Multiuser MIMO Scheme for Enhanced 3GPP HSDPA

Sungjin Kim†§, Hojin Kim‡, and Kwang Bok Lee† 1

†Seoul National University, School of Electrical Engineering, Seoul, Korea
‡Samsung AIT, Communication & Network Lab, Suwon, Korea
e-mail: kimsj@mobile.snu.ac.kr, hkim73@samsung.com, klee@snu.ac.kr

Abstract: We compare the achievable throughput of space-time multiple access (STMA) scheme versus that of time division multiple access multiple-input multiple-output (TDMA-MIMO) schemes, illustrated in the third generation partnership project (3GPP) MIMO technical report (TR). These schemes have been proposed to improve 3GPP high speed downlink packet access (HSDPA) channel by employing multiple antennas at both the transmitter and receiver. Our comparisons are focused on the performance in a multiuser environment, which is recently considered using the multiuser scheduling based on TDMA in both the single-input single-output (SISO) and the MIMO schemes for HSDPA. Motivated by those observations, we propose the hybrid STMA (H-STMA) scheme, in which the multiuser scheduling is operated in TDMA or STMA mode depending on the number of loaded users in the system. H-STMA is shown to outperform both TDMA-MIMO and STMA under HSDPA channel parameters. Moreover, we find that it has two practical advantages: the significant reduction of the receiver complexity and the required feedback amount.

1. Introduction

Recently, the high-rate data transmission has been one of key issues in wireless mobile communications [1]. Multiple-input multiple-output (MIMO) is an emerging technology offering high spectral efficiency with the increased link reliability and interference suppression [2], [3].

Based on the basic multiple antenna technologies, a lot of hybrid methods have been brought up for higher performance gain. MIMO can be separated into two structures which are open-loop and closed-loop systems. In open-loop MIMO (OL-MIMO), the transmitter has no channel information for data transmissions, and hence fixed transmit parameters are used. Closed-loop MIMO (CL-MIMO) exploits the channel state information for transmissions. Most previous MIMO schemes are based on point-to-point communications at a time, i.e., single-user MIMO (SU-MIMO). In multiuser MIMO (MU-MIMO) systems, all users are coordinated for communications by considering scheduling algorithms and quality of service (QoS) requirements of each user. In the case of CL-MIMO with multiple users, the complexity is of a concern, including feedback signaling, multiuser scheduling, and transmit/receive optimization, etc. Recently, the industrial organizations have proposed their MIMO techniques in 3rd generation partnership project (3GPP) standard. In 3GPP, various multi-antenna schemes are on active discussions, especially when combined with high speed downlink packet access (HSDPA).

In this paper, we review the proposed MIMO techniques which have been agreed to be included in 3GPP MIMO working item technical report (WI-TR). We also propose the MU-MIMO scheme with scheduling algorithms for 3GPP HSDPA. The proposed scheme is based on the multiuser scheduling which performs single user or multiuser transmissions adaptively.

The rest of the paper is organized as follows. In Section 2, we examine the characteristics of SU-MIMO and MU-MIMO as well as hybrid systems. Section 3 reviews the MIMO candidate solutions in 3GPP. In Section 4, scheduling based schemes are investigated in combination with our proposed MIMO scheme. Section 5 describes the hybrid scheduling algorithm based on the HSDPA MIMO candidate. Performance comparisons for various techniques are observed in Section 6. We conclude in Section 7.

2. System Model

In multiuser MIMO systems over wireless mobile channels, a radio base station (BS) communicates with K mobile stations (MSs) [4]. Each MS has the linear/non-linear reception entity and \( M_r \) receive antennas, while the BS has \( M_t \) transmit antennas. Based on the channel state information (CSI) fed back from MSs if necessary, BS performs appropriate space-time processing such as multiuser scheduling [5], power and modulation adaptation [6], beamforming [7], and/or space-time coding [8].

Assume that \( \mathbf{H}_k \) is the \( M_r \times M_t \) MIMO channel matrix from the BS to the \( k \)th MS, \( \mathbf{x} \) is the \( M_t \times 1 \) transmitted symbol vector, \( \mathbf{n}_k \) is the \( M_r \times 1 \) independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN) vector \( \sim \mathcal{CN}(0, \mathbf{I}_{M_r}) \), and \( \mathbf{y}_k \) is the \( M_r \times 1 \) received symbol vector. Then, the received signal for the \( k \)th MS in multiuser MIMO systems is mathematically described as

\[
\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{n}_k
\]  

where \( k = 1, \ldots, K \). The transmitter is subject to an average power constraint \( \text{Tr}(\mathbf{\Sigma}_x) \leq P \) where denotes the covariance matrix of the input signal. In our analysis, the

---

1This paper has been supported in part by the Samsung Advanced Institute of Technology (SAIT) and in part by National Research Laboratory (NRL) program.
mobile channel $H_k$ is modeled as a single path Rayleigh with i.i.d. entries $\sim \mathcal{CN}(0, 1)$ and block fading.

MIMO schemes for a broadcast channel (BC) represented as (1) can be brought into several scenarios, depending on the assumptions about the BS and the constraints put on transmit streams. First, we consider two methods of multiple access, which are time division multiple access multiple-input multiple output (TDMA-MIMO) [9] and space-time multiple access (STMA) [10]. TDMA-MIMO is a point-to-point communication, in which BS transmits to the one selected user at a time, in which BS transmits only to the user with the largest capacity at a time, and hence the sum-rate capacity is given as

$$C_{TDMA}(H_1, \ldots, H_K) = \max_{k=1,\ldots,K} C(H_k)$$

where we have used the definition that the maximum sum-rate of TDMA-MIMO is the largest single user capacity of the $K$ users. It is shown in [9] that the TDMA-MIMO sum-rate is lower bounded by

$$C_{TDMA}(H_1, \ldots, H_K) \geq \log(1 + P ||H_k||_{max}^2)$$

where $||H_k||_{max} = \max_{k=1,\ldots,K} ||H_k||$.

2.2. TDMA-MIMO Capacity

Consider the sum-rate capacity of the TDMA-MIMO, in which BS transmits only to the user with the largest capacity at a time, and hence the sum-rate capacity is given as

$$C_{TDMA}(H_1, \ldots, H_K) = \max_{k=1,\ldots,K} C(H_k)$$

where we have used the definition that the maximum sum-rate of TDMA-MIMO is the largest single user capacity of the $K$ users. It is shown in [9] that the TDMA-MIMO sum-rate is lower bounded by

$$C_{TDMA}(H_1, \ldots, H_K) \geq \log(1 + P ||H_k||_{max}^2)$$

2.3. STMA Capacity

It has been known that the sum capacity of STMA is achieved by using dirty paper coding (DPC) to simultaneously transmit to not only one user but also several users optimally selected. Intuitively, DPC processing for BC can be seen like the successive interference cancellation (SIC) with the minimum mean-square error (MMSE) QR decomposition at the transmitter side. The sum capacity of STMA, which can be achieved by DPC, is given in [9] by the upper bound

$$C_{STMA}(H_1, \ldots, H_K) = C_{DPC}(H_1, \ldots, H_K)$$

where we have employed the duality of the multiple antenna BC and MAC. STMA is shown to achieve the gain around two times over TDMA-MIMO, where the sum-power iterative water-filling proposed in [13] is used to obtain STMA capacity.

On the other hand, users located in different geometries must be considered in the realistic fading channel environment, where average SNRs of users in different locations are commonly not the same [5]. In that case, in order to ensure fairness between users we suggest to use, namely, the iterative water-filling with both the sum power constraint and long-term individual power constraints, which is designed using a fairness criterion
given by

\[
\{Q_k\}_{k=1}^{K} = \arg \max_{\{Q_k\}_{k=1}^{K}} \log I + \sum_{k=1}^{K} H_k^T Q_k H_k \quad (7)
\]

s.t. \(Q_k \geq 0, \sum_{k=1}^{K} \text{Tr}(Q_k) \leq P, \mathbb{E}[\text{Tr}(Q_k)] \leq \frac{P}{K}.

We propose a practical method for implementing (7), which is a two-stage approach. In our proposed scheme, we first find allowable powers \(\{p_k\}_{k=1}^{K}\) for each user using the sum-power iterative water-filling [13] with a normalized channel matrix set \(\{H_k/\sqrt{\mathbb{E}[\|H_k\|_F^2]}\}_{k=1}^{K}\).

Based on the individual-power iterative water-filling [14], both \(\{p_k\}_{k=1}^{K}\) as power constraints and \(\{H_k\}_{k=1}^{K}\) are then applied to optimize each input covariance matrix \(\{Q_k\}_{k=1}^{K}\). A detailed description is beyond the scope of this paper, and more related work is found in [15].

The ergodic capacity of STMA with \(K\) users in a fading channel, based on the result of (2), is upper bounded by

\[
\begin{align*}
C_{M_t, M_t, K} & \leq M_t \bar{C}_{1,1,K} \\
& = M_t \sum_{k=1}^{K} \left( \frac{K}{k} \right) (-1)^{k+1} e^{k/P} \Gamma_0[k/P] \\
\end{align*}
\]

(8)

where \(\Gamma_0[z] = \int_{z}^{\infty} t^2 e^{-t} dt\) is the incomplete gamma function.

In practice, DPC seems to be difficult to implement and has yet to be realized. One feasible approach is beamforming with the space-time multiuser scheduling such as the per-user unitary rate control (PUSRC), which will be described in Section 4.1.

3. TDMA-MIMO Proposals in 3GPP

In this section, we investigate advanced MIMO solutions, mainly focusing on MIMO candidates in 3GPP standardizations. In 3GPP, major industrial organizations proposed their schemes toward the official use of MIMO in future wireless communication systems.

3.1. PARC

Lucent initially proposed their multiple antenna solution, which is called the per-antenna rate control (PARC), in 3GPP MIMO TR [16], in which separately encoded data streams are transmitted from each antenna with equal power but possibly with different data rates while spreading code is reused through all streams. The data rates for each antenna are controlled by adaptively allocating transmit resources such as modulation order, code rate, and number of spreading codes. The postdecoding SINR of each transmit antenna is estimated at the receiver and then fed back to the transmitter, which is used to determine the data rate on each antenna.

The MMSE filtering and SIC are applied to the receiver, i.e., SIC reception. Let

\[
G_{k,m} = (H_k^H H_k + (1/P) I_{M_t-K})^{-1}
\]

(9)

where \(I_{M_t-K}\) is the \(M_t-K\) dimension square identity matrix and \(H_k\) is a deflated version of \(H_k\) in which columns 1, 2, ..., \(K\) have been zeroed. The received SINR of the \(m\)th stream becomes

\[
\gamma_{S,k,m} = \frac{P/M_t}{\|G_{k,m}\|_{mm}} - 1.
\]

(10)

The capacity is then

\[
C_k = \sum_{m=1}^{K} c_f(\gamma_{S,k,m})
\]

(11)

where \(c_f(\gamma) = \log(1 + \gamma)\). On the other hand, by replacing \(H_k\) by \(H_k\) in (4), the received SINR, denoted as \(\gamma_{M_t,k,m}\), and the capacity can be derived for MMSE reception.

3.2. S-PARC

The selective PARC (S-PARC) has been proposed by Ericsson, which is conceptually based on PARC scheme in the previous subsection [17]. There is a significant gap between the OL capacity and the CL capacity when signal-to-noise ratio (SNR) is low and/or the number of receive antennas is less than the number of transmit antennas. An alternative way is by S-PARC, which overcomes performance gap by the gain of antenna selection. More specifically, S-PARC adaptively selects the number of antennas, i.e., mode, and the best subset of antennas for the selected mode. Interestingly, S-PARC will operate like a single stream transmit diversity with transmit antenna selection if the number of the selected antennas is limited to one.

The antenna selection is suboptimal, so we first review the optimal power allocation across antennas before describing the operation of S-PARC. To allocate optimal power for each transmit antenna in PARC-like systems, we may use the water-filling with iterative method. From the similarity of PARC and MAC, the objective of such a problem in SU-MIMO channels can be represented by the simplified form of (6), which is given by

\[
C_{S-PARC}(H) = \max_{\{p_m \geq 0, \sum_{m=1}^{M_t} p_m \leq P\}} \log \left| I + \sum_{m=1}^{M_t} p_m h_m h_m^H \right|
\]

(12)

where \(H = [h_1, h_2, \ldots, h_{M_t}]\). Substantial progress for optimization has recently been made [6], [13], which is not yet completely answered because of its implication of the sum-power constraint, i.e., \(\sum_{m=1}^{M_t} p_m \leq P\).

Secondly, S-PARC performs the antenna selection based on a subset property, by which selections at the
prior mode are considered, so as to reduce the amount of feedback [18]. For the equivalent approach in multiuser scenario, greedy DPC is suggested in [19].

Fig. 1 shows the maximum achievable rate for comparison when $M_t = 4$ and $M_r = 2$, in which S-PARC outperforms PARC by about 1.8dB in all range of SNR. Note that this gain comes from antenna selection diversity.

### 3.3. PSRC

The per stream rate control (PSRC) is another proposed scheme by Lucent, in which each substream is weighted prior to transmission. As in transmit antenna array (TxAA), the amplitude and phase of the weight matrix is appropriately quantized to have a set of finite matrices. The MS can choose a single stream for transmission, which is the case of selection transmit diversity using TxAA. Note that PARC is a special case of PSRC when the weight matrix is the identity matrix.

### 4. STMA Proposal in 3GPP

#### 4.1. PU$^2$RC

We propose the multiuser MIMO scheme using the unitary basis matrix, which is called the per-user rate control (PU$^2$RC). The proposed system model is shown in Fig. 2. Applying to (1), we have the received signal vector as

$$y_k = H_k T s + n_k$$

where $L \leq M_t$ is the total number of transmit streams. Transmit beamforming is denoted as $x = Ts$ where $T = \{ t_1 \cdots t_L \}$ is the beamforming matrix. Since PU$^2$RC is a STMA scheme, each $s_l$ is allocated to users independently, i.e., it can be allocated to different users.

We let the beamforming matrix $T$ in (13) be a unitary matrix, i.e., $T^H T = I_L$, in order to improve the capacity obtained by the matched filter beamforming (hereafter denoted by unitary matched filter beamforming (UMF-BF)). Although UMF-BF is much simpler to implement than other transmit precoding methods such as dirty-paper coding (DPC), UMF-BF combined with the space-domain user diversity leads to significant capacity performance improvement. It follows from [20] and [21] that the sum rate of both UMF-BF and DPC scales as $M_t \log \log K M_r$, when $K$ is large. To utilize user diversity in the space and time domains, $T$ is obtained by

$$t_l = \arg \max_{v_{k,m}} c_f(\rho_k(v_{k,m})) \quad l = 1, \ldots, L$$

s.t. $T^H T = I_L$ (14)

where $v_{k,m}$ is the quantized version of the $m$th eigenvector of $(H_k^H H_k)$ by use of a subspace packing such as Grassmannian line packing, and $\rho_k(v_{k,m})$ is the received SINR function of $v_{k,m}$ for the $k$th user. $\rho_k(t_l)$ can be expressed as

$$\rho_k(t_l) \geq \frac{|b^H H_k t_l|^2}{|b^H H_k|^2 - |b^H H_k t_l|^2 + \frac{L}{2} |b|^2}$$

where $b$ is the receive beamforming vector for the $k$th user and $t_l$, and the equality holds if $L = M_t$. Note that the other transmit beamforming vectors, i.e., $\{ t_m \}_{m \neq l}$, are not used in (16), which easily become useless when $L \rightarrow M_t$ from the unitary property of $T$. Therefore, PU$^2$RC incorporating UMF-BF at the transmitter offers the following two advantages: simplified user diversity in the space domain, and effective calculation of received SINRs.

#### 4.2. Feedback Signaling for PU$^2$RC

The CL-MIMO obtains channel information at the transmitter through feedback channel. In this subsection, we describe the characteristics of feedback information in PU$^2$RC, and design feedback signaling protocols. In PU$^2$RC, two types of channel information are fed back to
the transmitter, which are the beamforming vectors and the corresponding channel qualities. More specifically, the beamforming vectors and the channel qualities are the quantized eigenvectors of each user, i.e., \( \{v_{k,m}\}_m \), and the received SINRs, i.e., \( \{\rho_k(v_{k,m})\}_m \), respectively, as described in the previous subsection. We consider quantized vectors from the set predefined by a subspace packing, where the beam selection is preferable to the eigen-decomposition which is practically difficult to implement. In particular, the set of selected vectors correspond to the maximum sum rate at the receiver and are optionally constrained to be orthonormal to each other.

According to the characteristics of feedback information described above, we now take into account three feedback protocols: full feedback, partial feedback, and hybrid feedback protocols. The information of the \( k \)th user for feedback signaling is given by

\[
F_{A,k} = \{g_k, \{\gamma_{M,k,m} = 1, \ldots, M\}\}
\]

(17)

\[
F_{B,k} = \{g_k, m_S, \gamma_{M,k,m_S}\}
\]

(18)

which represents the full feedback and the partial feedback protocols, respectively, where \( g_k \) is the index of the set of the selected vectors, and \( \{\gamma_{M,k,m}\}_m \) are the received SINRs estimated at the receiver based on \( g_k \). Note that all \( \{\gamma_{M,k,m}\}_m \) denote the post-decoding SINRs based on the MMSE reception. To reduce the burden of feedback, \( F_{B,k} \) in (18) contains the maximum SINR \( \gamma_{M,k,m_S} \) as well as its index \( m_S \), instead of SINRs for all vectors in (17), where

\[
\gamma_{M,k,m_S} = \max_{m=1,\ldots,M} \gamma_{M,k,m}.
\]

(19)

In practice, the feedback protocol \( F_{B,k} \) is organized as follows. A 1-bit is used to specify \( g_k \); a 2-bit denotes \( m_S \), and a 5-bit is assigned to \( \gamma_{M,k,m_S} \). The last 5-bit has been used for the SINR feedback signaling in the HSDPA specifications. Finally, the protocol for hybrid feedback is given by

\[
F_{C,k} = \{g_k, m_S, \gamma_{M,k,m_S}, \gamma_{S,k,m} \in S\}
\]

(20)

where SINRs are included for both MMSE and SIC receiver structures (i.e., \( \gamma_{M,k,m_S} \) and \( \gamma_{S,k,m_S} \)), respectively, while the number of SINRs in (20) and (17) are the same.

5. Hybrid STMA with PU²RC

In this section, multi-user MIMO scheme with scheduling is proposed for MIMO broadcast channel. Scheduling methodology is considered because all users cannot be served at the same time due to the limited resources (e.g., the number of antennas, transmit power, etc.). We exploit scheduling schemes using user diversity for MIMO systems when advanced receivers, i.e., SIC receivers, are utilized, and propose the effective hybrid scheduling methods for such systems. In MIMO systems, two basic scheduling methods have been considered [4]. One of them is that all the transmit antennas are assigned to a single user selected based on the single user multiplexing methods. Regardless of a receiver structure (whether SIC or not), its capacity is expressed as

\[
C_A(t) = \max_k \sum_m c_f(\gamma_{k,m}(t))
\]

(21)

where \( \gamma_{k,m}(t) \) can be either \( \gamma_{S,k,m}(t) \) or \( \gamma_{M,k,m}(t) \). The other one is that all users compete independently for each transmit antenna for performance enhancement. The capacity of this scheme heavily depends on a particular receiver structure so that it is expressed as

\[
C_{B,1}(t) = \max_Q \sum_m \min_{k \in Q_m} c_f(\gamma_{S,k,m}(t)),
\]

(22)

\[
C_{B,2}(t) = \sum_k \max_m c_f(\gamma_{M,k,m}(t)),
\]

(23)

for SIC receivers and linear receivers, respectively, where \( Q \) is a possible sub-set of all users, \( Q_{m+1} \) is deflated version of \( Q_m \) in which the user after decoding at the \( m \)th layer has been zeroed, and \( Q_1 = Q \). By the fact that the capacity of (22) is apparently equal to that of (21), i.e., \( C_{B,1}(t) = C_A(t) \), we may use the scheduler (21) for simplicity when advanced receivers are involved.

To achieve the maximum capacity through advanced receivers, hybrid schedulers can be used. One of hybrid schedulers, suggested in [22], is given by

\[
C_{H,1}(t) = \max\{C_A(t), C_{B,2}(t)\},
\]

(24)

in which both metrics of \( C_A(t) \) with SIC receivers and \( C_{B,2}(t) \) with linear receivers are used to select the best user. It is seen in [22] that only one metric is sufficient for the hybrid scheduling if we simply switch the scheduling policies between \( C_A(t) \) and \( C_{B,2}(t) \) after the threshold point determined by the number of scheduled users. In practice, it is desirable to choose the point \( K_{sw} \) satisfying \( E\{C_A(t)\} = E\{C_{B,2}(t)\} \), so that the rule of the modified hybrid scheduler is then

\[
C_{H,2}(t) = \begin{cases} 
C_A(t), & K <= K_{sw} \\
C_{B,2}(t), & K > K_{sw} 
\end{cases}
\]

(25)

Since it is often difficult to perfectly know how many users are to be scheduled before the activation of the scheduling method, we propose the hybrid scheme in which reception is to be constrained as single user SIC (SU-SIC), which is given by

\[
C_{H,3}(t) = \max_{\{S_j\}} \sum_j \max_{m \in S_j} c_f(\gamma_{H,k,m}(t))
\]

(26)

where \( \gamma_{H,k,m}(t) \) is the received SINR obtained by SU-SIC receiver, e.g., the SINR in (20), \( S_j \) denotes the \( j \)th sub-group of transmit antennas with constraints \( \bigcup_j S_j = \{1, \ldots, M\} \).
6. Numerical Results

In this section, we provide simulation results of performance comparisons for MIMO transmission schemes and scheduling. In Fig. 3, it is shown that the proposed PU²RC scheme outperforms PARC-MMSE because the PU²RC has about 2dB gain of transmit beamforming with a 4-bit feedback over the system without beamforming, and achieves additional user diversity gain over PARC-MMSE with and without user diversity by about 3.5dB and 7dB, respectively. The additional gain over PARC-MMSE with user diversity is user diversity gain in the space domain, which cannot be exploited in PARC schemes. The number of users is assumed to be $K = 10$.

In Fig. 4, we show the performance of PU²RC with partial feedback. In Figs. 4 and 5, we illustrate results in terms of the number of users at average SNR = 10dB, and a 1-bit feedback is used in PU²RC for transmit beamforming. PU²RC outperforms PARC-MMSE when feedback information for all transmit antennas are transmitted from user terminals back to the BS, and with the assumption of the partial feedback, i.e., the SINR of the selected basis or antenna vector, PU²RC still has significant gain over PARC and S-PARC. This is because with the partial feedback, S-PARC only exploits one transmit antenna, which results in a limited capacity gain over the number of users, while PU²RC can transmit as many data streams as transmit antennas at its maximum.

In Fig. 5, the throughput performance of PU²RC with hybrid scheduling is examined. It shows that when the number of users is less than 7, the performance of PARC, using SIC receivers, is better than that of PU²RC, but for higher number of users PU²RC, using linear receivers, outperforms the PARC. As expected from this result, the hybrid scheduling scheme between PU²RC and PARC performs better than both schemes, independent of the number of users. Note that the performances of all schemes are upper bounded by PU²RC with the optimal beamforming.

7. Conclusions

In this paper, we proposed the advanced MIMO scheme using hybrid scheduling algorithm for 3GPP HSDPA. We investigated two categorized transmission methods such as SU-MIMO and PU²RC for MU-MIMO. The proposed scheme has practical advantages, which are the reduction of receiver complexity and the amount of feedback. For future work, the efficient resource allocation scheme is required for SU-MIMO and
MU-MIMO communications with feedback signaling. Moreover, spatial channel modeling needs to be combined together with the above MIMO schemes for fair comparisons in real system environments.

8. Acknowledgment

The authors would like to thank Dr. Hyeon Woo Lee, Dr. Sungho Choi, and Dr. Juho Lee for their comments, as well as 3GPP MIMO delegates for their valuable discussions during the meetings.

REFERENCES


