Coordinated Transmissions to Direct and Relayed Users in Wireless Cellular Systems

Thai, Chan T.D.; Popovski, Petar; Kaneko, Megumi; Carvalho, Elisabeth De

Published in:
I E E E International Conference on Communications

DOI (link to publication from Publisher):
10.1109/icc.2011.5962696

Publication date:
2011

Document Version
Accepted author manuscript, peer reviewed version

Link to publication from Aalborg University

Citation for published version (APA):
Abstract—The ideas of wireless network coding at the physical layer promise high throughput gains in wireless systems with relays and multi–way traffic flows. This gain can be ascribed to two principles: (1) joint transmission of multiple communication flows and (2) usage of a priori information to cancel the interference. In this paper we use these principles to devise new transmission schemes in wireless cellular systems that feature both users served directly by the base stations (direct users) and users served through relays (relayed users). We present four different schemes for coordinated transmission of uplink and downlink traffic in which one direct and one relayed user are served. These schemes are then used as building blocks in multi-user scenarios, where we present several schemes for scheduling pairs of users for coordinated transmissions. The optimal scheme involves exhaustive search of the best user pair in terms of overall rate. We propose several suboptimal scheduling schemes, which perform closely to the optimal scheme. The numerical results show a substantial increase in the system–level rate with respect to the systems with non–coordinated transmissions.

Index Terms—Cooperative communications, relaying, analog network coding, interference cancellation, a priori information.

I. INTRODUCTION

Recently there have been extensive studies on cooperative, relay–based transmission schemes for extending cellular coverage or increasing diversity. Several basic relaying transmission techniques have been introduced, such as amplify-and-forward (AF) [3], decode-and-forward (DF) [4] and compress-and-forward (CF) [5]. These transmission techniques have been applied in one-, two- or multi-way relaying scenarios.

In particular, two–way relaying scenarios [1], [2], [6] have attracted a lot of attention, since it has been demonstrated that in these scenarios one can apply techniques based on network coding in order to obtain a significant throughput gain. There are two basic principles used in designing throughput–efficient schemes with wireless network coding:

1) Aggregation of communication flows: instead of transmitting each flow independently, the principle of network coding is used in which flows are sent/processed jointly;

2) Intentional cancellable interference: in analog network coding, flows are allowed to interfere, knowing a priori that the interference can be cancelled by the destination.

The motivation for this work was to generalize the two basic principles from above and devise novel transmission schemes in multi-user scenarios. We consider scenarios based on cellular networks with relays, where direct and relayed users are served in uplink/downlink.

Assume for example that a direct user wants to send a packet to the Base Station (BS), while the BS has a packet to send to a relayed user. In a conventional cellular system, these packets are sent over separate UL/DL phases. Instead, the BS may first send the packet which is received at the Relay Station (RS). While the RS forwards this packet to its intended relayed user, the direct user sends its packet to the BS, thus saving the required transmission time compared to the conventional method. We term such a scheme coordinated direct/relay (CDR) transmission scheme. Transmission schemes that are related to some of the schemes proposed in this paper have appeared before in the literature [8], [9], or to relayed users [7]. However, in this paper we have used the principles described above to generalize the transmission schemes to in total four transmission schemes, which represent a superset of the existing schemes. As in the case of wireless network coding, our schemes take advantage of the combining of uplink and downlink traffic flows. Furthermore, we consider multi–user (> 2) scenarios, in which the proposed CDR schemes are used as building blocks for creating novel scheduling schemes. We consider the rate–optimal scheme, which requires exhaustive search across the pairs of users and is complex. Therefore, we propose several suboptimal schemes and the results show that they perform closely to the optimal one, while all the proposed schemes show significant rate gains with respect to the reference system in which the CDR schemes are not employed.

This paper is organized as follows. Section II describes the system model for two-user and multi-user networks. The two-user and multi-user schemes are described and analyzed in Section III and IV, respectively. Section V presents the numerical results and Section VI concludes the paper.

II. SYSTEM MODEL

The basic setup for a CDR scheme is the scenario with one base station (BS), one relay (RS), and two users (MS1 and MS2), see Fig. 1. All transmissions have a unit power and normalized bandwidth of 1 Hz. Each of the complex channels $h_i, i \in \{1, 2, 3, 4, 5\}$, is reciprocal, known at the receiver and Rayleigh–faded with parameter $\sigma = 1/\sqrt{2}$. We use the following notation, with a slight abuse: $x_i$ may denote...
a packet or a single symbol, and it will be clear from the context. For example, the packet that BS wants to send to the MS1 is denoted by $x_1$; but if we want to express the signal received, then we use expressions of type $y = h x_1 + z$, where all variables denote symbols (received, sent, or noise). We introduce further notation: $x_4$ is the packet sent from BS to MS2, while the packets that BS needs to receive are $x_3$ from MS1 and $x_2$ from MS2. Note that the example on Fig. 1 does not show traffic patterns that involve $x_3$ and $x_4$.

The basic time unit is one time slot. A direct transmission takes one slot. One transmission through the relay takes also one slot: in the downlink, the first half of the slot is for the transmission BS-RS, while the second slot is for the transmission RS-MS. The uplink transmission is similar. Relaying with amplify-and-forward (AF) is used, and therefore the transmission BS-MS has the same duration with the transmission RS-MS (and vice versa in the uplink). The received signal and Additive White Gaussian Noise (AWGN) at BS, RS, MS1 and MS2 in time slot $j$ is denoted by $y_{ij}$ and $z_{ij} \sim CN(0, n)$, $i \in \{B, R, 1, 2\}, j \in \{1, 2\}$. The instantaneous Signal-to-Noise Ratio (SNR) for the $i$-th channel is $\gamma_i = |h_i|^2/n$ and its capacity is denoted as $C(\gamma_i) = \log_2(1 + \gamma_i)$. The direct channel BS-MS1 is assumed weak and MS1 relies only on the amplified/forwarded signal from RS in order to decode the signal from BS. At the RS, the received signal is scaled to comply with transmit constraint.

In the scenarios with more than two users, where scheduling also needs to be applied, there are $k$ relayed users and $k$ direct users. The transmissions are organized in sessions. In a session, each user has a packet for an uplink or downlink with probability of $p_u$, or $1 - p_u$, respectively. An example of traffic pattern is shown in Table I which lists the user numbers in 4 traffic types: direct downlink (DD), direct uplink (DU), relayed downlink (RD) and relayed uplink (RU). Direct users 1, 4, 5 request a downlink, direct users 2, 3 request an uplink, relayed users 6, 7, 9 request a downlink and relayed users 8, 10 request an uplink. One session consists of multiple frames, each frame consists of two slots in which a CDR transmission is performed. We slightly abuse the notation for the wireless channels by not explicitly indexing the channel with the particular user: $h_i$, $i \in \{1, 2, 3, 4, 5\}$ thus refers to a channel in the set currently considered for a certain coordination. All channels are assumed known at BS and constant during each frame, but vary independently from frame to frame.

III. SCHEDULING IN TWO-USER SCHEMES

We propose four types of two-user schemes, each combining user pairs (DU, RD), (DD, RU), (DD, RD) and (DU, RU) for which there are packets to be transmitted. These schemes are compared with reference ones in terms of sum-rate.

<table>
<thead>
<tr>
<th>Frame</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>DD</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>DU</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>RD</td>
<td>6</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>RU</td>
<td>8</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

TABLE I

WAITING LIST

Fig. 1. Reference scheme 1 (a) and coordinated scheme 1 (b).

A. Reference Schemes

In the reference schemes there are only orthogonal transmissions and no interference. Four reference schemes corresponding to four user pairs described above have the same time slot structure, only the order and direction of transmissions for the relayed user are different. The reference scheme is denoted $E_i$, $i \in \{1, 2, 3, 4\}$.

Fig. 1 (a) describes the transmissions in reference scheme 1. In the first half slot, BS transmits $x_1$ and RS receives $y_{R1} = h_1 x_1 + z_{R1}$, in the second half-slot, RS scales $y_{R1}$ with amplification factor $g_{E1} = h_1^2/n$ so that the transmit power is 1 and transmits $\sqrt{g_{E1}} y_{R1}$ and MS1 receives $y_{S1} = h_2 \sqrt{g_{E1}} y_{R1} + z_{S1} = h_2 \sqrt{g_{E1}} h_1 x_1 + h_2 \sqrt{g_{E1}} z_{R1} + z_{S1}$, in the second slot, BS receives $y_{B2} = h_3 x_2 + z_{B2}$.

B. Proposed Basic Coordinated Schemes

In CDR schemes, the three transmissions are scheduled as one transmission in one slot and two simultaneous transmissions in the other slot although the order and direction of the transmissions are different. The transmissions are arranged so that the interference is reduced or canceled. There are four basic coordinated schemes denoted as $S_i$, $i \in \{1, 2, 3, 4\}$.

Coordinated Scheme $S_1$ (Fig. 1b), BS transmits $x_1$ to RS in the first slot, RS receives $y_{R1} = h_1 x_1 + z_{R1}$. In the second slot, RS scales the received signal with the amplification factor $g_{S1} = h_2^2/n$ and transmits it. At the same time, MS2 transmits $x_2$. MS1 therefore receives signal $y_{S1} = h_2 \sqrt{g_{S1}} y_{R1} + h_4 x_2 + z_{S1} = h_2 \sqrt{g_{S1}} h_1 x_1 + h_2 \sqrt{g_{S1}} + z_{R1} + h_4 x_2 + z_{S1}$ and BS receives $y_{B2} = h_1 \sqrt{g_{S1}} y_{R1} + h_3 x_2 + z_{B2} = ...
h_1 \sqrt{g_{S1}} h_x x_1 + h_1 \sqrt{g_{S1}} z_{R1} + h_3 x_2 + z_{B2}. Since BS knows \( x_1 \) and the channels, it cancels the component in \( x_1 \) in \( y_{B2} \), gets \( y_{B2} = h_3 x_2 + h_1 \sqrt{g_{S1}} z_{R1} + z_{B2} \) and decodes \( x_2 \) with SNR
\[
\gamma_{S12} = \frac{|h_3|^2}{|h_1|^2 g_{S1} n + n} = \frac{|h_3|^2 (|h_1|^2 n + n^2)}{2|h_1|^2 n + n^2} = \frac{2}{2|h_1|^2 + 1},
\]
(2)
MS1 decodes \( x_1 \) treating \( x_2 \) as interference with SINR
\[
\gamma_{S11} = \frac{|h_3|^2 g_{S1}|h_2|^2}{|h_1|^2 g_{S1} n + n} = \frac{\gamma_{S12} |h_2|^2}{\gamma_{S12} + 1} \frac{\gamma_{S12} + 1}{\gamma_{S12} + 1}.
\]
(3)
The sum-rate is therefore \( C_{S1} = C(\gamma_{S11}) + C(\gamma_{S12}) \).

**Coordinated Scheme S2** (Fig. 2): MS1 transmits \( x_3 \) and BS transmits \( x_4 \) simultaneously in the first slot. BS receives \( y_{B1} = h_2 x_3 + h_1 x_4 + z_{R1} \) and MS2 receives \( y_{Z2} = h_2 x_3 + h_4 x_4 + z_{S2} \). In the second slot, BS scales the received signal with the amplification factor \( g_{S2} = |h_1|^2 + |h_2|^2 + |n| + n \) and transmits it. BS receives \( y_{B2} = h_1 \sqrt{g_{S2}} y_{R1} + z_{B2} \) and MS2 receives \( y_{Z2} = h_1 \sqrt{g_{S2}} y_{R1} + z_{S2} \). Since BS knows \( x_4 \) and the channels, it cancels \( x_3 \) component in \( y_{Z2} \), gets \( y_{B2} = h_1 \sqrt{g_{S2}} (h_2 x_3 + z_{R1}) + z_{S2} \) and decodes \( x_3 \). At MS2, \( y_{Z2} \) and \( y_{Z2} \) form a virtual 2-antenna received signal \( y = H x + z \), with \( y = [y_{Z2} y_{Z2}]^T \), where \( x = [x_4 x_3]^T \), \( z = [z_{S2} h_5 \sqrt{g_{S2}} r_{S1} + z_{S2}]^T \), and
\[
H = \begin{bmatrix} h_3 & h_4 \\ h_3 \sqrt{g_{S2}} h_5 & h_4 \sqrt{g_{S2}} h_5 \end{bmatrix}.
\]
(4)
We can apply MMSE receiver using (13) in the Appendix to have the sum-rate
\[
C_{S2} = C \left( \frac{\gamma_{S12} \gamma_{S12}}{2|h_1|^2 + 1} \right) + C_{S2},
\]
(5)
in which \( C_{S2} = C \left( \frac{\gamma_{S12} \gamma_{S12}}{2|h_1|^2 + 1} \right) \) and
\[
C_{S3} = C \left( \frac{\gamma_{S12} \gamma_{S12}}{2|h_1|^2 + 1} \right) + C_{S3}.
\]
(6)

**Coordinated Scheme S4** (Fig. 2): MS1 transmits \( x_3 \) and MS2 transmits \( x_2 \) in the first slot, BS transmits what it received in the second slot. The transmissions are
\[
y_{R1} = h_2 x_3 + h_5 x_4 + z_{R1}, \quad g_{S4} = \frac{1}{|h_2|^2 + |h_5|^2 + |n|^2}.
\]
(3)
y_{B1} = h_3 x_2 + z_{B1}, \( y_{B2} = h_1 \sqrt{g_{S4}} y_{R1} + z_{B2} \). BS decodes \( x_2 \) and \( x_3 \) from \( y_{B1} \) and \( y_{B2} \). Similar to the previous schemes, we have \( y = H x + z \), with \( y = [y_{B1} y_{B2}]^T \), where \( x = [x_2 x_3]^T \), \( z = [z_{B1} h_1 \sqrt{g_{S4}} r_{S1} + z_{B2}]^T \), and
\[
H = \begin{bmatrix} h_3 & h_4 \\ h_3 \sqrt{g_{S2}} h_5 & h_4 \sqrt{g_{S2}} h_5 \end{bmatrix}.
\]
(4)
We can apply MMSE receiver using (13) in the Appendix to have the sum-rate
\[
C_{S4} = C \left( \frac{\gamma_{S12} \gamma_{S12}}{2|h_1|^2 + 1} \right) + C_{S3},
\]
(7)
and
\[
C_{S4} = C \left( \frac{\gamma_{S12} \gamma_{S12}}{2|h_1|^2 + 1} \right) + C_{S3}.
\]
(8)

**C. Coordinated Schemes with User Priority**

In the reference schemes, information for the relayed and direct user is transmitted separately. On average, the BS-relayed user rate is lower than the BS-direct user rate because of AF relaying. However, in some of the coordinated schemes, the BS-relayed user rate may be higher than or approximately equal to the BS-direct user rate. In addition, with the same amount of resource (transmit power, time slots...) a coordinated scheme can increase the sum-rate by allocating more resource to the direct user than the amount which is used in basic coordinated schemes. In this section, we introduce a prioritizing factor denoted as \( \lambda, -1 \leq \lambda \leq 1 \). If \( \lambda > 0 \), the direct user has more priority than in the basic coordinated scheme and the relayed user has less priority and vice versa.

When \( \lambda > 0 \), the scheme begins with two time slots as in a basic coordinated scheme however the length of each time slot is \( 1 - \lambda \) symbol time instead of 1 symbol (as \( T_{S31} \) and \( T_{S31} \) in Fig. 3). The residual time with a length of \( 2\lambda \) is used for an additional transmission between BS and the direct user (\( T_{S31} \)). When \( \lambda < 0 \), the additional transmissions are from BS to RS and RS to the relayed user for a relayed downlink or in opposite order for a relayed uplink. If \( \lambda = 1 \) or \( \lambda = -1 \), the whole time is used for the direct user or the relayed user respectively, which is not considered.
The average rate for coordinated scheme $i$ with prioritizing factor $\lambda$ is

$$C_{\lambda i}^D = \begin{cases} (1 - \lambda)(C_{S1i} + C_{S12}) + 2\lambda C(\gamma_3) & \text{for } \lambda > 0 \\ (1 + \lambda)(C_{S1i} + C_{S12}) - \lambda C(\gamma_2) & \text{for } \lambda < 0 \end{cases}$$

with $\gamma_R = \frac{\gamma_1\gamma_2}{\gamma_1 + \gamma_2 + 1}$.

IV. SCHEDULING IN MULTI-USER SCHEMES

This part presents different ways of scheduling transmissions in a session. At the beginning of a session, BS receives requests for uplink/downlink and channel information from some users. It schedules appropriate transmissions in the first frame according to one of the schemes below. After that, it receives channel information and schedules for the new frame and so on until all requests in the session are fulfilled.

Note that one frame has two coordinated schemes, such that we use the notation $S_i, j$ to denote the fact that the frame contains the coordinated schemes $S_i$ and $S_j$.

A. Multi-user Reference Scheme

A frame in the multi-user reference scheme contains 4 time slots for 4 traffic types (DD, RD, DU, RU). The users of each traffic type are served according to first-in-first-out discipline. If there is not a packet corresponding to a slot, it is left empty. A slot for a relayed user is divided into two small slots: one for the transmission between BS and RS and one for the transmission between RS and the relayed user. In the example of table I, the first, second and third frames are (1, 2, 6, 8), (4, 3, 7, 10) and (5, $\emptyset$, 9, $\emptyset$) respectively. In slot $\emptyset$ there is no transmission.

B. Coordinated Schemes

Instead of transmitting packets of 4 traffic types separately in a frame, we can use the coordinated schemes described in part III-B to combine the transmissions. In order to always have 4 traffic types (DD, RD, DU, RU) in a frame, only two types of combining are possible. Those are $S1$ (DU, RD) and $S2$ (DD, RU), and $S3$, 4, which includes $S3$ (DD, RD) and $S4$ (DU, RU). The packets are thus transmitted frame by frame according to first-in-first-out discipline using one of these two combining types. If $S3$, 4 is used in the example in table I, the frames are $[(1, 6), (2, 8)], [(4, 7), (3, 10)], [(5, 9), \emptyset]$. In a frame, the transmissions of $S3$ or $S4$ be can performed first; and after that are those of the other.

C. Proposed Multi-user Schemes

As users experience different channel qualities within each frame, the achieved performance will highly depend on the chosen user combinations, in each scheme. With an exhaustive search among the users and sum-rate estimating for each case and both combining types $S1, 2$ and $S3, 4$, we can find a combination of four users and a combining type which has the highest sum-rate in the current frame.

The complexity for CDR with exhaustive search, however, is prohibitively high since we have to calculate the sum-rate for every permutation of the packets. It is necessary to propose some sub-optimal schemes which requires a lower complexity without a significant rate loss. Such a reduction of the search space would be, for example, if only combination of $S3$ and $S4$ in a frame is considered (thus, the $S3$, 4 combining type), without considering the possibility to use the schemes $S1$ and $S2$.

Another suboptimal scheme is Best Direct User CDR (BDCDR): first, the direct downlink user which has the best channel to BS (max($\gamma_1$)) is picked. The downlink relayed user which has the best combination with that direct user is chosen after that. Then we have a sub-optimal coordinated scheme of $S3$. We can do similarly for $S1, S2, S4$ and choose the higher of $S12$ and $S34$. In Best Relayed User CDR scheme (BRCDR), the relayed downlink user which has the best relayed channel to BS (max($\gamma_2$)) picked first. The steps after that are processed in a similar way to BDCDR.

V. NUMERICAL RESULTS

Computer simulations with network scenarios and parameters as presented in part II is conducted to illustrate the sum-rate for the reference and proposed schemes.

In case of two users, Fig. 4 shows the rate for the relayed user ($C_{i1}$, $i \in \{E, S1, S2, S3, S4\}$), and the direct user ($C_{i2}$) changing as a function of prioritizing factor in reference and coordinated schemes. The case when $\lambda = 0$ is correspondent.
to the basic reference and coordinated schemes. When $\lambda$ increases, the direct user become more prioritized and in the opposite direction the relayed user does. Prioritizing the direct user gives more improvement than prioritizing the relayed user in terms of sum–rate because of AF relaying communication between the relayed user and BS. $C_{S42}$ is the highest among $C_{S12}$ because as seen in (8), always higher than $C_{E2} = C(\gamma_3)$, that BS exploits the information in the second time slot beside the information it receives in the first time slot which is equal to $C(\gamma_3)$. $C_{S32}$, always lower than $C(\gamma_3)$, is the lowest among $C_{S12}$ because MS2 only receives information in time slot 2 which is interfered by the transmission of R5. On contrary, $C_{S31}$ is the highest among $C_{S11}$ since MS1 receives information from R5 without any interference in a full time slot compared to half slot in reference scheme. Because $\gamma_{S31} = \gamma_R$, when $\lambda < 0$, $C_{S31}$ does not change therefore decreasing a negative $\lambda$ does not bring any benefit but a decrease in the sum–rate. $C_{S11}$ is the lowest due to the interference from MS2 at MS1 over the inter-user channel.

In multi-user case with $k = 10$, Fig. 5 compares the sum–rate for different CDR schemes with $p_h = \frac{1}{2}, \lambda = 0$. The CDR scheme with exhaustive search has the highest sum–rate due to the optimal combination of transmissions in each frame. The BDCDR scheme with much lower complexity achieves a slightly lower sum–rate since the rate for all scheme increases with $\gamma_3$ which is optimal in a Best Direct scheme and the channels are independently distributed. CDR with only $S3, 4$ has much lower sum–rate since there is no choice among the users. The $S1, 2$ CDR even has a lower sum–rate than the reference scheme due to useless inter-user interference in $S1$.

VI. CONCLUSION

In this paper, we have proposed coordinated schemes for a network with a direct user, a relayed user, a base station and a relay station. These schemes are inspired by the principles used in physical–layer network coding: uplink and downlink flows are aggregated, while the interference is not avoided by orthogonalization, but rather through the usage of information that is known a priori. The proposed schemes are shown to have higher sum–rate than the conventional schemes. These schemes are then used as building blocks in multi–user scenarios, where we present several schemes for scheduling pairs of users for coordinated transmissions. In order to avoid scheduling complexity, we propose several suboptimal scheme, which perform closely to the optimal scheme. We have also discussed the trade–off between the network sum–rate and user prioritization. As a future work, we intend to analyze the proposed schemes in frameworks with proportional fair scheduling and propose related schemes in multi–channel systems, such as OFDMA.

APPENDIX

Consider a 2x2 MIMO system with transmit vector $x = [x_1 \ x_2]^T$, receive vector $y$ and channel matrix $H$. We have transmission vector equation $y = Hx + z = h_1 x_1 + h_2 x_2 + z$. Denote $H = \begin{bmatrix} h_1 & h_2 \end{bmatrix}$

$$\Gamma = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{12} & \gamma_{22} \end{bmatrix} = \frac{1}{n} \begin{bmatrix} |h_{11}|^2 & |h_{21}|^2 \\ |h_{12}|^2 & |h_{22}|^2 \end{bmatrix},$$

$$\gamma_a = \frac{|h_{11} h_{21} + h_{12} h_{22}|^2}{n^2}, \quad \gamma_b = \frac{|h_{11} h_{22} - h_{21} h_{12}|^2}{n^2},$$

$$N = E[zz^H] = n \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix}.$$  

By noise whitening [10], we have SINR matrix $\text{SINR} = h_1^H K_2^{-1} h_1$ in which SINR for the first signal stream is

$$\text{SINR}_1 = \frac{\alpha \gamma_{11} + \gamma_{12} + \gamma_b}{\gamma_{21} + \gamma_{22} + \alpha}.$$  

In case the noise power matrix is in another form $E[zz^H] = n \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix}$, we have

$$\text{SINR}_1 = \frac{\gamma_{11} + \alpha \gamma_{12} + \gamma_b}{\gamma_{21} + \alpha \gamma_{22} + \alpha}.$$  

ACKNOWLEDGMENT

This work is supported by the Danish Research Council for Technology and Production, grant nr. 09 – 065035.

REFERENCES


