Improving the Rates in Wireless Relay Systems through Superposition Coding

Petar Popovski, *Member, IEEE* and Elisabeth de Carvalho, *Member, IEEE*

***Abstract*—We introduce a new two–step relaying scheme based on superposition coding, *SC–relaying*. In Step 1, the source A_S broadcasts a message, created by superposition coding, to the relay A_R and the destination A_D. After decoding the information from A_S, A_R relays a part of this message in Step 2, by using a codebook that is adapted to the link $A_R - A_D$. The information–theoretic design shows that the proposed scheme can improve the spectral efficiency, attaining almost optimal value. We demonstrate that the SC–relaying can be ported even to the case of uncoded systems, where it is shown how to select the transmission scheme in order to maximize the throughput between the source and the destination.**

I. Introduction

Recently, there have been significant research efforts related to wireless relay systems, most notable being the schemes for cooperative diversity [1] as well as the information–theoretic investigations of the strategies for cooperative relaying [2], [3]. Fig. 1 depicts a wireless relay system, where A_S being the source, A_D is the information sink and the relay A_R is assisting the communication between A_S and A_D. We consider the systems where the link $A_S A_R$ is stronger $A_S A_D$, which is the interesting case for relay–based systems [4].

We propose a 2–step relaying scheme based on superposition coding, shortly *SC–relaying*. SC–relaying belongs to the class of Decode–and–Forward (DF) schemes [5][6]. In Step 1, A_S broadcasts to A_R and A_D, while in Step 2 A_R transmits to A_D. The data transmitted by A_R in Step 2 is a function of the data that A_R has received from A_S in Step 1. The superposition coding [7] has been introduced for efficient broadcasting. The usage of superposition coding for relaying has been investigated in [8], where the role of the superposition coding is to increase the diversity and thus decrease the outage probability. On the other hand, the objective of our SC–relaying scheme is to increase the spectral efficiency and is therefore necessary to assume Channel State Information (CSI) at the transmitters [6], such that the transmission rates can be adapted, allowing for unequal durations of Step 1 and Step 2. The treatment of the issues related to outage probability, when the CSI is entirely/partially not known is outside the scope of this letter. In order to improve the spectral efficiency, the node A_R recodes the signal and adapts it to the achievable rate on the link $A_R A_D$. Due to the particular choice of the transmission format in Step 1, the signal received in Step 2 can be easily utilized to help A_D decode the signal received in Step 1. Note that such a combining is not straightforwardly possible for the Decode–and–Forward (DF) scheme from [1]. Furthermore, the spectrally–efficient relaying used in [6] states that A_D uses maximum ratio combining of the transmission by A_S and A_R, but doe not specify the encoding procedure. On the other hand, SC–relaying represents a scheme that uses standard single link codebooks and is designed to enable simple combining of the transmission by A_S and A_R at A_D, and therefore portability of the ides to the cases of uncoded/lightly coded systems.

The authors are with Department of Electronic Systems, Aalborg University, Niels Jernes Vej 12, DK-9220 Aalborg, Denmark. Email: {petarp,edc}@es.aau.dk

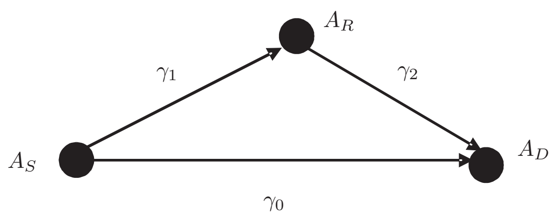

Fig. 1. The wireless relay scenario. γ_0, γ_1, and γ_2 denote the SNRs.

The letter is organized as follows. Section II introduces the system model and the assumptions. The operation of the SC–relaying scheme is described in Section III. The design of the scheme under information–theoretic assumptions, along with some numerical results, is given in Section IV. Design of the SC–relaying scheme for uncoded systems and guidelines for practical adaptive modulation and coding (AMC) are described in Section IV. The last section concludes the paper and outlines topics for future work.

II. System Model

There is a single channel with a normalized bandwidth of 1 Hz and a link with SNR of γ can send up to:

$$C(\gamma) = \log(1+\gamma) \text{ [bit/s]} \tag{1}$$

With the normalized bandwidth, the terms "rate" and "spectral efficiency" become equivalent, while the time can be expressed in number of symbols. We use $x_V[m]$ to denote the $m-$th complex baseband transmitted symbol from node $V \in \{A_S, A_R\}$. A complex–valued vector is denoted by $\mathbf{x}$. If the node V is transmitting, then the $m-$th received symbol at the node $W \in \{A_R, A_D\}$ is given by:

$$y_W[m] = h_i x_V[m] + z_W[m] \tag{2}$$

where $i=0$ if $V = A_S, W = A_D$, $i=1$ if $V = A_S, W = A_R$, and $i=2$ if $V = A_R, W = A_D$. Each h_i denotes a complex channel coefficient. $z_W[m]$ is the complex additive Gaussian noise $\mathcal{CN}(0, \sigma^2)$. The transmitted symbols have

Step 1: A_S broadcasts the signal $$\sqrt{1-\alpha}\mathbf{x}_b + \sqrt{\alpha}\mathbf{x}_s$$ • A_R receives $\mathbf{y}_R = h_1(\sqrt{1-\alpha}\mathbf{x}_b + \sqrt{\alpha}\mathbf{x}_s) + \mathbf{z}_R$ and decodes successfully first $\mathbf{x}_b$ and then $\mathbf{x}_s$. • A_D receives $\mathbf{y}_{D1} = h_0(\sqrt{1-\alpha}\mathbf{x}_b + \sqrt{\alpha}\mathbf{x}_s) + \mathbf{z}_{D1}$ and does not decode the signals, but keeps $\mathbf{y}_{D1}$ in memory.
Step 2: A_R transmits the signal $\mathbf{x}_r$, which has identical information content as $\mathbf{x}_s$ from Step 1, but is *re–coded* with a different codebook. • A_D receives $\mathbf{y}_{D2} = h_2\mathbf{x}_r + \mathbf{z}_{D2}$, decodes $\mathbf{x}_r$ and uses the information to locally generate $\mathbf{x}_s$. • A_D obtains the vector $\mathbf{y}_{D3} = \mathbf{y}_{D1} - h_0\sqrt{\alpha}\mathbf{x}_s = h_0\sqrt{1-\alpha}\mathbf{x}_b + \mathbf{z}_{D1}$ which is used to decode $\mathbf{x}_b$.

TABLE I
DESCRIPTION OF THE SC–RELAYING SCHEME.

$E\{x_V[m]\} = 0$ and a normalized power $E\{|x_V[m]|^2\} = 1$, such that the SNRs are given by $\gamma_i = \frac{|h_i|^2}{\sigma^2}$ for $i = 0, 1, 2$. The transmitted power is fixed for both A_S and A_R. However, we do not put a joint power constraint on the source and the relay, as our assumption is that the relay is deployed to use maximal power. Unless stated otherwise, the following will be always assumed:

$$\gamma_1 > \gamma_0 \qquad \gamma_2 > \gamma_0 \tag{3}$$

All the channels are considered to be constant at least during Step 1 and Step 2. Each node is assumed to know all the relevant SNRs $\gamma_0, \gamma_1, \gamma_2$ before the start of the transmission of A_S. For the information–theoretic design, we assume that the packet lengths are sufficiently large, such that long error–correcting codes can be applied to yield a practical zero probability of error if the rate is chosen to be below the channel capacity.

III. DESCRIPTION OF THE SC–RELAYING SCHEME

The two steps of the SC–relaying scheme are described in Table I. In the description, α is the *power–division coefficient* for superposition coding and $0 \le \alpha \le 1$. $\mathbf{x}_b$ is called the *basic* message and $\mathbf{x}_s$ is the *superposed* message. The duration of Step 1 is N symbols, such that both $\mathbf{x}_b$ and $\mathbf{x}_s$ are vectors of N complex components.

Let R_b and R_s be the transmission rates of the basic and the superposed message, respectively, while R_2 is the transmission rate applied by A_R in Step 2. Then R_b and R_s should be chosen such that, after Step 1, A_R should be able to decode both $\mathbf{x}_b$ and $\mathbf{x}_s$ from $\mathbf{y}_R$. Furthermore, R_2 should be chosen such that A_D can reliably decode $\mathbf{x}_r$ (and therefore $\mathbf{x}_s$) from $\mathbf{y}_{D2}$. Finally, R_b should be also chosen in such a way that A_D can reliably decode $\mathbf{x}_b$ from $\mathbf{y}_{D3}$.

Let us assume that R_b, R_s, R_2 are appropriately chosen, as required in the description above. Then, the overall transmission rate R_{sc} from A_S to A_D can be calculated as follows. The total number of bits that A_S transmits in Step 1 is given by $N(R_b + R_s)$. The total number of bits that A_R needs to send in Step 2 is NR_s, as $\mathbf{x}_r$ contains identical information with $\mathbf{x}_s$ from Step 1. Since A_R transmits at the rate R_2, the Step 2 has a duration of $N_2 = \frac{NR_s}{R_2}$ symbols and in general $N_2 \neq N$. The amount of data received by A_D after the two steps is $N(R_b + R_s)$, such that the overall data rate is:

$$R_{\text{sc}} = \frac{N(R_b + R_s)}{N + N_2} = \frac{(R_b + R_s)R_2}{R_2 + R_s} \tag{4}$$

If $N_2 < N$, then it can be seen that the overall rate is improved and this is exactly because we make provisions for the adapted duration of the second step.

IV. INFORMATION–THEORETIC DESIGN

A. Choice of the Rates and α

Central to the proper information–theoretic design of the proposed scheme is the selection of the coefficient α and the rates R_b and R_s. From the condition that A_R should decode $\mathbf{x}_b$ and $\mathbf{x}_s$, we obtain the following conditions:

$$R_b \le C\left(\frac{(1-\alpha)\gamma_1}{1+\alpha\gamma_1}\right) = R_{b_1}^U(\alpha) \quad R_s \le C(\alpha\gamma_1) = R_s^U(\alpha) \tag{5}$$

From the requirement that A_D should decode $\mathbf{x}_b$ from $\mathbf{y}_{D1}$ after knowing $\mathbf{x}_s$, we obtain:

$$R_b \le C((1-\alpha)\gamma_0) = R_{b_2}^U(\alpha) \tag{6}$$

Another required constraint comes from the fact that A_D should decode $\mathbf{x}_r$, sent by A_R in Step 2, which implies:

$$R_2 \le C(\gamma_2) \tag{7}$$

It is easy to see that, to maximize the overall spectral efficiency, we should set $R_2 = C(\gamma_2)$.

We now proceed towards the discussion on selecting α in order to optimize R_{sc}. The optimization problem can be stated as follows: *Given $\gamma_0, \gamma_1, \gamma_2$, find $\alpha \in [0, 1]$ that maximizes (4) while satisfying the conditions (5) and (6).*

Note that the direct transmission corresponds to $\alpha = 0$ and the multi–hop to $\alpha = 1$, which makes them special instances of the SC–relaying. First, it can be easily seen that R_2 should have the maximal possible value $R_2 = C(\gamma_2)$. Since R_{sc} strictly increases with R_b, then R_b should be set to the maximal possible value $R_b = \min\left\{R_{b_1}^U, R_{b_2}^U\right\}$. We consider two different cases:

$$R_{b_1}^U \le R_{b_2}^U \Leftrightarrow \alpha \ge \frac{1}{\gamma_0} - \frac{1}{\gamma_1} = \alpha_0 \tag{8}$$

$$R_{b_1}^U \ge R_{b_2}^U \Leftrightarrow \alpha \le \frac{1}{\gamma_0} - \frac{1}{\gamma_1} = \alpha_0 \tag{9}$$

To cover the case $\gamma_0 < \frac{\gamma_1}{1+\gamma_1}$, we should write $\alpha_0 = \min\left\{\frac{1}{\gamma_0} - \frac{1}{\gamma_1}, 1\right\}$. Clearly, if $\gamma_0 > \frac{\gamma_1}{1+\gamma_1}$ and (3) is satisfied, then $0 < \alpha_0 < 1$.

In (8) $\min\left\{R_{b_1}^U, R_{b_2}^U\right\} = R_{b_1}^U = R_b$ and $\alpha \in [\alpha_0, 1]$. Here $R_b + R_s = C(\gamma_1)$ and is independent of α. On the other hand, $R_{\text{sc}} = \frac{C(\gamma_1)C(\gamma_2)}{C(\gamma_2)+C(\alpha\gamma_1)}$ strictly decreases with α, such that it is optimal to set $\alpha = \alpha_0$. In (9) $\min\left\{R_{b_1}^U, R_{b_2}^U\right\} = R_{b_2}^U = R_b$ and $\alpha \in [0, \alpha_0]$. Here $R_{\text{sc}} = \frac{C(\gamma_2)[C((1-\alpha)\gamma_0)+C(\alpha\gamma_1)]}{C(\gamma_2)+C(\alpha\gamma_1)}$ and the optimal α needs to be found by setting $\frac{\mathrm{d}R_{\text{sc}}}{\mathrm{d}\alpha} = 0$. The optimal value α_{opt}, in general, is a function of all three values $\gamma_0, \gamma_1, \gamma_2$, but it cannot be expressed in a closed form. Nevertheless, it can be shown that, for given γ_0 and γ_1, there is a threshold value $f_2(\gamma_0,\gamma_1)$, such that when $\gamma_2 \geq f_2(\gamma_0,\gamma_1)$, then the choice $\alpha=\alpha_0$ is optimal, see [9]. The function $f_2(\gamma_0,\gamma_1)$ can be determined by setting $\frac{\mathrm{d}R_{\text{sc}}}{\mathrm{d}\alpha}\big|_{\alpha=\alpha_0}=0$, which results in:

$$f_2(\gamma_0,\gamma_1)=(1+\gamma_1)^{\frac{\gamma_0(1+\gamma_1)}{\gamma_0(1+\gamma_1)-\gamma_1}}\cdot(\gamma_0/\gamma_1)-1 \tag{10}$$

When γ_1 is relatively high, it follows from (10) that $f_2(\gamma_0,\gamma_1) \gg \gamma_0-1$ for small values of γ_0, while $f_2(\gamma_0,\gamma_1)\approx\gamma_0$ for larger γ_0. However, the numerical results show that choosing $\alpha=\alpha_0$ whenever $\gamma_2>\gamma_0$ leads only to slight suboptimality in the spectral efficiency. When $\alpha=\alpha_0$, the SC–relaying has a rate:

$$R_{\text{sc}}(\alpha_0)=\frac{C(\gamma_1)C(\gamma_2)}{C(\gamma_2)+\log_2\left(\frac{\gamma_1}{\gamma_0}\right)} \tag{11}$$

Instead of optimizing α, we will compare the performance of the SC–relaying with $\alpha=\alpha_0$ to an optimal reference information–theoretic scheme, described in Section IV-C.

B. Reference Schemes

Here we consider three alternatives to the SC–relaying scheme:

1) *Direct transmission* between A_S and A_D at a rate $R_{\text{dir}}=C(\gamma_0)$.
2) *Multi–hop transmission*, where all the data from Step 1 is relayed by A_R in Step 2, but the transmission rate in Step 2 is adapted to γ_2. Here A_D ignores the transmission of A_S in Step 1. The rate of the multi–hop transmission is easily found to be:

$$R_{\text{mh}}=\frac{C(\gamma_1)C(\gamma_2)}{C(\gamma_1)+C(\gamma_2)} \tag{12}$$

3) *Decode–and–forward (DF)* transmission, as in [1], where A_R decodes the packet from A_S and retransmits the packet in Step 2 using the *same* modulation/codebook as used in Step 1. Due to the same transmission format, now A_D can do Maximum Ratio Combining of the signal received from A_S in Step 1 with the signal received from A_R in Step 2. The transmission rate by A_S should be chosen such that A_R can decode A_R in Step 1 and A_D can decode the combined signal, such that:

$$R_{\text{DF}}=\frac{1}{2}\min\{C(\gamma_1),C(\gamma_0+\gamma_2)\} \tag{13}$$

The following observations are in order. First, it is interesting to note[1] that the direct transmission corresponds to $\alpha=0$ and the multi–hop transmission corresponds to $\alpha=1$. The choice $\alpha=\alpha_0$ makes $R_{\text{sc}}>R_{\text{mh}}$ whenever $\gamma_1>\gamma_0$ and $\gamma_2>\gamma_0$. Regarding the direct transmission, for given γ_0,γ_1, the minimal $\gamma_2=\gamma_{2,0}$ for which $R_{\text{sc}}(\alpha_0)\geq R_{\text{dir}}$ is determined through:

$$\log_2(1+\gamma_{2,0})=\log_2(1+\gamma_0)\cdot\frac{\log_2(\gamma_1/\gamma_0)}{\log_2\left[(1+\gamma_1)/(1+\gamma_0)\right]} \tag{14}$$

[1]There is no analogous correspondence with the DF scheme.

C. The Optimal Reference Scheme

The SC–relaying scheme is not optimal in an information–theoretic sense. In this section we are seeking for the scheme that offers maximal achievable rate by only putting the constraint that in Step 1 only A_S transmits and in Step 2 only A_R transmits. As pointed out in [6], such constraints are appropriate for the practical half–duplex relay channel. An additional constraint is that the scheme is decode–and–forward, such that A_R can reliably receive the message in Step 1. Hence, after A_S broadcasts the message in Step 1 at a rate $C(\gamma_1)$, the relay decodes it, while the uncertainty of A_D about the message sent by A_S is $N[C(\gamma_1)-C(\gamma_0)]$. Thus, if we use an optimized transmission at A_R, then in Step 2 there should be $N[C(\gamma_1)-C(\gamma_0)]$ bits in total transmitted to A_D. In Step 2 A_R transmits at a rate $C(\gamma_2)$, which makes the overall rate:

$$\begin{aligned} R_{\text{opt}} &= \frac{NC(\gamma_1)}{N+N\frac{C(\gamma_1)-C(\gamma_0)}{C(\gamma_2)}}=\frac{C(\gamma_1)C(\gamma_2)}{C(\gamma_2)+C(\gamma_1)-C(\gamma_0)} \\ &= \frac{C(\gamma_1)C(\gamma_2)}{C(\gamma_2)+\log_2\left(\frac{1+\gamma_1}{1+\gamma_0}\right)} \end{aligned} \tag{15}$$

By comparing (15) with (11) we see that when γ_0,γ_1 are (reasonably) greater than 1, then $R_{\text{sc}}\approx R_{\text{opt}}$.

For the optimal reference scheme, A_R should use a codebook created e. g. by random binning (Slepian–Wolf coding) [7], which is complex in practice. On the other hand, SC–relaying can use standard single–user codebooks, as $\mathbf{x}_b$ and $\mathbf{x}_s$ are coded independently. This makes the ideas of SC–relaying applicable in the case of lightly–coded or uncoded systems, unlike the approach with random binning.

When compared to the optimal reference scheme, it can be noted that there are two sources of suboptimality for the SC–relaying. *First*, in SC–relaying the node A_R relays $N\log_2\left(\frac{\gamma_1}{\gamma_0}\right)$ bits, which is too many compared to the minimal necessary, since:

$$C(\gamma_1)-C(\gamma_0)=\log_2\left(\frac{1+\gamma_1}{1+\gamma_0}\right)<\log_2\left(\frac{\gamma_1}{\gamma_0}\right) \tag{16}$$

whenever $\gamma_1>\gamma_0$. *Second*, in the SC–relaying, if after Step 2, we allow A_D to decode $\mathbf{x}_r$ by using $\mathbf{y}_{D1}$ as a side information, then $\mathbf{x}_r$ can be sent at a rate larger than $C(\gamma_2)$, which would increase the overall rate. Instead we have designed the SC–relaying to decode $\mathbf{x}_r$ from the transmission of A_R and then simply combine this information with the previous reception $\mathbf{y}_{D1}$ by subtraction. Such an operation makes the scheme easily portable to the case of uncoded/lightly–coded systems, as it will be seen in Section V. As a final note, the closeness of the SC–relaying to the optimal scheme, suggests a small gain from transmitting $\mathbf{x}_s$ at a rate $R_2>C(\gamma_2)$ and then decoding it by using $\mathbf{y}_{D1}$ as a side information.

D. Numerical Results

Here we provide numerical illustration for the information–theoretic design of the SC–relaying. Fig. 2 shows the gain in terms of spectral efficiency brought by the SC–relaying scheme. In this evaluation scenario, A_S,A_R,A_D are lying on

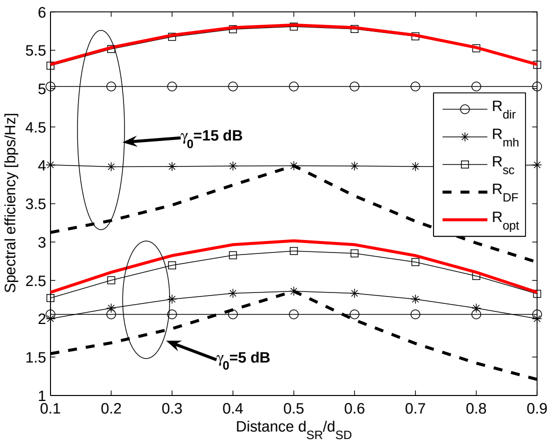

Fig. 2. Comparison of the spectral efficiency for different positions of A_R on the line $A_S A_D$. d_{ij} is the distance between A_i and A_j.

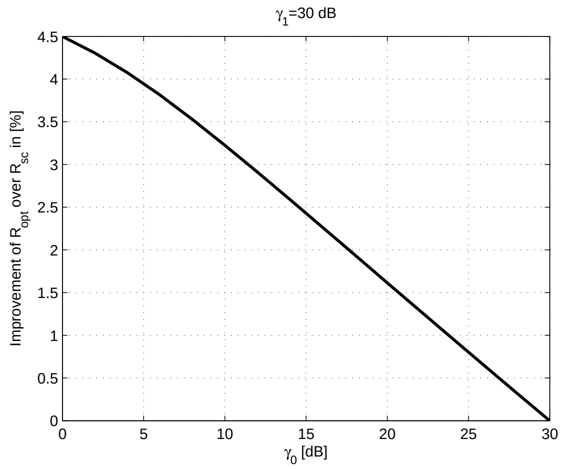

Fig. 3. Percentage of improvement of the optimized relaying over the SC–relaying, where the latter uses $\alpha = \alpha_0$. Here $\gamma_1 = 30$ dB, while for each γ_0 it is $\gamma_2 = \gamma_{2,0}$, where $\gamma_{2,0}$ is given by (14)

a same line, the distance between A_S and A_D is d_{SD} and the distance between A_S and A_R is d_{SR}. Two different values for γ_0 are used, $\gamma_0 = 5$ dB and $\gamma_0 = 15$ dB. For given γ_0 and given ratio $\frac{d_{SR}}{d_{SD}}$, the other two SNRs are determined as $\gamma_1 = \gamma_0 \left(d_{SD}/d_{SR}\right)^{\nu}$ and $\gamma_2 = \gamma_0 \left(d_{SD}/(d_{SD} - d_{SR})\right)^{\nu}$, where $\nu = 3$ is the path loss coefficient. In such a way, we always obtain $\gamma_1 > \gamma_0$ and $\gamma_2 > \gamma_0$. The major gain over R_{dir} and R_{mh} is brought by the SC–relaying, while the improvement of R_{opt} over R_{sc} is relatively small. When γ_0 increases, the SC–relaying performs very close to the optimal scheme, but the gain of the SC–relaying over the direct link is decreased. For the considered scenario, the DF scheme has the worst performance in terms of achievable rate. Note that the DF scheme approaches the multi–hop scheme (and slightly outperforms it) when the values of γ_1 and γ_2 are approximately equal, which happens around the middle position.

Fig. 3 illustrates the impact of the suboptimal choice $\alpha = \alpha_0$ for the SC–relaying. Note that when $\gamma_1 > \gamma_0$ and $\gamma_2 > \gamma_0$, the value of R_{opt} provides an upper bound for the spectral efficiency of the SC–relaying. Fig. 3 shows the value of $\frac{R_{\text{opt}} - R_{\text{sc}}(\alpha_0)}{R_{\text{sc}}(\alpha_0)} \cdot 100$ [%]. We have fixed $\gamma_1 = 30$ dB. In principle, for a given γ_0, we can do the evaluation for an arbitrary positive value of γ_2. Nevertheless, we have evaluated R_{opt} and R_{sc} for $\gamma_2 = \gamma_{2,0}$ (given in 14) due to the following. When $\gamma_0 < \gamma_2 < \gamma_{2,0}$, then the direct link is preferred over the SC–relaying and $R_{\text{opt}} > R_{\text{dir}} > R_{\text{sc}}(\alpha_0)$. The difference $R_{\text{opt}} - R_{\text{dir}}$ increases as γ_2 increases. The value of $\gamma_{2,0}$ is the minimal value of γ_2 where the SC–relaying is selected over the direct link. It can be easily shown that, for $\gamma_2 > \gamma_{2,0}$, the difference $R_{\text{opt}} - R_{\text{sc}}(\alpha_0)$ monotonically decreases with γ_2. Hence, the calculation with $\gamma_2 = \gamma_{2,0}$ provides an upper bound to the value of $\frac{R_{\text{opt}} - R_{\text{sc}}(\alpha_0)}{R_{\text{sc}}(\alpha_0)}$. It can be seen that the usage of α_0 is only slightly suboptimal, which renders the further optimization of α unnecessary.

V. Practical Design of the SC–Relaying

Analogously to the case in which practical adaptive modulation and coding (AMC) approximates the ideal AMC for given SNR (given by the capacity expression (1)), the practical implementation of the SC–relaying is an approximation to the scheme designed in the previous section. The objective of this section is to demonstrate that the idea of SC–relaying can indeed bring gain in the spectral efficiency even when the information–theoretic assumptions are removed. The result will be a method for adapting the rate between A_S and A_D by selection of the transmission method.

A. Note on the Modulation Levels

In order to present an exemplary design, we consider a system with uncoded transmission. We assume that the system has two different modulations QPSK and 16–QAM, respectively, such that when the transmission is done over a direct link, there is a single SNR threshold for switching between the two modulations.

The constellation used for QPSK is $\{e^{j\frac{\pi}{4}}, e^{j\frac{3\pi}{4}}, e^{j\frac{5\pi}{4}}, e^{j\frac{7\pi}{4}}\}$. Regarding the 16–QAM constellation, instead the usual one with Gray encoding, in this example we have opted to use another type of 16–QAM, created by superposition of two independent QPSK–modulated packets. This is because we can use the same modulation to do the function of superposition coding in Step 1 and as a 16–QAM modulation over a direct link [2]. A 16–QAM symbol is represented as:

$$x_{16} = \sqrt{1 - \alpha_{16}}\, x_b + \sqrt{\alpha_{16}}\, x_s \tag{17}$$

where $\alpha_{16} = \frac{1}{5}$, while x_b, x_s are QPSK symbols. With such α_{16}, the constellation points for x_{16} correspond to a standard 16–QAM constellation with average power of 1.

Let $P_{\text{QPSK}}(\gamma)$ denote the probability that a packet of N QPSK symbols is received correctly when transmitted at SNR of γ. When 16–QAM is used for transmission over a direct link with γ, let $P_{16,b}(\gamma)$ denote the probability that the basic QPSK message in the superposition–based 16–QAM transmission

[2] This is done solely for simplicity, we could have as well used the usual 16–QAM for transmission over the direct links; however, the conclusions will not change and we in fact will show that by having only these two modulations, we can still reap the benefits of SC–relaying.

is received correctly. The superposed message is discarded if the basic QPSK message is not received correctly. The expected throughput of QPSK and 16–QAM transmission is, respectively:

$$\tau_{QPSK}(\gamma) = 2P_{\text{QPSK}}(\gamma) \quad (18)$$

$$\tau_{16-QAM}(\gamma) = 2P_{\text{16,b}}(\gamma) + 2P_{\text{16,b}}(\gamma)P_{\text{QPSK}}(\alpha_{16}\gamma) \quad (19)$$

Note that $P_{\text{16,b}}(\gamma) \neq P_{\text{QPSK}}((1-\alpha_{16})\gamma)$, as the interference from the superposed QPSK message is not Gaussian. It is also worth noting that the throughput of the described uncoded 16–QAM scheme is slightly better than 16–QAM with Gray coding and only a single packet, since with the described scheme in some instances at least 2 bits per symbol succeed to be transmitted.

We fix the transmission of the source A_S to consist of $N = 200$ symbols. If the two modulations are used to design adaptive modulation scheme based on maximal throughput over a direct link, then the threshold for switching between the two is made at γ_t, such that $\tau_{QPSK}(\gamma_t) = \tau_{16-QAM}(\gamma_t)$ and, for $N = 200$, we have found $\gamma_t = 16.2$ dB. Furthermore, we have found that for $\gamma \geq 20$ dB, the transmission with 16–QAM is virtually error–free and $\tau_{16-QAM}(\gamma) = 4$.

Note that the DF scheme is excluded from this evaluation, as in this particular example it would have performed almost identically with the multi–hop scheme (with a negligibly higher reliability for the 16–QAM signal received at the destination).

B. Application of the Modulation Schemes in SC–relaying

Now consider the relay system from Fig. 1 and assume that $\gamma_1 = \gamma_2 = 20$ dB, such that the 16–QAM transmissions $A_S A_R$ and $A_R A_D$ can be assumed to be error–free. The question we want to ask is: *For given γ_0, how should the source select the transmission mode in order to maximize the throughput?* For the chosen example, the multi–hop transmission has $R_1 = R_2 = 4$, such that the throughput is $R_{\text{mh}} = 2$. Furthermore, the direct transmission uses adaptive modulation: QPSK for $\gamma_0 < 16.2$ dB and 16–QAM otherwise.

Let us consider the operation of the SC–relaying for the given parameters. After Step 1, A_R takes the $2N = 400$ bits of the superposed messages and transmits them by using the 16–QAM modulation. In fact, it creates two messages of 200 bits each and uses the superposition. The transmission during Step 2 takes $\frac{400}{4} = \frac{N}{2}$ symbols. Having that a 16–QAM transmission of 200 symbols can be treated as error–free at $\gamma_2 = 20$ dB, the same holds for the transmission of 100 symbols. After A_D decodes the transmission of A_R, it creates the superposed message and subtracts it from the signal originally received from A_S and it should next decode a QPSK signal at a SNR of $(1-\alpha_{16})\gamma_0$. If K 2–step transmissions are made, then the total number of bits received successfully at A_D is:

$$K \cdot 2N + K \cdot 2N \cdot P_{\text{QPSK}}\left((1-\alpha_{16})\gamma_0\right) \quad (20)$$

The duration of a single 2–step transmission is $N + \frac{N}{2} = \frac{3N}{2}$ symbols, such that the expected throughput for the SC–relaying scheme is

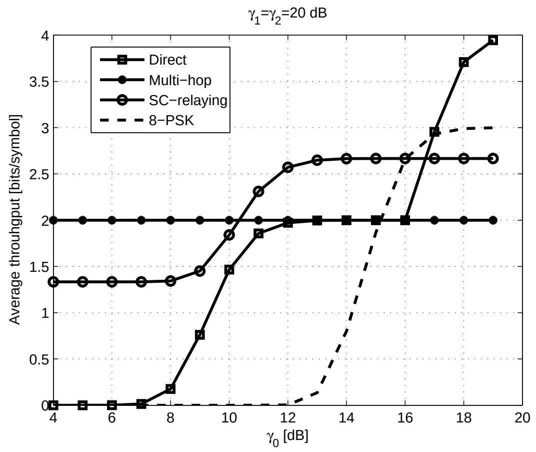

Fig. 4. Achieved throughput for different transmission modes for the case of uncoded modulations. The transmission of the source has a duration of $N = 200$ symbols.

$$\tau_{\text{sc}}(\gamma_0) = \frac{K \cdot 2N + K \cdot 2N \cdot P_{\text{QPSK}}((1-\alpha_{16}\gamma_0)\gamma_0)}{K\frac{3N}{2}} = \frac{4 + 4P_{\text{QPSK}}((1-\alpha_{16})\gamma_0)}{3} \quad (21)$$

Fig. 4 depicts the average throughput offered by the different transmission modes versus the SNR of the direct link. It can be seen that there is a region of values of γ_0 for which the SC–relaying with α_{16} offers the best throughput and thus it should be the preferred transmission mode. Hence, the value of α for this system should be selected as:

$$\alpha = \begin{cases} 1 & \gamma_0 < 10.4 \text{ dB} \\ \alpha_{16} = \frac{1}{5} & 10.4 \text{ dB} < \gamma_0 < 16.7 \text{ dB} \\ 0 & \gamma_0 > 16.7 \text{ dB} \end{cases} \quad (22)$$

In this simple example we have seen that the SC–relaying can bring benefits even when α has only three discrete values $(0, \frac{1}{5}, 1)$. By additionally optimizing the choice of α e. g. with respect to the throughput, we can obtain further improvements. For example, if $\gamma_1 = 30$ dB, then the superposed message might need to be chosen to be modulated with 16–QAM (recall that the rate of the superposed message in the information–theoretic design grows with γ_1). Note that, from the point of view of signalling overhead, it is better to have a smaller set of usable values α, since the value should be reported to A_D and used in the decoding.

One might argue that if there are more intermediate modulation levels available (between QPSK and 16–QAM), then the direct link might not need to use SC–relaying. On Fig. 4 we have depicted the throughput offered by 8–PSK and it is clear that there is still gain from using the SC–relaying. However, it should be noted that the introduction of more modulation levels also implies finer adaptation on the links $A_S A_R$ and $A_R A_D$ and richer set of available values for α, which should be taken into account when assessing the performance of the SC–relaying.

It can be concluded that the SC–relaying makes the problem of adaptive modulation and coding multidimensional, since it is not anymore function only of the SNR on the link (in this case γ_0), but is parameterized by γ_1, γ_2 and the available set of values for α.

VI. Conclusions

We have considered the problem of relay–based transmission in a wireless system that consists of a source A_S, relay A_R and destination A_D. We have introduced a novel 2–step transmission scheme based on superposition coding, termed SC–relaying. In Step 1, A_S broadcasts a message that is obtained using superposition coding. A part of this message is relayed in Step 2, where A_R adapts the transmission rate with respect to the link quality towards A_D. The key point of the proposed scheme is that the signal format in the two steps is chosen in a way that enables a simple combining at A_D. For the information–theoretic design, the numerical results show that the proposed scheme can bring a notable gain in spectral efficiency over the conventional multi–hop transmission with adaptive modulation, while performing very closely to the information-theoretic optimal scheme. Regarding the practical usage of the SC–relaying, we have demonstrated that it can bring gains even for the uncoded transmissions and we have shown that the introduction of such a relaying scheme generalizes the problem of adaptive modulation and coding. As a future work, we will investigate the optimization of the SC–relaying in case of multiple available transmission channels, e. g. in OFDM.

Acknowledgment

The authors would like to thank Samsung Electronics, Korea for supporting this work.

References

[1] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[2] G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” *IEEE Trans. Inf. Theory*, vol. IT-51, no. 9, pp. 3037–3063, 2005.

[3] Y. Liang and V. V. Veeravalli, “Gaussian orthogonal relay channels: Optimal resource allocation and capacity,” *IEEE Trans. Inf. Theory*, vol. IT-51, no. 9, pp. 3284–3289, 2005.

[4] T. Cover and A. E. Gamal, “Capacity theorems for the relay channel,” *IEEE Trans. Inf. Theory*, vol. IT-25, pp. 572–584, 1979.

[5] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[6] D. Gunduz and E. Erkip, “Opportunistic cooperation by dynamic resource allocation,” in *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, Apr. 2007, pp. 1446–1454.

[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons Inc., 1991.

[8] A. K. Goparaju, S. Wei, and Y. Liu, “On superposition coding based cooperative diversity schemes,” in *Proc. Asilomar Conference on Signals, Systems and Computers,*, Pacific Grove, California, USA,, Oct. 2005.

[9] P. Popovski and E. de Carvalho, “Spectrally–efficient wireless relaying based on superposition coding,” in *Proc. 65th IEEE Vehicular Technology Conference (VTC Spring 2007),*, Dublin, Ireland, Apr. 2007.