Throughput and Diversity Gain of Buffer-Aided Relaying

Nikola Zlatanov††, Robert Schober††, and Petar Popovski††
††The University of British Columbia, ††Aalborg University

Abstract—In this paper, we consider a simple network consisting of a source, a half-duplex decode-and-forward relay, and a destination. In contrast to most of the existing literature, we assume that the relay is equipped with a buffer and show that this can lead to substantial performance gains. We propose a simple protocol which chooses either the source-relay or the relay-destination link for transmission depending on the instantaneous channel state information. For this simple protocol, we derive the throughput for adaptive rate transmission and the outage probability for fixed rate transmission. Our results show that, unlike conventional relaying, buffer-aided relaying yields a diversity gain of two. In addition, throughput gains of up to 100% compared to conventional relaying are possible.

I. INTRODUCTION

The capacity of a three-node transmission systems consisting of a source, a destination, and a single half-duplex decode-and-forward (DF) relay was investigated in [1]. Later, [2] and [3] have sparked renewed interest in relay-assisted communication. Since then the simple three-node system has become a building block for more sophisticated relay networks and has led to the creation of a host of cooperative communication techniques [4]-[7].

It was shown in [1] that the capacity of a three-node DF network without a direct source-destination link is given by the minimum of the source-relay link capacity and the relay-destination link capacity. Furthermore, in [4] the outage probability of this simple network was obtained. In [4]-[7] and most of the existing literature on relaying, it is assumed that relays receive a packet from the source in one time slot and forward it to the destination in the next time slot, which will be referred to as “conventional relaying” in the following. In this paper, we will abandon this paradigm and give relays the freedom to decide in which time slot to receive and in which time slot to transmit. This new approach requires the relays to have buffers. Relays with buffers were also considered in [8] and [9]. In [8], the buffer at the relay was used to enable the relay to receive for a fixed number of time slots before retransmitting the received information in a fixed number of time slots. In [9], relay selection was considered and buffers enable the selection of the relay with the best source-relay channel for reception and the best relay-destination channel for transmission. However, like in conventional relay selection without buffers [10], the source transmits in every other time slot. Thus, the relay selection scheme in [9] achieves a signal-to-noise ratio (SNR) gain compared to conventional relay selection but no diversity gain. Both [8] and [9] do not fully exploit the flexibility offered by relays with buffers since the schedule of when the source transmits and when a relay transmits is a priori fixed.

In this paper, we assume that one of the three nodes of the network acquires the channel state information (CSI) of the source-relay and relay-destination links and subsequently exploits this CSI to decide which link is used for transmission in the considered time slot. We propose a link selection criterion and analyze the throughput and the outage probability of the proposed buffer-aided relaying scheme assuming an infinite buffer size. The impact of finite buffer sizes is studied via simulations. We show that buffer-aided relaying achieves significant throughput and diversity gains compared to conventional relaying without buffers since a link is only utilized for transmission when its CSI is good. These gains come at the expense of an additional delay that is introduced by the buffer.

Thus, the proposed scheme is most suitable for applications which require high power and bandwidth efficiency but can tolerate delays.

The remainder of the paper is organized as follows. In Section II, the considered system model is presented. The optimal link selection threshold is derived in Sections III, and the outage probability is analyzed in Section IV. In Section V, numerical results are presented, and some conclusions are drawn in Section VI.

II. SYSTEM MODEL

We consider a three-node communication system comprising a source $S$, a relay $R$, and a destination $D$. The source sends packets to the half-duplex relay, which decodes these packets, possibly stores them in its buffer, and eventually sends them to the destination. We assume that time is divided into slots of equal lengths. The instantaneous link SNRs of the $S$-$R$ and $R$-$D$ channels in the $i$th time slot are denoted by $s(i)$ and $r(i)$, respectively. We assume that the link SNRs are constant during one time slot but change from one time slot to the next due to e.g. the mobility of the involved nodes and/or frequency hopping. $s(i)$ and $r(i)$ are modeled as mutually independent, non-negative random processes with expected values $\Omega_S = E\{s(i)\}$ and $\Omega_R = E\{r(i)\}$, respectively, where $E\{\cdot\}$ denotes expectation. The average link SNRs $\Omega_S$ and $\Omega_R$ can be decomposed as $\Omega_S = \gamma \Omega_S$ and $\Omega_R = \gamma \Omega_R$, where $\gamma = P_T/\sigma_n^2$ denotes the transmit SNR with $P_T$ and $\sigma_n^2$ representing the transmit power and the variance of the additive white Gaussian noise at the relay and the destination, respectively, and $\Omega_S$ and $\Omega_R$ are the variances of the source-relay and relay-destination channel gains, respectively. Here, we assume that the source and the relay transmit with identical power $P_T$, i.e., power allocation is not considered. We note that in general we do not require $s(i)$ and $r(i)$ to be fully temporally uncorrelated. However, this assumption will be made in some parts of the paper to facilitate analysis.

In a given time slot $i$, one of the nodes acquires knowledge of $s(i)$ and $r(i)$ and makes, based on this knowledge, a decision whether the source or the relay should transmit. Subsequently, this node broadcasts its decision to the other nodes and transmission in time slot $i$ begins. Throughout this
paper, we assume that the source node has always data to transmit and the buffer at the relay is not limited in size. The effect of limited buffer size will be investigated in Section V via simulations.

In the following, we discuss two different modes of transmission.

A. Adaptive Rate Transmission

We first consider adaptive transmission, where the source and the relay know $s(i)$ and $r(i)$, respectively, and adjust their transmission rates such that they transmit with the maximum possible rate without causing outages.

**Source transmits:** If the source transmits in a given time slot $i$, it transmits with rate

$$S_{SR}(i) = \log_2(1 + s(i)). \quad (1)$$

In the following, for convenience we normalize the number of bits transmitted in one time slot to the number of symbols per time slot. Thus, the data rate is identical to the normalized number of bits per time slot. Hence, the relay receives $S_{SR}(i)$ data bits from the source and appends them to the queue in its buffer. The number of bits in the buffer of the relay at the end of the $i$-th time slot is denoted by $Q(i)$ and given by

$$Q(i) = Q(i-1) + S_{SR}(i). \quad (2)$$

**Relay transmits:** If the relay transmits in a given time slot, the number of bits transmitted by the relay is given by

$$R_{RD}(i) = \min\{\log_2(1 + r(i)), Q(i-1)\}, \quad (3)$$

where we take into account that the maximal number of bits that can be send by the relay is limited by the number of bits in the buffer. The number of data bits remaining in the buffer at the end of time slot $i$ is given by

$$Q(i) = Q(i-1) - R_{RD}(i), \quad (4)$$

which is always non-negative because of (3).

**Throughput:** Because of the half-duplex constraint, we have $R_{RD}(i) = 0$ and $S_{SR}(i) = 0$ if the source transmits (and the relay listens) and the relay transmits, respectively. Thus, assuming the source has always data to transmit, the average (normalized) number of bits that arrive at the destination per time slot is given by

$$\tau = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} R_{RD}(i), \quad (5)$$

i.e., $\tau$ is the throughput of the considered communication system. In Section III, we will introduce a link selection criterion and optimize for maximization of $\tau$.

B. Fixed Rate Transmission

If the destination selects the link used for transmission, the destination can estimate the CSI of the $R$-$D$ link and acquire the CSI of the $S$-$R$ link from the relay through a feedback channel. In this case, the source and the relay are not aware of $s(i)$ and $r(i)$, respectively, and as a consequence, are not able to adapt their transmission rates to the channel state. Thus, the source and the relay transmit with fixed rates $R_S$ and $R_R$, respectively. Therefore, if the source is selected for transmission in time slot $i$, we have $S_{SR}(i) = R_S$ and $R_{RD}(i) = 0$. In contrast, if the relay is selected for transmission in time slot $i$, we have $R_{RD}(i) = \min\{R_R, Q(i-1)\}$ and $S_{SR}(i) = 0$. The queue state of the buffer still changes according to (2) and (4).

For fixed rate transmission, in time slot $i$ an outage occurs if the source is selected for transmission and $\log_2(1 + s(i)) < R_S$ or if the relay is selected for transmission and $\log_2(1 + r(i)) < \min\{R_R, Q(i-1)\}$. The outage probability of fixed rate transmission will be derived in Section IV.

III. ADAPTIVE RATE TRANSMISSION

In this section, we derive a link selection criterion and the corresponding throughput for adaptive rate transmission.

A. Problem Formulation

Let $d_i \in \{0, 1\}$ denote a binary decision variable. We set $d_i = 1$ if the $R$-$D$ channel is selected for transmission in time slot $i$, i.e., the relay transmits and the destination receives. Similarly, we set $d_i = 0$ if the $S$-$R$ channel is selected for transmission in time slot $i$, i.e., the source transmits and the relay receives. Exploiting $d_i$, the number of bits send from the source to the relay and from the relay to the destination in time slot $i$ can be written in compact form as

$$S_{SR}(i) = (1 - d_i)S(i) \quad (6)$$

and

$$R_{RD}(i) = d_i \min\{R(i), Q(i-1)\}, \quad (7)$$

respectively, where $S(i) = \log_2(1 + s(i))$ and $R(i) = \log_2(1 + r(i))$. Consequently, the throughput in (5) can be rewritten as

$$\tau = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} d_i \min\{R(i), Q(i-1)\}. \quad (8)$$

In order to limit complexity, we consider link selection criteria which only exploit the CSI in the considered time slot, i.e.,

$$d_i = \begin{cases} 1 & \text{if } F(r(i)) \geq \rho F(s(i)) \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

where $\rho$ and $F(\cdot)$ are referred to as decision threshold and decision function, respectively. The decision function is a non-negative, smooth, and increasing function, i.e., $F(x + \epsilon) > F(x)$ for $\epsilon > 0$. We assume that the inverse of $F(\cdot)$ exists and denote it by $F^{-1}(\cdot)$. In order to maximize the throughput in (8) for a given decision function $F(\cdot)$, $\rho$ has to be optimized, cf. Section III-B. We note that as will be explained later, if $\rho$ is optimized, the performance of the link selection criterion in (9) cannot be improved by also taking into account the queue state for link selection, cf. Remark 1. On the other hand, the choice of the decision function influences the achievable throughput and the computational complexity required to find the optimal $\rho$, cf. Sections III-C, and V.

B. Optimal $\rho$ for General $F(\cdot)$

Let us first define the mean arrival rate, $A$, and the mean departure rate, $D$, of the queue in the buffer of the relay as

$$A = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} (1 - d_i)S(i) \quad (10)$$

and
and
\[ D = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} d_i \min\{R(i), Q(i-1)\}, \] (11)
respectively. We note that the departure rate of the queue is equal to the throughput. The queue is said to be unstable or absorbing if \( A > D = \tau \). The following theorem characterizes the optimal \( \rho \) in terms of the state of the queue in the buffer of the relay.

**Theorem 1:** For a given decision function \( F(\cdot) \), the optimal \( \rho, \rho_{\text{opt}} \), that maximizes the throughput is found at the point where the queue in the buffer of the relay switches from an absorbing to a non-absorbing queue, i.e., the queue is non-absorbing for \( \rho = \rho_{\text{opt}} - \epsilon \) and absorbing for \( \rho = \rho_{\text{opt}} + \epsilon \), where \( \epsilon \to 0^+ \) denotes an arbitrarily small positive number.

**Proof:** From (9) and (10), we observe that \( A \) is an increasing function of \( \rho \). As \( \rho \) increases, we have \( d_i = 0 \) for all time slots. Thus, \( A \) assumes its minimum value \( \rho \to \rho_{\text{opt}} \), and the assumption that \( s = \rho \) is valid and equality holds if and only if the queue is non-absorbing. If we set \( \rho = \infty \), we have \( d_i = 0 \) for all time slots. Thus, \( A \) assumes maximum value and \( \tau \) assumes its minimum value \( \tau = 0 \). The queue at the relay absorbs all bits transmitted by the source. By gradually decreasing \( \rho \), we decrease \( A \), increase \( \tau \), and decrease the amount of data that gets absorbed by the queue. When we reach \( \rho = \rho_{\text{opt}} \) for which \( A = \tau_{\rho = \rho_{\text{opt}}} \) but \( A > \tau_{\rho = \rho_{\text{opt}} + \epsilon} \), the throughput \( \tau \) is maximized. This is exactly the point when the queue at the relay switches from an absorbing queue to a non-absorbing queue. If we decrease \( \rho \) further, \( A = \tau \) is still valid and the queue is non-absorbing but \( A \) has a smaller value. This completes the proof. \[ \blacksquare \]

**Theorem 2:** Assuming non-negative, ergodic, and stationary random processes \( s(i) \) and \( r(i) \), for a given decision function \( F(\cdot) \), the optimal \( \rho, \rho_{\text{opt}} \), that maximizes the throughput is the solution to
\[ E\{1-d_i\}S(i) = E\{d_iR(i)\}, \] (12)
where \( d_i \) is defined in (9).

**Proof:** We know from Theorem 1 that \( A = \tau \) holds for \( \rho = \rho_{\text{opt}} \). Let us denote the set of indices \( i \) with \( R(i) \leq Q(i-1) \) by \( C \) and the set of indices \( i \) with \( R(i) > Q(i-1) \) by \( \bar{C} \). Let cardinalities of \( C \) and \( \bar{C} \) be \( M = |C| \) and \( N-M = |\bar{C}| \), respectively. If the queue is absorbing, on average the number of bits that are arriving at the relay exceeds the number of bits leaving the relay. Thus, \( R(i) \leq Q(i-1) \) holds almost always and \( M/N \to 0 \). Consequently, the throughput in (8) simplifies to
\[ \tau = \lim_{N \to \infty} \frac{1}{N} \sum_{i \in C} d_i R(i) + \lim_{N \to \infty} \frac{1}{N} \sum_{i \in \bar{C}} d_i Q(i-1) \]
\[ = \lim_{N \to \infty} \frac{N-M}{N} E\{d_i R(i)\} = E\{d_i R(i)\}, \] (13)
where we have exploited the ergodicity of \( r(i) \). Since according to Theorem 1, the queue is absorbing for \( \rho = \rho_{\text{opt}} - \epsilon \), \( \epsilon \to 0^+ \), and \( F(\cdot) \) is a smooth function, (13) is also valid for \( \rho = \rho_{\text{opt}} \). Hence, (12) follows directly from Theorem 1, (13), and the assumption that \( s(i) \) is ergodic. This completes the proof. \[ \blacksquare \]

**Remark 1:** According to Theorem 2, for the optimal value of \( \rho \), the queue is non-absorbing but is almost always filled to such a level that the number of bits in the queue exceed the number of bit that can be transmitted over the relay-destination channel. Thus, for the optimal value of \( \rho \), the performance of the link selection criterion in (9) cannot be improved by also taking into account the state of the queue at the relay.

We are now ready to provide a formula that can be used for computation of the optimal \( \rho \).

**Lemma 1:** Denote the probability density functions (pdfs) of \( s(i) \) and \( r(i) \) by \( f_s(s(i)) \) and \( f_r(r(i)) \), respectively. Then, the optimal \( \rho \) is the solution of
\[ \int_0^\infty \log_2(1+s)f_s(s)ds - \int_0^\infty \log_2(1+r)f_r(r)dr - \int_0^\infty \log_2(1+r)f_r(r)dr f_s(s)(12) \]
where the integral limits are given by \( G(r) = F^{-1}(F(r)/\rho) \) and \( H(s) = F^{-1}(sF(s)) \).

**Proof:** The proof follows directly from Theorem 2 and basic probability theory. \[ \blacksquare \]

**C. Throughput in Rayleigh Fading**

In this subsection, we assume Rayleigh fading, i.e., the link SNRs \( s(i) \) and \( r(i) \) are exponentially distributed with means \( \Omega_S \) and \( \Omega_R \), respectively. Their pdfs are given by \( f_s(s) = e^{-s/\Omega_S}/\Omega_S \) and \( f_r(r) = e^{-r/\Omega_R}/\Omega_R \). In the following, we will determine \( \rho_{\text{opt}} \) and the corresponding maximum throughput \( \tau_{\text{max}} \) for two different decision functions \( F(\cdot) \).

**1) \( F(x) = x \):** In this case, the limits \( G(r) \) and \( H(s) \) in (14) are given by \( G(r) = r/\rho \), and \( H(s) = ps \). Thus, after some elementary manipulations, (14) simplifies to
\[ \exp\left(\frac{1}{\Omega_S}\right) E_i\left(\frac{1}{\Omega_S}\right) - \exp\left(\frac{1}{\Omega_R}\right) E_i\left(\frac{1}{\Omega_R}\right) \]
\[ - \frac{\Omega_R}{\Omega_R + \rho \Omega_S} \exp\left(\frac{\Omega_R + \rho \Omega_S}{\Omega_S \Omega_R}\right) E_i\left(\frac{\Omega_R + \rho \Omega_S}{\Omega_S \Omega_R}\right) \]
\[ + \frac{\rho \Omega_S}{\Omega_R + \rho \Omega_S} \exp\left(\frac{\Omega_R + \rho \Omega_S}{\rho \Omega_S \Omega_R}\right) E_i\left(\frac{\Omega_R + \rho \Omega_S}{\rho \Omega_S \Omega_R}\right) = 0 \]
where \( E_i(x) = \int_x^\infty e^{-t/t} dt \), \( x > 0 \), denotes the exponential integral function. The optimal value of \( \rho \) for \( F(x) = x \), \( \rho_{\text{opt}}, \) can be obtained via a simple one dimensional search from (15).

The corresponding maximal throughput can be obtained from \( \tau_{\text{max},1} = E\{d\log_2(1+r)\} \) and can be computed as
\[ \tau_{\text{max},1} = \frac{1}{\ln(2)} \exp\left(\frac{1}{\Omega_S}\right) E_i\left(\frac{1}{\Omega_S}\right) \]
\[ - \frac{\rho \Omega_S}{\ln(2) \Omega_R + \rho \Omega_S} \exp\left(\frac{\Omega_R + \rho \Omega_S}{\rho \Omega_S \Omega_R}\right) E_i\left(\frac{\Omega_R + \rho \Omega_S}{\rho \Omega_S \Omega_R}\right), \]
where \( \rho = \rho_{\text{opt},1} \).

**Remark 2:** If \( \Omega_S = \Omega_R \), the solution to (15) is \( \rho_{\text{opt},1} = 1 \).
2) $F(x) = \log_2(1 + x)$: In this case, we have $G(r) = (1 + r)^2 - 1$ and $H(s) = (1 + s)^{\rho - 1}$ in (14). Thus, after some manipulations, we obtain from (14)

$$\int_0^\infty \left[ \exp \left( \frac{(r + 1)^2}{\Omega_S} - 1 \right) \right] \ln \left( (r + 1)^{r} \right) dr + e^{\frac{\Omega}{\Omega_S}} E_i \left( \frac{(r + 1)^2}{\Omega_S} \right) \times \frac{1}{\Omega_S} \exp \left( - \frac{r}{\Omega_S} \right) ds \quad \text{(17)}$$

$$- \int_0^\infty \left[ \exp \left( \frac{(s + 1)^2 - 1}{\Omega_R} \right) \ln \left( (s + 1)^{s} \right) \right] \times \frac{1}{\Omega_S} \exp \left( - \frac{s}{\Omega_S} \right) ds = 0.$$  

The optimal $\rho$, $\rho_{\text{opt},2}$, can be found numerically from (17). The corresponding maximum throughput is obtained as

$$\tau_{\text{max},2} = \left[ \frac{1}{\ln(2)} \int_0^\infty \left[ \exp \left( - \frac{r}{\Omega_S} \right) \right] \times \ln \left( (s + 1)^{s} \right) + e^{\frac{\Omega}{\Omega_S}} E_i \left( \frac{(s + 1)^s}{\Omega_R} \right) \right] \times \frac{1}{\Omega_S} \exp \left( - \frac{s}{\Omega_S} \right) ds,$$  

(18)

where $\rho = \rho_{\text{opt},2}$.

**Remark 3:** If $\Omega_S = \Omega_R$ then the solution to (17) is $\rho_{\text{opt},2} = 1$.

**Remark 4:** Since for the second decision function, $F(x) = \log_2(1 + x)$, the channel capacity is used as selection criterion, $\tau_{\text{max},2} \geq \tau_{\text{max},1}$ holds. In fact, $\tau_{\text{max},2}$ is the maximum achievable throughput for any decision function. However, the achievable gain compared to the first decision function, $F(x) = x$, is small, cf. Fig. 1, and the search complexity for the optimal $\rho$ is higher since (17) requires numerical integration.

IV. FIXED RATE TRANSMISSION

For fixed rate transmission, for simplicity, we only consider the decision function $F(x) = x$. The link is still selected according to (9). We assume that the relay adds the decoded bits to its queue regardless of whether or not they have been decoded correctly.

A. Optimal Decision Threshold

The following theorem provides the optimal decision threshold, $\rho_{\text{opt},3}$.

**Theorem 3:** Consider Rayleigh fading links and assume the source and the relay transmit with rates $R_S$ and $R_R$, respectively. In this case, for decision function $F(x) = x$, the optimal decision threshold maximizing the throughput is given by

$$\rho_{\text{opt},3} = \frac{\Omega_R R_R}{(\Omega_S R_S)}.$$

(19)

**Proof:** Using similar arguments as in the proof of Theorem 1 for the adaptive transmission case, it can be shown that the throughput is maximized when the queue switches from an absorbing queue to a non-absorbing queue. Using then similar arguments as in the proof of Theorem 2, we find that the optimal $\rho$ is the solution to

$$R_S(1 - E\{d_i\}) - E\{d_i\} R_R = 0.$$  

(20)

On the other hand, we obtain from (9) $E\{d_i\} = Pr\{r(i) > \rho\}$. Combining this with (20) leads directly to (12), which completes the proof. 

**Lemma 2:** The maximum throughput in the fixed rate case is given by

$$\tau_{\text{max},3} = R_S R_R/((3 + 2)R_S)$$

(21)

**Remark 5:** Although the maximum throughput for fixed rate transmission does not depend on $\Omega_S$ and $\Omega_R$, the corresponding outage probability does.

B. Outage Probability

For fixed rate transmission, outages are unavoidable. Thus, the outage probability is a relevant performance metric. To facilitate our analysis, we assume in this subsection that $(s(i)$ and $r(i)$ are temporally uncorrelated.

**Theorem 4:** Consider a buffer-aided relaying system where the source and the relay transmit with fixed rate, the decision function is $F(x) = x$ is used to select one of the links for transmission, and the source–relay and relay-destination channels are affected by temporally uncorrelated Rayleigh fading. In this case, the outage probability decays on a log-log scale with a slope of two with increasing transmit SNR $\gamma$, i.e., the diversity order is two.

**Proof:** An outage event occurs if the source transmits in time slot $i$ but the source-relay link is in outage or if the source-relay link was not in outage in time slot $i$ but the relay-destination link is in outage when the packet is retransmitted by the relay in time slot $j > i$. For simplicity, we use the notation $s = s(i)$, $r = r(i)$, $\bar{r} = r(j)$, and $\bar{s} = s(j)$. Note that $r$, $s$, $\bar{r}$, and $\bar{s}$ are mutually independent. Thus, the outage probability can be expressed as

$$P_{\text{out}} = Pr\{\log_2(1 + s) < R_S; \rho < r\} \text{ OR }$$

$$\{\log_2(1 + \bar{r}) < R_S; \rho < r\} \text{ AND }$$

$$\{\log_2(1 + \bar{s}) < \min\{R_S, Q(j - 1)\}; \bar{r} > \bar{\rho}\} \}$$

$$= \frac{Pr\{\log_2(1 + s) < R_S \text{ AND } \rho < r\}}{Pr\{\rho < r\}} +$$

$$Pr\{\log_2(1 + \bar{r}) < \min\{R_S, Q(j - 1)\} \text{ AND } \bar{r} > \bar{\rho}\}$$

$$- Pr\{\log_2(1 + s) < R_S \text{ AND } \rho < r\}$$

$$\times Pr\{\log_2(1 + \bar{s}) < \min\{R_S, Q(j - 1)\} \text{ AND } \bar{r} > \bar{\rho}\}$$

$$= P_1 + (1 - P_1)$$

$$\times \frac{Pr\{\rho < \bar{r} < 2^{\min\{R_S, Q(j - 1)\}} - 1\}}{Pr\{\bar{r} > \bar{\rho}\}},$$

(22)
where $P_1$ is given by

$$P_1 = \Pr\{r/\rho < s < 2R^S - 1\} / \Pr\{s > r/\rho\}$$

$$= 1 + \frac{\Omega_R}{\rho \Omega_S} \exp\left( -\frac{\Omega_R + \rho \Omega_S}{\Omega_S \Omega_R} (2R^S - 1) \right) - \frac{\Omega_R + \rho \Omega_S}{\rho \Omega_S} \exp\left( -\frac{2R^S - 1}{\Omega_S} \right).$$

(23)

Since $Q(j - 1)$ is unknown, we cannot obtain the exact outage probability. However, exploiting the fact that $P_1 < 1$ and the inequality $\Pr\{\rho S < \bar{r} < 2\min(R^S, Q(j - 1)) - 1\} \leq \Pr\{\rho S < \bar{r} < 2R^S - 1\}$, we obtain from (22) the upper bound

$$P_{out} \leq P_1 + P_2 - P_1 P_2,$$

where

$$P_2 = \Pr\{\rho S < \bar{r} < 2R^S - 1\} / \Pr\{\bar{r} > \rho S\}$$

$$= 1 + \frac{\rho \Omega_S}{\Omega_R} \exp\left( -\frac{\Omega_R + \rho \Omega_S}{\rho \Omega_S \Omega_R} (2R^S - 1) \right) - \frac{\Omega_R + \rho \Omega_S}{\rho \Omega_S} \exp\left( -\frac{2R^S - 1}{\Omega_R} \right).$$

(25)

Exploiting now $\Omega_S = \gamma \bar{\Omega}_S$ and $\Omega_R = \gamma \bar{\Omega}_R$ and the Taylor series expansion $e^{-x} = 1 - t + t^2/2$, we obtain for high transmit SNRs (i.e., $\gamma \rightarrow \infty$) from (24) the asymptotic bound $P_{out} \leq (\bar{\Omega}_R/\bar{\Omega}_S + \rho)(2R^S - 1)^2 + (1/\rho + \bar{\Omega}_S/\bar{\Omega}_R)(2R^S - 1)^2 \bar{\Omega}_R \gamma^2$. (26)

From (26), we directly observe that $P_{out}$ decays with slope two on a log-log scale for high SNR $\gamma$, which concludes the proof.

**Remark 6:** The tightness of the upper bound in (24) depends on the choice of $\rho$. For $\rho \geq \rho_{opt,3}$, the queue in the buffer is absorbing or at the boundary between the absorbing and non-absorbing state and $Q(j - 1) < R_R$ will occur with negligible probability. Thus, in this case, the upper bound in (24) is tight.

**Remark 7:** The diversity order of conventional, non-buffer-aided DF relaying is equal to one [4]. This means buffer-aided relaying achieves a diversity gain compared to conventional relaying.

**V. NUMERICAL AND SIMULATION RESULTS**

In this section, we evaluate the performance of buffer-aided DF relaying and compare it with that of conventional relaying. Throughout this section, we assume temporally independent Rayleigh fading.

**A. Throughput for Adaptive Rate Transmission**

In Fig. 1, we show the ratio of the optimal throughput of buffer-aided relaying, $\tau_{max}$, and the throughput of conventional relaying, $\tau_{conv}$, which is given by $\tau_{conv} = \min\{\exp(\Omega_S^{-1})E(\Omega_S^{-1}), \exp(\Omega_R^{-1})E(\Omega_R^{-1})\}/(2 \log_2(2))$, as a function of $\Omega_R/\Omega_S$ for several different values of $\Omega_S$. For buffer-aided relaying, we considered decision functions $F(x) = \log_2(1 + x)$ and $F(x) = x$ and calculated the corresponding throughputs based on (18) and (16), respectively.

The results in Fig. 1 have been confirmed by computer simulations. These simulations are not shown here for clarity of presentation.

Clearly, buffer-aided relaying leads to substantial throughput gains compared to conventional relaying. Both considered decision functions lead to a very similar performance, although at very high ratios $\Omega_R/\Omega_S$ the optimal decision function $F(x) = \log_2(1 + x)$ yields a small throughput gain. The ratio $\tau_{max}/\tau_{conv}$ approaches two as $\Omega_R/\Omega_S \rightarrow 0$ and $\Omega_R/\Omega_S \rightarrow \infty$. For $\Omega_R/\Omega_S \rightarrow 0$, buffer-aided relaying uses the source-relay link very rarely since comparatively large amounts of data can be transferred to the relay in a single time slot. Thus, the relay can almost always transmit as compared to half of the time in conventional relaying. On the other hand, for $\Omega_R/\Omega_S \rightarrow \infty$, it is the relay-destination channel that is used very rarely and the source can transmit almost all the time, which results in twice the throughput as for conventional relaying.

**B. Outage Probability**

For a fair comparison of the outage probabilities of buffer-aided relaying with decision function $F(x) = x$ and conventional relaying we require both schemes to have the same throughput. This is achieved by choosing $R_S = R_R = R_0$ for both relaying schemes. For buffer-aided relaying the decision threshold is chosen as specified in Theorem 3. The outage probabilities of buffer-aided and conventional relaying as functions of the transmit SNR $\gamma$ are shown in Fig. 2 for $\Omega_R = 1$, $R_0 = 2$ b/s/Hz, and different values of $\Omega_S$. The outage probability of conventional relaying was calculated based on the results in [4]. For the outage probability of buffer-aided relaying, we show the upper bound in (24) and computer simulations. As predicted in Remark 6, the upper bound is tight since the optimal $\rho$ was used and the event $Q(j - 1) < R_R$ has a negligible effect. Furthermore, as expected from Theorem 4, buffer-aided relaying achieves a diversity order of two, whereas conventional relaying achieves only a diversity order of one, which underlines the superiority of buffer-aided relaying.

**C. Limited Buffer Size**

So far, we have assumed that the size of the buffer at the relay is infinite. Clearly, in practice, this will not be the case. Thus, we investigate in Fig. 3 the effect of a limited buffer size $L$ on the outage probability. Here, we assume that one packet is transmitted in each time slot and $L$ denotes the number of packets that can be stored at the relay. We assume $R_S = R_R = R_0 = 2$ b/s/Hz and $\Omega_S = \Omega_R = 1$. The scheme with finite buffer size operates exactly the same way as the scheme with infinite buffer size except when the buffer is empty (full) the source (relay) is forced to transmit in the next time slot. In addition to the simulations for finite buffer size, we also show analytical results for buffer-aided relaying with infinite buffer size and conventional relaying. For all buffer-aided schemes the decision function $F(x) = x$ is adopted. As expected, as $L$ increases, buffer-aided relaying with finite buffer size approaches the performance of buffer-aided relaying with infinite buffer size. Note that the maximum delay of the considered system is equal to the buffer size. Therefore, even for moderate delays of 30 packets, significant performance gains compared to conventional relaying are possible. However, for any finite buffer size, the slope of the outage probability curves will approach unity for sufficiently
Fig. 1. Ratio $\tau_{\text{max}}/\tau_{\text{conv}}$ vs. $\Omega_R/\Omega_S$ for several different values of $\Omega_S$.

Fig. 2. Outage probability of buffer-aided (BA) and conventional relaying vs. $\gamma$. $\bar{\Omega}_S = 0.1; 1; 10$.

Fig. 3. Outage probability of buffer-aided (BA) relaying with limited buffer size $L$. $R_S = R_R = R_0 = 2$ b/s/Hz and $\bar{\Omega}_S = \bar{\Omega}_R = 1$.

VI. Conclusions

In this paper, we proposed a novel relaying protocol for relays with buffers. In contrast to conventional relaying, where the source and the relay transmit in successive time slots regardless of the channel state, in the proposed scheme, always the node with the stronger link is selected for transmission. We derived the throughput for adaptive rate transmission and the outage probability for fixed rate transmission. Furthermore, we could show that the diversity order of the proposed buffer-aided relaying scheme is two, which results in significant performance gains compared to conventional relaying. Thus, buffers seem an efficient means to improve the performance of relay networks for non-delay limited applications.

Interesting topics for future work include using the considered simple three-node network as a building block for larger networks, as has been done for relays without buffers before, and the analysis of buffer-aided relaying with buffers of finite size.

REFERENCES