Achievable Rates in Low-Power Relay Links Over Fading Channels

Xiaodong Cai, *Member, IEEE*, Yingwei Yao, *Member, IEEE*, and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—Relayed transmissions enable low-power communications among nodes (possibly separated by a large distance) in wireless networks. Since the capacity of general relay channels is unknown, we investigate the achievable rates of relayed transmissions over fading channels for two transmission schemes: the block Markov coded and the time-division multiplexed (TDM) transmissions. The normalized achievable minimum energy per bit required for reliable communications is derived, which also enables optimal power allocation between the source and the relay. The time-sharing factor in TDM transmissions is optimized to improve achievable rates. The region where relayed transmission can provide a lower minimum energy per bit than direct transmission, as well as the optimal relay placement for these two transmission schemes, are also investigated. Numerical results delineate the advantages of relayed, relative to direct, transmissions.**

***Index Terms*—Fading channels, low-power transmissions, minimum energy per bit, relay channels, spectral efficiency.**

I. INTRODUCTION

AD HOC connectivity is an integral part of many emerging applications, including wireless sensor and home networks [1]. Unlike well-structured wireless cellular systems, where all communications are controlled by the base station, ad hoc wireless networks do not have an established infrastructure. Without an inherent infrastructure, nodes in ad hoc wireless networks handle the control of communications by themselves. Peer-to-peer communications are enabled either through a direct link, or via relayed transmissions, where intermediate nodes serve as relays to send packets toward their final destination. Relayed transmissions not only improve error-probability performance [2]–[6], but also have the potential to increase the capacity of ad hoc wireless networks [7]. Since relayed transmissions can mitigate the effects of path loss, they can save transmit power, and also reduce interference among nodes. This allows for frequency reuse, which increases the network capacity. Relayed transmissions also provide a means of effecting diversity to combat channel fading introduced by the wireless interface. A node can transmit its signals directly and through multiple relay paths to its destination. Since the destination receives signals transmitted through multiple independent paths, relaying introduces a form of diversity that is known as cooperative diversity [8]–[11], because it is enabled via cooperating relay nodes.

Although the capacity of the Gaussian degraded relay channel was obtained in [12], only an upper bound on the capacity and achievable rates of general relay channels were derived in [12]. The capacity and achievable rates of Gaussian relay channels with multiple relays were studied in [13]–[17]; and the achievable rates of Gaussian relay channels with time or frequency-division multiplexed (TDM/FDM) transmissions were considered in [18]–[20]. Ergodic capacity of relay links over multi-input multi-output (MIMO) fading channels was also investigated recently in [21]. Since relayed transmissions can considerably save the transmit power, while enabling communications between nodes separated by a large distance, they are particularly attractive for low-power communications, such as those encountered with wireless sensor networks, where small, battery-powered sensors are deployed at each node. Upper and lower bounds on the minimum energy per bit required for reliable communications over Gaussian relay channels were derived in [22].

In this paper, we investigate the achievable rates of relayed transmissions over *fading* channels in the low-power regime, using the tools developed in [23]. We consider the block Markov coding (BMC) introduced by [12], as well as TDM transmissions, and study the spectral efficiency of these two schemes over wireless fading links. We derive the normalized achievable minimum energy per bit, $(E_b/N_0)_{\min}$, required for reliable communications, which, in turn, allows us to optimize the power allocated between the source and the relay. We also optimize the time-sharing factor in TDM transmissions to improve spectral efficiency, and select the relay placement optimally to minimize the $(E_b/N_0)_{\min}$.

The rest of the paper is organized as follows. In Section II, we first derive a lower bound on the $(E_b/N_0)_{\min}$, and then obtain the achievable $(E_b/N_0)_{\min}$, the spectral efficiency of BMC, and the optimal relay placement. TDM transmission schemes are investigated in Section III; and numerical results are presented in Section IV. Finally, we draw summarizing conclusions in Section V.

Paper approved by H. Leib, the Editor for Communication and Information Theory of the IEEE Communication Society. Manuscript received October 7, 2003; revised July 28, 2004. This paper was prepared through collaborative participation in the Communications and Networks Consortium sponsored by the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. This paper was presented in part at the 38th Conference on Information Sciences and Systems, Princeton University, NJ, March 2004.

X. Cai was with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA. He is now with the Department of Electrical and Computer Engineering, University of Miami, Coral Gables, FL 33124 USA (e-mail: x.cai@miami.edu).

Y. Yao was with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA. He is now with the Department of Electrical and Computer Engineering, University of Illinois, Chicago, IL 60607 USA (e-mail: yyao@ece.uic.edu).

G. B. Giannakis is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: georgios@ece.umn.edu).

Digital Object Identifier 10.1109/TCOMM.2004.840634

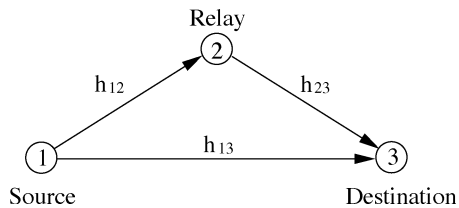


Fig. 1. System model of relayed transmissions.

II. Block Markov Transmissions

Fig. 1 depicts transmissions over a relay channel, where the source (node 1) encodes the incoming information-bearing bits, and transmits the resulting codewords. The relay (node 2) transmits a codeword based on the received signal from the source; the destination (node 3) jointly decodes the received signals from both source and relay. We consider frequency-flat fading channels, and denote the channel coefficient between nodes i and j as h_{ij}. Throughout the paper, we assume that the channel coefficients $\{h_{ij}\}$ are ergodic random processes. In practice, nodes are surrounded by many local scatterers, and the distance between any two nodes is much larger than the carrier wavelength; hence, the corresponding channels are uncorrelated. We also assume that h_{12} is known at the relay perfectly, and likewise h_{13} and h_{23} at the destination. The capacity of the Gaussian degraded relay channel was derived in [12]; however, the capacity of general relay channels is unknown. Since we are interested in low-power transmissions, instead of analyzing the channel capacity for each signal-to-noise ratio (SNR) value, we will use the tools developed in [23] to study the performance of relayed transmissions over *fading* channels in the low-power regime.

The two key performance measures in the low-power regime are the normalized minimum energy per bit $(E_b/N_0)_{\min}$ required for reliable communication, and the slope of the spectral efficiency at the point $(E_b/N_0)_{\min}$ [23]. If the channel capacity is denoted by $C(\text{SNR})$ nats/dimension, and the derivative of $C(\text{SNR})$ at SNR $= 0$ is $\dot{C}(0)$, then the $(E_b/N_0)_{\min}$ is given by [23]

$$\left(\frac{E_b}{N_0}\right)_{\min} = \frac{\log_e(2)}{\dot{C}(0)} \tag{1}$$

and the slope of spectral efficiency in bits/seconds/Hertz/(3 dB) at $(E_b/N_0)_{\min}$ is expressed as [23]

$$\mathcal{S} = \frac{2\left[\dot{C}(0)\right]^2}{-\ddot{C}(0)} \tag{2}$$

where $\ddot{C}(0)$ denotes the second derivative of $C(\text{SNR})$ at SNR $= 0$.

In this section, after deriving a lower bound on the $(E_b/N_0)_{\min}$, we will investigate the achievable rates based on the BMC scheme introduced by [12]. This will lead us to optimal relay selection.

A. Lower Bound on $(E_b/N_0)_{\min}$

Let us denote the transmitted symbols from the source and from the relay as $\sqrt{P_1}x_1$ and $\sqrt{P_2}x_2$, respectively. We assume that $E[|x_1|^2] = E[|x_2|^2] = 1$, where $E[\cdot]$ denotes the expectation over the random variables in parentheses, and let the transmit power at the source and the relay be P_1 and P_2, respectively. The received symbols at the relay and the destination can thus be written, respectively, as

$$\begin{aligned} y_2 &= \sqrt{P_1}h_{12}x_1 + n_2 \\ y_3 &= \sqrt{P_1}h_{13}x_1 + \sqrt{P_2}h_{23}x_2 + n_3 \end{aligned} \tag{3}$$

where n_2 and n_3 denote additive white Gaussian noise (AWGN) with zero mean and variance N_0. It has been shown in [12] that the capacity C of the relay channel, depicted in Fig. 1, is upper bounded as

$$C \le \sup_{p(x_1,x_2)} \min\left\{I(x_1; y_2, y_3|x_2),\ I(x_1, x_2; y_3)\right\} \tag{4}$$

where $p(x_1, x_2)$ is the joint probability distribution of x_1 and x_2. Let $P := P_1 + P_2$, $\beta := P_1/P$, SNR $:= P/N_0$, and $\rho := E[x_1 x_2^*]$. Based on (3) and applying the ergodic capacity expressions of, e.g., [24] for fading channels, we obtain

$$\begin{aligned} C_1 &:= \sup_{p(x_1,x_2)} I(x_1; y_2, y_3|x_2) \\ &= E\left[\log\left(1 + \beta\left(1 - |\rho|^2\right)\left(|h_{12}|^2 + |h_{13}|^2\right)\text{SNR}\right)\right] \\ C_2 &:= \sup_{p(x_1,x_2)} I(x_1, x_2; y_3) \\ &= E\Big[\log\Big(1 + \Big(\beta|h_{13}|^2 + 2\Re\left(\rho h_{13}h_{23}^*\right)\sqrt{\beta(1-\beta)} \\ &\qquad + (1-\beta)|h_{23}|^2\Big)\text{SNR}\Big)\Big] \end{aligned} \tag{5}$$

where supremum is taken over the set of $p(x_1, x_2)$ with fixed β and ρ, $\Re(\cdot)$ denotes the real part of the variable in parentheses, the superscript * denotes conjugation, C_1 and C_2 are achieved when x_1 and x_2 are jointly Gaussian distributed, and the units of C_1 and C_2 are nats/s/Hz.

The capacity of Gaussian relay channels was studied with a fixed β in [12]. The optimal power allocation specified by an optimal β was derived in [16] for the Gaussian degraded channel. For the fading channels under consideration, similar to [16], we will optimize the power allocation to achieve optimal performance, based on certain partial channel state information (CSI) available at the transmitter, specified in the following assumption.

AS1: Source and relay know the first-, second-, and fourth-order moments of h_{12}, h_{13}, and h_{23}.

With this partial CSI available at the source and the relay, we have

$$C \le \max_{\rho,\beta} \min\{C_1, C_2\}. \tag{6}$$

Due to the expectation operation involved in C_1 and C_2, it is difficult to find the optimal value of β which maximizes transmission rate for each SNR value. Since we are interested in low-power communications, we will instead pursue a lower bound on the $(E_b/N_0)_{\min}$ based on (5), (6), and *AS1*.

Targeting (1) for our problem, the first derivatives of C_1 and C_2 at SNR $= 0$ can be found from (5) as

$$\begin{aligned} \dot{C}_1(0, \rho, \beta) &= \beta(\bar{\alpha}_{12} + \bar{\alpha}_{13})(1 - |\rho|^2) \\ \dot{C}_2(0, \rho, \beta) &= \bar{\alpha}_{13}\beta + \bar{\alpha}_{23}(1 - \beta) \\ &\quad + 2\Re\left(\rho\bar{h}_{13}\bar{h}_{23}^*\right)\sqrt{\beta(1-\beta)} \end{aligned} \tag{7}$$

where $\bar{\alpha}_{ij} := E[|h_{ij}|^2]$, and $\bar{h}_{ij} := E[h_{ij}]$. Since $C_1 = C_2 = 0$ at SNR $= 0$, from (1) and (6), a lower bound on the $(E_b/N_0)_{\min}$ can be written as

$$\left(\frac{E_b}{N_0}\right)_{\min}^{(\text{lb})} = \min_{\rho,\beta} \max\left\{\frac{\log_e(2)}{\dot{C}_1(0,\rho,\beta)}, \frac{\log_e(2)}{\dot{C}_2(0,\rho,\beta)}\right\}. \quad (8)$$

It is seen from (8) that finding $(E_b/N_0)_{\min}^{(\text{lb})}$ is equivalent to maximizing the $\min\{\dot{C}_1(0,\rho,\beta), \dot{C}_2(0,\rho,\beta)\}$.

From (7), we deduce that for $\dot{C}_2(0,\rho,\beta)$ to be maximized, the phase of ρ should equal that of $\bar{h}_{13}^*\bar{h}_{23}$; and then, we can express $\dot{C}_2(0,\rho,\beta)$ as

$$\dot{C}_2(0,\rho,\beta) = \bar{\alpha}_{13}\beta + \bar{\alpha}_{23}(1-\beta) + 2\left|\bar{h}_{13}\bar{h}_{23}^*\right| |\rho| \sqrt{\beta(1-\beta)}. \quad (9)$$

It is clear that $\dot{C}_1(0,\rho,\beta)$ is a monotonically decreasing function of $|\rho|$ since $|\rho| \in [0\ 1]$, and that $\dot{C}_2(0,\rho,\beta)$ is a linearly increasing function of $|\rho|$. In order to choose $|\rho|$ to maximize the $\min\{\dot{C}_1(0,\rho,\beta), \dot{C}_2(0,\rho,\beta)\}$, we need to consider two cases. First, if $\dot{C}_2(0,0,\beta) \geq \dot{C}_1(0,0,\beta)$, or equivalently, if $\bar{\alpha}_{23}(1-\beta) \geq \beta\bar{\alpha}_{12}$, we have $\rho = 0$. Second, if $\dot{C}_2(0,0,\beta) < \dot{C}_1(0,0,\beta)$, since $\dot{C}_2(0,1,\beta) \geq \dot{C}_1(0,1,\beta) = 0$, $\dot{C}_1(0,\rho,\beta)$ and $\dot{C}_1(0,\rho,\beta)$ have a cross point over the interval $|\rho| \in [0,1]$; and then we can find $|\rho|$ by setting $\dot{C}_1(0,\rho,\beta) = \dot{C}_2(0,\rho,\beta)$, which yields the equation shown at bottom of the page. Note that condition $\dot{C}_2(0,0,\beta) < \dot{C}_1(0,0,\beta)$ guarantees that $|\rho|$ in (10) satisfies $0 < |\rho| < 1$. After finding ρ, we can use a one-dimensional search to obtain the optimal value of β that minimizes the $(E_b/N_0)_{\min}^{(\text{lb})}$.

When $\bar{h}_{13} = 0$ or/and $\bar{h}_{23} = 0$, we can obtain the optimal β and the corresponding $(E_b/N_0)_{\min}^{(\text{lb})}$ in closed form. In this case, it is apparent that $\rho = 0$ is optimal, because the third term in $\dot{C}_2(0,\rho,\beta)$ is equal to zero, and $\dot{C}_1(0,\rho,\beta)$ is maximized at $\rho = 0$ [c.f. (7)]. With $\rho = 0$, $\dot{C}_1(0,\rho,\beta)$ and $\dot{C}_2(0,\rho,\beta)$ in (7) become

$$\begin{aligned}\dot{C}_1(0,0,\beta) &= (\bar{\alpha}_{12} + \bar{\alpha}_{13})\beta \\ \dot{C}_2(0,0,\beta) &= \bar{\alpha}_{23} + (\bar{\alpha}_{13} - \bar{\alpha}_{23})\beta. \quad (11)\end{aligned}$$

If $\bar{\alpha}_{23} \leq \bar{\alpha}_{13}$, it is clear from (11) that $\beta = 1$ is optimal, which implies that relayed transmission cannot provide a lower $(E_b/N_0)_{\min}$ than the direct link; and we have $(E_b/N_0)_{\min} = \log_e(2)/\bar{\alpha}_{13}$ for the direct link. If $\bar{\alpha}_{23} > \bar{\alpha}_{13}$, $\dot{C}_2(0,0,\beta)$ is a decreasing function of β, while $\dot{C}_1(0,0,\beta)$ is an increasing function of β; and thus, $\min\{\dot{C}_1, \dot{C}_2\}$ is maximized, when $\dot{C}_1(0,0,\beta) = \dot{C}_2(0,0,\beta)$. Setting $\dot{C}_1(0,0,\beta) = \dot{C}_2(0,0,\beta)$, we obtain

$$\beta_{\text{opt}} = \frac{\bar{\alpha}_{23}}{\bar{\alpha}_{12} + \bar{\alpha}_{23}} \quad (12)$$

if $\bar{\alpha}_{23} > \bar{\alpha}_{13}$, and

$$\left(\frac{E_b}{N_0}\right)_{\min}^{(\text{lb})} = \frac{\log_e(2)(\bar{\alpha}_{12} + \bar{\alpha}_{23})}{\bar{\alpha}_{23}(\bar{\alpha}_{12} + \bar{\alpha}_{13})} \quad (13)$$

if $\bar{h}_{13} = 0$ or/and $\bar{h}_{23} = 0$, and $\bar{\alpha}_{23} > \bar{\alpha}_{13}$. Note that our $(E_b/N_0)_{\min}$ for fading channels depends on the first- and second-order moments of channel coefficients, while $(E_b/N_0)_{\min}$ of [22] for Gaussian channels depends on channel coefficients. Furthermore, perfect CSI is required at the source and the relay in [22], whereas we only need the first- and second-order moments of the channel coefficients at the source and the relay for optimal power allocation.

It was shown in [12] that for the Gaussian degraded channel, the upper bound in (4) is actually the channel capacity; however, it is unknown whether this upper bound is achievable for general relay channels. Hence, it is not clear if we can achieve the lower bounds on the $(E_b/N_0)_{\min}$ found using numerical search when $\bar{h}_{13} \neq 0$ and $\bar{h}_{23} \neq 0$, and from (13) for the case where $\bar{h}_{13} = 0$ or/and $\bar{h}_{23} = 0$. In the next subsection, we will study the achievable $(E_b/N_0)_{\min}$ and the slope of the spectral efficiency for the BMC advocated in [12].

B. Achievable Rates of BMC

The information-bearing bit stream (message) at the source is parsed into blocks, each with m symbols. Let $w_i \in [1, 2^{mR}]$ be the message to be sent by the source during the ith block. The set $\mathcal{M} = \{1, 2, \cdots, 2^{mR}\}$ of messages is randomly partitioned into $\mathcal{Q} = \{Q_1, Q_2, \cdots, Q_{2^{mR_0}}\}$ with $R_0 < R$, as in [12]. A random code book $\mathcal{X} = \{\mathbf{x}_1(w|q), \mathbf{x}_2(q); w \in [1, 2^{mR}], q \in [1, 2^{mR_0}]\}$ is generated based on a joint probability distribution $p(x_1, x_2)$ [12]. We suppose that entries of codewords $\mathbf{x}_1(w|q)$ and $\mathbf{x}_2(q)$ are independent, identically distributed, and obey a Gaussian distribution with zero mean and unit variance. After the relay correctly decodes the message from the source during the $(i-1)$st block, it transmits a codeword $\mathbf{x}_2(q_i)$ in the ith block. We refer the reader to [12] for a detailed description of the block Markov encoding and decoding. The signal vectors transmitted by the source and the relay simultaneously in the ith block are given, respectively, by

$$\begin{aligned}\tilde{\mathbf{x}}_1(i) &= \sqrt{P_1}\mathbf{x}_1(w_i|q_i) + b\sqrt{(1-\beta_2)P_2}\mathbf{x}_2(q_i) \\ \tilde{\mathbf{x}}_2(i) &= \sqrt{\beta_2 P_2}\mathbf{x}_2(q_i). \quad (14)\end{aligned}$$

Notice that different from (3), P_1 and P_2 are the transmit powers of $\mathbf{x}_1(w_i|q_i)$ and $\mathbf{x}_2(q_i)$. With $\beta_2 \in (0,1)$ determining the fraction of power P_2 allocated to the relay, P_1 and P_2 are not equal to the transmit power at the source and relay, if $\beta_2 \neq 0$. The constant b satisfying $|b| = 1$ aims to improve the combining of $\mathbf{x}_2(q_i)$ from the relay and the source at the destination. The received blocks at the relay and the destination can be expressed, respectively, as

$$\begin{aligned}\mathbf{y}_2(i) &= h_{12}\tilde{\mathbf{x}}_1(i) + \mathbf{n}_2(i) \\ \mathbf{y}_3(i) &= h_{13}\tilde{\mathbf{x}}_1(i) + h_{23}\tilde{\mathbf{x}}_2(i) + \mathbf{n}_3(i) \quad (15)\end{aligned}$$

where $\mathbf{n}_2(i)$ and $\mathbf{n}_3(i)$ are AWGN blocks with zero mean and covariance $N_0\mathbf{I}_m$. From the rate formulas of BMC [12], we ob-

$$|\rho| = \frac{-\left|\bar{h}_{13}\bar{h}_{23}^*\right|\sqrt{\beta(1-\beta)}}{\beta(\bar{\alpha}_{12} + \bar{\alpha}_{13})} + \frac{\sqrt{\left|\bar{h}_{13}\bar{h}_{23}^*\right|^2 \beta(1-\beta) - (\bar{\alpha}_{12} + \bar{\alpha}_{13})\beta\left[\bar{\alpha}_{23}(1-\beta) - \bar{\alpha}_{12}\beta\right]}}{\beta(\bar{\alpha}_{12} + \bar{\alpha}_{13})} \quad (10)$$

tain the rate $R_1 = (1/m)I(\mathbf{x}_1; \mathbf{y}_2|\mathbf{x}_2)$ for the link between the source and the relay, and the rate $R_2 = (1/m)I(\mathbf{x}_1, \mathbf{x}_2; \mathbf{y}_3)$ for the multiple-access channel between the source, the relay, and the destination, where we omitted the indexes of the transmitted codewords and the received blocks; and the achievable rate of BMC is given by $R = \max\{R_1, R_2\}$ [12, eq. (12)]. Letting $P := P_1 + P_2$, $\beta_1 := P_1/P$, $\text{SNR} := P/N_0$, and supposing that the codewords are long enough to capture the ergodic nature of the fading processes, we obtain

$$R_1 = E\left[\log\left(1 + \beta_1|h_{12}|^2\text{SNR}\right)\right] \tag{16}$$

and

$$R_2 = E\left[\log\left(1 + \left(\beta_1|h_{13}|^2 + (1-\beta_1) \times \left|\sqrt{\beta_2}h_{23} + b\sqrt{(1-\beta_2)}h_{13}\right|^2\right)\text{SNR}\right)\right]. \tag{17}$$

The transmission rate between the source and the destination can be found as

$$R = \max_{\beta_1,\beta_2,b}\{R_1, R_2\}. \tag{18}$$

The rate R in (18) is not the capacity, but the achievable rate between the source and the destination for BMC. Because the optimal values of β_1, β_2, and b in (18) cannot be found in closed form, we will look for the triplet $\{\beta_1, \beta_2, b\}$ to minimize the *achievable* minimum energy per bit. With slight misuse of notation, we still denote this achievable normalized minimum energy per bit as $(E_b/N_0)_{\min}$. The first derivatives of R_1 and R_2 at $\text{SNR} = 0$ can be found from (16) and (17) as

$$\begin{aligned}\dot{R}_1(0) &= \beta_1\bar{\alpha}_{12}\\ \dot{R}_2(0) &= \beta_1\bar{\alpha}_{13} + (1-\beta_1)E\left[\left|\sqrt{\beta_2}h_{23} + b\sqrt{1-\beta_2}h_{13}\right|^2\right].\end{aligned} \tag{19}$$

The $(E_b/N_0)_{\min}$ can then be written as [c.f. (1)]

$$\left(\frac{E_b}{N_0}\right)_{\min} = \min_{\beta_1,\beta_2,b}\max\left\{\frac{\log_e(2)}{\dot{R}_1(0)}, \frac{\log_e(2)}{\dot{R}_2(0)}\right\} \tag{20}$$

and finding $(E_b/N_0)_{\min}$ is equivalent to maximizing the $\min\{\dot{R}_1(0), \dot{R}_2(0)\}$. Let us denote part of the second term of $\dot{R}_2(0)$ in (19) as

$$\begin{aligned}f(\beta_2, b) &= E\left|\sqrt{\beta_2}h_{23} + b\sqrt{1-\beta_2}h_{13}\right|^2\\ &= \beta_2\bar{\alpha}_{23} + (1-\beta_2)\bar{\alpha}_{13} + 2\sqrt{\beta_2(1-\beta_2)}\Re\left(b\bar{h}_{13}\bar{h}_{23}^*\right).\end{aligned} \tag{21}$$

It is apparent from (21) that

$$b = \frac{\bar{h}_{13}^*\bar{h}_{23}}{\left|\bar{h}_{13}^*\bar{h}_{23}\right|} \tag{22}$$

so that $f(\beta_2, b)$ is maximized. With b given by (22), it is easy to show that the first-order derivative of $f(\beta_2, b)$ with respect to β_2, $f'(\beta_2, b)$, is a monotonically decreasing function over $\beta_2 \in [0\,1]$, and that $\lim_{\beta\to 0} f'(\beta_2, b) = \infty$ while $\lim_{\beta\to 1} f'(\beta_2, b) = -\infty$. Hence, it follows readily that $f(\beta_2, b)$ has a unique maximum at

$$\beta_2 = \frac{1}{2}\left[1 + \frac{\bar{\alpha}_{23} - \bar{\alpha}_{13}}{\sqrt{4\left|\bar{h}_{23}^*\bar{h}_{13}\right|^2 + (\bar{\alpha}_{23} - \bar{\alpha}_{13})^2}}\right] \tag{23}$$

and the maximum value of $f(\beta_2, b)$ is given by

$$f_{\max} = \frac{1}{2}(\bar{\alpha}_{23} + \bar{\alpha}_{13}) + \frac{1}{2}\sqrt{(\bar{\alpha}_{23} - \bar{\alpha}_{13})^2 + 4\left|\bar{h}_{23}^*\bar{h}_{13}\right|^2}. \tag{24}$$

If $\bar{h}_{13} \neq 0$ and $\bar{h}_{23} \neq 0$, it can be readily shown that $f_{\max} > \max\{\bar{\alpha}_{13}, \bar{\alpha}_{23}\}$. Thus, with b and β_2 given by (22) and (23), $\dot{R}_2(0)$ in (19) is a decreasing function of β_1. On the other hand, $\dot{R}_1(0)$ in (19) is an increasing function of β_1. Hence, to maximize the $\min\{\dot{R}_1(0), \dot{R}_2(0)\}$, we must have $\dot{R}_1(0) = \dot{R}_2(0)$; and the optimal value of β_1 can be found using this relationship as

$$\beta_1 = \frac{f_{\max}}{\bar{\alpha}_{12} + f_{\max} - \bar{\alpha}_{13}}. \tag{25}$$

Using (25), the $(E_b/N_0)_{\min}$ can be obtained in closed form

$$\left(\frac{E_b}{N_0}\right)_{\min} = \frac{\log_e(2)(\bar{\alpha}_{12} + f_{\max} - \bar{\alpha}_{13})}{\bar{\alpha}_{12}f_{\max}}. \tag{26}$$

We now turn our attention to the case where $\bar{h}_{13} = 0$ or/and $\bar{h}_{23} = 0$. In this case, we can write $\dot{R}_2(0)$ in (19) as

$$\dot{R}_2(0) = \beta_1\bar{\alpha}_{13} + (1-\beta_1)\left[\beta_2\bar{\alpha}_{23} + (1-\beta_2)\bar{\alpha}_{13}\right]. \tag{27}$$

If $\bar{\alpha}_{23} \leq \bar{\alpha}_{13}$, it follows from (27) that the maximum value of $\dot{R}_2(0)$ is $\bar{\alpha}_{13}$, which implies that relayed transmission cannot provide a lower $(E_b/N_0)_{\min}$ than the direct link. From $\dot{R}_1(0)$ in (19), we deduce that we must have $\bar{\alpha}_{12} > \bar{\alpha}_{13}$ in order for the relay link to have a lower $(E_b/N_0)_{\min}$ than the direct link. Hence, when $\bar{h}_{13} = 0$ or/and $\bar{h}_{23} = 0$, the necessary condition for a relayed transmission with BMC to provide a lower $(E_b/N_0)_{\min}$ than the direct link is as follows.

C1: The expected values of $|h_{12}|^2$ and $|h_{23}|^2$ are greater than the expected value of $|h_{13}|^2$, i.e., $\bar{\alpha}_{12} > \bar{\alpha}_{13}$ and $\bar{\alpha}_{23} > \bar{\alpha}_{13}$.

Condition *C1* guarantees that the quality of the links, measured by the second moment of each channel coefficient between the source and the delay and between the relay and the destination, is better than that of the link between the source and the destination.

It is clear from (27) that under *C1*, $\dot{R}_2(0)$ is maximized when $\beta_2 = 1$, and the transmitted blocks in (14) reduce to

$$\tilde{\mathbf{x}}_1 = \sqrt{P_1}\mathbf{x}_1, \quad \tilde{\mathbf{x}}_2 = \sqrt{P_2}\mathbf{x}_2. \tag{28}$$

We will henceforth focus on this simpler transmission scheme, for which $\dot{R}_2(0)$ can be found as

$$\dot{R}_2(0) = \bar{\alpha}_{23} - (\bar{\alpha}_{23} - \bar{\alpha}_{13})\beta_1 \tag{29}$$

regardless of the values of $\bar{h}_{13}$ and $\bar{h}_{23}$, while $\dot{R}_1(0)$ is the same as that given in (19). When $\bar{h}_{13} \neq 0$ and $\bar{h}_{23} \neq 0$, this transmission scheme does not achieve the $(E_b/N_0)_{\min}$ in (26), as we will show later. However, analyzing this scheme provides useful insights on relayed transmissions, since in practical scenarios, it is likely that there is no line of sight (LOS) between the source and the destination, and thus, we have $\bar{h}_{13} = 0$. Using the transmitted signals given by (28) instead of by (14) when $\bar{h}_{13} = 0$ does not affect $(E_b/N_0)_{\min}$, but only reduces the complexity of the transmitter at the source node. Note that with this simpler transmission scheme, *C1* is also necessary for relayed transmissions to achieve a lower $(E_b/N_0)_{\min}$ than direct transmission when $\bar{h}_{13} \neq 0$ and $\bar{h}_{23} \neq 0$.

With the transmission signals given by (28), and under the condition *C1*, $\dot{R}_2(0)$ in (29) is a linearly decreasing function of β_1. Since $\dot{R}_1(0)$ is a linearly increasing function of β_1, the optimal value of β_1 can be found by setting $\dot{R}_1(0) = \dot{R}_2(0)$. This leads to

$$\beta_1 = \frac{\bar{\alpha}_{23}}{\bar{\alpha}_{12} + \bar{\alpha}_{23} - \bar{\alpha}_{13}} \tag{30}$$

and the corresponding $(E_b/N_0)_{\min}$ can be expressed as

$$\left(\frac{E_b}{N_0}\right)_{\min} = \frac{\log_e(2)(\bar{\alpha}_{12} + \bar{\alpha}_{23} - \bar{\alpha}_{13})}{\bar{\alpha}_{12}\bar{\alpha}_{23}}. \tag{31}$$

Since we have $f_{\max} > \bar{\alpha}_{23}$, it follows that the $(E_b/N_0)_{\min}$ in (31) for the transmitted signals given by (28) is larger than that in (26) for the transmitted signals in (14), when $E[h_{13}] \neq 0$ and $E[h_{23}] \neq 0$. If there is no relay, the $(E_b/N_0)_{\min}$ of the direct transmission is given by $(E_b/N_0)_{\min}^{(d)} = \log_e(2)/\bar{\alpha}_{13}$. Under *C1*, it follows that $(E_b/N_0)_{\min} < (E_b/N_0)_{\min}^{(d)}$; hence, relayed transmissions reduce the achievable minimum energy per bit.

Using (2), we can find the slopes of R_1 and R_2 at SNR $= 0$ from (16) and (17) with $\beta_2 = 1$ as

$$\begin{aligned} \mathcal{S}_1 &= \frac{2}{\kappa_{12}} \\ \mathcal{S}_2 &= \frac{2\left[\beta_1\bar{\alpha}_{13} + (1-\beta_1)\bar{\alpha}_{23}\right]^2}{\beta_1^2\overline{\alpha_{13}^2} + (1-\beta_1)^2\overline{\alpha_{23}^2} + 2\beta_1(1-\beta_1)\bar{\alpha}_{13}\bar{\alpha}_{23}} \end{aligned} \tag{32}$$

where $\overline{\alpha_{ij}^2} := E[|h_{ij}|^4]$ and $\kappa_{ij} := E[|h_{ij}|^4]/(E[|h_{ij}|^2])^2$ is the kurtosis of the random variable h_{ij}. Since we set $\dot{R}_1(0) = \dot{R}_2(0)$ in finding the $(E_b/N_0)_{\min}$ in (31), the slope of R at SNR $= 0$, or equivalently, at $(E_b/N_0)_{\min}$, is given by

$$\mathcal{S} = \min\{\mathcal{S}_1, \mathcal{S}_2\}. \tag{33}$$

The kurtosis of a random variable is a measure of the "peakedness" of its probability density function (pdf) [25], or a measure of the amount of fading in the context of wireless channels [26, p. 18]. Kurtosis achieves its minimum value of one when the underlying random variable is actually deterministic, or equivalently, when the pdf is a delta function. The larger the kurtosis is, the more the pdf spreads out, and the more severe the fading is. We have the following lemma regarding $\mathcal{S}_2$ in (32).

Lemma 1: For any $\beta_1 \in (0,1)$, the slope of R_2 at SNR $= 0$ in (32) satisfies

$$\mathcal{S}_2 \geq \min\left\{\frac{2}{\kappa_{13}}, \frac{2}{\kappa_{23}}\right\}. \tag{34}$$

Proof: The slope $\mathcal{S}_2$ can also be written as

$$\mathcal{S}_2 = \frac{2G}{\kappa_{13}} \tag{35}$$

where G is defined as

$$G := \frac{\left[\beta_1 + (1-\beta_1)\frac{\bar{\alpha}_{23}}{\bar{\alpha}_{13}}\right]^2}{\beta_1^2 + (1-\beta_1)^2\frac{\overline{\alpha_{23}^2}}{\overline{\alpha_{13}^2}} + 2\beta_1(1-\beta_1)\frac{\bar{\alpha}_{13}\bar{\alpha}_{23}}{\overline{\alpha_{13}^2}}}. \tag{36}$$

Denoting the difference between the numerator and the denominator of G as G_d, we have

$$\begin{aligned} G_d = 2\beta_1(1-\beta_1)\bar{\alpha}_{13}\bar{\alpha}_{23}\left(\frac{1}{\bar{\alpha}_{13}^2} - \frac{1}{\overline{\alpha_{13}^2}}\right) \\ +(1-\beta_1)^2\left(\frac{\bar{\alpha}_{23}^2}{\bar{\alpha}_{13}^2} - \frac{\overline{\alpha_{23}^2}}{\overline{\alpha_{13}^2}}\right). \end{aligned} \tag{37}$$

The first term of G_d is ≥ 0 since $\overline{\alpha_{13}^2} \geq \bar{\alpha}_{13}^2$. If $\kappa_{13} \geq \kappa_{23}$, the second term of G_d is also ≥ 0. Hence, we have $G_d \geq 0$, which implies that $G \geq 1$ and $\mathcal{S}_2 \geq 2/\kappa_{13}$. Similarly, if $\kappa_{23} \geq \kappa_{13}$, we have $\mathcal{S}_2 \geq 2/\kappa_{23}$; and (34) follows. ■

If $\kappa_{12} \leq \kappa_{13}$ and $\kappa_{23} \leq \kappa_{13}$, then from (32), (33), and *Lemma 1*, we have $\mathcal{S} \geq 2/\kappa_{13}$ for any $\beta_1 \in (0,1)$, which implies that relayed transmissions also offer a larger spectral efficiency slope than direct transmissions, since the spectral efficiency slope in direct transmissions is $\mathcal{S}_d = 2/\kappa_{13}$. In summary, we have established the following proposition.

Proposition 1: If condition *C1* is satisfied, block-Markov-coded transmissions over fading relay channels with transmitted blocks, as in (28), can afford a lower $(E_b/N_0)_{\min}$ than direct transmissions, and provide a larger spectral efficiency slope at the $(E_b/N_0)_{\min}$ if conditions $\kappa_{12} \leq \kappa_{13}$ and $\kappa_{23} \leq \kappa_{13}$ are also satisfied.

Conditions $\kappa_{12} \leq \kappa_{13}$ and $\kappa_{23} \leq \kappa_{13}$ suggest that the link quality, measured by the kurtosis of the channel coefficients between the source and the relay, and between the relay and the destination, is identical to or better than that of the link between the source and the destination. If these conditions and *C1* are satisfied, then relayed transmissions provide a larger spectral efficiency slope at $(E_b/N_0)_{\min}$. Note that conditions $\kappa_{12} \leq \kappa_{13}$ and $\kappa_{23} \leq \kappa_{13}$ are sufficient for improving spectral efficiency, but they may not be necessary. In practice, these conditions hold if all channel coefficients have the same pdf, or if there is no LOS between the source and the destination, but LOS is present between the source and the relay, and between the relay and the destination.

C. Relay Selection

In a wireless network, there may be more than one node available to relay the information of a source to its destination. While it is possible to have multiple nodes serving as relays for a particular source and destination pair (as in [14] and [17]), we here focus on the single-relay case and look for the best relay node to minimize the $(E_b/N_0)_{\min}$. To this end, we first parameterize the channels using experimentally validated physical path-loss models of wireless propagation. As discussed in [27, p. 102], the path loss is proportional to $(d/d_0)^n$, where d is the distance from the transmitter to the receiver, n is the path-loss exponent whose value (typically ranging from three to six) depends on the specific environment, and d_0 is the close-in reference. For simplicity, we neglect shadowing effects. Based on this model, we obtain $\bar{\alpha}_{ij} = \nu d_{ij}^{-n}$, where ν is a constant and d_{ij} denotes the distance between nodes i and j. The $(E_b/N_0)_{\min}$ in (31) becomes

$$\left(\frac{E_b}{N_0}\right)_{\min} = \frac{\log_e(2)}{\nu}\left(d_{12}^n + d_{23}^n - \frac{d_{12}^n d_{23}^n}{d_{13}^n}\right). \quad (38)$$

The optimal relay is the one that minimizes either (31) when $\{\bar{\alpha}_{ij}\}$ are known, or (38) when distances $\{d_{ij}\}$ are known.

We now look for the region where the relay can provide lower $(E_b/N_0)_{\min}$ than direct transmission. Based on the path-loss model, the $(E_b/N_0)_{\min}$ of direct transmissions is $(E_b/N_0)_{\min}^{(d)} = \log_e(2)d_{13}^n/\nu$. First, it is clear that we must have $d_{12} < d_{13}$, so that the link between the source and relay requires a lower $(E_b/N_0)_{\min}$ than direct transmission, because the relay needs to detect its received signals correctly. For the $(E_b/N_0)_{\min}$ in (38) to be less than $(E_b/N_0)_{\min}^{(d)}$, we require

$$d_{12}^n + d_{23}^n - \frac{d_{12}^n d_{23}^n}{d_{13}^n} < d_{13}^n \quad (39)$$

which implies that

$$\left[1 - \left(\frac{d_{12}}{d_{13}}\right)^n\right]\left[1 - \left(\frac{d_{23}}{d_{13}}\right)^n\right] > 0. \quad (40)$$

Since we have $d_{12} < d_{13}$, (40) implies that we must also have $d_{23} < d_{13}$. Hence, with the optimal power allocation specified by (30), the region where relayed transmissions provide a lower $(E_b/N_0)_{\min}$ is specified by $d_{12} < d_{13}$ and $d_{23} < d_{13}$. Combining this result with condition *C1*, we establish the following proposition.

Proposition 2: With the transmitted signals given by (28), the necessary and sufficient condition under which relayed transmissions with optimal power allocation have a lower $(E_b/N_0)_{\min}$ than direct transmission is $d_{12} < d_{13}$ and $d_{23} < d_{13}$.

With equal power allocation and the transmitted signals given by (28), we can find the $(E_b/N_0)_{\min}$ from (19) and (20) by setting $\beta_1 = 1/2$ and $\beta_2 = 1$, which leads to

$$\left(\frac{E_b}{N_0}\right)_{\min} = \max\left\{\frac{2\log_e(2)d_{12}^n}{\nu}, \frac{2\log_e(2)}{\nu\left(d_{13}^{-n} + d_{23}^{-n}\right)}\right\}. \quad (41)$$

Letting the $(E_b/N_0)_{\min}$ in (41) be less than the $(E_b/N_0)_{\min}^{(d)}$, we obtain $d_{12} < (1/2)^{1/n}d_{13}$ and $d_{23} < d_{13}$. The region where relayed transmissions have a lower $(E_b/N_0)_{\min}$ than direct transmissions is plotted in Fig. 2. Relative to the equal power allocation, it is clear that the optimal power allocation enlarges the preferred region for relay links.

If the relay can be placed at any location, it is interesting to seek the placement where relayed transmissions provide the smallest $(E_b/N_0)_{\min}$, which also yields the largest performance gain relayed transmissions can offer. It turns out that the optimal location is at the middle point on the line connecting the source with the destination, which is formally stated in the following proposition.

Proposition 3: With the transmitted signals given by (28) and optimal power allocation specified by (30), the $(E_b/N_0)_{\min}$ of relayed transmissions is minimized by placing the relay at the middle point on the line connecting the source with the destination. The $(E_b/N_0)_{\min}$ in this case is $-10\log 10[(1/2)^{n-1} - (1/2)^{2n}]$ dB lower than that of direct transmissions.

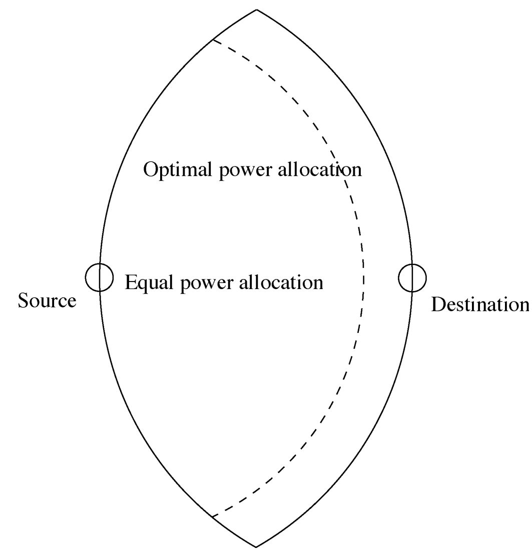


Fig. 2. Regions where relayed transmissions have a lower $(E_b/N_0)_{\min}$ than direct transmissions.

Proof: Using (38), the ratio between the $(E_b/N_0)_{\min}$ of relayed transmissions and that of direct transmissions can be written as

$$J = \frac{\left(\frac{E_b}{N_0}\right)_{\min}}{\left(\frac{E_b}{N_0}\right)_{\min}^{(d)}} = \left(\frac{d_{12}}{d_{13}}\right)^n + \left(\frac{d_{23}}{d_{13}}\right)^n - \left(\frac{d_{12}}{d_{13}}\right)^n\left(\frac{d_{23}}{d_{13}}\right)^n \quad (42)$$

which we wish to minimize with respect to d_{12} and d_{23}. To minimize J, it is sufficient to only consider the locations on the line connecting the source with the destination, because for a fixed d_{12}, if we move the relay away from this line, d_{23} will increase, which, in turn, will increase J in (42) under condition *C1*. Supposing that the relay is placed on this line, we define $\lambda = d_{12}/d_{13}$, and therefore, $d_{23}/d_{13} = 1 - \lambda$; and rewrite (42) as

$$J = \lambda^n + (1-\lambda)^n - \lambda^n(1-\lambda)^n. \quad (43)$$

Taking the first derivative of J with respect to λ, we obtain

$$\frac{dJ}{d\lambda} = n\lambda^{n-1} - n(1-\lambda)^{n-1} - n(1-\lambda)^n\lambda^{n-1} + n(1-\lambda)^{n-1}\lambda^n. \quad (44)$$

It is clear that we have $0 < \lambda < 1$, and then it is easy to verify that we have $(d^2J/d\lambda^2) > 0$, which implies that $dJ/d\lambda$ is an increasing function of λ. It is seen from (44) that one solution of $(dJ/d\lambda) = 0$ is $\lambda = 1/2$. Since $(dJ/d\lambda)|_{\lambda=0} = -n$ and $(dJ/d\lambda)|_{\lambda=1} = n$, $\lambda = 1/2$ is the only solution of $(dJ/d\lambda) = 0$ when $0 < \lambda < 1$. Hence, J is minimized when $\lambda = 1/2$, and the minimum value of J is $J_{\min} = (1/2)^{n-1} - (1/2)^{2n}$. ■

If the transmit power is equally split between source and relay, it follows from (41) that

$$J = \max\left\{2\lambda^n, \frac{2}{1 + (1-\lambda)^{-n}}\right\} \quad (45)$$

which is minimized when $2\lambda^n = 2/(1+(1-\lambda)^{-n})$, or equivalently, $\lambda^{-n} - (1-\lambda)^{-n} = 1$. Hence, the optimal λ that minimizes J in this equal power-allocation case is $< 1/2$, and it is easy to show that the minimum of J satisfies $1/(2^{n-1}+1/2) < J_{\min} < 1/2^{n-1}$. Thus, if the relay is optimally placed, relayed transmissions with equal power allocation or with optimal allocation have almost identical $(E_b/N_0)_{\min}$, which is approximately $3(n-1)$ dB lower than that of direct transmissions. However, if the relay is at an arbitrary location, optimal power allocation will provide considerably lower $(E_b/N_0)_{\min}$ than equal power allocation, as we will show in Section IV.

III. TDM Transmissions

In BMC, we assume that the relay and the source *simultaneously* transmit their signals over the same frequency bandwidth, while the relay has the capability of nullifying the interference of its own transmitted signals to its received signals. Although BMC provides high rates, it may not be practically feasible; and thus, different orthogonal channels need to be assigned to the source and the relay. In this section, we consider TDM transmissions, where the source and the relay transmit over the same frequency bandwidth but in different time slots.

In TDM transmissions, we allow the transmission time of the source and the relay to be different, which provides another degree of freedom to improve the transmission rate, depending on the quality of different links. Letting the time-sharing factor $0 < \mu < 1$ denote the percentage of the time allocated to the source, we divide every time slot of period T into two subslots of duration μT and $(1-\mu)T$. In the first subslot, the source transmits, while in the second subslot, there are two possibilities: either both the relay and the source transmit simultaneously, or only the relay transmits. We will next consider the first case, and then focus on the second one. Using the same method described in Section II-B to generate codewords $\mathbf{x}_1$ and $\mathbf{x}_2$ but with different codeword lengths, the block transmitted by the source in the first subslot is given by

$$\tilde{\mathbf{x}}_1(1) = \sqrt{\frac{P_1}{\mu}}\mathbf{x}_1 \tag{46}$$

while the blocks transmitted by the source and the relay in the second subslot are

$$\begin{aligned}\tilde{\mathbf{x}}_1(2) &= b\sqrt{\frac{(1-\beta_2)P_2}{1-\mu}}\mathbf{x}_2\\ \tilde{\mathbf{x}}_2(2) &= \sqrt{\frac{\beta_2 P_2}{1-\mu}}\mathbf{x}_2\end{aligned} \tag{47}$$

where the total transmit power is $P = P_1 + P_2$. Then, similar to (16) and (17), the transmission rate can be expressed as

$$\begin{aligned}R_1 &= \mu E\left[\log\left(1+\frac{\beta_1|h_{12}|^2}{\mu}\mathrm{SNR}\right)\right]\\ R_2 &= \mu E\left[\log\left(1+\frac{\beta_1|h_{13}|^2}{\mu}\mathrm{SNR}\right)\right]+(1-\mu)\\ &\quad\times E\left[\log\left(1+\frac{(1-\beta_1)\left|\sqrt{\beta_2}h_{23}+b\sqrt{1-\beta_2}h_{13}\right|^2}{1-\mu}\mathrm{SNR}\right)\right]\end{aligned} \tag{48}$$

where we obtained R_2 using the random bin argument of [12] and [28]. The first derivatives of R_1 and R_2 at $\mathrm{SNR}=0$ can be found to be identical to (19). Hence, the $(E_b/N_0)_{\min}$ of TDM transmissions is the same as that of BMC. As we discussed in Section II-B, when $\bar{h}_{13} = 0$ or/and $\bar{h}_{23} = 0$, the $(E_b/N_0)_{\min}$ is achieved by transmitting $\mathbf{x}_2$ only at the relay. This can be also achieved in TDM transmissions by setting $\beta_2 = 1$ in (47); effectively, the source does not transmit in the second subslot. The optimal β_1 and the $(E_b/N_0)_{\min}$ of TDM transmissions, in this case, coincide with those in (30) and (31) for BMC. We will next maximize the slope of the spectral efficiency for this TDM scheme.

With $\beta_2 = 1$ in (47), the slopes of R_1 and R_2 at $\mathrm{SNR}=0$ can be obtained from (48) as

$$\mathcal{S}_1 = \frac{2\mu}{\kappa_{12}}, \quad \mathcal{S}_2 = \frac{2\left[\beta_1\bar{\alpha}_{13}+(1-\beta_1)\bar{\alpha}_{23}\right]^2}{\frac{\beta_1^2}{\mu}\overline{\alpha_{13}^2}+\frac{(1-\beta_1)^2}{1-\mu}\overline{\alpha_{23}^2}}. \tag{49}$$

Using the optimal value of β_1 from (30), the slope $\mathcal{S}_2$ in (49) reduces to

$$\mathcal{S}_2 = \frac{2\bar{\alpha}_{12}^2}{\frac{\overline{\alpha_{13}^2}}{\mu}+\frac{(\bar{\alpha}_{12}-\bar{\alpha}_{13})^2}{1-\mu}\kappa_{23}}. \tag{50}$$

Since we have $\dot{R}_1(0) = \dot{R}_2(0)$, the slope of R can be found from

$$\mathcal{S} = \max_{\mu}\min\{\mathcal{S}_1, \mathcal{S}_2\}. \tag{51}$$

It is easy to show that the slope $\mathcal{S}_2$ in (50) has a unique maximum over $\mu \in (0,1)$, and that the maximum of $\mathcal{S}_2$ is given by

$$\mathcal{S}_2^{\max} = \frac{2\bar{\alpha}_{12}^2}{\left[\sqrt{\overline{\alpha_{13}^2}}+(\bar{\alpha}_{12}-\bar{\alpha}_{13})\sqrt{\kappa_{23}}\right]^2} \tag{52}$$

when

$$\mu = \left[1+\left(\frac{\bar{\alpha}_{12}}{\bar{\alpha}_{13}}-1\right)\sqrt{\frac{\kappa_{23}}{\kappa_{13}}}\right]^{-1} =: \mu_2. \tag{53}$$

The slope $\mathcal{S}_1$ at $\mu = \mu_2$ is given by $\mathcal{S}_1^{\mu_2} = 2\mu_2/\kappa_{12}$. If $\mathcal{S}_1^{\mu_2} \geq \mathcal{S}_2^{\max}$, it follows from (51) that $\mathcal{S} = \mathcal{S}_2^{\max}$; otherwise, the slope $\mathcal{S}$ is given by $\mathcal{S} = 2\mu_1/\kappa_{12}$, where μ_1 is found by setting $\mathcal{S}_1 = \mathcal{S}_2$, or equivalently, $\ddot{R}_1(0) = \ddot{R}_2(0)$. Letting $\ddot{R}_1(0) = \ddot{R}_2(0)$, we obtain

$$\mu_1 = \left[1+\frac{(\bar{\alpha}_{12}-\bar{\alpha}_{13})^2\kappa_{23}}{\overline{\alpha_{12}^2}-\overline{\alpha_{13}^2}}\right]^{-1}. \tag{54}$$

It can be verified that when $\mathcal{S}_1^{\mu_2} < \mathcal{S}_2^{\max}$, we have $\overline{\alpha_{12}^2} > \overline{\alpha_{13}^2}$; and thus, we have $0 < \mu_1 < 1$. We summarize our results on the slope of the spectral efficiency in the following proposition.

Proposition 4: In TDM transmissions, with the optimal β_1 given by (30), the slope of R at $\mathrm{SNR}=0$ and the optimal value of μ achieving this slope are given by $\mathcal{S} = \mathcal{S}_2^{\max}$ and $\mu = \mu_2$, if $\mathcal{S}_1^{\mu_2} \geq \mathcal{S}_2^{\max}$, or $\mathcal{S} = 2\mu_1/\kappa_{12}$ and $\mu = \mu_1$, if $\mathcal{S}_1^{\mu_2} < \mathcal{S}_2^{\max}$.

In TDM transmissions, the source and the relay transmit over different time slots, which may drop the spectral efficiency by 50%, compared with BMC. In Section IV, we will rely on numerical results to compare the spectral efficiency of TDM and BMC. In Section II, we have shown that the slope of the spectral

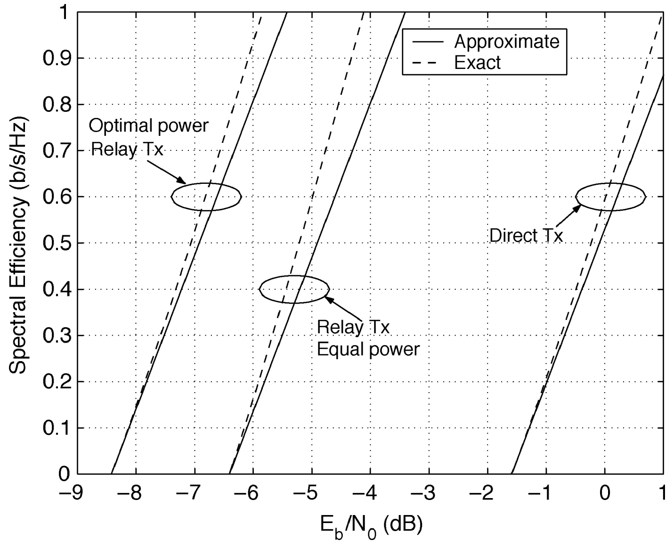


Fig. 3. Spectral efficiency of block-Markov-coded transmissions.

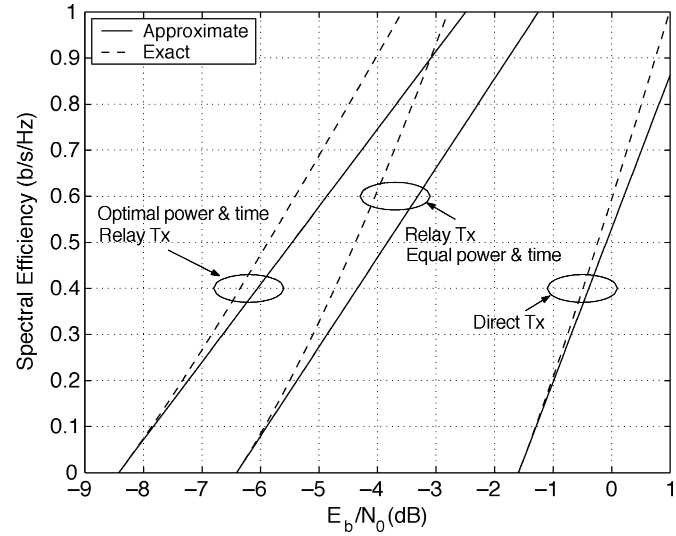


Fig. 4. Spectral efficiency of TDM transmissions.

efficiency of BMC is greater than that of direct transmissions. Here, we will prove that under mild conditions, the slope of the spectral efficiency of TDM transmissions is greater than 50% of that of direct transmissions. Note that although the spectral efficiency slope of TDM transmissions may be less than that of direct transmissions, the spectral efficiency of TDM is still higher than that of direct transmissions over a wide E_b/N_0 region, since the $(E_b/N_0)_{\min}$ for TDM transmissions is considerably lower, as we will verify numerically in Section IV.

The spectral efficiency slope of direct transmissions is given by $\mathcal{S}_d = 2/\kappa_{13}$. If $\kappa_{12} = \kappa_{23} = \kappa_{13}$, it is seen from (52) that we have $\mathcal{S}_2^{\max} = 2/\kappa_{23} > \mathcal{S}_1^{\mu_2}$; and thus, we have from *Proposition 4* that $\mathcal{S} = 2\mu_1/\kappa_{12}$ with $\mu_1 = [1 + (\bar{\alpha}_{12} - \bar{\alpha}_{13})/(\bar{\alpha}_{12} + \bar{\alpha}_{13})]^{-1} > 1/2$, which implies that $\mathcal{S} > 1/\kappa_{13} = \mathcal{S}_d/2$. We now consider more general conditions $\kappa_{12} \leq \kappa_{13}$ and $\kappa_{23} \leq \kappa_{13}$. For every value of μ, it is seen from (49) and (50) that both $\mathcal{S}_1$ and $\mathcal{S}_2$ will increase if κ_{12} and κ_{23} decrease. We thus deduce from (51) that $\mathcal{S}$ must also increase, which implies that $\mathcal{S} > \mathcal{S}_d/2$ when $\kappa_{12} \leq \kappa_{13}$ and $\kappa_{23} \leq \kappa_{13}$. We summarize our analysis of the spectral efficiency slope in the following lemma.

Lemma 2: If $\kappa_{12} \leq \kappa_{13}$ and $\kappa_{23} \leq \kappa_{13}$, then the spectral efficiency slope of TDM transmissions is larger than 50% of that of direct transmissions.

IV. Numerical Results

In this section, we will present numerical results to illustrate the advantage of relayed transmissions. We consider Rayleigh fading channels; i.e., h_{12}, h_{13}, and h_{23} are zero-mean, complex Gaussian distributed random variables. Thus, the kurtosis of the channel coefficients is $\kappa_{ij} = 2$. The variance of h_{13} is normalized to $\bar{\alpha}_{13} = 1$. The variances of h_{12} and h_{23} are chosen according to the following physical channel model. The relay is placed on the line connecting the source with the destination, and the path-loss exponent defined in Section II-C is equal to four. Hence, the variances of h_{12} and h_{13} can be calculated as $\bar{\alpha}_{12} = (d_{13}/d_{12})^4$ and $\bar{\alpha}_{23} = [d_{13}/(d_{13} - d_{12})]^4$, respectively.

Figs. 3 and 4 depict the spectral efficiency of BMC and TDM transmissions, respectively. We use $d_{12} = d_{13}/3$ in Figs. 3 and 4; and thus, we have $\bar{\alpha}_{12} = 81$ and $\bar{\alpha}_{12} = 81/16$. The exact spectral efficiency is obtained from (16) and (17) with $\beta_2 = 1$ for BMC, and from (48) with $\beta_2 = 1$ for TDM. In Fig. 3 for BMC, we use β_1 in (30) for optimal power allocation and $\beta_1 = 1/2$ for equal power allocation; in Fig. 4 for TDM, we use β_1 in (30) and μ in *Proposition 4* for optimal resource allocation, and $\beta_1 = \mu = 1/2$ for equal resource allocation. The approximate spectral efficiency plotted in Figs. 3 and 4 is calculated as $R = [E_b/N_0 - (E_b/N_0)_{\min}]\mathcal{S}/2$, while the slope S will be shown in Fig. 7. It is seen from Figs. 3 and 4 that the gap between approximate and exact curves is small when the spectral efficiency is less than 0.3 b/s/Hz, which confirms that our resource allocation based on the $(E_b/N_0)_{\min}$ and the slope achieves near-optimal rates when spectral efficiency is relatively low. It also is observed that relayed transmissions with optimal power allocation achieve about 7 dB gain relative to the direct transmissions, while optimal power allocation provides about 2 dB advantage relative to equal power allocation. The slope of R in TDM transmissions with optimal power allocation in this case is found to be $\mathcal{S} = 0.506$ b/s/Hz/(3 dB); although it is less than that of direct transmissions, the spectral efficiency of TDM transmissions is considerably higher than that of direct transmissions over a wide E_b/N_0 region, as verified by Fig. 4.

We next consider how the relay location affects performance. Fig. 5 depicts the $(E_b/N_0)_{\min}$ versus relay location, where the lower bound and the $(E_b/N_0)_{\min}$ when power is optimally allocated are given by (13) and (31), respectively, and $(E_b/N_0)_{\min} = \max\{2\log_e(2)/\bar{\alpha}_{12}, 2\log_e(2)\ /(\bar{\alpha}_{23} + \bar{\alpha}_{13})\}$ when power is equally allocated. It is seen that relayed transmissions with optimal power allocation offer considerably

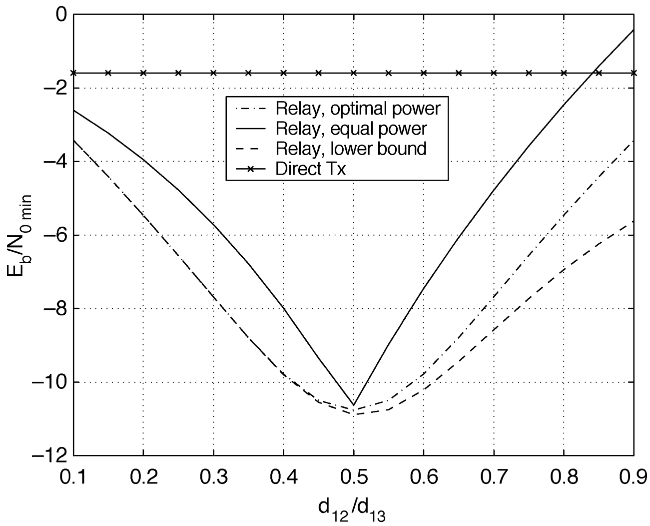


Fig. 5. Minimum energy per bit for relayed and direct transmissions.

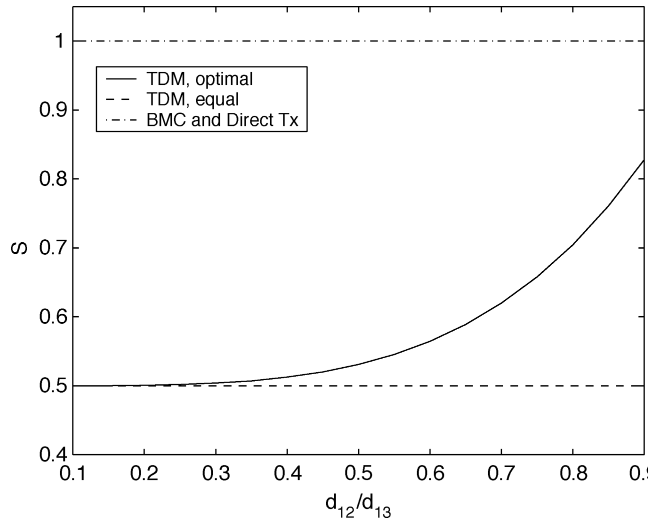


Fig. 7. Slopes of spectral efficiency for direct transmissions, BMC, and TDM.

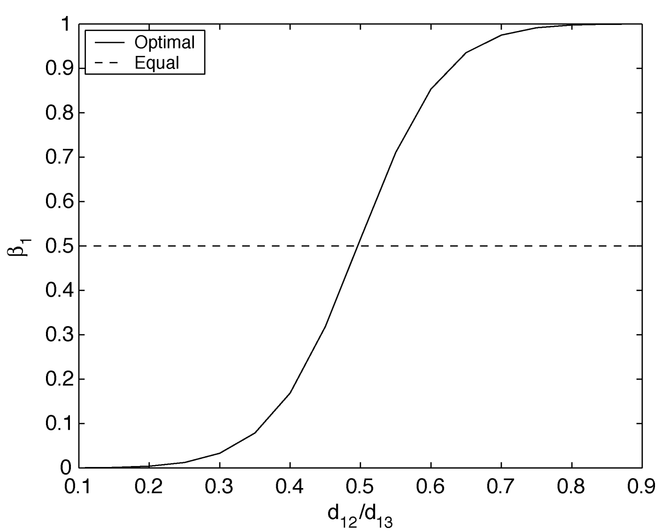


Fig. 6. Power allocation between the source and the relay.

lower $(E_b/N_0)_{\min}$ than direct transmissions, and that the lowest $(E_b/N_0)_{\min}$ is achieved at $d_{12}/d_{13} = 1/2$, which is approximately 9 dB lower than that of direct transmissions, as we discussed in Section II-C. The optimal power allocation also outperforms the equal power allocation considerably at most locations. When d_{12} is larger than a certain value, the $(E_b/N_0)_{\min}$ of relayed transmissions with equal power allocation is higher than that of direct transmissions. This is because the relay is out of the region described in Fig. 2. When $d_{12}/d_{13} < 1/2$, the $(E_b/N_0)_{\min}$ of relayed transmissions with optimal power allocation is very close to its lower bound. But when $d_{12}/d_{13} > 1/2$, it becomes considerably higher than its lower bound. Since in the BMC or TDM transmissions we considered in Sections II and III the relay needs to decode the information from the source correctly, the quality of the link between source and relay has a large impact on the overall performance. If $d_{12}/d_{13} < 1/2$, this link has reliable quality, and thus, the $(E_b/N_0)_{\min}$ reaches its lower bound. However, when the quality of this link worsens, this decoding and forwarding relay transmission strategy incurs larger performance loss; in this case, the facilitating transmission strategy in [12, Th. 6] may be a better alternative. Fig. 6 shows how the total transmit power is optimally allocated between the source and the relay, where β_1 for the optimal power allocation is calculated from (30). It is observed that more power is allocated to the weak link so that the overall performance is optimized. Fig. 7 depicts the slope of spectral efficiency. The slope for direct transmissions is $\mathcal{S}_d = 2/\kappa_{13} = 1$ b/s/Hz/(3 dB), while the slope for BMC is found from *Lemma 1* to be $\mathcal{S} = 1$ b/s/Hz/(3 dB). The slope for TDM transmissions is obtained from *Proposition 4* as $\mathcal{S} = 2\mu_1/\kappa_{12} = \mu_1$ with μ_1 given in (54), which is $> \mathcal{S}_d/2$, as discussed in *Lemma 2* and confirmed by Fig. 7. When d_{12} increases, the link between the relay and the destination improves, and thus, more transmission time is allocated to the source, which increases the slope of spectral efficiency. It is worth reiterating here that although the slope for TDM transmissions is less than that for direct transmissions, TDM offers considerably higher spectral efficiency than direct transmissions over a wide E_b/N_0 region, due to its lower $(E_b/N_0)_{\min}$.

V. Conclusions

We have investigated the achievable rates of relayed transmissions over fading channels for both BMC and TDM transmissions. Optimal power allocation between the source and the relay was derived to minimize the minimum energy per bit $(E_b/N_0)_{\min}$ required for reliable communications. When the link between the source and the relay is relatively reliable, the achievable $(E_b/N_0)_{\min}$ is close to its lower bound; otherwise, it is relatively larger than the lower bound. The optimal relay location minimizing the $(E_b/N_0)_{\min}$ was found to be the middle point of the line connecting the source with the destination. Under mild conditions, it was shown that the slope of the spectral efficiency at zero SNR for BMC is larger than that of direct transmissions, while the slope for TDM transmissions is larger than 50% of that of direct transmissions. Since relayed transmissions provide a considerably lower $(E_b/N_0)_{\min}$ requirement relative to direct transmissions, they offer relatively higher spectral efficiency over a wide E_b/N_0 region.

While the spectral efficiency of relayed transmissions over fading channels is difficult to maximize, we optimized the transmission system based on the $(E_b/N_0)_{\min}$ for BMC, and based on both the $(E_b/N_0)_{\min}$ and the slope of the spectral efficiency at zero SNR for TDM transmissions. Another possible approach to improving spectral efficiency could be based on a first-order approximation of the spectral efficiency in the low-SNR regime. As we mentioned in Section IV, the facilitating transmission strategy in [12, Th. 6] may also be worth pursuing.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their careful reading and critique of the manuscript. Their comments and suggestions improved the quality of this paper.[1]

REFERENCES

[1] A. J. Goldsmith and S. B. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Commun.*, vol. 9, pp. 8–27, Aug. 2002.

[2] P. A. Anghel and M. Kaveh, "Exact symbol error probability of a cooperative network in a Rayleigh fading environment," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1416–1421, Sep. 2004.

[3] J. Boyer, D. D. Falconer, and H. Yanikomeroglu, "A theoretical characterization of the multihop wireless communications channel with diversity," in *Proc. GLOBECOM*, San Antonio, TX, Nov. 2001, pp. 841–845.

[4] V. Emamian and M. Kaveh, "Combating shadowing effects for systems with transmitter diversity by using collaboration among mobile users," in *Proc. Int. Symp. Commun.*, Taipei, Taiwan, Nov. 2001, pp. 105.1–105.4.

[5] M. O. Hasna and M.-S. Alouini, "Performance analysis of two-hop relayed transmissions over Rayleigh fading channels," in *Proc. IEEE Veh. Technol. Conf.*, Birmingham, AL, 2002, pp. 1992–1996.

[6] A. Ribeiro, X. Cai, and G. B. Giannakis, "Symbol error probabilities for general cooperative links," *IEEE Trans. Wireless Commun.*, to be published.

[7] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, pp. 388–404, Mar. 2000.

[8] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[9] J. N. Laneman and G. Wornell, "Distributed space–time coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 59, pp. 2415–2525, Oct. 2003.

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

[11] ——, "User cooperation diversity—Part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, pp. 1938–1948, Nov. 2003.

[12] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, pp. 572–584, Sep. 1979.

[13] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: The relay case," in *Proc. INFOCOM*, New York, NY, Jun. 2002, pp. 1577–1586.

[14] ——, "On the asymptotic capacity of Gaussian relay networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Lausanne, Switzerland, 2002, p. 195.

[15] P. Gupta and P. R. Kumar, "Toward an information theory of large networks: An achievable rate region," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1877–1894, Aug. 2003.

[16] A. Reznik, S. R. Kulkarni, and S. Verdú, "Capacity and optimal resource allocation in the degraded Gaussian relay channel with multiple relays," in *Proc. 40th Allerton Conf. Commun. Contr.*, Monticello, IL, Oct. 2002, pp. 377–386.

[17] B. Schein and R. G. Gallager, "The Gaussian parallel relay network," in *Proc. IEEE Int. Symp. Inf. Theory*, Sorrento, Italy, Jun. 2000, p. 22.

[18] A. Høst-Madsen, "On the capacity of wireless relaying," in *Proc. IEEE Veh. Technol. Conf.*, Vancouver, BC, Canada, Sep. 2002, pp. 1333–1337.

[19] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "On capacity of Gaussian "cheap" relay channel," in *Proc. GLOBECOM*, San Francisco, CA, Dec. 2003, pp. 1776–1780.

[20] Y. Liang and V. V. Veeravalli, "Gaussian frequency division relay channels: Optimal bandwidth allocation and capacity," in *Proc. Conf. Inf. Sci. Syst.*, Princeton, NJ, Mar. 2004, pp. 585–590.

[21] B. Wang, J. Zhang, and A. Høst-Madsen, "On ergodic capacity of MIMO relay channel," in *Proc. Conf. Inf. Sci. Syst.*, Princeton, NJ, Mar. 2004, pp. 603–608.

[22] A. El Gamal and S. Zahedi, "Minimum energy communication over a relay channel," in *Proc. Int. Symp. Inf. Theory*, Yokohama, Japan, 2003, p. 344.

[23] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, pp. 1319–1343, Jun. 2002.

[24] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

[25] S. Shamai and S. Verdú, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inf. Theory*, vol. 47, pp. 1302–1327, May 2001.

[26] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channels: A Unified Approach to Performance Analysis*. New York: Wiley, 2000.

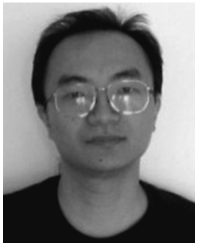
[27] T. S. Rappaport, *Wireless Communications*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[28] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. IT-19, pp. 471–480, Jul. 1973.

[1]The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government.

Xiaodong Cai (S'00–M'01) received the B.S. degree from Zhejiang University, Hangzhou, China, the M.Eng. degree from the National University of Singapore, Singapore, and the Ph.D. degree from the New Jersey Institute of Technology, Newark, in 2001, all in electrical engineering.

From February 2001 to June 2001, he was a Member of Technical Staff with Lucent Technologies, NJ. From July 2001 to October 2001, he was a Senior System Engineer with Sony Technology Center, San Diego, CA. From November 2001 to August 2004, he was a Postdoctoral Research Associate with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis. Since September 2004, he has been an Assistant Professor with the Department of Electrical and Computer Engineering, University of Miami, Coral Gables, FL. His research interests lie in the areas of communication theory, signal processing, and networking. Current research focuses on cooperative communications and collaborative signal processing for wireless ad hoc and sensor networks.



Yingwei Yao (S'98–M'03) received the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 2002.

He was a Postdoctoral Researcher with the University of Minnesota, Minneapolis, from 2002 to 2004. Currently, he is an Assistant Professor of Electrical and Computer Engineering with the University of Illinois at Chicago. His research interests include statistical signal processing, wireless communication theory, and wireless networks.

Georgios B. Giannakis (S'84–M'86–SM'91–F'97) received the Diploma in electrical engineering from the National Technical University of Athens, Greece, in 1981. He received the MSc. degree in electrical engineering in 1983, M.Sc. degree in mathematics in 1986, and the Ph.D. degree in electrical engineering in 1986, from the University of Southern California (USC), Los Angeles.

After lecturing for one year at USC, he joined the University of Virginia, Charlottesville, in 1987, where he became a Professor of Electrical Engineering in 1997. Since 1999, he has been with the University of Minnesota, Minneapolis, as a Professor in the Department of Electrical and Computer Engineering, and holds an ADC Chair in Wireless Telecommunications. His general interests span the areas of communications and signal processing, estimation and detection theory, time-series analysis, and system identification, subjects on which he has published more than 200 journal papers, 350 conference papers, and two edited books. Current research focuses on transmitter and receiver diversity techniques for single- and multiuser fading communication channels, complex-field and space–time coding, multicarrier, ultra-wideband wireless communication systems, cross-layer designs, and distributed sensor networks.

Dr. Giannakis is the corecipient of six Best Paper Awards from the IEEE Signal Processing (SP) Society (1992, 1998, 2000, 2001, 2003, and 2004). He also received the Society's Technical Achievement Award in 2000. He co-organized three IEEE-SP Workshops, and guest co-edited four special issues. He has served as Editor in Chief for the IEEE SIGNAL PROCESSING LETTERS, as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS, as secretary of the SP Conference Board, as member of the SP Publications Board, as member and vice-chair of the Statistical Signal and Array Processing Technical Committee, and as chair of the SP for Communications Technical Committee. He has served as a member of the Editorial Board for the PROCEEDINGS OF THE IEEE, and the steering committee of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He is a member of the IEEE Fellows Election Committee, and has served on the IEEE-SP Society's Board of Governors.