Fading Relay Channels: Performance Limits and Space–Time Signal Design

Rohit U. Nabar, *Member, IEEE*, Helmut Bölcskei, *Senior Member, IEEE*, and
Felix W. Kneubühler, *Student Member, IEEE*

Abstract—**Cooperative diversity is a transmission technique, where multiple terminals pool their resources to form a virtual antenna array that realizes spatial diversity gain in a distributed fashion. In this paper, we examine the basic building block of cooperative diversity systems, a simple fading relay channel where the source, destination, and relay terminals are each equipped with single antenna transceivers. We consider three different time-division multiple-access-based cooperative protocols that vary the degree of broadcasting and receive collision. The relay terminal operates in either the amplify-and-forward (AF) or decode-and-forward (DF) modes. For each protocol, we study the ergodic and outage capacity behavior (assuming Gaussian code books) under the AF and DF modes of relaying. We analyze the spatial diversity performance of the various protocols and find that full spatial diversity (second-order in this case) is achieved by certain protocols provided that appropriate power control is employed. Our analysis unifies previous results reported in the literature and establishes the superiority (both from a capacity, as well as a diversity point-of-view) of a new protocol proposed in this paper. The second part of the paper is devoted to (distributed) space–time code design for fading relay channels operating in the AF mode. We show that the corresponding code design criteria consist of the traditional rank and determinant criteria for the case of colocated antennas, as well as appropriate power control rules. Consequently space–time codes designed for the case of colocated multiantenna channels can be used to realize cooperative diversity provided that appropriate power control is employed.**

I. INTRODUCTION

TRANSMISSION over wireless channels suffers from random fluctuations in signal level known as fading and from cochannel interference. Diversity is a powerful technique to mitigate fading and improve robustness to interference. In classical diversity techniques, the data signal is conveyed to the receiver over multiple (ideally) independently fading signal paths (in time/frequency/space). Appropriate combining at the receiver realizes diversity gain, thereby improving link reliability. Spatial or antenna diversity techniques are particularly attractive since they provide diversity gain without incurring an expenditure of transmission time or bandwidth. Signal design for multiantenna systems with colocated antennas (also known as space–time coding) aimed at extracting spatial diversity gain has been studied extensively in the literature [1]–[4].

A new way of realizing spatial diversity gain (in a distributed fashion) has recently been introduced in [5]–[8] under the name of *user cooperation diversity* or *cooperative diversity*. Here,
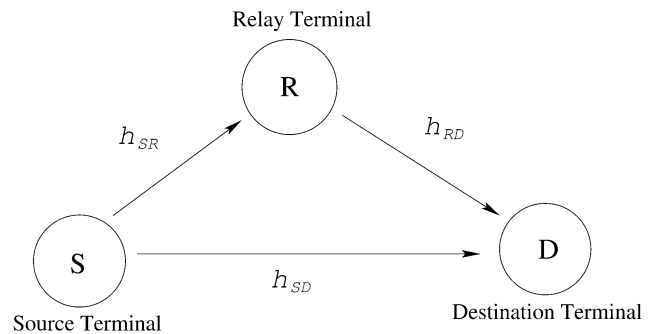
Fig. 1. Schematic of fading relay channel.

multiple terminals (sensors) in a network cooperate to form a virtual antenna array realizing spatial diversity in a distributed fashion. In [9], it has been demonstrated that uplink capacity can be increased via user cooperation diversity. A variety of cooperation protocols for channels with a single relay terminal have been studied and analyzed in [10]–[13]. In [14], it is shown that for channels with multiple relays, cooperative diversity with appropriately designed codes realizes full spatial diversity gain. We note that many cooperative diversity schemes can be cast into the framework of network coding [15]–[17]. Finally, we refer to [18], [19] for fundamental results on nonfading relay channels and to [20] and [21] for recent results on scaling laws in large (relay) networks.

Contributions and relation to previous work. The first part of this paper is devoted to the information-theoretic performance limits of three different time-division multiple-access (TDMA)-based transmission protocols for the single relay channel shown in Fig. 1. The protocols we consider implement varying degrees of broadcasting and receive collision in the network.[1] In each of the protocols, the relay terminal is allowed to either amplify-and-forward (AF) or decode-and-forward (DF) the signal received from the source terminal. The second part of the paper deals with (distributed) space–time code design for the fading relay channel operating in the AF mode. Our detailed contributions in relation to previous work reported in [5]–[14] are summarized as follows.

- We establish a *unified framework* for the results on fading relay channels reported in [5]–[14], *propose a new protocol which is superior to existing protocols* for the single-relay fading channel, and put the performance gains achievable in the distributed multiantenna case into

---

[1]The degree of broadcasting is determined by the number of nodes listening to a broadcasted message.

the context of traditional multiple-input–multiple-output (MIMO) gains.

- Assuming a *Gaussian codebook*, we derive *closed form expressions* for *the mutual information* associated with each of the protocols analyzed. Based on these results, we compare the performance of the different protocols in terms of achievable rates and establish the superiority of protocols implementing maximum degrees of broadcasting and receive collision.
- Based on an outage capacity analysis, we investigate the *diversity performance* of the proposed protocols. In particular, we find that *full spatial diversity* is achieved by certain protocols provided that *appropriate power control* is employed.
- For an AF single-relay fading channel, we *derive the design criteria for (distributed) space–time codes*. Our results indicate that optimal space–time code design in the single-relay case consists of satisfying *the classical rank and determinant criteria* for colocated antennas [2], as well as appropriate *power control rules* between the terminals. It is shown that the power control rule arising in the context of (distributed) space–time code design is equivalent to the power control rule obtained through an outage capacity analysis. Finally, we note that the differences between [14] and the space–time code design problem considered in this paper will be explained in greater detail in Section V.

Organization of the paper. The rest of this paper is organized as follows. Section II describes AF and DF single-relay channels and introduces the three different TDMA-based protocols, as well as the corresponding channel and signal models. Sections III and IV provide an information-theoretic comparison of the different protocols for the AF and DF cases, respectively. Section V deals with (distributed) space–time signal design for AF single-relay fading channels. We conclude in Section VI.

Notation. The superscripts ^{T}, ^{H}, and $*$ stand for transposition, conjugate transposition and element-wise conjugation, respectively. $\mathcal{E}$ denotes the expectation operator, $\mathbf{I}_m$ is the $m \times m$ identity matrix, $\mathbf{0}$ stands for an all zeros matrix of appropriate dimensions, and $\|\mathbf{a}\|$ is the Euclidean norm of the vector $\mathbf{a}$. A circularly symmetric complex Gaussian random variable is a random variable $Z = X + jY \sim \mathcal{CN}(0, \sigma^2)$, where X and Y are independent identically distributed (i.i.d.) $\mathcal{N}(0, \sigma^2/2)$.

II. PROTOCOL DESCRIPTIONS AND CHANNEL AND SIGNAL MODELS

A. General Setup and Protocol Descriptions

Consider the fading relay channel shown in Fig. 1. Data is to be transmitted from the source terminal S to the destination terminal D with the assistance of the relay terminal R. All terminals are equipped with single antenna transmitters and receivers. Throughout this paper, we assume that a terminal cannot transmit and receive simultaneously. The relay terminal assists in communication with the destination terminal by either amplifying-and-forwarding (AF) or decoding-and-forwarding (DF) the received signal. In the AF mode, the relay terminal

TABLE I
THREE DIFFERENT TDMA-BASED PROTOCOLS. S, R, AND, D STAND FOR THE SOURCE, RELAY, AND DESTINATION TERMINALS, RESPECTIVELY. $A \rightarrow B$ SIGNIFIES COMMUNICATION BETWEEN TERMINALS A AND B

Time Slot/Protocol	I	II	III
1	S→R, D	S→R, D	S→R
2	S→D, R→D	R→D	S→D, R→D

simply amplifies and retransmits the signal received from the source terminal (the signal received at the relay terminal is corrupted by fading and additive noise). No demodulation or decoding of the received signal is performed in this case. In the DF mode, the signal received from the source terminal is demodulated and decoded before retransmission. The signal models associated with the AF and DF transmission modes are discussed in greater detail in Section II-B. We note that in practice the AF mode when compared with the DF mode requires significantly lower implementation complexity at the relay terminal.

For each of the two forwarding modes (AF and DF) we shall next describe three different cooperative protocols, which implement varying degrees of broadcasting and receive collision in the network. The degree of broadcasting is given by the number of nodes simultaneously (i.e., in the same time slot) listening to the source node (i.e., 2 if both R and D listen, 1 if only R or D listens). Furthermore, receive collision is said to be maximum if the destination node receives information simultaneously from both S and R.

Protocol I: The source terminal communicates with the relay and destination terminals during the first time slot. In the second time slot, both the relay and source terminals communicate with the destination terminal. This protocol realizes maximum degrees of broadcasting and receive collision.

Protocol II: In this protocol, the source terminal communicates with the relay and destination terminals over the first time slot. In the second time slot, only the relay terminal communicates with the destination terminal. This protocol realizes a maximum degree of broadcasting and exhibits no receive collision.

Protocol III: The third protocol is identical to Protocol I apart from the fact that the destination terminal chooses not to receive the direct[2] $S \rightarrow D$ signal during the first time slot for reasons that will be motivated later in this section. This protocol does not implement broadcasting but realizes receive collision.

The protocols are summarized in Table I. Protocols II and III were first proposed in [8] and [22], respectively. Protocol I appears to be new. Note that while the signal conveyed to the relay and destination terminals over the two time slots is the same under Protocol II, Protocols I and III can potentially convey different signals to the relay and destination terminals. This fact will be exploited in Section V in the context of (distributed) space–time code design for fading relay channels.

Additional comments on the three protocols described above are in order. The conditions and setup for Protocol I are self-evident. Protocol II is logical in a scenario where the source terminal engages in data reception from another terminal in the network over the second time slot thereby rendering it unable

---

[2]$A \rightarrow B$ signifies the link between terminals A and B.

to transmit. Similarly, for Protocol III the destination terminal may be engaged in data transmission to another terminal during the first time slot. Hence, the transmitted signal is received only at the relay terminal and buffered for subsequent forwarding. We assume that the source terminal expends the same amount of power over the two time slots. In Protocol II, the source terminal is silent over the second time slot, which implies that this protocol is more efficient than Protocols I and III in terms of battery life.

B. Channel and Signal Models

Throughout this paper, we assume frequency-flat fading, no channel knowledge in the transmitters, perfect channel state information in the receivers and perfect synchronization. Perfect channel state information in the receivers implies that the $S \rightarrow R$ channel is known to the relay terminal, while the individual $S \rightarrow D, R \rightarrow D$, and $S \rightarrow R$ channels are known to the destination terminal. Depending on the relaying mode (AF or DF), knowledge of a specific individual channel gain may not be required at the relay/destination terminal. Such a relaxation of the assumption on channel knowledge will be highlighted in the corresponding discussion. The assumption on synchronization is most critical since synchronization becomes increasingly challenging in larger networks. Protocols II and III are essentially derivatives of Protocol I. We shall, therefore, first provide the input-output relation for Protocol I for both the AF and DF modes and then specialize to Protocols II and III.

Input–output relation for Protocol I in the AF mode. The signals transmitted by the source terminal during the first and second time slots are denoted as $x_1[n]$ and $x_2[n]$, respectively. In the following, we consider symbol-by-symbol transmission so that the time index n can be dropped and we simply write x_1 and x_2 for the symbols transmitted in the first and second time slots, respectively. We assume that $\mathcal{E}\{x_i\} = 0$ and $\mathcal{E}\{|x_i|^2\} = 1$ for $i = 1, 2$. The data symbols may be chosen from a complex-valued finite constellation such as quadrature amplitude modulation (QAM) or from a Gaussian codebook. The signal received at the destination terminal in the first time slot is given by

$$y_{D,1} = \sqrt{E_{SD}}h_{SD}x_1 + n_{D,1} \quad (1)$$

where E_{SD} is the average signal energy received at the destination terminal over one symbol period through the $S \rightarrow D$ link (having accounted for path loss and shadowing between the source and destination terminals), h_{SD} is the random,[3] complex-valued, unit-power channel gain between source and destination terminals and $n_{D,1} \sim \mathcal{CN}(0, N_o)$ is additive white noise. The signal received at the relay terminal during the first time slot is given by

$$y_{R,1} = \sqrt{E_{SR}}h_{SR}x_1 + n_{R,1} \quad (2)$$

where E_{SR} is the average signal energy over one symbol period received at the relay terminal (having accounted for path loss and shadowing between the source and relay terminals), h_{SR} is the random, complex-valued, unit-power channel gain between

the source and relay terminals and $n_{R,1} \sim \mathcal{CN}(0, N_o)$ is additive white noise. Note that in general $E_{SD} \neq E_{SR}$ due to differences in path loss and shadowing between the $S \rightarrow R$ and $S \rightarrow D$ links.

The relay terminal normalizes the received signal by a factor of $\sqrt{\mathcal{E}\{|y_{R,1}|^2\}}$ (so that the average energy is unity) and retransmits the signal during the second time slot. The destination terminal receives a superposition of the relay transmission and the source transmission during the second time slot according to

$$y_{D,2} = \sqrt{E_{SD}}h_{SD}x_2 + \sqrt{E_{RD}}h_{RD}\frac{y_{R,1}}{\sqrt{\mathcal{E}\{|y_{R,1}|^2\}}} + n_{D,2} \quad (3)$$

where E_{RD} is the average signal energy over one symbol period received at the destination terminal through the $R \rightarrow D$ link (having accounted for path loss and shadowing between the relay and destination terminals), h_{RD} is the random, complex-valued, unit-power channel gain between the relay and destination terminals and $n_{D,2} \sim \mathcal{CN}(0, N_o)$ is additive white noise. We note that (3) contains the additional assumption of constant E_{SD} and h_{SD} over the two time slots. Using $\mathcal{E}\{|y_{R,1}|^2\} = E_{SR} + N_o$, we can rewrite (3) as

$$y_{D,2} = \sqrt{E_{SD}}h_{SD}x_2 + \sqrt{\frac{E_{SR}E_{RD}}{E_{SR} + N_o}}h_{SR}h_{RD}x_1 + \tilde{n} \quad (4)$$

where the effective noise term $\tilde{n}|h_{RD} \sim \mathcal{CN}(0, N_o')$ with $N_o' = N_o(1 + (E_{RD}|h_{RD}|^2)/(E_{SR} + N_o))$. Finally, we assume that the receiver normalizes $y_{D,2}$ by a factor[4] $\omega = (1 + (E_{RD}|h_{RD}|^2)/(E_{SR} + N_o))^{1/2}$. This normalization does not alter the signal-to-noise ratio (SNR) but simplifies the ensuing presentation. The effective input-output relation for Protocol I in the AF mode can now be summarized as[5]

$$\mathbf{y}_1 = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (5)$$

where $\mathbf{y}_1 = [y_{D,1} \ y_{D,2}/\omega]^T$ is the received signal vector, $\mathbf{H}$ is the effective 2×2 channel matrix given by

$$\mathbf{H} = \begin{bmatrix} \sqrt{E_{SD}}h_{SD} & 0 \\ \frac{1}{\omega}\sqrt{\frac{E_{SR}E_{RD}}{E_{SR}+N_o}}h_{SR}h_{RD} & \frac{\sqrt{E_{SD}}}{\omega}h_{SD} \end{bmatrix} \quad (6)$$

$\mathbf{x} = [x_1 \ x_2]^T$ is the transmitted signal vector, and $\mathbf{n}$ (when conditioned on the channel $\mathbf{H}$) is circularly symmetric complex Gaussian noise with $\mathcal{E}\{\mathbf{n}|\mathbf{H}\} = \mathbf{0}$ and $\mathcal{E}\{\mathbf{n}\mathbf{n}^H|\mathbf{H}\} = N_o\mathbf{I}_2$. We shall make use of the fact that $\mathbf{n}$ conditioned on $\mathbf{H}$ is Gaussian when calculating the mutual information for the AF-based protocols in Section III.

Input–output relation for Protocol I in the DF mode. In the DF mode, still assuming Protocol I, the signal received at the destination terminal during the first time slot is identical to that for the AF mode and is, hence, given by (1). The signal received at the relay terminal is given by (2). Unlike the AF mode, the relay terminal now demodulates and decodes the signal received during the first time slot. Assuming that the signal is decoded correctly and retransmitted, we obtain

$$y_{D,2} = \sqrt{E_{SD}}h_{SD}x_2 + \sqrt{E_{RD}}h_{RD}x_1 + n_{D,2}.$$

---

[3]Unless specified otherwise, we do not make any assumptions on the precise distribution of the channel gains.

[4]Recall that we assumed perfect channel state information in the receiver.
[5]The subscript 1 in $\mathbf{y}_1$ reflects the fact that we are dealing with Protocol I.

The effective input-output relation in the DF mode for Protocol I can be summarized as

$$\mathbf{y}_1 = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{7}$$

where $\mathbf{y}_1 = [y_{D,1}\ y_{D,2}]^T$ is the received signal vector, $\mathbf{H}$ is the effective 2×2 channel matrix given by

$$\mathbf{H} = \begin{bmatrix} \sqrt{E_{\mathrm{SD}}}h_{\mathrm{SD}} & 0 \\ \sqrt{E_{\mathrm{RD}}}h_{\mathrm{RD}} & \sqrt{E_{\mathrm{SD}}}h_{\mathrm{SD}} \end{bmatrix} \tag{8}$$

$\mathbf{x} = [x_1\ x_2]^T$ is the transmitted signal vector, and $\mathbf{n}$ is additive white Gaussian noise with $\mathcal{E}\{\mathbf{n}\} = \mathbf{0}$ and $\mathcal{E}\{\mathbf{n}\mathbf{n}^H\} = N_o\mathbf{I}_2$. From (8) it is clear that knowledge of h_{SR} is not required at the destination terminal in the DF mode.

Input-output relation for Protocols II and III. The corresponding input–output relations for Protocols II and III in the AF and DF modes may be derived from (5) and (7), respectively. For Protocol II, the received signal for either forwarding mode can be written as

$$\mathbf{y}_2 = \mathbf{h}x_1 + \mathbf{n} \tag{9}$$

where $\mathbf{h}$ denotes the first column of $\mathbf{H}$ (chosen appropriately from (6) or (8) depending on the transmission mode) and $\mathbf{n}$ (conditioned on $\mathbf{h}$ in the AF mode) is the 2×1 additive white complex Gaussian noise vector with $\mathcal{E}\{\mathbf{n}\,|\,\mathbf{h}\} = \mathbf{0}$ and $\mathcal{E}\{\mathbf{n}\mathbf{n}^H\,|\,\mathbf{h}\} = N_o\mathbf{I}_2$. Similarly, the signal received at the destination terminal under Protocol III (the received signal is scalar in this case) satisfies

$$y_3 = \mathbf{g}^T\mathbf{x} + n \tag{10}$$

where $\mathbf{g}^T$ is the second row of $\mathbf{H}$ (chosen appropriately from (6) or (8) depending on the transmission mode) and n (conditioned on $\mathbf{g}$ in the AF mode) is scalar $\mathcal{CN}(0, N_o)$ additive white noise.

Note that the different protocols convert the spatially distributed antenna system into effective single-input–multiple-output (SIMO) (with Protocol II), multiple-input–single-output (MISO) (with Protocol III), and MIMO (with Protocol I) channels allowing the fundamental gains of multiple-antenna systems such as *diversity gain*, *array gain* and *interference canceling gain* to be exploited in a distributed fashion. We emphasize that *multiplexing gain* (i.e., a linear increase in achievable rate with the number of antennas in MIMO channels [23]–[26]) is conspicuously absent, since time is expended to create a virtual MIMO channel thereby negating any multiplexing gain. Further, note that the general structure and statistics of the effective channels created by the different protocols are different from the classical i.i.d. circularly symmetric complex Gaussian behavior widely used in the MIMO literature [2], [24], [25].

III. INFORMATION-THEORETIC PERFORMANCE OF PROTOCOLS IN THE AF MODE

In this section, we analyze the information-theoretic performance of the three different AF-based protocols introduced in Section II.

A. Mutual Information of AF-Based Protocols

In the following, we employ an ergodic block-fading channel model (with independent blocks) and assume an i.i.d. Gaussian codebook with covariance matrix $\mathbf{R_{xx}} = \mathcal{E}\{\mathbf{xx}^H\} = \mathbf{I}_2$. Moreover, we assume that the destination terminal has perfect knowledge of h_{SD}, h_{SR}, and h_{RD}. The mutual information for Protocols I-III is obtained from (5), (9), and (10) as[6]

$$I_j^{\mathrm{AF}} = \frac{1}{2}\log_2\det\left(\mathbf{I}_2 + \frac{1}{N_o}\mathbf{A}_j\mathbf{A}_j^H\right)\ \mathrm{bps/Hz}, \quad j = 1, 2, 3 \tag{11}$$

where $\mathbf{A}_1 = \mathbf{H}, \mathbf{A}_2 = \mathbf{h}, \mathbf{A}_3 = \mathbf{g}$, and the factor $1/2$ accounts for the fact that information is conveyed to the destination terminal over two time slots. If coding is performed over an infinite number of independent channel realizations, the capacity of each of the three protocols, C_j^{AF} $(j = 1, 2, 3)$, is given by the ergodic capacity $C_j^{\mathrm{AF}} = \mathcal{E}\{I_j^{\mathrm{AF}}\}$ with the expectation carried out with respect to the random channel. We emphasize that C_j^{AF} is the capacity of the single-relay fading channel *in conjunction* with Protocol j. If coding is performed only within one block the Shannon capacity is zero. In this case, we resort to the $p\%$ outage capacity [27], [28], $C_{j,p,\mathrm{out}}^{\mathrm{AF}}$ $(j = 1, 2, 3)$, defined as

$$P\left(I_j^{\mathrm{AF}} \leq C_{j,p,\mathrm{out}}^{\mathrm{AF}}\right) = p\% \tag{12}$$

or equivalently, the rate $C_{j,p,\mathrm{out}}^{\mathrm{AF}}$ is guaranteed to be supported for $(100 - p)\%$ of the channel realizations. In the following, we compare the different protocols in the AF mode both from a capacity (ergodic and outage) and a diversity point-of-view.

B. Comparison From a Capacity Point-of-View

We begin with a comparison of Protocols I and II. Note that $I_1^{\mathrm{AF}} = I(\mathbf{y}_1; \mathbf{x})$, where $I(\mathbf{y}_1; \mathbf{x})$ is the mutual information between the vectors $\mathbf{y}_1$ and $\mathbf{x}$ as defined in (5). Applying the chain rule for mutual information [29], we have

$$I_1^{\mathrm{AF}} = I(\mathbf{y}_1; x_2) + I(\mathbf{y}_1; x_1\,|\,x_2),$$

where $I(\mathbf{y}_1; x_2)$ is the mutual information between $\mathbf{y}_1$ and x_2, while $I(\mathbf{y}_1; x_1\,|\,x_2)$ is the conditional mutual information between $\mathbf{y}_1$ and x_1 given x_2. It is easy to verify that $I(\mathbf{y}_1; x_1\,|\,x_2) = I(\mathbf{y}_2; x_1)$, where $\mathbf{y}_2$ is defined in (9). Noting that $I(\mathbf{y}_2; x_1) = I_2^{\mathrm{AF}}$ it then follows that

$$I_1^{\mathrm{AF}} = I(\mathbf{y}_1; x_2) + I_2^{\mathrm{AF}}.$$

Since $I(\mathbf{y}_1; x_2) \geq 0$ it follows immediately that $I_1^{\mathrm{AF}} \geq I_2^{\mathrm{AF}}$, which shows that the achievable rate for Protocol I is higher than that for Protocol II. We have, therefore, shown the intuitive result that the information rate is reduced if the source terminal does not transmit to the destination terminal in the second time slot. We note, however, that the superiority of Protocol I comes at the cost of increased receiver complexity which is due to the fact that in the second time slot the destination terminal receives the superposition of the signals from source and relay terminals whereas Protocol II is collision-free in the second time slot. This result establishes the importance of receive collision for

---

[6]Recall that the noise is conditionally (on the channel) Gaussian.

achieving high throughput. In the context of multiaccess fading channels, a similar observation has been made by Gallager in [30].

We shall next compare Protocols II and III and start by noting that for $\mathbf{h}$ and $\mathbf{g}$ defined in (9) and (10), respectively, we have $\|\mathbf{h}\|^2 \geq \|\mathbf{g}\|^2$ (since $\omega^2 \geq 1$), and consequently, $I_2^{\mathrm{AF}} \geq I_3^{\mathrm{AF}}$. We can, therefore, summarize our results as follows:

$$I_1^{\mathrm{AF}} \geq I_2^{\mathrm{AF}} \geq I_3^{\mathrm{AF}} \tag{13}$$

establishing the superiority of Protocol I over the other two protocols in terms of achievable rate. We emphasize that the ordering in (13) applies to ergodic and outage capacities for all three protocols.[7] Note that $\omega \approx 1$ implies $I_2^{\mathrm{AF}} \approx I_3^{\mathrm{AF}}$. The factor ω may be viewed as a noise amplification factor. In order to have $\omega \approx 1$, we need $(E_{\mathrm{RD}}|h_{\mathrm{RD}}|^2)/(E_{\mathrm{SR}} + N_o) \approx 0$, which is the case if the S $\rightarrow$ R link is good (i.e., $E_{\mathrm{SR}}/N_o \gg 1$) and much stronger than the R $\rightarrow$ D link. Physically, this may occur when the source terminal is located very close to the relay terminal resulting in high SNR for the S $\rightarrow$ R link. On the other hand, if $E_{\mathrm{RD}}|h_{\mathrm{RD}}|^2 \gg (E_{\mathrm{SR}} + N_o)$ the noise amplification will be substantial and the performance of Protocol III will deteriorate significantly compared with Protocol II. The reason for this is intuitively clear. In Protocol II, the destination terminal receives the source transmission over the first time slot without any added amplified noise from the relay terminal, whereas in Protocol III the information transmitted in the first time slot arrives at the destination terminal through the noise-amplifying relay link. Hence, Protocol II is expected to outperform Protocol III if the noise amplification is large. Due to the assumption of i.i.d. (across time slots) codebooks, the information transmitted in the second time slot of Protocol III on the S $\rightarrow$ D link is independent of the corrupted signal transmitted in the first time slot and can, therefore, not compensate for the poor relay link.

Finally, we shall interpret the ordering in (13) in terms of traditional MIMO gains. From (11), we can see that the price to be paid for cooperative transmission over two time slots is a reduction in spectral efficiency (compared with a MIMO system with colocated antennas) accounted for by the factor $1/2$ in front of the log term. As evidenced by (11), Protocol I is the only protocol that can realize a multiplexing gain in the classical sense and, hence, recover (to a certain extent) from this 50% loss in spectral efficiency. We note, however, that the effective channel is not i.i.d. complex Gaussian as is the case in traditional MIMO systems. This implies that in general we may not recover fully from the loss in spectral efficiency. The corresponding difference in performance can be attributed to the fact that we are dealing with a distributed system where the individual terminals have to cooperate through noisy links. A more detailed quantitative discussion of this performance difference is in many cases possible but seems beyond the scope of this paper. Protocols II and III do not provide multiplexing gain, which explains their inferior performance when compared with Protocol I. Finally, the fact that Protocol II is superior to Protocol III can be attributed to the fact that Protocol II corresponds to a SIMO system realizing array gain, whereas Protocol III corresponds to a MISO system devoid of array

---

[7]Recall that the source terminal was assumed to expend the same amount of power over the two time slots. Allowing a flexible allocation of transmit power across the two time slots can lead to an ordering different from (13).

gain (recall that we assumed perfect channel knowledge in the receivers and no channel knowledge in the transmitters). Maximizing the degree of broadcasting and receive collision (as is done in Protocol I) will in general result in a higher number of degrees-of-freedom (and, hence, higher achievable rates in the degrees-of-freedom limited case) reflected by the creation of an effective MIMO channel.

C. Diversity Performance

We shall next analyze and compare the different protocols from a diversity point-of-view. Following the approach in [31] and [8], we shall interpret the outage probability at a certain transmission rate as the packet-error rate (PER). The diversity order is then given by the magnitude of the slope of the PER as a function of SNR (on a log-log scale). To be more precise, we define the diversity order for transmission rate R as

$$d(R) = \lim_{\mathrm{SNR} \to \infty} \frac{-\log P_e(R, \mathrm{SNR})}{\log \mathrm{SNR}} \tag{14}$$

where $P_e(R, \mathrm{SNR})$ denotes the PER or outage probability at transmission rate R as a function of SNR. Equivalently, a scheme achieving diversity order $d(R)$ at rate R has an error probability that behaves as $P_e(R, \mathrm{SNR}) \propto \mathrm{SNR}^{-d(R)}$ at high SNR. In the remainder of this subsection, we assume that the channel gains h_{SD} and h_{SR} are independent $\mathcal{CN}(0,1)$, which corresponds to Rayleigh fading on these two links. Furthermore, we take the channel between the relay terminal and the destination terminal to be additive white Gaussian noise (AWGN) (i.e., $h_{\mathrm{RD}} = 1$). We note that the latter assumption is conceptual and simplifies the performance analysis significantly. The general case seems difficult to deal with analytically. Physically, this assumption could correspond to a scenario where the destination and relay terminals are static and have line-of-sight connection, while the source terminal is moving.

We start by investigating Protocol III and noting that I_3^{AF} can be lower-bounded as

$$I_3^{\mathrm{AF}} \geq \frac{1}{2} \log_2 \left(1 + \beta_3^{\mathrm{AF}}(|h_{\mathrm{SD}}|^2 + |h_{\mathrm{SR}}|^2) \right) \tag{15}$$

where

$$\beta_3^{\mathrm{AF}} = \min \left\{ \frac{1}{\omega^2} \frac{E_{\mathrm{SD}}}{N_o}, \frac{1}{\omega^2} \frac{E_{\mathrm{SR}} E_{\mathrm{RD}}}{(E_{\mathrm{SR}} + N_o) N_o} \right\} \tag{16}$$

and $\omega^2 = 1 + (E_{\mathrm{RD}})/(E_{\mathrm{SR}} + N_o)$ under the simplifying assumption, $h_{\mathrm{RD}} = 1$, made above. It follows that the outage probability at transmission rate R can be upper-bounded according to

$$P\left(I_3^{\mathrm{AF}} \leq R\right) \leq P\left(|h_{\mathrm{SD}}|^2 + |h_{\mathrm{SR}}|^2 \leq \frac{2^{2R} - 1}{\beta_3^{\mathrm{AF}}}\right).$$

Recalling that $|h_{\mathrm{SD}}|$ and $|h_{\mathrm{SR}}|$ are independent Rayleigh distributed and using the approximation $e^{-(2^{2R}-1)/(\beta_3^{\mathrm{AF}})} \approx 1 - (2^{2R}-1)/(\beta_3^{\mathrm{AF}})$ for β_3^{AF} sufficiently large, we obtain

$$P\left(I_3^{\mathrm{AF}} \leq R\right) \leq \left(\frac{2^{2R} - 1}{\beta_3^{\mathrm{AF}}}\right)^2 \tag{17}$$

which using (14) shows that second-order diversity is achieved in the effective SNR β_3^{AF}. We emphasize, however, that the diversity performance being determined by the effective SNR β_3^{AF}

implies that careful power control among terminals is necessary to ensure that the error rate decays according to a second-order diversity behavior. In order to further illustrate the necessity for power control, consider the case where E_{SD} and E_{SR} are kept constant, and E_{RD} is increased. For $E_{RD} > (E_{SD}(E_{SR} + N_o))/E_{SR}$, we have $\beta_3^{AF} = E_{SD}/(\omega^2 N_o)$ which shows that error probability performance is not improved (according to a second-order diversity behavior) by further increasing E_{RD}. Instead, careful balancing between E_{SD}, E_{SR}, and E_{RD} is necessary which can be achieved through power control. We note that these energy levels are in general not independently controllable. In particular, for fixed path loss on the $S \rightarrow D$ and $S \rightarrow R$ links, the quantities E_{SD} and E_{SR} will be a function of the transmit power only and, hence, cannot be adjusted separately. We conclude that depending on the propagation conditions a certain target effective SNR may not be achievable. We proceed to analyze Protocol II by defining

$$\beta_2^{AF} = \min\left\{ \frac{E_{SD}}{N_o}, \frac{1}{\omega^2} \frac{E_{SR} E_{RD}}{(E_{SR} + N_o) N_o} \right\}$$

and noting that for sufficiently large β_2^{AF}

$$P\left(I_2^{AF} \leq R \right) \leq \left(\frac{2^{2R} - 1}{\beta_2^{AF}} \right)^2 .$$

Hence, Protocol II achieves second-order diversity in the effective SNR β_2^{AF}. For Protocol I, we finally obtain

$$I_1^{AF} = \frac{1}{2} \log_2 \left(1 + \left(1 + \frac{1}{\omega^2} \right) \frac{E_{SD}}{N_o} |h_{SD}|^2 \right.$$
$$\left. + \frac{1}{\omega^2} \frac{E_{SR} E_{RD}}{(E_{SR} + N_o) N_o} |h_{SR}|^2 + \left(\frac{E_{SD}|h_{SD}|^2}{\omega N_o} \right)^2 \right) .$$

Ignoring the term proportional to $|h_{SD}|^4$, we can upper bound the outage probability for Protocol I according to

$$P\left(I_1^{AF} \leq R \right) \leq P\left(|h_{SD}|^2 + |h_{SR}|^2 \leq \frac{2^{2R} - 1}{\beta_1^{AF}} \right)$$

where

$$\beta_1^{AF} = \min\left\{ \left(1 + \frac{1}{\omega^2} \right) \frac{E_{SD}}{N_o}, \frac{1}{\omega^2} \frac{E_{SR} E_{RD}}{(E_{SR} + N_o) N_o} \right\} . \quad (18)$$

For β_1^{AF} sufficiently large it, therefore, follows that:

$$P\left(I_1^{AF} \leq R \right) \leq \left(\frac{2^{2R} - 1}{\beta_1^{AF}} \right)^2$$

which shows that Protocol I achieves second-order diversity in the effective SNR β_1^{AF}. Finally, by inspection we obtain the following ordering of effective SNRs

$$\beta_1^{AF} \geq \beta_2^{AF} \geq \beta_3^{AF} \quad (19)$$

which demonstrates the superiority of Protocol I over the other two protocols from an effective SNR point-of-view. We summarize the results of this section by noting that all three protocols achieve second-order diversity in their effective SNRs. Recall that in traditional MIMO systems the presence of array gain is reflected by an increased receive SNR when compared with the case where no array gain is present. Consequently, the ordering in (19) can be interpreted as reflecting the amount of array gain realized by the individual protocols in the AF mode.

IV. INFORMATION-THEORETIC PERFORMANCE OF PROTOCOLS IN THE DF MODE

In this section, we analyze the information-theoretic performance of the three different protocols in the DF mode. Throughout our analysis, we assume Gaussian code books with $\mathbf{R_{xx}} = \mathbf{I_2}$.

A. Achievable Rates for DF-Based Protocols

Let us start by analyzing Protocol I. We assume that the destination terminal has perfect knowledge of h_{SD} and h_{RD} (knowledge of h_{SR} is required only in the relay terminal). Assume that the transmission rates over the first and second time slots are R_1 and R_2, respectively. For the relay terminal to be able to decode the transmitted signal correctly R_1 must satisfy

$$R_1 \leq \log_2 \left(1 + \frac{E_{SR}}{N_o} |h_{SR}|^2 \right) . \quad (20)$$

In the following, we assume that if (20) is satisfied the relay terminal produces an error-free estimate of the transmitted signal. The channel in (7) may then be interpreted as a vector (multiple receive ports) multiple-access channel (MAC) [29], which imposes constraints on the individual rates R_1 and R_2, as well as the sum-rate $R_1 + R_2$ for successful decoding at the destination terminal. Particularly, R_1 and R_2 must satisfy [32]

$$R_1 \leq \log_2 \left(1 + \frac{1}{N_o} \|\mathbf{h}\|^2 \right) \quad (21)$$

$$R_2 \leq \log_2 \left(1 + \frac{E_{SD}}{N_o} |h_{SD}|^2 \right) \quad (22)$$

$$R_1 + R_2 \leq \log_2 \det\left(\mathbf{I_2} + \frac{1}{N_o} \mathbf{HH}^H \right) \quad (23)$$

where $\mathbf{H}$ is defined in (8) and $\mathbf{h}$ denotes the first column of $\mathbf{H}$. Equations (21)–(23) define the capacity region of the vector MAC (see Fig. 2). Any rate pair (R_1, R_2) satisfying these constraints is achievable (over the vector MAC). For the sake of convenience, in the following discussion, we shall denote the right-hand side (RHS) of (20), (21), (22), and (23) as $R_{relay}^{max}, R_1^{max}, R_2^{max}$, and R_{total}, respectively. R_{total} is the maximum sum-rate supported by the MAC, or equivalently, in our setup the maximum sum-rate over the two time slots. This sum-rate upper bounds the achievable total spectral efficiency for Protocol I in the DF mode. Note that R_1 must satisfy $R_1 \leq \min\{R_{relay}^{max}, R_1^{max}\}$, which implies that under certain channel conditions (namely when the $S \rightarrow R$ link is weak), R_{total} may not be achievable. Defining the maximum achievable sum-rate for Protocol I in the DF mode as $R_{sum}^{(1)}$ it follows that:

$$R_{sum}^{(1)} = \begin{cases} R_{total}, & R_{relay}^{max} \geq R_{total} - R_2^{max} \\ R_{relay}^{max} + R_2^{max}, & R_{relay}^{max} < R_{total} - R_2^{max}. \end{cases} \quad (24)$$

Hence, R_{total} is not achievable if the $S \rightarrow R$ link is weak and becomes the bottleneck during the first time slot. Denoting the maximum achievable sum-rate for Protocol II in the DF mode as $R_{sum}^{(2)}$ and noting that $R_2 = 0$ (source terminal is silent during the second time slot), it is easy to verify that

$$R_{sum}^{(2)} = \min\{R_{relay}^{max}, R_1^{max}\} \quad (25)$$
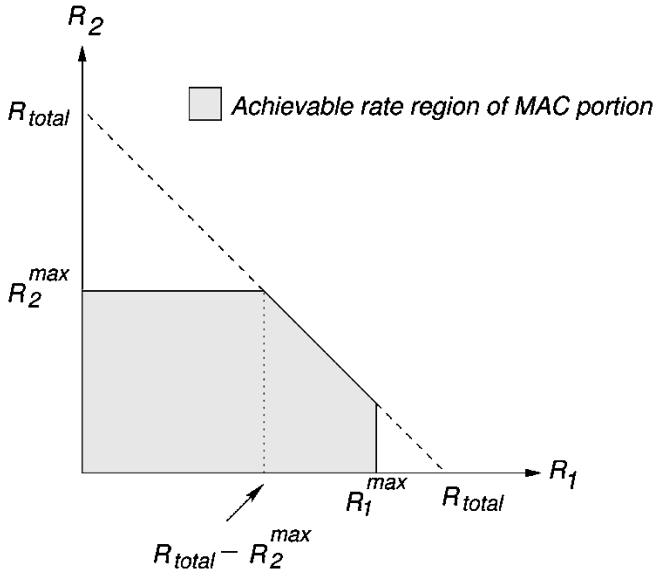
Fig. 2. Achievable rate region of the multiple-access portion of the fading relay channel for Protocol I in the DF mode.

where $R_{\text{relay}}^{\max}$ and $R_1^{\max}$ are defined in (20) and (21), respectively, i.e., with respect to Protocol I. Finally, we note that for Protocol III the transmission rate over the first time slot is constrained by

$$R_1 \leq \min\left\{ R_{\text{relay}}^{\max}, \log_2\left(1 + \frac{E_{\text{RD}}}{N_o}|h_{\text{RD}}|^2\right) \right\}.$$

Recalling that $\|\mathbf{g}\| = \|\mathbf{h}\|$ in the DF mode, we can show that the maximum sum-rate for Protocol III is given by

$$R_{\text{sum}}^{(3)} = \begin{cases} R_1^{\max}, & R_{\text{relay}}^{\max} \geq R_1^{\max} - R_2^{\max} \\ R_{\text{relay}}^{\max} + R_2^{\max}, & R_{\text{relay}}^{\max} < R_1^{\max} - R_2^{\max} \end{cases} \quad (26)$$

where once again we alert the reader to the fact that $R_1^{\max}$ and $R_2^{\max}$ are defined with respect to Protocol I in (21) and (22), respectively. In the following, we shall be interested in the sum-rate achievable by the different protocols in the DF mode. The corresponding (sum) mutual information associated with the three protocols in the DF mode is given by

$$I_j^{\text{DF}} = \frac{R_{\text{sum}}^{(j)}}{2} \text{ bps/Hz}, \quad j = 1, 2, 3 \quad (27)$$

where again the factor $1/2$ reflects the fact that transmission occurs over two time slots. The outage capacity for each of the protocols in the DF mode follows from the corresponding definition for the AF-based protocols [cf. (12)]. In order to compute the ergodic capacity for the DF protocol, we need to calculate the ergodic information rate supported by the $\text{S} \rightarrow \text{R}$ link, $\bar{R}_{\text{relay}}^{\max} = \mathcal{E}\{R_{\text{relay}}^{\max}\}$, as well as the ergodic capacity region for the MAC portion of the relay channel. For the latter see [33] for details. The ergodic capacity for the three protocols in the DF mode is then obtained through a relative comparison of $\bar{R}_{\text{relay}}^{\max}$ and the ergodic rate region in a similar fashion as for the case of given channel realizations discussed above.

B. Comparison From a Capacity Point-of-View

Let us start by comparing Protocols I and III. It follows by inspection of (24) and (26) that $I_1^{\text{DF}} \geq I_3^{\text{DF}}$, with equality if $R_{\text{relay}}^{\max} < R_1^{\max} - R_2^{\max}$.

We proceed by comparing Protocols II and III. Note from (25) and (26) that if the $\text{S} \rightarrow \text{R}$ link is strong so that $R_{\text{relay}}^{\max} \geq R_1^{\max}$, then both Protocols II and III achieve the same sum-rate of $R_1^{\max}$. However, if the relay channel is poor so that $R_{\text{relay}}^{\max} < R_1^{\max}$ then Protocol III outperforms Protocol II. Therefore, $I_3^{\text{DF}} \geq I_2^{\text{DF}}$ with equality if $R_{\text{relay}}^{\max} \geq R_1^{\max}$. Note that this result is in direct contrast to the comparative spectral efficiencies for Protocols II and III in the AF mode, where we saw that $I_2^{\text{AF}} \geq I_3^{\text{AF}}$. In summary, the mutual information for the three protocols satisfies

$$I_1^{\text{DF}} \geq I_3^{\text{DF}} \geq I_2^{\text{DF}}. \quad (28)$$

Consequently, the relation between the ergodic and outage capacities follows the same ordering. Again, the superiority of Protocol I can be attributed to the fact that it realizes multiplexing gain in the classical sense and, hence, recovers from (some of) the loss due to the use of two time slots for transmission. Finally, we note that again allowing a flexible allocation of transmit power across the two time slots can lead to an ordering which is different from (28).

C. Comparison From a Diversity Point-of-View

Let us next compare the three different protocols in the DF mode in terms of their diversity performance. Again, we make the conceptual assumption of the $\text{R} \rightarrow \text{D}$ link being AWGN. The general case seems rather difficult to analyze. Furthermore, for the sake of simplicity of exposition we assume that the $\text{R} \rightarrow \text{D}$ link is stronger than the $\text{S} \rightarrow \text{D}$ and $\text{S} \rightarrow \text{R}$ links (i.e., $E_{\text{RD}} \gg E_{\text{SD}}$ and $E_{\text{RD}} \gg E_{\text{SR}}$) so that $R_{\text{relay}}^{\max} < R_1^{\max} - R_2^{\max}$ (with probability close to 1). This assumption is reasonable when the relay terminal is located close to the destination terminal. We start by analyzing the diversity performance of Protocol II. Under the above assumptions, we have

$$I_2^{\text{DF}} = \frac{R_{\text{relay}}^{\max}}{2}.$$

It follows that for $(E_{\text{SR}}/N_o) \gg 1$ and a transmission rate of R the outage probability can be upper-bounded as

$$P\left(I_2^{\text{DF}} \leq R\right) \leq \frac{2^{2R} - 1}{\beta_2^{\text{DF}}}$$

where $\beta_2^{\text{DF}} = E_{\text{SR}}/N_o$. We can, therefore, conclude that Protocol II extracts only first-order diversity in the DF mode. This is also true if the $\text{R} \rightarrow \text{D}$ channel is Rayleigh fading as shown in [8] and follows intuitively from the fact that the information rate for Protocol II in the DF mode can never exceed that supported by the Rayleigh fading $\text{S} \rightarrow \text{R}$ link in either case [cf. (25)]. For Protocol III, under the simplifying assumptions made above, we obtain

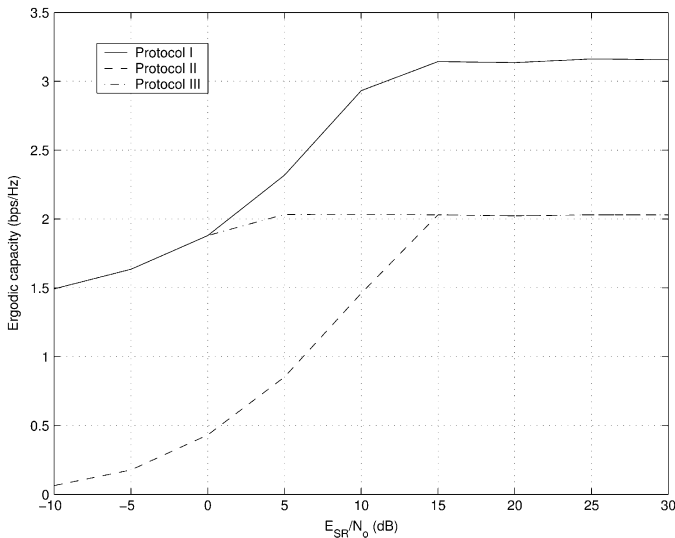$$I_3^{\text{DF}} = \frac{R_{\text{relay}}^{\max} + R_2^{\max}}{2}.$$

Fig. 3. Comparison of the ergodic capacities of the three protocols in the DF mode for $E_{\mathrm{SD}}/N_o = E_{\mathrm{RD}}/N_o = 10$ dB. Protocol II is severely limited by the relay channel for low E_{SR}/N_o. Protocol I benefits from "multiplexing gain" in the high E_{SR}/N_o regime recovering some of the loss due to TDMA-based transmission.

Fig. 4. Comparison of the ergodic capacities of the three protocols in the AF mode for $E_{\mathrm{SD}}/N_o = E_{\mathrm{RD}}/N_o = 10$ dB. For low E_{SR}/N_o, noise amplification in the relay terminal becomes significant resulting in the performance of Protocol III being significantly worse than that of Protocols I and II.

Defining $\beta_3^{\mathrm{DF}} = \min\{E_{\mathrm{SR}}/N_o, E_{\mathrm{SD}}/N_o\}$, it is easy to verify that for β_3^{DF} large the outage probability corresponding to Protocol III for transmission rate R can be upper-bounded according to

$$P(I_3^{\mathrm{DF}} \leq R) \leq \left(\frac{2^{2R}-1}{\beta_3^{\mathrm{DF}}}\right)^2$$

which demonstrates that Protocol III extracts second-order diversity in the effective SNR β_3^{DF}. Finally, we note that under the simplifying assumptions made above on the $\mathrm{S} \rightarrow \mathrm{R}, \mathrm{R} \rightarrow \mathrm{D}$, and $\mathrm{S} \rightarrow \mathrm{D}$ links, $I_1^{\mathrm{DF}} = I_3^{\mathrm{DF}}$. Consequently, Protocol I will also extract second-order diversity in the DF mode.

Finally, we note that if the $\mathrm{R} \rightarrow \mathrm{D}$ link is assumed fading and the $\mathrm{S} \rightarrow \mathrm{R}$ link is static with E_{SR} large, so that the $\mathrm{S} \rightarrow \mathrm{R}$ link is not a bottleneck over the first time-slot, then it is straightforward to show that all three protocols are capable of extracting second-order diversity in the DF mode.

D. Numerical Results

We conclude our discussion of the performance limits of the individual protocols with numerical results quantifying some of our analytical findings. Figs. 3 and 4 show the ergodic capacities (found through Monte Carlo simulation) for the three different protocols in the DF and AF modes, respectively, as a function of E_{SR}/N_o with $E_{\mathrm{SD}}/N_o = E_{\mathrm{RD}}/N_o = 10$ dB. The complex channel gains $h_{\mathrm{SR}}, h_{\mathrm{SD}}$, and h_{RD} are assumed i.i.d. $\mathcal{CN}(0,1)$. Fig. 3 verifies the ordering in (28) which holds irrespectively of the fading statistics (Rayleigh/Ricean/AWGN) of the individual channels. We can furthermore see that Protocol II is severely throughput limited compared with Protocols I and III when the $\mathrm{S} \rightarrow \mathrm{R}$ link is poor. Moreover, Protocols I and III perform equally well in this case. At high E_{SR}/N_o (i.e., when the $\mathrm{S} \rightarrow \mathrm{R}$ link is no longer a bottleneck), Protocols II and III perform equally well, but are clearly outperformed by Protocol

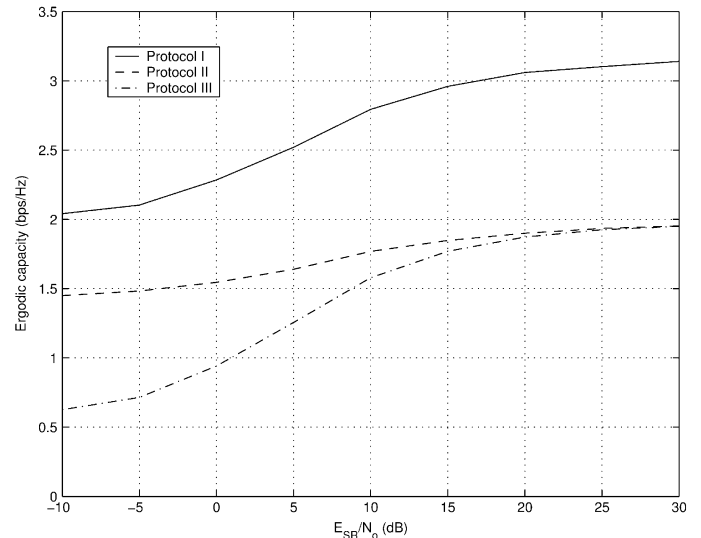I which benefits from "multiplexing gain," recovering some of the factor $1/2$ loss due to TDMA-based transmission.

Fig. 4 verifies the ordering in (13) and shows that the results are different in the AF mode. For low E_{SR}/N_o (i.e., noise in the relay terminal undergoes large amplification), Protocol III performs significantly worse than Protocols I and II. When the noise amplification is low, i.e., E_{SR}/N_o is high, Protocols II and III perform equally well and are significantly outperformed by Protocol I, which again benefits from "multiplexing gain."

V. SPACE–TIME SIGNAL DESIGN FOR AF-BASED PROTOCOLS

In this section, we examine (distributed) space–time signal design for the fading relay channel assuming AF-based protocols. The problem of (distributed) space–time signal construction does not apply to Protocol II since the effective channel resembles a SIMO channel. Employing multiple relay terminals, [14] discusses space–time code design for Protocol II. In contrast to the setup considered in this paper, [14] deals with space–time coding across relay terminals with the relays transmitting (linear or nonlinear) functions of the signal received from the source terminal to realize spatial diversity gain in a distributed fashion. In the following, for the sake of simplicity, we restrict ourselves to space–time code construction for Protocol III. We assume that the $\mathrm{S} \rightarrow \mathrm{D}$ and $\mathrm{S} \rightarrow \mathrm{R}$ channels are independent $\mathcal{CN}(0,1)$ block-fading with block length T, and the $\mathrm{R} \rightarrow \mathrm{D}$ link is static with $h_{\mathrm{RD}} = 1$. The latter assumption is again conceptual and can be relaxed in certain special cases (see the discussion on space–time block codes later in this section). The source terminal transmits elements of the first and second rows of the $2 \times T$ space–time codeword $\mathbf{C}$ serially over the direct and the relay-assisted channels, respectively. Stacking the signals received at the destination terminal to form a $T \times 1$ vector $\mathbf{Y}$, we obtain the following input–output relation:

$$\mathbf{Y}^T = \tilde{\mathbf{g}}^T \mathbf{\Lambda} \mathbf{C} + \tilde{\mathbf{n}}^T \tag{29}$$

where

$$\tilde{\mathbf{g}}^T = [h_{\mathrm{SR}} \, h_{\mathrm{SD}}], \quad \boldsymbol{\Lambda} = \begin{bmatrix} \frac{1}{\omega}\sqrt{\frac{E_{\mathrm{SR}}E_{\mathrm{RD}}}{E_{\mathrm{SR}}+N_o}} & 0 \\ 0 & \frac{\sqrt{E_{\mathrm{SD}}}}{\omega} \end{bmatrix}$$

$\omega^2 = 1 + E_{\mathrm{RD}}/(E_{\mathrm{SR}} + N_o)$ and $\tilde{\mathbf{n}}$ denotes a $T \times 1$ zero-mean circularly symmetric complex Gaussian noise vector with $\mathcal{E}\{\tilde{\mathbf{n}}\tilde{\mathbf{n}}^H\} = N_o\mathbf{I}_T$. Assuming maximum-likelihood (ML) decoding, the destination terminal constructs an estimate of the transmitted space–time codeword according to

$$\hat{\mathbf{C}} = \arg\min_{\mathbf{C}} \|\mathbf{Y}^T - \tilde{\mathbf{g}}^T\boldsymbol{\Lambda}\mathbf{C}\|^2 \tag{30}$$

where the minimization is performed over all possible codeword matrices $\mathbf{C}$. We recall that this decoding rule requires that the receiver not only knows the $\mathrm{S} \to \mathrm{D}$ channel, but also has perfect knowledge of the $\mathrm{S} \to \mathrm{R}$ channel. From (30), it follows immediately that the decoding complexity in the relay case is inherited from the underlying space–time code. For a given channel realization $\tilde{\mathbf{g}}$ the probability that a transmitted codeword $\mathbf{C}$ is mistaken for another codeword $\mathbf{E}$ is obtained as the pairwise error probability (PEP)

$$P(\mathbf{C} \to \mathbf{E}\,|\,\tilde{\mathbf{g}}) = Q\left(\sqrt{\frac{\|\tilde{\mathbf{g}}^T\boldsymbol{\Lambda}(\mathbf{C}-\mathbf{E})\|^2}{2N_o}}\right).$$

Applying the standard Chernoff bound and using a result from [34], we can upper bound the PEP averaged over all channel realizations, $P(\mathbf{C} \to \mathbf{E}) = \mathcal{E}_{\tilde{\mathbf{g}}}\{P(\mathbf{C} \to \mathbf{E}\,|\,\tilde{\mathbf{g}})\}$, as

$$P(\mathbf{C} \to \mathbf{E}) \leq \prod_{i=1}^{2} \frac{1}{1 + \frac{\lambda_i}{4N_o}} \tag{31}$$

where λ_i $(i = 1, 2)$ are the eigenvalues of the 2×2 matrix $\boldsymbol{\Lambda}(\mathbf{C}-\mathbf{E})(\mathbf{C}-\mathbf{E})^H\boldsymbol{\Lambda}$. Applying Ostrowski's theorem [35] to the matrix $\boldsymbol{\Lambda}(\mathbf{C}-\mathbf{E})(\mathbf{C}-\mathbf{E})^H\boldsymbol{\Lambda}$, the eigenvalues λ_i $(i = 1, 2)$ can be lower-bounded as

$$\lambda_i \geq \min\left\{\frac{E_{\mathrm{SD}}}{\omega^2}, \frac{1}{\omega^2}\frac{E_{\mathrm{SR}}E_{\mathrm{RD}}}{E_{\mathrm{SR}}+N_o}\right\}\gamma_i$$

where γ_i $(i = 1, 2)$ denotes the eigenvalues of $(\mathbf{C}-\mathbf{E})(\mathbf{C}-\mathbf{E})^H$. Consequently, the average PEP can be upper-bounded as

$$P(\mathbf{C} \to \mathbf{E}) \leq \prod_{i=1}^{2} \frac{1}{1 + \frac{\beta_3^{\mathrm{AF}}}{4}\gamma_i} \tag{32}$$

where β_3^{AF} was defined in (16). For $\gamma_i > 0$ $(i = 1, 2)$ and $\beta_3^{\mathrm{AF}} \gg 1$, we obtain

$$P(\mathbf{C} \to \mathbf{E}) \leq \left(\prod_{i=1}^{2}\gamma_i\right)^{-1}\left(\frac{\beta_3^{\mathrm{AF}}}{4}\right)^{-2} \tag{33}$$

which shows that relay-assisted communication using a space–time code achieving second-order diversity in the case of colocated antennas achieves second-order diversity in the effective SNR β_3^{AF}. This result conforms with the diversity behavior exhibited by the PER through the outage capacity analysis in Section III-C. Moreover, we note that the coding gain achieved in the relay case is equal to the coding gain achieved in the case of colocated antennas. Hence, optimum signal design for the relay case consists of constructing a

space–time code satisfying the classical rank and determinant criteria [2], as well as ensuring proper power control (diversity is achieved in the effective SNR β_3^{AF}).

Comments on orthogonal designs. In the case of orthogonal space–time block codes (OSTBCs) [3], [4], we can refine (33). Assuming Alamouti transmission, $(\mathbf{C}-\mathbf{E})(\mathbf{C}-\mathbf{E})^H = \gamma\mathbf{I}_2$. Therefore, the eigenvalues of $\boldsymbol{\Lambda}(\mathbf{C}-\mathbf{E})(\mathbf{C}-\mathbf{E})^H\boldsymbol{\Lambda}$ take on a particularly simple form

$$\lambda_1 = \frac{E_{\mathrm{SD}}}{\omega^2}\gamma, \quad \lambda_2 = \frac{E_{\mathrm{SR}}E_{\mathrm{RD}}}{\omega^2(E_{\mathrm{SR}}+N_o)}\gamma.$$

Consequently, the upper bound on $P(\mathbf{C} \to \mathbf{E})$ can be evaluated directly without applying Ostrowski's theorem to yield

$$P(\mathbf{C} \to \mathbf{E}) \leq \frac{1}{1 + \frac{E_{\mathrm{SD}}}{4N_o\omega^2}\gamma}\frac{1}{1 + \frac{E_{\mathrm{SR}}E_{\mathrm{RD}}}{4N_o\omega^2(E_{\mathrm{SR}}+N_o)}\gamma}$$

$$\leq \left(\frac{1}{1 + \frac{\beta_3^{\mathrm{AF}}}{4}\gamma}\right)^2. \tag{34}$$

For β_3^{AF} large, we get

$$P(\mathbf{C} \to \mathbf{E}) \leq \gamma^{-2}\left(\frac{\beta_3^{\mathrm{AF}}}{4}\right)^{-2}$$

which conforms with (33) and shows that second-order diversity in the effective SNR β_3^{AF} can indeed be achieved. Furthermore, (34) also illustrates the need for appropriate power control. For example, it is clear that simply increasing E_{SD}/N_o in (34) will decrease the PEP according to a first-order rather than a second-order diversity behavior. We conclude by noting that the Alamouti scheme can be applied to our setup without altering the simplified decoding procedure specified in [3].

PEP for fading $\mathrm{R} \to \mathrm{D}$ *link*. So far, we have considered the case where the $\mathrm{R} \to \mathrm{D}$ link is static. In certain cases, the average PEP for a fading $\mathrm{R} \to \mathrm{D}$ link $(h_{\mathrm{RD}} \sim \mathcal{CN}(0, 1))$ becomes analytically tractable. For example, continuing with the Alamouti scheme, the PEP for a given realization of h_{RD} (averaged over the $\mathrm{S} \to \mathrm{R}$ and $\mathrm{S} \to \mathrm{D}$ channels) can be upper-bounded as

$$P(\mathbf{C} \to \mathbf{E}\,|\,h_{\mathrm{RD}}) \leq \frac{1}{1 + \frac{E_{\mathrm{SD}}/(4N_o)}{1 + \frac{E_{\mathrm{RD}}|h_{\mathrm{RD}}|^2}{E_{\mathrm{SR}}+N_o}}\gamma}$$

$$\times \frac{1}{1 + \frac{E_{\mathrm{SR}}E_{\mathrm{RD}}|h_{\mathrm{RD}}|^2/(4(E_{\mathrm{SR}}+N_o)N_o)}{1 + \frac{E_{\mathrm{RD}}|h_{\mathrm{RD}}|^2}{E_{\mathrm{SR}}+N_o}}\gamma}.$$

Now, assuming $E_{\mathrm{SR}}/N_o \gg 1$ and $E_{\mathrm{SR}} \gg E_{\mathrm{RD}}$, we have $\beta_3^{\mathrm{AF}} \approx \min\{E_{\mathrm{SD}}/N_o, E_{\mathrm{RD}}/N_o\}$. For $\beta_3^{\mathrm{AF}} \gg 1$ the average PEP, $P(\mathbf{C} \to \mathbf{E}) = \mathcal{E}_{h_{\mathrm{RD}}}\{P(\mathbf{C} \to \mathbf{E}\,|\,h_{\mathrm{RD}})\}$, can be upper-bounded as

$$P(\mathbf{C} \to \mathbf{E}) \leq \left(\frac{\gamma\beta_3^{\mathrm{AF}}}{4}\right)^{-2}\Gamma\left(0, \frac{1}{\gamma\beta_3^{\mathrm{AF}}/4}\right)\exp\left(\frac{1}{\gamma\beta_3^{\mathrm{AF}}/4}\right) \tag{35}$$

where $\Gamma(a, x)$ is the *incomplete gamma function* defined as $\Gamma(a, x) = \int_x^{\infty} t^{a-1}e^{-t}\,dt$. Using the series representation [36]

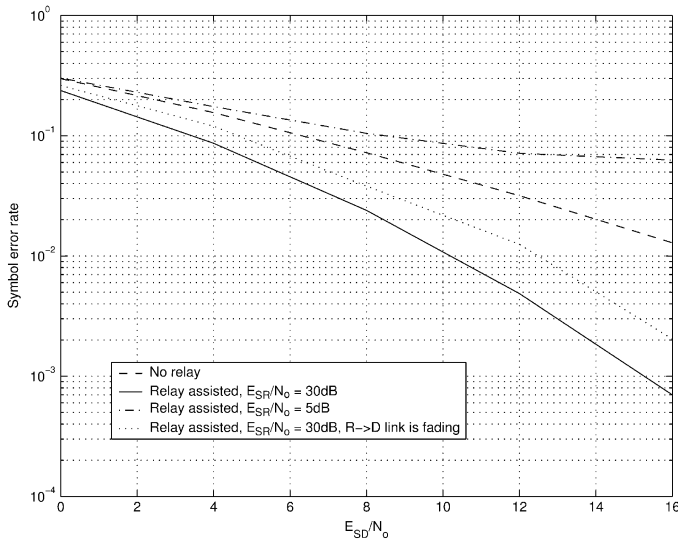$$\Gamma(0, x) = e^{-x}\sum_{n=0}^{\infty}\frac{L_n(x)}{n+1}$$

Fig. 5. Uncoded symbol-error rate for (distributed) Alamouti scheme as a function of E_{SD}/N_o for Protocol III in the AF mode (unless otherwise specified the $R \to D$ link was assumed AWGN). Second-order diversity is achieved under suitable channel conditions. Fading on the $R \to D$ link results in a coding gain loss.

where $L_n(x) = \sum_{m=0}^n (-1)^m (n!x^m)/(m!(n-m)!m!)$ is the Laguerre polynomial of order n it can be shown that asymptotically in β_3^{AF} the RHS of (35) decays proportional to $(\beta_3^{AF})^{-2}$. Moreover, since $\Gamma(0, 1/(\gamma\beta_3^{AF}/4)) \exp(1/(\gamma\beta_3^{AF}/4)) > 1$ for β_3^{AF} large, we can conclude that second-order diversity is achieved and the fading nature of h_{RD} incurs a coding gain loss only. These observations are verified in the simulation result below.

Numerical results. In the following simulation example, we assume Protocol III in the AF mode employing the Alamouti scheme as described above. Fig. 5 depicts the uncoded symbol-error rate (Monte Carlo simulation) with and without the assistance of the relay as a function of E_{SD}/N_o for a 4-QAM constellation. We assume $E_{SD} = E_{RD}$, i.e., the $S \to D$ and $R \to D$ links are balanced (this can be achieved through power control) and for now that the $R \to D$ link is static, i.e., $h_{RD} = 1$. The uncoded symbol-error rate for relay assisted communication is shown for two different values of E_{SR}/N_o. Note that for $E_{SR}/N_o = 30$ dB, $\omega \approx 1$ over the plotted range of E_{SD}/N_o. Furthermore, for $\omega \approx 1$ and $E_{SD} = E_{RD}, \beta_3^{AF} = (E_{SD}/N_o)\min\{1, (E_{SR}/N_o)/(1 + E_{SR}/N_o)\}$ which for E_{SR}/N_o large yields $\beta_3^{AF} \approx E_{SD}/N_o$. Hence, for E_{SR}/N_o large relay-assisted communication indeed achieves second-order diversity in E_{SD}/N_o as reflected in Fig. 5 by the slope of symbol-error probability as a function of E_{SD}/N_o. However, performance degrades dramatically as the $S \to R$ link deteriorates to $E_{SR}/N_o = 5$ dB. We observe an error flooring effect, which can be attributed to the amplified noise received at the destination terminal through the $R \to D$ link. This noise amplification offsets any gain resulting from using the relay channel so that using the direct link only yields superior performance. In fact, for E_{SR}/N_o sufficiently small and E_{RD}/N_o large the effective SNR is given by $\beta_3^{AF} = E_{SR}/N_o$, which shows that the link performance is governed by E_{SR}/N_o. Finally, for the sake of comparison, we plot on the same graph the symbol-error rate for the Alamouti

scheme for the case where the $R \to D$ link is Rayleigh fading and $E_{SR}/N_o = 30$ dB. The simulation result conforms with the discussion and observations made subsequent to (35) showing that the error rate decays according to a second-order diversity behavior with a coding gain loss.

We conclude this section by noting that in a similar manner, it is easy to show that space–time codes designed for the colocated antenna case are capable of extracting full spatial diversity gain under Protocol I. In fact, it is straightforward to see that the distributed Alamouti scheme discussed above, when applied to Protocol I will realize second-order diversity, while extracting additional coding gain (due to the increased channel energy captured over the first time slot).

VI. CONCLUSION

We studied three different TDMA-based cooperative protocols for a simple fading relay channel with AF and DF modes of relaying. For each of the protocols, assuming Gaussian code books, we derived the ergodic and outage capacities and established the importance of maximizing the degree of broadcasting and receive collision. We analyzed the diversity performance of the different protocols through an outage probability analysis. The corresponding results indicate that full spatial diversity (second-order in this case) is achieved by certain protocols provided that appropriate power control is employed. Finally, we considered space–time code design for AF-based relay channels and found that the code design criteria for the relay case consist of the traditional rank and determinant criteria for colocated antennas combined with appropriate power control rules. These power control rules were found to be the same as those arising in the diversity performance analysis. Our results show that space–time codes designed for the case of colocated antennas can be used to realize cooperative diversity provided that appropriate power control is employed. We conclude by noting that the idea of mapping cooperative protocols onto effective point-to-point MIMO channels can be easily extended to larger networks and more complex transmission schemes. The resulting effective MIMO channels will have larger dimensionality and significantly different statistics compared with the classical i.i.d. Gaussian fading channel.

REFERENCES

[1] J. Guey, M. P. Fitz, M. R. Bell, and W. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," in *Proc. IEEE Vehicular Technology Conf.*, vol. 1, Atlanta, GA, Apr./May 1996, pp. 136–140.

[2] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space–time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[4] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space–time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.

[5] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—Part I: System description," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1938, Nov. 2003.

[6] ——, "User cooperation diversity—Part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[7] J. N. Laneman, G. W. Wornell, and D. N. C. Tse, "An efficient protocol for realizing cooperative diversity in wireless networks," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Washington, DC, June 2001, p. 294.

[8] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, submitted for publication.

[9] A. Sendonaris, E. Erkip, and B. Aazhang, "Increasing uplink capacity via user cooperation diversity," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Cambridge, MA, Aug. 1998, p. 156.

[10] J. N. Laneman and G. W. Wornell, "Energy-efficient antenna sharing and relaying for wireless networks," in *Proc. IEEE Wireless Communications Networking Conf. (WCNC)*, vol. 1, Chicago, IL, Sept. 2000, pp. 7–12.

[11] A. Stefanov and E. Erkip, "Cooperative coding for wireless networks," in *Proc. Int. Workshop Mobile Wireless Communications Networks*, Stockholm, Sweden, Sept. 2002, pp. 273–277.

[12] T. E. Hunter and A. Nosratinia, "Cooperative diversity through coding," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Lausanne, Switzerland, June 2002, p. 220.

[13] A. Host-Madsen, "On the capacity of wireless relaying," in *Proc. IEEE Vehicular Technology Conf.*, vol. 3, Vancouver, Canada, Sept. 2002, pp. 1333–1337.

[14] J. N. Laneman and G. W. Wornell, "Distributed space–time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2415–2425, Oct. 2003.

[15] R. Ahlswede, N. Cai, S. Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1204–1216, July 2000.

[16] N. Cai and R. W. Yeung, "Secure network coding," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Lausanne, Switzerland, June 2002, p. 323.

[17] R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Trans. Networking*, vol. 11, pp. 782–795, Oct. 2003.

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

[19] M. Gastpar, G. Kramer, and P. Gupta, "The multiple-relay channel: Coding and antenna clustering capacity," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Lausanne, Switzerland, July 2002, p. 136.

[20] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, pp. 388–404, Mar. 2002.

[21] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: The relay case," in *Proc. IEEE INFOCOM*, vol. 3, New York, June 2002, pp. 1577–1586.

[22] R. U. Nabar and H. Bölcskei, "Space–time signal design for fading relay channels," in *Proc. IEEE GLOBECOM*, vol. 4, San Francisco, CA, Dec. 2003, pp. 1952–1956.

[23] A. J. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmission/directional reception," U.S. Patent 5 345 599, 1994.

[24] G. J. Foschini, "Layered space–time architecture for wireless communication in a fading environment when using multielement antennas," *Bell Labs. Tech. J.*, pp. 41–59, 1996.

[25] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Tel.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[26] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, pp. 225–234, Feb. 2002.

[27] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol*, vol. 43, pp. 359–378, May 1994.

[28] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.

[29] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[30] R. G. Gallager, "An inequality on the capacity region of multiaccess multipath channels," in *Communications and Cryptography—Two Sides of One Tapestry*, R. E. Blahut, D. J. Costello Jr., U. Maurer, and T. Mittelholzer, Eds. Norwell, MA: Kluwer, 1994, pp. 129–139.

[31] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1073–1096, May 2003.

[32] B. Suard, G. Xu, and T. Kailath, "Uplink channel capacity of space-division-multiple-access schemes," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1468–1476, July 1998.

[33] S. Shamai and A. Wyner, "Information-theoretic considerations for symmetric cellular, multiple-access fading channels—Part II," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1895–1911, Nov. 1997.

[34] G. Turin, "The characteristic function of Hermitian quadratic forms in complex normal variables," *Biometrika*, vol. 47, pp. 199–201, 1960.

[35] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[36] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, 5th ed. Orlando, FL: Academic, 1994.

Rohit U. Nabar (S'02–M'03) was born in Bombay, India, on December 18, 1976. He received the B.S. degree (*summa cum laude*) in electrical engineering from Cornell University, Ithaca, NY, in 1998 and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 2000, and 2003, respectively. His doctoral research focused on signaling for real-world MIMO channels.

He is currently a Postdoctoral Research Assistant with the Communication Theory Group, Swiss Federal Institute of Technology (ETH), Zürich, Switzerland. His research interests include signal processing and information theory for wireless communications, MIMO wireless systems, and ad hoc wireless networks.

Dr. Nabar was a recipient of the Dr. T. J. Rodgers Stanford Graduate Fellowship.
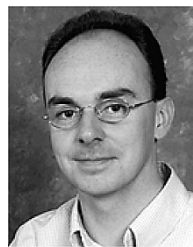
Helmut Bölcskei (S'94–A'97–SM'02) was born in Austria on May 29, 1970. He received the Dr. Techn. degree in electrical engineering from the Vienna University of Technology, Vienna, Austria, in 1997.

In 1998, he was a Postdoctoral Researcher at Vienna University of Technology. From 1999 to 2001, he was a Postdoctoral Researcher in the Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA. During that period, he was also a Consultant for Iospan Wireless, Inc. From 2001 to 2002, he was an Assistant Professor of Electrical Engineering at the University of Illinois at Urbana-Champaign. Since February 2002, he has been an Assistant Professor of Communication Theory at the Swiss Federal Institute of Technology (ETH), Zürich, Switzerland. He was a Visiting Researcher at Philips Research Laboratories, Eindhoven, The Netherlands, ENST, Paris, France, and the Heinrich Hertz Institute, Berlin, Germany. His research interests include communication and information theory with special emphasis on wireless communications.

Dr. Bölcskei received a 2001 IEEE Signal Processing Society Young Author Best Paper Award and was an Erwin Schrödinger Fellow (1999–2001) of the Austrian National Science Foundation (FWF). He is an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the *EURASIP Journal on Applied Signal Processing*.

Felix W. Kneubühler (S'03) was born in Basel, Switzerland. He received the Ing. Dipl. degree in communication systems from Ecole Polytechnique Federale de Lausanne (EPFL), Lausanne, Switzerland, in 2002. During his undergraduate studies, he was an exchange student at Carnegie Mellon University, Pittsburgh, PA, and the Institute Eurecom, Sophia-Antipolis, France. He is currently working toward the Dr. Sc. Techn. degree in the Communication Theory Group, Swiss Federal Institute of Technology (ETH) Zürich, Switzerland.

His research interests include communication and information theory for wireless communications, multiuser communications, and ad hoc networks.

Mr. Kneubühler was a recipient of the Eurecom Award for the Best Student in Mobile Communications.