RESEARCH   **Open Access**

Cooperative relaying in OFDMA networks based on the joint use of hierarchical modulation and link adaptation

Anis Jdidi and Tijani Chahed[*]

Abstract

In cooperative relaying, the base station broadcasts a signal to the destination, the relay overhears it and forwards it in a next time slot to the destination. The latter combines the two received copies of the signal in order to reconstruct the original one. In this study, we consider a relay-based OFDMA network, as is the case of long term evolution advanced systems. We, first, make use of hierarchical modulation to send additional information to the relay so as to enable it to reconstruct a more robust copy of the original signal, and, second, adapt the transmission from the relay to the destination by taking advantage of the typically good radio conditions between them so as to reduce the cost of additional resources needed by relaying and hence improve the overall system capacity. We model such a system at the user level, considering a realistic arrival and departure setting and quantify the gains thus achieved in a cross-layer manner, in terms of throughput and blocking probability. We, eventually, propose an enhancement to the afore-mentioned scheme that takes advantage of the good radio conditions of users who are close to the base station to send them an additional stream using hierarchical modulation.

1 Introduction

In order to improve the system performance in terms of capacity and/or coverage, relays have been proposed to OFDMA-based systems in general and long term evolution advanced (LTE-A) systems in particular [1]. In such a system, a relay station (RS) relays information from the base station (BS) to the Subscriber Station (SS), i.e., user, who suffers from poor radio conditions in its direct link with the BS. This allows the SS to take advantage of the spatial diversity gain without using multiple antennas.

We consider in this study cooperative relaying, wherein the SS can combine the two signals received from the direct BS-SS and indirect BS-RS-SS links so as to reconstruct a robust copy of the original signal.

To relay information to the destination, the RS typically uses one of two relaying techniques: amplify and forward (AF) or decode and forward (DF) [2]. AF is a simple scheme, where the RS only amplifies the

amplitude of the received signal from the BS and then forwards it to the destination. One of the major drawbacks of this technique is the amplification of the noise. With DF, the RS first decodes the signal overheard from the BS and then re-encodes it and transmits it to the destination. Apart from the extra cost and delay resulting from the decoding and encoding steps, this technique may be inefficient if the RS is not able to successfully decode the received signal from the BS. This shortcoming can be efficiently overcome by the use of hierarchical modulation (HM).

In effect, with HM, also termed superposition coding or yet embedded constellation [3], the idea is to send an additional data stream to a user with good radio conditions on a subcarrier initially allocated to a user with worse radio conditions. Combining HM with relaying, the idea is to use the fact that the RS enjoys typically better radio conditions with the BS compared to the destination to send an additional information to the RS on the same subcarrier that was initially allocated to the destination. The RS, as it enjoys good radio conditions, would be able to decode the two signals: the primary one sent to the destination and the additional

* Correspondence: tijani.chahed@it-sudparis.eu
Telecom SudParis, UMR CNRS 5157, 9 rue C. Fourier, 91011 Evry Cedex, France

Springer

information sent to it, and to generate a more robust copy of the initial signal compared to the case of classical relaying where the RS will use only the overheard signal to decode and encode a new copy of the initial signal.

Moreover, as the link between the RS and SS is typically of good radio conditions, one can also use high rate modulation to forward this more robust copy of the signal from the RS to the SS. In doing so, we will be able to reduce the amount of resources needed for relaying which will allow to increase the overall system capacity.

In this study, we propose and model the joint use of HM and link adaptive cooperative relaying in an OFDMA system at the user level, for a realistic dynamic setting where users come to the system at random time epochs and leave it after a finite duration corresponding to the transfer of a file of a given size. We relate this so-called *flow-level* modeling to the lower layer resource allocation mechanisms in a cross-layer manner. We also propose an enhancement that takes advantage of the good radio conditions of users who are close to the BS to send them an additional stream using HM. This in turn allows to enhance the system performance even further.

Other works studied the use of HM in cooperative relaying systems. In [4], Peng et al. propose hierarchical cooperative relaying to support both unicast and multicast services. The BS utilizes two-dimensional source coding schemes and transmits the corresponding data streams within two time slots in one frame. In [5,6], the authors propose to send an additional information using HM via the indirect link in order to improve the bit error rate (BER) at reception. In [2,7], the authors propose to optimize the use of HM in order to also minimize the BER at the destination. In [8], Jeong-Chul et al. study the use of HM in the case of a multiple relay scheme. And in [9], Larsson and Vojcic investigate the case of a cooperative transmit diversity based on the use of HM in the case of two sources that communicate directly with one destination. These studies focus solely on the physical layer, notably on the BER improvement, and do not address the capacity issue, along with the gains achieved, at the *user level*, which is the focus of our present study.

Moreover, these studies do not consider also the problem of the additional resources needed for relaying, and which we combat, in this study, using the aforementioned link adaptation. In [10], Popovski and de Carvalho do indeed adapt the transmission of the relay to the radio conditions of the RS-SS link, but their work does not make use of cooperative relaying in the sense that the destination does not combine two copies of

information, one from the direct link and one from the indirect one, in order to reconstruct the original signal. In [11], Whang et al. develop a multiuser cooperative scheme, where HM is used between two destinations; the RS decodes the two overheard signals, without additional information, and relays them to the destinations using HM also.

In addition, all of these studies consider a static scenario with a fixed number of users in the system, whereas in our study, we consider, as indicated previously, a dynamic setting where, again, (data) users come to the system and leave it after a finite duration, corresponding to the transfer of their files.

The remainder of this study is organized as follows. In Section 2, we describe the relay-based OFDMA system and the way resources are allocated in such a setting. In Section 3, we describe HM. In Section 4, we detail the use of HM and our link adaptation relaying proposal. In Section 5, we present our flow-level modeling and derive several performance metrics. In Section 6, we describe the enhancement proposed to our scheme. Section 7 contains numerical and simulation results. Section 8 eventually concludes the article.

2 Relay-based OFDMA system

Let us first describe a downlink OFDMA cell with a single base station. Using OFDMA, the total bandwidth W is divided into N orthogonal subcarriers which can be shared by several users in the same time slot. Depending on the signal-to-noise ratio (SNR), each user will be assigned a coding and modulation scheme, as dictated by the adaptive modulation and coding (AMC) feature used in IEEE802.16 WiMAX or LTE; with a larger constellation size and coding scheme for users with higher SNR [12].

Let $\text{SNR}_{s,n}$ be the SNR experienced by user $s, s = 1,...,S$, where S is the number of users in the system, on subcarrier $n, n = 1,...,N$. It is given by

$$\text{SNR}_{s,n} = \frac{P_{s,n}|h_{s,n}|^2}{(\sigma_{s,n})^2} \qquad (1)$$

where $P_{s,n}$ is the transmit power of user s on subcarrier n, $h_{s,n}$ is its channel gain which includes the effects of fast fading, path loss and shadowing and $\sigma_{s,n}$ is the variance of the additive white Gaussian noise (AWGN) with zero mean.

The number of bits $b_{s,n}$ that can be sent by user s on subcarrier n is given by [13]:

$$b_{s,n} = \log_2\left(1 + \frac{\text{SNR}_{s,n}}{\Gamma}\right) \qquad (2)$$

where Γ represents the SNR-gap which is used to ensure some quality of service (QoS) for an uncoded QAM system and is given by:

$$\Gamma = \frac{1}{3}\left[\mathcal{Q}^{-1}\left(\frac{P_e}{4}\right)\right]^2 \quad , \Gamma \geq 1 \qquad (3)$$

where $\mathcal{Q}^{-1}(.)$ is the inverse standard Marcum Q-function and P_e is the target BER.

Let us now, consider that I RSs are added to the plain OFDMA architecture, described above, in order to assist users who cannot decode successfully their signals on the direct link (BS-SSs). Let P^R denotes the probability that the user to be served does indeed require the help of an RS to be able to successfully decode the original signal.

Let us further denote by P the probability that the user who needs the assistance of an RS is indeed covered by an RS. This probability depends on the number and the position of RSs in the cell. P is given by:

$$P = \min\left(\frac{I}{I_{\text{threshold}}}, 1\right) \qquad (4)$$

where $I_{\text{threshold}}$ denotes the minimum number of RSs in the cell in order to cover it all, and hence be able to assist any user who needs relaying, at any time.

In order to determine $I_{\text{threshold}}$, we assume that the cell is represented by a disc with, say, two regions ($J = 2$) corresponding to users with good radio conditions (inner region) and those with worse ones (outer region). RSs are placed at the edge of the first region as shown in Figure 1. Assuming that the areas covered by RSs are tangent to each other, we obtain:

$$I_{\text{threshold}} = \frac{\pi}{\alpha} \qquad (5)$$

where α is equal to $\cos^{-1}\left(1 - \frac{(r^2 - r^1)^2}{2(r^1)^2}\right)$ and r^j denotes the radius of region $j, j = 1, 2$.
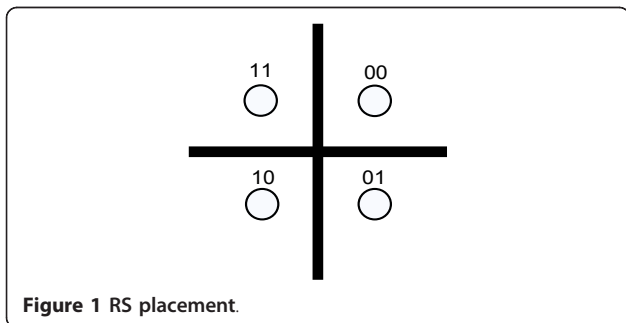
3 Hierarchical modulation

As indicated above, HM is a technique that allows to send an additional stream to a user with good radio conditions on a subcarrier that was initially allocated to a first user generally with worse radio conditions. In what follows, indices k and k^* stand, respectively, for user with bad radio conditions and user with good ones ($h_{k,n} << h_{k^*,n}$).

Let us consider, for instance, that each user (k and k^*) would receive a 4-QAM constellation signal (see, Figure 2). Using HM, we actually combine two such constellations into a single one, as shown in Figure 3. In the resulting 4/16-QAM constellation, each symbol is composed of 4 bits: the two most significant bits–to the left–represent the basic information sent to user k (also termed the first level of hierarchy) and the other 2 bits represent the additional information sent to user k^* (the second level of hierarchy).

As of decoding, to user k^*, the received signal is clear as it enjoys good radio conditions and hence using successive interference cancelation (SIC) [13,14], the constellation of user k can be easily filtered out before decoding its own signal.

User k sees the signal of user k^* as extra (Gaussian) noise surrounding his own signal. This extra noise is not very significant because of the small amount of power allocated to user k^* [13]. And hence the possibility to transmit simultaneously to two users on the same subcarrier without degrading the capacity of the link.

Generally, we denote by L/M-QAM the constellation that results from the use of HM, where L ($L = 2^{2l}$) and M ($M = 2^{2m}$) represent the constellation size for,

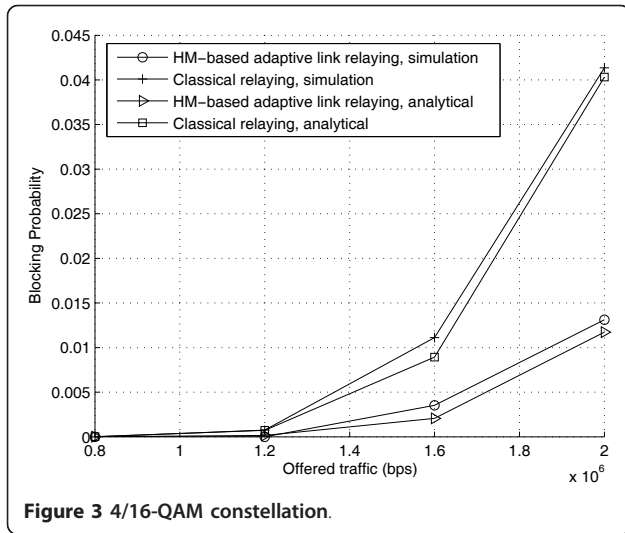**Figure 1 RS placement.**

Figure 2 4-QAM constellation.
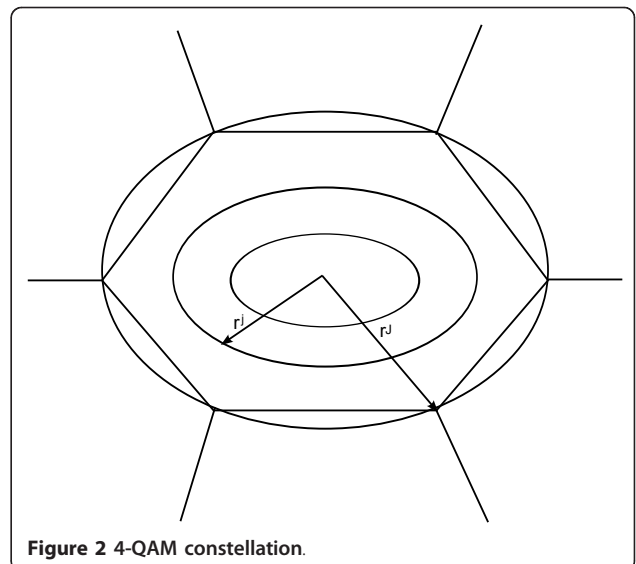
Figure 3 4/16-QAM constellation.

respectively, the first and second levels of hierarchy. In this case, l bits are sent to user k and $(m - l)$ bits are sent to user k^* [12].

For the square constellation 4/M-QAM, in order to ensure a decoding without error, the amounts of power allocated to the two users, on the shared subcarrier n, are given by [15]:

$$P_{k^*,n} = \frac{d^1_{k^*,n}\left(2^{b_{k^*,n}} - 1\right)}{6} \text{ and } P_{k,n} = \frac{d^2_{k,n}\left(2^{b_{k,n}} - 1\right)}{6} \quad (6)$$

where $b_{k,n}$ and $b_{k^*,n}$ correspond, respectively, to the number of bits that we want to transmit to users k and k^*.

At decoding, $b_{k,n}$ and $b_{k^*,n}$ are given, respectively, by the Shannon formula:

$$b_{k^*,n} = \log_2\left(1 + \frac{P_{k^*,n}|h_{k^*,n}|^2}{\Gamma\left(\sigma_{k^*,n}\right)^2}\right)$$

and

$$b_{k,n} = \log_2\left(1 + \frac{P_{k,n}|h_{k,n}|^2}{\Gamma\left((\sigma_{k,n})^2 + P_{k^*,n}|h_{k,n}|^2\right)}\right) \quad (7)$$

where, again, $h_{k/k^*,n}$ and $\sigma_{k/k^*,n}$, denote, respectively, the channel gain and the variance of the AWGN with zero mean of users k and k^*.

$d^1_{k^*,n}$ and $d^2_{k,n}$ represent, respectively, the minimum distance between the flectious symbol in the constellation of users k and k^*, respectively, on subcarrier n (please refer to, Figure 3). In our case, $d^1_{k^*,n}$ is related to $d^2_{k,n}$ by:

$$d^2_{k,n} = d'_n + \left(\frac{\sqrt{M}}{2} - 1\right)d^1_{k^*,n} \quad (8)$$

where $d^1_{k^*,n}$ (respectively, d'_n) is given by:

$$\sqrt{\frac{6\Gamma(\sigma_{k^*,n})^2}{(h_{k^*,n})^2}} \left(\text{resp.} \sqrt{\frac{6\Gamma(\sigma_{k,n})^2}{(h_{k,n})^2}}\right) \quad (9)$$

Note that the modulation parameter D, defined by $D = \frac{d^1}{d'}$, indicates the degree of protection of each stream. This parameter varies between 0 and 1: if $D = 0$, the total sent signal represents a 4-QAM constellation whereas if $D = 1$, it represents 16-QAM constellation.

4 Use of hierarchical modulation and link adaptive relaying

We now describe the operation of HM and link adaptive relaying in the relay-based OFDMA system described in Section 2. We assume, without loss of generality, that users are of two types: users of type 1 enjoy good radio conditions and can, hence, decode the signal properly, and users of type 2 which have lower radio conditions and which can or not decode their signal successfully. When they cannot, we will make use of an RS, based on DF, to help them decode their signal.

We assume that the BS has knowledge of the channel state information (CSI) of all links: BS-RSs, BS-SSs, and RSs-SSs. We denote, respectively, by SNR_k, SNR_{sr}, and SNR_{rd}, the SNRs of the BS-SS, BS-RS, and RS-SS links, as shown in Figure 4. Index k stands for a given user of type 2 who cannot decode his information bits on the direct link (BS-SS).

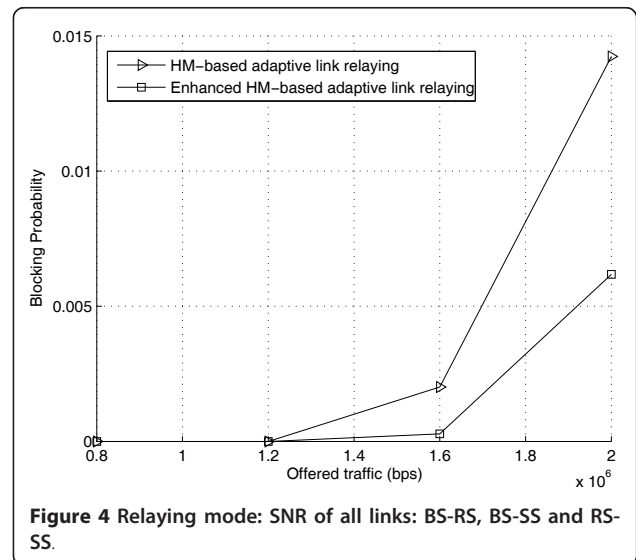In classical cooperative relaying using DF, the BS broadcasts a M-QAM signal A to a user k with low

Figure 4 Relaying mode: SNR of all links: BS-RS, BS-SS and RS-SS.

radio conditions, the RS overhears it, decodes it, encodes a new copy $\bar{A}$ and forwards it in the next time slot to this given user.

We suppose that the destination uses a simple maximal ratio combining (MRC) combiner in order to decode its signal. In this case and using DF-based relaying, the SNR achieved by a given user k after the two steps described above is given by:

$$\overline{SNR_k} = \min((SNR_k + SNR_{coop}), SNR_{sr}) \qquad (10)$$

where SNR_{coop} is equal to SNR_{rd} if $SNR_k < SNR_{rd}$. Note that if $SNR_k > SNR_{sr}$, relaying is useless and $\overline{SNR_k} = SNR_k$.

Let us define by BER_{coop} the BER achieved over the indirect link (BS-RS-SS). It is given by:

$$BER_{coop} = 1 - (1 - BER_{sr})(1 - BER_{rd}) \qquad (11)$$

where BER_{sr} and BER_{rd} stand, respectively, for the BER of the BS-RS and RS-SS links.

The BER_k of a given user k for an M-QAM system over an AWGN channel is related to its SNR_k through the following equation [16]:

$$BER_k = 0.2 \exp\left(-\frac{3SNR_k}{2(M-1)}\right) \qquad (12)$$

Based on Equations (10), (11), and (12), we see that the quality of the decoding at the destination depends, notably, on the quality of the signal decoded and encoded at the RS (depends notably on SNR_{sr} and SNR_{rd}). If the overheard signal cannot be decoded correctly by the RS, the encoded copy, $\bar{A}$, will also be corrupted, and hence the relaying will be inefficient. And so, in order to improve the decoding at the RS, we propose the use of HM.

Using HM, the BS will be able to send two different signals, which we term A and B, to the SS and RS by adapting the transmission to each entity. Signal A represents the original signal sent to the SS with an amount of power equal to θP_n and signal B represents the additional information sent to the RS with an amount of power equal $(1 - \theta)P_n$. θ is a real number that verifies $0 < \theta < 1$ and can be obtained using Equations (6) and P_n represents the total amount of power allocated to the shared subcarrier n. In this case, and as described in Section 3, the RS, owing to its good radio conditions, will be able to decode the two signals A and B while the SS will be able to receive only signal A.

Please note that B can be a simple copy of the initial signal A or control information that allows to ensure the success of decoding at the RS, as considered in [5]. The latter case corresponds to a coded cooperation scheme, with distributed channel coding [17], and allows to ensure better error protection.

We suppose, for simplicity, that signal B contains a second copy of the original signal A. This results in a spatial diversity gain at the RS. Using MRC, the RS will now be able to decode the original signal with an SNR equal to $SNR_k + SNR_{sr}$. Since the BER is inversely proportional to the received SNR (Equation (12)), and because of the higher received SNR ($SNR_k + SNR_{sr} > SNR_k$), the RS will thus be able to decode and encode a more robust signal $\bar{A}$ with a lower BER. This, in turn, will ensure a decoding without error at the destination as well.

Moreover, using classical relaying, the RS relays the signal to the destination using the same rate (same constellation size M) as that of the overheard signal, which is typically low because of the worst good radio conditions of the direct link (BS-SS). The RS-SS link is, however, typically of good radio conditions, and may support higher bit rates. We thus propose to take advantage of this fact and to forward the relayed signal $\bar{A}$ at a higher bit rate, or equivalently, in a lower number of time slots than classically needed, assuming for instance Time Division Multiplexing (TDM) scheduling.

Formally, if $BER_{rd} \leq 10^{-2}$, Equation (12) can be inverted and yields the maximum constellation size M' that the RS can use for a given SNR_{rd} and pre-fixed target BER_{rd} as [16]:

$$\log_2(M') = \log_2\left(1 + \frac{3SNR_{rd}}{2K_0}\right) \qquad (13)$$

where $K_0 = -\ln(5BER_{rd})$. Since the RS-SS link presents good radio conditions (higher SNR_{rd}), the RS is thus able to use a higher constellation size M' to serve the destination without affecting the target BER (BER_{rd} in this case). This, in turn, yields a higher spectral efficiency by increasing the signal rate on the RS-SS link.

In this study, we take one time slot as our smallest time unit. In this case, we shall assign, using TDM, $\beta = \dfrac{\log_2(M')}{\log_2(M)}$ consecutive time slots to each user to be served ($\beta > 1$) so that the RS to SS relaying operation is done using one time slot. In this one time slot, the RS actually relays β different (robust) signals to the target destination.

Example: To illustrate this proposal, let us assume that users with low radio conditions are assigned 4-QAM constellations for their direct link with the BS and that the BS-RS and RS-SS links allow the use of 16-QAM constellations. In classical relaying, the RS overhears the original signal using 4-QAM, in a first time slot, and generates a signal to the SS using 4-QAM also. This second signal, which can or cannot be decoded-encoded successfully at the RS, will be sent to the SS in a second time slot using the same constellation size as

that of the overheard signal, 4-QAM in this case. At the SS as well, it may or may not be decoded successfully.

In order to use HM and link adaptive relaying, we now allow each user to be served for β = 2 time slots. In each time slot, the RS will receive additional information for a different signal from the BS using HM, it will decode the original signal and the additional information (a copy of the same signal in our case) and encode a more robust copy of this signal. It then relays the 2 robust copies of the two different signals received in the previous two time slots to the destination using a higher bit rate, as allowed by the 16-QAM constellation. This corresponds to one time slot for the relaying of the two robust signals (see, Figure 5b), as opposed to the case of classical relaying where two time slots are needed to relay two different signals (not even robust copies) to the destination, as shown in Figure 5a. This gain in time slots improves the overall system capacity.

5 Flow level modeling of the HM-based link adaptive relaying scheme

We now model the HM-based link adaptive relaying scheme described in Section 4.

5.1 Model

As indicated above, based on the channel conditions of users, different regions can be constructed by grouping different mean channel responses of users into a finite set of values $j, j = 1,...,J$; each region containing users with similar radio conditions and, thus, similar coding and modulation scheme (see, Figure 6) [12].

Let users arrive to the system following a Poisson process with mean intensity λ. Users in region j are granted $N^j(\mathbf{y})$ subcarriers for a (finite) mean time duration T^j. $\mathbf{y}$ is a vector with entry j representing the number y^j of users present in region j, $j = 1,...,J$. The service duration T^j depends on the quantity of resources the users get, which in turn depends on the number and type of users that are simultaneously in progress in the system as well as how well they can take advantage of the resources they are granted, i.e., their radio conditions.

Recall that users are served using TDM. This ensures fairness in time among all users in the cell. In this case, the number of users in our system can be modeled as a Continuous Time Markov Chain (CTMC) with state vector $\mathbf{y}$, again, denoting the number of users y^j in each region of the system, $j = 1,..., J$. This model is completely described by a Processor Sharing (PS) queue with $S = \sum_j y^j$ being the total number of users in the system.

We also apply an admission control scheme on the maximal number of users s_{max} that can be admitted to the system. This guarantees every admitted user some minimal number of subcarriers n_{min} equal to $\dfrac{N}{s_{max}}$ which, in turn, guarantees some minimal throughput for each user in the system.

Let us denote by $\bar{\rho}$ the total cell load and by $\bar{\rho}^j$ the load corresponding to region $j, j = 1,...,J$. We have:

$$\bar{\rho} = \sum_{j=1}^{J} \bar{\rho}^j \tag{14}$$

where $\bar{\rho}^j = \rho^j/c^j$; $\rho^j = \rho\pi\left((r^j)^2 - (r^{j-1})^2\right)$; $\rho = \lambda E[F]$. r^j is the radius of region j, $j = 1,..., J$, $E[F]$ is the mean file size and c^j is the rate that a user s of type j can achieve when he is alone in the system. In a scheme without
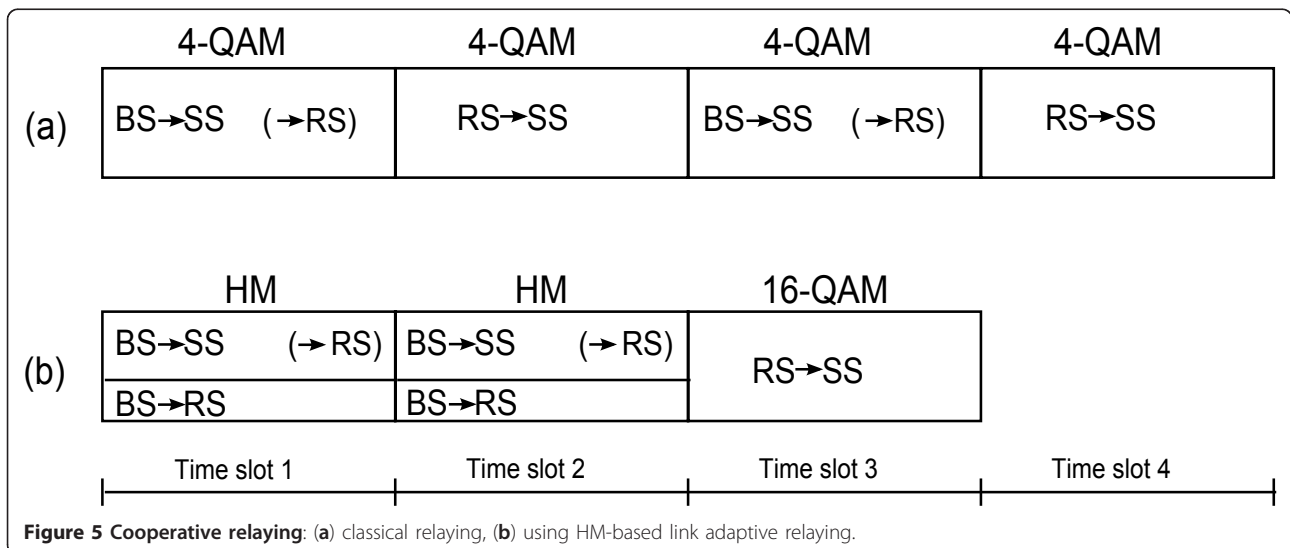

Figure 5 Cooperative relaying: (**a**) classical relaying, (**b**) using HM-based link adaptive relaying.

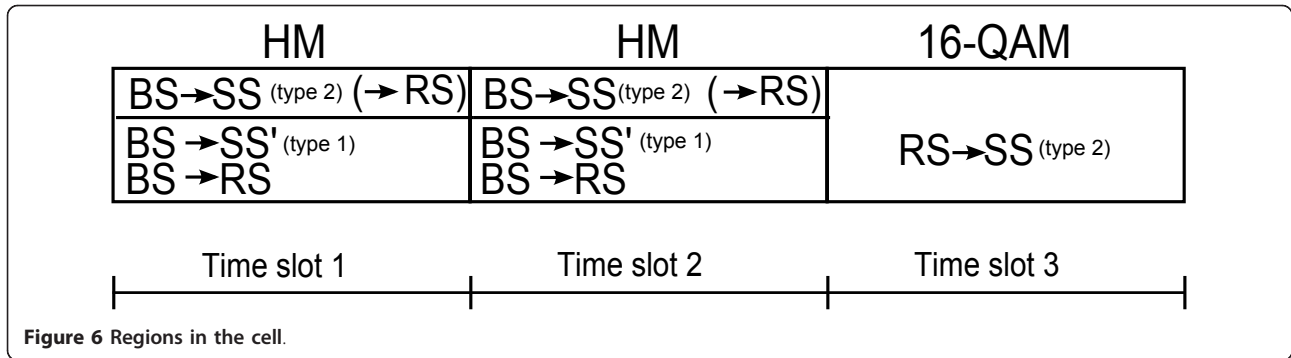| HM	HM	16-QAM
BS➔SS (type 2) (➔RS) BS➔SS' (type 1) BS➔RS	BS➔SS (type 2) (➔RS) BS➔SS' (type 1) BS➔RS	RS➔SS (type 2)
Time slot 1	Time slot 2	Time slot 3

Figure 6 Regions in the cell.

using relays and without HM, the maximum achieved throughput by user s in region j, is given by [13]:

$$c_s^j = \sum_{n=1}^{N} \frac{W}{N} \log_2 \left(1 + \frac{\mathrm{SNR}_{s,n}^j}{\Gamma} \right) \qquad (15)$$

where W denotes the total bandwidth, N denotes the total number of subcarriers and $\mathrm{SNR}_{s,n}^j$ is the instantaneous SNR experienced by user s in region j on subcarrier n.

Let us consider, as in Section 4 and, again, without loss of generality, that we have two types of users in the system ($J = 2$): users of type 1 who enjoy good radio conditions and who are able to decode successfully their signals, and users of type 2, who present worse radio conditions and may not be able to decode successfully their signals through the direct link (BS-SS) only.

5.1.1 Using classical relaying

Considering TDM scheduling, time slots are classified into three possibilities: one time slot to serve a user of type 1, one time slot to serve a user of type 2 if $P^R = 0$ and 2 time slots to serve the same user if $P^R = 1$. And so, given P and P^R, users of type 2 get a rate equal to $((1 - P^R) + PP^R)c^2$ for $\frac{\gamma^{-2}((1 - P^R) + PP^R)}{\gamma^{-1} + (1 - P^R)\gamma^{-2} + 2P^R\gamma^{-2}}$ of the time where $\bar{\gamma}^j, j = 1, 2$, denotes the mean number of users of each type in the system, and is equal to $\frac{\bar{\rho}^j}{1 - \bar{\rho}}$.

Without the need for relaying (successful decoding on the direct link), users of type 2 get a rate equal to c^2 for $\frac{\gamma^{-2}}{\gamma^{-1} + \gamma^{-2}}$ of the time.

Consequently, the capacity c^2 of users of type 2 changes, and the capacity given by Equation (15) will now be multiplied by the following term:

$$\frac{(\gamma^{-1} + \gamma^{-2})((1 - P^R) + PP^R)}{\gamma^{-1} + (1 - P^R)\gamma^{-2} + 2P^R\gamma^{-2}} \qquad (16)$$

The change in c^2 will, in turn, impact the performance of users of type 1 as they will now have one time slot less every time the transmission to users of type 2 uses a relay. Users of type 1 obtain a rate equal to c^1 for $\frac{\gamma^{-1}}{\gamma^{-1} + (1 - P^R)\gamma^{-2} + 2P^R\gamma^{-2}}$ of the time.

On the other hand, without the need for relaying, they obtain a rate equal to c^1 for $\frac{\gamma^{-1}}{\gamma^{-1} + \gamma^{-2}}$ of the time. Consequently, the capacity for users of type 1, given by Equation (15) under TDM will now be multiplied by:

$$\frac{\gamma^{-1} + \gamma^{-2}}{\gamma^{-1} + (1 - P^R)\gamma^{-2} + 2P^R\gamma^{-2}} \qquad (17)$$

5.1.2 Using HM-based link adaptive relaying

Considering the HM-based adaptive relaying scheme described previously, the TDM scheduling will also be altered due to the need for resources for relaying. As mentioned previously, the use of a resource allocation that is adapted to the radio conditions of the RS-SS link allows us to save some resources (in terms of time slots and power) compared to the classical relaying scheme and, in this case, during only one time slot, the RS forwards β different signals to the destination instead of only one signal. Recall that β is an integer equal to $\frac{c^R}{c^2}$, where c^R is the throughput achieved at the RS-SS link and c^2 is the capacity of users of type 2. Recall that the ratio $\frac{c^R}{c^2}$ is equal to the ratio $\frac{\log_2(M')}{\log_2(M)}$ (where M and M' stand for the constellation size without and with the use of link adaptation in the RS-SS link, respectively).

Here too, time slots can be divided into three possibilities: β time slots to serve a user of type 2 if $P^R = 0$, $\beta + 1$ time slots to serve a user of type 2 if $P^R = 1$ and β time slots to serve a user of type 1. And so, users of type 2 will get a rate equal to $(\beta PP^R + \beta(1 - P^R))c^2$ for $\frac{\gamma^{-2}((1 - P^R) + PP^R)}{\beta\gamma^{-1} + \beta(1 - P^R)\gamma^{-2} + (\beta + 1)R^R\gamma^{-2}}$ of the time where $\gamma^{-j}, j = 1, 2$, denotes, again, the mean number of users of each type in the system.

Hence, using HM-based link adaptive relaying, the capacity for users of type 2, c^2, given by Equation (15), will now be multiplied by:

$$\frac{\beta PP^R + \beta(1 - P^R)}{\beta\gamma^{-1} + \beta(1 - P^R)\gamma^{-2} + (\beta + 1)P^R\gamma^{-2}}(\gamma^{-1} + \gamma^{-2}) \quad (18)$$

As of users of type 1, their capacity c^1, given by Equation (15), will now be multiplied by a new term given by:

$$\frac{\beta(\gamma^{-1} + \gamma^{-2})}{\beta\gamma^{-1} + \beta(1 - P^R)\gamma^{-2} + (\beta + 1)P^R\gamma^{-2}} \quad (19)$$

5.2 Performance metrics
The steady-state probabilities are given by [18]:

$$\pi(x) = \frac{(\bar{\rho})^x}{1 + \bar{\rho} + \cdots + (\bar{\rho})^{s_{max}}} \quad (20)$$

for $0 \le x \le s_{max}$.

Using Equation (20), the mean transfer time T^j for a user in region j, $j = 1,..., J$, for a file of mean size $E[F]$, is given by:

$$T^j = E[F]\frac{1 - (s_{max} + 1)\bar{\rho}^{s_{max}} + s_{max}\bar{\rho}^{(s_{max}+1)}}{c^j(1 - \bar{\rho})(1 - \bar{\rho}^{s_{max}})} \quad (21)$$

and the mean transfer time T in all the cell is given by

$$T = \frac{\sum_{j=1}^{J} T^j\lambda^j}{\sum_{j=1}^{J} \lambda^j} \quad (22)$$

where λ^j is the mean arrival intensity to region j.

The blocking probability B of a new flow, based on our admission control algorithm, is given by:

$$B = \frac{\bar{\rho}^{s_{max}}}{1 + \bar{\rho} + \cdots + \bar{\rho}^{s_{max}}} \quad (23)$$

6 Enhancement
We propose, now, to take advantage of the good radio conditions of users who are close to the BS to send them a new additional stream using HM.

So far, the HM-based adaptive relaying scheme did send the additional information to the RS using a low modulation rate (4-QAM in our example above). The BS-RS link is, however, typically of good radio conditions, which allows the use of a higher modulation rate, for instance 16-QAM, to transmit the second level of hierarchy in the HM operation.

In this case, the second level of hierarchy of the HM operation can be used to transmit, in addition to the additional information B sent to the RS and meant to the user with bad radio conditions (type 2), a new extra

signal C to a user with good radio conditions (type 1), as shown in Figure 7.

This can be achieved by sharing the number of sub-carriers N equally between the two signals in this second level of hierarchy. And so, over the two time slots ($\beta = 2$), a user of type 1 will now receive at an additional rate given by:

$$\beta c^{add1} = \beta \sum_{n=1}^{N/2} \frac{W}{N}\log_2\left(1 + \frac{p^1_{k*,n}\left|h^1_{k*,n}\right|^2}{\Gamma(\sigma^1_{k*,n})^2}\right) \quad (24)$$

where $p_{k*,n}$ denotes the amount of power given by Equation (6).

The capacity of a user of type 2 will remain unchanged compared to the case of HM-based link adaptive relaying, whereas the rate that a user of type 1 can achieve will vary. In fact, considering TDM, the time is again divided into three parts: β time slots to serve a user of type 2 if $P^R = 0$ with rate equal to c^2, $\beta + 1$ time slots to serve a user of type 2 with rate equal to βc^2 and users of type 1 with rate equal to βc^{add1} if $P^R = 1$ and β time slots to serve a user of type 1 with rate equal to c^1.

Users of type 1 will, now, be served with a rate equal to $\beta c^1 + \beta c^{add1}$ for $\dfrac{\gamma^{-1}\beta\left(1 + \dfrac{c^{add1}}{c^1}P^R\right)}{\beta\gamma^{-1} + \beta(1 - P^R)\gamma^{-2} + (\beta + 1)P^R\gamma^{-2}}$ of the time. Consequently, the capacity for users of type 1, given by Equation (15), will now be multiplied by:

$$\frac{(\gamma^{-1} + \gamma^{-2})\beta\left(1 + \dfrac{c^{add1}}{c^1}P^R\right)}{\beta\gamma^{-1} + \beta(1 - P^R)\gamma^{-2} + (\beta + 1)P^R\gamma^{-2}} \quad (25)$$
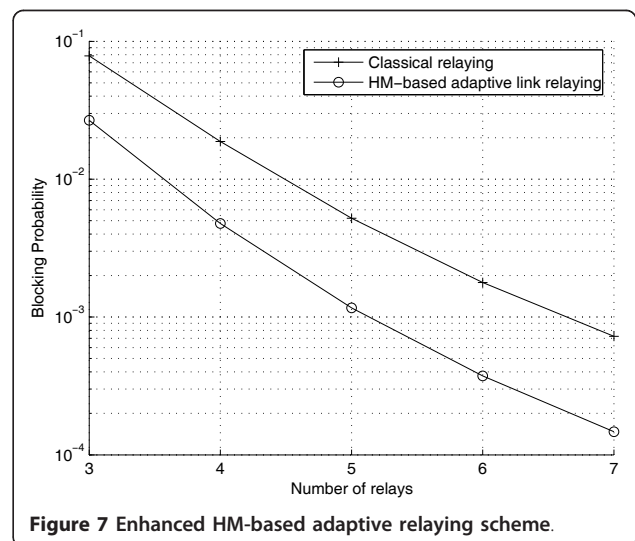

Figure 7 Enhanced HM-based adaptive relaying scheme.

where c^{add1} is the additional rate that a user of type 1 can achieve in each of the β time slots that are allocated to user of type 2.

7 Model validation and performance evaluation

We now validate, in this Section, our analytical model and investigate the flow-level performance of the HM-based link adaptive relaying scheme, in terms of the mean transfer time and blocking rates, and quantify the gain achieved using the enhancement to this scheme.

Please note that, the physical layer performance, in terms of BER, is beyond the scope of our study.

For these purposes, we have implemented a new simulator for a downlink cooperative relay-based OFDMA system, where users follow a realistic setting: they come to the system according to a random, Poisson process, and leave it after a finite duration, upon the completion of their file transfers. We have, also, implemented, as indicated previously, an admission control scheme based on a maximal number $s_{\max}$ of active users in the system; $s_{\max}$ is obtained by setting a minimal threshold $n_{\min}$ on the number of subcarriers per admitted user.

For the channel implementation, we have used a multi-path Model E channel, which corresponds to a typical office environment for non line-of-sight (NLOS) conditions with an exponential decay power profile [19].

Table 1 shows the parameters used in our numerical applications as well as simulations.

Note that Γ = 8.8 dB ensures a probability of error equal to 10^{-6} at decoding. We suppose also that the base station has a perfect knowledge of the CSI on all links.

We assume that users of type 2 need RS assistance (i. e., P^R = 1) and that we have a sufficient number of RSs in the cell to cover theses users (i.e., P = 1). We also assume that the BS and RS use, respectively, 4-QAM and 16-QAM signal modulation to serve users of type 2. Users of type 1 use 16-QAM signal modulation. Based on TDM, each user will be served for β = 2 time slots at each cycle.

Figures 8, 9, and 10 show the mean transfer time for all users in the system, users of type 1 and users of type 2, respectively, as a function of an increasing offered traffic. The latter corresponds to a cell load ranging from 0.3 to 0.8, i.e., from moderately lightly to heavy-loaded system. The curves show both analytical and
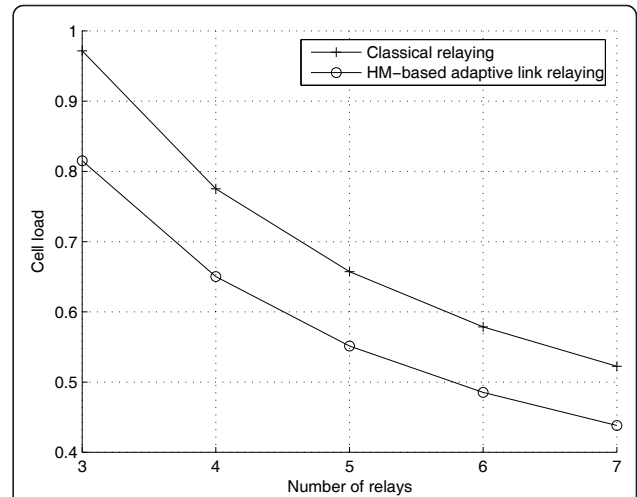
Table 1 System parameters

Carrier frequency	3.5 GHz
System channel bandwidth	1.25 MHz
FFT	128
OFDMA symbol duration	102,9 μs
Noise	-160 dbm/Hz
Γ	8.8 dB
File size distribution	Exponential with mean 40 Kb.

Figure 8 Mean transfer time for all users: HM-based link adaptive relaying scheme versus classical relaying.
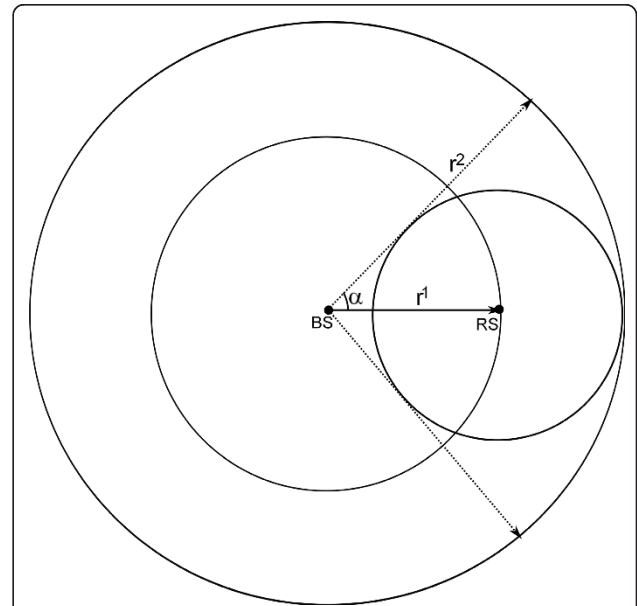

Figure 9 Mean transfer time for users of type 1: HM-based link adaptive relaying scheme versus classical relaying.
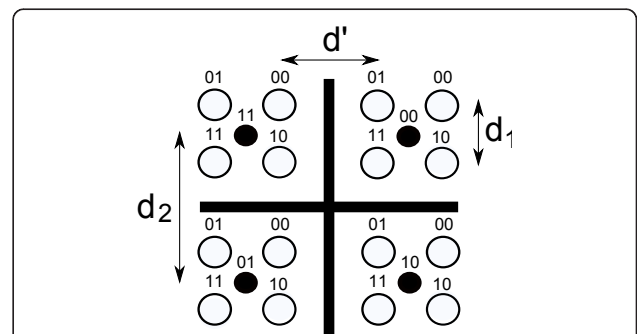

Figure 10 Mean transfer time for users of type 2: HM-based link adaptive relaying scheme versus classical relaying.

simulation results for the cases of the HM-based link adaptive relaying scheme versus classical relaying.

We observe, first, a good match between the analytical model and the simulation results for the two schemes.

We observe, second, that the HM-based link adaptive relaying scheme allows a gain, in terms of the mean transfer time, for users of both types, as compared to the classical relaying scheme. This is due to the use of HM and also to the use of the resource allocation strategy that adapts the modulation scheme to the radio conditions on the RS-SS link. Indeed, in this case, relays use less resources, in terms of time, to assist users with worse radio conditions; the resources thus freed will be used by other users in the system to finish their services earlier which allows to increase the overall system capacity.
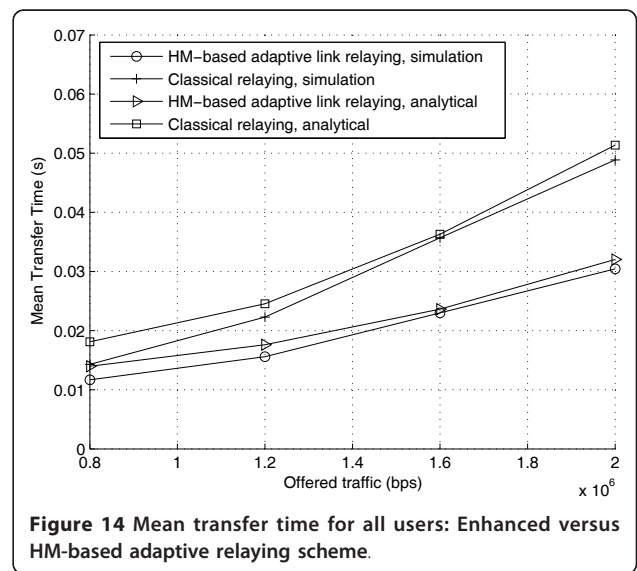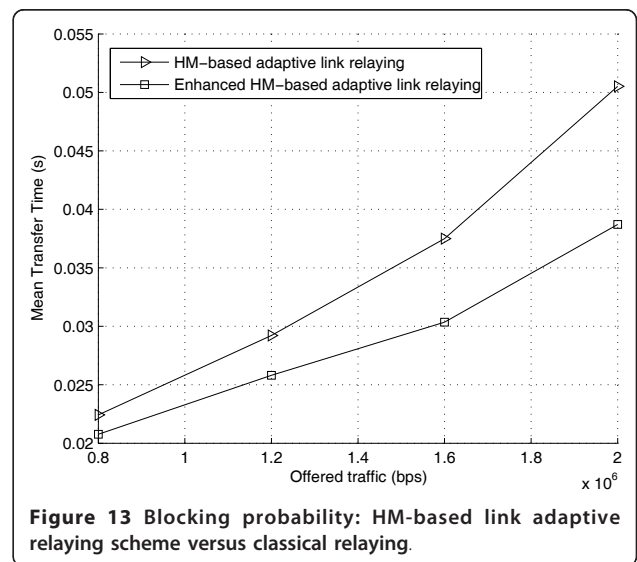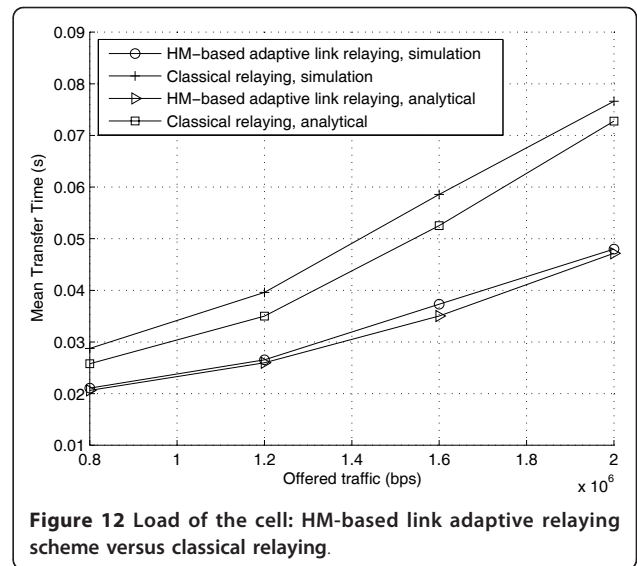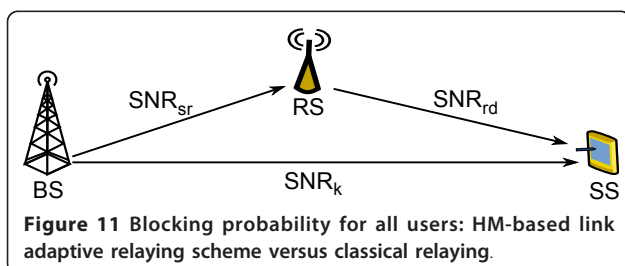
We observe, third, that the gain is larger as the system becomes more loaded, i.e., larger number of users, and this, again, for all users. This is due to multiuser diversity.
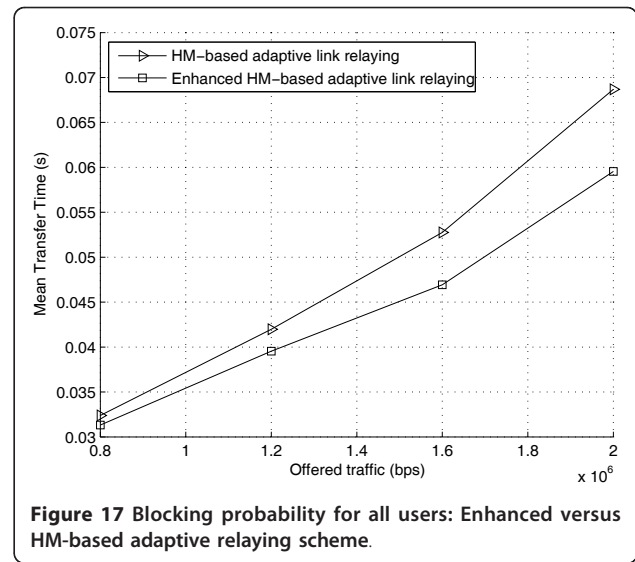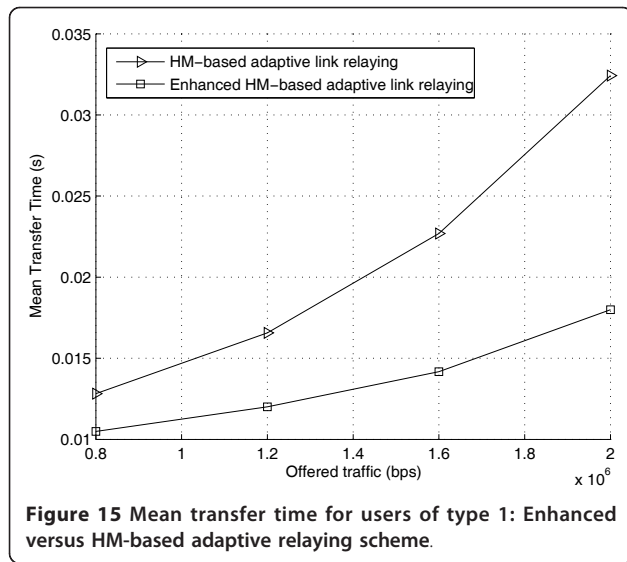
Figure 11 shows the blocking probability for the two schemes. Here too, the blocking probability is better for the HM-based link adaptive relaying scheme, which means a higher admission rate, because again users of both types finish their service earlier and hence leave more room to yet other users to be admitted to the system.

We eventually turn to the impact of the number of RSs on the performance of our relay-based system. Indeed, in the previous figures, we assumed that the number of RSs was large enough so as to assist all users with bad radio conditions (i.e., $P = 1$). In a real case, P can be less than 1. We show, in Figures 12 and 13, the load in the system (given by Equation (14)) as well as the blocking rate as a function of the number of RSs, respectively, both for classical relaying and HM-based adaptive link relaying.

We observe that both measures decrease when the number of relays increases. This is due to the additional resources that users with bad radio conditions can get using relaying, with a clear advantage to the case of HM and adaptive link relaying over the classical one.

We now investigate the performance of the proposed enhancement as compared to the HM-based adaptive relaying scheme. We show in Figures 14, 15, and 16 the mean transfer time for all the users in the system, as

Figure 11 Blocking probability for all users: HM-based link adaptive relaying scheme versus classical relaying.

Figure 12 Load of the cell: HM-based link adaptive relaying scheme versus classical relaying.

Figure 13 Blocking probability: HM-based link adaptive relaying scheme versus classical relaying.

Figure 14 Mean transfer time for all users: Enhanced versus HM-based adaptive relaying scheme.

Figure 15 Mean transfer time for users of type 1: Enhanced versus HM-based adaptive relaying scheme.

Figure 17 Blocking probability for all users: Enhanced versus HM-based adaptive relaying scheme.

well as for users of type 1 and users of type 2, respectively. We observe that the enhanced scheme achieves the lowest mean transfer time, and this is due to the additional throughput received by users of good radio conditions who will now finish their service earlier (Figure 15) and free resources to the other users in the system (Figure 16).

The same result holds for the blocking rate also, as shown in Figure 17 for the same reasons as detailed above for the case of HM-based relaying over the classical scheme.

8 Conclusion

We considered, in this study, the joint use of HM and adaptive link relaying in a cooperative relay-based

Figure 16 Mean transfer time for users of type 2: Enhanced versus HM-based adaptive relaying scheme.

OFDMA system, as is the case of LTE-A systems, and modeled it in a cross-layer manner, taking into account realistic user dynamics in terms of arrivals and departures after a finite service duration. HM allows the reconstruction of a more robust signal at the RS owing to the sending of an additional information from the BS to the RS. The link adaptation between the RS and SS allows the reduction of the cost of additional resources needed for relaying and hence improves the overall system capacity. Our numerical results enabled us to validate our analytical model and to quantify the gains obtained by such a joint HM and adaptive relaying scheme, in terms of mean transfer time and blocking rate, as compared to the classical relaying scheme. We also proposed an enhancement to the HM-based adaptive relaying scheme which takes advantage of the good radio conditions of the BS-RS link so as to send extra information to other users in the cell using higher modulation rate.

One future study perspective is on the study of the problem of power allocation in the case of cooperative communication, and its impact on system performance, again, considering a dynamic user scenario, at the flow level.
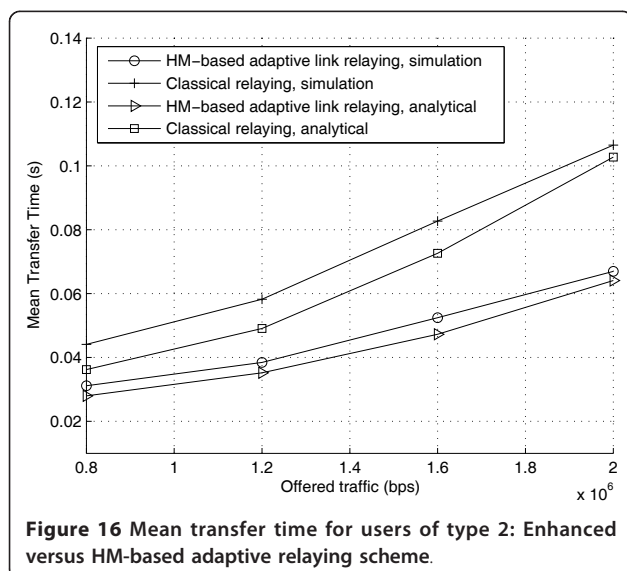
References
1. IEEE: *IEEE Baseline document v4 for draft standard, Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems*, Multihop Relay Specification, (2008)
2. HX Nguyen, HH Nguyen, T Le-Ngoc, Signal transmission with unequal error protection in wireless relay networks. IEEE Trans Veh Technol. **59**, 2166–2178 (2010)

3. T Cover, Broadcast channels. IEEE Trans Inf Theory. **IT-18**(no. 1), 2–14 (1972)
4. M Peng, Y Liu, D Wei, W Wang, HH Chen, Hierarchical cooperative relay based heterogeneous networks. IEEE Wirel Commun. **18**, 48–56 (2011)
5. C Hausl, J Hagenauer, Relay communication with hierarchical modulation. IEEE Commun Lett. **11**, 64–66 (2007)
6. L Zheng, K Wang, W Wang, Performance analysis of coded cooperation with hierarchical modulation. ICC 4978–4982 (2008)
7. MK Chang, SY Lee, Performance analysis of cooperative communication system with hierarchical modulation over Rayleigh fading channel. IEEE Trans Wirel Commun. **8**, 2848–2852 (2009)
8. S Jeong-Chul, S Jin-Hyuk, K Jee-Hoon, S Hyoung-Kyu, New transmission scheme with hierarchical modulation in dual-hop wireless relaying systems. ATNAC 1–5 (2009)
9. EG Larsson, BR Vojcic, Cooperative transmit diversity based on superposition modulation. IEEE Com-mun Lett. **9**, 778–780 (2005)
10. P Popovski, E de Carvalho, Improving the rates in wireless relay systems through superposition coding. IEEE Trans Wirel Commun. **7**, 4831–4836 (2008)
11. RJ Whang, H Liu, EK Hong, Multiuser cooperative relay communication employing hierarchical modulation, (VTC Spring, 2010), pp. 1–5
12. A Jdidi, SE Elayoubi, T Chahed, H Besbes, Modeling the flow-level performance of hierarchical modulation in OFDMA-based networks. Telecommun Syst. **55**, 1784–1793 (2010)
13. D Tse, P Viswanath, *Fundamentals of Wireless Communication*, (Cambridge University Press, Cambridge, 2005)
14. B Liu, H Liu, DPC-based hierarchical broadcasting: design and implementation. IEEE Trans Veh Technol. **57** (2008)
15. S Pietrzyk, GJM Janssen, Subcarrier and power allocation for qos-aware ofdma systems using embedded modulation, in *Proceedings IEEE ICC*. **6**, 20–24 (2004)
16. T Nechiporenko, C Tellambura, KT Phan, HH Nguyen, Performance analysis of adaptive M-QAM for Rayleigh fading cooperative systems, in *IEEE International Conference on Communications*, pp. 3393–3399 (2008)
17. TE Hunter, A Nosratinia, Diversity through coded cooperation. IEEE Trans Wirel Commun. **5**(2), 283–289 (2006)
18. T Bonald, A Proutiere, Wireless downlink data channels: User performance and cell dimensioning, in *Proc of MOBICOM*, pp. 339–352 (2003)
19. Z Shen, JG Andrews, BL Evans, Adaptative resource allocation in multiuser OFDM system with proportional rate constraints. IEEE Trans Wirel Commun. **4**(6), 2726–2737 (2005)