This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2797128, IEEE Access

A Multiple-Relay Communication Protocol for Achieving Fairness in Dense Networks

Rodolfo Torrea-Duran[1,2], Fernando Rosas[1,3],
Sofie Pollin[1,4], Luc Vandendorpe[5], and Marc Moonen[1,2]
[1]KU Leuven, Department of Electrical Engineering (ESAT), Leuven, Belgium
[2]STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics
[3] MICAS, Microelectronics and Sensors
[4] TELEMIC, Telecommunications and Microwaves
[5] Université Catholique de Louvain (UCL), ICTEAM Institute, Digital Communications Group, LLN, Belgium
{Rodolfo.TorreaDuran, Fernando.Rosas, Sofie.Pollin, Marc.Moonen}@esat.kuleuven.be, Luc.Vandendorpe@uclouvain.be

***Abstract*—Increasing the density of base stations deployment is regarded as a means to satisfy the growing demand of wireless connectivity over a shared bandwidth, however increasing the disparity in the service received by each user. Fairness can be realized by dynamically allocating resources to users using detailed channel state information at the transmitter (CSIT), which constitutes an expensive overhead, especially for dense networks. One solution to improve the fairness without CSIT is to introduce spatial diversity through the use of relays in combination with physical-layer network coding (PNC). Most relaying-PNC solutions focus on the particular case of the two-way relay channel or they require additional relays and a large transmission time. In this paper we propose a multiple-relay communication protocol (MRCP) for achieving fairness in dense networks. It exploits spatial diversity without requiring additional relays since it uses the base stations as relays. Furthermore, MRCP is applicable to an arbitrary number of base stations and users, while keeping a small transmission time. We show that our approach achieves the highest max-min fairness among users and almost full diversity with asymmetric transmissions.**

***Index Terms*—Fairness, physical-layer network coding, relaying, dense networks.**

I. Introduction

It is expected that by 2021 there will be globally 11.6 billion mobile devices with an average mobile connection speed three times larger than that of 2016 [1]. This growth contrasts with the limited bandwidth that needs to be shared among an increasing number of users.

Network densification [2], [3] has been proposed as a promising technique to satisfy the previous demands over a shared bandwidth. This is realized by increasing the density of base stations deployed, resulting in dense clusters of base stations severely affected by inter-cell interference.

The most widely used approach to tackle inter-cell interference consists in orthogonalizing resources for users with high interference (typically in the cell edge), which means to assign disjoint resources (time or frequency) to users connected to different base stations, while reusing resources for users with low interference (typically in the cell center). However, when orthogonalizing resources, the amount of resources that each user receives decreases with the number of users. Furthermore, since additional base stations (such as femtocells) are usually deployed in an unplanned way and with restricted access, the inter-cell interference becomes extremely high for some users and extremely low for others. Therefore, providing a fair resource allocation for all the users while efficiently managing a shared bandwidth is of key importance in dense networks [4].

Traditionally, fairness in cellular networks is realized by dynamically allocating time, frequency [5], [6], or power [7], [8] to users according to strategies such as max-min or proportional fairness [9], [10], [11]. However, such a dynamic allocation can only be achieved with detailed channel state information at the transmitter (CSIT) from the whole network. Furthermore, given the large number of base stations and users in dense networks and the large variability of the wireless channel, obtaining reliable CSIT for a fair resource assignment remains an expensive overhead [12]. Moreover, sharing all this information through a backhaul link might be costly or even prohibitive for base stations in dense networks [3], [13].

One solution to improve fairness over a shared bandwidth without CSIT is to introduce spatial diversity through multiple-input multiple-output (MIMO) systems. which can average out the channel variability over different signal paths. However, physical constraints such as the number of radio frequency chains necessary at each device to support MIMO and the required separation between antennas of at least half of a wavelength, make this technology unsuitable for the widespread small-size and energy-limited devices found in dense networks.

In order to maintain the benefits of MIMO systems without the previous limitations, cooperative communication can be exploited with the aid of relays. Relays can achieve spatial diversity by forwarding received signals in dedicated time-slots without the need for a backhaul connection [14], [15]. However, relays require certain coordination with the base stations in order to use the bandwidth efficiently. For this purpose, relays have been used in combination with physical-layer network coding (PNC), first proposed in [16]. PNC exploits the linear superposition of wireless signals by combining transmissions overheard from different sources, hence providing a balance between orthogonalization and reuse of resources.

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2797128, IEEE Access

Much of the available literature on relaying-PNC focuses on the particular case of the two-way relay channel (TWRC) [17], [18]. In fact, most contributions deal with ways to combine the signals to achieve higher rates in a TWRC. For example, in [19] base stations and users transmit the symbols multiplied by the inverse of the channel so that the relay can receive the coherently added symbols. In [20] this scheme is extended to support multiple users simultaneously. In [21], a modulo operation is performed at the relay instead of transmitting a simple superposition of signals, and in [22], [23], [24], relays use a lattice-based approach to combine signals with a compress-and-forward strategy in [22] and with a compute-and-forward strategy in [23], [24].

To further increase the spatial diversity, multiple-input multiple-output (MIMO) techniques for the TWRC have been recently explored. For instance, in [25] a lattice-based PNC scheme with massive MIMO base stations and relays is used to reduce the number of time-slots required for transmission. In [26] the authors combine interference alignment with PNC to increase the sum-rate and the achievable degrees of freedom. In [27], a linear vector PNC scheme is used for spatial multiplexing MIMO TWRC without channel state information at the transmitter. Finally, in [28] PNC is used together with cross-layer information from different nodes in a MIMO network with asymmetric TWRC.

However, the spatial diversity achieved by these systems is limited since in a TWRC the users can only receive signals from the relay and not from the base station. Hence, a more general relaying-PNC configuration has been proposed in [29], where two users can achieve spatial diversity by receiving signals from a base station and a relay node. Yet, all of the previous approaches require additional infrastructure (i.e. relay nodes), with the operational and economical costs that their deployment implies. Furthermore, they all consider only a limited number of nodes, leaving the extension to an arbitrary number of base stations and users an open issue.

A different way of exploiting spatial diversity in a larger network using PNC principles is through space-time network coding (STNC) [30]. With STNC, each of the L base stations of a cluster broadcasts its own signal in dedicated time-slots during the transmission phase, and during the relaying phase M relays broadcast a combination of the received symbols in dedicated time-slots, hence achieving diversity order of $M+1$. This allows to achieve spatial diversity from different transmissions. In a TWRC users *exchange* information. Therefore, the symbol transmitted by one user serves as a-priori information to deduct it from the received signal in order to obtain the symbol transmitted by the other user. In contrast, with STNC, users have no such a-priori information. The achievable diversity comes from the relays transmissions.

A step to increase the diversity of the previous STNC scheme was taken in [31] and in [32], where the authors propose that each relay decodes the transmission not only from the base stations but from the previous relays. However, a main disadvantage of these type of schemes is that, apart from requiring additional relays, the price to achieve such diversity is the use of $L+M$ time-slots. This makes the previous approaches unsuitable for dense networks.

To tackle these problems, in this paper we present a multiple-relay communication protocol (MRCP) for achieving fairness in dense networks. Different from the approaches in [30], [31], [32], our approach exploits spatial diversity from a cluster of L base stations with only $L+1$ time-slots, while it avoids the need for a backhaul connection and relay deployment by using currently-deployed base stations as relays. With MRCP, all the base stations transmit in dedicated time-slots during the first L time-slots, and in the $L+1$-th time-slot they all transmit a combination of the overheard symbols. The last time-slot comes with an imperfect synchronization due to the multiple simultaneous transmissions. However, recent advances in delay-tolerant codes [33] or joint frequency and timing synchronization [34] show that this issue can be mitigated.

We then analyze the capabilities of the proposed system in terms not only of spectral efficiency and diversity, but also in terms of fairness, a critical metric for dense networks which has been mostly overlooked in relaying-PNC schemes. When compared to other communication schemes, we show that MRCP achieves the highest max-min fairness among users and a bit error rate (BER) that achieves almost full diversity.

The main contributions of this paper can be summarized as follows:

1) A new communication protocol for dense networks that achieves fairness by exploiting spatial diversity with a reduced number of time-slots compared to state-of-the-art. This is achieved without the need of backhaul connection nor additional relays.
2) The computation of closed-form expressions of the spectral efficiency and BER for any number of base stations and users using the proposed approach. This allows to analyze and predict the achievable gains for any network size and to compare it to other approaches.
3) The analysis of the studied approaches in terms of fairness. To our best knowledge, all the relaying-PNC schemes have overlooked the potential of achieving fairness from the combination of transmissions from different sources. Here we analyze the fairness based on a number of metrics for an integrated view of the system performance.

The rest of this paper is organized as follows. First, Section II describes the system model with the baseline and proposed approaches. Then, Section III describes the extension to a larger system size. Next, Section IV shows the performance evaluation of the proposed approach and its benchmarking against other communication schemes. Finally, Section V presents the main conclusions.

Notation: Matrices and vectors are denoted by upper and lower case boldface letters, respectively; $\mathbf{A}^T$ means transpose of $\mathbf{A}$; $\mathbf{A}^H$ means Hermitian of $\mathbf{A}$; $|\mathbf{A}|$ means determinant of $\mathbf{A}$; $\mathbb{E}\{\}$ is the expected value operator.

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2797128, IEEE Access

II. System Model

Consider a dense cluster composed by L base stations, each of which has data to be delivered to a specific user over a shared bandwidth.[1] We denote as s_l the symbol from the l-th base station (BSl) intended to the l-th user (Ul). It is assumed that symbols for different users are uncorrelated. The short distance between base stations allows all base stations and users to overhear all the transmitted symbols. Hence, the base stations can act as relays by retransmitting the overheard signals. Our goal is to find an efficient way to perform these transmissions.

It is assumed that all the base stations and users are half-duplex (i.e. they cannot transmit and receive simultaneously) and, if available, the backhaul link is only used for control information, not for data sharing. This last condition excludes the use of communication schemes that require coordinated base stations, e.g. space-time block codes like Alamouti codes. In this paper, we focus on single-antenna systems, leaving an extension to multiple antennas for future work.

For the following sections, let us define P_m as the transmit power of BSm and introduce the variable $\eta_m = \frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2} P_m$, where $\sigma_{\mathrm{s}}^2 = \mathbb{E}\{|s_1|^2\} = \mathbb{E}\{|s_2|^2\} = \cdots = \mathbb{E}\{|s_L|^2\}$ and σ_{n}^2 is the received noise power, which is assumed to be equal for all the users. We define the signal-to-noise ratio (SNR) from BSm at Ul as $\gamma_{ml} = \eta_m |h_{ml}|^2$, where h_{ml} is the channel gain from BSm to Ul and we define the SNR from BSm at BSn as $\xi_{mn} = \eta_m |g_{mn}|^2$, where g_{mn} is the channel gain from BSm to BSn.

Since fairness is a key parameter in this paper, different metrics are necessary to provide an integrated view of the system performance. Therefore, we propose to use the following:

1) Jain's Fairness Index (JFI), which is one of the most widely used quantitative fairness measures [4], first proposed in [35]. It has the advantages of being independent of the population size and the scale, continuous, and bounded between 0 and 1. It is defined as

$$\mathrm{JFI} = \frac{\left(\sum_{l=1}^{L} x_l\right)^2}{L \sum_{l=1}^{L} x_l^2}. \tag{1}$$

In this paper x_l represents spectral efficiency per time-slot of Ul. A large JFI value represents a fairer system. However, the JFI does not help in identifying unfairly treated users.

2) The *minimum spectral efficiency*, defined here as the spectral efficiency of the user with the lowest signal-to-interference-plus-noise ratio (SINR), provides an indication of the max-min fairness of the system. Although this metric does not reflect the overall system performance, it can pinpoint the performance of those unfairly treated users.

3) Finally, we use the BER as a measure of the diversity of the system.

[1]The shared bandwidth can represent a (set of) subcarrier(s) in an OFDM system.

For the sake of clarity, in this section we focus on the case where there are only two base stations and two users (i.e. $L = 2$), leaving the extension to larger systems for Section III. In the following, Section II-A presents the analysis of the spectral efficiency and bit error rate (BER) of three baseline approaches. Then, Sections II-B1 and II-B2 present the analysis of the proposed transmission and reception strategies.

A. Baseline Approaches

This section presents the analysis of three baseline approaches, which will be used to benchmark the performance of the proposed approach.

1) TDMA-2: In a basic time division multiple access (TDMA) approach, the communication is done in turns, i.e. first BS1 transmits s_1 to U1 while BS2 is inactive, and then BS2 transmits s_2 to U2 while BS1 is inactive, hence requiring 2 time-slots. This approach represents a full orthogonalization of resources and we refer to it as TDMA-2, where the 2 refers to $L = 2$. The weakness of this approach is the absence of spatial diversity and of fairness: if the channel between a user and its base station is in a deep fading, the transfer of information will be limited.

The spectral efficiency per time-slot of Ul using TDMA-2 is given as

$$C_{\mathrm{U}l}^{\mathrm{TDMA-2}} = \frac{1}{2}\mathbb{E}\{\log_2(1+\gamma_{ll})\}, \tag{2}$$

where the expected value is calculated over the distribution of the channel coefficients. Note that the $\frac{1}{2}$ scaling factor is due to the number of required time-slots for each transmission round. It can be seen that the spectral efficiency in equation (2) can be quite different for different users, which shows that this is not a fair approach.

Assuming BPSK transmissions and Rayleigh fading statistics (assumptions holding throughout the whole paper), the BER of Ul is computed in terms of the Gaussian Q-function using the results of appendix A as

$$\begin{aligned}\mathrm{BER}_{\mathrm{U}l}^{\mathrm{TDMA-2}} &= \mathbb{E}\left\{Q(\sqrt{2\gamma_{ll}})\right\} \\ &= \frac{1}{\pi}\int_0^{\pi/2}\left(\frac{1}{1+\frac{\bar{\gamma_{ll}}}{\sin^2\phi}}\right)d\phi \approx \frac{1}{\bar{\gamma_{ll}}},\end{aligned} \tag{3}$$

where $\bar{\gamma_{ml}} = \mathbb{E}\{\gamma_{ml}\} = \eta_m \mathbb{E}\{|h_{ml}|^2\}$ and the approximation holds in the high SNR regime [36]. Equation (3) shows the lack of diversity of this approach since the BER decays as $1/\bar{\gamma_{ll}}$.

2) DIV-2: One way of increasing the spatial diversity, and hence the fairness, is to exploit the overhearing capabilities of the system in order to share the transmitted symbols between the base stations. In this way, each symbol can reach its destination following more than one signal path. For instance, BS1 transmits s_1 to both U1 and BS2 in time-slot 1, then

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2797128, IEEE Access

BS2 transmits s_2 to both U2 and BS1 in time-slot 2. Finally, BS1 relays s_2 in time-slot 3 and BS2 relays s_1 in time-slot 4. We refer to this approach as *Diversity*-2 (DIV-2). The relayed symbols are assumed to be the noisy versions of the original symbols.

Assuming a channel coherence time larger than four time-slots, the received signals for Ul in the four time-slots are given as:

$$\begin{aligned} y_{\mathrm{U}l}^{(1)} &= \sqrt{P_1} h_{1l} s_1 + n_{\mathrm{U}l}^{(1)} \\ y_{\mathrm{U}l}^{(2)} &= \sqrt{P_2} h_{2l} s_2 + n_{\mathrm{U}l}^{(2)} \\ y_{\mathrm{U}l}^{(3)} &= \sqrt{P_1} h_{1l} z_{21} + n_{\mathrm{U}l}^{(3)} \\ y_{\mathrm{U}l}^{(4)} &= \sqrt{P_2} h_{2l} z_{12} + n_{\mathrm{U}l}^{(4)} \end{aligned} \tag{4}$$

with

$$\begin{aligned} z_{12} &= s_1 + \frac{n_{\mathrm{BS2}}^{(1)}}{\sqrt{P_1} g_{12}} \\ z_{21} &= s_2 + \frac{n_{\mathrm{BS1}}^{(2)}}{\sqrt{P_2} g_{21}}, \end{aligned} \tag{5}$$

where $y_{\mathrm{U}l}^{(t)}$ is the received signal of Ul in time-slot t, and $n_{\mathrm{U}l}^{(t)}$ and $n_{\mathrm{BS}m}^{(t)}$ are the AWGN noise of Ul and of BSm, respectively, in time-slot t (note that only half of the time-slots are useful for Ul).

The received signal $\mathbf{y}_l = \begin{bmatrix} y_{\mathrm{U}l}^{(1)} & y_{\mathrm{U}l}^{(2)} & y_{\mathrm{U}l}^{(3)} & y_{\mathrm{U}l}^{(4)} \end{bmatrix}^T$, can be expressed in matrix form as

$$\begin{bmatrix} y_{\mathrm{U}l}^{(1)} \\ y_{\mathrm{U}l}^{(2)} \\ y_{\mathrm{U}l}^{(3)} \\ y_{\mathrm{U}l}^{(4)} \end{bmatrix} = \begin{bmatrix} \sqrt{P_1} h_{1l} & 0 \\ 0 & \sqrt{P_2} h_{2l} \\ 0 & \sqrt{P_1} h_{1l} \\ \sqrt{P_2} h_{2l} & 0 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} n_{\mathrm{U}l}^{(1)} \\ n_{\mathrm{U}l}^{(2)} \\ n_{\mathrm{U}l}^{(3)} + \frac{\sqrt{P_1} h_{1l}}{\sqrt{P_2} g_{21}} n_{\mathrm{BS1}}^{(2)} \\ n_{\mathrm{U}l}^{(4)} + \frac{\sqrt{P_2} h_{2l}}{\sqrt{P_1} g_{12}} n_{\mathrm{BS2}}^{(1)} \end{bmatrix}. \tag{6}$$

Using a minimum mean-square error (MMSE) receiver (as described in section II-B), the spectral efficiency per time-slot of Ul using DIV-2 is computed as

$$C_{\mathrm{U}l}^{\mathrm{DIV-2}} = \frac{1}{4} \mathbb{E} \left\{ \log_2 \left(1 + \gamma_{ll} + \frac{\gamma_{ml}}{\frac{\gamma_{ml}}{\xi_{lm}} + 1} \right) \right\}, \tag{7}$$

for $m \neq l$ (i.e. $m = 1$ if $l = 2$ and $m = 2$ if $l = 1$), which shows the gain of Ul from the transmission of both base stations. From equation (7), it is observed that when the channel gain between BSl and BSm is strong compared to the channel gain between BSm and Ul[2], i.e. $\xi_{lm} \gg \gamma_{ml}$, DIV-2 achieves high fairness from the transmission of both base stations.

Using this assumption and the results of appendix A and assuming independent and identically distributed Rayleigh fading paths as well as equally-likely transmitted symbols, the BER for Ul is computed as

$$\begin{aligned} \mathrm{BER}_{\mathrm{U}l}^{\mathrm{DIV-2}} &= \mathbb{E} \left\{ Q \left(\sqrt{2 \left(\gamma_{ll} + \frac{\gamma_{ml}}{\frac{\gamma_{ml}}{\xi_{lm}} + 1} \right)} \right) \right\} \\ &= \frac{1}{\pi} \int_0^{\pi/2} \left(\frac{1}{1 + \frac{\bar{\gamma_{1l}}}{\sin^2 \phi}} \right) \left(\frac{1}{1 + \frac{\bar{\gamma_{2l}}}{\sin^2 \phi}} \right) d\phi \approx \frac{1}{\bar{\gamma_{1l}} \bar{\gamma_{2l}}}, \end{aligned} \tag{8}$$

Again, the approximation holds in the high SNR regime. Equation (8) shows that DIV-2 achieves full diversity, i.e. a diversity gain of order L, at the cost of doubling the number of time-slots compared to TDMA-2.

3) INTF-2: A last baseline approach consists in both base stations transmitting simultaneously regardless of the resulting interference. This approach represents a full reuse of resources and we refer to it as *Interference*-2 (INTF-2). The spectral efficiency per time-slot of Ul using INTF-2 is computed as

$$C_{\mathrm{U}l}^{\mathrm{INTF-2}} = \mathbb{E} \left\{ \log_2 \left(1 + \frac{\gamma_{ll}}{1 + \gamma_{ml}} \right) \right\} \tag{9}$$

for $m \neq l$. Although it uses each time-slot for transmission, equation (9) shows that this approach is not fair as it only benefits the user with the smallest interference.

To compute the BER of Ul with INTF-2, we use equation (52) from appendix A and integrate over the joint PDF of γ_{1l} and γ_{2l}, i.e. $p(\gamma_{1l}, \gamma_{2l})$. Since γ_{1l} and γ_{2l} are statistically independent, $p(\gamma_{1l}, \gamma_{2l}) = p(\gamma_{1l}) p(\gamma_{2l})$, which for Rayleigh fading statistics results in

$$\begin{aligned} &\mathrm{BER}_{\mathrm{U}l}^{\mathrm{INTF-2}} = \mathbb{E} \left\{ Q \left(\sqrt{2 \left[\frac{\gamma_{ll}}{1 + \gamma_{ml}} \right]} \right) \right\} = \\ &\frac{1}{\bar{\gamma_{1l}} \bar{\gamma_{2l}}} \int_0^{\infty} \int_0^{\infty} Q \left(\sqrt{2 \left[\frac{\gamma_{ll}}{1 + \gamma_{ml}} \right]} \right) \exp \left(- \left[\frac{\gamma_{1l}}{\bar{\gamma_{1l}}} + \frac{\gamma_{2l}}{\bar{\gamma_{2l}}} \right] \right) d\gamma_{1l} d\gamma_{2l}. \end{aligned} \tag{10}$$

Equation (10) shows that the diversity gain of this approach is lower than the diversity gain achieved by TDMA-2, especially when $\gamma_{ml} > \gamma_{ll}$.

B. Proposed Approach with Two Base Stations: MRCP-2

This section presents the proposed approach with $L = 2$, which combines a three-time-slot transmission strategy and an MMSE reception strategy. We refer to this approach as multiple-relay communication protocol-2 (MRCP-2).

1) Three-Time-Slot Transmission Strategy: The proposed approach exploits spatial diversity while using only three time-slots. In the first time-slot, BS1 transmits s_1 to U1, U2, and BS2. Then, in the second time-slot, BS2 transmits s_2 to U1, U2, and BS1. At this point, each base station has overheard the symbols transmitted by the other base station. Hence, in the third time-slot, each base station relays the received symbols

[2]This is a reasonable assumption since base stations are usually equipped with more powerful receivers (i.e. with greater sensitivity and smaller noise figure) and they often count with line-of-sight (LOS) between them.

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2797128, IEEE Access

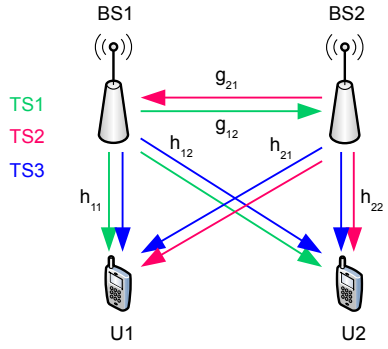

Fig. 1. System model for the case of two base stations and two users using MRCP-2. The parameter h_{ml} defines the channel gain of Ul from BSm and g_{mn} defines the channel gain of BSn from BSm. TSt defines the t-th time-slot.

(s_2 for BS1 and s_1 for BS2) to U1 and U2. The relayed symbols are assumed to be the noisy versions of the original symbols, i.e. we assume an amplify-and-forward (AF) strategy. This is shown in Fig. 1.

Assuming a channel coherence time larger than three time-slots, the received signals for Ul in the three time-slots are given as:

$$\begin{aligned} y_{\mathrm{U}l}^{(1)} &= \sqrt{P_1}h_{1l}s_1 + n_{\mathrm{U}l}^{(1)} \\ y_{\mathrm{U}l}^{(2)} &= \sqrt{P_2}h_{2l}s_2 + n_{\mathrm{U}l}^{(2)} \\ y_{\mathrm{U}l}^{(3)} &= \sqrt{P_1}h_{1l}z_{21} + \sqrt{P_2}h_{2l}z_{12} + n_{\mathrm{U}l}^{(3)}, \end{aligned} \tag{11}$$

with

$$\begin{aligned} z_{12} &= s_1 + \frac{n_{\mathrm{BS2}}^{(1)}}{\sqrt{P_1}g_{12}} \\ z_{21} &= s_2 + \frac{n_{\mathrm{BS1}}^{(2)}}{\sqrt{P_2}g_{21}}, \end{aligned} \tag{12}$$

where $y_{\mathrm{U}l}^{(t)}$ is the received signal of Ul in time-slot t, and $n_{\mathrm{U}l}^{(t)}$ and $n_{\mathrm{BS}m}^{(t)}$ are the AWGN noise of Ul and of BSm, respectively, in time-slot t.

This approach is intermediate between TDMA-2 and INTF-2, as it exploits the transmissions of both base stations without requiring too many additional time-slots.

The received signal $\mathbf{y}_l = \begin{bmatrix} y_{\mathrm{U}l}^{(1)} & y_{\mathrm{U}l}^{(2)} & y_{\mathrm{U}l}^{(3)} \end{bmatrix}^T$, can be expressed in matrix form as

$$\begin{bmatrix} y_{\mathrm{U}l}^{(1)} \\ y_{\mathrm{U}l}^{(2)} \\ y_{\mathrm{U}l}^{(3)} \end{bmatrix} = \begin{bmatrix} \sqrt{P_1}h_{1l} & 0 \\ 0 & \sqrt{P_2}h_{2l} \\ \sqrt{P_2}h_{2l} & \sqrt{P_1}h_{1l} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} n_{\mathrm{U}l}^{(1)} \\ n_{\mathrm{U}l}^{(2)} \\ n_{\mathrm{U}l}^{(3)} + \frac{\sqrt{P_1}h_{1l}}{\sqrt{P_2}g_{21}}n_{\mathrm{BS1}}^{(2)} + \frac{\sqrt{P_2}h_{2l}}{\sqrt{P_1}g_{12}}n_{\mathrm{BS2}}^{(1)} \end{bmatrix}. \tag{13}$$

Considering U1, equation (13) can be expressed in vector form as

$$\mathbf{y}_1 = \mathbf{a}_1 s_1 + \mathbf{a}_2 s_2 + \mathbf{n}_1 = \mathbf{a}_1 s_1 + \mathbf{w}_1, \tag{14}$$

where $\mathbf{a}_1 = \begin{bmatrix} \sqrt{P_1}h_{1l} & 0 & \sqrt{P_2}h_{2l} \end{bmatrix}^T$, $\mathbf{a}_2 = \begin{bmatrix} 0 & \sqrt{P_2}h_{2l} & \sqrt{P_1}h_{1l} \end{bmatrix}^T$, $\mathbf{w}_1 = \mathbf{a}_2 s_2 + \mathbf{n}_1$ is defined as the interference-plus-noise vector of U1, and $\mathbf{n}_1 = \begin{bmatrix} n_{\mathrm{U1}}^{(1)} & n_{\mathrm{U1}}^{(2)} & n_{\mathrm{U1}}^{(3)} + \frac{\sqrt{P_1}h_{11}}{\sqrt{P_2}g_{21}}n_{\mathrm{BS1}}^{(2)} + \frac{\sqrt{P_2}h_{21}}{\sqrt{P_1}g_{12}}n_{\mathrm{BS2}}^{(1)} \end{bmatrix}^T$.

2) MMSE Reception Strategy: The MMSE receiver is well known for being an efficient low-complexity linear receiver, which provides sufficient statistics to detect the input signal when it follows a Gaussian distribution [37]. The MMSE receiver of U1 can be derived in two steps. The first step consists in whitening the colored noise term $\mathbf{w}_1$ from equation (14) with its covariance matrix $\mathbf{R}_{w_1}$. This is achieved by filtering the received signal in (14) with the matrix $\mathbf{R}_{w_1}^{-1/2}$:

$$\mathbf{R}_{w_1}^{-1/2}\mathbf{y}_1 = \mathbf{R}_{w_1}^{-1/2}\mathbf{a}_1 s_1 + \mathbf{R}_{w_1}^{-1/2}\mathbf{w}_1. \tag{15}$$

The second step consists in performing maximal ratio combining (MRC) over the remaining signal, which is an optimal processing when the additive noise is white [37]. This is achieved by taking the inner product of the signal of (15) and the vector $\mathbf{R}_{w_1}^{-1/2}\mathbf{a}_1$:

$$\begin{aligned} \hat{z} = (\mathbf{R}_{w_1}^{-1/2}\mathbf{a}_1)^H\mathbf{R}_{w_1}^{-1/2}\mathbf{y}_1 &= \mathbf{a}_1^H\mathbf{R}_{w_1}^{-1}\mathbf{a}_1 s_1 + \mathbf{a}_1^H\mathbf{R}_{w_1}^{-1}\mathbf{w}_1 \\ &= \hat{z}_{\mathrm{sig}} + \hat{z}_{\mathrm{noise}}. \end{aligned} \tag{16}$$

From equation (16), the signal power can be calculated as

$$\begin{aligned} \mathbb{E}\{|\hat{z}_{\mathrm{sig}}|^2\} &= \mathbb{E}\{(\mathbf{a}_1^H\mathbf{R}_{w_1}^{-1}\mathbf{a}_1 s_1)(\mathbf{a}_1^H\mathbf{R}_{w_1}^{-1}\mathbf{a}_1 s_1)^H\} \\ &= (\mathbf{a}_1^H\mathbf{R}_{w_1}^{-1}\mathbf{a}_1)^2\sigma_{\mathrm{s}}^2 \end{aligned} \tag{17}$$

and the noise power as

$$\begin{aligned} \mathbb{E}\{|\hat{z}_{\mathrm{noise}}|^2\} &= \mathbb{E}\{(\mathbf{a}_1^H\mathbf{R}_{w_1}^{-1}\mathbf{w}_1)(\mathbf{a}_1^H\mathbf{R}_{w_1}^{-1}\mathbf{w}_1)^H\} \\ &= \mathbf{a}_1^H\mathbf{R}_{w_1}^{-1}\mathbf{a}_1. \end{aligned} \tag{18}$$

Using equations (17) and (18), the SNR of U1 using the MMSE receiver is found to be

$$\mathrm{SNR}_{\mathrm{U1}}^{\mathrm{MRCP-2}} = \frac{\mathbb{E}\{|\hat{z}_{\mathrm{sig}}|^2\}}{\mathbb{E}\{|\hat{z}_{\mathrm{noise}}|^2\}} = \mathbf{a}_1^H\mathbf{R}_{w_1}^{-1}\mathbf{a}_1\sigma_{\mathrm{s}}^2. \tag{19}$$

3) Performance Analysis: Let us now derive the spectral efficiency of U1 using the proposed transmission strategy. By assuming s_1 and s_2 as independent symbols with a Gaussian distribution, the spectral efficiency per time-slot of U1 is calculated as

$$C_{\mathrm{U1}}^{\mathrm{MRCP-2}} = \frac{1}{3}\left[H(\mathbf{y}_1) - H(\mathbf{w}_1)\right] = \frac{1}{3}\log_2\frac{|\mathbf{R}_{y_1}|}{|\mathbf{R}_{w_1}|}, \tag{20}$$

where $H(\mathbf{y}_1)$ and $H(\mathbf{w}_1)$ are the entropies of $\mathbf{y}_1$ and $\mathbf{w}_1$, and $\mathbf{R}_{y_1}$ and $\mathbf{R}_{w_1}$ are the covariance matrices of $\mathbf{y}_1$ and $\mathbf{w}_1$. The $\frac{1}{3}$ scaling factor corresponds to the number of time-slots used for each transmission round. The last equality in (20) is obtained by using the well-known expression for the entropy of a multivariate complex Gaussian distribution [38]. Moreover,

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2797128, IEEE Access

$\mathbf{R}_{w_1}$ is given as

$$\mathbf{R}_{w_1} = \mathbb{E}\{(\mathbf{a}_2 s_2 + \mathbf{n}_1)(\mathbf{a}_2 s_2 + \mathbf{n}_1)^H\} = \mathbf{a}_2\mathbf{a}_2^H\sigma_{\mathrm{s}}^2 + \mathbf{N}_1, \tag{21}$$

where $\mathbf{N}_1 = \mathrm{diag}\left[\sigma_{\mathrm{n}}^2, \sigma_{\mathrm{n}}^2, \sigma_{\mathrm{n}}^2\left(1 + \frac{P_1\mathbb{E}\{|h_{11}|^2\}}{P_2\mathbb{E}\{|g_{21}|^2\}} + \frac{P_2\mathbb{E}\{|h_{21}|^2\}}{P_1\mathbb{E}\{|g_{12}|^2\}}\right)\right]$, and $\mathbf{R}_{y_1}$ is given as

$$\begin{aligned}\mathbf{R}_{y_1} &= \mathbb{E}\{(\mathbf{a}_1 s_1 + \mathbf{w}_1)(\mathbf{a}_1 s_1 + \mathbf{w}_1)^H\} \\ &= \mathbf{a}_1\mathbf{a}_1^H\sigma_{\mathrm{s}}^2 + \mathbf{a}_2\mathbf{a}_2^H\sigma_{\mathrm{s}}^2 + \mathbf{N}_1.\end{aligned} \tag{22}$$

Hence, equation (20) can be re-written as

$$\begin{aligned}C_{\mathrm{U1}}^{\mathrm{MRCP-2}} &= \frac{1}{3}\log_2\left|\mathbf{R}_{w_1}^{-1}(\mathbf{R}_{w_1} + \mathbf{a}_1\mathbf{a}_1^H\sigma_{\mathrm{s}}^2)\right| \\ &= \frac{1}{3}\log_2\left(1 + \mathbf{a}_1^H\mathbf{R}_{w_1}^{-1}\mathbf{a}_1\sigma_{\mathrm{s}}^2\right).\end{aligned} \tag{23}$$

Comparing equation (19) with equation (23), it is clear that the MMSE receiver is able to achieve the channel capacity. Following the same procedure, the spectral efficiency per time-slot of U2 is calculated as

$$C_{\mathrm{U2}}^{\mathrm{MRCP-2}} = \frac{1}{3}\log_2\left(1 + \mathbf{a}_2^H\mathbf{R}_{w_2}^{-1}\mathbf{a}_2\sigma_{\mathrm{s}}^2\right), \tag{24}$$

where in this case $\mathbf{R}_{w_2} = \mathbf{a}_1\mathbf{a}_1^H\sigma_{\mathrm{s}}^2 + \mathbf{N}_2$.

Using equations (21), (22), and (23), and the results presented in appendix B, the spectral efficiency per time-slot of Ul using MRCP-2 is computed as

$$C_{\mathrm{U}l}^{\mathrm{MRCP-2}} = \frac{1}{3}\mathbb{E}\left\{\log_2\left(1 + \gamma_{ll} + \frac{\gamma_{ml}}{\frac{\gamma_{ll}}{\gamma_{ml}+1} + \frac{\gamma_{ll}}{\xi_{ml}} + \frac{\gamma_{ml}}{\xi_{lm}} + 1}\right)\right\} \tag{25}$$

where $m \neq l$. Equation (25) indicates a gain from the transmission of both base stations. When the channel gains between base stations are stronger than the channel gains between base stations and users, i.e. $\xi_{ml} \gg \gamma_{ll}$ and $\xi_{lm} \gg \gamma_{ml}$, MRCP-2 shows symmetry between γ_{ll} and γ_{ml}, which highlights the fairness of the proposed approach.

Let us now calculate the BER of Ul. By noting that γ_{1l} and γ_{2l} are statistically independent and defining $\gamma_{\mathrm{MRCP}} = \gamma_{ll} + \frac{\gamma_{ml}}{\frac{\gamma_{ll}}{\gamma_{ml}+1} + \frac{\gamma_{ll}}{\xi_{ml}} + \frac{\gamma_{ml}}{\xi_{lm}} + 1}$, it is shown that

$$\mathrm{BER}_{\mathrm{U}l}^{\mathrm{MRCP-2}} = \mathbb{E}\left\{Q\left(\sqrt{2\gamma_{\mathrm{MRCP}}}\right)\right\} =$$

$$\frac{1}{\bar{\gamma_{1l}}\bar{\gamma_{2l}}}\int_0^\infty\int_0^\infty Q\left(\sqrt{2\gamma_{\mathrm{MRCP}}}\right)\exp\left(-\left[\frac{\gamma_{1l}}{\bar{\gamma_{1l}}} + \frac{\gamma_{2l}}{\bar{\gamma_{2l}}}\right]\right)d\gamma_{1l}d\gamma_{2l}. \tag{26}$$

Note that $\gamma_{\mathrm{MRCP}} \leq \gamma_{1l} + \gamma_{2l}$, the difference being small if $\gamma_{1l} \gg \gamma_{2l}$ or if $\gamma_{2l} \gg \gamma_{1l}$. Therefore, the diversity gain provided by this approach is larger than the diversity provided by TDMA-2 but smaller than the diversity provided by DIV-2.

III. Extension to an Arbitrary Number of Base Stations and Users

This section extends the results of Section II to a larger system size. In the following, Section III-A presents the analysis of the extended baseline approaches, and then Section III-B presents the analysis of the extended proposed approach.

A. Extended Baseline Approaches

This section presents the analysis of the extended baseline approaches without diversity (TDMA-L), with diversity (DIV-L), and with interfering continuous transmissions (INTF-L). These will be used to benchmark the performance of the extended proposed approach. These benchmarks are chosen to show the achievable trade-off between achieving full spatial diversity (DIV-L) and reducing the transmission time, i.e. maximizing the degrees of freedom (INTF-L). In other words, these benchmarks correspond to the two extremes of the diversity-multiplexing trade-off, whereas the proposed approach achieves a trade-off between both extremes.

1) TDMA-L: In this approach, every base station performs one transmission while the others remain inactive, taking L time-slots to transmit all L symbols. We refer to this approach as TDMA-L. The spectral efficiency per time-slot of Ul is computed as

$$C_{\mathrm{U}l}^{\mathrm{TDMA-L}} = \frac{1}{L}\mathbb{E}\{\log_2(1 + \gamma_{ll})\}. \tag{27}$$

Equation (27) shows that the spectral efficiency per time-slot of each user decreases with L. Also, each user can have a very different spectral efficiency depending on γ_{ll}, which shows that this is not a fair approach.

The BER of Ul remains the same as in the case where $L = 2$:

$$\mathrm{BER}_{\mathrm{U}l}^{\mathrm{TDMA-L}} = \frac{1}{\pi}\int_0^{\pi/2}\left(\frac{1}{1 + \frac{\bar{\gamma_{ll}}}{\sin^2\phi}}\right)d\phi \approx \frac{1}{\bar{\gamma_{ll}}}, \tag{28}$$

From equation (28), it can be seen that the BER is independent of L and hence no diversity gain is achieved.

2) DIV-L: In this approach, each base station first transmits a symbol to its intended user and then relays each received symbol in a dedicated time-slot. Specifically, BS1 transmits s_1 to all users and base stations in the first time-slot, then BS2 transmits s_2 to all users and base stations in the second time-slot, until BSL transmits s_L to all users and base stations in time-slot L. Similarly in time-slots $L+1$ to $2L$, but this time BS1 relays s_2, BS2 relays s_3, and so on until BSL relays s_1. Finally, in time-slots $(L-1)L+1$ to L^2, BS1 relays s_L, BS2 relays s_1, and so on until BSL relays s_{L-1}. Hence, one complete round of transmission takes L^2 time-slots. We refer to this approach as DIV-L. The relayed symbols are assumed to be the noisy versions of the original symbols. In comparison to TDMA-L, DIV-L achieves fairness and diversity from all the base stations at the cost of increasing the transmission time to L^2.

Assuming a channel coherence time larger than L^2 time-slots, the received signals for Ul in the L^2 time-slots are given

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2797128, IEEE Access

as:

$$\begin{aligned}
y_{\mathrm{U}l}^{(1)} &= \sqrt{P_1}h_{1l}s_1 + n_{\mathrm{U}l}^{(1)} \\
y_{\mathrm{U}l}^{(2)} &= \sqrt{P_2}h_{2l}s_2 + n_{\mathrm{U}l}^{(2)} \\
&\vdots \\
y_{\mathrm{U}l}^{(L)} &= \sqrt{P_L}h_{Ll}s_L + n_{\mathrm{U}l}^{(L)} \\
y_{\mathrm{U}l}^{(L+1)} &= \sqrt{P_1}h_{1l}z_{21} + n_{\mathrm{U}l}^{(L+1)} \\
y_{\mathrm{U}l}^{(L+2)} &= \sqrt{P_2}h_{2l}z_{32} + n_{\mathrm{U}l}^{(L+2)} \\
&\vdots \\
y_{\mathrm{U}l}^{((L-1)L)} &= \sqrt{P_L}h_{Ll}z_{(L-2)L} + n_{\mathrm{U}l}^{((L-1)L)} \\
y_{\mathrm{U}l}^{((L-1)L+1)} &= \sqrt{P_1}h_{1l}z_{L1} + n_{\mathrm{U}l}^{((L-1)L+1)} \\
y_{\mathrm{U}l}^{((L-1)L+2)} &= \sqrt{P_2}h_{2l}z_{12} + n_{\mathrm{U}l}^{((L-1)L+2)} \\
&\vdots \\
y_{\mathrm{U}l}^{(L^2)} &= \sqrt{P_L}h_{Ll}z_{(L-1)L} + n_{\mathrm{U}l}^{(L^2)}
\end{aligned} \tag{29}$$

with

$$z_{lm} = s_l + \frac{n_{\mathrm{BS}m}^{(l)}}{\sqrt{P_l}g_{lm}}, \tag{30}$$

$\forall m \neq l$.

The received signal $\mathbf{y}_l = \begin{bmatrix} y_{\mathrm{U}l}^{(1)} & y_{\mathrm{U}l}^{(2)} & \cdots & y_{\mathrm{U}l}^{(L^2)} \end{bmatrix}^T$, can be expressed in matrix form as in equation (31).

Using the methodology described in section III-B, the spectral efficiency per time-slot of Ul using DIV-L is computed as

$$C_{\mathrm{U}l}^{\mathrm{DIV-L}} = \frac{1}{L^2}\mathbb{E}\left\{\log_2\left(1 + \gamma_{ll} + \sum_{\substack{m=1 \\ m\neq l}}^{L} \frac{\gamma_{ml}}{\frac{\gamma_{ml}}{\xi_{lm}} + 1}\right)\right\}, \tag{32}$$

When the channel gains between base stations are strong compared to the channel gains between base stations and users, i.e. $\xi_{lm} \gg \gamma_{ml}$ $\forall m \neq l$ and for the particular case in which $\gamma_{1l} = \gamma_{2l} = \cdots = \gamma_{Ll} = \bar{\gamma}$, then

$$1 + \gamma_{ll} + \sum_{\substack{m=1 \\ m\neq l}}^{L} \frac{\gamma_{ml}}{\frac{\gamma_{ml}}{\xi_{lm}} + 1} \approx L\bar{\gamma}. \tag{33}$$

This results in

$$C_{\mathrm{U}l}^{\mathrm{DIV-L}} \approx \frac{1}{L^2}\left[\log_2(L) + \log_2(\bar{\gamma})\right]. \tag{34}$$

As $\log_2(L)/L^2$ tends to zero faster than $1/L$, it is seen that $C_{\mathrm{U}l}^{\mathrm{DIV-L}}$ decreases faster with L than $C_{\mathrm{U}l}^{\mathrm{TDMA-L}}$. However, DIV-L achieves a high fairness from the transmission of multiple base stations.

Using the results from appendix A, the BER of Ul is computed as

$$\mathrm{BER}_{\mathrm{U}l}^{\mathrm{DIV-L}} = \frac{1}{\pi}\int_0^{\pi/2}\prod_{m=1}^{L}\left(\frac{1}{1 + \frac{\bar{\gamma_{ml}}}{\sin^2\phi}}\right)d\phi \approx \frac{1}{\prod_{m=1}^{L}\bar{\gamma_{ml}}}, \tag{35}$$

where the approximation holds in the high SNR regime. Again, equation (35) shows the full diversity achieved by DIV-L.

3) INTF-L: In this case all the base stations transmit continuously regardless of the resulting interference. Hence, a full transmission round takes only one time slot. We refer to this approach as INTF-L. The spectral efficiency per time-slot of Ul using INTF-L is computed as

$$C_{\mathrm{U}l}^{\mathrm{INTF-L}} = \mathbb{E}\left\{\log_2\left(1 + \frac{\gamma_{ll}}{1 + \sum_{\substack{m=1 \\ m\neq l}}^{L}\gamma_{ml}}\right)\right\}. \tag{36}$$

Although this approach uses all the time-slots for transmitting to each user, it is unfair due to the different interference levels perceived by each user. For the particular case in which $\gamma_{1l} = \gamma_{2l} = \cdots = \gamma_{Ll} = \bar{\gamma}$ and for a large L, using a first order Taylor series expansion $\log(1+x) \approx x$ shows that

$$C_{\mathrm{U}l}^{\mathrm{INTF-L}} \approx \mathbb{E}\left\{\frac{\gamma_{ll}}{1 + L\bar{\gamma}}\right\} = \frac{1}{\bar{\gamma}^{-1} + L}. \tag{37}$$

This shows that the spectral efficiency decreases as $1/L$.

The BER of Ul using this approach is computed as

$$\mathrm{BER}_{\mathrm{U}l}^{\mathrm{INTF-L}} = \mathbb{E}\left\{Q\left(\sqrt{2\gamma_{\mathrm{INTF}}}\right)\right\} = \frac{1}{\prod_{m=1}^{L}\bar{\gamma_{ml}}}\int_0^{\infty}\cdots\int_0^{\infty}Q\left(\sqrt{2\gamma_{\mathrm{INTF}}}\right)\exp\left(-\sum_{m=1}^{L}\frac{\gamma_{ml}}{\bar{\gamma_{ml}}}\right)d\gamma_{1l}\cdots d\gamma_{Ll}, \tag{38}$$

where $\gamma_{\mathrm{INTF}} = \gamma_{ll}/\left(1 + \sum_{\substack{m=1 \\ m\neq l}}^{L}\gamma_{ml}\right)$. Since $\gamma_{ll} \geq \gamma_{\mathrm{INTF}}$, equation (38) shows that the diversity gain is lower than that of TDMA-L and it decreases with L.

B. Proposed Approach with L Base Stations: MRCP-L

This section extends the proposed approach defined in Section II-B to an arbitrary number of base stations and users. It also shows the computation of closed-form expressions of the spectral efficiency and BER of the proposed approach.

1) (L+1)-Time-Slot Transmission Strategy: First each base station transmits its symbol to all the users and base stations in its own dedicated time-slot. At this point, each base station has overheard the transmissions from all the other base stations. Therefore, at the end of this round, one extra time slot is added, where each base station relays a linear combination of all the received symbols. Specifically, in time-slot l, BSl transmits s_l to Ul and to all other base stations and users for

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

$$\begin{bmatrix} y_{\mathrm{U}l}^{(1)} \\ y_{\mathrm{U}l}^{(2)} \\ \vdots \\ y_{\mathrm{U}l}^{(L^2)} \end{bmatrix} = \begin{bmatrix} \sqrt{P_1}h_{1l} & 0 & 0 & \cdots & 0 & 0 \\ 0 & \sqrt{P_2}h_{2l} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \sqrt{P_L}h_{Ll} \\ 0 & \sqrt{P_1}h_{1l} & 0 & \cdots & 0 & 0 \\ 0 & 0 & \sqrt{P_2}h_{2l} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \sqrt{P_1}h_{1l} \\ \sqrt{P_2}h_{2l} & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{P_L}h_{Ll} & 0 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_L \end{bmatrix} + \begin{bmatrix} n_{\mathrm{U}l}^{(1)} \\ n_{\mathrm{U}l}^{(2)} \\ \vdots \\ n_{\mathrm{U}l}^{(L)} \\ n_{\mathrm{U}l}^{(L+1)} + \frac{\sqrt{P_1}h_{1l}}{\sqrt{P_2}g_{21}} n_{\mathrm{BS}1}^{(2)} \\ \vdots \\ n_{\mathrm{U}l}^{(L^2)} + \frac{\sqrt{P_L}h_{Ll}}{\sqrt{P_{L-1}}g_{(L-1)L}} n_{\mathrm{BS}L}^{(L-1)} \end{bmatrix}. \quad (31)$$

$l = 1, 2, \cdots, L$ and in time-slot $L+1$ BSl relays $\sum_{\substack{m=1 \\ m \neq l}}^{L} s_m$ for $l = 1, 2, \cdots, L$. The relayed symbols are assumed to be the noisy versions of the original symbols, i.e. we assume an AF strategy. Hence, one complete round of transmission takes $L+1$ time-slots. We refer to this approach as MRCP-L.

Assuming that the channel coherence time is larger than $L+1$ time-slots, the received signals for Ul in $L+1$ time-slots is given as

$$\begin{aligned} y_{\mathrm{U}l}^{(1)} &= \sqrt{P_1}h_{1l}s_1 + n_{\mathrm{U}l}^{(1)} \\ y_{\mathrm{U}l}^{(2)} &= \sqrt{P_2}h_{2l}s_2 + n_{\mathrm{U}l}^{(2)} \\ &\vdots \\ y_{\mathrm{U}l}^{(L)} &= \sqrt{P_L}h_{Ll}s_L + n_{\mathrm{U}l}^{(L)} \\ y_{\mathrm{U}l}^{(L+1)} &= \sum_{m=1}^{L} \sqrt{P_m}h_{ml} \sum_{\substack{r=1 \\ r \neq m}}^{L} z_{rm} + n_{\mathrm{U}l}^{(L+1)} \end{aligned} \quad (39)$$

with

$$z_{lm} = s_l + \frac{n_{\mathrm{BS}m}^{(l)}}{\sqrt{P_l}g_{lm}}, \quad (40)$$

The received signal $\mathbf{y}_l = \begin{bmatrix} y_{\mathrm{U}l}^{(1)} & y_{\mathrm{U}l}^{(2)} & \cdots & y_{\mathrm{U}l}^{(L)} & y_{\mathrm{U}l}^{(L+1)} \end{bmatrix}^T$ can be expressed in matrix form as equation (41). The vector form of the received signal is then

$$\mathbf{y}_l = \sum_{m=1}^{L} \mathbf{a}_m s_m + \mathbf{n}_l = \mathbf{a}_l s_l + \mathbf{w}_l, \quad (42)$$

where $\mathbf{a}_l$ is the lth column of matrix $\mathbf{A}$, $\mathbf{n}_l =$

$$\begin{bmatrix} n_{\mathrm{U}l}^{(1)} & n_{\mathrm{U}l}^{(2)} & \cdots & n_{\mathrm{U}l}^{(L)} & n_{\mathrm{U}l}^{(L+1)} + \sum_{m=1}^{L} \sqrt{P_m}h_{ml} \sum_{\substack{r=1 \\ r \neq m}}^{L} \frac{n_{\mathrm{BS}m}^{(r)}}{\sqrt{P_r}g_{rm}} \end{bmatrix}^T, \quad (43)$$

and $\mathbf{w}_l = \sum_{\substack{m=1 \\ m \neq l}}^{L} \mathbf{a}_m s_m + \mathbf{n}_l$ is defined as the interference-plus-noise vector of Ul.

The MMSE reception strategy can be derived similarly as in equations (15) and (16). Then, the SNR of Ul results in

$$\mathrm{SNR}_{\mathrm{U}l}^{\mathrm{MRCP-L}} = \mathbf{a}_l^H \mathbf{R}_{w_l}^{-1} \mathbf{a}_l \sigma_{\mathrm{s}}^2, \quad (44)$$

where $\mathbf{R}_{w_l}$ is the covariance matrix of $\mathbf{w}_l$, whose explicit expression can be found in equation (67) for $l = 1$.

2) Performance Analysis: Let us now derive the spectral efficiency per time-slot of Ul using the proposed extended approach. For statistically independent Gaussian symbols, the covariance matrix $\mathbf{R}_{w_l}$ is computed as

$$\begin{aligned} \mathbf{R}_{w_l} &= \mathbb{E}\left\{ \left(\sum_{\substack{m=1 \\ m \neq l}}^{L} \mathbf{a}_m s_m + \mathbf{n}_l \right) \left(\sum_{\substack{m=1 \\ m \neq l}}^{L} \mathbf{a}_m s_m + \mathbf{n}_l \right)^H \right\} \\ &= \sum_{\substack{m=1 \\ m \neq l}}^{L} \mathbf{a}_m \mathbf{a}_m^H \sigma_{\mathrm{s}}^2 + \mathbf{N}_l \end{aligned} \quad (45)$$

where $\mathbf{N}_l = \mathrm{diag}\left[\sigma_{\mathrm{n}}^2, \sigma_{\mathrm{n}}^2, \cdots, \sigma_{\mathrm{n}}^2 (1 + D_l)\right]$, with

$$D_l = \sum_{m=1}^{L} P_m \mathbb{E}\{|h_{ml}|^2\} \sum_{\substack{r=1 \\ r \neq m}}^{L} \frac{1}{P_r \mathbb{E}\{|g_{rm}|^2\}} \quad (46)$$

and the covariance matrix $\mathbf{R}_{y_l}$ is

$$\begin{aligned} \mathbf{R}_{y_l} &= \mathbb{E}\{(\mathbf{a}_l s_l + \mathbf{w}_l)(\mathbf{a}_l s_l + \mathbf{w}_l)^H\} \\ &= \sum_{m=1}^{L} \mathbf{a}_m \mathbf{a}_m^H \sigma_{\mathrm{s}}^2 + \mathbf{N}_l. \end{aligned} \quad (47)$$

Hence, the spectral efficiency per time-slot achieved by Ul is written as

$$C_{\mathrm{U}l}^{\mathrm{MRCP-L}} = \frac{1}{L+1} \log_2 \frac{|\mathbf{R}_{y_l}|}{|\mathbf{R}_{w_l}|}. \quad (48)$$

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2797128, IEEE Access

$$\begin{bmatrix} y_{\mathrm{U}l}^{(1)} \\ \vdots \\ y_{\mathrm{U}l}^{(L)} \\ y_{\mathrm{U}l}^{(L+1)} \end{bmatrix} = \underbrace{\begin{bmatrix} \sqrt{P_1}h_{1l} & 0 & \cdots & 0 \\ 0 & \sqrt{P_2}h_{2l} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{P_L}h_{Ll} \\ \sum\limits_{\substack{m=1 \\ m\neq 1}}^{L} \sqrt{P_m}h_{ml} & \sum\limits_{\substack{m=1 \\ m\neq 2}}^{L} \sqrt{P_m}h_{ml} & \cdots & \sum\limits_{\substack{m=1 \\ m\neq L}}^{L} \sqrt{P_m}h_{ml} \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} s_1 \\ \vdots \\ s_L \end{bmatrix} + \begin{bmatrix} n_{\mathrm{U}l}^{(1)} \\ \vdots \\ n_{\mathrm{U}l}^{(L)} \\ n_{\mathrm{U}l}^{(L+1)} + \sum\limits_{m=1}^{L} \sqrt{P_m}h_{ml} \sum\limits_{\substack{r=1 \\ r\neq m}}^{L} \frac{n_{\mathrm{BS}m}^{(r)}}{\sqrt{P_r}g_{rm}} \end{bmatrix} \tag{41}$$

Solving equation (48) as shown in appendix C, it can be shown that $C_{\mathrm{U}l}^{\mathrm{MRCP-L}}$ results in

$$C_{\mathrm{U}l}^{\mathrm{MRCP-L}} = \frac{1}{L+1}\mathbb{E}\left\{ \log_2 \left(1 + \gamma_{ll} + \frac{\overbrace{\Upsilon_l}^{\Gamma}}{\sum\limits_{\substack{m=1 \\ m\neq l}}^{L} \frac{\Upsilon_m}{1+\gamma_{ml}} + D_l + 1} \right) \right\}, \tag{49}$$

where

$$\Upsilon_l = \left(\sum_{\substack{r=1 \\ r\neq l}}^{L} \sqrt{\gamma_{rl}} \right)^2. \tag{50}$$

Finally, using the results presented in appendix A, the BER of Ul is calculated as

$$\begin{aligned} \mathrm{BER}_{\mathrm{U}l}^{\mathrm{MRCP-L}} = \frac{1}{\prod\limits_{m=1}^{L} \bar{\gamma_{ml}}} \int_0^{\infty} \cdots \int_0^{\infty} Q\left(\sqrt{2\left(\gamma_{ll}+\Gamma\right)}\right) \\ \exp\left(-\sum_{m=1}^{L} \frac{\gamma_{ml}}{\bar{\gamma_{ml}}} \right) d\gamma_{1l} d\gamma_{2l} \cdots d\gamma_{Ll}. \end{aligned} \tag{51}$$

Note that $\gamma_{ll} + \Gamma \leq \sum_{m=1}^{L} \gamma_{ml}$. Therefore, the diversity gain provided by this approach is larger than that of TDMA-L but smaller than that of DIV-L.

Note that by using a decode-and-forward (DF) strategy instead of an AF strategy, the spectral efficiency would be limited by the decoding errors at each overhearing base station. Hence, this approach would require a selection of the base stations most suitable to relay the overheard symbols. Although interesting to analyze, this strategy is out of the scope of this paper.

IV. Performance Evaluation

In this section we compare the approaches analyzed in previous sections in terms of the JFI in Section IV-A, in terms of minimum spectral efficiency per time-slot in Section IV-B, and in terms of BER in Section IV-C. Throughout this section, we use the JFI as a quantitative measure of the fairness of the system and we use the minimum spectral efficiency per time-slot as a measure of the max-min fairness of the system. Finally, we also consider the BER as a measure of the diversity. Our evaluations consider BPSK modulation and a Rayleigh fading channel model with $\mathbb{E}\{|h_{ml}|^2\} = 1$ and $\mathbb{E}\{|g_{mn}|^2\} = 10\text{dB}\ \forall l, m, n$. Unless stated otherwise, we consider the same η_m from all but one base station, i.e. $\eta_1 = \eta_3 = \cdots = \eta_L = \bar{\eta}$ and $\eta_2 = 0\text{dB}$. This is done in order to evaluate the system performance with strongly asymmetric transmissions, typical of dense networks.

A. JFI Evaluation

The JFI can be interpreted as a quantitative measure of the "equality" in allocating resources to users. If all the users get the same amount of resources, then the system is 100% fair. If only a few users get most of the resources, then the JFI approaches zero. From this, it is easy to see that DIV and MRCP achieve the highest JFI since all the users benefit almost equally from the transmission of multiple base stations. In contrast, TDMA and INTF show a low fairness as they benefit only the user with the best conditions. This explains why when one user has a low SINR, the fairness of TDMA and INTF is very low for $L = 2$, while the effect of that unfairly treated user fades out with the rest of the users when L increases as seen in Fig. 2.

B. Spectral Efficiency Evaluation

Max-min fairness can be achieved with a high minimum spectral efficiency. In this sense, it is seen from equations (25) and (49), and confirmed from the numerical evaluations in this section that MRCP provides the highest max-min fairness. This can be seen for $L = 2$ in Figs. 3 and 4 and for different values of L in Fig. 6. This gain in minimum spectral efficiency increases not only with L, but also with $\bar{\eta}$ as can be seen in Fig. 5. Only for low values of $\bar{\eta}$, the minimum spectral efficiency of INTF is higher due to the low interference level between users, which is not common in dense networks.

TDMA and INTF are not fair approaches in most conditions. For example, the minimum spectral efficiency of TDMA shows a flooring when $\eta_1 > \eta_2$ because the increase in η_1 benefits one user until its spectral efficiency reaches the spectral efficiency of the other user (see Figs. 3 and 4). A similar lack of fairness can be seen for $L = 10$ in Fig. 5. In

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2797128, IEEE Access

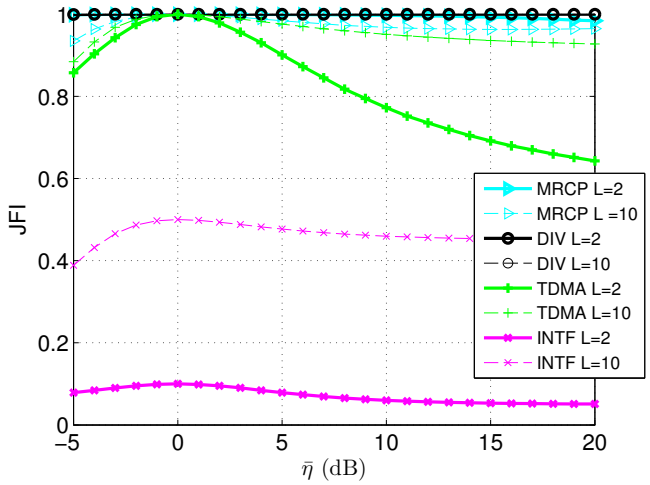

Fig. 2. JFI for $L = 2$ and $L = 10$. We consider the same η_m from all but one base station, i.e. $\eta_1 = \eta_3 = \cdots = \eta_L = \bar{\eta}$ and $\eta_2 = 0$dB.

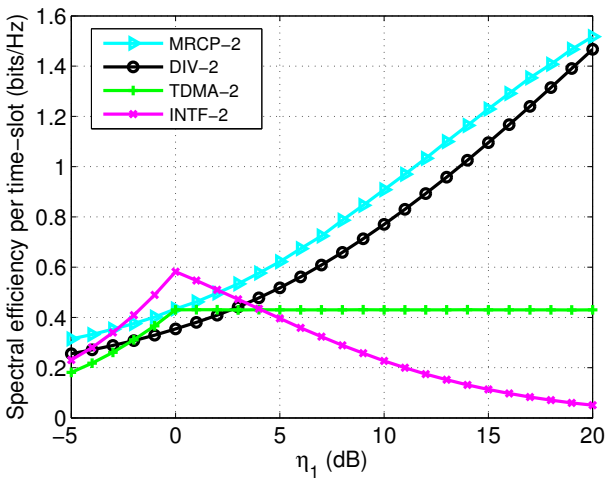

Fig. 3. Minimum spectral efficiency per time-slot for $\eta_2 = 0$dB.

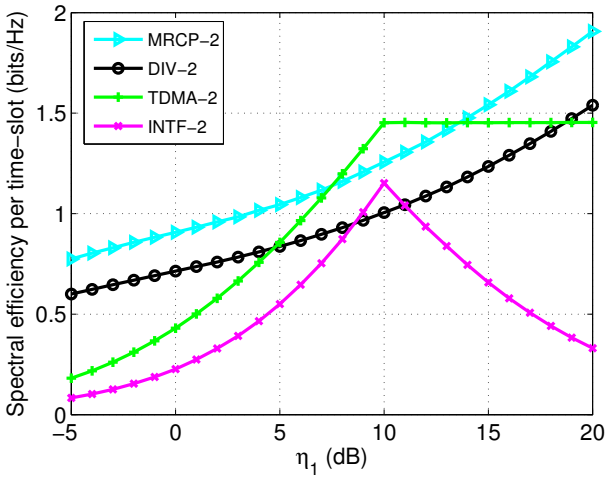

Fig. 4. Minimum spectral efficiency per time-slot for $\eta_2 = 10$dB.

the case of INTF, an increase in η_1 or $\bar{\eta}$ only increases the interference level, reducing the spectral efficiency.

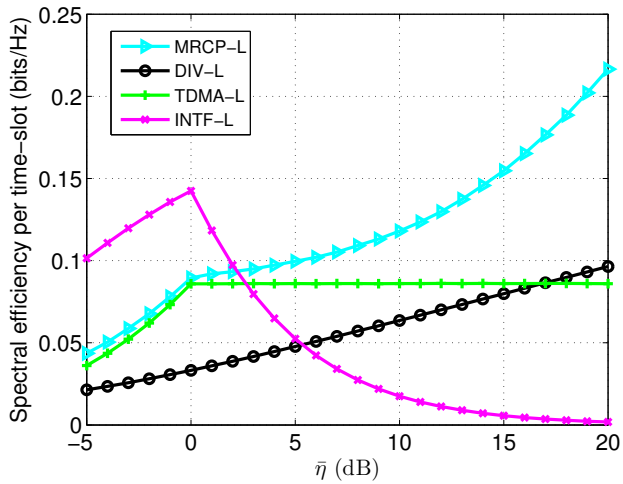

Fig. 5. Minimum spectral efficiency per time-slot for L = 10. We consider the same η_m from all but one base station, i.e. $\eta_1 = \eta_3 = \cdots = \eta_L = \bar{\eta}$ and $\eta_2 = 0$dB.

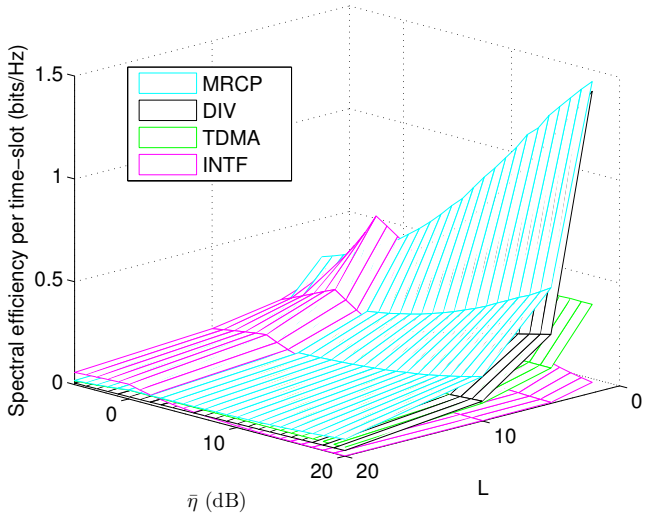

Fig. 6. Minimum spectral efficiency of the studied approaches. We consider the same η_m from all but one base station, i.e. $\eta_1 = \eta_3 = \cdots = \eta_L = \bar{\eta}$ and $\eta_2 = 0$dB.

C. BER Evaluation

Let us first consider symmetric conditions with equal η_m from *all* the base stations (i.e. $\eta_1 = \eta_2 = \cdots = \eta_L = \bar{\eta}$), MRCP presents almost the same BER as DIV for $L = 2$, and it degrades smoothly for $L = 5$, 10, and 20 (see Fig. 7). In the limit, with a very large value of L, MRCP achieves a similar BER as TDMA since the diversity gain comes only from time-slot $L + 1$. Needless to say, INTF is the approach with the poorest BER since increasing the number of base stations only decreases the diversity gain. Since DIV exploits full diversity, it has the best BER performance compared to the other approaches and its diversity gain increases with L (see Fig. 7). This comes at a price of increasing the number of time-slots to L^2.

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2797128, IEEE Access

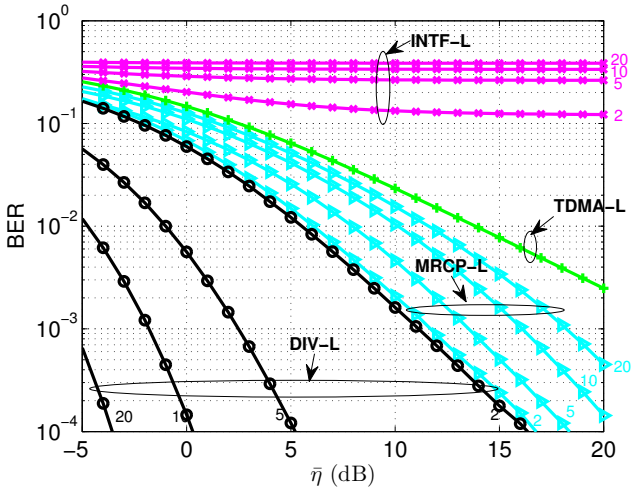

Fig. 7. BER for BPSK for $L = 2, 5, 10, 20$ (small numbers). For simplicity we consider the same η_m from all the base stations $\eta_1 = \eta_2 = \cdots = \eta_L = \bar{\eta}$.

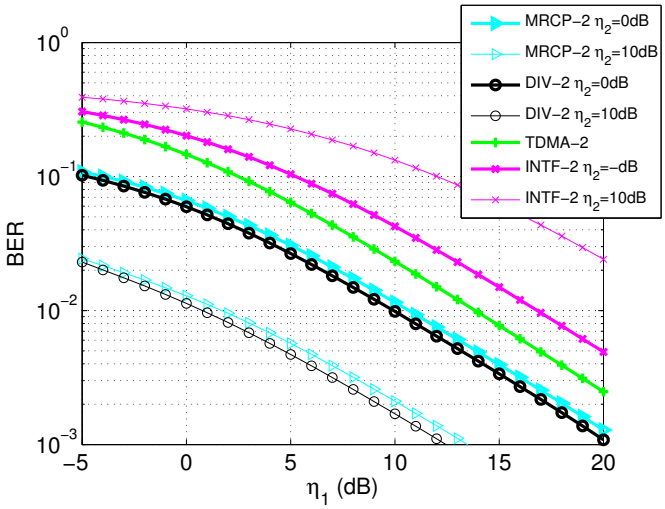

Fig. 8. BER for BPSK for $L = 2$ and η_2 = 0dB, 10dB.

Dealing with asymmetric conditions, i.e. different values of η_m, is a more realistic scenario for dense networks. In this case, the performance of MRCP and DIV are both the highest compared to the other approaches due to the exploited spatial diversity, with a marginal difference between them (see Fig. 8). Also, as the value of η_2 increases from 0 to 10dB, the performance gap between MRCP and DIV remains the same, while the performance gap between them and TDMA and INTF increases.

Note that the BER of TDMA is the same for all values of L due to the lack of diversity (see Fig. 7). Also, the BER of TDMA for $L = 2$ is the same for η_2 =0dB and η_2 =10dB as this is the BER of U1, which experiences no interference from U2 (see Fig. 8).

D. Final remarks

From the previous analysis of JFI, minimum spectral efficiency, and BER, we can conclude INTF and TDMA are both unfair approaches. On the other hand, DIV, closely followed by MRCP, can achieve the highest JFI fairness by exploiting the spatial diversity from the transmission of different base stations. However, the JFI fairness achieved by DIV comes at the cost of increasing the number of time-slots quadratically, while MRCP achieves the highest max-min fairness with less time-slots. This highlights the ability of MRCP to find a trade-off between spatial diversity and transmission time. Regarding the BER, DIV and MRCP are able to achieve the highest diversity gains with asymmetric transmissions. As an additional note, we believe that the proposed approach could be enhanced with the use of power control. Furthermore, considering multiple-antenna base stations and users can further increase the performance of the proposed approach. The study of these techniques is a topic of ongoing work.

V. Conclusions

In this paper we have proposed MRCP, a multiple-relay communication protocol for achieving fairness in dense networks. The proposed approach exploits spatial diversity without requiring a backhaul connection or additional infrastructure as it uses currently-deployed base stations as relays. In this way, the proposed approach is able to find an efficient balance between spatial diversity and transmission time. Also, different from other relaying-PNC schemes, MRCP is applicable to any number of base stations and users, while keeping a small transmission time.

We have shown that MRCP reaches almost maximum JFI fairness and the highest max-min fairness when compared to other communication schemes. This comes with a BER that reaches almost full diversity with asymmetric transmissions, typical of dense networks.

Appendix A

For coherent binary signals with BPSK modulation, the average BER of an L-branch MRC receiver with equally-likely transmitted symbols is given by the Gaussian Q-function [39]

$$\text{BER} = \mathbb{E}\{Q(\sqrt{2\gamma_{\text{tot}}})\} = \int_0^\infty Q(\sqrt{2\gamma_{\text{tot}}}) p_{\gamma_{\text{tot}}}(\gamma_{\text{tot}}) d\gamma_{\text{tot}}, \tag{52}$$

where $p(\gamma_{\text{tot}})$ is the probability density function (PDF) of γ_{tot}, and γ_{tot} is the total SNR per symbol at the output of receiver l:

$$\gamma_{\text{tot}} = \sum_{m=1}^{L} \gamma_{ml} = \sum_{m=1}^{L} \eta_l |h_{ml}|^2, \tag{53}$$

and where the Q-function is defined as (only for $x \geq 0$) [39]

$$Q(x) = \frac{1}{\pi} \int_0^{\pi/2} \exp\left(-\frac{x^2}{2\sin^2\phi}\right) d\phi. \tag{54}$$

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2797128, IEEE Access

The BER in (52) then results in

$$\mathrm{BER} = \underbrace{\int_0^\infty \cdots \int_0^\infty}_{L} \frac{1}{\pi} \int_0^{\pi/2} \prod_{m=1}^{L} \exp\left(\frac{-\gamma_{ml}}{\sin^2\phi}\right) p(\gamma_{ml}) d\phi d\gamma_{1l} d\gamma_{2l} ... d\gamma_{Ll}, \tag{55}$$

where it is assumed that γ_{ml} for $m = 1, ...L$ are statistically independent. By using the definition of the moment generating function (MGF) of γ_{ml} [39]:

$$M_{\gamma_{ml}}\left(\frac{-1}{\sin^2\phi}\right) = \int_0^\infty p(\gamma_{ml}) \exp\left(\frac{-\gamma_{ml}}{\sin^2\phi}\right) d\gamma_{ml}, \tag{56}$$

equation (55) becomes

$$\begin{aligned} \mathrm{BER} &= \frac{1}{\pi} \int_0^{\pi/2} \prod_{m=1}^{L} M_{\gamma_{ml}}\left(\frac{-1}{\sin^2\phi}\right) d\phi \\ &= \frac{1}{\pi} \int_0^{\pi/2} \prod_{m=1}^{L} \left(\frac{1}{1+\frac{\bar{\gamma_{ml}}}{\sin^2\phi}}\right) d\phi, \end{aligned} \tag{57}$$

where the MGF for a Rayleigh channel is defined as $M_{\gamma_{ml}}(s) = (1 - s\bar{\gamma_{ml}})^{-1}$.

An upper bound of the BER in equation (57) is obtained by applying the Chernoff bound $Q(x) \leq \exp(-x^2/2)$ [36]. Assuming BPSK and i.i.d. Rayleigh fading, this results in

$$\begin{aligned} &\mathrm{BER} \leq \\ &\int_0^\infty \exp\left(-\sum_{m=1}^{L} \gamma_{ml}\right) p(\gamma_{1l}) p(\gamma_{2l}) \cdots p(\gamma_{Ll}) d\gamma_{1l} d\gamma_{2l} \cdots d\gamma_{Ll} \\ &= \prod_{m=1}^{L} \frac{1}{1+\bar{\gamma_{ml}}} \approx \frac{1}{\prod_{m=1}^{L} \bar{\gamma_{ml}}}. \end{aligned} \tag{58}$$

where the approximation holds only for high SNR values.

APPENDIX B

In this section, a simpler expression for the SNR of U1 given by

$$\mathrm{SNR}_{\mathrm{U1}} = \mathbf{a}_1^H \mathbf{R}_{w_1}^{-1} \mathbf{a}_1 \sigma_{\mathrm{s}}^2 \tag{59}$$

is derived. Using the Sherman-Morrison-Woodbury formula

$$(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1}, \tag{60}$$

where $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ are arbitrary matrices, and substituting the following matrices $\mathbf{A} = \mathbf{N}_1$, $\mathbf{B} = \mathbf{a}_2$, $\mathbf{C} = \mathbf{a}_2^H \sigma_{\mathrm{s}}^2$, $\mathbf{D} = -1$, then $\mathbf{R}_{w_1}^{-1}$ from equation (21) becomes

$$\mathbf{R}_{w_1}^{-1} = \mathbf{N}_1^{-1} - \sigma_{\mathrm{s}}^2 \mathbf{N}_1^{-1} \left(1 + \sigma_{\mathrm{s}}^2 \mathbf{a}_2^H \mathbf{N}_1^{-1} \mathbf{a}_2\right)^{-1} \mathbf{a}_2 \mathbf{a}_2^H \mathbf{N}_1^{-1}. \tag{61}$$

Now, substituting $\mathbf{R}_{w_1}^{-1}$ in equation (59) leads to

$$\begin{aligned} &\mathrm{SNR}_{\mathrm{U1}} \\ &= \mathbf{a}_1^H \left(\mathbf{N}_1^{-1} - \sigma_{\mathrm{s}}^2 \mathbf{N}_1^{-1} \left(1 + \sigma_{\mathrm{s}}^2 \mathbf{a}_2^H \mathbf{N}_1^{-1} \mathbf{a}_2\right)^{-1} \mathbf{a}_2 \mathbf{a}_2^H \mathbf{N}_1^{-1}\right) \mathbf{a}_1 \sigma_{\mathrm{s}}^2 \\ &= \sigma_{\mathrm{s}}^2 \mathbf{a}_1^H \mathbf{N}_1^{-1} \mathbf{a}_1 - \sigma_{\mathrm{s}}^2 \left(1 + \sigma_{\mathrm{s}}^2 \mathbf{a}_2^H \mathbf{N}_1^{-1} \mathbf{a}_2\right)^{-1} |\mathbf{a}_2^H \mathbf{N}_1^{-1} \mathbf{a}_1|^2 \sigma_{\mathrm{s}}^2. \end{aligned} \tag{62}$$

Using the definition of vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ from Section II-B1 leads to

$$\begin{aligned} \mathbf{a}_1^H \mathbf{N}_1^{-1} \mathbf{a}_1 &= \frac{P_1 |h_{11}|^2}{\sigma_{\mathrm{n}}^2} + \frac{P_2 |h_{21}|^2}{\sigma_{\mathrm{n}}^2 \left(1 + \frac{\sqrt{P_1} h_{11}}{\sqrt{P_2} g_{21}} + \frac{\sqrt{P_2} h_{21}}{\sqrt{P_1} g_{12}}\right)} \\ |\mathbf{a}_2^H \mathbf{N}_1^{-1} \mathbf{a}_1|^2 &= \frac{P_1 P_2 |h_{11}|^2 |h_{21}|^2}{\sigma_{\mathrm{n}}^4 \left(1 + \frac{\sqrt{P_1} h_{11}}{\sqrt{P_2} g_{21}} + \frac{\sqrt{P_2} h_{21}}{\sqrt{P_1} g_{12}}\right)^2} \\ \mathbf{a}_2^H \mathbf{N}_1^{-1} \mathbf{a}_2 &= \frac{P_2 |h_{21}|^2}{\sigma_{\mathrm{n}}^2} + \frac{P_1 |h_{11}|^2}{\sigma_{\mathrm{n}}^2 \left(1 + \frac{\sqrt{P_1} h_{11}}{\sqrt{P_2} g_{21}} + \frac{\sqrt{P_2} h_{21}}{\sqrt{P_1} g_{12}}\right)}. \end{aligned} \tag{63}$$

Substituting equations (63) in (62) leads to

$$\begin{aligned} &\mathrm{SNR}_{\mathrm{U1}} = \sigma_{\mathrm{s}}^2 \left(\frac{P_1 |h_{11}|^2}{\sigma_{\mathrm{n}}^2} + \frac{P_2 |h_{21}|^2}{\sigma_{\mathrm{n}}^2 \left(1 + \frac{\sqrt{P_1} h_{11}}{\sqrt{P_2} g_{21}} + \frac{\sqrt{P_2} h_{21}}{\sqrt{P_1} g_{12}}\right)}\right) - \\ &\left(1 + \sigma_{\mathrm{s}}^2 \left(\frac{P_2 |h_{21}|^2}{\sigma_{\mathrm{n}}^2} + \frac{P_1 |h_{11}|^2}{\sigma_{\mathrm{n}}^2 \left(1 + \frac{\sqrt{P_1} h_{11}}{\sqrt{P_2} g_{21}} + \frac{\sqrt{P_2} h_{21}}{\sqrt{P_1} g_{12}}\right)}\right)\right)^{-1} \\ &\frac{\sigma_{\mathrm{s}}^4 P_1 P_2 |h_{11}|^2 |h_{21}|^2}{\sigma_{\mathrm{n}}^4 \left(1 + \frac{\sqrt{P_1} h_{11}}{\sqrt{P_2} g_{21}} + \frac{\sqrt{P_2} h_{21}}{\sqrt{P_1} g_{12}}\right)^2} \\ &= \frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2} P_1 |h_{11}|^2 + \frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2} P_2 |h_{21}|^2 \times \\ &\left(\frac{1 + \frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2} P_2 |h_{21}|^2}{\frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2} P_1 |h_{11}|^2 + \left(\frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2} P_2 |h_{21}|^2 + 1\right) \left(\frac{\sqrt{P_1} h_{11}}{\sqrt{P_2} g_{21}} + \frac{\sqrt{P_2} h_{21}}{\sqrt{P_1} g_{12}} + 1\right)}\right). \end{aligned} \tag{64}$$

Following an analogous derivation, $\mathrm{SNR}_{\mathrm{U2}}$ is given by

$$\begin{aligned} &\mathrm{SNR}_{\mathrm{U2}} = \frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2} P_2 |h_{22}|^2 + \frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2} P_1 |h_{12}|^2 \times \\ &\left(\frac{1 + \frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2} P_1 |h_{12}|^2}{\frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2} P_2 |h_{22}|^2 + \left(\frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2} P_1 |h_{12}|^2 + 1\right) \left(\frac{\sqrt{P_1} h_{12}}{\sqrt{P_2} g_{21}} + \frac{\sqrt{P_2} h_{22}}{\sqrt{P_1} g_{12}} + 1\right)}\right). \end{aligned} \tag{65}$$

APPENDIX C

In order to find a closed form expression for equation (48), an expression for the determinants of the matrices $\mathbf{R}_{y_l}$ and $\mathbf{R}_{w_l}$ is needed. Expanding the matrix $\mathbf{R}_{y_l}$ results in equation (66) and expanding the matrix $\mathbf{R}_{w_l}$ results in $\mathbf{R}_{w_l} = \sum_{\substack{m=1 \\ m \neq l}}^{L} \mathbf{a}_m \mathbf{a}_m^H \sigma_{\mathrm{s}}^2 + \mathbf{N}_l$, presented in equation (67) for $l = 1$.

Finally, by using the Laplace expansion for the determinants $|\mathbf{R}_{y_l}|$ and $|\mathbf{R}_{w_l}|$, equation (68) is obtained.

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

$$\mathbf{R}_{y_l} = \mathbf{a}_1\mathbf{a}_1^H\sigma_{\mathrm{s}}^2 + \mathbf{a}_2\mathbf{a}_2^H\sigma_{\mathrm{s}}^2 + \cdots + \mathbf{a}_L\mathbf{a}_L^H\sigma_{\mathrm{s}}^2 + \mathbf{N}_l = \begin{bmatrix} 1+\gamma_{1l} & 0 & \cdots & 0 & \sqrt{\gamma_{1l}\Upsilon_1} \\ 0 & 1+\gamma_{2l} & \cdots & 0 & \sqrt{\gamma_{2l}\Upsilon_2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1+\gamma_{Ll} & \sqrt{\gamma_{Ll}\Upsilon_L} \\ \sqrt{\gamma_{1l}\Upsilon_1} & \sqrt{\gamma_{2l}\Upsilon_2} & \cdots & \sqrt{\gamma_{Ll}\Upsilon_L} & \Upsilon_1+\Upsilon_2+\cdots+\Upsilon_L+D_1+1 \end{bmatrix} \tag{66}$$

$$\mathbf{R}_{w_1} = \mathbf{a}_2\mathbf{a}_2^H\sigma_{\mathrm{s}}^2 + \cdots + \mathbf{a}_L\mathbf{a}_L^H\sigma_{\mathrm{s}}^2 + \mathbf{N}_l = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1+\gamma_{21} & \cdots & 0 & \sqrt{\gamma_{21}\Upsilon_2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1+\gamma_{L1} & \sqrt{\gamma_{L1}\Upsilon_L} \\ 0 & \sqrt{\gamma_{21}\Upsilon_2} & \cdots & \sqrt{\gamma_{L1}\Upsilon_L} & \Upsilon_2+\cdots+\Upsilon_L+D_1+1 \end{bmatrix} \tag{67}$$

$$\frac{|\mathbf{R}_{y_l}|}{|\mathbf{R}_{w_l}|} = 1+\gamma_{ll} + \left(\sum_{\substack{m=1\\m\neq l}}^{L}\sqrt{\gamma_{ml}}\right)^2 \frac{\prod_{\substack{m=1\\m\neq l}}^{L}(1+\gamma_{ml})}{\sum_{\substack{m=1\\m\neq l}}^{L}\prod_{\substack{q=1\\q\neq l\\q\neq m}}^{L}(1+\gamma_{ql})\left(\sum_{\substack{r=1\\r\neq m}}^{L}\sqrt{\gamma_{rl}}\right)^2 + (D_l+1)\prod_{\substack{m=1\\m\neq l}}^{L}(1+\gamma_{ml})} = 1+\gamma_{ll} + \frac{\left(\sum_{\substack{m=1\\m\neq l}}^{L}\sqrt{\gamma_{ml}}\right)^2}{\sum_{\substack{m=1\\m\neq l}}^{L}\frac{\left(\sum_{\substack{r=1\\r\neq m}}^{L}\sqrt{\gamma_{rl}}\right)^2}{1+\gamma_{ml}} + D_l + 1} \tag{68}$$

ACKNOWLEDGMENTS

This research work was carried out at the ESAT Laboratory of KU Leuven, in the frame of KU Leuven Research Council PFV/10/002 (OPTEC), and the Belgian Programme on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office "Belgian network on stochastic modelling, analysis, design and optimization of communication systems (BESTCOM)" 2012-2017. Rodolfo Torrea-Duran acknowledges the support of the Mexican National Council for Science and Technology (CONACYT). Fernando Rosas acknowledges the support of the European Union's H2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 702981. The scientific responsibility is assumed by its authors. A part of this work was published in [40].

REFERENCES

[1] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016-2021", Online: http : //www.cisco.com/c/en/us/solutions/index.html Feb. 2016.

[2] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. T. Sukhavasi, C. Patel, and S. Geirhofer, "Network densification: The dominant theme for wireless evolution into 5G, *IEEE Communications Magazine*, vol. 52, no. 2, pp. 8289, Feb. 2014.

[3] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-Dense Networks: A Survey, *IEEE Communications Surveys & Tutorials*, no. 99, May 2016.

[4] H. Shi, R. Venkatesha Prasad, E. Onur, I.G.M.M. Niemegeers, "Fairness in Wireless Networks: Measures, Issues, and Challenges", *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 5–24, Feb. 2014.

[5] A. UI Quddus C. Kosta, B. Hunt and R. Tafazolli, "On interference avoidance through inter-cell interference coordination (ICIC) based on OFDMA mobile systems", *IEEE Communications Surveys & Tutorials*, pp. 1–23, 2013.

[6] R. Kwan and C. Leung, "A survey of scheduling and interference mitigation in LTE", *Journal of Electrical and Computer Engineering*, 2010.

[7] A. Gjendemsjo, D. Gesbert, G. E. Oien, and S. G. Kiani, "Binary Power Control for Sum Rate Maximization over Multiple Interfering Links", *IEEE Transactions on Wireless Communications*, vol.7, no.8, pp.3164-3173, Aug. 2008.

[8] D. Julian, M. Chiang, D. O'Neill, and S. Boyd, "QoS and Fairness Constrained Convex Optimization of Resource Allocation for Wireless Cellular and Ad Hoc Networks", *IEEE INFOCOM*, pp.477-486, Jun. 2002.

[9] J. Wang, Y. Zhang, H. Hui, and N. Zhang, "QoS-aware proportional fair energy-efficient resource allocation with imperfect CSI in downlink OFDMA systems", *IEEE 26th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Hong Kong, pp.1116–1120, Sep. 2015.

[10] R. Fritzsche, P. Rost, and G. P. Fettweis, "Robust Rate Adaptation and Proportional Fair Scheduling with Imperfect CSI", https://arxiv.org/pdf/1504.03803.pdf, Apr. 2015.

[11] S. Chang Liew and Y. J. Zhang, "Proportional Fairness in Multi-Channel Multi-Rate Wireless NetworksPart I: The Case of Deterministic Channels

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2797128, IEEE Access

with Application to AP Association Problem in Large-Scale WLAN", *IEEE Transactions on Wireless Communications*, vol.7, no.9, pp.3446–3456, Sep. 2008.

[12] H. Minn, and D. Munoz, "Channel Knowledge Acquisition in Relay and Multipoint-to-Multipoint Transmission Systems", *IEEE Transactions on Vehicular Technology*, vol.64, no.4, pp.1416–1434, Apr. 2015.

[13] A. Gatherer, "Death by Starvation?: Backhaul and 5G", *IEEE ComSoc Technology News*, http://www.comsoc.org/ctn/death-starvation-backhaul-and-5g. Sep. 2015.

[14] A. Sendonaris, E. Erkip, and B. Aazhang, User cooperative diversity - part I: system description, *IEEE Transactions on Communications*, vol. 51, no. 11, pp. 1927-1938, Nov. 2003.

[15] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, Cooperative diversity in wireless networks: efficient protocols and outage behavior, *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062-3080, Dec. 2004.

[16] S. Zhang, S. Liew, and P. Lam, "Hot topic: physical layer network coding", *ACM MOBICOM*, 2006.

[17] R.H.Y. Louie, Y. Li, and B. Vucetic, "Practical Physical Layer Network Coding for Two-Way Relay Channels: Performance Analysis and Comparison", *IEEE Transactions on Wireless Communications*, vol. 9, no. 2, Feb. 2010.

[18] P. Popovski and H. Yomo, "Physical Network Coding in Two-Way Wireless Relay Channels", *IEEE International Conference on Communications (ICC)*, pp.707-712, June 2007.

[19] H.Yang, K. Lee, and J. Chun, "Zero-forcing Based Two-phased Relaying", *IEEE International Conference on Communications (ICC)*, 2007.

[20] H. Yang, B. Jung, and J. Chun, "Zero-forcing Based Two-phased Relaying with Multiple Mobile Stations", *Asilomar*, 2008.

[21] I.J. Baik and S. Y. Chung, "Network Coding for Two-Way Relay Channels Using Lattices", *IEEE International Conference on Communications (ICC)*, 2008.

[22] S. Smirani, M. Kamoun, M. Sarkiss, A. Zaidi, and P. Duhamel, "Achievable Rate Regions for Two-Way Relay Channel Using Nested Lattice Coding", *IEEE Transactions on Wireless Communications*, vol. 13, no. 10, Feb. 2014.

[23] C. Feng, D. Silva, and F. Kschischang, "An Algebraic Approach to Physical-Layer Network Coding", *IEEE Transactions on Information Theory*, vol. 59, no. 11, Nov. 2013.

[24] B. Nazer and M. Gastpar, "Compute-and-Forward: Harnessing Interference Through Structured Codes", *IEEE Transactions on Information Theory*, vol. 57, no. 10, Oct. 2011.

[25] J.S. Lemos and F. A. Monteiro, "Full-duplex Massive MIMO with Physical Layer Network Coding for the Two-Way Relay Channel", *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Jul. 2016.

[26] T-T Chan and T-M Lok, "Interference Alignment with Physical Layer Network Coding in MIMO Relay Channels", *IEEE International Conference on Communications (ICC)*, May 2016.

[27] J. Guo, T. Yang, J. Yuan, J. A. Zhang, "Linear Vector Physical-Layer Network Coding for MIMO Two-Way Relay Channels: Design and Performance Analysis", *IEEE Transactions on Communications*, vol. 63, no. 7, Jul. 2015.

[28] H. Zhang and L. Cai, "HePNC: A Cross-Layer Design for MIMO Networks with Asymmetric Two-Way Relay Channel", *IEEE Global Communications Conference*, Dec. 2015.

[29] S. Katti, I. Maric, A. Goldsmith, D. Katabi, and M. Medard, "Joint Relaying and Network Coding in Wireless Networks", *IEEE International Symposium on Information Theory (ISIT)*, Nice, France, 2007.

[30] H. Q. Lai, and K. J. R. Liu, "Space Time Network Coding", *IEEE Transactions on Signal Processing*, vol. 59, no. 4, Apr. 2011.

[31] K. Xiong, P. Fan, H. C. Yang,and K. B. Letaief, "Space Time Network Coding with Overhearing Relays", *IEEE Transactions on Signal Processing*, vol. 13, no. 7, Jul. 2014.

[32] Y. Zhang, K. Xiong, P. Fan, X. Di, and X. Zhou, "Outage Performance of Space Time Network Coding with Overhearing AF Relays", *IEEE Communications Letters*, vol. 19, no. 12, Dec. 2015.

[33] M. Damen and A. Hammons, "Delay-tolerent distributed-TAST codes for cooperative diversity", *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 2941–2956, Oct. 2007.

[34] J. Zhang and C. Shen and G. Deng and Y. Wang, "Timing and Frequency Synchronization for Cooperative Relay Networks", *IEEE Vehicular Technology Conference (VTC)*, Sep. 2013.

[35] R. Jain, D. Chiu, and W. Hawe, A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Systems", *Digital Equipment Corporation*, Technical Report DEC-TR-301, Tech. Rep., 1984.

[36] A. Goldsmith, "Wireless Communications," *Cambridge University Press*, 2005.

[37] D. Tse and P. Viswanath, "Fundamentals of Wireless Communications," *Cambridge University Press*, 2005.

[38] T. Cover and J. Thomas, "Elements of Information Theory," *John Wiley and Sons*, 1991.

[39] Simon M. K. and M. S. Alouini, "Digital Communication over Fading Channels: A Unified Approach to Performance Analysis", New York, *John Wiley and Sons*, 2000.

[40] R. Torrea-Duran, F. Rosas, Z.K.Z. Khan, S. Pollin, P. Tsiaflakis, and M. Moonen, "Double Relay Communication Protocol for Bandwidth Management in Cellular Systems", *European Signal Processing Conference (EUSIPCO 2015)*, Nice, France, Aug. 2015.

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.