

The Relay Selection and Transmission Trade-off in Cooperative Communication Systems

Shah, V.; Mehta, N.B.; Yim, R.

TR2010-106 August 2010

Abstract

A common and practical paradigm in cooperative communication systems is the use of a dynamically selected best relay to decode and forward information from a source to a destination. Such systems use two phases a relay selection phase, in which the system uses transmission time and energy to select the best relay, and a data transmission phase, in which it uses the spatial diversity benefits of selection to transmit data. In this paper, we derive closed-form expressions for the overall throughput and energy consumption, and study the time and energy trade-off between the selection and data transmission phases. To this end, we analyze a baseline non-adaptive system and several adaptive systems that adapt the selection phase, relay transmission power, or transmission time. Our results show that while selection yields significant benefits, the selection phases time and energy overhead can be significant. In fact, at the optimal point, the selection can be far from perfect, and depends on the number of relays and the mode of adaptation. The results also provide guidelines about the optimal system operating point for different modes of adaptation. The analysis also sheds new insights on the fast splitting-based algorithm considered in this paper for relay selection.

IEEE Transactions on Wireless Communications

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.

The Relay Selection and Transmission Trade-off in Cooperative Communication Systems

Virag Shah, *Student Member, IEEE*, Neelesh B. Mehta, *Senior Member, IEEE*, Raymond Yim, *Member, IEEE*

Abstract—A common and practical paradigm in cooperative communication systems is the use of a dynamically selected ‘best’ relay to decode and forward information from a source to a destination. Such systems use two phases – a relay selection phase, in which the system uses transmission time and energy to select the best relay, and a data transmission phase, in which it uses the spatial diversity benefits of selection to transmit data. In this paper, we derive closed-form expressions for the overall throughput and energy consumption, and study the time and energy trade-off between the selection and data transmission phases. To this end, we analyze a baseline non-adaptive system and several adaptive systems that adapt the selection phase, relay transmission power, or transmission time. Our results show that while selection yields significant benefits, the selection phase’s time and energy overhead can be significant. In fact, at the optimal point, the selection can be far from perfect, and depends on the number of relays and the mode of adaptation. The results also provide guidelines about the optimal system operating point for different modes of adaptation. The analysis also sheds new insights on the fast splitting-based algorithm considered in this paper for relay selection.

Index Terms—Cooperative communications, fading channels, relays, relay selection, splitting algorithms, trade-off, outage, adaptation, throughput, energy consumption

I. INTRODUCTION

RELAY-BASED multi-hop cooperation, in which a source node transmits information to the destination with the help of a relay selected from the available nodes in the system, has attracted considerable attention in the literature [1]–[3]. The relay is selected depending on its instantaneous channel gains on the basis of a real-valued metric that is a function of the relay-destination (RD) channel gain or the source-relay (SR) channel gain or both. The function depends on the cooperation protocol. Relay selection has been shown to help the system exploit the spatial diversity afforded by having geographically spaced multiple relays. It improves the symbol error probability [1] or increases the data transmission rate [4]. Another benefit is the reduction in energy consumption [5],

Manuscript received April 5, 2009; revised January 1, 2010 and April 14, 2010; accepted May 1, 2010. The associate editor coordinating the review of this letter and approving it for publication was Prof. E. Hossain.

V. Shah is with the Indian Institute of Technology (IIT), Bombay, India. He was with the Indian Institute of Science (IISc), Bangalore, India during the course of this work. (e-mail: virag4u@gmail.com).

N. B. Mehta is with the Electrical Communication Eng. Dept. at IISc. (e-mail: nbmehta@ece.iisc.ernet.in).

R. Yim is with the Mitsubishi Electric Research Labs (MERL), Cambridge, MA, USA. (e-mail: yim@merl.com).

The work by V. Shah and N. B. Mehta was partially supported by a Boeing-ANRC sponsored research project at IISc.

A part of this work has appeared in IEEE Global Communications Conference (Globecom), Honolulu, USA, Nov. 2009.

Digital Object Identifier 10.1109/TWC.2010.090487.

which is important, for example, for cooperative networks that use battery powered nodes. In general, the extent and nature of the benefits from selection depend on both the selection criteria and the cooperation scheme [6]–[10].

After the source broadcasts its data to the relays, these cooperative systems typically use two core phases to complete the transmission to the destination: (i) a *relay selection phase*, in which the ‘best’ relay with the highest metric is chosen by a selection mechanism, and (ii) a *data transmission phase*, in which data is transmitted to the destination by the selected relay. The selection phase is needed because the source does not know a priori which relay is the best one. Furthermore, since the metric is a function of local channel gains, each relay knows only its metric, and not that of the others.

In effect, the system expends time and energy in the selection phase and reaps the benefits of selection, in the form of increased throughput or lower energy consumption, in the data transmission phase. The simulations in [5], [11], which modeled several practical aspects of a contention-based selection process, indicate that the relative fraction of time and energy spent in the relay selection phase can be considerable. This overhead, and thus, the overall system performance, clearly depend on the selection mechanism. For example, in a centralized polling mechanism, the time for selection increases linearly with the number of available relays. In [11], the source node uses overhead handshaking messages to exhaustively track the rate that each candidate relay can support. The selection phase overhead can be reduced by using distributed mechanisms based on back-off timers [12], [13] or time-slotted splitting algorithms [14]–[16].

Furthermore, the selection process may not always find the best relay. For example, the system may terminate the selection phase after a pre-determined time even if the best relay has not been selected. This leads to an outage during the subsequent data transmission phase. While increasing the selection phase duration reduces this outage probability, it does so at the expense of the overall system throughput since a smaller fraction of time is used for data transmission. Doing so also increases the energy consumed in the relay selection phase. Thus, the two phases affect each other, and cannot be optimized in isolation.

A comprehensive system-level analysis and optimization that considers the trade-offs between the two phases is, thus, relevant and essential, and is the focus of this paper. While the individual performances of the two phases have been analyzed, interesting open problems remain about a system-level analysis. For example, the possibility of outage and rate adaptation was analyzed in [17], but the time and energy

overhead of multiple access selection was not accounted for. The analysis in [5] focused on minimizing transmit energy and maximizing network lifetime, and did not consider energy consumption during relay selection.

We consider a generic system-level model that explicitly models the two phases and brings out the inherent trade-offs between them. We develop analytical expressions for the overall throughput and energy consumption that account for both the relay selection overhead and the benefits from selection in the subsequent data transmission phase. As we shall see, our analysis helps us determine the optimal time duration of the relay selection phase. This point optimally trades off between using a large amount of time for selection, which increases the probability of benefiting from the best relay, and spending less time in selection, which decreases the selection overhead. Furthermore, by considering a non-adaptive system and various adaptive ones, we find that the optimal operating parameter settings for the two phases depend on the type of adaptation, if present, in the system. In the non-adaptive system, parameters such as transmit power, contention duration, and modulation scheme are kept fixed, while in the adaptive systems these parameters may vary dynamically.

In order to make system-level comparisons that are quantitative in nature, we consider a specific splitting-based algorithm for relay selection [14], [16]. We make this choice because the algorithm is fast, distributed, and scales well with the number of relays. In the asymptotic regime of a large number of nodes, the average number of slots required to select the best node is upper bounded by 2.5 [14], with the exact value being marginally less at 2.47 [16]. We shall see that even though this algorithm is fast, its overhead can be considerable. For example, the outage probability when the maximum time available is 3 slots (which is close to 2.47) is more than 20%. As our results show, from an overall system-level throughput perspective, the system is sometimes better off allocating more time to the selection phase, e.g., 7-8 slots, to improve the odds of selecting the best relay. Our results, thus, suggest that a joint design is also necessary in cooperative systems that use other selection mechanisms such as the back-off timer-based mechanism of [12], [13] or the handshaking message-based tracking approach used in [11].

A secondary contribution of the paper is its new insight into the splitting-based selection algorithm itself. Specifically, it develops, for use in the system-level analysis, new expressions for the outage probability of the selection algorithm. Hitherto, only the average number of slots required to select the best relay has been analyzed for this algorithm [14], [16].

The paper is organized as follows. The system model is set up in Sec. II. The non-adaptive and adaptive modes are analyzed in Sec. III and Sec. IV, respectively. The design implications are brought out in Sec. V, which also verifies the analysis with simulations. Our conclusions follow in Sec. VI. Several mathematical proofs are relegated to the Appendix.

II. SYSTEM MODEL

Figure 1 shows a cooperative relay network that contains a source node, a destination node, and n decode-and-forward

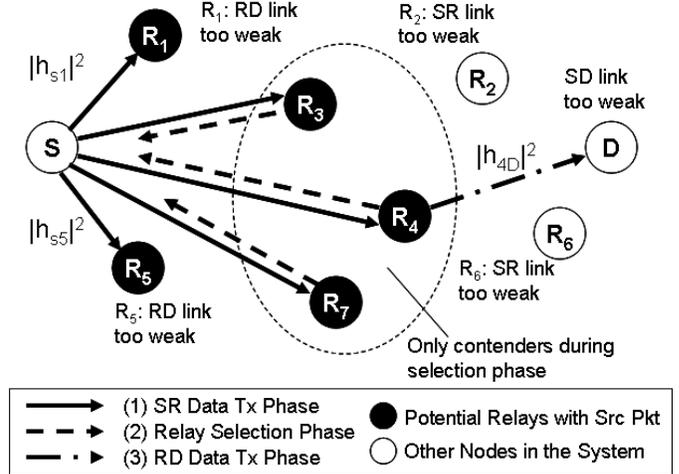


Fig. 1. A cooperative system consisting of a source (S), a destination (D), and $n = 7$ relays.

relays. The SR and RD channels are assumed to have flat frequency responses. We assume a block fading model, where the channel gains remain constant over the selection and data transmission phases. This assumption is valid when the total amount of time spent in the selection and transmission phase is smaller than the coherence time of the channel. Independent Rayleigh fading is assumed: the fading gains from the source to relay i , $|h_{si}|^2$, and from the relay to destination, $|h_{id}|^2$, are independent, exponentially distributed random variables with unit mean. Additive white Gaussian noise with zero mean and unit variance is assumed at each receiver. For analytical tractability, we assume that the direct source to destination link is very weak or blocked, as is done, for example, in [18], [19].¹

The analysis is developed for a *general and practically relevant case* in which the relays are distributed over a geographical region \mathcal{A} . Therefore, the SR and RD channel gains for different relays are not statistically identical as their means depend on the relay position. We assume that the source is located at the origin. The location of relay i , \mathbf{x}_i , is set randomly according to the probability distribution $\xi_{\mathcal{A}}(\mathbf{x}_i)$ defined over \mathcal{A} . Let \mathbf{x}_d be the position of the destination. The distances of relay i from the source and destination are denoted by $|\mathbf{x}_i|$ and $|\mathbf{x}_d - \mathbf{x}_i|$, respectively. Wherever possible, we also specialize the results for a simpler and more intuitive *symmetric case*, in which the SR (and RD) channels of different relays are statistically identical, since it leads to simpler closed-form expressions.

A. Cooperation Model

We first describe the relay cooperation model for the baseline non-adaptive system, and then describe the modifications associated with the various adaptive systems.

¹The reader is referred to [20] for an interesting analysis that includes the source-destination (SD) link, and a discussion of why doing so complicates the analysis. Note that the gains may not be significant even if the SR link is comparable to the other links. This is because the diversity order achieved will increase by one, which is relatively small compared to the n th order diversity already obtained by selecting from among n relays. An exact closed-form system-level trade-off analysis that accounts for the direct SD link remains an open problem.

- 1) *SR Data Transmission Phase*: The source broadcasts at power ρ_s for $T_d \in \mathbb{R}$ seconds and all the other relays listen. The SNR, γ_{si} , at relay i is equal to $\gamma_{si} = \frac{\rho_s |h_{si}|^2}{N_0} \frac{\kappa}{|\mathbf{x}_i|^\epsilon}$, where ϵ is the propagation exponent that takes values between 2 and 4 and κ is a constant [21]. To relate the data payload B , $|h_{si}|^2$, and ρ_s , we use the ideal Shannon capacity formula, and assume that relay i can decode the source's transmission only if

$$B \leq WT_d \log_2 \left(1 + \frac{\rho_s |h_{si}|^2}{N_0} \frac{\kappa}{|\mathbf{x}_i|^\epsilon} \right), \quad (1)$$

where W is the transmission bandwidth in Hz.²

- 2) *Relay Selection Phase*: In this phase, the relays contend for $T_c \in \mathbb{Z}$ slots. Each slot is of duration t_{slot} sec, which includes possible transmissions from contending nodes and feedback. For example, in the IEEE 802.11n standard, t_{slot} would be at least 130 μsec when $W = 10$ MHz [13], [24]. The selection algorithm is explained in detail below in Sec. II-E.
- 3) *RD Data Transmission Phase*: At the end of the relay selection phase, the selected relay transmits data to the destination at a power ρ_r for T_d seconds. If relay i is selected for transmission, then the destination decodes the data if $B \leq WT_d \log_2 \left(1 + \frac{\rho_r |h_{id}|^2}{N_0} \frac{\kappa}{|\mathbf{x}_d - \mathbf{x}_i|^\epsilon} \right)$.

An outage occurs when the destination cannot decode the source's message.

B. Adaptive Modes

A variety of adaptations are possible to improve the performance of the above non-adaptive system. We analyze the following different adaptations in this paper:

- 1) *Adaptive selection duration*: To improve system throughput, the relay data transmission can start as soon as the best relay is selected instead of waiting for a fixed duration of T_c slots.
- 2) *Adaptive relay transmit power*: Since local channel knowledge is available at the relays, the selected relay can adapt its transmit power (subject to a peak power constraint) in the RD data transmission phase, and improve the overall energy efficiency.
- 3) *Adaptive relay transmit rate*: The selected relay can instead adapt its transmit rate (not power) to reduce RD data phase duration and improve system throughput.

Several other adaptations, including hybrid versions of the above, are indeed possible. For example, the SR data transmission phase can itself be made adaptive by making it follow the relay selection phase; the source can then acquire additional knowledge of the best source-relay channel [25] and adapt accordingly. Power control can be used in the relay selection phase itself to improve energy efficiency [5]. We choose the above three systems as they cover several common forms of adaptation [26]–[29], and highlight the effect of individual adaptations on overall system performance.

²This simplification of the stair-case like practical data rate function, which depends on specific coding and modulation schemes, allows us to gain insights into the optimal trade-off between the selection and data transmission phases. Practical coding inefficiencies can also be easily incorporated into this model [22], [23].

C. Relay Selection Criterion

Selection is done on the basis of the suitability metric. Each relay only knows its metric and not that of others [18]. Only those relays that decode the source's message and have RD gains large enough to support transmission to the destination participate in the selection process. Other relays do not contend as this wastes energy without improving throughput. This is achieved by setting their metrics to 0.³

It follows from the discussion above that the local metric, after removing terms common to all relays, can be set to

$$\mu_i = \begin{cases} \frac{|h_{id}|^2}{|\mathbf{x}_d - \mathbf{x}_i|^\epsilon}, & \frac{|h_{si}|^2}{|\mathbf{x}_i|^\epsilon} \geq a_s \text{ and } \frac{|h_{id}|^2}{|\mathbf{x}_d - \mathbf{x}_i|^\epsilon} \geq a_r \\ 0, & \text{otherwise} \end{cases}, \quad (2)$$

where $a_s = \left(2^{\frac{B}{WT_d}} - 1 \right) \frac{N_0}{\kappa \rho_s}$ and $a_r = \left(2^{\frac{B}{WT_d}} - 1 \right) \frac{N_0}{\kappa \rho_r}$.

D. Complementary Cumulative Distribution Function (CCDF) of the Metric

We assume that the SNRs of the SR and RD channels of a relay, conditioned on its position, are exponentially distributed and independent. This is justifiable because the local scatterers that contribute to the fading in the SR channel are often different from those that contribute to the fading in the RD channel. Then, in general, based on (2), the CCDF of the metric takes the following form:

$$F_c(\mu) = \begin{cases} \int_{\mathcal{A}} \exp(-a_s |\mathbf{x}|^\epsilon - \mu |\mathbf{x}_d - \mathbf{x}|^\epsilon) \xi_{\mathcal{A}}(\mathbf{x}) d\mathbf{x}, & \mu \geq a_r \\ \int_{\mathcal{A}} \exp(-a_s |\mathbf{x}|^\epsilon - a_r |\mathbf{x}_d - \mathbf{x}|^\epsilon) \xi_{\mathcal{A}}(\mathbf{x}) d\mathbf{x}, & 0 < \mu < a_r \\ 1, & \mu = 0 \end{cases} \quad (3)$$

The exponential form of the integrand in (3) occurs because the SNRs of the SR and RD channels of a relay, conditioned on its position, are exponentially distributed and independent.⁴ In practice, the CCDF and its inverse are numerically computed and stored in a lookup table. The special case where the SR (and RD) channels of different relays are statistically identical is equivalent to a set up where all the relays are at a virtual position \mathbf{x}_0 and see independent fading to the source and destination. In this case, the CCDF simplifies to:

$$F_c(\mu) = \begin{cases} \exp(-a_s |\mathbf{x}_0|^\epsilon - \mu |\mathbf{x}_d - \mathbf{x}_0|^\epsilon), & \mu \geq a_r \\ \exp(-a_s |\mathbf{x}_0|^\epsilon - a_r |\mathbf{x}_d - \mathbf{x}_0|^\epsilon), & 0 < \mu < a_r \\ 1, & \mu = 0 \end{cases} \quad (4)$$

Notice that, in general, the CCDF is constant in the interval $(0, a_r)$, and jumps to 1 at $\mu = 0$. This is a consequence of a relay setting its metric to 0 if its SR or RD channel is weak.

³Another possible approach here is to model only the relays that have decoded the source's message as the contending relays. However, this is practically infeasible since the selection algorithm described next would then require every relay to know the number of relays that have decoded the source's message.

⁴Such an averaging over the position cannot be done if the relays somehow know their positions since the metrics of different relays will be statistically different. It can be shown that this can only improve the speed of relay selection. However, a detailed study of this is beyond the scope of this paper.

E. Relay Selection Algorithm

The goal of the selection algorithm is to choose the relay with the highest metric. At the beginning of a slot k , each relay (independently) calculates two thresholds $H_L(k)$ and $H_H(k)$, and transmits with power ρ_c if its metric μ satisfies $H_L(k) < \mu < H_H(k)$. The content of the transmission does not matter as long as every node remembers whether it transmitted or not in a slot. The source node then broadcasts one of three outcomes to all the relays: (i) *idle* when no relay transmitted, (ii) *success* when exactly one relay transmitted, or (iii) *collision* when multiple relays transmitted. The source can determine these outcomes based on, for example, the strength of the received power, which is measurable by many receivers today [15].⁵ The relays update their thresholds based on the outcome of the previous slot. The feedback messages are assumed to be received error-free given their low payload. If no relay has been selected within T_c slots, the algorithm terminates and the system declares an outage.

We first formally describe the algorithm, and then provide a brief explanation. We first define the following: (i) $H_{\min}(k)$, which is the largest value of the metric known up to slot k above which the best metric surely lies, and (ii) the complementary CDF (CCDF) of the metric μ_i : $F_c(\mu) = \Pr(\mu_i \geq \mu)$, and (iii) A *split* function $\text{split}(a, b) \triangleq F_c^{-1}\left(\frac{F_c(a)+F_c(b)}{2}\right)$, which ensures that, on average, only half the relays involved in the last collision transmit in the next slot [14].

In the first slot, the variables are initialized to: $H_L(1) = \max(a_r, F_c^{-1}(1/n))$, $H_H(1) = \infty$, and $H_{\min}(1) = 0$. In the $(k+1)^{\text{th}}$ slot, the variables are updated as follows:

- 1) If feedback (of the k^{th} slot) is an idle and no collisions have occurred thus far, then $H_H(k+1) = H_L(k)$, $H_{\min}(k+1) = 0$, and $H_L(k+1) = \max(a_r, F_c^{-1}(\frac{k+1}{n}))$.
- 2) If feedback is a collision, then $H_H(k+1) = H_H(k)$, $H_{\min}(k+1) = H_L(k)$, and $H_L(k+1) = \text{split}(H_L(k), H_H(k))$.
- 3) If feedback is an idle and a collision has occurred in the past, then $H_H(k+1) = H_L(k)$, $H_{\min}(k+1) = H_{\min}(k)$, and $H_L(k+1) = \text{split}(H_{\min}(k), H_L(k))$.

Brief explanation: Since the goal of the algorithm is to find the node with the highest metric, in the first slot it makes nodes with metrics between $F_c^{-1}(1/n)$ and ∞ transmit. This ensures that, on average, only one node transmits in this slot. If no node transmits in this slot, the thresholds are lowered, ensuring again that one node transmits on average. The thresholds are lowered further in subsequent slots, until a collision or a success occurs. In case of a collision, since at least 2 nodes transmitted in that slot, the threshold intervals of the subsequent slots are successively *split* until success occurs. If a success occurs, the node that transmitted in that slot is selected, since success occurs only when one node has transmitted. This node then forwards the data in the subsequent RD transmission phase.

⁵Note that this collision MAC model, while classical, is pessimistic as it ignores the differences in received power which arise due to different wireless channel fades and path losses. An analogous analysis of the impact of modifications such as capture on the system-level throughput is an interesting avenue for future work.

Note that our formulation above differs slightly from the original one in [14] because we introduce the $\max(\cdot)$ function to prevent relays with metrics below a_r from transmitting. This avoids the difficult problem of the source having to know how many nodes decoded the message, which is a random variable. By ensuring that only those relays that have decoded the source's message and have a strong enough RD channel contend, it also avoids wasting energy.

III. ANALYSIS: BASELINE NON-ADAPTIVE SYSTEM

We first analyze the throughput and energy consumption of the non-adaptive system. The adaptive modes are analyzed thereafter in Sec. IV.

A. Throughput Analysis

The destination can successfully decode only when all the following conditions are satisfied:

- 1) There is at least one relay that has decoded the message from the source,
- 2) Among the relays that have decoded the source message, at least one of them has a good enough link to the destination, and
- 3) The selection phase can select the best relay within T_c slots.

In our set up, the above conditions together are equivalent to the condition that the selection phase terminates successfully within T_c slots. Let $P_s(T_c)$ be the probability that the best node is successfully selected in the relay selection phase. Thus, $P_s(T_c)$ is also the probability that destination can successfully decode the source's data. Therefore, the throughput, η , for the non-adaptive system is

$$\eta = \frac{BP_s(T_c)}{2T_d + T_{c\text{tslot}}} \text{ bits/second}, \quad (5)$$

where $2T_d + T_{c\text{tslot}}$ is the total time taken over the three phases of the cooperation protocol. The time taken by the source to broadcast feedback is assumed to be negligible. (Alternately, it can be accounted for in t_{slot} .) To find the probability of successful selection, $P_s(T_c)$, we first state an intermediate result about the selection algorithm.

Lemma 1: The probability, $p(k, b)$, that exactly b slots are required to resolve a collision involving k relays is

$$p(k, b) = \frac{1}{2^k} \left(p(k, b-1) + \sum_{i=2}^k \binom{k}{i} p(i, b-1) \right), \quad \forall k, b > 1, \quad (6)$$

where the recursion is initialized by $p(k, 1) = k/2^k$, $\forall k > 1$, and $p(1, b) = 0$.

Proof: The proof is given in Appendix A. ■

The main result on the probability of success now follows.

Theorem 1: The probability of successful selection is given by (7) (on the next page), where $I_{\{x\}}$ is an indicator function that equals 1 if condition x is true and is 0 otherwise, $\alpha = F_c(a_r)$, $r = \lceil n\alpha \rceil - 1$, and $\lceil \cdot \rceil$ is the ceil function.

Proof: The proof is given in Appendix B. ■

In (7), the first and the second terms correspond to the first non-idle slot being a success. The remaining two terms correspond to the first non-idle slot (i^{th} slot) being a collision among $k > 1$ relays that is resolved in $T_c - i$ slots.

$$\begin{aligned}
 P_s(T_c) = & \sum_{i=1}^{\min(T_c, r)} \left(1 - \frac{i}{n}\right)^{n-1} + I_{\{T_c > r\}} n \left(\alpha - \frac{r}{n}\right) (1 - \alpha)^{n-1} + \sum_{i=1}^{\min(T_c-1, r)} \sum_{k=2}^n \binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{i}{n}\right)^{n-k} \sum_{j=1}^{T_c-i} p(k, j) \\
 & + I_{\{T_c > r+1\}} \sum_{k=2}^n \binom{n}{k} \left(\alpha - \frac{r}{n}\right)^k (1 - \alpha)^{n-k} \sum_{j=1}^{T_c-r-1} p(k, j). \quad (7)
 \end{aligned}$$

B. Energy Analysis

The total energy consumed is the sum of the energy consumed by the SR and RD data transmission phases and by the relay selection phase. We limit our attention to transmit energy consumption because it is traditionally the dominant component. A more detailed characterization, which accounts for receive power consumption and transmit non-idealities, is beyond the scope of this paper. The reader is referred to [30], for example, for a discussion of these issues.

To derive the expression for the average energy used in selection phase, E_c , the following result about the number of transmissions in the selection algorithm shall come in handy.

Lemma 2: The average number of relays that transmit in the first b slots that follow a collision among k relays is

$$N(k, b) = \frac{k}{2} + \frac{1}{2^k} \left(N(k, b-1) + \sum_{j=1}^k \binom{k}{j} N(j, b-1) \right), \quad \forall k, b > 1, \quad (8)$$

where the recursion starts with $N(k, 1) = k/2$, $\forall k > 1$, and $N(1, b) = N(k, 0) = 0$, $\forall k, b > 1$.

Proof: The proof is given in Appendix C. ■

The main result on the energy consumed by the selection phase now follows.

Theorem 2: The average total energy consumed, E_c , in the selection phase is given by (9) (on the next page).

Proof: The proof is given in Appendix D. ■

The energy consumed in the RD data transmission phase is $E_d = \rho_r T_d P_s(T_c)$ and in the SR phase is $\rho_s T_d$. The $P_s(T_c)$ factor occurs because a RD transmission occurs only if the selection phase is successful. Let E_s denote the energy consumed by the source in sending the feedback message. Then, the average energy consumed by the source during the selection phase is $E_s m(T_c)$, where $m(T_c)$ is the average number of slots per selection phase. An expression for it is derived in the next section and is not repeated here. Therefore, the total energy consumed is⁶

$$E_{\text{total}} = \rho_s T_d + E_s m(T_c) + E_c + \rho_r T_d P_s(T_c). \quad (10)$$

IV. THREE ADAPTIVE SYSTEMS

We now analyze the throughput and energy consumption of the three adaptive systems.

⁶Given the focus of the paper on the trade-offs between the selection and data transmission phases, we do not model the energy consumed by a node to locally acquire the channel knowledge and determine its metric.

A. Adaptive Selection Duration

In this system, the selection phase stops as soon as the best relay has been selected or it is obvious that no relay can be selected. It follows that the overall throughput of the system is

$$\eta = \frac{B P_s(T_c)}{2T_d + m(T_c) t_{\text{slot}}} \text{bits/second}, \quad (11)$$

Notice that the above equation has $m(T_c)$ in its denominator as compared to T_c in the throughput expression in (5) for the non-adaptive scheme. $m(T_c)$ is given as follows.

Lemma 3: The average number of slots required during the selection phase is

$$m(T_c) = T_c - \sum_{j=1}^{T_c-1} P_s(j) - I_{\{T_c > r+1\}} (T_c - r - 1) (1 - \alpha)^n. \quad (12)$$

Proof: The proof is given in Appendix E. ■

The above result clearly shows the throughput benefit of adaptation because $m(T_c) < T_c$. The selection phase energy consumption is the same as in the non-adaptive case (Theorem 2).

B. Adaptive Relay Transmit Power

We now consider a system in which the selected relay i adapts its transmit power in the RD data transmission phase to minimize energy consumption. From (1), it follows that a relay can adapt its transmit power to $\rho_r = \frac{(2^{\frac{B}{W T_d}} - 1) N_0 |\mathbf{x}_d - \mathbf{x}_i|^\epsilon}{\kappa |h_{id}|^2}$. If this power exceeds a peak power of P_{\max} , the relay does not contend. (This occurs when $\frac{|h_{id}|^2}{|\mathbf{x}_d - \mathbf{x}_i|^\epsilon} < \frac{(2^{\frac{B}{W T_d}} - 1) N_0}{\kappa P_{\max}} \triangleq a_r^{\text{pow}}$.) Therefore, the throughput of this system is then the same as that for the baseline non-adaptive system in Sec. III-A with a_r replaced by a_r^{pow} .

The energy consumed in the source transmission and selection phases is the same as before since ρ_s and ρ_c are fixed. The RD energy consumption changes since the selected relay can adapt its transmit power. It must be noted that the energy consumed in the RD data transmission phase is coupled with the time at which success occurs in the selection phase. For example, *the higher the best metric value, the fewer the number of initial idle slots required in the selection phase.* Both these observations are captured in the expression for E_d derived below.

Theorem 3: Let $E(u, v; k)$, $0 \leq u < v$, denote the average energy consumed for data transmission by the best relay selected from among k relays whose metrics lie in the interval

$$E_c = \rho_c t_{\text{slot}} \sum_{i=1}^{\min(T_c, r)} \left(\sum_{k=1}^n \binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{i}{n}\right)^{n-k} N(k, T_c - i) + \left(1 - \frac{i-1}{n}\right)^{n-1} \right) \\ + \rho_c t_{\text{slot}} I_{\{T_c > r\}} \left(\sum_{k=1}^n \binom{n}{k} \left(\alpha - \frac{r}{n}\right)^k (1 - \alpha)^{n-k} N(k, T_c - r - 1) + (n\alpha - r) \left(1 - \frac{r}{n}\right)^{n-1} \right). \quad (9)$$

$(F_c^{-1}(v), F_c^{-1}(u))$. It is given by

$$E(u, v; k) = \frac{2^{\frac{B}{W} T_d} - 1}{\kappa} N_0 T_d \\ \times \int_{F_c^{-1}(v)}^{F_c^{-1}(u)} \frac{-k F_c'(\mu) (v - F_c(\mu))^{k-1}}{\mu (v - u)^k} d\mu. \quad (13)$$

And, the average energy consumed, E_d , in the RD data transmission phase is given by

$$E_d = \sum_{i=1}^{\min(T_c, r^{\text{pow}})} \left(1 - \frac{i}{n}\right)^{n-1} E\left(\frac{i-1}{n}, \frac{i}{n}; 1\right) \\ + I_{\{T_c > r^{\text{pow}}\}} \left(\alpha^{\text{pow}} - \frac{r^{\text{pow}}}{n}\right) (1 - \alpha^{\text{pow}})^{n-1} E\left(\frac{r^{\text{pow}}}{n}, \alpha^{\text{pow}}; 1\right) \\ + \sum_{i=1}^{\min(T_c-1, r^{\text{pow}})} \sum_{k=2}^n \binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{i}{n}\right)^{n-k} E\left(\frac{i-1}{n}, \frac{i}{n}; k\right) \\ \times \sum_{j=1}^{T_c-i} p(k, j) + I_{\{T_c > r^{\text{pow}}+1\}} \sum_{k=2}^n \binom{n}{k} \left(\alpha^{\text{pow}} - \frac{r^{\text{pow}}}{n}\right)^k \\ (1 - \alpha^{\text{pow}})^{n-k} E\left(\frac{r^{\text{pow}}}{n}, \alpha^{\text{pow}}; k\right) \sum_{j=1}^{T_c-r^{\text{pow}}-1} p(k, j), \quad (14)$$

where $\alpha^{\text{pow}} = F_c(a_r^{\text{pow}})$, and $r^{\text{pow}} = \lceil n\alpha^{\text{pow}} \rceil - 1$.

Proof: The proof is given in Appendix F. ■

For symmetric channel gains, $E(u, v; k)$ takes the following simpler closed form:

$$E(u, v; k) = \frac{(2^{\frac{B}{W} T_d} - 1) N_0 T_d}{\kappa} \frac{k e^{-k a_s \mathbf{x}_0^\epsilon} |\mathbf{x}_d - \mathbf{x}_0|^\epsilon}{(v - u)^k} \\ \times \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} (v e^{a_s})^{k-1-j} \\ \times \left(\mathbb{E}((j+1) |\mathbf{x}_d - \mathbf{x}_0|^\epsilon F_c^{-1}(u)) \right. \\ \left. - \mathbb{E}((j+1) |\mathbf{x}_d - \mathbf{x}_0|^\epsilon F_c^{-1}(v)) \right), \quad (15)$$

where $F_c^{-1}(\mu) = -(\log(\mu) + a_s)$ and the Euler exponential integral $\mathbb{E}(z)$ is the principal value of $\int_{-\infty}^z \frac{e^t}{t} dt, \forall z > 0$ [31]. (See also Appendix F.)

The total average energy used is then $E_{\text{total}} = \rho_s T_d + E_s m(T_c) + E_c + E_d$, where E_c is given by Theorem 2.

C. Adaptive RD Data Transmission Phase Duration

In this system, the selected relay, if any, adapts its transmission rate to the destination since it knows its channel gain to the destination. Equivalently, it adapts its transmission duration, subject to a maximum duration of T_{max} (which is

a system parameter). Thus, a relay i contends only when $\frac{|h_{id}|^2}{|\mathbf{x}_d - \mathbf{x}_i|^\epsilon} \geq \frac{(2^{\frac{B}{W} T_{\text{max}}} - 1) N_0}{\kappa \rho_r} \triangleq a_r^{\text{rate}}$. Also, in case the selection phase results in an outage, no time is spent in the RD phase.

The throughput of this system is now equal to

$$\eta = \frac{B P_s(T_c)}{T_d + T_c t_{\text{slot}} + T_{\text{adp}}} \text{ bits/second.} \quad (16)$$

The probability of successful selection, $P_s(T_c)$, and, consequently, the throughput is again given by (7), with a_r replaced by a_r^{rate} . Lastly, the expression for T_{adp} is as follows.

Theorem 4: Let $T(u, v; k)$, $0 \leq u < v$, denote the average time required by the selected relay to transmit data when it is selected from k relays whose metrics lie in the interval $(F_c^{-1}(v), F_c^{-1}(u))$. It equals

$$T(u, v; k) = \frac{B}{W} \int_{F_c^{-1}(v)}^{F_c^{-1}(u)} \frac{k f(\mu) (v - F_c(\mu))^{k-1}}{\log_2 \left(1 + \frac{\rho_r \kappa}{N_0} \mu\right) (v - u)^k} d\mu. \quad (17)$$

The average duration of the adaptive RD data transmission phase is

$$T_{\text{adp}} = \sum_{i=1}^{\min(T_c, r^{\text{rate}})} \left(1 - \frac{i}{n}\right)^{n-1} T\left(\frac{i-1}{n}, \frac{i}{n}; 1\right) \\ + I_{\{T_c > r^{\text{rate}}\}} \left(\alpha^{\text{rate}} - \frac{r^{\text{rate}}}{n}\right) (1 - \alpha^{\text{rate}})^{n-1} T\left(\frac{r^{\text{rate}}}{n}, \alpha^{\text{rate}}; 1\right) \\ + \sum_{i=1}^{\min(T_c-1, r^{\text{rate}})} \sum_{k=2}^n \binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{i}{n}\right)^{n-k} T\left(\frac{i-1}{n}, \frac{i}{n}; k\right) \\ \times \sum_{j=1}^{T_c-i} p(k, j) + I_{\{T_c > r^{\text{rate}}+1\}} \sum_{k=2}^n \binom{n}{k} \left(\alpha^{\text{rate}} - \frac{r^{\text{rate}}}{n}\right)^k \\ \times (1 - \alpha^{\text{rate}})^{n-k} T\left(\frac{r^{\text{rate}}}{n}, \alpha^{\text{rate}}; k\right) \sum_{j=1}^{T_c-r^{\text{rate}}-1} p(k, j), \quad (18)$$

where $\alpha^{\text{rate}} = F_c(a_r^{\text{rate}})$, and $r^{\text{rate}} = \lceil n\alpha^{\text{rate}} \rceil - 1$.

Proof: The proof is given in Appendix G. ■

For symmetric channel gains, $T(u, v; k)$ simplifies to

$$T(u, v; k) = \frac{B}{W} \frac{k e^{-k a_s \mathbf{x}_0^\epsilon}}{|\mathbf{x}_d - \mathbf{x}_0|^\epsilon (v - u)^k} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} (v e^{a_s \mathbf{x}_0^\epsilon})^{k-1-j} \\ \times \log(2) \left(\psi_1 \left(\frac{\kappa \rho_r}{N_0 (j+1) |\mathbf{x}_d - \mathbf{x}_0|^\epsilon}, \log \left(1 + \frac{\kappa \rho_r}{N_0} F_c^{-1}(v) \right) \right) \right. \\ \left. - \psi_1 \left(\frac{\kappa \rho_r}{N_0 (j+1) |\mathbf{x}_d - \mathbf{x}_0|^\epsilon}, \log \left(1 + \frac{\kappa \rho_r}{N_0} F_c^{-1}(u) \right) \right) \right), \quad (19)$$

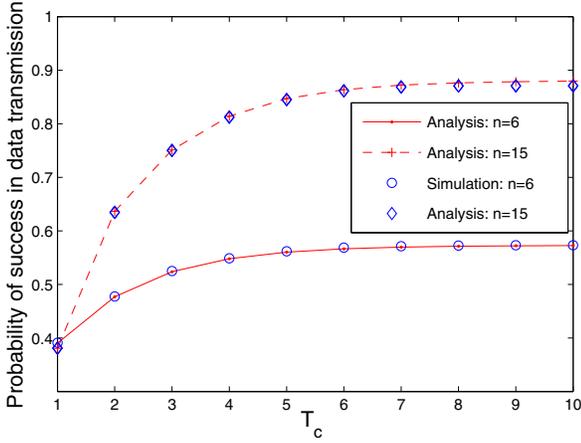


Fig. 2. Probability of successful selection, $P_s(T_c)$, as a function of number of slots in selection phase, T_c , for the non-adaptive system.

where $\psi_1(a, w) = \int_w^\infty \frac{1}{t} \exp\left(t + \frac{1-e^t}{a}\right) dt$, as defined in [19]. (See also Appendix G.)

Since a relay transmits with constant power ρ_r for an average duration of T_{adp} , the total energy consumed by the relays in this case is $E_{\text{total}} = \rho_s T_d + E_s m(T_c) + E_c + \rho_r T_{\text{adp}} P_s(T_c)$.

V. RESULTS AND SYSTEM DESIGN IMPLICATIONS

We now study the system-level trade-offs using our analytical results, and verify them independently using Monte Carlo simulations that use 10^5 samples.

Simulation set up: We developed a custom time-driven simulator in Matlab that models all the phases of cooperation and selection and does not use any results or short cuts from the analysis. The figures are plotted for a propagation exponent $\epsilon = 4$ and $\frac{B}{WT_d} = \log_2(1 + 10^{0.6})$, which implies that a relay or destination can decode successfully if its instantaneous received SNR is at least 6 dB. The source is located at the origin and the destination is placed at $(2, 0)$. The relays are uniformly distributed in a rectangular region, whose vertices are given by $(0.5, -1)$, $(0.5, 1)$, $(1.5, -1)$, and $(1.5, 1)$. The source and relay transmit powers are set so that the fading-averaged SNR at a unit distance from the transmitter is 9 dB. The figures are plotted for $n = 6$ and $n = 15$ nodes to show the effect of the number of nodes. Such node counts have also been used in the literature, see, for example, [2], [4], [10], [12], [17], [20]; even larger node deployments are likely in sensor networks.

A. Throughput vs. Selection Duration Trade-off

For a non-adaptive system, Fig. 2 shows that the probability of successful selection, $P_s(T_c)$, increases as the number of contention slots, T_c , or relays, n , increases. For large T_c , $P_s(T_c)$ saturates at $1 - (1 - \alpha)^n$ because $(1 - \alpha)^n$ is the probability that none of the n relays participate in the selection phase. Notice that the analysis and simulation results match well.

While increasing T_c improves $P_s(T_c)$, it also increases the total duration of the three phases. This important trade-off is shown in Fig. 3, which plots the normalized throughput,

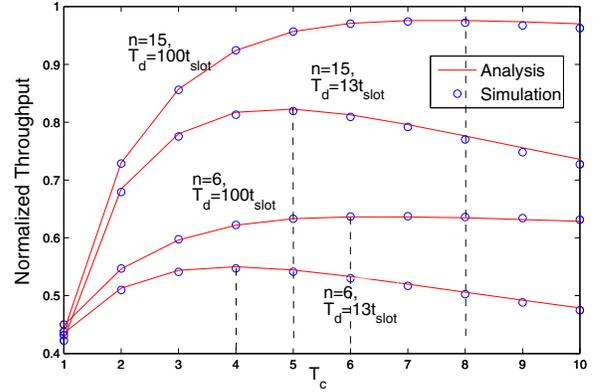


Fig. 3. Normalized throughput, $\frac{\eta}{WT_{\text{slot}}}$, as a function of number of slots in selection phase, T_c , for the non-adaptive system.

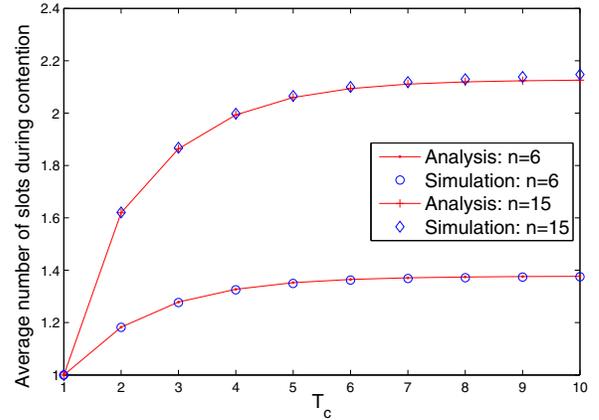


Fig. 4. Average number of slots required by the selection phase as a function of the maximum number of allowed contention slots for the adaptive selection duration system.

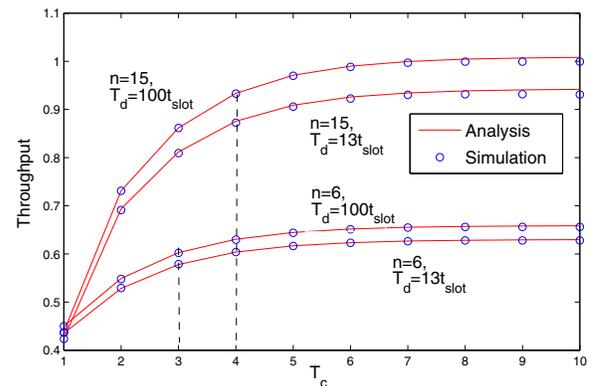


Fig. 5. Normalized throughput, $\frac{\eta}{WT_{\text{slot}}}$, as a function of the maximum number of allowed contention slots for the adaptive selection duration system.

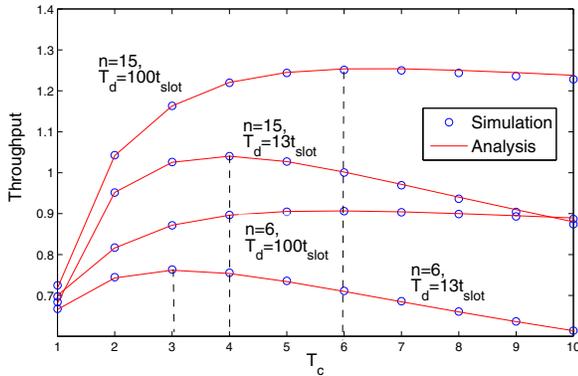


Fig. 6. Normalized throughput, $\frac{\eta}{Wt_{\text{slot}}}$, as a function of the number of selection phase slots, T_c , when the RD data transmission duration is adapted (with $T_{\text{max}} = T_d$).

$\frac{\eta}{Wt_{\text{slot}}}$, for the non-adaptive system. Increasing T_c when it is small improves throughput since the probability of outage decreases. However, for larger T_c , the throughput starts decreasing because $P_s(T_c)$ saturates. The figure shows that the optimal value for the selection phase duration depends on both n and T_d . In general, the throughput increases as n or T_d increase because the relative time overhead of the selection phase decreases.

Figure 4 plots the average number of slots needed in the selection phase, $m(T_c)$, when the selection phase duration is variable. Now, $m(T_c)$ is often much smaller than T_c , which improves throughput. Also, $m(T_c)$ increases with n because the odds that more relays contend increases. The corresponding throughput is plotted in Fig. 5. Unlike the non-adaptive system, the overall time needed to send the data does not decrease for large T_c ; it instead saturates. The figure provides guidelines on how to set the maximum selection phase duration. For example, to achieve 90% of the maximum throughput possible, the selection phase should have at least 3 and 4 slots for $n = 6$ and $n = 15$ relays, respectively. Otherwise, outage probability increases adversely.

Finally, the throughput when the RD transmission duration is adaptive (subject to a maximum duration of $T_{\text{max}} = T_d$) is plotted in Fig. 6. Compared to the non-adaptive scheme, the optimal throughput is higher; for example, it is 29% more for $n = 15$ and $T_d = 100t_{\text{slot}}$. And, the optimal selection phase duration decreases. For example, for $T_d = 13t_{\text{slot}}$, $T_c = 3$ and 4 slots are optimal for $n = 6$ and 15, respectively. Whereas, for $T_d = 100t_{\text{slot}}$, $T_c = 6$ slots is optimal for both $n = 6$ and 15. Notably, at the optimum point, $P_s(T_c)$ need not be close to its maximum value. For example, for $n = 6$, it is 91.2% and 98.2% of the maximum value of 0.57 for $T_d = 13t_{\text{slot}}$ and $100t_{\text{slot}}$, respectively. As in the non-adaptive system, the throughput increases with both n and T_d .

Comparison with Cooperation without Selection: Since relay selection incurs an overhead, it is worthwhile to compare the performance – for the same simulation set up – with a cooperative system that does not select relays on the basis of their channel states and, consequently, does not incur the overhead of selection. Now, each source randomly selects any one of the relays and transmits to the destination through it.

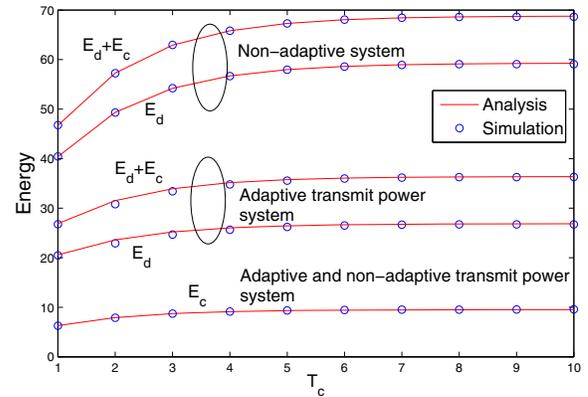


Fig. 7. Comparison of energy used by relays during the selection and RD data transmission phases for the non-adaptive and adaptive relay transmit power systems (with $n = 6$ and $T_d = 13t_{\text{slot}}$).

As before, an outage occurs if either the SR or the RD channel of the selected relay is weak. For the baseline non-adaptive system, the average throughput turns out to be 0.15 for all n and T_d . From Fig. 3, we see that it is just 26.7% and 23.0% of the optimal throughput with selection for $T_d = 13t_{\text{slot}}$ and $T_d = 100t_{\text{slot}}$, respectively. When the RD data transmission duration is adapted by the selected relay, the throughput of random selection increases to 0.28 for all T_d . From Fig. 6, we see that it is just 37.2% and 31.4% of the corresponding optimal throughput with selection. A key reason for the worse performance of random selection is its significantly lower probability of success. The performance of direct transmission is even worse given the relatively large distance between the source and the destination, and is not discussed here to conserve space.

B. Energy Consumption Overhead

Figure 7 compares the relative contribution of the relay selection phase to the energy consumed in the RD data transmission phase for the non-adaptive and the adaptive relay transmit power systems for $n = 6$ and $T_d = 13t_{\text{slot}}$. The energy consumed by the source in the SR transmission phase is a constant and is, therefore, not shown. Similarly, the energy, E_s , consumed by the source to feedback messages is typically negligible, and is not shown. In the adaptive relay transmit power scheme, the peak power, P_{max} , is set so that the maximum fading-averaged SNR at a unit distance from transmitter is also 9 dB. In both cases, the total energy consumed increases with the selection phase duration. This is because increasing the selection phase duration increases E_c since more relays have an opportunity to contend and transmit, and it also increases E_d since it is more likely that a relay transmits in the RD phase after it has been successfully selected.⁷

VI. CONCLUSIONS

We saw that a cooperative relay selection system expends time and energy in the selection phase and reaps the benefits of

⁷Note that even though the energy consumption is low for small T_c , so is the throughput and the energy efficiency, measured in terms of total energy required to transmit a bit of information.

selection in the form of increased throughput or lower energy consumption in the data transmission phase. We analyzed the interactions and trade-off between the selection and data transmission phases of both adaptive and non-adaptive cooperative relay systems. Even when a fast, distributed splitting-based selection algorithm was used, the optimal setting of the selection phase yielded significant gains over random selection, but still incurred a non-negligible time and energy cost. Thus, the amount of time allocated to the selection phase needs to be carefully chosen in order to maximize the overall system throughput. It depends on the number of relays, the data transmission duration, and the specific form of adaptation used. We also analyzed the relative energy overhead of the relay selection phase and showed that it can be significant especially when the selected relay adapts its transmit power. Our work encourages a careful analysis and optimization of this trade-off for other slower, but perhaps more energy-efficient, selection mechanisms as well.

APPENDIX

A. Proof of Lemma 1

For $b = 1$, the probability, $p(k, 1)$, that any one relay, among the k that collided, transmits in the next slot equals $k/2^k$. This is because the new interval contains only half of the probability mass of the interval that was split and the metrics of the relays are independent and identically distributed. When $b > 1$, the following three cases arise:

- 1) The next slot is idle: This case happens with probability $\binom{k}{0}/2^k$. Now, the collision among k relays needs to be resolved in exactly $b - 1$ slots, which occurs with probability $p(k, b - 1)$.
- 2) $i \geq 2$ relays collide in the next slot: This occurs with probability $\binom{k}{i}/2^k$. The probability that the collision can be resolved in exactly $b - 1$ slots is $p(i, b - 1)$.
- 3) The next slot is a success: In this case, the collision is resolved in strictly fewer than b slots. Therefore, the probability that it will be resolved in exactly b slots is 0.

B. Proof of Theorem 1

We first derive the probability that the i^{th} slot is the first non-idle slot and that $k \geq 1$ relays collide in it. When $i \leq r$, this happens with probability $\binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{i}{n}\right)^{n-k}$.

The case where $i = r + 1$ is slightly different. From (2), relays whose RD channel gains are below a_r set their metrics to be 0. This occurs with probability $1 - \alpha$, where $\alpha = F_c(a_r)$. Now, the probability that the first non-idle slot is the $(r + 1)^{\text{th}}$ slot with k relays involved in it is the probability that k relays have their metric between $F_c^{-1}\left(\frac{\alpha}{n}\right)$ and a_r , and the remaining $n - k$ relays have metric 0. Therefore, this probability equals $\binom{n}{k} \left(\alpha - \frac{\alpha}{n}\right)^k (1 - \alpha)^{n-k}$. No non-idle slot can occur after the $(r + 1)^{\text{th}}$ slot. If k relays are involved in the collision in the i^{th} slot, the probability that the collision is resolved in the remaining $T_c - i$ slots is $\sum_{j=1}^{T_c-i} p(k, j)$.

Depending on T_c and r , the following two cases are possible:

- 1) $T_c \leq r$: Successful data transfer occurs if in the first non-idle slot (say the i^{th} slot): (i) a success occurs,

which happens with probability $\binom{n}{1} \left(\frac{1}{n}\right) \left(1 - \frac{i}{n}\right)^{n-1}$, or (ii) $k \geq 2$ relays collide, which happens with probability $\binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{i}{n}\right)^{n-k}$, and are resolved within $T_c - i$ slots. Thus,

$$P_s(T_c) = \sum_{i=1}^{T_c} \left(1 - \frac{i}{n}\right)^{n-1} + \sum_{i=1}^{T_c-1} \sum_{k=2}^n \binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{i}{n}\right)^{n-k} \sum_{j=1}^{T_c-i} p(k, j). \quad (20)$$

- 2) $T_c > r$: Successful data transfer occurs if during the first non-idle slot ($i \leq r + 1$) (i) a success occurs, or (ii) if $k \geq 2$ relays collide and this is resolved within $T_c - i$ slots. If a collision occurs in the $(r + 1)^{\text{th}}$ slot, the interval that needs to be split has a probability mass $\alpha - \frac{\alpha}{n}$ since only relays with non-zero metrics contend. The expression for $P_s(T_c)$ becomes

$$P_s(T_c) = \sum_{i=1}^r \left(1 - \frac{i}{n}\right)^{n-1} + n \left(\alpha - \frac{\alpha}{n}\right) (1 - \alpha)^{n-1} + \sum_{i=1}^r \sum_{k=2}^n \binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{i}{n}\right)^{n-k} \sum_{j=1}^{T_c-i} p(k, j) + \sum_{k=2}^n \binom{n}{k} \left(\alpha - \frac{\alpha}{n}\right)^k (1 - \alpha)^{n-k} \sum_{j=1}^{T_c-r-1} p(k, j).$$

Equation (7) compactly expresses the above two results using the indicator function.

C. Proof of Lemma 2

When $b = 1$ and k relays are involved in a collision, the average numbers of relays that transmit in the next slot is $N(k, 1) = k/2$ since the probability that a relay's metric lies in the split interval is $1/2$. When $b > 1$, the following three cases arise in calculating $N(k, b)$: (i) The next slot is idle: This case happens with probability $\binom{k}{0}/2^k$. The average number of relays that transmit thereafter is $N(k, b - 1)$. (ii) The next slot is a collision among $i (\geq 2)$ relays: This happens with probability $\binom{k}{i}/2^k$. The average number of relays that transmit thereafter is $N(i, b - 1)$. (iii) The next slot is a success: No more relays transmit after this. Thus, $N(1, b) = 0$.

D. Proof of Theorem 2

If the first non-idle slot is the i^{th} slot and k relays are involved in it, the average total number of relay transmissions is, by definition of $N(\cdot, \cdot)$, is equal to $N(k, T_c - i) + k$. As in the proof of Theorem 1 in Appendix B, the probability that the first non-idle slot is the i^{th} slot and k relays are involved is $\binom{n}{k} \left(\frac{p_c}{n}\right)^k \left(1 - \frac{ip_c}{n}\right)^{n-k}$, for $i \leq r$, and is

$\binom{n}{k} \left(\alpha - \frac{r}{n}\right)^k (1 - \alpha)^{n-k}$, for $i = r + 1$. Thus, we get

$$\begin{aligned} E_c &= \rho_c t_{\text{slot}} \sum_{i=1}^{\min(T_c, r)} \sum_{k=1}^n \binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{i}{n}\right)^{n-k} (N(k, T_c - i) + k) \\ &\quad + \rho_c t_{\text{slot}} I_{\{T_c > r\}} \sum_{k=1}^n \binom{n}{k} \left(\alpha - \frac{r}{n}\right)^k (1 - \alpha)^{n-k} \\ &\quad \times (N(k, T_c - r - 1) + k). \quad (21) \end{aligned}$$

Simplifying further, we get

$$\begin{aligned} E_c &= \rho_c t_{\text{slot}} \sum_{i=1}^{\min(T_c, r)} \left(\sum_{k=1}^n \binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{i}{n}\right)^{n-k} N(k, T_c - i) \right. \\ &\quad \left. + \sum_{k=1}^n k \binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{i}{n}\right)^{n-k} \right) \\ &\quad + \rho_c t_{\text{slot}} I_{\{T_c > r\}} \left(\sum_{k=1}^n \binom{n}{k} \left(\alpha - \frac{r}{n}\right)^k (1 - \alpha)^{n-k} \right. \\ &\quad \left. \times N(k, T_c - r - 1) + \sum_{k=1}^n k \binom{n}{k} \left(\alpha - \frac{r}{n}\right)^k (1 - \alpha)^{n-k} \right). \quad (22) \end{aligned}$$

Using combinatorial identities and simplifying results in (9).

E. Proof of Lemma 3

Let M be the random variable that denotes the number of slots required by the selection phase. Thus, $m(T_c) = \mathbf{E}[M] = \sum_{j=0}^{T_c-1} \Pr(M > j)$. Now, $\Pr(M > 0) = 1$, since at least one slot is required. And, $\Pr(M > j) = 1 - \Pr(M \leq j)$, for $j \geq 1$. Thus, we get

$$m(T_c) = 1 + \sum_{j=1}^{T_c-1} (1 - \Pr(M \leq j)). \quad (23)$$

Let μ_{\max} denote the maximum metric's value. From the law of total probability, we have

$$\begin{aligned} \Pr(M \leq j) &= \Pr(M \leq j, \mu_{\max} \geq a_r) \\ &\quad + \Pr(M \leq j, \mu_{\max} < a_r). \quad (24) \end{aligned}$$

When $\mu_{\max} < a_r$, which occurs with probability $(1 - \alpha)^n$, the selection phase lasts for $M = r + 1$ slots. Otherwise, when $\mu_{\max} \geq a_r$, $\Pr(M \leq j)$ is exactly $P_s(j)$. Therefore, $\Pr(M \leq j) = P_s(j)$, for $1 \leq j \leq r$, and $\Pr(M \leq j) = P_s(j) + (1 - \alpha)^n$, for $r + 1 \leq j \leq T_c - 1$. Substituting this in (24) and (23) gives the required result.

F. Proof of Theorem 3

First, we derive the expression for $E(u, v; k)$. Given that a metric with CCDF $F_c(\cdot)$ lies in $(F_c^{-1}(v), F_c^{-1}(u))$, its probability distribution function (PDF) is $f_1(\mu) = -F_c'(\mu)/(v - u)$ and its cumulative distribution function (CDF) is $F_1(\mu) = (v - F_c(\mu))/(v - u)$. From (2), the energy transmitted given metric μ equals $\frac{\left(2^{\frac{B}{W T_d}} - 1\right) N_0}{\kappa \mu} T_d$. For the node with highest

metric selected from among k relays, it follows from order statistics [32] that the average energy consumption equals

$E(u, v; k) = \frac{\left(2^{\frac{B}{W T_d}} - 1\right) N_0}{\kappa} T_d \int_{F_c^{-1}(v)}^{F_c^{-1}(u)} \frac{1}{\mu} k f_1(\mu) F_1(\mu)^{k-1} d\mu$, which yields (13). For symmetric channel gains, it follows from (4) that $f_1(\mu) = \frac{|\mathbf{x}_d - \mathbf{x}_0|^{\epsilon} e^{-a_s |\mathbf{x}_0|^{\epsilon}} e^{-\mu |\mathbf{x}_d - \mathbf{x}_0|^{\epsilon}}}{v - u}$ and $F_1(\mu) = \frac{e^{-a_s |\mathbf{x}_0|^{\epsilon}}}{v - u} (e^{a_s |\mathbf{x}_0|^{\epsilon}} v - e^{-\mu |\mathbf{x}_d - \mathbf{x}_0|^{\epsilon}})$. Substituting these in the above integral and simplifying gives the desired result.

Given that the first non-idle slot is the i^{th} slot and k relays are involved in it, the average energy transmitted is $E\left(\frac{i-1}{n}, \frac{i}{n}; k\right)$, if $i \leq r$, and is $E\left(\frac{r}{n}, \alpha; k\right)$, if $i = r + 1$. The desired result follows by combining this with the probability that the i^{th} slot is the first non-idle slot, which was derived in Appendix B.

G. Proof of Theorem 4

First, we derive the expression for $T(u, v; k)$. The time taken by the relay with a metric μ to transmit to the destination equals $\frac{B}{W \log_2\left(1 + \frac{\rho r \kappa}{N_0} \mu\right)}$. As in Appendix F, the PDF and CDF of the metric, given that it lies in $(F_c^{-1}(v), F_c^{-1}(u))$, equal $f_1(\mu) = \frac{-F_c'(\mu)}{v - u}$ and $F_1(\mu) = \frac{v - F_c(\mu)}{v - u}$. Using order statistics [32], the average transmit time required by the best relay selected among k relays with metrics in the above interval is then

$$T(u, v; k) = \frac{B}{W} \int_{F_c^{-1}(v)}^{F_c^{-1}(u)} \frac{k f_1(\mu) F_1(\mu)^{k-1}}{\log_2\left(1 + \frac{\rho r \kappa}{N_0} \mu\right)} d\mu, \quad (25)$$

which yields (17). For the symmetric case, this can be further simplified by substituting $f_1(\mu) = \frac{|\mathbf{x}_d - \mathbf{x}_0|^{\epsilon} e^{-a_s |\mathbf{x}_0|^{\epsilon}} e^{-\mu |\mathbf{x}_d - \mathbf{x}_0|^{\epsilon}}}{v - u}$ and $F_1(\mu) = \frac{e^{-a_s |\mathbf{x}_0|^{\epsilon}}}{v - u} (e^{a_s |\mathbf{x}_0|^{\epsilon}} v - e^{-\mu |\mathbf{x}_d - \mathbf{x}_0|^{\epsilon}})$ in the above integral and simplifying. Doing so eventually leads to

$$\begin{aligned} T(u, v; k) &= \frac{B}{W} \frac{k |\mathbf{x}_d - \mathbf{x}_0|^{\epsilon} e^{-k a_s |\mathbf{x}_0|^{\epsilon}}}{(v - u)^k} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} \\ &\quad \times \left(v e^{a_s |\mathbf{x}_0|^{\epsilon}} \right)^{k-1-j} \int_{F_c^{-1}(v)}^{F_c^{-1}(u)} \frac{e^{-(j+1)\mu |\mathbf{x}_d - \mathbf{x}_0|^{\epsilon}}}{\log_2\left(1 + \frac{\rho r \kappa}{N_0} \mu\right)} d\mu. \quad (26) \end{aligned}$$

Further simplification using variable substitutions and $\psi_1(a, u) = \int_u^{\infty} \frac{1}{t} \exp\left(t + \frac{1-e^t}{a}\right) dt$ yields (19).

Given that the first non-idle slot is the i^{th} slot and k relays are involved in it, the average time required is $T\left(\frac{i-1}{n}, \frac{i}{n}; k\right)$, if $i \leq r$, and $T\left(\frac{r}{n}, \alpha; k\right)$, if $i = r + 1$. The probability of this event was derived in the proof of Theorem 1 in Appendix B. Combining this with (25) yields the desired result.

REFERENCES

- [1] Y. Zhao, R. Adve, and T. J. Lim, "Symbol error rate of selection amplify-and-forward relay systems," *IEEE Commun. Lett.*, vol. 10, pp. 757-759, Nov. 2006.
- [2] D. S. Michalopoulos and G. K. Karagiannidis, "Performance analysis of single relay selection in Rayleigh fading," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 3718-3724, Oct. 2008.
- [3] Y. Jing and H. Jafarkhani, "Single and multiple relay selection schemes and their achievable diversity orders," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 1414-1423, Mar. 2009.

- [4] Y. Zhao, R. Adve, and T. J. Lim, "Improving amplify-and-forward relay networks: optimal power allocation versus selection," in *Proc. IEEE Int. Symp. Inf. Theory*, pp. 1234-1238, 2006.
- [5] Z. Zhou, S. Zhou, J.-H. Cui, and S. Cui, "Energy-efficient cooperative communication based on power control and selective single-relay in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 3066-3078, Aug. 2008.
- [6] 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 (Spring)*, pp. 1489-1493, 2005.
- [7] Z. Lin and E. Erkip, "Relay search algorithms for coded cooperative systems," in *Proc. Globecom*, 2005.
- [8] V. Sreng, H. Yanikomeroglu, and D. Falconer, "Relayer selection strategies in cellular networks with peer-to-peer relaying," in *Proc. VTC (Fall)*, pp. 1949-1953, 2003.
- [9] T. C.-Y. Ng and W. Yu, "Joint optimization of relay strategies and resource allocations in cooperative cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 328-339, Feb. 2007.
- [10] R. Madan, N. B. Mehta, A. F. Molisch, and J. Zhang, "Energy-efficient cooperative relaying over fading channels with simple relay selection," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 3013-3025, Aug. 2008.
- [11] P. Liu, Z. Tao, Z. Lin, E. Erkip, and S. Panwar, "Cooperative wireless communications: a cross-layer approach," *IEEE Trans. Wireless Commun.*, vol. 13, pp. 84-92, Aug. 2006.
- [12] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Sel. Areas Commun.*, vol. 24, pp. 659-672, Mar. 2006.
- [13] V. Shah, N. B. Mehta, and R. Yim, "Optimal timer based selection schemes," *IEEE Trans. Commun.*, vol. 58, pp. 1814-1823, June 2010.
- [14] X. Qin and R. Berry, "Opportunistic splitting algorithms for wireless networks," in *Proc. INFOCOM*, pp. 1662-1672, Mar. 2004.
- [15] R. Yim, N. B. Mehta, and A. F. Molisch, "Fast multiple access selection through variable power transmission," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 1962-1973, Apr. 2009.
- [16] V. Shah, N. B. Mehta, and R. Yim, "Splitting algorithms for fast decentralized cooperative relay selection," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 1525-1535, Apr. 2010.
- [17] C. K. Lo, J. R. W. Heath, and S. Vishwanath, "The impact of channel feedback on opportunistic relay selection for hybrid-ARQ in wireless networks," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 1255-1268, Mar. 2009.
- [18] A. Bletsas, H. Shin, and M. Win, "Cooperative communications with outage-optimal opportunistic relaying," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 3450-3460, Sept. 2007.
- [19] N. B. Mehta, V. Sharma, and G. Bansal, "Queued cooperative wireless networks with rateless codes," *Proc. Globecom*, 2008.
- [20] E. Beres and R. Adve, "Selection cooperation in multi-source cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 118-127, Jan. 2008.
- [21] A. F. Molisch, *Wireless Communications*. Wiley-IEEE Press, 2005.
- [22] T. Starr, J. M. Cioffi, and P. J. Silverman, *Understanding Digital Subscriber Line Technology*. Prentice Hall PTR, 1 ed., 1999.
- [23] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2349-2360, 2005.
- [24] "Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," Tech. Rep. IEEE Std 802.11-2007, IEEE Computer Society, June 2007.
- [25] D. S. Michalopoulos and G. K. Karagiannidis, "PHY-layer fairness in amplify and forward cooperative diversity systems," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1073-1083, Mar. 2008.
- [26] J. K. Cavers, "Variable-rate transmission for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. COM-20, pp. 15-22, Feb. 1972.
- [27] A. J. Goldsmith and S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 1218-1230, Oct. 1997.
- [28] M.-S. Alouini and A. J. Goldsmith, "Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 1165-1181, July 1999.
- [29] T. Haustein, von Helmlolt, E. C. Jorswieck, Jungnickel, and V. V. Pohl, "Performance of MIMO systems with channel inversion," *Proc. VTC (Spring)*, pp. 35-39, 2002.
- [30] S. Cui, R. Madan, A. J. Goldsmith, and S. Lall, "Cross-layer energy and delay optimization in small-scale sensor networks," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 3688-3699, Oct. 2007.
- [31] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*. Academic Press, 4th ed., 1980.
- [32] H. A. David and H. N. Nagaraja, *Order Statistics*. Wiley Series in Probability and Statistics, 3 ed., 2003.



tion theory.

Virag Shah received his Bachelor of Engineering degree in Electronics from University of Mumbai, India in 2007. He received his Master of Engineering degree in Telecommunications from Indian Institute of Science, Bangalore, India in 2009. Since November 2009, he has been a Research Fellow at Indian Institute of Technology, Bombay, India working in the area of communication networks. His research interests include the design and analysis of algorithms for communication networks, network control, queuing theory, graph theory and information



Neelesh B. Mehta (S'98-M'01-SM'06) received his Bachelor of Technology degree in Electronics and Communications Engineering from the Indian Institute of Technology (IIT), 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 is now an Assistant Professor at the Dept. of Electrical Communication Engineering, Indian Institute of Science (IISc), Bangalore, India. Prior to joining IISc, he has held research positions in the Wireless

Systems Research group in AT&T Laboratories, Middletown, NJ, USA, Broadcom Corp., Matawan, NJ, USA, and Mitsubishi Electric Research Laboratories, Cambridge, MA, USA.

His research includes work on link adaptation, multiple access protocols, system-level performance analysis of cellular systems, MIMO and antenna selection, and cooperative communications. He was also actively involved in radio access network physical layer (RAN1) standardization activities in 3GPP. He has served on several TPCs, and was a TPC co-chair for the Transmission technologies track of VTC 2009 (Fall) and the Frontiers of Networking and Communications symposium of Chinacom 2008. He is an Editor of the IEEE Transactions on Wireless Communications and an executive committee member of the Bangalore chapter of the IEEE Signal Processing Society and the IEEE Bangalore section.



Raymond Yim (S'00-M'07) received his B. Eng. and M. Eng. degrees in electrical and computer engineering from McGill University in Canada, and Ph.D. from Harvard University. In 2006-2007, he was an assistant professor in electrical and computer engineering at F. W. Olin College of Eng., Needham, MA. Since 2008, he has been a research scientist at the Mitsubishi Electric Research Laboratories, Cambridge, MA. His research interests include the design and analysis of cross-layered architectures and protocols for wireless communication networks,

multiple access protocols, link adaptation, cooperative communications, smart antenna systems, interference management and intelligent transportation systems.