LARGE DIMENSIONAL ANALYSIS AND OPTIMIZATION FOR MASSIVE MIMO WIRELESS NETWORKS

by

LONGWEI WANG

Presented to the Faculty of the Graduate School of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON
December 2017

ACKNOWLEDGEMENTS

I would like to thank my supervising professor Dr. Qilian Liang for constantly motivating and encouraging me, and also for his invaluable advice during the course of my doctoral studies. I wish to thank my academic advisors Dr. Jonathan Bredow, Dr. Wan, Dr. Schizas, and Dr. Saibun Tjuatja for their interests in my research and for taking time to serve in my dissertation committee. I also thank my group members from the Wireless Communications and Networking (WCN) Laboratory. I learned a lot from this group about life, research, how to tackle new problems and how to develop techniques to solve them. Meanwhile, this work was supported in part by U.S. Office of Naval Research (ONR) under Grant N00014-13-1-0043 and N00014-17-1-2733. Last but not least, I would like to thank my family who have been extremely understanding and supportive of my studies.

Nov 15, 2017

ABSTRACT

LARGE DIMENSIONAL ANALYSIS AND OPTIMIZATION FOR MASSIVE MIMO WIRELESS NETWORKS

Longwei Wang, Ph.D.

The University of Texas at Arlington, 2017

Supervising Professor: Qilian Liang

The Last decade has seen a massive growth of wireless devices. Demands for high capacity and massive connectivity always increases. To meet the these requirements, massive multiple-input multiple-output (MIMO) technology is proposed for the next generation wireless systems. With massive antenna arrays at the BS, the channel vectors between the users and the BS are asymptotically orthogonal, and the intercell interference can be eliminated with simple linear processing method. The effects of fast fading, and uncorrelated noise tend to disappear as the number of BS station antennas grows large.

In this thesis, we mainly focus on the performance analysis and system optimization of massive MIMO cellular networks. We compare two cooperative multicell precoding methods, namely, centralized precoding with global CSI and distributed precoding with local individual CSI in the large dimensional regime. Two different massive MIMO scenarios are considered. When the number of antennas and that of users go large at the same rate, there is a constant gap between the two precoding cases. When the number of antennas goes large, while the number of users is

fixed, the performances of both schemes are the same. This means that the impact of local individual CSI vanishes. In the large dimension limit, certain terms such as signal-to-interference-plus-noise ratio (SINR) can be approximated depending only on the statistical CSI. Based on this result, the power optimization problem does not need to adapt as frequently as the instantaneous channel state information. By this means, performance evaluation and optimization is more computationally efficient. Partial interference alignment in finite SNR is proposed, which captures the trade-off between interference avoidance at other users and spatial multiplexing at the intended user. We try to project the interference into partial subspace of the small cell user instead of the full subspace, so that extra improvement on the rate of macro cell can be offered in finite SNR. Large dimensional analysis is provided to obtain the the number of transmit dimensions for small cell in finite SNR. We show that the number of dimensions the small cell can transmit is determined by the singular value distribution of channel matrix and the transmit SNR of the system.

In heterogeneous sensor networks, the sensing data may not be compatible with each other due to heterogeneous sensing modalities. We propose a probabilistic inference framework for fusing information from heterogeneous sensors. The inherent inter-sensor relationship is exploited to encode the original sensor data in a graph. The iterative belief propagation is used to fuse the local sensing belief. Then we consider the more general correlation case, in which the relation between two sensors is characterized by the correlation factor. The belief propagation provides intuitive insights as to how the local probabilistic update helps to reinforce beliefs when performing information fusion.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Motivations and Background

With the prosperity of wireless devices and applications today, the demand for higher capacity, low latency and massive connectivity in wireless networks has been increasing dramatically. Fig.1.1 shows the demand for mobile data traffic and the number of connected devices. Global mobile data traffic is expected to increase to 16 exabytes per month by 2018. The most important performance metric to be considered is the wireless throughput, which is determined by the bandwidth and the spectral efficiency.

To meet the required performance, we can either increase the bandwidth or improve the spectral efficiency. Lots of new technologies have been proposed for the promotion of the next generation mobile networks. In this thesis, we mainly focus on the improvement of spectral efficiency by multiple antennas systems.

In wireless scenarios, the channel fading which includes fast fading and slow fading, poses a fundamental challenge for reliable communication. The multiple an-

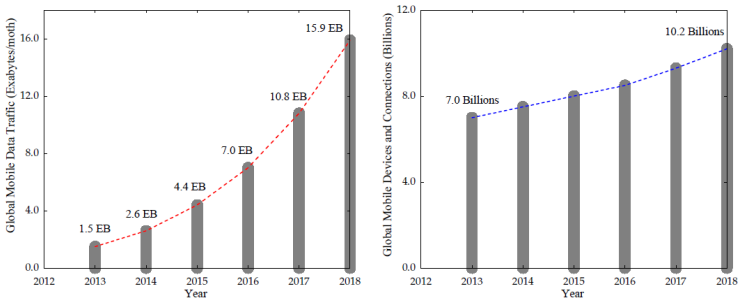

Figure 1.1. The growth of data traffic and wireless devices .

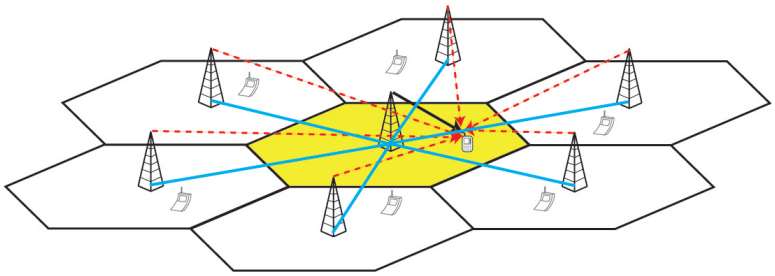

Figure 1.2. Multicell cellular networks and intercell interference .

tennas communication, also known as the multiple-input multiple-output (MIMO) technique can efficiently enhance the reliability of wireless communication by providing diversity gains. Furthermore, multiplexing gain can be obtained by sending multiple data streams simultaneously in multiple antennas communication. It will significantly improves the capacity of the system. MIMO systems have gained significant attention for the past decades, and large scale MIMO system is a emerging technique for next generation wireless standards.

In traditional cellular networks, inter-cell interference remains the main factor limiting the achievable performance. Network MIMO [51] [52] has been proposed as an effective way to alleviate the mutual interference between adjacent cells. However, the full potential of BS cooperation can be achieved at the expense of global CSI and data sharing among all the BSs [53] [54]. In this case, the BSs can cooperate as a single super BS serving all the users. The resulting transmission works in the same scenario as the MISO broadcast channel. It can completely eliminate the intercell interference. Centralized CSI is usually assumed in most related works in which the global CSI and all users data are collected in a central unit and the precoding is based on the globally shared data and CSI. However, due to imperfect backhaul and limited feedback, the estimated CSI for each BS is noisy and delayed. The CSI imperfection poses a severe negative impact on wireless systems, which will degrade

the performance of BS cooperation. [55] considers the case that precoding is performed in each BS with global CSI, but due to heterogeneous and imperfect backhaul among the BSs, the CSI will introduce quantization noise and delay, which makes the CSI imperfections in each BS are independent. They refer to this scenario as cooperative precoding with distributed CSI. The capacity bound of MISO broadcast channel can be achieved by dirty paper coding, however the capacity bound for broadcasting with distributed CSI is still unsolved. [58] characterizes the DoF region of two cooperating BSs. They show that there is a penalty with the CSI inconsistence from the high SNR point of view.

1.2 Massive MIMO and Mutiuser MIMO Cellular Networks

Massive MIMO, or large scale MIMO is a recently proposed technology. It can efficiently improve the spectral efficiency with fairly simple precoding and signal processing. Massive MIMO is used to serve many users simultaneously and the number of antennas deployed in the BSs is large. In Massive MIMO, we usually assume hundreds of antennas in the BSs serve tens or hundreds of users over the same band resource.

Massive MIMO has attracted lots of attention in the research community [12] [64]. There are many advantages in deploying large scale antennas in cellular networks. First, it can increase the network performance by simply deploying more antennas in the existing base stations. Second, working in the time division duplex (TDD) mode, the channel training cost is proportional to the number of users, instead of the number of BS antennas. Thirdly, when the number of antennas in the BSs is far more than the number of users served, the simple linear precoder and detectors, such as matched filter or maximum ratio transmission, are optimal. The intercell interference

can be averaged out, and the effect of channels estimation errors and noise will be canceled.

1.2.1 Mutiuser MIMO Cellular Networks and Basic Precoding Methods

Consider a single cellular network which consists of one BS and K active users. The BS is equipped with N antennas, while each user has a single-antenna [1]. The transmitted signal of BS, denoted as $\mathbf{x}$, is given by the following expression

$$\mathbf{x} = \sum_{k=1}^{K} \mathbf{w}_k s_k \tag{1.1}$$

$$= \alpha \mathbf{W}\mathbf{s} \tag{1.2}$$

where $\mathbf{w}_k$ denotes the precoding vector designed for user k, $\mathbf{s} = [s_1\ s_2\ \cdots\ s_K]^T$ denotes the transmitted signal vector of dimension K and α is the normalization term, which can be given as

$$\alpha = \frac{1}{\mathrm{tr}(\mathbf{W}\mathbf{W}^H)} \tag{1.3}$$

then the received signal of user k can be expressed as

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k \tag{1.4}$$

where n_k is the additive complex Gaussian noise with distribution $\mathcal{CN}(0, 1)$ and $\mathbf{h}_{mk}^H$ denotes the channel from BS to the user k.

By stacking the received signal together, we can obtain

$$\mathbf{y} = \mathbf{H}^H \mathbf{x} + \mathbf{n} \tag{1.5}$$

$$= \alpha \mathbf{H}^H \mathbf{W}\mathbf{s} + \mathbf{n} \tag{1.6}$$

where $\mathbf{y} = [y_1\ y_2\ \cdots\ y_K]^T$, $\mathbf{n} = [n_1\ n_2\ \cdots\ n_K]^T$, and $\mathbf{H} = [\mathbf{h}_1\ \mathbf{h}_2\ \cdots\ \mathbf{h}_K]$. For a single user, the received signal plus interference and noise form can be given as

$$y_k = \mathbf{h}_k^H \sum_{k=1}^{K} \mathbf{w}_k s_k + n_k \tag{1.7}$$

$$= \mathbf{h}_k^H \mathbf{w}_k s_k + \mathbf{h}_k^H \sum_{k' \neq k}^{K} \mathbf{w}_{k'} s_{k'} + n_k \tag{1.8}$$

Denote the signal to interference plus noise ratio (SINR) for user k as γ_k. It can be expressed as

$$\gamma_k = \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum\limits_{k' \neq k} |\mathbf{h}_{k'}^H \mathbf{w}_{k'}|^2 + 1} \tag{1.9}$$

where $\mathbf{w}_{k'}$ is the precoding vector from BS to user j. Then the corresponding achievable rate for user k is $\log(1 + \gamma_k)$.

Three basic linear precoders are maximum-ratio transmission (MRT) (also called conjugate beforming), ZF, and MMSE precoders [1]. These precoders are expressed as

$$\mathbf{W} = \begin{cases} \mathbf{H}, & \text{for MRT} \quad (1.10) \\ \mathbf{H}(\mathbf{H}^H \mathbf{H})^{-1}, & \text{for ZF} \quad (1.11) \\ \mathbf{H}(\mathbf{H}^H \mathbf{H} + \gamma \mathbf{I})^{-1}, & \text{for MMSE} \quad (1.12) \end{cases}$$

1.2.2 Challenges in Massive MIMO

Due to the limited coherence time of wireless channels, the number of orthogonal pilots available are limited, which will lead to the reuse of pilot in the channel training stage. This will introduce pilot interference, also called the pilot contamination. In this paper, we do not consider the pilot interference problem in massive MIMO as usual. Recent works [59] [60] show that the pilot reuse problem can be solved by using subspace channel estimation approach. By coordinating the pilot in multiple cells, the pilot contamination can be reduced efficiently [61]. Further, pilot

contamination precoding (PCP) can effectively remove the pilot contamination to achieve interference free cellular network with massive MIMO [62].

1.3 Why is Large Dimensions Useful?

1.3.1 Asymptotic Orthogonality and Favorable Propagation Property

The channel vectors between the BS and users are assumed to be flat fading, which accounts for small scale fading and large scale fading coefficients. The channel from BS to users can be expressed as

$$\mathbf{H} = \mathbf{G}\mathbf{D}^{\frac{1}{2}}, \tag{1.13}$$

where $\mathbf{G}$ is a $K \times N$ matirx corresponding to the fast fading whose elements are complex Gaussian variables. The k-th column-vector of $\mathbf{G}$ denotes the small scale fading between the k-th terminal and the BS. $\mathbf{D}$ is a $K \times K$ diagonal matrix, which is the slow-fading terms. The k-th diagonal entry of $\mathbf{D}$ denotes the large scale fading between the k-th terminal and the BS.

We will analyze the performance of multiuser cellular network under the massive MIMO scenario. The scenario that the number of antennas at the BS N is much larger than the number of users K, i.e., $N \gg K$ and $N \to \infty$, is considered. Under the most favorable propagation conditions, the column-vectors of the propagation matrix are asymptotically orthogonal [2],

$$(\frac{\mathbf{H}^H\mathbf{H}}{N})_{N\gg K} = \mathbf{D}^{\frac{1}{2}}(\frac{\mathbf{G}^H\mathbf{G}}{N})_{N\gg K}\mathbf{D}^{\frac{1}{2}}, \tag{1.14}$$

$$\approx \mathbf{D} \tag{1.15}$$

where the effect of small-scale Rayleigh fading vanishes as N grows.

The BS usually has CSI corresponding to all users based on uplink pilot transmission. Therefore, it is possible for the BS to perform power allocation to maximize the sum transmission rate. The sum capacity for the system is given as

$$C = \max_{\mathbf{P}} \log\det(\mathbf{I} + \alpha \mathbf{H}^H \mathbf{P} \mathbf{H}) \tag{1.16}$$

$$\approx \max_{\mathbf{P}} \log\det(\mathbf{I} + \alpha N \mathbf{P} \mathbf{D}) \tag{1.17}$$

where $\mathbf{P}$ is $K \times K$ diagonal matrix denoting the power allocation for the K users.

If the MRT precoder is used for the multiuser MIMO cellular network, the transmitted signal vector is

$$\mathbf{x} = \alpha \mathbf{G} \mathbf{D}^{\frac{1}{2}} \mathbf{P}^{\frac{1}{2}} \mathbf{s} \tag{1.18}$$

Then, the received signal vector at all users is

$$\mathbf{y} = \alpha \mathbf{H}^H \mathbf{G} \mathbf{D}^{\frac{1}{2}} \mathbf{P} \mathbf{s} + \mathbf{n} \tag{1.19}$$

$$= \alpha \mathbf{D}^{\frac{1}{2}} \mathbf{G}^H \mathbf{G} \mathbf{D}^{\frac{1}{2}} \mathbf{P} \mathbf{s} + \mathbf{n} \tag{1.20}$$

$$\approx \alpha N \mathbf{P} \mathbf{D} \mathbf{s} + \mathbf{n} \tag{1.21}$$

The sum capacity for the multiuser cellular network employing the MRT precoding is

$$C_{sum} = K \log(1 + \alpha N p_k) \tag{1.22}$$

since $\mathbf{P}$ and $\mathbf{D}$ are both diagonal matrices, the MRT precoding separates the signals from different users into different streams without inter-user interference. Inter-user interference has been suppressed asymptotically in the large array scenario. K is the multiplexing gain, and N is the array gain. Huge spectral efficiency can be obtained by scaling the number of users and the number of antennas deployed in the BS.

1.3.2 Large Dimensional Analysis

Besides spectral efficiency gain, another inspiring property of massive MIMO is that it enables the large dimensional analysis of system performance. The system performance may be concealed in the finite case due to random fluctuations. In the large dimension limit where both the number of antennas in the base station and the number of users grow to infinity with a constant ratio, the randomness in channels from BS to users disappears. Certain property can be revealed in the asymptotic limit. When the number of antennas and that of users go to infinity with a fixed ratio, different performance metrics such as achievable rates or bit error rates become possible to obtain based on the large random matrix theory. This is due to the fact that the eigenvalue or singular value distributions of large random matrices often converge to a fixed asymptotic distribution. The large dimensional analysis can provide deterministic results in the massive MIMO scenario which can facilitate the optimization and analysis of system performance.

In order to understand usefulness of the large dimensional analysis, a simple and motivating example is provided. Let us consider the point to point MIMO channel model [68], which can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{1.23}$$

where $\mathbf{y} = [y_1 \ y_2 \ \cdots \ y_N]^T$, $\mathbf{n} = [n_1 \ n_2 \ \cdots \ n_N]^T$ is complex Gaussian noise, and $\mathbf{H} = [\mathbf{h}_1 \ \mathbf{h}_2 \ \cdots \ \mathbf{h}_M]$ is the channel matrix with entry distribution $\mathcal{CN}(0, \frac{1}{N})$.

Assuming that the channel matrix H is perfectly known at the receiver but unknown at the transmitter, then the capacity for this channel can be given as

$$C = E\{\log\det(\mathbf{I} + \rho\ \mathbf{H}\mathbf{H}^H)\} \tag{1.24}$$

$$= E\{\sum_{i=1}^{N}\log(1+\rho\ \lambda_i(\mathbf{H}\mathbf{H}^H))\} \tag{1.25}$$

$$= \int_0^\infty \log(1+\rho\ x) d\mathrm{F}_{\mathbf{H}\mathbf{H}^H}(x) \tag{1.26}$$

when the system goes to large dimensions, the summation can be written in the integral form. ρ is the SNR of the channel and $\mathrm{F}_{\mathbf{H}\mathbf{H}^H}(x)$ denotes the empirical cumulative distribution function of the eigenvalues $\{\lambda_i(\mathbf{H}\mathbf{H}^H)\}$ of matrix $\mathbf{H}\mathbf{H}^H$, which is defined as

$$\mathrm{F}_{\mathbf{H}\mathbf{H}^H}(x) = \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}(\lambda_i(\mathbf{H}\mathbf{H}^H) \le x) \tag{1.27}$$

where $\mathbf{1}$ is the indicator function. The empirical eigenvalue distribution $F_{\mathbf{H}^H\mathbf{H}}(\lambda)$ converges almost to the deterministic limiting eigenvalue distribution $F_{\mathbf{H}^H\mathbf{H}}(\lambda)$. And $f_{\mathbf{H}^H\mathbf{H}}(\lambda)$ is the corresponding probability density function, which can be given according to the Marcenko-Pastur law [96],

$$f_{\mathbf{H}^H\mathbf{H}}(\lambda) = (1-\frac{M}{N})^+\delta(\lambda) + \frac{\sqrt{(\lambda-a)^+(b-\lambda)^+}}{2\pi\lambda} \tag{1.28}$$

where $a = (1-(1/\sqrt{\frac{N}{M}}))^2$ and $b = (1+(1/\sqrt{\frac{N}{M}}))^2$. Note that the density function has no-zero values in $\{\{0\}\bigcup\{[a,b]\}\}$.

Fig. 1.3 shows the Marcenko-Pastur distribution for $N/M = 2$ and the example of empirical distribution for $N = 30, 60, 120, 600$. It is shown that as N grows, the randomness of the eigenvalue distributions gets smaller and smaller and the empirical distributions converge to the theoretical Marcenko-Pastur density function, which can be exploited for the analytical evaluation in the large dimension regime.

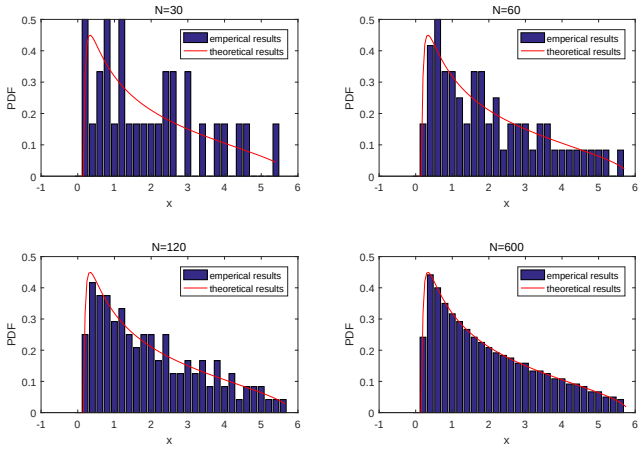

Figure 1.3. Empirical PDF when the dimensions of the channel are increased. .

1.4 Summary of Specific Contributions

In this thesis, we mainly focus on the performance analysis and system optimization of massive MIMO cellular networks. Large dimensional analysis provides an efficient way to quantify the network performance. We compare two cooperative multicell precoding methods, namely, centralized precoding with global CSI and distributed precoding with local individual CSI in the large dimensional regime. Two different massive MIMO scenarios are considered. When the number of antennas and that of users go large at the same rate, there is a constant gap between the two precoding cases. When the number of antennas goes large, while the number of users is fixed, the performances of both schemes are the same. This means that the impact of local individual CSI vanishes.

In the large dimension limit where both the number of antennas per base station and the number of users per cell go to infinity with a constant ratio, certain terms such as signal-to-interference-plus-noise ratio (SINR) can be approximated depending only on the statistical CSI. The randomness in the users' SINR disappears. Based on

this result, the power optimization problem does not need to adapt as frequently as the instantaneous channel state information. By this means, performance evaluation and optimization is more computationally efficient.

Interference alignment provides an efficient way to achieve optimal Degree-of-Freedom (DoF) in interference channels. However, it is not necessarily optimal in terms of achievable sum rate in the finite SNR regime, since the spatial diversity in the receiver sides is ignored. We propose an interference coordination strategy called partial interference alignment in small cell networks. The partial interference alignment adaptively select the dimensions of subspace the small cell user can transmit based on the power constraints. This partial interference alignment captures the trade-off between interference avoidance at other users and spatial multiplexing at the intended user. We try to project the interference into partial subspace of the small cell user instead of the full subspace, so that extra improvement on the rate of macro cell can be offered in finite SNR. We perform large dimensional analysis to obtain the the number of transmit dimensions for small cell in finite SNR. We show that the number of dimensions the small cell can transmit is determined by the singular value distribution of channel matrix and the transmit SNR of the system.

CHAPTER 2

Large Dimensional Analysis of Cooperative Multicell Precoding with Local Individual CSI

In this Chapter, we try to compare the average rate performance achieved by precoding with global CSI and that with local individual CSI in cooperative multicell transmission.

The following two different massive MIMO scenarios are considered:

1. The number of antennas in each BS N and the number of users K both go large at the same rate, $N \to \infty, K \to \infty, N/K = \beta$.
2. The number of antennas in each BS N goes large, while the number of users K is fixed, $N/K \to \infty$, and K is fixed.

Specifically, the main contributions of this work can be described as follows: The performances of both precoding methods are quantified in the large dimensional regime. The performance loss caused by limited CSI can be revealed. We will show that average rate performance for the two precoding schemes converges to constant values when the number of antennas at BSs N and the number of users K both go large with fixed ratio $N \to \infty, N/K = \beta$. The performance gap can be referred to as the price of local individual CSI. When the number of antennas is much larger than that of the users $N/K \to \infty$ and K is fixed, the average rate performances for both precoding schemes are the same. The impact of local individual CSI vanishes.

Asymptotic results are developed in the large dimensional regime. Due to the large antennas deployed in the BSs, the channel dimensions become large which facilitate the asymptotic analysis based on the results of large RMT [68] [69]. Since the

DoF characterize the performance of wireless system in the high SNR, in the low and medium SNR, the performance of cooperative multicell can not be easily revealed. However, closed form results can be derived in the large dimensional regime. Some insights can be obtained for the system design and optimization.

The remainder of this Chapter is organized as follows. The cooperative multicell transmission models with global and local individual CSI are provided in section 4.1. We derive the large dimensional approximations for the SINR of both precoding methods in the fixed ratio scenario in section 2.2. The performance analysis of second scenario: $N/K \to \infty$ with K fixed, is presented in section 2.3. We provide the numerical results and related discussions in section 4.3. Section 4.4 finally concludes the paper.

Throughout this thesis, boldface uppercase and lowercase letters are used to denote matrices and column vectors, respectively. We use $\mathbf{I}$ as the identity matrix. We denote $\mathrm{diag}(x_1 \cdots x_k)$ as the diagonal matrix of size k with the entries x_i in the diagonal. For a matrix $\mathbf{X}$, $\mathbf{X}^H$, $\mathrm{tr}(\mathbf{X})$, $\mathbf{X}^T$ denote the conjugate transpose, the trace, and transpose of matrix $\mathbf{X}$. The notation $\longrightarrow$ denotes the asymptotic convergence. The notation $\overset{a.s.}{\to}$ denotes that the term converges almost surely.

2.1 System Model and Problem Description

2.1.1 Cooperative Downlink System Model

We consider that there are M BSs cooperatively serving K user equipments in a multicell scenario. It is assumed that each BS is deployed with N antennas, where N is a relatively large due to the large scale MIMO scenario. And each user is equiped with only a single antenna. The users are clustered together. The precoding with global CSI is depicted in Fig. 4.2, in which all the locally estimated CSI in each BS is

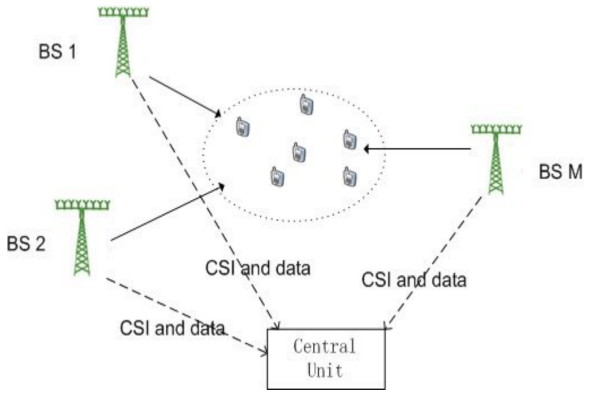

Figure 2.1. Cooperative downlink precoding with global CSI. .

sent to a central control unit. Then the precoding is performed based on the globally shared CSI. The M BSs: $\mathbf{BS}_1, \ldots, \mathbf{BS}_M$ serve all the users simultaneously. It requires synchronization among all the BSs. In Fig. 2.2, no information is exchanged among the BSs. The precoding is based only on the local individual CSI in each BS.

The channel vectors between the BSs and users are assumed to be flat fading, which contains fast-fading and slow-fading factors. The channel from BS i to the kth user can be expressed as [64]

$$\mathbf{h}_{ik} = \mathbf{D}_{ik}^{\frac{1}{2}} \mathbf{g}_{ik}, \tag{2.1}$$

where $\mathbf{g}_{ik}$ is the random vector corresponding to the fast fading whose elements are complex Gaussian variables, and $\mathbf{D}_{ik}$ is the slow-fading terms. In this paper, we assume the matrix $\mathbf{D}_{ik} = \beta_{ik}\mathbf{I}$. The diagonal elements are nonnegative scalar that is related with the distance between BS i and the kth user and the shadow fading. There is no correlation between the antenna elements in the BSs.

2.1.2 Cooperative Multicell Precoding with Global CSI

The multicell cooperative downlink transmission with global CSI is depicted in Fig. 4.2. The global CSI is sent to the central control unit via backhaul links.

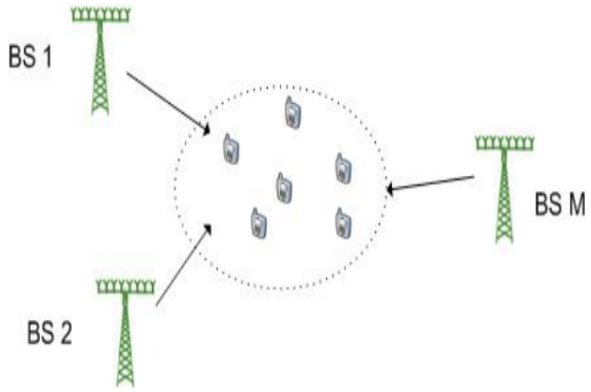

Figure 2.2. Cooperative downlink precoding with local individual CSI. .

Then the precoding is based on the shared CSI. The signals to be transmitted can be expressed as

$$\mathbf{s} = [s_1, \cdots, s_k, \cdots, s_K] \tag{2.2}$$

where s_k is the symbol sent to user k.

Let $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_K]^H \in \mathbb{C}^{K \times MN}$ be the global CSI channel vector from all the BSs to all the users. $\mathbf{h}_k^H = [\mathbf{h}_{1k}^H, \ldots, \mathbf{h}_{Mk}^H] \in \mathbb{C}^{1 \times MN}$ is the channel from all the BSs to user k, and $\mathbf{h}_{ik}^H \in \mathbb{C}^{1 \times N}$ denote the channel vector from $\mathbf{BS}_i$ to user k. Let $\mathbf{w}_k^g \in \mathbb{C}^{MN \times 1}$ denote the precoding vectors from all the BSs to user k with global CSI.

Then, in the user k side, the received signal r_k is

$$r_k^g = \mathbf{h}_k^H \mathbf{w}_k^g s_k + \sum_{k' \neq k} \mathbf{h}_k^H \mathbf{w}_{k'}^g s_{k'} + n_k \tag{2.3}$$

where $\mathbf{w}_k^g$ and $\mathbf{w}_{k'}^g$ are the precoding vectors for user k and k'. Since each user is equipped with single antenna, the received signal r_k is scalar and the noise term n_k is distributed as $n_k \in \mathcal{CN}(0, 1)$.

When the received signals of all users are stacked together, we obtain $\mathbf{r}^g = [r_1^g, \ldots, r_K^g]^T \in \mathbb{C}^{K \times 1}$, which can be expressed in a more integrated form,

$$\mathbf{r}^g = \mathbf{H}\mathbf{W}^g\mathbf{s} + \mathbf{n}, \tag{2.4}$$

where $\mathbf{W}^g = [\mathbf{w}_1^g, \ldots, \mathbf{w}_K^g] \in \mathbb{C}^{MN \times K}$ is the precoding matrix designed based on the global CSI $\mathbf{H}$, and $\mathbf{n} = [n_1, \ldots, n_K]^T \in \mathbb{C}^{K \times 1}$.

To keep a fair comparison between the precoding with local individual CSI and global CSI, equal power is allocated among all the users[1]. Also a balanced precoding method called SLNR precoding which is presented in section 2.1.4, is adopted to evaluate the performance of both CSI cases.

According to the SLNR precoding, the precoding vector of user k under the global CSI case, $\bar{\mathbf{w}}_k^g$ can be expressed as

$$\bar{\mathbf{w}}_k^g = (\mathbf{I} + \sum_{k' \neq k} \alpha \mathbf{h}_{k'} \mathbf{h}_{k'}^H)^{-1} \mathbf{h}_k, \tag{2.5}$$

We can write the normalized precoding vector as

$$\mathbf{w}_k^g = \frac{\bar{\mathbf{w}}_k^g}{\|\bar{\mathbf{w}}_k^g\|} \tag{2.6}$$

where $\bar{\mathbf{w}}_k^g$ is the precoding vector for user k before normalization.

The corresponding SINR for the precoding with global CSI at user k is given as

$$\gamma_k^g = \frac{|\mathbf{h}_k^H \mathbf{w}_k^g|^2}{\sum_{k' \neq k} |\mathbf{h}_k^H \mathbf{w}_{k'}^g|^2 + 1} \tag{2.7}$$

where we assume the noise is complex Gaussian distributed with $n_k \in \mathcal{CN}(0, 1)$, According to the continuous mapping theorem [63], the sum rate of all users can be given as

[1]Here, it is different from the case in [63], in which it considers the normalized minimum constraints. It is more suitable for fair comparison in this work

$$R_{sum}^{g} = \sum_{k=1}^{K} \log(1 + \gamma_k^g) \tag{2.8}$$

2.1.3 Cooperative Multicell Precoding with Local Individual CSI

When the precoding is performed in each BS with local individual CSI, it means that the precoding vector $\mathbf{W}_i^l$ designed by BS i is based only on $\mathbf{H}_i = [\mathbf{h}_{i1}, \cdots, \mathbf{h}_{iK}]$. Let $\mathbf{h}_{ik}^H \in \mathbb{C}^{1\times N}$ denote the channel vector from $\mathbf{BS}_i$ to user k. Then $\mathbf{H} = [\mathbf{H}_1^H, \cdots, \mathbf{H}_M^H]$, and let $\mathbf{W}^l = [{\mathbf{W}_1^l}^H, \cdots, {\mathbf{W}_M^l}^H]^H$ be the precoding matrix for the local individual CSI case with $\mathbf{W}_i^l = [\mathbf{w}_{i1}^l, \cdots, \mathbf{w}_{ik}^l, \cdots, \mathbf{w}_{iK}^l]$, $\mathbf{w}_{ik}^l \in \mathbb{C}^{N\times 1}$.

Let $\mathbf{w}_{ik}^l$ and s_k denote the precoding vector and signal transmitted from BS i to user k. The precoded signal for BS i can be given as

$$\mathbf{x}_i = \mathbf{W}_i^l \mathbf{s} = \sum_{k=1}^{K} \mathbf{w}_{ik}^l s_k \tag{2.9}$$

where the precoding vector $\mathbf{w}_{ik}^l$ is a normalized vector and $\mathbf{w}_{ik}^l \in \mathbb{C}^{N\times 1}$. For the local CSI case, the precoding matrix $\mathbf{W}_i^l$ of BS i is designed based on the local individual CSI of BS i, $\mathbf{H}_i$.

Then the received signal in user k can be written as

$$r_k^l = \sum_{i=1}^{M} \mathbf{h}_{ik}^H \mathbf{W}_i^l \mathbf{s} = \sum_{i=1}^{M} \mathbf{h}_{ik}^H \sum_{k=1}^{K} \mathbf{w}_{ik}^l s_k + n_k \tag{2.10}$$

By stacking all the users' signal, the received signal vector model can be written as

$$\mathbf{r}^l = [r_1^l, \cdots, r_K^l]^T = \sum_{i=1}^{M} \mathbf{H}_i^H \mathbf{W}_i^l \mathbf{s} + \mathbf{n} \tag{2.11}$$

According to the SLNR precoding, the precoding vector from BS i to user k under the local individual CSI case, $\bar{\mathbf{w}}_{ik}^l$ can be expressed as

$$\bar{\mathbf{w}}_{ik}^l = (\mathbf{I} + \sum_{k'\neq k} \alpha \mathbf{h}_{ik'} \mathbf{h}_{ik'}^H)^{-1} \mathbf{h}_{ik} \tag{2.12}$$

The corresponding normalized precoding vector can be given as

$$\mathbf{w}_{ik}^{l} = \frac{\bar{\mathbf{w}}_{ik}^{l}}{\|\bar{\mathbf{w}}_{ik}^{l}\|} \tag{2.13}$$

The corresponding SINR for the precoding with local individual CSI at user k is given as

$$\gamma_k^l = \frac{\sum_{i=1}^{M} |\mathbf{h}_{ik}^H \mathbf{w}_{ik}^l|^2}{\sum_{i=1}^{M} \sum_{k' \neq k} |\mathbf{h}_{ik}^H \mathbf{w}_{ik'}^l|^2 + 1} \tag{2.14}$$

where we assume the noise is complex Gaussian distributed with $n_k \in \mathcal{CN}(0,1)$, According to the continuous mapping theorem [63], the sum rate of all users can be given as

$$R_{sum}^l = \sum_{k=1}^{K} \log(1+\gamma_k^l) \tag{2.15}$$

Remark 1. *The difference between precoding with global CSI and that with local individual CSI can be summarized as: the precoding matrix* $\mathbf{W}_i^l$ *for the local individual CSI case is designed based on* $\mathbf{H}_i$*, which is the channel matrix from BS* i *to all the users.* $\mathbf{H}_i$ *is called* local individual CSI. $\mathbf{W}^l = [{\mathbf{W}_1^l}^H, \cdots, {\mathbf{W}_M^l}^H]^H$ *is the precoding matrix for the local individual CSI case.*

If the local CSI $\mathbf{H}_i$ *is sent to the central control unit, and the global CSI* $\mathbf{H} = [\mathbf{H}_1^H, \cdots, \mathbf{H}_M^H]^H$ *is available to the central unit, then precoding marix for the global CSI case* $\mathbf{W}^g = [\mathbf{w}_1^g, \ldots, \mathbf{w}_K^g] \in \mathbb{C}^{MN \times K}$ *is designed based on* $\mathbf{H}$*.* $\mathbf{W}^g$ *is called the precoding with* global CSI $\mathbf{H}$*.*

2.1.4 Balanced Precoding Method

For multiuser MIMO precoding, it is critical for the precoding to take into consideration of maximizing the intended signal and minimizing the interference caused

to other users. This can be achieved by maximizing the signal to leakage plus noise ratio (SLNR). The SLNR precoding method [65] was a linear method, which is proposed to achieve optimal multiuser performance for MISO channel. Take the local individual CSI for example. The SLNR is written as

$$\mathrm{SLNR}_{ik} = \frac{|\mathbf{h}_{ik}^H \mathbf{w}_{ik}^l|^2}{1 + \sum\limits_{k' \neq k} \alpha |\mathbf{h}_{ik'}^H \mathbf{w}_{ik}^l|^2} \tag{2.16}$$

When the parameter α is properly optimized, by maximizing the SLNR for each receiver, the Pareto boundary points can be achieved. We can obtain the corresponding closed formed precoding vector as

$$\mathbf{w}_{ik}^l = \frac{\bar{\mathbf{w}}_{ik}^l}{\|\bar{\mathbf{w}}_{ik}^l\|} \tag{2.17}$$

where $\bar{\mathbf{w}}_{ik}^l$ can be expressed as

$$\bar{\mathbf{w}}_{ik}^l = (\mathbf{I} + \sum_{k' \neq k} \alpha \mathbf{h}_{ik'} \mathbf{h}_{ik'}^H)^{-1} \mathbf{h}_{ik} \tag{2.18}$$

The precoding vectors can also be given in a more compact form,

$$\bar{\mathbf{w}}_{ik}^l = (\mathbf{I} + \alpha \mathbf{H}_{[ik]} \mathbf{H}_{[ik]}^H)^{-1} \mathbf{h}_{ik} \tag{2.19}$$

where $\mathbf{H}_{[ik]} = [\mathbf{h}_{ik'}], k' \neq k$ is the collection of channels vectors for BS i to all users except user k .

2.2 Large Dimensional Analysis When N and K Both Go large at the Same Rate

This section gives the performance analysis for the scenario that the number of antennas at BSs N and the number of users K both go large at the same rate $N \to \infty, N/K = \beta$. We try to derive the large dimensional approximations for the SINR of precoding with global CSI and local individual CSI. The derivations are

obtained using the theoretic results of large random matrix theory. We make the following assumptions in this section.

Assumption 1. *The regime of large numbers of antennas and users is defined as follows:*

- *Large dimensional regime is considered where N and K approach infinity, $N \to \infty$ and $K \to \infty$.*
- *The number of antennas and users goes large at the same rate, that is $N/K = \beta$.*

2.2.1 Large Dimensional Approximations for Precoding with Local Individual CSI

According to the SLNR precoding method, we can rewrite the SINR under local individual CSI case as

$$\gamma_k^l = \frac{\sum_{i=1}^{M} |\mathbf{h}_{ik}^H \bar{\mathbf{w}}_{ik}^l|^2 / \|\bar{\mathbf{w}}_{ik}^l\|^2}{\sum_{i=1}^{M} \sum_{k' \neq k} |\mathbf{h}_{ik}^H \bar{\mathbf{w}}_{ik'}^l|^2 / \|\bar{\mathbf{w}}_{ik'}^l\|^2 + 1} \tag{2.20}$$

where $\bar{\mathbf{w}}_{ik}^l = (\mathbf{I} + \alpha \mathbf{H}_{[ik]} \mathbf{H}_{[ik]}^H)^{-1} \mathbf{h}_{ik}$ and $\bar{\mathbf{w}}_{ik'}^l = (\mathbf{I} + \alpha \mathbf{H}_{[ik']} \mathbf{H}_{[ik']}^H)^{-1} \mathbf{h}_{ik'}$.

The following lemma is used to obtain the deterministic approximations of the SINR.

Lemma 1. *As $N \longrightarrow \infty, K \longrightarrow \infty$ with a fixed ratio $0 < N/K = \beta < \infty$ [2], we can obtain*

$$|\boldsymbol{h}_{ik}^H \bar{\boldsymbol{w}}_{ik}^l|^2 \xrightarrow{N \to \infty} (\psi_{ik}^{\circ})^2 \tag{2.21}$$

$$\|\bar{\boldsymbol{w}}_{ik}^l\|^2 \xrightarrow{N \to \infty} \bar{\psi}_{ik}^{\circ} \tag{2.22}$$

[2] In this paper, the scenario that hundreds of antennas serving tens of users is assumed. And the ratio β should be smaller than 10.

$$|\boldsymbol{h}_{ik}^H \bar{\boldsymbol{w}}_{ik'}^l|^2 \xrightarrow{N\to\infty} \frac{1}{N}\frac{\bar{\Phi}_{ik'k}^{\circ}}{(1+\alpha\Phi_{ik'k}^{\circ})^2} \tag{2.23}$$

where in (2.21), $\frac{1}{N}tr\Sigma_{ik}-\psi_{ik}^{\circ} \xrightarrow{N\to\infty} 0$ *and in (2.22),* $\frac{1}{N}tr\Sigma_{ik}^2-\bar{\psi}_{ik}^{\circ} \xrightarrow{N\to\infty} 0$. *In (2.23),* $\frac{1}{N}tr\Sigma_{ik'k}-\Phi_{ik'k}^{\circ} \xrightarrow{N\to\infty} 0$, $\frac{1}{N}tr\Sigma_{ik'k}^2-\bar{\Phi}_{ik'k}^{\circ} \xrightarrow{N\to\infty} 0$. *And* $\Sigma_{ik}=(\frac{1}{N}\boldsymbol{I}+\sum\limits_{k'\neq k}\frac{\alpha}{N}\boldsymbol{h}_{ik'}\boldsymbol{h}_{ik'}^H)^{-1}$, $\Sigma_{ik'k}$ *is defined as* $\Sigma_{ik'k}=(\frac{1}{N}\boldsymbol{I}+\sum\limits_{(\bar{k})\neq(k,k')}\frac{\alpha}{N}\boldsymbol{h}_{i\bar{k}}\boldsymbol{h}_{i\bar{k}}^H)^{-1}$. *The terms* ψ_{ik}°, $\bar{\psi}_{ik}^{\circ}$, $\Phi_{ik'k}^{\circ}$ *and* $\bar{\Phi}_{ik'k}^{\circ}$ *are only functions of the slow fading coefficients of the channel.*

The detailed proof of Lemma 3 is given in Appendix B. Then we present the main result of precoding with local individual CSI in the following theorem.

Theorem 1. *As* $N\longrightarrow\infty, K\longrightarrow\infty$ *go large at a fixed ratio* $0<N/K=\beta<\infty$ *, the large dimensional deterministic equivalents of SINR* γ_k *can be written as*

$$\begin{aligned}\gamma_k^l &\xrightarrow{N\to\infty} {\gamma_k^l}^{\infty}\\ &= \frac{\sum\limits_{i=1}^{M}\frac{(\psi_{ik}^{\circ})^2}{\bar{\psi}_{ik}^{\circ}}}{\sum\limits_{i=1}^{M}\sum\limits_{k'\neq k}\frac{\beta_{ik'}}{N}\frac{\bar{\Phi}_{ik'k}^{\circ}}{\bar{\psi}_{ik'}^{\circ}(1+\alpha\Phi_{ik'k}^{\circ})^2}+1}\end{aligned} \tag{2.24}$$

with

$$\frac{1}{N}tr\mathbf{D}_{ik}\Sigma_{ik}-\psi_{ik}^{\circ} \xrightarrow{N\to\infty} 0, \tag{2.25a}$$

$$\frac{1}{N}tr\mathbf{D}_{ik}\Sigma_{ik}^2-\bar{\psi}_{ik}^{\circ} \xrightarrow{N\to\infty} 0, \tag{2.25b}$$

$$\frac{1}{N}tr\mathbf{D}_{ik}\Sigma_{ik'k}-\Phi_{ik'k}^{\circ} \xrightarrow{N\to\infty} 0 \tag{2.25c}$$

$$\frac{1}{N}tr\mathbf{D}_{ik}\Sigma_{ik'k}^2-\bar{\Phi}_{ik'k}^{\circ} \xrightarrow{N\to\infty} 0 \tag{2.25d}$$

where $\Sigma_{ik'k}=(\frac{1}{N}\boldsymbol{I}+\sum\limits_{(\bar{k})\neq(k,k')}\frac{\alpha}{N}\boldsymbol{h}_{i\bar{k}}\boldsymbol{h}_{i\bar{k}}^H)^{-1}$ *and* $\Sigma_{ik}=(\frac{1}{N}\boldsymbol{I}+\sum\limits_{k'\neq k}\frac{\alpha}{N}\boldsymbol{h}_{ik'}\boldsymbol{h}_{ik'}^H)^{-1}$. *The terms* ψ_{ik}°, $\bar{\psi}_{ik}^{\circ}$, $\Phi_{ik'k}^{\circ}$ *and* $\bar{\Phi}_{ik'k}^{\circ}$ *are only functions of the slow fading coefficients of the channel. This means that the SINR* γ_k^l *can be approximated by its asymptotic term* ${\gamma_k^l}^{\infty}$ *without knowing the actual channel realizations.*

Proof. We can plug in the results in Lemma 3 and the Theorem 1 can be directly derived. □

2.2.2 Large Dimensional Approximations for Precoding with Global CSI

Recall that the SINR of $\mathbf{UE}_k$ is given by

$$\gamma_k = \frac{|\mathbf{h}_k^H \bar{\mathbf{w}}_k^g|^2 / \|\bar{\mathbf{w}}_k^g\|^2}{\sum\limits_{k' \neq k} |\mathbf{h}_k^H \bar{\mathbf{w}}_{k'}^g|^2 / \|\bar{\mathbf{w}}_{k'}^g\|^2 + 1} \tag{2.26}$$

where $\bar{\mathbf{w}}_k^g = (\mathbf{H}_{[k]}\mathbf{H}_{[k]}^H + \alpha\mathbf{I})^{-1}\mathbf{h}_k$.

Lemma 2. *As $N \longrightarrow \infty, K \longrightarrow \infty$ go large at a fixed ratio $0 < N/K = \beta < \infty$, we obtain*

$$|\boldsymbol{h}_k^H \bar{\boldsymbol{w}}_k^g|^2 \xrightarrow{N\to\infty} (m_k^\circ)^2 \tag{2.27}$$

with

$$\begin{aligned} m_k^\circ &= \frac{1}{N} tr \mathbf{D}_k \Sigma_k \\ &= \frac{1}{N} tr \mathbf{D}_k diag(\mathbf{\Psi}_1(\alpha), \mathbf{\Psi}_2(\alpha), \ldots, \mathbf{\Psi}_M(\alpha)) \end{aligned} \tag{2.28}$$

where $\Sigma_k = (\frac{1}{N}\boldsymbol{I} + \sum\limits_{k' \neq k} \frac{\alpha}{N} \boldsymbol{h}_{k'} \boldsymbol{h}_{k'}^H)^{-1}$ and

$$\mathbf{\Psi}_i(\alpha) = (\frac{1}{N} \sum_{k=1}^{K} \frac{1}{\sum\limits_{i=1}^{M} e_{ik}(\alpha) + 1} \mathbf{D}_{ik} + \alpha\mathbf{I})^{-1},$$

with $e_{ik}(\alpha) = \frac{1}{N} tr(\mathbf{D}_{ik}\mathbf{\Psi}_i)$ and $\mathbf{D}_k = diag(\mathbf{D}_{1k}, \mathbf{D}_{2k}, \ldots, \mathbf{D}_{Mk})$. The term $\mathbf{\Psi}_i(\alpha)$ can be iteratively obtained.

The term $\|\bar{\boldsymbol{w}}_k^g\|^2$ can be approximated as

$$\|\bar{\boldsymbol{w}}_k^g\|^2 \xrightarrow{N\to\infty} \bar{m}_k^\circ \tag{2.29}$$

where $\frac{1}{N} tr\mathbf{D}_k\Sigma_k - \bar{m}_k^\circ \xrightarrow{N\to\infty} 0$.

For the term $|\boldsymbol{h}_k^H \bar{\boldsymbol{w}}_{k'}^g|^2$*, it can be derived in the same way as Lemma 3,*

$$|\boldsymbol{h}_k^H \bar{\boldsymbol{w}}_{k'}^g|^2 \xrightarrow{N\to\infty} \frac{1}{N}\frac{\bar{\Omega}_{k'k}^\circ}{(1+\alpha\Omega_{k'k}^\circ)^2} \tag{2.30}$$

where $\frac{1}{N} tr\Sigma_{k'k} - \Omega_{k'k}^\circ \xrightarrow{N\to\infty} 0$, $\frac{1}{N} tr\Sigma_{k'k}^2 - \bar{\Omega}_{k'k}^\circ \xrightarrow{N\to\infty} 0$. *And* $\Sigma_{k'k} = (\frac{1}{N}\boldsymbol{I} + \sum_{(\bar{k})\neq(k,k')} \frac{\alpha}{N}\boldsymbol{h}_{\bar{k}}\boldsymbol{h}_{\bar{k}}^H)^{-1}$. m_k°, $\bar{m}_k^\circ$, Ω° *and* $\bar{\Omega}_{k'k}^\circ$ *are only functions of the second order slow fading coefficients of the channel.*

In Lemma 2, the term $|\mathbf{h}_k^H \bar{\mathbf{w}}_k^g|^2$ can be derived in the same way as in Theorem 1 in [67].

Theorem 2. *As* $N \longrightarrow \infty, K \longrightarrow \infty$ *go large at a fixed ratio* $0 < N/K = \beta < \infty$ [3]*, the large dimensional equivalents of SINR* γ_k^g *can be written as*

$$\begin{aligned}\gamma_k^g &\xrightarrow{N\to\infty} \gamma_k^{g\infty} \\ &= \frac{\frac{(m_k^\circ)^2}{\bar{m}_k^\circ}}{\sum\limits_{\bar{k}\neq k} \frac{\beta_{k'}}{N}\frac{\bar{\Omega}_{k'k}^\circ}{\bar{\psi}_{k'}^\circ(1+\alpha\Omega_{k'k}^\circ)^2} + 1}\end{aligned} \tag{2.31}$$

with

$$\frac{1}{N} tr\mathbf{D}_k\Sigma_k - m_k^\circ \xrightarrow{N\to\infty} 0, \tag{2.32a}$$

$$\frac{1}{N} tr\mathbf{D}_k\Sigma_k^2 - \bar{m}_k^\circ \xrightarrow{N\to\infty} 0, \tag{2.32b}$$

$$\frac{1}{N} tr\mathbf{D}_k\Sigma_{k'k} - \Omega_{k'k}^\circ \xrightarrow{N\to\infty} 0 \tag{2.32c}$$

$$\frac{1}{N} tr\mathbf{D}_k\Sigma_{k'k}^2 - \bar{\Omega}_{k'k}^\circ \xrightarrow{N\to\infty} 0 \tag{2.32d}$$

[3] In this paper, the scenario that hundreds of antennas serving tens of users is assumed. And the ratio β should be smaller than 10.

where $\Sigma_{k'k} = (\frac{1}{N}\boldsymbol{I} + \sum_{(\bar{k})\neq(k,k')} \frac{\alpha}{N}\boldsymbol{h}_{\bar{k}}\boldsymbol{h}_{\bar{k}}^H)^{-1}$ *and* $\Sigma_k = (\frac{1}{N}\boldsymbol{I} + \sum_{k'\neq k} \frac{\alpha}{N}\boldsymbol{h}_{k'}\boldsymbol{h}_{k'}^H)^{-1}$. m_k°, $\bar{m}_k^\circ$, Ω° *and* $\bar{\Omega}_{k'k}^\circ$ *are only functions of the second order slow fading coefficients of the channel. This means that the SINR* γ_k^g *can be approximated by its asymptotic term* $\gamma_k^{g\infty}$ *without knowing the actual channel realizations.*

Proof. We can plug in the results in Lemma 2 and the Theorem 2 can be directly derived. □

2.3 Large Dimensional Analysis When $N/K \to \infty$ and K is Fixed

In this section, we will compare the performances of the two CSI cases in the second scenario: $N/K \to \infty$ and K is fixed. We will show that the precoding with local individual CSI can achieve the same performance as that with global CSI. First, we will present an useful property of random vectors.

We first present the asymptotic results on gaussian vectors [24]. Let $\mathbf{x} = [x_1, \ldots, x_n]^T$ and $\mathbf{y} = [y_1, \ldots, y_n]^T$ be random vectors with their entries Gaussian distributed with variance σ_x^2 and σ_y^2. They are independent with each other. When the dimension of both vectors goes large,

$$\lim_{n\to\infty} \frac{\mathbf{x}^H\mathbf{x}}{n} \overset{a.s.}{\to} \sigma_x^2 \tag{2.33}$$

$$\lim_{n\to\infty} \frac{\mathbf{x}^H\mathbf{y}}{n} \overset{a.s.}{\to} 0 \tag{2.34}$$

where $\overset{a.s.}{\to}$ denotes that the term converges almost surely.

When the number of antennas at the BSs is much larger than that of the users, there is a favorable property for the channels between the users and BSs. The channels between the users and BSs are asymptotic orthogonal. Based on this property, we can derive that the precoding with local individual CS and that with global CSI are

equivalent when the number of antennas at the BSs is much large than that of the users. And we assume maximum ratio transmission (MRT) precoding in this scenario, since it also zero forces the interference due to asymptotic orthogonality. The result is provided in the following theorem.

Theorem 3. *The SINR of precoding with local individual CSI (2.14) and that with global CSI (2.7) are asymptotic equivalent when $N/K \to \infty$ and K is fixed.*

Proof. When the MRT precoding is adopted, the precoding vector for local individual CSI based scheme can be expressed as $\mathbf{w}_{ik}^l = \frac{\mathbf{h}_{ik}}{\|\mathbf{h}_{ik}\|}$, and precoding vector for the global CSI based scheme is $\mathbf{w}_k^g = \frac{\mathbf{h}_k}{\|\mathbf{h}_k\|}$.

We can plug the precoding vector $\mathbf{w}_k^g$ into the SINR term γ_k^g under global CSI. Then the γ_k^g can be rewritten as follows,

$$\begin{aligned}
\gamma_k^g &= \frac{|\mathbf{h}_k^H \bar{\mathbf{w}}_k^g|^2 / \|\bar{\mathbf{w}}_k^g\|^2}{\sum\limits_{k' \neq k} |\mathbf{h}_k^H \bar{\mathbf{w}}_{k'}^g|^2 / \|\bar{\mathbf{w}}_{k'}^g\|^2 + 1} \\
&= \frac{|\mathbf{h}_k^H \frac{\mathbf{h}_k}{\|\mathbf{h}_k\|}|^2}{\sum\limits_{k' \neq k} |\mathbf{h}_k^H \frac{\mathbf{h}_{k'}}{\|\mathbf{h}_{k'}\|}|^2 + 1} \\
&= \frac{|\mathbf{h}_k^H \frac{\mathbf{h}_k}{\|\mathbf{h}_k\|}|^2}{\sum\limits_{k' \neq k} |\frac{\mathbf{g}_k^H \mathbf{D}_k \mathbf{D}_{k'} \mathbf{g}_{k'}}{\|\mathbf{h}_{k'}\|}|^2 + 1}
\end{aligned}$$

since $\mathbf{g}_k$ and $\mathbf{g}_{k'}$ are independent, using the trace lemma (Lemma 3) in appendix B, we can see that $\frac{\mathbf{g}_k^H \mathbf{D}_k \mathbf{D}_{k'} \mathbf{g}_{k'}}{\|\mathbf{h}_{k'}\|} \xrightarrow{N\to\infty} 0$. The interference term in the denominator is zero. Then the SINR in (2.7) can be expressed as

$$\begin{aligned}
\gamma_k^g &\xrightarrow{N\to\infty} |\mathbf{h}_k^H \frac{\mathbf{h}_k}{\|\mathbf{h}_k\|}|^2 \\
&= \|\mathbf{h}_k\|^2 \\
&= \|\mathbf{D}_k^{\frac{1}{2}} \mathbf{g}_k\|^2 \\
&= \sum_{i=1}^{M} \|\mathbf{D}_{ik}^{\frac{1}{2}} \mathbf{g}_{ik}\|^2
\end{aligned} \tag{2.35}$$

Next, we will consider the SINR γ_k^l in (2.14). We can plug the precoding vector $\mathbf{w}_{ik}^l$ into the SINR term γ_k^l under local individual CSI. Then the γ_k^l can be rewritten as follows,

$$\begin{aligned}
\gamma_k^l &= \frac{\sum_{i=1}^{M} |\mathbf{h}_{ik}^H \mathbf{w}_{ik}^l|^2}{\sum_{i=1}^{M} \sum_{k'\neq k} |\mathbf{h}_{ik}^H \mathbf{w}_{ik'}^l|^2 + 1} \\
&= \frac{\sum_{i=1}^{M} |\mathbf{h}_{ik}^H \frac{\mathbf{h}_{ik}}{\|\mathbf{h}_{ik}\|}|^2}{\sum_{i=1}^{M} \sum_{k'\neq k} |\mathbf{h}_{ik}^H \frac{\mathbf{h}_{ik'}}{\|\mathbf{h}_{ik'}\|}|^2 + 1}
\end{aligned}$$

Based on the asymptotic orthogonality principle $\mathbf{h}_{ik}^H \frac{\mathbf{h}_{ik'}}{\|\mathbf{h}_{ik'}\|} \xrightarrow{N\to\infty} 0$, the interference term in the denominator is zero. Then the SINR γ_k in (2.14) can be expressed as

$$\gamma_k^l \xrightarrow{N\to\infty} \sum_{i=1}^{M} |\mathbf{h}_{ik}^H \mathbf{w}_{ik}^l|^2$$
$$= \sum_{i=1}^{M} |\mathbf{h}_{ik}^H \frac{\mathbf{h}_{ik}}{\|\mathbf{h}_{ik}\|}|^2$$
$$= \sum_{i=1}^{M} \|\mathbf{D}_{ik}^{\frac{1}{2}} \mathbf{g}_{ik}\|^2 \tag{2.36}$$

We can see that the asymptotic expressions γ_k^l and γ_k^g for both CSI cases (2.35)(2.36) are the same when $N/K \to \infty$ and K is fixed. Thus, Theorem 3 is proved.

□

Theorem 3 actually shows a very interesting advantageous property of the second scenario, $N/K \to \infty$ and K is fixed. When the number of antennas is much larger than that of the users served, precoding with local CSI in cooperative multicell transmission can achieve the rate performance as good as that with global CSI. The impact of limited CSI in this scenario vanishes. The asymptotic orthogonality actually decouples the interference among all the cells, which is the main reason for this property. This motivates us to perform the cooperative multicell transmission with only local CSI instead of global CSI, when much more antennas are deployed in the BSs compared with the number of users. In this way much signaling overhead for the CSI information exchange can be saved.

2.4 Simulation Results

The large dimensional approximation results developed in section 2.2 provide an efficient way to obtain the average rate for the precoding methods. In this section, we try to provide numerical simulations to verify the accuracy of the large dimensional results. The performance gap between both precoding schemes is also revealed. When

we increase the number of BSs, the rate performance of precoding with local individual CSI will decrease almost linearly. The numerical results are based on the Monte Carlo simulation. Many random Rayleigh fading channels realizations are generated to verify the results.

2.4.1 Asymptotic Gap for Precoding with Local Individual CSI

A multicell network consisting of $M = 2$ BSs is considered in this simulation. Each BS is subject to a fixed power constraint. All the users are clustered in the center of the BSs. The path-loss model from the BSs to the users is assumed to be $L(dB) = 15.3+37.6\log_{10} d$, where d represents the distance between BS and users. For the local individual CSI precoding case, each BS precode based on its own CSI, while for the precoding with global CSI, the two BSs can be regarded as a single BS ($M = 1$) by sharing the individual CSI. Then the precoding is performed as the multiuser MISO broadcast channel. Different ratio of antennas and users is considered:$N = 100, 200, 300, 400, 500, 600$ and $N/K = \beta = 2, 4, 5, 10$. The simulation results are obtained by averaging over a 2000 Rayleigh block fading channels.

Fig. 2.3 illustrates the convergence behavior of the average rate performance when the number of antennas deployed in the BSs goes large. When precoding with local individual CSI ($M = 2$), there is a constant performance loss in the large number of antennas regime. Also the rate per user is only related to the ratio of antennas and users $N/K = \beta$, which is irrespective of the actual channel realizations. The theoretical values in Fig. 2.3 are derived in Theorem 1 and Theorem 2, which are the large dimensional approximations for the average rate per user.

Fig. 2.4 illustrates the accuracy of large dimensional approximations developed in Section 2.2. For the local individual CSI precoding case, each BS precode based on its own CSI ($M = 2$), while for the precoding with global CSI, the two BSs can

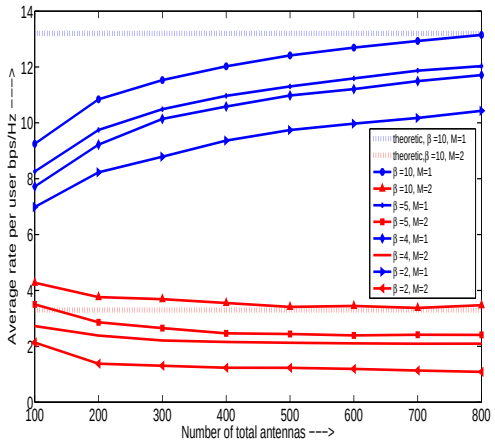

Figure 2.3. Average rate performance versus the number of antennas in the BSs. Different ratios of $N/K = \beta = 2, 4, 5, 10$ are considered. .

be regarded as a single BS ($M = 1$) by sharing the local individual CSI. We compare the analytical results in Section 2.2 with the Monte Carlo simulation results. Fig. 2.4 shows that for both precoding cases, the large dimensional approximations are accurate for different ratios of antennas and users $N/K = \beta$.

2.4.2 The Impact of Increased Number of BSs on the Performance of Local Individual CSI Based Precoding

We now investigate the impact of increased number of BSs on the performance of local individual CSI based precoding. The total number of antennas deployed in all the BSs is fixed, $N_{total} = 800$. We change the number of BSs to be $M = 1, 2, 4, 5, 8$. The antennas are distributed in each BS equally. In Fig. 2.5, when we keep the total number of antennas be fixed, the average rate performance of precoding with local individual CSI would degrade if the number of cells (BSs) is increased. Also, the average rate performance would saturate in high SNR.

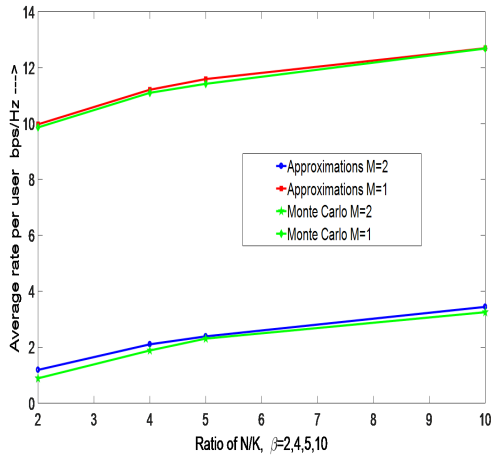

Figure 2.4. Average rate performance versus the ratio $N/K = \beta = 2, 4, 5, 10$. The total number of antennas is 600. .

When the number of cooperating BSs increases, the extent of distributedness increases. We consider different levels of SNR: high SNR, intermediate SNR and low SNR. As shown in Fig. 2.6, the average rate decreases linearly with the increased number of BSs when the total number of antennas is fixed, especially for the high SNR case.

2.5 Conclusions

The performance of precoding with global CSI is compared with that of precoding with local individual CSI in the large dimensional regime. It is easy to quantify the performance loss caused by limited CSI in the large dimensional regime. Using theoretical results from the large random matrix theory, we can derive large dimensional approximations of SINR for both precoding cases. The results depend only on the slow fading channel statistics instead of the fast fading channel realizations. Two different massive MIMO scenarios are considered. When the number of antennas and that of users goes large at the same rate, there is a constant gap between the two

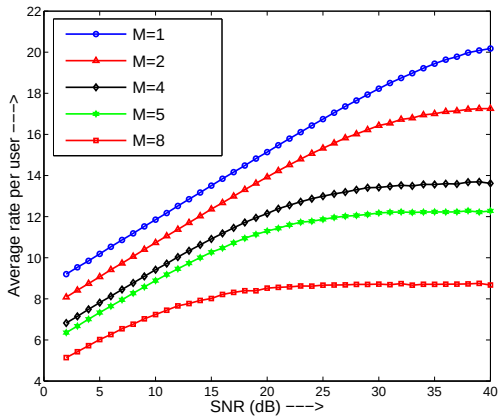

Figure 2.5. Average rate performance versus SNR. The total number of antennas is fixed, $N_{total} = 800$, while the number of BSs increases $M = 1, 2, 4, 5, 8$.

CSI cases. When the number of antennas goes large, while the number of users is fixed, the performances of both CSI cases are the same. This means that the impact of local individual CSI vanishes. Numerical results validates the theoretical analysis and show that there is constant performance loss due to the price of limited CSI. For future work, we can take advantage of the large dimensional results to optimize the network. Since the SINR is only related with the slow fading channel statistics, it is easier for us to optimize the system over a slow time scale, which could reduce the optimization update rate and signaling overhead.

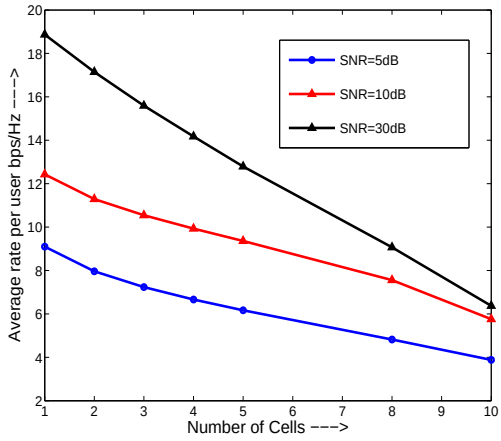

Figure 2.6. Average rate performance versus the number of BSs. Different SNR levels are considered. .

CHAPTER 3

Optimization for User Centric Massive MIMO Cell Free Network Via Large Dimensional Analysis

The traditional cellular network infrastructure suffers from severe intercell interference, especially for cell-edge users. The only reason for the severe interference caused at the cell edge users is that each user is only served by its own BS according to the traditional cellular network paradigm. BS cooperation [34] [52] can be an efficient way to mitigate inter BS interference. Multiple nearby BSs cooperatively transmitting to a user can convert the interference from the neighboring cells into useful signals. These BSs can be seen as a virtual cell serving the mobile users. This kind of BS cooperation can lead to a novel network paradigm that breaks the boundary of traditional cells. We call this *user centric cell free* networks. In [29], It allows the users to be served by multiple BSs, not restricted to its own BS as in the conventional cellular network, which can provide additional flexibility to tap the potential of the cellular network. Since each user can be flexibly served by multiple BSs, the resource of the network can be efficiently assigned in the global network perspective.

Note that this network paradigm does not change the existing cellular infrastructure. We just need to properly design the precoding vector for each user from multiple BSs and perform power allocation to achieve better network performance. Since the signal for a user is formed by the signals from multiple different BSs, it requires joint optimization of all the BSs, which is very complicated, especially for large scale networks. We primarily focus on a heuristic precoding method, which balances the maximization of signal to the intended user and minimization of inter-

ference caused to other users [65]. The sum rate optimization of cell free network requires global channel state information (CSI) and estimating the CSI consumes lots of resources, such as pilot training. Sharing the instantaneous CSI requires large amount of information exchange between the BSs and central control unit, especially in fast fading scenarios.

In massive MIMO networks, channel dimensions become large due to the massive antennas deployed in the BSs. In this case, certain terms in the SINR of users can be approximated with large scale fading irrespective of the small scale channel realization. In the large dimensional regime, the SINR is only related to CSI statistics. Then, the optimization for massive MIMO network can be performed based only on large scale channel statistics. The performance of large scale MIMO system has been extensively studied based on random matrix theory (RMT). In [66] [67], it is shown that the performance of regularized zero forcing and maximum ratio transmission converge in the large dimensional regime and the regularized zero forcing is superior compared with maximum ratio transmission. The asymptotic analysis for receiver design in the BS was investigated in [69] [40]. The theoretical results of the large dimensional analysis can predict the performance of massive MIMO wireless networks accurately. It is even accurate for some finite cases.

In this Chapter, we perform network optimization based on the theoretical results of random matrix theory. In previous MIMO network optimization, the power allocation or other system parameter optimization depends on the knowledge of instantaneous channel realizations [47]. The optimization should adapt timely to the change of instantaneous CSI. In fast fading scenarios, it would be impractical to do the optimization, especially for large scale mobile networks. When deploying large number of antennas in the BSs, certain parameters in the SINR expressions only depend on the channel statistical information, such as the path-loss and shadowing.

Then we don't need to follow the variation of instantaneous CSI for the network optimization so that much communication resource can be saved. Also, it is easier for us to optimize the system over a slow time scale, which could reduce the optimization update rate. This can save much system overhead which is critical for wireless networks.

Our main contributions are summarized as follows:

- We combine the benefits of massive MIMO and user centric cell free networks.
- Large dimensional analysis is carried out to approximate the SINR, which depends only on the channel statistical information.
- We perform optimization based on the channel statistics, which can greatly reduce the system overhead and computational update.

Organizations: The remainder of the chapter is organized as follows. In Section 4.1, we introduce the system model and formulate the sum rate optimization problem. The heuristic precoding structure that balances signal and interference is provided in Section III. In Section IV, we derive the large dimensional approximation of certain terms in the SINR and reformulate the power optimization problem that depends only on channel statistics. Numerical results and discussions are presented in Section V. We conclude this paper in Section VI.

3.1 System Model and Problem Formulation

3.1.1 Downlink Cell Free Network Model

We consider a cellular network consisting of M BSs and K users, which is shown in Fig. 3.2. Each BS is equipped with N antennas, and each user has a single antenna. We don't consider the scheduling problem. All the users can be served by several nearby BSs simultaneously, as compared with the traditional cellular network

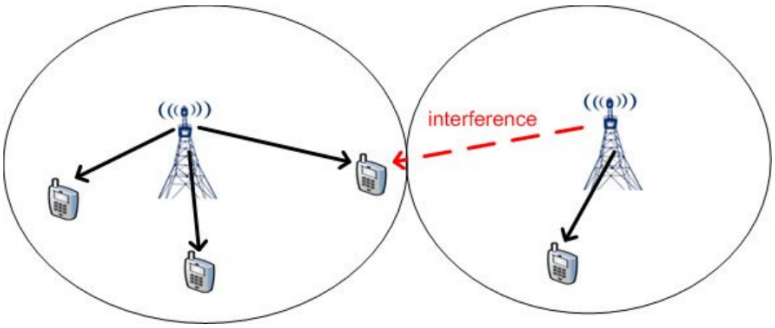

Figure 3.1. Traditional cell model.

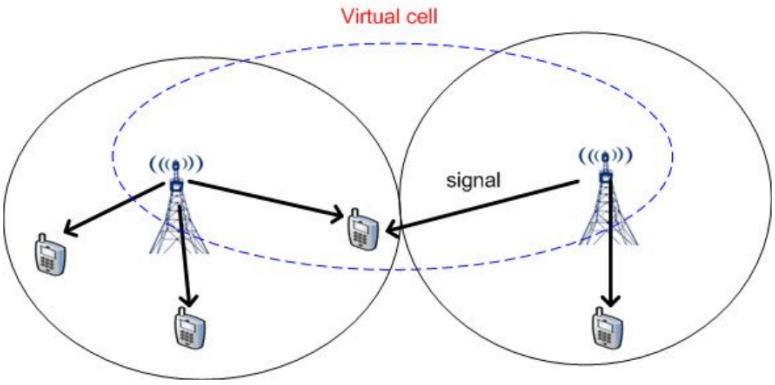

Figure 3.2. Cell free model.

that user can only be served by only one BS. The universal frequency reuse is adopted for the downlink transmission. There can be severe interference among the signals from the BSs to the users. The BSs are connected to a central controller by the backhaul network. The controller will compute the system operating parameters for the BSs. The focus of the controller is on the transmission interference coordination of the BSs side.

The transmitted signal of BS m, denoted as $\mathbf{x}_m$, is given by the following expression

$$\mathbf{x}_m = \sum_{k \in \pi_m} \mathbf{w}_{mk} \sqrt{p_{mk}} s_k, \tag{3.1}$$

where $\mathbf{w}_{mk}$ denotes the precoding vector from BS m to user k, π_m are the set of users that is served by BS m, p_{mk} is the power assigned to user k and then the received signal of user k can be expressed as

$$y_k = \underbrace{\sum_{m\in\rho_k} \mathbf{h}_{mk}^H \mathbf{w}_{mk}\sqrt{p_{mk}}s_k}_{desired\ signal} + \underbrace{\sum_{j\neq k}\sum_{\bar{m}\in\rho_j} \mathbf{h}_{\bar{m}k}^H \mathbf{w}_{\bar{m}j}\sqrt{p_{\bar{m}j}}s_j}_{interference} + n_k, \tag{3.2}$$

where n_k is the additive complex Gaussian noise with distribution $\mathcal{CN}(0,1)$. ρ_k and ρ_j is the set of BSs that serve users k and j. Let $\mathbf{h}_{mk}^H$ denotes the channel from BS m to the user$k$, which can be expressed as

$$\mathbf{h}_{mk} = \mathbf{D}_{mk}^{\frac{1}{2}}\mathbf{g}_{mk}, \tag{3.3}$$

where $\mathbf{g}_{mk}$ is the small scale fading from the mth BS antennas to user k. It is assumed to be fixed over one coherence time, and it is complex Gaussian distributed, $\mathbf{g}_{mk} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$. $\mathbf{D}_{mk} = \mathrm{diag}(\beta_{mk}, \ldots, \beta_{mk})$ is the large scale fading (path-loss and shadowing), which is constant over a much longer coherence time.

Denote the signal to interference plus noise ratio (SINR) for user k as γ_k. It can be expressed as

$$\gamma_k = \frac{\sum\limits_{m\in\rho_k} |\sqrt{p_{mk}}\mathbf{h}_{mk}^H \mathbf{w}_{mk}|^2}{\sum\limits_{j\neq k}\sum\limits_{\bar{m}\in\rho_j} |\sqrt{p_{\bar{m}j}}\mathbf{h}_{\bar{m}k}^H \mathbf{w}_{\bar{m}j}|^2 + 1} \tag{3.4}$$

where $\mathbf{w}_{\bar{m}j}$ is the precoding vector from BS $\bar{m}$ to user j. Then the corresponding achievable rate for user k is $\log(1+\gamma_k)$.

3.1.2 Problem Formulation

The sum rate of the cell free network can be written as

$$\begin{aligned} \underset{p_{mk},\mathbf{w}_{mk}}{\text{maximize}} \quad & R_{sum} = \sum_{k=1}^{K} \log(1+\gamma_k) \\ \text{subject to} \quad & \sum_{k\in\pi_m} p_{mk} \leq P_m \end{aligned} \tag{3.5}$$

where γ_k is the SINR for user k. Here we assume the power constraint in each BS is the same for the convenience of comparison of different schemes.

Since the SINR of each user is determined by the signal from multiple BSs and interference caused by other users. The transmission signal from each BS contributes useful signal power to the intended users, while at the same time it will cause interference to all other users. Thus the optimization of the sum rate requires a joint optimization of all the signals, which is very complicated. Especially when the scale of the network (users and BSs)goes large, the complexity of the optimization would be prohibitively high.

3.2 Heuristic Precoding Structure: Balancing Signal and Interference Leakage Power

In this section, we will introduce a heuristic precoding method for the maximization of the network sum rate.

3.2.1 Heuristic Precoding Structure

For multiuser MIMO precoding, it is critical for the precoding to take into consideration of maximizing the intended signal and minimizing the interference caused to other users. This can be achieved by maximizing the signal to leakage plus noise ratio (SLNR). The SLNR precoding method [65] was a linear method, which is proposed to achieve optimal multiuser performance for MISO channel. The SLNR is written as

$$\mathrm{SLNR}_{mk} = \frac{|\mathbf{h}_{mk}^H \mathbf{w}_{mk}|^2 p_{mk}}{1 + \sum\limits_{\bar{k} \neq k} \alpha |\mathbf{h}_{m\bar{k}}^H \mathbf{w}_{mk}|^2 p_{mk}} \tag{3.6}$$

When the parameter α is properly optimized, by maximizing the SLNR for each receiver, the Pareto boundary points can be achieved. We can obtain the corresponding closed formed precoding vector as

$$\mathbf{w}_{mk} = \frac{\bar{\mathbf{w}}_{mk}}{\|\bar{\mathbf{w}}_{mk}\|} \tag{3.7}$$

and $\bar{\mathbf{w}}_{mk}$ is expressed as

$$\bar{\mathbf{w}}_{mk} = (\mathbf{I} + \sum_{\bar{k} \neq k} \alpha \mathbf{h}_{m\bar{k}} \mathbf{h}_{m\bar{k}}^H)^{-1} \mathbf{h}_{mk} \tag{3.8}$$

which can be expressed in a more compact form,

$$\bar{\mathbf{w}}_{mk} = (\mathbf{I} + \alpha \mathbf{H}_{mk} \mathbf{H}_{mk}^H)^{-1} \mathbf{h}_{mk} \tag{3.9}$$

where the columns of $\mathbf{H}_{mk}$ are $\mathbf{h}_{m\bar{k}}, \bar{k} \neq k$.

The sum rate optimization could be represented into the following problem,

$$\begin{aligned} \underset{p_{mk}}{\text{maximize}} \quad & R_{sum} = \sum_{k=1}^{K} \log(1 + \gamma_k) \\ \text{subject to} \quad & \sum_{k \in \pi_m} p_{mk} \leq P_m \end{aligned} \tag{3.10}$$

hence, we only need to optimize over the power allocation and the weighting coefficients α. In this way, much computational overhead with respect to precoding optimization can be avoided.

3.2.2 A Special Case

Here, we will also consider a zero forcing based suboptimal algorithm for the problem (4.10). We first rewrite (3.9) according to matrix inversion lemma,

$$\bar{\mathbf{w}}_{mk} = (\mathbf{I} - \mathbf{H}_{mk}(\frac{1}{\alpha}\mathbf{I} + \mathbf{H}_{mk}\mathbf{H}_{mk}^H)^{-1}\mathbf{H}_{mk}^H)\mathbf{h}_{mk} \tag{3.11}$$

The zero forcing algorithm can be obtained by letting the weighting coefficient α go to infinity, then we get

$$\bar{\mathbf{w}}_{mk} = (\mathbf{I} - \mathbf{H}_{mk}(\mathbf{H}_{mk}\mathbf{H}_{mk}^H)^{-1}\mathbf{H}_{mk}^H)\mathbf{h}_{mk} \tag{3.12}$$

For each precoding vector $\bar{\mathbf{w}}_{mk}$, it satisfies $\mathbf{h}_{m\bar{k}}^H\bar{\mathbf{w}}_{mk} = 0, \bar{k} \neq k$. In this way, the precoding precancels all the interference caused to other users, and the SINR γ_k for user k can be given as,

$$\gamma_k = \sum_{m\in\rho_k} |\sqrt{p_{mk}}\mathbf{h}_{mk}^H\mathbf{w}_{mk}|^2 \tag{3.13}$$

where $\mathbf{w}_{mk} = \frac{\bar{\mathbf{w}}_{mk}}{\|\bar{\mathbf{w}}_{mk}\|}$ is the normalized precoding vector. Then the problem (3.10) for the zero forcing precoding has the following form,

$$\begin{aligned} \underset{p_{mk}}{\text{maximize}} \quad & R_{sum} = \sum_{k=1}^{K}\log(1 + \sum_{m\in\rho_k}|\sqrt{p_{mk}}\mathbf{h}_{mk}^H\mathbf{w}_{mk}|^2) \\ \text{subject to} \quad & \sum_{k\in\pi_m} p_{mk} \leq P_m \end{aligned} \tag{3.14}$$

The optimal power allocation for this problem can be solved by opportunistic water-filling, i.e. $p_{mk} = (p' - \frac{1}{\max_{\rho_k}(|\mathbf{h}_{mk}^H\mathbf{w}_{mk}|^2)})^+$. The power level p' is set to satisfy the power constraint $\sum_{k\in\pi_m} p_{mk} \leq P_m$.

3.3 Large Dimensional Analysis and Optimization based on Channel Statistics

The optimization problem in the previous section requires instantaneous CSI. The instantaneous CSI includes fast fading, which may change rapidly with time. It will cause the optimization algorithm to frequently adapt to the changing information. Since deploying large number of antennas in the BSs is one of the most important technologies in 5G networks. In this case, certain terms in the SINR expressions tend to depend only on large scale channel statistical information. Then we can optimize the power based only on the channel statistics.

3.3.1 Large Dimensional Analysis

In this section, we will perform a large dimensional analysis of the SLNR precoding scheme when the number of antennas N goes to infinity. Motivated by existing work [66] [67], we will show that in the large dimensional regime, the SINR of the user γ_k converges almost surely to deterministic approximations irrespective of instantaneous channel state information.

Note that $\mathbf{H}_{mk}$ is $N \times (K-1)$ with i.i.d standardized complex Gaussian entries, so $\frac{1}{N}\mathbf{H}_{mk}\mathbf{H}_{mk}^H$ has a uniformly bounded spectral norm for all large N. Let $\Sigma_{mk} = (\frac{1}{N}\mathbf{I} + \sum_{\bar{k}\neq k} \frac{\alpha}{N}\mathbf{h}_{m\bar{k}}\mathbf{h}_{m\bar{k}}^H)^{-1}$. Then we can obtain the following results:

Lemma 3. *As $N \longrightarrow \infty, K \longrightarrow \infty$ go large at a fixed ratio $0 < N/K = \beta < \infty$, the large dimensional deterministic equivalents of the terms in SINR can be written as*

$$|\boldsymbol{h}_{mk}^H \boldsymbol{w}_{mk}|^2 \xrightarrow{N\to\infty} \frac{(\psi_{mk}^{\circ})^2}{\bar{\psi}_{mk}^{\circ}} \tag{3.15}$$

$$|\boldsymbol{h}_{\bar{m}k}^H \boldsymbol{w}_{\bar{m}j}|^2 \xrightarrow{N\to\infty} \frac{\beta_{\bar{m}j}}{N} \frac{\bar{\Phi}_{\bar{m}kj}^{\circ}}{\bar{\psi}_{\bar{m}j}^{\circ}(1+\alpha\Phi_{\bar{m}kj}^{\circ})^2} \tag{3.16}$$

where in (3.15) $\frac{1}{N}tr\mathbf{D}_{mk}\Sigma_{mk} - \psi_{mk}^{\circ} \xrightarrow{N\to\infty} 0$ and $\frac{1}{N}tr\mathbf{D}_{mk}\Sigma_{mk}^2 - \bar{\psi}_{mk}^{\circ} \xrightarrow{N\to\infty} 0$. In (3.16), $\frac{1}{N}tr\mathbf{D}_{\bar{m}k}\Sigma_{\bar{m}kj} - \Phi_{\bar{m}kj}^{\circ} \xrightarrow{N\to\infty} 0$ and $\frac{1}{N}tr\mathbf{D}_{\bar{m}k}\Sigma_{\bar{m}kj}^2 - \bar{\Phi}_{\bar{m}kj}^{\circ} \xrightarrow{N\to\infty} 0$.

Proof. The detailed proof is given in the Appendix D.

□

Lemma 4. *As $N \longrightarrow \infty, K \longrightarrow \infty$ go large at a fixed ratio $0 < N/K = \beta < \infty$, the large dimensional deterministic equivalents of SINR can be written as*

$$\begin{aligned}\gamma_k &\xrightarrow{N\to\infty} \gamma_k^{\infty} \\ &= \frac{\sum\limits_{m\in\rho_k} \frac{(\psi_{mk}^{\circ})^2}{\bar{\psi}_{mk}^{\circ}} p_{mk}}{\sum\limits_{j\neq k}\sum\limits_{\bar{m}\in\rho_j} \frac{\beta_{\bar{m}j}}{N} \frac{\bar{\Phi}_{\bar{m}kj}^{\circ}}{\bar{\psi}_{\bar{m}j}^{\circ}(1+\alpha\Phi_{\bar{m}kj}^{\circ})^2} p_{\bar{m}j} + 1}\end{aligned} \tag{3.17}$$

with

$$\frac{1}{N} tr\mathbf{D}_{mk}\Sigma_{mk} - \psi_{mk}^{\circ} \xrightarrow{N\to\infty} 0 \tag{3.18}$$

$$\frac{1}{N} tr\mathbf{D}_{mk}\Sigma_{mk}^2 - \bar{\psi}_{mk}^{\circ} \xrightarrow{N\to\infty} 0 \tag{3.19}$$

$$\frac{1}{N} tr\mathbf{D}_{\bar{m}k}\Sigma_{\bar{m}jk} - \Phi_{\bar{m}jk}^{\circ} \xrightarrow{N\to\infty} 0 \tag{3.20}$$

$$\frac{1}{N} tr\mathbf{D}_{\bar{m}k}\Sigma_{\bar{m}jk}^2 - \bar{\Phi}_{\bar{m}jk}^{\circ} \xrightarrow{N\to\infty} 0 \tag{3.21}$$

where $\Sigma_{\bar{m}jk} = (\frac{1}{N}\boldsymbol{I} + \sum\limits_{(\bar{k})\neq(k,k')} \frac{\alpha}{N}\boldsymbol{h}_{\bar{m}\bar{k}}\boldsymbol{h}_{\bar{m}\bar{k}}^H)^{-1}$ and $\Sigma_{mk} = (\frac{1}{N}\boldsymbol{I} + \sum\limits_{\bar{k}\neq k} \frac{\alpha}{N}\boldsymbol{h}_{m\bar{k}}\boldsymbol{h}_{m\bar{k}}^H)^{-1}$. ψ_{mk}°, $\bar{\psi}_{mk}^{\circ}$, $\Phi_{\bar{m}jk}^{\circ}$ and $\bar{\Phi}_{\bar{m}jk}^{\circ}$ are only functions of the slow fading coefficients of the channel. This means that the SINR γ_k can be approximated by its asymptotic term γ_k^{∞} without knowing the actual channel realizations.

Proof. The derived asymptotic expression of SINR can be obtained based on the results derived in Lemma 1. □

Let $c_{mk} = \frac{(\psi_{mk}^{\circ})^2}{\bar{\psi}_{mk}^{\circ}}$ and $c_{\bar{m}kj} = \frac{\beta_{\bar{m}j}}{N}\frac{\bar{\Phi}_{\bar{m}kj}^{\circ}}{\bar{\psi}_{\bar{m}j}^{\circ}(1+\alpha\Phi_{\bar{m}kj}^{\circ})^2}$, then the SINR can be written as $\gamma_k = \frac{\sum\limits_{m\in\rho_k} c_{mk}p_{mk}}{\sum\limits_{j\neq k, j\in\mathcal{K}}\sum\limits_{\bar{m}\in\rho_j} c_{\bar{m}kj}p_{\bar{m}j}+\sigma_k^2}$.

3.3.2 Iterative Algorithm Based on Channel Statistics

Since the terms in the SINR are only related to the channel statistics, we can optimize the power allocations at a much slower time scale.

Let $p_{mk}, k \in \mathcal{K}, m \in \mathcal{M}$, denote the power allocation to be optimized. Then the problem can be reformulated as

$$\begin{aligned} \underset{p_{mk}}{\text{maximize}} \quad & R_{sum} = \sum_{k=1}^{K} \log(1+\gamma_k) \\ & \sum_{k \in \pi_m} p_{mk} \leq P_m, \forall m \in \mathcal{M} \\ & \gamma_k = \frac{\sum\limits_{m \in \rho_k} c_{mk} p_{mk}}{\sum\limits_{j \neq k} \sum\limits_{\bar{m} \in \rho_j} c_{\bar{m}kj} p_{\bar{m}j} + 1} \end{aligned} \tag{3.22}$$

The above problem is non-convex. It can not be easily solved in closed form. Since the rate of a user is determined by the signals of all the other users, the optimization is quite complicated due to the coupling between the users. If the power of one user is changed, the power levels of other users should be altered to get the optimal performance. We try to solve this problem in an iterative manner. The key idea behind the proposed optimization algorithm is that we first fix the interference from other users to user k and then update the power level for user k satisfying certain zero gradient conditions. We do it for all the other users in the same way. Then the interference levels of all users change. We compute the power levels again until they converges.

The lagrangian of problem (3.22) can be given as

$$\begin{aligned} \mathcal{L}(p_{mk}, \lambda) = \sum_{k=1}^{K} & (\log(1 + \frac{\sum\limits_{m \in \rho_k} c_{mk} p_{mk}}{\sum\limits_{j \neq k} \sum\limits_{\bar{m} \in \rho_j} c_{\bar{m}kj} p_{\bar{m}j} + 1}) \\ & - \sum_{m \in \mathcal{M}} \lambda_m (P_m - \sum_{k \in \pi_m} p_{mk}), \end{aligned} \tag{3.23}$$

where λ_m is the nonnegative Lagrange multiplier associated with the power constraint for each BS. Then the dual problem is given as,

$$\underset{\lambda}{\text{minimize}} \quad g(\lambda), \tag{3.24}$$

where $g(\lambda)$ is the Lagrange dual function, defined as,

$$g(\lambda) = \underset{p_{mk}}{\text{maximize}} \quad \mathcal{L}(p_{mk}, \lambda) \tag{3.25}$$

After differentiating the Lagrangian function with respect to p_{mk}, we obtain

$$\begin{aligned}
&\frac{\partial \mathcal{L}(p_{mk}, \lambda)}{\partial p_{mk}} = \\
&\frac{1}{c_{mk}p_{mk} + \sum\limits_{\bar{m}\in\rho_k/m} c_{\bar{m}k}p_{\bar{m}k} + \sum\limits_{j\neq k, j\in\mathcal{K}} \sum\limits_{\bar{m}\in\rho_j} c_{\bar{m}kj}p_{\bar{m}j} + 1} c_{mk} \\
&+ \sum_{j\neq k} \frac{1}{\sum\limits_{i\in\mathcal{K}} \sum\limits_{\bar{m}\in\rho_i} c_{\bar{m}kj}p_{\bar{m}i} + 1} c_{\bar{m}kj} - \lambda_m
\end{aligned} \tag{3.26}$$

Here, we make high SINR approximations in the process of differentiation regarding the Lagrangian function.

The optimal solution to the power allocation should satisfy the following conditions:

1. $\lambda_m(P_m - \sum\limits_{k\in\pi_m} p_{mk}) = 0, \forall m \in \mathcal{M}$
2. $\lambda_m \geq 0, \sum\limits_{k\in\pi_m} p_{mk} \leq P_m, \forall m \in \mathcal{M}$
3. $\frac{\partial \mathcal{L}(p_{mk}, \lambda)}{\partial p_{mk}} = 0$

Two auxiliary variables are introduced here,

$$t_{mk} = \sum_{j\neq k, j\in\mathcal{K}} \frac{1}{\sum\limits_{i\in\mathcal{K}} \sum\limits_{\bar{m}\in\mathcal{M}_i} c_{\bar{m}kj}p_{\bar{m}i} + 1} c_{\bar{m}kj} \tag{3.27}$$

$$r_{mk} = (\sum_{\bar{m}\in\rho_k/m} c_{\bar{m}k}p_{\bar{m}k} + \sum_{j\neq k, j\in\mathcal{K}} \sum_{\bar{m}\in\rho_j} c_{\bar{m}kj}p_{\bar{m}j} + 1)/c_{mk} \tag{3.28}$$

To update the power p_{mk}, we fix the power of all other users. We can rewrite $\frac{\partial\mathcal{L}(p_{mk},\lambda)}{\partial p_{mk}} = 0$ as $-t_{mk} + \lambda_m = 1/\ln 2/(p_{mk} + r_{mk})$.

Then we obtain

$$p_{mk} = [\frac{1/\ln 2}{t_{mk} + \lambda_m} - r_{mk}]^+ \tag{3.29}$$

The problem can be solved by iteratively updating the dual variable λ_m according to the following steps: With fixed λ_m, find the optimal power allocation p_{mk} (3.29) using zero gradient condition. With the power allocation p_{mk}, update the lagrange multiplier λ_m, t_{mk} and r_{mk}.

The iterative method can be described in **Algorithm 1** shown in the following page.

3.4 Numerical Results and Discussion

In this section, numerical results are provided to illustrate the performance of large dimensional optimization framework for the cell free network model. We consider a network model with three BSs and 30 users. Each BS is equipped with $N = 200$ antennas. We first show that the power optimization based on channel statistics can achieve performance close to that based on actual channel realizations. Then we will demonstrate that the user centric cell free model can efficiently boost the performance and maintain fairness of the network at the same time.

The simulation parameters are presented in Table 3.1. The path-loss model from the BSs to the users is assumed to be

Algorithm 1 large dimensional iterative algorithm based on channel statistics

1. For fixed precoding vector, initialize the power for all the

   vectors as equal power allocation $p_{mk} = \frac{P_m}{|\pi_m|}$.

   set $\lambda_m = \theta/2$, $[a_m, b_m] = [0, \theta]$

2. Update t_{mk} and r_{mk} according to (3.27) and (3.28)

3. Given t_{mk} and r_{mk}, calculate the new p_{mk}, given by

   $p_{mk} = [\frac{1/\ln 2}{t_{mk}+\lambda_m} - r_{mk}]^+$.

4. Update λ_m, $\lambda_m = (a_m + b_m)/2$

5. Calculate $\hat{P}_m$, $\hat{P}_m = \sum_{k \in \pi_m} p_{mk}$

6. if $\hat{P}_m > P_m$, then update a_m, $a_m = \lambda_m, \forall m \in \mathcal{M}$

7. if $\hat{P}_m < P_m$, then update b_m, $b_m = \lambda_m, \forall m \in \mathcal{M}$

8. Update until $\hat{P}_m$ converge

$$L(dB) = 15.3 + 37.6 \log_{10} d \tag{3.30}$$

where d represents the distance between BS and users in meters. And the log-normal shadowing is set to be with standard deviation of 10 dB.

Table 3.1. Simulation Parameters

Number of base stations	3
Number of users	30
Number of antennas in base stations	200
Number of antennas per user	1
Log-normal shadowing	10dB

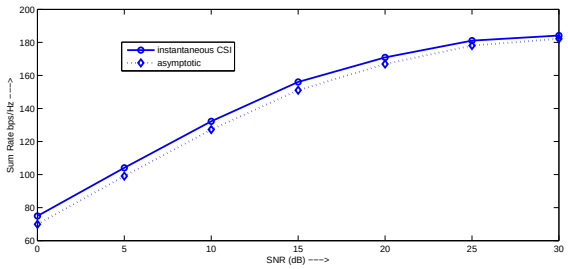

Figure 3.3. Asymptotic optimality: comparison of the optimization using channel statistics and that based on actual channel realizations .

3.4.1 Convergence and Asymptotic Optimality

We leverage the results in RMT to derive the deterministic equivalents for the SINR of users, as shown in (3.17). It is even very accurate in the finite regime. The SINR is only related with large scale channel statistics in the large dimensional regime. Then power allocation computed in **Algorithm 2** is based on the deterministic equivalents in (3.15)(3.16) irrespective of the actual channel realizations. In Fig. 3.3, we compare the optimization based on channel statistics and that based on instantaneous CSI. It can be observed that there is only marginal performance loss for the asymptotic case. The optimization based on the asymptotic results is close to that based on actual channel realizations, which verifies the accuracy of the large dimensional approximations.

We also considered the computational complexity of the optimization method based on channel statistics. Since the channel statistics varies at a much slower time scale compared with the fast fading. So the optimization does not need to be updated as frequently as the fast fading. In this way, much computational overhead can be saved.

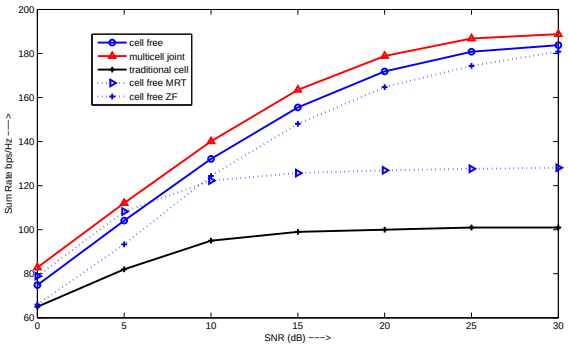

Figure 3.4. Performance comparison between three different network paradigms: multicell joint transmission, traditional cellular network and cell free network .

3.4.2 Performance Comparison

The performances of three different network paradigms are compared as follows:

- Multicell joint transmission: All the BSs cooperate as a super BS serving all the users simultaneously. The transmission is performed based on global CSI and data sharing.
- Traditional cellular network: Each user is served by the BS in its own cell. The BSs individually design the precoding vectors.
- Cell free network: It is our proposed network model. Each user is not restricted to be served by only one BS. It can be served by multiple BSs at the same time.

To keep a fair comparison for the three networks, the number of BSs and users are set to be the same. The power constraints are also the same for each BS.

Observe Fig. 3.4, our proposed algorithm based on cell free model has a rate increase with SNR similar to multicell joint transmission across all SNR regime due to effective interference and signal power balancing. The traditional cellular networks suffer from severe intercell interference in the high SNR. We also consider the cell free MRT and cell free ZF methods. From the figure, the MRT works well in low SNR, while the ZF can achieve performance close to our proposed algorithm in high SNR.

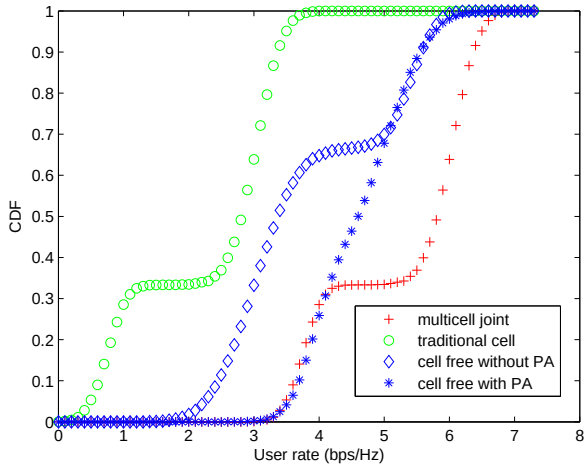

Figure 3.5. Cumulative distribution function (CDF) of users' rates: multicell joint transmission, traditional cellular network, cell free network with and without power allocation.

3.4.3 Network Fairness

Fig. 3.5 shows the cumulative distribution function (CDF) of users' rates for different methods: multicell joint transmission, traditional cellular network, cell free network with and without power allocation. The figure is obtained by evaluating users' rate over 20000 independent channel realizations at a SNR of 15 dB.

3.4.4 Intuitions and Insights for the Cell Free Model

We compared the performance of cell free network with that of the traditional cellular networks. In traditional cellular networks, each user is served by only one BS. The cell edge users would suffer from severe intercell interference. While in cell free networks, users can adaptively associate with multiple BSs, thus it can convert the otherwise interference into signal power. The cell free model provide some kind of flexibility for handling the intercell interference.

When the number of antennas deployed in the BS goes to infinity, with the number of users fixed, the intuition for the association rule is that the users should be associated with the nearest BS to achieve the best sum rate performance. Since the channels between difference users are asymptotically orthogonal, the interference terms in the SINR expression are all zeros. So we just allocate more transmit power to the users with large channel gains. However, this will lead to poor fairness performance. In the finite system regime, when the interference and signal power in the design of precoding is well balanced, we can efficiently exploit the flexibility of this cell free model. The large dimensional approximation can be accurate in the finite regime, thus facilitating the optimization of network. With proper power allocation, much higher rate gain can be obtained and good fairness can be guaranteed.

3.5 Conclusions

We investigated the rate optimization problem under the cell free network model in which users can be served by multiple nearby BSs. Due to coupled nature of interference, it is difficult to obtain the global optimal solution for the precoding. The complexity of problem would be prohibitively high to get the optimal solution. Thus, a heuristic precoding method that balances interference and signal is adopted. Then we solve the power allocation problem under the framework of large dimensional optimization. The SINR of the users can be approximated as deterministic equivalents that depend on statistical CSI in the massive MIMO scenario. Since the optimization is performed based only on large scale channel statistics, it does not need to adapt to the actual channel realization, which would greatly reduce the computational overhead. Numerical results show that the cell free network model can efficiently enhance the network rate performance and maintain good fairness among users.

CHAPTER 4

Partial Interference Alignment in Finite SNR Via Large Dimensional Analysis

The ever-growing demands for higher data rates has posed a big challenge for next generation wireless systems (5G) [71] [72]. It is predicted that the network capacity would be increased by a thousand times over the next decade. The current cellular network structure obviously can not support such high transmission capacity. Many candidate technologies have been proposed for 5G, among which massive MIMO and heterogeneous cellular networks (Hetnets) [73] [74] are two different potential approaches. Massive MIMO scales up the number of antennas equipped at BSs, which is much larger than that of users. It can improve the system performance in terms of both spectral and energy efficiency [75] [76] [77]. Small cell networks deploy dense low-power small cell base stations for the data offloading from the macro cell.

Traditional approach for the coexistence of the two-tiered network is to make the macro cell base stations and small cell base stations operate on disjoint channels [78] [79], so as to avoid mutual interference between the two tiers. However, this method results in low spectral efficiency. To enable maximum spectral reuse, the complete spectrum sharing where the macro base stations and small cell base station operate on the same channels is the most attractive way [80]. The cross-tier interference has been identified as the main limiting factor for the macro cell and small cell coexistence in this case. In this paper, we consider the open subscriber group (OSG) mode in Hetnet in which the dominant interferences in macro cell-small cell networks are the interference from macrocell BS to small cell users and the interference between small cell BS and the users of neighboring small cells.

Recently, significant research efforts have been dedicated to the study of macro cell and small cell coexistence via interference alignment (IA) [81] [82] [83]. IA is a signal processing approach that attempts to simultaneously align the interference on a lower dimension subspace at the receiver so that the desired signals can be transmitted on the interference-free dimensions. It can guarantee virtual interference free transmission by exploiting the spatial dimensions of MIMO channels, which is very suitable for the complete spectrum sharing scenario in Hetnet. Although the small cell and macro cell operate on the same channels, they can exploit additional space dimensions for interference free transmission [84] [85]. Also, IA based methods for interference management in cognitive radio network have been investigated in [86] [87] [88]. In [89] and [90], the authors show that IA is optimal in degrees-of-freedom (DoF) sense. IA can achieve the optimal DoF in high SNR. However, it is not necessarily optimal in terms of achievable sum capacity since the signal projection in the receiver side limits the spatial diversity of the system.

A natural question arises that when the network operate in the intermediate SNR regime, how can we maximize the system sum capacity? When it comes to intermediate SNR, there is a tradeoff in designing the precoding matrix. Since there is power imbalance between the macro base stations and small cell base stations, we allow some interference leakage caused to the small cell user to optimize the overall network performance. We first propose an interference coordination strategy based on the idea of interference alignment, which is called hard partial interference alignment. It adaptively selects the dimensions of subspace the small cell user can transmit based on the power constraints. This partial interference alignment captures the trade-off between interference avoidance at other users and spatial multiplexing at the intended user. We try to project the interference into partial subspace of the small cell user instead of the full subspace, so that extra improvement on the transmission rate of

macro cell can be offered in finite SNR. After the subspaces of the small cell users are determined, the soft partial interference alignment method is used to optimize the sum capacity by considering the signal and interference balancing. It is solved in an iterative way. The performance will be further improved when the water-filling (WF) method is used. The water-filling level is determined by the singular value distributions of the channel matrix in the large dimensional regime instead of the actual realization of the channels.

DoF provide SNR-asymptotic performance for the wireless systems. For the finite SNR cases, large dimensional analysis [91] [92] [93] provide a size-asymptotic approach to quantify the network performance. Size-asymptotic approach is a viable alternative to the SNR-asymptotic one, since it produces the results that hold at any realistic SNR. It has been applied to derive the achievable rates of the successive interference cancelation in MIMO broadcast channel [94]. Also large dimensional analysis is performed to analyze the rates of block diagonalization as well as dirty paper coding [95]. When the number of transmit and receive antennas goes to infinity at a fixed ratio, it will render us to perform closed form analysis by using the tools in random matrix theory [91]. In this paper, we perform large dimensional analysis to obtain the the number of transmit dimensions for small cell in finite SNR. We show that the number of dimensions the small cell can transmit is determined by the singular value distribution of the channel matrix and the transmit SNR of the system.

The rest of this chapter is organized as follows. The Hetnet downlink transmission model and the interference channel model are introduced in Section 4.1. A new interference coordination strategy called partial interference alignment is proposed and analyzed in Section 4.2. In Section 4.3, the large dimensional asymptotic analysis is performed to obtain the number of transmit dimensions for small cell user and the waterfilling levels. The numerical results and discussion are provided in Section

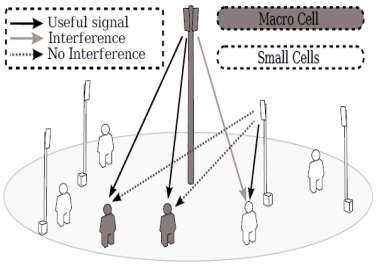

Figure 4.1. Downlink transmission in Hetnet. .

4.4 and Section 4.5 concludes this paper.

4.1 Network Model and Problem Formulation

4.1.1 OSG Cross-tier Interference in Downlink Heterogeneous Cellular Networks

We consider a two-tiered downlink Hetnet which consists of a macro cell and some small cell based stations (BS), as shown in Fig. 4.1. The macro cell BS is equipped with N antennas and the small cell BSs are equipped with M antennas. For the users, they all have K antennas. Due to power imbalance in heterogeneous networks, the transmitting power of macro cell BS is much larger than that of the small cell. Since the macrocell user is geographically far away from the small cell BS, the interference experienced by macro cell user from the small cell BS can be reasonably neglected. On the other hand, the interference from macro BS to small cell user must be properly mitigated to avoid a severe degradation of the network performance.

The Hetnet model can be abstracted as the interference network model, as shown in Fig. 4.2, where one macro cell BS, MC 0 coexists with two small cell BSs, SC 1 and SC 2. The red line denotes interference and the dark line represents intended signal. The interference mode: open subscribe group (OSG), where macro

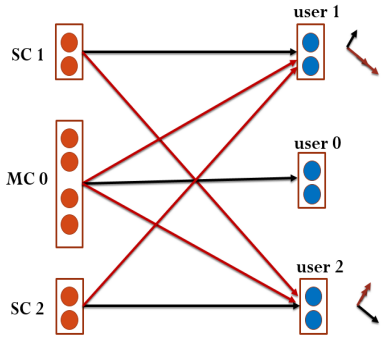

Figure 4.2. OSG interference mode in Hetnet. .

users receive negligible interference from small cell BSs due to range expansion is considered. The extension to more complex model is straightforward. The channel matrices between BS and user end are assumed to be constant for certain number of time slots and do not change during transmission.

The received signal in user 0, user 1 and user 2 can be expressed as $K \times 1$ vectors $\mathbf{r}_0$, $\mathbf{r}_1$ and $\mathbf{r}_2$,

$$\mathbf{r}_1 = \mathbf{H}_{11}\mathbf{V}_1\mathbf{P}_0\mathbf{s}_1 + \mathbf{H}_{12}\mathbf{V}_2\mathbf{P}_2\mathbf{s}_2 + \mathbf{H}_{10}\mathbf{V}_0\mathbf{P}_0\mathbf{s}_0 + \mathbf{n}_1 \tag{4.1}$$

$$\mathbf{r}_2 = \mathbf{H}_{21}\mathbf{V}_1\mathbf{P}_0\mathbf{s}_1 + \mathbf{H}_{22}\mathbf{V}_2\mathbf{P}_2\mathbf{s}_2 + \mathbf{H}_{20}\mathbf{V}_0\mathbf{P}_0\mathbf{s}_0 + \mathbf{n}_2 \tag{4.2}$$

$$\mathbf{r}_0 = \mathbf{H}_{00}\mathbf{V}_0\mathbf{P}_0\mathbf{s}_0 + \mathbf{n}_0 \tag{4.3}$$

where $\mathbf{H}_{00}$, $\mathbf{H}_{11}$ and $\mathbf{H}_{22}$ are intended signal channel matrices of user 0, user 1 and user 2. $\mathbf{H}_{10}$, $\mathbf{H}_{20}$, $\mathbf{H}_{21}$ and $\mathbf{H}_{12}$ are cross interference channel matrices. The entries of these matrices are independent and identically distributed, and can be expressed as

$$\mathbf{H}_{ij} = \mathbf{D}_{ij}^{\frac{1}{2}}\mathbf{G}_{ij}, \tag{4.4}$$

where $\mathbf{G}_{ij}$ denotes the small scale fading part of the channel from the jth BS to user i. It is assumed to be fixed over one coherence time, and it is complex Gaussian distributed. $\mathbf{D}_{ij}$ is the corresponding path loss part, which lasts over a relatively longer time. $\mathbf{n}_1$, $\mathbf{n}_2$ and $\mathbf{n}_3$ are additive white Gaussian noise vector with variance σ^2. The data symbols $\mathbf{s}_k \in d_k \times 1$ is intended for user k with d_k as the dimensions or the degrees of freedom achieved by user k. $\mathbf{P}_k$ is a diagonal matrix, denoted as $\mathbf{P}_k = \text{diag}(\sqrt{p_{k1}}, \ldots, \sqrt{p_{kd_k}}),\ k = 0, 1, 2$. Each entry in the diagonal can be seen as the power allocated to each corresponding symbol.

We also consider the precoding in concatenated with the power allocation. Each symbol is precoded with a precoder $\mathbf{V}_k$, which is unitary matrix and it is responsible for the symbol direction. The corresponding power constraint can be expressed as following,

$$\text{Tr}(\mathbf{V}_1\mathbf{P}_1\mathbf{V}_1^H) \leq p_{1,max} \tag{4.5}$$

$$\text{Tr}(\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H) \leq p_{2,max} \tag{4.6}$$

$$\text{Tr}(\mathbf{V}_0\mathbf{P}_0\mathbf{V}_0^H) \leq p_{0,max} \tag{4.7}$$

Remark 2 (Realistic SNR and power imbalance)**.** *Note that, here we consider finite SNR in this paper, so the power constraint is limited. Also considering the power imbalance in hetnet, we set* $p_{0,max} > p_{1,max}$ *.*

4.1.2 Problem Formulation

After the design of the interference avoidance and alignment matrix, the achievable rate of user k can be expressed as

$$R_k = \text{logdet}[\mathbf{I}_{d_k} + \mathbf{W}_k^H \mathbf{H}_{kk} \mathbf{V}_k \mathbf{V}_k^H \mathbf{H}_{kk}^H \mathbf{W}_k (\mathbf{W}_k \mathbf{Q}_k \mathbf{W}_k^H)^{-1}] \tag{4.8}$$

where $\mathbf{W}_k$ denotes the receiving matrix of user k, and $\mathbf{Q}_k$ is

$$\mathbf{Q}_k = \sum_{l \neq k} \mathbf{H}_{kl} \mathbf{V}_l \mathbf{P}_l^2 \mathbf{V}_l^H \mathbf{H}_{kl}^H + \sigma^2 \mathbf{I} \tag{4.9}$$

Since we consider the intermediate SNR case, it is necessary for us to maximize the sum capacity instead of the DoF. We aim to maximize the achievable network sum capacity subject to the power constraint, written as follows:

$$\begin{aligned} &\underset{\mathbf{P}_i, \mathbf{V}_i}{\text{maximize}} \quad R_{sum} = R_0 + R_1 + R_2 \\ &\text{subject to} \quad p_1 \leq p_{1,max}, \quad p_2 \leq p_{2,max}, \quad p_0 \leq p_{0,max} \end{aligned} \tag{4.10}$$

where R_0, R_1 and R_2 are the achievable rates of user 0, user 1 and user 2.

4.2 Interference Coordination Strategy at Intermediate SNR

In this section, we try to propose an interference coordination strategy for the coexistence of small cell and macro cell downlink transmissions based on the idea of interference alignment. The only difference is that since we consider the intermediate SNR case, we need to balance the the transmission dimensions of small cell users and macro cell user in order to maximize the sum capacity of the two-tier networks. We first review the conventional full interference alignment strategy, which can provide the best DoF performance in high SNR region. Since the intermediate SNR case is considered in this paper, there are power constraints in the transmitter sides of the small cell and macro cell. The transmit dimensions of small cell users are adaptively

selected depending on the power constraints and channel conditions. Further, after the dimensions of the user is chosen, balancing of desired signal maximization and interference alignment is considered in the design of precoding and receiving matrices. In this way, the network sum capacity can be maximized with limited power constraints.

4.2.1 Design of Full Interference Alignment Strategy

For small cell users, user 1 and user 2, there are cross small cell interference and interference from the macro cell BS. In light of the conventional interference alignment strategy, we should align the cross small cell interference and macro cell interference in each small cell receiver side, that is

$$\text{span}(\mathbf{H}_{10}\mathbf{V}_0) = \text{span}(\mathbf{H}_{12}\mathbf{V}_2) \tag{4.11}$$

$$\text{span}(\mathbf{H}_{20}\mathbf{V}_0) = \text{span}(\mathbf{H}_{21}\mathbf{V}_1) \tag{4.12}$$

where (4.11) aligns the interference caused in the user 1 side from transmitter 0 and transmitter 2, (4.12) aligns the interference caused in the user 2 side from transmitter 0 and transmitter 1.

For the full interference alignment strategy, the receiving matrix $\mathbf{V}_k \in M \times d_k$ is designed to eliminate the interference and extract the intended symbols, with $d_k = K$. The estimates of the data symbols can be obtained as

$$\hat{s}_k = \mathbf{W}_k^H r_k = \mathbf{W}_k^H(\mathbf{H}_{kk}\mathbf{V}_k\mathbf{P}_k\mathbf{s}_k + \sum_{l \neq k} \mathbf{H}_{kl}\mathbf{V}_l\mathbf{P}_l\mathbf{s}_l + \mathbf{n}_k) \tag{4.13}$$

In the conventional IA strategy, the following conditions must hold:

$$\mathbf{W}_k\mathbf{H}_{kl}\mathbf{V}_l = \mathbf{0}_{d\times d}, \forall l \neq k, \text{rank}(\mathbf{W}_k\mathbf{H}_{kk}\mathbf{V}_k) \geq d_k. \tag{4.14}$$

After the design of the interference avoidance and alignment matrix, the achievable rate of user k can be expressed as

$$R_k = \text{logdet}(\mathbf{I}_{d_k} + \mathbf{W}_k^H \mathbf{H}_{kk} \mathbf{V}_k \mathbf{V}_k^H \mathbf{H}_{kk}^H \mathbf{W}_k) \tag{4.15}$$

For the system model introduced in this paper, to achieve full interference alignment, the constraint that $N - 2K \geq K$ is assumed. Since the precoding and receiving matrices for maximizing the system sum capacity are NP-hard to obtain, as an alternative to finding optimal closed-form solutions, we provide an intuitive method called partial interference alignment to solve the problem. An iterative algorithm that balances the interference and signal by taking into account of the power imbalance in the Hetnet model is proposed to further optimize the network sum capacity.

In this paper, we decompose the precoding vector into two parts, which represent the spatial multiplexing at the intended user and interference avoidance at the other users. We first consider the macro cell link. For the transmitter side of user 0, the design of precoding matrix should take into account of the interference caused to the small cell users. This can be achieved by writing $\mathbf{V}_0$ as the product of two sub precoding matrices.

$$\mathbf{V}_0 = \mathbf{V}_0^{(1)} \mathbf{V}_0^{(2)} \tag{4.16}$$

where $\mathbf{V}_0^{(1)}$ is responsible for the interference avoidance at the small cell user and $\mathbf{V}_0^{(2)}$ is designed to maximize the multiplexing gain for the macro cell user.

The interference channels from MC 0 to user 1 and user 2 can staked as

$$\mathbf{H}_{210} = \begin{pmatrix} \mathbf{H}_{10} \\ \mathbf{H}_{20} \end{pmatrix} \tag{4.17}$$

where $\mathbf{H}_{210} \in 2K \times N$ due to the fact that $\mathbf{H}_{10}, \mathbf{H}_{20} \in K \times N$. Let $\mathbf{H}_{210} = \mathbf{U}_{\mathbf{H}_{210}} \, \mathbf{\Lambda}_{\mathbf{H}_{210}} \, \mathbf{V}^H_{\mathbf{H}_{210}}$ be the singular value decomposition of the interference channel matrix from MC 0 to user 1 and user 2, in which $\mathbf{V}_{\mathbf{H}_{210}} \in N \times N$. The first part of the precoding matrix of MC 0 $\mathbf{V}_0^{(1)}$ can be designed to be orthogonal to the left $2K$ columns of $\mathbf{V}_{\mathbf{H}_{210}}$, that is

$$\mathbf{V}_0^{(1)} = (\mathbf{I} - \mathbf{V}_{\mathbf{H}_{210}} \mathbf{V}^H_{\mathbf{H}_{210}}) \tag{4.18}$$

In this way, the interference caused to user 1 and user 2 can be eliminated. From this, we can see that some degree of freedom in the downlink transmission of the macro cell link is sacrificed for the interference avoidance at the small-cell users.

After the $\mathbf{V}_0^{(1)}$ is obtained, we next consider the design of $\mathbf{V}_0^{(2)}$. We can write the *equivalent* macro cell link channel matrix as $\widetilde{\mathbf{H}}_{00} = \mathbf{H}_{00}\mathbf{V}_0^{(1)}$. Then let $\widetilde{\mathbf{H}}_{00} = \mathbf{U}_{\widetilde{\mathbf{H}}_{00}} \, \mathbf{\Lambda}_{\widetilde{\mathbf{H}}_{00}} \, \mathbf{V}^H_{\widetilde{\mathbf{H}}_{00}}$, the $\mathbf{V}_0^{(2)}$ can be designed as

$$\mathbf{V}_0^{(2)} = \mathbf{V}_{\widetilde{\mathbf{H}}_{00}}(:, 1:K) \tag{4.19}$$

and the corresponding receiving matrix can be given as

$$\mathbf{W}_0 = \mathbf{U}_{\widetilde{\mathbf{H}}_{00}}(:, 1:K) \tag{4.20}$$

For the small-cell link, the precoding matrix can be designed in the same way. For the transmitter side of user 1, the design of precoding matrix should take into account of the interference caused to the small-cell user 2. This can be achieved by design $\mathbf{V}_1$ as

$$\mathbf{V}_1 = \mathbf{V}_1^{(1)}\mathbf{V}_1^{(2)} \tag{4.21}$$

Let $\mathbf{H}_{21} = \mathbf{U}_{\mathbf{H}_{21}}\ \mathbf{\Lambda}_{\mathbf{H}_{21}}\ \mathbf{V}_{\mathbf{H}_{21}}^H$ be the singular value decomposition of the interference channel matrix from SC 1 to user 2, in which $\mathbf{V}_{\mathbf{H}_{21}} \in M \times M$. The first part of the precoding matrix of SC 1 $\mathbf{V}_1^{(1)}$ can be designed by selecting the right $\min(K, M-K)$ columns of $\mathbf{V}_{\mathbf{H}_{21}}$. In this way, the interference caused to user 2 can be eliminated.

After the $\mathbf{V}_1^{(1)}$ is obtained, we next consider the design of $\mathbf{V}_1^{(2)}$. We can write the *equivalent* small-cell link channel matrix as $\widetilde{\mathbf{H}}_{11} = \mathbf{H}_{11}\mathbf{V}_1^{(1)}$. Then let $\widetilde{\mathbf{H}}_{11} = \mathbf{U}_{\widetilde{\mathbf{H}}_{11}}\ \mathbf{\Lambda}_{\widetilde{\mathbf{H}}_{11}}\ \mathbf{V}_{\widetilde{\mathbf{H}}_{11}}^H$, the $\mathbf{V}_1^{(2)}$ can be designed as

$$\mathbf{V}_1^{(2)} = \mathbf{V}_{\widetilde{\mathbf{H}}_{11}} \tag{4.22}$$

and the corresponding receiving matrix can be given as

$$\mathbf{W}_1 = \mathbf{U}_{\widetilde{\mathbf{H}}_{11}}^H \tag{4.23}$$

The rational for the design of the precoding for the full interference alignment scheme is that the nulled subspace of $\mathbf{H}_{210}$ is not likely to be correlated with $\mathbf{H}_{00}$.

In the following sections, we try to find the optimal precoding matrix when the transmitting power is limited, i.e. the intermediate SNR case.

4.2.2 Hard Partial Interference Alignment and selective dimensions of small cell user

From the operation of full Interference Alignment, it can be seen that the macro cell link has to sacrifice some dimensions to mitigate the interference caused to the small cell link. Since there is power imbalance among the the macro-cell BS and the small-cell BS, we can adaptively select the number of dimensions for the small cell links according to the power constraints of the transmitters.

To achieve the optimal sum capacity of the system, we try to partially align the signal of interference by considering the power constraints in each transmitter instead of full IA. Without loss of generality, we assume the degree of freedom achieved by user k is $d_k < K, k = 1, 2$.

After we project the precoding vector $\mathbf{V}_0$ into the null space of $\mathbf{H}_{210}$, we would sacrifice some degree of freedom in the downlink transmission of the macro cell BS.

We can select the subspace spanned by the first d_1 column vectors of $\mathbf{V}_{\mathbf{H}_{11}}$ corresponding to the largest d_1 singular values in the small cell link. Then, in order to align the interference caused by the macro cell link into the rest $K - d_1$ dimensions, we consider the interference channel matrix $\widetilde{\mathbf{H}}_{10} = \mathbf{U}_{\mathbf{H}_{11}}\mathbf{H}_{10}$ and $\widetilde{\mathbf{H}}_{20} = \mathbf{U}_{\mathbf{H}_{22}}\mathbf{H}_{20}$, which can be expressed as the block structure

$$\widetilde{\mathbf{H}}_{10} = \begin{pmatrix} \widetilde{\mathbf{H}}_1 \\ \widetilde{\mathbf{H}}_2 \end{pmatrix} \tag{4.24}$$

$$\widetilde{\mathbf{H}}_{20} = \begin{pmatrix} \widetilde{\mathbf{H}}_3 \\ \widetilde{\mathbf{H}}_4 \end{pmatrix} \tag{4.25}$$

where $\widetilde{\mathbf{H}}_{10}$ is divided into $d_1 \times N$ $\widetilde{\mathbf{H}}_{10}$ and $(K - d_1) \times N$ $\widetilde{\mathbf{H}}_2$ and $\widetilde{\mathbf{H}}_{20}$ is divided into $d_2 \times N$ $\widetilde{\mathbf{H}}_3$ and $(K - d_2) \times N$ $\widetilde{\mathbf{H}}_4$.

Since there is power imbalance in the transmitting power of macro-cell BS 0 and small-cell BS 1 and 2, a heuristic way to do this is to design the precoding matrix $\mathbf{V}_0$ to be orthogonal to the partial subspace of $\widetilde{\mathbf{H}}_{10}$ and $\widetilde{\mathbf{H}}_{20}$. In this way, more transmit dimensions can be reserved for the macro-cell link.

To align the interference caused by the macro link into the $d_1 + d_2$ dimensions, the precoding matrix $\mathbf{V}_0^{(1)}$ must satisfy:

$$\begin{pmatrix} \widetilde{\mathbf{H}}_1 \\ \widetilde{\mathbf{H}}_3 \end{pmatrix} \mathbf{V}_0^{(1)} = \widetilde{\mathbf{H}}_{13} \mathbf{V}_0^{(1)} = \mathbf{0} \tag{4.26}$$

Let $\widetilde{\mathbf{H}}_{13} = \mathbf{U}_{\widetilde{\mathbf{H}}_{13}} \, \mathbf{\Lambda}_{\widetilde{\mathbf{H}}_{13}} \, \mathbf{V}_{\widetilde{\mathbf{H}}_{13}}^H$, and let $\hat{\mathbf{V}}_{\widetilde{\mathbf{H}}_{13}}$ be the first d columns of $\mathbf{V}_{\widetilde{\mathbf{H}}_{13}}$. Then $\mathbf{V}_0^{(1)}$ can be obtained by projected into the null space of $\hat{\mathbf{V}}_{\widetilde{\mathbf{H}}_{13}}$,

$$\mathbf{V}_0^{(1)} = (\mathbf{I} - \hat{\mathbf{V}}_{\widetilde{\mathbf{H}}_{13}} \hat{\mathbf{V}}_{\widetilde{\mathbf{H}}_{13}}^H) \tag{4.27}$$

Note that the $\hat{\mathbf{V}}_{\widetilde{\mathbf{H}}_{13}} \hat{\mathbf{V}}_{\widetilde{\mathbf{H}}_{13}}^H$ is non-identity when we select part of the columns as the precoding vector, which is an effective one. And it is identity when we select the full columns. So the precoding matrix (4.27) is effective here.

We next consider the design of $\mathbf{V}_1$. It should cancel the interference caused to user 2. in order to align it with the interference caused by the macro cell link to the rest $K - d_2$ dimensions of user 2, we consider the channel matrix $\widetilde{\mathbf{H}}_{21} = \mathbf{U}_{\mathbf{H}_{22}} \mathbf{H}_{21}$, which can be expressed as the block structure

$$\widetilde{\mathbf{H}}_{21} = \begin{pmatrix} \widetilde{\mathbf{H}}_5 \\ \widetilde{\mathbf{H}}_6 \end{pmatrix} \tag{4.28}$$

where $\widetilde{\mathbf{H}}_{21}$ is divided into $d_2 \times N$ $\widetilde{\mathbf{H}}_5$ and $(K - d_2) \times N$ $\widetilde{\mathbf{H}}_6$

Then $\mathbf{V}_1^{(1)}$ can be obtained by projected into the null space of $\hat{\mathbf{V}}_{\widetilde{\mathbf{H}}_5}$, We can write the *equivalent* macro cell link channel matrix as $\widetilde{\mathbf{H}}_{11} = \mathbf{H}_{11}(\mathbf{I} - \hat{\mathbf{V}}_{\widetilde{\mathbf{H}}_5} \hat{\mathbf{V}}_{\widetilde{\mathbf{H}}_5}^H)$. Then let $\widetilde{\mathbf{H}}_{11} = \mathbf{U}_{\widetilde{\mathbf{H}}_{11}} \, \mathbf{\Lambda}_{\widetilde{\mathbf{H}}_{11}} \, \mathbf{V}_{\widetilde{\mathbf{H}}_{11}}^H$, the $\mathbf{V}_1^{(2)}$ can be designed as

$$\mathbf{V}_1^{(2)} = \mathbf{V}_{\widetilde{\mathbf{H}}_{11}} \tag{4.29}$$

and the corresponding receiving matrix can be given as

$$\mathbf{W}_1 = \mathbf{U}_{\widetilde{\mathbf{H}}_{11}}^H \tag{4.30}$$

Then the balanced precoding vector $\mathbf{V}_1$ can be expressed as

$$\mathbf{V}_1 = \mathbf{V}_1^{(1)}\mathbf{V}_1^{(2)} = (\mathbf{I} - \hat{\mathbf{V}}_{\widetilde{\mathbf{H}}_5}\hat{\mathbf{V}}_{\widetilde{\mathbf{H}}_5}^H)\mathbf{V}_{\widetilde{\mathbf{H}}_{11}} \tag{4.31}$$

This precoding balances the multiplexing gain at the macro link and interference avoidance at the small cell user. Also it adaptively select the number of dimensions the small cell link can transmit based on the power constraint of the system, which flexibly takes advantage of the interference alignment method.

4.2.3 Soft Partial Interference Alignment and Iterative Algorithm

In this section, we propose a soft IA algorithm in order to further improve the hard IA method. Instead of strictly aligning the interference in the transmitter side, we balance the interference and useful signals in each user side to improve the sum capacity of the system.

The total received interference at user k from the undesired BSs and noise is given by:

$$I_k = \mathrm{Tr}[\mathbf{W}_k\mathbf{Q}_k\mathbf{W}_k^H] \tag{4.32}$$

where

$$\mathbf{Q}_k = \sum_{l\neq k}\mathbf{H}_{kl}\mathbf{V}_l\mathbf{P}_l^2\mathbf{V}_l^H\mathbf{H}_{kl}^H + \sigma^2\mathbf{I} \tag{4.33}$$

is the interference covariance matrix for user k.

The intended signal power at user k from the desired BSs is given by:

$$S_k = \mathrm{Tr}[\mathbf{W}_k\mathbf{R}_k\mathbf{W}_k^H] \tag{4.34}$$

where

$$\mathbf{R}_k = \mathbf{H}_{kk}\mathbf{V}_k\mathbf{P}_k^2\mathbf{V}_k^H\mathbf{H}_{kk}^H \tag{4.35}$$

is the signal covariance matrix for user k.

By balancing the interference and useful signals, in each small cell user try to solve the following optimization problem,

$$\mathbf{W}_k^* = \arg \max_{\mathbf{W}_k\mathbf{W}_k^H=\mathbf{I}_{d_k}} \frac{\mathrm{Tr}[\mathbf{W}_k\mathbf{R}_k\mathbf{W}_k^H]}{\mathrm{Tr}[\mathbf{W}_k\mathbf{Q}_k\mathbf{W}_k^H]} \tag{4.36}$$

This is a standard trace ratio problem, which can be solved by the iterative optimization procedure. The column vectors of the obtained solution $\mathbf{W}_k^*$ are unitary and orthogonal to each other.

Denote $\mathbf{A}_k = \mathbf{R}_k + \mathbf{Q}_k$, then the optimization problem (4.37) is equivalent to the following,

$$\mathbf{W}_k^* = \arg \max_{\mathbf{W}_k\mathbf{W}_k^H=\mathbf{I}_{d_k}} \frac{\mathrm{Tr}[\mathbf{W}_k\mathbf{R}_k\mathbf{W}_k^H]}{\mathrm{Tr}[\mathbf{W}_k\mathbf{A}_k\mathbf{W}_k^H]} \tag{4.37}$$

where we have $0 \leq \mathrm{Tr}[\mathbf{W}_k\mathbf{R}_k\mathbf{W}_k^H]/\mathrm{Tr}[\mathbf{W}_k\mathbf{A}_k\mathbf{W}_k^H] \leq 1$.

We start to solve the optimization problem in an iterative manner. First compute the trace ratio value λ^n from the previous receiving matrix $\mathbf{W}_k^{n-1}$,

$$\lambda^n = \frac{\mathrm{Tr}[\mathbf{W}_k^{n-1}\mathbf{R}_k(\mathbf{W}_k^{n-1})^H]}{\mathrm{Tr}[\mathbf{W}_k^{n-1}\mathbf{A}_k(\mathbf{W}_k^{n-1})^H]} \tag{4.38}$$

Then the receiving matrix can be updated based on the obtained λ^n as

$$\mathbf{W}_k^n = \arg \max_{\mathbf{W}_k\mathbf{W}_k^H=\mathbf{I}_{d_k}} \mathrm{Tr}[\mathbf{W}_k(\mathbf{A}_k - \lambda^n\mathbf{R}_k)(\mathbf{W}_k)^H] \tag{4.39}$$

which can be solved by the eigenvalue decomposition method by selecting the column vectors corresponding to the largest d_k singular eigenvalues,

$$(\mathbf{A}_k - \lambda^n \mathbf{R}_k) = \mathbf{V}_k \ \mathbf{\Lambda}_k \ \mathbf{V}_k^H \tag{4.40}$$

$$\mathbf{W}_k = \mathbf{V}_k(:, 1 : d_k) \tag{4.41}$$

Similarly, for the precoding matrix $\mathbf{V}_k$, it can be updated in the reciprocal way. Its interference leakage to other users is given by:

$$I_k^{'} = \mathrm{Tr}[\mathbf{V}_k \mathbf{Q}_k^{'} \mathbf{V}_k^H] \tag{4.42}$$

where

$$\mathbf{Q}_k^{'} = \sum_{l \neq k} \mathbf{H}_{lk} \mathbf{W}_l^H \mathbf{P}_k^2 \mathbf{W}_l \mathbf{H}_{lk}^H \tag{4.43}$$

The intended signal power at user k from the desired BSs is given by:

$$S_k^{'} = \mathrm{Tr}[\mathbf{W}_k \mathbf{R}_k^{'} \mathbf{W}_k^H] \tag{4.44}$$

where

$$\mathbf{R}_k^{'} = \mathbf{H}_{kk} \mathbf{V}_k \mathbf{P}_k^2 \mathbf{V}_k^H \mathbf{H}_{kk}^H \tag{4.45}$$

The iterative method can be described in **Algorithm 2** shown in the following page.

4.3 Large Dimensional Analysis for DoF Balancing at Intermediate SNR

In this section, we perform large dimensional analysis of the asymptotic transmit dimensions for the small-cell users when the number of antennas goes large. We first make some assumptions on the configuration of the system as follows.

Algorithm 2 Iterative Algorithm to Solve the Soft Partial Interference Alignment

Step 1. Initialize the the receiving matrix $\mathbf{W}_k$ as the

randomly selected matrix with orthogonal columns;

Step 2. For $n = 1, 2, \ldots, N_{max}$

1). Compute the ratio λ^n based on the matrix $\mathbf{W}_k^{n-1}$:

$$\lambda^n = \frac{\mathrm{Tr}[\mathbf{W}_k^{n-1}\mathbf{R}_k(\mathbf{W}_k^{n-1})^H]}{\mathrm{Tr}[\mathbf{W}_k^{n-1}\mathbf{A}_k(\mathbf{W}_k^{n-1})^H]}$$

2). Formulate the difference problem for computing $\mathbf{W}_k^n$ as:

$$\mathbf{W}_k^n = \arg \max_{\mathbf{W}_k\mathbf{W}_k^H=\mathbf{I}_{d_k}} \mathrm{Tr}[\mathbf{W}_k^{n-1}(\mathbf{A}_k - \lambda^n\mathbf{R}_k)(\mathbf{W}_k^{n-1})^H]$$

3). Solve the problem by eigenvalue decomposition:

$$(\mathbf{A}_k - \lambda^n\mathbf{R}_k) = \mathbf{V}_k\ \mathbf{\Lambda}_k\ \mathbf{V}_k^H$$

4). reshape the column vectors of $\mathbf{W}_k^n$

3. Update until $\| \mathbf{W}_k^n - \mathbf{W}_k^{n-1} \| < \varepsilon$

Assumption 2. *The large dimension assumption in terms of numbers of antennas is defined as follows:*

- *The number of antennas in the transmit and receive side M, N and K approach infinity, $M \to \infty$, $N \to \infty$ and $K \to \infty$.*
- *The ratio of antennas is fixed, $M/K = \alpha$, $N/K = \xi$.*

4.3.1 Achievable Rate Under Water-Filling

4.3.1.1 The asymptotic water-filling level and singular value distribution

We first consider the problem of determining the asymptotic water-filling level. Then the asymptotic transmit dimensions of small cell user can be obtained directly.

The water level β can be derived by satisfying the power constraint (4.15),

$$\frac{1}{M}\sum_{n=1}^{M} p_{2,n} = \frac{1}{M}\sum_{n=1}^{M}(\beta - \frac{\sigma^2}{\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22},n}})^+ \tag{4.46}$$

In finite case, the power $p_{2,n}$ corresponding to the nth singular value can be given by

$$p_{2,n} = \max(0, \beta - \frac{\sigma^2}{\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22},n}}) \tag{4.47}$$

where β can be written as

$$\beta = \frac{1}{d}(Mp_{2,max} + \sum_{n=1}^{d}\frac{\sigma^2}{\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22},n}}) \tag{4.48}$$

where d is the number of actual transmit dimensions. The dimension with $\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22},n} < \sigma^2/\beta$ is not used. The number of actual transmit dimensions d can be given as the sum of indicator functions:

$$d = \sum_{n=1}^{M}\mathbf{1}(\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22},n}) \tag{4.49}$$

here, when $\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22},n} > \frac{\sigma^2}{\beta}$, $\mathbf{1}(\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22},n}) = 1$, else $\mathbf{1}(\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22},n}) = 0$.

We define the asymptotic transmit dimensions of small cell user as follows

$$d_\infty = \lim_{M\to\infty}\frac{d}{M} \tag{4.50}$$

then

$$\begin{aligned} d_\infty &= \lim_{M\to\infty}\frac{\sum_{n=1}^{M}\mathbf{1}(\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22},n})}{M} \\ &= \int_{\lambda_{min}}^{b}\mathbf{1}(\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22}})f(\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22}})\mathrm{d}\lambda \end{aligned} \tag{4.51}$$

where an approximation is made here. When the system goes to large dimension, the summation can be written in the integral form. The empirical eigenvalue distribution $F^{(M)}_{\mathbf{H}^H_{22}\mathbf{H}_{22}}(\lambda)$ converges almost to the deterministic limiting eigenvalue distribution $F_{\mathbf{H}^H_{22}\mathbf{H}_{22}}(\lambda)$. And $f^{(M)}_{\mathbf{H}^H_{22}\mathbf{H}_{22}}(\lambda)$ is the corresponding probability density function, which can be given according to the Marcenko-Pastur law [96],

$$f_{\mathbf{H}^H_{22}\mathbf{H}_{22}}(\lambda) = (1-\frac{K}{M})^+\delta(\lambda) + \frac{\sqrt{(\lambda-a)^+(b-\lambda)^+}}{2\pi\lambda} \tag{4.52}$$

where $a = (1-(1/\sqrt{\frac{M}{K}}))^2$ and $b = (1+(1/\sqrt{\frac{M}{K}}))^2$. Note that the density function has no-zero values in $\{\{0\}\bigcup\{[a,b]\}\}$.

In the large dimensional regime, the eigenvalues of random matrices converges in distributions. Let $f(\lambda)$ be an function of parameter λ, then if the eigenvalues fall into the interval $\lambda \in (a,b)$ and the distribution converges to $g(\lambda_j)$, the averaged sum of $f(\lambda_j)$ converges as the dimension goes large,

$$\frac{\sum_{j=1}^{d} f(\lambda_j)}{d} \xrightarrow{d\to\infty} E_\lambda[f(\lambda)] = \int_a^b f(\lambda)g(\lambda)d\lambda \tag{4.53}$$

Then we can obtain the asymptotic water-filling level as follows:

$$\begin{aligned}
\beta_\infty &= \lim_{M\to\infty}\frac{1}{d}(Mp_{2,max} + \sum_{n=1}^{d}\frac{\sigma^2}{\lambda_{\mathbf{H}^H_{22}\mathbf{H}_{22},n}}) \\
&= \frac{\lim_{M\to\infty} p_{2,max} + \frac{1}{M}\sum_{n=1}^{d}\frac{\sigma^2}{\lambda_{\mathbf{H}^H_{22}\mathbf{H}_{22},n}}}{d_\infty} \\
&= \frac{\lim_{M\to\infty} p_{2,max} + \int_{\lambda_{min}}^{b}\frac{\sigma^2}{\lambda_{\mathbf{H}^H_{22}\mathbf{H}_{22},n}} f(\lambda_{\mathbf{H}^H_{22}\mathbf{H}_{22}})\mathrm{d}\lambda}{\int_{\lambda_{min}}^{b}\mathbf{1}(\lambda_{\mathbf{H}^H_{22}\mathbf{H}_{22}})f(\lambda_{\mathbf{H}^H_{22}\mathbf{H}_{22}})\mathrm{d}\lambda}
\end{aligned} \tag{4.54}$$

Note that $\lambda_{min} = \frac{\sigma^2}{\beta_\infty}$. Then the water-filling level β_∞ can be derived by solving the fixed point equation.

From the expression in (4.54), the water-filling level is determined by the asymptotic singular value distribution of the channel matrix. When the number of antennas goes large, the water-filling level remains constant over different channel realizations.

4.3.2 Upper Bound for the Asymptotic Number of Transmit Dimensions of Small Cell User

After the water-filling level is obtained, the transmit dimensions can be obtained as

$$\begin{aligned} d_\infty &= \lim_{M\to\infty} \frac{1}{M} \sum_{n=1}^{M} \mathbf{1}_{[\frac{\sigma^2}{\beta_\infty},\infty]}(\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22},n}) \\ &= \lim_{M\to\infty} \int_{-\infty}^{+\infty} \mathbf{1}(\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22}}) f(\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22}}) \mathrm{d}\lambda \\ &= \lim_{M\to\infty} \int_{\max(\frac{\sigma^2}{\beta_\infty},a)}^{b} \mathbf{1}(\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22}}) f(\lambda_{\mathbf{H}_{22}^H\mathbf{H}_{22}}) \mathrm{d}\lambda \end{aligned} \tag{4.55}$$

The number of transmit dimensions can be exploited by the macro BS is determined by the SNR of the system and the number of transmit antennas in the BS.

4.4 Numerical Results and Discussion

In this section, we present numerical performance analysis for the proposed method. It is assumed that small scale fading part of the channels from the macro cell BS to user 0 H_{00} and the small cell user links H_{11} and H_{22} are drawn from the random matrix for which each element is generated by independent complex Gaussian variable $\mathcal{CN}(0,1)$, and the small scale fading part of cross-tier interference channel H_{10} and H_{20} are also independent complex Gaussian. So are the interference channel among the small cell links H_{12} and H_{21} The range of SNR per transmit antenna is

from −10dB to 20dB. We perform Monte-Carlo simulation with 500 random generated channels. The rest parameters can be summarized in the following table.

Table 4.1. Simulation parameters

Parameter	Assumption
Macro cell BS transmit power	45 dBm
Inter BS distance	500 m
Bandwidth	10 MHz
Macro cell BS transmit power	45 dBm
Path loss from macro cell BS to user	128+37log R
Path loss from small cell BS to user	140+36log R

4.4.1 Comparison with Other Interference Avoidance Scheme

The performance of the hard partial interference alignment is evaluated by comparing with conjugate beamforming and full interference alignment in Fig. 4.3. The conjugate beamforming is performed without considering the interference caused to other users. It only tries to maximize the intended signal in the receiver side. In the model of this paper, the procedure is carried out as this. Take the the user 1 link for example. Since the interference from other cell is ignored, we consider the user 1 link alone. Let $\mathbf{H}_{11} = \mathbf{U}_{\mathbf{H}_{11}} \ \mathbf{\Lambda}_{\mathbf{H}_{11}} \ \mathbf{V}_{\mathbf{H}_{11}}^{H}$ be the singular value decomposition of the user 1 link channel matrix. $\mathbf{U}_{\mathbf{H}_{11}}$ and $\mathbf{V}_{\mathbf{H}_{11}}$ are two unitary matrices with dimensions $K \times K$ and $M \times M$ and $\mathbf{\Lambda}$ is a diagonal matrix with singular values on the diagonal entries and zeros on other entries. The capacity can be achieved by design $\mathbf{V}_1$ and $\mathbf{W}_1$ as

$$\mathbf{V}_1 = \mathbf{V}_{\mathbf{H}_{11}} \tag{4.56}$$

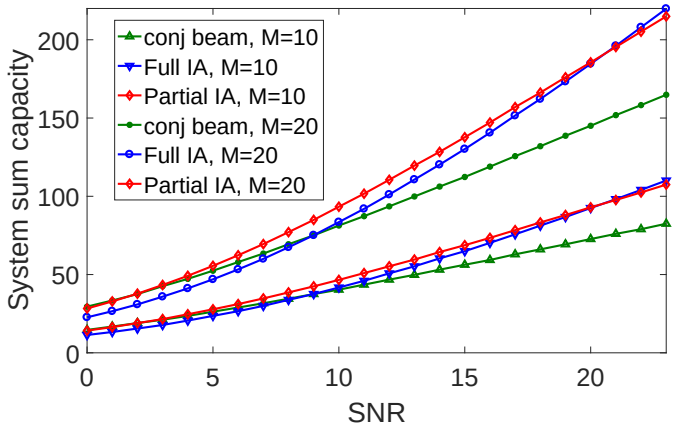

Figure 4.3. Performance comparison with other methods:hard partial interference alignment, full interference alignment method and conjugate beamforming. .

$$\mathbf{W}_1 = \mathbf{U}_{\mathbf{H}_{11}}^H \tag{4.57}$$

The proposed hard partial interference alignment achieves better performance than that of the full interference alignment method and conjugate beamforming in intermediate SNR. The intuition after this is the partial interference alignment adaptively select the number of subspaces according to the power imbalance among the macro cell BS and small cell BSs. In the intermediate SNR, it does not try to maximize the DoF of the system, while the full interference alignment method achieves the optimal sum DoF in the high SNR. In low SNR, the conjugate beamforming can obtain marginal superior performance than the partial interference alignment scheme.

Note that the average sum capacity performance can be further improved with water-filling optimization since it takes advantage of the diversity in the singular values, as can be seen in Fig. 4.4. The performances of the scheme with and without water-filling are asymptotic equivalent in high SNR.

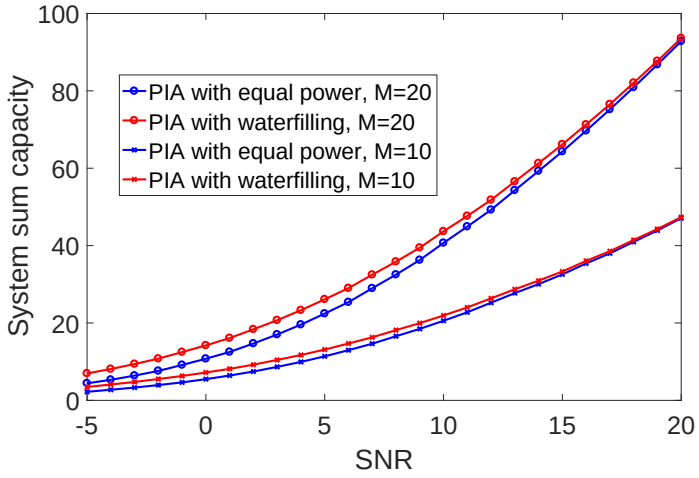

Figure 4.4. Partial interference alignment with and without WF..

4.4.2 Small Cell Transmit Dimensions and Soft Partial Interference Alignment

We further consider the effect of different number of transmit dimensions in small cell links. As shown in Fig. 4.5, we first study the relationship between sum capacity and transmit dimensions of the small cell link in low and intermediate SNR. From the figure, it is better to transmit with fewer dimensions when the transmit power is limited in small cell links. However, there is performance transition when we increase the system SNR, as shown in Fig. 4.6. It is more beneficial to select as many dimensions in small cell link as possible in high SNR. The guideline for the hard partial interference alignment scheme is that we should adaptively select the dimensions of the small cell link according to the system transmit power of Hetnet in order to achieve best system sum capacity.

After the transmit dimensions of small cell links are determined, the network sum capacity can be improved by soft partial interference alignment, as shown in Fig.4.7. The soft partial interference alignment works in an iterative way, which

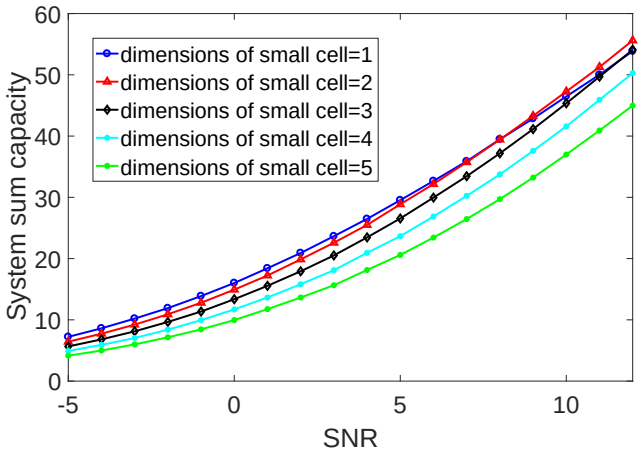

Figure 4.5. Performance of different dimensions selected in the small cell user links at low and intermediate SNR. .

takes into account of the signal and interference balancing among macro cell and small cell links.

4.4.3 Asymptotic Transmission Rate

Fig. 4.9 shows the asymptotic per antenna transmission rate when we increase the number of transmit antennas in the macro cell BS. This is actually determined by the fact that the singular value of the random channel matrices converges to certain deterministic distribution. Fig. 4.8 shows the empirical probability density function (PDF) of the eigenvalue $\lambda_{\mathbf{H}_{22}^H \mathbf{H}_{22}}$. When we increase the number of antennas in the links, the empirical PDF will coincide with the theoretical results.

In Fig. 4.9, observe that the transmission rate converges to a certain value when there are only tens of antennas. The large dimension results can be even accurate with not-so-large number of antennas.

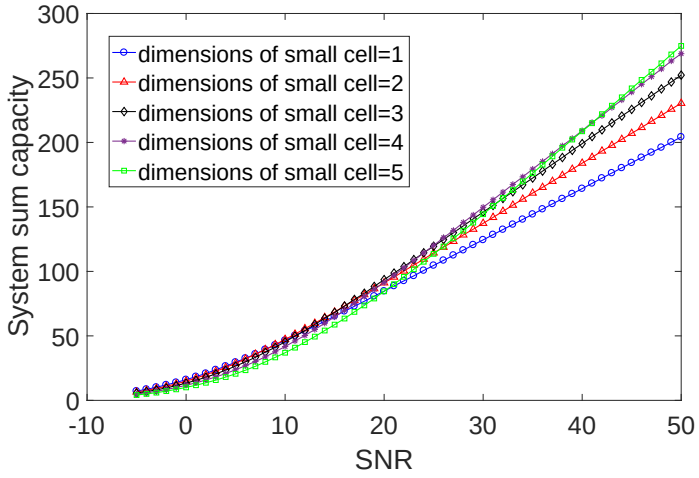

Figure 4.6. Performance of different dimensions selected in the small cell user links at high SNR. .

4.5 Conclusions

In this paper, we propose a new partial interference coordination strategy based on the idea of interference alignment that is suitable for the network sum capacity improvement in intermediate SNR. Interference alignment provides an efficient way to achieve optimal DoF in high SNR. However, it is not necessarily optimal in terms of achievable sum capacity in the low and intermediate SNR regime. We propose a interference coordination strategy called partial interference alignment, which adaptively selects the subspace for transmission based on the power constraints. This partial interference alignment captures the trade-off of interference avoidance at other users and spatial multiplexing at the intended user. The performance can be further improved by balancing the signal and interference among the small cell and macro cell links. Furthermore, large dimensional analysis is performed to show that the number of transmit dimensions which can be exploited by the small cell BS is determined by the SNR of the system and the eigenvalue distribution of the channel matrices. Sim-

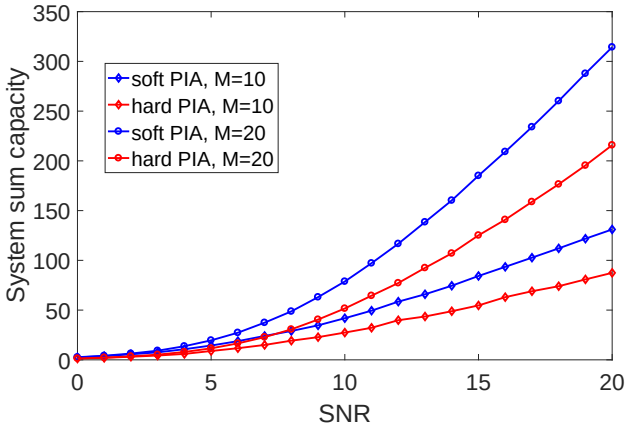

Figure 4.7. Partial interference alignment with and without signal and interference balancing . .

ulation results suggest that the proposed partial interference alignment can achieve better achievable sum capacity in the intermediate SNR regime.

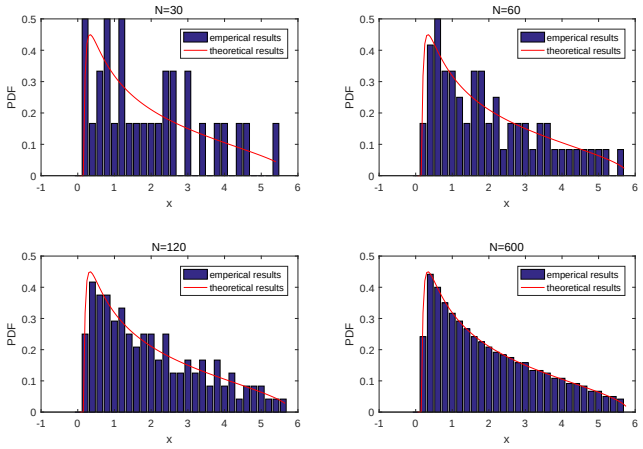

Figure 4.8. Empirical PDF when the dimensions of the channel are increased. .

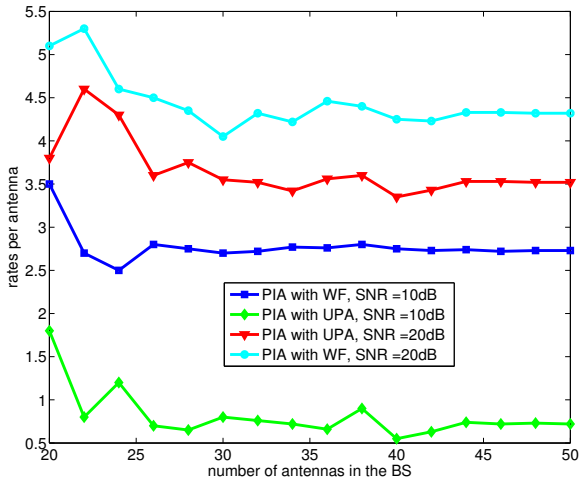

Figure 4.9. Asymptotic transmission rate. .

CHAPTER 5

Nature Encoded Sensing and Fusion for Heterogeneous Sensor Networks

Heterogeneous sensor networks [97] [98] [99] with multiple sensing modalities are gaining increasing popularity because they can provide several advantages for performance improvement in different realistic scenarios. Fusion of data from heterogeneous modalities, observing a certain phenomenon, has been shown to improve the performance of many surveillance and monitoring tasks. The key motivation is that sensors of different modalities will provide richer information than a single sensor, or even several sensors of the same modality. Take the audio-visual fusion [101] in human speech communication for example. Speech with the help of visual cues from the lip movements can enhance the intelligibility of speech. The aid of visual cues increase redundancy and makes the speech more robust to noise and interruptions.

Two sensors are said to be heterogeneous if their respective observation models cannot be described by the same probability density function [100]. This heterogeneity raises the question of how to integrate the data from such diversity of modalities. If the joint probability density function under each candidate hypothesis is known, we would easily obtain the optimal performance by the Neyman-Pearson rule for detection (binary hypothesis testing) or by the maximum a posteriori probability rule for both detection and classification (multiple hypothesis testing) [108]. However, in practice, this information may not be available. This usually happens when the dimensionality of the sample space is high and when we do not have enough training samples to have an accurate estimate of the joint PDF.

If the data from heterogeneous sensors are independent under certain hypothesis and the local individual sensing data is correctly received in the fusion center, the optimal fusion decision can be obtained by the conventional product rule. However, the problem becomes complicated when the condition that independence among the sensors does not hold [104]. In this paper, the scenario that the joint distribution between the sensors being unknown is considered. This is commonly seen in heterogeneous sensor networks, i.e., sensors with disparate sensing modalities. For example, it is not immediately clear how one could model the joint distributions between data of an audio and a video sensor monitoring a common target of interest.

The heterogeneous data processing [102] [103] has been extensively studied in the literature. The problem of binary hypothesis testing with heterogeneous sensors has been considered in [100], where a parametric framework using the statistical theory of copulas is developed. The application of copula theory for fusing correlated decisions has been recently considered in [104]. It has been shown that there is diversity gain and redundancy loss in the detection problem [105] and the influence of statistical dependence has been characterized. Previous fusion methods include the linear weighting methods [107], majority voting methods and product methods for the totally independent sensors. In linear weighting methods, individual decisions are weighted according to the reliability of the detector and then a threshold comparison is performed to obtain the global decision. The previous methods may result in information loss in many practical scenarios which will degrade the fusion performance.

In heterogeneous sensor network, due to the heterogeneity of the sensors, such as the cameras, videos and so on, they display different sensing capability, which would provide incorrect and conflict information sometimes. So we need to infer the original information from these sensed information. Also, sensors may suffer from

external attacks which will provide the wrong information to the fusion centers. We try to exploit the iterative information fusion to combat the conflict and attacks in heterogeneous sensor network.

In this Chapter, we propose a probabilistic inference framework for fusing information from heterogeneous sensors. The probabilistic framework allows data from different modalities to be processed in a unified information fusion space. The inherent inter-sensor relationship is exploited and it can be seen as a nature encoded sensing with heterogeneous sensors. Then iterative belief propagation is used to refine and fuse the local individual belief. Instead of estimating the joint PDF, we just need to abstract the local log-likelihood ratio from each sensor. Then the local log-likelihood ratio is sent to the fusion center for processing. We also consider the more general correlation case, in which the relation between two sensors is characterized by the correlation factor. The belief propagation provides intuitive insights as to how the probabilistic updates reinforce beliefs with the help of correlation factor.

5.1 Modeling Heterogeneous Sensors

In Figure 1, it shows a simple scenario of heterogeneous sensing. The sensor receives data by observing the common target of interest. Different shapes represent different sensing modalities. We can abstract the scenario as following problem. Considering the following detection problem,

$$H_1 : X_1 \sim N(\theta_1, \sigma_{X_1}), X_2 \sim N(\theta_1, \sigma_{X_2}), c(\Sigma_{X_1X_2})$$

$$H_0 : X_1 \sim N(\theta_0, \sigma_{X_1}), X_2 \sim N(\theta_0, \sigma_{X_2}), c(\Sigma_{X_1X_2})$$

where

$$\Sigma_{X_1X_2} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \tag{5.1}$$

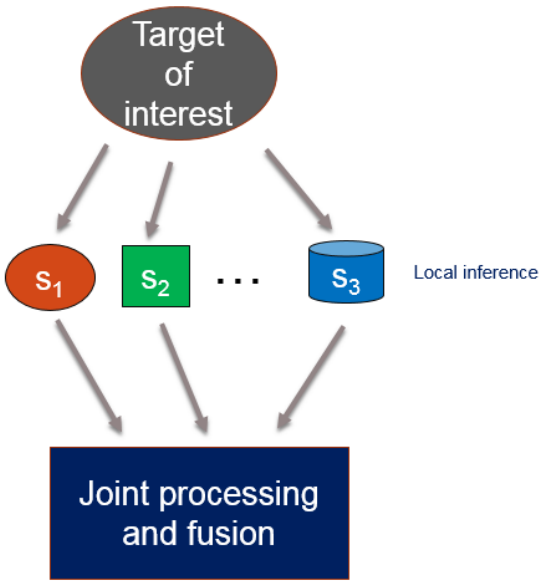

Figure 5.1. Heterogeneous sensor model.

since the sensors detect the same target, there is correlation among the sensing distributions. ρ is assumed to be the covariance between two sensor observations.

5.1.1 Log-likelihood Ratio-Based Test Statistic and Independent Case

An optimal test in both the Neyman-Pearson (NP) and the Bayesian sense computes the test statistic $\mathcal{T}$, the log-likelihood ratio (LLR) based on the joint PDF of the sensors under each hypothesis and decides in favor of H_0, when the ratio is larger than a threshold

$$\mathcal{T}(X_1, X_2) = \log \frac{p_0(X_1, X_2)}{p_1(X_1, X_2)} \geq \eta \tag{5.2}$$

To obtain the optimal detection performance, we need to know the joint distribution of all the heterogeneous sensor observations. However, this is often not feasible in most practical scenarios since it requires many training samples to estimate the joint

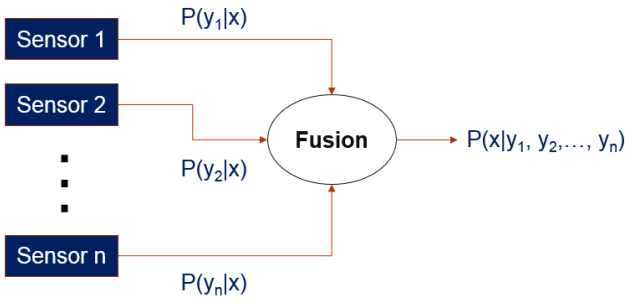

Figure 5.2. Probabilistic fusion framework .

distribution of sensors. And especially for the heterogeneous sensors, there is lack of an unified framework to process the data of different modalities. For the independent case, we can write the joint PDF as the product of each sensor's probability distribution, and the conventional product rule can obtain the optimal performance.

5.2 Belief Propagation and Nature Encoded Fusion

5.2.1 Local Information Abstraction

In this Section, we try to fuse and combine the outputs of entirely different sensor modalities under the probabilistic framework. The probabilistic approach theoretically provides a unified framework for data fusion in heterogeneous sensor networks. Instead of estimating the joint PDF of the heterogeneous sensor observations, we just need to calculate the local LLR of each sensors.

$$\mathrm{LLR}_i = \log \frac{p_0(X_i|H_0)}{p_1(X_i|H_1)} \tag{5.3}$$

which we refer to as the local belief of each heterogeneous sensor. The fusion framework is depicted in Fig. 5.2.

Since each sensor can provide the local likelihood ratio for the fusion center, by posing some constraints on the sensory signals, we can generate some prior information for each sensor observation. This falls into the Bayesian probabilistic inference framework. The Bayesian approach is theoretically sufficient for providing a unified framework for data fusion in heterogeneous sensor networks. This framework answers the question of how to weight or process outputs of diverse sensors, no matter they have different sensing modes.

We propose a nature encodes fusion method based on the belief propagation principle. The joint PDF of the sensors' observations is unknown. In our method, each sensor just perform local inference by calculating the local LLR. It is easily available since the local marginal distribution can be estimated individually. It is interesting to note that a simplified, specific form of this type of information processing occurs in the so called belief propagation. It is thus in a general form that is suitable for application to heterogeneous multi-sensor systems. This provides intuitive insight as to how the probabilistic updates help to reinforce the local beliefs when performing inference.

5.2.2 Inherent Constraints among the Sensors

We try to take advantage of the intermodal relation among the sensory signals by modeling it as a nature encoding process. Then we can update the LLR in the Bayesian way by posing some inherent constraints on the local inference of the heterogeneous sensors.

According to the Bayesian principle, the posterior LLR is the summation of likelihood LLR and the prior LLR. However, there is lack of prior information in the present model. In this paper, the prior of one sensor can be provided by the belief of other sensors through some inherent constraints among the sensors. Then the local

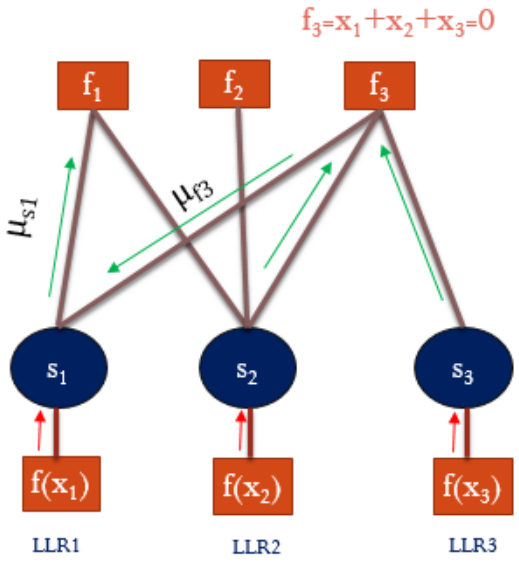

Figure 5.3. Factor graph for nature encoded fusion .

belief can be propagated among the sensors and finally a better fusion results can be obtained.

$$\mathrm{LLR}_{posterior} = \mathrm{LLR}_{likelihood} + \mathrm{LLR}_{prior} \tag{5.4}$$

When there is only one target, the sensing results should be the same for all the sensors. we can make use of this intermodal relation to fuse all the local belief. For any two of the sensors, we can generate a factor node by letting the logic summation of them be zero.

$$x_0 = x_1 = x_2 = \cdots = x_k$$

and the corresponding factors can be written as

$$f_1 = x_0 + x_1 = 0 \tag{5.5}$$

$$f_2 = x_1 + x_2 = 0 \tag{5.6}$$

$$\cdots \tag{5.7}$$

$$f_k = x_{k-1} + x_k = 0 \tag{5.8}$$

which we can write it in matrix form $\begin{bmatrix} 1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$. Each row corresponds to a constraint and the entries with 1 mean that the sensor is involved in that constraint.

We can also relax the constraints and let the summation of any even number of sensors be zero, such as

$$x_0 + x_1 + x_2 + x_3 = 0$$

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 0$$

which will be evaluated in the simulation.

5.2.3 Probability Calculation for the Fusion Process

For the modulo-2 addition, suppose there are two binary random variables x and y, with probability distributions $p(x) = \{p_0^x, p_1^x\}$ and $p(y) = \{p_0^y, p_1^y\}$, then the probability distribution for the modulo-2 addition is

$$P(x \oplus y = 0) = \frac{1}{2}((p_0^x + p_1^x)(p_0^y + p_1^y) + (p_0^x - p_1^x)(p_0^y - p_1^y)),$$

We can extend the result to the infinite number of random varibles $x_i, i = 0, 1, 2, \ldots, N$, then the distribution for the modulo-2 addition is

$$P(\sum_{i=1} \oplus x_i = 0) = \frac{1}{2}(\prod_{i=0}^{N}(p_0^{x_i} + p_1^{x_i}) + \prod_{i=0}^{N}(p_0^{x_i} - p_1^{x_i})) \tag{5.9}$$

$$P(\sum_{i=1} \oplus x_i = 1) = \frac{1}{2}(\prod_{i=0}^{N}(p_0^{x_i} + p_1^{x_i}) - \prod_{i=0}^{N}(p_0^{x_i} - p_1^{x_i})) \tag{5.10}$$

Then the log-likelyhood ratio (LLR) is defined as $L(x) = \ln(\frac{P(x=0)}{P(x=1)})$. We can write the corresponding probability in the LLR form as follows:

$$P(x = 0) = \frac{e^{L(x)}}{1 + e^{L(x)}} \tag{5.11}$$

$$P(x = 1) = \frac{1}{1 + e^{L(x)}} \tag{5.12}$$

Then we will give the LLR form of the modulo-2 variable $L(\sum_{i=1} \oplus x_i) = L(x_1 \oplus x_2 \oplus \ldots \oplus x_N)$. According to (3)(4) and (5)(6),

$$L(\sum_{i=1} \oplus x_i) = \ln(\frac{\prod_{i=0}^{N}(e^{L(x_i)} + 1) + \prod_{i=0}^{N}(e^{L(x_i)} - 1)}{\prod_{i=0}^{N}(e^{L(x_i)} + 1) - \prod_{i=0}^{N}(e^{L(x_i)} - 1)})$$

We can rewrite the equation as

$$L(\sum_{i=1} \oplus x_i) = \ln(\frac{1 + \prod_{i=0}^{N} \tanh(L(x_i))}{1 - \prod_{i=0}^{N} \tanh(L(x_i))})$$

here $\tanh(\frac{x}{2}) = \frac{e^x - 1}{e^x + 1}$.

Then let $\gamma = \prod_{i=0}^{N} \tanh(\frac{L(x_i)}{2})$, we have

$$\begin{aligned} \tanh(\frac{1}{2}L(\sum_{i=1} \oplus x_i)) &= \frac{e^{\ln(\frac{1+\gamma}{1-\gamma})} - 1}{e^{\ln(\frac{1+\gamma}{1-\gamma})} + 1} \\ &= \frac{(1+\gamma) - (1-\gamma)}{(1+\gamma) + (1-\gamma)} \\ &= \gamma \end{aligned}$$

So we can get

$$L(\sum_{i=1} \oplus x_i) = 2\tanh^{-1}(\prod_{i=0}^{N} \tanh(\frac{L(x_i)}{2})) \tag{5.13}$$

5.2.4 LLR Form for the Belief Propagation

Belief propagation algorithm uses message passing over the factor graph. There are two sets of variables over the graph: variable nodes and factor nodes. The variable nodes correspond to the sensor belief, while the factor nodes corresponds to the constraints in our settings. Now we will go to the details of applying belief propagation algorithm to the decoding of heterogeneous sensor networks. First, we will define the one to one correspondence of message distribution and the LLR form.

$$\lambda_{x_n \to f_m}(x_n) = L(\mu_{x_n \to f_m}(x_n))$$
$$\lambda_{f_m \to x_n}(x_n) = L(\mu_{f_m \to x_n}(x_n))$$

So it is straight forward to convert the updating rule from variable to factor message to the following form:

$$L(\mu_{x_n \to f_m}(x_n)) = L(\prod_{f_i \in f_n \setminus f_m} \mu_{f_i \to x_n}(x_n))$$
$$\lambda_{x_n \to f_m}(x_n) = \sum_{f_i \in f_n \setminus f_m} \lambda_{f_i \to x_n}(x_n)$$

Then we will consider the updating from the factor to the variable message. Since $f_m = \sum_{x_i \in x_m} \oplus x_i$ and the logic relation between x_n and f_m is

$$x_n = f_m \oplus \sum_{x_i \in x_m \setminus x_n} \oplus x_i$$

we know that the updating rule for factor to variable is

$$L(\mu_{f_m \to x_n}) = L(\sum_{\mathcal{X}_m \backslash x_n} f_m(x_m) \prod_{x_i \in x_m \backslash x_n} \mu_{x_i \to f_m}(x_i))$$

That is

$$\lambda_{f_m \to x_n}(x_n) = L(y_m \oplus \sum_{x_i \in \mathcal{X}_m \backslash x_n} \oplus x_i)$$

Based on the results (7), we can derive that

$$\begin{aligned} &\lambda_{f_m \to x_n}(x_n) \\ &= 2 \tanh^{-1}(\tanh(\lambda_{f_m}) \prod_{x_i \in \mathcal{X}_m \backslash x_n} \tanh(\frac{\lambda_{x_i \to f_m}(x_i))}{2})) \end{aligned} \tag{5.14}$$

5.3 General Correlation Case and Theoretical Analysis of the Fusion Problem

In this section, we will consider the more general correlation case, in which the relation between two sensors is characterized by the spatial correlation factor λ_{ij}. A probabilistic graphical model is proposed to fuse the prior information from other sensors. The corresponding belief updating rule is developed and the performance is analyzed theoretically.

5.3.1 Correlation Modeling

We further assume that the spatial interactions between adjacent nodes are pairwise. The correlation between sensors x_i and x_j can be modeled as

$$\psi_{ij}(x_i, x_j) = \exp(\lambda_{ij} g(x_i, x_j))$$

where $\lambda_{ij} \geq 0$ and it is proportional with the increased correlation between the sensors. And the indicator function $g(x_i, x_j)$ is defined as

$$g(x_i, x_j) = x_i x_j + (1 - x_i)(1 - x_j)$$

The joint posterior distribution can be given as

$$p(\mathbf{x}|\mathbf{y}) \propto \prod_{i=1} p(y_i|x_i) \prod_{j \in \mathcal{N}_i} \exp(\lambda_{ij} g(x_i, x_j))$$

which is the product of the local likelihood and the correlation factor. We aims to determine the belief of the target for each sensor $p(x_n|\mathbf{y})$, given the observations and the correlations. The problem can be given as the maximum a posteriori (MAP) estimation:

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x} \in \{0,1\}^n} p(\mathbf{x}|\mathbf{y})$$

5.3.2 Belief Propagation for Spatial Correlated Sensor Observations

The spatially correlated sensors can be illustrated in a factor graph, as depicted in Fig. 5.4. The sensors are represented as circles and the inter-sensor correlations are represented as squares.

Now we will fuse the sensors' information based on belief propagation algorithm. Each sensor fuse the multi-prior information from other sensors through the inter-sensor correlations. The final belief of each sensor can be reached by iteration of messages among the nodes.

Here, the variable nodes correspond to the sensor belief, while the factor nodes corresponds to the inter-sensor correlation in our settings. It tries to reveal the

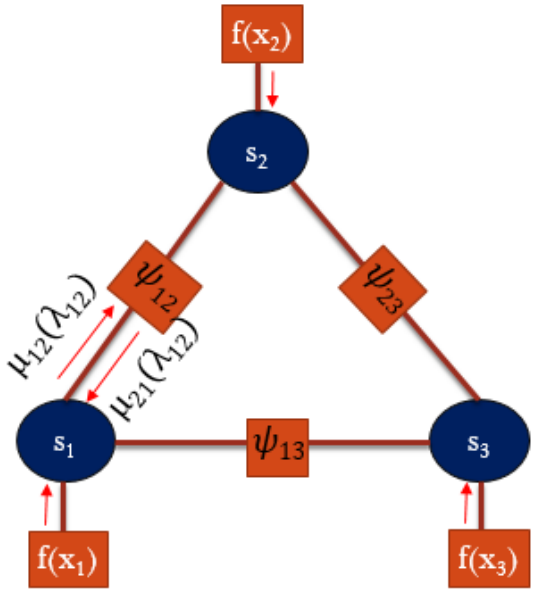

Figure 5.4. Factor graph for general correlated sensors .

distribution of variable nodes with the help of factor nodes. By performing message passing among each pair of the variable and factor nodes, the belief of the variable nodes can be updated. Usually, the updating rules are defined as:

- The message from variable nodes to factor nodes:

$$\mu_{x_n \to f_m} = \prod_{f_i \in \mathcal{F}_n \setminus f_m} \mu_{f_i \to x_n}(x_n) \tag{5.15}$$

- Updating rule from factor to variable:

$$\mu_{f_m \to x_n} = \sum_{\sim x_n} f_m \prod_{x_i \in \mathcal{X}_m \setminus x_n} \mu_{x_i \to f_m}(x_i) \tag{5.16}$$

where the factor node f_m and the variable node x_n are connected. $\mathcal{F}_n$ is the set of factor nodes that are connected to variable node x_n and likewise, $\mathcal{X}_m$ is the set of variable nodes that are connected to factor node f_m.

Note that, there are two types of factor nodes connected to each variable node. The first one is the local sensing factor,

$$\mu_{f \to x_n} = f(x_n) \tag{5.17}$$

the other one is the inter-sensor correlation $f_m = \psi_{ni}$,

$$\mu_{f_m \to x_n} = \sum_{\sim x_i} f(x_n, x_i) \mu_{x_i \to f_m}(x_i) \tag{5.18}$$

$$= \sum_{\sim x_i} \psi_{ni}(x_n, x_i) \mu_{x_i \to f_m}(x_i) \tag{5.19}$$

After each iteration, the belief of each variable node is computed by the product of all the information from the factor node. And then the probability is normalized and pass it to its corresponding factor node.

5.3.3 Kullback-Leibler Divergence and Theorectical Analysis

In the detection problem, when there are two hypothesis, the sensed data would exhibit distribution $f_0(x|H_0)$ under hypothesis H_0 and distribution $f_1(x|H_1)$ under hypothesis H_1. The Kullback-Leibler Divergence (KLD) is usually used to evaluate the performance of detection, which is defined as

$$D(H_0 \| H_1) = \mathrm{E}_{H_0}[\log \frac{f(\mathbf{x}|H_0)}{f(\mathbf{x}|H_1)}] \tag{5.20}$$

where E_{H_0} is the expectation taken with respect to the joint distribution of $\mathbf{x}$ under hypothesis H_0. KLD can be interpreted as the error exponent in the Neyman-Pearson framework, which means that the probability of miss detection goes to zero exponentially with the number of observations at a rate equal to KLD.

If the sensor observations are independent of each other, then $p_1(x,y) = p_1(x)p_1(y)$, $p_0(x,y) = p_0(x)p_0(y)$. Then KLD can be written as

$$\begin{aligned} &D(p_1(x,y)\|p_0(x,y)) \\ &\quad = D(p_1(x)\|p_0(x)) + D(p_1(y)\|p_0(y)) \end{aligned} \tag{5.21}$$

In the case of correlated modalities $p_1(x,y) \neq p_1(x)p_1(y)$, The KLD for the two hypothesis can be given according to the chain rule,

$$\begin{aligned} &D(p_1(x,y)\|p_0(x,y)) \\ &\quad = D(p_1(x)\|p_0(x)) + D(p_1(y|x)\|p_0(y|x)) \end{aligned} \tag{5.22}$$

there is a lemma in [104] which proves that conditioning does not reduce relative entropy, which means that $D(p_1(y)\|p_0(y)) \leq D(p_1(y|x)\|p_0(y|x))$. In this paper, we proposed the fusion method based on belief propagation by exploiting this kind of correlations between the sensor observations.

Also in [105], it gives a proposition that the KLD is nondecreasing with the increased number of sensors. For any $S' \subseteq S$, $D(S') \leq D(S)$.

5.4 Performance Analysis and Numerical Results

Numerical simulations are performed to illustrate our proposed fusion algorithm for the heterogeneous sensor network. To model the heterogeneous modalities, we make some abstractions and generate the sensing data based on different distributions. A sensor network with N sensors is considered. The correlation between sensors is quantified by the power exponential model.

$$\lambda_{ij} = e^{-d_{ij}^2} \tag{5.23}$$

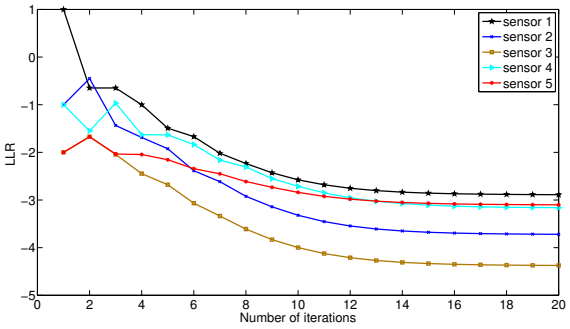

Figure 5.5. Iterations of belief propagation based fusion .

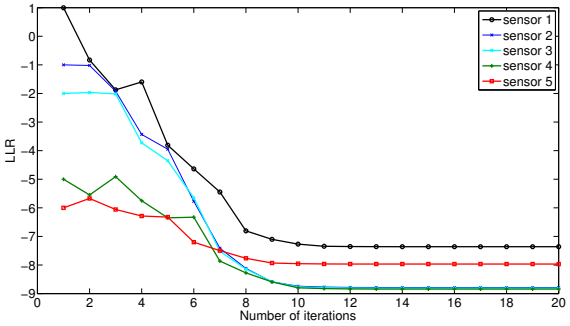

Figure 5.6. Iterations of belief propagation based fusion .

where d_{ij} is the distance between node i and j.

Each sensor's data follow Gaussian distribution with different mean values θ and variance σ^2. They can be adjusted to meet different SNR scenarios.

5.4.1 Fusion for the Conflicted Belief

We first consider the robustness of the belief propagation based fusion method. Two different initial LLR settings are evaluated, which are the $[1, -2, -1, -2, -1]$ and $[1, -1, -2, -5, -6]$. The convergence results for both cases are shown in Fig. 5.5 and Fig. 5.6. For both cases, the first sensor's initial belief is negative and is opposite to the rest of sensors. After a few iterations, all the sensors' belief become positive. The belief propagation forces the ambiguous belief to be correct thanks to the correlations

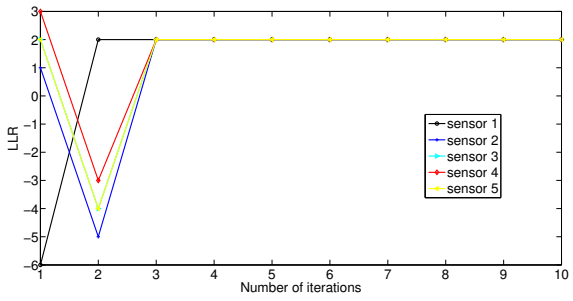

Figure 5.7. Initial LLR=[-6 1 2 3 2] .

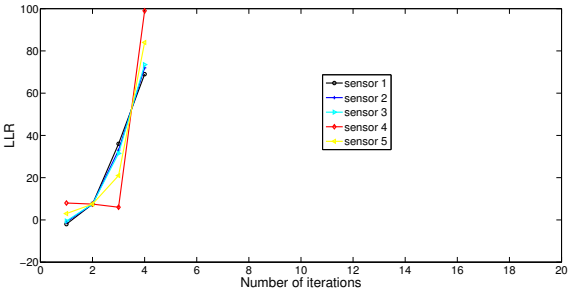

Figure 5.8. Initial LLR= [-0.2 -0.2 -0.5 8 3].

among the sensors. After the fusion, we can apply the majority voting method and the detection performance would be robust to the error and uncertain noise.

5.4.2 Impact of the Correlation Coefficients

The scenario that the correlation among the sensors λ_{ij} goes large is considered in Fig. 5.7 and Fig. 5.8. The LLR values of all the sensors are forced to be the same belief in Fig. 5.7. Even when the first sensor exhibits high negative LLR value, the iterative belief propagation can correct that belief and achieve a consensus fusion result. The reason why the case in Fig. 5.8 does not converge is that the factor graph contains loops in this example. Also we can see that the first three sensors' LLRs are negative and exhibit weak belief as to whether the hypothesis is H_0. Still they can be refined to be positive thanks to the correlation among the sensors. If majority voting is

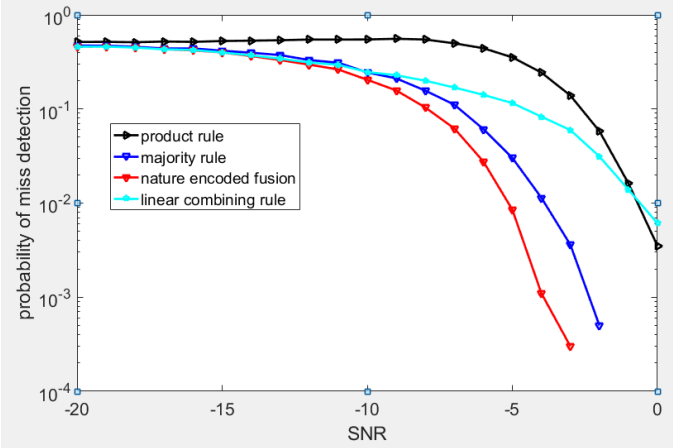

Figure 5.9. Performance comparison with product rule, linear combining rule, and majority voting.

applied after the refinement, we can achieve better detection performance. Correlation among the sensor observations makes some sensors' measurements redundant, which will provide more robustness in the fusion process.

5.4.3 Performance Comparison

We compare the performance of proposed nature encoded fusion with the majority voting and linear weighted combining method. The linear weighted combing is computed as

$$p_{fusion} = \frac{\sum_i w_i p_i}{\sum_i w_i} \tag{5.24}$$

Fig. 5.9 shows that the belief propagation based fusion can achieve superior detection performance compared with the other three fusion methods. It may result in information loss duo to hard processing of majority voting scheme. Also much training overhead is required to obtain the optimal weighting coefficient for the linear combining method. The product rule will cause error propagation which will degrade the fusion performance.

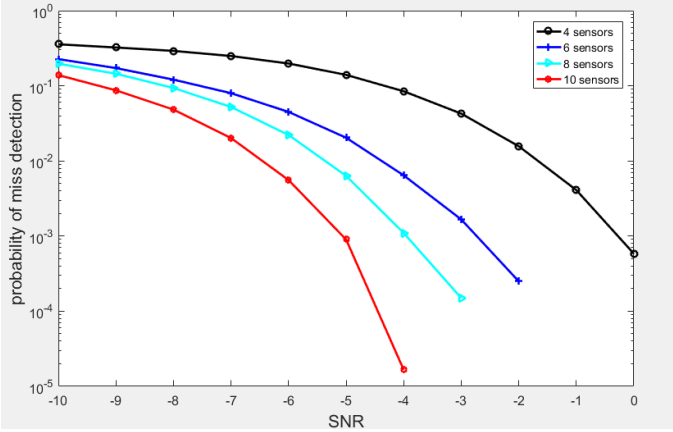

Figure 5.10. Detection performance with increasing number of sensors.

In Fig. 5.10, it shows that the more sensors deployed for the detection, the better the detection performance. It verifies the theoretical result that KLD is non-decreasing with the increased number of sensors.

5.5 Conclusions

In this Chapter, we propose a probabilistic inference framework for fusing information from heterogeneous sensors. The probabilistic framework allows data from different modalities to be processed in a unified information fusion space. The inherent inter-sensor relationship is exploited and it can be seen as a nature encoded sensing with heterogeneous sensors. Then iterative belief propagation is used to refine and fuse the local individual belief. Instead of estimating the joint PDF, we just need to abstract the local log-likelihood ratio from each sensor. Then the local log-likelihood ratio is sent to the fusion center for processing. Further we consider the more general correlation case, in which the relation between two sensors is characterized by the correlation factor. The belief propagation provides intuitive insights as to how the probabilistic updates reinforce beliefs with the help of correlation factor.

CHAPTER 6

Conclusions and Future Works

This chapter concludes the whole dissertation. It begins with a summary of the dissertation results and contributions, follows with a discussion of future research directions.

6.1 Summary

Generally speaking, there were two main themes of this thesis, which are summarized are follows:

In the first part, we mainly focus on the large dimensional performance analysis and optimization of massive MIMO cellular networks. Large dimensional analysis provides an efficient way to quantify the network performance. We compare two cooperative multicell precoding methods, namely, centralized precoding with global CSI and distributed precoding with local individual CSI in the large dimensional regime. In the large dimension limit, certain terms such as signal-to-interference-plus-noise ratio (SINR) can be approximated depending only on the statistical CSI. Based on this result, the power optimization problem does not need to adapt as frequently as the instantaneous channel state information.

In the second part, we consider the problem of fusion for heterogeneous sensor networks. We propose a probabilistic inference framework for fusing information from heterogeneous sensors. The probabilistic framework allows data from different modalities to be processed in a unified information fusion space. The inherent inter-

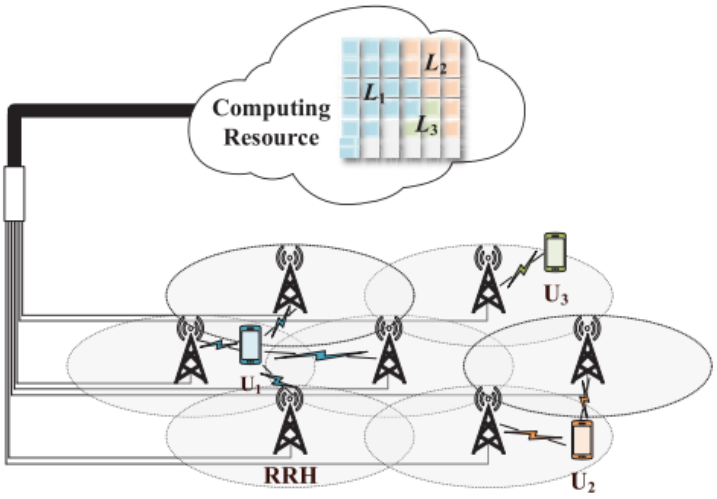

Figure 6.1. The downlink transmission of a C-RAN with pooling computing resource .

sensor relationship is exploited to encode the original sensor data in a graph. The iterative belief propagation is used to fuse the local sensing belief.

6.2 Future Works

6.2.1 Large Dimensional Analysis and Optimization for Cloud-Radio Access Network

Since performance evaluation and optimization is more computationally efficient in the large dimensional limit, we will apply the large dimensional analysis to the optimization of cloud-radio access network (C-RAN). C-RAN , as shown in Fig. 6.1 is a centralized, real time, and cloud computing-based architecture for radio access networks that supports 4G and future wireless communication standards. C-RAN has emerged as a promising solution to support exponentially increasing demand in data rate. The attractive capacity enhancement mainly comes from centralized and coordinated processing, which poses great challenges on computing capability in the

baseband unit pool. When the network scale goes large, the computational complexity for the centralized optimization would be prohibitively high.

The large dimensional analysis can efficiently reduce the updating rate of the optimization, which will be a good fit for the C-RAN. In our previous research, we assume synchronized fading among all the users and BSs, where the fading remains constant for all the users. For future work, we will consider more realized scenario, in which there is mismatch among the variations of users' channel fading. The asymptotic performance of the network will be analyzed in the large dimensional limits. We can quantify the performance loss due to the channel mismatch among the users. More robust optimization can be developed to counter the negative impact of channel mismatch.

6.2.2 Energy Efficient Fusion and Transmission for Heterogeneous Sensor Networks

For heterogeneous sensor networks, energy efficiency is another important problem to be considered. Heterogeneity between different sensor modalities can maximize the information for the target of interest. It can reduce the redundancy among the sensors. We can perform local compressive combination by exploiting the heterogeneity between different sensor modalities. Then energy efficiency can be achieved in the transmission to the fusion center. Therefore, one task of our future work is to develop efficient combination algorithm for the heterogeneous sensor networks and analyze the energy efficiency performance compared to other methods.

In our previous research, we assume information abstraction of LLR in each individual sensor. For the energy efficiency of the transmission, the multiple sensors joint belief can be extracted and compressed. Due to the diversity among sensors, the heterogeneous sensor network can provide more information given a certain number of sensors. A natural question arise that can we achieve more with less by having as

few observations as possible and yet enough to provide detection guarantees. We will characterize the information limit that the heterogeneous sensors can provide. The energy efficiency performance due to diversity will be studied.

APPENDIX A

Matrix inversion lemma [68]

Let $\mathbf{A}$ be an $N \times N$ invertible matrix and let $\mathbf{x}$ and $\mathbf{y}$ be two $N \times 1$ complex vectors such that $(\mathbf{A} + \mathbf{x}\mathbf{x}^H)$ is invertible. c is a scalar.Then we have

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

$$(\mathbf{A} + c\mathbf{x}\mathbf{x}^H)^{-1}\mathbf{x} = \frac{\mathbf{A}^{-1}\mathbf{x}}{1 + c\mathbf{x}^H\mathbf{A}^{-1}\mathbf{x}} \tag{A.2}$$

where $E[.]$ is the mathematical expectation.

APPENDIX B

Trace lemma [68]

Lemma 5. *Assume that* $\mathbf{A}$ *is a deterministic* $N \times N$ *matrix with uniformly bounded spectral norm (with respect to* N*). Let* $\mathbf{x} = \frac{1}{\sqrt{N}}[x_1, x_2, \cdots, x_N]$ *where* $\mathbf{x}$ *are i.i.d complex random variables with zero mean, unit variance and finite moment. Let* $\mathbf{y}$ *be a similar vector independent of* $\mathbf{x}$*. Both* $\mathbf{x}$ *and* $\mathbf{y}$ *are independent of* $\mathbf{A}$*. Then*

$$\mathbf{x}^H \mathbf{A} \mathbf{x} \xrightarrow{N \to \infty} \frac{1}{N} tr\mathbf{A} \tag{B.1}$$

$$\mathbf{x}^H \mathbf{A} \mathbf{y} \xrightarrow{N \to \infty} 0 \tag{B.2}$$

almost surely when $N \to \infty$*.*

APPENDIX C

Proof of lemma 1

We begin the proof by considering the signal power term, which can be written as

$$|\mathbf{h}_{ik}^H \mathbf{w}_{ik}^l|^2 = |\mathbf{h}_{ik}^H \frac{\bar{\mathbf{w}}_{ik}^l}{\|\bar{\mathbf{w}}_{ik}^l\|}|^2 = \frac{|\mathbf{h}_{ik}^H \bar{\mathbf{w}}_{ik}^l|^2}{\|\bar{\mathbf{w}}_{ik}^l\|^2} \tag{C.1}$$

We first consider the term $|\mathbf{h}_{ik}^H \bar{\mathbf{w}}_{ik}^l|^2$, which can be rewritten as

$$\begin{aligned} \mathbf{h}_{ik}^H \bar{\mathbf{w}}_{ik}^l &= \frac{1}{N}\mathbf{h}_{ik}^H (\frac{1}{N}\mathbf{I} + \sum_{k' \neq k} \frac{\alpha}{N}\mathbf{h}_{ik'}\mathbf{h}_{ik'}^H)^{-1}\mathbf{h}_{ik} \\ &= \frac{1}{N}\mathbf{h}_{ik}^H \Sigma_{ik}\mathbf{h}_{ik} \\ &= \frac{1}{N}\mathbf{g}_{ik}^H \mathbf{D}_{ik}^{\frac{1}{2}} \Sigma_{ik} \mathbf{D}_{ik}^{\frac{1}{2}} \mathbf{g}_{ik} \end{aligned} \tag{C.2}$$

where $\Sigma_{ik} = (\frac{1}{N}\mathbf{I} + \sum_{k' \neq k} \frac{\alpha}{N}\mathbf{h}_{ik'}\mathbf{h}_{ik'}^H)^{-1}$, Then by applying Lemma 3, we get

$$\mathbf{h}_{ik}^H \bar{\mathbf{w}}_{ik}^l - \frac{1}{N}\mathrm{tr}\mathbf{D}_{ik}\Sigma_{ik} \xrightarrow{N\to\infty} 0$$

according to Theorem 1 in [67],

$$\frac{1}{N}\mathrm{tr}\mathbf{D}_{ik}\Sigma_{ik} - \psi_{ik}^{\circ} \xrightarrow{N\to\infty} 0 \tag{C.3}$$

According to Lemma 3 and Theorem 2 in [67], the term $\|\bar{\mathbf{w}}_{ik}^l\|^2$ can be approximated by

$$\|\bar{\mathbf{w}}_{ik}^l\|^2 - \frac{1}{N}\mathrm{tr}\mathbf{D}_{ik}\Sigma_{ik}^2 \xrightarrow{N\to\infty} 0$$

$$\frac{1}{N}\mathrm{tr}\mathbf{D}_{ik}\Sigma_{ik}^2 - \bar{\psi}_{ik}^{\circ} \xrightarrow{N\to\infty} 0 \tag{C.4}$$

Then based on (C.3) and (C.4),

$$|\mathbf{h}_{ik}^H \mathbf{w}_{ik}^l|^2 \xrightarrow{N\to\infty} \frac{(\psi_{ik}^\circ)^2}{\bar{\psi}_{ik}^\circ} \tag{C.5}$$

Next, we consider the term $|\mathbf{h}_{ik}^H \bar{\mathbf{w}}_{ik'}^l|^2$, which can be written as

$$\begin{aligned} |\mathbf{h}_{ik}^H \bar{\mathbf{w}}_{ik'}^l|^2 &= \frac{1}{N^2}\mathbf{h}_{ik}^H(\Sigma_{ik'}\mathbf{h}_{ik'}\mathbf{h}_{ik'}^H\Sigma_{ik'})\mathbf{h}_{ik} \\ &= \frac{1}{N^2}\mathbf{h}_{ik'}^H(\Sigma_{ik'}\mathbf{h}_{ik}\mathbf{h}_{ik}^H\Sigma_{ik'})\mathbf{h}_{ik'} \end{aligned} \tag{C.6}$$

where $\Sigma_{ik'} = (\frac{1}{N}\mathbf{I} + \sum_{\bar{k}\neq k'} \frac{\alpha}{N}\mathbf{h}_{i\bar{k}}\mathbf{h}_{i\bar{k}}^H)^{-1}$ and $\mathbf{h}_{ik'}$ is independent of $\Sigma_{ik'}$ and $\mathbf{h}_{ik}$, then Lemma 3 can be applied here. We get

$$|\mathbf{h}_{ik}^H \bar{\mathbf{w}}_{ik'}^l|^2 - \frac{\beta_{ik'}}{N^2}\mathrm{tr}(\Sigma_{ik'}\mathbf{h}_{ik}\mathbf{h}_{ik}^H\Sigma_{ik'}) \xrightarrow{N\to\infty} 0$$

since $\mathrm{tr}(\Sigma_{ik'}\mathbf{h}_{ik}\mathbf{h}_{ik}^H\Sigma_{ik'}) = \mathbf{h}_{ik}^H\Sigma_{ik'}^2\mathbf{h}_{ik}$ Here $\Sigma_{ik'}$ includes $\mathbf{h}_{ik}$. In order to remove the dependence between them, the matrix inversion lemma in the appendix is applied here, then

$$\frac{1}{N^2}\mathbf{h}_{ik}^H\Sigma_{ik'}^2\mathbf{h}_{ik} = \frac{1}{N^2}\frac{\mathbf{h}_{ik}^H\Sigma_{ik'k}\Sigma_{ik'k}\mathbf{h}_{ik}}{(1+\frac{\alpha}{N}\mathbf{h}_{ik}^H\Sigma_{ik'k}\mathbf{h}_{ik})^2}$$

where $\Sigma_{ik'k} = (\frac{1}{N}\mathbf{I} + \sum_{(\bar{k})\neq(k,k')} \frac{\alpha}{N}\mathbf{h}_{i\bar{k}}\mathbf{h}_{i\bar{k}}^H)^{-1}$. Since $\mathbf{h}_{ik}$ is independent of $\Sigma_{ik'k}$, then we can apply Lemma 3, and get

$$\frac{1}{N^2}\mathbf{h}_{ik}^H\Sigma_{ik'}^2\mathbf{h}_{ik} - \frac{1}{N^2}\frac{\mathrm{tr}\mathbf{D}_{ik}\Sigma_{ik'k}^2}{(1+\frac{\alpha}{N}\mathrm{tr}\mathbf{D}_{ik}\Sigma_{ik'k})^2} \xrightarrow{N\to\infty} 0$$

According to Theorem 2 in [67], the term $\frac{1}{N}\mathrm{tr}\mathbf{D}_{ik}\Sigma_{ik'k}^2$ and $\frac{1}{N}tr\mathbf{D}_{ik}\Sigma_{ik'k}$ can be approximated as

$$\frac{1}{N}\mathrm{tr}\mathbf{D}_{ik}\Sigma_{ik'k} - \Phi_{ik'k}^\circ \xrightarrow{N\to\infty} 0$$

$$\frac{1}{N}\mathrm{tr}\mathbf{D}_{ik}\Sigma_{ik'k}^2 - \bar{\Phi}_{ik'k}^\circ \xrightarrow{N\to\infty} 0$$

where $\Phi^{\circ}_{k'k}$ and $\bar{\Phi}^{\circ}_{ik'k}$ are only functions of the slow fading coefficients of the channel.

Then $|\mathbf{h}_{ik}^H \mathbf{w}_{ik'}^l|^2$ can be approximated as

$$|\mathbf{h}_{ik}^H \mathbf{w}_{ik'}^l|^2 \xrightarrow{N\to\infty} \frac{\beta_{ik'}}{N} \frac{\bar{\Phi}^{\circ}_{ik'k}}{\bar{\psi}^{\circ}_{ik'}(1+\alpha\Phi^{\circ}_{ik'k})^2} \tag{C.7}$$

REFERENCES

[1] Hien Quoc Ngo, Massive MIMO: Fundamentals and System Designs, *Linkping Studies in Science and Technology Dissertations*, No. 1642.

[2] Lu Lu, Geoffrey Ye Li, et al, An Overview of Massive MIMO:Benefits and Challenges, *IEEE Journal of Selected Topics in Signal Processing*, VOL. 8, NO. 5, OCTOBER 2014.

[3] J. G Andrews, X. Zhang, G. D Durgin, A. K Gupta, Are We Approaching the Fundamental Limits of Wireless Network Densification ?, *arXiv preprint arXiv:1512.00413*.

[4] T.Nakamuraetal., Trends in small cell enhancements in LTE advanced, *IEEE Commun. Mag.*, vol. 51, no. 2, pp. 98-105, Feb. 2013.

[5] Cisco System, A Cisco White Paper, Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2012-2017, San Jose, CA, USA, Feb. 2013.

[6] Ericsson, Ericsson White paper, 5G Radio Access, Research and Vision, Jun. 2013.

[7] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, Multi-cell MIMO cooperative networks: A new look at interference, *IEEE J. Sel.Areas Commun.*, vol. 28, no. 9, pp. 1380-1408, Dec. 2010.

[8] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, Network coordination for spectrally efficient communications in cellular systems, *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 56-61, Aug. 2006.

[9] R. W. Heath, T. Wu, Y. H. Kwon, and A. C. K. Soong, Multiuser MIMO in distributed antenna systems with out-of-cell interference, *IEEE Trans.Signal Process.*,vol. 59, no. 10,pp.4885-4899,Oct.2011.

[10] K. Hosseini, W. Yu, and R. Adve, Large-scale MIMO versus network MIMO for multicell interference mitigation, *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, pp. 930-941, Oct 2014.

[11] S. A. Ramprashad and G. Caire, Cellular vs. network MIMO: A comparison including the channel state information overhead, in *Proc. IEEE Annu. Int. Symp. Personal, Indoor, Mobile Radio Commun.* (PIMRC), Sep. 2009, pp. 878-884.

[12] T. L. Marzetta, Noncooperative cellular wireless with unlimited numbers of base station antennas, *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590-3600, Nov. 2010.

[13] R. Mueller, M. Vekhapera, and L. Cottatellucci, Blind pilot decontamination, in *Proc. 17th Int. ITG WSA*, Mar. 2013, pp. 1-6.

[14] R. Mueller, M. Vehkapera, and L. Cottatellucci, Analysis of pilot decontamination based on power control, in *Proc. IEEE Veh. Technol. Conf.*, Jun. 2013, pp. 1-5.

[15] M. Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, and M. Tanno, Coordinated multipoint transmission/reception techniques for LTE-Advanced, *IEEE Wireless Commun.*, vol. 17, no. 3, pp. 26-34, Jun. 2010.

[16] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback, *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4509-4537, July 2012.

[17] J. Hoydis, S. ten Brink, and M. Debbah, Massive MIMO in UL/DL of cellular networks: How many antennas do we need? *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160-171, Feb. 2013.

[18] R. Zakhour and D. Gesbert, Distributed multicell-MISO precoding using the layered virtual SINR framework, IEEE Trans. Wireless Commun., vol. 9, no. 8, pp. 2444-2448, Aug. 2010.

[19] R. Zhang and S. Cui, Cooperative interference management with MISO beamforming, *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5450- 5458, 2010.

[20] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Sel.Areas Commun.*, vol. 28, no. 9, pp. 1380-1408, Dec. 2010.

[21] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, Network coordination for spectrally efficient communications in cellular systems, *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 56-61, Aug. 2006.

[22] N. Jindal, MIMO broadcast channels with finite-rate feedback, IEEE Trans. Inf. Theory, vol. 52, no. 11, pp. 5045C5060, Nov. 2006.

[23] M. Maddah-Ali and D. Tse, Completely stale transmitter channel state information is still very useful, IEEE Trans. Inf. Theory, vol. 58, no. 7, pp. 4418C4431, Jul. 2012.

[24] T. L. Marzetta, Noncooperative cellular wireless with unlimited numbers of base station antennas, *IEEE Trans. Wireless. Commun.*, vol. 9, no. 11, pp. 3590-3600, Nov. 2010.

[25] H. Q. Ngo, E. G. Larsson and T. L. Marzetta "Energy and spectral efficiency of very large multiuser MIMO systems", *IEEE Trans. Commun.*, vol. 61, no. 4, pp.1436 -1449 2013

[26] F. Rusek, Scaling up MIMO: Opportunities and challenges with very large arrays, *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp.40 -60 2013

[27] Ibrahim, H., ElSawy, H., Nguyen, U.T. and Alouini, M.-S, Modeling virtualized downlink cellular networks with ultra-dense small cells, *IEEE International Conference on Communications (ICC 2015)*, pp. 5360 - 5366.

[28] Q. Ren, J. Fan, X. Luo and Z. Xu , Analysis of spectral and energy efficiency in ultra-dense network, *IEEE International Conference on Communication Workshop (ICCW 2015)*, pp. 2812 - 2817.

[29] H. Quoc Ngo, Ashikhmin, A., Larsson, E.G., H. Yang, and Marzetta, T.L. Cell-free massive MIMO: uniformly great service for everyone, *Proc. 16th IEEE International Workshop on Signal Processing Advances in Wireless Commun.*, June 2015.

[30] S. Lakshminaryana, M. Debbah and M. Assaad, Coordinated Multi-cell Beamforming for Massive MIMO: A Random Matrix Theory Approach, *IEEE Transactions on Information Theory* , vol. 53, issue 8, pp. 4509 - 4537, July 2015.

[31] R. Zhang and S. Cui, Cooperative interference management with MISO beamforming, *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5450- 5458, 2010.

[32] Hoydis, J., S. ten Brink, and M. Debbah, Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need?, *IEEE Journal on Selected Areas in Communications - Large-Scale Multiple Antenna Wireless Systems*, vol. 32, issue 2, pp. 160-171, Feb. 2013.

[33] Wagner, S., R. Couillet, M. Debbah, and D. T. M. Slock, Large System Analysis of Linear Precoding in Correlated MISO Broadcast Channels under Limited Feedback, *IEEE Transactions on Information Theory*, vol. 58, issue 7, pp. 4509 - 4537, 07/2012.

[34] M. Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, and M. Tanno, Coordinated multipoint transmission/reception techniques for LTE-Advanced, *IEEE Wireless Commun.*, vol. 17, no. 3, pp. 26-34, Jun. 2010.

[35] K. Hosseini, W. Yu, and R. Adve, Large-scale MIMO versus network MIMO for multicell interference mitigation, *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, pp. 930-941, Oct 2014.

[36] D. Palomar and J. Fonollosa, Practical algorithms for a family of water- filling solutions, *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 686- 695, Feb. 2005.

[37] W. Yu, Multiuser water-filling in the presence of crosstalk, in *Proc. Inf. Theory Appl. Workshop*, 2007, pp. 414-420.

[38] R. Couillet and M. Debbah Random Matrix Methods for Wireless Communications, 2011 : Cambridge University Press

[39] H. Huh, A. M. Tulino and G. Caire Network MIMO with linear zero-forcing beamforming: Large system analysis, impact of channel estimation, and reduced-complexity scheduling, *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp.2911 -2934 2012

[40] S. Lakshminaryana, J. Hoydis, M. Debbah and M. Assaad "Asymptotic analysis of distributed multi-cell beamforming", *Proc. 2010 IEEE Personal Indoor and Mobile Radio Communications*, pp.2105 -2110

[41] A. Wiesel, Y. C. Eldar and S. Shamai "Linear precoding via conic optimization for fixed MIMO receivers", *IEEE Trans. Signal Process*, vol. 54, no. 1, pp.161 -176 2006

[42] D. Cao, S. Zhou and Z. Niu "Optimal base station density for energy-efficient heterogeneous cellular networks", *IEEE International Conference on Communications (ICC 14)*, pp.4379 -4383

[43] H. ElSawy and E. Hossain "On stochastic geometry modeling of cellular uplink transmission with truncated channel inversion power control", *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp.4454 -4469 2014.

[44] H. S. Jo, Y. J. Sang, P. Xia and J. G. Andrews "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis", *IEEE Transactions on Wireless Communications*, vol. 11, no. 10, pp.3484 -3495 2012

[45] A. Ashikhmin and T. Marzetta "Pilot contamination precoding in multi-cell large scale antenna systems", *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp.1137-1141

[46] M. J. M. Peacock, I. B. Collings and M. L. Honig "Eigenvalue distributions of sums and products of large random matrices via incremental matrix expansions", *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp.2123 -2138 2008

[47] M. Hong, R. Sun, H. Baligh and Z. Luo "Joint base station clustering and beamformer design for partial coordinated transmission in heterogeneous networks", *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp.226 -240 2013

[48] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton and J. Najim "On the capacity achieving covariance matrix for Rician MIMO channels: an asymptotic approach", *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp.1048 -1069 2010

[49] R. Couillet, M. Debbah and J. W. Silverstein "A deterministic equivalent for the analysis of correlated MIMO multiple access channels", *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp.3493 -3514 2011

[50] J. Jose, A. Ashikhmin, T. L. Marzetta and S. Vishwanath "Pilot contamination and precoding in multi-cell TDD systems", *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp.2640 -2651 2011

[51] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Sel.Areas Commun.*, vol. 28, no. 9, pp. 1380-1408, Dec. 2010.

[52] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, Network coordination for spectrally efficient communications in cellular systems, *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 56-61, Aug. 2006.

[53] O. Somekh, O. Simeone, Y. Bar-Ness, A. M. Haimovich, and S. Shamai (Shitz), Cooperative multicell zero-forcing beamforming in cellular downlink channels, *IEEE Trans. Inf. Theory*, vol. 2009, no. 7, pp. 3206?3219, July 2009.

[54] S. Jing, D. N. C. Tse, J. Hou, J. B. Soriaga, J. E. Smee, and R. Padovani, multi-cell downlink capacity with coordinated processing, *EURASIP J. Wireless Commun. Netw.*, vol. 2008, pp. 1?9, 2008.

[55] P de Kerret, D Gesbert, U Salim, Regularized ZF in Cooperative Broadcast Channels under Distributed CSIT: A Large System Analysis, *2015 IEEE International Symposium on Information Theory (ISIT).*

[56] N. Jindal, MIMO broadcast channels with finite-rate feedback, *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045C5060, Nov. 2006.

[57] M. Maddah-Ali and D. Tse, Completely stale transmitter channel state information is still very useful, *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4418C4431, Jul. 2012.

[58] P De Kerret, D Gesbert, Degrees of freedom of the network MIMO channel with distributed CSI, *IEEE Transactions Information Theory,* 2012

[59] R. Mueller, M. Vekhapera, and L. Cottatellucci, Blind pilot decontamination, in *Proc. 17th Int. ITG WSA*, Mar. 2013, pp. 1-6.

[60] R. Mueller, M. Vehkapera, and L. Cottatellucci, Analysis of pilot decontamination based on power control, in *Proc. IEEE Veh. Technol. Conf.*, Jun. 2013, pp. 1-5.

[61] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, Inter-cell interference in non-cooperative TDD large scale antenna systems, *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 192-201, Feb. 2013.

[62] A. Ashikhmin and T. L. Marzetta, Pilot contamination precoding in multi-cell large scale antenna systems, in *IEEE International Symposium on Information Theory (ISIT)*, Cambridge, MA, Jul. 2012.

[63] J Zhang, CK Wen, S Jin, X Gao, Large system analysis of cooperative multi-cell downlink transmission via regularized channel inversion with imperfect CSIT, *IEEE Trans. Wireless. Commun.*, vol. 22, no. 4, pp.1432 -1634 2013

[64] F. Rusek, Scaling up MIMO: Opportunities and challenges with very large arrays, *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp.40 -60 2013

[65] R. Zhang and S. Cui, Cooperative interference management with MISO beamforming, *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5450- 5458, 2010.

[66] Hoydis, J., S. ten Brink, and M. Debbah, Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need?, *IEEE Journal on Selected Areas in Communications - Large-Scale Multiple Antenna Wireless Systems*, vol. 32, issue 2, pp. 160-171, Feb. 2013.

[67] Wagner, S., R. Couillet, M. Debbah, and D. T. M. Slock, Large System Analysis of Linear Precoding in Correlated MISO Broadcast Channels under Limited Feedback, *IEEE Transactions on Information Theory*, vol. 58, issue 7, pp. 4509 - 4537, 07/2012.

[68] R. Couillet and M. Debbah Random Matrix Methods for Wireless Communications, 2011 : Cambridge University Press

[69] H. Huh, A. M. Tulino and G. Caire Network MIMO with linear zero-forcing beamforming: Large system analysis, impact of channel estimation, and reduced-

complexity scheduling, *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp.2911 -2934 2012

[70] H. Q. Ngo, E. G. Larsson and T. L. Marzetta, Energy and spectral efficiency of very large multiuser MIMO systems, *IEEE Trans. Commun.*, vol. 61, no. 4, pp.1436 -1449 2013

[71] H. Ibrahim, H. El Sawy, U. Nguyen, and M. Slim Alouini. Modeling virtualized downlink cellular networks with ultra-dense small cells. In *Communications (ICC), 2015 IEEE International Conference on*, pages 5360–5366. IEEE, 2015.

[72] M Karakayali, G. J Foschini, and R. A Valenzuela. Network coordination for spectrally efficient communications in cellular systems. *Wireless Communications, IEEE*, 13(4):56–61, 2006.

[73] V. Chandrasekhar, J. G Andrews, and A. Gatherer. Femtocell networks: a survey. *Communications Magazine, IEEE*, 46(9):59–67, 2008.

[74] M. Andrews, V. Capdevielle, A. Feki, and P. Gupta. Autonomous spectrum sharing for mixed lte femto and macro cells deployments. In *INFOCOM IEEE Conference on Computer Communications Workshops, 2010*, pages 1–5. IEEE, 2010.

[75] T. L Marzetta. Noncooperative cellular wireless with unlimited numbers of base station antennas. *Wireless Communications, IEEE Transactions on*, 9(11):3590–3600, 2010.

[76] J. Hoydis, S. Ten Brink, and M. Debbah. Massive mimo in the ul/dl of cellular networks: How many antennas do we need? *Selected Areas in Communications, IEEE Journal on*, 31(2):160–171, 2013.

[77] S. Wagner, R. Couillet, M. Debbah, and D. Slock. Large system analysis of linear precoding in correlated miso broadcast channels under limited feedback. *Information Theory, IEEE Transactions on*, 58(7):4509–4537, 2012.

[78] D. Lopez-Perez, S. Güvenç, G. De la Roche, M. Kountouris, T. QS Quek, and J. Zhang. Enhanced intercell interference coordination challenges in heterogeneous networks. *Wireless Communications, IEEE*, 18(3):22–30, 2011.

[79] Q. Ren, J. Fan, X. Luo, Z. Xu, and Y. Chen. Analysis of spectral and energy efficiency in ultra-dense network. In *Communication Workshop (ICCW), 2015 IEEE International Conference on*, pages 2812–2817. IEEE, 2015.

[80] V. Singh, M. Lentz, B. Bhattacharjee, R. La, and M. Shayman. Dynamic frequency resource allocation in heterogeneous cellular networks.

[81] V. R Cadambe and S. Ali Jafar. Interference alignment and degrees of freedom of the-user interference channel. *Information Theory, IEEE Transactions on*, 54(8):3425–3441, 2008.

[82] B. Da and R. Zhang. Exploiting interference alignment in multi-cell cooperative ofdma resource allocation. In *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, pages 1–5. IEEE, 2011.

[83] G. Liu, M. Sheng, X. Wang, W. Jiao, Y. Li, and J. Li. Interference alignment for partially connected downlink mimo heterogeneous networks. *Communications, IEEE Transactions on*, 63(2):551–564, 2015.

[84] Y. Lan, I. Lai, C. Lee, and T. Chiueh. Efficient active precoder identification for receivers with inter-cell interference in heterogeneous networks. *Wireless Communications, IEEE Transactions on*, 14(9):5009–5021, 2015.

[85] B. Guler and A. Yener. Uplink interference management for coexisting mimo femtocell and macrocell networks: An interference alignment approach. *Wireless Communications, IEEE Transactions on*, 13(4):2246–2257, 2014.

[86] R. Zhang and S. Cui. Cooperative interference management with miso beamforming. *Signal Processing, IEEE Transactions on*, 58(10):5450–5458, 2010.

[87] M. Amir, A. El-Keyi, and M. Nafie. Opportunistic interference alignment for multiuser cognitive radio. In *Information Theory (ITW 2010, Cairo), 2010 IEEE Information Theory Workshop on*, pages 1–5. IEEE, 2010.

[88] F. Negro, I. Ghauri, and D. Slock. Spatial interweave for a mimo secondary interference channel with multiple primary users. In *Proceedings of the 4th International Conference on Cognitive Radio and Advanced Spectrum Management*, page 69. ACM, 2011.

[89] P. De Kerret and D. Gesbert. Degrees of freedom of the network mimo channel with distributed csi. *Information Theory, IEEE Transactions on*, 58(11):6806–6824, 2012.

[90] W. Shin, W. Noh, K. Jang, and H. Choi. Hierarchical interference alignment for downlink heterogeneous networks. *Wireless Communications, IEEE Transactions on*, 11(12):4549–4559, 2012.

[91] R. Couillet and M. Debbah. *Random matrix methods for wireless communications*. Cambridge University Press, 2011.

[92] S. Lakshminarayana, M. Assaad, and M. Debbah. Coordinated multicell beamforming for massive mimo: A random matrix approach. *Information Theory, IEEE Transactions on*, 61(6):3387–3412, 2015.

[93] S. Lakshminaryana, J. Hoydis, M. Debbah, and M. Assaad. Asymptotic analysis of distributed multi-cell beamforming. In *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*, pages 2105–2110. IEEE, 2010.

[94] C. Guthy, W. Utschick, and M. Honig. Large system analysis of sum capacity in the gaussian mimo broadcast channel. *Selected Areas in Communications, IEEE Journal on*, 31(2):149–159, 2013.

[95] H. Huang and V. Lau. Partial interference alignment for-user mimo interference channels. *Signal Processing, IEEE Transactions on*, 59(10):4900–4908, 2011.

[96] V. Marchenko and L. Pastur. Distribution of eigenvalues for some sets of random matrices. *Matematicheskii Sbornik*, 114(4):507–536, 1967.

[97] Tong, Y., et al. Adaptive fusion algorithm of heterogeneous sensor networks under different illumination conditions. *Signal Processing* 126 (2016): 149-158.

[98] Zhang Z, Zhang Z, Liu S. Cross domain boosting for information fusion in heterogeneous sensor-cyber sources[J]. *Signal Processing*, 2016, 126: 180-186.

[99] Kushwaha, M., Oh, S., Amundson, I., Koutsoukos, X., Ledeczi, A. (2008, August). Target tracking in heterogeneous sensor networks using audio and video sensor fusion. In *Multisensor Fusion and Integration for Intelligent Systems*, 2008.

[100] Iyengar, S. G., Varshney, P. K., Damarla, T. (2011). A parametric copula-based framework for hypothesis testing using heterogeneous data. *IEEE Transactions on Signal Processing*, 59(5), 2308-2319.

[101] Wu, Pingping, et al. A Novel Lip Descriptor for Audio-Visual Keyword Spotting Based on Adaptive Decision Fusion. *IEEE Transactions on Multimedia* 18.3 (2016): 326-338.

[102] Wu, Y., Chang, E. Y., Chang, K. C. C., Smith, J. R. (2004, October). Optimal multimodal fusion for multimedia data analysis. In *Proceedings of the 12th annual ACM international conference on Multimedia* (pp. 572-579). ACM.

[103] Bramon, R., Boada, I., Bardera, A., Rodriguez, J., Feixas, M., Puig, J., Sbert, M. (2012). Multimodal data fusion based on mutual information. *IEEE transactions on visualization and computer graphics*, 18(9), 1574-1587.

[104] Iyengar, S. G., Niu, R., Varshney, P. K. (2012). Fusing dependent decisions for hypothesis testing with heterogeneous sensors. *IEEE Transactions on Signal Processing*, 60(9), 4888-4897.

[105] He, Hao, and Pramod K. Varshney. A Coalitional Game for Distributed Inference in Sensor Networks with Dependent Observations. *IEEE Transactions on Signal Processing* 64.7 (2016): 1854-1866.

[106] Checka, N., Wilson, K., Rangarajan, V., Darrell, T. (2003, June). A probabilistic framework for multi-modal multi-person tracking. In *Computer Vision and Pattern Recognition Workshop, 2003.* CVPRW'03. Conference on (Vol. 9, pp. 100-100). IEEE.

[107] Glotin H, Vergyr D, Neti C, et al. Weighting schemes for audio-visual fusion in speech recognition[C] *IEEE International Conference on Acoustics, Speech, and Signal Processing,*. IEEE, 2001, 1: 173-176.

[108] Kay, S., Q. Ding, and M. Rangaswamy. Sensor integration by joint PDF construction using the exponential family. *IEEE Transactions on Aerospace and Electronic Systems* 49.1 (2013): 580-593.