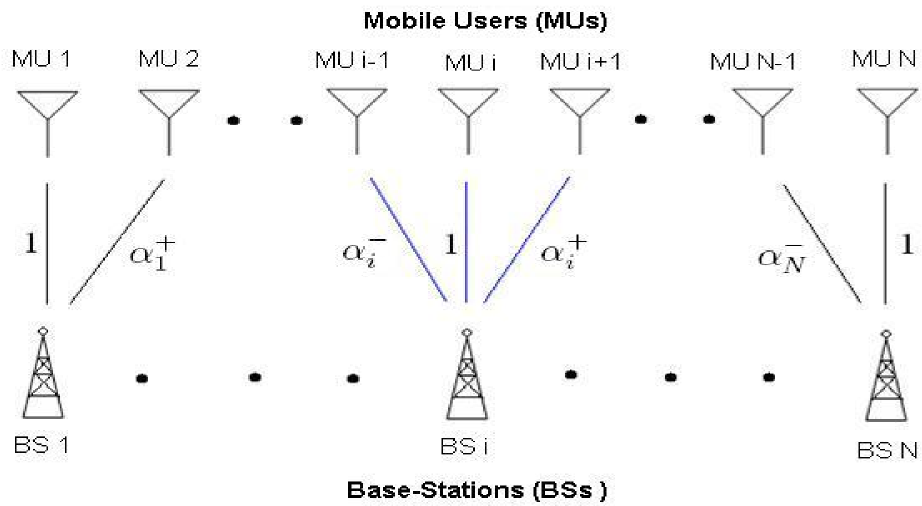

Figure 2.3: Wyner's linear cellular array model

The difficulty in computing the capacity region is as follows: in point-to-point MIMO channels, one can use the advantage of cooperation of receiving antennas, but in MIMO BC the receiving antennas do not cooperate. Point-to-point MIMO channels can be parallelized by using beamforming at the transmitter and the receiver and water-filling can be done over these parallel channels, but similar technique is no longer practical for MIMO BC, due to the lack of cooperation at the receivers.

First result for MIMO BC is given by Caire and Shamai in [9] for two single-antenna users case. They propose an optimal scheme using Costa's DPC strategy [8]. Costa showed that if DPC is used, the capacity of a single user channel where interference is known by the transmitter is equal to the capacity where interference does not exist. In MIMO BC, the transmitter can calculate the amount of interference created by the transmitted signals for other users. So, the users can be ordered and encoded knowing the interference caused by previously encoded signals. However, the method in [9] is difficult to extend to more than 2 users case. Asymptotic results for BC capacity are presented and zero-forcing

(ZF)- DPC method is shown to be suboptimal. However, for high SNR regime (as transmission power goes to infinity) for the channels with full row rank, ZF-DPC method is shown to achieve the capacity.

Yu and Cioffi generalize this result and find the optimal capacity as the saddle point of the mutual information maximized over signal covariance matrix and minimized over noise covariance matrix in [15]. But, this result is only valid when the noise covariance matrix is non-singular.

The general result for more than 2 users case case is found by a different approach in [6] and [7]. The BC capacity region for more than two users is computationally complex. Because of this, the duality of MAC and BC is exploited and the sum capacity of the downlink is proven to be equal to the capacity of the dual MAC as explained in Section 2.3. The sum capacity (C_{sum}) under a total power constraint P_T on the users is found as:

$$C_{\text{sum}} = \sup_{\mathbf{D}} \log \det \left(\mathbf{I} + \mathbf{H}\mathbf{D}\mathbf{H}^H \right) \tag{2.3}$$

where $\mathbf{D}$ is a KxK diagonal matrix with uplink user powers on the diagonals with $\text{Tr}[\mathbf{D}] \leq P_T$. They prove the entire achievable region with DPC for downlink is exactly identical to the MAC capacity region. In [16], the capacity region is characterized for the MIMO BC under a wide range of input covariance constraints, and for both of the total power and the per-antenna power constraints. The capacity region is achieved by the transmission scheme which is a combination of beamforming with DPC.

2.3 Uplink-Downlink Duality

BC optimization problems are not convex in general, whereas MAC problems are often convex problems. The nonconvexity makes BC problems computationally complex. However, there is a connection between MAC and BC problems known as MAC-BC (uplink-downlink) duality which helps BC problems to be solved easily by its dual in the MAC. The MIMO BC and dual MIMO MAC is shown in Fig. 2.4.

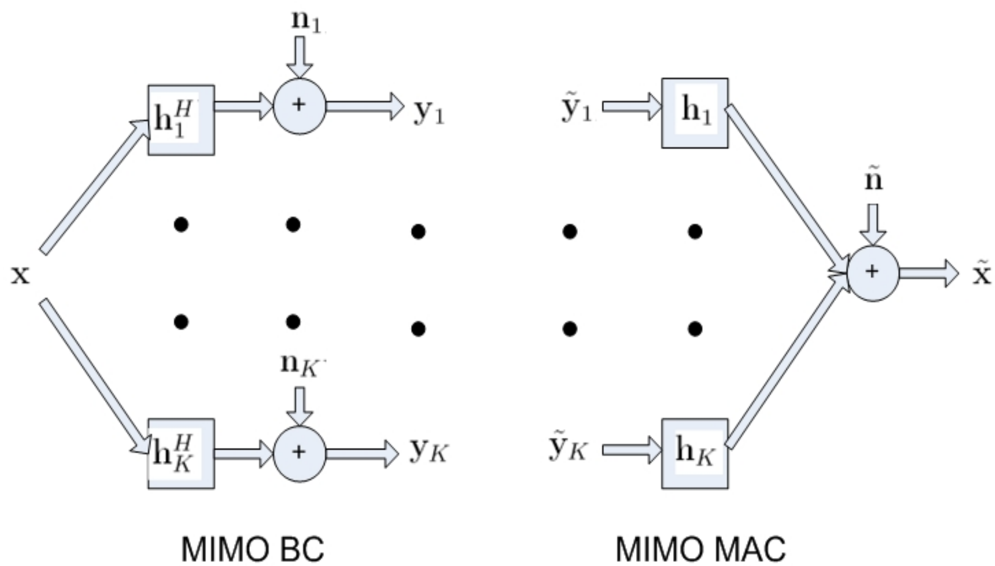

Figure 2.4: MIMO BC and dual MAC

Duality states that any achievable rate vector with user power constraints $P_1, \ldots, P_K$ in the MAC is also achievable for BC with total power constraint $P_T = \sum_{i=1}^{K} P_i$ [7]:

$$C_{\mathrm{BC}}(P_T, \mathbf{H}) = C_{\mathrm{MAC}}(P_1, \ldots, P_K, \mathbf{H}^H) \tag{2.4}$$

where C_{BC} and C_{MAC} denote BC and MAC capacity, respectively.

As shown in the following section, the SINR expressions for the BC problems are coupled by beamformers, whereas in the dual MAC they are not coupled. In the dual uplink, the beamformer vectors are found as the SINR maximizing minimum-mean-square-error (MMSE) filters [6]. The uplink beamformers are identical with downlink beamformers upto a scaling factor [2]. So, the optimization problem is solved for the dual MAC problem with low complexity and this solution is converted to the solution for BC problem easily.

Chapter 3

BEAMFORMING ALGORITHMS IN THE LITERATURE

In this chapter, some of the downlink beamforming algorithms proposed in the literature which are related to the algorithm proposed in this work will be summarized. First, an algorithm for computing the optimal beamforming vectors under sum power constraint on the antennas will be presented. Then two different approaches for finding the optimal beamforming vectors under per-antenna power constraints will be summarized.

3.1 A Beamforming Technique Under Sum-Power Constraint on the Antennas

Several algorithms have been proposed for computing the optimum power allocation over users and optimum beamformers under sum-power constraints on the antennas. There are various algorithms in the literature but we now summarize the algorithm in [1], since it will be used as a benchmark for comparison with the proposed method.

An iterative algorithm computing the beamforming vectors and user power allocation, while simultaneously satisfying individual SINR constraints on the users with minimum total transmit power is proposed in [1]. The optimization is performed firstly in the dual uplink and then this result is used for finding downlink beamformers and power allocations. The algorithm does not require any computationally complex operations such as matrix inversion or eigenvalue decomposition.

Uplink and Downlink Problem Formulations

The beamformers of K users are adjusted so that target SINRs $\gamma_1, \ldots, \gamma_K$ are achieved with minimum total power. In the downlink, the users are coded with the index order $\overline{\pi} = \overline{\pi}_1, \ldots, \overline{\pi}_K$ where user with index $\overline{\pi}_1$ is encoded first and the user with index $\overline{\pi}_K$ is encoded last. The interference caused by the users indexed by $\overline{\pi}_1, \ldots, \overline{\pi}_{i-1}$ to the user i is known before the transmission. One can use DPC to cancel the interference caused by the previously encoded users. The user indexed by $\overline{\pi}_1$ is effected by the interference from all users, the user indexed by $\overline{\pi}_2$ is effected by the interference from users with index $\overline{\pi}_3, \ldots, \overline{\pi}_K$, and so on. The user indexed by $\overline{\pi}_K$ sees no interference. The SINR expression for downlink becomes:

$$\mathrm{SINR}^{\mathrm{DL}}_{\overline{\pi}_i}(\mathbf{w}_{\overline{\pi}}, \mathbf{p}, \overline{\pi}) = \frac{p_{\overline{\pi}_i} \left| \mathbf{w}^H_{\overline{\pi}_i} \mathbf{h}_{\overline{\pi}_i} \right|^2}{\sum_{k=i+1}^{K} p_{\overline{\pi}_k} \left| \mathbf{w}^H_{\overline{\pi}_k} \mathbf{h}_{\overline{\pi}_i} \right| + \sigma^2}, \ \forall i. \tag{3.1}$$

where $\mathbf{p}$ is the power vector whose entries are the allocated transmit powers for K users, $[p_1, \ldots, p_K]^T$. The total power constraint is P_T. Then, the downlink problem is stated as:

$$P^{\mathrm{DL}}(\pi) = \min_{\mathbf{w}_1, \ldots, \mathbf{w}_K, \mathbf{p} > 0} \sum_{i=1}^{K} p_i \tag{3.2}$$

$$\text{subject to } \mathrm{SINR}^{\mathrm{DL}}_i(\mathbf{w}_i, \mathbf{p}, \overline{\pi}) \geq \gamma_i, \quad 1 \leq i \leq K \tag{3.3}$$

$$\|\mathbf{w}_i\|_2 = 1, \quad 1 \leq i \leq K \tag{3.4}$$

$$\sum_{i=1}^{K} p_i < P_T. \tag{3.5}$$

The dual uplink problem can be formulated similar to the downlink problem. In the dual uplink, user powers are represented as $\lambda_1, \ldots, \lambda_K$. The dual of DPC, in the uplink is successive interference cancellation (SIC). In dual uplink, SIC with decoding order π which is the reverse of downlink encoding order is applied. The user with index π_1 is decoded first and the user with index π_K is decoded last. Note that $\overline{\pi}_1 = \pi_K$ and $\overline{\pi}_K = \pi_1$. SIC cannot be used in downlink beamforming at the receiver side since the receivers do not cooperate with each other due to mobility and complexity constraints. By SIC, the interference caused by the previously decoded users is subtracted and interference-plus-noise covariance matrix, $\mathbf{Z}_{\pi_i}$, of user with index π_i becomes

$$\mathbf{Z}_{\pi_i}(\boldsymbol{\lambda}, \pi) = \sigma^2\mathbf{I} + \sum_{k \in \{\pi_{i+1}, \ldots, \pi_K\}} \lambda_k \mathbf{h}_k \mathbf{h}_k^H, \quad 1 \leq i \leq K. \tag{3.6}$$

The interference-plus-noise covariance matrix indicates the interference and noise correlation between the N antennas at the BS. Using this information, the beamformers can be formed in order to reduce the effect of interference and noise. The uplink beamformers denoted as $\hat{\mathbf{w}}_i$ for the ith user are assumed to be unity-norm, i.e. $\|\hat{\mathbf{w}}_i\|_2 = 1$.

The SINR for user i in the dual uplink is defined as follows:

$$\mathrm{SINR}_i^{\mathrm{UL}}(\hat{\mathbf{w}}_i, \boldsymbol{\lambda}, \pi) = \frac{\lambda_i \left|\hat{\mathbf{w}}_i^H \mathbf{h}_i\right|^2}{\hat{\mathbf{w}}_i^H \mathbf{Z}_i(\boldsymbol{\lambda}, \pi)\hat{\mathbf{w}}_i}. \tag{3.7}$$

The uplink optimization problem is described as:

$$P^{\mathrm{UL}}(\pi) = \min_{\hat{\mathbf{w}}_1, \ldots, \hat{\mathbf{w}}_K, \boldsymbol{\lambda} > 0} \sum_{i=1}^{K} \lambda_i \tag{3.8}$$

$$\text{subject to } \mathrm{SINR}_i^{\mathrm{UL}}(\hat{\mathbf{w}}_i, \boldsymbol{\lambda}, \pi) \geq \gamma_i, \quad 1 \leq i \leq K \tag{3.9}$$

$$\|\hat{\mathbf{w}}_i\|_2 = 1, \quad 1 \leq i \leq K. \tag{3.10}$$

$$\sum_{i=1}^{K} \lambda_i < P_T. \tag{3.11}$$

Since uplink SINR expressions are not coupled by beamformers, the SINR functions can be individually maximized for fixed $\boldsymbol{\lambda}$ and π by the MMSE solution:

$$\hat{\mathbf{w}}_i^{\mathrm{MMSE}}(\boldsymbol{\lambda}, \pi) = \beta \mathbf{Z}_i^{-1}(\boldsymbol{\lambda}, \pi)\mathbf{h}_i, \quad 1 \le i \le K \tag{3.12}$$

where β is a normalization constant to assure that $\|\hat{\mathbf{w}}_i\|_2 = 1$. Using $\hat{\mathbf{w}}_i^{\mathrm{MMSE}}$ in the SINR expression, we obtain

$$\mathrm{SINR}_i^{\mathrm{UL}}(\boldsymbol{\lambda}, \pi) = \lambda_i \mathbf{h}_i^H \mathbf{Z}_i^{-1}(\boldsymbol{\lambda}, \pi)\mathbf{h}_i. \tag{3.13}$$

The easy part of the uplink beamforming is the simple expression for beamformers in (3.12). That is, the SINR expressions in uplink are not coupled wih beamformers and they are individually maximized. This does not hold for SINR expressions in downlink beamforming. However, it is shown that the optimal downlink beamformers $\mathbf{w}_i$'s are identical to dual uplink beamfomers $\hat{\mathbf{w}}_i$'s [6], [7].

To achieve a higher SINR value, a user must use more power, however this causes the interference to get higher for other users. To satisfy their SINR value, the other users will want to transmit with more power, which in turn causes the total transmit power to increase. Therefore, it is easily seen that the optimum P^{UL} is achieved when the SINR constraints are active, that is, $\mathrm{SINR}_i^{\mathrm{UL}}$'s are met with equality.

Uplink and Downlink Solution

First, we solve the uplink problem in (3.8). It can be shown that interference-plus-noise covariance matrix has a recursive structure

$$\mathbf{Z}_{\pi_{i-1}}(\boldsymbol{\lambda}, \pi) = \mathbf{Z}_{\pi_i}(\boldsymbol{\lambda}, \pi) + \lambda_{\pi_i}\mathbf{h}_{\pi_i}\mathbf{h}_{\pi_i}^H, \quad 1 \le i \le K \tag{3.14}$$

where $\mathbf{Z}_{\pi_K}(\boldsymbol{\lambda}, \pi) = \sigma^2 \mathbf{I}$. Exploiting this structure the following update formula can be used to compute the matrix inverse in (3.12). For a nonsingular matrix $\mathbf{A}$ and vectors $\mathbf{c}$,$\mathbf{d}$,

$$(\mathbf{A} + \mathbf{c}\mathbf{d}^H)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1} + \mathbf{c}\mathbf{d}^H\mathbf{A}^{-1}}{1 + \mathbf{c}^H\mathbf{A}^{-1}\mathbf{d}}. \tag{3.15}$$

Using the formula above, we obtain

$$\mathbf{Z}_{\pi_{i-1}}(\boldsymbol{\lambda}, \pi)^{-1} = \mathbf{Z}_{\pi_i}(\boldsymbol{\lambda}, \pi)^{-1} + \frac{\lambda_{\pi_i}\mathbf{Z}_{\pi_i}(\boldsymbol{\lambda}, \pi)^{-1}\mathbf{h}_{\pi_i}\mathbf{h}_{\pi_i}^H\mathbf{Z}_{\pi_i}(\boldsymbol{\lambda}, \pi)^{-1}}{1 + \gamma_{\pi_i}}; \tag{3.16}$$

$$\mathbf{Z}_{\pi_K}(\boldsymbol{\lambda}, \pi)^{-1} = \frac{1}{\sigma^2}\mathbf{I}. \tag{3.17}$$

This derivation leads to the following algorithm that obtaines $\hat{\mathbf{w}}^{\min}$ and $\boldsymbol{\lambda}^{\min}$ as the optimum:

1. $\mathbf{Z}_{\pi_K}(\boldsymbol{\lambda}, \pi) \leftarrow \mathbf{I}/\sigma^2$
2. for $\quad i = K \quad$ to $\quad i = 1$
3. $\quad \lambda_{\pi_i}^{\min} \leftarrow \gamma_{\pi_i}/\mathbf{h}_{\pi_i}^H\mathbf{Z}_{\pi_i}^{-1}\mathbf{h}_{\pi_i}$
4. $\quad \hat{\mathbf{w}}_{\pi_i}^{\min} \leftarrow \mathbf{Z}_{\pi_i}^{-1}\mathbf{h}_{\pi_i}/\left\|\mathbf{Z}_{\pi_i}^{-1}\mathbf{h}_{\pi_i}\right\|_2$
5. $\quad \mathbf{Z}_{\pi_i-1}^{-1} \leftarrow \mathbf{Z}_{\pi_i}^{-1} - \lambda_{\pi_i}^{\min}\mathbf{Z}_{\pi_i}^{-1}\mathbf{h}_{\pi_i}\mathbf{h}_{\pi_i}^H\mathbf{Z}_{\pi_i}^{-1}/(1 + \gamma_{\pi_i})$
6. end

The uplink algorithm is easy-to-compute because of the structure of MMSE beamformer in (3.12). Additionally, it does not require any matrix inversion since covariance matrices can be computed recursively with the use of (3.16) .

Before switching to downlink solution we must establish the duality between uplink and downlink beamforming. As stated earlier, the uplink and downlink beamformers are identical. Considering this, the mutual cross-talk between the users, denoted by a matrix $\mathbf{\Psi}_\pi$, is given as

$$[\mathbf{\Psi}_\pi]_{\pi_i,k} = \begin{cases} \left|\mathbf{w}_k^H \mathbf{h}_{\pi_i}\right|^2, & k \in \{\pi_1, \ldots, \pi_{i-1}\} \\ 0, & k \in \{\pi_i, \ldots, \pi_K\} \end{cases}, \forall i. \tag{3.18}$$

Downlink interference observed at the ith user is the ith row of $\mathbf{\Psi}_\pi$, whereas the uplink interference for ith user is the ith column of $\mathbf{\Psi}_\pi$. Denoting $\mathbf{D}_\pi$ as the diagonal normalization matrix and defining it as

$$[\mathbf{D}_\pi]_{i,k} = \begin{cases} \frac{\gamma_{\pi_k}}{\left|\mathbf{w}_{\pi_k}^H \mathbf{h}_{\pi_k}\right|^2}, & k = i \\ 0, & k \neq i \end{cases} \tag{3.19}$$

we can write uplink and downlink expressions in matrix form as

$$(\mathbf{I} - \mathbf{D}_\pi \mathbf{\Psi}_\pi)\, \mathbf{p}_\pi = \sigma^2 \mathbf{D}_\pi \mathbf{1}, \quad \text{(downlink)} \tag{3.20}$$

$$\left(\mathbf{I} - \mathbf{D}_\pi \mathbf{\Psi}_\pi^T\right) \boldsymbol{\lambda}_\pi = \sigma^2 \mathbf{D}_\pi \mathbf{1}, \quad \text{(uplink).} \tag{3.21}$$

For fixed π, the matrices $(\mathbf{I} - \mathbf{D}_\pi \mathbf{\Psi}_\pi)$ and $\left(\mathbf{I} - \mathbf{D}_\pi \mathbf{\Psi}_\pi^T\right)$ are nonsingular. Since $\mathbf{\Psi}_\pi$ has a cascaded structure, it is easy to solve the characteristic equation $\det(\tau \mathbf{I} - \mathbf{D}_\pi \mathbf{\Psi}_\pi) = 0$ (here τ represents the eigenvalue) by Gaussian elimination. For this case, the determinant is the product of the diagonal elements of $\mathbf{D}_\pi \mathbf{\Psi}_\pi$ and since $\tau_K = 0$, the determinant becomes 0. Therefore, the maximal eigenvalue of $\mathbf{D}_\pi \mathbf{\Psi}_\pi$ is 0 [1]. This guarantees that there exist positive solutions to $\mathbf{p}$ and $\boldsymbol{\lambda}$ as

$$\mathbf{p} = \sigma^2 (\mathbf{I} - \mathbf{D}_\pi \mathbf{\Psi}_\pi)^{-1} \mathbf{D}_\pi \mathbf{1}, \text{ (downlink)} \tag{3.22}$$

$$\boldsymbol{\lambda} = \sigma^2 \left(\mathbf{I} - \mathbf{D}_\pi \mathbf{\Psi}_\pi^T\right)^{-1} \mathbf{D}_\pi \mathbf{1}, \text{ (uplink).} \tag{3.23}$$

and $\mathbf{p}$ and $\boldsymbol{\lambda}$ achieve the SINR targets. The minimum required total power for the uplink is

$$\sum_{i=1}^{K} \lambda_k = \sigma^2 \mathbf{1}^T \left(\mathbf{D}_\pi^{-1} - \mathbf{\Psi}_\pi^T\right)^{-1} \mathbf{1} \tag{3.24}$$

$$= \sigma^2 \mathbf{1}^T \left(\mathbf{D}_\pi^{-1} - \mathbf{\Psi}_\pi\right)^{-1} \mathbf{1} = \sum_{i=1}^{K} p_k. \tag{3.25}$$

As shown above, uplink and downlink both require same total power for achieving same SINR targets. Additionally, the same beamformers are used in uplink and downlink. This illustrates the duality between uplink and downlink. From (3.24), the uplink coupling matrix is the transpose of the downlink coupling matrix. This is due to the fact that downlink precoding order is the reverse of uplink decoding order.

Using the duality result, the downlink problem can be easily solved. Having computed the beamformers, the following algorithm finds the optimum downlink power allocation.

1. compute the beamformers $\mathbf{w}_1^{\min}, \ldots, \mathbf{w}_K^{\min}$ by using the uplink algorithm
2. for $\quad i = 1 \;$ to $\; i = K$
3. $\quad p_{\pi_i}^{\min} \leftarrow \left(\gamma_{\pi_i} / \left|\mathbf{w}_{\pi_i}^H \mathbf{h}_{\pi_i}\right|^2\right) \left(\sum_{k \in \{\pi_1, \ldots, \pi_{i-1}\}} p_k^{\min} \left|\mathbf{w}_k^H \mathbf{h}_{\pi_i}\right|^2 + \sigma^2\right)$
4. end.

3.2 Beamforming Techniques Under Per-Antenna Power Constraints

Downlink beamforming techniques have been generally developed under total power constraints. Minimizing transmit power under SINR and total power constraints is analytically easy to solve, but in practice it is far from reality. Since, every antenna has its own amplifier and limited by linear region of the amplifier, per-antenna power constraint based optimization is more practical [2]. Furthermore for a system with BS cooperation, per-antenna power constraint are more natural than total power constraint.

Beamforming under per-antenna power constraints has been previously analyzed in terms of ZF beamforming [12] and beamforming with transmit power margin minimization. ZF beamforming algorithm in [12] attempts to maximize the minimum common rate achieved by all users under per-antenna power constraints. The beamforming algorithm in [2] computes the optimal beamforming vectors minimizing the power margin under individual SINR constraints at the users and per-antenna power constraints. In the sequel, we summarize both of the algorithms.

3.2.1 Equal-Rate Zero-Forcing Transmission

In [12], the ZF scheme is implemented in an elegant manner. The idea behind this approach is that the decrease in capacity is caused from inter-cell interference. Even if the signal power is high, the capacity can be very small because of interference from other users. One approach to mitigate interference is to select the beamforming vectors such that the transmission for each user does not cause interference for any other user. For this method to be applicable, the channel matrix including all users' channel vectors must be full-rank. Although ZF is easy to implement, it is a suboptimal method in terms of achieving the capacity when the signal-to-noise ratio (SNR) is small [12].

In ZF method, the beamformer vector of a user is chosen to be orthogonal to other user's channel vectors. In other words, the beamformer vectors of users do not lie in the subspace spanned by other users' channels. The beamformer vectors are assumed to be unit-norm. As a result, beamforming vectors should satisfy

$$\mathbf{h}_i^H \mathbf{w}_j = 0, \ \forall i \neq j, \ \text{and} \tag{3.26}$$

$$\|\mathbf{w}_i\|_2 = p_i, \ \forall i, \tag{3.27}$$

where p_i is the power allocated to user i. The orthogonality requirement for the beamformers can be satisfied with a series of operations. The channel vector can be decomposed into the sum of two orthogonal vectors

$$\mathbf{h_i} = \mathbf{a}_i + \mathbf{a}_i^{'} \tag{3.28}$$

where $\mathbf{a}_i^{'}$ denotes the component of $\mathbf{h}_i$ which lies in the subspace spanned by other users' channels. The user i's beamforming vectors are confined in the row space of the vector $\mathbf{a}_i^H$.

In order to find $\mathbf{a}_i$, the following procedure is applied. $\mathbf{G} \in \mathbb{C}^{(K-1)\text{x}N}$ matrix whose rows are equal to the rows of $\mathbf{H}$ except the row $\mathbf{h}_i^H$, is formed. Then the orthonormal basis for the range of $\mathbf{G}$ is calculated. The orthonormal basis vectors form the rows of matrix $\mathbf{G}^{'}$. After subtracting the projection of $\mathbf{h}_i^H$ with each of the rows of $\mathbf{G}^{'}$ from $\mathbf{h}_i^H$, we have the vector $\mathbf{a}_i^H$. Mathematically, it is described as follows:

$$\mathbf{a}_i^H = \mathbf{h}_i^H - (\mathbf{h}_i^H(\mathbf{G}_1^{'})^T)\mathbf{G}_1^{'} - \ldots - (\mathbf{h}_i^H(\mathbf{G}_{K-1}^{'})^T)\mathbf{G}_{K-1}^{'} \tag{3.29}$$

where $\mathbf{G}_i^{'}$ is the ith row of $\mathbf{G}^{'}$.

For finding the basis vector for the row space of $\mathbf{a}_i^H$, singular value decomposition (SVD) theorem is used. Since $\mathbf{a}_i^H$ is a row vector, SVD of $\mathbf{a}_i^H$ is simply

$$\mathbf{a}_i^H = \sqrt{\frac{\eta_i}{p_i}}\mathbf{w}_i^H \tag{3.30}$$

where $\eta_i = \mathbf{a}_i^H\mathbf{a}_i$.

The ith user's received signal is given as

$$y_i = \mathbf{h}_i^H\left(\sum_{j=1}^{K} d_j\mathbf{w}_j\right) + n_i \tag{3.31}$$

$$= \left(\mathbf{a}_i + \mathbf{a}_i^{'}\right)^H (d_i\mathbf{w}_i) + n_i \tag{3.32}$$

since $\mathbf{a}_i^{'}$ satisfies $\left(\mathbf{a}_i^{'}\right)^T \mathbf{w}_j = 0,\ j \neq i$, we can write

$$y_i = \mathbf{a}_i^H (d_i \mathbf{w}_i) + \mathbf{n}_i. \tag{3.33}$$

Using (3.30), the equation becomes

$$\begin{aligned} y_i &= \sqrt{\frac{\eta_i}{p_i}} \mathbf{w}_i^H d_i \mathbf{w}_i + n_i & (3.34) \\ &= \sqrt{\eta_i p_i} d_i + n_i. & (3.35) \end{aligned}$$

Doing this transformation we obtain a Gaussian channel whose rate equals $\log_2(1 + \eta_i p_i)$. The power used by transmit antenna t is found by summing up all the contributions to all users from that antenna as: $\sum_{i=1}^{K} \left(\mathbf{w}_i \mathbf{w}_i^H\right)_{t,t}$.

The problem is stated as an optimization problem where the objective is to maximize the minimum rate of users satisfying per-antenna power constraints. The problem is defined as:

$$\begin{aligned} & \max \quad r_0 & (3.36) \\ \text{subject to } \log_2(1 + \eta_i p_i) &\geq r_0, \quad i = 1, \ldots, K,\ r_0 > 0 & (3.37) \\ \sum_{i=1}^{K} \left(\mathbf{w}_i \mathbf{w}_i^H\right)_{t,t} &\leq P_t, \quad t = 1, \ldots, N & (3.38) \end{aligned}$$

where P_t is the maximum available power for antenna t.

Logarithm function is a concave function and the region between logarithm function and the hypercube defined by r_0 is a convex region as shown in Fig. 3.1. Thus, the problem is a convex optimization problem since the constraint set is convex. Therefore, one can use standard convex optimization packages to solve this problem. However, it should be noted that Matlab's optimization toolbox does not support logarithmic functions in the constraint set. Therefore, the optimization problem of interest can not be solved using Matlab. However, a powerful optimization package Yalmip [17] which does not have such limitations on constraint sets can be used to find the optimal power allocation.

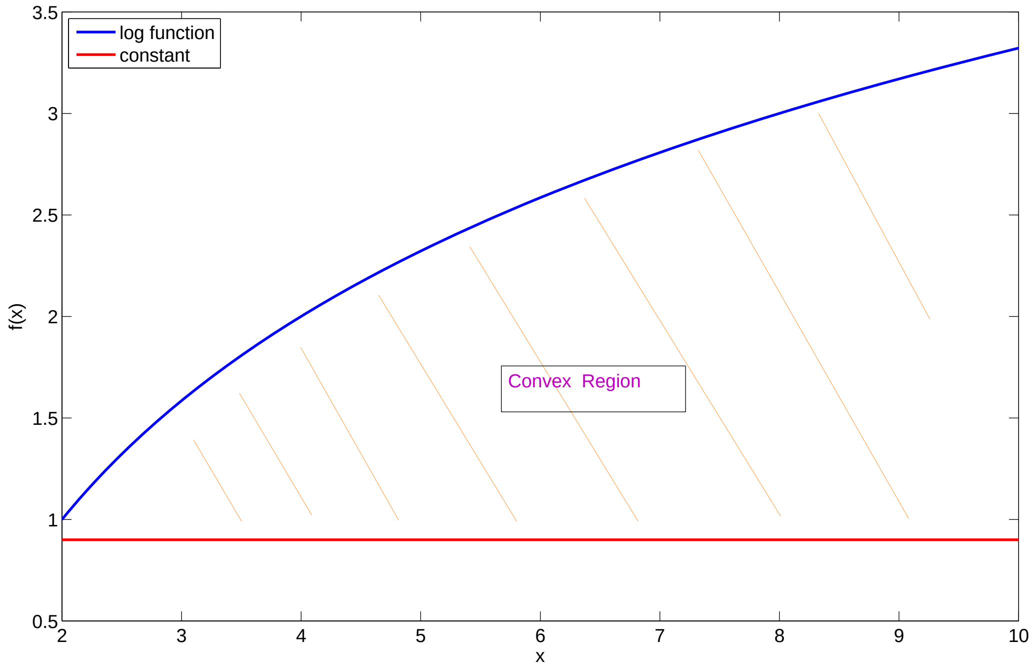

Figure 3.1: The convexity of the area between logarithm function and a constant

One disadvantage in using ZF beamforming vectors in the downlink is that ZF beamforming is near optimal only when SNR is high [12]. Since ZF aims to null out the interference, it is not optimal in terms of maximizing SINR of users. As a result, in certain cases as will be demonstrated in Chapter 5, it uses significantly higher power than other beamforming algorithms minimizing the transmit power under SINR constraints. As shown in [9], the capacity region with per-antenna power constraints is achieved by using DPC with MMSE BF with proper power allocation to users. Therefore, ZF beamforming is suboptimal.

3.2.2 Power Margin Minimization in the Downlink

In the work by Yu and Lan [2], an efficient iterative algorithm that computes the optimum beamforming vectors minimizing the power margin under per-antenna power and individual SINR constraints is proposed. The received signal is

$$y_i = \mathbf{h}_i^H \left(\sum_{j=1}^{K} d_j \mathbf{w}_j \right) + n_i, \; i = 1, \ldots, K. \qquad (3.39)$$

The SINR for each user is expressed as:

$$\text{SINR}_i = \frac{|\mathbf{h}_i^H \mathbf{w}_i|^2}{\sum_{j\neq i} |\mathbf{h}_i^H \mathbf{w}_j|^2 + \sigma^2} \quad i = 1, ..., K. \tag{3.40}$$

The downlink problem is stated as follows:

$$\min_{\alpha, \mathbf{w}_1, ..., \mathbf{w}_K} \quad \alpha \sum_{i=1}^{N} P_i \tag{3.41}$$

$$\text{subject to} \left[\sum_{j=1}^{K} \mathbf{w}_j \mathbf{w}_j^H\right]_{i,i} \leq \alpha P_i, \quad i = 1, ..., N \tag{3.42}$$

$$\frac{|\mathbf{h}_i^H \mathbf{w}_i|^2}{\sum_{j\neq i} |\mathbf{h}_i^H \mathbf{w}_j|^2 + \sigma^2} \geq \gamma_i, \quad i = 1, ..., K \tag{3.43}$$

In this optimization problem the optimal $\mathbf{w}_i$'s are not unique since $\tilde{\mathbf{w}}_i = \mathbf{w}_i e^{j\theta_i}$ also satisfy the constraints with same objective function value. As a result, we use the convention that $\mathbf{w}_i$'s are chosen such that $\mathbf{w}_i^H \mathbf{h}_i$ is real valued. The optimization problem in (3.41)-(3.43) is not convex, but 'Strong Duality' (explained in Appendix B) holds for this problem [2], [18]. Therefore by solving the convex Lagrangian dual problem (explained in Appendix B), the optimal beamforming vectors can be easily found.

The Lagrangian function for the downlink problem is found as:

$$\begin{aligned} L(\alpha, \mathbf{w}_i, \mathbf{Q}, \lambda_i) = \alpha \sum_{i=1}^{N} P_i + \sum_{i=1}^{N} q_i \left(\left[\sum_{j=1}^{K} \mathbf{w}_j \mathbf{w}_j^H\right]_{i,i} - \alpha P_i \right) \\ - \sum_{i=1}^{K} \lambda_i \left(\frac{1}{\gamma_i} |\mathbf{h}_i^H \mathbf{w}_i|^2 - \sum_{j\neq i} |\mathbf{h}_i \mathbf{w}_j|^2 - \sigma^2 \right) \end{aligned} \tag{3.44}$$

where λ_is and $\mathbf{Q} = \text{diag}(q_1, \ldots, q_N)$ are the dual variables corresponding to SINR and per-antenna power constraints, respectively and $\mathbf{\Phi} = \text{diag}(P_1, \ldots, P_N)$. (3.44) can be written in a more compact form:

$$\begin{aligned} L(\alpha, \mathbf{w}_i, \mathbf{Q}, \lambda_i) = \sum_{i=1}^{K} \lambda_i \sigma^2 - \alpha \left\{ \text{Tr}(\mathbf{Q}\mathbf{\Phi}) - \text{Tr}(\mathbf{\Phi}) \right\} \\ + \sum_{i=1}^{K} \mathbf{w}_i^H \left(\mathbf{Q} + \sum_{j\neq i} \lambda_j \mathbf{h}_j \mathbf{h}_j^H - \frac{\lambda_i}{\gamma_i} \mathbf{h}_i \mathbf{h}_i^H \right) \mathbf{w}_i \end{aligned} \tag{3.45}$$

We can state the dual objective function for the Lagrangian problem as

$$g(\mathbf{Q}, \lambda_i) = \min_{\mathbf{w}_i, \alpha} L(\alpha, \mathbf{w}_i, \mathbf{Q}, \lambda_i). \tag{3.46}$$

Since there is not any constraint on the beamformer $\mathbf{w}_i$ and α is a positive number, $g(\mathbf{Q}, \lambda_i) = -\infty$ if $\mathrm{Tr}(\mathbf{Q\Phi}) \geq \mathrm{Tr}(\Phi)$ or $\mathbf{Q} + \sum_{j \neq i} \lambda_i \mathbf{h}_j \mathbf{h}_j^H - \frac{\lambda_j}{\gamma_i} \mathbf{h}_i \mathbf{h}_i^H$ is not positive semidefinite. For the Lagrangian dual $g(\mathbf{Q}, \lambda_i)$ to give a meaningful lower bound to the optimal value of the original problem, it must be finite, so $\mathbf{Q}$ and λ_i should be chosen accordingly. The Lagrangian dual problem can be stated as follows:

$$\max_{\mathbf{Q}} \max_{\lambda_i} \quad \sum_{i=1}^{K} \lambda_i \sigma^2 \tag{3.47}$$

$$\mathbf{Q} + \sum_{j=i}^{K} \lambda_j \mathbf{h}_j \mathbf{h}_j^H \succeq \Big(1 + \frac{1}{\gamma_i}\Big) \lambda_i \mathbf{h}_i \mathbf{h}_i^H \tag{3.48}$$

$$\mathrm{Tr}(\mathbf{Q\Phi}) \leq \mathrm{Tr}(\mathbf{\Phi}), \ \mathbf{Q} \text{ diagonal}, \ \mathbf{Q} \succeq 0. \tag{3.49}$$

It is shown in [2] that the Lagrangian dual problem in (3.47) is equivalent to the following dual problem with same SINR constraints.

$$\max_{\mathbf{Q}} \min_{\lambda_i, \hat{\mathbf{w}}_i} \sum_{i=1}^{K} \lambda_i \sigma^2 \tag{3.50}$$

$$\text{subject to} \quad \frac{\lambda_i \sigma^2 |\hat{\mathbf{w}}_i^H \mathbf{h}_i|^2}{\sum_{j \neq i} \lambda_j \sigma^2 |\hat{\mathbf{w}}_i^H \mathbf{h}_j|^2 + \hat{\mathbf{w}}_i^H \sigma^2 \mathbf{Q} \hat{\mathbf{w}}_i} \geq \gamma_i, i = 1, ..., K \tag{3.51}$$

$$\mathrm{Tr}\left(\mathbf{Q\Phi}\right) \leq \mathrm{Tr}\left(\mathbf{\Phi}\right), \ \mathbf{Q} \text{ diagonal}, \ \mathbf{Q} \succeq 0. \tag{3.52}$$

where $\hat{\mathbf{w}}_i$ is the dual uplink beamformer, $\lambda_i \sigma^2$ is the dual uplink power and $\sigma^2 \mathbf{Q}$ is the noise covariance matrix. The minimization of total uplink power under minimum SINR constraints is achieved by MMSE beamforming vectors:

$$\hat{\mathbf{w}}_i = \left(\sum_{j=1}^{K} \lambda_j \mathbf{h}_j \mathbf{h}_j^H + \mathbf{Q} \right)^{-1} \mathbf{h}_i. \tag{3.53}$$

Since minimum total power is obtained when SINR targets are met with equality,

$$\Big(1 + \frac{1}{\gamma_i}\Big) \lambda_i \mathbf{h}_i^H \Big(\sum_{j=1}^{K} \lambda_j \mathbf{h}_j \mathbf{h}_j^H + \mathbf{Q} \Big)^{-1} \mathbf{h}_i = 1. \tag{3.54}$$

Therefore, the optimal λ_i should be the unique fixed-point of the following equation

$$\lambda_i^* = \frac{1}{\left(1+\frac{1}{\gamma_i}\right)\mathbf{h}_i^H\left(\sum_{j=1}^K \lambda_j^* \mathbf{h}_j \mathbf{h}_j^H + \mathbf{Q}\right)^{-1}\mathbf{h}_i}, \forall i. \tag{3.55}$$

Then, the dual uplink problem is reduced to

$$\max_{\mathbf{Q}} \sum_{i=1}^K \lambda_i \sigma^2 \tag{3.56}$$

$$\text{subject to } \lambda_i^* = \frac{1}{\left(1+\frac{1}{\gamma_i}\right)\mathbf{h}_i^H\left(\sum_{j=1}^K \lambda_j^* \mathbf{h}_j \mathbf{h}_j^H + \mathbf{Q}\right)^{-1}\mathbf{h}_i}, \forall i. \tag{3.57}$$

$$\mathrm{Tr}\left(\mathbf{Q}\boldsymbol{\Phi}\right) \leq \mathrm{Tr}\left(\boldsymbol{\Phi}\right),\ \mathbf{Q} \text{ diagonal},\ \mathbf{Q} \succeq 0. \tag{3.58}$$

It is shown in [2] that this is a concave optimization problem which can be solved by subgradient projection algorithm where diag$\left\{\sum_{i=1}^K \mathbf{w}_i \mathbf{w}_i^H\right\}$ is the subgradient of $\mathbf{Q}$. Since the problem is concave, the subgradient projection method converges to the globally optimum $\mathbf{Q}$.

Once λ_i and $\mathbf{Q}$ is found, the optimal beamforming vectors of downlink problems are found by taking the derivative of the Lagrangian in (3.45) with respect to $\mathbf{w}_i$ and equating it to 0:

$$\partial L/\partial \mathbf{w}_i = \left(\mathbf{Q} + \sum_{j\neq i} \lambda_j \mathbf{h}_j \mathbf{h}_j^H - \frac{\lambda_i}{\gamma_i}\mathbf{h}_i \mathbf{h}_i^H\right)\mathbf{w}_i = 0. \tag{3.59}$$

If we add $\left(1+\frac{1}{\gamma_i}\right)\lambda_i \mathbf{h}_i \mathbf{h}_i^H \mathbf{w}_i$ to both sides and solve for $\mathbf{w}_i$:

$$\mathbf{w}_i = \left(\sum_{j=1}^K \lambda_j \mathbf{h}_j \mathbf{h}_j^H + \mathbf{Q}\right)^{-1}\left(1+\frac{1}{\gamma_i}\right)\lambda_i \mathbf{h}_i \mathbf{h}_i^H \mathbf{w}_i. \tag{3.60}$$

where $\mathbf{h}_i^H \mathbf{w}_i$ expression is assumed to be real valued and positive. We can easily see that $\mathbf{w}_i$ is a scalar multiple of $\hat{\mathbf{w}}_i$. Denoting the scalar as $\sqrt{\delta_i}$:

$$\sqrt{\delta_i} = \sigma^2\left(1+\frac{1}{\gamma_i}\right)\lambda_i \mathbf{h}_i^H \mathbf{w}_i. \tag{3.61}$$

As seen in the expression, the scalar is a function of $\mathbf{w}_i$. At this point,we exploit the condition that SINRs are met with equality:

$$\frac{1}{\gamma_i}\left|\mathbf{w}_i^H\mathbf{h}_i\right|^2 = \sum_{i\neq j}\left|\mathbf{w}_j^H\mathbf{h}_i\right|^2 + \sigma^2. \tag{3.62}$$

If $\mathbf{w}_i = \sqrt{\delta_i}\hat{\mathbf{w}}_i$ is substituted in the above expression, we obtain K equations with K unknowns as:

$$\mathbf{G}\left[\delta_1 \ldots \delta_K\right]^T = \mathbf{1}\sigma^2 \tag{3.63}$$

where matrix $\mathbf{G}$ is defined as

$$\mathbf{G}_{i,j} = \begin{cases} \frac{1}{\gamma_i}|\hat{\mathbf{w}}_i^H\mathbf{h}_i|^2 & \text{if } i = j \\ -|\hat{\mathbf{w}}_j^H\mathbf{h}_i|^2 & \text{else} \end{cases} \tag{3.64}$$

Based on these results and definitions, the iterative algorithm is stated as follows:

1. Set $n = 1$ and initialize $\mathbf{Q}^{(1)}$,
2. Solve the following equation by fixed-point iteration for fixed $\mathbf{Q}^{(n)}$:
   $(\lambda_i^*)^{(n)} = \frac{1}{\left(1+\frac{1}{\gamma_i}\right)\mathbf{h}_i^H\left(\sum_{j=1}^K\left(\lambda_j^*\right)^{(n)}\mathbf{h}_j\mathbf{h}_j^H+\mathbf{Q}^{(n)}\right)^{-1}\mathbf{h}_i}$,
3. Calculate optimal uplink beamformers using λ_i^* and downlink beamformers
   $\hat{\mathbf{w}}_i^{(n)} = \left(\sum_{i=1}^K\left(\lambda_j^*\right)^{(n)}\mathbf{h}_j\mathbf{h}_j^H + \mathbf{Q}^{(n)}\right)^{-1}\mathbf{h}_i$
   $\mathbf{w}_i^{(n)} = \sqrt{\delta}_i\hat{\mathbf{w}}_i^{(n)}$ where $\delta = \mathbf{G}^{-1}\mathbf{1}\sigma^2$
4. Update $\mathbf{Q}^{(n+1)}$ by subgradient projection method summarized in Appendix C with step size t_n (possible choices of t_n are explained in Appendix C):
   $\mathbf{Q}^{(n+1)} = P\left\{\mathbf{Q}^{(n)} + t_n\text{diag}\left\{\sum_{i=1}^K\mathbf{w}_i^{(n)}\left(\mathbf{w}_i^{(n)}\right)^H\right\}\right\}$
   where P denotes the projection of the subgradient of the function onto the constraint set composed of the constraints: $\text{Tr}(\mathbf{Q}\mathbf{\Phi}) \leq \text{Tr}(\mathbf{\Phi})$ and $\mathbf{Q} \succeq 0$.

As stated in [2], one can modify the algorithm to include the case where DPC is used together with beamforming.

One of the possible problems with the problem formulation in [2] is that, the objective is to minimize the power margin as opposed to the total transmit power. When the system is asymmetric, the resulting beamforming vectors may use significantly larger transmit power compared to the optimal beamforming vectors minimizing the total transmit power under same SINR constraints.

This formulation can also return infeasible per-antenna power levels, that is power values exceeding the maximum power level. The only assumption about α is its positiveness. For the problem to be feasible, the transmission powers must exactly be lower or equal to the per-antenna power constraints. But, in fact the optimization may yield $\alpha > 1$, which means the problem is infeasible in terms of power constraints.

The power margin minimization problem also requires some complex optimization functionalities such as subgradient projection method. For this reason, the time for convergence is very high in some cases. The power margin minimization problem is also very difficult to be implemented in a distributed manner. For these reasons, there is a need for another algorithm that is easy-to-implement, not time consuming and can be easily implemented distributively. As a result, we reformulate the optimization problem in [2] where the objective is to minimize the total transmit power.

Chapter 4

PROPOSED BEAMFORMING ALGORITHMS

As discussed in Chapter 3, the beamforming optimization problem in [2] focus on minimizing the worst case power margin for each antenna which is defined as the ratio of the power transmitted on each antenna to the corresponding power constraint. When the system is asymmetric, i.e. users have different power and/or SINR constraints or the channel for users are different, optimizing the power margin may result in excessive use of power to satisfy the SINR constraints. While formulating the problem as a power margin minimization problem provides an alternative viewpoint for formulating the well-known duality between uplink and downlink within the Lagrangian dual problem framework, from a system designer's point of view, it is more critical to provide efficient use of resources (transmit power in this case) rather than minimizing the power margin.

As a result, we reformulate the optimization problem considered in [2] to optimize the total transmit power. Using Lagrangian dual framework, we provide an iterative algorithm for computing the optimum beamforming vectors. For implementing the algorithm in a practical system with BS cooperation, we need to limit the amount of information exchange between BS required to compute beamforming vectors. As a result, we investigate the distributed implementation of the proposed algorithm using only limited local information exchange between BSs. In this chapter, we present the proposed centralized beamforming vector computation algorithm and its distributed implementation.

4.1 Centralized Algorithm

We follow the dual problem formulation in [2]. We define the objective as the total power used and keep the same constraints. This change causes the Lagrangian dual function and the algorithmic solution to change. In the reformulated optimization, downlink problem is stated as follows:

$$\min_{\mathbf{w}_1,\ldots,\mathbf{w}_K} \quad \sum_{i=1}^{K} \|\mathbf{w}_i\|^2 \tag{4.1}$$

$$\text{subject to} \left[\sum_{j=1}^{K} \mathbf{w}_j \mathbf{w}_j^H\right]_{i,i} \leq P_i, \quad i = 1, ..., N \tag{4.2}$$

$$\frac{|\mathbf{h}_i^H \mathbf{w}_i|^2}{\sum_{j\neq i} |\mathbf{h}_i^H \mathbf{w}_j|^2 + \sigma^2} \geq \gamma_i, \quad i = 1, ..., K \tag{4.3}$$

The Lagrangian for the downlink problem is found as:

$$\begin{aligned} L(\mathbf{w}_i, \mathbf{Q}, \lambda_i) = & \sum_{i=1}^{K} |\mathbf{w}_i|^2 + \sum_{i=1}^{N} q_i \Big(\Big[\sum_{j=1}^{K} \mathbf{w}_j \mathbf{w}_j^H\Big]_{i,i} - P_i \Big) \\ & - \sum_{i=1}^{K} \lambda_i \Big(\frac{1}{\gamma_i} |\mathbf{h}_i^H \mathbf{w}_i|^2 - \sum_{j\neq i} |\mathbf{h}_i \mathbf{w}_j|^2 - \sigma^2 \Big) \end{aligned} \tag{4.4}$$

and (4.4) can be written in a more compact form:

$$\begin{aligned} L(\mathbf{w}_i, \mathbf{Q}, \lambda_i) = & \sum_{i=1}^{K} \lambda_i \sigma^2 - \mathrm{Tr}(\mathbf{Q}\boldsymbol{\Phi}) \\ & + \sum_{i=1}^{K} \mathbf{w}_i^H \Big(\mathbf{Q} + \mathbf{I} + \sum_{j\neq i} \lambda_j \mathbf{h}_j \mathbf{h}_j^H - \frac{\lambda_i}{\gamma_i} \mathbf{h}_i \mathbf{h}_i^H \Big) \mathbf{w}_i. \end{aligned} \tag{4.5}$$

The dual objective function for the Lagrangian problem becomes

$$g(\mathbf{Q}, \lambda_i) = \min_{\mathbf{w}_i} L(\mathbf{w}_i, \mathbf{Q}, \lambda_i). \tag{4.6}$$

The dual problem is stated as

$$\max_{\mathbf{Q},\lambda_i} g(\mathbf{Q}, \lambda_i) \tag{4.7}$$

$$\mathbf{Q} \succeq 0 \tag{4.8}$$

$$\lambda_i > 0, \ \forall i. \tag{4.9}$$

For the dual problem to give a meaningful lower bound on the optimal value of the original problem, $g(\mathbf{Q}, \lambda_i)$ must be bounded away from $-\infty$. As a result, $\mathbf{Q}$ and λ_i should be such that $\mathbf{Q}+\mathbf{I}+\sum_{j\neq i} \lambda_j \mathbf{h}_j \mathbf{h}_j^H - \frac{\lambda_i}{\gamma_i} \mathbf{h}_i \mathbf{h}_i^H$ is positive semi-definite.

One can show that strong duality holds of the optimization problem in (4.1) using the same approach in [2]. As a result, the nonconvex optimization problem in (4.1) can be solved by its convex dual problem in (4.7).

The optimal beamforming vectors of downlink problems are found by taking the derivative of the Lagrangian with respect to $\mathbf{w}_i$ and equating it to 0:

$$\partial L/\partial \mathbf{w}_i = \left(\mathbf{Q} + \mathbf{I} + \sum_{j\neq i} \lambda_j \mathbf{h}_j \mathbf{h}_j^H - \frac{\lambda_i}{\gamma_i} \mathbf{h}_i \mathbf{h}_i^H \right) \mathbf{w}_i = 0. \tag{4.10}$$

If we add $\left(1 + \frac{1}{\gamma_i}\right) \lambda_i \mathbf{h}_i \mathbf{h}_i^H \mathbf{w}_i$ to both sides and solve for $\mathbf{w}_i$:

$$\mathbf{w}_i = \left(\sum_{j=1}^{K} \lambda_j \mathbf{h}_j \mathbf{h}_j^H + \mathbf{Q} \right)^{-1} \left(1 + \frac{1}{\gamma_i} \right) \lambda_i \mathbf{h}_i \mathbf{h}_i^H \mathbf{w}_i. \tag{4.11}$$

where $\mathbf{h}_i^H \mathbf{w}_i$ expression is assumed to be real valued and positive (since optimal $\mathbf{w}_i$'s are not unique as explained earlier). We can easily see that $\mathbf{w}_i$ is a scalar multiple of $\hat{\mathbf{w}}_i$. Denoting the scalar as $\sqrt{\delta_i}$:

$$\mathbf{w}_i = \underbrace{\left(\sum_{j=1}^{K} \lambda_j \mathbf{h}_j \mathbf{h}_j^H + \mathbf{Q} \right)^{-1} \mathbf{h}_i}_{\hat{\mathbf{w}}_i} \underbrace{\left(1 + \frac{1}{\gamma_i} \right) \lambda_i \mathbf{h}_i^H \mathbf{w}_i}_{\sqrt{\delta_i}}. \tag{4.12}$$

Solving this equation for λ_i here by multiplying both sides with $\mathbf{h}_i^H$ and cancelling $\mathbf{h}_i^H \mathbf{w}_i$ on both sides, we obtain the following fixed-point equation [19]:

$$\lambda_i^* = \frac{1}{\left(1 + \frac{1}{\gamma_i}\right) \mathbf{h}_i^H \left(\sum_{j=1}^K \lambda_j^* \mathbf{h}_j \mathbf{h}_j^H + \mathbf{Q} + \mathbf{I}\right)^{-1} \mathbf{h}_i}, \forall i. \tag{4.13}$$

To minimize transmit power, SINR constraints must be met with equality.

$$\frac{1}{\gamma_i} \left|\mathbf{w}_i^H \mathbf{h}_i\right|^2 = \sum_{i \neq j} \left|\mathbf{w}_j^H \mathbf{h}_i\right|^2 + \sigma^2. \tag{4.14}$$

If we substitute (4.12) into the above, we can solve for $\sqrt{\delta}_i$'s as in [2].

Therefore, the Lagrangian dual problem is stated as

$$\max_{\mathbf{Q}} \min_{\lambda_i, \hat{\mathbf{w}}_i} \sum_{i=1}^K \lambda_i \sigma^2 - \text{Tr}(\mathbf{Q}\mathbf{\Phi}) \tag{4.15}$$

$$\text{subject to } \lambda_i = \frac{1}{\left(1 + \frac{1}{\gamma_i}\right) \mathbf{h}_i^H \left(\sum_{j=1}^K \lambda_j \mathbf{h}_j \mathbf{h}_j^H + \mathbf{Q} + \mathbf{I}\right)^{-1} \mathbf{h}_i}, \forall i. \tag{4.16}$$

$$\mathbf{Q} \text{ diagonal}, \mathbf{Q} \succeq 0. \tag{4.17}$$

This dual problem has two parts. An inner minimization part and an outer maximization part. The inner minimization was shown to be solved via fixed-point iterations. It is shown in [2] that $\min_{\lambda_i, \hat{\mathbf{w}}_i} \sum_{i=1}^K \lambda_i \sigma^2$ is a concave function of $\mathbf{Q}$. Since $\text{Tr}(\mathbf{Q}\mathbf{\Phi})$ is a convex function, $-\text{Tr}(\mathbf{Q}\mathbf{\Phi})$ is a concave function of $\mathbf{Q}$. Since the addition of two concave functions is concave, then $\min_{\lambda_i, \hat{\mathbf{w}}_i} \sum_{i=1}^K \lambda_i \sigma^2 - \text{Tr}(\mathbf{Q}\mathbf{\Phi})$ is also a concave function of $\mathbf{Q}$. Following the same procedure in [2], we can show that $\text{diag}\left\{\sum_{i=1}^K \mathbf{w}_i^{(n)} \left(\mathbf{w}_i^{(n)}\right)^H\right\} - \mathbf{\Phi}$ is a subgradient of the inner minimization part. Therefore, the outer maximization can be solved with subgradient projection method. The subgradient projection here reduces to just comparison of the diagonals of the subgradient matrix with $\mathbf{0}$ because of the only constraint $\mathbf{Q} \succeq 0$. This method is guaranteed to converge to the optimum value, since the inner part is a concave function of $\mathbf{Q}$ [20].

The proposed algorithm that solves the downlink beamforming problem is summarized as follows:

1. First check the SINR values for feasibility using the procedure given in Appendix A. If the problem is not feasible, terminate the algorithm.

2. Set n=1 and initialize $\mathbf{Q}^{(1)}$.

3. Solve fixed-point equation with fixed $\mathbf{Q}^{(n)}$:
$(\lambda_i^*)^{(n)} = \frac{1}{\left(1+\frac{1}{\gamma_i}\right)\mathbf{h}_i^H\left(\sum_{j=1}^{K}\left(\lambda_j^*\right)^{(n)}\mathbf{h}_j\mathbf{h}_j^H+\mathbf{Q}^{(n)}+\mathbf{I}\right)^{-1}\mathbf{h}_i}.$

4. Compute uplink beamformer vectors using fixed-point values:
$\hat{\mathbf{w}}_i^{(n)} = \left(\sum_{j=i}^{K}\left(\lambda_j^*\right)^{(n)}\mathbf{h}_j\mathbf{h}_j^H + \mathbf{Q}^{(n)} + \mathbf{I}\right)^{-1}\mathbf{h}_i.$

5. Compute downlink beamformer vectors using $\mathbf{G}$ matrix defined in (3.64):
$\mathbf{w}_i^{(n)} = \sqrt{\delta_i}\hat{\mathbf{w}}_i^{(n)}.$

6. Update $\mathbf{Q}^{(n)}$ with step size t_n (possible choices of t_n are explained in Appendix C):
$\mathbf{q}_i^{(n+1)} = \max\left(\mathbf{0}, \mathbf{q}_i^{(n)} + t_n\left[\left(\sum_{i=1}^{K}\mathbf{w}_i^{(n)}\left(\mathbf{w}_i^{(n)}\right)^H\right)_{i,i} - P_i\right]\right).$

7. $n \leftarrow n+1$.

8. Go to Step 2 and repeat the procedure until convergence.

Beamforming with DPC:

One can also incorporate the case where DPC along with beamforming is used. The encoding order $\overline{\pi}$ is assumed to be fixed. Based on this assumption, the SINR constraint becomes:

$$\frac{|\mathbf{h}_{\overline{\pi}_i}^H\mathbf{w}_{\overline{\pi}_i}|^2}{\sum_{j\in\{\overline{\pi}_{i+1},\ldots,\overline{\pi}_K\}}|\mathbf{h}_{\overline{\pi}_i}^H\mathbf{w}_{\overline{\pi}_j}|^2+\sigma^2} \geq \gamma_{\overline{\pi}_i}, \quad i=1,\ldots,K \tag{4.18}$$

The fixed-point iteration and computation of $\mathbf{w}_i$ is modified according to the SINR constraints given in (4.18).

4.2 Suboptimal Distributed Algorithm

In the previous section, we propose a centralized algorithm that solves the power minimization problem subject to per-antenna power constraints and SINR constraints. In this section, the distributed version of the proposed algorithm is presented. However, for the distributed algorithm, we assume that the cellular scenario under investigation is Wyner's linear array with $N = K$. The proposed algorithm requires high amount of message-passing in the case of BS cooperation, therefore the optimal distributed scheme with limited amount of information exchange between BSs is not possible and we propose a suboptimal distributed algorithm.

If we observe the expression for the uplink beamformer $\hat{\mathbf{w}}_i$, we see that the expression has a special structure. To compute $\hat{\mathbf{w}}_i$ distributively, it can be modeled as in the framework in [21]. In [21], a distributed solution for computing the MMSE beamformer vector by formulating it as a dual linear minimum-mean-squared-error (LMMSE) estimation problem is proposed.

We first formulate the computation of $\hat{\mathbf{w}}_i$ in the framework in [21]. Having defined λ_i and $\mathbf{Q}$ in the previous section and defining $\mathbf{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_K)$, $\tilde{\mathbf{H}} = \mathbf{H}\mathbf{\Lambda}^{1/2}$, $\mathbf{\Gamma} = \text{diag}(q_1 + 1, \ldots, q_K + 1)$ and $\tilde{\mathbf{n}} \sim N\left(0, [q_1 + 1, \ldots, q_K + 1]^T\right)$, we can write observation equation for the dual LMMSE estimation problem as

$$\mathbf{y} = \tilde{\mathbf{H}}\mathbf{x} + \tilde{\mathbf{n}}. \tag{4.19}$$

Let $\hat{\mathbf{x}}$ be the estimate of $\mathbf{x}$ given the observation $\mathbf{y}$. A LMMSE estimator for $\hat{\mathbf{x}}$ is the following [21], [22]:

$$\hat{\mathbf{x}} = \tilde{\mathbf{H}}^T \left(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^T + \mathbf{\Gamma}\right)^{-1} \mathbf{y}. \tag{4.20}$$

The above expression is equivalent to

$$\hat{\mathbf{x}} = \left(\tilde{\mathbf{H}}^T\tilde{\mathbf{H}} + \mathbf{\Gamma}\right)^{-1} \tilde{\mathbf{H}}^T\mathbf{y}. \tag{4.21}$$

We can manipulate $\hat{\mathbf{w}}_i$ expression so that it looks like the above expression and can be solved via its dual LMMSE problem as in [21]:

$$\hat{\mathbf{w}}_i = \left(\sum_{j=i}^{K} \lambda_j \mathbf{h}_j \mathbf{h}_j^H + \mathbf{Q} + \mathbf{I}\right)^{-1} \mathbf{h}_i = \left(\mathbf{H}^T \mathbf{\Lambda} \mathbf{H} + \mathbf{Q} + \mathbf{I}\right)^{-1} \mathbf{h}_i,\ \forall i \qquad (4.22)$$

$$= \left(\tilde{\mathbf{H}}^T \tilde{\mathbf{H}} + \mathbf{\Gamma}\right)^{-1} \mathbf{h}_i \mathbf{\Lambda}^{1/2} \mathbf{\Lambda}^{-1/2}, \forall i. \qquad (4.23)$$

Let $\hat{\mathbf{W}} = [\hat{\mathbf{w}}_1, \ldots, \hat{\mathbf{w}}_K]$, the above equation can be stated in a compact form:

$$\hat{\mathbf{W}} = \left(\tilde{\mathbf{H}}^T \tilde{\mathbf{H}} + \mathbf{\Gamma}\right)^{-1} \mathbf{H} \mathbf{\Lambda}^{1/2} \mathbf{\Lambda}^{-1/2} \qquad (4.24)$$

$$= \left(\tilde{\mathbf{H}}^T \tilde{\mathbf{H}} + \mathbf{\Gamma}\right)^{-1} \tilde{\mathbf{H}} \mathbf{\Lambda}^{-1/2} \qquad (4.25)$$

The above equation states that we can form a new observation equation with $\hat{\mathbf{W}}$ as data matrix to be estimated and $\mathbf{\Lambda}^{-1/2}$ as observation and apply the above estimator for finding $\hat{\mathbf{W}}$:

$$\mathbf{\Lambda}^{-1/2} = \tilde{\mathbf{H}} \hat{\mathbf{W}} + \tilde{\mathbf{N}} \qquad (4.26)$$

where the columns of $\tilde{\mathbf{N}}$ are the noise vectors in the observation equations of dual LMMSE problem.

The LMMSE estimates which correspond to beamforming matrix $\hat{\mathbf{W}}$ can be computed via forward-backward algorithm based on Kalman smoothing in [23]. The forward and backward Kalman filters are initially estimated by BS 1 and BS N as seen in Fig. 4.1. The output of BS 1 is passed to BS 2 and the output of BS N is passed to BS $N-1$. The BSs which received the information compose their new messages with the data they take from the neighbour (correction term) and the data they estimate (prediction term). When all the BSs have received messages from their right and left neighbours, the messages at the BSs are combined and the data is estimated. After K iterations (message-passing from the first BS to the last BS finishes), the uplink beamforming vectors are estimated. In the kth iteration, the ith BS estimates $(\hat{\mathbf{w}}_k)_i$.

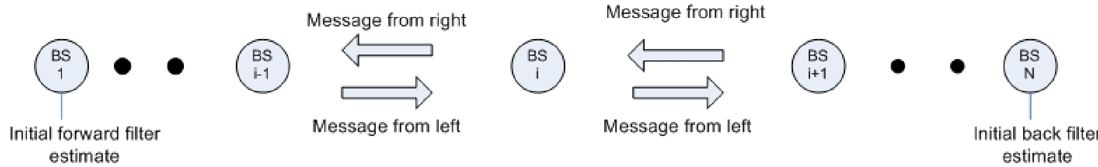

Figure 4.1: Message-passing between neighbour BSs

The fixed-point equation can be written as follows:

$$\lambda_i^* = \frac{1}{\left(1 + \frac{1}{\gamma_i}\right) \mathbf{h}_i^H \hat{\mathbf{w}}_i} \tag{4.27}$$

The above equation implies that every node can compute its fixed-point by same method after computing $\hat{\mathbf{w}}_i$'s. We start with a all-zero $\boldsymbol{\lambda}$ vector. BS N ($N = K$) computes $\boldsymbol{\lambda}_K$ and sends this vector to BS $N-1$, having this information BS $N-1$ computes $\boldsymbol{\lambda}_{K-1}$ and it sends this vector to BS $N-2$. The process is continued back and forward until convergence. After convergence the $\boldsymbol{\lambda}$ vector is passed to the all BS from their neighbours.

In the centralized algorithm the downlink beamformers are scaled versions of uplink beamformers with $\sqrt{\delta}$. Since $\delta = \mathbf{G}^{-1}\mathbf{1}\sigma^2$, the $\mathbf{G}$ matrix must be computed distributively in order to find downlink beamformers. If we observe the structure of $\mathbf{G}$ matrix, we see that the entries except from the diagonals and the band diagonals are significantly smaller than the diagonals. Additionally, BS i can compute $\mathbf{G}_{i,i+1}$ and $\mathbf{G}_{i,i-1}$ with the local information gathered while computing the uplink beamformers. Therefore, we can approximate the $\mathbf{G}$ matrix as tridiagonal. To find the beamforming vectors, the distributed formula for computing the inverse of tridiagonal matrices called Thomas algorithm in [24] is used. First BS N ($N = K$) computes δ_K and it passes this value to BS $N-1$ and BS $N-1$ computes δ_{K-1} and it passes it to BS $N-2$, and the process goes on similarly. After message-passing finishes from BS N to BS 1, δ_i's are computed. Since the inverse of $\mathbf{G}$ matrix cannot be exactly computed, an error is introduced at this stage and this error propagates through the iterations.

Having computed the scalar δ_i's all BSs can compute the downlink beamformers, more precisely BS i can compute $(\mathbf{w}_1)_i, \ldots, (\mathbf{w}_K)_i$.

The last step (subgradient projection) is also done distributively. the projected term in the expression is nothing but the difference between the power used by the BS and the power constraint of that BS. All the BSs are informed priorly what the projection step size is. Once they are converged, they inform each other and the iteration stops with convergence.

Chapter 5

NUMERICAL RESULTS

In this chapter, we provide some numerical examples that compare the proposed methods (centralized and distributed algorithms) with the existing methods in terms of performance, reliability and efficiency for various cellular array scenarios. In the numerical computations, for the sake of simplicity and to gain insight, we consider simple array models, such as Wyner's general circular array, Wyner's symmetric circular array and Wyner's linear array with the same interference factor.

5.1 Centralized Algorithm

To verify that the proposed algorithm is working properly, we compare it with the method proposed by [10] that finds the minimum total power subject to SINR constraints under different channel scenarios. By providing the proposed algorithm with sufficiently large per-antenna power constraints to satisfy the SINR requirements, we computed the beamforming vectors minimizing total power. In Fig. 5.1, the total transmit power as a function of SINR constraint on all users are plotted for Wyner's symmetric circular array with 10 cells for different interference parameters.

As shown in Fig. 5.1, the results obtained by the proposed method perfectly match the results obtained by the method in [1]. It is also observed that the transmit power for $\alpha = 0.5$ is larger than that of $\alpha = 0.75$ (which corresponds to higher interference from neighbour BS). This is due to the fact that as BSs cooperate, they can take the advantage of large channel gains from other users.

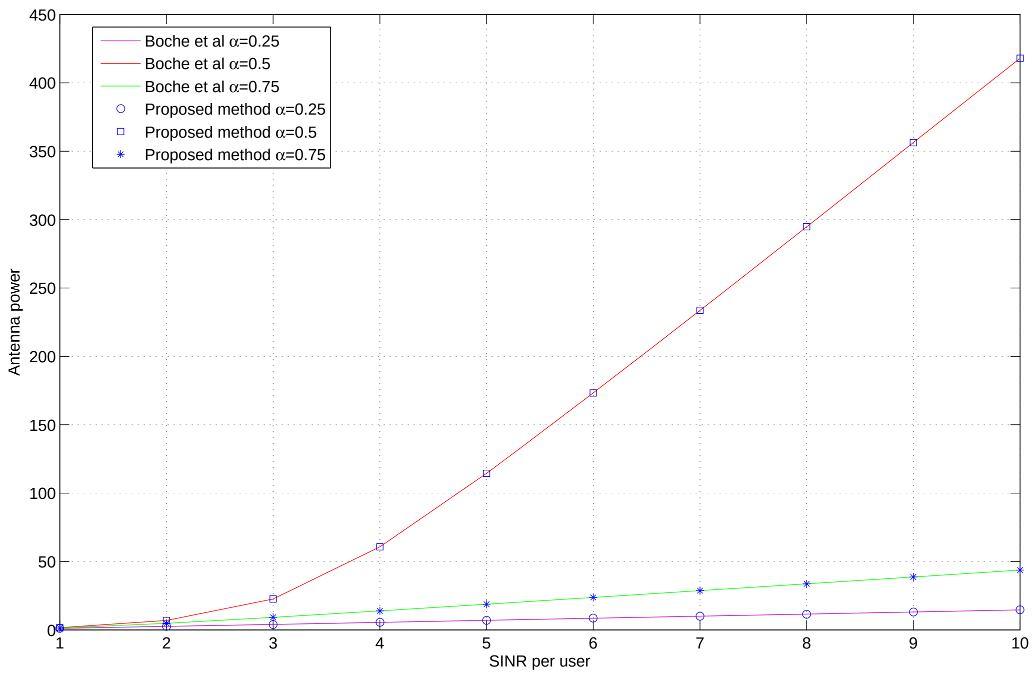

Figure 5.1: Comparison of the proposed method with the method in [1] (denoted as Boche et.al.) in terms of total transmit power for Wyner's symmetric circular array with $N = 10$ and various α values

The proposed method is then compared with the method in [2]. As stated in Chapter 4, the proposed method is expected to use lower total transmit power compared to the method in [2] when the channel is not symmetric. The average antenna power for the beamforming vector computed by using the two algorithms is first compared for Wyner's circular and linear arrays with $N = 10$ cells and interference parameter $\alpha = 0.3$. The per-antenna power constraints are chosen large enough to satisfy the SINR constraints. Since circular array has a symmetrical structure, the total transmit power (which is proportional to average power) for both methods are the same. However, for the linear arrays due to the asymmetry for the cells at the edge, there is a slight difference in the total transmit power when SINR constraints for the users are large as seen in Fig. 5.2.

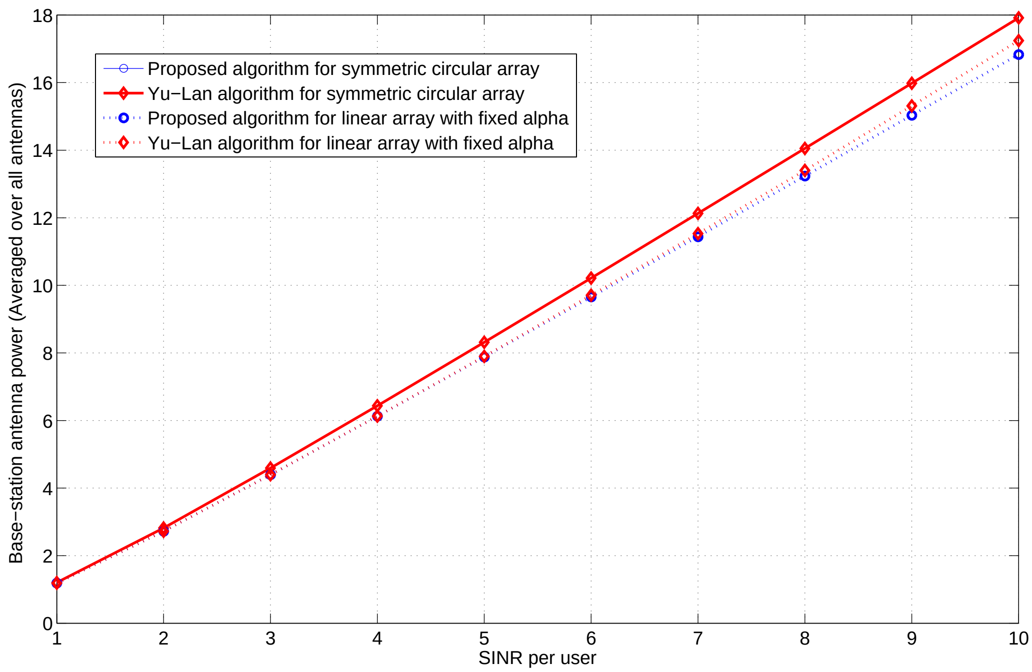

Figure 5.2: Proposed algorithm vs. Yu-Lan algorithm in [2] for symmetric circular array and linear array for N= 10 and $\alpha = 0.3$

In Fig. 5.3 the average power is compared with an asymmetrical Wyner's circular array with $N = 10$ cells where the interference factors α_i^+ and α_i^- are chosen randomly. It is observed that when the SINR constraints are large, the difference between the total transmit power is significant. For example for target spectral efficiency of 3 bits/sec/Hz for each user (corresponding to SINR=7), Yu-Lan algorithm uses approximately 1 dB more transmit power than the proposed algorithm.

The difference between Yu-Lan's and proposed algorithms in terms of optimization criteria is illustrated in Fig. 5.4. In this figure the maximum and minimum of the antenna powers as well as the average antenna power for the two methods are compared for a general circular array with $N = 10$ cells and random interference parameters. The total transmit power is lower in the proposed method especially for high SINR (corresponding to rate) targets since proposed method aims to minimize total transmit power. However, if the variation in the level of power transmitted by each antenna (corresponding to difference between maximum and minimum antenna powers) is considered, it is observed that the variation is less in Yu-Lan's method. This is expected as Yu-Lan's method aims to minimize the power margin, trying to establish a power balance over the antennas.

Proposed Algorithm

N	3	5	7	10	15	20	30	50
Circular Array $\alpha = 0.3$	3	3	3	3	3	3	3	3
Linear Array $\alpha = 0.3$	3	3	3	3	3	3	3	3
Random Circular Array	3	3	3	3	3	3	3	3

Yu and Lan Algorithm

N	3	5	7	10	15	20	30	50
Circular Array $\alpha = 0.3$	3	3	3	3	3	3	3	3
Linear Array $\alpha = 0.3$	29	132	261	392	509	985	614	599
Random Circular Array	31	142	101	108	399	474	576	732

Table 5.1: Number of iterations for convergence for various array scenarios for SINR target 5 for all users

The proposed algorithm is compared with the algorithm in [2] in terms of convergence rate and computational complexity. In Table 5.1, the number of iterations performed for convergence is compared for the proposed and Yu-Lan's algorithms for different array scenarios with SINR=5 for all users. It is observed that the number of iterations required for the symmetric array structure (circular array with $\alpha = 0.3$) are the same for both methods. However, for asymmetric array scenarios (linear array with $\alpha = 0.3$, general circular array with randomly chosen interference parameters) and as the size of the array gets large, number of iterations required by Yu-Lan's method gets significantly larger. In fact, the number of iterations required for the proposed algorithm seems to be independent

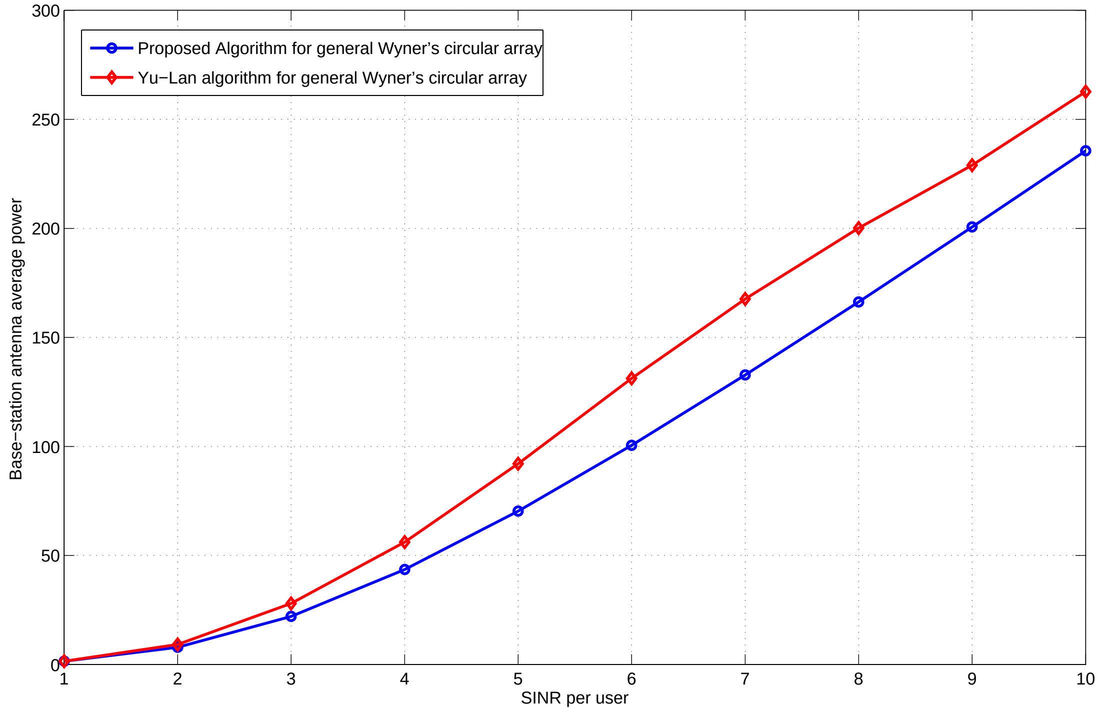

Figure 5.3: Proposed algorithm vs. Yu-Lan algorithm in [2] for asymmetric Wyner's circular array with $N = 10$

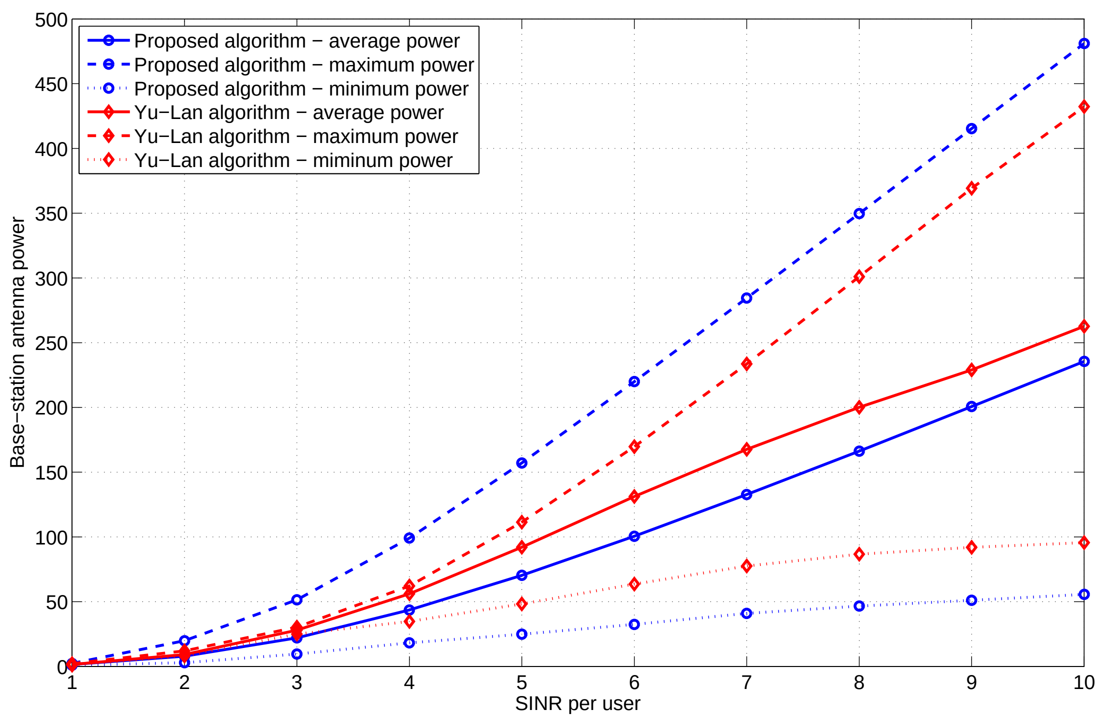

Figure 5.4: Comparison of proposed and Yu-Lan algorithm in terms of maximum, average and minimum BS antenna powers for an asymmetric Wyner's circular array with $N = 10$ and random interference parameters

of the array structure and size. In that sense, we can say that proposed algorithm is more robust and very efficient.

In Fig. 5.5, the computational speed of the two algorithms is compared for the symmetric circular array and asymmetric linear array with $N = 10$ cells and $\alpha = 0.3$. Both algorithms are implemented in Matlab 7.0 version running on an Intel 1.70 GHz Pentium M processor. From the figure, it is observed that elapsed time for Yu-Lan's algorithm is significantly larger than the proposed method especially for aymmetic array with large number of cells. This is due to the fact that the number of iterations for Yu-Lan algorithm increases as the size of the array increases. In addition, subgradient projection is required for Yu-Lan algorithm which is computationally complex as it involves solving an optimization problem, whereas for the proposed algorithm, subgradient projection simplifies into comparison with an all zero vector.

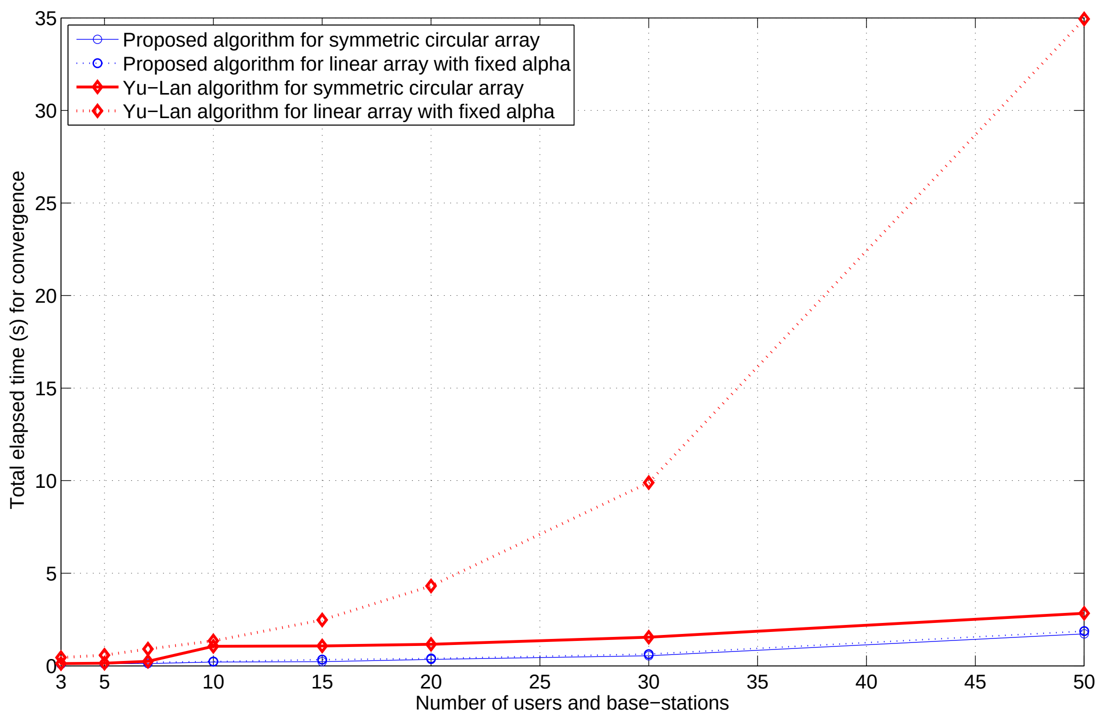

Figure 5.5: Time for convergence of proposed algorithm vs. Yu-Lan algorithm for symmetric circular and linear array with $N = 10$ and $\alpha = 0.3$

The proposed algorithm is compared with the ZF algorithm [12] that maximizes the minimum rate subject to per-antenna power constraints in Fig. 5.6 for

Wyner's symmetric circular array scenario with $N = 5$ cells and different interference factors. For this comparison, firstly we find the minimum rate that ZF algorithm achieves under certain per-antenna constraints. Then,we find the total transmit power to achieve these rates with the proposed algorithm. The sum of per-antenna power constraints for ZF algorithm and the total power returned by the proposed algorithm are finally compared. The aim of the ZF method is to null the interference and as a result it orthogonalize the channel. In this respect, to achieve the given SINR constraints when the interference is low (interference factor α is small), the total power used in ZF algorithm is almost the same with that of the proposed algorithm. When interference is higher, the difference between used total powers for the two algorithms gets bigger. An important point to mention is that the ZF algorithm cannot be used when the channel is rank deficient since the channel cannot be perfectly orthogonalized.

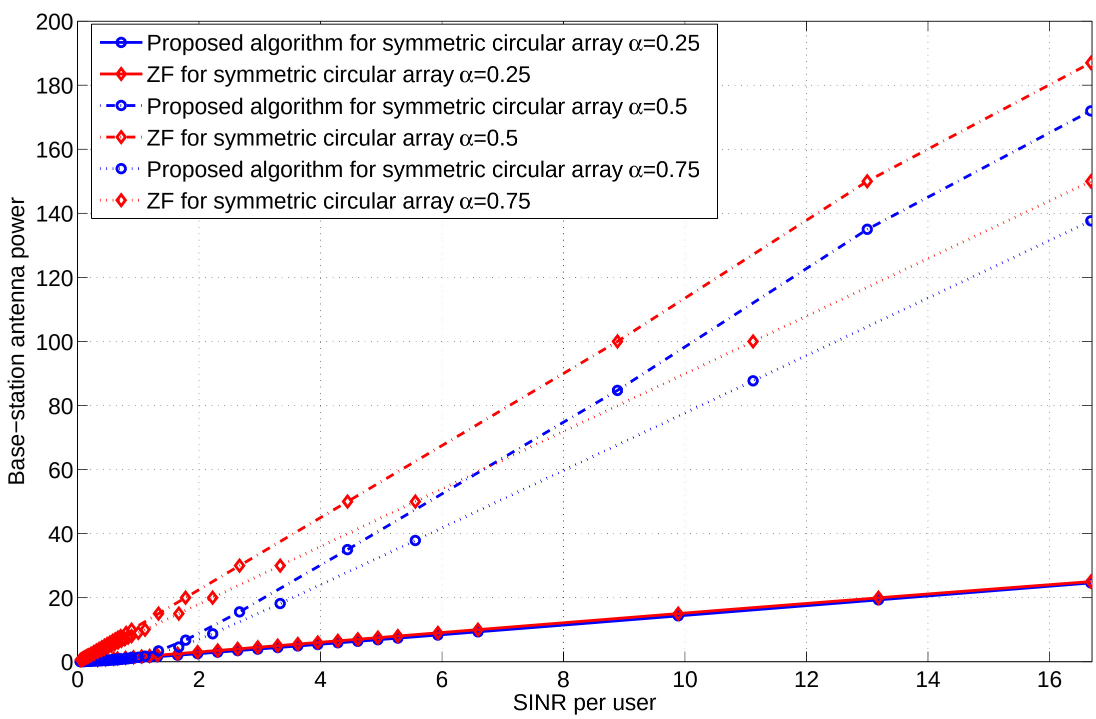

Figure 5.6: Comparison of proposed algorithm and ZF algorithm for a symmetric Wyner's circular array with $N = 5$ and various α values

As mentioned before, the optimal scheme achieving the capacity region of MIMO BC is DPC with beamforming. In Fig. 5.7, the performance of the transmission scheme with beamforming only is compared with a transmission scheme with beamforming and DPC for Wyner's symmetric circular array with $N = 9$ cells and $\alpha = 0.25$. As observed in the figure, the beamforming only scheme uses higher average power than that of DPC employing scheme. The transmission powers of $N = 9$ BS antennas vary since DPC scheme knows the interference of priorly encoded users and encodes the following users accordingly and the lastly encoded user sees no interference, therefore minimal power is used for that user. The variation among the transmission powers for DPC employing scheme increases as SINR increases which is an expected result.

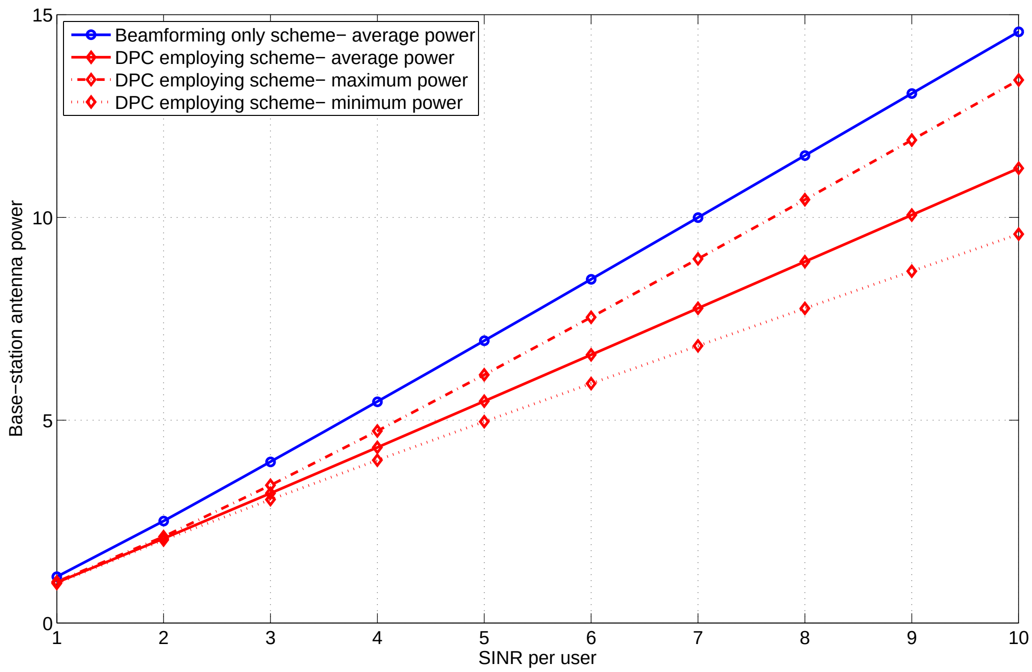

Figure 5.7: Comparison of beamforming only scheme and beamforming with DPC scheme for a symmetric Wyner's circular array with $N = 9$ and $\alpha = 0.25$

5.2 Distributed Algorithm

The scenario we consider is a Wyner's linear array with $N = 3$ and fixed $\alpha = 0.5$ and the target SINR for users is 2. The channel is:

$$\mathbf{H} = \begin{bmatrix} 1 & 0.5 & 0 \\ 0.5 & 1 & 0.5 \\ 0 & 0.5 & 1 \end{bmatrix}.$$

The solution of fixed-point equations are found as: $\boldsymbol{\lambda}^* = [3.7957; 5.4666; 3.7957]^T$. Then the observation channel matrix is computed as:

$$\tilde{\mathbf{H}} = \begin{bmatrix} 1.9483 & 0.9741 & 0 \\ 1.1690 & 2.3381 & 1.1690 \\ 0 & 0.9741 & 1.9483 \end{bmatrix}.$$

$K = 3$ so there will be 3 message-passing cycles/ iterations through the network. After one cycle of message-passing through the whole network we find one of the following three matrices. After 3 cycles, we find all of them. The second rows of the following matrices are the transposes of the uplink beamforming vectors $\hat{\mathbf{w}}_1$, $\hat{\mathbf{w}}_2$ and $\hat{\mathbf{w}}_3$, respectively. They are the Kalman filter estimates defined in [23] for the uplink beamformers at the BSs. The ith column is the Kalman estimate of BS i.

$$\left[\begin{array}{ccc} 0 & 0.1972 & -0.0432 \\ \hline 0.1972 & -0.0432 & -0.0113 \\ \hline -0.0432 & -0.0113 & 0 \end{array}\right].$$

The second iteration returns the following matrix:

$$\left[\begin{array}{ccc} 0 & -0.0224 & 0.1443 \\ \hline -0.0224 & 0.1443 & -0.0224 \\ \hline 0.1443 & -0.0224 & 0 \end{array}\right].$$

The third/ last iteration returns the following matrix:

$$\left[\begin{array}{ccc} 0 & -0.0113 & -0.0432 \\ \hline -0.0113 & -0.0432 & 0.1972 \\ \hline -0.0432 & 0.1972 & 0 \end{array}\right].$$

Since the scenario is a linear array scenario, the (1,1)th and (3,3)th elements of the matrix is 0 meaning that BS 1 and BS 3 do not send any information to BS 3 and BS 1, respectively. At the end of every message-passing stage, the BSs store the data in their corresponding column and in the 2nd row i.e, BS i stores the data in the $(2, i)$th entry of the matrix. Finally BS i forms the uplink beamforming vector $\hat{\mathbf{w}}_i$. The uplink beamforming matrix is as follows:

$$[\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2, \hat{\mathbf{w}}_3] = \left[\begin{array}{c|c|c} 0.1972 & -0.0224 & -0.0113 \\ -0.0432 & 0.1443 & -0.0432 \\ -0.0113 & -0.0224 & 0.1972 \end{array}\right].$$

The optimal $\left[\sqrt{\delta_1}, \sqrt{\delta_2}, \sqrt{\delta_3}\right] = [10.2513; 14.3027; 10.2513]$. But the distributed algorithm finds them as $\left[\sqrt{\delta_1}, \sqrt{\delta_2}, \sqrt{\delta_3}\right]$= [9.8408;14.1093;9.8408]. The optimal downlink beamformers are found as

$$\mathbf{w} = \left[\begin{array}{c|c|c} 2.0217 & -0.3198 & -0.1159 \\ -0.4424 & 2.0640 & -0.4424 \\ -0.1159 & -0.3198 & 2.0217 \end{array}\right].$$

but the downlink beamformers found by the distributed algorithm are

$$\mathbf{w} = \left[\begin{array}{c|c|c} 1.9407 & -0.3154 & -0.1112 \\ -0.4247 & 2.0361 & -0.4247 \\ -0.1112 & -0.3154 & 1.9407 \end{array}\right].$$

The middle user satisfies its SINR requirement, but the other users have a SINR = 1.8690. The duality gap is computed as -0.7955.

The distributed algorithm is a suboptimal algorithm due to the approximation made in $\mathbf{G}$ matrix defined in previous section and duality gap occurs between the original and Lagrangian dual problems. In Fig. 5.8 we compare the duality gap for the proposed suboptimal distributed algorithm for Wyner's linear array with $N = 3$ cells and different interference factors. We observe that, as SINR and the interference factor α values increase, the duality gap gets larger. The duality gap can be used as an indicator of how far the result for the distributed algorithm is from the optimum.

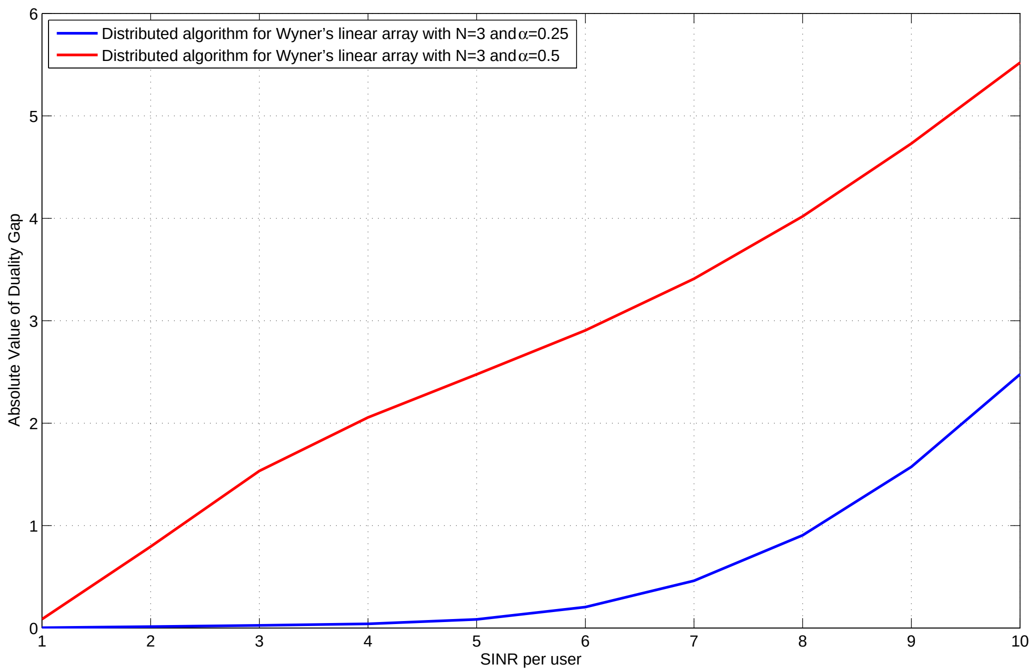

Figure 5.8: The absolute value of the duality gap for distributed algorithm for Wyner's linear array with $N = 3$ and $\alpha = 0.25$ and $\alpha = 0.5$

Chapter 6

CONCLUSIONS

In this thesis, iterative algorithms that compute downlink beamforming vectors minimizing total transmit power under individual SINR and per-antenna power constraints for systems with BS cooperation were investigated. Per-antenna power constraints are logical for systems with BS cooperation, as the antennas of the global transmitter are geographically seperated and each antenna is limited by the linearity region of its power amplifier.

In [2], an elegant iterative algorithm computing beamforming vectors minimizing power-margin under individual SINR and per-antenna power constraints based on Lagrangian dual framework is presented. While this formulation provides an alternative view of well known duality between uplink and downlink from the perspective of Lagrangian dual framework, from a system designer's point of view, it may be more critical to optimize total transmit power than minimizing power margin.

As a result, we reformulated the optimization problem in [2] as minimization of total transmit power and following the Lagrangian dual framework, we proposed an iterative algorithm that computes the optimal beamforming vectors. The performance of the proposed algorithm was compared with other algorithms under per-antenna power constraints. It is observed that, compared with other algorithms in the literature, less transmission power is used in the proposed algorithm to achieve the same set of SINR targets especially when the system is asymmetric. It is also observed that the proposed algorithm is computationally more efficient.

However, for the algorithm to be implemented in practical systems with BS cooperation, it needs to be implementable in a distributed fashion. When system has certain structure like Wyner's linear array, it is shown to be implemented in a suboptimal way with limited information exchange between BSs.

As a future work, truly distributed implementation of the proposed algorithm with better performance will be investigated.

APPENDIX A

Test of SINR Feasibility

In [19] conditions to test feasibility of SINR constraints are derived. For a given SINR value γ_0 and a channel matrix $\mathbf{H} \in \mathbb{R}^{K\text{x}N}$, a precoder/beamformer matrix $\mathbf{T} \in \mathbb{R}^{N\text{x}K}$ exists if:

$$\min_i \frac{\left|[\mathbf{HT}]_{i,i}\right|^2}{\sum_{i\neq j}\left|[\mathbf{HT}]_{i,j}\right|^2 + \sigma^2} \geq \gamma_0. \tag{A.1}$$

Since σ^2 is positive, for simplicity of SINR feasibility analysis one can easliy show that

$$\frac{\left|[\mathbf{HT}]_{i,i}\right|^2}{\sum_{i\neq j}\left|[\mathbf{HT}]_{i,j}\right|^2 + \sigma^2} < \frac{\left|[\mathbf{HT}]_{i,i}\right|^2}{\sum_{i\neq j}\left|[\mathbf{HT}]_{i,j}\right|^2}, \ \forall i. \tag{A.2}$$

A feasible $\mathbf{T}$ satisfying SINR constraints exists if

$$\min_i \frac{\left|[\mathbf{HT}]_{i,i}\right|^2}{\sum_{i\neq j}\left|[\mathbf{HT}]_{i,j}\right|^2} \geq \gamma_0. \tag{A.3}$$

which is shown to be equivalent to

$$\gamma_0 \leq \frac{1}{\frac{K}{\text{rank}(\mathbf{H})} - 1}. \tag{A.4}$$

Note that, $\text{rank}(\mathbf{H}) \leq \min(K, N)$, so when $N < K$, $\text{rank}(\mathbf{H}) \leq N < K$ then there is a SINR limit. When $\text{rank}(\mathbf{H}) = K$, then $\gamma_0 < \infty$ meaning that any SINR is feasible.

APPENDIX B

Lagrangian Dual Formulation

Consider an optimization problem in standard form with variable $\mathbf{x} \in \mathbb{R}^n$:

$$\text{minimize } f_0(\mathbf{x}) \tag{B.1}$$

$$\text{subject to } f_i(\mathbf{x}) \leq 0,\ i = 1, \ldots, m \tag{B.2}$$

$$h_i(\mathbf{x}) = 0,\ i = 1, \ldots, k \tag{B.3}$$

Denote the optimal value of the above problem as $\mathbf{x}^*$. The Lagrangian of the above problem is defined [18] as :

$$L(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{q}) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{k} q_i h_i(\mathbf{x}) \tag{B.4}$$

where λ_i's and q_i's are called the Lagrangian multipliers for the constraints $f_i(\mathbf{x})$ and $h_i(\mathbf{x})$, respectively. They are also called dual variables and satisfy the conditions $\lambda_i \geq 0$ and $\mathbf{q} \in \mathbb{R}^k$.

The Lagrangian dual function is defined as the minimum of Lagrangian function over $\mathbf{x}$ values:

$$g(\boldsymbol{\lambda}, \mathbf{q}) = \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{x}) = \inf_{\mathbf{x}} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{k} q_i h_i(\mathbf{x}) \right) \tag{B.5}$$

The dual Lagrangian is $-\infty$ if the Lagrangian is unbounded below.

Consider an arbitrary feasible point, $\tilde{\mathbf{x}}$ for the basic optimization problem. We have $\sum_{i=1}^{m} \underbrace{\lambda_i}_{\geq 0} \underbrace{f_i(\tilde{\mathbf{x}})}_{\leq 0} + \sum_{i=1}^{k} q_i \underbrace{h_i(\tilde{\mathbf{x}})}_{=0}$, therefore

$L(\tilde{\mathbf{x}}, \boldsymbol{\lambda}, \mathbf{x}) = f_0(\tilde{\mathbf{x}}) + \underbrace{\sum_{i=1}^{m} \lambda_i f_i(\tilde{\mathbf{x}}) + \sum_{i=1}^{k} q_i h_i(\tilde{\mathbf{x}})}_{\leq 0} \leq f_0(\tilde{\mathbf{x}})$. Then,

$g(\boldsymbol{\lambda}, \mathbf{q}) = \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{q}) \leq L(\tilde{\mathbf{x}}, \boldsymbol{\lambda}, \mathbf{q}) \leq f_0(\tilde{\mathbf{x}}), \forall \tilde{\mathbf{x}}$.

Therefore, $g(\boldsymbol{\lambda}, \mathbf{q}) \leq \mathbf{x}^*$ if $\boldsymbol{\lambda} \geq 0$.

When $g(\boldsymbol{\lambda}, \mathbf{q}) = -\infty$, the dual problem does not give a meaningful lower bound on the optimal value. Therefore, the following dual problem is stated choosing $\boldsymbol{\lambda}$ and $\mathbf{q}$ such that Lagrangian dual function is finite.

$$\text{maximize } g(\boldsymbol{\lambda}) \tag{B.6}$$

$$\text{subject to } \lambda_i \geq 0 \tag{B.7}$$

Let's illustrate this with an example. For the following optimization problem :

$$\text{minimize } \mathbf{c}^T \mathbf{x} \tag{B.8}$$

$$\text{subject to } \mathbf{a}_i^T \mathbf{x} \leq b_i, \; i = 1, \ldots, m \tag{B.9}$$

$$\mathbf{d}_i^T \mathbf{x} = 0, \; i = 1, \ldots, k \tag{B.10}$$

the Lagrangian is $L(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{q}) = \mathbf{c}^T \mathbf{x} + \sum_{i=1}^{m} \lambda_i \left(\mathbf{a}_i^T \mathbf{x} - b_i\right) + \sum_{i=1}^{k} q_i \mathbf{d}_i^T \mathbf{x}$. Therefore, Lagrangian dual function is stated as

$$\begin{aligned} g(\boldsymbol{\lambda}, \mathbf{q}) &= \min_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{q}) \\ &= \min_{\mathbf{x}} \left(-\sum_{i=1}^{m} \lambda_i b_i + \left(\mathbf{c} + \sum_{i=1}^{m} \lambda_i \mathbf{a}_i + \sum_{i=1}^{k} q_i \mathbf{d}_i\right)^T \mathbf{x} \right) \\ &= \begin{cases} -\sum_{i=1}^{m} \lambda_i b_i & \text{, if } \left(\mathbf{c} + \sum_{i=1}^{m} \lambda_i \mathbf{a}_i + \sum_{i=1}^{k} q_i \mathbf{d}_i\right) = \mathbf{0} \\ -\infty & \text{, otherwise.} \end{cases} \end{aligned} \tag{B.11}$$

Then, the Lagrangian dual problem is defined as follows:

$$\text{maximize } -\sum_{i=1}^{m} \lambda_i b_i \tag{B.12}$$

$$\text{subject to } \mathbf{c} + \sum_{i=1}^{m} \lambda_i \mathbf{a}_i + \sum_{i=1}^{k} q_i \mathbf{d}_i = \mathbf{0}, \tag{B.13}$$

$$\boldsymbol{\lambda} \succeq 0. \tag{B.14}$$

Duality is classified into two categories: weak duality and strong duality. The optimal value, d^*, of the Lagrangian dual problem is the best lower bound on the optimal value, p^*, of the original (primal) optimization problem (B.1). That is stated as by following inequality

$$d^* \leq p^*. \tag{B.15}$$

This equality is valid even if the primal problem is not convex. This situation is called weak duality [18].

The difference $p^* - d^*$ is called optimal duality gap and is a measure for the difference between optimal value of original problem and the optimal value for the Lagrangian dual function. Weak duality is sometimes used to find a lower bound for difficult-to-solve optimization problems.

If the above inequality is satisfied with equality, i.e.,

$$d^* = p^*, \tag{B.16}$$

then the duality gap is 0 and it is stated that 'strong duality' holds (the best lower bound is obtained). Strong duality holds for optimization problems in some certain conditions. Slater's conditions are used for test of strong duality [18]. When the primal problem is convex, and Slater's conditions holds for this problem, then strong duality holds:

$$g(\boldsymbol{\lambda}^*, \mathbf{q}^*) = d^* = p^*. \tag{B.17}$$

APPENDIX C

Subgradient Projection

First, we define what projection and subgradient are. For C, a closed convex subset of $\mathbb{R}^n$ and $\mathbf{x} \in \mathbb{R}^n$, projection of $\mathbf{x}$ onto the set C is defined as:

$$P_C(\mathbf{x}) = \arg\min_{\mathbf{z} \in C} \|\mathbf{z} - \mathbf{x}\|_2 \tag{C.1}$$

where z is the unique vector that minimizes $\|\mathbf{z} - \mathbf{x}\|_2$.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function, then $\mathbf{d} \in \mathbb{R}^n$ is a subgradient of f at $\mathbf{x} \in \mathbb{R}^n$ if

$$f(\mathbf{z}) \geq f(\mathbf{x}) + (\mathbf{z} - \mathbf{x})^T \mathbf{d}, \ \forall \mathbf{z} \in \mathbb{R}^n \tag{C.2}$$

holds.

Subgradient projection method solves the following convex optimization problem

$$\text{minimize } f(\mathbf{x}) \tag{C.3}$$

$$\text{subject to } \mathbf{x} \in C \tag{C.4}$$

where C is a convex set. The subgradient projection method is given by

$$\mathbf{x}^{(k+1)} = P_C(\mathbf{x}^{(k)} - t_k \mathbf{d}^{(k)}) \tag{C.5}$$

where $\mathbf{d}^{(k)}$ is a subgradient of f at $\mathbf{x}^{(k)}$ and t_k is the step size.

The subgradient projection method is guaranteed to converge since when a point is projected onto the set C, we get closer to any optimal point in C [25].

Various types of step size rules can be used in the subgradient projection method.

- **Constant step size**: $t_k = c$ is a constant (hence independent of k).
- **Constant step length**: $t_k = c/\left\|\mathbf{d}^{(k)}\right\|_2$ meaning that $\left\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\right\|_2 = c$.
- **Square summable but not summable**: Step size satisfies $\sum_{i=1}^{\infty} t_k^2 < \infty$ and $\sum_{i=1}^{\infty} t_k = \infty$, e.g. $t_k = \frac{1}{k}$.
- **Nonsummable diminishing**: Step size satisfies $\lim_{k\to\infty} t_k = 0$ and $\sum_{k=1}^{\infty} t_k = \infty$. $t_k = p/\sqrt{k}$,where $p > 0$ is an example for this type.

Bibliography

[1] M. Schubert and H. Boche, "Iterative multiuser uplink and downlink beamforming under SINR constraints," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2324–2334, July 2005.

[2] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2646–2660, June 2007.

[3] A. Paulraj and C. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Magazine*, vol. 14, no. 6, pp. 49–83, Nov. 1997.

[4] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov. 1999.

[5] F. Rashid-Farrokhi, K. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1437–1450, Oct. 1998.

[6] P. Viswanath and D. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

[7] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[8] M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[9] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.

[10] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[11] N. Jindal, W. Rhee, S. Vishwanath, S. Jafar, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1570–1580, Apr. 2005.

[12] G. Foschini, K. Karakayali, and R. Valenzuela, "Coordinating multiple antenna cellular networks to achieve enormous spectral efficiency," *IEE Proceedings on Communications*, vol. 153, p. 548, Aug. 2006.

[13] J. He and M. Salehi, "Autonomic coordinated beamforming for multiuser MIMO-BC," *Annual Conference on Information Sciences and Systems*, pp. 129–133, Mar. 2006.

[14] A. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-accesschannel," *IEEE Transactions on Information Theory*, vol. 40, no. 6, pp. 1713–1727, Nov. 1994.

[15] W. Yu and J. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 1875–1892, Sept. 2004.

[16] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936–3964, Sept. 2006.

[17] J. Löfberg, "Yalmip : A toolbox for modeling and optimization in matlab," *Proceedings of the CACSD Conference*, Sept. 2004.

[18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.

[19] A. Wiesel, Y. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Transactions on Signal Processing*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

[20] N. Shor, K. Kiwiel, and A. Ruszcaynski, *Minimization Methods for Non-differentiable Functions*. Springer-Verlag , NY, USA, 1985.

[21] B. Ng, J. Evans, S. Hanly, and D. Aktas, "Transmit beamforming with cooperative base stations," *Proceedings of International Symposium on Information Theory*, pp. 1431–1435, Sept. 2005.

[22] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Upper Saddle River, NJ, USA, 1993.

[23] B. Ng, J. Evans, and S. Hanly, "Distributed linear multiuser detection in cellular networks based on Kalman smoothing," *IEEE Global Telecommunications Conference*, vol. 1, pp. 134–138, Dec. 2004.

[24] S. Conte and C. De Boor, *Elementary Numerical Analysis*. McGraw-Hill , NY, USA, 1972.

[25] S. Boyd, L. Xiao, and A. Mutapcic, "Subgradient Methods," *Stanford University Course Notes for EE392o*, Oct. 2003.