

Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2009, Article ID 197682, 10 pages
doi:10.1155/2009/197682

Research Article

Optimal Multiuser MIMO Linear Precoding with LMMSE Receiver

Fang Shu, Wu Gang, and Li Shao-Qian

National Key Lab of Communication, University of Electronic Science and Technology of China, Chengdu 611731, China

Correspondence should be addressed to Fang Shu, fangshu@uestc.edu.cn

Received 21 January 2009; Revised 11 May 2009; Accepted 19 June 2009

Recommended by Cornelius van Rensburg

The adoption of multiple antennas both at the transmitter and the receiver will explore additional spatial resources to provide substantial gain in system throughput with the spatial division multiple access (SDMA) technique. Optimal multiuser MIMO linear precoding is considered as a key issue in the area of multiuser MIMO research. The challenge in such multiuser system is designing the precoding vector to maximize the system capacity. An optimal multiuser MIMO linear precoding scheme with LMMSE detection based on particle swarm optimization is proposed in this paper. The proposed scheme aims to maximize the system capacity of multiuser MIMO system with linear precoding and linear detection. This paper explores a simplified function to solve the optimal problem. With the adoption of particle swarm optimization algorithm, the optimal linear precoding vector could be easily searched according to the simplified function. The proposed scheme provides significant performance improvement comparing to the multiuser MIMO linear precoding scheme based on channel block diagonalization method.

Copyright © 2009 Fang Shu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

In recent years, with the increasing demand of transmitting high data rates, the (Multiple-Input Multiple-Output) MIMO technique, a potential method to achieve high capacity has attracted enormous interest [1, 2]. When multiple antennas are equipped at both base stations (BSs) and mobile stations (MSs), the space dimension can be exploited for scheduling multi-user transmission besides time and frequency dimension. Therefore, the traditional MIMO technique focused on point-to-point single-user MIMO (SU-MIMO) has been extended to the point-to-multipoint multi-user MIMO (MU-MIMO) technique [3, 4]. It has been shown that time division multiple access (TDMA) systems can not achieve sum rate capacity of MU-MIMO system of broadcast channel (BC) [5] while MU-MIMO with spatial division multiple access (SDMA) could, where one BS communicates with several MSs within the same time slot and the same frequency band [6, 7]. MU-MIMO based on SDMA improves system capacity taking advantage of multi-user diversity and precanceling of multi-user interference at the transmitter.

Traditional MIMO technique focuses on point-to-point transmission as the STBC technique based on space-time coding and the VBLAST technique based on spatial multiplexing. The former one can efficiently combat channel fading while its spectral efficiency is low [8, 9]. The latter one could transmit parallel data streams, but its performance will be degraded under spatial correlated channel [10, 11]. When the MU-MIMO technique is adopted, both the multi-user diversity gain to improve the BER performance and the spatial multiplexing gain to increase the system capacity will be obtained. Since the receive antennas are distributed among several users, the spatial correlation will effect less on multi-user MIMO system. Besides, because the multi-user MIMO technique utilizes precoding at the transmit side to precanceling the cochannel interference (CCI), so the complexity of the receiver can be significantly simplified. However multi-user CCI becomes one of the main obstacles to improve MU-MIMO performance. The challenge is that the receiving antennas that are associated with different users are typically unable to coordinate with each other. By mitigating or ideally completely eliminating CCI, the BS exploits the channel state information (CSI) available at the

transmitter to cancel the CCI at the transmitter. It is essential to have CSI at the BS since it allows joint processing of all users' signals which results in a significant performance improvement and increased data rates.

The sum capacity in a multiuser MIMO broadcast channel is defined as the maximum aggregation of all the users' data rates. For Gaussian MIMO broadcast channels (BCs), it was proven in [12] that Dirty Paper Coding (DPC) can achieve the capacity region. The optimal precoding of multi-user MIMO is based on dirty paper coding (DPC) theory with the nonlinear precoding method. DPC theory proves that when a transmitter has advance knowledge of the interference, it could design a code to compensate for it. It is developed by Costa which can eliminate the interference by iterative precoding at the transmitter and achieve the broadcast MIMO channel capacity [13, 14]. The famous Tomlinson-Harashima precoding (THP) is the non-linear precoding based on DPC theory. It is first developed by Tomlinson [15] and Miyakawa and Harashima [16] independently and then has become the Tomlinson-Harashima precoding (THP) [17–20] to combat the multi-user cochannel interference (CCI) with non-linear precoding. Although THP performs well in a multi-user MIMO scenario, deploying it in real-time systems is difficult because of its high complexity of the precoding at the transmitter. Many suboptimal MU-MIMO linear precoding techniques have emerged recently, such as the channel inversion method [21] and the block diagonalization (BD) method [22–24]. Channel inversion method [25] employs some traditional MIMO detection criterions, such as the Zero Forcing (ZF) and Minimum Mean Squared Error (MMSE), precoding at the transmitter to suppress the CCI. Channel inversion method based on ZF can suppress CCI completely; however it may lead to noise amplification since the precoding vectors are not normalized. Channel inversion method based on MMSE compromises the noise and the CCI, and outperforms ZF algorithm, but it still cannot obtain good performance. BD method decomposes a multi-user MIMO channel into multiple single user MIMO channels in parallel to completely cancel the CCI by making use of the null space. With BD, each users precoding matrix lies in the null space of all other users channels, and the CCI could be completely canceled. The generated null space vectors are normalized vectors, which could avoid the noise amplification problem efficiently. So BD method performs much better than channel inversion method. However, since BD method just aims to cancel the CCI and suppress the noise, its precoding gain is not optimized.

It is obvious that the CCI, the noise, and the precoding gain are the factors affecting on the performance of the preprocessing MU-MIMO. The above linear precoding methods just take one factor into account without entirely consideration. A rate maximization linear precoding method is proposed in [26]. This method aims to maximize the sum rate of the MU-MIMO system with linear preprocessing. However, the optimized function in [26] is too complex to compute. In this paper, we solve the optimal linear precoding with linear MMSE receiver problem in a more simplified way.

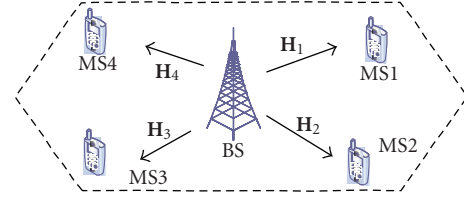


FIGURE 1: The configuration of MU-MIMO system

An optimal MU-MIMO linear precoding scheme with linear MMSE receiver based on particle swarm optimization (PSO) is proposed in this paper. PSO algorithm has been used in many complex optimization tasks, especially in solving the optimization of continuous space [27, 28]. In this paper, PSO is firstly introduced into MIMO research to solve some optimization issues. The adoption of PSO to MIMO system provides a new method to solve the MIMO processing problem. In this paper, we first analyze the optimal linear precoding vector with linear MMSE receiver and establish a simplified function to measure the optimal linear precoding problem. Then, we employ the novel PSO algorithm to search the optimal linear precoding vector according to the simplified function. The proposed scheme obtains significant MU-MIMO system capacity and outperforms the channel block diagonalization method.

This paper is organized into seven parts. The system model of MU-MIMO is given in Section 2. Then the analysis of optimal linear precoding with linear MMSE receiver is given in Section 3. The particle swarm optimization algorithm is given in Section 4. In Section 5, the proposed optimal linear precoding MU-MIMO scheme with LMMSE detection based on particle swarm optimization is introduced. In Section 6, the simulation results and comparisons are given. Conclusions are drawn in the last section. The channel block diagonalization algorithm is given in the appendix.

2. System Model of MU-MIMO

The MU-MIMO system could transmit data streams of multiple users of the same cellular at the same time and the same frequency resources as Figure 1 shows.

We consider an MU-MIMO system with one BS and K MS, where the BS is equipped with M antennas and each MS with N antennas, as shown in Figure 2. The point-to-multipoint MU-MIMO system is employed in downlink transmission.

Because MU-MIMO aims to transmit data streams of multiple-users at the same time and frequency resources, we discuss the algorithm at single-carrier, for each subcarrier of the multicarrier system, and it is processed as same as the single-carrier case. Since OFDM technique deals the frequency selective fading as flat fading, we model the channel as the flat fading MIMO channel:

$$\mathbf{H}_k = \begin{bmatrix} h_{1,1} & \cdots & h_{1,M} \\ \vdots & \ddots & \vdots \\ h_{N,1} & \cdots & h_{N,M} \end{bmatrix} \quad (1)$$

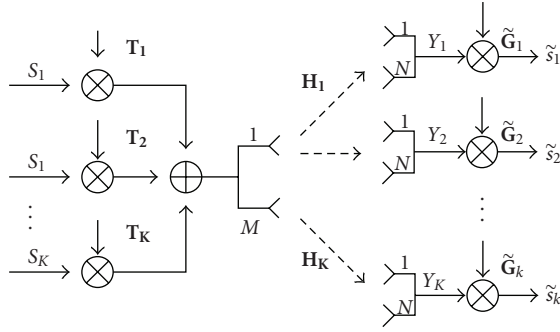


FIGURE 2: The block diagram of MU-MIMO system

where \mathbf{H}_k is the MIMO channel matrix of user k . $h_{i,j}$ indicates the channel impulse response coupling the j th transmit antenna to the i th receive antenna. Its amplitude obeys independent and identically Rayleigh-distribution.

Data streams of K ($K \leq M$) users are precoded by their precoding vectors \mathbf{T}_k ($k = 1 \dots K$) before transmission. \mathbf{T}_k is the $M \times 1$ normalized precoding vector for user k with $\mathbf{T}_k^H \mathbf{T}_k = 1$. The received signal at the k th user is

$$\begin{aligned} y_k &= \mathbf{H}_k \mathbf{T}_k \sqrt{p_k} s_k + \mathbf{H}_k \sum_{i=1, i \neq k}^K \mathbf{T}_i \sqrt{p_i} s_i + \mathbf{n}_k \\ \sum_{k=1}^K p_k &= p_0 \end{aligned} \quad (2)$$

where y_k is the received signal of user k . The elements of additive noise \mathbf{n}_k obey distribution $CN(0, N_0)$ that are spatially and temporarily white. p_k is the transmit signal power of the k th data stream, and p_0 is the total transmit power.

The received signal at the k th user can also be expressed as

$$\begin{aligned} y_k &= \mathbf{H}_k \mathbf{W} \mathbf{s} + \mathbf{n}_k \\ \mathbf{W} &= [\mathbf{T}_1 \quad \mathbf{T}_2 \quad \dots \quad \mathbf{T}_K] \\ \mathbf{s} &= [\sqrt{p_1} s_1 \quad \sqrt{p_2} s_2 \quad \dots \quad \sqrt{p_K} s_K]^T \end{aligned} \quad (3)$$

where \mathbf{s} is the transmitted symbol vector with K data streams, \mathbf{W} is the precoding matrix with K precoding vectors, and $[\cdot]^T$ denotes the matrix transposition:

$$\tilde{\mathbf{H}}_k = \mathbf{H}_k \mathbf{W} \quad (4)$$

The channel matrix $\tilde{\mathbf{H}}_k$ can be assumed as the virtual channel matrix of user k after precoding. At the receiver, a linear receiver $\tilde{\mathbf{G}}_k$ is exploited to detect the transmit signal for the user k . The detected signal of the k th user is

$$\hat{s}_k = \tilde{\mathbf{G}}_k y_k. \quad (5)$$

The linear receiver $\tilde{\mathbf{G}}_k$ can be designed by ZF or MMSE criteria, and linear MMSE will obtain better performance.

In order to simplify the analysis, the power allocation is assumed as equal $\beta = p_k/N_0 = p_0/KN_0$, and linear MMSE MIMO detection is used in this paper as

$$\tilde{\mathbf{G}}_k = \tilde{\mathbf{h}}_k^H (\tilde{\mathbf{H}}_k \tilde{\mathbf{H}}_k^H + \beta \mathbf{I}_N)^{-1}, \quad (6)$$

where $(\cdot)^{-1}$ indicates the inverse of the matrix, $(\cdot)^H$ denotes the matrix conjugation transposition, and \mathbf{I}_N is the $N \times N$ identity matrix:

$$\tilde{\mathbf{h}}_k = \mathbf{H}_k \mathbf{T}_k = [\tilde{\mathbf{H}}_k]_k = [\mathbf{H}_k \mathbf{W}]_k, \quad (7)$$

where $[\cdot]_k$ denotes the k th column of the matrix. Then the detected SINR for the user k with the linear detection is

$$\begin{aligned} \text{SINR}_k &= \frac{\beta |\tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_k|^2}{\sum_{i=1, i \neq k}^K \beta |\tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_i|^2 + \|\tilde{\mathbf{G}}_k\|_2^2} \\ &= \frac{\beta |\tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_k|^2 / \|\tilde{\mathbf{G}}_k\|_2^2}{\sum_{i=1, i \neq k}^K \beta |\tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_i|^2 / \|\tilde{\mathbf{G}}_k\|_2^2 + 1}, \end{aligned} \quad (8)$$

where $\|\cdot\|_2$ denotes the matrix two-norm.

Because the nonnormalized precoding vector will amplify the noise at the receiver, the precoding vectors \mathbf{T}_k are assumed to be normalized as follow:

$$\|\mathbf{T}_k\|^2 = 1 \quad (9)$$

for $k = 1, \dots, K$.

3. Optimal Multiuser MIMO Linear Precoding

We assume that the MIMO channel matrices \mathbf{H}_k ($k = 1, \dots, K$) are available at the BS. It can be achieved either by channel reciprocity characteristics in time-division-duplex (TDD) mode or by feedback in frequency-division-duplex (FDD) mode. And the channel matrix \mathbf{H}_k is known at the receiver k through channel estimation. We just discuss the equal power allocation case in this paper. The optimal power allocation is achieved through water-filling according to the SINR of each user.

The MIMO channel of user k can be decomposed by the singular value decomposition (SVD) as

$$\mathbf{H}_k = \mathbf{U}_k \Sigma_k \mathbf{V}_k^H. \quad (10)$$

If we apply $\mathbf{T}_k = [\mathbf{V}_k]_1$ to precode for user k , it obtains the maximal precoding gain as follow.

Lemma 1. One has

$$\|\mathbf{H}_k \mathbf{T}_k\|_2 = \|\mathbf{H}_k [\mathbf{V}_k]_1\|_2 = \lambda_k^{\max}, \quad (11)$$

where $[\mathbf{V}_k]_1$ denotes the first column of \mathbf{V}_k , and λ_k^{\max} denotes the maximal singular value of \mathbf{H}_k .

Proof. One has

$$\begin{aligned}\|\mathbf{H}_k \mathbf{T}_k\|_2 &= \sqrt{(\mathbf{H}_k \mathbf{T}_k)^H (\mathbf{H}_k \mathbf{T}_k)} \\ &= \sqrt{(\mathbf{H}_k [\mathbf{V}_k]_1)^H (\mathbf{H}_k [\mathbf{V}_k]_1)} \\ &= \sqrt{(\lambda_k^{\max} [\mathbf{U}_k]_1)^H (\lambda_k^{\max} [\mathbf{U}_k]_1)}.\end{aligned}\quad (12)$$

So

$$\|\mathbf{H}_k \mathbf{T}_k\|_2 = \lambda_k^{\max}, \quad (13)$$

where $[\mathbf{U}_k]_1$ denotes the first column of unitary matrix \mathbf{U}_k . \square

Thus, precoding with the singular vector corresponding to the maximal singular value is an initial thought to obtain good performance. However, if the singular vector is directly used at the transmitter as the precoding vector, the CCI will be large, and the performance will be degraded severely. Only for the special case that the MIMO channel among all these users are orthogonal that the CCI will be zero if we directly use the singular vector of each user as its precoding vector. But in realistic case, the transmit users' channels are always nonorthogonal, and so the singular vector could not be utilized directly. We have drawn some analysis as follow.

(1) *Ideal channel case.* The ideal channel case is that the MIMO channels of transmitting users' are orthogonal. There is

$$|([\mathbf{V}_k]_1)^H [\mathbf{V}_i]_1| = 0 \quad (i \neq k) \quad (14)$$

If we apply $\mathbf{T}_k = [\mathbf{V}_k]_1$ to precode for user k , the maximal precoding gain will be obtained as (13) shows, and the CCI will turn to zero as follow.

Lemma 2. *One has*

$$\frac{\sum_{i=1, i \neq k}^K \beta \left| \tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_i \right|^2}{\left\| \tilde{\mathbf{G}}_k \right\|_2^2} = 0. \quad (15)$$

Proof. One has

$$\sum_{i=1, i \neq k}^K \left| \tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_i \right|^2 = \sum_{i=1, i \neq k}^K \left| \tilde{\mathbf{G}}_k \mathbf{H}_k [\mathbf{V}_i]_1 \right|^2, \quad (16)$$

$$\tilde{\mathbf{G}}_k = (\mathbf{H}_k \mathbf{T}_k)^H \left(\tilde{\mathbf{H}}_k \tilde{\mathbf{H}}_k^H + \frac{1}{\beta} \mathbf{I}_M \right)^{-1}, \quad (17)$$

$$\mathbf{H}_k \mathbf{T}_k = \mathbf{H}_k [\mathbf{V}_k]_1 = \lambda_k^{\max} [\mathbf{U}_k]_1, \quad (18)$$

$$\tilde{\mathbf{G}}_k = (\lambda_k^{\max} [\mathbf{U}_k]_1)^H \left((\mathbf{H}_k \mathbf{W})(\mathbf{H}_k \mathbf{W})^H + \frac{1}{\beta} \mathbf{I}_M \right)^{-1}, \quad (19)$$

$$(\mathbf{H}_k \mathbf{W})(\mathbf{H}_k \mathbf{W})^H = \mathbf{U}_k \Sigma_k \mathbf{V}_k^H \mathbf{W} \mathbf{W}^H \mathbf{V}_k \Sigma_k \mathbf{U}_k^H. \quad (20)$$

Because we assume that $|([\mathbf{V}_k]_1^H [\mathbf{V}_i]_1)| = 0$ ($i \neq k$), and $[\mathbf{V}_k]_1$ is the unit vector with $|([\mathbf{V}_k]_1^H [\mathbf{V}_k]_1)| = 1$, so $\mathbf{W} = [[\mathbf{V}_1]_1, [\mathbf{V}_2]_1, \dots, [\mathbf{V}_K]_1]$ is an unitary matrix with

$$\mathbf{W} \mathbf{W}^H = \mathbf{I}_M,$$

$$\begin{aligned}\tilde{\mathbf{G}}_k &= (\mathbf{H}_k \mathbf{T}_k)^H \left(\tilde{\mathbf{H}}_k \tilde{\mathbf{H}}_k^H + \frac{1}{\beta} \mathbf{I}_M \right)^{-1} \\ &= (\lambda_k^{\max} [\mathbf{U}_k]_1)^H \left(\mathbf{U}_k \Sigma_k^2 \mathbf{U}_k^H + \frac{1}{\beta} \mathbf{I}_M \right)^{-1} \\ &= \lambda_k^{\max} [\mathbf{U}_k]_1^H \left(\Sigma_k^2 + \frac{1}{\beta} \mathbf{I}_M \right) \mathbf{U}_k^H \right)^{-1}.\end{aligned}\quad (21)$$

Since $([\mathbf{U}_k]_1)^H (\mathbf{U}_k^H)^{-1} = [1, 0, \dots, 0]$, so

$$\begin{aligned}\tilde{\mathbf{G}}_k &= \lambda_k^{\max} [\mathbf{U}_k]_1^H (\mathbf{U}_k^H)^{-1} \left(\Sigma_k^2 + \frac{1}{\beta} \mathbf{I}_M \right)^{-1} \mathbf{U}_k^{-1} \\ &= \frac{\lambda_k^{\max}}{(\lambda_k^{\max})^2 + 1/\beta} [\mathbf{U}_k^H]_1^1 \\ &= \frac{\lambda_k^{\max}}{(\lambda_k^{\max})^2 + 1/\beta} [\mathbf{U}_k]_1^H,\end{aligned}\quad (22)$$

where $[\mathbf{U}_k^H]_1^1$ denotes the first row of \mathbf{U}_k^H .

Also

$$\begin{aligned}\sum_{i=1, i \neq k}^K \left| \tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_i \right|^2 &= \sum_{i=1, i \neq k}^K \left| \frac{\lambda_k^{\max}}{(\lambda_k^{\max})^2 + 1/\beta} [\mathbf{U}_k]_1^H \mathbf{H}_k \mathbf{T}_i \right|^2 \\ &= \sum_{i=1, i \neq k}^K \left| \frac{\lambda_k^{\max}}{(\lambda_k^{\max})^2 + 1/\beta} [\mathbf{U}_k]_1^H \mathbf{U}_k \Sigma_k \mathbf{V}_k^H [\mathbf{V}_i]_1 \right|^2.\end{aligned}\quad (23)$$

Since $|([\mathbf{V}_k]_1)^H [\mathbf{V}_i]_1| = 0$, so $\mathbf{V}_k^H [\mathbf{V}_i]_1 = [0, a_1, \dots, a_{K-1}]^T$. Combining $([\mathbf{U}_k]_1)^H \mathbf{U}_k = [1, 0, \dots, 0]$, there is

$$\sum_{i=1, i \neq k}^K \left| \tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_i \right|^2 = \sum_{i=1, i \neq k}^K \left| \tilde{\mathbf{G}}_k \mathbf{H}_k [\mathbf{V}_i]_1 \right|^2 = 0. \quad (24)$$

so

$$\frac{\sum_{i=1, i \neq k}^K \beta \left| \tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_i \right|^2}{\left\| \tilde{\mathbf{G}}_k \right\|_2^2} = 0. \quad (25)$$

\square

After linear MMSE detection at the receiver, user k obtains the maximal SINR as follows.

Lemma 3. *One has*

$$\text{SINR}_k = \frac{\beta \left| \tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_k \right|^2}{\left\| \tilde{\mathbf{G}}_k \right\|_2^2} = \beta (\lambda_k^{\max})^2. \quad (26)$$

Proof. One has

$$\text{SINR}_k = \frac{\beta(\tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_k)^H (\tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_k)}{|\tilde{\mathbf{G}}_k \tilde{\mathbf{G}}_k^H|}. \quad (27)$$

According to (13) and (22)

$$\begin{aligned} & (\tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_k)^H (\tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_k) \\ &= \left(\frac{\lambda_k^{\max}}{(\lambda_k^{\max})^2 + 1/\beta} [\mathbf{U}_k]_1^H \lambda_k^{\max} [\mathbf{U}_k]_1 \right)^H \\ & \quad \times \left(\frac{\lambda_k^{\max}}{(\lambda_k^{\max})^2 + 1/\beta} [\mathbf{U}_k]_1^H \lambda_k^{\max} [\mathbf{U}_k]_1 \right) \\ &= \left(\frac{(\lambda_k^{\max})^2}{(\lambda_k^{\max})^2 + 1/\beta} \right)^2 \\ |\tilde{\mathbf{G}}_k \tilde{\mathbf{G}}_k^H| &= \left(\frac{\lambda_k^{\max}}{(\lambda_k^{\max})^2 + 1/\beta} \right)^2 \end{aligned} \quad (28)$$

$$\text{SINR}_k = \beta(\lambda_k^{\max})^2. \quad \square$$

(2) *Ill channel case* The ill channel case is that all these transmitting users' channels are highly correlated. There is

$$|([\mathbf{V}_k]_1)^H [\mathbf{V}_i]_1| = 1 \quad (i \neq k). \quad (29)$$

If we still apply $\mathbf{T}_k = [\mathbf{V}_k]_1$ to precode for user k , the multiuser CCI will be very large, and the system performance will be degraded severely. The SINR after MMSE detection with equal power allocation for user k is as follows.

Lemma 4. *One has*

$$\text{SINR}_k = \frac{\beta(\lambda_k^{\max})^2}{\sum_{i=1, i \neq k}^K \beta(\lambda_k^{\max})^2 + 1}. \quad (30)$$

Proof. Since we have proven that when $\mathbf{T}_k = [\mathbf{V}_k]_1$ to precode for user k , then $\beta|\tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_k|^2 / \|\tilde{\mathbf{G}}_k\|_2^2 = \beta(\lambda_k^{\max})^2$, so

$$\begin{aligned} \text{SINR}_k &= \frac{\beta|\tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_k|^2 / \|\tilde{\mathbf{G}}_k\|_2^2}{\sum_{i=1, i \neq k}^K \beta|\tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_i|^2 / \|\tilde{\mathbf{G}}_k\|_2^2 + 1} \\ &= \frac{\beta(\lambda_k^{\max})^2}{\sum_{i=1, i \neq k}^K \beta|\tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_i|^2 / \|\tilde{\mathbf{G}}_k\|_2^2 + 1}. \end{aligned} \quad (31)$$

According to (19)

$$\begin{aligned} \tilde{\mathbf{G}}_k &= (\lambda_k^{\max} [\mathbf{U}_k]_1)^H \\ & \quad \times \left(\mathbf{U}_k \Sigma_k \mathbf{V}_k^H \mathbf{W} \mathbf{W}^H \mathbf{V}_k \Sigma_k \mathbf{U}_k^H + \frac{1}{\beta} \mathbf{I}_M \right)^{-1}. \end{aligned} \quad (32)$$

Since we assume that $|([\mathbf{V}_k]_1)^H [\mathbf{V}_i]_1| = 1$ ($i \neq k$), so $\mathbf{W} = [[\mathbf{V}_1]_1, [\mathbf{V}_2]_1, \dots, [\mathbf{V}_K]_1]$ is

$$\mathbf{W} = \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix}. \quad (33)$$

Let the diagonal matrix $\bar{\Sigma}_k = \Sigma_k \mathbf{V}_k^H \mathbf{W} \mathbf{W}^H \mathbf{V}_k \Sigma_k$, and so there is

$$\tilde{\mathbf{G}}_k = (\lambda_k^{\max} [\mathbf{U}_k]_1)^H \left(\mathbf{U}_k \left(\bar{\Sigma}_k + \frac{1}{\beta} \mathbf{I}_M \right) \mathbf{U}_k^H \right)^{-1}. \quad (34)$$

Since $([\mathbf{U}_k]_1)^H (\mathbf{U}_k^H)^{-1} = [1, 0, \dots, 0]$, so

$$\begin{aligned} \tilde{\mathbf{G}}_k &= \frac{\lambda_k^{\max}}{\bar{\lambda}_k + 1/\beta} [\mathbf{U}_k^H]^1 \\ &= \frac{\lambda_k^{\max}}{\bar{\lambda}_k + 1/\beta} [\mathbf{U}_k]_1^H, \end{aligned} \quad (35)$$

where $\bar{\lambda}_k$ indicates the first diagonal element of the diagonal matrix $\bar{\Sigma}_k$. So there is

$$\begin{aligned} \tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_i &= \frac{\lambda_k^{\max}}{\bar{\lambda}_k + 1/\beta} [\mathbf{U}_k]_1^H \mathbf{H}_k \mathbf{T}_i \\ &= \frac{\lambda_k^{\max}}{\bar{\lambda}_k + 1/\beta} [\mathbf{U}_k]_1^H \mathbf{U}_k \Sigma_k \mathbf{V}_k^H [\mathbf{V}_i]_1. \end{aligned} \quad (36)$$

Since $|([\mathbf{V}_k]_1)^H [\mathbf{V}_i]_1| = 1$, so $\mathbf{V}_k^H [\mathbf{V}_i]_1 = [1, a_1, \dots, a_{K-1}]^T$. Combining $([\mathbf{U}_k]_1)^H \mathbf{U}_k = [1, 0, \dots, 0]$, there is

$$\begin{aligned} |\tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_i|^2 &= \left(\frac{(\lambda_k^{\max})^2}{\bar{\lambda}_k + 1/\beta} \right)^2, \\ \|\tilde{\mathbf{G}}_k\|_2^2 &= |\tilde{\mathbf{G}}_k \tilde{\mathbf{G}}_k^H| = \left(\frac{\lambda_k^{\max}}{\bar{\lambda}_k + 1/\beta} \right)^2, \\ \frac{\sum_{i=1, i \neq k}^K \beta |\tilde{\mathbf{G}}_k \mathbf{H}_k \mathbf{T}_i|^2}{\|\tilde{\mathbf{G}}_k\|_2^2} &= \sum_{i=1, i \neq k}^K \beta (\lambda_k^{\max})^2. \end{aligned} \quad (37)$$

So the SINR for user k is

$$\text{SINR}_k = \frac{\beta(\lambda_k^{\max})^2}{\sum_{i=1, i \neq k}^K \beta(\lambda_k^{\max})^2 + 1} \quad \square \quad (38)$$

(3) *Practical case.* The practical case is that the transmitting users' channels are neither orthogonal nor ill. There is

$$\begin{aligned} |([\mathbf{V}_k]_1)^H [\mathbf{V}_i]_1| &\neq 0 \quad (i \neq k) \\ |([\mathbf{V}_k]_1)^H [\mathbf{V}_i]_1| &\neq 1 \quad (i \neq k). \end{aligned} \quad (39)$$

The practical case is usually in realistic environment. If we apply \mathbf{T}_k ($\mathbf{T}_k \neq [\mathbf{V}_k]_1$) to precode for user k , then

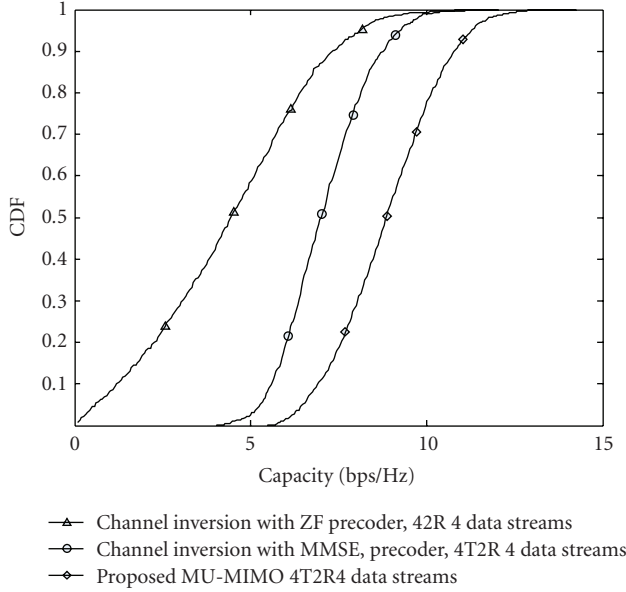


FIGURE 3: The system capacity CDF comparison of the two schemes.

$\xi_k = |\mathbf{T}_k^H [\mathbf{V}_k]_1|$ can be the parameter to measure the precoding gain, and $\rho_i = |\mathbf{T}_k^H [\mathbf{V}_i]_1|$ can be the parameter to measure the CCI. The SINR for user k according to the above analysis can be approximated denoted as

$$\begin{aligned} \text{SINR}_k &\approx \frac{\beta (\lambda_k^{\max} \xi_k)^2}{\sum_{i \neq k, i=1}^K \beta (\lambda_k^{\max} \rho_i)^2 + 1} \\ &= \frac{|\lambda_k^{\max} (\mathbf{T}_k^H [\mathbf{V}_k]_1)|^2}{\sum_{i \neq k, i=1}^K |\lambda_k^{\max} (\mathbf{T}_k^H [\mathbf{V}_i]_1)|^2 + 1/\beta}. \end{aligned} \quad (40)$$

The system capacity is related to SINR of the transmit users k , ($k = 1, \dots, K$). So in order to obtain the system capacity, we should obtain the SINR_k . Thus, when the optimal precoding vector is obtained by the PSO algorithm, the system capacity could be calculated by (41).

The system capacity of the MU-MIMO system can be indicated as

$$C_{\text{MU}} = \sum_{k=1}^K \log_2(1 + \text{SINR}_k). \quad (41)$$

We aim to maximize the system capacity of the MU-MIMO system in this paper. The optimal MU-MIMO linear precoding vector for the MU-MIMO system is the vector that can maximize the SINR at each receiver as

$$\begin{aligned} \bar{\mathbf{T}}_k &= \arg \max_{\mathbf{T}_k \in \mathbf{U}} \\ &\sum_{k=1}^K \log_2 \left(1 + \frac{|\lambda_k^{\max} (\mathbf{T}_k^H [\mathbf{V}_k]_1)|^2}{\sum_{i \neq k, i=1}^K |\lambda_k^{\max} (\mathbf{T}_k^H [\mathbf{V}_i]_1)|^2 + 1/\beta} \right) \end{aligned} \quad (42)$$

where \mathbf{U} denotes the unitary vector that $\mathbf{U}^H \mathbf{U} = \mathbf{I}$. From the above equation, it is clear that if we want to maximize

the system capacity of MU-MIMO, then the SINR of each user should be maximized. The SINR of user k is associated with three parameters as the singular vector correspond to the maximal singular value of all users and the noise.

4. The Particle Swarm Optimization Algorithm

Particle swarm optimization algorithm was originally proposed by Kennedy and Eberhart[27] in 1995. It searches the optimal problem solution through cooperation and competition among the individuals of population.

Imagine a swarm of bees in a field. Their goal is to find in the field the location with the highest density of flowers. Without any prior knowledge of the field, the bees begin in random locations with random velocities looking for flowers. Each bee can remember the location that is found the most flowers and somehow knows the locations where the other bees found an abundance of flowers. Torn between returning to the location where it had personally found the most flowers, or exploring the location reported by others to have the most flowers, the ambivalent bee accelerates in both directions to fly somewhere between the two points. There is a function or method to evaluate the goodness of a position as the fitness function. Along the way, a bee might find a place with a higher concentration of flowers than it had found previously. Constantly, they are checking the concentration of flowers and hoping to find out the absolute highest concentration of flowers.

Suppose that the size of swarm and the dimension of search space are C and D , respectively. Each individual in the swarm is referred to as a particle. The location and velocity of particle i ($i = 1, \dots, C$) are represented as the vector $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{iD}]^T$ and $\mathbf{v}_i = [v_{i1}, v_{i2}, \dots, v_{iD}]^T$. Each bee remembers the location where it personally encountered the most flowers which is denoted as $\mathbf{P}_i = [p_{i1}, p_{i2}, \dots, p_{iD}]^T$, which is the flight experience of the particle itself. The highest concentration of flowers discovered by the entire swarm is denoted as $\mathbf{P}_g = [p_{g1}, p_{g2}, \dots, p_{gD}]^T$, which is the flight experience of all particles. Each particle is searching for the best location according to \mathbf{P}_i and \mathbf{P}_g . The particle i updates its location and velocity according to the following two formulas [27]:

$$\begin{aligned} v_{id}^{t+1} &= wv_{id}^t + c_1\varphi_1(p_{id}^t - x_{id}^t) + c_2\varphi_2(p_{gd}^t - x_{id}^t) \\ x_{id}^{t+1} &= x_{id}^t + v_{id}^{t+1} \end{aligned} \quad (43)$$

where t is the current iteration number; v_{id}^t and $x_{id}^t + 1$ denote the velocity and location of the particle i in the d th dimensional direction. p_{id}^t is the individual best location of particle i in the d th dimensional direction, p_{gd}^t is the population best location in the d th dimensional direction. φ_1 and φ_2 are the random numbers between 0 and 1, c_1 and c_2 are the learning factors, and w is the inertia factor. Learning factors determine the relative "pull" of \mathbf{P}_i^t and \mathbf{P}_g^t that usually content $c_1 = c_2 = 2$. Inertia factor determines to what extent the particle remains along its original course unaffected by the pull of \mathbf{P}_g^t and \mathbf{P}_i^t that is usually between 0 and 1. After this process is carried out for each particle in the swarm, the

process is repeated until reaching the maximal iteration or the termination criteria are met.

5. The Optimal Linear Precoding Multiuser MIMO with LMMSE Detection Based on Particle Swarm Optimization

With the adoption of PSO algorithm and the simplified function (40), the optimal linear precoding vector $\bar{\mathbf{T}}_k$ ($k = 1, \dots, K$) could be easily searched.

The proposed optimal MU-MIMO linear precoding scheme based on PSO algorithm will search the optimal precoding vector for each user following 6 steps.

- (1) The BS obtains λ_k^{\max} , $[\mathbf{V}_k]_1$ and β of each user.
- (2) The BS employs the PSO algorithm to search the optimal linear precoding vector for each user. For user k , the PSO algorithm sets the maximal iteration number I and a group of M dimensional particles with the initial velocity $\mathbf{v}_{i,k}^1 = [v_{i1,k}^1, v_{i2,k}^1, \dots, v_{iM,k}^1]^T$ and the initial location $\mathbf{x}_{i,k}^1 = [x_{i1,k}^1, x_{i2,k}^1, \dots, x_{iM,k}^1]^T$ for particle i ($i = 1, \dots, C$). In order to accelerate the searching process, the initial location $\mathbf{x}_{i,k}^1$ could be initialized as $[\mathbf{V}_k]_1$, while the initial velocity $\mathbf{v}_{i,k}^1$ could be produced randomly. The real and imaginary parts of the initial velocity obey a normal distribution with mean zero and standard deviation one.
- (3) The BS begins to search with the initial location $\mathbf{x}_{i,k}^1$ and velocity $\mathbf{v}_{i,k}^1$. The goodness of the location is measured by the following equation:

$$f_{i,k}^t = \frac{|\lambda_k^{\max} [(\mathbf{x}_{i,k}^t)^H [\mathbf{V}_k]_1]|^2}{\sum_{j=1, j \neq k}^K |\lambda_k^{\max} [(\mathbf{x}_{i,k}^t)^H [\mathbf{V}_j]_1]|^2 + 1/\beta}, \quad (44)$$

where the fitness function $f_{i,k}^t$ indicates the obtained SINR for user k precoded by $\mathbf{x}_{i,k}^t$. The PSO algorithm finds $\mathbf{P}_{i,k}^t$ and $\mathbf{P}_{g,k}^t$ that are individual best location and population best location measured by (44) for the next iteration. $\mathbf{P}_{i,k}^t$ denotes the individual best location which means the best location of particle i at the t th iteration of the k th user. $\mathbf{P}_{g,k}^t$ denotes the population best location which means the best location of all particles at the t th iteration of the k th user.

- (4) For the t th iteration, the algorithm finds a $\mathbf{P}_{i,k}^t$ and a $\mathbf{P}_{g,k}^t$. The location and velocity for each particle will be updated according to (43) for the next iteration. In order to obtain the normalized optimal precoding vector to suppress the noise, the location $\mathbf{x}_{i,k}^t$ should be normalized in each iteration.
- (5) When reaching the maximal iteration number I , the algorithm stops, and $\mathbf{P}_{g,k}^I$ is the obtained optimal precoding vector for user k .
- (6) For an MU-MIMO system with K users, the scheme will search the precoding vectors according to the above criteria for each user.

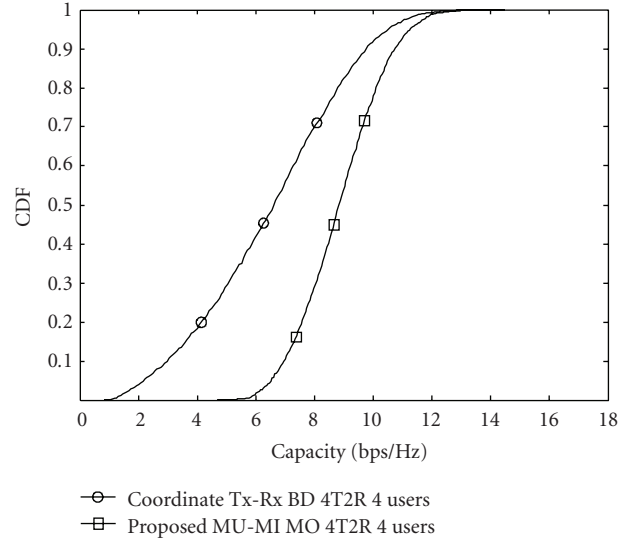


FIGURE 4: The system capacity CDF comparison of the two schemes.

6. Simulation Results

We simulated the proposed MU-MIMO scheme, the BD algorithm in [22] (Coordinate Tx-Rx BD), and the channel inversion algorithm in [25] in this paper to compare their performance under the same simulation environment.

Figure 3 is the system capacity comparison of the cumulative distribution function (CDF) of the channel inversion algorithm with ZF precoder and MMSE precoder and the proposed MU-MIMO algorithm when $M = 4, N = 2, p_0/N_0 = 5$ dB with equation power allocation and MMSE detection at the receiver. For channel inversion method, the BS transmits 4 data streams and 2 users simultaneously with 2 data stream for each user. For the proposed MU-MIMO, the BS transmit 4 data streams and 4 users simultaneously with 1 data stream for each user.

Figure 4 is the system capacity comparison of the CDF of the coordinated Tx-Rx BD algorithm and the proposed MU-MIMO algorithm when $M = 4, K = 4, N = 2, p_0/N_0 = 5$ dB with equation power allocation and MMSE detection at the receiver.

Figure 5 is the system capacity comparison of the CDF of the coordinated Tx-Rx BD algorithm and the proposed MU-MIMO algorithm when $M = 4, K = 4, N = 4, p_0/N_0 = 5$ dB with equation power allocation and MMSE detection at the receiver. Both the simulation results of the proposed MU-MIMO scheme with PSO algorithm from Figure 3 to Figure 5 are based on the PSO parameters with the particle number $C = 20$ and the iteration number $I = 20$. It could be seen that the proposed MU-MIMO scheme can effectively increase the system capacity compared to the BD algorithm and channel inversion algorithm.

Figure 6 is the average BER performance of the proposed MU-MIMO scheme and the coordinated Tx-Rx BD algorithm with $M = 4, K = 4, N = 4$. Figure 7 is the average BER performance of the proposed MU-MIMO scheme and the coordinated Tx-Rx BD algorithm with $M = 4, K = 4, N = 2$.

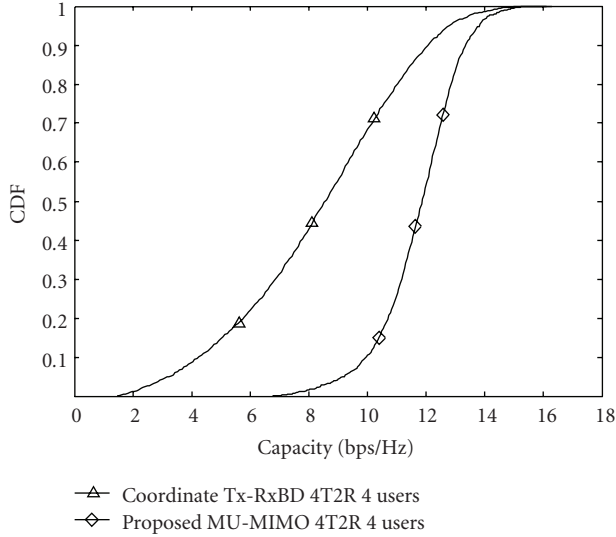


FIGURE 5: The system capacity CDF comparison of the two schemes.

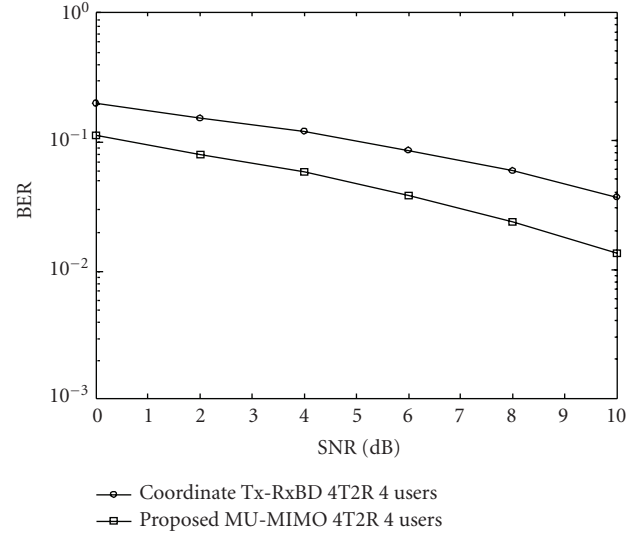


FIGURE 7: The BER comparison of the two schemes.

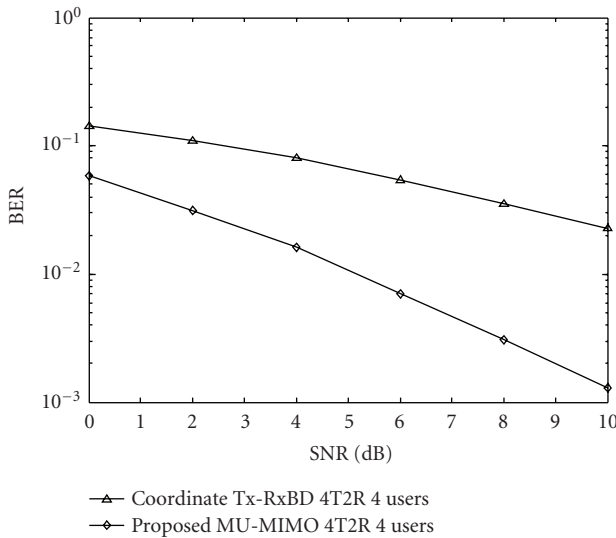
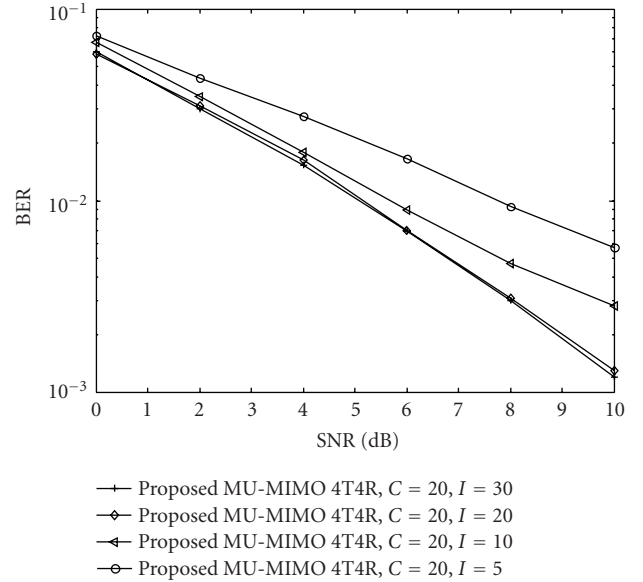


FIGURE 6: The BER comparison of the two schemes.

FIGURE 8: The BER comparison of the two schemes with different C and I .

Both the schemes adopt equal power allocation, MMSE detection, QPSK, and no channel coding. The proposed MU-MIMO scheme, with PSO algorithm from Figures 6 and 7 are based on the PSO parameters with the particle number $C = 20$ and the iteration number $I = 20$.

From the simulation results, it is clear that the proposed MU-MIMO linear precoding with LMMSE detection based on particle swarm optimization scheme outperforms the BD algorithm and the channel inversion algorithm. The reason lies in that the BD algorithm just aims to utilize the normalized precoding vector to cancel the CCI and suppress the noise. The channel inversion algorithm also aims to suppress CCI and noise. So the users' transmit signal covariance matrices of these schemes are generally not

optimal that are caused by the inferior precoding gain. The proposed MU-MIMO optimal linear precoding scheme aims to find the optimal precoding vector to maximize each users' SINR at each receiver to improve the total system capacity.

Figure 8 shows the BER performance of the proposed MU-MIMO scheme with the same particle size and different iteration size when $M = 4$, $K = 4$, $N = 4$. It adopts equal power allocation, MMSE detection, QPSK, and no channel coding. The particle number C is 20, and the iteration number scales from 5 to 30. We could see that when the iteration number is small, the proposed scheme could not obtain the best performance. With the increase of the iteration number, more performance as well as the computational complexity will increase too. However, when the iteration number is

larger than 20 for this case, the algorithm could not obtain more performance gain. Generally, for different case, the best iteration number is different. The iteration number is related to the transmit antenna number M at the BS and the transmit user number K ($K \leq M$). With the increasing of M or K , the iteration number should increase in order to obtain the best performance.

7. Conclusion

This paper solves the optimal linear precoding problem with LMMSE detection for MU-MIMO system in downlink transmission. A simplified optimal function is proposed and proved to maximize the system capacity. With the adoption of the particle swarm optimization algorithm, the optimal linear precoding vector with LMMSE detection for each user could be searched. The proposed scheme can obtain significant system capacity improvement compared to the multi-user MIMO scheme based on channel block diagonalization under the same simulation environment.

Appendix

Coordinated Tx-Rx BD Algorithm

Coordinated Tx-Rx BD algorithm is the improved BD algorithm. It could solve the antenna constraint problem in traditional BD algorithm and extends the BD algorithm to arbitrary antenna configuration. For a coordinated Tx-Rx BD algorithm with M transmit antennas at the BS, N receive antennas at the MS, and K users to be transmitted simultaneously, the algorithm follows 6 steps.

- (1) For $j = 1, \dots, K$, compute the SVD

$$\mathbf{H}_j = \mathbf{U}_j \boldsymbol{\Sigma}_j \mathbf{V}_j^H. \quad (\text{A.1})$$

- (2) Determine m_j , which is the number of subchannels for each user. In order to compare the two schemes fairly, $m_j = 1$ for each user.
- (3) For $j = 1, \dots, K$, let \mathbf{A}_j be the first m_j columns of \mathbf{U}_j . Calculate $\bar{\mathbf{H}}_j = \mathbf{A}_j^H \mathbf{H}_j$.

$$\tilde{\mathbf{H}}_j = [\bar{\mathbf{H}}_1^T \quad \dots \quad \bar{\mathbf{H}}_{j-1}^T \quad \bar{\mathbf{H}}_{j+1}^T \quad \dots \quad \bar{\mathbf{H}}_K^T]^T \quad (\text{A.2})$$

- (4) For $j = 1, \dots, K$, compute $\tilde{\mathbf{V}}_j^{(0)}$, the right null space of $\tilde{\mathbf{H}}_j$ as

$$\tilde{\mathbf{H}}_j = \tilde{\mathbf{U}}_j \tilde{\boldsymbol{\Sigma}}_j [\tilde{\mathbf{V}}_j^{(1)} \quad \tilde{\mathbf{V}}_j^{(0)}]^H, \quad (\text{A.3})$$

where $\tilde{\mathbf{V}}_j^{(1)}$ holds the first L_j right singular vectors, $\tilde{\mathbf{V}}_j^{(0)}$ holds the last $N - L_j$ right singular vectors and $L_j = \text{rank}(\mathbf{H}_j)$.

- (5) Compute the SVD

$$\mathbf{H}_j \tilde{\mathbf{V}}_j^{(0)} = \mathbf{U}_j \begin{bmatrix} \boldsymbol{\Sigma}_j & 0 \\ 0 & 0 \end{bmatrix} [\mathbf{V}_j^{(1)} \quad \mathbf{V}_j^{(0)}]^H. \quad (\text{A.4})$$

- (6) The precoding matrix \mathbf{W} for the transmit users with average power allocation is

$$\mathbf{W} = [\tilde{\mathbf{V}}_1^{(0)} \mathbf{V}_1^{(1)} \quad \tilde{\mathbf{V}}_2^{(0)} \mathbf{V}_2^{(1)} \quad \dots \quad \tilde{\mathbf{V}}_K^{(0)} \mathbf{V}_K^{(1)}]. \quad (\text{A.5})$$

Acknowledgments

The project was supported by the National Natural Science Foundation of China (60702073) and the Key Laboratory of Universal Wireless Communications Lab. (Beijing University of Posts and Telecommunications), Ministry of Education, China.

References

- [1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.
- [2] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, 2003.
- [3] Q. H. Spencer, C. B. Peel, A. L. Swindlehurst, and M. Haardt, "An introduction to the multi-user MIMO downlink," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 60–67, 2004.
- [4] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, "Precoding in multiantenna and multiuser communications," *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1305–1316, 2004.
- [5] N. Jindal and A. Goldsmith, "Dirty-paper coding versus TDMA for MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 51, no. 5, pp. 1783–1794, 2005.
- [6] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1912–1921, 2003.
- [7] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2668, 2003.
- [8] E. A. Jorswieck and A. Sezgin, "Impact of spatial correlation on the performance of orthogonal space-time block codes," *IEEE Communications Letters*, vol. 8, no. 1, pp. 21–23, 2004.
- [9] S. N. Diggavi, "On achievable performance of spatial diversity fading channels," *IEEE Transactions on Information Theory*, vol. 47, no. 1, pp. 308–325, 2001.
- [10] P. Tarasak, H. Minn, and V. K. Bhargava, "Linear prediction receiver for differential space-time block codes with spatial correlation," *IEEE Communications Letters*, vol. 7, no. 11, pp. 543–545, 2003.
- [11] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Transactions on Communications*, vol. 50, no. 2, pp. 225–234, 2002.
- [12] M. H. M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, 1983.

- [13] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian MIMO broadcast channel," in *Proceedings of IEEE International Symposium on Information Theory*, p. 174, June 2004.
- [14] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, 2003.
- [15] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electronics Letters*, vol. 7, no. 5-6, pp. 138–139, 1971.
- [16] H. Miyakawa and H. Harashima, "Matched-transmission technique for channels with intersymbol interference," *IEEE Transactions on Communications*, vol. 20, pp. 774–779, 1972.
- [17] V. Stankovic and M. Haardt, "Successive optimization Tomlinson-Harashima precoding (SO THP) for multi-user MIMO systems," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '05)*, vol. 3, pp. 1117–1120, March 2005.
- [18] W. Miao, L. Xiao, Y. Li, S. Zhou, and J. Wang, "Joint stream-wise THP transceiver design for the multiuser MIMO downlink," in *Proceedings of IEEE Wireless Communications & Networking Conference (WCNC '08)*, pp. 330–334, 2008.
- [19] K. Takeda, H. Tomeba, and F. Adachi, "BER performance of joint THP/pre-DFE," in *Proceedings of IEEE Vehicular Technology Conference (VTC '08)*, pp. 1016–1020, 2008.
- [20] P. L. Athanasios, "Tomlinson-Harashima precoding with partial channel knowledge," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 5–9, 2005.
- [21] T. Haustein, C. von Helmolt, E. Jorswieck, V. Jungnickel, and V. Pohl, "Performance of MIMO systems with channel inversion," in *Proceedings of the 55th IEEE Vehicular Technology Conference (VTC '02)*, vol. 1, pp. 35–39, 2002.
- [22] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, 2004.
- [23] L.-U. Choi and R. D. Murch, "A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach," *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 20–24, 2004.
- [24] K.-K. Wong, R. D. Murch, and K. B. Letaief, "A joint-channel diagonalization for multiuser MIMO antenna systems," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 773–786, 2003.
- [25] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication—part I: channel inversion and regularization," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 195–202, 2005.
- [26] M. Stojnic, H. Vikalo, and B. Hassibi, "Rate maximization in multi-antenna broadcast channels with linear preprocessing," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '04)*, vol. 6, pp. 3957–3961, November 2004.
- [27] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proceedings of IEEE International Conference on Neural Networks*, vol. 5, pp. 1942–1948, 1995.
- [28] J. Robinson and Y. Rahmat-Samii, "Particle swarm optimization in electromagnetics," *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 2, pp. 397–407, 2004.