Sparse Beamforming Design for Network MIMO System with Limited Backhaul

by

Binbin Dai

A thesis submitted in conformity with the requirements for the degree of Master of Applied Science
Graduate Department of Electrical and Computer Engineering
University of Toronto

© Copyright 2014 by Binbin Dai
Abstract

Sparse Beamforming Design for Network MIMO System with Limited Backhaul

Binbin Dai
Master of Applied Science
Graduate Department of Electrical and Computer Engineering
University of Toronto
2014

This thesis considers a downlink multicell cooperation model in which the base-stations (BSs) are connected to a central processor via rate-limited backhaul links and each scheduled user is cooperatively served by a potentially overlapped cluster of BSs. Two different problem formulations are considered respectively: optimal tradeoff between the total transmit power and backhaul capacity under fixed user rate constraints, and utility maximization for given per-BS power and backhaul constraints. Motivated by the compressive sensing literature, we propose to approximate the backhaul rate as a function of the weighted $\ell_2$-norm square of the beamformers. This novel idea allows the optimal tradeoff to be converted into a weighted power minimization problem, which then can be solved efficiently using the well-known uplink-downlink duality approach; it also makes the utility maximization problem solvable through a generalized weighted minimum mean square error approach. The effectiveness of the proposed algorithms is justified through numerical simulation results.
To my mother and brother.
Acknowledgements

It is my privilege to have the opportunity to study at the University of Toronto and work with a group of incredible fellows. First, I thank my supervisor Prof. Wei Yu for his patience, support and devotion throughout my study. It is my great honor to have him as not only an advisor of study but also a mentor of my life. I have benefited a lot from his passion and insights on research.

I would also like to thank my thesis defense committee members, Prof. Ashish Khisti, Prof. Ravi Adve and Prof. Jianwen Zhu, for taking time to read my thesis and provide valuable comments.

It is a lot of pleasure working in BA4162. I am highly inspired by the discussions with Gokul Sridharan, Yuhan Zhou, Pratik Patil, Siyu Liu, and Yicheng Lin. Many thanks to my dear friends, Louis Tan, Soroush Tabatabaei, Chunpo Pan, Alice Gao, etc, for bringing me many joys over the past two years.

To my mother and brother, for your unconditional love and support.

Last but not least, I express my gratitude to Dan Fang, for all her endeavors during the past 4 years we spent together in filling up my life with wonderful adventures.
# Contents

1 **Introduction**  
   1.1 Thesis Outline ........................................ 7  
   1.2 Notations ............................................. 7

2 **Sparse Beamforming Design for Resources Minimization**  
   2.1 System Model and Problem Formulation ................. 9  
   2.2 Sparse Beamforming Design Algorithms ................ 12  
      2.2.1 Reweighted Power Minimization .................... 12  
      2.2.2 Comparison with Algorithms in [1] ................. 16  
   2.3 Simulation Results .................................... 18  
   2.4 Conclusion ............................................. 22

3 **Sparse Beamforming Design for Utility Maximization**  
   3.1 System Model and Problem Formulation ................. 25  
   3.2 Proposed Algorithm ................................... 28  
   3.3 Numerical Analysis .................................... 33  
      3.3.1 Iterative Link Removal ............................ 34  
      3.3.2 Iterative User Pool Shrinking .................... 35  
      3.3.3 Convergence Check ............................... 37  
      3.3.4 Performance Evaluation ......................... 38  
   3.4 Conclusion ............................................. 42
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Simulation parameters for 7-cell wrapped-around homogenous networks</td>
<td>19</td>
</tr>
<tr>
<td>3.1</td>
<td>Simulation parameters for two-tier heterogeneous networks</td>
<td>32</td>
</tr>
<tr>
<td>3.2</td>
<td>Average maximum backhaul and 50th-percentile rate comparison</td>
<td>39</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Fixed clustering scheme: a fixed subset of BSs are grouped into a larger cluster to coordinately serve all the users within the coverage. .......................... 3
1.2 User-centric clustering: each user individually selects a favorable subset of neighboring BSs as its serving cluster. ........................................... 4

2.1 Network MIMO system with limited sum backhaul capacity. ................ 10
2.2 7-cell wrap-around topology with 3 users/cell randomly distributed. ...... 20
2.3 Transmit powers of the 7 BSs for serving user 3 in cell 2 under SINR=15dB and \( \eta = 0 \). ................................................................. 21
2.4 The tradeoff between the total number of cooperation links and the total transmit power across the 7 BSs under SINR = 5dB. ....................... 22
2.5 The tradeoff between the total number of cooperation links and the total transmit power across the 7 BSs under SINR = 15dB. ....................... 23

3.1 Network MIMO system with limited per-BS backhaul capacity. .......... 26
3.2 7-cell wrapped around two-tier heterogeneous network. ................. 33
3.3 Power distributions of the strongest 8 BSs for serving user 3 in cell 2 with \((C_{\text{macro}}, C_{\text{pico}}) = (245, 70)\) Mbps, \( \alpha_k = 1, \forall k, L_c = 8 \). ............................................. 34
3.4 User rates trajectories for all the 30 users in cell 2 with \((C_{\text{macro}}, C_{\text{pico}}) = (245, 70)\) Mbps, \( \alpha_k = 1, \forall k, L_c = 8 \). Each curve corresponds to a user rate in cell 2. ................................................................. 35
3.5 Total number of survived users at each iteration with $\alpha_k = 1, \forall k, L_c = 8$. 36

3.6 Convergence behavior of the sum rate for the proposed algorithms with $\alpha_k = 1, \forall k, L_c = 4$. 37

3.7 Convergence behavior of $\sum_k \log(\bar{R}_k)$ for the proposed algorithms, where $\bar{R}_k$ is the long term average rate for user $k$ and $L_c = 4$. 38

3.8 Cumulative distribution of user rates in Mbps with $L_c = 8$. 40

3.9 Probability density distribution of macro-BS backhaul consumption in Mbps with $L_c = 8$. 41

3.10 Probability density distribution of pico-BS backhaul consumption in Mbps with $L_c = 8$. 42
Chapter 1

Introduction

Wireless cellular networks are increasingly deployed with progressively smaller cell sizes in order to support the ever increasing demand for high-speed data. As a consequence, intercell interference is now the main physical-layer bottleneck in cellular networks. Multicell cooperation, which allows neighboring base-stations (BSs) to cooperate with each other for joint precoding and joint processing of user data, is a promising technology for intercell interference mitigation [2]. This emerging architecture, also known as network multiple-input multiple-output (MIMO) [3], has the potential to significantly improve the overall throughput of the cellular network.

The idealized implementation of multicell cooperation, where all BSs in the entire network cooperate and share the data for all users, is clearly impractical. One way to implement multicell cooperation in practice is to connect all the BSs with a central processor (CP) via rate-limited backhaul links. For downlink transmission, the CP then only needs to distribute the user’s data to its serving BSs. Roughly speaking, there exist two schemes to determine the set of serving BSs for each user: fixed clustering and user-centric clustering, as illustrated in Fig. 1.1 and 1.2. In fixed clustering scheme, a fixed set of neighboring BSs are grouped together into a larger cluster to cooperatively serve the users within their coverage area. Although fixed clustering scheme has already shown
reasonable performance gain, in such a scheme users at the cluster edge still suffer from considerable inter-cluster interference which limits the benefit of network MIMO [4]. In user-centric clustering where the BS clusters are not fixed but are determined for each user individually, each user dynamically selects a set of favorable BSs; these BSs then cooperatively serve the user using joint precoding techniques. The benefit of user-centric clustering is that it has no explicit cluster edges.

Another advantage of user-centric clustering is that the hand-off problem for high-mobility users can be easily taken care of since the neighboring BSs for a moving user usually overlap from instant to instant. Instead, for fixed clustering scheme, the serving BSs for a user may totally change once the user is moving from one cluster to another. User-centric clustering scheme can keep the scheduled user always being served by some neighboring BSs.

To determine the best set of serving BSs for each user is not a straightforward task, because from the users’ perspective, each user wishes to be served cooperatively by as many BSs as possible, while from the BSs’ perspective, serving more users consumes more power and backhaul capacity. There exists therefore a tradeoff between the user rates, the transmit power, and the backhaul capacity. Further, the beamformer design problem for the network MIMO system with user-centric clustering is also nontrivial, because the sets of BSs serving different users may overlap as illustrated in Fig. 1.2. The traditional zero-forcing (ZF) beamforming [5] and minimum mean square error (MMSE) beamforming [6] designs specifically developed for the single cell case need to be generalized.

This thesis solves the above problem by designing a sparse beamforming vector for each user, where the nonzero entries correspond to the serving BSs. Specifically, two different problem formulations are considered respectively in Chapter 2 and Chapter 3:

- Optimal tradeoff between the total transmit power and the sum backhaul capacity over all the BSs under fixed signal-to-interference-and-noise (SINR) constraints at the remote users;
Figure 1.1: Fixed clustering scheme: a fixed subset of BSs are grouped into a larger cluster to coordinately serve all the users within the coverage.

- Network utility maximization under per-BS power constraints and per-BS backhaul constraints.

Both problem formulations involve a nonconvex $\ell_0$-norm structure corresponding to the backhaul consumption. Motivated by the compressive sensing literature [7], this thesis proposes to iteratively approximate the discrete $\ell_0$-norm as the weighted $\ell_2$-norm square of the beamformers, which can be interpreted as transmit powers from the BSs. This key observation allows the optimal power-backhaul tradeoff problem considered in Chapter 2 to be converted into a weighted power minimization problem with SINR constraints, which can then be solved using the uplink-downlink duality approach [8, 9] with low computational complexity; it also enables the per-BS backhaul constraint to be approximated into a weighted per-BS power constraint so that the utility maximization problem

![Central Processor Diagram](image-url)
considered in Chapter 3 is solvable through a generalized version of the weighted MMSE (WMMSE) approach, which was originally developed for the non-cooperating MIMO interfering broadcast channel [10,11].

The sparse beamforming design problem under fixed user rate constraints has already been addressed in the literature. In a recent work [12], the authors propose to approximate the discrete $\ell_0$-norm through a series of smooth exponential functions. Alternatively, [13] uses the $\ell_1$-norm of the beamforming vector to approximate the cluster size, which is further improved by reweighting in [1]. In both [13] and [1], the cluster size is determined from the $\ell_2$-norm of the beamformers at each BS, and the resulting optimization problem becomes a second-order cone programming (SOCP) problem [14], which can be solved numerically by the interior-point method. To reduce the compu-
tational complexity of the interior-point method, the authors of [1] further propose a second algorithm, which first solves the sum power minimization problem, then iteratively removes the links corresponding to the least link transmit power.

The work in Chapter 2 differs from the previous literature by approximating the discrete backhaul capacity (not just cluster size) with the weighted $\ell_2$-norm-square of the beamformers. One of the main contributions of this thesis is to show that by working with the weighted $\ell_2$-norm-square of the beamformers (which is equivalent to power), instead of the $\ell_2$-norm itself, the optimal tradeoff between the transmit power and backhaul capacities can be reformulated as a weighted power minimization, for which the well-known uplink-downlink duality approach can be used to solve the problem efficiently. The newly proposed algorithm can therefore be thought of as combining iterative reweighting with the low-complexity feature of the weighted power minimization formulation. Further, a new weight updating rule in $\ell_1$ reweighting is adopted, which is different from the weight updating method of [1], where the weights are chosen as inversely proportional to the beamformer entries. The new weight updating rule enables the proposed algorithm to achieve a better tradeoff between the sum power and the sum backhaul capacity, as compared to the algorithms in [1].

Network utility optimization problem for network MIMO system also has been considered in previous literature. For instance, sum rate maximization for fixed clustering scheme is studied in [15] where the block diagonalization precoding method originally designed for the MIMO broadcast channels is generalized to accommodate inter-cluster interference mitigation. Utility maximization is considered respectively in [16,17] for predetermined user-centric clustering and in [18] for dynamic user-centric clustering. The authors in [16] propose to approximate the nonconvex rate expression using the first-order Taylor expansion to transform the problem into a convex optimization problem while [17] resorts to the generalized version of WMMSE approach to find a local optimal solution. Joint beamforming and user-centric clustering design is investigated in [18].
by imposing a $\ell_2$-norm approximation of the cluster size as a penalized item onto the traditional weighted sum rate (WSR) maximization problem. Placing the cluster size constraint onto the objective function results in the power constraints separable between the BSs, which makes the existing block coordinate descent (BCD) algorithm [19] applicable. From the system design perspective, however, this also makes it hard to control the backhaul consumption at each BS since one has to carefully choose the price terms to make the final beamforming vector have the desired sparsity. Furthermore, [18] restricts the candidate BSs serving each user within each cell. This restriction shares the common drawback as fixed clustering in Fig. 1.1 that the users at the cluster edge may still suffer from considerable inter-cell interference.

In contrast to the existing works, one of the main contributions of this thesis is that in Chapter 3 an explicit per-BS backhaul constraint is considered in the network utility maximization formulation so that the system can easily control the backhaul consumption at each BS and the user-centric clustering is formed dynamically for each user. We first approximate the discrete backhaul constraint by its weighted $\ell_1$-norm and then take advantage of the WMMSE approach to propose an iterative algorithm aiming for a local optimum. To further improve the efficiency of the proposed algorithm, we propose to exclude those users with negligible rates out of the user scheduling pool and those BSs with negligible transmit power out of the candidate BS set iteratively. The effectiveness of the proposed algorithm is justified by numerical simulation results in terms of convergence speed, traffic load balancing, system throughput improvement and efficiency in backhaul resource utilization.

The idea of compressive sensing has been applied to various scenarios in communication system design. In [20], the authors design sparse MMSE receivers for the uplink multicell cooperation model using the $\ell_1$-norm approximation, while [21] uses similar ideas for joint power and link admission control in an interference channel. Moreover, [22] applies the idea to the green cloud radio access network (Cloud-RAN) to jointly minimize
the transmit power from the BSs and the transport power from the backhaul links. In this thesis, we adopt the compressive sensing idea to deal with the cluster formulation problem in network MIMO system, where the discrete $\ell_0$-norm is approximated by the reweighted $\ell_2$-norm square of the beamformers. By adopting this approximation approach, we show that the network MIMO system designs with limited backhaul can be much simplified.

1.1 Thesis Outline

The rest of this thesis is organized as follows. Chapter 2 aims to design a sparse beamforming vector for each user to achieve the optimal tradeoff between the sum power and sum backhaul capacity over all the BSs under fixed user SINR constraints. Chapter 3 proposes a low-complexity sparse beamforming design algorithm to maximize the network utility under per-BS power constraint and per-BS backhaul constraint. Finally in Chapter 4, a summary of this thesis is presented and a few possible extendible works are also provided.

1.2 Notations

In this thesis, we use lower-case bold letters (e.g. $\mathbf{w}$) to denote vectors and upper-case bold letters (e.g. $\mathbf{H}$) to denote matrices. All unbold letters represent scalars. $\mathbb{R}$ and $\mathbb{C}$ stand for real and complex domain respectively. The matrix inverse, conjugate transpose and $\ell_p$-norm of a vector is denoted as $(\cdot)^{-1}$, $(\cdot)^H$ and $|\cdot|_p$. $\mathcal{CN}(\cdot, \cdot)$ refers to a complex Gaussian random variable with mean in the first entry and variance (or covariance matrix) in the second entry. We denote $\text{diag}\{\cdots\}$ as a block diagonal matrix with the entries as the block diagonal elements and $\text{Re}\{\cdot\}$ as the real part of the entry. $\mathbb{E}[\cdot]$ is referred as the expectation of a random variable.
Chapter 2

Sparse Beamforming Design for Resources Minimization

Transmit beamformers aim at spatially separating the data streams on the same frequency/time slot to different users. Optimal transmit beamformers designs based on minimizing the transmit power under fixed user quality-of-service (SINR or equivalently data rate) have been attracting interest in research for decades. For instance, optimal transmit beamformers to minimize the total power subject to SINR constraint for each user have been found in [23] while the more practical per-antenna power minimization problem has been solved by [8] using Lagrangian Duality approach in optimization theory, which has also been generalized to the multicell setup in [9] for both sum power minimization and per-antenna power minimization problems.

The beamforming and resource allocation problem for the network MIMO system differs from that of the conventional cellular network is that the capacity limits of the backhaul links need to be explicitly taken into account. In this thesis, we adopt a data-sharing strategy for the network MIMO system, where the user data is made available at every one of the user’s serving BSs. In this setting, there exists a tradeoff between the backhaul capacities and transmit power under fixed user rate constraints: the higher
backhaul capacities, the more BSs could cooperate to serve each user so that interference between different users could be more efficiently mitigated and then less transmit power will be needed. Backhaul capacities and transmit power have to be jointly considered in network MIMO system resource allocation problems.

Pure power minimization in conventional cellular network can be formulated as a convex optimization problem. However, this is not true for joint power and backhaul capacity optimization in network MIMO systems since the backhaul consumption is a function of the cluster formation, which is a discrete variable. Finding the global optimum for a mixed continuous and discrete optimization problem is highly nontrivial. In this chapter, we only focus on finding a “good” local optimum numerically in an efficient manner.

2.1 System Model and Problem Formulation

Consider the downlink of a network MIMO system with $L$ BSs connected to a CP via limited backhaul links with a total capacity limit $C$, as depicted in Fig. 2.1. Suppose that each BS has $M$ antennas and there are $K$ single antenna users, and the CP has access to all user data and channel state information (CSI) in the system. Although a fully cooperative network MIMO system, where every single user is served by all the $L$ BS’s, can dramatically reduce the intercell interference, it also requires very high backhaul capacity, because the CP needs to make every user’s data available at every BS. A more practical architecture is that each user selects only a subset of serving BSs (which are potentially overlapping) and then CP only distributes the user’s data to its serving BSs.

Let $\mathbf{w}_k \in \mathbb{C}^{ML \times 1} = [\mathbf{w}_k^1, \mathbf{w}_k^2, \cdots, \mathbf{w}_k^L]$ be the transmit beamformer over all the BSs for user $k$, where $\mathbf{w}_k^l \in \mathbb{C}^{M \times 1}$ is the transmit beamformer from BS $l$ ($l = 1, 2, \cdots, L$) to user $k$ ($k = 1, 2, \cdots, K$). Note that $\mathbf{w}_k^l = \mathbf{0}$ if BS $l$ is not part of user $k$’s serving cluster.
The received signal $y_k \in \mathbb{C}$ at user $k$ can be written as:

$$y_k = h_k^H w_k s_k + \sum_{j \neq k}^K h_k^H w_j s_j + n_k$$  \hspace{1cm} (2.1)$$

where $h_k \in \mathbb{C}^{ML \times 1}$ denotes the CSI vector from all the BSs to user $k$, $s_k \sim \mathcal{CN}(0, 1)$ and $n_k \sim \mathcal{CN}(0, \sigma^2)$ are the intended signal and the receiver noise for user $k$ respectively.

Assuming the data streams between different users and the noise are mutually independent, the SINR for user $k$ can be expressed as:

$$\text{SINR}_k = \frac{|h_k^H w_k|^2}{\sum_{j \neq k} |h_k^H w_j|^2 + \sigma^2}. \hspace{1cm} (2.2)$$
The achievable rate for user $k$ is then

$$R_k = \log(1 + \text{SINR}_k). \tag{2.3}$$

Since each user’s data only needs to be made available at its serving BSs, the backhaul capacity consumption $C_k$ needed for serving user $k$ can be represented as

$$C_k = \| [\|w^1_k\|_2, \|w^2_k\|_2, \cdots, \|w^L_k\|_2] \|_0 R_k \tag{2.4}$$

where $\| \cdot \|_0$ denotes the $\ell_0$-norm of a vector\footnote{Strictly speaking, $\| \cdot \|_0$ is not a norm by definition since it does not satisfy the homogenous property. However, we still call it $\ell_0$-norm by convention.}, i.e. the number of nonzero entries in the vector.

We can now formulate an optimization problem that relates various network resources and the system throughput. The network resources considered in this thesis consist of the backhaul capacities and the transmit powers at the BSs. Clearly more resources lead to higher throughput. But at fixed user throughput, there is also a tradeoff between the backhaul capacity and the transmit power. Intuitively, higher backhaul capacity allows for more BSs to cooperate, which leads to less interference; hence less transmit power is needed to achieve a target user rate.

Formally, the tradeoff between the total transmit power and the sum backhaul capacity over all BSs under fixed user data rates can be formulated as the following optimization problem:

$$\begin{align*}
\text{minimize} & \quad \{w^l_k\} \\
\text{subject to} & \quad \text{SINR}_k \geq \gamma_k, \quad \forall k
\end{align*} \tag{2.5}$$

where $\gamma_k$ is the SINR target for user $k$ and $\eta \geq 0$ is a constant indicating the tradeoff
between sum backhaul capacity and sum power. By setting different $\eta$’s, we can get different total transmit power and sum backhaul capacity pair to achieve the same set of SINR targets. Note that at the optimal point, the SINR constraint in Problem (2.5) must meet with equality. To see this, suppose at the optimum there exists a user $k$ with SINR strictly larger than the target $\gamma_k$, i.e. $\text{SINR}_k > \gamma_k$, then we can reduce the transmit power for user $k$ to further decrease the objective value but still keep all the SINRs above the targets. Therefore, the rate in the objective function can be treat as a constraint $R_k = \log_2(1 + \gamma_k)$.

Problem (2.5) is nonconvex due to the discrete nature of $\ell_0$-norm, this thesis focuses on a numerical solution to this problem. It is worth noting that the above problem formulation is not the only possibility here. In the next chapter, we will consider a different problem formulation to maximize the network utility under given radio resources.

### 2.2 Sparse Beamforming Design Algorithms

The optimization problem (2.5) is nonconvex due to the $\ell_0$-norm representation of the backhaul rate. Finding the global optimal solution to (2.5) is difficult. Motivated by the compressive sensing literature, we propose to solve (2.5) heuristically by iteratively relaxing the $\ell_0$-norm as its weighted $\ell_1$-norm. In this section, we first introduce our proposed algorithm and then compare it with existing algorithms.

#### 2.2.1 Reweighted Power Minimization

First, we make an observation that if the $\ell_2$-norm in (2.4) is replaced by $\ell_2$-norm square, the overall $\ell_0$-norm remains the same. Thus, the backhaul consumption $C_k$ can also be written as

$$C_k = \|\|w_k^1\|_2^2, \|w_k^2\|_2^2, \cdots, \|w_k^L\|_2^2\|_0 R_k$$  \hspace{1cm} (2.6)
The basic idea of $\ell_1$-heuristics in compressive sensing is to replace the $\|\cdot\|_0$ norm by $\|\cdot\|_1$ norm in the optimization problem. Applying this idea to (2.6) and further introducing the appropriate weights, $C_k$ can now be approximated as the weighted $\ell_2$-norm square of the beamformers, and the problem (2.5) can now be relaxed as

$$\minimize_{\{w^l_k\}} \sum_k \left( \sum_l \rho^l_k \|w^l_k\|_2^2 \right) R_k + \eta \sum_k \sum_l \|w^l_k\|_2^2$$

subject to $\text{SINR}_k \geq \gamma_k$, $\forall k$  

where $\rho^l_k$ is the weight associated with BS $l$ and user $k$.

Observe that the problem (2.7) can be further rearranged into the following form

$$\minimize_{\{w^l_k\}} \sum_{k,l} \alpha^l_k \|w^l_k\|_2^2$$

subject to $\text{SINR}_k \geq \gamma_k$, $\forall k$  

where $\alpha^l_k = \rho^l_k R_k + \eta$. Since the $\ell_2$-norm square of the beamforming vectors are just the transmit powers at the BSs, the above optimization problem is essentially a weighted power minimization problem.

The weighted power minimization problem (2.8) has been extensively studied in the literature. The key point is that it can be solved efficiently using the well-known uplink-downlink duality approach. One of the main contributions of this thesis is thus the observation that this particular relaxation of $C_k$ as weighted $\ell_2$-norm square results in a problem formulation whose structure can be efficiently exploited by numerical algorithms.

Uplink-downlink duality for weighted power minimization is developed for single cell case in [8] and subsequently generalized to the multicell setting in [9]. This thesis further generalizes duality to the case where the weight associated with each BS-user pair may be different.

Note that the solution to (2.8) for a fixed weight $\rho^l_k$ does not necessarily provide suf-
Chapter 2. Sparse Beamforming Design for Resources Minimization

However, by iteratively updating the weights $\rho_k^l$ and by solving problem (2.8) repeatedly with updated $\rho_k^l$, a sparse network-wide beamforming vector for each user can be obtained eventually, where those entries corresponding to the BSs outside of the optimal serving cluster go to zero in the limit.

The optimal reweighting function to update the weight is still an open problem. Although the authors in [7] suggested a few reweighting functions, the optimal choice may differ from case to case. However, a general rule is to update the weight inversely proportional to the corresponding entry. In this thesis, we adopt the following reweighting function to update $\rho_k^l$

$$\rho_k^l = \frac{1}{\|\mathbf{w}_k^l\|^2_p + \epsilon}$$  \hspace{1cm} (2.9)

where $p$ is some positive exponent and $\epsilon$ is adaptively chosen to be $\epsilon = \max\{\min_{k,l} \|\mathbf{w}_k^l\|_2^2, \tau\}$ and $\tau$ is some small positive constant, and $\mathbf{w}_k^l$ is the beamforming vector from the previous iteration. We will show numerically that with a properly chosen $p$, the reweighting function (2.9) improves upon the performance of previous algorithms.

To completely characterize the proposed algorithm, it is still necessary to give the solution to (2.8) based on the following generalization of uplink-downlink duality.

**Proposition 2.2.1** The downlink weighted power minimization problem (2.8) is equivalent to the following uplink sum power minimization problem in the sense that they have the same optimal solution up to a scalar factor, i.e., $\mathbf{w}_k = \sqrt{\delta_k} \hat{\mathbf{w}}_k, \forall k$:

$$\min_{\lambda_k, \hat{w}_k} \sum_k \lambda_k$$

$$\text{subject to} \quad \frac{\lambda_k |\hat{w}_k^H \mathbf{h}_k|^2}{\sum_{j \neq k} \lambda_j |\hat{w}_j^H \mathbf{h}_j|^2 + \hat{w}_k^H \mathbf{B}_k \hat{w}_k} \geq \gamma_k$$  \hspace{1cm} (2.10)

where $\hat{\mathbf{w}}_k \in \mathbb{C}^{ML \times 1}$ can be interpreted as the receiver beamforming vector of the dual uplink channel and $\lambda_k \geq 0$ has the interpretation of dual uplink power, which is also the Lagrangian dual variable associated with the SINR constraint in (2.8), and $\mathbf{B}_k$ is the dual
uplink noise covariance matrix defined as \( B_k = \text{diag}\{\alpha_k^1I_M, \alpha_k^2I_M, \ldots, \alpha_k^L I_M\}, \forall k \).  

**Proof** We list the detailed proof in Appendix A.

The optimal solution to (2.10) is the MMSE receiver [24], which can be readily written as

\[
\hat{w}_k = \left( \sum_j \lambda_j h_j h_j^H + B_k \right)^{-1} h_k \tag{2.11}
\]

where the dual variable \( \lambda_j \) is to be determined. In addition, to find the optimal solution \( w_k \) to problem (2.8), it is also needed to find the scalar \( \delta_k \) relating \( \hat{w}_k \) to \( w_k \). Note that it’s easy to see that the SINR constraints in both (2.8) and (2.10) must be achieved with equality at the optimal point. This observation provides a way to find \( \lambda_j \), then \( \delta_k \).

Substituting (2.11) into the SINR constraint in problem (2.10) with equality, we can get

\[
\lambda_k = \frac{\gamma_k}{h_k^H \left( \sum_{j \neq k} \lambda_j h_j h_j^H + B_k \right)^{-1} h_k} \tag{2.12}
\]

which can be easily verified by matrix inversion lemma [25]. The expression in (2.12) implies that \( \lambda_k \) can be found numerically by fixed-point method, whose convergence is guaranteed by the fact that the function in (2.12) is a standard function [27]; see [9,28].

Now, by substituting \( w_k = \sqrt{\delta_k} \hat{w}_k \) into the SINR constraint in (2.8) with equality, we get \( K \) linear equations with \( K \) unknowns \( \delta_k, k = 1, 2, \ldots, K \):

\[
\frac{1}{\gamma_k} \delta_k |\hat{w}_k^H h_k|^2 = \sum_{j \neq k} \delta_j |\hat{w}_k^H h_j|^2 + \sigma^2, \quad \forall k. \tag{2.13}
\]

Therefore, \( \delta_k \) can be obtained by solving a system of linear equations:

\[
\delta = F^{-1} \mathbf{1} \sigma^2 \tag{2.14}
\]

\[\text{MMSE receiver in general is not unique. The expression in (2.11) is optimal up to a scalar. See results in [26].}\]
where \( \boldsymbol{\delta} = [\delta_1, \delta_2, \ldots, \delta_K] \), \( \mathbf{F} \) is defined as: \( F_{ii} = \frac{1}{\gamma_i} |\hat{\mathbf{w}}^H_i \mathbf{h}_i|^2 \), and \( F_{ij} = -|\hat{\mathbf{w}}^H_j \mathbf{h}_i|^2 \) for \( i \neq j \), and \( \mathbf{1} \) denotes the all-one vector.

We summarize the proposed algorithm in the following:

**Algorithm 2.1 Sparse Beamforming Design**

*Fix the tradeoff scalar \( \eta \):*

**Initialization:** \( \rho^l_k = 1 \quad \forall k, l \);

**Repeat:**

1. Find the optimal dual variable \( \lambda_k \) according to (2.12) using fixed-point method;
2. Compute the optimal dual uplink receiver beamforming \( \hat{\mathbf{w}}_k \), \( \forall k \) according to (2.11);
3. Update \( \mathbf{w}_k = \sqrt{\delta_k} \hat{\mathbf{w}}_k \), \( \forall k \) with \( \delta_k \) found by (2.14);
4. Update \( \rho^l_k \) according to (2.9).

**Until** convergence

To find a different tradeoff point between total transmit power and sum backhaul, change \( \eta \) and repeat the above steps.

### 2.2.2 Comparison with Algorithms in [1]

As mentioned earlier, the reweighted \( \ell_1 \)-norm approach has already been used to solve the sparse beamforming problem in the previous work [1]. However, [1] uses a formulation which is slightly different from (2.5); it studies the problem of minimizing the total number of BS-user cooperation links, i.e.

\[
N_{\text{total}} = \sum_k ||[\|\mathbf{w}_1^k\|_2, \|\mathbf{w}_2^k\|_2, \cdots, \|\mathbf{w}_L^k\|_2]||_0 \tag{2.15}
\]

subject to SINR and power constraints\(^3\). Moreover, [1] approximates the \( \ell_0 \)-norm in (2.15) directly from its weighted \( \ell_1 \)-norm, which results in a form of weighted \( \ell_2 \)-norm of the beamforming vectors. In contrast, we propose to approximate the backhaul rate in (2.4) using its \( \ell_2 \)-norm square which leads to the weighted \( \ell_2 \)-norm square of the beamforming vectors. In [1], the authors considered per-BS power constraint. For fair comparison with the proposed algorithm, sum power constraint is stated here.
Chapter 2. Sparse Beamforming Design for Resources Minimization

After approximating the $\ell_0$ norm in (2.15) by the weighted $\ell_1$-norm, the resulting optimization problem in [1] becomes an SOCP as follows:

$$\begin{align*}
\text{minimize} & \quad \sum_{k,l} \rho_{k,l} t_{k,l}^l \\
\text{subject to} & \quad \|w_{k,l}^l\|_2 \leq t_{k,l}^l, \quad \forall k, l \\
& \quad \sum_{k,l} (t_{k,l}^l)^2 \leq P, \\
& \quad \text{SINR}_{k} \geq \gamma_{k}, \forall k
\end{align*}$$

(2.16)

where $P$ denotes the total power budget, and the SINR constraint can also be cast into a SOCP form [28].

The key advantage of the algorithm proposed in this thesis is that (2.16) can only be solved numerically using, for example, the interior-point method. The complexity of such a general purpose solver is much higher than the uplink-downlink duality approach proposed in this thesis.

This SOCP approach is referred to as “Algorithm 1” in [1]. In Algorithm 1, sparse beamforming vectors are obtained by repeatedly solving (2.16) with iteratively updated weights $\rho_{k}^l$. In [1], the authors choose to update the weights according to $\rho_{k}^l = \frac{1}{t_{k,l} + \epsilon}$, $\forall k, l$, where $\epsilon$ is some fixed small positive value. Note that this reweighting function can be seen as a special case of (2.9) with $p = 1/2$ and with fixed regularization constant $\epsilon$. Numerically, this thesis shows that by adopting an optimized $p$, the performance of Algorithm 1 can be improved.

To address the complexity issue arising from the need to use the interior-point method to solve (2.16), [1] also proposes an “Algorithm 2”, which first solves the following sum
power minimization problem

$$\begin{align*}
\text{minimize} & \quad \sum_{k,l} \|w^l_k\|_2^2 \\
\text{subject to} & \quad \text{SINR}_k \geq \gamma_k, \forall k
\end{align*}$$

(2.17)

then manually removes the BS-user cooperation links corresponding to the least transmit powers until the desired tradeoff point is achieved.

Note that the problem (2.17) has the same computational complexity as (2.8), as the uplink-downlink duality approach are used in both cases. Rather than manually deleting the BS-user cooperation links with the least transmit powers, the algorithm proposed in this thesis dynamically selects the BS cluster for each user by iteratively updating the weights in (2.8). Numerically, we will show that the iterative reweighting approach gives superior performance as compared to manual link removal. It is also interesting to point out that Algorithm 2 of [1] can be thought of as a special case of the algorithm proposed in this paper, where all the $\alpha_k^l$’s in (2.8) are set to be 1, except for the ones corresponding to the least transmit powers, which are set to $+\infty$ in each iteration. At this point, the proposed algorithm can been treated as an unified work of Algorithm 1 and 2 in [1] since it shares both the flavor of iterative reweighting in Algorithm 1 and the flavor of power minimization in Algorithm 2.

## 2.3 Simulation Results

The effectiveness of the proposed algorithm is validated through simulations based on a 7-cell wrapped-around homogenous cellular network with parameters listed in Table 2.1.

To illustrate how the sparse beamforming vectors are formed by the proposed algorithm, we first set the tradeoff constant $\eta = 0$ (i.e. minimizing the total backhaul only) and the SINR target $\gamma_k = 15$dB for every user. Fig. 2.3 shows the transmit power distri-
Chapter 2. Sparse Beamforming Design for Resources Minimization

Table 2.1: Simulation parameters for 7-cell wrapped-around homogenous networks

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Cellular Layout</th>
<th>Hexagonal 7-cell wrapped-around</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS-to-BS distance</td>
<td>0.8 km</td>
<td></td>
</tr>
<tr>
<td>Number of Tx antennas/BS</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Number of Rx antennas/user</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Number of users/cell</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Background Noise</td>
<td>−162 dBm/Hz</td>
<td></td>
</tr>
<tr>
<td>Distance-dependent path loss</td>
<td>$128.1 + 37.6 \log_{10}(d)$</td>
<td></td>
</tr>
<tr>
<td>Log-normal shadowing</td>
<td>8 dB</td>
<td></td>
</tr>
<tr>
<td>Rayleigh small scale fading</td>
<td>0 dB</td>
<td></td>
</tr>
<tr>
<td>Reweighting function parameter</td>
<td>$p = 4/3, \tau = 10^{-10}$</td>
<td></td>
</tr>
</tbody>
</table>

From Fig. 2.3, we see that as the iterations progress, BSs 2 and 4 form a serving cluster for the user, while all the other BSs eventually drop their transmit power to zero.

In Figs. 2.4 and 2.5, we compare the performance of the proposed algorithm with Algorithms 1 and 2 in [1] in terms of the tradeoff between total transmit power and sum backhaul capacity over all BS’s under SINR=5dB and 15dB respectively. For the proposed algorithm, we simulate a series of different tradeoff constant $\eta$'s to get different points along the tradeoff curve. Specifically, by setting $\eta = 0$ we get the minimal backhaul capacity but maximal power as the right extreme point on the tradeoff curve, while by setting $\eta = +\infty$ we get the minimal transmit power but maximal backhaul capacity as the left extreme point on the tradeoff curve. For Algorithm 1, we set different total power budgets $P$, and for each $P$, we solve problem (2.16) iteratively to find the minimum total number of cooperation links needed, which gives one point on the tradeoff curve. For Algorithm 2, we iteratively remove the BS-user cooperation link corresponding to the least transmit power and find the corresponding minimal total transmit power as in (2.17) until the problem becomes infeasible. Note that in this case since all the users have the same fixed SINR target, optimizing the total backhaul capacity is equivalent to optimizing the total number of cooperation links, as the two are related by a constant $R_k = \log(1 + \text{SINR})$. 

bution over all 7 BS’s for serving user 3 in cell 2 (as shown in the topology in Fig. 2.2).
From Figs. 2.4 and 2.5 we can see that all the algorithms can achieve the full cooperation case with 147 total number of BS-user links, in which every one of the total 21 users is served by all the 7 BSs with the least total transmit power, i.e. about -38dBm/Hz for SINR=5dB case and -27dBm/Hz for SINR=15dB case respectively. By jointly optimizing the backhaul capacity and transmit power, more tradeoff points can be obtained to support the same user SINR targets. The simulation results show that overall the proposed algorithm has almost the same tradeoff performance as Algorithms 1 and 2 of [1] at SINR=5dB, whereas at SINR=15dB, the proposed algorithm outperforms the algorithms of [1]. More specifically, as can be seen in Fig. 2.5, to serve users at the SINR of 15dB, under a total transmit power of -14dBm/Hz across the 7 BSs (corresponding to about -23dBm/Hz average per-BS transmit power), our proposed algorithm can reduce the total number of cooperation links by 5 as compared to the solution provided by Algorithm 2 of [1]. Likewise, at a typical average cluster size of 3 for each user, which corresponds to a total number of active cooperation links of 63, our proposed algorithm can reduce the amount of total transmit power by more than 2dB as compared with Al-
Algorithm 2 of [1]. Note that Algorithm 1 of [1] cannot achieve fewer than 66 cooperation links, which appears to be a significant disadvantage of the SOCP formulation.

As we can see, most of the performance gain of the proposed algorithm comes from the 15dB high SINR case\(^4\). This is because, to achieve high SINR target for each served user, transmit cooperation is of paramount importance to ensure that the inter-user interference is sufficiently mitigated at the receiver. Therefore, there exists more cooperation operating points at high SINR regime and the proposed algorithm can select a better BS clustering scheme to achieve a better tradeoff between the backhaul capacity and the transmit power.

To illustrate the effectiveness of the reweighting function adopted in this paper, we also plot the performance of Algorithm 1 of [1] but with weights updated according to (2.9) for \( p = 4/3 \), denoted as “Algorithm 1 with \( p = 4/3 \)” in Fig. 2.5. We see that “Algorithm 1 with \( p = 4/3 \)” significantly improves the original Algorithm 1, and it now

\(^4\)Similar phenomenon has also been observed in [1] for Algorithm 1 and Algorithm 2 compared with the baseline considered therein.
Finally, it should be emphasized that our proposed algorithm is significantly less complex than Algorithm 1 in [1]. Thus, even at the same performance, the proposed algorithm still has considerable advantage.

2.4 Conclusion

This chapter investigates the tradeoff between total transmit power and sum backhaul capacity over all BSs in a network MIMO system with limited cooperation. By adopting a compressive sensing approach of using reweighted $\ell_1$-norm to approximate the $\ell_0$-norm, we turn the original nonconvex problem into a series of convex weighted power minimization problem, which can be solved using a low-complex uplink-downlink duality approach. The proposed algorithm can efficiently find a sparse network-wide beamforming vector.
for each user where the entries corresponding to the non-serving BSs eventually go down to zero in the iterative process. Simulation results show that the proposed algorithm can achieve a better tradeoff between total transmit power and sum backhaul capacity than existing methods in the high SINR regime.
The previous chapter of this thesis considers the resource minimization problem under fixed user rate constraints. But as variable-rate applications and adaptive modulations are now prevalent, a more important problem formulation needs to be considered in wireless network system design is to maximize the network utility under given radio resource limits. Among the family of network utility functions, WSR has been widely applied to various network control and optimization problems (see a recent review in [29]). Generally speaking, WSR maximization problem is NP hard [30], for which obtaining the global optimality is currently only available through exhaustive search [31, 32]. Low complexity algorithms have also been proposed to solve the problem either in an approximated way or to pursue a good suboptimal solution. For instance, [16] approximates the rate expression using first-order Taylor series expansion to transform the problem into a convex optimization problem while [33] seeks a local optimum by finding a solution to the Karush-Kuhn-Tucker (KKT) conditions of the WSR problem. Alternatively, motivated by the relationship between mutual information and MMSE in information theory, the authors in [10] propose to solve the equivalent WMMSE minimization problem to find
a local optimum of the WSR maximization problem, which is further generalized to a more wider class of utility functions in [11].

The idea of solving WSR maximization problem through its equivalent WMMSE minimization problem has been expanded to various areas in heterogeneous network designs, including joint BS association and beamforming design in [34], joint BS clustering and beamforming design in [18] and joint BS activation and linear transceiver design in [35]. In this chapter, we fully utilize the equivalence between WSR maximization and WMMSE minimization and further apply it into the network MIMO system, where each BS is connected to the CP via a rate-limited backhaul link as depicted in Fig. 3.1. As opposed to the traditional WSR maximization problem which mainly considers the continuous transmit power as the only constraint, the additional backhaul constraint in a discrete $\ell_0$-norm formulation makes the WSR maximization problem in network MIMO system even more challenging. In this chapter, we propose a novel algorithm which can approximately solve the WSR maximization problem under per-BS power constraint and per-BS backhaul constraint with low computational complexity.

### 3.1 System Model and Problem Formulation

Consider a downlink cellular network with $L$ BSs and $K$ users, where each BS has $M$ transmit antennas while each user has $N$ receive antennas. The received signal at user $k$, denoted as $y_k \in \mathbb{C}^{N \times 1}$, can be modeled as

$$y_k = H_k x + n_k$$  \hspace{1cm} (3.1) 

where $x \in \mathbb{C}^{M_t \times 1}$ represents the transmitted signal across all the $M_t = LM$ transmit antennas among the whole network, $H_k \in \mathbb{C}^{N \times M_t}$ denotes the CSI matrix from all the BSs to user $k$, $n_k \in \mathbb{C}^{N \times 1}$ is the received noise and is assumed as $n_k \sim \mathcal{CN}(0, \sigma^2 I)$. For the ease of analysis, it is assumed that each user can be potentially served by any subset
of the $L$ BSs\(^1\). Under this assumption, the transmitted vector signal $\mathbf{x}$ for linear transmit beamforming scheme can be written as

$$\mathbf{x} = \sum_{k=1}^{K} \mathbf{w}_k s_k$$  \hspace{1cm} (3.2)

where $\mathbf{w}_k \in \mathbb{C}^{M_t \times 1} = [\mathbf{w}_{k1}^1, \mathbf{w}_{k2}^2, \cdots, \mathbf{w}_{kL}^L]$ and $s_k \in \mathbb{C}$ denote the beamforming vector and intended data for the $k$’th user, respectively. Again, as in previous chapter, $\mathbf{w}_k^l \in \mathbb{C}^{M \times 1}$ denotes the transmit beamformer from BS $l$ to user $k$ and $\mathbf{w}_k^l = \mathbf{0}$ if BS $l$ is not part of user $k$’s serving cluster. Here, we only consider single stream transmission and assume

\(^1\)In practice, to reduce the computational complexity and overhead channel, one may only consider a few strongest BSs around each user as the candidate serving BSs, which is also the case in the simulations to be presented later on. However, this consideration will not affect the following analysis.
$s_k \in \mathcal{CN}(0,1)$. By substituting (3.2) into (3.1), $y_k$ can be rewritten as

$$y_k = H_k w_k s_k + \sum_{j \neq k}^{K} H_k w_j s_j + n_k. \quad (3.3)$$

Assuming that MMSE receiver $u_k$, defined as

$$u_k = \left( \sum_{j=1}^{K} H_k w_j w_j^H H_k^H + \sigma^2 I \right)^{-1} H_k w_k, \quad (3.4)$$

is applied at the user side, then the achievable data rate for user $k$ will be:

$$R_k = \log \left( 1 + w_k^H H_k^H \left( \sum_{j \neq k}^{K} H_k w_j w_j^H H_k^H + \sigma^2 I \right)^{-1} H_k w_k \right). \quad (3.5)$$

In this chapter, we consider utility maximization under limited radio resources including transmit power and backhaul link capacity. By previously defined notations, the transmit power constraint at BS $l$ can be readily formulated as:

$$\sum_k \|w_k^l\|_2^2 \leq P_l \quad (3.6)$$

where $P_l$ is the power budget at the $l$th BS. Differing from the last chapter which considers sum backhaul constraint across all the BSs, we now explicitly model the per-BS backhaul constraints. In our problem formulation, each BS is connected to the CP via an individually rate-limited backhaul link with capacity constraint $C_l$, as depicted in Fig. 3.1. The per-BS backhaul constraint can be cast as:

$$\sum_k \|\|w_k^l\|_2^2\|_0 R_k \leq C_l \quad (3.7)$$

where $R_k$ is the data rate for user $k$ as defined in (3.5). Intuitively, the backhaul consumption at the $l$th BS is the accumulated user rates served by BS $l$ and $\|\|w_k^l\|_2^2\|_0$.
characterizes whether BS \( l \) serves user \( k \) or not by definition:

\[
\|\|w^l_k\|_2^2\|_0 = \begin{cases} 
0, & \text{if } \|w^l_k\|_2^2 = 0 \\
1, & \text{otherwise}
\end{cases}
\] (3.8)

As for the utility function, we take the WSR utility as the objective but need to point out that the proposed algorithm can be easily extend to any utility function that holds an equivalence relationship with WMMSE minimization (see [11] for a sufficient condition for the equivalence).

We now can formulate the WSR maximization problem for the network MIMO system under per-BS power constraint and per-BS backhaul constraint as follows:

\[
\text{maximize} \quad \sum_k \alpha_k R_k \\
\text{subject to} \quad \sum_k \|w^l_k\|_2^2 \leq P_l, \ \forall l \\
\quad \sum_k \|\|w^l_k\|_2^2\|_0 R_k \leq C_l, \ \forall l
\] (3.9) (3.10) (3.11)

with the unknown \( R_k \) defined in (3.5).

### 3.2 Proposed Algorithm

Conventional WSR maximization problem under a power constraint is a well-known nonconvex problem, for which finding its global optimality is already quite challenging even without the additional discrete backhaul constraint (3.11). In this section, we propose a low-complexity algorithm to solve the above mixed continuous and discrete optimization problem in an \textit{approximate} manner.

First, similar to the previous chapter, we propose to approximate the discrete back-
haul constraint (3.11) into a continuous format by its weighted $\ell_1$-norm:

$$\sum_k \beta_k l_k R_k \|\mathbf{w}_k^l\|_2^2 \leq C_l$$

where $\beta_k l_k$ is a constant weight associated with BS $l$ and user $k$ and will be updated iteratively according to

$$\beta_k l_k = \frac{1}{\|\mathbf{w}_k^l\|_2^2 + \tau}, \forall k, l$$

with $\|\mathbf{w}_k^l\|_2^2$ from the previous iteration and some small constant regularization factor $\tau > 0$.

Even with the above approximation, however, the optimization problem (3.9) with constraint (3.11) being replaced by (3.12) is still intractable due to the fact that the rate $R_k$ in the constraint is an optimization variable. To address this difficulty, we propose to solve the problem (3.9) with fixed rate $\hat{R}_k$ iteratively and update $\hat{R}_k$ simultaneously with $\beta_k l_k$ as

$$\hat{R}_k = R_k, \forall k,$$

where $R_k$ is the achievable user rate obtained from the previous iteration.

Under fixed $\beta_k l_k$ and $\hat{R}_k$, the optimization problem (3.9) now reduces to

$$\text{maximize } \sum_k \alpha_k R_k$$

subject to

$$\sum_k \|\mathbf{w}_k^l\|_2^2 \leq P_l, \forall l$$

$$\sum_k \beta_k l_k \hat{R}_k \|\mathbf{w}_k^l\|_2^2 \leq C_l, \forall l$$

where the backhaul constraint (3.17) can now be interpreted as a weighted per-BS power constraint bearing a resemblance to the traditional per-BS power constraint (3.16).

Although the approximated problem (3.15) is still nonconvex, we observe in this thesis that it is possible to reformulate it as an equivalent WMMSE minimization problem...
in order to arrive at a local optimum of (3.15). The equivalence between WSR maximization and WMMSE minimization was first discovered in [10,11] for MIMO broadcast interfering channel without cooperation between the transmitters. We generalize this result to the case of network MIMO system where each user can be cooperatively served by a potentially overlapped subset of BSs, as elaborated by the following proposition:

**Proposition 3.2.1** The WSR maximization problem (3.15) is equivalent to the following WMMSE minimization problem:

\[
\begin{align*}
\text{minimize} & \quad \sum_{k=1}^{K} \rho_k e_k \\
\text{subject to} & \quad \sum_k \|w_k^l\|_2^2 \leq P_l, \quad \forall l \\
& \quad \sum_k \beta_k^l \hat{R}_k \|w_k^l\|_2^2 \leq C_l, \quad \forall l
\end{align*}
\]

(3.18)

in the sense that any (local) optimal point \( \{w_k^*\} \) in problem (3.15) will also be the optimal point for problem (3.18) and vice versa, where \( \rho_k \) denotes the MSE weight for user \( k \) and \( e_k \) is the corresponding MSE defined as

\[
e_k = E\left[ (u_k^H y_k - s_k) (u_k^H y_k - s_k)^H \right] = u_k^H \left( \sum_{j=1}^{K} H_k w_j H_j^H + \sigma^2 I \right) u_k - 2 \text{Re}\{u_k^H H_k w_k\} + 1 \quad (3.19)
\]

under receiver \( u_k \).

**Proof** We provide the detailed proof in Appendix B.

The advantage of transforming the WSR maximization problem into its equivalent WMMSE minimization problem is that the latter is a convex problem in term of the receive beamformer \( u_k \) under fixed \( w_k \), and also a convex problem in term of the transmit beamformer \( w_k \) under fixed \( u_k \). The optimal receiver for fixed transmit beamformer is
the MMSE receiver given in (3.4), while the optimization problem to find the optimal transmit beamformer under fixed receiver is a quadratically constrained quadratic programming (QCQP) problem, which can be solved using standard convex optimization solvers, such as CVX [36]. It is this crucial observation that allows WMMSE minimization problem to be solved using an iterative algorithm as listed in Algorithm 3.1, which finds a local optimum to the WSR maximization problem. The updating rule for the MSE weight $\rho_k$ is elaborated in Appendix B.

**Algorithm 3.1** WMMSE

**Initialization:** $w_k, \forall k$;

**Repeat:**

1. Compute receive beamformer $u_k$ and MSE $e_k$ according to (3.4) and (3.19) respectively under fixed $w_k, \forall k$;

2. Update MSE weight $\rho_k$ according to $\rho_k = \frac{\alpha_k}{e_k}, \forall k$;

3. Find the optimal transmit beamformer $w_k$ under fixed $u_k$ and $\rho_k, \forall k$.

**Until** convergence

A straightforward way of using Algorithm 3.1 in the overall utility maximization problem (3.9) would involve two loops: an inner loop to solve the approximated WSR maximization problem (3.15) with fixed weight $\beta^l_k$ and rate $\hat{R}_k$ using WMMSE approach in Algorithm 3.1; and an outer loop to update the weight $\beta^l_k$ together with the rate $\hat{R}_k$ according to (3.13) and (3.14). However, this two-loop iterative algorithm could be computationally intensive. To reduce the computational complexity, we further propose to combine the two loops into a single loop algorithm which updates the weight $\beta^l_k$ and rate $\hat{R}_k$ inside of the WMMSE iteration after the transmit beamformer update step, as illustrated in Algorithm 3.2.

Algorithm 3.2 has the same complexity order as the conventional WMMSE approach in Algorithm 3.1 since it only introduces two more additional steps in each iteration to update $\beta^l_k$ and $\hat{R}_k$, both of which are closed-form functions of the transmit beamformers.
Algorithm 3.2 Low Complexity Sparse Beamforming Design for WSR Maximization

**Initialization:** $\beta^l_k$, $\hat{R}_k$, $w_k$, $\forall l, k$;

**Repeat:**

1. Compute receive beamformer $u_k$ and MSE $e_k$ according to (3.4) and (3.19) respectively under fixed $w_k$, $\forall k$;

2. Update MSE weight $\rho_k$ according to $\rho_k = \alpha_k/e_k$, $\forall k$;

3. Find the optimal transmit beamformer $w_k$ under fixed $u_k$ and $\rho_k$, $\forall k$.

4. Compute the achievable rate $R_k$ according to (3.5), $\forall k$;

5. Update $\beta^l_k$ according to (3.13) and $\hat{R}_k = R_k$, $\forall l, k$.

**Until** convergence

<table>
<thead>
<tr>
<th>Table 3.1: Simulation parameters for two-tier heterogeneous networks.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellular</td>
<td>Hexagonal 7-cell wrapped-around</td>
</tr>
<tr>
<td>Channel Bandwidth</td>
<td>10 MHz</td>
</tr>
<tr>
<td>Distance between cells</td>
<td>0.8 km</td>
</tr>
<tr>
<td>Number of users/cell</td>
<td>3</td>
</tr>
<tr>
<td>Number of macro-BSs/cell</td>
<td>1</td>
</tr>
<tr>
<td>Number of pico-BSs/cell</td>
<td>3</td>
</tr>
<tr>
<td>Number of Tx antennas/macro-BS</td>
<td>4</td>
</tr>
<tr>
<td>Number of Tx antennas/pico-BS</td>
<td>2</td>
</tr>
<tr>
<td>Number of Rx antennas/user</td>
<td>2</td>
</tr>
<tr>
<td>Max. Tx Power at macro-BS</td>
<td>43 dBm</td>
</tr>
<tr>
<td>Max. Tx Power at pico-BS</td>
<td>30 dBm</td>
</tr>
<tr>
<td>Antenna Gain</td>
<td>15 dBi</td>
</tr>
<tr>
<td>Background Noise</td>
<td>$-169$ dBm/Hz</td>
</tr>
<tr>
<td>Distance-dependent path loss from macro-BS to user</td>
<td>$128.1 + 37.6 \log_{10}(d)$</td>
</tr>
<tr>
<td>Distance-dependent path loss from pico-BS to user</td>
<td>$140.7 + 36.7 \log_{10}(d)$</td>
</tr>
<tr>
<td>Log-normal shadowing</td>
<td>8 dB</td>
</tr>
<tr>
<td>Rayleigh small scale fading</td>
<td>0 dB</td>
</tr>
<tr>
<td>Reweighting function parameter</td>
<td>$\tau = 10^{-10}$</td>
</tr>
</tbody>
</table>

In the next section, we will consider two more techniques to further reduce the complexity of WMMSE approach per-iteration.
Chapter 3. Sparse Beamforming Design for Utility Maximization

3.3 Numerical Analysis

In this section, numerical experimental results are presented to justify the effectiveness of the proposed algorithm. We consider a 7-cell wrapped around two-tier heterogeneous network with the simulation parameters listed in Table 3.1. Each cell forms a regular hexagon with a single macro-BS in the center and 3 pico-BSs equally separated within the cell as illustrated in Fig. 3.2. To simplify the discussions, we set all the macro-BSs to have equal backhaul constraints and the same for the pico-BSs, denoted as \((C_{\text{macro}}, C_{\text{pico}})\), and correspondingly denote the proposed algorithm as “Proposed Algorithm \((C_{\text{macro}}, C_{\text{pico}})\)Mbps”.

The proposed algorithm is tested under the power constraints listed in Table 3.1 with various sets of \((C_{\text{macro}}, C_{\text{pico}})\) constraints.

In what follows, we propose two more techniques, \textit{iterative link removal} and \textit{iterative user pool shrinking}, to further improve the efficiency of the proposed algorithm in each iteration. The computational complexity for each iteration mainly comes from the optimal
transmit beamformer design under fixed receiver beamformers, which is reformulated as a SOCP and solved by the CVX. For each SOCP problem, the computational complexity using the interior-point method is approximately $O((KLM)^3)$ [37]. The proposed iterative link removal technique aims at reducing the number of potential transmit antennas $M$ serving each user while the iterative user pool shrinking is intended to decrease the total number of users $K$ considered in each iteration.

### 3.3.1 Iterative Link Removal

As indicated previously, instead of considering all the $L$ BSs in the entire network as the candidates serving each user, in simulations we only consider the strongest $L_c$ ($L_c \leq L$) BSs, according to the maximum transmit power compensated by the path-loss, around each user as its candidate set so as to reduce the size of the initial transmit beamformer. Moreover, similar to the phenomenon observed in Fig. 2.3, the transmit power from some
Figure 3.4: User rates trajectories for all the 30 users in cell 2 with $(C_{\text{macro}}, C_{\text{pico}}) = (245, 70)$Mbps, $\alpha_k = 1, \forall k$, $L_c = 8$. Each curve corresponds to a user rate in cell 2.

of the candidate BSs will drop down very close to 0 as the iteration goes on. We plot the power distributions of the strongest 8 BSs for the third user in the second cell as an example in Fig. 3.3. As we can see from Fig. 3.3, after around 20 iterations only two BSs maintain at a reasonable transmit power level and eventually form as the cluster to serve user 3 in cell 2. By taking advantage of this, we propose to iteratively remove the $l$th BS from the $k$’th user’s candidate cluster once the transmit power from BS $l$ to user $k$, i.e. $\|w_k^l\|_2^2$, is below a certain threshold, say $-90$dBm/Hz.

### 3.3.2 Iterative User Pool Shrinking

In order to get a good local optimum to the WSR maximization problem, the proposed algorithm implements user scheduling implicitly by initially considering all the users among the entire network into the optimization procedure but only ending up with a small amount of users with non-negligible rates, which correspond to the scheduled users.
for current set of weights $\alpha_k$'s. For instance, in Fig. 3.4, we plot the user rate trajectories for all the 30 users in cell 2. As we can see, only 9 out of 30 users are being served with non-negligible rates in the end. Therefore, to reduce the computational complexity in each WMMSE iteration, we propose to check the achievable user rates at Step 4 iteratively for the proposed Algorithm 3.2 and ignore those users with negligible rates (below some threshold, say 0.01 bps/Hz) for the next iteration. To demonstrate the effectiveness of this technique, we plot the total number of survived users whose rates are above the threshold at each iteration in Fig. 3.5, from which we can see that after 7 iterations the number of users to be potentially scheduled can be reduced to half of the initial entire user pool. This idea can also be applied to the traditional WMMSE approach in Algorithm 3.1.

Figure 3.5: Total number of survived users at each iteration with $\alpha_k = 1, \forall k, L_c = 8$. 
Figure 3.6: Convergence behavior of the sum rate for the proposed algorithms with $\alpha_k = 1, \forall k, L_c = 4$.

### 3.3.3 Convergence Check

Before we evaluate the effectiveness of the proposed algorithm in improving the system performance, we first check its convergence behavior. For conventional WMMSE algorithm which only alternates between the receive beamformers, MSE weights and transmit beamformers, it is guaranteed to converge to a stationary point (see proof in [11]). As for the proposed algorithm with additional two alternating elements, i.e. weight $\beta_k^l$ and approximated rate $\hat{R}_k$, and with the above mentioned two expediting techniques being taken into account, the convergence can be observed numerically within a reasonable number of iterations. We first set $\alpha_k = 1, \forall k$, and plot the sum rate convergence in Fig. 3.6, where it’s easy to see that the proposed algorithm can converge roughly in around 20 iterations. We also further update the weights $\alpha_k$’s according to the proportional fairness criteria and plot the log-utilities of the proposed algorithm in Fig. 3.7, from which we can also find that the overall algorithm can converge within roughly 40 iterations.
Figure 3.7: Convergence behavior of $\sum_k \log(\bar{R}_k)$ for the proposed algorithms, where $\bar{R}_k$ is the long term average rate for user $k$ and $L_c = 4$.

### 3.3.4 Performance Evaluation

To justify the effectiveness of the proposed algorithm in improving system performance, we consider the user-centric clustering scheme where each user is served by the strongest $S$ BSs ($S = 1, 2, 3, 4$) as the baseline schemes. Note that baseline schemes do not have explicit backhaul constraints\(^2\), the necessary backhaul requirements to support the baselines are only obtainable after the completion of the algorithms. In order to make a fair comparison, we first measure the backhaul requirements for the baseline schemes then set the backhaul constraints correspondingly for the proposed algorithm. However, it is worth pointing out again that the proposed algorithm only solves the original problem (3.9) \textit{approximately} due to the weighted $\ell_1$-norm relaxation of the discrete backhaul constraint in (3.12). In other words, the original backhaul constraint (3.7) may not be

---

\(^2\)To the best knowledge of the author, there is no available literature yet considering an explicit backhaul constraint as the proposed algorithm in network MIMO system design.
Table 3.2: Average maximum backhaul and 50th-percentile rate comparison.

<table>
<thead>
<tr>
<th></th>
<th>Avg. Max. Backhaul ((\bar{C}<em>{macro}, \bar{C}</em>{pico}))Mbps</th>
<th>50th-Percentile Rate in Mbps</th>
<th>Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongest 1 BS</td>
<td>(268, 73)</td>
<td>4.9</td>
<td>42%</td>
</tr>
<tr>
<td>Proposed Algorithm</td>
<td>(266, 76)</td>
<td>6.9</td>
<td></td>
</tr>
<tr>
<td>Strongest 2 BSs</td>
<td>(690, 108)</td>
<td>6.2</td>
<td>34%</td>
</tr>
<tr>
<td>Proposed Algorithm</td>
<td>(689, 112)</td>
<td>8.3</td>
<td></td>
</tr>
<tr>
<td>Strongest 3 BSs</td>
<td>(1197, 173)</td>
<td>8.1</td>
<td>28%</td>
</tr>
<tr>
<td>Proposed Algorithm</td>
<td>(1135, 173)</td>
<td>10.3</td>
<td></td>
</tr>
<tr>
<td>Strongest 4 BSs</td>
<td>(1832, 300)</td>
<td>10.2</td>
<td>25%</td>
</tr>
<tr>
<td>Proposed Algorithm</td>
<td>(1790, 290)</td>
<td>12.7</td>
<td></td>
</tr>
</tbody>
</table>

Exactly met at convergence, where we would allow an \(\epsilon\) excess over the budget\(^3\). Therefore, in order to keep a similar backhaul consumption as the baseline schemes, we set \(C_l/(1 + \epsilon)\) as the backhaul constraint in (3.12) in simulations with \(\epsilon = 0.1\) and list the actual average maximum backhaul consumption for all algorithms in Table 3.2, defined as

\[
\bar{C}_{max} = \frac{\sum_t \max_l C_l(t)}{L}
\]

(3.20)

where \(C_l(t)\) is the backhaul consumption at BS \(l\) for time slot \(t\) and each time slot \(t\) corresponds to a set of weights \(\alpha_k\)'s in WSR maximization. To ease the discussions, we denote the proposed algorithms and baseline schemes as “Proposed Algorithm (\(\bar{C}_{macro}, \bar{C}_{pico}\))Mbps” and “Strongest S BSs (\(\bar{C}_{macro}, \bar{C}_{pico}\))Mbps” respectively in the following comparisons with (\(\bar{C}_{macro}, \bar{C}_{pico}\)) obtained from Table 3.2.

Fig. 3.8 shows the cumulative distributions of the long average user rates with the \(\alpha_k\)'s updated according to the proportional fairness criteria. It’s easy to see that the proposed algorithm can significantly improve the overall user rates compared with connecting the users with fixed number of strongest BSs. For instance, 42% of improvement has been observed at the 50th-percentile rate from the “Proposed Algorithm (266, 76)Mbps” com-

\(^3\)Smaller \(\epsilon\) may result in slower convergence speed for the proposed algorithm.
Figure 3.8: Cumulative distribution of user rates in Mbps with $L_c = 8$.

pared with the baseline “Strongest 1 BS (268, 73)Mbps”. Another more interesting comparison occurs between the “Proposed Algorithm (689, 112)Mbps” and the “Strongest 3 BSs (1197, 173)Mbps” scheme, where both of them have similar user rate distributions. However, the proposed algorithm only requires about 60% of the backhaul of the baseline scheme. To further investigate the reason behind this significant reduction in backhaul consumption while still keeping the similar performance, we plot the probability density distributions of the backhaul consumptions for the macro-BSs and the pico-BSs in Fig. 3.9 and 3.10 respectively.

From Fig. 3.9, we can see that the real-time backhaul consumption at macro-BSs for the proposed algorithm is centered at 620Mbps, which is very near to its average maximum backhaul rate 689Mbps. This means the macro-BS backhaul links are most of time fully utilized by the proposed algorithm. However, for the “Strongest 3 BSs (1197, 173)Mbps” scheme, the macro-BS backhaul consumption is widely spread and the central point is only about half of the average maximum backhaul, which indicates the backhaul consumption for the “Strongest 3 BSs (1197, 173)Mbps” scheme is bursty. Similar conclusion can be


Figure 3.9: Probability density distribution of macro-BS backhaul consumption in Mbps with $L_c = 8$.

drawn for the pico-BS backhaul consumption in Fig. 3.10. The reason for the bursty backhaul consumption in connecting the users with strongest BSs is the load balancing issue since a substantial number of users will connect to the high-power macro-BSs, while the proposed algorithm with explicit backhaul constraints can efficiently solve this problem by adaptively forming the clusters according to the available backhaul budget. This has an effect of balancing the data traffic. At this point, we can conclude that the gain of the proposed algorithm mostly comes from the efficient utilization of the backhaul budget in contrast to the bursty consumption of the backhaul in connecting users with strongest BSs. Therefore, explicit backhaul constraint not only helps the system easily control the backhaul consumption but also leads to the fully utilization of the backhaul resource.
Figure 3.10: Probability density distribution of pico-BS backhaul consumption in Mbps with $L_c = 8$.

3.4 Conclusion

In this chapter, a low-complexity algorithm is proposed to iteratively solve the network utility maximization problem for the network MIMO system under per-BS power constraint and per-BS backhaul constraint. We first adopt the reweighted $\ell_1$-norm approximation technique in compressive sensing to approximate the discrete backhaul constraint into a continuous format and then accommodate the approximated problem into an equivalent WMMSE minimization problem to pursue a local optimum. Simulation results show that the proposed algorithm can converge within a reasonable number of iterations and can fully utilize the backhaul resource to balance the traffic load and to improve the overall system throughput.
Chapter 4

Conclusions

The major difference between network MIMO system design and traditional cellular network design is the additional backhaul constraints for limited information sharing between transmitters, which is also one of main difficulties in setting up network MIMO system since one has to jointly consider the tradeoff between user rates, transmit power and backhaul capacities. Of all the efforts in network MIMO system design, a substantial portion has been working on tackling the backhaul rate constraint, which is not only a function of the (continuous) user rate but also related to the (discrete) number of associated users. Motivated by the recent breakthrough in compressive sensing, we propose to approximate the backhaul rate through the weighted $\ell_2$-norm square of beamforming vectors, whose effectiveness is shown in this thesis from various perspectives.

In chapter 2, we consider the optimal tradeoff between the total transmit power and the sum backhaul capacity across all the BSs in the network under fixed user receive SINR constraints. By applying the compressive sensing idea, we show that the tradeoff problem can be effectively converted into an iteratively weighted power minimization problem, which can be solved efficiently using the well-known uplink-downlink duality approach. We also show numerically the importance of reweighting function in achieving sparsity.
Chapter 3 further applies the compressive sensing idea to the network utility maximization problem in network MIMO system. By iteratively approximating and updating the user rates in backhaul consumption, we transform the per-BS backhaul constraint into a weighted per-BS power constraint and use a generalized WMMSE approach to find a local optimum. In order to expedite the execution of the proposed algorithm, we propose to iteratively drop out the users with negligible rates and the BS-user link with trivial transmit power. Numerically results show that the proposed algorithm can converge within limited number of iterations and can fully utilize the backhaul resource to balance the traffic load and improve the overall throughput.

4.1 Future Work

There are quite a few directions worth investigating in the future. First, although only the sum power and sum backhaul capacity constraints are considered in Chapter 2, the proposed algorithm can be potentially generalized to the more practical setup where each BS has its individual transmit power and backhaul link budget. The only difference is that there will be more dual variables associated with each per-BS power constraint and each per-BS backhaul constraint to be optimized. Even though the subgradient method used in [8,9] will also be applicable here, since the proposed algorithm in Chapter 2 is in an iterative manner, a more efficient method is in need to improve the overall efficiency of the proposed algorithm.

Second, for the utility maximization problem considered in Chapter 3, a significant gain has been observed from the proposed algorithm. To better understand where this gain comes from, we can first fix the scheduled users as the same for both the proposed algorithm and the baseline to see if it’s due to the implicit user scheduling in WMMSE approach. Then check the difference of the clustering schemes between the proposed algorithm and the baseline to find out whether the cluster size or the cluster pattern
Chapter 4. Conclusions

plays a more important role in improving the performance. In-depth understanding of the performance gain can help to devise simpler algorithms.

Third, in this thesis, we only consider user data sharing as the main consumption in backhaul capacities. In practice, CSI sharing also takes a considerable amount of backhaul capacities and needs to be taken into account.

Moreover, although we are assuming a single CP connected with all the BSs in the entire network, a more practical cooperation scheme in reality would be a hierarchical CP system, where each CP in the first-tier only controls a limited set of BSs and is controlled by high-tier CPs. At this point, distributive algorithm which only requires local cluster information will be needed to reduce the information exchange between the CPs and the incurred delay.

Another more challenging topic to be studied would be the information-theoretical analysis on the limits of the network MIMO system. In network MIMO system with a cloud center, the BSs can be effectively considered as relays between the CP and the users. However, since the capacity region for general relay channel is still an open problem in information theory, the available results for the fundamental limits of network MIMO system with limited backhaul are still quite limited.
Appendix A

Proof for Proposition 2.2.1

Problem (2.8) is nonconvex due to the SINR constraint. However, strong duality still holds for (2.8) since the SINR constraint can be reformulated as a SOC constraint [28]. We establish the equivalence between problem (2.8) and (2.10) by first deriving the dual of problem (2.8) and then relating the dual problem with problem (2.10).

The Lagrangian of problem (2.8) can be written as

\[
\mathcal{L}(w_k, \lambda_k) = \sum_k w_k^H B_k w_k - \sum_k \lambda_k \left( \frac{1}{\gamma_k} w_k^H h_k h_k^H w_k - \sum_{j \neq k} w_j^H h_k h_k^H w_j - \sigma^2 \right) \tag{A.1}
\]

where \( \lambda_k \) is the Lagrangian variable associated with the SINR constraint and \( B_k \in \mathbb{R}^{LM \times LM} \) is a diagonal matrix defined as \( B_k = \text{diag}\{\alpha_1^k I_M, \alpha_2^k I_M, \cdots, \alpha_L^k I_M\}, \forall k \).

The dual objective function is then given by

\[
g(\lambda_k) = \min_{w_k} \mathcal{L}(w_k, \lambda_k) \tag{A.2}
\]

\[
g(\lambda_k) = \begin{cases} 
\sum_k \lambda_k \sigma^2, & \text{if } B_k - \frac{\lambda_k}{\gamma_k} h_k h_k^H + \sum_{j \neq k} \lambda_j h_j h_j^H \succeq 0, \forall k \\
-\infty, & \text{otherwise}
\end{cases}
\]
Therefore, we can formulate the dual problem of the primal problem (2.8) as follows:

$$\begin{align*}
\text{maximize} & \quad \sum_k \lambda_k \\
\text{subject to} & \quad B_k + \sum_j \lambda_j h_j h_j^H \preceq \left(1 + \frac{1}{\gamma_k}\right) \lambda_k h_k h_k^H, \forall k
\end{align*}$$ (A.3)

By [8, Lemma 1], the constraint (A.4) is equivalent to

$$\left(1 + \frac{1}{\gamma_k}\right) \lambda_k h_k^H \left(B_k + \sum_j \lambda_j h_j h_j^H\right)^{-1} h_k \leq 1, \forall k$$ (A.5)

Hence, the dual problem (A.3) can be rewritten as

$$\begin{align*}
\text{maximize} & \quad \sum_k \lambda_k \\
\text{subject to} & \quad \left(1 + \frac{1}{\gamma_k}\right) \lambda_k h_k^H \left(B_k + \sum_j \lambda_j h_j h_j^H\right)^{-1} h_k \leq 1, \forall k,
\end{align*}$$ (A.6)

Now, by substituting the optimal MMSE receiver (2.11) into the SINR constraint of problem (2.10), the problem (2.10) can be reformulated as:

$$\begin{align*}
\text{minimize} & \quad \sum_k \lambda_k \\
\text{subject to} & \quad \left(1 + \frac{1}{\gamma_k}\right) \lambda_k h_k^H \left(B_k + \sum_j \lambda_j h_j h_j^H\right)^{-1} h_k \geq 1, \forall k,
\end{align*}$$ (A.7)

Comparing problem (A.6) and (A.7), we can find that the only difference between those two problems is the reverse sign of the objective function and the constraint. However, it’s easy to see that both problems will achieve the optimality at the equality of the
constraint, i.e.

\[
\left(1 + \frac{1}{\gamma_k}\right) \lambda_k h_k^H \left(B_k + \sum_j \lambda_j h_j h_j^H\right)^{-1} h_k = 1, \forall k \tag{A.8}
\]

Hence, problem (A.6) and (A.7) are equivalent to each other due to the uniqueness of the fixed-point [28]. Therefore, the primal problem (2.8) is equivalent to (2.10).
Appendix B

Proof for Proposition 3.2.1

The equivalence between WSR maximization and WMMSE minimization in network MIMO system with partial cooperation was first proved in [17] by transforming the MIMO interference channel with partial message sharing into an equivalent MIMO interference channel with individual message knowledge and generalized linear constraints. In this Appendix, we provide an alternative proof by directly writing down the KKT conditions of the two problems (3.15) and (3.18) and relating the equivalence between them.

First, it is easy to see that the optimal receiver $u_k$ for problem (3.18) with given transmit beamformer is the MMSE receiver defined in (3.4). By substituting (3.4) into (3.19), the optimal MSE for user $k$ can be formulated as a function of the transmit beamformer only:

$$e_k = \left(1 + w_k^H H_k^H \left( \sum_{j \neq k} H_k w_j w_j^H H_k^H + \sigma^2 I \right)^{-1} \right)^{-1}$$

(B.1)

in which we have applied the matrix inversion lemma [25]. By relating the above MSE equation (B.1) and the rate expression in (3.5), one can immediately find that the optimal
MSE and the achievable rate has the following relationship:

\[ R_k = -\log(e_k). \]  

(B.2)

Second, the consideration that each user is only served by a subset of BSs can be characterized by imposing an additional set of constraints:

\[ \mathbf{E}_k \mathbf{w}_k = \mathbf{0}, \forall k \]  

(B.3)

where \( \mathbf{E}_k \in \mathbb{R}^{M_t \times M_t} \) is a diagonal matrix defined as

\[
(E_k)_{j,j} = \begin{cases} 
0, & \text{if antenna } j \text{ serves user } k \\
1, & \text{otherwise} 
\end{cases}
\]  

(B.4)

Since both the problem (3.15) and (3.18) have the same set of constraints, it’s sufficient to check their equivalence by the first order optimality condition. Hence, by utilizing the relationship in (B.2) and setting the partial derivatives of the Lagrangians for problem (3.15) and (3.18) being zero respectively, we have

\[
\frac{\partial \mathcal{L}_{\text{wsr}}}{\partial \mathbf{w}_k} = \sum_k \frac{\alpha_k}{e_k} \frac{\partial e_k}{\partial \mathbf{w}_k} + \frac{\partial \left( \sum_l \lambda_l \left( \sum_k \| \mathbf{w}_k^l \|_2^2 - P_l \right) \right)}{\partial \mathbf{w}_k} 
\]

\[
+ \frac{\partial \left( \sum_l \mu_l \left( \sum_k \beta_k^l \hat{R}_k \| \mathbf{w}_k^l \|_2^2 - C_l \right) \right)}{\partial \mathbf{w}_k} + \frac{\partial \left( \sum_k \nu_k^T \mathbf{E}_k \mathbf{w}_k \right)}{\partial \mathbf{w}_k} = 0
\]  

(B.5)

\[
\frac{\partial \mathcal{L}_{\text{mmse}}}{\partial \mathbf{w}_k} = \sum_k \frac{\rho_k}{e_k} \frac{\partial e_k}{\partial \mathbf{w}_k} + \frac{\partial \left( \sum_l \lambda_l \left( \sum_k \| \mathbf{w}_k^l \|_2^2 - P_l \right) \right)}{\partial \mathbf{w}_k} 
\]

\[
+ \frac{\partial \left( \sum_l \mu_l \left( \sum_k \beta_k^l \hat{R}_k \| \mathbf{w}_k^l \|_2^2 - C_l \right) \right)}{\partial \mathbf{w}_k} + \frac{\partial \left( \sum_k \nu_k^T \mathbf{E}_k \mathbf{w}_k \right)}{\partial \mathbf{w}_k} = 0
\]  

(B.6)

where \( \lambda_l \in \mathbb{C}, \mu_l \in \mathbb{C}, (l = 1, 2, \ldots, L) \) and \( \nu_k \in \mathbb{R}^{M_t \times 1}, (k = 1, 2, \ldots, K) \) denote the
Lagrangian variable associated with the per-BS power constraint, per-BS backhaul constraint and the partial cooperation constraint (B.3) respectively.

Now suppose a set of transmit beamformers $\{\mathbf{w}_k^*\}$ make the equality hold in equation (B.5), then the same set of $\{\mathbf{w}_k^*\}$ will also satisfy the equation of (B.6) if the MSE weight $\rho_k$ is set as

$$\rho_k = \alpha_k/e_k^*$$

(B.7)

where the $e_k^*$ is the optimal MSE obtained by substituting $\{\mathbf{w}_k^*\}$ into equation (B.1). Similarly, if another set of $\{\mathbf{w}_k^*\}$ is the solution to the equation (B.6), then the same set of $\{\mathbf{w}_k^*\}$ will also satisfy the equation of (B.5) if we let $\alpha_k = \rho_k e_k^*$. Therefore, the WSR maximization problem (3.15) and the WMMSE minimization problem (3.18) are equivalent to each other.
Bibliography


[10] S. Christensen, R. Agarwal, E. Carvalho, and J. Cioffi, “Weighted Sum-Rate Max-

proach to Distributed Sum-Utility Maximization for a MIMO Interfering Broadcast


[13] Y. Zeng, E. Gunawan, Y. L. Guan, and J. Liu, “Joint Base Station Selection and Linear Precoding for Cellular Networks with Multi-Cell Processing,” in *IEEE TEN-


