Impact of Receive Antenna Selection on Scheduling for Orthogonal Space Division Multiplexing

Boon Chin Lim, Witold A. Krzymien*, Christian Schlegel

Department of Electrical & Computer Engineering, University of Alberta
Edmonton, Alberta, Canada

{bcylim, wak, schlegel}@ece.ualberta.ca

*) also with TRLabs, Edmonton, Alberta, Canada

Abstract—We demonstrate that judicious receive antenna selection (RAS) can be used to create room for scheduling more users while minimizing the rate loss of existing users for multi-user MIMO wireless downlinks that employ block diagonalization (BD) to achieve orthogonal space division multiplexing (OSDM), when each user terminal has multiple antennas. When coupled with proper user selection, the overall sum rate may also be increased. In general, the method is applicable to both iterative and non-iterative zero-forcing approaches that use null-space projection. Strategies for the inclusion of RAS together with user selection algorithms in BD-OSDM to help increase the number of scheduled users are presented.

Key words: Multi-user MIMO, user scheduling, antenna selection, downlink beamforming.

1. Introduction

For the downlink of a wireless base station equipped with multiple antennas where coordination is feasible among the transmit chains but not among the mobile user terminals, simultaneous transmissions to multiple users are possible when channel state information is available at the transmitter. The optimum scheme that can achieve downlink sum capacity is dirty paper coding, which has very high computational complexity. A reduced complexity sub-optimal alternative is beamforming, which has been shown to achieve a large fraction of the capacity in multi-user systems when the number of users is large. However, the determination of its optimal weight vectors is still a tedious non-convex optimization problem. A sub-optimal beamforming technique is zero-forcing beamforming where the weight vectors are chosen to enforce zero co-channel interference (CCI) among all users. Transmit zero-forcing beamforming can be implemented so that data for each user is encoded individually and spatial multiplexing is achieved. When each user terminal has multiple antennas however, creating parallel channels with zero CCI at the same terminal is sub-optimal since each terminal is able to coordinate the processing of its receivers. It would therefore be better to impose orthogonality between users only and not between intra-terminal antennas. This is commonly referred to as block diagonalization (BD) and examples of its use to achieve orthogonal space division multiplexing (OSDM) are found in [1] – [3]. For \( K \) users, each with \( M_{R_j} \) receive antennas \((j = 1,\ldots,K)\), the BD processing constraints \([1], [3]\) impose \( M_R \leq M_T \), where \( M_R = \sum_{j=1}^{K} M_{R_j} \) and \( M_T \) are the total number of receive and transmit antennas respectively. However, the pool of potential users is normally large, which implies \( M_R \gg M_T \) and this necessitates scheduling with a subset of \( K' \) users being chosen so that \( (M_R = \sum_{j=1}^{K} M_{R_j}) \leq M_T \). Given the linear tradeoff between \( K' \) and \( M_{B_j} \), a straightforward way of increasing \( K' \) is to de-activate some antennas at each user, as long as the constraint \( (M'_{R} = \sum_{j=1}^{K} M_{B_j}) \leq M_T \) is met, where \( M_{B_j} \) is the number of antennas remaining at each user \( j \). When done arbitrarily, this method of increasing \( K' \) may result in excessive rate reduction for the users that were originally scheduled.

Other ways of increasing \( K' \) are possible with the coordinated transmit-receive (CTR) scheme described in [1] or the iterative null-space directed SVD (Nu-SVD) scheme described in [2]. Both schemes allow the activation of different numbers of spatial modes per user, subject to a simultaneous total of \( M_T \) spatial modes. In this way, a maximum of \( K' = M_T \) users can be served. However, when channel conditions are poor or when QoS requirements are demanding, power constraints may not allow the activation of all \( M_T \) spatial modes while meeting the QoS requirement of each user.

In this paper, we demonstrate that the rate loss for currently scheduled users can be minimized for BD-OSDM schemes via judicious receive antenna selection (RAS) when increasing the number of scheduled terminals via receive antenna de-activation. In some cases, the transmit power for a user may also be reduced when the RAS process resulted in rate gain for that user. Next, for schemes like CTR and Nu-SVD, RAS helps improve the spatial mode gains, which then allows a system to approach the activation of \( M_T \) spatial modes and hence the scheduling of \( K' = M_T \) users. In general, this approach is useful for both iterative and non-iterative zero-forcing methods that use null-space projection techniques. Strategies for the inclusion of RAS together with user selection algorithms in BD-OSDM to increase the number of users served simultaneously will be discussed. The rest of the paper is
organized as follows. In Section 2, the system model is presented. In Section 3, the theory and algorithms for applying RAS in BD-OSDM to increase the number of scheduled users are developed. In Section 4, the numerical results are presented and Section 5 contains conclusions.

2. System Model

We focus on the multi-user MIMO downlink of a base station (BS) serving a group of $K$ geographically distributed users via spatial multiplexing that is achieved using linear pre- and post-processing at the transmitter and receivers. The BS has $M_T$ transmit-chains and antennas while each user $j$ has $M_R$ multiple antennas, each coupled with a receive-chain. The total number of receive antennas is $M_R = \sum_{j=1}^{K} M_R$ and BD-OSDM constraints require $M_T \geq M_R$. The overall ($M_R \times M_T$) channel matrix is $H$ while each user’s ($M_R \times M_T$) channel sub-matrix is denoted as $H_j$. Each data vector $d_j$ of arbitrary dimension ($m_j \times 1$) is pre-coded by a ($M_T \times m_j$) matrix $T_j$ to result in a ($M_T \times 1$) transmission vector $s_j = T_j d_j$. The overall ($M_T \times 1$) transmission vector is $s = \sum_{j=1}^{K} s_j T_j d_j$ and the received ($M_R \times 1$) signal vector $y_j$ at user $j$ is shown in (1) while the overall ($\sum_{j=1}^{K} M_R \times 1$) received vector $y$ is shown in (2)

$$y_j = \sqrt{E_s} / M_T H_j \sum_{i=1}^{K} T_j H_j d_i + n_j,$$

$$y = \sqrt{E_s} / M_T H d + n,$$

where $H = [H_1 \ T \ H_2 \ T \ ... \ H_K \ T]$, $T = [T_1 \ T_2 \ ... \ T_K]$, $d = [d_1 \ d_2 \ ... \ d_K]$ and $n = [n_1 \ n_2 \ ... \ n_K]$; $[.]^T$ is matrix transposition. This paper assumes a quasi-static, flat fading Rayleigh channel that is constant over several transmission blocks. The entries of $H$ are zero mean jointly circular Gaussian with variances scaled by path loss and shadow fading and $n$ is ($M_R \times 1$) with covariance $E[n_j n_j^H] = N_0 I_{M_R}$. All data streams $a_i$ are i.i.d. $\sim \mathcal{CN}(0, I)$ and $R_{dd} = E[d d^H] = \text{diag}(\gamma_1^2, \ldots, \gamma_K^2)$. $E_s$ is the total average transmit energy per symbol and $E_s / M_T$ is the average energy transmitted by each antenna per channel use. To constrain the total transmit power, $R_{dd} = \mathcal{E} \mathbb{E}[s s^H]$ must satisfy $\text{tr}(R_{dd}) = M_T$. Channel state information at the transmitter (CSIT) is assumed available, e.g., via time-division duplex or a feedback channel.

3. Applying RAS in BD-OSDM Schemes

3.1. Pertinent Points of BD-OSDM Processing

We begin by highlighting the pertinent points of block diagonalized orthogonal space division multiplexing (BD-OSDM). To provide an illustration, we will use the BD and CTR schemes discussed in [1]. In general however, the techniques developed in this paper are applicable to both iterative and non-iterative BD-OSDM schemes that make use of nullspace projection. Examples of these are found in [1] – [3].

To eliminate co-channel interference (CCI) between users, BD imposes $H_j T_j = 0$ for $i \neq j$. The channel rate for such a system with a power constraint is

$$C_{\text{bd}}(H) = \max_{\text{tr}(R_{dd}) = M_T, R_{dd} \succeq \mathbb{X}} \log_2 \det \left( I_{M_R} + \rho R_{dd} H_j^H H_j \right),$$

$$= \max_{\text{tr}(R_{dd}) = M_T, R_{dd} \succeq \mathbb{X}} \sum_{j=1}^{K} \log_2 \det \left( I_{M_R} + \rho H_j T_j R_{dd} T_j^H H_j^H \right),$$

where $\rho = E_s / M_T N_0$ and $[.]^H$ indicates Hermitian transpose. Next, we define $H_j = [H_j \ 0 \ 0 \ ... \ 0]$ and the columns of $H_j$ form an orthonormal basis for the left nullspace of $H_j$. This formulation allows direct access to the spatial modes of the projected channels and enables waterfilling to maximize each user’s throughput. Hence

$$C_{\text{bd}}(H) = \max_{\text{tr}(R_{dd}) = M_T, R_{dd} \succeq \mathbb{X}} \log_2 \det \left( I_{M_R} + \rho \mathbb{S} R_{dd} \right),$$

where $\mathbb{S} = \text{diag}(\Sigma_1, \Sigma_2, \ldots, \Sigma_K)$.

3.2. Impact of RAS in BD-OSDM

It has been shown in [5] that RAS is a necessary part of increasing the achievable channel sum rate and minimizing the transmission power for BD-OSDM systems. The benefits are realized even for cases where optimal beamforming with SVD-based techniques is done at each user to achieve individual projected channel capacity. Judicious implementation of RAS via algorithms such as those in [6] and [7] improves the spatial mode gains of the projected channels $H_j$ in two ways.

First, the removal of antennas with high inter-terminal correlation increases the orthogonality among the user channel sub-matrices $H_j$. Since BD-OSDM achieves zero inter-terminal interference by projecting each $H_j$ into the corresponding null-space of $H_j$, improving the orthogonality among users has the effect of decreasing the degree of orthogonality between the spaces spanned by $H_j$ and null($H_j$). This is advantageous when increases in the spatial mode gains for other users’ projected channels $H_{kj} = H_j V_j$ outweigh the rate loss at one user due to
RAS. Since the CTR scheme in [1] also makes use of the BD process, this advantage is applicable to it as well.

Second, each receive antenna removal at a particular terminal provides an additional degree of freedom to all other terminals. For example if one antenna is removed from a user \( k \), then the dimension of the projected channels \( H_{kj} \) of any other user \( j \) is

\[
M_{kj} \times (M_T - \sum_{i=1}^{k} M_{ci} + 1).
\]

(7)

The number of columns in \( H_{kj} \) is increased by one and this has the effect of adding more transmission resources to all users other than \( k \) (more details in [5]). This should be done using RAS algorithms such as those in [6] and [7] to minimize the loss at user \( k \). Noting that the capacity loss for user \( k \) arising from the removal of an antenna with high intra-terminal correlation is low in percentage terms, weeding out such antennas throughout the system can result in higher overall sum rates due to (7). When RAS is applied to all terminals, there is mutual benefit to be shared among users and this translates to sum rate increase. It was shown in [5] that the achievable sum rate increases significantly even though sum capacity loss is simultaneously present due to RAS. An example using a 4-user BD-OSDM system, each equipped with 4 antennas is given in Fig. 1. Notice that the gap between the sum capacity and the sum rate curves narrows when RAS is applied in the system. That this second advantage as captured in (7) is directly applicable to BD but not to CTR since spatial mode selection is more appropriate in that case. Note also that the RAS may be guided by different objectives besides being based on sum rate maximization. For example it may be guided by individual QoS requirements and depending on each user’s current channel rates, tradeoffs among the users may be made via RAS to better meet each user’s QoS needs. Users whose available channel rates are above their QoS requirements may deactivate more antennas. This will boost other users’ channel rates by re-directing more transmission resources to them. Again, the de-activation should be done using RAS algorithms to maximize the returns.

3.3. Impact of RAS on Scheduling in BD-OSDM

The rate gains due to RAS may be leveraged upon to enable the scheduling of more users. Let \( M'_{bj} \) be the number of receive antennas at user \( j \) after RAS, and \( M'_{b} = \sum_{i=1}^{k} M'_{ci} \leq M_T \). When \( M'_{b} < M_T \), it is possible to support an additional number of \( M_T - M'_{b} \) receive antennas, which may come from different additional users. Incorporation of these new users is possible especially if the user channel rates of the original schedule were increased due to RAS. This is illustrated using a 4-user BD-OSDM system, each user equipped with 4 antennas. First, RAS is performed on this 4-user system and as shown in [5], it results in channel rate increase. This is shown in Fig. 1 where the sum rate is increased and in Fig. 2 where the individual channel rates are increased. It generally resulted in \( M'_b < M_T \), which allows for the incorporation of more users. For illustration, we consider the impact of adding one more user. As shown in Fig. 2, the addition of a fifth user causes rate loss at all four existing users. This loss is reduced however when another round of RAS is done on the 5-user system. In fact, the resulting user rate is still above the original 4-user system without RAS. Note also that the channel rate of the fifth user is fairly good. When spare resources are still available, this incremental process of adding one more user may be repeated until the sum rate shows no more gain, or when individual rates of the currently scheduled users fall below their individual QoS requirements.

Next, for OSDM schemes like CTR [1] and Nu-SVD [2], RAS improves the spatial mode gains of the projected channels so that the number of activated modes may approach the maximum of \( M_T \) while meeting individual
QoS requirements. In turn, this raises the likelihood of being able to schedule the maximum number of $M_T$ users if desired. Numerical results for CTR with RAS given in Section 4 will illustrate this point.

3.4. Strategies for Integration of RAS and Scheduling

As shown above, the existence of highly correlated antennas at both intra- and inter-terminal levels causes poor sum rate performance in BD-OSDM. This means that a judicious choice must be made when introducing new users to an existing user subset. Next, we note that the potential pool of users is normally large, i.e., $M_s \gg M_T$, and this necessitates scheduling where a subset of $K'$ users is chosen so that $(M'_k = \Sigma_{j=1}^{c_{K'}} M_{R}) \leq M_T$. User subset selection may be done using selection schemes as in [8] or [9]. There are two schemes in [8] and the first one involves the repeated use of BD pre-coding for incremental user selection while the sum rate is maximized at each iteration. The second scheme is less complex but still requires repeated use of SVD and Gram-Schmidt orthogonalization procedures. The schemes in [9] are derived from antenna selection algorithms. A composite channel matrix is formed by combining all user sub-matrices and then a capacity maximization approach is adopted for user selection. Numerical results show good BD-OSDM performance for user subsets chosen by schemes from [9]. By avoiding the need for pre-coding procedures, the schemes in [9] are generally less complex. This is advantageous as it reduces the complexity when bringing the RAS and user scheduling functions together in BD-OSDM.

We propose an incremental RAS-plus-scheduling (termed IRASS for convenience) scheme to facilitate the scheduling of more users in BD-OSDM while minimizing the rate loss at currently scheduled users. Essentially, IRASS picks the next best user for inclusion in the current user subset if spare resources are available after RAS is done on the current subset. This is followed by another round of RAS on the new subset to improve its BD-OSDM performance. This may be repeated if spare resources are still available after RAS. However the stopping criterion may also be based on sum rate maximization or on individual user QoS requirements. This scheme was actually demonstrated earlier in Section 3.3. As mentioned before, care must be exercised when adding new users and this may be done using the algorithms like those in [8] or [9]. However, invoking these algorithms on an ad hoc basis may result in the de-selection of a currently scheduled user. To avoid this, one may include the requirement of scheduling more users from the onset when choosing the first user subset from the original pool of $K$ users. For example, one may rank all users in the pool so that the next users are ready in line for inclusion if possible. This is computationally intensive if the pool is very large and is likely not needed since the inclusion of too many new users will cause too much rate loss to the current user subset. Hence, ranking of the next few users may be sufficient in most situations. Note that this method of ranking the users once from the onset helps reduce complexity but is an approximation since RAS will affect the channel sub-matrices of the currently chosen users. In any case, the algorithms from [9] are less complex and facilitate the ranking of more users, beyond the initial subset.

Next, we will consider various approaches to RAS and the RAS algorithms. The material here is mostly drawn from [5]. In general, RAS done at a global level (across all users in the subset) will give better results than one done at the individual user level. In line with this, decremental antenna selection (DAS) is generally better than incremental antenna selection (IAS). This is a good situation for BD-OSDM since the number of correlated antennas to be removed via RAS operation is generally small compared to the total number of receive antennas. Hence the DAS approach will generally incur less complexity than the IAS approach. Examples of RAS algorithms can be found in [6] and [7]. The complexity of the DAS algorithm in [6] is upper bounded at $O(M'_k M'_T)$ for this application when RAS is done across all user antennas. This becomes $O(M'_k)$ when one starts from $M'_k = M_T$. The DAS algorithm in [7] is applicable to situations with $M'_k \leq M_T$ and is therefore applicable here. It has near optimal performance and has better performance than the DAS algorithms in [6] for transmit zero-forcing beamforming sum rate maximization. Its complexity is determined essentially by the iterative computation of $(\mathbf{HH}^H)^{-1}$ where $\mathbf{H}$ is the current composite channel matrix of all users in the selected subset. It has an upper bound complexity of around $O(M'^2)$ for typical user numbers and with RAS done across all user antennas. We introduce an improvement here by reducing the complexity to around $O(M'^{1.1})$. This is done via partitioned matrix inversion identities together with the fact that switching a pair of rows in $\mathbf{H}$ corresponds to switching a pair of rows and columns in $(\mathbf{HH}^H)^{-1}$ with the same corresponding indices. Note that RAS done on the global basis requires a maximum of $M'_k$ iterations, and a BD-OSDM process is needed for rate evaluation at each iteration.

Given that user terminals are normally geographically distributed, intra-terminal antenna correlations are usually higher than those found between terminals. This begs the question if complexity reduction may be achieved by performing RAS at a local user level. The motivation arises from the fact that the DAS complexity per use reduces to $O(M'^2 M'_T)$ for [6] and to $O(M'^{1.1})$ for [7]. One idea is to identify and eliminate the worst antenna at each user using a RAS algorithm and evaluate the corresponding BD-OSDM sum rate. The reduced user antenna set that results in the highest sum rate will be retained. This process is repeated until the sum rate does not improve any more with further RAS. Noting that this scheme requires re-examinations of all remaining antennas and that BD-OSDM pre-coding is needed for each examination, the advantage diminishes when the number of antennas to be eliminated is
more than the number of antennas per terminal. In fact, since the number of antennas to be eliminated is generally not high, the global RAS approach will arrive at the solution in less than $M_T^2$ iterations. Hence it is better to adopt the global RAS approach in general.

4. Numerical Results

The presence of spatial fading correlation in $H_j$ is captured by modeling the channel as $H_j = R_j^{1/2} H R_j^{1/2}$, where $H_j$ is the i.i.d. spatially white channel and $R_j$ and $R_t$ are positive definite Hermitian matrices that specify the receive and transmit correlations respectively. We have assumed that the BS antennas are well spaced so that $R_t = 1_{M_t}$. An exponential correlation model is used for the user terminals. Each element $r_{ij}$ in $R_j$ is $r_{ij} = \rho^{j-\frac{d_{ij}}{d_{ij}}}$, where $\rho$ is the maximum correlation between two antennas at each terminal. To avoid the trivial case where some users are dropped solely due to high channel loss, all users are placed at the same distance from the base station. The entries of $H_j$ are then $CN(0,1)$. The results here are obtained with RAS done on a global basis, i.e., across all users. Fig. 3 shows the BD-OSDM performance of an 8-user system when a ninth user is added. Although not shown, the ninth user’s rate is close to that shown in Fig. 3(b). Fig. 4 shows the CTR performance for an 8-user system. As shown, the sum rate and the individual channel rate are improved by RAS, which makes the scheduling of $M_T$ modes and hence $M_T$ users more realizable.

5. Conclusions

We have shown that receive antenna selection (RAS) can be used to create room for the scheduling of more users while minimizing the rate loss of existing users for multi-user MIMO wireless downlinks that employ block diagonalization to achieve orthogonal space division multiplexing (BD-OSDM). Strategies for the inclusion of RAS together with user selection algorithms in BD-OSDM have been presented. Numerical results have illustrated the benefits of RAS as it increases the overall system sum rate while minimizing the individual rate losses when making additional room available for the scheduling of more users.

Acknowledgment

The authors gratefully acknowledge the funding for this work provided by DSO National Laboratories (Singapore), Natural Sciences and Engineering Research Council (NSERC) of Canada, Alberta Informatics Circle of Research Excellence (iCORE), Rohit Sharma Professorship, and TRLabs.

References