Padmanabhan P. (2015) Thesis -
MIMO IBC Precoder design for weighted sum rate maximization. University of Oulu, Department of Communications Engineering, Master's Degree Program in Wireless Communications Engineering. Master's thesis, 47 p.

ABSTRACT

A downlink multi cell Multi User-Multiple Input Multiple-Output Multi-User-Multiple-Input-Multiple-Output (MU-MIMO) interference broadcast channel Interference Broadcast Channel (IBC) is taken into account, where IBC network is limited by multi user interference. The objective is to design a downlink Downlink (DL) multiantenna communication with base stations Base-station (BS) that perform cooperative precoding in a distributed fashion, where precoders maximizes the sum throughput of all users served by the coordinating BS. The precoders are designed in such a way that the interference is reduced and the system through put is achieved. The base problem and algorithm requires, local knowledge of the channel and converge to the stationary points of the weighted sum rate maximization weighted sum rate maximization (WSRM) problem. The problem is non convex and we use successive convex approximation Successive Convex approximation (SCA) to convert the problem to convex form by taking the Taylor series approximation. Keywords:

TABLE OF CONTENTS

FOREWORD

This masters thesis is focused on distributed precoder design for MIMOs IBCs. I would like to thank Centre for wireless Communication (CWC), Department of Communication Engineering (DCE) and University of Oulu for providing me the chance to do the thesis work.

I would like to express my sincere gratitude to Prof. Markku Juntti for providing me an opportunity to work on this thesis topic and also for his guidance and support throughout the work. I am very grateful to Dr. Le Nam Tran for his meticulous support and encouragement during the thesis work.

My special thank goes to My husband Mr. Ganesh Venkatraman for always motivating me and giving me valuable thoughts on my studies and also for his guidance and inspiration to do this research. Without him, I would not have been able to accomplish successfully.

Finally, I would also like to extend my sincere appreciation to my lovely son Master. Adhithya and my Parents Mr. A. Padmanabhan and Mrs. G. Santhi for always being a supportive pillar. I also thank my family members for their love and support.

LIST OF ABBREVIATIONS AND SYMBOLS

ADMM alternative-distributed-method-of-multipliers

AP-GP Arithmetic Geometric Mean

AWGN Additive White Gaussian Noise

BER Bit Error Rate

BS Base-station

CSI Channel State Information

CSIT Channel State Information at Transmitter

CSIR Channel State Information at Receiver

DL Downlink

DPC Dirty Paper Coding

DoF Degree of Freedom

IBC Interference Broadcast Channel

IC Interference Channel

IID Independently and Identically distributed

ICI Inter Cell Interference

KKT karush-khun-tucker

MU Multi User

MIMO multiple-input-multiple-output

MU-MIMO Multi-User-Multiple-Input-Multiple-Output

MISO multiple-input-single-output

MSE Mean Square Error

MMSE Minimum Mean Square Error

MRT Maximum Ratio Transmission

OFDM Orthogonal Frequency Division Multiplexing

PL Path Loss

QoS Quality of Service

SOCP second-order-cone-programming

SDMA space-division multiple access

SVD singular-value-decomposition

SNR signal-to-noise-ratio

SIR Signal to Interference ratio

SINR signal-to-interference-noise-ratio

SOC second-order-cone

SCA Successive Convex approximation

UL Uplink

WSRM weighted sum rate maximization

ZF Zero Forcing

1. INTRODUCTION

Wireless communication is gaining more importance due to its quick and easy accessibility and various other advantages like improved data rate and range extensions through multiplexing and diversity schemes. Our day to day life requires wireless services and applications with the advent of smart phones, since the demand is more the rate requirement is increasing exponentially. Availability of the spectrum is limited because of greater number of users and their demands, so our spectral usage needs to be improved. multiple-input-multiple-output (MIMO) communication provides higher system performance compared to the single antenna systems, more number of users can share the available spectrum [1].

Let us consider MIMO-IBC with multiple BSs that can transmit signals to the users in its own cell and causes interference to the neighbouring cell users. We design a linear transmit precoders, which is employed at the BS to extract the available signal gains from multiple antenna systems. In order to design effecient precoders the knowledge of Channel State Information at Transmitter (CSIT) is essential, for doing this pilots from each user are sent orthogonally to their respective serving BS.

In this thesis we study a MU-MIMO system, where multiple BSs serves users. The data is transmitted over a shared wireless network but since we have multiple BSs frequency reuse schemes are introduced to obtain maximum utilization of resources. The available link has certain restrictions and limitations due to interference from the neighboring BS like Inter Cell Interference (ICI) due to frequency reuse. In the DL, known as the MIMO broadcast channel, the BS sends different information streams to the users and in the Uplink (UL), the BS receives different information from the users. We consider the transmit precoder design in which a vector of information symbols is multiplied with a precoder matrix before the antenna array transmission. MU-MIMO in DL is interesting because, MIMO sum capacity can scale with the minimum of the number of BS antennas and the sum of the number of users times the number of antennas per user. This means that MU-MIMO can achieve MIMO capacity gains with a multiple antenna BS and a bunch of single antenna mobile users.

In the existing problem, the WSRM with linear transmit precoding for multicell multi-input single-output multiple-input-single-output (MISO) DL is considered. WSRM scheme is non convex and there exists beamformer designs which are based on achieving the necessary optimal conditions of the WSRM scheme. There are previous papers and results on WSRM scheme that they work very close to the optimal design using the iterative algorithm considering Karush-Kuhn-Tucker KKT equations.

In this report, we analyze the existing problem of WSRM in MIMO DL. The beamformers are based on SCA method. In the existing algorithm we see the approximation of WSRM with second-order-cone-programming (SOCP) method. This algorithm takes less time for convergence and also obtain optimal beamformers with the objective of maximum sum rate. The limitations of the existing problem of WSRM is overcome with different algorithms which follows the same aim as the existing work of approximaing the WSRM. The simulation results shows us that the new algorithm used to overcome the limitation performs better in means of convergence rate.

In this paper we propose a distributed precoder design algorithm for WSRM problem, that is based on alternative-distributed-method-of-multipliers (ADMM) approach, KKT based approach for Arithmetic Geometric Mean (AP-GP) and iterative minimization of weighted Mean Square Error (MSE). The proposed distributed algorithm requires only the local channel knowledge and converges to a stationary point of the weighted sum-rate maximization problem. The effectiveness of the proposed algorithm is evaluated in the numerical experiments and is discussed in the simulation results section.

2. BACKGROUND REVIEW

1.. MIMO- Basics

MIMO is a break through in wireless communication, that uses spatial dimension provided by the N transmit and M receive antennas to combat multipath fading. MIMO has become an essential element of wireless communication standards including IEEE 802.11n (Wi-Fi), IEEE 802.11ac (Wi-Fi), and Long Term Evolution (4G).

A MIMO network with N number of transmit and M number of receive antennas an extension of antenna array communication, that provides high spectral efficiency (The spectral efficiency can be defined by the total number of information bits per second per Hertz transmitted from one array to another), improves reliability and sensitivity to fading; reduced by spatial diversity that is provided by the multiple paths . There are several advantages of having MIMO antennas like gain, spatial and transmit diversity.

The figure represents a typical MIMO system where multiple data streams are transmitted in a single channel at the same time and multiple radios collect the multipath signals.

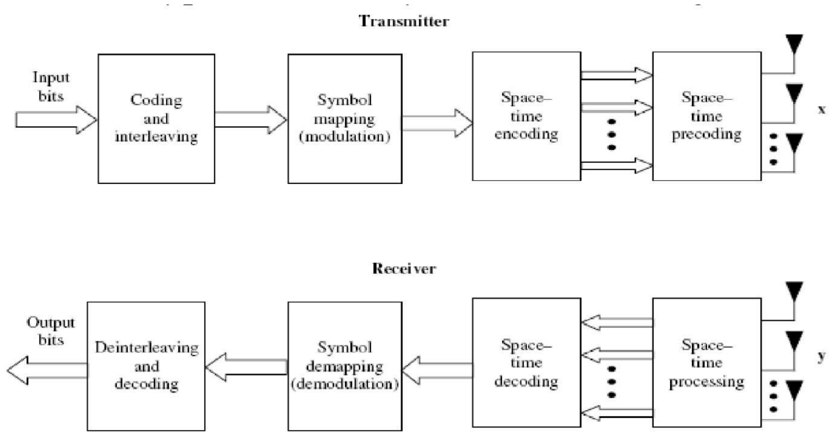


Figure 1.1: MIMO System Model

A typical MIMO channel and system model equation is as represented as in equation 1. The wireless channel part can be seen in keenly with the received vector y can be represented in terms of the channel H.

$$\mathbf{y} = H\mathbf{x} + n \tag{1}$$

where, transmitted signal vector $x = x_1, x_2, \ldots, x_n$, recieved vector $y = y_1, y_2, \ldots, y_n$ and the channel matrix can be represented as, H =

$$\begin{pmatrix} h_{11} & h_{12} & \ldots & h_{1M} \\ h_{21} & h_{22} & \ldots & h_{2M} \\ \ldots & \ldots & \ldots & \ldots \\ h_{N1} & h_{N2} & \ldots & h_{NM} \end{pmatrix}$$

.

The capacity for MIMO defined in [**?**], network capacity can be written as

$$C = \max_{f(x)} I(X;Y) \tag{2}$$

where, $I(X;Y)$ is the mutual information of the channel represented as,

$$I(X;Y) = \int \log(\frac{f(y|x)}{f(y)}) dF(x,y) \tag{3}$$

where the integral is taken over the random variables X and Y, $f(x)$ and $f(y)$ is

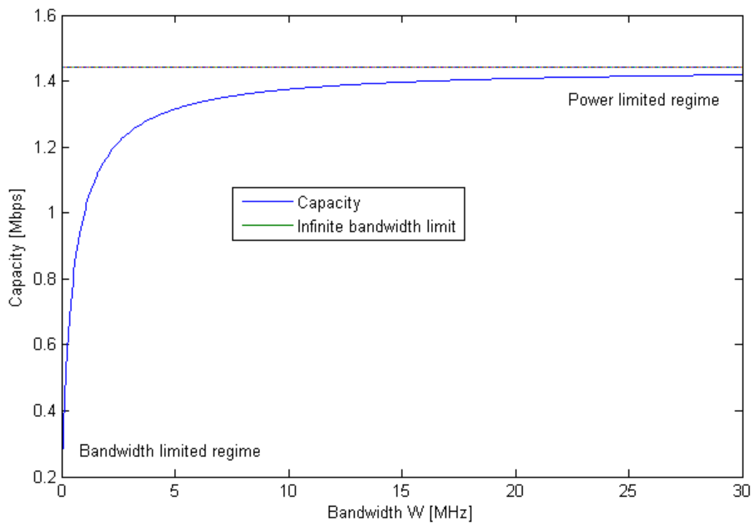


Figure 1.2: MIMO Capacity

denoted as the probability density function of X and Y, $F(x,y)$ is the cumulative distribution function of X and Y respectively. The mutual Information can also be written in terms of differential output of channel entropy and the conditional entropy.

$$I(X;Y) = h(Y) - h(Y|X) \tag{4}$$

For a time-invariant Additive White Gaussian Noise (AWGN) channel with received signal-to-noise-ratio (SNR) γ, the maximizing input distribution is Gaussian, which results in the channel capacity

$$C = \alpha \log(1 + \gamma) \tag{5}$$

At high SNR, the capacity of the i.i.d. Rayleigh fast fading channel scales like $n_min \log$ SNR bits/s/Hz, where n_min is the minimum of the number of transmit antennas n_t and receive antennas n_r. Thus, at high SNR we obtain degree-of-freedom gain. At low SNR, the capacity is approximately n_r SNR $\log_2$e bits/s/Hz. Thus at low SNR what we obtain is the recieve beamforming power gain. At all SNR the capacity increases linearly with n_min due to the linear combination of degree of freedom gain and beamforming power gain.

2.. MIMO-IC

MIMO communication systems has been showing a great potential in increasing the average throughput in the wireless communication scenario. Due to the performance gain in channel capacity and spectral efficiency in point to point MIMO systems has made the inclusion of singe user MIMO in different communication stansards. SU-MIMO has proved its efficiency to enhance the perormance in wireless networks. But, in cellular systems the available spectrum is very costly and is scarce so such systems has to deal with the inter-cell interference which doesnot exists in isolated point to point systems.

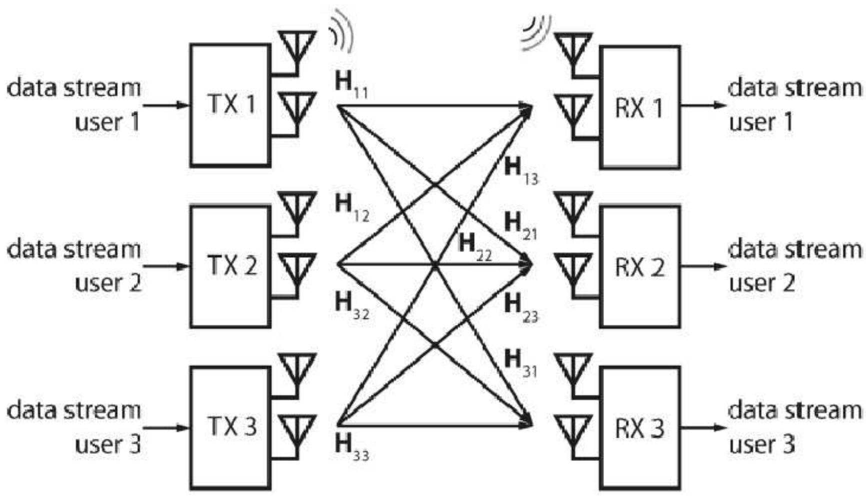


Figure 2.3: MIMO-IC System Model

Interference is a major set back and a limiting factor in the wireless communication networks. The problem of interference is in general dealt with planning of the (mostly static) radio resource management. Now a days we have a popularity of wireless devices having different wireless communication standards, the the ability to produce a desired or intended result of such interference avoidance solutions is limited. These days, standardization bodies are including interference coordination strategies in next generation cellular communication standards. A systematic study of the performance of cellular communication systems where each cell communicates multiple streams to its users while causing interference from and to the neighboring cells due to transmission over a common shared resource known as, MIMO-Interference Channel (IC). A K-user MIMO-IC model consists of a network of K transmit-receive pairs where each transmitter communicates multiple data streams to its respective receiver. In doing so, it generates interference at all other receivers present in the system.

3.. MIMO-IBC

The linear transceiver design problem is considered in a MIMO-IBC, in which a set of BSs send data to their intended users. Both the BSs and the users are equipped with multiple antennas, and they share the same time/frequency resource for transmission. The objective is to maximize the minimum rate among all the users in the network, in order to achieve fairness.

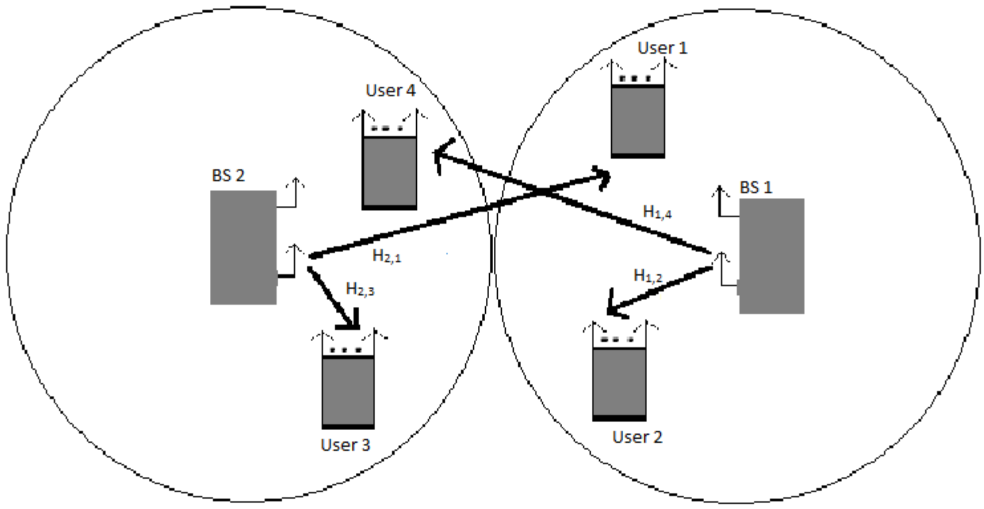


Figure 3.4: MIMO-IBC System Model

Interference is the main limiting factor in wireless transmission. BS consisting of multiple antennas are able to serve multiple users simultaneously, which is the principle of space-division multiple access (SDMA) or MU-MIMO. However, Multi User (MU) systems have precise requirements for CSIT which is more difficult to acquire than Channel State Information (CSI) at the Rx Channel State Information at Receiver (CSIR). Hence we focus here on the more challenging DL (though the UL is also non-trivial in the case of Mobile Terminals (MTs) with multiple antennas). In cellular systems, one can distinguish between the cell center where a single cell design is appropriate (due to high Signal to Interference ratio (SIR) and the cell edge where a multi-cell approach is mandatory. The MU-MIMO DL problem for the cell center users is called the (MIMO) Broadcast Channel (BC). or the cell edge users, the recent introduction of Interference Alignment (IA) has shown that approaching high system capacity through agressive frequency reuse should in principle be possible. Whereas precise capacities for cellular systems remain unknown, IA allows to reach the optimal high SNR rate prelog, called Degree of Freedom Degree of Freedom (DoF) (or spatial multiplexing factor, or number of streams). That is, before accounting for CSI acquisition.

Optimal precoder design for WSRM in MIMO interference networks is studied. For this well known non-convex optimization problem, convex approximations based on interference alignment are developed, for multi-beam cases. Considering that each user treats interference from other users as noise. It is well known that, due to interference coupling, the problem is a non-convex optimization and is hard to solve. In the high SNR regime, there has been recent progress on maximizing the sum degrees of freedom, exploiting the idea of interference alignment. It has been shown that maximizing the sum degrees of freedom is still an NP hard problem.

4.. Mathematical Preliminaries

4.a. Convex Optimization

2.4.1.1. Convexity

Convexity is also called convex analysis, which is an area in mathematics where one studies about convex sets and convex functions. Convexity is also the mathematical core of optimization [1], where it plays an important role in statistics, differential equations and mathematical economics. Convexity can also be called as convex analysis. Some examples of convex sets are triangle, rectangle, polyhedron and quadratic equations.[1].

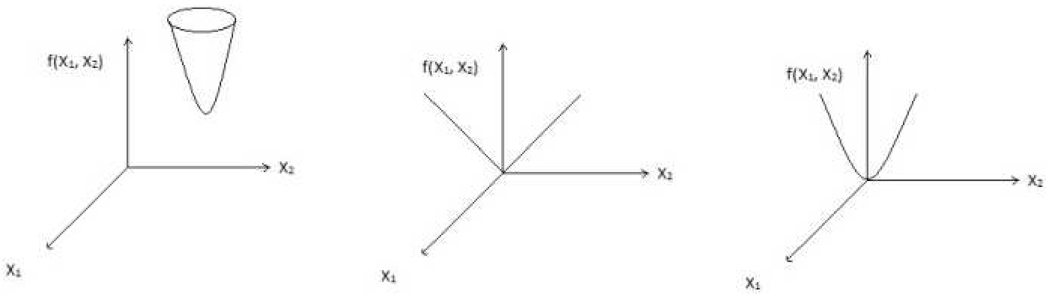


Figure 4.5: Convex Functions

4.b. Convex Sets

Consider $\mathcal{K}$ is a set of $\mathcal{R}^n$ is said to be convex when the line segment through the points x, y belongs to $\mathcal{K}$. In general, the points must lie inside the set and the set is connected such that without leaving the set we can pass through any two points. As mentioned above there are several examples for convex sets like ellipsoid, hypercubes etc. We can also define that the intersection of any convex sets is a convex set.[2]. A set that is not convex is called non convex set or also called as concave set. We can define a non convex set by considering a line segment joining the points x, y lies outside the set $\mathcal{K}$.

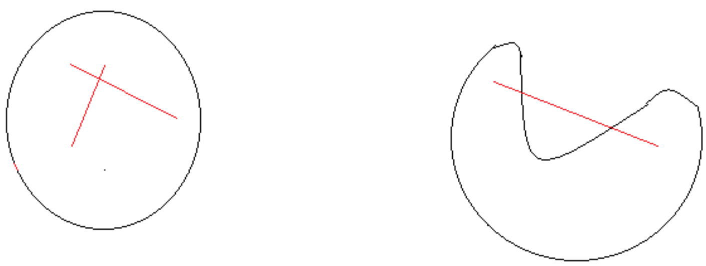


Figure 4.6: Convex Sets

2.4.2.1. Convex Functions

Convex functions are continuous function and is convex if and only if the region above the graph as shown in figure is convex set. A function f is convex if, $\forall\, x, y \in \mathcal{K}$, $\forall\, \theta \in$ [0, 1]:

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y) \tag{6}$$

The function f is said to be strictly convex if, $\forall\, x \neq y \in \mathcal{K}$, $\forall\, \theta \in$ (0, 1):

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta)f(y) \tag{7}$$

A function f is said to be concave if the function -f is strictly convex. Strict convexity means that the graph of f lies below the segment $\mathcal{S}$. Certain examples of strict convex functions are exponential and quadratic function. Several operations on these function preserves the convexity like summation, multiplication of convex functions.

2.4.2.2. Optimization Problem

A generic optimization problem is similar to linear programming problem, that can be solved quickly depending on the variables and the constraints. The objective and the constraints must be linear for the problem to be convex. In general, a convex optimization problem has all of the constraints as convex functions, and the objective is a convex function of minimizing, or a concave function of maximizing.

The standard form of optimization problem can be written as,

$$\underset{x}{\text{minimize}} \qquad f(x) \tag{8}$$
$$\text{subject to} \qquad x \in \mathcal{X}. \tag{9}$$

where, $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ is the objective function that has to be minimized with respect to x and $\mathcal{X} \subset \mathbb{R}^n$ is the feasible set. A maximization problem can be written by negating the objective function.

2.4.2.3. Convex Problem

A convex optimization problem is one which has both the objective and the given set of constraint as convex. In general linear functions are convex so the linear programming problem is convex problem. A general convex optimization problem can be written as

$$\underset{x}{\text{minimize}} \quad f(x) \tag{10}$$
$$\text{subject to} \quad g_i(x) \leq 0, i = 1, ...m \tag{11}$$
$$h_j(x) = 0, j = 1, ...p. \tag{12}$$

where, $f(x) : \mathbb{R}^n \to \mathbb{R}$ is the objective function and $\mathcal{X} \subset \mathbb{R}^n$ is the feasible set and is called convex when $\mathcal{X}$ is closed convex set and $f(x)$ is convex on $\mathbb{R}^n$. The second equation is the inequality constraint and the third is the equality constraint in the above optimization problem.

The optimality condition for a convex problem; assume a feasible point $\mathbf{x}^*$ but it is necessary to know if this point is the optimal solution. We can consider the gradient of a function f that shows the ascent direction of the function. Gradient of f at a point x divides the space into three regions one where the function increases, one where the function decreases and the one where we cannot figure out using the gradient alone. For $\mathbf{x}^*$ to be optimal and the feasible region doesn't lie in the half space where the function decreases, if not then the point $\mathbf{x}^*$ is not the optimal point. Convexity makes this condition sufficient for optimality.

To find solution for an unconstrained objective, we differentiate the objective function with respect to the optimization variable x and equate to zero as $\nabla f(x) = 0$. However for an constrained problem, we solve the Lagrangian as

$$\underset{\lambda}{\text{maximize}} \quad \underset{x}{\text{minimize}} \quad L(x, \lambda_1, \ldots, \lambda_m) = f(x) + \lambda_1 g_1(x) + \ldots + \lambda_m g_m(x) + \mu_1 h_1(x) + \ldots + \mu_p h_p(\tag{13}$$

where $\lambda_i \geq 0$ and μ_j are Lagrange multipliers. In order to solve the constrained problem, similar to the unconstrained problem we take the partial derivative of the Lagrangian with respect to the optimization variable.

When the objective is convex, and the equality conditions are affine and the inequality conditions are convex then one of the possible solutions, in some minimum principle is equivalent to KKT optimality conditions. To solve a convex optimization problem as above with the KKT approach we need the lagrange multiplier. There exists $\lambda_1, \lambda_2, \ldots, \lambda_m, \mu_1, \ldots, \mu_p$, called the Lagrange multipliers, for each point $x \in X$ that minimizes f over $\mathcal{X}$. The Lagrange multipliers must satisfy certain conditions:

1. $x \quad minimizes \quad L(z, \lambda_1, \lambda_2, \ldots, \lambda_m), \forall z \in X,$

2. $\lambda_1 \geq 0, \lambda_2 \geq 0, \ldots, \lambda_m \geq 0$,
3. Complementary Slackness: $\lambda_1 g_1 = 0, \lambda_2 g_2 = 0, \ldots, \lambda_m g_m = 0$.

When the problem is has a convex objective and non convex constraint set then KKT method is not a feasible approach.

2.4.2.4. Non Convex Problem

A non convex problem is the one that has objective or any one of its constraint as non convex. Such problems have multiple feasible regions and multiple local minima in a region. An example for a non convex function would be a sine wave. Let us consider a non convex problem,

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(x) && (14) \\ \text{subject to} \quad & g_i(x) \leq 0, i = 1, \ldots, m && (15) \\ & h_j(x) = 0, j = 1, \ldots, p. && (16) \end{aligned}$$

with variable $x \in \mathbb{R}^n$

Let us consider a non convex problem when the objective function $f(x)$ is convex and the inequality constraint $g_i(x)$, $i = 1, \ldots, n$ is differentiable convex function and $g_i(x)$ $i = (n+1), \ldots, m$ are differentiable function and the linearity constraint is affine. These non convex problems are common in wireless communication and this is our interest. In order to solve this problem the non convex part of the objective is approximated around a convex function to solve in an iterative manner.

As mentioned in the paper gordon.p. wright the inner approximation algorithm for the minimization problem can be done in the following steps,

step 0. Set a starting point for the variable and constraint $x^0 \in F$ and set $h^0 = g_0(x^0)$. Let $A^0 = \{x | h^0 = g_0(x) \text{ and } x \in F\}$, where F can be defined as the feasible region.

step 1. In the k^{th} iteration replace the constraint $g_i(x) \leq 0$, $i = (n+1), \ldots, m$, by $\bar{g}_i(x, x^k) \leq 0$, where $\bar{g}_i(x, x^k)$ is a differentiable convex function and $x^k \in \mathcal{A}^{k-1}$. Each function $\bar{g}_i(x, x^k)$ must have the following properties,
1. $g_i(x) \leq \bar{g}_i(x, x^k) \quad \forall x \in F^k$
2. $g_i(x) = g_i(\bar{x^k}, x^k)$
3. $\delta g_i(x^k)/\delta x_j = \delta \bar{g}_i(x^k, x^k)/\delta x_j \quad j = 1, \ldots, n$

The feasible region $F^k = \{x|g_i(x) \leq 0 \forall i = 1,\ldots,n$ and $\bar{g}_i(x,x^k) \leq 0 \forall i = n+1,\ldots,m\}$ should satisfy slaters constraint qualification condition for convex programs.

step 2. Solve the approximation convex program

$$\underset{x}{\text{minimize}} \quad g_0(x) \tag{17}$$
$$\text{subject to} \quad g_i(x) \leq 0, i = 1,\ldots,n \tag{18}$$
$$\bar{g}_i(x,x^k) \leq 0, i = n+1,\ldots,m. \tag{19}$$

Let $h^k = min\{g_0(x)|x \in F^k\}$.

step 3. If $h^k = h^(k-1)$, then x^k is a KKT solution for the minimization problem. Otherwise, let $a^k = \{x|h^k = g_0(x)$ and $x \in F^k\}$ and return to step 1.

The above algorithm proposed by WRIGHT can be used to optimize non linear programs even when the constraint is a non convex function. The objective is replaced with a new variable and is added into the constraint set. As mentioned in the paper, the solution for the algorithm is not only the KKT point but also the global minimum for approximating convex problem that is interior to the feasible region.

3. PRECODER DESIGN

1.. Introduction to Precoder design

Precoding can be explained as the beamforming method used for multi-stream transmission in MIMO communication. Multiple data streams are emitted from the transmit antenna having weights to maximize the throughput of receiver output. In this technique, transmitter sends coded information to the receiver for analyzing the channel, where the receiver is a simple detector, example a matched filter, and does not need the channel side information. Thus reducing the effects of channel used for the communication. Precoding can be explained for both point to point systems and multi-user MIMO system.

In point-to-point MIMO system, transmitter is equipped with multiple antennas that communicates with a receiver that has multiple antennas. The precoding in point to point case assumes a narrowband slowly fading channel, that is achieved through Orthogonal Frequency Division Multiplexing (OFDM), and the channel capacity and throughput can be maximized depending on the CSI available in the system. If the transmitter has statistical information and the receiver knows the channel matrix then eigen beamforming achieves the MIMO channel capacity, where, the transmitter emits multiple streams in eigen directions of the channel covariance matrix. If the channel matrix is completely known, singular-value-decomposition singular-value-decomposition (SVD) precoding is known to achieve the MIMO channel capacity, where, the channel matrix is diagonalized by taking an SVD and removing the two unitary matrices through pre- and post-multiplication at the transmitter and receiver, respectively, thus, one data stream per singular value can be transmitted without having any interference.

In MU-MIMO, a multi-antenna transmitter communicates simultaneously with multiple receivers with one or more antennas known as SDMA. Precoding algorithms for SDMA systems can be sub-divided into linear and nonlinear precoding types. The capacity achieving algorithms are nonlinear approach but linear precoding approaches usually achieve reasonable performance with much lower complexity. Linear precoding strategies include Maximum Ratio Transmission (MRT), Zero Forcing (ZF) precoding, and transmit Wiener precoding, there are also precoding strategies for low-rate feedback of channel state information. Nonlinear precoding is designed based on the concept of Dirty Paper Coding (DPC), which shows that any known interference at the transmitter can be subtracted without the penalty of radio resources if the optimal precoding scheme can be applied on the transmit signal.

DPC is a coding technique that pre-cancels known interference without power penalty. Only the transmitter needs to know this interference, but full CSI is required every-where to achieve the weighted sum capacity. DPC is known as the capacity achieving scheme in the MIMO channel. But DPC is a Non-linear technique for interference cancellation having higher degree of complexity. Thus, to overcome this we can go for a linear technique considering the problem of WSRM with linear transmit precoding for multicell MIMO downlink. But, the WSRM problem, for single-antenna receivers are considered to be NP-hard. Although optimal beamformers can be obtained and they may not be practically useful since the complexity of finding optimal designs grows exponentially with the problem size. Hence, the need of computationally conducive suboptimal solutions to the WSRM problem still remains. Since the WSRM problem is nonconvex and NP-hard, there exists a class of beamformer designs which are based on achieving the necessary optimal conditions of the WSRM problem.

1.a. System Model

Consider a downlink MIMO IBC system with $\mathcal{B}$ coordinated BS of N transmit antennas each and K single antenna receivers. The set of all K users is denoted by $\mathcal{U} = \{1, 2, \ldots, K\}$. We assume that data for the k^{th} user is transmitted only from one BS, which is denoted by $b_k \in \mathcal{B}$, where $\mathcal{B} \triangleq \{1, 2, \ldots, \mathcal{B}\}$ is the set of all BS. The set of all users served by BS b is denoted by $\mathcal{U}_b$. Under flat fading channel conditions, the input-output relation for k^{th} user is given as

$$y_k = \mathbf{h}_{b_k,k}^H \mathbf{x}_k + n_k \tag{20}$$

where h_k is the channel coefficient showing the channel response between transmit and recieve antenna, ie., from BS b to user k and $n \sim \mathcal{CN}(0, \sigma^2)$ is complex circularly symmetric zero mean gaussian noise with variance σ^2. The receiver requires information about the channel h_k to suppress the effect of noise n. By doing this, complexity is increased, but the receiver has to be simple that, the BS can predict the channel.

Under linear precoding the transmitted vector x can be written as

$$\mathbf{x}_k = \sum_{i=1}^{K} \mathbf{w}_i d_i \tag{21}$$

where, d_i is normalized data symbol and w_i is the linear precoding vector. The signal-to-interference-noise-ratio (SINR) (γ_k) can be written as

$$\gamma_k = \frac{||\mathbf{h}_{b_k,k}\mathbf{w}_k||^2}{\sigma^2 + \sum_{i=1, i \neq k}^{K} ||\mathbf{h}_{b_i,k}\mathbf{w}_i||^2} \tag{22}$$

For reliable communication achievable commuinication rate can be written as

$$C = \mathbf{log}_2(1 + \gamma_k) \tag{23}$$

where, C is the achievable capacity. The transmission is power limited with a total power constraint as

$$\sum_{i=1}^{K} \|\mathbf{w}_i\|^2 \leq \mathbf{P}. \tag{24}$$

The precoder design can be classified into Centralized and Distributed approach. In the centralized approach all the BS share the CSI. The centralized controller is equipped to know and calculate the expected sum rate. In distributed approach, practical difficulties of distributing CSI over the backhaul network and high complexity of joint precoding design motivates the analysis. The beamforming and power allocation strategies can be computed locally using only local CSI in a distributed design. In general, the goal of precoding is to maximize the signal power at the intended terminal while minimizing the interference caused at others.

4. CENTRALIZED PRECODER DESIGN

A centralized coordinated DL transmission requires CSI to be feedback from the users to their serving BS, and aggregated at the central coordination node to form the channel matrix for precoding, so that interference can be mitigated.

1.. System Model for WSRM

Consider a downlink MIMO IBC system with $\mathcal{B}$ coordinated BS of N transmit antennas each and K single antenna receivers. The set of all K users is denoted by $\mathcal{U} = \{1, 2, \ldots, K\}$. We assume that data for the k^{th} user is transmitted only from one BS, which is denoted by $b_k \in \mathcal{B}$, where $\mathcal{B} \triangleq \{1, 2, \ldots, \mathcal{B}\}$ is the set of all BSs. The set of all users served by BS b is denoted by $\mathcal{U}_b$. Under flat fading channel conditions, the signal received by the k^{th} user is

$$h_k x_k = \mathbf{w}_k^{\mathrm{H}} \mathbf{h}_{b_k,k} \mathbf{x}_k + \mathbf{w}_k^H \sum_{i=1, i\neq k}^{K} \mathbf{h}_{b_i,k} \mathbf{x}_i + \mathbf{w}_k^H \mathbf{n}_k \tag{25}$$

where $\mathbf{h}_{b_i,k} \in \mathbb{C}^{1\times N}$ is the channel (row) vector from BS b_i to user $k, \mathbf{w}_k \in \mathbb{C}^{N\times 1}$ is the beamforming vector (beamformers) from BS b_k to user k, d_k is the normalized complex data symbol, and $n_k \sim \mathcal{CN}(0, \sigma^2)$ is complex circularly symmetric zero mean gaussian noise with variance σ^2. The term $\sum_{i=1, i\neq k}^{K} \mathbf{h}_{b_i,k} \mathbf{w}_i d_i$ in (**??**) includes both intra- and inter-cell interference. The total power transmitted by BS b is $\sum_{k\in\mathcal{U}_b} \|\mathbf{w}_k\|^2$. The SINR γ_k of user k is

$$\gamma_k = \frac{\|\mathbf{h}_{b_k,k} \mathbf{w}_k\|^2}{\sigma^2 + \sum_{i=1, i\neq k}^{K} \|\mathbf{h}_{b_i,k}\|} \tag{26}$$

Here, we are interested in the problem of WSRM under per-BS power constraints, which is formulated as

$$\max_{\sum_{k\in\mathcal{U}_b} \|\mathbf{w}_k\|^2 \leq P_b, \forall\, b\in\mathcal{B}} \quad \sum_{k=1}^{K} \alpha_k \log(1 + \gamma_k) \tag{27}$$

where α_k's are positive weighting factors which are typically introduced to maintain a certain degree of fairness among users.

2.. Problem Formulation

To achieve a tractable solution for the Low-Complexity beamfomer design, we can write the problem as shown in the following equation,

$$\max_{\sum_{k\in\mathcal{U}_b}\|\mathbf{w}_k\|^2\leq P_b,\forall b\in\mathcal{B}} \quad \prod_k (1+\gamma_k)^{\alpha_k} \tag{28}$$

The equations can be re-written as

$$\begin{aligned} \max_{w_k,t_k} \quad & \prod_k t_k && \text{(29a)}\\ \text{subject to} \quad & \gamma_k \geq {t_k}^{1/\alpha_k} - 1, \forall k \in \mathcal{U}, && \text{(29b)}\\ & \sum_{k\in\mathcal{U}_b} \|\mathbf{w}_k\|_2^2 \leq P_b, \forall b \in \mathcal{B}. && \text{(29c)} \end{aligned}$$

In equation (29b) and (29c) it can be seen that all constraints are active at the optimum, otherwise, we can obtain a larger objective by increasing t_k without violating the constraints. We can reformulate (29b) by reintroducing additional slack variable β_k,

$$\begin{aligned} \underset{w_k,t_k,\beta_k}{\text{maximize}} \quad & \prod_k t_k && \text{(30a)}\\ \text{subject to} \quad & \Re\mathbf{h}_{b_k,k}\mathbf{w}_k \geq \sqrt{{t_k}^{1/\alpha_k} - 1}\beta_k, \forall k \in \mathcal{U} && \text{(30b)}\\ & \Im(\mathbf{h}_{b_k,k}\mathbf{w}_k) = 0, \forall k \in \mathcal{U}, && \text{(30c)}\\ & (\sigma^2 + \sum_{i=1,i\neq k}^{K} \|\mathbf{h}_{b_i,k}\mathbf{w}_k\|^2)^{1/2} \leq \beta_k, \forall k \in \mathcal{U}, && \text{(30d)}\\ & \sum_{k\in\mathcal{U}_b} \|\mathbf{w}_k\|_2^2 \leq P_b, \forall b \in \mathcal{B}. && \text{(30e)} \end{aligned}$$

The relation between (29b) and (30c) can be seen as, first, by forcing the imaginary part of $\mathbf{h}_{b_k,k}\mathbf{w}_k$ to zero in (30c) does not affect the optimality of (29c) since phase rotation on $\mathbf{w}_k$ will result in the same objective while satisfying all constraints. Second we can show that all the constraints in (30) hold with equality at the optimum.

5. DISTRIBUTED PRECODER DESIGN

In the distributed precoder design, the precoders are designed independently at each BS. The convex formulation in the above equation requires a centralized controller to perform the precoder design for all users belonging to the coordinating BSs. To design the precoders independently at each BS with lesser information exchange via backhaul, iterative decentralized methods are used. In general, the decentralized approaches are mainly based on primal or the alternating method of multipliers ADMM schemes.

1.. Problem Formulation

We consider the convex sub problem with fixed receive beamformers w_k based on Taylors series approximation for the non convex constraint. The objective function of equation (29a) can be decoupled across each BS as

$$\sum_k \log(1+\gamma_k) \geq \log(1+\gamma_k) \tag{31a}$$

$$\prod_k (1+\gamma_k) \geq (1+\gamma_k) \tag{31b}$$

thus the centralized problem can be represented in equivalent form as

$$\begin{aligned} \underset{w_k,\beta_k}{\text{maximize}} \quad & \sum_{k=1}^{K} \log(1+\gamma_k) && \text{(32a)} \\ \text{subject to} \quad & \text{(30b)} - \text{(30e)} && \text{(32b)} \end{aligned}$$

where, $\log(1+\gamma_k)$ represents the vector of weighted sum of users $k \in \mathcal{U}_b$. Let us consider $\mathcal{U}_b$ as, the users from my BS and $\mathcal{U}_{\bar{b}}$ as the users from the interfering BS. The set of all K users is denoted by $\mathcal{U} = \{1, 2, \ldots, K\}$. We assume that data for the k^{th} user is transmitted only from one BS, which is denoted by $b_k \in \mathcal{B}$, where $\mathcal{B} \triangleq \{1, 2, \ldots, \mathcal{B}\}$ is the set of all BS. The set of all users served by BS b is denoted by $\mathcal{U}_b$.

The problem can be decentralized or distributed with the BS specific interference vector term, which can be fixed or can be used as a variable depending on the choice of our decomposition method. The given convex problem can be decomposed to parallel iterative subproblems cordinated by primal or dual decomposition update. The coupling variables are updated in each iteration by interchanging the given information within the subproblem.

2.. Alternating Direction Method of Multipliers (ADMM) Schemes

In the Primal decomposition method, the convex problem can be solved for optimal transmit precoders in an iterative manner for fixed BS. Using a master-slave model the problem can be solved, where the slave problem is solved in the corresponding BS for optimal tranmit precoders. Upon finding the optimal transmit precoder in the subproblem, the master problem is used to update the BS specific interference term for the next iteration by using the dual variables corresponding to the interference constraint. Thus these steps are continued further until a global optimal solution is obtained. The primal problem is similar to the minimum power precoder design as shown in paper "g. scutari". Hence, the master problem treats the interference term as variable where as, in the slave the interference term is considered as a constant for every iteration.

The ADMM method is used to decouple the precoder design across multiple BSs in-order to solve the convex problem. In contrary with the primal decomposition method, dual composition method reduces the constraint by considering the objective functions in the sub problem. The ADMM approach is prefered over dual decomposition method due to its robustness and improved convergence. In this method we hold a local and a global copy of the signal interference. At optimality the copies remain equal. ADMM method can be formulated as following,

$$
\begin{aligned}
\underset{\substack{w_k, t_k, \beta_k,\\ \delta_{b,k}^{b_k}, \delta_{b_k,i}^{b_k}}}{\text{maximize}} \quad & \sum_k \log t_k + \sum_{k\in\mathcal{U}_{b_k}} \sum_{b\in\bar{\mathcal{B}}_{b_k}} \lambda_{b,k}^{b_k}(\delta_{b,k}^{b_k} - \delta_{b,k}^{G}) + \sum_{i\in\bar{\mathcal{U}}_{b_k}} \lambda_{b_k,i}^{b_k}(\delta_{b_k,i}^{b_k} - \delta_{b_k,i}^{G}) - \\
& \frac{\rho}{2} \sum_{k\in\mathcal{U}_{b_k}} \sum_{b\in\bar{\mathcal{B}}_{b_k}} \|\delta_{b,k}^{b_k} - \delta_{b,k}^{G}\|_2^2 - \frac{\rho}{2} \sum_{i\in\bar{\mathcal{U}}_{b_k}} \|\delta_{b_k,i}^{b_k} - \delta_{b_k,i}^{G}\|_2^2 && \text{(33a)}\\
\text{subject to} \quad & \mathbf{h}_{b_k,k}\mathbf{w}_k \geq \beta_k \sqrt{({t_k}^{1/\alpha_k} - 1)}, \forall k \in \mathcal{U}_{b_k} && \text{(33b)}\\
& \mathrm{Im}(\mathbf{h}_{b_k,k}\mathbf{w}_k) = 0, \forall k \in \mathcal{U}_{b_k}, && \text{(33c)}\\
& \sigma^2 + \sum_{\substack{i\in\mathcal{U}_{b_k}\\ i\neq k}} \|\mathbf{h}_{b_k,k}\mathbf{w}_i\|^2 + \sum_{b\in\bar{\mathcal{B}}_{b_k}} \delta_{b,k}^{(b_k)} \leq \beta_k, \forall k \in \mathcal{U}_{b_k}, && \text{(33d)}\\
& \delta_{b,j}^{(b_k)} \geq \sum_{i\in\mathcal{U}_b} \|\mathbf{h}_{b_k,k}\mathbf{w}_i\|^2, \forall j \in \bar{\mathcal{U}}_{b_k} && \text{(33e)}\\
& \sum_{k\in\mathcal{U}_b} \|\mathbf{w}_k\|_2^2 \leq P_b, \forall k \in \mathcal{U}_{b_k}. && \text{(33f)}
\end{aligned}
$$

In the equation above the objective function of the precoder design is to maximize the sum rate (WSRM)problem. The ADMM method uses the sub gradient update method, where we assume the interference term from the adjascent BS. The $\delta^{(b_k)}$ is the local interference term taken into account from the adjascent BS, similarly the $\delta^{(G)}$ is the global interference term taken into account from the adjascent BS.At optimality the local and the global interference terms becomes same. The equation (33b) and the equation (33c) remains the same as the previous formulation. In the equation (33d) we observe that there is the local interference term coming from the adjascent BS adding into the noise plus the interference from the same BS. Similarly equation (33e) shows the sum of all interefrence terms from the adjascent BS.

Algorithm 1 ADMM Method

Input: $\alpha_k, \mathbf{h}_{b_k,k}, \forall b \in \mathcal{B}, \forall k \in \mathcal{U}_b$.
Output: $\mathbf{w}_k, \forall k \in \{1, 2, \ldots, K\}$
Initialization: $i = 0$ and $\mathbf{w}_k$ satisfying (5.5f)
precoders randomly satisfies the power constant $\sum_{k \in \mathcal{U}_b} \|\mathbf{w}_k\|_2^2 \leq P_b, \forall k \in \mathcal{U}_b$
initialize global interference vector $\delta^{(0)} = 0^T$
initialize interference threshold $\lambda \forall b \in \mathcal{B}$
for each BS $b \in \mathcal{B}$ perform the following procedure
repeat
  initialize $j = 0$
  repeat
    solve the precoders $\mathbf{w}_k$ and local interference $\delta^{b_i,k}$ using the equation.
    exchange δ^b across the co-ordinating BS $\mathcal{B}$
    update dual variable λ using equation
    update the global interference
  until convergence $j \geq J_{max}$
  update the beamformer $\mathbf{w}_k$ with the recent precoder.
  exchange the receive precoder $\mathbf{w}_k \forall k \in \mathcal{U}_b$ among the BS.
  Update β_k with the precoder $\mathbf{w}_k$ i=i+1
until convergence $i \geq I_{max}$

Let us consider a two base station scenario such that the global interference term can be updated as follows,

$$\delta_{b,k}^{G} = \frac{1}{2}(\delta_{b,k}^{(b)} + \delta_{b,k}^{(b_k)}) \tag{34}$$

where $\delta_{b,k}^{(b)}$ refers to the actual interference caused by BS b and $\delta_{b,k}^{(b_k)}$ refers to the local interference caused by the BS b_k.

The update for the local interference term is made through iteration of the objective function in each BS. Once the local interference iterations are done then the global interference and the dual variable update are made in the main problem. The dual variable corresponds to the interference terms in the BS b_k and is updated with the sub gradient method. ρ gives the dual update step length depending upon the system performance and convergence behaviour. The convergence nature of the distributed algorithm depends upon the choice made on step size. The distributed precoder design based on ADMM is shown in Algorithm 1. The convergence for distributed algorithm is discussed in the Appendix

$$\lambda_{b,k}^{(n+1)} = \lambda_{b,k}^{(n)} - \rho(\delta_{b,k}^{(b_k)^*} - \delta_{b,k}^{(G)^*}) \tag{35}$$

3.. Karush Kuhn Tucker KKT based Decomposition

In order to decentralize the precoder design across the coordinating BSs in $\mathcal{B}$ we consider the KKT based decentralization approach for AP-GP method, MSE and Rate reformulation method. Our optimization problem can be solved using the KKT conditions. In this KKT approach, the transmit precoders and subgradiant updates are performed at the same instant within few number of iterations. In the ADMM method we have signaling overhead due to information exchange about the coupling variable, thus if the number of iterations required for the convergence is large, then it many not be practically viable.

The KKT approach is practical due to the limited signaling requirement in each iteration. This approach has been considered in "ref. ganesh 8 and 9". We understand that the distributed approach discussed above using ADMM may not be viable due to the signalling overhead involved in exchanging the coupling variable when the number of iterations required to converge is large that depends on the size of the system.

3.a. KKT for AP - GP Method without Rate Constraint

In this section we discuss a way to decentralize the precoder design across the corresponding BS in $\mathcal{B}$ based on AP-GP method. In contrast to the centralized and ADMM method, the problem is solved using the KKT conditions. The weighted sum rate maximization problem with Quality of Service (QoS) constraints subject to convex transmit power constraint $\mathcal{P}$ is considered. The problem in the general form is written as,

$$
\begin{aligned}
\underset{\mathbf{w}_k,\gamma_k,\beta_k}{\text{maximize}} \quad & \sum_k \log_e(1+\gamma_k) && \text{(36a)}\\
\text{subject to} \quad & a_k : \|\mathbf{h}_{b_k,k}\,\mathbf{w}_k\|^2 \geq \sqrt{\gamma_k\beta_k}, && \text{(36b)}\\
& b_k : \sigma^2 + \sum_{\substack{i=1,\\ i\neq k}}^{K} \|\mathbf{h}_{b_i,k}\mathbf{w}_i\|^2 \leq \beta_k, \forall i \in \bar{\mathcal{U}}_{b_k}, && \text{(36c)}\\
& c_k : \sum_{k\in\mathcal{U}_b} \|\mathbf{w}_k\|_2^2 \leq P_b, \forall b \in \mathcal{B}. && \text{(36d)}
\end{aligned}
$$

The variable a_k, b_k, c_k are dual variables corresponding to equations from (36b) - (36d) which is the constraints. The optimization variables are the precoder vector $\mathbf{w}_k \in \mathbb{C}^{N_T}\forall k$, where k= 1,2,. . .,K.

In the above problem the equation (36b) can be seen as a non convex constraint and we can decompose the right hand side of the constraint $\sqrt{\gamma_k\beta_k}$ into $\frac{1}{2\phi_k}\gamma_k + \frac{\phi_k}{2}\beta_k$, where ϕ_k can be written as, $\phi_k = \frac{\gamma_k}{\beta_k}$. Now the constraint in (36b) can be written as

$$
\begin{aligned}
a_k : \Re\{\mathbf{h}_{b_k,k}\mathbf{w}_k\} \geq \frac{1}{2\phi_k}\gamma_k + \frac{\phi_k}{2}\beta_k, \forall k \in \mathcal{U},\\
\Im\{\mathbf{h}_{b_k,k}\mathbf{w}_k\} == 0, \forall k \in \mathcal{U}. \qquad \text{(37a)}
\end{aligned}
$$

Now we can replace the equation in (36b) with (37). The Lagrangian function for the given problem can be written as

$$
\begin{aligned}
L(\gamma_k,\beta_k,\mathbf{w}_k) = &-\log_e(1+\gamma_k)\log_2 e + a_k\left(\frac{1}{2\phi_k}\gamma_k + \frac{\phi_k}{2}\beta_k - \mathbf{h}_{b_k,k}\mathbf{w}_k\right)\\
&+b_k\left(\sigma^2 + \sum_{\substack{i=1,\\ i\neq k}}^{K}\|\mathbf{h}_{b_i,k}\mathbf{w}_i\|^2 - \beta_k\right) + c_k\left(\sum_{k\in\mathcal{U}_b}\|\mathbf{w}_k\|_2^2 - \beta_k\right). \qquad (38)
\end{aligned}
$$

By evaluating the Lagrangian function in (38) with respect to the primal and dual variables we obtain an iterative solution as

Since the dual variable $a_k^{(i+1)}$ is dependent on $\phi_k^{(i)}$, one has to be optimized first to optimize the other. Here, $a_k^{(i+1)}$ is fixed to evaluate $\phi_k^{(i)}$. In every iteration the dual variables are linearly interpolated between the fixed iterate $a_k^{(i)}$.

The KKT expression in (39) is solved in an iterative manner by initializing the dual variable a_k with ones to have equal priorities among the user set in the system. The transmit precoder $\mathbf{w}_k$ in equation (39) depends on the BS specific dual variable c_k, which is found by the bisection search satisfying the total power constraint in (36d).

Inorder to obtain a possible practical decentralized precoder design, we consider that the BS b knows the initial channel $\mathbf{h}_{b_k,k} \forall k$. Once, after receiving the updated transmit precodersfrom all BSs in $\mathcal{B}$, each user evaluates its corresponding precoder vector and is notify that to the BS via UL precoded pilots. On obtaining the pilots the BS updates all the values. Using the current updated values the $a_k^{(i)}, b_k^{(i)}, \mathbf{w}_k^{(i)}, \gamma_k^{(i)}$ are valuates using (39) and the updated dual variables are exchanged between the BS to evaluate the transmit precoders for the next iteration.

The users belonging to a particular BSs perform all the processing that is required and will update the precoders based on the feedback information from the user, inorder to avoid back haul exchanges within the BS. Once the transmit precoders are obtained from the BS, every user update their dual variables $a_k^{(i)}$ and $b_K^{(i)}$ and the transmit precoder $\mathbf{w}_k$ and rate γ_k is updated. After recieving the updates the BS use the effective channel to update the transmit precoders. Algorithm 2 gives a practical way for updating the transmit precoders for the KKT based AP-GP reformulated WSRM problem. The convergence analysis for the algorithm is discussed in the Appendix.

3.b. KKT for AP - GP Method with Rate Constraint

In this section we discuss a way to decentralize the precoder design across the corresponding BS in $\mathcal{B}$ based on AP-GP method with rate constraint. In this method also the problem is solved using the KKT conditions. The weighted sum rate maximization problem with QoS constraints subject to convex transmit power constraint $\mathcal{P}$ is considered. Let us consider the same convex optimization problem as in (4.8) the objective function (4.8a) and the constraint equation set from (4.9ba) and (4.8c) to (4.8d) remains the same. In addition we add a rate constraint to the equations, making a total of four constraints to the objective function.

$$\begin{aligned} \underset{w_k,\gamma_k,\beta_k}{\text{maximize}} \quad & \sum_k \log_2 e \, \log(1+\gamma_k) && \text{(40a)} \\ \text{subject to} \quad & (37) \quad and \quad (36c) - (36d) \\ & d_k : \log_2 e \, \log_e(1+\gamma_k) \geq R_0 && \text{(40b)} \end{aligned}$$

The variable a_k, b_k, c_k, d_k are dual variables corresponding to equations in (40). The optimization variables is the transmit precoder vector $\mathbf{w}_k \in \mathbb{C}^{N_T} \forall k$, where k= 1,2,. . .,K.

$$
\begin{aligned}
a_k^{(i)} &\longrightarrow \frac{\phi_k^{(i)}}{1+\gamma_k^{(i-1)}} \\
b_k^{(i)} &\longrightarrow \frac{a_k^{(i)}\phi_k^{(i)}}{2} \\
\mathbf{w}_k^{(i)} &\longrightarrow \frac{a_k^{(i)}}{2}\left(\sum_{i\neq K} b_i^{(i)}\mathbf{h}_{b_k,i}^H\mathbf{h}_{b_k,i} + c_k\mathbf{I}_{N_T}\right)^{-1}\mathbf{h}_{b_k,k}^H \\
\gamma_k^{(i)} &\longrightarrow 2\phi_k^{(i)}\left(\Re\{\mathbf{h}_{b_k,k}\mathbf{w}_k\} - \frac{\phi_k^{(i)}\beta_k}{2}\right)
\end{aligned}
\tag{39}
$$

List 5.1: Update Procedure

Algorithm 2 KKT for AP-GP Method with and without Rate Constraint

Input: $\alpha_k, \mathbf{h}_{b_k,k}, \forall b \in \mathcal{B}, \forall k \in \mathcal{U}_b$.
Output: $\mathbf{w}_k, \forall k \in \{1, 2, \ldots, K\}$
Initialization: $i = 1$, dual variables $a_k{}^{(0)} = 1$, and I_{max} for certain value
repeat
  for each BS $b \in \mathcal{B}$ perform the following procedure
  update $\mathbf{w}_k^{(i)}$ using (5.13) and perform the downlink pilot transmission
  evaluate $\gamma_k^{(i)}, \phi_k^{(i)}, a_k^{(i)}$ using the equations.
  if Rate constraint exists **then**
    evaluate $d_k^{(i)}$ using equation (5.25)
    update dual variable $a_k^{(i)}$ using the equation (5.24)
  end if
  using the precoded uplink pilots $\mathbf{w}_k^{(i)}$ and $a_k^{(i)}$ are notified to all BS in $\mathcal{B}$
until convergence or $i \geq I_{max}$

The Lagrangian function for the problem can be written as,

$$L(\gamma_k, \beta_k, \mathbf{w}_k) = -\log_2 e \log_e(1+\gamma_k) + a_k \left(\frac{1}{2\phi_k}\gamma_k + \frac{\phi_k}{2}\beta_k - \mathbf{h}_{b_k,k}\mathbf{w}_k \right)$$
$$+ b_k \left(\sigma^2 + \sum_{\substack{i=1,\\ i\neq k}}^{K} \|\mathbf{h}_{b_i,k}\mathbf{w}_i\|^2 - \beta_k \right) + c_k \left(\sum_{k\in\mathcal{U}_b} \|\mathbf{w}_k\|_2^2 - P_b \right) + d_k \left(R_0 - \log_2 e \log_e(1+\gamma_k) \right) \quad (41)$$

By evaluating the Lagrangian function in (41) with respect to the primal and dual variables we obtain an iterative solution as

$$\begin{aligned}
a_k^{(i)} &\longrightarrow \frac{2\phi_k^{(i)}(1+d_k^{(i-1)})}{1+\gamma_k^{(i-1)}} \\
b_k^{(i)} &\longrightarrow \frac{a_k^{(i)}\phi_k^{(i)}}{2} \\
\mathbf{w}_k^{(i)} &\longrightarrow \frac{a_k^{(i)}}{2} \left(\sum_{i\neq K} b_i^{(i)} \mathbf{h}_{b_k,i}^H \mathbf{h}_{b_k,i} + c_k \mathbf{I}_{N_T} \right)^{-1} \mathbf{h}_{b_k,k}^H \\
\gamma_k^{(i)} &\longrightarrow 2\phi_k^{(i)} \left(\Re\{\mathbf{h}_{b_k,k}\mathbf{w}_k\} - \frac{\phi_k^{(i)}\beta_k}{2} \right) \\
d_k^{(i)} &\longrightarrow d_k^{(i-1)} + \rho \left(R_0' - \log(1+\gamma_k^{(i)}) \right)
\end{aligned} \quad (42)$$

List 5.2: Update Procedure

Since the dual variable $a_k^{(i)}$ is dependent on $\phi_k^{(i)}$, one has to be optimized first to optimize the other. Here, $a_k^{(i+1)}$ is fixed to evaluate $\phi_k^{(i)}$. In every iteration the dual variables are linearly interpolated between the fixed iterate $a_k^{(i)}$. The dual variable $d_k^{(i+1)}$ is iterated with a fixed $d_k^{(i)}$ using a step size ρ. Similar to ρ depends upon the system model and its behaviour. The stepsize ρ must be small or diminishing such that the convergence is gauranteed.

The KKT expression in (42) is solved in an iterative manner by initializing the dual variable a_k with ones to have equal priorities among the user set in the system. The transmit precoder $\mathbf{w}_k$ in equation (42) depends on the BS specific dual variable c_k, which is found by the bisection search satisfying the total power constraint in (36d). It can be noted that the fixed SCA operating point is given by $d_k^{(i)}$ which is also considered in the expression in (42).

Inorder to obtain a possible practical decentralized precoder design, we consider that the BS b knows the initial channel $\mathbf{h}_{b_k,k}\forall k$. Once, after receiving the updated transmit precodersfrom all BSs in $\mathcal{B}$, each user evaluates its corresponding precoder vector and is notify that to the BS via UL precoded pilots. On obtaining the pilots the BS updates all the values. Using the current updated values the $a_k^{(i)}, b_k^{(i)}, \mathbf{w}_k^{(i)}, \gamma_k^{(i)}$ are valuates using (42) and the updated dual variables are exchanged between the BS to evaluate the transmit precoders for the next iteration. The SCA operating point is also updated using the current rate γ_k.

The users belonging to a particular BSs perform all the processing that is required and will update the precoders based on the feedback information from the user, inorder to avoid back haul exchanges within the BS. Once the transmit precoders are obtained from the BS, every user update their dual variables $a_k^{(i)}$, $b_K^{(i)}$ and $d_k^{(i)}$ and the transmit precoder $\mathbf{w}_k$ and rate γ_k is updated. After recieving the updates the BS use the effective channel to update the transmit precoders. Algorithm 2 gives a practical way for updating the transmit precoders for the KKT based AP-GP with rate constraint reformulated WSRM problem. In the algorithm it can be observed that we have an inner loop if there is a rate constraint and if there is no rate constraint. According to our problem we can switch between the algorithms. The convergence analysis for the algorithm is discussed in the Appendix.

3.c. KKT for MSE without Rate Constraint

In this section we discuss a way to decentralize the precoder design across the corresponding BS in $\mathcal{B}$ based on MSE Reformulation without rate constraint. In this method also the problem is solved using the KKT conditions. The weighted sum rate maximization problem with QoS constraints subject to convex transmit power constraint $\mathcal{P}$ is solved by exploiting the relationship between the MSE and the achievable SINR when the Minimum Mean Square Error (MMSE) receivers are used at the terminals reference ganesh 4 and 5. The MSE ϵ_k for a data symbol d_k is given as

$$\epsilon_k = \mathbb{E}\left[(d_k' - d_k)^2\right] = |1 - u_k^H \mathbf{h}_{b_k,k}\mathbf{w}_k|^2 + \sum_{i\in\bar{\mathcal{U}}_b} \|u_i^H \mathbf{h}_{b_k,i}\mathbf{w}_i\|^2 + \bar{N}_0 \qquad (43)$$

where d_k' is the estimated data symbol or the esitmate of transmit symbol. It must be noted that for fixed recievers, equation (43) is a convex function in terms of transmit beamformer $w_k\forall k$. The receive beamformers $u_k\forall k$ can be solved directly by evaluating the roots of the gradients of the Lagrangian of our main problem in equation 1. The optimal receive beamformers turn out to be equal to the well known MMSE recievers. The optimal reciever is in fact a scaled version of MMSE reciever for k users given as

$$R_k = \sum_{i=1}^{K} \mathbf{h}_{b_k,k}\mathbf{w}_k\mathbf{w}_k^H\mathbf{h}_{b_k,k}^H + N_0 I_{N_R}$$
$$u_k = R_k^{-1}\mathbf{h}_{b_k,k}\mathbf{w}_k \tag{44}$$

The MMSE receiver in (44) can also be used without compromising the performance. Using the (44) in the MSE expression in (43) and when MMSE receive beamformers are applied for each spatial data stream, the corresponding SINR is inversely related to the MSE as,

$$\epsilon_k = (1+\gamma_k)^{-1}. \tag{45}$$

We apply the above equations and reformulate our WSRM problem. Also, we note that in the problem formulation above, the receive beamformers is no longer considered as optimization variable.

Let us consider the objective function that can be written as maximize $\sum_{i=1}^{K} \log(1+\gamma_k) \rightleftarrows$ minimize $\sum_{i=1}^{K} \log \epsilon_k$. The alternative formulation is non convex so we can take the SCA approach (reference) by relaxing the constraint by a sequence of convex subsets by a sequence of convex subsets using first order taylors series expansion around the fixed MSE point $\bar{\epsilon}_k$ as, minimize $\sum_{i=1}^{K}\{\log \bar{\epsilon}_k + \dfrac{\epsilon - \bar{\epsilon}_k}{\log_2 \bar{\epsilon}_k}\}$. Where, $\log_2 \bar{\epsilon}_k$ is a constant and $\frac{\epsilon_k}{\log_2 \bar{\epsilon}_k}$ is a variable. So with all this results, the equivalent optimization problem can be formulated with the objective function as, minimize $-\sum_{i=1}^{K} \log \epsilon_k$. Using these approximations for the rate constraint, the problem is solved for optimal transmit precoders $\mathbf{w}_k$, MSEs ϵ_k and the user rates over each sub channel for a fixed revceive beamformer. The optmization sub problem to find the transmit precoders for a fixed recieve beamformer $\mathbf{u}_k$ is

$$\underset{w_k,\epsilon_k,t_k}{\text{maximize}} \quad \sum_{i=1}^{K} \frac{\epsilon_k}{\bar{\epsilon}_k} \tag{46a}$$
$$\text{subject to} \quad a_k : \epsilon_k \geq |1 - u_k^H \mathbf{h}_{b_k,k}\mathbf{w}_k|^2 + \sum_{i\in\bar{\mathcal{U}}_b} \|u_i^H \mathbf{h}_{b_k,i}\mathbf{w}_i\|^2 + N_0 \tag{46b}$$
$$b_k : \sum_{k\in\mathcal{U}_b} \|\mathbf{w}_k\|_2^2 \leq P_b, \forall b \in \mathcal{B} \tag{46c}$$

where, a_k, b_k are the dual variables corresponding to the constraints in the equation (46b) and (46c).

The Lagrangian for the above problem in (46) is

$$L(\epsilon_k, \mathbf{w}_k) = \sum_{i=1}^{K} \frac{\epsilon_k}{\bar{\epsilon_k}} + a_k \left[|1 - \mathbf{u}_k^H \mathbf{h}_{b_k,k} \mathbf{w}_k|^2 + \sum_{i \in \bar{\mathcal{U}_b}} \|\mathbf{u}_i^H \mathbf{h}_{b_k,i} \mathbf{w}_i\|^2 + N_0 - \epsilon_k \right] + b_k \left[\sum_{k \in \mathcal{U}_b} \|\mathbf{w}_k\|_2^2 - \right. \tag{47}$$

By evaluating the Lagrangian function in (47) with respect to the primal and dual variables we obtain an iterative solution as

$$\begin{aligned} a_k^{(i)} &\longrightarrow \frac{1}{\bar{\epsilon_k}^{(i-1)}} \\ \mathbf{w_k}^{(i)} &\longrightarrow a_k \left(a_k \textstyle\sum_{i=1}^{K} \mathbf{h}_{b_i,k}^H \mathbf{u}_i \mathbf{u}_i^H \mathbf{h}_{b_k,k} + b_k I \right)^{-1} \mathbf{u}_k^H \mathbf{h}_{b_k,k} \\ \epsilon_k^{(i)} &\longrightarrow |1 - u_k^{H(i)} \mathbf{h}_{b_k,k} \mathbf{w}_k^{(i)}|^2 + \textstyle\sum_{i \in \bar{\mathcal{U}_b}} \|u_i^{H(i)} \mathbf{h}_{b_k,i} \mathbf{w}_i^{(i)}\|^2 + N_0 \end{aligned} \tag{48}$$

List 5.3: Update Procedure

The KKT expression is solved in an iterative way by initializing the transmit beamformer w_k with single user beamforming and the MMSE vector. The dual variables a_k is initialized with ones to have equal priorities for all the users in system. Then the transmit precoder is evaluated by making use of the equation in (48). The transmit precoder depends upon the BS specific dual variable b_k which can be found by the bisection search by satisfying the total power constraint. The fixed SCA operating point is given by $\bar{\epsilon}_k^{(i+1)} = \epsilon_k^{(i)}$.

Inorder to obtain a distributed precoder design, an assumption is made that each BS b knows the corresponding equivalent channel $\mathbf{u}_k^H \mathbf{h}_{b_k,k} \forall k \in \mathcal{U}$, which includes the receivers, $\mathbf{u}_k$ through precoded uplink pilot signaling. Once the updated transmitted precoder is received from all BSs in $\mathcal{B}$, each user evaluates the MMSE receiver in equation (4.20) and is updated to the BSs via the uplink precoded pilots. Upon recieving the pilot symbol the BS update the MSE as,

$$\epsilon_k^{(i)} = 1 - \mathbf{u}_k^{(i)} \mathbf{h}_{b_k,k} \mathbf{w}_k^{(i)} \tag{49}$$

Using the current MSE value $a - k^{(i)}$ is evaluated using (48) and the updated dual variables are exchanged between the BS to evaluate the transmit precoders $\mathbf{w}_k^{(i+1)}$ for the next iteration. The SCA operating point is also updated with the current MSE value.

The users belonging to a particular BSs perform all the processing that is required and will update the precoders based on the feedback information from the user, inorder to avoid back haul exchanges within the BS. Once the transmit precoders are obtained from the BS, every user update their dual variables $a_k^{(i)}$ and the transmit precoder $\mathbf{w}_k$ and rate ϵ_k is updated. After recieving the updates the BS use the effective channel to update the transmit precoders. Algorithm 3 gives a practical way for updating the transmit precoders for the KKT based MSE without rate constraint for reformulated WSRM problem. The convergence analysis for the algorithm is discussed in the Appendix.

Algorithm 3 KKT for MSE with and without Rate Constraint

Input: $\alpha_k, \mathbf{h}_{b_k,k}, \forall b \in \mathcal{B}, \forall k \in \mathcal{U}_b$.
Output: $\mathbf{w}_k, \forall k \in \{1, 2, \ldots, K\}$
Initialization: $i = 1$, dual variables $a_k{}^{(0)} = 1$, and I_{max} for certain value
repeat
  for each BS $b \in \mathcal{B}$ perform the following procedure
  update $\mathbf{w}_k^{(i)}$ using (5.33) and perform the downlink pilot transmission
  evaluate $\epsilon_k^{(i)}, a_k^{(i)}$ using the equations (5.41) and (5.42)
  if Rate constraint exists **then**
    evaluate $c_k^{(i)}$ using equation (5.43)
    update dual variable $a_k^{(i)}$ using the equation (5.42)
  end if
  using the precoded uplink pilots $\mathbf{w}_k^{(i)}$ and $a_k^{(i)}$ are notified to all BS in $\mathcal{B}$
until convergence or $i \geq I_{max}$

3.d. KKT for MSE with Rate Constraint

In this section we discuss a way to decentralize the precoder design across the corresponding BS in $\mathcal{B}$ based on MSE Reformulation with rate constraint. The problem is solved using the KKT conditions. The weighted sum rate maximization problem with QoS constraints subject to convex transmit power constraint $\mathcal{P}$ is solved by exploiting the relationship between the MSE and the achievable SINR when the MMSE receivers are used at the terminals reference ganesh 4 and 5. The problem KKT for MSE with rate constraint can be seen similar to KKT for MSE without rate constraint. We add a rate constraint to the existing KKT for MSE adding a total of three constraints to the problem in (4.18). Rest every assumptions remains the same in this formulation,

$$\begin{aligned} \underset{w_k,\epsilon_k,t_k}{\text{maximize}} \quad & \sum_{i=1}^{K} \frac{\epsilon_k}{\bar{\epsilon_k}} && \text{(50a)} \\ \text{subject to} \quad & (46b) - (46c) \\ & c_k : -\log \epsilon_k \geq R_0. && \text{(50b)} \end{aligned}$$

where, a_k, b_k, c_k are the dual variables corresponding to the constraints in the equation (50). We can observe that the the constraint in (50b) is non convex, so upon taking the first order taylors series approximation we obtain

$$c_k : -\log \bar{\epsilon}_k + \frac{\epsilon_k - \bar{\epsilon}_k}{\log \bar{\epsilon}_k} \geq R_0' \tag{51}$$

so that we can replace the constraint in (50b) with (51).

The Lagrangian for the above problem in (50) is given as

$$\begin{aligned} L(\epsilon_k, \mathbf{w}_k) = \sum_{i=1}^{K} \frac{\epsilon_k}{\bar{\epsilon_k}} + a_k \left[|1 - \mathbf{u}_k^H \mathbf{h}_{b_k,k} \mathbf{w}_k|^2 + \sum_{i \in \bar{\mathcal{U}_b}} \|\mathbf{u}_i^H \mathbf{h}_{b_k,i} \mathbf{w}_i\|^2 + N_0 - \epsilon_k \right] \\ + b_k \left[\sum_{k \in \mathcal{U}_b} \|\mathbf{w}_k\|_2^2 - P_b \right] + c_k \left[R_0' + \log \epsilon_k - \frac{\epsilon_k - \bar{\epsilon}_k}{\log \bar{\epsilon}_k} \right]. \end{aligned} \tag{52}$$

By evaluating the Lagrangian function in (52) with respect to the primal and dual variables we obtain an iterative solution as

$$\begin{aligned} a_k^{(i)} &\longrightarrow -\frac{c_k^{(i)}}{\log \bar{\epsilon_k}} + \frac{1}{\bar{\epsilon_k}} \\ \mathbf{w_k}^{(i)} &\longrightarrow a_k \left(a_k \textstyle\sum_{i=1}^{K} \mathbf{h}_{b_i,k}^H \mathbf{u}_i \mathbf{u}_i^H \mathbf{h}_{b_k,k} + b_k I \right)^{-1} \mathbf{u}_k^H \mathbf{h}_{b_k,k} \\ \epsilon_k^{(i)} &\longrightarrow |1 - u_k^{H(i)} \mathbf{h}_{b_k,k} \mathbf{w}_k^{(i)}|^2 + \textstyle\sum_{i \in \bar{\mathcal{U}_b}} \|u_i^{H(i)} \mathbf{h}_{b_k,i} \mathbf{w}_i^{(i)}\|^2 + N_0 \\ c_k^{(i+)} &\longrightarrow c_k^{(i)} + \alpha \left(R_0' + \log \bar{\epsilon}_k - \frac{\epsilon_k - \bar{\epsilon}_k}{\log \epsilon_k} \right) \end{aligned} \tag{53}$$

List 5.4: Update Procedure

The KKT expression is solved in an iterative way by initializing the transmit beamformer w_k with single user beamforming and the MMSE vector. The dual variables

a_k is initialized with ones to have equal priorities for all the users in system. Then the transmit precoder is evaluated by making use of the equation in (53). The transmit precoder depends upon the BS specific dual variable b_k which can be found by the bisection search by satisfying the total power constraint. The dual variable c_k is updated by SCA approximation. The fixed SCA operating point is given by $\bar{\epsilon}_k^{(i+1)} = \epsilon_k^{(i)}$.

Inorder to obtain a distributed precoder design, an assumption is made that each BS b knows the corresponding equivalent channel $\mathbf{u}_k^H \mathbf{h}_{b_k,k} \forall k \in \mathcal{U}$, which includes the receivers, $\mathbf{u}_k$ through precoded uplink pilot signaling. Once the updated transmitted precoder is received from all BSs in $\mathcal{B}$, each user evaluates the MMSE receiver in equation (53) and is updated to the BSs via the uplink precoded pilots. Upon recieving the pilot symbol the BS update the MSE as,

$$\epsilon_k^{(i)} = 1 - \mathbf{u}_k^{(i)} \mathbf{h}_{b_k,k} \mathbf{w}_k^{(i)} \tag{54}$$

Using the current MSE value $a_k^{(i)}$ is evaluated using (53) and the updated dual variables are exchanged between the BS to evaluate the transmit precoders $\mathbf{w}_k^{(i+1)}$ for the next iteration. The SCA operating point is also updated with the current MSE value.

The users belonging to a particular BSs perform all the processing that is required and will update the precoders based on the feedback information from the user, inorder to avoid back haul exchanges within the BS. Once the transmit precoders are obtained from the BS, every user update their dual variables $a_k^{(i)}$ and the transmit precoder $\mathbf{w}_k$ and rate ϵ_k is updated. After updating the precoders and rate the SCA update is made for the dual variable c_k and is updated. After receiving the updates the BS use the effective channel to update the transmit precoders. Algorithm 3 gives a practical way for updating the transmit precoders for the KKT based MSE with rate constraint for reformulated WSRM problem. In the algorithm there is an inner loop to find if there is a rate constraint and if the answer is yes, then the SCA update is made for the rate constraint with the dual variable c_k. The convergence analysis for the algorithm is discussed in the Appendix.

In general, all the above algorithms will converge to a feasible solution if the QoS constraints are feasible. For each non-feasible rate constraint, the sum rate variable $\gamma_k \forall k = 1, 2, \ldots, K$ will increase until the rate constraints are satisfied. When the problem is non feasible from the start, then the algorithm oscillates among a group of non-feasible rate constraints. The behavior of these algorithms can be seen in the next section Numerical Results, where each scenario is given as example. The convergence analysis of the distributed algorithm is discussed in Appendix

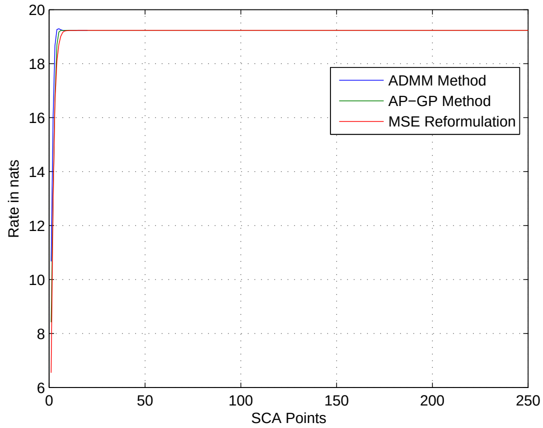
Theorem 1. *Every limit point of the sequence generated by above algorithms is a stationary point.*

Proof. See Appendix. $\square$

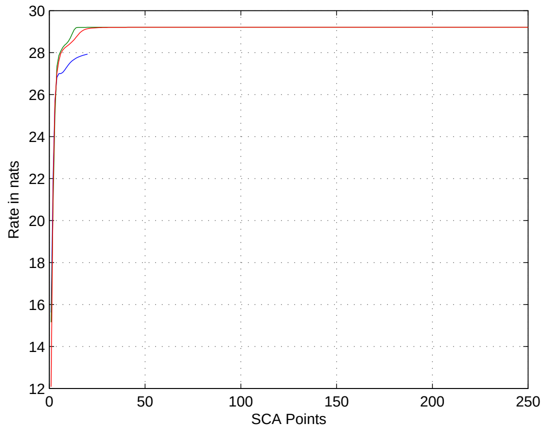
6. SIMULATION RESULTS

The Simulation work for this thesis work was carried out using MAT Lab. For this thesis work, we consider the Path Loss (PL) model that varies uniformly across all the users in the system with the channels drawn from the i.i.d samples. The behavior of the proposed decentralized algorithm is considered for different cell system. The systems convergence behavior and robustness is studied.

To begin with we consider a MIMO model with operating point at 10 dB SNR, we consider a system with $N_B = 2$ BSs, each equipped with $N_T = 4$ transmit antennas serving $k = 3$ users each. The allocations for BS and users are made by selecting the BS with the lowest PL component. Fig. 1 (a) shows the performance of distributed schemes for single antenna recieve system. From the figure a comparission can be made in terms of total rate after each SCA update. In figure 1 we can see that all the distributed algorithms converge to the same point

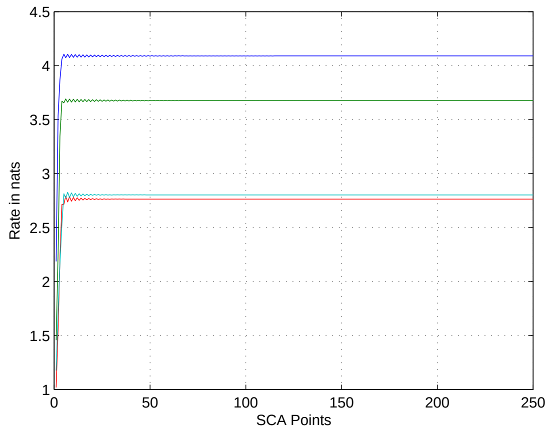


(a) Convergence of sum rate with $N_{\mathrm{T}} = 8, B = 2, K = 4$, Mobile users

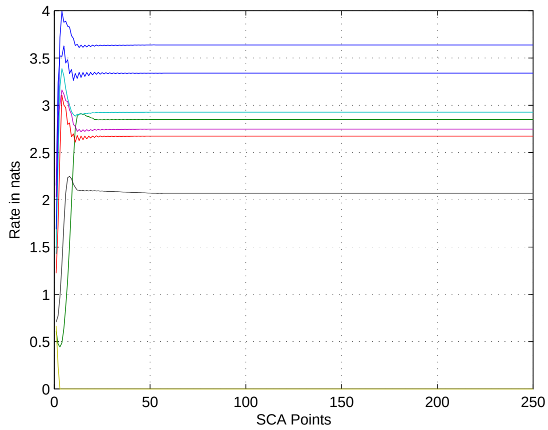


(b) Convergence of sum rate with $N_{\mathrm{T}} = 8, B = 2, K = 8$, Mobile users

Figure 0.7: Sum rate convergence for mobile users

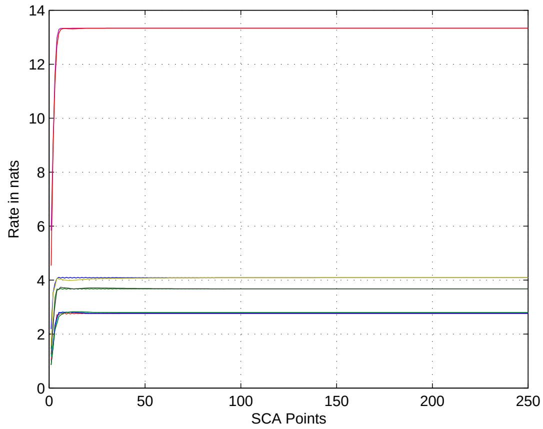


(a) Convergence of sum rate with $N_{\mathrm{T}} = 8, B = 2, K = 4$, Mobile users

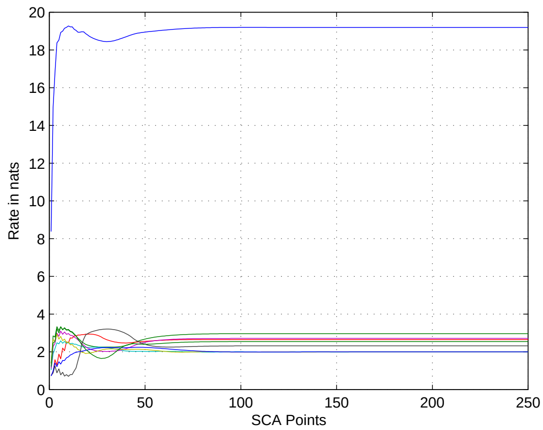


(b) Convergence of sum rate with $N_{\mathrm{T}} = 8, B = 2, K = 8$, Mobile users

Figure 0.8: Sum rate convergence for mobile users



(a) Convergence of sum rate with $N_{\mathrm{T}} = 8, B = 2, K = 4$, Mobile users



(b) Convergence of sum rate with $N_{\mathrm{T}} = 8, B = 2, K =$, Mobile users

Figure 0.9: Sum rate convergence for mobile users

7. SUMMARY AND CONCLUSION

In this thesis, distributed precoder design for WSRM problem in MIMO-IBC networks over **iid!** (**iid!**) channels was studied. We considered mainly two different precoding techniques, the centralized and the distributed approach with practical aspects.

8. REFERENCES

[1] Alamouti S.M. (1998) A simple transmit diversity technique for wireless communications. Selected Areas in Communications, IEEE Journal on 16, pp. 1451–1458.

9. APPENDICES

1.. Convergence Proof for Centralized Algorithm

2.. Convergence Proof for Distributed Algorithm

The convergence of the distributed algorithms as proposed in Algorithm 1, Algorithm 2 and Algorithm 3 follows the same discussion as in Convergence proof for centralized algorithm, if the sub problem in (4.1) converges to the centralized solution. Slater's condition is a sufficient condition for strong duality to hold for a convex optimization problem, Slater's condition states that the feasible region must have an interior point which is satisfied by (4.1) where, by having a non empty set interior and a compact set as required for the convergence 'reference'. Let us consider each distributed algorithms into our convergence analysis.

In Primal decomposition, the master subproblem uses subgradient to update the coupling variable (interference vectors) in consensus with the objective function, the convergence of the sub problem is gauranteed as the iteration index tends to infinite for a diminishing step size parameter 'reference'.

In ADMM method, we prove the convergence by considering the problem discussed in 'reference' $'boydhttps://web.stanford.edu/boyd/papers/pdf/admm_slides.pdf'$ by writing the problem as

$$\begin{aligned} \underset{x,z}{\text{maximize}} \quad & f(x) + g(z) && \text{(55a)} \\ \text{subject to} \quad & Ax = z && \text{(55b)} \\ & x \in \mathbb{C}_1, z \in \mathbb{C}_2 && \text{(55c)} \end{aligned}$$

Let us consider the decomposition via KKT conditions for AP-GP method without rate constraint, presented in section 2.A, updates all the optimization variables at once

3.. KKT condition for AP-GP method and MSE Reformulation with and without Rate Constraint

In order to solve for an iterative precoder design algorithm, the KKT expressions for the problem in (36) and (40) are obtained by differentiating the Lagrangian by assuming their constraint.

3.a. AP-GP without Rate Constraint

By differentiating the Lagrangian by assuming the equality constraint for (36). At stationary points, the following conditions are satisfied.

$$\nabla_{\gamma_k} \; : \; -\frac{-\log_2 e}{1+\gamma_k} + \frac{a_k}{2\phi_k} \; = \; 0 \tag{56a}$$

$$\nabla_{\beta_k} \; : \; -\frac{a_k \phi_k}{2} - b_k \; = \; 0 \tag{56b}$$

$$\nabla_{\mathbf{w}_k} \; : 2\mathbf{w}_k \left(\sum_{i \neq K} b_i \mathbf{h}_{b_k,i}^H \mathbf{h}_{b_k,i} + c_k \mathbf{I}_{N_T} \right) \; = \; a_k \mathbf{h}_{b_k,k}^H. \tag{56c}$$

Along with the primal constraints given in (36), the complementary slackness must also be satisfied at stationary point. On solving the above equations in (7.1) with the complementary slackness conditions, we get an iterative algorithm for finding the transmit beamformer as shown in (39).

3.b. AP-GP with Rate Constraint

By differentiating the Lagrangian by assuming the equality constraint for (36) and (40b). At stationary points, the following conditions are satisfied.

$$\nabla_{\gamma_k} \; : \; \frac{a_k}{2\phi_k} - \frac{1}{1-\gamma_k} - d_k \; = \; 0 \tag{57a}$$

$$\nabla_{\beta_k} \; : \; -\frac{a_k \phi_k}{2} - b_k \; = \; 0 \tag{57b}$$

$$\nabla_{\mathbf{w}_k} \; : 2\mathbf{w}_k \left(\sum_{i \neq K} b_i \mathbf{h}_{b_k,i}^H \mathbf{h}_{b_k,i} + c_k \mathbf{I}_{N_T} \right) \; = \; a_k \mathbf{h}_{b_k,k}^H. \tag{57c}$$

Along with the primal constraints given in (36) and (40b), the complementary slackness must also be satisfied at stationary point. On solving the above equations in (7.2) with the complementary slackness conditions, we get an iterative algorithm for finding the transmit beamformer as shown in (42).

3.c. MSE with Rate Constraint

By differentiating the Lagrangian by assuming the equality constraint for (46). At stationary points, the following conditions are satisfied.

$$\nabla_{\epsilon_k} \;:\; \frac{1}{\bar{\epsilon_k}} - a_k \;=\; 0 \tag{58a}$$

$$\nabla_{\mathbf{w}_k} \;: \mathbf{w}_k \left(a_k \sum_{i \neq K} \mathbf{h}_{b_k,i}^H \mathbf{u}_i^H \mathbf{u}_i \mathbf{h}_{b_k,i} + b_k \mathbf{I}_{N_T} \right) \;=\; a_k \mathbf{u}_k^H \mathbf{h}_{b_k,k}. \tag{58b}$$

Along with the primal constraints given in (46), the complementary slackness must also be satisfied at stationary point. On solving the above equations in (7.3) with the complementary slackness conditions, we get an iterative algorithm for finding the transmit beamformer as shown in (48).

3.d. MSE with Rate Constraint

In order to solve for an iterative precoder design algorithm, the KKT expressions for the problem in (46) and (50b) are obtained by differentiating the Lagrangian by assuming their constraint.

$$\nabla_{\epsilon_k} \;:\; \frac{1}{\bar{\epsilon_k}} - \frac{c_k}{\log \bar{\epsilon_k}} - a_k \;=\; 0 \tag{59a}$$

$$\nabla_{\mathbf{w}_k} \;: \mathbf{w}_k \left(a_k \sum_{i \neq K} \mathbf{h}_{b_k,i}^H \mathbf{u}_i^H \mathbf{u}_i \mathbf{h}_{b_k,i} + b_k \mathbf{I}_{N_T} \right) \;=\; a_k \mathbf{u}_k^H \mathbf{h}_{b_k,k}. \tag{59b}$$

Along with the primal constraints given in (46) and (50b), the complementary slackness must also be satisfied at stationary point. On solving the above equations in (7.4) with the complementary slackness conditions, we get an iterative algorithm for finding the transmit beamformer as shown in (53).