Spectral Efficient Signaling for Half-duplex Relay Channels

Boris Rankov and Armin Wittneben
Swiss Federal Institute of Technology (ETH) Zurich
Communication Technology Laboratory
Sternwartstrasse 7, 8092 Zurich, Switzerland
E-Mail: {rankov, wittneben}@nari.ee.ethz.ch

***Abstract*— We study two-hop communication protocols where one or two relay terminals assist in the communication between two transceiver terminals. All terminals operate in half-duplex mode, i.e., may not receive and transmit simultaneously at the same time and frequency. This leads to a loss in spectral efficiency due to the pre-log factor 1/2 in corresponding expressions for the achievable rate (capacity). We propose and analyze two relaying protocols that avoid the pre-log factor 1/2 but still work with half-duplex relays. Firstly, we consider a relaying protocol where two half-duplex relays, either amplify-and-forward (AF) or decode-and-forward (DF), alternately forward messages from a source terminal to a destination terminal (*two-path relaying*). It is shown that the protocol can recover a significant portion of the half-duplex loss. Secondly, we propose a relaying protocol where a bidirectional connection between two transceiver terminals is established via one half-duplex AF or DF relay (*two-way relaying*). It is shown that the sum rate of the two-way half-duplex AF relay channel achieves the rate of the one-way full-duplex AF relay channel, whereas the sum rate of the two-way half-duplex DF relay channel achieves the rate of the one-way full-duplex DF relay channel only in certain cases.**

I. Introduction

The design and analysis of cooperative transmission protocols for wireless networks has recently attracted a lot of interest. Of particular interest are two-hop channels where a relay terminal assists in the communication between a source terminal and a destination terminal. For example, in [1] the authors consider a relay network with one source and one destination both equipped with M antennas and K half-duplex relays each equipped with $N \geq 1$ antennas. In the absence of a direct link between source and destination and the use of amplify-and-forward (AF) relays the authors show that the capacity scales as $\frac{M}{2}\log(\mathrm{SNR})$ for high signal-to-noise ratios (SNR) when the number of relays K grows to infinity. The pre-log factor $\frac{1}{2}$ is induced by the half-duplex signaling and causes a substantial loss in spectral efficiency. Further half-duplex relaying protocols with a pre-log factor $\frac{1}{2}$ can be found in [2] and the references therein. One way to avoid the pre-log factor $\frac{1}{2}$ is to use a full-duplex relay that may receive and transmit at the same time and frequency [3], but such a relay is difficult to implement. Large differences in the signal power of the transmitted and the received signal drive the relay's analog amplifiers in the receive chain into saturation and cause problems for the cancelation of the self-interference.

Previous Work. In [4] the authors address the half-duplex loss by proposing a spatial reuse of the relay slot. They consider a base station that transmits K messages to K users and their corresponding relays in K orthogonal time slots. In time slot $K+1$ each relay retransmits its received signal, causing interference to the other users. The capacity of a single connection (base station to user) has then a pre-log factor $\frac{K}{K+1}$ instead of $\frac{1}{2}$. Another solution is presented in [5] where the authors propose a transmission scheme with two half-duplex AF relays that alternately forward messages from a source to a destination. In order to decrease the inter-relay interference, one relay performs interference cancelation. This cooperation scheme turns the equivalent channel between source and destination into a frequency-selective channel. A maximum likelihood sequence estimator at the destination is applied to extract the introduced diversity, an idea which is known as delay diversity [6]. In [7] we proposed two half-duplex relaying schemes that mitigate the loss in spectral efficiency due to the half-duplex operation of the relays. *Firstly*, [7] considers a similar relaying scheme as in [5] but with the difference that both relays are only allowed to amplify-and-forward their received signals (no cancelation of the inter-relay interference at one of the relays as in [5]), whereas the destination employs successive decoding with partial or full cancelation of the inter-relay interference. *Secondly*, [7] proposes a relaying protocol where a synchronous bidirectional connection between two terminals (e.g., two wireless routers) is established using one AF half-duplex relay. The scheme was also extended to bidirectional communication between multiple terminal pairs assisted by multiple nonregenerative relays.

Contribution of this Work. We extend the protocols described in [7] to the case where decode-and-forward (DF) half-duplex relays are used and compare them with the AF protocols. For the first protocol (two-path relaying) it is shown that with DF relays the half-duplex loss can be recovered for strong inter-relay channels as well as for weak inter-relay channels whereas the AF protocol works well for weak to moderate inter-relay channels. In the second protocol (two-way relaying) we show that the sum rate of the two-way half-duplex AF relay channel achieves the rate of the corresponding one-way full-duplex AF relay channel whereas the two-way half-duplex DF relay channel achieves the rate of the corresponding full-duplex DF relay channel when the relay is near to one of the transmitting terminals.

The paper is organized as follows. Section II introduces two-path relaying and derives the achievable rates of this protocol for AF and DF relays. Section III discusses two-way relaying and the achievable sum rates again for AF and DF relays. Some numerical examples are given in Section IV.

II. Protocol I: Two–path Relaying

We consider transmissions of messages from a source S to a destination D via two relays R_1 and R_2, which may not receive and transmit simultaneously. We assume that there is no direct connection between S and D, e.g., due to shadowing or too large separation between S and D. A message is transmitted in two time slots. In the first slot the source transmits the message to relay R_1 or R_2 and in the second slot the message is forwarded to the destination. The length of one slot is equal to the length of one codeword (frame) and is NT, where T is the sampling interval and N the number of symbols in each frame. In odd time slots, $k = 1, 3, 5, \ldots$, relay R_1 receives and R_2 transmits. (Except for $k = 1$, where R_2 does not transmit), whereas in even time slots, $k = 2, 4, 6, \ldots$, it is the other way around. This cooperation protocol avoids the pre-log factor $\frac{1}{2}$ since the source transmits a new message in every time slot and has not to be quiet in each second time slot. However, since the relays do not operate in orthogonal channels, there will be interference between R_1 and R_2 and it is not clear *a priori* whether this inter-relay interference cancels the achieved gain achieved by the increased pre-log factor.

Assume that a sequence of K messages is to be transmitted. In time slot $k \in \{1, 2, \ldots, K\}$ the source S chooses randomly a message (index) $M[k] \in \{1, 2, \ldots, 2^{NR[k]}\}$ according to a uniform distribution with $R[k]$ being the achievable rate for frame k. The message $M[k]$ is then mapped to a codeword $s[k] = (s[k, 1], s[k, 2], \ldots, s[k, N])$ of length N where the symbols $\{s[k, n]\}_{k,n}$ are independent and identically distributed (i.i.d.) according to $\mathcal{CN}(0, P_\mathrm{s})$ with P_s the average transmit power of the source.

A. Amplify-and-forward

We review here the protocol described in [7] for AF relays. Let S be node 0, R_1 node 1, R_2 node 2 and D node 3. The channel gain between node i and node j at the discrete time $[k, n] := (kN + n)T$ is denoted as $h_{ij}[k, n]$, with $\mathbb{E}|h_{01}|^2 = \mathbb{E}|h_{02}|^2 = \nu_1^2$, $\mathbb{E}|h_{13}|^2 = \mathbb{E}|h_{23}|^2 = \nu_2^2$ and $\mathbb{E}|h_{12}|^2 = \nu_{12}^2$. Due to notational simplicity, we assume channel reciprocity for the inter-relay channel and equal fading variances from the source to both relays and equal fading variances from both relays to the destination, respectively. We assume that $\{h_{ij}[k, n]\}_{k,n}$ are independent, stationary and ergodic processes. The source S knows the *fading distribution* of the channel gains in the network but not the *fading realizations*. The relay nodes R_1 and R_2 do not have any channel knowledge and the destination knows the *fading realizations* of all channel gains in the network. The receive signal of relay R_p at time instant $(kN + n)T$ with $p = 2 - \mathrm{mod}\,(k, 2) \in \{1, 2\}$ and $q = 2 - \mathrm{mod}\,(k - i, 2) \in \{1, 2\}$ is given as

$$r[k,n] = h_{0p}[k,n]s[k,n] + n_\mathrm{r}[k,n] + \sum_{i=1}^{k-1}\Big(h_{0q}[k-i,n]s[k-i,n] + n_\mathrm{r}[k-i,n]\Big)f_i[k,n] \quad (1)$$

where

$$f_i[k,n] := \prod_{j=1}^{i} h_{12}[k+1-j,n]g[k-j] \quad (2)$$

denotes the *inter-relay interference factor*. The relay noise samples $\{n_\mathrm{r}[k, n]\}_{k,n}$ are i.i.d. according to $\mathcal{CN}(0, \sigma_\mathrm{r}^2)$. The transmit signal of relay R_p is a scaled version of its received signal: $t[k + 1, n] = g[k]r[k, n]$, where $g[k]$ is the scaling coefficient of either relay R_1 or R_2 and for $k = 2, 3, 4, \ldots$ chosen as[1]

$$g^2[k] = \frac{P_\mathrm{r}}{\frac{1}{N}\sum_{n=1}^{N}|r[k,n]|^2} \approx \frac{P_\mathrm{r}}{P_\mathrm{s}\nu_1^2 + P_\mathrm{r}\nu_{12}^2 + \sigma_\mathrm{r}^2} := g^2 \quad (3)$$

and in the first time slot $k = 1$

$$g[1] = \frac{P_\mathrm{r}}{\frac{1}{N}\sum_{n=1}^{N}|r[1,n]|^2} \approx \frac{P_\mathrm{r}}{P_\mathrm{s}\nu_1^2 + \sigma_\mathrm{r}^2} \geq g^2 \quad (4)$$

where P_r is the average transmit power of each relay. The approximations in (3) and (4) are exact for $N \to \infty$ by the law of large numbers.

Destination D observes at $((k + 1)N + n)T$ the signal

$$d[k+1,n] = h_{p3}[k+1,n]g[k]r[k,n] + n_\mathrm{d}[k+1,n] \quad (5)$$

where the noise samples $\{n_\mathrm{d}[k, n]\}_{k,n}$ are i.i.d. according to $\mathcal{CN}(0, \sigma_\mathrm{d}^2)$ and $r[k, n]$ is given in (1). To decode $s[k]$ from (5) the destination receiver first subtracts the previously decoded codewords $s[k–1], \ldots, s[1]$ from the received signal $d[k+1, n]$, because these codewords appear as accumulated inter-relay interference at the destination. However, the influence of the codewords transmitted several time slots before k is weak since they were attenuated several times by the inter-relay channel h_{12}, which acts as forgetting factor for the decoding process, see (1) and (2).

After perfect cancelation of m previously decoded codewords $s[k-1], s[k-2], \ldots, s[k-m]$ the destination signal is given by (6) at the bottom of the page where $f_i[k, n] := 1$

[1]We omit here to denote by the index p whether the receive signal, transmit signal, scaling coefficient and relay noise belong to R_1 or R_2

$$d_m[k+1,n] = h_{p3}[k+1,n]g[k]\left(h_{0p}[k,n]s[k,n] + \sum_{i=m+1}^{k-1} h_{0q}[k-i,n]s[k-i,n]f_i[k,n] + \sum_{i=0}^{k-1} n_\mathrm{r}[k-i,n]f_i[k,n]\right) + n_\mathrm{d}[k+1,n] \quad (6)$$

for $i = 0\ \forall k, n$. For $m = k-1$ all previously transmitted codewords are canceled (*full interference cancelation*). For $m = 0$ all codewords up to $s[k-1]$ appear as inter-frame interference when $s[k]$ is decoded. For $0 < m < k-1$ only the last m transmitted codewords are canceled and $s[1], s[2], \ldots, s[k-m-1]$ remain as interference terms (*partial interference cancelation*). The ergodic rate in time slot $k+1$ measured in b/s/Hz follows as

$$R[k{+}1] = \mathbb{E}\log\left(1 + \frac{P_{\mathrm{s}}|h_{p3}gh_{0p}|^2}{\sigma_{\mathrm{d}}^2 + |h_{p3}g|^2\big(P_{\mathrm{s}}I_1[k] + \sigma_{\mathrm{r}}^2 I_2[k]\big)}\right) \quad (7)$$

with inter-frame interference

$$I_1[k] = \sum_{i=m+1}^{k-1} |h_{0q}|^2 |f_i[k]|^2 \quad (8)$$

and relay noise interference

$$I_2[k] = \sum_{i=0}^{k-1} |f_i[k]|^2. \quad (9)$$

Note that $f_i[k]$ models the inter-relay interference factor as random variable whose statistics depends on k whereas $f_i[k,n]$ denotes its realization in time slot k at symbol time n[2]. After the first time slot, i.e., for $k = 1$, we have $I_1[1] = 0$ and $I_2[1] = 1$. Clearly, $R[1] = 0$ because after transmission of the first frame no signal is received by the destination yet. The expectation is taken with respect to the statistics of h_{0p}, h_{p3} for $p \in \{1, 2\}$ and h_{12} and depends on the channel model that is used for the fading variables. It can be shown that $R[k+1]$ is a non-increasing sequence when choosing $g[1] = g$ [8]. Therefore, after transmission of a sequence of K messages we get the average rate

$$\overline{R}_K = \frac{1}{K+1}\sum_{k=1}^{K} R[k{+}1] \quad (10)$$

$$\geq \frac{K}{K+1} R[K{+}1] \quad (11)$$

$$\geq \frac{K}{K+1} \lim_{k\to\infty} R[k] \quad (12)$$

where the pre-log $\frac{K}{K+1} \approx 1$ for large K. The disadvantage of signaling according to (7) is that the source has to adapt the rate for each frame. However, the lower bounds in (11) and (12) suggest to use a fixed-rate scheme at the source, either $R[K{+}1]$ or $\lim_{k\to\infty} R[k]$. By using $\lim_{k\to\infty} R[k]$ the rate is independent of the number of messages K to be transmitted. In order to simplify the computation of $\lim_{k\to\infty} R[k]$ we lower bound the rate (7). For $k = 2, 3, \ldots, K$ it is [7]

$$\begin{aligned} R[k+1] \geq\ & \mathbb{E}\log\Big(P_{\mathrm{s}}|h_{p3}gh_{0p}|^2 + \sigma_{\mathrm{d}}^2 + \\ & |h_{p3}g|^2\big(P_{\mathrm{s}}I_{1,k'}[k] + \sigma_{\mathrm{r}}^2 I_{2,k'}[k]\big)\Big) \\ & - \log\Big(\sigma_{\mathrm{d}}^2 + \nu_2^2 g^2\big(P_{\mathrm{s}}\overline{I}_1[k] + \sigma_{\mathrm{r}}^2\overline{I}_2[k]\big)\Big) \\ =\ & R_{\mathrm{low}}[k{+}1] \end{aligned} \quad (13)$$

[2]Similar for the channels: h_{ij} is the random variable whose statistics remain the same for all time and $h_{ij}[k,n]$ its realization at $(kN+n)T$

with

$$I_{1,k'}[k] = \sum_{i=m+1}^{k'} |h_{0q}|^2 |f_i[k]|^2 \quad (14)$$

$$I_{2,k'}[k] = \sum_{i=0}^{k'} |f_i[k]|^2 \quad (15)$$

and

$$\overline{I}_1[k] = \sum_{i=m+1}^{k-1} \mathbb{E}|f_i[k]|^2 = \frac{q^{m+1} - q^k}{1-q} \quad (16)$$

$$\overline{I}_2[k] = \sum_{i=0}^{k-1} \mathbb{E}|f_i[k]|^2 = \frac{1 - q^k}{1-q} \quad (17)$$

where $q = g^2\nu_{12}^2$. For $k \to \infty$ and $q < 1$ we get $\lim_{k\to\infty}\overline{I}_1[k] = \frac{q^{m+1}}{1-q}$ and $\lim_{k\to\infty}\overline{I}_2[k] = \frac{1}{1-q}$ and the lower bound (13) becomes independent of the actual frame number k. Note that for a stationary inter-relay channel h_{12} the statistics of $I_{1,k'}[k]$ and $I_{2,k'}[k]$ become independent of k (but dependent on k'). Numerical results in Section IV show that fixed-rate signaling according to $\lim_{k\to\infty} R_{\mathrm{low}}[k]$ or $R_{\mathrm{low}}[K+1]$ in each frame induces only a small loss compared to variable-rate signaling according to (7), but has the advantage that the source does not have to adapt the rate for each frame. The parameter k' can be used to improve the lower bound: the larger k' the better the lower bound, but the more involved becomes the evaluation of the expectation of the first $\log$-term in (13).

B. Decode-and-forward

At the discrete time $[k,n] := (kN+n)T$ the receive signal of relay $\mathtt{R}_p$ with $p = 2 - \mathrm{mod}\,(k,2) \in \{1,2\}$ and $q = 2 - \mathrm{mod}\,(k-i,2) \in \{1,2\}$ is given as

$$r[k,n] = h_{0p}[k,n]s[k,n] + h_{12}[k,n]s[k{-}1,n] + n_{\mathrm{r}}[k,n]. \quad (18)$$

Relay $\mathtt{R}_p$ may choose two different decoding strategies: a) the relay decodes $s[k,n]$ treating $h_{12}[k,n]s[k{-}1,n]$ as interference or b) the relay decodes first $s[k{-}1,n]$ treating $h_{0p}[k,n]s[k,n]$ as interference, subtracts $h_{12}[k,n]s[k{-}1,n]$ from the received signal $r[k,n]$, then decodes $s[k,n]$ interference-free. The achievable rate with strategy a) is given by

$$R_{\mathrm{a}} = \min\left\{C\left(\frac{P_{\mathrm{s}}|h_{0p}|^2}{\sigma_{\mathrm{r}}^2 + P_{\mathrm{r}}|h_{12}|^2}\right), C\left(\frac{P_{\mathrm{r}}|h_{p3}|^2}{\sigma_{\mathrm{d}}^2}\right)\right\} \quad (19)$$

where $C(x) = \mathbb{E}\log(1+x)$. The first term in (19) determines the maximum rate in the first hop (source to relay $\mathtt{R}_p$) when the inter-relay signal is treated as interference and the second term denotes the maximum rate in the second hop (from relay $\mathtt{R}_p$ to destination). The achievable rate with strategy b) follows as

$$\begin{aligned} R_{\mathrm{b}} = \min\Bigg\{ & C\left(\frac{P_{\mathrm{s}}|h_{0p}|^2}{\sigma_{\mathrm{r}}^2}\right), C\left(\frac{P_{\mathrm{r}}|h_{p3}|^2}{\sigma_{\mathrm{d}}^2}\right), \\ & C\left(\frac{P_{\mathrm{r}}|h_{12}|^2}{\sigma_{\mathrm{r}}^2 + P_{\mathrm{s}}|h_{0q}|^2}\right)\Bigg\}. \end{aligned} \quad (20)$$

The first term in (20) denotes the maximum rate in the first hop, when there is no interference from the other relay, the second term is again the maximum rate in the second hop. The third term is the maximum rate from one relay to the other relay when the source signal is treated as interference. Strategy a) works well when the inter-relay channel is not too strong since the inter-relay interference in the first capacity expression of (19) is small. Strategy b) works well, when the inter-relay channel is strong, since the rate in (20) is not dominated by the third expression and the relay may decode the new message coming from the source interference-free.

III. Protocol II: Two-way Relaying

In the relaying scheme described in the previous section we needed two relays to circumvent the pre-log factor $\frac{1}{2}$ in the achievable rate. Another solution arises, when we assume that two nodes $\mathtt{T}_1$ and $\mathtt{T}_2$ want to establish a synchronous bidirectional connection (for example two wireless routers), i.e., both nodes communicate in both directions through a common half-duplex relay $\mathtt{R}$.

A. Amplify-and-forward

The proposed relaying scheme works as follows: in time slot k both nodes $\mathtt{T}_1$ and $\mathtt{T}_2$ transmit their symbols to relay $\mathtt{R}$ in the same time slot and the same bandwidth. The relay scales the received signal in order to meet its average power constraint and retransmits the signal in the next time slot. The received signal at node $\mathtt{T}_i$, $i=1,2$, in time slot $k+1$ is[3]

$$\begin{aligned} y_i[k+1] &= h_i[k+1]g[k]h_j[k]x_j[k] + h_i[k+1]g[k]h_i[k]x_i[k] \\ &\quad + h_i[k+1]g[k]\cdot n_{\mathrm{r}}[k] + n_i[k+1] \end{aligned} \tag{21}$$

where $i=2, j=1$ for $\mathtt{T}_1 \rightarrow \mathtt{R} \rightarrow \mathtt{T}_2$ and $i=1, j=2$ for $\mathtt{T}_1 \leftarrow \mathtt{R} \leftarrow \mathtt{T}_2$[4]. The i.i.d. symbols $x_1[k] \sim \mathcal{CN}(0,P_1)$ and $x_2[k] \sim \mathcal{CN}(0,P_2)$ are the transmit symbols of node $\mathtt{T}_1$ and $\mathtt{T}_2$, respectively. h_1 is the channel gain between $\mathtt{T}_1$ and relay $\mathtt{R}$ and h_2 the channel gain between $\mathtt{T}_2$ and relay $\mathtt{R}$[5]. Additive white Gaussian noise (AWGN) at the relay is denoted by $n_{\mathrm{r}} \sim \mathcal{CN}(0,\sigma_{\mathrm{r}}^2)$ and $n_i \sim \mathcal{CN}(0,\sigma_i^2)$ denotes AWGN at node $\mathtt{T}_i$. Since nodes $\mathtt{T}_1$ and $\mathtt{T}_2$ know their own transmitted signals they can subtract the back-propagating self-interference in (21) prior to decoding, assuming perfect knowledge of the corresponding channel coefficients. The sum rate of this protocol is given by

$$R_{\mathrm{sum}}^{\mathrm{AF}} = C_{\mathrm{AF}}^{\rightarrow} + C_{\mathrm{AF}}^{\leftarrow} \tag{22}$$

with

$$C_{\mathrm{AF}}^{\rightarrow} = \frac{1}{2} C\left(\frac{P_1|h_2 g h_1|^2}{\sigma_2^2 + \sigma_{\mathrm{r}}^2 |h_2 g|^2}\right) \tag{23}$$

$$C_{\mathrm{AF}}^{\leftarrow} = \frac{1}{2} C\left(\frac{P_2|h_2 g h_1|^2}{\sigma_1^2 + \sigma_{\mathrm{r}}^2 |h_1 g|^2}\right). \tag{24}$$

[3] k denotes here discrete symbol time

[4] $\mathtt{A} \rightarrow \mathtt{B}$ indicates information flow from node $\mathtt{A}$ to node $\mathtt{B}$.

[5] Again we assume reciprocity for both channels.

The relay gain is chosen as

$$g = \left(\frac{P_{\mathrm{r}}}{P_1|h_1|^2 + P_2|h_2|^2 + \sigma_{\mathrm{r}}^2}\right)^{\frac{1}{2}} \tag{25}$$

where P_{r} is the average transmit power of relay $\mathtt{R}$. Note that the transmission in each direction suffers still from the pre-log factor $\frac{1}{2}$. However, the half-duplex constraint can here be exploited to establish a bidirectional connection between two nodes and to increase the sum rate of the network.

B. Decode-and-forward

We consider now a bidirectional communication between terminals $\mathtt{T}_1$ and $\mathtt{T}_2$ via a half-duplex DF relay $\mathtt{R}$. Terminal $\mathtt{T}_1$ encodes its message w into the codeword $\underline{x}_1(w)$ with rate R_1 and average power P_1. Terminal $\mathtt{T}_2$ encodes its message v into the codeword $\underline{x}_2(v)$ with rate R_2 and average power P_2. Both codewords are then transmitted in N channel uses to the relay $\mathtt{R}$, which decodes the messages. The relay then encodes both messages again using the codebooks of terminal $\mathtt{T}_1$ and $\mathtt{T}_2$ and broadcasts the sum signal $\underline{x}_1(w)+\underline{x}_2(v)$ with sum power P_{r} in N channel uses to both terminals. Since terminal $\mathtt{T}_1$ knows $\underline{x}_1(w)$ and its channel gain to the relay it can subtract the back-propagating self-interference from the the received signal. The same is true for terminal $\mathtt{T}_2$. The sum rate of this protocol is given by

$$R_{\mathrm{sum}}^{\mathrm{DF}} = \max_{\beta} \min\left\{C_{\mathrm{MA}}, C_{\mathrm{DF}}^{\rightarrow}(\beta) + C_{\mathrm{DF}}^{\leftarrow}(\beta)\right\} \tag{26}$$

where

$$C_{\mathrm{DF}}^{\rightarrow}(\beta) = \frac{1}{2}\min\left\{C\left(\frac{P_1|h_1|^2}{\sigma_{\mathrm{r}}^2}\right), C\left(\frac{\beta P_{\mathrm{r}}|h_2|^2}{\sigma_2^2}\right)\right\} \tag{27}$$

$$C_{\mathrm{DF}}^{\leftarrow}(\beta) = \frac{1}{2}\min\left\{C\left(\frac{P_2|h_2|^2}{\sigma_{\mathrm{r}}^2}\right), C\left(\frac{(1-\beta)P_{\mathrm{r}}|h_1|^2}{\sigma_1^2}\right)\right\} \tag{28}$$

$$C_{\mathrm{MA}} = \frac{1}{2} C\left(\frac{P_1|h_1|^2 + P_2|h_2|^2}{\sigma_{\mathrm{r}}^2}\right). \tag{29}$$

Notice that $\beta \in [0,1]$ is the fraction of the relay power P_{r} allocated to codeword $\underline{x}_1(w)$ and $1-\beta$ the fraction of P_{r} used for codeword $\underline{x}_2(v)$. If the relay does not have any knowledge of the channels to terminal $\mathtt{T}_1$ and $\mathtt{T}_2$ it chooses $\beta = \frac{1}{2}$. For the case the relay has some channel knowledge about h_1 and h_2 (it may learn it during the previous transmission from the terminals to the relay) β may be chosen such, that the sum rate is maximized. Clearly, the power allocation that maximizes the sum rate depends on the network geometry and the degree of channel knowledge (instantaneous channel gains, second order statistics, etc). For simplicity we consider a linear network topology as shown in Fig1 and assume that the relay has only knowledge of the pathloss coefficients. Further we choose $P_1 = P_2 = \frac{P}{2}$, $P_{\mathrm{r}} = P$ and $\sigma_{\mathrm{r}}^2 = \sigma_{\mathrm{d}}^2 = 1$. The optimal power allocation is then given by [8]

$$\beta^* = \begin{cases} \frac{d_2^\alpha}{2d_1^\alpha}, & d_1 \geq d_2; \\ 1 - \frac{d_1^\alpha}{2d_2^\alpha}, & d_1 < d_2 \end{cases} \tag{30}$$

where α is the pathloss exponent, d_1 the distance from terminal $\mathtt{T}_1$ to relay $\mathtt{R}$ and d_2 the distance between relay $\mathtt{R}$ and terminal

$\mathtt{T}_2$. As the relay moves towards terminal $\mathtt{T}_1$, i.e., $d_1 \to 0$, more relay power is spent for the $\mathtt{T}_1 \to \mathtt{R} \to \mathtt{T}_2$ transmission and less power for the reverse $\mathtt{T}_1 \leftarrow \mathtt{R} \leftarrow \mathtt{T}_2$ transmission. The reason is that the link capacity from terminal $\mathtt{T}_2$ to relay $\mathtt{R}$ is small and dominates the overall capacity for the $\mathtt{T}_1 \leftarrow \mathtt{R} \leftarrow \mathtt{T}_2$ transmission, irrespective of the relay power allocated to that transmission. When the relay moves towards terminal $\mathtt{T}_2$ it is the other way around.

IV. Numerical Examples

A. Two-path Relaying

We evaluate the achievable rates of the relaying schemes described in Section II by Monte Carlo simulations. The channel gains are i.i.d. (in space and time) complex normal with zero mean and channel variances ν_{12}^2, ν_1^2 and ν_2^2. We choose for all examples $\nu_1^2 = \nu_2^2 = 1$. The AWGN variances are chosen as $\sigma_{\mathrm{r}}^2 = \sigma_{\mathrm{d}}^2 = \sigma^2$ and the transmit powers $P_{\mathrm{s}} = P_{\mathrm{r}} = P$. We simulated 5000 random channels for each value in Fig.2 and Fig.3. The SNR is defined as $\mathrm{SNR} = \frac{P}{\sigma^2}$.

In Fig.2 we see that for $\nu_{12}^2 = 0.5$ (inter-relay channel is 3dB weaker than the other channels) the two-path relaying protocol with two alternating half-duplex (AHD) AF relays and full cancelation of the inter-relay interference achieves an average rate $\overline{R}_K$ that is near to the rate of one full-duplex relay and outperforms clearly the case where only one half-duplex relay is used. The notation $(2P, 2P)$ means that in the first time slot the total transmit power of the network is $2P$ and in the second time slot $2P$ (every node transmits with power P). For the lower bound (13) we have chosen $k^{'} = 0$ and $I_1[k] = 0$ (full interference cancelation). Further we observe from Fig.2 that the performance loss of the fixed-rate schemes based on $\lim_{k\to\infty} R_{\mathrm{low}}[k]$ or $R_{\mathrm{low}}[K+1]$ is small compared to the performance of the variable-rate scheme (7).

In Fig.3 we compare the rates achievable by two-path DF relaying with and without interference cancelation. We see that for strong inter-relay channels the relay should perform interference cancelation before decoding the message coming from the source and for weak inter-relay channels it is better to treat the signal from the other relay as interference. Note that the performance of the DF scheme performs only well for weak or strong inter-relay channels where the AF scheme performs well for weak to moderate inter-relay channels. All schemes consume a network power of $2P$ in each time slot.

B. Two-way Relaying

Again we evaluate the achievable rates of the relaying schemes described in Section III by Monte Carlo simulations. We consider a linear one-dimensional network geometry as in Fig.1. The channel gains are modeled as $h_i = \frac{\xi_i}{(1+d_i)^{\alpha/2}}$ with i.i.d. $\xi_i \sim \mathcal{CN}(0,1)$ (Rayleigh fading), where $1 + d_i$ is the distance between terminal $\mathtt{T}_i$ and relay $\mathtt{R}$ and $\alpha = 3$ the path loss exponent. The AWGN variances are chosen as $\sigma_1^2 = \sigma_2^2 = \sigma_{\mathrm{r}}^2 = \sigma^2$ and the transmit powers $P_1 = P_2 = P/2$ and $P_{\mathrm{r}} = P$ (i.e., the network consumes in each time slot an average power of P). We simulated 5000 random channels for each value in Fig.4 and Fig.5. The SNR is defined as

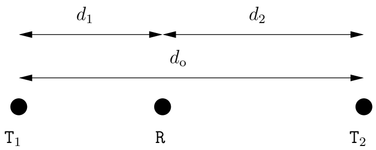

Fig. 1. One-dimensional linear relay network

$\mathrm{SNR} = \frac{P}{\sigma^2(1+d_i)^{\alpha}}$ and is chosen to be 20dB, when the relay is halfway between terminal $\mathtt{T}_1$ and terminal $\mathtt{T}_2$.

In Fig.4 we see that the sum rate of the two-way AF relay channel even outperforms slightly the rate achievable by the one-way full-duplex AF relay channel. The reason is that due to our power normalization the source as well as the relay power is $P/2$ in the one-way full-duplex protocol whereas in the two-way half-duplex protocol the powers are $P/2$ for each source terminal and the relay power is P, i.e., the relay gain may be chosen higher.

In Fig.5 we observe that the sum rate of the two-way DF relay channel achieves the rate of the full-duplex DF relay channel when the relay is in the vicinity of terminal $\mathtt{T}_1$ or terminal $\mathtt{T}_2$. The sum rate is improved by using the optimal power allocation given in (30). When the relay moves towards the halfway position between the terminals the sum rate drops down. The reason is that the sum rate is dominated here by the multiple access sum rate. In Fig.5 we also see the single-user rates. The relay first decodes the message from the stronger user, subtracts the decoded message from the received signal and then decodes the message from the weaker user interference-free. Note that when the relay is halfway between the two terminals the decoding order is switched. For comparison, the figure shows also the sum rate when the terminals do not exchange their messages via a relay but communicate with the relay in a consecutive uplink/downlink scheme: In the first time slot both terminals transmit their messages to the relay (base station) and in the second time slot the terminals receive new information from the relay (base station). The maximum sum rate for this communication scheme is given by

$$C_{\mathrm{u}} = \max_{\beta} \min \{C_{\mathrm{MA}}, C_{\mathrm{BC}}\} \tag{31}$$

with C_{MA} given in (29) and

$$C_{\mathrm{BC}} = \frac{1}{2} C\left(\frac{\beta P_{\mathrm{r}} |h_1|^2}{\sigma_1^2}\right) + \frac{1}{2} C\left(\frac{(1-\beta) P_{\mathrm{r}} |h_2|^2}{\sigma_2^2}\right). \tag{32}$$

The β that maximizes (31) is given by the standard water-filling algorithm [9]. The capacity C_{MA} is due to the multiple access cut which determines the maximum sum rate for the simultaneous transmission from terminals $\mathtt{T}_1$ and $\mathtt{T}_2$ to relay $\mathtt{R}$. The capacity C_{BC} is due to the broadcast cut which determines the maximum rate that is achievable from the relay to the terminals in interference free downlink channels[6] . We observe

[6]We assume orthogonal downlink channels in order to compare this sum rate with the two-way relay transmission, where due to the cancelation of the back-propagating self-interference the channels from the relay to the terminals are also orthogonal

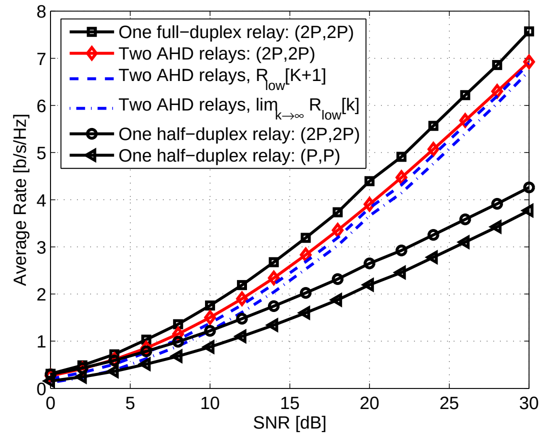

Fig. 2. Two-path relaying with amplify-and-forward and $\nu_{12}^2 = 0.5$

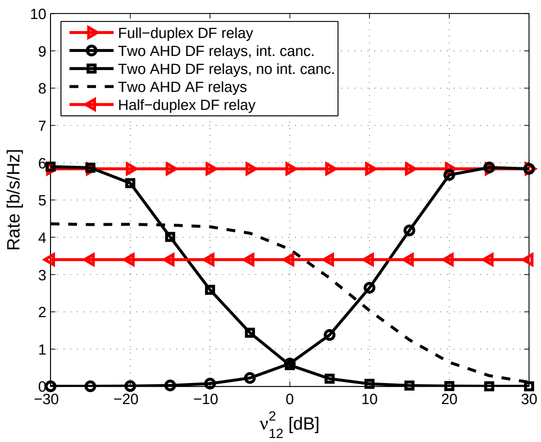

Fig. 3. Two-path relaying with decode-and-forward and SNR=20dB

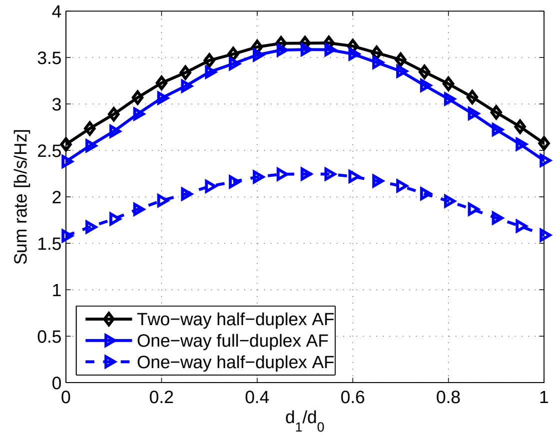

Fig. 4. Sum rate for two-way relaying with amplify-and-forward

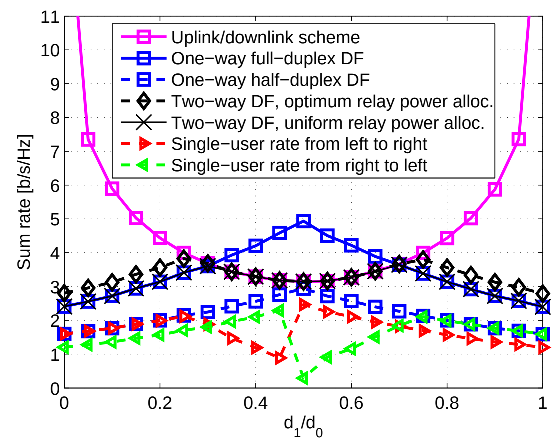

Fig. 5. Sum rate for two-way relaying with decode-and-forward

that the uplink/downlink sum rate is not tight when the relay moves to one of the terminals. The reason is that the broadcast cut does not consider that the high rate information obtained from the nearby terminal in the previous time slot cannot be transferred with the same rate to the distant terminal due to its weak channel.

V. Conclusions

We studied two half-duplex relaying protocols with increased spectral efficiency compared to conventional half-duplex relay protocols. For the first protocol (two-path relaying) it was shown that a large portion of the half-duplex loss can be recovered depending on the strength of the inter-relay channel and the operation of the relay (AF or DF). In the second relaying protocol we established a bidirectional connection between two nodes using one half-duplex relay (two-way relay channel). It was shown that the sum rate of the two-way half-duplex AF relay channel achieves the rate of the full-duplex AF relay channel. The two-way half-duplex DF relay channel achieves the rate of the full-duplex DF relay channel only in certain cases, depending on the network geometry, but always outperforms the half-duplex DF relay channel.

References

[1] H. Bölcskei, R. U. Nabar, O. Oyman, and A. J. Paulraj, “Capacity scaling laws in MIMO relay networks,” *IEEE Trans. Wireless Commun.*, 2006. To appear.

[2] N. J. Laneman, D. N. Tse, and G. W. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *IEEE Trans. Inform. Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[3] T. M. Cover and A. El Gamal, “Capacity theorems for the relay channel,” *IEEE Trans. Inform. Theory*, vol. 25, pp. 572–584, Sept. 1979.

[4] O. Munoz, A. Augustin, and J. Vidal, “Cellular capacity gains of cooperative MIMO transmission in the downlink,” in *Proc. IZS*, (Zurich, Switzerland), pp. 22–26, Feb. 2004.

[5] T. J. Oechtering and A. Sezgin, “A new cooperative transmission scheme using the space-time delay code,” in *Proc. ITG Workshop on Smart Antennas*, (Munich, Germany), pp. 41–48, Mar. 2004.

[6] A. Wittneben, “A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation,” in *Proc. ICC*, vol. 3, pp. 1630–1634, 1993.

[7] B. Rankov and A. Wittneben, “Spectral efficient protocols for nonregenerative half-duplex relaying,” in *Proc. Allerton Conf. Comm., Contr. and Comp.*, (Monticello, IL), Oct. 2005.

[8] B. Rankov and A. Wittneben, “Spectral efficient protocols for half-duplex relay channels,” in *IEEE Trans. Wireless Commun.*, 2005. in preparation.

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