Multi-flow scheduling for coordinated direct and relayed users in cellular systems

Thai, Chan; Popovski, Petar; Kaneko, Megumi; De Carvalho, Elisabeth

Published in:
IEEE Transactions on Communications

DOI (link to publication from Publisher):
10.1109/TCOMM.2012.120512.110669

Publication date:
2013

Document Version
Early version, also known as pre-print

Link to publication from Aalborg University

Citation for published version (APA):
Multi-Flow Scheduling for Coordinated Direct and Relayed Users in Cellular Systems

Chan Dai Truyen Thai, Petar Popovski, Senior Member, IEEE, Megumi Kaneko, Member, IEEE, and Elisabeth de Carvalho, Member, IEEE

Abstract—There are two basic principles used in wireless network coding to design throughput-efficient schemes: (1) aggregation of communication flows and (2) interference is embraced and subsequently cancelled or mitigated. These principles inspire design of many novel multi-flow transmission (MFT) schemes. Such are the Coordinated Direct/Relay (CDR) schemes, where each basic transmission involves two flows to a direct and a relayed user. Usage of MFT schemes as building blocks of more complex transmission schemes essentially changes the problem of scheduling, since some of the flows to be scheduled are coupled in a signal domain and they need to be assigned a communication resource simultaneously. In this paper we define a novel framework that can be used to analyze MFT schemes and assess the system-level gains. The framework is based on cellular wireless users with two-way traffic and it sets the basis for devising composite time-multiplexed MFT schemes, tailored to particular optimization criteria. Those criteria can be formulated by adapting well-known schedulers in order to incorporate MFT schemes. The results show rate advantages brought by the CDR schemes in pertinent scenarios. Another key contribution is the proposed framework, which can be used to evaluate any future multi-flow transmission scheme.

Index Terms—Wireless relay, wireless network coding, interference cancellation, coordinated transmission.

I. INTRODUCTION

A. Motivation

Wireless network coding has recently emerged as one of the key generic techniques that can boost the throughput performance of wireless networks. A canonical scenario that demonstrates the benefit of wireless Network Coding (NC) is the scenario with Two-Way Relaying (TWR). 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) Embracing the interference that can be subsequently cancelled or mitigated. For example, in analog network coding, flows are allowed to interfere, knowing a priori that the interference can be cancelled by the destination.

Using these two principles can give rise to novel transmission techniques. In [1], [2] we have shown that the communication flows of a direct and a relayed user can be jointly served, which can bring very visible performance benefits with respect to the reference (conventional) way of serving the same communication flows. The four schemes discussed in [2] are described on Fig. 1. Assume for example that a direct user wants to send a packet to Base Station (BS) B, 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, as seen in scheme $S_{du}$ of Fig. 1 the BS may first send the packet which is received at Relay Station (RS) L. 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.

Despite the fact that each individual scheme from [2] brings throughput benefit with respect to the related reference schemes, the system-level aspects of these novel transmission...
schemes remain largely unclear. In [2] there was a preliminary study on how the individual CDR schemes can be used as building blocks for scheduling schemes to serve multiple users. However, the multi–user scenario in [2] is rather limited and cannot provide satisfactory answers in comparing the CDR schemes with other state-of-the-art schemes, notably the two-way relaying [3].

One of the main objectives of this paper is to define a suitable framework for analyzing and comparing the multi–flow transmission (MFT) schemes, such as CDR and two-way relaying. Rigorously speaking, any time–division scheme in which e.g. the communication flows of multiple users are served in the downlink, is a MFT scheme. However, here we use the term MFT scheme to denote a transmission scheme in which the multiple flows are essentially coupled in the signal domain, they use the wireless medium simultaneously, and cannot be decoupled via time division. This is true, for example, for a two–way relaying scheme with wireless network coding based on amplify–and–forward [4]. That is an MFT scheme and the signal–domain coupling is seen in the fact that the achievable rate region is not the triangle obtained by time–division between two one–way relaying schemes.

In attempting to define the suitable analytical framework, we need to change the scheduling task from its usual definition (“at a given time, allocate the single communication resource to a certain communication flow “) to a definition that can deal with MFT schemes (“at a given time, allocate the single communication resource to a group of communication flows”). With such prerequisites, the problem at hand can be defined as follows. Suppose that \( M \) flows should be served in a given scheduling epoch. In that case the performance of any proposed algorithm to serve these flows can be assessed by an \( M \)-dimensional achievable rate region. The central question that we will address is: how much can we enlarge the region of achievable rates if the algorithm that serves the flows can leverage on MFT schemes? In particular, we investigate the benefits of applying the CDR schemes.

The work in this paper treats the case in which any relaying operation is conducted by using amplify–and–forward. Using other relaying techniques may change the analysis and the conclusions, but, the analysis framework and the problem at hand, sublimed in the central question above, remains the same. This framework consists of scenarios and methods for combining MFT schemes into a single composite scheme that serves multiple flows. Another contribution is that we apply insights from practice in order to put constraints on the communication flows and restrict the analysis of the achievable rate regions to tractable, two-dimensional sub-regions. Finally, we consider multi–user scenarios for which we show how several canonical schedulers (Round Robin, Maximum Sum–Rate ( Opportunistic Scheduler), Proportional Fair scheduler) can be formulated and applied when MFT schemes are used as building blocks for transmission.

Surveying the related work and regarding the combination of the transmissions of two users, transmission schemes that are somewhat related to the schemes treated in this paper have appeared before [5], [6], or to relayed users [7], [8]. In another aspect, regarding the joint resource allocation of uplink and downlink of a user, [9], [10] and references therein discuss that the total amount of resource for uplink and downlink can be dynamically adjusted and the uplink and downlink satisfy the user as in a common service. In this paper, we jointly allocate uplink/downlink resources to different user types and analyze the rate performance under different scheduling policies. On the other hand, several advanced versions of TWR NC have been proposed e.g. the optimization of the different phase durations in TWR NC is considered in [12] and the space distribution of users is exploited to optimize the TWR NC scheme in [11]. Therefore in this paper, we treat TWR NC as a building block that represents the state-of-the-art.

The paper is organized as follows. Section II presents the network model of the paper. We present the individual schemes in section III. Section IV describes the composite schemes and presents the framework for analyzing the achievable rate region. Section V compares the reference schemes and CDR schemes using different schedulers while Section VI presents and analyzes the numerical results. Section VII concludes the paper.

II. SYSTEM MODEL

We first introduce the relevant notation and system concepts by considering a cellular network with one base station (B), one relay (L), one relayed user (R) and one direct user (D) see Fig. 1. The direct channel BS–R is assumed weak and R relies only on the amplified/forwarded signal from the RS in order to decode the signal from the BS. All transmissions are in one frequency with a normalized bandwidth of 1 Hz. All stations are single-antenna and half-duplexed. Each of the complex channels \( h_{ij} \), \( i, j \in \{B, L, R, D\} \), is reciprocal, known at the receiver. We assume all the channels are known at the BS as in [14], [15]. Each user requests an uplink and a downlink transmissions to the BS. We assume that the data to/from each user is infinitely backlogged so that there are always data to transmit as in many works regarding downlink [16] and Two-way Relaying optimization [17] and scheduling [18], [19]. Thus the achievable rate for a user at a certain time is equal to the information theoretic capacity, i.e. \( C(\gamma) = \log_2(1 + \gamma) \), where \( \gamma \) denotes an instantaneous received signal-to-noise ratio (SNR) of the channel used. Therefore the maximal rate received at a station over channel \( l \) is \( C(\gamma) \).

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 the BS wants to send to user R 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 \( y, x_1 \) and \( z \) denote symbols (received, sent and noise respectively). We introduce further notations: \( x_4 \) is the packet sent from the BS to D, while the packets that the BS needs to receive are \( x_3 \) from user R and \( x_2 \) from D. \( x_1, x_2, x_3 \) and \( x_4 \) are therefore corresponding to 4 traffic types: relayed downlink (Rd), direct uplink (Du), relayed uplink (Ru) and direct downlink (Dd) respectively. Throughout this paper, small \( u \) and \( d \) denote uplink and downlink while capitalized \( D \) and \( R \) denote the direct and relayed users respectively. All relaying transmissions in this paper are Amplify-and-Forward (AF).

The transmissions are organized in scheduling epochs with fixed duration. The channels are constant throughout the
scheduling epoch. In this paper we will assume that in each
epoch two users will be served, each having uplink/downlink
traffic, which corresponds to four flows per epoch. If there
are more than two users in the system, then two of them
are selected according to a certain scheduling criterion and,
again only four flows are served in an epoch. During a
scheduling epoch several different transmission schemes can
be multiplexed in time, including both MFT schemes and
single–flow transmission (SFT) schemes. The part of the
epoch during which a fixed transmission scheme is used
is termed frame. While the duration of a scheduling epoch
is fixed, the duration of each frame within the scheduling
epoch is variable and subject to optimization. The transmission
schemes corresponding to different frames will be introduced in
the next section.

In an individual scheme, the received signal and Additive
White Gaussian Noise (AWGN) at the BS, the RS, user R
and user D in time slot \( j \) is denoted by \( y_{il}[j] \) and \( z_{il}[j] \sim
\mathcal{CN}(0, \sigma^2) \), \( i \in \{B, L, R, D\}, j \in \{1, 2\} \). The average
transmit power at all stations is 1, \( \mathbb{E}[\|s\|^2] = 1 \), \( s \) is a symbol
when transmission is done at the BS, user R or D, \( s \) is a relayed
signal when transmission is done at the RS. At the RS, the
received signal is scaled to comply with transmit constraint.

Regarding the notation, for a compact notation, we will
write the matrices by using ";" to separate different rows of a
matrix. For example, the \( 2 \times n \) matrix with \( 1 \times n \) row vectors
\( \mathbf{a}_1, \mathbf{a}_2 \) can be written as \( [\mathbf{a}_1; \mathbf{a}_2] \) instead of
\( \mathbf{a}_1 \mathbf{a}_2 \).

III. TRANSMISSION SCHEMES: THE BUILDING BLOCKS

In this section we introduce the transmission schemes,
where each scheme is a candidate to be used during a frame
that is a part of a certain scheduling epoch. The candidate
schemes are of two types, SFT schemes and MFT schemes.
We first describe the four MFT schemes based on coordinated
direct and relay (CDR) transmission. The other schemes that
can be used in a frame feature conventional one–way direct
and relay transmission, as well as the two–way–relaying.
It should be noted that different transmission schemes corre-
spond to different set of communication flows. For example,
the first scheme, denoted \( S_{1u} \) on Fig. 1 serves two flows, the
BS to user R and user D to the BS, respectively. Another
scheme would be an SFT scheme in which only the flow user
D to the BS is served - hence, the set of flows in this latter
case is a subset, but not identical to the set of flows used in
\( S_{1u} \).

A. Multi–Flow Transmission Schemes with CDR

Each of schemes \( S_{1u}, S_{ud}, S_{dd}, S_{uu} \) combines two
communication flows, one associated with a direct and another
with a relayed user, respectively. There are four possible
flows associated with these two users, which we have already
denoted as Dd, Du, Rd, Ru. Each of the four schemes is
an MFT scheme that has a duration of \( 2N \) symbols. \( S_{1u} \)
combines Du and Rd, subscript \( du \) means that the relayed user
(the first user) has a downlink message and the direct user
(the second user) has an uplink message. \( S_{ud} \) combines Dd and
Ru, \( S_{dd} \) combines Dd and Rd and \( S_{uu} \) combines Du and Ru.
The time interval of \( 2N \) used by a given scheme is divided
into two time slots, each having \( N \) symbols. In each slot,
one single transmission or two simultaneous transmissions are
performed. The transmissions in each scheme are arranged so
that the interference is reduced or cancelled. We present each
scheme in details below.

1) Coordinated Scheme \( S_{1u} \): BS transmits \( x_1 \) to the RS
in the first slot, the RS receives \( y_{il}[1] = h_{BL}x_1 + z_{il}[1] \)
(Fig. 1). In the second slot, the RS scales the received
signal with the amplification factor \( \alpha_{S_{1u}} = \frac{1}{\sqrt{h_{BL}^2 + \sigma^2}} \) and
transmits it. At the same time, D transmits \( x_2 \). User R
therefore receives signal
\[
y_{R}[2] = h_{LR}\sqrt{\alpha_{S_{1u}} y_{il}[1]} + h_{RD}x_2 + z_{R}[2] = h_{LR}\sqrt{\alpha_{S_{1u}} h_{BL}x_1} + h_{RD}\sqrt{h_{BL}^2 + \sigma^2} + z_{R}[2]
\]
and the RS receives
\[
y_{B}[2] = h_{BL}\sqrt{\alpha_{S_{1u}} y_{il}[1]} + h_{BD}x_2 + z_{B}[2] = h_{BL}\sqrt{\alpha_{S_{1u}} h_{BL}x_1} + h_{BD}\sqrt{h_{BL}^2 + \sigma^2} + z_{B}[2]
\].
Since the BS knows \( x_1 \) and all channels, it cancels
the component in \( x_1 \) in \( y_{B}[2] \), gets \( \tilde{y}_{B}[2] = h_{BD}x_2 + h_{BL}\sqrt{\alpha_{S_{1u}} z_{B}[1]} + z_{B}[2] \) and decides \( x_2 \) with
\[
\text{SNR}_{\gamma_{D}} = \frac{h_{BD}^2 (h_{BL}^2 + \sigma^2)}{2h_{BL}\sqrt{h_{BL}^2 + \sigma^2} + \sigma^2} \quad x_{2} = \frac{\gamma_{D} y_{B}}{h_{BL}\sqrt{\alpha_{S_{1u}}[h_{BL}^2 + \sigma^2]}} + \frac{\gamma_{D} z_{B}[1]}{h_{BL}\sqrt{\alpha_{S_{1u}}[h_{BL}^2 + \sigma^2]}}
\]
the RS transmits what it received in the first slot (Fig. 2). The BS transmits $y_R(x_3)$, and user D transmits $x_2$ in the second slot. In the second slot, the RS transmits what it received in the first slot (Fig. 1). The transmissions are $y_d[1] = h_{LR}x_3 + h_{LD}x_2 + z_d[1]$, $\alpha_{S_{uo}} = \gamma_{BL}y_d[1] + y_R[1] = h_{BD}x_2 + z_B[1]$, $y_B[2] = h_{BL}\alpha_{S_{uo}}y_d[1] + z_B[2]$. The BS decodes $y_d$ and $x_2$ from $y_B[1]$ and $y_B[2]$. To perform this, we have the same maximal achievable rate $R^{S_{uo}} = \frac{1}{2} C \left( \frac{\gamma_{BL}y_d[1] + z_d[1]}{\gamma_{BL} + \gamma_{LD} + 1} \right)$ and $R^{S_{uo}} = \frac{1}{2} C \left( \frac{\gamma_{BD}y_B[1] + z_B[1]}{\gamma_{BD} + \gamma_{LD} + 1} \right)$.

4) Coordinated Scheme $S_{uo}$: User R transmits $x_3$ and user D transmits $x_2$ in the first slot. In the second slot, the RS transmits what it received in the first slot (Fig. 1). The transmissions are $y_d[1] = h_{LR}x_3 + h_{LD}x_2 + z_d[1]$, $\alpha_{S_{uo}} = \gamma_{BL}y_d[1] + y_R[1] = h_{BD}x_2 + z_B[1]$, $y_B[2] = h_{BL}\alpha_{S_{uo}}y_d[1] + z_B[2]$. The BS decodes $y_d$ and $x_2$ from $y_B[1]$ and $y_B[2]$. Similar to the previous schemes, we have $y = Hx + z$, with $y = [y_d[1] \ y_B[2]]^T$, $x = [x_2 \ x_3]^T$, $z = [z_d[1] \ z_B[1]]^T$, and $H = \frac{h_{BD}}{\sqrt{\alpha_{S_{uo}}}x_3 + \sqrt{\alpha_{S_{uo}}}x_2}$. We can apply MMSE receiver for both users to have the sum-rate. We have the rates $R^{S_{uo}} = \frac{1}{2} C \left( \frac{\gamma_{BL}y_d[1] + z_d[1]}{\gamma_{BL} + \gamma_{LD} + 1} \right)$ and $R^{S_{uo}} = \frac{1}{2} C \left( \frac{\gamma_{BD}y_B[1] + z_B[1]}{\gamma_{BD} + \gamma_{LD} + 1} \right)$.

B. Reference Transmission Schemes

In this part we describe other transmission schemes that can be used to build a composite reference scheme. The motivation comes from the following: If a designer is not aware about the CDR schemes receives the task to serve $M$ communication flows in a given epoch, which schemes are at his/her disposal? Clearly, the first candidates are the usual single–flow schemes, which we, for convenience, denote as $S_{od}$, $S_{uo}$, $S_{do}$ and $S_{uo}$. Scheme $S_{od}$ is a one-hop transmission from the BS to user D (Fig. 2). The BS transmits $x_2$ (Du) and user D receives $y_d = h_{BD}x_2 + z_d$ and decodes $x_2$ with the maximal achievable rate of $R^{S_{od}} = C(\gamma_{BD})$. In scheme $S_{uo}$, user D transmits $x_2$ (Du) and the BS receives. Similarly, we have the same maximal achievable rate of $R^{S_{uo}} = C(\gamma_{BD})$. In scheme $S_{do}$, first the BS transmits $x_1$ (Rd), the RS receives $y_d = h_{BL}x_1 + z_L$, amplifies with amplification factor $\alpha_{S_{do}} = \frac{h_{BL}}{\sqrt{\gamma_{BL}}}$ and transmits. User R receives and decodes $x_3$ with the maximal achievable rate of $R^{S_{do}} = C \left( \frac{\gamma_{BL}y_d[1] + z_d[1]}{\gamma_{BL} + \gamma_{LD} + 1} \right)$. The transmission in the opposite direction (Ru) is conducted in scheme $S_{uo}$. We have the same maximal achievable rate $R^{S_{uo}} = R^{S_{uo}}$.

As a part of the reference transmission schemes, we consider the $TWR$ NC based on AF. In accordance with the other schemes, this will be denoted $S_{TWR}$ and it consists of the transmission from the BS to user R (Rd) and the transmission from user R to the BS (Ru) using TWR NC (Fig. 2). First, in the Multiple Access phase, the BS and user R simultaneously transmit $x_1$ (Rd) and $x_3$ (Ru) respectively in $N$ symbols, the RS receives $y_d = h_{BL}x_1 + h_{LR}x_3 + z_L$, amplifies it with amplification factor $\alpha_{S_{TWR}} = \frac{1}{\sqrt{\gamma_{BL} + \gamma_{LD} + 1}}$ and transmits it in the Broadcast phase in $N$ symbols. User R receives, cancels the contribution of $x_1$ and decodes $x_1$ with SINR $R^{S_{TWR}} = \frac{\gamma_{BL}y_d[1] + z_d[1]}{\gamma_{BL} + \gamma_{LD} + 1}$. The BS receives, cancels the contribution of $x_1$ and decodes $x_3$ with SINR $\gamma_{TWR} = \frac{\gamma_{BL}y_d[1] + z_d[1]}{\gamma_{BL} + \gamma_{LD} + 1}$. We have the rates corresponding to the relayed uplink and downlink $R^{S_{TWR}} = \frac{NC(\gamma_{TWR})}{2N} = \frac{1}{2} C (\gamma_{TWR})$ and $R^{S_{TWR}} = \frac{1}{2} C (\gamma_{TWR})$.
vector \( \mathbf{r} = [R_{Dd}, R_{Du}, R_{Rd}, R_{Ru}]^T \) corresponding to the four communication flows \( Dd, Du, Rd, \) and \( Ru, \) respectively. We will use the notation \( R_i \) and \( R_{Dd} \) interchangeably, also for the other three flows. The total rate achieved by a given flow in the scheduling epoch consists of contributions from each of the individual schemes that serves that flow, e.g. \( R_{Du} \) has contributions from \( S_{du}, S_{uu}, S_{ou}. \) For a compact representation, denote \( \mathbf{A} \) as in (3) in which the element at row \( i \) and column \( j \) is the rate contribution of the individual scheme \( S_{ij}, j \in \{du, ud, dd, uu, od, do, ou, do,uo, \) TWR \} which is equivalent to column \( 1, 2, ..., 9 \) respectively, to the \( i-\text{th} \) flow \( (R_i, i \in \{1,...,n\}). \) Hence, the four-dimensional region of the rates that can be achieved by the composite scheme is described as \( 0 \leq \mathbf{r} \leq \mathbf{A} \mathbf{\theta} \) where \( \mathbf{0}_p \) or \( \mathbf{1}_p \) is the column vector \([0,...,0]^T\) or \([1,...,1]^T\) with \( p \) elements. The operator \( \leq \) implies that for each vector component \( \leq \) is satisfied. The region is obtained by varying \( \mathbf{\theta} \) such that \( \mathbf{0}_p \leq \mathbf{\theta} \leq \mathbf{1}_p \) and \( \mathbf{1}_p^T \mathbf{\theta} = 1. \)

The reference combined scheme \( \mathbf{S}_{ref} \) consisting of only reference individual schemes is a special case of \( \mathbf{S}_{all} \) when \( \theta_{du} = \theta_{ud} = \theta_{dd} = \theta_{uu}. \) Similarly, we have CDR composite schemes \( \mathbf{S}_{duud} \) when all \( \theta_i = 0 \) except \( \theta_{du}, \theta_{ud} \) and \( \mathbf{S}_{duud} \) when all \( \theta_i = 0 \) except \( \theta_{ud}, \theta_{uu}. \) While \( \mathbf{S}_{du} \) is the multiplexing of all defined reference and CDR schemes, \( \mathbf{S}_{CDR} \) is the multiplexing of all CDR schemes and \( \mathbf{S}_{ref} \) is the multiplexing of all reference schemes. In the following parts, we will analyze and compare these composite schemes; the gain by the CDR schemes is seen in the enlarged rate region offered by \( \mathbf{S}_{all} \) compared to \( \mathbf{S}_{ref} \).

### B. Rate Analysis with Uplink/Downlink Coefficient

A comprehensive analysis of the achievable rate region for the composite transmission scheme involves consideration of four-dimensional regions, which is not always tractable or sufficiently informative. In order to collapse the rate region to subregions of practical interest, we resort to the features of the communication flows served in an epoch. By assumption, each of the users (direct or relayed) scheduled in an epoch, has both uplink and downlink traffic. In a conventional approach, the uplink demand and downlink demand of a user in a wireless network consisting of a base station (or an access-point) and multiple users (mobiles) are independently allocated in different resource portions (TDD, FDD or CDMA). This happens despite the fact that the uplink and the downlink traffic of a given user can exhibit dependence that emerges from e.g. the application-layer behavior and needs. For example, the downlink/uplink ratio is often fixed and depends on the type of application e.g. gaming and calls have ratio of 1:1, web browsing has a ratio of about 5:1 [13]. Generally, we can assume that the downlink rate demand and the uplink rate demand of a user satisfies, as in [10], \( R_i^d = \beta_i R_i^u, \) \( i \in \{D,R\}, \) where \( \beta_i \) is termed uplink/downlink ratio (UDC) for the \( i-\text{th} \) user.

We assume that the traffic demands by the \( i-\text{th} \) user are posed by specifying the UDC, defined as follows. In a certain scheduling epoch, all uplink/downlink flows of users are served and the rates are selected such that UDC is achieved for each of the served users. More formally, let \( (R_{D,d,i}(n), R_{R,i}(n)) \) be the downlink/uplink rates achieved for the \( i-\text{th} \) user in \( n-\text{th} \) scheduling epoch. Then with any \( n, \) \( R_{B,R,i}(n) \) is \( \beta_i. \)

In the following we analyze the rate region under the UDC constraints for the composite scheme \( \mathbf{S}_{all}, \) let us represent the matrix \( \mathbf{A} \) as \( \mathbf{A} = [\mathbf{a}_1^T; \mathbf{a}_2^T; \mathbf{a}_3^T; \mathbf{a}_4^T] \) where e.g. \( \mathbf{a}_p^T \) is the \( p \)-th row of \( \mathbf{A} \) in (3). We are now interested in getting the bound for the two-dimensional vector that contains only the downlink rates \( r_d = [R_{Dd}, R_{Rd}]^T. \) Let the corresponding vector of uplink rates be \( \mathbf{r}_u = [R_{Du}, R_{Ru}]^T, \) such that \( r_d = \beta D \mathbf{r}_u. \)

Considering the direct user, if \( \mathbf{a}_p^T \mathbf{\theta} \leq \beta D \mathbf{a}_p^T \mathbf{\theta} \), we select \( R_{Dd} = \mathbf{a}_p^T \mathbf{\theta} \) and \( R_{Du} = \frac{\mathbf{a}_p^T \mathbf{\theta}}{\beta D} \). If \( \mathbf{a}_p^T \mathbf{\theta} > \beta D \mathbf{a}_p^T \mathbf{\theta} \), we select \( R_{Dd} = \beta D \mathbf{a}_p^T \mathbf{\theta} \) and \( R_{Du} = \mathbf{a}_p^T \mathbf{\theta} \). Thus \( R_{Dd} \) is \( \min \{\mathbf{a}_p^T \mathbf{\theta}, \beta D \mathbf{a}_p^T \mathbf{\theta}\} \). \( R_{Ru} \) is derived similarly. Now the two-dimensional region is determined by:

\[
0 \leq r_d \leq \left[ \min \{\mathbf{a}_1^T \mathbf{\theta}, \beta D \mathbf{a}_2^T \mathbf{\theta}\}, \min \{\mathbf{a}_3^T \mathbf{\theta}, \beta D \mathbf{a}_4^T \mathbf{\theta}\} \right]^T.
\]

(4)

To see how the rate region looks like we consider the simple case of \( S_{duud}. \) In this case, \( \theta_{du} = \theta_{ud} = \theta_{dd} = \theta_{uu} = 0, \) and \( \theta_{dud} = 1 - \theta_{uu}. \) Replacing \( \mathbf{\theta} \) with \( \theta_{du}, 1 - \theta_{du}, 0, 0, 0, 0, 0, 0, 0, 0, 0 \) and \( \mathbf{A} \) in (3) to (4), we have

\[
0 \leq r_d \leq \left[ \min \{R_{Dd}^u (1 - \theta_{du}), \beta D R_{Ru} S_{duu} (1 - \theta_{du}) \}, \right.
\]

\[
\min \{R_{Dd}^u (1 - \theta_{du}), \beta D R_{Ru} S_{duu} (1 - \theta_{du}) \}^T.
\]

(5)

In each min function in the right hand side, there are two elements thus there are totally 4 cases to consider. Denoting
m_1 = \beta_D R_{Du}^{S_{du}} and m_2 = \beta_R R_{Dr}^{S_{du}} we have.

- The first case is when \( R_{Du}^{S_{du}}(1 - \theta_{du}) \leq \beta_D R_{Du}^{S_{du}} \theta_{du} \) and \( R_{Dr}^{S_{du}} \theta_{du} \leq \beta_R R_{Dr}^{S_{du}}(1 - \theta_{du}) \). This is equivalent to \( \frac{1}{m_1 + 1} \leq \theta_{du} \leq \frac{m_2}{m_2 + 1} \). The rate region is now determined by \( 0 \leq r_d \leq \left[ R_{Du}^{S_{du}}(1 - \theta_{du}), R_{Dr}^{S_{du}} \theta_{du} \right]^T \). If \( \theta_{du} = \frac{1}{m_1 + 1} \) and the line segment \( O_a \) is determined.

- The second case is when \( R_{Du}^{S_{du}}(1 - \theta_{du}) > \beta_D R_{Du}^{S_{du}} \theta_{du} \) and \( R_{Dr}^{S_{du}} \theta_{du} \leq \beta_R R_{Dr}^{S_{du}}(1 - \theta_{du}) \). This is equivalent to \( 0 \leq \theta_{du} < \theta_a \) and the rate region is now determined by \( 0 \leq r_d \leq \left[ \beta_D R_{Du}^{S_{du}} \theta_{du}, R_{Dr}^{S_{du}} \theta_{du} \right]^T \). Both \( R_{Du} \) and \( R_{Rd} \) increases with \( \theta_{du} \) and the line segment \( OO_a \) is determined.

- The third case is when \( R_{Du}^{S_{du}}(1 - \theta_{du}) \leq \beta_D R_{Du}^{S_{du}} \theta_{du} \) and \( R_{Dr}^{S_{du}} \theta_{du} > \beta_R R_{Dr}^{S_{du}}(1 - \theta_{du}) \). This is equivalent to \( \theta_b < \theta_{du} \leq 1 \) and the rate region is now determined by \( 0 \leq r_d \leq \left[ R_{Du}^{S_{du}}(1 - \theta_{du}), \beta_R R_{Dr}^{S_{du}}(1 - \theta_{du}) \right]^T \). Both \( R_{Du} \) and \( R_{Rd} \) decreases with \( \theta_{du} \) and the line segment \( OO_b \) is determined.

- The forth case is when \( R_{Du}^{S_{du}}(1 - \theta_{du}) > \beta_D R_{Du}^{S_{du}} \theta_{du} \) and \( R_{Dr}^{S_{du}} \theta_{du} > \beta_R R_{Dr}^{S_{du}}(1 - \theta_{du}) \). This is equivalent to \( \frac{m_2}{m_2 + 1} \leq \theta_{du} < \frac{1}{m_1 + 1} \). If (6) is not satisfied, the diagonal upper bound of this rate region part is a line segment similar to case 1. If (6) is satisfied, there is not any valid value of \( \theta_{du} \).

If (6) is satisfied, the first three cases determine a triangle in which if we decrease \( R_{Rd} \) of point \( O_a \) correspondingly, we get the segment \( O_a R_{Du}, 1 \) and if we decrease \( R_{Du} \) of point \( O_b \) correspondingly, we get the segment \( O_b R_{Rd}, 2 \). The convex rate region is therefore formed as in Fig. 4. If (6) is not satisfied, the last three cases form a similar rate region.

The rate region of scheme \( S_{all} \) can be found by the algorithm described through the Pseudocode 1 in which \( k \) is the index vector of the outer points of all achievable rate pairs \((x, y)\).

### V. Multiple-User Scheduling

Considering a pair of one direct user and one relayed user with four traffic types (Dd, Du, Rd, and Ru), we have collapsed the four-dimensional rate region to two-dimensional rate region by introducing UDCs. A natural question is what are the multiple-user schedulers that need to be used? In the following we describe how three commonly used schedulers can be adapted to the scheduling task defined in this paper: Round Robin with Equal Rates (RR ER) (Opportunistic Scheduler), Maximum Sum-Rate (MSR) and Proportional Fairness (PF).

We will make an additional distinction among the multiuser schemes. The composite schemes \( S_{all}, S_{CDR}, S_{ref} \) will be termed multi-user-epoch (MUE) schemes since two users are served in each epoch. As a reference, we consider single-user-epoch (SUE) schemes, in which a single user is served in an epoch. Thus, SUE is a composite scheme which in a given epoch can be either multiplexing of \( S_{od}, S_{ou} \) (if the user is direct) or \( S_{do}, S_{uo}, S_{TWR} \), if the user is relayed.

#### A. Round Robin with Equal Rates

In order to emulate the original concept of round robin, here the two users (direct, relayed) are selected not based on the channel gains or achievable rates, but arbitrarily (e.g. by user ID). Without losing generality, we can say that in the first epoch the scheduler picks the first direct user and the first relayed user to make a pair and apply the composite CDR scheme. However, differently from a single-user scheduling scheme, here we have additional degrees of freedom also after the users are selected. These degrees of freedom are instantiated by changing the time-sharing vector \( \theta \). In order to further restrict the rate region, one possibility is to select \( \theta \) such that the downlink rates of the direct and the relayed user are equal. Recall that the uplink/downlink rates are related via UDC, such that the uplink rates of these users will have a ratio of \( \beta_D : \beta_R \). In the next epoch, that user pair is put aside and the second direct user and the second relayed user are selected. If there are \( n_D \) direct users and \( n_R \) relayed users, the number of such epochs is \( \min(n_D, n_R) \). After that the scheduler picks the next direct user (if \( n_D > n_R \)) or the next relayed user (if \( n_R > n_D \)) to serve as a single user in the next epoch, using an appropriate SUE. However, in order to make a fair time allocation, this single-user epoch is half-length of the
two-user epoch described before. In order words, two direct users or two relayed users are put into a two-user epoch strictly divided by 2. Consequently, the number of two-user epochs in a scenario is \( \lceil \frac{x}{2} \rceil \), in which \( \lceil x \rceil \) is the nearest integer \( \geq x \), regardless of how many direct users and relayed users there are.

For a certain user pair and considering the UDCs, we have the conditions \( R_{Dd} = R_{Dd}, R_{Dd} = \beta D(R_{du} + R_{dd}) \) and \( R_{Rd} = \beta RR_{Ru} \). In other words, we have to select \( \theta \) so that the scheme \( S_m, m \in \{ \text{all, CDR, ref} \ldots \} \) can provide the rates for \( Dd, Du, Rd \), and \( Ru \) with ratio \( 1 : \beta D : 1 : \frac{1}{\beta R} \). Considering the direct user and reasoning as in part IV-B, we select \( R_{Dd} = \min \{ a^T_1 \theta, \beta_D a^T_2 \theta \} \) and \( R_{Rd} = \min \{ a^T_1 \theta, \beta_R a^T_4 \theta \} \). Since \( R_{Dd} = R_{Rd} \), we select the maximum among different \( \theta \) as in (7) in which \( R_{(i,j)p}^S \) and \( a^T_{(i,j)q} \) are the rate of traffic type \( p, p \in \{ Dd, Du, Rd, Ru \} \) and the \( q \)-th row of the matrix \( \mathbf{A} \) corresponding to the user pair \((i, j)\) offered by scheme \( S_m \), respectively and \( \theta_m \) is the time segment ratio vector corresponding to scheme \( S_m \), e.g. \( S_{CDR} \) has a vector of \( \theta_{CDR} = [\theta_{dd} \theta_{dd} \theta_{dd} \theta_{uu} \theta_{uu} \theta_{uu} \theta_{uu} \theta_{uu}]^T \). This problem is to find an optimal vector \( \theta_m \) which gives the highest downlink rates for a selected user pair. The procedure continues until all users are served as described above. After all users pairs are served, either only direct or relayed users are left, which are served through SUE. The average downlink rate offered by CDR with RR ER can be worse than that of SUE with RR ER, since the equal rate constraint limits the rate of the stronger user. This will be examined and discussed in the section with Numerical Results.

B. Maximum Sum-Rate

The Maximum Sum-Rate scheduler always selects the pair with the highest sum-rate to serve considering UDCs. The achievable downlink rate of direct user \( i \) is \( \min \{ a^T_{(i,j)1} \theta, \beta_D a^T_{(i,j)2} \theta \} \) and the achievable downlink rate of a relayed user is \( \min \{ a^T_{(i,j)3} \theta, \beta_R a^T_{(i,j)4} \theta \} \). Therefore the user pair selected in an epoch favors the users with good channels and is described as in (8).

C. Proportional Fairness

In this part we use Proportional Fairness (PF) [18], [22]–[24] as a metric for selecting a user pair in an epoch. In an epoch, PF selects the user with the highest ratio of its instantaneous rate and its average rate to serve. In the following, we consider UDC with PF which is long-term. This is viable because PF is actually related to the competition of flows from different users, while the UDC requirements capture the relation between the flows belonging to the same user.

The average rate of user \( i \) at the beginning of session \( t+1 \) if during epoch \( t \), user \( i \) is provided a rate of \( R_i(t) \) is \( \bar{R}_i(t+1) = 0 \) if \( t = 0 \) and \( \bar{R}_i(t+1) = \frac{n_i-1}{n_i} \bar{R}_i(t) + \frac{1}{n_i} R_i(t) \) if \( t > 0 \). Here, by user rate, we refer to his downlink rate and \( n_i \) is the number of epochs. Its uplink rate is scaled down by a corresponding ratio \( \beta_D \) or \( \beta_R \) due to UDCs.

Normally, in a PF scheduler, at a certain time, only one user, which has the highest PF, is selected. However, there are some proposals for a scenario in which several users can be selected at the same time since there are several available resource portions such as in OFDM system where there are several sub-carriers which can be assigned to different users at the same time. Among those, [23] proved that a scheduler is proportionally fair for a multi-carrier transmission system if and only if it satisfies

\[
P = \arg \max_{S} \prod_{i \in U} \left( 1 + \frac{R_i(t)}{(n_e - 1)R_i(t - 1)} \right)
\]

in which \( S \) is any scheduler, \( U \) is the considered user set, \( R_i(t) \) is the total rate provided to user \( i \) at session \( t \), and \( R_i(t - 1) \) is the average of user \( i \) after the previous session.

In our case, we have to select two users in an epoch, therefore we can treat our scenario as an OFDM system with two sub-carriers. Thus the product in (9) has only two factors. The downlink rate provided by a composite scheme to a user can be seen as the rate a user can achieve when using a certain sub-carrier. The difference is that we cannot select two direct users or two relayed users and the resources are not fixed as the sub-carriers in OFDM but optimized using another degree of freedom.

Here in each epoch we select two users thus when we consider a user pair, the current rates of other users are 0. In a session, the achievable downlink rates of a direct user and a relayed user in the user pair \((i, j)\) are \( \min \{ a^T_1 \theta, \beta_D a^T_2 \theta \} \) and \( \min \{ a^T_1 \theta, \beta_R a^T_4 \theta \} \) respectively. Therefore the user pair selected in an epoch according to PF is given in (10).

VI. NUMERICAL RESULTS

A. Two-User Scenario

In this section we calculate the rate region for a two-user scenario with fixed channels. Fig. 5 shows the downlink rate regions \( (R_{Dd}, R_{Dd}) \) of different combined schemes \( S_{\text{all}} \) (all schemes), \( S_{\text{ref}} \) (all conventional schemes), \( S_{\text{dual}} \) (CDR schemes 1 and 2), \( S_{\text{dual}} \) (CDR schemes 3 and 4) and \( S_{\text{CDR}} \) (all CDR schemes) in case of channels \( \gamma = [\gamma_{BL} \gamma_{LR} \gamma_{BD} \gamma_{RD} \gamma_{LD}] = [10 10 10 -10 10] \) and \( \beta_D = \beta_R = 4 \). To calculate the rate region, the values of \( \theta_i, i \in \{ \text{du}, \ldots, \text{TWR} \} \) use resolution \( \Delta \theta_i = 0.08 \). The rate region of \( S_{\text{all}} \) certainly contains all of the other rate regions because it is the general case consisting of all value of \( \theta_i, i \in \{ \text{du}, \ldots, \text{TWR} \} \).

The rate region of \( S_{\text{CDR}} \) is larger than the union of the rate regions of \( S_{\text{dual}} \) and \( S_{\text{dual}} \) while the rate region \( S_{\text{all}} \) is also larger than the union of the rate regions of \( S_{\text{CDR}} \) and \( S_{\text{ref}} \). This is because \( \theta_m \) has to be selected such that UDCs are satisfied. \( S_{\text{CDR}} \) provides almost equal rates for the direct and relayed users while \( S_{\text{ref}} \) provide a low rate for the relayed user by prioritizing the direct user. This is because the CDR schemes feature joint transmissions which contribute to the rates of both users.

B. Multiple-User Scenario

In this section we calculate the average downlink rate and the PF coefficient for users with randomized positions in a network. In the simulation, we run \( n_{\text{iteration}} = 1000 \) scenarios. In a scenario, \( k = 20 \) users are randomly put in a circular cell with uniform distribution. The cell has a normalized radius.
of $R = 1$. The BS is at the center of the cell which is also the origin of the coordinate system $O(0,0)$. 4 relay stations are placed at the angle $\phi = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$ respectively and away from the BS with a normalized distance of $R_r = 0.7$. A user is determined to be a direct user if it is inside the circle with center at $O$ and radius of $R_r = 0.7$ as a similar principle in [21]. Any other user is determined as relayed user, attached to the nearest relay station. The number of direct users $n_D$ and the number of relayed users $n_R = k - n_D$ are not necessarily equal. All channels are modeled as $h = \frac{h_f}{\sqrt{d^\alpha}}$ in which $h_f$ is a Rayleigh fading coefficient with variance $\sigma_f = 1$, $d$ is the distance between the transmitter and the receiver and path loss coefficient $\alpha = 3$. We compare the performance of the composite MUE schemes $S_m$, $m \in \{\text{all, CDR, ref}\}$ and composite SUE considering different schedulers RR ER, MSR and PF. In a scenario, there are $n_v = 200$ epochs. In an epoch with RR ER scheduler, a user pair or a single user is served. After that a new pair or user is selected until no user is left. If a user pair is selected, the linear optimization problem presented in V-A is solved to find the optimal $\theta_m$. The rates will be calculated accordingly based on the formulas in sections III, V. If a single user is selected, since all channels are reciprocal and therefore the maximal downlink rate and maximal uplink rate are the same for both users, the single-user epoch is divided into two parts with ratio $\beta_i : 1$, $i \in \{D, R\}$. In an epoch with MSR scheduler, the user pair which has the highest total downlink rate provided by $S_m$ is selected.

In an epoch with with PF scheduler, a time fraction vector which can maximize the PF coefficient as defined in (10) using a composite scheme $S_m$ is determined for each relayed-direct user pair. The pair with the highest maximum PF coefficient is selected. In case of SUE, the user with the highest maximum PF coefficient is selected as $i_o = \arg \max_{i \in U} \frac{1}{1 + \frac{\min(\bar{R}^{T}a^{T}_m \theta, \beta_D a^{T}_m \theta) + \min(\bar{R}^{T}a^{T}_m \theta, \beta_R a^{T}_m \theta)}}{n_v - 1}R_i(t - 1)$ in which $U$ is the user set, $i$, $R_i$ are the provided rate and average rate of user $i$ in an epoch.

The average downlink rate of all users when different composite schemes and different schedulers are used is shown in Fig. 6. In this simulation, all downlink rates of all users are summed up and divided by the total time used by all users in a given scenario. For each of schedulers RR ER, MSR and PF, the order from the best to the worst is $S_{m \text{all}}, S_{\text{CDR}}, S_{\text{ref}}$. SUE performs worse than $S_{\text{ref}}$ does since it lacks $S_{\text{TVR}}$. As expected, the MSR scheduler has the highest sum rate, while the PF scheduler is better than the RR ER scheduler because it tries to maximize the network rate taking into account the PF coefficient.

Fig. 7 shows the results of PF coefficients of the schemes and we can see that the comparison is similar to the results of the average user’s downlink rate. The results for MSR scheduler are not shown, but MSR performs poorest in terms
of fairness. In Fig. 8, we can see the time fraction pairs $\theta_{\text{ud}}$, $\theta_{\text{uu}}$, $\theta_{\text{od}}$, $\theta_{\text{ou}}$, $\theta_{\text{TWR}}$ which are almost symmetric with the axe of symmetry $\beta = 0 \text{ dB}$. This is because in each pair, the time fractions provide opposite traffic types and the single fraction $\theta_{\text{TWR}}$ provides a symmetric traffic type.

The results point out that in a certain scenario, no individual scheme is consistently better than the others, seen in non-zero time fractions allocated to the other schemes. Although the general composite scheme is the best, the results point out that if we have to select a simpler composite scheme, comprising a lower number of individual schemes, then the CDR schemes are preferable because they not only improve the rates but also the fairness compared to the state-of-the-art conventional schemes. This is valid for all schedulers. Moreover, the result show that with CDR schemes, it is still correct that MSR gives the highest rate and RR ER gives the lowest rate among the considered schedulers. Especially the application of the PF scheduler along with the CDR schemes supports rate pairs that exhibit fairness and are not attainable by the reference

VII. CONCLUSION

In this paper we have investigated different multi-flow transmission (MFT) schemes in a wireless cellular scenario with combined direct/relay users. An example of a MFT scheme is two-way relaying and each MFT scheme can be used as a building block of more complex transmission schemes for serving multiple users. We have introduced a framework based on time-multiplexed composite transmission scheme that is able to integrate various MFT schemes, as well as the conventional single-flow transmission schemes. We have illustrated how this framework can be utilized to analyze the rate regions of scenarios in which multiple users have two-way traffic. The analysis is made by adapting well-known scheduling policies (Round Robin, Max Rate, and
Proportional Fair) to the proposed framework. Future work includes identification of other scenarios that can benefit from the MFT schemes, as well as considerations of MFT schemes with relaying methods that are more advanced compared to the amplify-and-forward.

REFERENCES


Chan Dai Truen Thai received his B.S. from Posts and Telecommunications Institute of Technology, Vietnam, in 2003; his MSc. from Korea Advanced Institute of Science and Technology (KAIST), South Korea, in 2008 and Ph.D. from Aalborg University, Denmark, in 2012.

Petar Popovski (S’97–A’98–M’04–SM’10) received the Dipl.-Ing. in electrical engineering and Magister Ing. in communication engineering from Sts. Cyril and Methodius University, Skopje, Macedonia, in 1997 and 2000, respectively and Ph.D. from Aalborg University, Denmark, in 2004. He was Assistant Professor (2004-2009) and Associate Professor (2009-2012) at Aalborg University. From 2008 to 2009 he held part-time position as a wireless architect at Oticon A/S. Since 2012 he is a Professor at Aalborg University. He has more than 130 publications in journals, conference proceedings and books and has more than 25 patents and patent applications. He has received the Young Elite Researcher award and the SAPERE AUDE career grant from the Danish Council for Independent Research. He has received six best paper awards, including three from IEEE. Dr. Popovski serves on the editorial board of several journals, including IEEE COMMUNICATIONS LETTERS (Senior Editor), IEEE TRANSACTIONS ON COMMUNICATIONS, and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. His research interests are in the broad area of wireless communication and networking, information theory and protocol design.

Megumi Kaneko received her B.S. and MSc. degrees in communication engineering in 2003 and 2004 from Institut National des Télécommunications (INT), France, jointly with a MSc. from Aalborg University, Denmark, where she received her Ph.D. degree in 2007. From January to July 2007, she was a visiting researcher in Kyoto University, Kyoto, Japan, and a JSPS post-doctoral fellow from April 2008 to August 2010. She is currently an Assistant Professor in the Department of Systems Science, Graduate School of Informatics, Kyoto University. Her research interests include wireless communication, cross-layer protocol design and communication theory. She received the 2009 Ericsson Young Scientist Award, the IEEE Globecon 2009 Best Paper Award, the 2011 Funai Young Researcher’s Award, the WPMC 2011 Best Paper Award and the 2012 Telecom System Technology Award.

Elisabeth de Carvalho received a Ph.D. in electrical engineering from Telecom ParisTech, France, in 1999. After her Ph.D. she was a post-doctoral fellow at Stanford University working in the field of ADSL communications. Then she worked in industry in the field of ADSL and wireless LAN. She joined Aalborg University in 2005 where she has led several national and international research projects in wireless communications. Her general area of expertise is in digital signal processing for wireless communications with multiple antennas. She is a co-author of the text book A Practical Guide to the MIMO Radio Channel.