Spectral Efficient Protocols for Half-Duplex Fading Relay Channels

Boris Rankov and Armin Wittneben

Abstract— **We study two-hop communication protocols where one or several relay terminals assist in the communication between two or more terminals. All terminals operate in half-duplex mode, hence the transmission of one information symbol from the source terminal to the destination terminal occupies two channel uses. This leads to a loss in spectral efficiency due to the pre-log factor one-half in corresponding capacity expressions. We propose two new half-duplex relaying protocols that avoid the pre-log factor one-half. Firstly, we consider a relaying protocol where a bidirectional connection between two terminals is established via one amplify-and-forward (AF) or decode-and-forward (DF) relay (*two-way relaying*). We also extend this protocol to a multi-user scenario, where multiple terminals communicate with multiple partner terminals via several orthogonalize-and-forward (OF) relay terminals, i.e., the relays orthogonalize the different two-way transmissions by a distributed zero-forcing algorithm. Secondly, we propose a relaying protocol where two relays, either AF or DF, alternately forward messages from a source terminal to a destination terminal (*two-path relaying*). It is shown that both protocols recover a significant portion of the half-duplex loss.**

Index Terms— **Relay channel, fading, half-duplex, spectral efficiency**

I. Introduction

THE ANALYSIS and design 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 (two channel uses) and causes a substantial loss in spectral efficiency. This loss is especially significant in high-SNR communication regimes[1]. 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 to the cancelation of the self-interference.

Related 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 $K+1$ all relays retransmit their 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 $\frac{1}{2}$. A similar scheme was proposed in [5] where the authors study the range extension potential of fixed half-duplex decode-and-forward relays in a cellular network. In order to prevent a loss in spectral efficiency (throughput) due to the half-duplex constraint of the relays the authors propose to reuse existing channels in the cellular network for the relay transmissions, thereby causing co-channel interference. They propose a relay channel selection scheme that keeps the level of co-channel interference low and show that significant throughput improvements can be achieved. Another solution is presented in [6] 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 [7]. However, the authors did not study the achievable rate of this scheme. A similar protocol with two alternating relays was introduced in [8] where the transmitting terminals, i.e., source, relay one and relay two, use orthogonal direct sequence spreading codes. By that the destination is able to separate the signals from the source and both relays. Also a relay may separate the signals transmitted by the source and the other relay. It was shown that with this system full diversity order of two is achievable without sacrificing bandwidth (besides the spreading). However, the system utilizes three different codes (code multiplex) to avoid interference between the transmissions. During the revision of this paper we discovered the work in [9] that was submitted recently and investigates a similar problem as we do in our work. The authors consider two relays that alternately receive and transmit data and where the direct channel is also used for transmissions. The relays

Manuscript received February 1, 2006; revised July 1, 2006. Part of this material was presented at the Allerton Conference 2005 and the Asilomar Conference 2005.

The authors are with Communication Technology Laboratory, ETH Zurich (e-mail: {rankov, wittneben}@nari.ee.ethz.ch).

Digital Object Identifier 10.1109/JSAC.2007.070213.

[1] In the low-SNR regime the capacity scales linearly with the signal-to-noise ratio and being power efficient is usually more important than being spectral efficient.

operate as half-duplex decode-and-forward transceivers with interference cancelation and the destination employs a V-BLAST receiver to resolve signal collisions between relay and source signals.

Contribution of this Work. We propose two half-duplex relaying protocols that mitigate the loss in spectral efficiency due to the half-duplex operation of the relays. *Firstly*, we propose a relaying protocol where a bidirectional connection (two-way protocol) between two terminals (e.g., two wireless routers) is established using one half-duplex AF or DF relay. Hereby, the achievable rate in one direction suffers still from the pre-log factor $\frac{1}{2}$ but since two connections are realized in the same physical channel we can achieve a sum-rate that is above the single user rate for the half-duplex relay channel. We extend the two-way protocol to a multi-user scenario, where multiple terminals communicate with multiple partner terminals via several orthogonalize-and-forward (OF) relays. *Secondly*, we consider a similar relaying scheme as in [6] and [8] but with the difference that the source and the relay operate in the same physical channel without using orthogonal spreading codes and that our AF relays only amplify-and-forward their received signals (no cancelation of the inter-relay interference at one of the relays as in [6]). We propose to employ successive decoding at the destination with partial or full cancelation of the inter-relay interference. We also analyze the achievable rate when DF relays are used. It is shown that this protocol can recover a significant portion of the half-duplex loss (pre-log factor $\frac{1}{2}$) for both AF and DF relays.

Notation and Organization of the Paper. We use bold upper letters to denote matrices and bold lower letters to denote vectors. Further $(\cdot)^*$, $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$ stand for conjugation, transposition and Hermitian transposition, respectively. $\mathcal{E}\{\cdot\}$ denotes the expectation operator, rk$(\cdot)$ the rank of a matrix and $\mathbf{a} \odot \mathbf{b}$ denotes elementwise multiplication of two vectors. $\mathbf{a} = \mathrm{diag}(\mathbf{A})$ denotes the vector that contains the diagonal elements of the matrix $\mathbf{A}$ and $\mathbf{A} = \mathrm{diag}(\mathbf{a})$ denotes the diagonal matrix $\mathbf{A}$ that contains on its diagonal the elements of $\mathbf{a}$. $\mathbf{I}_n$ is an $n \times n$ identity matrix. We shall use $|\cdot|$ to denote the magnitude of a complex scalar. We denote by $\mathrm{mod}(x, 2)$ the modulus after division, i.e., the remainder of division of x by 2. We denote by $\{h[k]\}_k$ a sequence of random variables indexed by the integer values k. A circularly symmetric complex Gaussian random variable Z is a random variable $Z = X + jY \sim \mathcal{CN}(m, \sigma^2)$, where X and Y are i.i.d. $\mathcal{N}\left(m, \frac{\sigma^2}{2}\right)$. Throughout this paper we use complex baseband notation and all logarithms are taken to the base 2.

The paper is organized as follows. Section II reviews the spectral efficiency of three half-duplex relaying protocols, namely amplify-and-forward (AF), decode-and-forward (DF) and orthogonalize-and-forward (OF). In Section III we introduce the *two-way relaying* protocol and analyze the achievable sum-rates for AF, DF and OF relays. Section IV introduces the *two-path relaying* protocol and derives the achievable rates for AF and DF. Numerical examples are given in Section V.

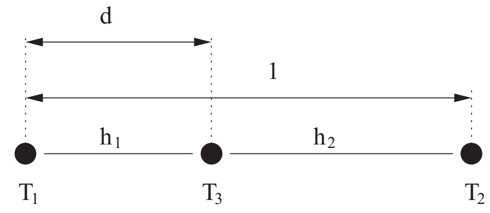

Fig. 1. One-dimensional linear relay network

II. Spectral Efficiency of Half-duplex Relaying Protocols

In this section we review three protocols used in wireless relaying. We consider the case where one source terminal communicates with one destination terminal with the help of one relay terminal. We look at the spectral efficiency that is achievable when the relay uses an *amplify-and-forward* (AF) strategy or a *decode-and-forward* (DF) strategy. We assume that there is no direct connection between the source and the destination (for example due to shadowing or too large separation) [1], [10] and that all terminals operate in half-duplex fashion [11], [12]. It is well known that this leads to an unavoidable loss in spectral efficiency due to the pre-log factor one-half. We also consider a network of terminals where a number of terminals communicates with a number of partner terminals with the help of a certain number of relays. In that case we look at the spectral efficiency of a scheme that was introduced in [13] and which we call here *orthogonalize-and-forward* (OF). Again, the assumptions that there is no direct connection between the terminals and the half-duplex constraint lead to a factor one-half loss in spectral efficiency.

A. Amplify-and-forward

A simple and popular relaying strategy is given by the amplify-and-forward scheme. The source terminal $\mathtt{T}_1$ transmits in the first time slot an information symbol to the relay terminal $\mathtt{T}_3$, see Fig.1. The relay amplifies the received symbol (including noise) according to its available average transmit power and forwards a scaled signal in the second time slot to the destination terminal $\mathtt{T}_2$. The relay receives in time slot k

$$y_3[k] = h_1[k]x_1[k] + n_3[k] \tag{1}$$

where h_1 is the complex channel gain between source and relay (first hop), $x_1 \sim \mathcal{CN}(0, P_1)$ the transmit symbol of the source, and $n_3 \sim \mathcal{CN}(0, \sigma_3^2)$ the additive white Gaussian noise at the relay. The relay scales $y_3[k]$ by[2]

$$g[k] = \sqrt{\frac{P_3}{P_1|h_1[k]|^2 + \sigma_3^2}} \tag{2}$$

where P_3 is the average transmit power of the relay [15]. The destination receives in time slot $k+1$

$$y_2[k{+}1] = h_2[k{+}1]g[k]h_1[k]x_1[k] + h_2[k{+}1]g[k]n_3[k] + n_2[k{+}1] \tag{3}$$

where h_2 is the complex channel gain between relay and destination (second hop), and $n_2 \sim \mathcal{CN}(0, \sigma_2^2)$ is the additive

[2]Depending on the amount of channel knowledge at the relay, different choices for the relay gain are possible [14].

white Gaussian noise at the destination. The information rate of this scheme for i.i.d. fading channels $\{h_1[k]\}_k$ and $\{h_2[k]\}_k$ is given by

$$R_{\mathrm{AF}} = \frac{1}{2}\,\mathcal{E}\left\{\log\left(1+\frac{P_1|h_2 g h_1|^2}{\sigma_2^2+\sigma_3^2|h_2 g|^2}\right)\right\} \tag{4}$$

where the expectation is the with respect to the channels h_1 and h_2[3]. The pre-log factor one-half follows because the relay operates in half-duplex mode and two channel uses are needed to transmit the information from the source to the destination.

B. Decode-and-forward

Another relaying strategy is given by the decode-and-forward scheme. The relay decodes the message sent by the source, re-encodes it by using the same or a different codebook and forwards the message to the destination. The relay receives in time slot k

$$y_3[k] = h_1[k]x_1[k] + n_3[k]. \tag{5}$$

After decoding and retransmission the destination receives in time slot $k+1$

$$y_2[k+1] = h_2[k+1]x_3[k+1] + n_2[k+1] \tag{6}$$

where $x_3 \sim \mathcal{CN}(0, P_3)$ is the transmit symbol of the relay[4]. The information rate of this scheme for i.i.d. fading channels $\{h_1[k]\}_k$ and $\{h_2[k]\}_k$ is given by

$$R_{\mathrm{DF}} = \frac{1}{2}\min\left\{\mathcal{E}\left\{\log\left(1+\frac{P_1|h_1|^2}{\sigma_3^2}\right)\right\},\right. \\ \left.\mathcal{E}\left\{\log\left(1+\frac{P_3|h_2|^2}{\sigma_2^2}\right)\right\}\right\}. \tag{7}$$

Note that since the direct connection between source and destination is not available it is not possible to use a decode-and-forward scheme based on superposition coding as described in [3], [12], where the signals from the relay and the source coherently add up at the destination. Furthermore, the rate given in (7) is exactly the ergodic capacity of the half-duplex relay channel with no direct connection, which can be easily seen by applying the cut-set upper bound [3] and by inspecting that (7) is equal to this upper bound.

C. Orthogonalize-and-forward

Consider now a network with $2N + K$ terminals. The network is divided into three sets: N source terminals in $\mathcal{T}_1$ want to transmit messages to N destination terminals in $\mathcal{T}_2$ with the help of K relay terminals in $\mathcal{T}_3$, see Fig.2. The random variables of the channel are:

- $\mathbf{x}_1$: $N \times 1$ dimensional vector that contains the transmit symbols of all source terminals in $\mathcal{T}_1$. The elements of $\mathbf{x}_1$ are i.i.d. $\mathcal{CN}(0, P_1)$,
- $\mathbf{y}_2$: $N \times 1$ dimensional vector of received symbols at the destination terminals in $\mathcal{T}_2$,
- $\mathbf{y}_3$: $K \times 1$ dimensional vector of received symbols at the relay terminals in $\mathcal{T}_3$,

[3]Note that g is a function of h_1.

[4]Note that due to simplicity we used in (5) and (6) symbolwise notation but in practice coding and decoding is done blockwise.

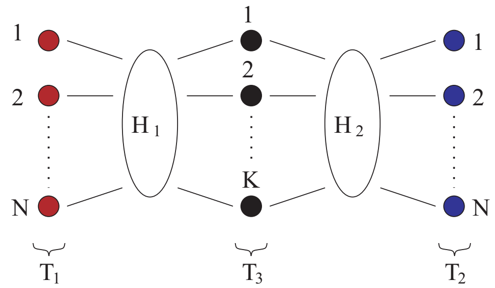

Fig. 2. Muli-user relaying with distributed zero-forcing relays (orthogonalize-and-forward)

- $\mathbf{H}_1$: $K \times N$ dimensional channel matrix between terminals in $\mathcal{T}_1$ and $\mathcal{T}_3$,
- $\mathbf{H}_2$: $N \times K$ dimensional channel matrix between terminals in $\mathcal{T}_3$ and $\mathcal{T}_2$,
- $\mathbf{n}_2$: $N \times 1$ dimensional vector of i.i.d. additive white Gaussian noise terms at the destination terminals in $\mathcal{T}_2$, with zero-mean and variance σ_2^2,
- $\mathbf{n}_3$: $K \times 1$ dimensional vector of i.i.d. additive white Gaussian noise terms at the relay terminals in $\mathcal{T}_3$, with zero-mean and variance σ_3^2.

All random variables are independent of each other. The relay terminals in $\mathcal{T}_3$ receive in time slot k

$$\mathbf{y}_3[k] = \mathbf{H}_1[k]\mathbf{x}_1[k] + \mathbf{n}_3[k]. \tag{8}$$

Each relay $\mathtt{T}_{3i} \in \mathcal{T}_3$, $i = 1, 2, \ldots, K$ scales its observation (which is a superposition of the transmit symbols of all sources $\mathtt{T}_{1i} \in \mathcal{T}_1$, $i = 1, 2, \ldots, N$) by a coefficient $g_i[k]$ such that an average sum power constraint among the relays is fulfilled and such that the overall channel between the sources and the destinations becomes diagonal. In order to achieve this each relay has to know all channel coefficients in the network, i.e., each relay has complete knowledge of $\mathbf{H}_1$ and $\mathbf{H}_2$. The destination terminals in $\mathcal{T}_2$ receive in time slot $k+1$

$$\mathbf{y}_2[k+1] = \mathbf{H}_2[k+1]\mathbf{G}[k]\mathbf{H}_1[k]\mathbf{x}_1[k] \\ +\mathbf{H}_2[k+1]\mathbf{G}[k]\mathbf{n}_3[k] + \mathbf{n}_2[k+1]. \tag{9}$$

We choose the diagonal gain matrix $\mathbf{G}[k] = \mathrm{diag}\left((g_1[k], g_2[k], \ldots, g_K[k])^{\mathrm{T}}\right)$ such that the transmissions between terminals in $\mathcal{T}_1$ and $\mathcal{T}_2$ become interference-free. For this purpose we define the interference matrix $\mathbf{H}_{\mathrm{int}}$ with dimensions $K \times N(N-1)$ where the columns of $\mathbf{H}_{\mathrm{int}}$ are defined by[5]

$$\mathbf{h}_p^{(1)} \odot \mathbf{h}_q^{(2)} \tag{10}$$

for all $p, q \in \{1, 2, \ldots, N\}$ and $p \neq q$, where $\mathbf{h}_p^{(1)}$ is the vector containing the channel gains from the pth node in $\mathcal{T}_1$ to all relays in $\mathcal{T}_3$ (i.e., the pth column of $\mathbf{H}_1$) and $\mathbf{h}_q^{(2)}$ is the vector containing the channel gains from the qth node in $\mathcal{T}_2$ to all relays in $\mathcal{T}_3$ (i.e., the qth column of $\mathbf{H}_2^{\mathrm{T}}$). The two-hop matrix channel $\mathbf{H}_2\mathbf{G}\mathbf{H}_1$ becomes diagonal if $\mathbf{g} = \mathrm{diag}(\mathbf{G}) = (g_1, g_2, \ldots, g_K)^{\mathrm{T}}$ lies in the null space of the interference matrix $\mathbf{H}_{\mathrm{int}}$ (*zero-forcing*). Let $r = \mathrm{rk}(\mathbf{H}_{\mathrm{int}}) =$

[5]For ease of notation we drop the time index k

$\min\{K, N(N-1)\}$ be the rank of the matrix $\mathbf{H}_{\text{int}}$ and define the singular value decomposition

$$\mathbf{H}_{\text{int}} = \mathbf{U}\mathbf{D}\left[\mathbf{V}^{(r)} \;\; \mathbf{V}^{(0)}\right]^{\text{H}} \tag{11}$$

where $\mathbf{V}^{(r)}$ contains the first r right singular vectors of $\mathbf{H}_{\text{int}}$ and $\mathbf{V}^{(0)}$ the last $K-r$ right singular vectors. The columns of $\mathbf{V}^{(0)}$ form an orthonormal basis for the null space of $\mathbf{H}_{\text{int}}$, i.e., $\mathbf{V}^{(0)} = \text{null}(\mathbf{H}_{\text{int}})$. A sufficient condition for the null space to be non-empty is

$$K \geq N(N-1)+1 \tag{12}$$

and we refer to it as *minimum relay configuration*. The orthogonalize-and-forward gain vector $\mathbf{g}$ is obtained by projecting any gain vector onto the null space of $\mathbf{H}_{\text{int}}$. In [13] it was shown that for $K \to \infty$ the choice $\mathbf{g}_\infty = \text{diag}(\mathbf{H}_2\mathbf{H}_1)$ diagonalizes the two-hop matrix channel $\mathbf{H}_2\mathbf{G}\mathbf{H}_1$. For finite number of relays $K < \infty$ we choose

$$\mathbf{g} = c\mathbf{Z}\mathbf{Z}^{\text{H}}\mathbf{g}_\infty \tag{13}$$

where $\mathbf{Z} = \mathbf{V}^{(0)}$ and c is chosen such that the average sum-power constraint of the relays is met. In order to compute $\mathbf{g}$ each relay has to know all channel gains of the network. For the scheme to work the channels have to be constant at least during two time slots, since the gain vector is chosen such, that it lies in the nullspace of the interference matrix, which depends on the channels in the first time slot and on the channels in the second time slot. There are different ways how to realize that. For example, each relay learns its first-hop channels and its second-hop channels and then broadcasts this information to the other relays. After that each relay is able to calculate the zero-forcing gain vector which is valid until the next channel update. Another possibility would be to use dedicated relays that are for example connected to a wired backbone and where the channel estimation and signal processing is done globally at a central unit. In [16] it is shown how this channel information may be distributed when the relays have access to a powerline distribution system. However, if the channel updates are communicated through the wireless channel, some loss in spectral efficiency has to be accepted [17].

Using $\mathbf{g}$ the product channel $\mathbf{H}_{12} = \mathbf{H}_2\mathbf{G}\mathbf{H}_1$ with $\mathbf{G} = \text{diag}(\mathbf{g})$ becomes diagonal and the transmissions between terminals in $\mathcal{T}_1$ and $\mathcal{T}_2$ interference-free. However, the noise terms at the destination terminals in $\mathcal{T}_2$ are still correlated. For separate (noncooperating) decoding at the terminals in $\mathcal{T}_2$ the network's sum-rate follows as

$$R_{\text{sum}}^{\text{OF}} = \frac{1}{2}\,\mathcal{E}\left\{\log\det\left(\mathbf{I}_N + P_1\mathbf{D}_2^{-1}\mathbf{H}_{12}\mathbf{H}_{12}^{\text{H}}\right)\right\} \tag{14}$$

where the diagonal matrix $\mathbf{D}_2$ contains the diagonal elements of the noise covariance matrix

$$\mathbf{R}_2 = \mathcal{E}\left\{\widetilde{\mathbf{n}}_2\widetilde{\mathbf{n}}_2^{\text{H}}\right\} = \sigma_3^2\mathbf{H}_2\mathbf{G}\mathbf{G}^{\text{H}}\mathbf{H}_2^{\text{H}} + \sigma_2^2\mathbf{I}_N \tag{15}$$

and where $\widetilde{\mathbf{n}}_2 = \mathbf{H}_2\mathbf{G}\mathbf{n}_3 + \mathbf{n}_2$ is the overall noise at the terminals in $\mathcal{T}_2$. The factor one-half in (14) is again due to the use of two time slots.

III. Protocol I: Two-way Relaying

The relaying protocols discussed in the previous section suffer from a loss in spectral efficiency due to the half-duplex constraint of the terminals. In order to increase the spectral efficiency of such a relay network we propose a bidirectional (two-way) communication between two terminals whereas the relay assists in the two-way communication. An example for such a scenario may be the communication between two wireless routers that communicate with each other through the help of a relay terminal due to the lack of a direct connection[6]. The two-way communication problem was first studied by Shannon in [18], which is considered as the first study of a network information theory problem. In [19] the achievable rate regions for the general full-duplex two-way relay channel (including direct link) is studied. In this section we show how the half-duplex relaying strategies discussed in the previous section can be extended to the two-way case and analyze the achievable sum-rates.

A. Amplify-and-forward

Terminal $\mathtt{T}_1$ wants to transmit a message to terminal $\mathtt{T}_2$ and vice versa. Since there is no direct connection between the terminals all traffic goes trough relay terminal $\mathtt{T}_3$. The proposed relaying scheme works as follows: in time slot k both terminals $\mathtt{T}_1$ and $\mathtt{T}_2$ transmit their symbols to relay $\mathtt{T}_3$. The relay receives in time slot k[7]

$$y_3[k] = h_1[k]x_1[k] + h_2[k]x_2[k] + n_3[k] \tag{16}$$

where $x_2 \sim \mathcal{CN}(0, P_2)$ is the transmit symbol of terminal $\mathtt{T}_2$. The relay scales the received signal by

$$g[k] = \sqrt{\frac{P_3}{P_1\left|h_1[k]\right|^2 + P_2\left|h_2[k]\right|^2 + \sigma_3^2}} \tag{17}$$

in order to meet its average transmit power constraint. It then broadcasts the signal in the next time slot to both destinations. The input-output relation for the $\mathtt{T}_1 \to \mathtt{T}_3 \to \mathtt{T}_2$ communication direction is as

$$\begin{aligned} y_2[k+1] = {} & h_2[k+1]g[k]h_1[k]x_1[k] + h_2[k+1]g[k]h_2[k]x_2[k] \\ & + h_2[k+1]g[k]n_3[k] + n_2[k+1] \end{aligned} \tag{18}$$

and for the $\mathtt{T}_1 \leftarrow \mathtt{T}_3 \leftarrow \mathtt{T}_2$ direction

$$\begin{aligned} y_1[k+1] = {} & h_1[k+1]g[k]h_2[k]x_2[k] + h_1[k+1]g[k]h_1[k]x_1[k] \\ & + h_1[k+1]g[k]n_3[k] + n_1[k+1]. \end{aligned} \tag{19}$$

Since nodes $\mathtt{T}_1$ and $\mathtt{T}_2$ know their own transmitted symbols they can subtract the *back-propagating self-interference* in (18) and (19) prior to decoding, assuming perfect knowledge of the corresponding channel coefficients. The sum-rate of this protocol is then given by

$$\begin{aligned} R_{\text{sum}}^{\text{AF}} = {} & \frac{1}{2}\,\mathcal{E}\left\{\log\left(1 + \frac{P_1|h_2gh_1|^2}{\sigma_2^2 + \sigma_3^2|h_2g|^2}\right)\right\} \\ & + \frac{1}{2}\,\mathcal{E}\left\{\log\left(1 + \frac{P_2|h_2gh_1|^2}{\sigma_1^2 + \sigma_3^2|h_1g|^2}\right)\right\}. \end{aligned} \tag{20}$$

[6]Note that even if a direct channel of sufficient quality would be available it couldn't be used in a two-way protocol with half-duplex terminals, since both terminals transmit simultaneously and also receive simultaneously.

[7]We assume channel reciprocity for h_1 and h_2.

The transmission in each direction suffers still from the pre-log factor one-half. However, the half-duplex constraint can here be exploited to establish a bidirectional connection between two terminals and to increase the sum-rate of the network.

B. Decode-and-forward

We consider now a two-way communication between terminals $\mathtt{T}_1$ and $\mathtt{T}_2$ via a half-duplex DF relay $\mathtt{T}_3$. In time slot k both terminals $\mathtt{T}_1$ and $\mathtt{T}_2$ transmit their symbols to relay $\mathtt{T}_3$. The relay receives in time slot k (multiple access phase)

$$y_3[k] = h_1[k]x_1[k] + h_2[k]x_2[k] + n_3[k], \tag{21}$$

decodes the symbols $x_1[k]$ and $x_2[k]$ and transmits $x_3[k{+}1] = \sqrt{\beta}x_1[k] + \sqrt{1-\beta}x_2[k]$ in the next time slot (broadcast phase). The input-output relation for the $\mathtt{T}_1 \rightarrow \mathtt{T}_3 \rightarrow \mathtt{T}_2$ communication direction is then

$$y_2[k] = h_2[k{+}1]x_3[k{+}1] + n_2[k] \tag{22}$$

and for the $\mathtt{T}_1 \leftarrow \mathtt{T}_3 \leftarrow \mathtt{T}_2$ direction

$$y_1[k] = h_1[k{+}1]x_3[k{+}1] + n_1[k]. \tag{23}$$

The relay uses an average transmit power of βP_3 for the forward direction and $(1{-}\beta)P_3$ for the backward direction. Since terminal $\mathtt{T}_1$ knows $x_1[k]$ and terminal $\mathtt{T}_2$ knows $x_2[k]$ these symbols (back-propagating self-interference) can be subtracted at the respective terminals prior to decoding of the symbol transmitted by the partner terminal. Again we assume channel reciprocity and that the relay can decode x_1 and x_2 without errors. The sum-rate of this protocol is then given by

$$R_{\text{sum}}^{\text{DF}} = \max_{\beta} \min\Big(R_{\text{MA}}, R_1(\beta) + R_2(1{-}\beta)\Big) \tag{24}$$

where

$$R_{\text{MA}} = \frac{1}{2}C\left(P_1|h_1|^2 + P_2|h_2|^2\right) \tag{25}$$

$$R_1(\beta) = \frac{1}{2}\min\Big(C\left(P_1|h_1|^2\right), C\left(\beta P_3|h_2|^2\right)\Big) \tag{26}$$

$$R_2(1{-}\beta) = \frac{1}{2}\min\Big(C\left(P_2|h_2|^2\right), C\left((1{-}\beta)P_3|h_1|^2\right)\Big) \tag{27}$$

where $C(x) = \mathcal{E}\{\log(1+x)\}$. If the relay does not have any knowledge of the channels[8] in the broadcast phase 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, when the channels are changing slowly) 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 Fig.1 and assume that the relay has only knowledge of the second order statistics (pathloss coefficients). Further we choose $P_1 = P_2 = P_3 = P$, i.e., all nodes transmit with equal power. The power allocation that maximizes the ergodic sum rate (24) is then given by [20]

$$\beta^* = \begin{cases} \min\left(\frac{d^\alpha-(1-d)^\alpha}{2P} + \frac{1}{2}, \frac{(1-d)^\alpha}{d^\alpha}\right), & d \geq \frac{1}{2}; \\ \max\left(\frac{d^\alpha-(1-d)^\alpha}{2P} + \frac{1}{2}, 1 - \frac{d^\alpha}{(1-d)^\alpha}\right), & d < \frac{1}{2} \end{cases} \tag{28}$$

where α is the pathloss exponent, d the normalized distance between terminal $\mathtt{T}_1$ and relay $\mathtt{T}_3$ and $1-d$ the normalized distance between relay $\mathtt{T}_3$ and terminal $\mathtt{T}_2$, see Fig.1. As the relay moves towards terminal $\mathtt{T}_1$, i.e., $d \rightarrow 0$, more relay power is spent for the forward $\mathtt{T}_1 \rightarrow \mathtt{T}_3 \rightarrow \mathtt{T}_2$ transmission and less power for the backward $\mathtt{T}_1 \leftarrow \mathtt{T}_3 \leftarrow \mathtt{T}_2$ transmission. The reason is that the link capacity from terminal $\mathtt{T}_2$ to relay $\mathtt{T}_3$ is small and dominates the overall capacity for the $\mathtt{T}_1 \leftarrow \mathtt{T}_3 \leftarrow \mathtt{T}_2$ 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. Note that (24)–(27) imply Gaussian codebooks and error-free decoding at all terminals. In order to study the effects of decoding errors and error propagation one would have to consider specific symbol alphabets (instead of Gaussian codebooks) an specific coding schemes. However, this is beyond the scope of the current paper.

C. Orthogonalize-and-forward

We apply the two-way communication scheme to the network with N two-hop communication links and K relays. Like in Section II we divide the network into three sets of terminals: terminals in $\mathcal{T}_1$ want to transmit messages to terminals in $\mathcal{T}_2$ and vice versa and the set $\mathcal{T}_3$ contains the relay terminals. The input-output relation for the $\mathcal{T}_1 \rightarrow \mathcal{T}_3 \rightarrow \mathcal{T}_2$ communication is

$$\begin{aligned} \mathbf{y}_2[k{+}1] = {} & \mathbf{H}_2[k{+}1]\mathbf{G}[k]\mathbf{H}_1[k]\mathbf{x}_1[k] \\ & + \mathbf{H}_2[k{+}1]\mathbf{G}[k]\mathbf{H}_2^{\mathrm{T}}[k]\mathbf{x}_2[k] \\ & + \mathbf{H}_2[k{+}1]\mathbf{G}[k]\mathbf{n}_3[k] + \mathbf{n}_2[k{+}1] \end{aligned} \tag{29}$$

and for the $\mathcal{T}_1 \leftarrow \mathcal{T}_3 \leftarrow \mathcal{T}_2$ communication is

$$\begin{aligned} \mathbf{y}_1[k{+}1] = {} & \mathbf{H}_1^{\mathrm{T}}[k{+}1]\mathbf{G}[k]\mathbf{H}_2^{\mathrm{T}}[k]\mathbf{x}_2[k] \\ & + \mathbf{H}_1^{\mathrm{T}}[k{+}1]\mathbf{G}[k]\mathbf{H}_1[k]\mathbf{x}_1[k] \\ & + \mathbf{H}_1^{\mathrm{T}}[k{+}1]\mathbf{G}[k]\mathbf{n}_3[k] + \mathbf{n}_1[k{+}1]. \end{aligned} \tag{30}$$

The $N \times 1$ vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ comprise the transmitted symbols from terminals in $\mathcal{T}_1$ and $\mathcal{T}_2$, respectively. Each element (symbol) of $\mathbf{x}_1$ and $\mathbf{x}_2$ is taken from a Gaussian codebook with average transmit power P_1 and P_2, respectively. The $N \times 1$ vectors $\mathbf{y}_1$ and $\mathbf{y}_2$ comprise the received symbols of terminals in $\mathcal{T}_1$ and $\mathcal{T}_2$ in the second time slot. In the following we discuss how to choose the diagonal gain matrix $\mathbf{G}$ such that the transmissions between terminals in $\mathcal{T}_1$ and $\mathcal{T}_2$ become interference-free in both communication directions. Comparing (29) with (9), we see that a receiving terminal in $\mathcal{T}_2$ suffers additionally from back-propagating interference caused by its neighbor terminals located in the same set. Each terminal knows only its own symbol (which was transmitted by this terminal in time slot k) and can subtract this contribution from the received signal in time slot $k+1$. The symbols transmitted by the neighbor terminals are unknown and cannot

[8]For example, when the channels change i.i.d. from time slot to time slot, the channel knowledge learned during the multiple access phase may not be used for the broadcast phase

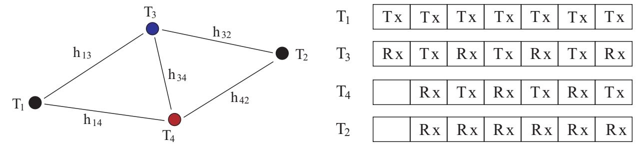

Fig. 3. Two-path relaying with alternating relays

be subtracted (since the terminals in each set do not cooperate and therefore have no knowledge about the other's transmitted symbols). The gain matrix $\mathbf{G}$ has therefore to be chosen such that the channels $\mathbf{H}_2\mathbf{G}\mathbf{H}_2^{\mathrm{T}}$ and $\mathbf{H}_1^{\mathrm{T}}\mathbf{G}\mathbf{H}_1$ become diagonal too. For this purpose the interference matrix has to be extended to a $K \times 2N(N-1)$ dimensional matrix

$$\mathbf{H}_{\mathrm{int,bi}} = [\mathbf{H}_{\mathrm{int}},\, \mathbf{H}_{\mathrm{int},1},\, \mathbf{H}_{\mathrm{int},2}] \tag{31}$$

where (32) and (33) below are both $K \times \frac{N(N-1)}{2}$ dimensional matrices. In order to have a non-empty nullspace the minimum relay configuration becomes

$$K \geq 2N(N-1)+1. \tag{34}$$

The network's two-way sum-rate follows as

$$\begin{aligned} R_{\mathrm{sum}}^{\mathrm{OF}} = &\frac{1}{2}\mathcal{E}\left\{\log\det\left(\mathbf{I}_N + P_1\mathbf{D}_2^{-1}\mathbf{H}_{12}\mathbf{H}_{12}^{\mathrm{H}}\right)\right\} + \\ &\frac{1}{2}\mathcal{E}\left\{\log\det\left(\mathbf{I}_N + P_2\mathbf{D}_1^{-1}\mathbf{H}_{21}\mathbf{H}_{21}^{\mathrm{H}}\right)\right\} \end{aligned} \tag{35}$$

where $\mathbf{H}_{12} = \mathbf{H}_2\mathbf{G}\mathbf{H}_1$ is the equivalent (diagonal) channel from "left to right" and $\mathbf{H}_{21} = \mathbf{H}_1^{\mathrm{T}}\mathbf{G}\mathbf{H}_2^{\mathrm{T}}$ the equivalent (diagonal) channel from "right to left". The orthogonalize-and-forward gain matrix $\mathbf{G}$ contains the elements of $\mathbf{g}$, which is the projection of $\mathbf{g}_\infty = \mathrm{diag}(\mathbf{H}_2\mathbf{H}_1)$ on the nullspace of $\mathbf{H}_{\mathrm{int,bi}}$. The diagonal matrices $\mathbf{D}_1$ and $\mathbf{D}_2$ contain the diagonal elements of the noise covariance matrices

$$\mathbf{R}_1 = \mathcal{E}\left\{\widetilde{\mathbf{n}}_1\widetilde{\mathbf{n}}_1^{\mathrm{H}}\right\} = \sigma_3^2\mathbf{H}_1^{\mathrm{T}}\mathbf{G}\mathbf{G}^{\mathrm{H}}\mathbf{H}_1^* + \sigma_1^2\mathbf{I}_N, \tag{36}$$

where $\widetilde{\mathbf{n}}_1 = \mathbf{H}_1^{\mathrm{T}}\mathbf{G}\mathbf{n}_3 + \mathbf{n}_1$ is the overall noise at the terminals in $\mathcal{T}_1$ and

$$\mathbf{R}_2 = \mathcal{E}\left\{\widetilde{\mathbf{n}}_2\widetilde{\mathbf{n}}_2^{\mathrm{H}}\right\} = \sigma_3^3\mathbf{H}_2\mathbf{G}\mathbf{G}^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}} + \sigma_2^2\mathbf{I}_N \tag{37}$$

respectively. Note that the two-way OF scheme reduces to the two-way AF scheme introduced in Section III-A when the number of source-destination pairs is one, i.e., $N=1$.

IV. Protocol II: Two–path Relaying

The protocol discussed in the previous section required bidirectional traffic between $\mathtt{T}_1$ and $\mathtt{T}_2$ in order to circumvent the half-duplex loss in spectral efficiency. For the second protocol we assume a unidirectional traffic model but propose to use two half-duplex relays that assist in the communication. Again we assume that no direct connection between terminals $\mathtt{T}_1$ and $\mathtt{T}_2$ is available. Transmission of messages from a source $\mathtt{T}_1$ to a destination $\mathtt{T}_2$ is done via two relays $\mathtt{T}_3$ and $\mathtt{T}_4$, which do not receive and transmit simultaneously. A message is transmitted in two time slots. In the first slot the source transmits the message to relay $\mathtt{T}_3$ or $\mathtt{T}_4$ and in the second time slot the message is forwarded to the destination, see Fig.3. The length of one time 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 $\mathtt{T}_3$ receives whereas $\mathtt{T}_4$ transmits. (Except for $k=1$, where $\mathtt{T}_4$ does not transmit). In even time slots, $k = 2, 4, 6, \ldots$, it is the other way around. This cooperation protocol avoids the pre-log factor one-half since the source transmits a new message in every time slot and has not to be *silent* in each second time slot. However, since the relays do not operate in orthogonal channels (as in [8]), there will be interference between $\mathtt{T}_3$ and $\mathtt{T}_4$ and it is not clear *a priori* whether this inter-relay interference cancels the 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 $\mathtt{T}_1$ 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 $x_1[k] = (x_1[k,1], x_1[k,2], \ldots, x_1[k,N])$ of length N where the symbols $\{x_1[k,n]\}_{k,n}$ are i.i.d. according to $\mathcal{CN}(0, P_1)$ with P_1 being the average transmit power of the source. The channel gain between node i and node j at the discrete time $[k,n] := (kN+n)T$ is denoted as $h_{ij}[k,n]$. 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, i.e., $\mathcal{E}\left\{|h_{13}|^2\right\} = \mathcal{E}\left\{|h_{14}|^2\right\} = \nu_1^2$, $\mathcal{E}\left\{|h_{32}|^2\right\} = \mathcal{E}\left\{|h_{42}|^2\right\} = \nu_2^2$ and $\mathcal{E}\left\{|h_{34}|^2\right\} = \nu_{34}^2$. We assume that the $\{h_{ij}[k,n]\}_{k,n}$ are independent, stationary and ergodic fading processes. The source $\mathtt{T}_1$ is aware of the *fading distribution* of the channel gains in the network but not of the *fading realizations*. The destination knows the *fading realizations* of all channel gains in the network. For the relay the assumptions about channel knowledge vary with the protocols (AF, DF).

A. Amplify-and-forward

The receive signal of relay $\mathtt{T}_p$ at time instant $(kN+n)T$ with $p = 4 - \mathrm{mod}\,(k,2) \in \{3,4\}$ and $q = 4 - \mathrm{mod}\,(k-i,2) \in \{3,4\}$ is given as

$$\begin{aligned} y_p[k,n] = &h_{1p}[k,n]x_1[k,n] + n_p[k,n] \\ &+ \sum_{i=1}^{k-1}\Big(h_{1q}[k-i,n]x_1[k-i,n] \\ &+ n_q[k-i,n]\Big)f_i[k,n] \end{aligned} \tag{38}$$

where

$$f_i[k,n] := \prod_{j=1}^{i} h_{34}[k-j,n]g[k-j] \tag{39}$$

denotes the *inter-relay interference factor*. The relay noise samples $\{n_i[k,n]\}_{k,n}$, $i \in \{3,4\}$ are i.i.d. according to $\mathcal{CN}(0, \sigma_{\mathrm{R}}^2)$ and the destination noise samples $\{n_2[k,n]\}_{k,n}$ are i.i.d. according to $\mathcal{CN}(0, \sigma_{\mathrm{D}}^2)$. The transmit signal of relay $\mathtt{T}_p$ is a scaled version of its received signal: $t_p[k+1,n] = g[k]y_p[k,n]$, where $g[k]$ is the relay scaling coefficient and for $k = 2, 3, 4, \ldots$ chosen as

$$g^2[k] = \frac{P_{\mathrm{R}}}{\frac{1}{N}\sum_{n=1}^{N}|y_p[k,n]|^2} \approx \frac{P_{\mathrm{R}}}{P_1\nu_1^2 + P_{\mathrm{R}}\nu_{34}^2 + \sigma_{\mathrm{R}}^2} := g^2 \tag{40}$$

$$\mathbf{H}_{\mathrm{int},1} = \left[\mathbf{h}_1^{(1)} \odot \mathbf{h}_2^{(1)}, \ldots, \mathbf{h}_1^{(1)} \odot \mathbf{h}_N^{(1)}, \mathbf{h}_2^{(1)} \odot \mathbf{h}_3^{(1)}, \ldots, \mathbf{h}_2^{(1)} \odot \mathbf{h}_N^{(1)}, \ldots, \mathbf{h}_{N-1}^{(1)} \odot \mathbf{h}_N^{(1)}\right] \quad (32)$$

$$\mathbf{H}_{\mathrm{int},2} = \left[\mathbf{h}_1^{(2)} \odot \mathbf{h}_2^{(2)}, \ldots, \mathbf{h}_1^{(2)} \odot \mathbf{h}_N^{(2)}, \mathbf{h}_2^{(2)} \odot \mathbf{h}_3^{(2)}, \ldots, \mathbf{h}_2^{(2)} \odot \mathbf{h}_N^{(2)}, \ldots, \mathbf{h}_{N-1}^{(2)} \odot \mathbf{h}_N^{(2)}\right] \quad (33)$$

---

and in the first time slot $k=1$

$$g^2[1] = \frac{P_{\mathrm{R}}}{\frac{1}{N}\sum_{n=1}^{N}|y_3[1,n]|^2} \approx \frac{P_{\mathrm{R}}}{P_1\nu_1^2 + \sigma_{\mathrm{R}}^2} \geq g^2 \quad (41)$$

where $P_3 = P_4 = P_{\mathrm{R}}$ is the average transmit power of each relay. The approximations in (40) and (41) are exact for $N \to \infty$ by the law of large numbers. Destination $\mathtt{T}_2$ observes at $((k+1)N+n)T$ the signal

$$y_2[k{+}1,n] = h_{p2}[k{+}1,n]gy_p[k,n] + n_2[k{+}1,n] \quad (42)$$

where $y_p[k,n]$ is given in (38). To decode $x_1[k]$ from (42) the destination receiver first subtracts the previously decoded codewords $x_1[k{-}1],\ldots,x_1[1]$ from the received frame $y_2[k{+}1]$, 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_{34} and g, which together act as forgetting factor for the decoding process, see (38) and (39). After perfect cancelation of m previously decoded codewords $x_1[k{-}1], x_1[k{-}2],\ldots,x_1[k{-}m]$ the destination signal is given by

$$\begin{aligned} y_2^{(m)}[k{+}1,n] = h_{p2}[k{+}1,n]g\Bigg(& h_{1p}[k,n]x_1[k,n] \\ & + \sum_{i=m+1}^{k-1} h_{1q}[k{-}i,n]x_1[k{-}i,n]f_i[k,n] \\ & + \sum_{i=0}^{k-1} n_q[k{-}i,n]f_i[k,n]\Bigg) + n_2[k{+}1,n] \end{aligned} \quad (43)$$

where $f_i[k,n] := 1$ 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 $x_1[k{-}1]$ appear as inter-frame interference when $x_1[k]$ is decoded. For $0<m<k-1$ only the last m transmitted codewords are canceled and $x_1[1], x_1[2],\ldots,x_1[k{-}m{-}1]$ remain as interference terms (*partial interference cancelation*). The ergodic rate in time slot $k+1$ measured in b/s/Hz follows as

$$R[k{+}1] = \mathcal{E}\left\{\log\left(1 + \frac{P_1|h_{p2}gh_{1p}|^2}{\sigma_{\mathrm{D}}^2 + |h_{p2}g|^2\big(P_1 I_1[k] + \sigma_{\mathrm{R}}^2 I_2[k]\big)}\right)\right\} \quad (44)$$

with $I_1[k]$ denoting the accumulated inter-frame interference given by

$$I_1[k] = \sum_{i=m+1}^{k-1} |h_{1q}|^2 |f_i[k]|^2 \quad (45)$$

and $I_2[k]$ the accumulated noise interference given by

$$I_2[k] = \sum_{i=0}^{k-1} |f_i[k]|^2. \quad (46)$$

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[9]. 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_{1p}, h_{p2} for $p \in \{3,4\}$ and h_{34} and depends on the channel model that is used for the fading variables. After transmission of a sequence of K messages we get the average rate

$$\begin{aligned} \overline{R}_K &= \frac{1}{K+1}\sum_{k=1}^{K} R[k{+}1] && (47) \\ &\geq \frac{K}{K+1} R[K{+}1] && (48) \\ &\geq \frac{K}{K+1} \lim_{k\to\infty} R[k] && (49) \end{aligned}$$

where the pre-log is $\frac{K}{K+1} \approx 1$ for large K. The inequalities (48) and (49) are motivated by the observation that the average interference power between the relays is upper bounded by the average relay transmit power, i.e, the rate does not diminish as k grows large. In order to see this we look at a lower bound on (44) and demand that it is greater than zero:

$$\begin{aligned} R[k+1] &> \mathcal{E}\left\{\log\left(\frac{P_1|h_{p2}gh_{1p}|^2}{\sigma_{\mathrm{D}}^2 + |h_{p2}g|^2\big(P_1 I_1[k] + \sigma_{\mathrm{R}}^2 I_2[k]\big)}\right)\right\} \\ &> \mathcal{E}\left\{\log\left(P_1|h_{p2}gh_{1p}|^2\right)\right\} - \sigma_{\mathrm{D}}^2 \\ &\quad - \mathcal{E}\left\{|h_{p2}g|^2\big(P_1 I_1[k] + \sigma_{\mathrm{R}}^2 I_2[k]\big)\right\} \\ &\overset{!}{>} 0. \end{aligned} \quad (50)$$

From that it follows

$$\begin{aligned} \mathcal{E}\left\{\log\left(P_1|h_{p2}gh_{1p}|^2\right)\right\} &> \sigma_{\mathrm{D}}^2 + g^2\mathcal{E}\left\{|h_{p2}|^2\right\} \\ &\quad \cdot \mathcal{E}\left\{\big(P_1 I_1[k] + \sigma_{\mathrm{R}}^2 I_2[k]\big)\right\} \end{aligned} \quad (51)$$

$$\begin{aligned} \frac{2^{\mathcal{E}\left\{\log\left(P_1|h_{p2}gh_{1p}|^2\right)\right\}} - \sigma_{\mathrm{D}}^2}{g^2\nu_2^2} &> P_1\overline{I}_1[k] + \sigma_{\mathrm{R}}^2\overline{I}_2[k] \\ &= P_{\mathrm{R}}\nu_{34}^2 \end{aligned} \quad (52)$$

where

$$\overline{I}_1[k] = \sum_{i=m+1}^{k-1} \mathcal{E}\left\{|f_i[k]|^2\right\} \quad (53)$$

$$\overline{I}_2[k] = \sum_{i=0}^{k-1} \mathcal{E}\left\{|f_i[k]|^2\right\}. \quad (54)$$

We evaluate the expectation in (52) for Gaussian fading first hop and second hop channels with zero mean and unit variance

[9]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 $(kN+n)T$

(double Rayleigh fading channel), i.e., $\nu_1^2 = \nu_2^2 = 1$:

$$\frac{c \cdot P_1 \cdot g^2 - \sigma_\mathrm{D}^2}{g^2} \approx cP_1 > P_\mathrm{R}\nu_{34}^2 \tag{55}$$

where $c = 1.12$. We obtain the following condition for the relay transmit power

$$P_\mathrm{R} < \frac{1.12 \cdot P_1}{\nu_{34}^2}. \tag{56}$$

This means that as long as the relay transmit power fulfils (56) the lower bound is larger than zero and therefore also the rate (44). The disadvantage of signaling according to (44) is that the source has to adapt the rate for each frame. However, the lower bounds in (48) and (49) 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 (44) in a different way from the previous lower bound. For $k = 2, 3, \ldots, K$ it is as in (57-59) below with

$$I_{1,k'}[k] = \sum_{i=m+1}^{k'} |h_{1q}|^2 |f_i[k]|^2 \tag{60}$$

$$I_{2,k'}[k] = \sum_{i=0}^{k'} |f_i[k]|^2 \tag{61}$$

and

$$\overline{I}_1[k] = \sum_{i=m+1}^{k-1} \mathcal{E}\left\{|f_i[k]|^2\right\} = \frac{u^{m+1} - u^k}{1-u} \tag{62}$$

$$\overline{I}_2[k] = \sum_{i=0}^{k-1} \mathcal{E}\left\{|f_i[k]|^2\right\} = \frac{1-u^k}{1-u} \tag{63}$$

where $u = g^2\nu_{34}^2$. The first inequality (58) follows due to $I_1[k] \geq I_{1,k'}[k] = \sum_{i=m+1}^{k'} |h_{1q}|^2|f_i[k]|^2$ and $I_2[k] \geq I_{2,k'}[k] = \sum_{i=0}^{k'} |f_i[k]|^2$ for $k' < k-1$. The second inequality (59) follows by applying Jensen's inequality [21] on the second $\log-$ term. For $k \to \infty$ and $u < 1$ we get

$$\lim_{k\to\infty} \overline{I}_1[k] = \frac{u^{m+1}}{1-u} \tag{64}$$

and

$$\lim_{k\to\infty} \overline{I}_2[k] = \frac{1}{1-u} \tag{65}$$

and the lower bound (59) becomes independent of the actual frame number k. Note that for a stationary inter-relay channel h_{34} the statistics of $I_{1,k'}[k]$ and $I_{2,k'}[k]$ become independent of k (but dependent on k'). Numerical results in Section V 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 (44), 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 (59). Note that the purpose of the lower bound in (59) is not to give qualitative insights into the performance of the two-path AF protocol but to show that one could use a fixed-rate signaling scheme based on such a lower bound. The lower bound works fine when the relay power is not too large, but can be loose, if the relay power is too large[10]. For the numerical examples in Section V we have chosen the relay power to be equal to the source power which leads to a tight lower bound (59). Another possibility to obtain a rate for fixed-rate signaling would be to determine the ergodic rate (44) for a specific channel model and then to compute (58) or (59).

B. Decode-and-forward

At the discrete time $[k, n] := (kN + n)T$ the receive signal of relay $\mathtt{T}_p$ with $p = 4 - \mathrm{mod}\,(k, 2) \in \{1, 2\}$ is given as

$$y_p[k,n] = h_{1p}[k,n]x_1[k,n] + h_{34}[k,n]x_1[k-1,n] + n_p[k,n]. \tag{66}$$

Relay $\mathtt{T}_p$ may choose two different decoding strategies: a) the relay decodes $x_1[k,n]$ treating $h_{34}[k,n]x_1[k-1,n]$ as interference or b) the relay decodes first $x_1[k-1,n]$ treating $h_{1p}[k,n]x_1[k,n]$ as interference, subtracts $h_{34}[k,n]x_1[k-1,n]$ from the received signal $y_p[k,n]$, then decodes $x_1[k,n]$ interference-free. The achievable rate with strategy a) is given by

$$R_\mathrm{a} = \min\left(C\left(\frac{P_1|h_{1p}|^2}{\sigma_\mathrm{R}^2 + P_\mathrm{R}|h_{34}|^2}\right), C\left(\frac{P_\mathrm{R}|h_{p2}|^2}{\sigma_\mathrm{D}^2}\right)\right) \tag{67}$$

where again $C(x) = \mathcal{E}\{\log(1+x)\}$. Note that we assume the same statistics for h_{13} and h_{14} as well as for h_{32} and h_{42}, hence the rate in (67) is independent of p. The first term in (67) determines the maximum rate in the first hop (source $\mathtt{T}_1$ to relay $\mathtt{T}_p$) when the inter-relay signal is treated as interference and the second term denotes the maximum rate in the second hop (relay $\mathtt{T}_p$ to destination $\mathtt{T}_2$). The achievable rate with strategy b) follows as

$$R_\mathrm{b} = \min\left(C\left(\frac{P_1|h_{1p}|^2}{\sigma_\mathrm{R}^2}\right), C\left(\frac{P_\mathrm{R}|h_{p2}|^2}{\sigma_\mathrm{D}^2}\right),\right.$$
$$\left. C\left(\frac{P_\mathrm{R}|h_{34}|^2}{\sigma_\mathrm{R}^2 + P_1|h_{1p}|^2}\right)\right). \tag{68}$$

The first term in (68) 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. Strategy b) works well, when the inter-relay channel is strong, since the rate in (68) is not limited by the third expression and the relay may decode the new message coming from the source interference-free.

V. Simulation Results

A. Two-way Relaying

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

[10]The lower bound may even become negative when the relay power is much larger than the source power.

$$
\begin{aligned}
R[k+1] &= \mathcal{E}\left\{\log\left(\frac{\sigma_{\mathrm{D}}^2 + |h_{p2}g|^2\big(P_1 I_1[k] + \sigma_{\mathrm{R}}^2 I_2[k]\big) + P_1|h_{p2}gh_{1p}|^2}{\sigma_{\mathrm{D}}^2 + |h_{p2}g|^2\big(P_1 I_1[k] + \sigma_{\mathrm{R}}^2 I_2[k]\big)}\right)\right\} && (57)\\
&\geq \mathcal{E}\left\{\log\left(\frac{\sigma_{\mathrm{D}}^2 + |h_{p2}g|^2\big(P_1 I_{1,k'}[k] + \sigma_{\mathrm{R}}^2 I_{2,k'}[k]\big) + P_1|h_{p2}gh_{1p}|^2}{\sigma_{\mathrm{D}}^2 + |h_{p2}g|^2\big(P_1 I_1[k] + \sigma_{\mathrm{R}}^2 I_2[k]\big)}\right)\right\} && (58)\\
&\geq \mathcal{E}\left\{\log\left(P_1|h_{p2}gh_{1p}|^2 + \sigma_{\mathrm{D}}^2 + |h_{p2}g|^2\big(P_1 I_{1,k'}[k] + \sigma_{\mathrm{R}}^2 I_{2,k'}[k]\big)\right)\right\}\\
&\quad - \log\left(\sigma_{\mathrm{D}}^2 + \nu_2^2 g^2\big(P_1 \overline{I}_1[k] + \sigma_{\mathrm{R}}^2 \overline{I}_2[k]\big)\right) && (59)\\
&= R_{\mathrm{low}}[k+1]
\end{aligned}
$$

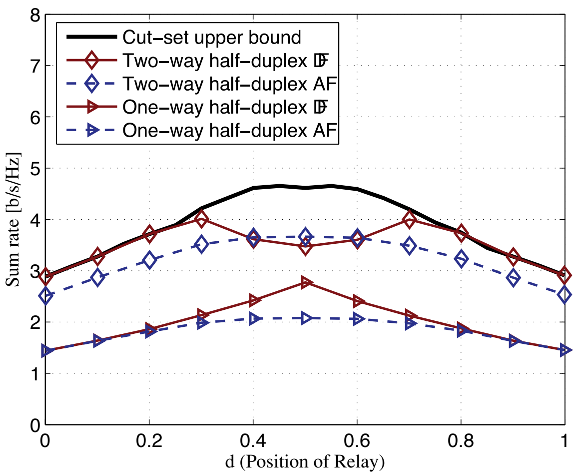

Fig. 4. Sum-rate for two-way half-duplex AF and DF relaying protocols

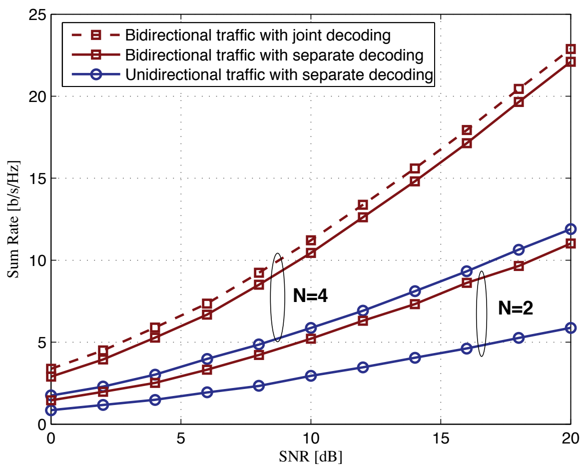

Fig. 5. Sum-rate for two-way half-duplex OF relaying protocol

Fig.1, where the distance between terminal $\mathtt{T}_1$ and $\mathtt{T}_2$ is normalized to one. The channel gains are modeled as $h_1 = \frac{\xi_1}{d^{\alpha/2}}$ and $h_2 = \frac{\xi_2}{(1-d)^{\alpha/2}}$ with i.i.d. $\xi_i \sim \mathcal{CN}(0,1)$ (Rayleigh fading), where d is the normalized distance between terminal $\mathtt{T}_1$ and relay $\mathtt{T}_3$, and $\alpha = 3$ the path loss exponent. The noise variances at the terminals are chosen as $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 1$ and the transmit powers $P_1 = P_2 = P_3 = 10$. We simulated 10000 random channels for each value in Fig.4 and Fig.5. In Fig.4 we compare the sum-rate of the two-way AF and two-way DF protocol with their one-way counterparts and the cut-set upper bound [21] applied to the two-way half-duplex relay channel with no direct link:

$$
\begin{aligned}
C_{\mathrm{sum}}^{\mathrm{u}} = \max_{0\leq\beta\leq 1}\Bigg(&\min\left(C\left(P_1|h_1|^2\right), C\left(\beta P_3|h_2|^2\right)\right)\\
&+\min\left(C\left(P_2|h_2|^2\right), C\left((1-\beta)\beta P_3|h_1|^2\right)\right)\Bigg) \quad (69)
\end{aligned}
$$

We observe that the DF protocol achieves the cut-set upper bound on the sum-rate when the relay is in the proximity of terminal $\mathtt{T}_1$ or terminal $\mathtt{T}_2$. Both two-way protocols, AF and DF, achieve sum-rates that are substantially larger than the rates of their one-way counterparts. In the one-way AF and DF strategies both protocols achieve the highest rate when the relay is exactly in the middle and DF achieves a higher maximum than AF. For the two-way case the DF protocol is worse than the AF protocol when the relay is in the middle. The reason is that the DF scheme has to cope with a multiple-access channel and the maximum sum-rate is achieved here for an asymmetric channel situation, i.e., when one terminal experiences a stronger channel gain than the other terminal.

In Fig.5 we compare the sum-rates of one-way and two-way OF relaying vs. signal-to-noise ratio (SNR). The noise variances are chosen to be $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2$ and the transmit powers $P_1 = P_2 = P_3 = P$. We simulated 10000 random channels for each value in Figs.5. The SNR is defined as $\mathrm{SNR} = \frac{P}{\sigma^2}$. As mentioned in Section III-C the noise terms at the destinations are spatially correlated. In order to see the effect of this noise correlation, we also plotted the sum rate when the destinations could perform joint decoding (i.e., considering the spatial noise correlations). We see that the loss due to separate decoding is small, since the transmissions are orthogonalized and neglecting the noise correlations at the receiving terminals does not degrade the performance significantly. We observe that two-way OF relaying with $N = 2$ terminal pairs and $K = 5$ relays (minimum relay configuration), i.e., 9 terminals in total, achieves almost the same sum-rate as one-way OF relaying with $N = 4$ terminal pairs and $K = 13$ relays, i.e., 21 terminals in total.

B. Two-path Relaying

We evaluate the achievable rates of the relaying schemes described in Section IV again by Monte Carlo simulations.

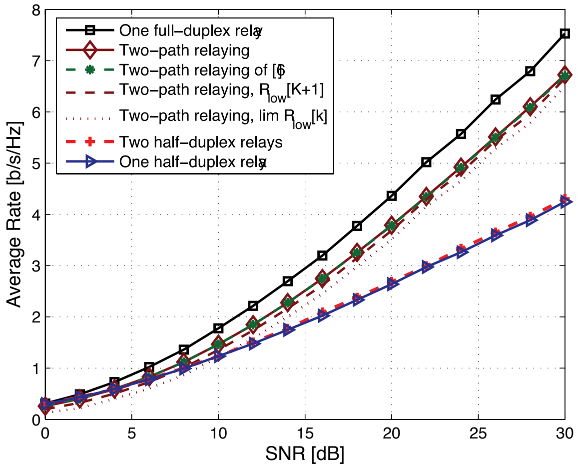

Fig. 6. Average rate vs. SNR for two-path half-duplex AF relaying protocol

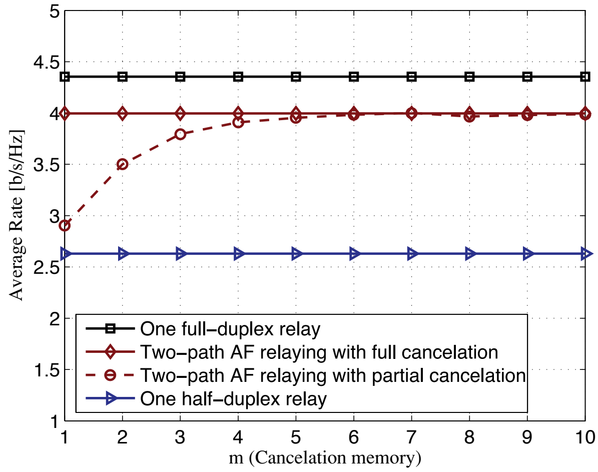

Fig. 8. Average rate vs. cancelation memory for two-path half-duplex AF relaying protocol

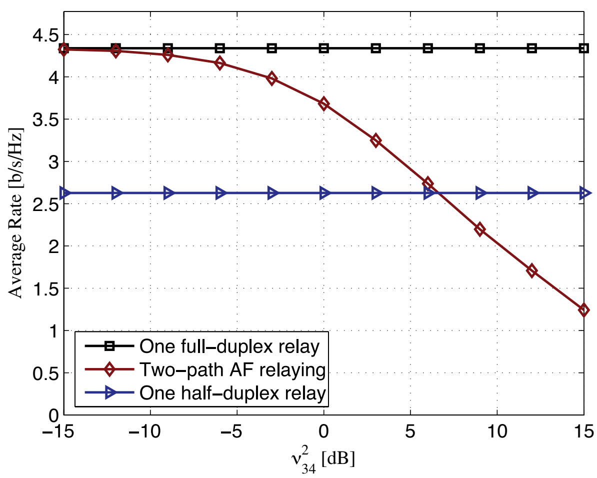

Fig. 7. Average rate vs. inter-relay channel gain for two-path half-duplex AF relaying protocol

We assume i.i.d (in space and time) channel gains $h_{13}, h_{14} \sim \mathcal{CN}(0, \nu_1^2)$, $h_{32}, h_{42} \sim \mathcal{CN}(0, \nu_2^2)$ and $h_{34} \sim \mathcal{CN}(0, \nu_{34}^2)$. For simplicity we assume $\nu_1^2 = \nu_2^2 = 1$ (symmetric network). The noise variances are chosen to be $\sigma_{\mathrm{R}}^2 = \sigma_{\mathrm{D}}^2 = \sigma^2$ and the transmit powers $P_1 = P_3 = P_4 = P$. We simulated 10000 random channels for each value in Figs.6–9. The SNR is defined as $\mathrm{SNR} = \frac{P}{\sigma^2}$.

In Fig.6 we see that for $\nu_{34}^2 = 0.5$ (inter-relay channel gain is 3dB weaker than the other channels) the two-path relaying protocol with AF relays and full cancelation of the accumulated inter-frame 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 when both source and relay transmit with power $2P$[11]. We also compared the performance of the two-path scheme with the case where two-half-duplex relays are used simultaneously, i.e., in the first time slot the source transmits with power $2P$ to both relays and in the second time slot the relays transmit each with power P. Both schemes outperform the two-path strategy in the low-SNR regime by 3dB but are inferior in the high SNR regime due to the pre-log factor one-half. For the lower bound (59) we have chosen $k^{'} = 0$ and $I_1[k] = 0$ (full interference cancelation). Further we observe from Fig.6 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 (44). We also plotted the average rate of the scheme proposed in [6], where the inter-relay interference is canceled at one of the relays. We observe that both schemes achieve practically the same performance. In our scheme the number of interference terms (relay noise samples) is larger than in [6], but the "older" relay noise terms have a small influence due to the repeated attenuation through the inter-relay channel.

Fig.7 shows the achievable rate (44) for different variances ν_{34}^2 of the inter-relay channel gain and $P/\sigma^2 = 100$. When the inter-relay channel gain is not too strong, the two-path AF relaying strategy performs very well. For inter-relay channel gains that are considerably stronger than the source-relay and relay-destination channel gains the two-path strategy does not work well due to the accumulated noise interference of the two relays at the destination.

In Fig.8 we compare the impact of full and partial interference cancelation on the average rate for $P/\sigma^2 = 100$ and $\nu_{34}^2 = 1/2$, cf. (43). For the simulation the number of transmitted frames K was chosen to be 30. We see that after cancelation of about five to six previously transmitted codewords the performance is the same as with full cancelation of the inter-frame interference. The inter-relay channel acts as forgetting factor (39) and lessens the inter-frame interference caused by the previously transmitted codewords.

In Fig.9 we compare the rates achievable by two-path DF relaying with and without interference cancelation at the relays for $P/\sigma^2 = 100$. 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-

[11] All schemes consume a network power of $2P$ per time slot.

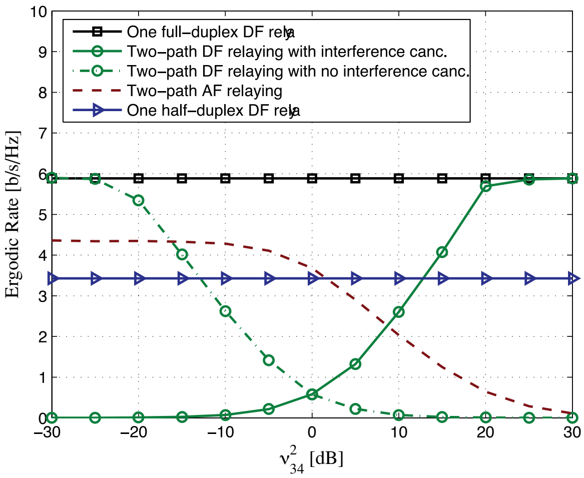

Fig. 9. Ergodic rate vs. inter-relay channel gain for two-path half-duplex DF relaying protocol

relay channels where the AF scheme performs well for weak to moderate inter-relay channels.

VI. Conclusions

We studied two half-duplex relay protocols with increased spectral efficiency compared to conventional half-duplex relay protocols. In the first protocol we established a bidirectional connection between two or more terminals using one or several half-duplex relays (*two-way relaying*). For the second protocol (*two-path relaying*) two relays assisted in the communication between two terminals. It was shown that both protocols may recover a large portion of the half-duplex loss for different relaying strategies. Further work may consider the extension of the protocols to multiple-input multiple-output (MIMO) terminals. For the two-path protocol one may also investigate the case where a direct connection between the source and destination terminals is available.

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] H. Hu, H. Yanikomeroglu, D. D. Falconer, and S. Periyalwar, "Range extension without capacity penalty in cellular networks with digital fixed relays," in *IEEE GLOBECOM*, vol. 5, (Dallas, (TX)), pp. 3053–3057, Nov. 2004.

[6] 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.

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

[8] A. Ribeiro, X. Cai, and G. B. Giannakis, "Opportunistic multipath for bandwidth-effcicent cooperative networking," in *IEEE ICASSP*, vol. 4, (Montreal, (CA)), pp. 549–552, May 2004.

[9] Y. Fan and J. Thompson, "Recovering multiplexing loss through successive relaying," *IEEE Trans. Wireless Commun.*, June 2006. submitted.

[10] B. Schein and R. Gallager, "The gaussian parallel relay network," in *Proc. IEEE Int. Symposium on Inf. Theory*, (Sorrento, (Italy)), p. 22, June 2000.

[11] M. Khojastepour, B. Aazhang, and A. Sabharwal, "On the capacity of 'cheap' relay networks," in *Proc. Conference on Information Sciences and Systems (CISS)*, (Princeton, NJ), Apr. 2003.

[12] A. Host-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 2020–2040, June 2005.

[13] A. Wittneben and B. Rankov, "Distributed antenna systems and linear relaying for gigabit MIMO wireless," in *IEEE Veh. Tech. Conf.*, (Los Angeles, LA), pp. 3624–3630, Sept. 2004.

[14] M. O. Hasna and M.-S. Alouini, "A performance study of dual-hop transmissions with fixed gain relays," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1963–1968, Nov. 2004.

[15] J. N. 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.

[16] M. Kuhn, S. Berger, I. Hammerström, and A. Wittneben, "Power-line enhanced cooperative wireless communications," *IEEE J. Select. Areas Commun.*, 2006. To appear.

[17] A. Wittneben and I. Hammerstroem, "Multiuser zero forcing relaying with noisy channel state information," in *Proc. IEEE Wirel. Comm. and Netw. Conf. (WCNC)*, (New Orleans, LA), Mar. 2005.

[18] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math. Stat. and Prob.*, vol. 1, pp. 611–644, 1961.

[19] B. Rankov and A. Wittneben, "Achievable rate regions for the two-way relay channel," in *Proc. IEEE Int. Symposium on Inf. Theory*, (Seattle, (WA)), 2006. To appear.

[20] B. Rankov, "Achieving MIMO gains in cooperative wireless systems." PhD Thesis, ETH Zurich.

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

Boris Rankov received the Dipl.-Ing. FH degree in electrical engineering from Zurich University of Applied Sciences Winterthur, Switzerland in November 1998 and the Dipl.-Ing. ETH degree from ETH Zurich, Switzerland in April 2002. Since August 2002 he is working towards his doctoral degree at the Communication Technology Laboratory (CTL), ETH Zurich, Switzerland. His main research interests are in the field of information theory and signal processing for wireless communications, in particular cooperative strategies in wireless networks and distributed MIMO systems.

Armin Wittneben is Full Professor of Wireless Communication at ETH Zurich/Switzerland and director of the Institute of Communication Technology. He graduated from the Technical University Darmstadt (TUD), Germany, where he obtained the Dipl.-Ing. degree in Electrical Engineering in 1983. In 1989 he received the doctorate in Electrical Engineering from the same school.

From 1989 until 1998 he was with ASCOM/Switzerland and in charge of the research activities of this company in wireless communications. In this period the department demonstrated among others the first Hiperlan/I modem and the fastest modem worldwide for digital powerline communication. In 1997 he completed his habilitation thesis and received the Venia Legendi in communication technology from the TUD. In 1998 he accepted a position as Full Professor of Communications at the Saarland University, Germany. In 2002 he joined ETH Zurich. His research work focuses on cooperative wireless communication, communication theory, digital signal processing, MIMO wireless and ultra wideband wireless.