Low Complexity User Selection Algorithms for Multi-user MIMO systems

Submitted by	Mentored by
Gaurav Gupta	Dr. A.K. Chaturvedi
gauravgg@iitk.ac.in	akc@iitk.ac.in
Dept. of EE	Dept. of EE

Indian Institute of Technology Kanpur

Abstract

In this paper we are going to consider two multi-user MIMO (MU-MIMO) models namely, single cell Broadcast (BC) and multi-cell interference model. In MU-MIMO systems the base stations (BS) are required to select a subset of users with the aim of maximizing the sum capacity. The optimal solution is achieved by using Brute-force search method which becomes computationally complex as the total number of users in a cell increases. We propose a conditional entropy based low complexity user selection algorithm for MU-MIMO BC system and finally try to extend the work for interference channels. Simulation results show that the conditional entropy based algorithm offer gains in terms of sum capacity and /or complexity compared to existing user selection algorithms for BC model.

1 Introduction

MIMO systems with multiple transmit and receive antennas have drawn a lot of attention in the past decade. First simulation studies that revealed potentially large capacities were done in 1980s and later the capacity is explored analytically [**?**]. While it is optimal to select a user with the best channel in terms of capacity in single-input single-output (SISO) systems, it was shown that it is optimal to transmit to multiple users simultaneously in multiuser MIMO (MU-MIMO) system [**?**]. Since in BC model the base station is broadcasting to multiple users simultaneously with different data to different users, there is inter-user interference in MU-MIMO systems. There are techniques for inter-user interference cancellation with less complexity, such as Zero-forcing beamforming (ZFBF), Block diagonalization (BD) and successive zero-forcing precoding (SZF) [**?**, **?**, **?**]. We will be using BD scheme which is based on determination of null space of channel space using singular value decomposition (SVD). The data meant to be transmitted to a particular user is multiplied by a precoding matrix which lies in the null space of channel spaces of other users being served simultaneously. Hence the number of users which can be simultaneously supported are limited by the number of transmit and receive antennas due to rank and nullity constraints. This constraint generally leads to degradation in sum capacity. Therefore, the complete cancellation of inter-user interference which is the case with ZFBF and BD is not always feasible. In [**?**] it has been shown that in several cases, for all values of SNR, the sum capacity achieved by SZF is larger than with BD.

The authors of [**?**] proposed a Interference Alignment (IA) technique for MU-MIMO interference channel which aims at maximization of Degrees of Freedom (DoF) by aligning the interference received at each users from other BS such that the interference occupy half of the signal space leaving rest half signal space at the receiver for desired signal. The required signal, then, can be recovered using a either a MMSE or ZF post processing at the receiver.

In a system where the number of users is large, the brute-force determination of the optimal subset of users is prohibitive because of large number of possible subsets and higher computational complexity of SVD. To reduce the computation load many suboptimal algorithms have been proposed for BC model [**?**, **?**, **?**] and [**?**] for IA model.

The work presented in this paper will be concentrated on maximization of sum capacity for both models considered. For BC model, the simulation results provided will show the superiority of the conditional entropy based algorithm in terms of sum capacity and/or complexity. The conditional entropy algorithm will also be extended to the antenna selection problem in BC model.

2 System Model

In this section a brief discussion on the BC model with BD and SZF schemes, the expressions for sum capacity and received signal will be presented. In this paper, the channels are assumed to be Rayleigh fading with complete channel state information of the system at each node.

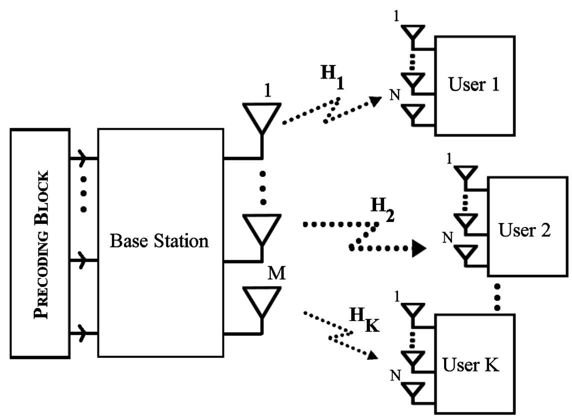

Figure 1: MU-MIMO BC system model

2.1 Broadcast model

As shown in Fig. **??**, the BC model is single celled with BS having M transmit antennas and K_T users each with N receiver antennas. The received signal for the kth user can be split into desired signal, interference from other users and AWGN originating at receiver, which is given by

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{T}_k s_k + \sum_{i=1, i \neq k}^{K} \mathbf{H}_k \mathbf{T}_i s_i + \mathbf{n}_k \tag{1}$$

where $\mathbf{H}_k$ $(k = 1, ..., K_T)$ is an $N \times M$ channel matrix for kth user, $\mathbf{n}_k$ is an $N \times 1$ complex additive white Gaussian noise (AWGN) vector with zero mean and unit variance i.i.d. entries, and $\mathbf{y}_k$ is the $N \times 1$ vector received by kth user. The transmitted vector $\mathbf{x}$ of size $M \times 1$, is given by

$$\mathbf{x} = \sum_{i=1}^{K} \mathbf{T}_i \mathbf{s}_i \tag{2}$$

where K is the number of simultaneous users served by the MU-MIMO system, $\mathbf{s}_i$ is the $t \times 1$ ith data vector preprocessed with $M \times t$ precoding matrix $\mathbf{T}_i$. For BD scheme the interference term is completely eliminated which impose the following condition on the simultaneous supported users by the BS due to non-zero dimension of null-space of effective channel [**?**]

$$M > (K-1) \times N \tag{3}$$

From (**??**) K, the maximum number of simultaneous supported users is given by $\lceil \frac{M}{N} \rceil$, where $\lceil g \rceil$ is the minimum integer number not smaller than g. The capacity of kth user with allocated power P_k is given by water-filling [**?**]

$$\mathcal{R}_k = \sum_{i=1}^{r} \log_2 (1 + P_i \lambda_i) \tag{4}$$

where r is the rank of the channel matrix $\mathbf{H}_k$ and λ_i's $(i \leq i \leq r)$ are the positive eigenvalues of $\mathbf{H}_k \mathbf{H}_k^H$, and P_i is the power allocated to each eigenmode of channel obtained using water-filling with power constraint P_k.

Unlike BD, the SZF precoding leads to partial inter-user interference cancellation. The motivation behind considering the SZF scheme is that its achievable sum capacity is strictly larger than that of BD for all SNR regimes [**?**]. With SZF transmission scheme, the received signal at the kth receiver is given by (**??**) as in [**?**]

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{T}_k \mathbf{s}_k + \mathbf{H}_k \sum_{i<k} \mathbf{T}_i \mathbf{s}_i + \mathbf{n}_k \tag{5}$$

With the assumption that $\mathbf{s}_i$ are Gaussian and independent [**?**], it follows that the interference term $\sum_{i<K} \mathbf{T}_i \mathbf{s}_i$ is also Gaussian. The maximum achievable rate of each user is given by

$$\mathcal{R}_k = \log_2 \frac{\det \left(\mathbf{I} + \mathbf{H}_k (\sum_{i=1}^{k} \bar{\mathbf{V}}_0^{i-1} \mathbf{Q}_i (\bar{\mathbf{V}}_0^{i-1})^H) \mathbf{H}_k^H \right)}{\det \left(\mathbf{I} + \mathbf{H}_k (\sum_{i=1}^{k-1} \bar{\mathbf{V}}_0^{i-1} \mathbf{Q}_i (\bar{\mathbf{V}}_0^{i-1})^H) \mathbf{H}_k^H \right)} \tag{6}$$

where $\bar{\mathbf{V}}_0^{i-1}$ is a matrix, columns of which spans the null space of $\bar{\mathbf{H}}^{i-1}$ and the covariance matrices $\mathbf{Q}_i$ of precoded user signals are defined such that $\mathbf{T}_i\mathbf{T}_i^H = \bar{\mathbf{V}}_0^{i-1}\mathbf{Q}_i(\bar{\mathbf{V}}_0^{i-1})^H$.

Finally, the achievable sum capacity of both the schemes for given user set $\mathcal{S}$ is given by

$$\mathcal{R}\left(\mathcal{S}\right) = \max_{\{\mathbf{Q}_k\}_{k\in\mathcal{S}}:\mathbf{Q}_k\geq 0\,\sum_i \mathrm{Tr}(\mathbf{Q}_i)\leq P} \sum_{k\in\mathcal{S}} \mathcal{R}_k \tag{7}$$

Before we propose the conditional entropy based scheduling algorithm, let us briefly discuss the notion of conditional entropy.

3 Conditional Entropy

In this section, we will discuss conditional entropy and its aspects for multivariate random variables.

Let us consider a SU-MIMO system for which we can write

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{8}$$

Here we consider the transmit vector $\mathbf{x}$ to be zero mean with covariance $E[\mathbf{x}\mathbf{x}^H] = \mathbf{Q}$ and satisfying power constraint $E[\mathbf{x}^H\mathbf{x}] \leq P$. Hence $\mathbf{y}$ be zero mean with covariance $E[\mathbf{y}\mathbf{y}^H] = \mathbf{H}\mathbf{Q}\mathbf{H}^H + \mathbf{I}_N$. For $\mathbf{y}$ and $\mathbf{x}$ we can write

$$\begin{aligned}\mathcal{I}\left(\mathbf{y};\mathbf{x}\right) &= \mathcal{H}(\mathbf{y}) - \mathcal{H}(\mathbf{y}|\mathbf{x}) \\ &= \mathcal{H}(\mathbf{y}) - \mathcal{H}(\mathbf{n})\end{aligned} \tag{9}$$

where $\mathcal{H}$ represents differential entropy and $\mathcal{I}$ represents mutual information. So in order to maximize $\mathcal{I}\left(\mathbf{y};\mathbf{x}\right)$ we have to maximize $\mathcal{H}(\mathbf{y})$. Given covariance of $\mathbf{y}$, the distribution which maximizes the entropy of $\mathbf{y}$ is circular symmetric complex Gaussian. This is only possible when $\mathbf{x}$ is also circular symmetric complex Gaussian random variable [**?**], then we can write

$$\mathcal{H}(\mathbf{y}) = \log_2\det\left(\pi e\left(\mathbf{H}\mathbf{Q}\mathbf{H}^H + \mathbf{I}_N\right)\right) \tag{10}$$

Now consider a set of n $\mathbf{y}$'s, $\mathbf{y}_i\ (1 \leq i \leq n)$ and $\tilde{\mathbf{y}} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \cdots, \mathbf{y}_n^T]^T$, $\tilde{\mathbf{y}}$ will also be a circular symmetric complex Gaussian random variable with zero mean and $E[\tilde{\mathbf{y}}\tilde{\mathbf{y}}^H] = \mathbf{\Sigma}$, such that $\mathbf{\Sigma} = [\mathbf{\Sigma}_{ij}]$ where

$$\mathbf{\Sigma}_{ij} = E[\mathbf{y}_i\mathbf{y}_j^H] = \begin{cases}\mathbf{H}_i\mathbf{Q}\mathbf{H}_i^H + \mathbf{I}_N & \text{if } i = j \\ \mathbf{H}_i\mathbf{Q}\mathbf{H}_j^H & \text{if } i \neq j\end{cases} \tag{11}$$

The entropy of $\tilde{\mathbf{y}}$ will be $\mathcal{H}(\tilde{\mathbf{y}}) = \log_2\det\left(\pi e\mathbf{\Sigma}\right)$ and by using (**??**) to write $\mathbf{\Sigma}$ in decomposed form, we have

$$\begin{aligned}\mathcal{H}(\tilde{\mathbf{y}}) &= \log_2\det\left(\mathbf{I}_{nN} + \begin{bmatrix}\mathbf{H}_1 \\ \vdots \\ \mathbf{H}_n\end{bmatrix}\mathbf{Q}\begin{bmatrix}\mathbf{H}_1^H & \cdots & \mathbf{H}_n^H\end{bmatrix}\right) + \log_2\left(\pi e\right)^{nN} \\ &= \log_2\det\left(\mathbf{I}_M + \mathbf{Q}\mathbf{H}_1^H\mathbf{H}_1 + \cdots + \mathbf{Q}\mathbf{H}_n^H\mathbf{H}_n\right) + \log_2\left(\pi e\right)^{nN}\end{aligned} \tag{12}$$

where (**??**) has been written using matrix determinant identity

$$\det\left(\mathbf{I}_M + \mathbf{A}\mathbf{B}\right) = \det\left(\mathbf{I}_N + \mathbf{B}\mathbf{A}\right) \tag{13}$$

where $\mathbf{A}$ and $\mathbf{B}$ are $M \times N$ and $N \times M$ matrices, respectively. Using chain rule [**?**] we can generalize the conditional entropy to n random variables as

$$\mathcal{H}\left(\mathbf{y}_i|\tilde{\mathbf{y}}_i\right) = \mathcal{H}\left(\tilde{\mathbf{y}}\right) - \mathcal{H}\left(\tilde{\mathbf{y}}_i\right) \tag{14}$$

where $\tilde{\mathbf{y}}_i = [\mathbf{y}_1^T, \cdots, \mathbf{y}_{i-1}^T, \mathbf{y}_{i+1}^T, \cdots, \mathbf{y}_n^T]^T$. From (**??**) we can now write the sum conditional entropy for the general case, which is given by

$$\mathcal{H}_{SC} = \sum_{i=1}^{n}\mathcal{H}\left(\mathbf{y}_i|\tilde{\mathbf{y}}_i\right) \tag{15}$$

Table 1: Conditional Entropy based Multiuser MIMO scheduling Algorithm

1) Initially, let $\Gamma = \{1, 2, ..., K_T\}$ and $\mathcal{S} = \emptyset$. Let $\mathbf{\Omega} = \frac{P}{M}\mathbf{I}_M$, $\mathbf{\Omega}_n = \mathbf{\Omega}$. Let $\mathcal{R}_{temp} = 0$;
2) for $i = 1 : K$
   for $j = 1 : i - 1$
      For each $k \in \Gamma$, $\mathbf{\Omega}_{s_j,k} = \mathbf{\Omega}_{s_j} - \mathbf{\Omega}_{s_j}\mathbf{H}_k^H \left(\mathbf{I}_N + \mathbf{H}_k\mathbf{\Omega}_{s_j}\mathbf{H}_k^H\right)^{-1}\mathbf{H}_k\mathbf{\Omega}_{s_j}$;
   end-for
   $p = \arg\max_{k\in\Gamma}\left\{\log_2\det\left(\mathbf{I}_N + \mathbf{H}_k\mathbf{\Omega}_n\mathbf{H}_k^H\right) + \sum_{s\in\mathcal{S}}\log_2\det\left(\mathbf{I}_N + \mathbf{H}_s\mathbf{\Omega}_{s,k}\mathbf{H}_s^H\right)\right\}$;
   $\mathcal{R} = \mathcal{R}\left(\mathcal{S}_{temp}\right); \mathcal{S}_{temp} = \mathcal{S} + \{p\}$;
   if $\mathcal{R} < \mathcal{R}_{temp}$
      break;
   else
      $\mathcal{S} = \mathcal{S}_{temp}; \Gamma = \Gamma - \{p\}; \mathcal{R}_{temp} = \mathcal{R}$;
   end-if
   $\mathbf{\Omega}_{s_i} = \mathbf{\Omega}_n$;
   $\mathbf{\Omega}_n = \mathbf{\Omega}_n - \mathbf{\Omega}_n\mathbf{H}_p^H\left(\mathbf{I}_N + \mathbf{H}_p\mathbf{\Omega}_n\mathbf{H}_p^H\right)^{-1}\mathbf{H}_p\mathbf{\Omega}_n$;
   end-for

4 Conditional Entropy based Scheduling Algorithm

We want the mutual information between the users which are supported simultaneously to be as small as possible, and their entropy to be as large as possible, because the mutual information gives an indication of the information which is being shared between them, which we want minimum and the entropy of the received signal to be large as it indicates the information between the desired user and the transmitter (**??**). Therefore, we will maximize the sum conditional entropy (**??**)

Suppose that after kth user selection step, the user set $\mathcal{S} = \{s_1, ..., s_k\}$ has been selected. After adding the $(k+1)$th user with channel matrix $\mathbf{H}_t$ where $s_{k+1} = t \notin \mathcal{S}$, the sum conditional entropy in (**??**) can be written as

$$\mathcal{H}_{SC}\left(\mathcal{S}, t\right) = \mathcal{H}\left(\begin{bmatrix}\mathbf{H}(\mathcal{S})\\ \mathbf{H}_t\end{bmatrix}\right) - \mathcal{H}\left(\mathcal{S}\right) + \sum_{i=1}^{k}\left(\mathcal{H}\left(\begin{bmatrix}\mathbf{H}(\mathcal{S})\\ \mathbf{H}_t\end{bmatrix}\right) - \mathcal{H}\left(\mathcal{S}_i\right)\right) \tag{16}$$

where $\mathcal{S}_i = \mathcal{S} + \{t\} - \{s_i\}$. Also the joint entropy of user set $\mathcal{S}$ and user t can be written as

$$\mathcal{H}\left(\begin{bmatrix}\mathbf{H}(\mathcal{S})\\ \mathbf{H}_t\end{bmatrix}\right) = \log_2\det\left(\mathbf{I}_M + \frac{P}{M}\mathbf{H}(\mathcal{S})^H\mathbf{H}(\mathcal{S})\right) + \log_2\det\left(\mathbf{I}_N + \mathbf{H}_t\left(\frac{M}{P}\mathbf{I}_M + \mathbf{H}(\mathcal{S})^H\mathbf{H}(\mathcal{S})\right)^{-1}\mathbf{H}_t^H\right) \tag{17}$$

where we have used (**??**) and have ignored the (πe) term because it is a constant and hence will not play any role in maximizing entropy. Therefore $\mathcal{H}_{SC}\left(\mathcal{S}, t\right)$ becomes

$$\begin{aligned}\mathcal{H}_{SC}(\mathcal{S}, t) &= \log_2\det\left(\mathbf{I}_N + \mathbf{H}_t\left(\frac{M}{P}\mathbf{I}_M + \mathbf{H}(\mathcal{S})^H\mathbf{H}(\mathcal{S})\right)^{-1}\mathbf{H}_t^H\right)\\ &\quad + \sum_{i=1}^{k}\log_2\det\left(\mathbf{I}_N + \mathbf{H}_{s_i}\left(\frac{M}{P}\mathbf{I}_M + \mathbf{H}(\mathcal{S}_i)^H\mathbf{H}(\mathcal{S}_i)\right)^{-1}\mathbf{H}_{s_i}^H\right)\end{aligned} \tag{18}$$

In order to maximize the sum conditional entropy, we will select the next user as $s_{k+1} = \arg\max_{t\notin S}\mathcal{H}_{SC}(\mathcal{S}, t)$. For this a series of matrix inversions are to be computed at each user selection step. This can be done through matrix recursion update formula [**?**]. With $\mathbf{A}$ as an $M \times N$ positive definite matrix and $\mathbf{B}$ as an $N \times M$ matrix,

$$\left(\mathbf{A} + \mathbf{B}^H\mathbf{B}\right)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}^H\left(\mathbf{I}_N + \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^H\right)^{-1}\mathbf{B}\mathbf{A}^{-1} \tag{19}$$

Now we will derive the recursion for the first term of (**??**) and the same arguments will be applied for the terms in the summation of the same equation. Let us define $\mathbf{\Omega}_k$ by

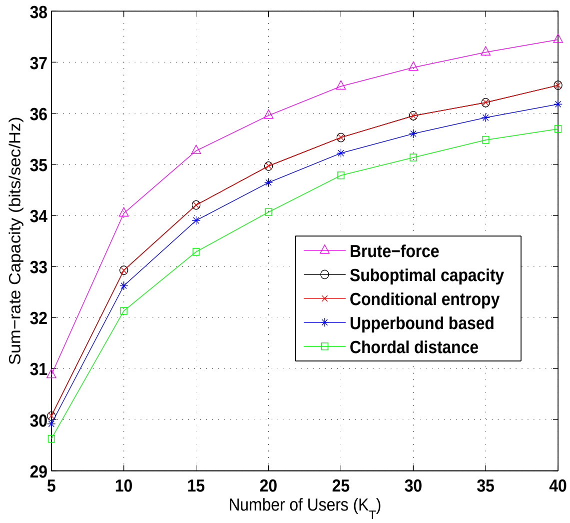

Figure 2: Sum capacity versus total number of users when $M = 6, N = 2, K = 3$ for SNR = 20 dB.

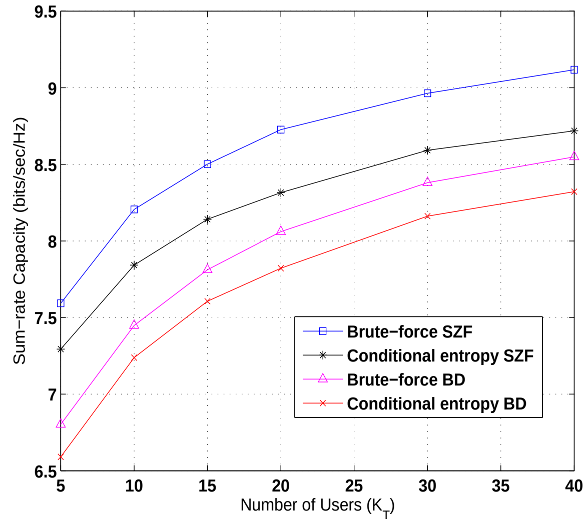

Figure 3: Sum capacity versus total number of users when $M = 4, N = 2, K = 2$ for SNR = 5 dB.

$$\mathbf{\Omega}_k = \left(\frac{M}{P}\mathbf{I}_M + \mathbf{H}(\mathcal{S})^H\mathbf{H}(\mathcal{S})\right)^{-1} \tag{20}$$

We can write the effective channel at the $(k+1)$th step as

$$\mathbf{H}_{\text{eff}} = \begin{bmatrix} \mathbf{H}(\mathcal{S}) \\ \mathbf{H}_t \end{bmatrix} \tag{21}$$

We now update (**??**) by replacing $\mathbf{H}(\mathcal{S})$ with $\mathbf{H}_{\text{eff}}$ from (**??**) to obtain $\mathbf{\Omega}_{k+1}$ as

$$\mathbf{\Omega}_{k+1} = \left(\mathbf{\Omega}_k^{-1} + \mathbf{H}_t^H\mathbf{H}_t\right)^{-1} \tag{22}$$

On substituting $\mathbf{A}$ for $\mathbf{\Omega}_k^{-1}$ and $\mathbf{B}$ for $\mathbf{H}_t$ in (**??**), the recursion is then given by

$$\mathbf{\Omega}_{k+1} = \mathbf{\Omega}_k - \mathbf{\Omega}_k\mathbf{H}_t^H\left(\mathbf{I}_N + \mathbf{H}_t\mathbf{\Omega}_k\mathbf{H}_t^H\right)^{-1}\mathbf{H}_t\mathbf{\Omega}_k \tag{23}$$

The resulting conditional entropy based scheduling algorithm is described in Table **??**. In Step 1, the algorithm is initialized while in Step 2, the algorithm first selects the user with maximum entropy and in successive steps it selects the user which maximizes the sum conditional entropy till the maximum number of simultaneously supportable users limit is reached. The proposed algorithm differs for BD and SZF only in the evaluation of $\mathcal{R}$, similarly it will work for all transmission schemes with their corresponding way of evaluation of $\mathcal{R}$.

5 Simulation Results

Fig. **??**, shows the sum capacity versus the total number of users (K_T) for $(M,N) = (6,2)$ at SNR $= 20$ dB for BD scheme. We can see that the sum capacity of conditional entropy based algorithm is better than upperbound [**?**] based algorithm and chordal distance [**?**] based algorithm.

In Fig. **??**, simulation results are provided for $(M,N) = (4,2)$ at SNR $= 5$ dB for SZF and BD schemes. The sum capacity achieved by the proposed algorithm with SZF is close to the optimal solution for SZF. It can be observed that, as expected, the sum rate achieved by conditional entropy based algorithm for SZF is strictly better than that for BD. At low SNR's, for e.g. 5 dB, there is a significant difference between the sum capacity achieved by the proposed algorithm for SZF and BD. The reason why SZF performs better than BD is that it can handle more interference than BD, hence it can support more users even in the low SNR regime.

In Fig. **??**, the average CPU run-time versus K_T of all these algorithms for $(M,N) = (12,6)$ is shown. It can be observed that the suboptimal capacity based algorithm [**?**] takes the longest time as K_T increases. The chordal distance based algorithm takes more time than the conditional entropy based algorithm. Further, it should be noted that even though the sum capacity plots of the suboptimal capacity based algorithm and the conditional entropy based algorithm overlap, run time of the conditional entropy based algorithm is significantly lower. It may be noted that the lower run-time of upperbound based algorithm comes

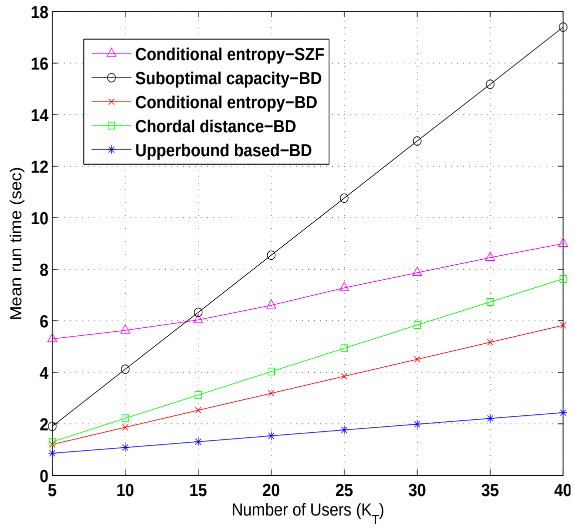

Figure 4: Mean run-time versus total number of users when $M = 12, N = 6, K = 2$

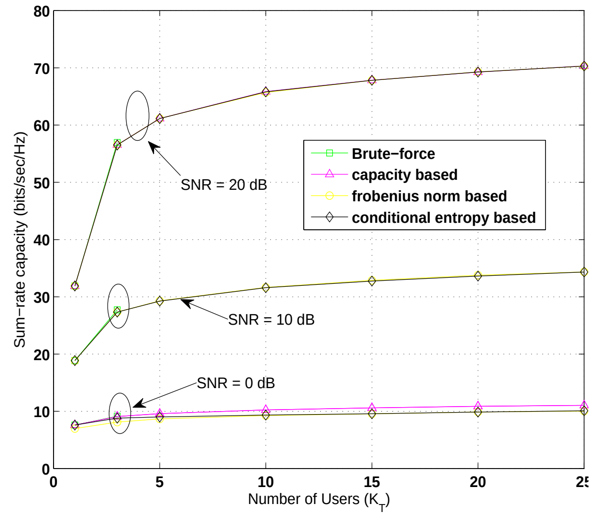

Figure 5: Sum capacity versus total number of users when $M = 12, N = 4$.

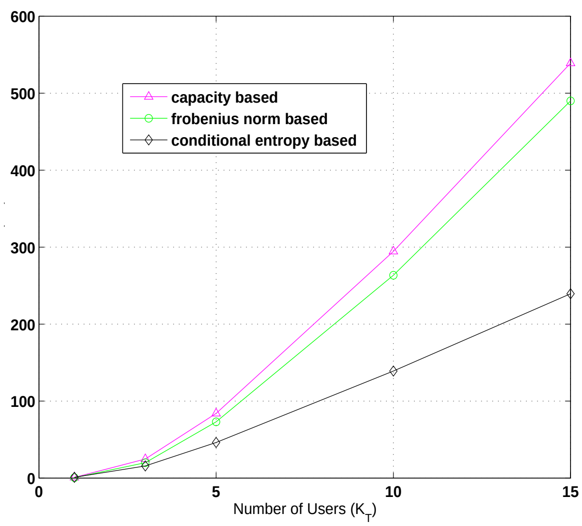

Figure 6: Mean run time versus total number of users when $M = 12, N = 3$ for SNR = 20 dB

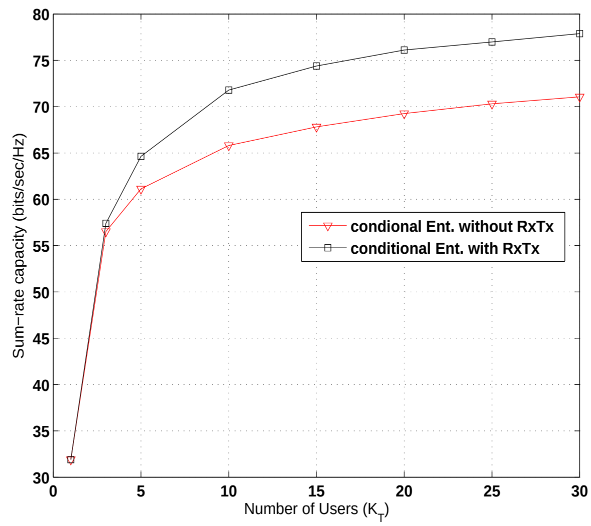

Figure 7: Sum capacity versus total number of users when $M = 12, N = 3$ for SNR = 20 dB

at the cost of its lower sum capacity. Lastly, since the rate of increase of run time with K_T is constant for all algorithms, the time gap increases with K_T.

6 Joint User and Antenna Selection

Exploiting the multimode diversity can significantly increase the sum capacity of the system. Since in fixed mode transmission all users will get equal number of streams which is not optimal, as varying the user streams can help in relaxing the total number of users to be supported simultaneously (**??**). It was shown in [**?**] using a receiver combining matrix increases the sum capacity of the system. However, computation of such optimal matrices requires frequent use of SVD which makes the method very computationally expensive. It was shown in [**?**] that antenna selection can be viewed as a special case of mode selection with receiver equalizer matrices as antenna selection matrices. The authors of [**?**] proposed two antenna and user selection algorithms namely frobenius norm based and capacity based algorithm. We will use our conditional entropy metric to select antennas, but before that let us introduce the system model and capacity equations for joint antenna and user selection model.

6.1 System Model

The system model used is MU-MIMO BC model but with one additional $L_k \times N$ matrix called receiver equalizer $\mathbf{R}_k^H$, where L_k denotes the number of transmission modes for user k. The received signal at the kth receiver is given by

$$\mathbf{y}_k = \mathbf{R}_k^H \mathbf{H}_k \mathbf{T}_k \mathbf{x}_k + \mathbf{R}_k^H \mathbf{H}_k \sum_{j=1, j\neq k}^{K} \mathbf{T}_j \mathbf{x}_j + \mathbf{R}_k^H \mathbf{n}_k \tag{24}$$

The principal idea of the BD is to design precoding matrices $\mathbf{T}_k$ such that the interference received is equal to zero. To ensure that the null space is not empty the necessary condition is $M \geq \sum_{j=1}^{K} L_j$ [**?**]. In capacity optimization problem, the number of modes L_k at kth receiver is selected such that the sum capacity is maximized as

$$C_{\text{max}} = \max_{L_k, \mathbf{T}_k, \mathbf{R}_k, \mathbf{Q}_k} \sum_{k=1}^{K} \log_2 \left| \mathbf{I} + \frac{1}{\sigma_n^2} \mathbf{R}_k^H \mathbf{H}_k \mathbf{T}_k \mathbf{Q}_k \mathbf{T}_k^H \mathbf{H}_k^H \mathbf{R}_k \right| \tag{25}$$

where $\sum_{k=1}^{K} L_k \leq M, 0 \leq L_k \leq N, \forall k$ and $\sum_{k=1}^{K} \text{trace}\left(\mathbf{Q}_k\right) \leq P$ such that K is the number of users selected. Now, the conditional entropy based algorithm is used to select antenna one by one till the final constraint on the antennas stated above is reached. The formulation of the algorithm remains same except here instead of users, the antennas are being selected at the kth step. Let us consider K users are selected with $\mathcal{A}_k$ and $\mathcal{S}_k$ denotes the set of index of unselected and selected antennas for MS k, where $\mathcal{A}_k$ and $\mathcal{S}_k$ are subsets of $\{1, 2, ..., N\}$. Let L_k = card$(\mathcal{S}_k)$ denote the number of selected antennas for MS k. Let $\mathcal{G}$ denote the set of active users with $L_k \geq 1$. The algorithm can be modelled in steps as

---

1. Stage $i = 0$: Set all antennas of all receivers inactive by making $L_1 = \ldots = L_K = 0, \mathcal{G} = \phi, \mathcal{S}_1 = \cdots = \mathcal{S}_K = \phi$, and $\mathcal{A}_1 = \cdots = \mathcal{A}_K = \{1, 2, ..., N\}$.
2. Stage $i = 1$: choose the best antenna $\bar{j}$ of the best user $\bar{k}$ with the largest frobenius norm

$$\left(\bar{k}, \bar{j}\right) = \arg \max_{k=1,\ldots,K; \quad j=1,\ldots,N} \|\mathbf{h}_{k,j}\|_F^2$$

   where $\mathbf{h}_{k,j}$ denotes the jth row of $\mathbf{H}_k$. Activate antenna $\bar{j}$ of the user $\bar{k}$. Let $H_{\text{temp}} = \|\mathbf{h}_{\bar{k},\bar{j}}\|_F^2$.
3. Stage $i = i + 1, i \leq M$.
   (a) For every unselected antenna j of every user k, activate it temporarily and calculate the sum conditional entropy as in (**??**).
   (b) Find the best antenna $\bar{j}$ and the best user $\bar{k}$

$$\left(\bar{k}, \bar{j}\right) = \arg \max_{k=1,\ldots,K; \quad j=1,\ldots,N} \mathcal{H}_{SC,k,j}$$

   (c) If $H_{\text{temp}} \leq \mathcal{H}_{SC,\bar{k},\bar{j}}$, activate the antenna $\bar{j}$ of user $\bar{k}$ and make $H_{\text{temp}} = \mathcal{H}_{SC,\bar{k},\bar{j}}$. Return to (3). Else, quit.

The simulation results for the antenna selection algorithm are provided in Fig. **??**. It is clear that the sum capacity achieved by the frobenius norm based, capacity based and the conditional entropy based algorithms are overlapping. However, from Fig. **??** we can see the differences in the run-time of these algorithms and it could be observed that the conditional entropy based algorithm which is giving the same performance as of the other two algorithms is taking the least run-time.

The performance of the conditional entropy algorithm can further be enhanced if the receiver equalization technique proposed in [**?**] is used after selecting the user-set $\mathcal{G}$. The dimensions required for the computation of $\mathbf{R}_k$ will be the cardinality of the set $\mathcal{S}_k$. The computation complexity of the algorithm will increase obviously, but since the computation of $\mathbf{R}_k$ will be done only after selecting the user-set, therefore as the number of users increases in the cell, the order of the complexity will not increase due to the computation of $\mathbf{R}_k$. Fig. **??** shows the comparison of the sum capacity versus total number of users (K_T) with and without receiver processing for the conditional entropy based algorithm.

7 Conclusion

In this paper, we have proposed a suboptimal conditional entropy based user selection algorithm. Simulation results show that the algorithm is offering significant gains in sum-capacity and/or complexity for MU-MIMO BC and SZF scheme. The proposed algorithm with suitable modification is also offering reduced complexity search in joint antenna and user selection problem as compared to the existing algorithms with sum capacity greater than or equal to them.

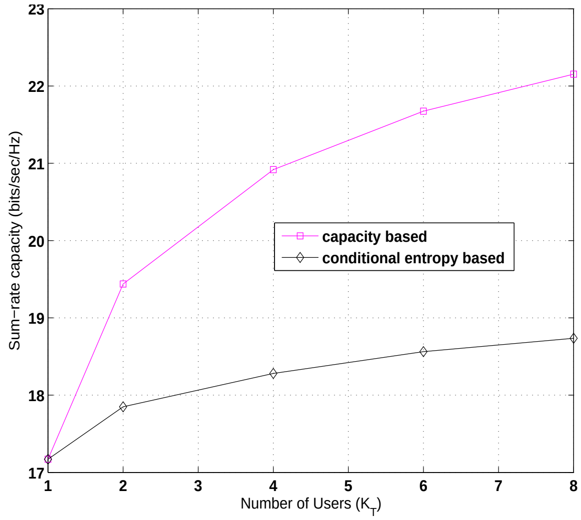

Figure 8: Sum capacity versus total number of users when $M = 4, N = 4$ for SNR $= 10$ dB

8 Multi-cell Interference model

The authors of [?] have proposed a IA scheme in which the interference from the other transmitters is aligned perfectly and span half the dimension of the signal space leaving other half for desired signal. It was shown in [?] that DoF achieved in a multi-user cell by the IA scheme is independent of the number of users in each cell. Now this result gives an opportunity to use scheduling in each cell to exploit multi-user diversity to increase the achievable sum capacity and since the DoF achieved will not change, we will have DoF as well as diversity gains. There are no existing IA techniques for greater than three cells, and since our focus in on scheduling method we will discussing scheduling for 3-cell interference model only. The authors of [?] have proposed a suboptimal algorithm which selects users by computing sum-capacity at each step using a coordinate ascent approach. Since the computation of sum capacity involves computation of pre-coding matrices and post-processing matrices at receiver which could either be MMSE or ZF, their computation cost turns out to be large. If we consider the optimal pre-post processing proposed in [?], they involves frequent use of SVD which is computationally heavy. So the conditional entropy algorithm with the following formulation is employed to select the users. In the algorithm, the notation $\mathcal{S}_{i,j}$ denotes that the ith user in the ordered set $\mathcal{S} = \{k_1, k_2, k_3\}$ is replaced by the user j. For example $\mathcal{S}_{1,j} = \{j, k_2, k_3\}$.

The metric which will be used in the following algorithm is given by

$$\mathcal{H}^{i}_{SC,k} = \mathcal{H}\left(\begin{bmatrix}\mathbf{H}_k^{(i,i)^H} & \mathbf{H}_k^{(i,j_1)^H} & \mathbf{H}_k^{(i,j_2)^H}\end{bmatrix}^H\right) - \mathcal{H}\left(\begin{bmatrix}\mathbf{H}_k^{(i,j_1)^H} & \mathbf{H}_k^{(i,j_2)^H}\end{bmatrix}^H\right) \tag{26}$$

where $j_1, j_2 \in \{1,2,3\} - i$ and the channels notation are same as [?].

---

1. Select user k_i^0 such that
   $k_i^0 = \arg \max_{k \in \{1,\ldots,K_i\}} \|\mathbf{H}_k^{(i,i)}\|_F^2 \quad$ for $i = 1,2,3$.

2. Initialize a user set $\mathcal{S} = \left(k_1^0, k_2^0, k_3^0\right)$ and $i = 1$.

3. for $i = 1 : 3$
   Find a user k_i^* such that
   $k_i^* = \arg \max_{k \in \{1,\ldots,K_i\}} \mathcal{H}^i_{SC,k}$
   Update $\mathcal{S} = \mathcal{S}_{i,k_i^*}$.

The simulation result for the above formulation of the conditional entropy metric is shown in the Fig. **??**. It can be seen that the sum capacity resulting from the conditional entropy based algorithm is lower than the sum capacity achieved by the capacity based algorithm. The reason can be that, the algorithm while selecting the user in the ith cell takes the conditional entropy of the user in that cell only, and we cannot take sum conditional entropy here like in BC model because the conditional entropy for the selected users in other cell will be constant on varying user in ith cell. Also changing interfering channels will change the precoding matrices of other two BS also [?]. This will effect their capacity which is not taken into effect by the present formulation. So a new metric has to be designed which could take care of effects induced on the users selected in other cells.

9 Future Work

As we have seen that the conditional entropy algorithm which maximizes the conditional entropy of user channel given interfering channels is not able to capture the effects of IA scheme. Therefore, there is a need for either modifying the existing metric or finding a new metric such that it would be able to quantize the effects of changing interference channels of user in the ith cell to the selected users in the remaining two cells.

References

[1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[2] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1691–1706, Jul. 2003.

[3] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, 2004.

[4] L. Choi and R. D. Murch, "A transmit preprocessing technique for multiuser MIMO system using a decomposition approach," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 20–24, Jan. 2004.

[5] A. D. Dabbagh and D. J. Love, "Precoding for multiple antenna gaussian broadcast channels with successive zero-forcing," *IEEE Trans. Signal Process.*, vol. 55, no. 7, p. 38373850, Jul. 2007.

[6] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the k-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3425–3441, Aug. 2008.

[7] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, and B. L. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3658–3663, Sep. 2006.

[8] X. Zhang and J. Lee, "Low complexity MIMO scheduling with channel decomposition using capacity upperbound," *IEEE Trans. Commun.*, vol. 56, no. 6, pp. 871–876, Jun. 2008.

[9] K. Ko and J. Lee, "Multiuser MIMO user selection based on chordal distance," *IEEE Trans. Commun.*, vol. 60, no. 3, pp. 649–654, Mar. 2012.

[10] H. Park, S.-H. Park, H. Sung, and I. Lee, "Scheduling methods with mimo interference alignment for mutually interfering broadcast channels," in *Proc. IEEE Globecom*, Dec. 2010, pp. 1–5.

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[12] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*. New York: Wiley, 1977.

[13] Z. Pan, K.-K. Wong, and T.-S. Ng, "Generalized multiuser orthogonal space-division multiplexing," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1969–1973, Nov. 2004.

[14] R. Chen, Z. Shen, J. G. Andrews, and R. W. Heath, "Multimode transmission for multiuser mimo systems with block diagonalization," *IEEE Trans. Signal Process.*, vol. 56, no. 7, p. 32943302, Jul. 2008.

[15] H. Sung, S.-H. Park, K.-J. Lee, and I. Lee, "Linear precoder designs for k-user interference channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, p. 291301, Jan. 2010.