Effects of multi-user MIMO scheduling freedom on cellular downlink system throughput
Fernandes, Pedro; Kyritsi, Persefon; Berger, Lars Torsten; Mártires, Jorge

Published in:
Proc. IEEE 60th Vehicular Technology Conference, Sept. 2004

Publication date:
2004

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Citation for published version (APA):
Effects of Multi User MIMO Scheduling Freedom on Cellular Downlink System Throughput

Pedro Fernandes
Instituto Superior Técnico
Technical University of Lisbon
Lisbon, Portugal

Persefoni Kyritsi
Department of Mathematics
Stanford University
Stanford, CA-USA

Lars T. Berger, Jorge Mártires
Department of Communication Technology
Aalborg University
Aalborg, Denmark

Abstract—This paper studies different degrees of channel quality-based scheduling freedom for the downlink of a DS-CDMA system where the base station (BS) and each mobile station (MS) have 2 antennas. The system combines opportunistic user time slot scheduling, opportunistic spatial multiplexing, and adaptive modulation and coding (AMC) with limited and unlimited allowable constellation sizes. System performance is investigated in a time dispersive propagation environment, with varying degrees of antenna correlation. Three transmission schemes with increasing degrees of resource allocation freedom are identified. The full freedom scheme allows for time slots and transmit antennas to be independently allocated to different users and is compared to two schemes of more limited flexibility in the antenna allocation. For a system with little multi user diversity, transmit diversity is favorable, while dual antenna transmission with independent antenna-to-user allocation becomes increasingly important when a higher degree of multi user diversity can be exploited.

Keywords- MIMO, proportional fair scheduling, spatial multiplexing, AMC

I. INTRODUCTION

Since the revolutionary publication by Telatar [1], the idea of multiple input multiple output (MIMO) antenna systems has been the topic of extensive research activity worldwide because of their potential to achieve high spectral efficiencies, and therefore high data rates. Different open [2] and closed loop [3] strategies have since been proposed to enable the transmission of parallel data streams over the MIMO radio channel. These techniques discuss the single user capacity of MIMO systems, i.e. they assume that the multiple transmitted data streams are destined to a single user, that is equipped with a multiple antenna enabled terminal.

In the context of multi-user MIMO systems, the concept of spatial multiplexing to different users has been introduced among others in [4]. Here one can think of a system where different antennas of the same transmit array are sending independent data streams for different users and potentially more than one BS antenna can be used for the same user.

As data traffic becomes more important, the focus shifts from delay stringent circuit switched communication to delay tolerant packet data communication. In this context, opportunistic transmission has been proposed as a way to exploit multi-user diversity. Using channel aware scheduling [5], opportunistic transmission techniques grant different users access to the system resources when their instantaneous channel conditions are favorable, thus enhancing the overall multi user system throughput [6], [7]. Only a limited amount of channel quality information per user is required at the transmitter side and due to its computationally practical nature and its enormous potential, channel quality based scheduling has recently been introduced to the downlink packet data transmission of emerging 3G evolution systems, i.e. 1xEV-DV [8], and HSDPA [9].

In this paper, we combine the concepts of MIMO transmission and opportunistic scheduling, and investigate the effects of increasing the scheduling freedom on the system throughput. Section II describes the simulations at link and system levels and states its main assumptions, while Sections III and IV illustrate the results and the conclusions, respectively.

II. SYSTEM MODEL

We have simulated a specific MIMO downlink DS-CDMA system, where independent data streams are sent through different antennas, potentially to different users.

In conventional orthogonal spread DS-CDMA systems like WCDMA and CDMA2000 the downlink data destined for any user is spread by a Walsh-Hadamard code unique to this particular user and scrambled by a sequence unique to the cell. Since the number of orthogonal spreading codes is limited, codes become a scarce resource within a cell and it might be desirable to have the option of reusing them for simultaneous data transmission from different antennas. Therefore, the studied system simulates transmissions where the total available power is used to target a group of users that time share a single spreading code.

Within this group, multiple access is achieved via

a) opportunistic spatial multiplexing, i.e. separate data streams are transmitted simultaneously from different antennas, therefore using different propagation channels, and/or

b) opportunistic time scheduling, i.e. different time-slots are allocated to different (sets of) users within this group.
Opportunistic spatial multiplexing enables simultaneous transmissions using the same spreading code, and opportunistic time scheduling takes advantage of the independence of the user fading statistics to provide multi user diversity [10].

A. Link level description

This section includes the transmission (spreading, scrambling and modulation), the propagation through a MIMO channel, the reception, and the CQI feedback.

At the BS, the data for all the users is spread using a single spreading code with a spreading factor of 16 and scrambled using a long Gold sequence particular to this BS. Quadrature amplitude modulation (QAM) is used with a constellation size dependent on the feedback from the MSs, as explained in section II.B.

The transmission frame has a 2 ms-long slotted structure similar to that of a Transmit Time Interval (TTI) used in HSDPA. It consists of 3 slots, 160 symbols per slot. The slot SINR is calculated as the geometrical average of the symbol SINRs as proposed in [11].

The MIMO channel model employed is of the correlation-based stochastic type, as described in [12] in compliance with the ITU-Vehicular A power delay profile (PDP). In order to illustrate the effect of the channel characteristics, different correlation levels were considered, namely low (0.3) and high (0.8) correlation at both ends of the communication link, and mixed high (transmitter) and low (receiver) correlation. Path loss, shadowing and other cell interference are determined by the geometry G factor, which is randomly selected to satisfy the probability distribution function over the cell area given by [13]. A fraction of the total transmit power (75%) is allocated to data transmission, while the rest is used for pilot and control signaling. We consider slowly moving users (3 km/h) and can therefore maintain the G factor value constant throughout the simulated transmission period.

Each MS estimates the channel transfer characteristics with the help of pilot signals (such as in UMTS), and, for the purpose of this analysis, we assume perfect channel estimation. The MS receiver can then use this knowledge to decode the originally transmitted data streams. Each user is equipped with a zero-forcing (ZF) receiver [14].

The MSs feed back to the BS channel quality indication (CQI) values on every possible antenna allocation option. The fed back CQI is the predicted SINR for each case of concurrent activity of the transmit antennas. Alternatively, the MSs could have fed back the wideband complex transfer matrix of the channel. However, feeding back only CQI limits the amount of feedback data and is less sensitive to errors and/or channel variations. It is assumed that the feedback link is error free and that the channel does not change significantly within the scheduling delay (which includes data reception and feedback of CQI from the MS, and user, antenna and constellation size selection at the BS). This assumption is not seriously violated for slowly moving MSs, such as the ones in our simulated scenario.

B. System level description

A smart packet scheduler located at the BS uses the CQI feedback (combined with knowledge about past transmissions) to allocate the BS’s antenna resources to the MSs dynamically on a time-slot basis. Moreover, the BS runs an AMC mechanism: the size of the transmitted QAM constellation size is adjusted according to the fed back CQI in order to maximize the achievable data rate for the given channel quality. The fed back SINR is mapped to a certain constellation size and data rate as in [15] under certain target performance constraints. The allowable constellation size is either unlimited or upper-bounded to 16 QAM. The symbol rate is adjusted independently for each of the BS transmit antennas similar to per antenna rate control (PARC) [16]. The total transmit power is kept constant independently of the modulation size used and of the number of active transmitting antennas.

In systems with opportunistic scheduling, the user is selected dynamically based on a scheduling algorithm. We have selected the Proportional-Fair Resource algorithm (P-FR) [8], because it achieves a better balance between cell throughput and user fairness [17]. The priority of user j to be scheduled at time instant n, Pr(j,n), is determined by

\[
Pr(j,n) = \frac{R_j(n)}{\overline{R}_j(n)}
\]

where \(R_j(n)\) stands for the potentially achievable throughput of the channel between the BS and the MS j at time slot n, and \(\overline{R}_j(n)\) is the mean allocated data to user j from the beginning of the transmission until time slot n. \(\overline{R}_j(n)\) is given by

\[
\overline{R}_j(n) = \left(1 - \frac{1}{t_c}\right)\overline{R}_j(n-1) + \frac{1}{t_c}R_j(n-1)
\]

\(R_j(n-1)\) equals the allocated data rate to user j at time slot n-1 if he was scheduled and 0 otherwise, as described in [8]. The parameter \(t_c\in[1,\infty)\) tunes the scheduler between a stringent \((t_c\to\infty)\) and a delay-tolerant \((t_c\to1)\) behavior. It is generally accepted the value of \(10^3\) [8]. At any time slot n, the scheduler selects to transmit to the user j with the maximum priority \(\max_j\{Pr(j,n)\}\).

Specifically for MIMO systems with independent stream transmission and antenna allocation, (1) can be generalized to express the priority of a scheduling option, which now includes the selection of a subset of users \(J\) and a subset of antennas \(I\): a

\[
Pr(I,J,n) = \sum_{(i,j)} \frac{R_j(n)}{R_j(n)}
\]
where $R_i(n)$ is the achievable data rate of the channel between the BS antenna $i \in I$, and MS $j \in J$ at time slot $n$ and $(i,j)$ stands for each of the BS antenna-user links that define the particular scheduling option. The scheduler selects the subset of users and antennas with the maximum priority \( \max_{(i,j)} \{ \text{Pr}(I,J,n) \} \).

In this paper, we consider offered services that are equally delay tolerant. If not, advanced scheduling algorithms can take into account the quality of service requirements of each user by modifying their priorities, e.g. through the variation of $t_c$. Such investigations are beyond the scope of this paper.

The transmitter (BS) and the receivers (MSs) were equipped with 2-element antenna arrays. The array size was kept small to simulate a realistic configuration. In a cell, the number of users sharing a code ranges from 1 to 10. The scheduling techniques have shown stable performance for these loading levels.

Under these conditions, there are three transmission options:

- **SMP-D1**: Dual stream to one user. This is associated with the feedback of a single CQI value. Here, $I = \{\text{ant.1 and ant.2}\}$ and $J = \{j\}$, $j \in [1,10]$, then (3) becomes

$$\text{Pr}(I,J,n) = \frac{R_{\text{ant.1,j}}(n)}{R_j(n)} + \frac{R_{\text{ant.2,j}}(n)}{R_j(n)}$$ (4)

- **SMP-D2**: Dual stream to two users, one stream per user (spatial stream- and user- multiplexing). This is associated with the feedback of 2 CQI values, each corresponding to the operation of one transmit antenna, while the other antenna is sending data to a different user and acts as an interferer. Here, $I = \{\text{ant.1 and ant.2}\}$ and $J = \{j_1, j_2\}$, $j_1, j_2 \in [1,10]$; (3) becomes

$$\text{Pr}(I,J,n) = \max \left\{ \frac{R_{\text{ant.1,j}_1}(n)}{R_{j_1}(n)} + \frac{R_{\text{ant.2,j}_2}(n)}{R_{j_2}(n)}, \frac{R_{\text{ant.1,j}_2}(n)}{R_{j_2}(n)} + \frac{R_{\text{ant.2,j}_1}(n)}{R_{j_1}(n)} \right\}$$ (5)

- **SMP-S1,2**: Single stream to one user. This is associated with the feedback of 2 CQI values, each corresponding to the operation of one transmit antenna keeping the other silent. Here, $I = \{\text{ant.1 or ant.2}\}$ and $J = \{j\}$, $j \in [1,10]$; (3) becomes

$$\text{Pr}(I,J,n) = \max \left\{ \frac{R_{\text{ant.1,j}}(n)}{R_j(n)}, \frac{R_{\text{ant.2,j}}(n)}{R_j(n)} \right\}$$ (6)

We study three transmission schemes that progressively allow more transmission options.

- **Scheme 1** only allows SMP-D1.
- **Scheme 2** allows SMP-D1 and SMP-D2
- **Scheme 3** allows all possible options.

As we allow more and more options, we increase the scheduling freedom in our system. An overview of the parameter settings employed in the link and system level simulations is given in TABLE 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Link Level</strong></td>
<td></td>
</tr>
<tr>
<td>Carrier frequency</td>
<td>2.15 GHz</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>5 MHz</td>
</tr>
<tr>
<td>Spreading factor</td>
<td>16</td>
</tr>
<tr>
<td>Modulation</td>
<td>QAM</td>
</tr>
<tr>
<td>Maximum constellation size</td>
<td>16</td>
</tr>
<tr>
<td>Time Unit</td>
<td>TTI (2 ms)</td>
</tr>
<tr>
<td>PDP</td>
<td>ITU Vehicular A</td>
</tr>
<tr>
<td>MS speed</td>
<td>3 km/h</td>
</tr>
<tr>
<td>Receiver</td>
<td>ZF</td>
</tr>
<tr>
<td>Channel model</td>
<td>Correlation-based stochastic type</td>
</tr>
<tr>
<td>Tx. power correlation coefficient</td>
<td>0.3 (Environment 1)</td>
</tr>
<tr>
<td></td>
<td>0.8 (Environment 2)</td>
</tr>
<tr>
<td></td>
<td>0.8 (Environment 3)</td>
</tr>
<tr>
<td>Rx. power correlation coefficient</td>
<td>0.3 (Environment 1)</td>
</tr>
<tr>
<td></td>
<td>0.3 (Environment 2)</td>
</tr>
<tr>
<td></td>
<td>0.8 (Environment 3)</td>
</tr>
<tr>
<td>Power allocated to data transmission</td>
<td>75 % of total transmit power</td>
</tr>
<tr>
<td>Channel Estimation</td>
<td>Ideal</td>
</tr>
<tr>
<td>CQI</td>
<td>SINR</td>
</tr>
<tr>
<td>Feedback Link</td>
<td>Error and delay free</td>
</tr>
<tr>
<td><strong>System Level</strong></td>
<td></td>
</tr>
<tr>
<td>Link Adaptation</td>
<td>AMC, user and antenna selection</td>
</tr>
<tr>
<td>Packet scheduler</td>
<td>Generalized P-FR</td>
</tr>
<tr>
<td>Time constant</td>
<td>10'</td>
</tr>
<tr>
<td>Number of BS and MS antennas</td>
<td>2</td>
</tr>
<tr>
<td>Number of users</td>
<td>[1,10]</td>
</tr>
</tbody>
</table>
III. RESULTS

A. SINR probability distribution and AMC

Figure 1. QAM constellation sizes capacities and SINR dynamic range.

Figure 1. shows how the post-detection SINR is mapped to a given modulation. This curve is similar to the one in [15], and is shown here for illustration purposes. Moreover, the plot shows the probability density function of the observed post-detection SINRs, which will be used later to explain the effect of limited maximum modulation size. We observe that about 50% of the observed SINRs would have resulted in modulation sizes larger than 16QAM ($M_{\text{MAX}}=16$).

B. Scheduling freedom

The cellular throughput achieved by each scheme (for limited and unlimited modulation) was investigated in the low transmit and receive correlation environment ($\rho_{\text{TX}}=\rho_{\text{RX}}=0.3$).

Figure 2. shows the spectral efficiency increase when multiuser diversity can be exploited: the more users there are in the cell, the higher the probability of encountering one in favorable channel conditions is. For a large number of users, the largest gain is obtained by allowing independent antenna allocation because the event of two favorable channels destined for two different users becomes more likely. For a small number of users, multi-user diversity does not gain as much as the freedom to switch off one antenna and reduce stream interference.

Figure 3. shows that, similarly to the case of unlimited modulation, turning one antenna off becomes an attractive option when there is not a lot of multi-user diversity to be exploited in the system. When the number of users increases, the penalty to be paid by limiting the maximum modulation size is up to 2b/s/Hz. Indeed Figure 1. showed that a significant set of SINRs would be affected by the capacity upper bound of the 16 QAM constellation size (4 b/s/Hz).

C. Effect of channel spatial correlation

Figures 3-5 show the probability of selection of each option in different correlation environments. For this set of curves, we have $M_{\text{MAX}}=16$.

Figure 4. shows that if the transmit and receive correlations are low, dual stream transmission to a single or two users (SMP-D1 or SMP-D2) is more likely than single stream transmission (SMP-S1,2) for any number of users despite the advantages of transmitting with only one antenna: receive diversity gain ($1\times2$ SIMO) and double transmit power for that antenna. As the number of users increases, dual stream transmission to two users (SMP-D2) is more likely than dual stream transmission to a single user (SMP-D1) because the probability that the two best spatial channels aim at the same user decreases with the number of users.

Figure 5. Probability of selection of each option for low transmit and receive spatial correlation.
MIMO configuration. The link level includes channel and these conditions. Not profitable for a reasonable number of users and that scheduling itself studied under various degrees of antenna reception emulation while the system level includes the system level simulations have been performed for a 2×2 downlink scheduling algorithm based on the P-FR. Link and the high correlation scenarios is shown in Figure 5.

Figure 5. Probability of selection of each option for high transmit and low receive spatial correlation.

Figure 6. shows that when the correlation at both ends of the communication link is high, the transmitter is more likely to select single stream transmission SMP-S1,2 than any other option. This holds true even in the realm of large number of users. Transitional behavior between the low correlation and the high correlation scenarios is shown in Figure 5.

Figure 6. Probability of selection of each option for high transmit and receive spatial correlation.

IV. CONCLUSIONS

This paper has studied the behavior of a proposed MIMO downlink scheduling algorithm based on the P-FR. Link and system level simulations have been performed for a 2×2 MIMO configuration. The link level includes channel and reception emulation while the system level includes the scheduling itself studied under various degrees of antenna allocation freedom. It is found that switching one antenna off is not profitable for a reasonable number of users and that independent antenna allocation achieves the higher gain in these conditions.

Furthermore, the probabilities of selection to transmit with each scheduling option for the full freedom scheme have been shown. These probabilities are very dependent on the spatial correlation and on the number of users in the cell.

REFERENCES