MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Energy-Efficient Cooperative Relaying Over Fading Channels with Simple Relay Selection

Ritesh Madan, Neelesh Mehta, Andreas Molisch, Jin Zhang

TR2008-070 November 2008

Abstract

We consider a cooperative wireless network where a set of nodes cooperate to relay in parallel the information from a source to a destination using a decode-and-forward approach. The source broadcasts the data to the relays, some or all of which cooperatively beamform to forward the data to the destination. We generalize the standard approaches for cooperative communications in two key respects: (i)we explicitly model and factor in the cost of acquiring channel state information (CSI), and (ii)we consider more general selection rules for the relays and compute the optimal one among them. In particular, we consider simple relay selection and outage criteria that exploit the inherent diversity of relay networks and satisfy a mandated outage constraint. These criteria include as special cases serveral relay selection criteria proposed in the literature. We obtain expressions for the total energy consumption for general relay selection and outage criteria for the non-homogeneous case, in which different relay links have different mean channel power gains, and the homogeneous case, in which the relay links statistics are identical. We characterize the structure of the optimal transmission scheme. Numerical results show that the cost of training and feedback of CSI is significant. The optimal strategy is to use a varying subset (and number) or relay nodes to cooperatively beamform at any given time. Depending on the relative location of the relays, the source, and the destination, numerical computations show energy savings of about 16% when an optimal relay selection rule is used. We also study the impact of shadowing correlation on the energy consumption for a cooperative relay network.

IEEE Transaction on Wireless Communication

Energy-Efficient Cooperative Relaying over Fading Channels with Simple Relay Selection

Ritesh Madan, *Member, IEEE*, Neelesh B. Mehta, *Senior Member, IEEE*, Andreas F. Molisch, *Fellow, IEEE*, and Jin Zhang, *Fellow, IEEE*

***Abstract*—We consider a cooperative wireless network where a set of nodes cooperate to relay in parallel the information from a source to a destination using a decode-and-forward approach. The source broadcasts the data to the relays, some or all of which cooperatively beamform to forward the data to the destination. We generalize the standard approaches for cooperative communications in two key respects: (i) we explicitly model and factor in the cost of acquiring channel state information (CSI), and (ii) we consider more general selection rules for the relays and compute the optimal one among them. In particular, we consider simple relay selection and outage criteria that exploit the inherent diversity of relay networks and satisfy a mandated outage constraint. These criteria include as special cases several relay selection criteria proposed in the literature. We obtain expressions for the total energy consumption for general relay selection and outage criteria for the non-homogeneous case, in which different relay links have different mean channel power gains, and the homogeneous case, in which the relay links statistics are identical. We characterize the structure of the optimal transmission scheme. Numerical results show that the cost of training and feedback of CSI is significant. The optimal strategy is to use a varying subset (and number) of relay nodes to cooperatively beamform at any given time. Depending on the relative location of the relays, the source, and the destination, numerical computations show energy savings of about 16% when an optimal relay selection rule is used. We also study the impact of shadowing correlation on the energy consumption for a cooperative relay network.**

***Index Terms*—Cooperative networks, relay networks, Rayleigh fading, broadcast, beamform, channel state information, energy optimization, outage, virtual branch analysis, optimization.**

I. Introduction

Cooperative communication networks, in which wireless nodes cooperate with each other in transmitting information, promise significant gains in overall throughput and energy efficiency [1]–[3]. The networks exploit the diversity inherent in multiple spatially distributed wireless links, and require only nodes with single antennas. We consider networks where nodes are powered by batteries that can supply only a finite amount of energy. In such networks, energy efficiency is critical as it affects network lifetime [4]. At the same time, the outage probability in fading channels must be kept below a specified level to make the network useful for data delivery. For such networks, we analyze the energy consumption of a class of cooperative communication schemes, where the number of relays that actively relay data vary with time as the channel state changes.

The broadcast nature of the wireless channel can be exploited to save energy by transmitting (broadcasting) to multiple relay nodes simultaneously; some or all of the relays decode the signal. Cooperative beamforming algorithms can then be used to save energy in transmitting data from the relays to the destination [5]–[9]. In this approach, the relays, with knowledge of the required channel state information (CSI), linearly weight their transmit signals so that they add up coherently at the destination.

While cooperative beamforming is motivated by maximal ratio transmission by a transmitter equipped with multiple antennas [21], [28]–[31], obtaining and exploiting CSI in a distributed manner is an additional challenge for cooperative beamforming. The cost of obtaining CSI for relay cooperation and its implications has not been considered in the literature cited above. Transmitter-side CSI (without factoring in the cost) was assumed in [3], [7], [10] and [11]. The authors in [5] considered amplify-and-forward instead of decode-and-forward, and also did not include the cost of acquiring CSI. Also, the model in [6] did not consider channel fading (and outage) or the cost of obtaining the channel phase information at the transmitters. In the papers cited above, all relay nodes that decode the signal cooperate in the beamforming. A fully distributed power allocation scheme, where the power at each relay node is decided only on the basis of its instantaneous channel gain to the destination, was described in [14]. Strategies for power allocation in Gaussian relay networks to maximize the net throughput have also been considered in the literature; see, for example [15].

In this paper, we explicitly model and factor in the cost of acquiring CSI. By means of a training process to obtain the CSI at the destination and a feedback process from the destination to the relays, the relays obtain the CSI necessary to enable energy efficient cooperation. Note that not factoring in the overhead for obtaining CSI leads to the trivial result that all the available nodes should relay the information to the destination. However, for fading channels, it is often energy-intensive to feedback CSI reliably to all the relays. When we

Manuscript received February 2006; revised April 4, 2007 and December 13, 2007; accepted February 24, 2008. The associate editor coordinating the review of this paper and approving it for publication was A. Yener.

R. Madan is with Qualcomm-Flarion Technologies, NJ (e-mail: rkmadan@stanfordalumni.org).

N. B. Mehta is with the Dept. of Electrical Communication Engineering, Indian Institute of Science (IISc), Bangalore, India (e-mail: nbmehta@ece.iisc.ernet.in).

A. F. Molisch and J. Zhang are with Mitsubishi Electric Research Labs (MERL), Cambridge, MA (e-mail: {molisch, jzhang}@merl.com). A. F. Molisch is also with the Department of Electroscience, Lund University, Sweden.

Digital Object Identifier 10.1109/TWC.2008.06090.

minimize the total energy consumption for data transmission and CSI acquisition, there is a tradeoff between decreasing energy consumption for data transmission by using more relays and decreasing overhead for CSI acquisition by using less relays. This naturally raises the problem of computing an *optimal relay selection rule*. The relay selection rules in [12], [13], [16]–[19] are restrictive in the sense that they either always use all the available relays or always use just one relay. Of the four simple relay selection criteria described in [12], two criteria select a single relay based on mean channel power gains, while the other two select all the relays. A single relay node is selected based on average CSI, such as distance or path loss, in [13], [18], [19], and based on instantaneous fading states of the links from the source to the relay and the relay to the destination in [16].

The main contributions of this paper are as follows. We analyze the total energy cost of data transmission for cooperative beamforming, and we optimize it over a more general class of relay selection rules than those considered in the literature. In particular, we optimize a general, yet simple, class of relay selection rules that select the *number of relays* based on the set of nodes that decode data from the source, but not on their instantaneous fading states. Both single relay selection and selection of all relays are clearly special cases of this rule. Note that the actual set of relays used *does* depend on the instantaneous channel states as the relays with the best instantaneous channels to the destination are chosen. Conditioned on the number of relays that decode, our selection rules also achieve the same *diversity order* as the instantaneous state-based rules (see, for example, [20]). Furthermore, it suggests a correspondingly simple rule for outage to save energy – outage is also declared based only on the set of nodes. While this leaves open the possibility that outage is declared even though a small number of relay nodes that decode the source data have high channel power gains, we show that even for our simple scheme, the resulting energy savings are substantial. Compared to criteria that depend on instantaneous channel states, this approach is clearly sub-optimal. However, it greatly reduces the complexity of implementing the selection rule at the destination and enables a unified tractable analytical treatment and optimization of relay selection and outage to minimize *total* energy consumption. In effect, the optimal relay selection and outage rules are functions only of the average CSI (mean channel power gains) of the various nodes, which is akin to [12], [13], [18], [19].

The outline of the paper is as follows. Section II describes the system model and the cooperative communication scheme. In Section III, we develop a general analytical framework for the homogeneous case, in which the relay links have identical statistics, and for the non-homogeneous case, in which they are not. We derive key properties of the optimal relay cooperation and outage rules that minimize the energy consumption. Section IV presents numerical results that illustrate the various trade-offs, and is followed by our conclusions in Section V.

II. System Model

Figure 1 shows a schematic of the relay network. It contains one source node, one destination node, and N relays. The channels from the source to the relays as well as from the relays to the destination are frequency non-selective channels that undergo independent Rayleigh fading. Thus, the channel power gains from source to relays (S-R), denoted by h_i, and from relays to destination (R-D), denoted by g_i, are independent, exponentially distributed random variables with means $\bar{h}_i$ and $\bar{g}_i$, respectively, where $i = 1, \ldots, N$. *The mean channel power gains depend upon the distance between the corresponding nodes and shadowing; in general, the means are not identical*. We also assume that all links are reciprocal. This condition is fulfilled in time division duplexing systems where the round-trip duplex time is much shorter than the coherence time of the channel, or in frequency division duplexing systems where the frequency duplex separation is smaller than the coherence bandwidth [21].

At all nodes, the additive white Gaussian noise has a power spectral density of N_0. All the transmissions in the system have a bandwidth of B Hz and occur with a fixed rate of r bits/symbol. Fixing r simplifies the design of the relays as they do not need to remodulate their transmissions using a different signal constellation. We assume that a node can decode data only if the received signal power exceeds a threshold, which is a function of the rate, r and the total noise power. We use the Shannon capacity formula to illustrate the relationship between the threshold and r; in particular, we choose the threshold as $N_0 B(2^r - 1)$.[1]

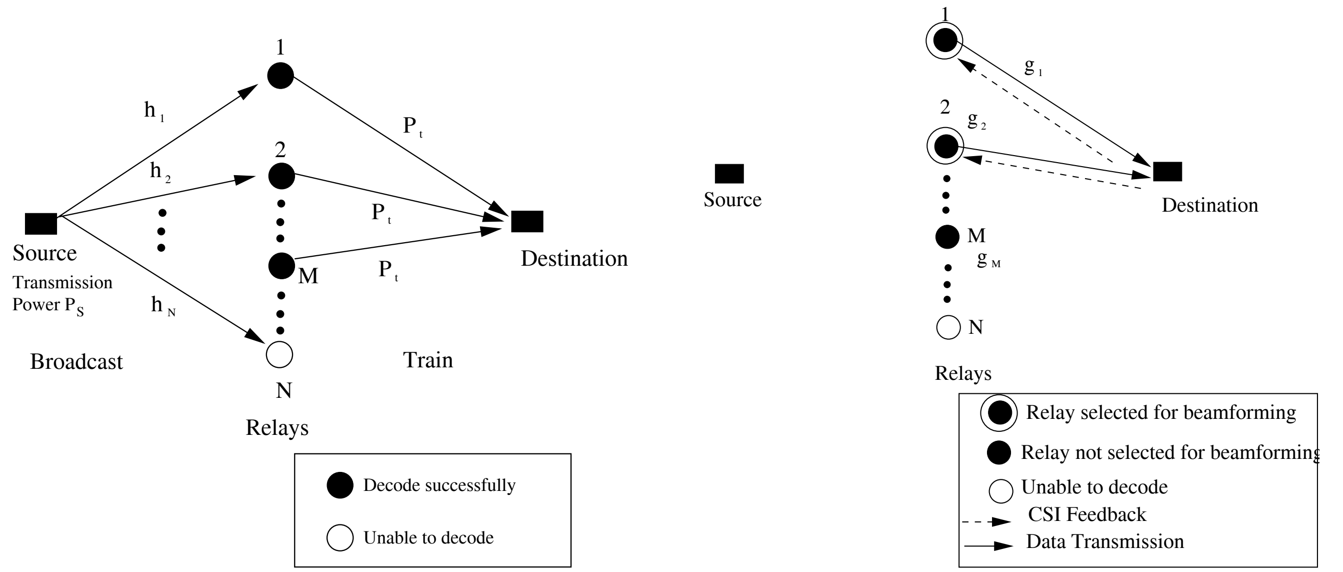
A. Cooperative Communication Scheme

We describe the steps for the transmission of a single message from the source to the destination, via the relay nodes. The source first broadcasts the data. The relays that receive the data with a sufficiently high SNR successfully decode the data. Subsequently, the relays that decode successfully send a training sequence to the destination, which estimates the respective channel states. The destination decides which relay nodes should actually participate in the data transmission, and feeds back the CSI to those nodes. In the final step, the selected nodes send the linearly weighted transmit signal to the destination. We assume the channel from the source to the destination is weak enough such that we can neglect the signal received by the destination directly from the source. However, the same analysis framework can be used to analyze this extension, as well. We now describe the steps in detail.

Broadcast: The source, which does not know the relay channel gains *a priori*, broadcasts data to the relays using a fixed transmission power, P_S, at a fixed rate, r bits/symbol, for T_d symbol durations. The received power at relay i is $h_i P_S$. This node can decode the data correctly only if the received power exceeds the threshold $N_0 B(2^r - 1)$. Thus, depending on the channel states, only a subset, $\mathcal{M} \subseteq \{1, \ldots, N\}$, of the relays successfully receives the data from the source. We use M to denote the size of $\mathcal{M}$.

Training: Only the M relays that receive data successfully from the source send training sequences at a rate r bits per symbol and power P_t to the destination. This enables the destination to estimate the instantaneous channel power gains, $\{g_i, i \in \mathcal{M}\}$, from the relays to the destination. P_t is taken to

[1]Similar threshold formulas exist for MFSK and MQAM with and without error correction coding [22].



(a) Broadcast and training (b) CSI feedback and data transmission

Fig. 1. Cooperative communication scheme steps.

be sufficient for the destination to accurately estimate the gains of channels whenever they are used for data transmission. Note that due to the presence of diversity in the system, the probability of using a bad channel is very low. Also, we assume that training transmissions use $T_t = M$ symbol durations, which is the minimum possible value.[2]

Feedback of CSI: Based on the channel power gains, $\{g_i, i \in \mathcal{M}\}$, the destination either declares an outage with probability $p_{\text{out}}(\mathcal{M})$ or it selects a subset of $\mathcal{M}$, consisting of $K(\mathcal{M})$ relays with the best channel power gains to the destination, and feeds back to them the required CSI. The CSI requirements are discussed in the next step.

As mentioned above, the *number of relays*, $K(\mathcal{M})$, selected for data transmission (when outage is not declared) is only based on $\mathcal{M}$ and not on the instantaneous channel states, $\{g_i, i \in \mathcal{M}\}$. However, the actual set of relays (of size $K(\mathcal{M})$) used at each step does depend on the instantaneous channel power gains. Similarly, outage is declared by the destination with a probability $p_{\text{out}}(\mathcal{M})$ that is a function of the set $\mathcal{M}$ and is independent of the channel power gains.

Let us reorder the channel power gains of the relays that decode the data from the source successfully in the descending order, $g_{[1]} > \ldots > g_{[M]}$, where $[i]$ denotes the index of the relay with the i^{th} largest gain. As shown in the next step, it is sufficient for the destination to feedback the sum of channel power gains, $\sum_{i=1}^{K(\mathcal{M})} g_{[i]}$, to all the selected $K(\mathcal{M})$ nodes, and the channels power gain ($g_{[i]}$) and its phase to only the corresponding selected relay $[i]$. The feedback, at a rate of r, takes T_f symbol durations. If c symbols are required to feedback each channel power gain and phase, then $T_f(K(\mathcal{M})) = c(1 + K(\mathcal{M}))$. Using the SNR threshold formula based on Shannon capacity, the minimum feedback power to reach relay i is $N_0 B(2^r - 1)/g_{[i]}$ and the minimum feedback power to broadcast the sum of channel power gains to all the $K(\mathcal{M})$ relays is determined by node $[K(\mathcal{M})]$ (with the worst channel) and is $N_0 B(2^r - 1)/g_{[K(\mathcal{M})]}$.

The $M = 1$ case (when one relay decodes the data) needs special attention because the minimum power at which the relay needs to transmit to reach the destination is proportional to the inverse of the channel power gain. As is well known, infinite average power is necessary for channel inversion with zero outage over a Rayleigh fading channel [23]. Therefore, for this special case, the node is allowed to transmit only if its channel power gain exceeds a threshold. Thus, it does *not* transmit, with a probability of δ, even if the destination has not declared an outage. We will assume that δ is a fixed system parameter. To summarize, when $M \geq 2$, the relays cannot transmit if the destination declares an outage, while for $M = 1$, the relay cannot transmit when the destination declares an outage or when the gain is below a threshold. Another minor difference is that when the destination does allow only one relay to transmit, it has to feedback to this relay only the gain of the channel from the relay to itself, which takes c, not $2c$, symbol durations.

Cooperative Beamforming: Given the knowledge of the CSI, the optimal transmission power at each selected relay i can be shown to be [6] $\frac{g_{[i]}}{\left(\sum_{j=1}^{K(\mathcal{M})} g_{[j]}\right)^2} N_0 B(2^r - 1)$. The $K(\mathcal{M})$ nodes cooperate, i.e., transmit coherently, to send data at a rate r bits/symbol to the destination for T_d symbol durations.

To ensure that the channel does not change during data transmission, the longest possible period for training, feedback of CSI, and message transmission, $N + c(N+1) + T_d$, must be less than the coherence time of the channel. We model only the energy required for radio transmission and not the energy consumed for receiving. This is justifiable as the radio transmission is the dominant component of energy consumption for long range transmissions [22]. Feedback quantization is taken to be sufficiently fine to not affect beamforming performance.

[2]Several mechanisms can be used to enable uncoordinated training among relay nodes. One approach, which is easily implementable, but is time inefficient, is to pre-assign a training slot for each relay. However, only relays that decode transmit a training sequence. MAC-based training mechanisms can be used to reduce this time-inefficiency.

The relay transmissions are assumed to be coherent and synchronized. An alternate model is to assume that the relays forward the message to the destination using orthogonal channels. While this makes the synchronization easier, it is not as spectrally efficient. Preliminary mechanisms for ensuring synchronization among simple distributed nodes for cooperative beamforming were proposed in [8], which showed energy savings even with imperfect synchronization. Imperfect synchronization can also be overcome by employing a more sophisticated Rake receiver at the destination to capture all the energy of the received signal. A detailed analysis of the impact of quantization and imperfect synchronization is beyond the scope of this paper.

III. Mathematical Analysis

Let $p(\mathcal{M}, P_S)$ denote the probability that exactly the set $\mathcal{M}$ of relays successfully decode the data broadcast by the source, when the source broadcast power is P_S. For a given relay selection rule $K(.)$, let $P_f(K(\mathcal{M}), \mathcal{M})$ denote the average power consumed in feeding back the CSI to the selected relays, and $P_d(K(\mathcal{M}), \mathcal{M})$ denote the average power consumed by the relays to coherently transmit data, both conditioned on the events that $\mathcal{M}$ is the set of relays that decode data from the source and that the destination does not declare outage. Here, all the averages are with respect to the joint distribution of the h_i's and g_i's. Note that for $M = 1$, these quantities also take into account the possibility that the single relay is not allowed to transmit because its relay gain to the destination is below a threshold.

We analyze the average total energy required to transmit a message from the source to the destination as a function of the broadcast power P_S, the outage rule given by the $p_{\text{out}}(\mathcal{M})$'s, and the relay selection rule K. The aim is to determine the optimal source broadcast power, P_S, outage probabilities, $p_{\text{out}}(\mathcal{M})$, and the relay selection rule, $K(\mathcal{M})$, for $\mathcal{M} \subseteq \{1, \ldots, N\}$, that minimize the average energy consumption per message subject to an outage constraint.

Data does not reach the destination when (i) no relays receive data from the source (with probability $p(\emptyset, P_S)$), or (ii) when only one relay, i, receives data and does not transmit to the destination (with probability $p_{\text{out}}(\{i\})p(\{i\}, P_S) + \delta p(\{i\}, P_S)(1 - p_{\text{out}}(\{i\}))$), or (iii) when exactly the set $\mathcal{M}$ of relays, with $M \geq 2$, receives data and outage is declared by the destination (with probability $p(\mathcal{M}, P_S)p_{\text{out}}(\mathcal{M})$). In the above discussion, recall that when one relay node, i, receives data, it does not transmit if at least one of the following two independent events occurs: (1) the channel gain from the relay to the destination is too low, or (2) the destination declares outage with probability $p_{\text{out}}(\{i\})$ (independent of the channel power gain). Therefore, the constraint that the destination receives data from the source with a probability that exceeds $(1 - P_{\text{fail}})$ can be written as

$$P_{\text{fail}} \geq p(\emptyset, P_S) + \sum_{i=1}^{N} \delta p(\{i\}, P_S)(1 - p_{\text{out}}(\{i\})) + \sum_{\mathcal{M} \subseteq \{1,\ldots,N\}} p(\mathcal{M}, P_S)p_{\text{out}}(\mathcal{M}). \quad (1)$$

The energy consumed in broadcasting a message from the source to all the relays is $T_d P_S$. The $\mathcal{M}$ relays, which receive the data, transmit training sequences to the destination. This consumes energy MP_t. If the destination decides that the relay(s) will transmit, it needs to feedback CSI to the $K(\mathcal{M})$ selected relays. This consumes an average energy of $T_f(K(\mathcal{M}))P_f(K(\mathcal{M}), \mathcal{M})$. The relays beamform to transmit the message to the destination, which consumes an average energy of $T_d P_d(K(\mathcal{M}), \mathcal{M})$. The total average energy consumption, $E(p_{\text{out}}, K, P_S)$, is given by

$$E(p_{\text{out}}, K, P_S) = T_d P_S + \sum_{\mathcal{M} \subseteq \{1,\ldots,N\}} p(\mathcal{M}, P_S)\Big(MP_t + (1 - p_{\text{out}}(\mathcal{M}))\Big(T_f(K(\mathcal{M}))P_f(K(\mathcal{M}), \mathcal{M}) + T_d P_d(K(\mathcal{M}), \mathcal{M})\Big)\Big). \quad (2)$$

Reorder the S-R gains $\{h_i\}_{i=1}^N$ in descending order $h_{[1]} > \ldots > h_{[N]}$. Note that the reordering permutation depends on the instantaneous values of $h_1, \ldots, h_N$. Relay i receives data if and only if $h_i \geq \frac{N_0 B(2^r - 1)}{P_S} \triangleq \gamma_r$. Then, we have the following relations:

$$\begin{aligned} &p(\mathcal{M}, P_S) \\ &= \Pr\big(\{h_i \geq \gamma_r : i \in \mathcal{M}\} \cap \{h_i < \gamma_r : i \notin \mathcal{M}\}\big) \\ &= \Pr\big(\{M \text{ nodes receive data}\} \cap \\ &\qquad \{[1], \ldots, [M] \text{ belong to set } \mathcal{M}\}\big). \end{aligned} \quad (3)$$

We now derive closed-form expressions for the total energy consumption for the homogeneous and non-homogeneous cases. In the homogeneous case, all S-R (and R-D) channels have the same mean channel power gain, i.e., $\bar{h}_i = \bar{h}$, and $\bar{g}_i = \bar{g}$, for $i = 1, \ldots, N$, respectively. We shall see that the analysis simplifies a lot, and leads to simple and efficient algorithms for computing the optimal transmission scheme. Also, the tradeoffs are very clear, and help us gain good intuition for system design.

A. Homogeneous Channels

Using symmetry arguments, we can show that there exists an optimal transmission strategy for which $p_{\text{out}}(\mathcal{M})$ is the same for all sets $\mathcal{M}$ of the same cardinality. Hence, in this section, without loss of generality, we will restrict ourselves to relay selection rules $K(\mathcal{M})$ and outage rules $p_{\text{out}}(\mathcal{M})$ that depend only on M. By symmetry, we also have: $p(\mathcal{M}, P_S) = p(M, P_S)$, $P_f(K(\mathcal{M}), \mathcal{M}) = P_f(K(M), M)$, and $P_d(K(\mathcal{M}), \mathcal{M}) = P_d(K(M), M)$, and $K^*(\mathcal{M})$ depends only on M. (Recall that $M = |\mathcal{M}|$.)

The constraint that the destination receives data from the source with a probability greater than or equal $(1 - P_{\text{fail}})$ in (1) can be now written as

$$P_{\text{fail}} \geq p(0, P_S) + \delta p(1, P_S)(1 - p_{\text{out}}(1)) + \sum_{M=1}^{N} p(M, P_S)p_{\text{out}}(M). \quad (4)$$

The total average energy consumed, $E(p_{\text{out}}, K, P_S)$, in (2) now simplifies to

$$E(p_{\text{out}}, K, P_S) = T_d P_S + \sum_{M=1}^{N} p(M, P_S) M P_t + \sum_{M=1}^{N} p(M, P_S)(1 - p_{\text{out}}(M))\Big(T_f(K(M)) P_f(K(M), M) + T_d P_d(K(M), M)\Big). \quad (5)$$

We now derive expressions for the energy terms corresponding to feedback and data transmission. The $M > 1$ and $M = 1$ cases are treated separately as their transmission criteria differ slightly. This is followed by the optimal transmission strategy characterization.

1) $M > 1$ Case: The statistics of g_i are independent of $\mathcal{M}$ because all channel power gains are independent of each other. Arrange the R-D channel power gains in descending order $g_{[1]} > \ldots > g_{[M]}$. As mentioned, the destination selects the $K(M)$ best relays with indices $[1], \ldots, [K(M)]$ and broadcasts the sum of the channel power gains to all of them, and the individual channel power gains and phases only to the corresponding relays. Hence, the average power consumption for feedback of CSI is

$$P_f(K(M), M) = \frac{N_0 B (2^r - 1)}{K(M) + 1} \mathbb{E}\left[\frac{1}{g_{[K(M)]}} + \sum_{i=1}^{K(M)} \frac{1}{g_{[i]}}\right]. \quad (6)$$

The term $(K(M) + 1)$ in the denominator arises as the energy is consumed over $(K(M) + 1)$ slots. The average power consumed by the relays to cooperatively beamform and transmit data is

$$P_d(K(M), M) = N_0 B (2^r - 1) \mathbb{E}\left[\frac{1}{g_{\text{sum}}}\right], \quad \text{where} \quad g_{\text{sum}} = \sum_{i=1}^{K(M)} g_{[i]}. \quad (7)$$

Expressions for $p(M, P_S)$, $P_f(K(M), M)$, and $P_d(K(M), M)$: We use the virtual branch analysis techniques in [26] to derive the expressions. Let

$$h_{[i]} = \sum_{n=i}^{N} \frac{W_n}{n}, \; i = 1, \ldots, N,$$
$$g_{[i]} = \sum_{n=i}^{M} \frac{V_n}{n}, \; i = 1, \ldots, M.$$

It can be shown that for Rayleigh fading, V_i are i. i. d. random variables and have an exponential distribution with mean $\bar{g}$. Similarly, W_i are also i. i. d. and have an exponential distribution with mean $\bar{h}$. It can then be shown that

$$p(M, P_S) = \frac{N!}{M!} \sum_{j=M+1}^{N} \frac{e^{-\frac{\gamma_T M}{h}} - e^{-\frac{\gamma_T j}{h}}}{(j - M) \prod_{l \geq M+1, l \neq j} (l - j)}.$$

Also, it follows from the change of variables above that

$$\mathbb{E}\left[\frac{1}{g_{[i]}}\right] = \mathcal{E}_{1+M-i}\left(\frac{\bar{g}}{i}, \ldots, \frac{\bar{g}}{M}\right),$$
$$\mathbb{E}\left[\frac{1}{g_{\text{sum}}}\right] = \mathcal{E}_M\left(\bar{g}, \ldots, \bar{g}, \frac{\bar{g} K(M)}{K(M) + 1}, \ldots, \frac{\bar{g} K(M)}{M}\right).$$

Here, $\mathcal{E}_n(\bar{y}_1, \ldots, \bar{y}_n)$ denotes the mean of $1/(Y_1 + \ldots Y_n)$, where $\bar{y}_i$ is the mean of exponential random variable Y_i. A closed-form expression for $\mathcal{E}_n(\bar{y}_1, \ldots, \bar{y}_n)$ is derived in Appendix A.

2) $M = 1$ Case: Let i denote the single relay that decodes the data from the source. This case is different because the relay also does not transmit if the instantaneous gain, g_i, is too low. When outage is not declared, the node inverts the channel to transmit data to the destination at rate r. The average power consumed to feedback CSI is given by

$$P_f(K(1), 1) = N_0 B (2^r - 1) \int_{\alpha_i}^{\infty} \frac{1}{\bar{g} x} e^{-\frac{x}{\bar{g}}} dx = -\frac{N_0 B (2^r - 1)}{\bar{g}} \text{Ei}\left(\frac{-\alpha}{\bar{g}}\right), \quad (8)$$

where $\alpha = -\bar{g} \log_e(1 - \delta)$ and Ei is the standard exponential integral function [24] given by $\text{Ei}(u) = \int_{-\infty}^{u} \frac{e^x}{x} dx$. Using similar arguments,

$$P_d(K(1), 1) = -\frac{N_0 B (2^r - 1)}{\bar{g}} \text{Ei}\left(\frac{-\alpha}{\bar{g}}\right).$$

Hence, the average energy consumption for CSI feedback and data transmission can now be computed from (6) and (7), respectively.

3) Optimal Transmission Strategy: It follows from (6), (7), and (8) that $P_f(K(M), M)$ and $P_d(K(M), M)$ do not depend on $K(M')$ for $M' \neq M$. It follows from this separable structure of the total average energy consumption per message in (2) that the optimal relay selection rule $K^*(\mathcal{M})$ can be computed for each set $\mathcal{M}$ separately. Thus,

$$K^*(M) = \underset{1 \leq K(M) \leq M}{\arg\min} \Big[T_f(K(M)) P_f(K(M), M) + T_d P_d(K(M), M)\Big]. \quad (9)$$

The optimal outage strategy can be shown to have a simple structure if for $M > 1$,

$$\frac{1}{(1 - \delta)}\big(c P_f(K^*(1), 1) + T_d P_d(K^*(1), 1)\big) \geq c(1 + K^*(M)) P_f(K^*(M), M) + T_d P_d(K^*(M), M), \quad (10)$$

i.e., the optimal feedback and data power consumption conditioned on $M > 1$ (nodes can beamform to transmit with zero outage) is less than or equal to $\frac{1}{1-\delta}$ times that conditioned on $M = 1$. The following lemma then follows.

Lemma 3.1: The optimal outage strategy has the following structure if (10) is satisfied. For some $0 < M^* < N$,

$$\begin{aligned} &p^*_{\text{out}}(M) = 1, \quad M < M^*, \\ &0 \leq p^*_{\text{out}}(M^*) \leq 1, \\ &p^*_{\text{out}}(M) = 0, \quad M > M^*. \end{aligned} \quad (11)$$

Moreover,

$$P_{\text{fail}} = p(0, P_S) + \sum_{M=1}^{N} p(M, P_S) p^*_{\text{out}}(M) + \delta p(1, P_S)(1 - p^*_{\text{out}}(1)). \quad (12)$$

Proof: The proof is given in Appendix B. ■

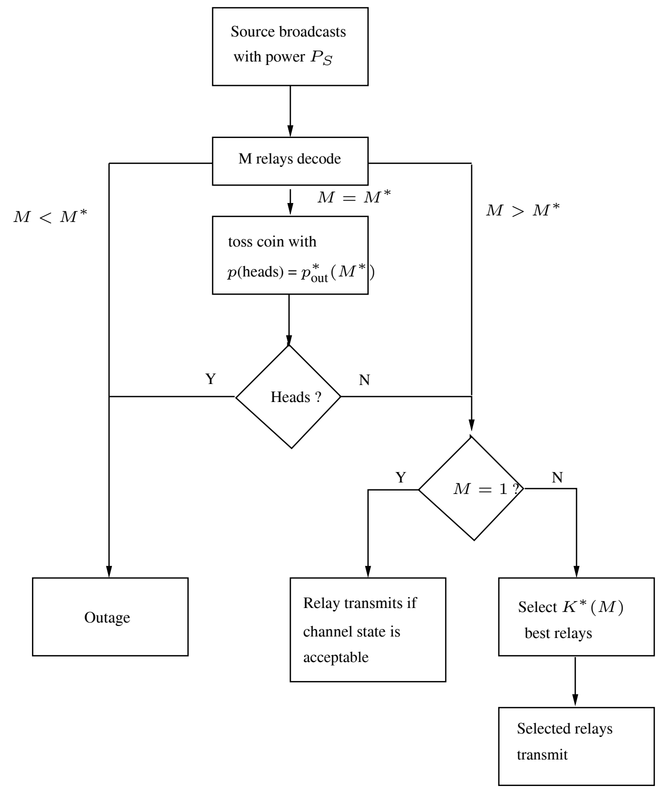


Fig. 2. Homogeneous relays: Structure of optimal transmission policy.

The structure of the optimal transmission strategy is illustrated in Fig. 2. If the number of relays, M, that decode the data successfully is less than the threshold, M^*, then the destination *always* declares outage. If $M = M^*$, the destination declares outage with probability $p^*_{\text{out}}(M^*)$. If $M > M^*$, the destination selects $K^*(M)$ relays and *never* declares outage unless $M = 1$, in which case the destination allows the relay to transmit only if its channel power gain exceeds a threshold determined by δ. In the event that the destination allows the selected relay(s) to transmit, it feedbacks CSI to the relay(s), which then transmit data with sufficient power to the destination.

The following lemma (proof omitted due to lack of space) about convexity gives a sufficient condition that allows an efficient bisection search, of complexity $O(\log_2(M))$, to be used to determine $K^*(M)$ for each M.

Lemma 3.2: For $M \geq 3$, if for all $v = 2, \ldots, M-1$,

$$\mathbb{E}\left[3\left(\frac{1}{g_{[v+1]}} - \frac{1}{g_{[v]}}\right) - \left(\frac{1}{g_{[v+1]}} - \frac{1}{g_{[v-1]}}\right)\right] \geq 0,$$

then there exists a convex function $z : \mathbb{R} \to \mathbb{R}$ such that $z(v) = T_f(v)P_f(v, M) + T_d P_d(v, M)$, for all $v \in \{2, \ldots, M\}$.

4) Computational Algorithms: In order to optimize the transmission scheme, the optimal outage probabilities, $p^*_{\text{out}}(M)$, broadcast power, P^*_S, and relay selection rule, $K^*(M)$, need to be computed. There are two main computations. First, from Lemma 3.2 and the discussion preceding it, we see that the search for $K^*(M)$ can be done efficiently for each M under certain conditions; in general the search has worst case linear complexity in M. Note that $K^*(M)$ is independent of P_S. Second, finding the optimal P_S involves an easy one-dimensional optimization over the range of P_S. For each value of P_S, the optimal outage rule can be computed efficiently using Lemma 3.1 and the corresponding total energy consumption (when the optimal relay selection rule $K^*(M)$ is used) can be computed using expressions derived in this section.

B. General Non-Homogeneous Channels

For energy consumption for CSI feedback and beamforming, we again treat the $M \geq 2$ and $M = 1$ cases separately.

1) $M > 1$ Case: The average power consumption for feedback of CSI now becomes

$$P_f(K(\mathcal{M}), \mathcal{M}) = \frac{N_0 B(2^r - 1)}{K(\mathcal{M}) + 1}\mathbb{E}\left[\frac{1}{g_{[K(\mathcal{M})]}} + \sum_{i=1}^{K(\mathcal{M})} \frac{1}{g_{[i]}}\right]. \tag{13}$$

The average total power consumed by the relays to beamform data to the destination is

$$P_d(K(\mathcal{M}), \mathcal{M}) = N_0 B\,(2^r - 1)\,\mathbb{E}\left[\frac{1}{g_{\text{sum}}}\right], \quad \text{where } g_{\text{sum}} = \sum_{i=1}^{K(\mathcal{M})} g_{[i]}. \tag{14}$$

We now use virtual branch analysis [25] to obtain expressions for $p(\mathcal{M}, P_S)$, $P_f(K(\mathcal{M}), \mathcal{M})$ and $P_d(K(\mathcal{M}), \mathcal{M})$. This is significantly more involved for the non-homogeneous case. Let $\mathcal{S}(n)$ denote the set of permutations of the set $\{1, \ldots, n\}$. Define the event A_σ that the permutation σ is the descending order of the N S-R gains and the event B_σ that the permutation σ is the descending order of M R-D gains:

$$\begin{aligned} A_\sigma &= \left(h_{[1]} = h_{\sigma(1)}, \ldots, h_{[N]} = h_{\sigma(N)}\right), \quad \sigma \in \mathcal{S}(N), \\ B_\sigma &= \left(g_{[1]} = g_{\sigma(1)}, \ldots, g_{[M]} = g_{\sigma(M)}\right), \quad \sigma \in \mathcal{S}(M). \end{aligned} \tag{15}$$

For Rayleigh fading, it follows from [25] that the probabilities of these events are given by

$$\begin{aligned} \Pr(A_\sigma) &= \prod_{i=1}^{N} \frac{1}{\bar{h}_{\sigma_i}} \left[\sum_{m=1}^{i} \frac{1}{\bar{h}_{\sigma_m}}\right]^{-1}, \ \sigma \in \mathcal{S}(N), \\ \Pr(B_\sigma) &= \prod_{i=1}^{M} \frac{1}{\bar{g}_{\sigma_i}} \left[\sum_{m=1}^{i} \frac{1}{\bar{g}_{\sigma_m}}\right]^{-1}, \ \sigma \in \mathcal{S}(M). \end{aligned} \tag{16}$$

We use the following change of variables:

$$\begin{aligned} h_{[i]} &= \sum_{n=i}^{N} W_n, \ i = 1, \ldots, N, \\ g_{[i]} &= \sum_{n=i}^{M} V_n, \ i = 1, \ldots, M. \end{aligned} \tag{17}$$

Conditioned on the event A_σ, the random variables $W_1, \ldots, W_N$ can be shown to be independent, and their joint probability density function is given by [25]

$$f_W(w_1,\ldots,w_N) = \sum_{\sigma\in\mathcal{S}(N)} \Pr(A_\sigma) f_{W|A_\sigma}(w_1,\ldots,w_N) = \sum_{\sigma\in\mathcal{S}(N)} \Pr(A_\sigma)\prod_{i=1}^{N} f_{W_i|A_\sigma}(w_i), \tag{18}$$

where the probability $f_{W_i|A_\sigma}(x)$ conditioned on A_σ is

$$f_{W_i|A_\sigma}(x) = \begin{cases} \frac{1}{\hat{h}_i(\sigma)} e^{-\frac{x}{\hat{h}_i(\sigma)}}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}, \quad i = 1,\ldots,N, \tag{19}$$

and $\hat{h}_i(\sigma) = \left[\sum_{m=1}^{i} \frac{1}{\bar{h}_{\sigma(m)}}\right]^{-1}$. Using (17) and (19), we get $f_{h_{[i]}|A_\sigma}(x) = \Phi_{1+N-i}\left(\hat{h}_i(\sigma),\ldots,\hat{h}_N(\sigma),x\right)$, where $\Phi_n(\bar{y}_1,\ldots,\bar{y}_n,x)$ denotes the probability distribution function of $Y_1+\cdots+Y_n$, and $\bar{y}_i$ is the mean of Y_i. An identical analysis can also done for $g[i]$ and V_i to show that $f_{g_{[i]}|B_\sigma}(x) = \Phi_{1+M-i}\left(\hat{g}_i(\sigma),\ldots,\hat{g}_M(\sigma),x\right)$, where $\hat{g}_i(\sigma) = \left[\sum_{m=1}^{i} \frac{1}{\bar{g}_{\sigma(m)}}\right]^{-1}$.

We can now evaluate $p(\mathcal{M},P_S)$ in (3) as follows:

$$\begin{aligned} p(\mathcal{M},P_S) &\overset{(a)}{=} \Pr\Big(\{h_{[M]} \geq \gamma_r, h_{[M+1]} < \gamma_r\} \cap \big\{A_\sigma : \sigma\in\mathcal{S}(N), \{\sigma(1),\ldots,\sigma(M)\} = \mathcal{M}\big\}\Big) \\ &\overset{(b)}{=} \sum_{\sigma:\{\sigma(1),\ldots,\sigma(M)\}=\mathcal{M}} \Pr(A_\sigma)\Pr\left(\{h_{[M+1]} < \gamma_r | A_\sigma\} \cap \{h_{[M+1]} + W_M > \gamma_r \mid A_\sigma\}\right) \\ &\overset{(c)}{=} \sum_{\sigma:\{\sigma(1),\ldots,\sigma(M)\}=\mathcal{M}} \Pr(A_\sigma)\int_0^{\gamma_r}\int_{\gamma_r-y}^{\infty} f_{h_{[M+1]}|A_\sigma}(y) \times f_{W_M|A_\sigma}(x)\,dx\,dy \\ &= \sum_{\sigma:\{\sigma(1),\ldots,\sigma(M)\}=\mathcal{M}} \Pr(A_\sigma)\int_0^{\gamma_r} \Phi_{N-M}\left(\hat{h}_{M+1}(\sigma),\ldots,\hat{h}_N(\sigma),y\right) \times \exp\left(\frac{y-\gamma_r}{\hat{h}_M(\sigma)}\right) dy. \end{aligned}$$

Step (a) holds because the first set corresponds to the event that M relays decode successfully, while the second set corresponds to the event that the relays in the set $\mathcal{M}$ are the ones with the best channel power gains from the source. Since the events A_σ are disjoint, step (b) follows from the law of total probability. Step (c) follows since, conditioned on A_σ, the random variables $h_{[M+1]}$ and W_M are independent [25]. Using the closed-form expression for Φ_{N-M}, derived in Appendix A in (26), we can now write the the expression for $p(M,P_S)$ as follows [24, 3.351.1]:

$$p(\mathcal{M},P_s) = \sum_{\sigma:\{\sigma(1),\ldots,\sigma(M)\}=\mathcal{M}} \Pr(A_\sigma) e^{-\frac{\gamma_r}{\hat{h}_M(\sigma)}} \sum_{j=1}^{m}\sum_{i=1}^{n_j} c_{ji} \times \left(\frac{1}{\left(\frac{1}{a_j}-\frac{1}{\hat{h}_M(\sigma)}\right)^i} - \sum_{k=0}^{i-1} \frac{e^{-\gamma_r\left(\frac{1}{a_j}-\frac{1}{\hat{h}_M(\sigma)}\right)}\gamma_r^k}{k!\left(\frac{1}{a_j}-\frac{1}{\hat{h}_M(\sigma)}\right)^{i-k}}\right). \tag{20}$$

For the special case, where the $\hat{h}_i(\sigma)$'s are distinct for each $\sigma\in S(N)$, we can simplify the above expression using (27) as follows:

$$p(\mathcal{M},P_S) = \sum_{\sigma:\{\sigma(1),\ldots,\sigma(M)\}=\mathcal{M}} \Pr(A_\sigma)\times \sum_{j=M+1}^{N} \frac{\hat{h}_M(\sigma)\hat{h}_j(\sigma)^{N-M-1}\left(e^{-\frac{\gamma_r}{\hat{h}_M(\sigma)}} - e^{-\frac{\gamma_r}{\hat{h}_j(\sigma)}}\right)}{(\hat{h}_M(\sigma)-\hat{h}_j(\sigma))\prod_{k=M+1,k\neq j}^{N}(\hat{h}_j(\sigma)-\hat{h}_k(\sigma))}. \tag{21}$$

To compute the average feedback power $P_f(K(\mathcal{M}),\mathcal{M})$ in (13), we need to compute $\mathbb{E}\left[\frac{1}{g_{[i]}}\right]$. This is done as follows:

$$\begin{aligned} \mathbb{E}\left[\frac{1}{g_{[i]}}\right] &= \sum_{\sigma\in\mathcal{S}(M)} \Pr(B_\sigma)\mathbb{E}\left[\frac{1}{g_{[i]}}\middle| B_\sigma\right] \\ &= \sum_{\sigma\in\mathcal{S}(M)} \Pr(B_\sigma)\mathbb{E}\left[\frac{1}{(V_i+\ldots+V_M)}\middle| B_\sigma\right] \\ &= \sum_{\sigma\in\mathcal{S}(M)} \Pr(B_\sigma)\mathcal{E}_{1+M-i}\left(\hat{g}_i(\sigma),\ldots,\hat{g}_M(\sigma)\right), \end{aligned} \tag{22}$$

where $\mathbb{E}[.|B_\sigma]$ denotes the expectation conditioned on the event B_σ, and $\Pr(B_\sigma)$ is given by (16). For computing the data power in equation (14), we need to evaluate $\mathbb{E}\left[\frac{1}{g_{\text{sum}}}\right]$. Exploiting the independence of the V_i's conditioned on B_σ, we can show that

$$\mathbb{E}\left[\frac{1}{g_{\text{sum}}}\right] = \sum_{\sigma\in\mathcal{S}(M)} \Pr(B_\sigma)\mathcal{E}_M\left(\hat{g}_1(\sigma), 2\hat{g}_2(\sigma),\ldots, K(\mathcal{M})\hat{g}_{K(\mathcal{M})}(\sigma),\ldots,K(\mathcal{M})\hat{g}_M(\sigma)\right). \tag{23}$$

2) $M=1$ Case: The analysis is identical to that for the homogeneous case. As before, we have

$$P_f(K(\{i\}),\{i\}) = P_d(K(\{i\}),\{i\}) = -\frac{N_0 B(2^r-1)}{\bar{g}_i}\text{Ei}\left(\frac{-\alpha_i}{\bar{g}_i}\right). \tag{24}$$

3) Optimal Transmission Strategy: The parameters p^*_{out}, K^*, and P^*_S for an optimal transmission scheme are the solution to the following optimization problem with corresponding variables p_{out}, K, and P_S:

$$\begin{aligned} \min.\quad & E(p_{\text{out}},K,P_S) \\ \text{s.t.}\quad & p(\emptyset,P_S) + \sum_{i=1}^{N}\delta p(\{i\},P_S)(1-p_{\text{out}}(\{i\})) + \sum_{\mathcal{M}\subseteq\{1,\ldots,N\}} p(\mathcal{M},P_S)p_{\text{out}}(\mathcal{M}) \leq P_{\text{fail}}, \\ & 0 \leq p_{\text{out}}(\mathcal{M}) \leq 1,\quad K(\mathcal{M})\in\{1,\ldots,M\}, \quad \forall\mathcal{M}\subseteq\{1,\ldots,N\}. \end{aligned}$$

As before,

$$K^*(\mathcal{M}) = \underset{1\leq K(\mathcal{M})\leq M}{\arg\min}\Big[T_f(K(\mathcal{M}))P_f(K(\mathcal{M}),\mathcal{M}) + T_d P_f(K(\mathcal{M}),\mathcal{M})\Big].$$

The following lemma on the outage probabilities, p^*_{out}, of the optimal scheme follows:

Lemma 3.3: Label the sets $\mathcal{M} \subseteq \{1,\ldots,N\}$ as $\mathcal{M}_1,\ldots,\mathcal{M}_{2^N}$ such that for all $i \le j$,

$$\begin{aligned}&\beta_{\mathcal{M}_i}(T_f(K^*(\mathcal{M}_i))P_f(K^*(\mathcal{M}_i),\mathcal{M}_i) + T_dP_d(K^*(\mathcal{M}_i),\mathcal{M}_i))\\ &\ge \beta_{\mathcal{M}_j}(T_f(K^*(\mathcal{M}_j))P_f(K^*(\mathcal{M}_j),\mathcal{M}_j)\\ &\qquad + T_dP_d(K^*(\mathcal{M}_j),\mathcal{M}_j)),\end{aligned}$$

where $\beta_{\mathcal{M}} = 1$ if $|\mathcal{M}| > 1$, and equals $1/(1-\delta)$ otherwise. Then optimal outage strategy is given as follows. For some $0 < n^* < 2^N$:

$$\begin{aligned}&p^*_{\text{out}}(\mathcal{M}_n) = 1,\ n < n^*\\ &0 \le p^*_{\text{out}}(\mathcal{M}_n) \le 1,\ n = n^*\\ &p^*_{\text{out}}(\mathcal{M}_n) = 0,\ n > n^*.\end{aligned} \quad (25)$$

Moreover,

$$\begin{aligned}&p(\emptyset,P_S) + \sum_{i=1}^{N} \delta p(\{i\},P_S)(1-p^*_{\text{out}}(\{i\}))\\ &\qquad + \sum_{\mathcal{M}\subseteq\{1,\ldots,N\}} p(\mathcal{M},P_S)p^*_{\text{out}}(\mathcal{M}) = P_{\text{fail}}.\end{aligned}$$

Proof: The proof is given in Appendix C. ■

From the above Lemma, to determine the optimal outage rule, order the sets $\mathcal{M}_i$ of relay nodes in a decreasing order of average feedback and data energy consumption times β_i. We then set $p_{\text{out}}(\mathcal{M}) = 1$ for as many states as possible, starting from the first state. Also, as expected, the outage constraint in (1) is satisfied with equality by the optimal scheme.

4) Computational Aspects: The computational algorithms in the non-homogenous case are very similar to those in the homogeneous case, but have a higher computational complexity because the relay selection rule, $K(\mathcal{M})$, and the outage rule $p_{\text{out}}(\mathcal{M})$ are now functions of the *set* of relays that successfully decodes the message broadcast by the source. Moreover, the outage rule is now a threshold rule based on an appropriate ordering of the sets $\mathcal{M}_i$'s. Optimization of P_S again involves a one-dimensional search. For each value of P_S, we can again compute the optimal outage rule, and hence compute the average total energy consumption corresponding to an optimal relay selection rule $K^*(\mathcal{M})$ (which is independent of P_S) using expressions derived in this Section.

IV. Numerical Results

A. Homogeneous Relays: Tradeoffs

Consider a cooperative relay network with $N = 10$ relays, rate, $r = 2$ bits/symbol, T_d =100 symbol durations, $\delta = 0.005$, and $P_{\text{fail}} = 0.01$. Unless otherwise mentioned, the mean channel power gains are $\bar{h} = \bar{g} = 1$. Assuming that 8 bits are required to feedback each channel power gain and phase, we have $c = 4$. For the sake of illustration, we assume that the training power, P_t, is such that it equals the power needed for transmitting from a relay to the destination at rate r and with an outage of 0.1 (which is higher than P_{fail}).[3] All the computed energy values are normalized with respect to N_0B.

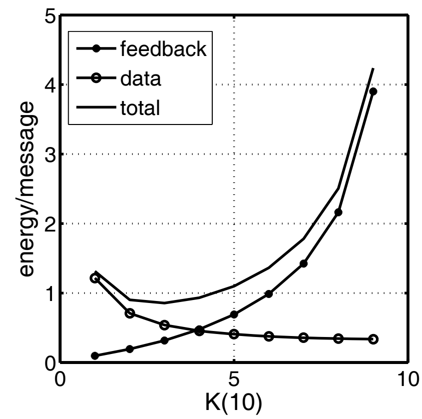


Fig. 3. Homogeneous relays: Effect of relay selection rule on energy for CSI feedback and data transmission when $M = 10$ relays decode data broadcast by source.

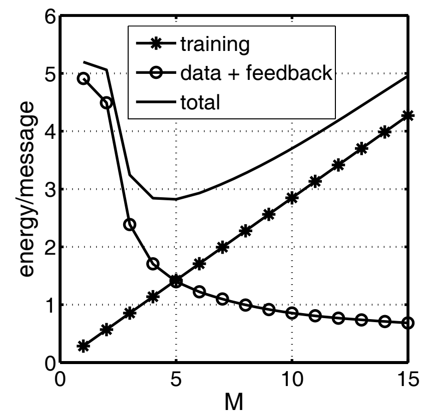
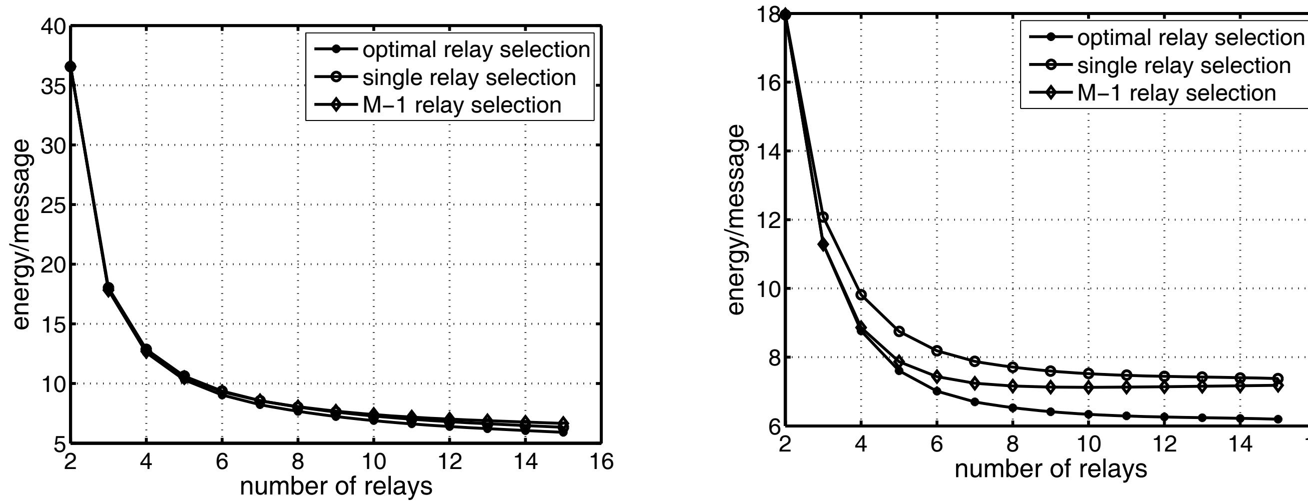


Fig. 4. Homogeneous relays: Energy for feedback and data transmission for optimal relay selection rule as a function of the number of relays that decode data broadcast by source.

Figure 3 shows the variation with $K(M)$ of the energy for feedback of CSI and energy for data transmission from the relays to the destination, when $M = 10$ relays receive the data broadcast from the source. As the training power does not change when M is fixed, it is not shown. In this case, $K(10)$ denotes the number of relays selected by the destination. We see that as $K(10)$ increases, the energy consumption for CSI feedback increases because the destination has to feedback to more relays with progressively worse channels. At the same time, the energy consumption for data transmission decreases because more relays now beamform to forward the data to the destination. Also, the total feedback and data power consumption as a function of K can be fitted to a convex function, as in Lemma 3.2. The computational results were verified by Monte-Carlo simulations of the system using 10^8 samples.

Figure 4 shows the variation of the energy consumed, as a function of M, for training and for cooperative beamforming and feedback of CSI for an optimal relay selection rule. As more relays decode the data from source, the power consumption for feedback and data transmission decreases

[3]This is justifiable because transmit diversity enables us to use a relay only when its channel to the destination is good. Hence, if the training sequence received at the destination has low power, it means that the channel is bad and, hence, will not be used for data transmission.



(a) S-R and R-D channels both have the same mean gains, $\bar{h} = \bar{g} = 1.0$

(b) S-R channel power gains have mean, $\bar{h} = 6$, and R-D channel power gains have mean, $\bar{g} = 0.5$

Fig. 5. Homogeneous relays: Energy consumption as a function of number of relays for different relay selection rules.

due to greater diversity in the system. However, this also increases the training overhead. The optimal relay selection rule (not shown in the figure) turns out to be the following: for $M \leq 2$, $K^*(M)$ equals 1, which is the conventional single relay selection. However, single relay selection is sub-optimal for larger M as for $3 \leq M \leq 6$, $K^*(M) = 2$, and for $7 \leq M \leq 15$, $K^*(M) = 3$. Hence, only a small subset of the relays – that changes depending on the fading on the relay links – is active at any given time. Therefore, while the energy cost of acquiring CSI limits the number of relays that cooperate at any instant, it is still beneficial to cooperate.

B. Homogeneous Relays: Relay Selection Rules

In Fig. 5, we show the energy consumption per message of the following three rules for selecting relays as a function of M, in two different SNR regimes: *optimal relay selection*, in which the best $K^*(M)$ relays are chosen, *single relay selection*, in which only one relay with the highest R-D gain is chosen, and $(M-1)$ *relay selection*, in which the best $(M-1)$ relays are always chosen. (The rule which selects all M relays is not shown as it will incur infinite average energy consumption for CSI feedback.) Figure 5(a) shows the energy consumption per message when all the S-R and R-D channels have mean gains of 1. As we saw in the previous subsection, in this regime, the energy to broadcast from the source to the relays is a significant component of the total energy. Hence, the performance of all relay selection rules is about the same. Figure 5 (b) plots the energy consumption per message when the S-R channels have mean channel power gains of 6, while the R-D channels have mean channel power gains of 0.5. This corresponds to the case where the relays are closer to the source. Relay selection now has a bigger impact on the total energy consumption. The optimal relay selection rule consumes approximately 16% less energy than the other two selection rules when 15 relays are present. Thus we can clearly see the gains obtained by varying the number of relays as a function of system parameters, and as the number of relays that decode varies with time.

C. General Non-Homogeneous Relays

We now consider the general case in which the mean channel power gains from the source to the relays, $\{\bar{h}_i\}_{i=1}^N$, and from the relays to the destination,$\{\bar{g}_i\}_{i=1}^N$, are not identical. As mentioned, this occurs due to different pathloss and lognormal shadowing for the different relays. Moreover, the lognormal shadowing for different links may be correlated [27]. In general, the (fading-averaged) means take the form $\bar{h}_i = 10^{0.1X_i}$ and $\bar{g}_i = 10^{0.1Y_i}$, where X_i and Y_i are Gaussian random variables with means μ_{X_i} and μ_{Y_i}, and standard deviations σ_{X_i} and σ_{Y_i}. Moreover, X_i and Y_i are correlated, with a correlation coefficient ρ.[4] Varying ρ provides an excellent mechanism to study the impact of the differences in the S-R and R-D link gains. Its importance was recognized in [16], which proposed relay selection metrics that took both link states into account.

Shadow fading varies at a much slower rate than Rayleigh fading. Thus the averages over Rayleigh fading states are computed given the shadow fading means. For each instantiation of shadow fading values, the optimal transmission scheme is evaluated, which involves determining, using the results of Sec. III, the optimal broadcast power, the optimal relay selection rule, and the optimal outage rule. The fading-averaged energy consumed is then computed from the formulae derived in Sec. III. Note that without our analysis framework, for a given instance of the shadowing values, brute-force Monte-Carlo simulations for all possible transmission schemes would have been the only recourse to find the optimal transmission scheme.

Figure 6 plots the cumulative distribution function (CDF) of the total fading-averaged energy consumed for $N = 3$ relays

[4]The general case in which the shadowing on all the relay links are correlated can also be easily analyzed from the results derived in this paper.

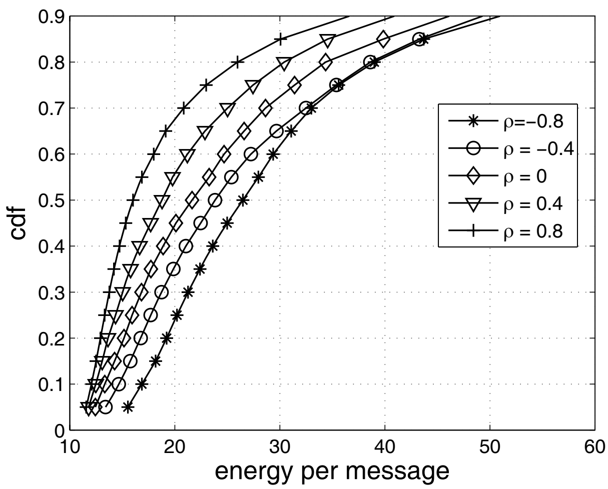


Fig. 6. Non-homogeneous relays: CDF of average energy consumed per message as a function of shadowing correlation between source-relay and relay-destination links ($N = 3$).

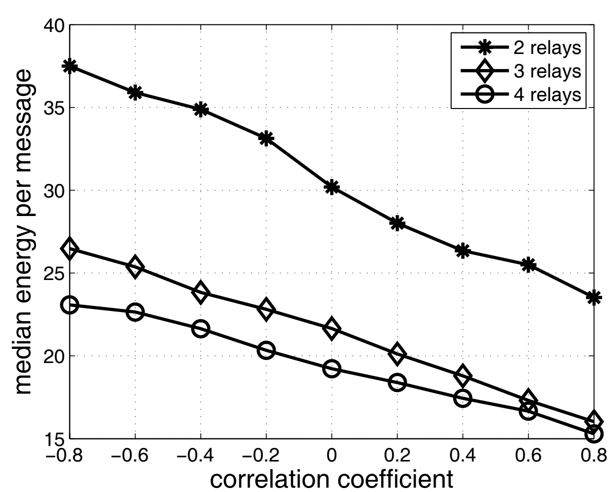


Fig. 7. Non-homogeneous relays: Median energy consumption for different numbers of relays as a function of shadowing correlation between source-relay and relay-destination links.

for different correlation coefficients, ρ. Plotting the CDF is very instructive because the CDF provides information about the entire energy distribution. For all the relays, the shadowing parameters were set as follows: $\mu_{X_i} = \mu_{Y_i} = 0$ dB and $\sigma_{X_i} = \sigma_{Y_i} = 6$ dB. As the correlation coefficient increases from -1.0 (perfect anti-correlation) to 1.0 (perfect correlation), the energy consumed decreases – the CDF curves shift to the left. This can be intuitively explained as follows. In the anti-correlated case, if a relay node decodes the message from the source, i.e., has a high S-R channel power gain, it is likely to have a low R-D channel power gain. Thus, with high probability, a large amount of energy is required to transmit a message from the relays to the destination. In contrast, for the perfectly correlated case, the relay nodes that decode the broadcast message are also likely to be the best ones for forwarding the message to the destination.

The effect of the number of relays on energy consumed is shown in Fig. 7, which plots the median energy as a function of ρ for $N = 2$, 3, and 4 relays. As was observed in the homogeneous case, the energy consumed decreases as the total number of relays increases, due to the greater diversity in the cooperative relay network.

V. Conclusions

We analyzed the total energy consumption for a general class of cooperative beamforming-based transmission schemes. The overhead energy consumption for obtaining the CSI was explicitly modeled. The relay selection rules considered in this paper generalize the ones considered in the literature, and yet are amenable to analysis and optimization. The properties of the optimal outage and relay selection rules derived in this paper make their computation efficient and feasible. Online computation can now be done for a small number of nodes by using Ei function tables. For the homogeneous case, the optimal transmission scheme has a very simple structure and can be computed online efficiently for a large number of relay nodes. The numerical results illustrated the tradeoff between decreasing energy consumption for data transmission and decreasing overhead energy consumption for CSI acquisition. Thus, the right amount of instantaneous local cooperation should be used to obtain maximum energy savings. The optimal cooperative communication scheme reduces energy consumption compared to non-cooperative schemes and cooperative schemes that use either a single relay or all available relays; energy savings up to 16% were observed after CSI energy overhead was accounted for.

We note that the current paper has been primarily concerned with the minimization of the total expended energy. The relay selection rule also influences the spectral efficiency in a very simple (linear) fashion, as an increase in the number of selected relays leads to larger (temporal) overhead for the transmission. Consequently, tradeoffs are possible between energy consumption and spectral efficiency.

Another model of interest to which our analysis easily applies is one in which the energy consumed by different components in the system is weighted differently. This models, for example, different levels of importance of energy efficiency in different system components. The weights will affect the optimal relay selection rule – the smaller the weight for CSI acquisition and feedback, the greater the number of relays selected for beamforming.

Acknowledgment

The authors thank Dr. P. Orlik for many helpful discussions.

Appendix A: Closed-Form Evaluation of $\mathcal{E}_n$

Recall that the probability distribution function of $X = Y_1 + \ldots Y_n$ is denoted by $\Phi_n(\bar{y}_1, \ldots, \bar{y}_n, x)$, where $\bar{y}_i$ is the mean of Y_i. Since Y_i are independent exponential random variables, the Laplace transform, $\mathcal{L}f$, of $f_X(x)$ is given by $(\mathcal{L}f)(s) = \prod_{i=1}^{n} \frac{1}{\bar{y}_i s + 1}$. Hence, we can evaluate $\Phi_n(\bar{y}_1, \ldots, \bar{y}_n, x)$ using a partial fraction expansion. In particular, for $n_1 + \ldots + n_m = n$, if

$$\{\bar{y}_1, \ldots, \bar{y}_n\} = \{\underbrace{a_1, \ldots, a_1}_{n_1 \text{ terms}}, \underbrace{a_2, \ldots, a_2}_{n_2 \text{ terms}}, \ldots, \underbrace{a_m, \ldots, a_m}_{n_m \text{ terms}}\},$$

then, for $x \geq 0$,

$$\Phi_n(a_1,\ldots,a_1,\ldots,a_m,\ldots,a_m,x) = \sum_{i=1}^{n_1} \frac{c_{1i}x^{i-1}e^{-x/a_1}}{(i-1)!} + \ldots + \sum_{i=1}^{n_m} \frac{c_{mi}x^{i-1}e^{-x/a_m}}{(i-1)!}, \quad (26)$$

where

$$c_{ji} = \frac{1}{(n_j-i)!\prod_{k=1}^m \prod_{l=1}^{n_k} a_l} \times \frac{d^{n_j-i}}{ds^{n_j-i}}(s+1/a_j)^{n_j} \prod_{k=1}^m \prod_{l=1}^{n_k} \frac{1}{s+1/a_l}\bigg|_{s=-1/a_j}.$$

When all the $\bar{y}_i$'s are distinct, we have, for $x \geq 0$,

$$\Phi_n(\bar{y}_1,\ldots,\bar{y}_n,x) = \sum_{j=1}^{n} \frac{\bar{y}_j^{n-2}e^{-x/\bar{y}_j}}{\prod_{k=1,k\neq j}^n (\bar{y}_j - \bar{y}_k)}. \quad (27)$$

The mean of $1/X$, denoted by $\mathcal{E}_n(\bar{y}_1,\ldots,\bar{y}_n)$, equals $\mathbb{E}\left[\frac{1}{X}\right] = \mathcal{E}_n(\bar{y}_1,\ldots,\bar{y}_n) = \int_0^\infty \frac{1}{x}\Phi_n(\bar{y}_1,\ldots,\bar{y}_n,x)\ dx$.

It is well known that the above integral is infinite for $n=1$ (see, for example, [21]). Now, for $n \geq 2$, we can write for $0 < \eta < \min_j a_j$ [24, 3.351.3]

$$\begin{aligned}\mathcal{E}_n(y_1,\ldots,y_n) &= \int_0^\eta \frac{1}{x}\sum_{j=1}^m c_{j1}e^{-x/a_j}dx + \int_\eta^\infty \frac{1}{x}\sum_{j=1}^m c_{j1}e^{-x/a_j}dx \\ &\quad + \int_0^\infty \sum_{j=1}^m \sum_{i=2}^{n_j} \frac{c_{ji}x^{i-2}e^{-x/a_j}}{(i-1)!}dx \\ &= \int_0^\eta \frac{1}{x}\sum_{j=1}^m c_{j1}e^{-x/a_j}dx + \sum_{j=1}^m c_{j1}\mathrm{Ei}\left(\frac{-\eta}{a_j}\right) \\ &\quad + \sum_{j=1}^m \sum_{i=2}^{n_j} \frac{c_{ji}a_j^{i-1}}{(i-1)},\end{aligned}$$

where Ei is the standard exponential integral function [24] given by $\mathrm{Ei}(u) = \int_{-\infty}^u \frac{e^x}{x}dx$. Using the initial value theorem and continuity of $\Phi(y_1,\ldots,y_n,x)$ as a function of x, it follows that $\Phi(y_1,\ldots,y_n,0) = \sum_{j=1}^m c_{j1} = 0$ for $n \geq 2$. Hence, truncating the Taylor series expansion to k terms incurs an error

$$\left|\int_0^\eta \frac{1}{x}\sum_{j=1}^m c_{j1}e^{-x/a_j}dx - \sum_{j=1}^m c_{j1}\left(\sum_{l=1}^k (-1)^l \frac{\eta^l}{a_j^l l! l}\right)\right| \leq \sum_{j=1}^m \frac{c_{j1}}{(k+1)!(k+1)}\left(\frac{\eta}{a_j}\right)^{k+1}.$$

The error $\to 0$ as $k \to \infty$. Thus, the function $\mathcal{E}_n(y_1,\ldots,y_n)$ is finite for $n \geq 2$. Moreover, it can be analytically approximated in terms of the Ei function to an arbitrary degree of accuracy by choosing k large enough.

APPENDIX B: PROOF OF LEMMA 3.1

Proof: We will show that for $L \geq M > 1$

$$\begin{aligned}&T_f(K^*(M))P_f(K^*(M),M) + T_dP_d(K^*(M),M) \\ &\quad \geq T_f(K^*(L))P_f(K^*(L),L) + T_dP_d(K^*(L),L). \quad (28)\end{aligned}$$

Let $\hat{K}$ be another relay selection rule such that $\hat{K}(M) = K^*(M)$ and $\hat{K}(L) = K^*(M)$, where $L > M$. Then, we have

$$\begin{aligned}&T_f(K^*(M))P_f(K(M),M) + T_dP_d(K(M),M) \\ &\overset{(a)}{=} T_f(\hat{K}(M))P_f(\hat{K}(M),M) + T_dP_d(\hat{K}(M),M) \\ &\overset{(b)}{\geq} T_f(\hat{K}(L))P_f(\hat{K}(L),L) + T_dP_d(\hat{K}(L),L) \\ &\overset{(c)}{\geq} T_f(K^*(L))P_f(K^*(L),L) + T_dP_d(K^*(L),L).\end{aligned}$$

Step (a) is by assumption. Step (b) follows from (6) and (7), and the following two inequalities: $\mathbb{E}_M\left[\frac{1}{g_{[i]}}\right] \geq \mathbb{E}_L\left[\frac{1}{g_{[i]}}\right]$ and $\mathbb{E}_M\left[\frac{1}{\sum_{i=1}^K g_{[i]}}\right] \geq \mathbb{E}_L\left[\frac{1}{\sum_{i=1}^K g_{[i]}}\right]$, which hold whenever $K \leq L$ and $L \geq M$. Here, $\mathbb{E}_M\left[\frac{1}{g_{[i]}}\right]$ denotes the mean of $\frac{1}{g_{[i]}}$, where $g_{[i]}$ is the ith ordered channel gain of M i.i.d. channel power gains. Step (c) follows because K^* is an optimal relay selection rule.

Combining this with (10), we see that the optimal feedback and data power consumption conditioned on M is a decreasing function of M. This is sufficient to prove the Lemma. We prove this for the general case of non-homogenous channels in Appendix C, and do not prove the special case here for brevity and to avoid repetition. ■

APPENDIX C: PROOF OF LEMMA 3.3

Proof: Let

$$c_{\mathcal{M}_i} = T_f(K^*(\mathcal{M}_i))P_f(K^*(\mathcal{M}_i),\mathcal{M}_i) + T_dP_d(K^*(\mathcal{M}_i),\mathcal{M}_i),$$

$$b = P_{\text{fail}} - p(\emptyset,P_S) - \sum_{i=1}^N \delta p(\{i\},P_S).$$

Given $P_f(K^*(\mathcal{M}_i),\mathcal{M}_i)$ and $P_d(K^*(\mathcal{M}_i),\mathcal{M}_i)$, for all sets $\mathcal{M}_i$, optimizing $p_{\text{out}}(\mathcal{M}_i)$ is equivalent to following linear programming problem with variables $p_{\text{out}}(\mathcal{M}_i)$ (see (1) and (2)):

$$\begin{aligned}\text{minimize} \quad & \sum_{i=1}^{2^N} \frac{p(\mathcal{M}_i,P_S)}{\beta_{\mathcal{M}_i}}((1-p_{\text{out}}(\mathcal{M}_i))\beta_{\mathcal{M}_i}c_{\mathcal{M}_i}, \\ \text{subject to} \quad & \sum_{i=1}^{2^N} \frac{p(\mathcal{M}_i,P_S)}{\beta_{\mathcal{M}_i}}p_{\text{out}}(\mathcal{M}_i) \leq b, \\ & 0 \leq p_{\text{out}}(\mathcal{M}_i) \leq 1, \quad \text{for all } i=1,\ldots,2^N,\end{aligned}$$

where $\beta_{\mathcal{M}_i}$ are as defined in the Lemma. The first constraint holds with equality because the objective is a strictly decreasing function of each of the $p_{\text{out}}(\mathcal{M}_i,P_S)$'s. By assumption, $\beta_{\mathcal{M}_i}c_{\mathcal{M}_i} \geq \beta_{\mathcal{M}_j}c_{\mathcal{M}_j}$ for $i < j$.

Consider an outage scheme p_{out} in which

$$p_{\text{out}}(\mathcal{M}_l) = \epsilon_l \quad \text{and} \quad p_{\text{out}}(\mathcal{M}_j) = \epsilon_j, \quad 0 \leq \epsilon_l < 1,\ 0 < \epsilon_j \leq 1, \quad (29)$$

for some $l < j$ such that $\beta_{\mathcal{M}_l}c_{\mathcal{M}_l} > \beta_{\mathcal{M}_j}c_{\mathcal{M}_j}$. It can be shown that changing the outage probabilities to p^1_{out}, where $p^1_{\text{out}}(\mathcal{M}_k) = p_{\text{out}}(\mathcal{M}_k)$, for $k \neq l,j$, and

$$p^1_{\text{out}}(\mathcal{M}_l) = \beta_{\mathcal{M}_l}\min\left(1, \epsilon_l + \frac{\epsilon_j p(\mathcal{M}_j,P_S)\beta_{\mathcal{M}_l}}{p(\mathcal{M}_l,P_S)\beta_{\mathcal{M}_j}}\right),$$

$$p^1_{\text{out}}(\mathcal{M}_j) = \epsilon_j - \frac{\beta_{\mathcal{M}_j}p(\mathcal{M}_l,P_S)(p^1_{\text{out}}(\mathcal{M}_l)-\epsilon_l)}{\beta_{\mathcal{M}_l}p(\mathcal{M}_j,P_S)},$$

leads to

$$\sum_{i=1}^{2^N} p(\mathcal{M}_i, P_S)((1 - p^1_{\text{out}}(\mathcal{M}_i))c_{\mathcal{M}_i} < \sum_{i=1}^{2^N} p(\mathcal{M}_i, P_S)((1 - p_{\text{out}}(\mathcal{M}_i))c_{\mathcal{M}_i}.$$

Hence, any p_{out}, for which (29) holds for some $i < j$ and $\beta_{\mathcal{M}_l}\mathcal{M}_l > \beta_{\mathcal{M}_j}\mathcal{M}_j$, cannot be optimal. If

$$\beta_{\mathcal{M}_l}\mathcal{M}_l = \beta_{\mathcal{M}_j}\mathcal{M}_j,$$

then p_{out} and p^1_{out} have the same performance. Hence, the result. ■

REFERENCES

[1] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Commun. Mag.*, vol. 42, pp. 68–73, 2004.
[2] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity–part I: system description," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1938, 2003.
[3] 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*, vol. 50, pp. 3062–3080, 2004.
[4] Y. Chen and Q. Zhao, "On the lifetime of wireless sensor networks," *IEEE Commun. Lett.*, vol. 9, pp. 976–978, Nov. 2005.
[5] A. Wittneben, I. Hammerstroem, and M. Kuhn, "Joint cooperative diversity and scheduling in low mobility wireless networks," in *Proc. Globecom*, pp. 780–784, 2004.
[6] A. E. Khandani, J. Abounadi, E. Modiano, and L. Zheng, "Cooperative routing in wireless networks," in *Proc. Allerton Conf. on Commun., Control and Computing*, 2003.
[7] B. Rankov and A. Wittneben, "Distributed spatial multiplexing in a wireless network," in *Proc. Asilomar Conf. on Signals, Systems, and Computers*, pp. 1932–1937, 2004.
[8] G. Barriac, R. Mudumbai, and U. Madhow, "Distributed beamforming for information transfer in sensor networks," in *Proc. 3rd International Symposium on Information Processing in Sensor Networks (IPSN)*, 2004.
[9] H. Ochiai, P. Mitran, H. V. Poor, and V. Tarokh, "Collaborative beamforming for distributed wireless ad hoc sensor networks," *IEEE Trans. Signal Processing*, vol. 53, pp. 4110–4124, Nov. 2005.
[10] M. M. Abdallah and H. C. Papadopoulos, "Beamforming algorithms for decode-and-forward relaying in wireless networks," in *Proc. Conf. on Inform. Sciences and Systems (CISS)*, 2005.
[11] P. Larsson, "Large-scale cooperative relay network with optimal coherent combining under aggregate relay power constraints," in *Proc. Future Telecommun. Conf. (FTC)*, 2003.
[12] J. Luo, R. S. Blum, L. J. Cimini, L. J. Greenstein, and A. M. Haimovich, "Link-failure probabilities for practical cooperative relay networks," in *Proc. VTC*, pp. 1489–1493, 2005.
[13] A. Stefanov and E. Erkip, "Cooperative space-time coding for wireless networks," *IEEE Trans. Commun.*, vol. 53, pp. 1804–1809, Nov. 2005.
[14] M. Chen, S. Serbetli, and A. Yener, "Distributed Power Allocation for Parallel Relay Networks," in *Proc. IEEE Globecom*, pp. 1177–1181, 2005.
[15] I. Maric and R. D. Yates, "Forwarding strategies for Gaussian parallel-relay networks," in *Proc. ISIT*, pp. 269–275, 2004.
[16] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Select. Areas Commun.*, vol. 24, pp. 659–672, Mar. 2006.
[17] Z. Lin, E. Erkip, and A. Stefanov, "Cooperative regions for coded cooperative systems," in *Proc. Globecom*, pp. 21–25, 2004.
[18] Z. Lin and E. Erkip, "Relay search algorithms for coded cooperative systems," in *Proc. Globecom*, pp. 1314–1319, 2005.
[19] A. Stefanov and E. Erkip, "Cooperative coding for wireless networks," *IEEE Trans. Commun.*, vol. 52, pp. 1470–1476, Sept. 2004.
[20] N. B. Mehta and A. F. Molisch, "Antenna selection in MIMO systems," in *MIMO System Technology for Wireless Communications* (G. Tsulos, ed.), ch. 6, CRC Press, 2006.
[21] A. F. Molisch, *Wireless Communications*. Wiley/IEEE Press, 2005.
[22] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, 2005.
[23] A. J. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.
[24] L. S. Gradshteyn and L. M. Ryzhik, *Tables of Integrals, Series and Products*. Academic Press, 1994.
[25] M. Z. Win and J. H. Winters, "Analysis of hybrid selection/maximal-ratio combining of diversity branches with unequal SNR in Rayleigh fading," in *Proc. VTC*, pp. 215–220, 1999.
[26] M. Z. Win and J. H. Winters, "Virtual branch analysis of symbol error probability for hybrid selection/maximal-ratio combining in Rayleigh fading," *IEEE Trans. Commun.*, pp. 1926–1934, Nov. 2001.
[27] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electron. Lett.*, vol. 27, pp. 2145–2146, Nov. 1991.
[28] S. A. Jafar, S. Srinivasa, "The Optimality of Transmit Beamforming: A Unified View," *IEEE Trans. Inform. Theory*, vol. 53, pp. 1558–1564, 2007.
[29] T. K. Y. Lo, "Maximum ratio transmission," *IEEE Trans. Commun.*, vol. 47, pp. 1458–1461, 1999.
[30] D. Gesbert, M. Shafi, D.-S. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 281–302, 2003.
[31] J.B. Andersen, "Array gain and capacity for known random channels with multiple element arrays at both ends," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2172–2178, 2000.

Ritesh Madan is at Qualcomm-Flarion Technologies, Bridgewater, NJ. He received a Ph.D. in 2006 and a M.S. in 2003, from Stanford University, and a B.Tech. from IIT Bombay in 2001, all in Electrical Engineering. He was a recipient of the Sequoia Capital Stanford Graduate Fellowship. He has held visiting research positions at Mitsubishi Electric Research Labs (MERL), Cambridge, MA, and at the Tata Institute of Fundamental Research (TIFR), Mumbai. His research interests include wireless networks, convex optimization, networked control, and dynamic programming.

Neelesh B. Mehta (S98-M01-SM06) received his Bachelor of Technology degree in Electronics and Communications Engineering from the Indian Institute of Technology, Madras in 1996, and his M.S. and Ph.D. degrees in Electrical Engineering from the California Institute of Technology, Pasadena, CA in 1997 and 2001, respectively. He was a visiting graduate student researcher at Stanford University in 1999 and 2000.

He is now an Assistant Professor at the Dept. of Electrical Communication Engineering, Indian Institute of Science (IISc), Bangalore, India. Until 2002, he was a research scientist in the Wireless Systems Research group in AT&T Laboratories, Middletown, NJ. In 2002-2003, he was a Staff Scientist at Broadcom Corp., Matawan, NJ, and was involved in GPRS/EDGE cellular handset development. From 2003-2007, he was a Principal Member of Technical Staff at the Mitsubishi Electric Research Laboratories, Cambridge, MA, USA. His research includes work on link adaptation, multiple access protocols, WCDMA downlinks, system-level analysis and simulation of cellular systems, MIMO and antenna selection, and cooperative communications. He is also actively involved in radio access network physical layer (RAN1) standardization activities in 3GPP. He is a TPC co-chair for the Frontiers of Networking and Communications symposium of Chinacom 2008, and is a TPC member of Globecom 2007, Globecom 2008, WCNC 2008, and VTC 2008 (Spring).

Andreas F. Molisch (S'89, M'95, SM00, F'05) received the Dipl. Ing., Dr. techn., and habilitation degrees from the Technical University Vienna (Austria) in 1990, 1994, and 1999, respectively. From 1991 to 2000, he was with the TU Vienna, becoming an associate professor there in 1999. From 2000-2002, he was with the Wireless Systems Research Department at AT&T (Bell) Laboratories Research in Middletown, NJ. Since then, he has been with Mitsubishi Electric Research Labs, Cambridge, MA, where he is now a Distinguished Member of Technical Staff and Chief Wireless Standards Architect. He is also Professor and Chairholder for Radio Systems at Lund University, Sweden.

Dr. Molisch has done research in the areas of SAW filters, radiative transfer in atomic vapors, atomic line filters, smart antennas, and wideband systems. His current research interests are measurement and modeling of mobile radio channels, UWB, cooperative communications, and MIMO systems. Dr. Molisch has authored, co-authored or edited four books (among them the recent textbook *Wireless Communications* Wiley-IEEE Press), eleven book chapters, some 110 journal papers, and numerous conference contributions.

Dr. Molisch is an editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, co-editor of recent or upcoming special issues in the JOURNAL OF WIRELESS COMMUNICATIONS AND MOBILE COMPUTING, IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, and the PROCEEDINGS OF THE IEEE. He has been member of numerous TPCs, vice chair of the TPC of VTC 2005 spring, general chair of ICUWB 2006, TPC co-chair of the wireless symposium of Globecomm 2007, and TPC chair of Chinacom2007. He has participated in the European research initiatives "COST 231," "COST 259," and "COST273," where he was chairman of the MIMO channel working group, he was chairman of the IEEE 802.15.4a channel model standardization group, and is also chairman of Commission C (signals and systems) of URSI (International Union of Radio Scientists). Dr. Molisch is a Fellow of the IEEE, Fellow of the IET, an IEEE Distinguished Lecturer, and recipient of several awards.

Jinyun Zhang received her B.Sc. degree in radio electronics from Tsinghua University, Beijing, China in 1970. Following her graduation, she was with Tsinghua University until 1984. She received her Ph.D. degree in electrical engineering from University of Ottawa, Canada in 1991. Dr. Zhang then joined Nortel Networks, where she held various management positions and engineering positions of increasing responsibility in the areas of digital signal processing, wireless communication and optical networks. Since 2001, Dr. Zhang has been the Manager of the Digital Communications & Networking Group at Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA, USA. Currently, she is leading many new broadband wireless communications and networking research projects that include UWB, ZigBee ad hoc networking, MIMO, broadband multimedia home networking, wireless sensor networks, high speed WLAN, cooperative communications, WiMAX and next generation mobile communications systems.

Dr. Zhang has authored and co-authored more than 110 publications, invented and co-invented more than 80 patents and patent applications, and made numerous contributions to international wireless communications standards. Dr. Zhang is an Associate Editor of IEEE TRANSACTIONS ON BROADCASTING, and has served as a Technical Program Committee member for various IEEE conferences. Dr. Zhang is a Fellow of the IEEE and a member of the IEEE IT, LEO, AP, BT, COMM, SP, VT, and ITS Societies.