Buffer-Aided Relaying with Adaptive Link Selection

Nikola Zlatanov, Robert Schober, and Petar Popovski

Abstract

In this paper, we consider a simple network consisting of a source, a half-duplex decode-and-forward relay, and a destination. We propose a new relaying protocol employing adaptive link selection, i.e., in any given time slot, based on the channel state information of the source-relay and the relay-destination link a decision is made whether the source or the relay transmits. In order to avoid data loss at the relay, adaptive link selection requires the relay to be equipped with a buffer such that data can be queued until the relay-destination link is selected for transmission. We study both delay constrained and delay unconstrained transmission. For the delay unconstrained case, we characterize the optimal link selection policy, derive the corresponding throughput, and develop an optimal power allocation scheme. For the delay constrained case, we propose to starve the buffer of the relay by choosing the decision threshold of the link selection policy smaller than the optimal one and derive a corresponding upper bound on the average delay. Furthermore, we propose a modified link selection protocol which avoids buffer overflow by limiting the queue size. Our analytical and numerical results show that buffer-aided relaying with adaptive link selection achieves significant throughput gains compared to conventional relaying protocols with and without buffers where the relay employs a fixed schedule for reception and transmission.

I. INTRODUCTION

The classical three-node relay channel was originally considered by van der Meulen [1]. Cover and El Gamal [2] investigated the capacity of a memoryless relay channel consisting of a source, a destination, and a single full-duplex relay. Recently, renewed interest in relay-assisted communication was sparked by [3] and [4]. Since then, the simple three-node system has become a building block for more sophisticated relay networks and a host of cooperative communication techniques have been proposed [5]-[12]. The capacity of a three-node relay channel comprised of a source, a destination, and a single half-duplex decode-and-forward (DF) relay was investigated in [5]. Under the assumption of full channel state information (CSI) at every node, it was shown in [5] that the capacity of a three-node DF network without a direct source-destination link is given by the minimum of the source-relay and the relay-destination link capacities for

N. Zlatanov and R. Schober are with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada, E-mail: zlatanov@ece.ubc.ca, rschober@ece.ubc.ca

Petar Popovski is with the Department of Electronic Systems, Aalborg University, Denmark, E-mail: petarp@es.aau.dk.

the case when the relay employs a fixed schedule for reception and transmission. In [5]-[12] and most of the existing literature on half-duplex relaying, it is assumed that the relays receive a packet from the source in one time slot and forward it to the destination in the next time slot. Relay protocols operating under this restriction are referred to as "conventional" relay protocols in the following. In this paper, we 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 [13] and [14]. In [13], 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 [14], relay selection was considered and buffers enable the selection of the relay with the best source-relay channel for reception and the relay with the best relay-destination channel for transmission. However, similar to conventional relay selection without buffers [15], the source transmits in every other time slot. Thus, both [13] and [14] 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 consider buffer-aided relaying with adaptive link selection, where in any given time slot based on the CSI of the source-relay and the relay-destination link a decision is made whether the source or the relay transmits. We consider the cases of delay constrained and delay unconstrained transmission. For the delay unconstrained case, we optimize the link selection protocol and the power allocated to the source and the relay. Interestingly, the optimal link selection policy requires only knowledge of the instantaneous CSI of the considered time slot and the statistical CSI of the involved links. However, the instantaneous CSI of past and future time slots and the state of the relay's buffer are not required for optimal link selection, which facilitates the implementation of the optimal policy. For the delay constrained case, we propose two alternative link selection protocols and provide an upper bound for the average delay of one of them. Our analytical and simulation results show, in good agreement, that buffer-aided relaying with adaptive link selection can achieve large performance gains compared to conventional relaying with or without buffer [13], as long as a certain delay can be tolerated.

The remainder of the paper is organized as follows. In Section II, the considered system and channel models are presented. The proposed link selection protocol for buffer-aided relaying is introduced and optimized in Section III, and optimal power allocation for source and relay is discussed in Section IV. In Section V, we propose two protocols for buffer-aided relaying with delay constraints. Numerical results are presented in Section VI, and some conclusions are drawn in Section VII.

II. SYSTEM MODEL

We consider a three-node communication system comprising a source $\mathcal{S}$, a half-duplex relay $\mathcal{R}$, and a destination $\mathcal{D}$, cf. Fig. 1. The source can communicate with the destination only through the relay, i.e., there is no direct $\mathcal{S}$-$\mathcal{D}$ link. The source sends packets to the relay, which decodes these packets, possibly stores them in its buffer, and eventually sends them to the destination. Throughout this paper, we assume that the source has always data to transmit.

A. Channel Model

We assume that time is divided into slots of equal lengths. In the ith time slot, the transmit powers of source and relay are denoted by $P_S(i)$ and $P_R(i)$, respectively, and the instantaneous (squared) channel gains of the $\mathcal{S}$-$\mathcal{R}$ and $\mathcal{R}$-$\mathcal{D}$ links are denoted by $h_S(i)$ and $h_R(i)$, respectively. $h_S(i)$ and $h_R(i)$ are modeled as mutually independent, non-negative, stationary, and ergodic random processes with expected values $E\{h_S(i)\} \triangleq \bar{\Omega}_S$ and $E\{h_R(i)\} \triangleq \bar{\Omega}_R$, where $E\{\cdot\}$ denotes expectation. We assume that the channel gains 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.

The instantaneous link signal-to-noise ratios (SNRs) of the $\mathcal{S}$-$\mathcal{R}$ and $\mathcal{R}$-$\mathcal{D}$ channels in the ith time slot are given by $s(i) \triangleq \gamma_S(i)h_S(i)$ and $r(i) \triangleq \gamma_R(i)h_R(i)$, respectively. Here, $\gamma_S(i) = P_S(i)/\sigma_{n_S}^2$ and $\gamma_R(i) = P_R(i)/\sigma_{n_R}^2$ denote the transmit SNRs of the source and the relay, respectively, and $\sigma_{n_S}^2$ and $\sigma_{n_R}^2$ are the variances of the additive white Gaussian noise (AWGN) at the relay and the destination, respectively. The average link SNRs are denoted by $\Omega_S \triangleq E\{s(i)\}$ and $\Omega_R \triangleq E\{r(i)\}$.

B. Link Adaptive Transmission Protocol

In the proposed link adaptive transmission protocol, one of the nodes of the network is responsible for deciding whether the source or the relay should transmit in a given time slot i. This node is referred to as the central node in the following. The central node broadcasts its decision to the other nodes before transmission in time slot i begins. If they are selected for transmission, the source and the relay adapt their transmission rates to the capacity of the respective link and transmit codewords spanning one time slot. We assume that source and relay employ capacity-achieving codes. For both link selection and rate adaptation, the nodes require CSI knowledge as will be detailed in the following.

CSI requirements: The central node requires knowledge of the instantaneous channel gains $h_S(i)$ and $h_R(i)$. In addition, regardless of which node is the central node, if the $\mathcal{S}$-$\mathcal{R}$ link is selected, the source and

the relay require knowledge of $h_S(i)$ for rate adaptation and decoding, respectively. On the other hand, if the $\mathcal{R}$-$\mathcal{D}$ link is selected, the relay and the destination require knowledge of $h_R(i)$ for rate adaptation and decoding, respectively. Nodes can obtain the instantaneous channel gains through estimation based on pilot symbols emitted by other nodes and/or CSI feedback from other nodes. Furthermore, we assume that the central node knows the noise variances $\sigma_{n_S}^2$ and $\sigma_{n_R}^2$, and regardless of which node is the central node, source, relay, and destination know $\sigma_{n_S}^2$, $(\sigma_{n_S}^2, \sigma_{n_R}^2)$, and $\sigma_{n_R}^2$, respectively. Also, if the transmit power is a priori fixed, i.e., $P_S(i) = P_S$ and $P_R(i) = P_R$, $\forall i$, the central node requires knowledge of P_S and P_R, and source, relay, and destination require knowledge of P_S, (P_S, P_R), and P_R, respectively. If power allocation is employed, $P_S(i)$ and $P_R(i)$ are computed by the nodes based on their respective knowledge of the instantaneous channel gains and statistical CSI knowledge (cf. Section IV, (28), (29)).

Which node serves as the central node depends on the network architecture. For example, in the downlink of a cellular network, the source (base station) can serve as the central node as it typically acquires the full CSI of all links anyways and can afford the complexity of performing adaptive link selection and power allocation. On the other hand, if the relay serves as the central node, source and destination have to acquire only the CSI of the $\mathcal{S}$-$\mathcal{R}$ and $\mathcal{R}$-$\mathcal{D}$ links, respectively.

For convenience we normalize the number of bits transmitted in one time slot by the number of symbols per time slot. Thus, throughout the remainder of this paper, the number of bits refers to the number of bits divided by the number of symbols in a codeword. In the following, we discuss the network dynamics when source and relay transmit, respectively.

Source transmits: If the source is selected for transmission in time slot i, it transmits with rate

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

Hence, the relay receives $S_{\mathcal{SR}}(i)$ data bits from the source and appends them to the queue in its buffer. The (normalized) number of bits in the buffer of the relay at the end of time slot i is denoted by $Q(i)$ and given by

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

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

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

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 its buffer and the instantaneous capacity of the $\mathcal{R}$-$\mathcal{D}$ link. The number of bits remaining in the buffer at the end of time slot i is given by

$$Q(i) = Q(i-1) - R_{\mathcal{R}\mathcal{D}}(i), \tag{4}$$

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

Because of the half-duplex constraint, we have $R_{\mathcal{R}\mathcal{D}}(i) = 0$ when the source transmits (and the relay listens), and we have $S_{\mathcal{S}\mathcal{R}}(i) = 0$ when the relay transmits.

C. Throughput

Since we assume the source has always data to transmit, the average (normalized) number of bits that arrive at the destination per time slot is given by

$$\tau = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} R_{\mathcal{R}\mathcal{D}}(i), \tag{5}$$

i.e., τ is the throughput of the considered communication system. The goal of the following sections is the maximization of τ by optimizing the adaptive link selection protocol and the transmit power allocated to source and relay.

D. Conventional Relaying

For comparison purpose, we provide the throughput of two baseline schemes. Thereby, we assume that the transmit powers at the source and the relay are fixed, i.e., $P_S(i) = P_S$, $P_R(i) = P_R$, $\forall i$.

1) Conventional relaying without buffer: The instantaneous throughput of conventional relaying without buffer, where the relay receives a packet in one time slot and transmits it in the next, is given in [5], and the corresponding average throughput is

$$
\begin{aligned}
\tau_{\text{conv},1} &\triangleq \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N/2} \min\{\log_2(1 + s(2i-1)), \log_2(1 + r(2i))\} \\
&= \frac{1}{2} E\{\min\{\log_2(1 + s(i)), \log_2(1 + r(i))\}\},
\end{aligned} \tag{6}
$$

where the ergodicity of $s(i)$ and $r(i)$ was exploited. Note that conventional relaying without buffer introduces a delay of one time slot since the relay has to wait until the entire codeword is received and decoded before sending the codeword to the destination.

2) Conventional relaying with buffer [13]: In conventional relaying with buffer as proposed in [13], the relay receives data from the source in the first $N/2$ (N is even) time slots and sends this cumulative information to the destination in the next $N/2$ slots. The corresponding maximum achievable average throughput is obtained for $N \to \infty$ and given by

$$\tau_{\text{conv},2} \triangleq \lim_{N \to \infty} \frac{1}{N} \min \left\{ \sum_{i=1}^{N/2} \log_2(1 + s(i)), \sum_{i=N/2}^{N} \log_2(1 + r(i)) \right\}$$
$$= \frac{1}{2} \min\{E\{\log_2(1 + s(i))\}, E\{\log_2(1 + r(i))\}\}. \tag{7}$$

Comparing (6) and (7), we observe that $\tau_{\text{conv},2} \geq \tau_{\text{conv},1}$ holds [13]. However, to realize this performance gain, the relay has to be equipped with a buffer of infinite size and an infinite delay is introduced.

3) Rayleigh fading: For the numerical results shown in Section VI, we consider the case where the $\mathcal{S}$-$\mathcal{R}$ and $\mathcal{R}$-$\mathcal{D}$ links are both Rayleigh faded, i.e., the probability density functions (pdfs) of $s(i)$ and $r(i)$ are given by $f_s(s) = e^{-s/\Omega_S}/\Omega_S$ and $f_r(r) = e^{-r/\Omega_R}/\Omega_R$, respectively. In this case, $\tau_{\text{conv},1}$ and $\tau_{\text{conv},2}$ can be obtained in closed form as

$$\tau_{\text{conv},1} = \frac{1}{2 \ln(2)} \exp\left(\frac{\Omega_R + \Omega_S}{\Omega_S \Omega_R}\right) E_1\left(\frac{\Omega_R + \Omega_S}{\Omega_S \Omega_R}\right) \tag{8}$$

and

$$\tau_{\text{conv},2} = \frac{1}{2 \ln(2)} \min \left\{ \exp\left(\frac{1}{\Omega_S}\right) E_1\left(\frac{1}{\Omega_S}\right), \exp\left(\frac{1}{\Omega_R}\right) E_1\left(\frac{1}{\Omega_R}\right) \right\}, \tag{9}$$

respectively, where $E_1(x) = \int_x^{\infty} e^{-t}/t \, dt$, $x > 0$, denotes the exponential integral function.

III. ADAPTIVE LINK SELECTION

To gain some insight, we assume throughout this section constant source and relay powers, i.e., $P_S(i) = P_S$, $P_R(i) = P_R$, $\forall i$, and a buffer of unlimited size at the relay. For this case, we derive the optimal link selection policy and the corresponding throughput. Optimal power allocation and the effect of a limited buffer size will be discussed in Sections IV and V, respectively.

A. Problem Formulation

Let $d_i \in \{0, 1\}$ denote a binary decision variable. We set $d_i = 1$ if the $\mathcal{R}$-$\mathcal{D}$ 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 $\mathcal{S}$-$\mathcal{R}$ 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

$$S_{\mathcal{SR}}(i) = (1 - d_i)S(i) \tag{10}$$

and

$$R_{\mathcal{RD}}(i) = d_i \min\{R(i), Q(i-1)\}, \tag{11}$$

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

$$\tau = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} d_i \min\{R(i), Q(i-1)\}. \tag{12}$$

The considered problem can now be stated as follows: Find the optimal link selection policy, i.e., the optimal sequence d_i, $i \geq 1$, which maximizes throughput τ.

Remark 1: Our problem formulation is quite general in the sense that we have introduced no restrictions concerning the required knowledge regarding the channel and the queue states. In other words, the optimal decision at time j, d_j, potentially depends on $h_S(i)$, $h_R(i)$, and $Q(i)$, $i \geq 1$, and thus requires non-causal channel knowledge. Fortunately, as will be shown in the ensuing section, this is not the case and the optimal policy turns out to be rather simple and easy to implement.

B. Optimal Link Selection Policy

Let us first define the average arrival rate of bits per slot into the queue of the buffer, A, and the average departure rate of bits per slot out of the queue of the buffer, D, as [16]

$$A \triangleq \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} (1 - d_i)S(i) \tag{13}$$

and

$$D \triangleq \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} d_i \min\{R(i), Q(i-1)\}, \tag{14}$$

respectively. We note that the average departure rate of the queue is equal to the throughput. The queue is said to be absorbing if $A > D = \tau$. The following theorem characterizes the optimal link selection policy in terms of the state of the queue in the buffer of the relay.

Theorem 1: A necessary condition for the optimal link selection policy, which maximizes the through-put, is that the queue in the buffer of the relay is at the edge of non-absorbtion, i.e., the queue is non-absorbing but is at the boundary of a non-absorbing and an absorbing queue.

Proof: Please refer to Appendix A. ∎

Exploiting Theorem 1, we can establish a useful condition that the optimal link selection policy has to fulfill and a simplified expression for the throughput. This is the subject of the following theorem.

Theorem 2: Assuming non-negative, ergodic, and stationary random processes $s(i)$ and $r(i)$, for the optimal link selection policy the identity

$$E\{(1 - d_i)S(i)\} = E\{d_i R(i)\} \tag{15}$$

holds and the throughput is then given by

$$\tau = E\{d_i R(i)\}. \tag{16}$$

Proof: Please refer to Appendix B. ∎

Remark 2: According to Theorem 2, for the optimal link selection policy, 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 bits that can be transmitted over the $\mathcal{R}$-$\mathcal{D}$ channel. In particular, as shown in Appendix B, condition (15) automatically ensures that for $N \to \infty$, $\frac{1}{N}\sum_{i=1}^{N} d_i R(i) = \frac{1}{N}\sum_{i=1}^{N} d_i \min\{R(i), Q(i-1)\}$ is valid, i.e., the impact of event $R(i) > Q(i-1)$, $i = 1, \ldots, N$, is negligible.

We are now ready to derive the optimal link selection policy for buffer-aided relaying. According to Theorem 2, the policy that maximizes the throughput τ in (16) can only be found inside the set of policies that produce a queue which satisfies (15), and not outside of this set of policies. Thus, for $N \to \infty$, we formulate the following optimization problem:

$$
\begin{aligned}
\underset{d_i}{\text{Maximize}} : \quad & \tfrac{1}{N}\sum_{i=1}^{N} d_i R(i) \\
\text{Subject to} : \quad & \text{C1} : \tfrac{1}{N}\sum_{i=1}^{N}(1 - d_i)S(i) = \tfrac{1}{N}\sum_{i=1}^{N} d_i R(i) \\
& \text{C2} : \tfrac{1}{N} d_i(1 - d_i) = 0 \quad \forall i
\end{aligned}
\tag{17}
$$

where constraint C1 ensures that the search for the optimal policy is conducted only among those policies that satisfy (15) and C2 ensures that $d_i \in \{0, 1\}$. We note that C1 and C2 do not exclude the case that the relay is chosen for transmission if $R(i) > Q(i-1)$. However, according to Remark 2, C1 ensures that the influence of event $R(i) > Q(i-1)$ is negligible. Therefore, an additional constraint dealing with this event is not required. The solution of problem (17) leads to the following theorem.

Theorem 3: The optimal policy maximizing the throughput of buffer-aided relaying with adaptive link selection is given by

$$d_i = \begin{cases} 1 & \text{if } \mathcal{F}(r(i)) \geq \rho \mathcal{F}(s(i)) \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

where ρ is referred to as the decision threshold and the optimal decision function is given by

$$\mathcal{F}(x) = \log_2(1 + x). \tag{19}$$

The optimal decision threshold ρ, ρ_{opt}, has to satisfy (15).

Proof: Please refer to Appendix C. ∎

Remark 3: Interestingly, we observe from Theorem 3 that the optimal decision, d_i, at time slot i, depends only on the instantaneous SNRs, $s(i)$ and $r(i)$, of that time slot. d_i does not depend on the state of the queue, $Q(i)$, in any time slot nor on the instantaneous SNRs in previous or future time slots. This makes the proposed optimal selection policy highly practical. We note that the optimal decision threshold, ρ_{opt}, depends on the statistical CSI of both involved links as will be established in the next section. The independence of the optimal link selection policy from non-causal instantaneous CSI is caused by the relay being operated at the edge of non-absorbtion, i.e., the relay node is practically fully backlogged. Non-causal knowledge would only help buffer management (i.e., ensuring that there is a sufficient number of bits in the buffer for upcoming time slots), which is not required in the considered regime.

Remark 4: In this paper, we assume that the transmitting nodes have perfect CSI and apply adaptive rate transmission. However, we note that this is not necessary for achieving the maximum throughput in (16). In fact, the proposed adaptive link selection protocol (18) also achieves the maximum throughput in (16) if source and relay transmit long codewords that span (ideally infinitely) many time slots (and consequently many channel states). In this case, both the source and the relay can transmit with constant rate $\tau = E\{(1 - d_i)S(i)\} = E\{d_i R(i)\}$ and rate adaptation is not necessary. The first codeword is transmitted by the source without link adaptation and decoded by the relay. For all subsequent codewords, link adaptation is performed based on (18) and source and relay transmit parts of a long codeword whenever they are selected for transmission. The disadvantage of this approach is that the long codewords inherently introduce (ideally infinitely) long delays and the generalization of this approach to the delay constrained case is difficult. Therefore, in this paper, we consider adaptive rate transmission and assume that one codeword spans only one time slot (and consequently one channel state).

C. Generalization of the Decision Function and Optimal Decision Threshold

The optimal decision function in (19) is intuitively pleasing since it basically means that link selection is performed based on the instantaneous link capacities. Nevertheless, for some fading distributions, the optimal decision function may lead to complicated expressions for the throughput and the optimal decision threshold ρ_{opt}. In such cases, simpler suboptimal decision functions such as $\mathcal{F}(x) = x$ may be preferable as they generally lead to a similar performance as $\mathcal{F}(x) = \log_2(1+x)$ but are analytically more tractable. Thus, in the following, we generalize the decision function, $\mathcal{F}(x)$, to be any non-negative, smooth, and increasing function, i.e., $\mathcal{F}(x+\epsilon) > \mathcal{F}(x)$ for $\epsilon > 0$. We assume that the inverse of $\mathcal{F}(\cdot)$ exists and denote it by $\mathcal{F}^{-1}(\cdot)$.

We note that Theorems 1 and 2 also hold for suboptimal decision functions $\mathcal{F}(x)$. For a given $\mathcal{F}(x)$, the corresponding decision threshold ρ has to satisfy (15). Thus, the optimal decision threshold, ρ_{opt}, for a given (optimal or suboptimal) decision function, $\mathcal{F}(x)$, can be computed based on the following lemma.

Lemma 1: Denote the pdfs of $s(i)$ and $r(i)$ by $f_s(s)$ and $f_r(r)$, respectively. For a given decision function, $\mathcal{F}(x)$, the optimal ρ_{opt} is the solution of

$$\int_0^\infty \left[\int_{G(r)}^\infty \log_2(1+s) f_s(s) ds \right] f_r(r) dr = \int_0^\infty \left[\int_{H(s)}^\infty \log_2(1+r) f_r(r) dr \right] f_s(s) ds, \qquad (20)$$

where the integral limits are given by $G(r) \triangleq \mathcal{F}^{-1}(\mathcal{F}(r)/\rho)$ and $H(s) \triangleq \mathcal{F}^{-1}(\rho \mathcal{F}(s))$.

Proof: The left hand side of (15) is the expectation of variable $(1-d_i) \log_2(1+s(i))$. This variable is nonzero only when $d_i = 0$. From (18) we observe that $d_i = 0$ if $\rho \mathcal{F}(s(i)) > \mathcal{F}(r(i))$, which is equivalent to $s(i) > \mathcal{F}^{-1}(\mathcal{F}(r(i))/\rho)$. Therefore, the domain of integration for calculating the expectation of $(1-d_i) \log_2(1+s(i))$ is $s(i) > \mathcal{F}^{-1}(\mathcal{F}(r(i))/\rho)$ and $r(i) > 0$, which leads to the left hand side of (20). Using a similar approach, the right hand side of (20) is obtained from the right hand side of (15). This concludes the proof. ∎

Remark 5: Eq. (20) reveals that the optimal decision threshold, ρ_{opt}, depends indeed on the statistical properties of both involved links as was already alluded to in Remark 3.

D. Rayleigh Fading

For concreteness, we provide in this subsection expressions for ρ_{opt} and the corresponding maximum throughput τ_{max} for Rayleigh fading links. Thereby, the optimal and a suboptimal decision function $\mathcal{F}(\cdot)$ are considered.

1) $\mathcal{F}(x) = x$: In this case, the limits $G(r)$ and $H(s)$ in (20) are given by $G(r) = r/\rho$ and $H(s) = \rho s$. Thus, after some elementary manipulations, (20) simplifies to

$$\exp\left(\frac{1}{\Omega_S}\right) E_1\left(\frac{1}{\Omega_S}\right) - \exp\left(\frac{1}{\Omega_R}\right) E_1\left(\frac{1}{\Omega_R}\right) - \frac{\Omega_R}{\Omega_R + \rho\Omega_S} \exp\left(\frac{\Omega_R + \rho\Omega_S}{\Omega_S\Omega_R}\right) E_1\left(\frac{\Omega_R + \rho\Omega_S}{\Omega_S\Omega_R}\right)$$
$$+ \frac{\rho\Omega_S}{\Omega_R + \rho\Omega_S} \exp\left(\frac{\Omega_R + \rho\Omega_S}{\rho\Omega_S\Omega_R}\right) E_1\left(\frac{\Omega_R + \rho\Omega_S}{\rho\Omega_S\Omega_R}\right) = 0. \tag{21}$$

The optimal value of ρ for $\mathcal{F}(x) = x$, $\rho_{\mathrm{opt},1}$, can be obtained from (21) via a simple one dimensional search.

The corresponding maximal throughput can be obtained from $\tau_{\mathrm{max},1} = E\{d_i \log_2(1 + r(i))\}$ and is given by

$$\tau_{\mathrm{max},1} = \frac{1}{\ln(2)} \exp\left(\frac{1}{\Omega_R}\right) E_1\left(\frac{1}{\Omega_R}\right) - \frac{1}{\ln(2)} \frac{\rho\Omega_S}{\Omega_R + \rho\Omega_S} \exp\left(\frac{\Omega_R + \rho\Omega_S}{\rho\Omega_S\Omega_R}\right) E_1\left(\frac{\Omega_R + \rho\Omega_S}{\rho\Omega_S\Omega_R}\right), \tag{22}$$

where $\rho = \rho_{\mathrm{opt},1}$.

2) $\mathcal{F}(x) = \log_2(1 + x)$: In this case, we have $G(r) = (1 + r)^{\frac{1}{\rho}} - 1$ and $H(s) = (1 + s)^\rho - 1$ in (20). Thus, after some manipulations, we obtain from (20)

$$\int_0^\infty \left[\exp\left(-\frac{(r+1)^{\frac{1}{\rho}} - 1}{\Omega_S}\right) \ln\left((r+1)^{\frac{1}{\rho}}\right) + e^{\frac{1}{\Omega_S}} E_1\left(\frac{(r+1)^{\frac{1}{\rho}}}{\Omega_S}\right)\right] \times \frac{1}{\Omega_R} \exp\left(-\frac{r}{\Omega_R}\right) dr$$
$$- \int_0^\infty \left[\exp\left(-\frac{(s+1)^\rho - 1}{\Omega_R}\right) \ln\left((s+1)^\rho\right) + e^{\frac{1}{\Omega_R}} E_1\left(\frac{(s+1)^\rho}{\Omega_R}\right)\right] \times \frac{1}{\Omega_S} \exp\left(-\frac{s}{\Omega_S}\right) ds = 0. \tag{23}$$

The optimal ρ, $\rho_{\mathrm{opt},2}$, can be found numerically from (23). The corresponding maximum throughput is obtained as

$$\tau_{\mathrm{max},2} = \frac{1}{\ln(2)} \int_0^\infty \left[\exp\left(-\frac{(s+1)^\rho - 1}{\Omega_R}\right) \times \ln\left((s+1)^\rho\right) + e^{\frac{1}{\Omega_R}} E_1\left(\frac{(s+1)^\rho}{\Omega_R}\right)\right] \frac{1}{\Omega_S} \exp\left(-\frac{s}{\Omega_S}\right) ds, \tag{24}$$

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

3) Special case ($\Omega_S = \Omega_R$): For the special case $\Omega_S = \Omega_R = \Omega$, we obtain from (21) and (23) $\rho_{\mathrm{opt},1} = \rho_{\mathrm{opt},2} = 1$, and the corresponding maximal throughput is

$$\tau_{\max} = \tau_{\mathrm{max},1} = \tau_{\mathrm{max},2} = \frac{1}{\ln(2)} \exp\left(\frac{1}{\Omega}\right) E_1\left(\frac{1}{\Omega}\right) - \frac{1}{2\ln(2)} \exp\left(\frac{2}{\Omega}\right) E_1\left(\frac{2}{\Omega}\right). \tag{25}$$

Comparing this throughput with the throughput achievable with conventional buffer-aided relaying without adaptive link selection, cf. (9), the gain of adaptive link selection can be characterized by

$$\tau_{\max}/\tau_{\mathrm{conv},2} = 2 - \exp\left(\frac{2}{\Omega}\right) E_1\left(\frac{2}{\Omega}\right) / \left[\exp\left(\frac{1}{\Omega}\right) E_1\left(\frac{1}{\Omega}\right)\right] \geq 1, \tag{26}$$

where the ratio $\tau_{\max}/\tau_{\mathrm{conv},2}$ monotonically increases from 1 to 1.5 as Ω decreases from ∞ to zero. We note that the results in Section VI reveal that the gain of adaptive link selection is minimum for $\Omega_S = \Omega_R$, cf. Fig. 2.

Remark 6: From (18) it is easy to see that for $\rho = 1$ the decision functions $\mathcal{F}(x) = \log_2(1 + x)$ and $\mathcal{F}(x) = x$ are equivalent in the sense that they lead to the same decisions. Hence, both decision functions lead to identical throughputs $\tau_{\max} = \tau_{\max,1} = \tau_{\max,2}$. However, for $\Omega_S \neq \Omega_R$, $\rho_{\mathrm{opt},1} \neq \rho_{\mathrm{opt},2} \neq 1$ holds, and the decision functions are no longer equivalent and $\tau_{\max,2} > \tau_{\max,1}$ holds.

IV. POWER ALLOCATION

So far, we have assumed that the source and relay transmit powers are fixed. In this section, we jointly optimize the power allocation and link selection policies for buffer-aided relaying.

A. Problem Formulation and Optimal Power Allocation

Our goal is to jointly optimize the link selection variable d_i and the powers $P_S(i)$ and $P_R(i)$ in each time slot i such that the throughput is maximized. For convenience, we optimize in the following the transmit SNRs without fading $\gamma_S(i)$ and $\gamma_R(i)$, which may be viewed as normalized powers, instead of the powers $P_S(i) = \gamma_S(i)\sigma_{n_S}^2$ and $P_R(i) = \gamma_R(i)\sigma_{n_R}^2$ themselves. For a fair comparison, we limit the average (normalized) power consumed by the source and the relay to Γ. This leads for $N \to \infty$ to the following optimization problem:

$$
\begin{aligned}
\underset{\gamma_S(i) \geq 0, \gamma_R(i) \geq 0, d_i}{\text{Maximize}:} \quad & \frac{1}{N}\sum_{i=1}^{N} d_i \log_2(1 + \gamma_R(i)h_R(i)) \\
\text{Subject to}: \quad & \text{C1}: \frac{1}{N}\sum_{i=1}^{N}(1 - d_i)\log_2(1 + \gamma_S(i)h_S(i)) = \frac{1}{N}\sum_{i=1}^{N} d_i \log_2(1 + \gamma_R(i)h_R(i)) \\
& \text{C2}: \frac{1}{N}d_i(1 - d_i) = 0 \\
& \text{C3}: \frac{1}{N}\sum_{i=1}^{N}(1 - d_i)\gamma_S(i) + \frac{1}{N}\sum_{i=1}^{N} d_i\gamma_R(i) \leq \Gamma
\end{aligned}
\tag{27}
$$

where constraints C1 and C2 are identical to the constraints in (17) and C3 is the joint source-relay power constraint. The solution of Problem (27) is summarized in the following theorem.

Theorem 4: The optimal (normalized) powers $\gamma_S(i)$ and $\gamma_R(i)$ and decision variable d_i maximizing the throughput of buffer-aided relaying with adaptive link selection while satisfying an average source-relay

power constraint are given by

$$
\gamma_S(i) = \begin{cases} \rho/\lambda - 1/h_S(i) & \text{if } h_S(i) > \lambda/\rho \\ 0 & \text{otherwise} \end{cases} \tag{28}
$$

$$
\gamma_R(i) = \begin{cases} 1/\lambda - 1/h_R(i) & \text{if } h_R(i) > \lambda \\ 0 & \text{otherwise} \end{cases} \tag{29}
$$

$$
d_i = \begin{cases} 1 & \text{if } \left[\ln\left(\frac{h_R(i)}{\lambda}\right) + \frac{\lambda}{h_R(i)} - 1 > \rho \ln\left(\frac{\rho}{\lambda} h_S(i)\right) + \frac{\lambda}{h_S(i)} - \rho \text{ AND } h_R(i) > \lambda \text{ AND } h_S(i) > \frac{\lambda}{\rho} \right] \\ \qquad \text{OR } \left[h_R(i) > \lambda \text{ AND } h_S(i) \le \frac{\lambda}{\rho} \right] \\ 0 & \text{otherwise} \end{cases} \tag{30}
$$

where the optimal ρ, ρ_{opt}, and the optimal λ, λ_{opt}, have to satisfy C1 and C3 in (27) for $N \to \infty$ with equality.

Proof: Please refer to Appendix D. ∎

Remark 7: Similar to the case without power allocation discussed in Section III, it is also possible for the case with power allocation to simplify the link selection policy in (30). For example, a simple suboptimal link selection policy which only depends on the instantaneous channel gains $h_S(i)$ and $h_R(i)$ but is independent of the transmit powers $\gamma_S(i)$ and $\gamma_R(i)$ may be adopted at the expense of some loss in performance. However, in this paper, we do not pursue suboptimal link selection policies for the case of power allocation because of space limitation.

B. Finding the Optimal λ and ρ

The following lemma establishes two equations from which the optimal λ and ρ can be found.

Lemma 2: Denote the pdfs of $h_S(i)$ and $h_R(i)$ by $f_{h_S}(h_S)$ and $f_{h_R}(h_R)$, respectively. Let the transmit powers of the source and the relay in time slot i be given by (28) and (29), respectively, and the link selection variable d_i by (30). Then, the optimal ρ, ρ_{opt}, and the optimal λ, λ_{opt}, maximizing the throughput of buffer-aided relaying with adaptive link selection and power allocation have to fulfill the following two equations

$$
\int_0^\lambda \left[\int_{\lambda/\rho}^\infty \log_2\left(\frac{\rho h_S}{\lambda}\right) f_{h_S}(h_S) dh_S \right] f_{h_R}(h_R) dh_R + \int_\lambda^\infty \left[\int_{L_1}^\infty \left(\frac{\rho h_S}{\lambda}\right) f_{h_S}(h_S) dh_S \right] f_{h_R}(h_R) dh_R
$$
$$
= \int_0^{\lambda/\rho} \left[\int_\lambda^\infty \log_2\left(\frac{h_R}{\lambda}\right) f_{h_R}(h_R) dh_R \right] f_{h_S}(h_S) dh_S + \int_{\lambda/\rho}^\infty \left[\int_{L_2}^\infty \log_2\left(\frac{h_R}{\lambda}\right) f_{h_R}(h_R) dh_R \right] f_{h_S}(h_S) dh_S , \tag{31}
$$

$$\int_0^\lambda \left[\int_{\lambda/\rho}^\infty \left(\frac{\rho}{\lambda} - \frac{1}{h_S} \right) f_{h_S}(h_S) dh_S \right] f_{h_R}(h_R) dh_R + \int_\lambda^\infty \left[\int_{L_1}^\infty \left(\frac{\rho}{\lambda} - \frac{1}{h_S} \right) f_{h_S}(h_S) dh_S \right] f_{h_R}(h_R) dh_R$$

$$+ \int_0^{\lambda/\rho} \left[\int_\lambda^\infty \left(\frac{1}{\lambda} - \frac{1}{h_R} \right) f_{h_R}(h_R) dh_R \right] f_{h_S}(h_S) dh_S + \int_{\lambda/\rho}^\infty \left[\int_{L_2}^\infty \left(\frac{1}{\lambda} - \frac{1}{h_R} \right) f_{h_R}(h_R) dh_R \right] f_{h_S}(h_S) dh_S$$

$$= \Gamma \tag{32}$$

where

$$L_1 = -\frac{\lambda}{\rho W(-e^{(h_R-\lambda)/(\rho h_R)-1}(\lambda/h_R)^{1/\rho})}, \quad L_2 = -\frac{\lambda}{W(-e^{\rho-1-\lambda/h_S}(\lambda/(\rho h_S))^\rho)}. \tag{33}$$

Here, $W(\cdot)$ is the Lambert W-function [17], which is available as built-in function in software packages such as Mathematica. The maximum throughput is given by the left (and right) hand side of (31).

Proof: Please refer to Appendix E. ∎

To find λ_{opt} and ρ_{opt} meeting (31) and (32), a two dimensional search over all $\lambda > 0$ and $\rho > 0$ has to be conducted or built-in root-finding functions of commercially available software such as Mathematica can be used. The optimal λ and ρ can be found offline since (31) and (32) only depend on the statistical properties of the $\mathcal{S}$-$\mathcal{R}$ and the $\mathcal{R}$-$\mathcal{D}$ links. Since these statistical properties change on a much slower time scale than the instantaneous channel gains, λ_{opt} and ρ_{opt} can be updated with a low rate.

C. Rayleigh Fading

For the special case of Rayleigh fading with $f_{h_S}(h_S) = e^{-h_S/\bar\Omega_S}/\bar\Omega_S$ and $f_{h_R}(h_R) = e^{-h_R/\bar\Omega_R}/\bar\Omega_R$, (31) can be simplified to

$$\frac{1}{\ln(2)} \left[\left(1 - e^{-\lambda/\bar\Omega_R} \right) E_1 \left(\frac{\lambda}{\rho\bar\Omega_S} \right) + \int_\lambda^\infty \left\{ e^{-L_1/\bar\Omega_S} \ln \left(\frac{\rho L_1}{\lambda} \right) + E_1 \left(\frac{L_1}{\bar\Omega_S} \right) \right\} \frac{e^{-h_R/\bar\Omega_R}}{\bar\Omega_R} dh_R \right]$$

$$= \frac{1}{\ln(2)} \left[\left(1 - e^{-\lambda/(\rho\bar\Omega_S)} \right) E_1 \left(\frac{\lambda}{\bar\Omega_R} \right) + \int_{\lambda/\rho}^\infty \left\{ e^{-L_2/\bar\Omega_R} \ln \left(\frac{L_2}{\lambda} \right) + E_1 \left(\frac{L_2}{\bar\Omega_R} \right) \right\} \frac{e^{-h_S/\bar\Omega_S}}{\bar\Omega_S} dh_S \right], \tag{34}$$

and (32) can be simplified to

$$\left(1 - e^{-\lambda/\bar\Omega_R} \right) \left(\frac{\rho}{\lambda} e^{-\lambda/(\rho\bar\Omega_S)} - \frac{E_1 \left(\frac{\lambda}{\rho\bar\Omega_S} \right)}{\bar\Omega_S} \right) + \int_\lambda^\infty \left\{ \frac{\rho}{\lambda} e^{-L_1/\bar\Omega_S} - \frac{E_1 \left(\frac{L_1}{\bar\Omega_S} \right)}{\bar\Omega_S} \right\} \frac{e^{-h_R/\bar\Omega_R}}{\bar\Omega_R} dh_R$$

$$+ \left(1 - e^{-\lambda/(\rho\bar\Omega_S)} \right) \left(\frac{1}{\lambda} e^{-\lambda/\bar\Omega_R} - \frac{E_1 \left(\frac{\lambda}{\bar\Omega_R} \right)}{\bar\Omega_R} \right) + \int_{\lambda/\rho}^\infty \left\{ \frac{1}{\lambda} e^{-L_2/\bar\Omega_R} - \frac{E_1 \left(\frac{L_2}{\bar\Omega_R} \right)}{\bar\Omega_R} \right\} \frac{e^{-h_S/\bar\Omega_S}}{\bar\Omega_S} dh_S = \Gamma, \tag{35}$$

where L_1 and L_2 are given in (33) and the maximum throughput is given by the left (and right) hand side of equation (34).

V. DELAY LIMITED TRANSMISSION

So far, we have assumed that there is no delay constraint and that the size of the buffer at the relay is infinite. In practice, there is usually some constraint on the delay and on the buffer size. In this section, we investigate how these constraints affect the performance of the proposed relaying protocol. For simplicity, we assume fixed transmit powers, i.e., $P_S(i) = P_S$, $P_R(i) = P_R$, $\forall i$, and consequently, policy (18) is used for link selection.

Since we assume that the source is backlogged and has always information to transmit, for the considered three-node network, the transmission delay is caused only by the buffer at the relay. Let $T(i)$ denote the delay of a bit that is transmitted by the source in time slot i and received at the destination in time slot $i + T(i)$, i.e., the considered bit is stored for $T(i)$ time slots in the buffer. Then, according to Little's law [18] the average delay $E\{T(i)\}$ (i.e., the average time that a bit is stored in the buffer) is given by

$$E\{T(i)\} = E\{Q(i)\}/A, \tag{36}$$

where $E\{Q(i)\}$ is the average queue length at the buffer and A is the average arrival rate into the queue. $E\{Q(i)\}$ is given in bits and A is given in bits/slot. Thus, the average delay $E\{T(i)\}$ is given in time slots. From (36), we observe that the delay can be controlled via the arrival rate and the queue size. In the following, we will present two different approaches to adjust the arrival rate and the queue size. The first approach is to "starve" the buffer, i.e., we intentionally limit the arrival rate by choosing $\rho < \rho_{\mathrm{opt}}$. The second approach is to limit the buffer size by forcing the relay to transmit if the buffer gets full.

A. Satisfying an Average Delay Constraint by "Starving" the Buffer

Starving the buffer is a common approach for limiting average delays in queueing systems [19], [20]. In our case, we can decrease the average arrival rate by selecting $\rho < \rho_{\mathrm{opt}}$ which leads to $E\{(1 - d_i)S(i)\} < E\{d_i R(i)\}$. In the following theorem, we establish an upper bound for the resulting average delay.

Theorem 5: Let $\rho < \rho_{\mathrm{opt}}$ in (18) such that $\xi = E\{(1 - d_i)S(i)\}/E\{d_i R(i)\} < 1$. In this case, assuming slot-by-slot uncorrelated fading the average delay in slots is bounded by

$$E\{T(i)\} \leq \frac{1}{2} \frac{1}{E\{(1 - d_i)S(i)\}} \frac{E\{(1 - d_i)S^2(i)\} + \xi(2 - \xi)E\{d_i R^2(i)\}}{E\{d_i R(i)\} - E\{(1 - d_i)S(i)\}} \tag{37}$$

and the throughput is given by $\tau = E\{(1 - d_i)S(i)\}$.

Proof: Please refer to Appendix F. ∎

Remark 8: Exploiting (37) the required value of $\rho < \rho_{\text{opt}}$ to ensure a desired average delay can be found. For example, assuming $\mathcal{F}(x) = x$ and Rayleigh distributed $\mathcal{S}$-$\mathcal{R}$ and $\mathcal{R}$-$\mathcal{D}$ channel gains the expected values required in (37) can be obtained as

$$E\{(1-d_i)S(i)\} = \frac{1}{\ln(2)}\left[\exp\left(\frac{1}{\Omega_S}\right)E_1\left(\frac{1}{\Omega_S}\right) - \frac{\Omega_R}{\Omega_R + \rho\Omega_S}\exp\left(\frac{\Omega_R + \rho\Omega_S}{\Omega_S\Omega_R}\right)E_1\left(\frac{\Omega_R + \rho\Omega_S}{\Omega_S\Omega_R}\right)\right] \tag{38}$$

$$E\{d_iR(i)\} = \frac{1}{\ln(2)}\left[\exp\left(\frac{1}{\Omega_R}\right)E_1\left(\frac{1}{\Omega_R}\right) - \frac{\rho\Omega_S}{\Omega_R + \rho\Omega_S}\exp\left(\frac{\Omega_R + \rho\Omega_S}{\rho\Omega_S\Omega_R}\right)E_1\left(\frac{\Omega_R + \rho\Omega_S}{\rho\Omega_S\Omega_R}\right)\right] \tag{39}$$

$$E\{(1-d_i)S^2(i)\} = \int_0^\infty\left[\int_{r/\rho}^\infty(\log_2(1+s))^2\frac{e^{-s/\Omega_S}}{\Omega_S}ds\right]\frac{e^{-r/\Omega_R}}{\Omega_R}dr \tag{40}$$

$$E\{d_iR^2(i)\} = \int_0^\infty\left[\int_{s\rho}^\infty(\log_2(1+r))^2\frac{e^{-r/\Omega_R}}{\Omega_R}dr\right]\frac{e^{-s/\Omega_S}}{\Omega_S}ds. \tag{41}$$

Inserting now (38)-(41) into (37), the value of ρ guaranteeing a certain average delay can be found by slowly increasing ρ from zero until the right hand side of (37) equals the desired average delay $E\{T(i)\}$.

If the buffer size is limited, there is a non-zero probability that the bits arriving into the buffer have to be dropped because the buffer is full, even if the buffer is starved. However, the probability of this event happening can be minimized by properly choosing the buffer size compared to the desired average delay. This issue is addressed in the following lemma.

Lemma 3: Denote the maximum queue size by Q_{max}. Then, we can bound the probability that the queue in the buffer exceeds Q_{max} as

$$\Pr\{Q(i) > Q_{\text{max}}\} \leq E\{Q(i)\}/Q_{\text{max}}. \tag{42}$$

Proof: The proof follows directly from Markov's inequality. ∎

We can guarantee any prescribed probability of dropped bits, $\Pr\{Q(i) > Q_{\text{max}}\}$, by selecting appropriate values for ρ and Q_{max} based on (42) and (64) given in Appendix F. The resulting throughput is given by

$$\begin{aligned}\tau &= E\{(1-d_i)S(i)|Q(i) < Q_{\text{max}}\}\Pr\{Q(i) < Q_{\text{max}}\} = E\{(1-d_i)S(i)\}\Pr\{Q(i) < Q_{\text{max}}\} \\ &= E\{(1-d_i)S(i)\}(1 - \Pr\{Q(i) > Q_{\text{max}}\}),\end{aligned} \tag{43}$$

for which a lower bound can be found by combining (42), (43), and (64).

B. Satisfying the Delay Constraint by Limiting the Queue Size

For the scheme proposed in the previous subsection dropped bits are unavoidable. In this subsection, we propose an alternative approach which allows us to avoid dropped bits. Let the buffer size again be limited to Q_{max} bits. The proposed scheme employs the following link selection protocol:

1) If $Q_{\max} - Q(i-1) > S(i)$, select d_i based on (18).

2) Otherwise, set $d_i = 1$.

Hence, if there is enough room in the buffer to accommodate the bits possibly sent from the source to the relay, the link selection protocol introduced in Section III is employed. On the other hand, if there exists the possibility of a buffer overflow, the relay transmits to reduce the amount of data in the buffer.

Remark 9: Although conceptually simple, a theoretical analysis of the throughput of the queue size limiting protocol is difficult. In contrast to the buffer starving protocol discussed in Section V-A, for the queue size limiting protocol, the average arrival rate A, depends on the frequency with which the buffer has to be emptied due to a full queue. The frequency of these events depends in turn on the average arrival rate. This mutual dependence of average arrival rate and emptying the buffer makes a meaningful theoretical analysis difficult. Thus, we will resort to simulations to evaluate the performance of the queue size limiting protocol in Section VI.

Remark 10: We note that both proposed protocols for the delay constrained case are heuristic in nature. The search for other protocols with possibly superior performance is an interesting topic for future work. The proposed protocols for the delay constrained and the delay unconstrained case can serve as benchmark and performance upper bound for these new protocols, respectively.

VI. NUMERICAL AND SIMULATION RESULTS

In this section, we evaluate the performance of buffer-aided DF relaying with adaptive link selection and compare it with that of conventional relaying. Throughout this section, we assume Rayleigh fading. All results shown in this section have been confirmed by computer simulations. However, the simulations are not shown in all instances for clarity of presentation.

A. Delay Unconstrained Transmission

First, we assume that there are no delay constraints and investigate the achievable throughputs with and without power allocation.

1) Throughput of Buffer-Aided Relaying with Adaptive Link Selection: We first consider the case of fixed transmit powers. In Fig. 2, we show the ratio of the optimal throughput of buffer-aided relaying with adaptive link selection, $\tau_{\max}$, and the throughput of conventional relaying with a buffer, $\tau_{\mathrm{conv},2}$, given in (9), as a function of Ω_R/Ω_S for several different values of Ω_S. The corresponding optimal decision thresholds, ρ_{opt}, for buffer-aided relaying with adaptive link selection are shown in Fig. 3. For buffer-aided relaying

with adaptive link selection, we considered the decision functions $\mathcal{F}(x) = x$ and $\mathcal{F}(x) = \log_2(1 + x)$ and calculated the corresponding throughputs based on (22) and (24), respectively. The optimal decision thresholds were obtained from (21) and (23), respectively. Clearly, buffer-aided relaying with adaptive link selection leads to substantial throughput gains compared to conventional relaying. Both considered decision functions lead to very similar performances, although at very high ratios Ω_R/Ω_S the optimal decision function $\mathcal{F}(x) = \log_2(1 + x)$ yields a small throughput gain. The ratio $\tau_{\max}/\tau_{\mathrm{conv},2}$ approaches two as $\Omega_R/\Omega_S \to 0$ and $\Omega_R/\Omega_S \to \infty$. For $\Omega_R/\Omega_S \to 0$, the source-relay link is selected very rarely for transmission (as $\rho_{\mathrm{opt}} \to 0$) 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 $\Omega_R/\Omega_S \to \infty$, it is the relay-destination channel that is used very rarely as $\rho_{\mathrm{opt}} \to \infty$ and the source can transmit almost all the time, which results in twice the throughput compared to conventional relaying.

2) Throughput with Power Allocation: In Figs. 4 and 5, we investigate the gains achievable with power allocation (PA) for a system with $\bar{\Omega}_S = 0.1$ and $\bar{\Omega}_R = 1.9$. Thereby, we compare the performances of buffer-aided relaying with adaptive link selection and conventional relaying with and without a buffer. For buffer-aided relaying with adaptive link selection and power allocation the throughput, power allocation, and link selection policy were obtained as described in Theorem 4 and Lemma 2 in Section IV. For conventional relaying with buffer, a similar optimal power allocation scheme as for buffer-aided relaying with adaptive link selection was adopted with variable powers $\gamma_j(i) = \max\{0, 1/\alpha - 1/h_j(i)\}$, $j \in \{S, R\}$, where α is chosen such that $E\{\gamma_j(i)\} = \Gamma$. For comparison, in Figs. 4 and 5, we also show the performance of both considered relaying schemes without power allocation, i.e., we set $\gamma_S(i) = \gamma_R(i) = \Gamma$. Furthermore, to highlight the gain compared to conventional relaying without buffer and without power allocation, in Fig. 4, we normalized the throughput with respect to $\tau_{\mathrm{conv},1}$ as given by (8). Figs. 4 and 5 show that optimal power allocation can improve performance of both buffer-aided relaying with adaptive link selection and conventional relaying. For example, for $\Gamma = 0$ dB buffer-aided relaying with adaptive link selection and power allocation leads to a throughput gain of 95 % compared to conventional relaying with buffer and power allocation. Nevertheless, the gain achievable by adaptive link selection is more significant than the gain from power allocation. For example, at $\Gamma = 20$ dB, adaptive link selection yields a throughput gain of 1 bit/slot compared to conventional relaying with buffer.

B. Delay Constrained Transmission

We now turn our attention to delay limited transmission and investigate the performance of the two proposed protocols for this case. In the following we assume uncorrelated fading. Furthermore, we assume fixed transmit powers for the source and relay, and adopt the suboptimal decision function $\mathcal{F}(x) = x$.

1) Starving the Buffer: In Fig. 6, we show the ratio of the throughput of buffer-aided relaying with starved buffer, $\tau_{\max}$, and the throughput of conventional relaying without buffer, $\tau_{\mathrm{conv},1}$, as given in (8), as a function of the upper bound on the average delay. The corresponding decision thresholds, ρ, are shown in Fig. 7. For the theoretical results shown in Figs. 6 and 7 the throughput and the upper bound on the average delay for buffer-aided relaying with starved buffer were obtained based on Theorem 5 and Remark 8. A comparison of the theoretical results with the simulation results also shown in Fig. 6 reveals that the derived upper bound on the delay is tight, especially for large delays. On the other hand, for very small delays, buffer-aided relaying with starved buffer becomes inefficient since the starving of the buffer decreases the average queue size and increases the probability that the relay is selected for transmission when $R(i) > Q(i-1)$. In fact, Fig. 6 shows that for very small delays buffer-aided relaying with starved buffer may be even outperformed by conventional relaying without buffer. As expected, the throughput of buffer-aided relaying increases with increasing tolerable delay and ρ approaches the optimal value for the delay constrained case, ρ_{opt} (computed from (21)), for large delays. The required delay to achieve a throughput gain compared to conventional relaying increases with increasing Ω_S-Ω_R ratio since for large Ω_S/Ω_R and small tolerable average delays, the arrival rate into the buffer has to be severely limited (i.e, ρ has to be chosen very small) which has a negative impact on the throughput which, by the conservation of flow, is equal to the arrival rate.

In Fig. 8, we show the probability of a dropped bit as a function of the buffer size $Q_{\max}$ for three different average delays and $\Omega_S = \Omega_R = 1$. These results were obtained via simulations since the bound obtained in (42) is relatively loose due to the looseness of Markov's inequality. Fig. 8 shows that the probability of dropping a bit rapidly decreases with increasing buffer size and decreasing average delay.

2) Limiting the Queue Size: In Fig. 9, we show the throughput achieved by limiting the queue size, τ_{limit}, and the throughput achieved by starving the buffer, τ_{starve}, normalized by the throughput of conventional relaying without buffer, $\tau_{\mathrm{conv},1}$, as given in (8), for symmetric and asymmetric link qualities. For comparison, Fig. 9 also contains the throughput of delay unconstrained buffer-aided relaying with adaptive link selection, which constitutes an upper bound for the throughput in the delay constrained

case, and conventional relaying with buffer as proposed in [13]. For conventional relaying with buffer, the relay drops information bits if the achievable rate of the $\mathcal{S}$-$\mathcal{R}$ link in the first $N/2$ time slots exceeds the achievable rate of the $\mathcal{R}$-$\mathcal{D}$ link in the second $N/2$ time slots. All results shown in Fig. 9 have been obtained by simulations. Fig. 9 reveals that the performance of both delay constrained buffer-aided relaying protocols is comparable for large delays and approaches that of the delay unconstrained protocol. For small delays, limiting the buffer size yields a higher throughput than starving the buffer. However, both proposed protocols may be outperformed by conventional relaying with and without buffer for very small delays as for the proposed simple protocols, the relay may be selected for transmission even if $R(i) > Q(i-1)$. While this event has negligible effect for delay unconstrained transmission since the optimal link selection policy ensures that the queue is sufficiently long such that $R(i) > Q(i-1)$ is avoided (cf. Remark 2), this is no longer true for the delay constrained case. Therefore, for the delay constrained case, more sophisticated protocols should be developed that take into account that $R(i) > Q(i-1)$ may occur.

VII. CONCLUSIONS

In this paper, we proposed a novel adaptive link selection protocol for relays with buffers. In contrast to conventional relaying, where the source and the relay transmit according to a pre-defined schedule regardless of the channel state, in the proposed scheme, always the node with the stronger link is selected for transmission. For delay unconstrained transmission, we derived the optimal link selection policy for the cases of fixed and variable source and relay transmit powers. Remarkably, in both cases, the optimal policy for a given time slot only depends on the instantaneous CSI of that time slot and the statistical CSI of the involved links. This makes the optimal policies attractive for implementation. For delay constrained transmission, we proposed two different methods to control the delay introduced by the buffer at the relay. Furthermore, for the case when the buffer is starved, we derived upper bounds on the average delay and the number of dropped bits for limited buffer size. Our analytical and simulation results showed that buffer-aided relaying with adaptive link selection is a promising approach to significantly increase the throughput compared to conventional relay-assisted transmission. Interesting extensions of the presented work include using the considered simple three-node network as a building block for larger networks, studying the impact of imperfect CSI, and deriving the outage probability for fixed rate transmission.

APPENDIX

A. Proof of Theorem 1

We first note that, because of the law of the conservation of flow, $A \geq \tau$ is always valid and equality holds if and only if the queue is non-absorbing. Assume first we have a link selection policy with average arrival rate A and throughput τ with $A > \tau$, i.e., the queue is absorbing. For this policy, we denote the set of indices with $d_i = 1$ by $\bar{I}$ and the set of indices with $d_i = 0$ by I, i.e., for $N \to \infty$ we have

$$A = \frac{1}{N} \sum_{i \in I} (1 - d_i) S(i) > \tau = \frac{1}{N} \sum_{i \in \bar{I}} d_i \min\{R(i), Q(i-1)\}. \tag{44}$$

From (44) we observe that the considered protocol cannot be optimal as it can be improved by moving some of the indices i in I to $\bar{I}$ which leads to an increase of τ at the expense of a decrease of A. However, once the point $A = \tau$ is reached, moving more indices i from I to $\bar{I}$ will decrease both A and τ because of the conservation of flow. Thus, a necessary condition for the optimal policy is that the queue is non-absorbing but the queue is at the edge of non-absorbtion, i.e., the queue is at the boundary of a non-absorbing and an absorbing queue. This completes the proof.

B. Proof of Theorem 2

We denote the sets of indices i for which $d_i = 1$ and $d_i = 0$ holds by $\bar{I}$ and I, respectively. ϵ denotes a subset of $\bar{I}$ and $|\cdot|$ is the cardinality of a set. Throughout the remainder of this proof $N \to \infty$ is assumed.

If the queue in the buffer of the relay is absorbing, $A > \tau$ holds and on average the number of bits arriving at the queue exceed the number of bits leaving the queue. Thus, $R(i) \leq Q(i-1)$ holds almost always and as a result the throughput can be written as

$$\tau = \frac{1}{N} \sum_{i \in \bar{I}} \min\{R(i), Q(i-1)\} = \frac{1}{N} \sum_{i \in \bar{I}} R(i). \tag{45}$$

Now, we assume that the queue is at the edge of non-absorption. That is $A = \tau$ holds but moving a small fraction ϵ, where $|\epsilon|/N \to 0$, of indices from $\bar{I}$ to I will make the queue an absorbing queue with $A > \tau$. For this case, we wish to determine whether or not

$$\frac{1}{N} \sum_{i \in \bar{I}} R(i) > \tau = \frac{1}{N} \sum_{i \in \bar{I}} \min\{R(i), Q(i-1)\} = A = \frac{1}{N} \sum_{i \in I} S(i) \tag{46}$$

holds. To test this, we move a small fraction ϵ, where $|\epsilon|/N \to 0$, of indices from $\bar{I}$ to I, thus making the queue an absorbing queue. As a result, (45) holds, and (46) becomes

$$\frac{1}{N} \sum_{i \in \bar{I} \setminus \epsilon} R(i) = \tau = \frac{1}{N} \sum_{i \in \bar{I} \setminus \epsilon} \min\{R(i), Q(i-1)\} < A = \frac{1}{N} \sum_{i \in I \cup \epsilon} S(i). \tag{47}$$

From the above we conclude that if (45) holds, then based on (46) and (47), for $|\epsilon|/N \to 0$, we must have

$$\frac{1}{N} \sum_{i \in \bar{I}} R(i) > \frac{1}{N} \sum_{i \in I} S(i) \tag{48}$$

and

$$\frac{1}{N} \sum_{i \in \bar{I} \setminus \epsilon} R(i) < \frac{1}{N} \sum_{i \in I \cup \epsilon} S(i). \tag{49}$$

However, for (48) and (49) to jointly hold, we require that the particular considered moving of indices from $\bar{I}$ to I has caused a discontinuity in $\frac{1}{N} \sum_{i \in \bar{I}} R(i)$ or/and a discontinuity in $\frac{1}{N} \sum_{i \in I} S(i)$ as $|\epsilon|/N \to 0$ is assumed. Since the capacities of the $\mathcal{S}$-$\mathcal{R}$ and $\mathcal{R}$-$\mathcal{D}$ links are such that $\lim_{N \to \infty} \sum_{i \in \epsilon} S(i)/N \to 0$ and $\lim_{N \to \infty} \sum_{i \in \epsilon} R(i)/N \to 0$, such discontinuities are not possible. Therefore, at the edge of non-absorption (46) is not true and we must have instead

$$\frac{1}{N} \sum_{i \in \bar{I}} R(i) = \tau = \frac{1}{N} \sum_{i \in \bar{I}} \min\{R(i), Q(i-1)\} = A = \frac{1}{N} \sum_{i \in I} S(i) \tag{50}$$

Using the the ergodicity of $s(i)$ and $r(i)$, (50) can be expressed as (15), and the throughput can be written as (16). This concludes the proof.

C. Proof of Theorem 3

The Lagrangian for Problem (17) is given by

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^{N} d_i R(i) - \mu \frac{1}{N} \sum_{i=1}^{N} [d_i R(i) - (1 - d_i) S(i)] - \frac{1}{N} \sum_{i=1}^{N} \beta_i d_i (1 - d_i), \tag{51}$$

where μ and β_i are Lagrange multipliers. Differentiating $\mathcal{L}$ with respect to d_i and setting the result to zero leads to

$$d_i = \frac{(-1 + \mu) R(i) + \mu S(i) + \beta_i}{2 \beta_i}. \tag{52}$$

For $d_i(1 - d_i) = 0$ to hold, we need either $d_i = 0$ or $d_i = 1$, which leads to two possible values for β_i:

$$d_i = 0 \Rightarrow \beta_{i,1} = (1 - \mu) R(i) - \mu S(i) \tag{53}$$

$$d_i = 1 \Rightarrow \beta_{i,2} = -\beta_{i,1} \tag{54}$$

For the maximum of $\mathcal{L}$ in (51), $\beta_i < 0$, $\forall i$, has to hold. Furthermore, $0 < \mu < 1$ has to hold since for $\mu \leq 0$ and $\mu \geq 1$ we have always $d_i = 1$ and $d_i = 0$, respectively, irrespective of the values of $R(i)$ and

$S(i)$. Hence, we have

$$
d_i = \begin{cases} 1 & \text{if } (1-\mu)R(i) - \mu S(i) \geq 0 \\ 0 & \text{if } (1-\mu)R(i) - \mu S(i) < 0 \end{cases} \tag{55}
$$

which is identical to (18) with $\mathcal{F}(x) = \log_2(1+x)$ if we set $\rho = \mu/(1-\mu)$. μ or equivalently ρ are chosen such that constraint C1 of Problem (17) is met. This completes the proof.

D. Proof of Theorem 4

To solve Problem (27), we form the Lagrangian

$$
\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N} d_i \log_2(1 + \gamma_R(i)h_R(i)) - \mu\frac{1}{N}\sum_{i=1}^{N}\left[d_i \log_2(1 + \gamma_R(i)h_R(i)) - (1 - d_i)\log_2(1 + \gamma_S(i)h_S(i)) \right]
$$
$$
- \nu\frac{1}{N}\sum_{i=1}^{N}\left[(1 - d_i)\gamma_S(i) + d_i\gamma_R(i) \right] - \frac{1}{N}\sum_{i=1}^{N}\beta_i d_i(1 - d_i), \tag{56}
$$

where the Lagrange multipliers μ, β_i, and ν are chosen such that C1, C2, and C3 are satisfied, respectively. By differentiating $\mathcal{L}$ with respect to $\gamma_S(i)$, $\gamma_R(i)$, and d_i, and setting the results to zero, we obtain three equations. Solving this system of equations for $\gamma_S(i)$, $\gamma_R(i)$, and d_i, and taking into account that $\beta_i < 0$, $0 < \mu < 1$, and $\nu > 0$, we obtain (28), (29), and (30) after letting $\rho = \mu/(1-\mu)$ and $\lambda = \nu\ln(2)/(1-\mu)$, which are chosen such that constraints C1 and C3 are meet with equality. This completes the proof.

E. Proof of Lemma 2

Since $s(i)$ and $r(i)$ are ergodic random processes, for $N \to \infty$, the normalized sums in C1 and C3 in (27) can be replaced by expectations. Therefore, the left hand side of C1 is the expectation of variable $(1 - d_i)\log_2(1 + \gamma_S(i)h_S(i))$. This variable is nonzero only when both $(1 - d_i)$ and $\gamma_S(i)$ are nonzero. The domain over which $(1 - d_i)$ and $\gamma_S(i)$ are jointly nonzero can be obtained from (28) and (30) and is given by

$$
(h_S(i) > \lambda/\rho \text{ AND } h_R(i) < \lambda) \text{ OR } (h_S(i) > L_1 \text{ AND } h_R(i) > \lambda) \tag{57}
$$

where L_1 is given by (33). Variable $(1 - d_i)\log_2(1 + \gamma_S(i)h_S(i))$ has to be integrated over domain (57) to obtain its average. This leads to the left side of (31).

Similarly, the right hand side of C1 is the expectation of the variable $d_i \log_2(1 + \gamma_R(i)h_R(i))$. This variable is nonzero only when both d_i and $\gamma_R(i)$ are nonzero. The domain over which d_i and $\gamma_R(i)$ are jointly nonzero can be obtained from (29) and (30) and is given by

$$
(h_R(i) > \lambda \text{ AND } h_S(i) < \lambda/\rho) \text{ OR } (h_R(i) > L_2 \text{ AND } h_S(i) > \lambda/\rho) \tag{58}
$$

where L_2 is given by (33). Variable $d_i \log_2(1+\gamma_R(i)h_R(i))$ has to be integrated over domain (58) to obtain its average. This leads to the right side of (31).

Following a similar procedure, we can obtain (32) from C3 in (27). This completes the proof.

F. Proof of Theorem 5

For $\xi = E\{(1-d_i)S(i)\}/E\{d_iR(i)\} < 1$ the queue is non-absorbing, and thus, because of the law of conservation of flow, the throughput is equal to the arrival rate, i.e., $\tau = E\{(1-d_i)S(i)\}$.

To arrive at an upper bound for the average queue size, we first introduce two auxiliary results from the literature. Let

$$q(i) = \max\{q(i-1) - u(i), 0\}, \tag{59}$$

where $u(i)$ is a slot by slot uncorrelated random variable with $E\{u(i)\} > 0$. Also, let $a(i)$ and $b(i)$ be non-negative slot by slot uncorrelated random variables with $E\{b(i)\} > E\{a(i)\}$ and set $u(i) = b(i) - a(i)$. Then, equality [21]

$$E\{u^2(i)\} - 2E\{u(i)\}E\{q(i)\} = E\{(\max\{u(i) - q(i), 0\})^2\} \tag{60}$$

and inequality [22]

$$E\{(\max\{u(i) - q(i), 0\})^2\} \geq \left(1 - \frac{E\{a(i)\}}{E\{b(i)\}}\right) E\{b^2(i)\} \tag{61}$$

hold. Furthermore, combining (60) and (61), the following bound is obtained [22]

$$E\{q(i)\} \leq \frac{1}{2} \frac{E\{(b(i) - a(i))^2\} - (1-\xi)^2 E\{b^2(i)\}}{E\{b(i)\} - E\{a(i)\}}, \tag{62}$$

where $\xi = E\{a(i)\}/E\{b(i)\}$.

By rewriting the queue size as

$$Q(i) = \max\{Q(i-1) - d_iR(i) + (1-d_i)S(i), 0\}, \tag{63}$$

we observe that (63) is in the form of (59) if we let $q(i) = Q(i)$, $a(i) = (1-d_i)S(i)$, $b(i) = d_iR(i)$ and $\xi = E\{(1-d_i)S(i)\}/E\{d_iR(i)\}$. Thus, assuming that $s(i)$ and $r(i)$ are slot by slot uncorrelated, we can exploit (62) and upper bound the average size of the queue as

$$
\begin{aligned}
E\{Q(i)\} &\leq \frac{1}{2} \frac{E\{(1-d_i)S^2(i)\} + E\{d_iR^2(i)\} - 2E\{(1-d_i)d_iS(i)R(i)\} - (1-\xi)^2 E\{d_iR^2(i)\}}{E\{d_iR(i)\} - E\{(1-d_i)S(i)\}} \\
&= \frac{1}{2} \frac{E\{(1-d_i)S^2(i)\} + \xi(2-\xi)E\{d_iR^2(i)\}}{E\{d_iR(i)\} - E\{(1-d_i)S(i)\}}.
\end{aligned}
\tag{64}
$$

Since the average arrival rate is given by $A = E\{(1-d_i)S(i)\}$, we obtain (37) from (64) and Little's law (36). This completes the proof.

REFERENCES

[1] E. C. V. D. Meulen, "Three-Terminal Communication Channels," *Advances in Applied Probability*, vol. 3, pp. 120–154, 1971.

[2] T. Cover and A. El Gamal, "Capacity Theorems for the Relay Channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 572–584, Sep. 1979.

[3] A. Sendonaris, E. Erkip, and B. Aazhang, "User Cooperation Diversity — Part I. System Description," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1938, Nov. 2003.

[4] ——, "User Cooperation Diversity — Part II. Implementation Aspects and Performance Analysis," *IEEE Trans. Commun.*, vol. 51, pp. 1939–1948, Nov. 2003.

[5] A. Host-Madsen and J. Zhang, "Capacity Bounds and Power Allocation for Wireless Relay Channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 2020 –2040, Jun. 2005.

[6] N. Laneman, D. Tse, and G. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[7] N. Laneman and G. Wornell, "Distributed Space–Time Block Coded Protocols for Exploiting Cooperative Diversity in Wireless Networks," *IEEE Trans. Inform. Theory*, vol. IT-49, pp. 2415–2425, Oct. 2003.

[8] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative Communication in Wireless Networks," *IEEE Communications Magazine*, vol. 42, pp. 74 – 80, Oct. 2004.

[9] K. Azarian, H. E. Gamal, and P. Schniter, "On the Achievable Diversity-Multiplexing Tradeoff in Half-Duplex Cooperative Channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 4152–4172, Dec. 2005.

[10] T. Hunter and A. Nosratinia, "Diversity Through Coded Cooperation," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 283 – 289, Feb. 2006.

[11] G. Kramer, M. G. P., and Gupta, "Cooperative Strategies and Capacity Theorems for Relay Networks," *IEEE Trans. Inform. Theory*, vol. 51, pp. 3037 – 3063, Sep. 2005.

[12] R. Nabar, H. Bölcskei, and F. Kneubühler, "Fading Relay Channels: Performance Limits and Space-Time Signal Design," *IEEE J. Select. Areas Commun.*, vol. 22, pp. 1099–1109, Aug. 2004.

[13] B. Xia, Y. Fan, J. Thompson, and H. V. Poor, "Buffering in a Three-Node Relay Network," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 4492–4496, Nov. 2008.

[14] A. Ikhlef, D. Michalopoulos, and R. Schober, "Max-Max Relay Selection for Relays with Buffers," *IEEE Trans. Wireless Commun.*, vol. 11, 2012.

[15] A. Bletsas, A. Khisti, D. Reed, and A. Lippman, "A Simple Cooperative Diversity Method Based on Network Path selection," *IEEE J. Select. Areas Commun.*, vol. 24, pp. 659–672, Mar. 2006.

[16] L. Georgiadis, M. J. Neely, and L. Tassiulas, *Resource Allocation and Cross Layer Control in Wireless Networks*. Now Publishers Inc, 2006.

[17] R. Corless, G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth, "On the LambertW function," *Advances in Computational Mathematics*, vol. 5, no. 1, pp. 329–359, 1996.

[18] J. D. C. Little, "A Proof of the Queueing Formula: $L = \lambda w$," *Operations Research*, vol. 9, no. 3, pp. 383–388, May - Jun. 1961.

[19] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*. Morgan & Claypool, 2010.

[20] D. Gross, J. F. Shortle, J. M. Thompson, and C. M. Harris, *Fundamentals of Queueing Theory*, 4th ed. John Wiley & Sons, Inc, 2008.

[21] J. Kingman, "Some Inequalities for the Queue GI/G/1," *Biometrika*, vol. 49, pp. 315–324, 1962.

[22] J. Daley, "Inequalities for Moments of Tails of Random Variables With a Queueing Application," *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 41, pp. 139–143, 1977.
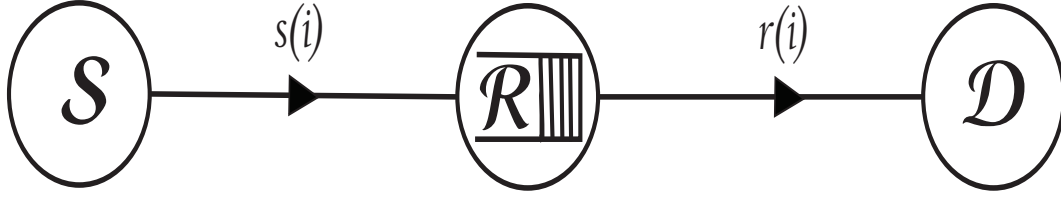
Fig. 1.   System model comprising a source $\mathcal{S}$, a half-duplex relay equipped with a buffer $\mathcal{R}$, and a destination $\mathcal{D}$. $s(i)$ and $r(i)$ are the instantaneous signal-to-noise ratios (SNRs) of the $\mathcal{S}$-$\mathcal{R}$ and $\mathcal{R}$-$\mathcal{D}$ links in the ith time slot, respectively.
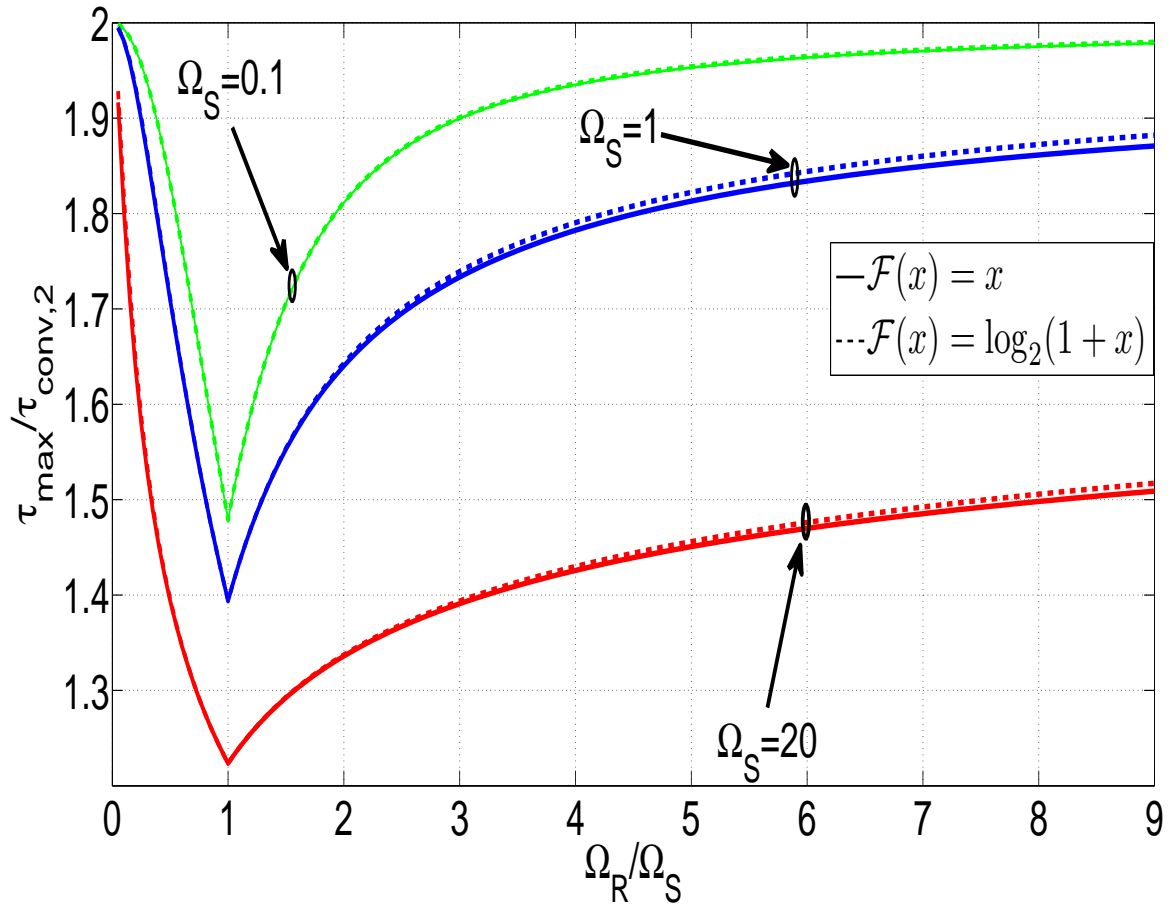

Fig. 2.   Throughput ratio $\tau_{\max}/\tau_{\mathrm{conv},2}$ vs. Ω_R/Ω_S for buffer-aided relaying with adaptive link selection and fixed transmit powers for source and relay.
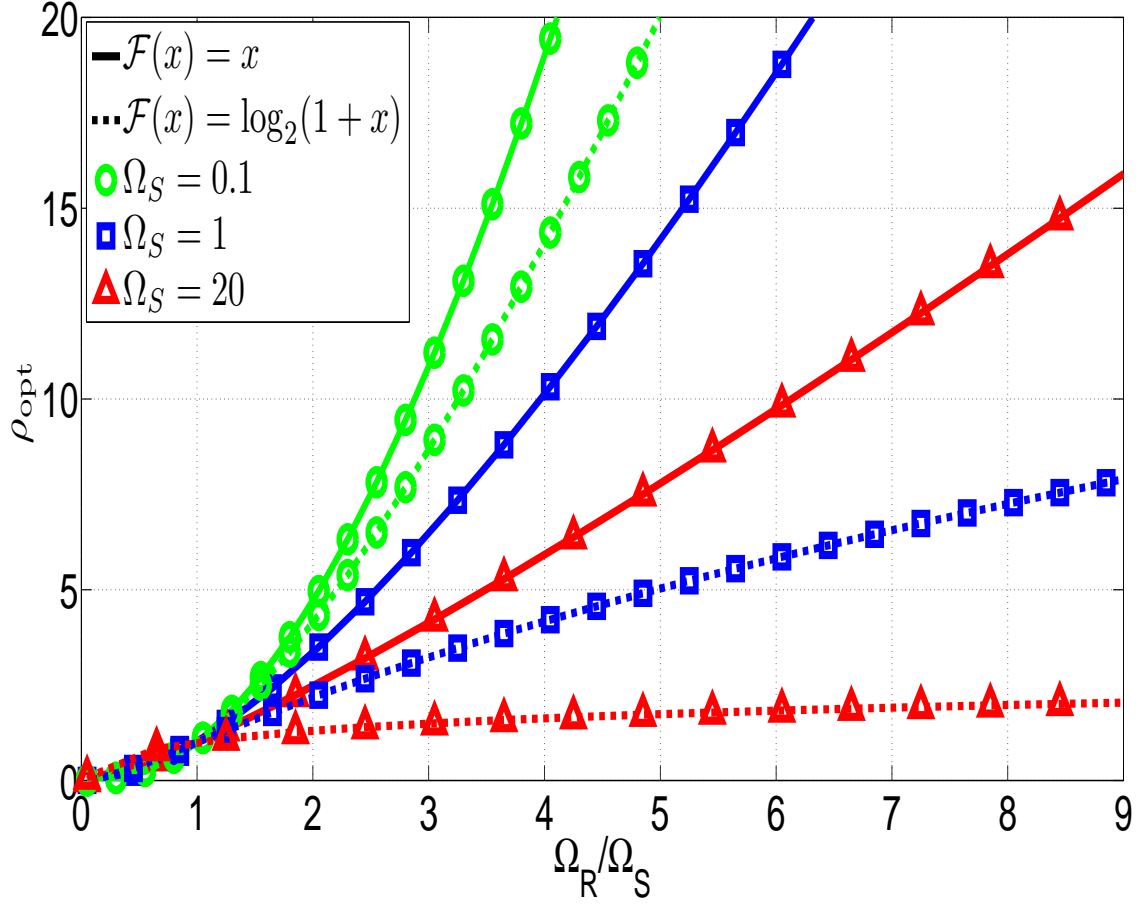
Fig. 3. Optimal decision threshold ρ_{opt} vs. Ω_R/Ω_S for buffer-aided relaying with adaptive link selection and fixed transmit powers for source and relay.
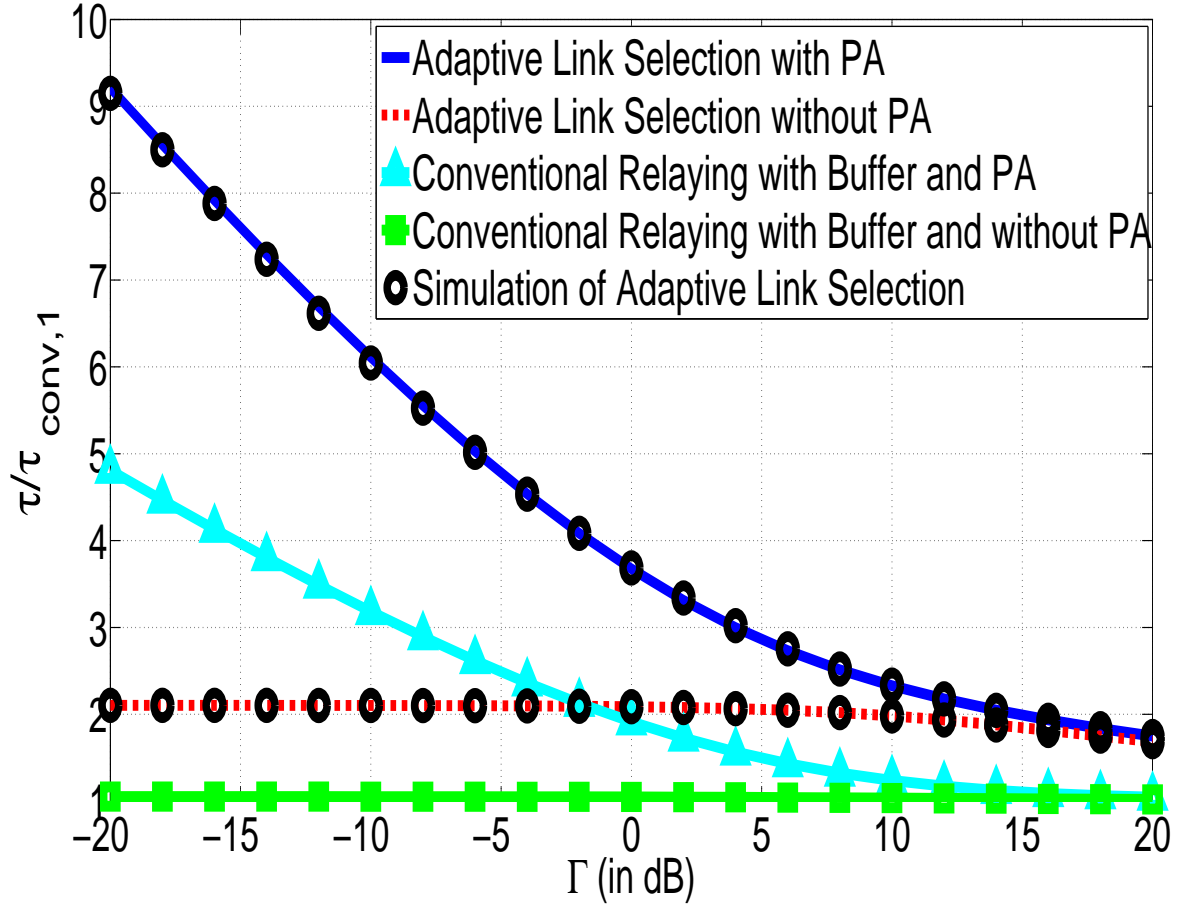
Fig. 4. Throughput normalized to $\tau_{\mathrm{conv},1}$ vs. Γ for buffer-aided relaying with adaptive link selection and conventional relaying with buffer. The performance of both schemes with and without power allocation (PA) is shown. $\bar{\Omega}_S = 0.1$ and $\bar{\Omega}_R = 1.9$
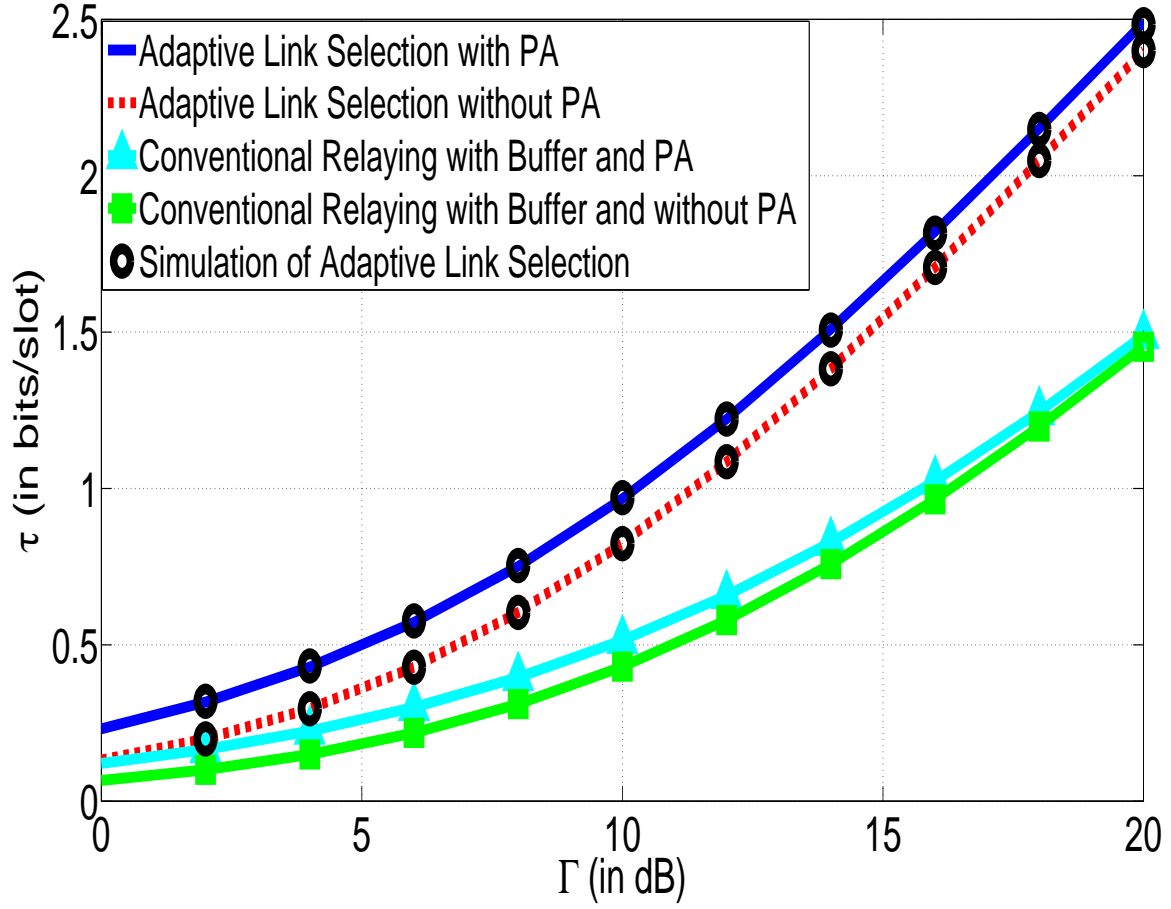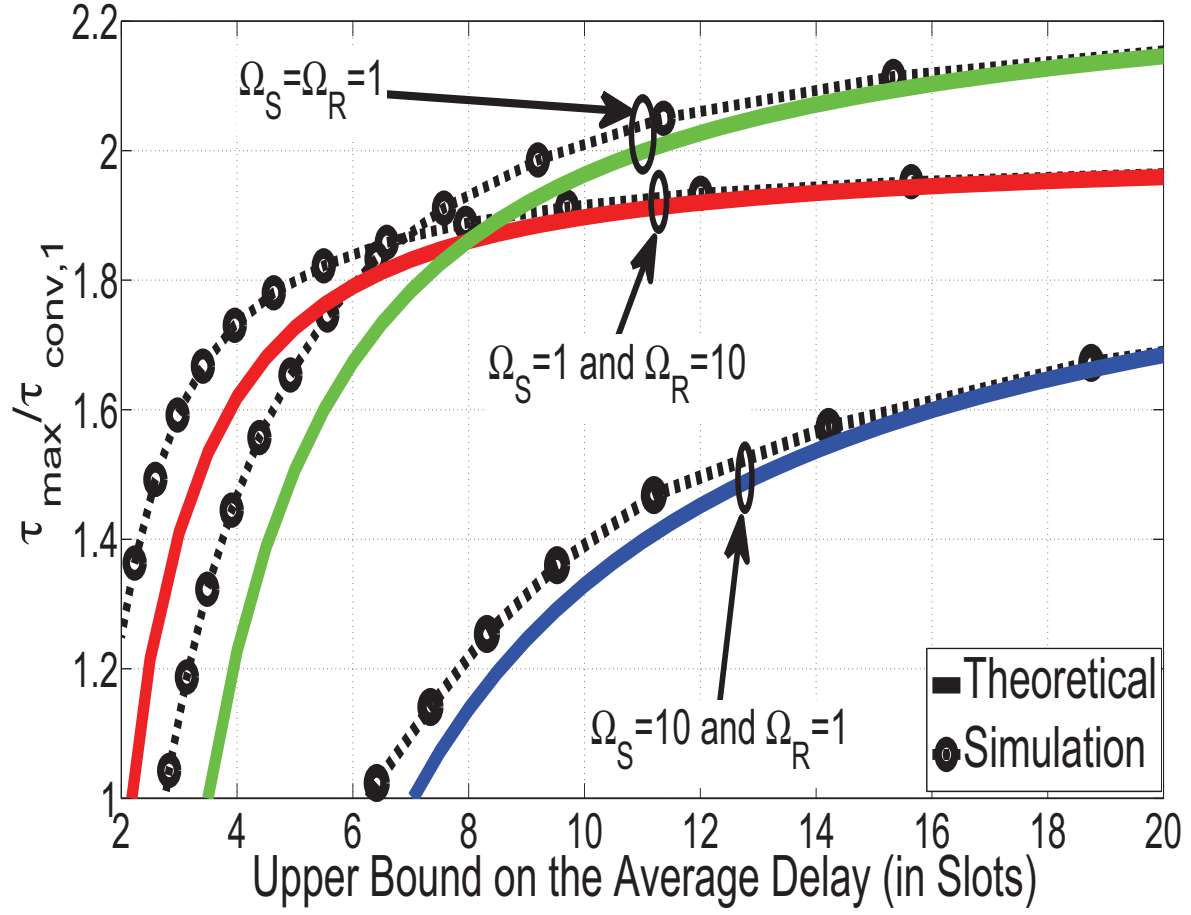
Fig. 5. Throughput vs. Γ for buffer-aided relaying with adaptive link selection and conventional relaying with buffer. The performance of both schemes with and without power allocation (PA) is shown. $\bar{\Omega}_S = 0.1$ and $\bar{\Omega}_R = 1.9$

Fig. 6. Throughput ratio $\tau_{\max}/\tau_{\mathrm{conv},1}$ vs. upper bound on the average delay for buffer-aided relaying with adaptive link selection and starved buffer. $\mathcal{F}(x) = x$.
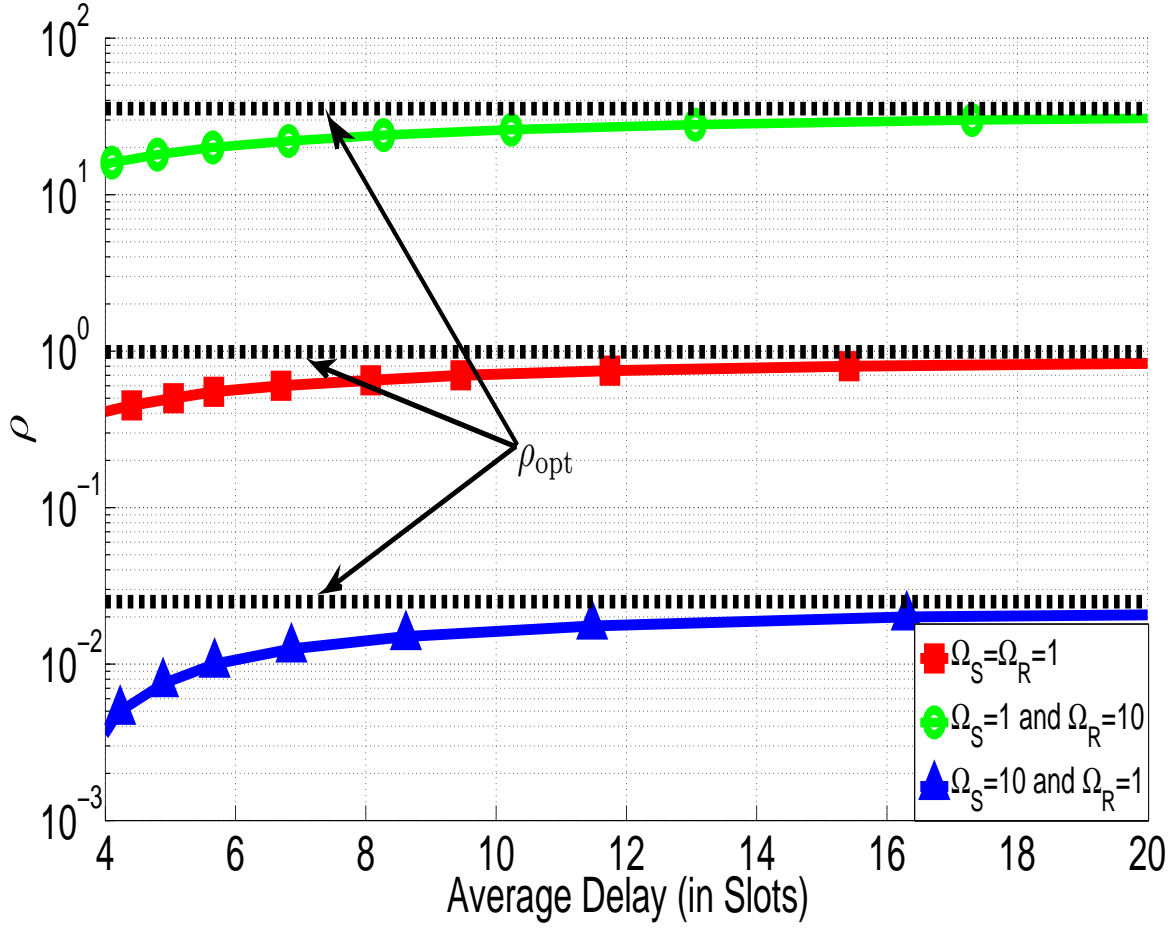
Fig. 7. Decision threshold ρ vs. upper bound on the average delay for buffer-aided relaying with adaptive link selection and starved buffer. $\mathcal{F}(x) = x$.
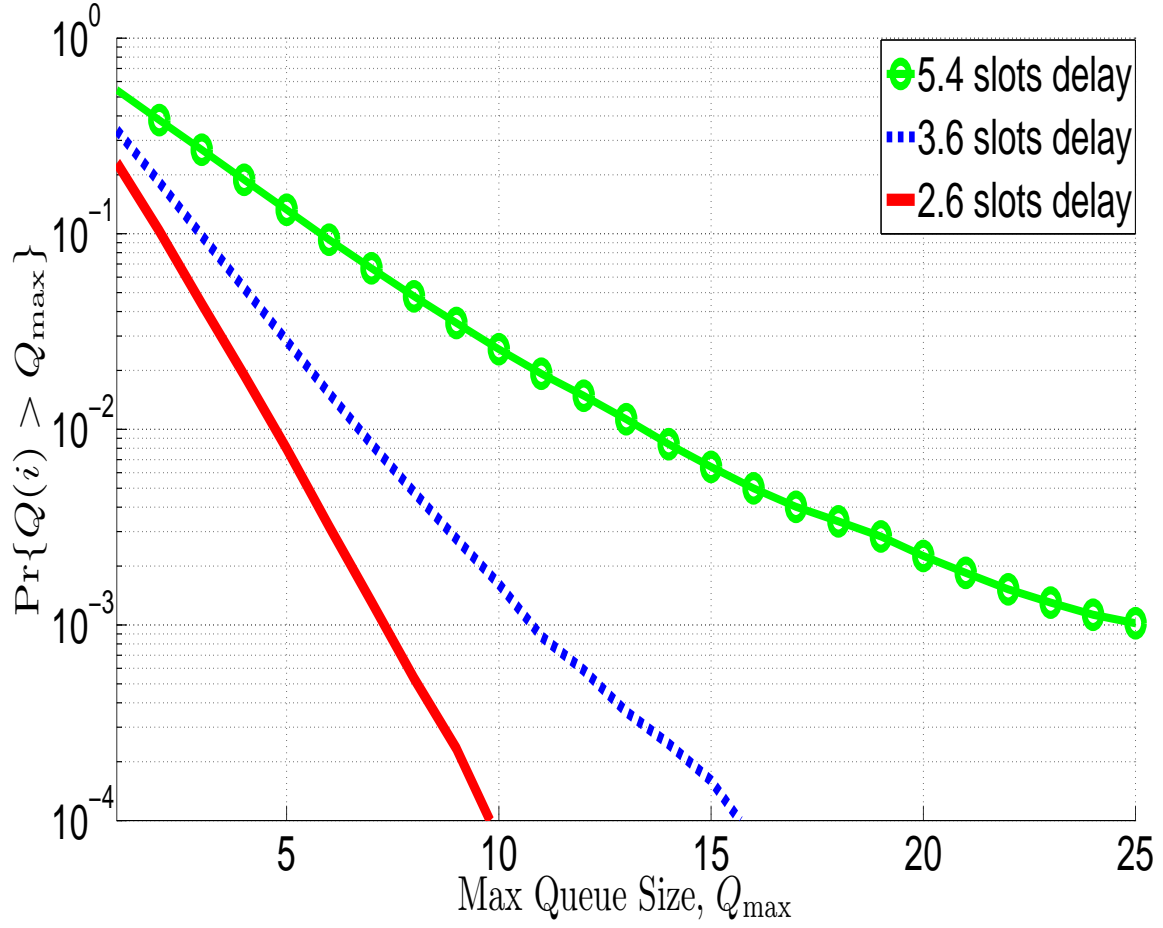
Fig. 8. Probability of a dropped bit vs. queue size $Q_{\max}$ for buffer-aided relaying with adaptive link selection and starved buffer. $\Omega_S = \Omega_R = 1$. $\mathcal{F}(x) = x$.
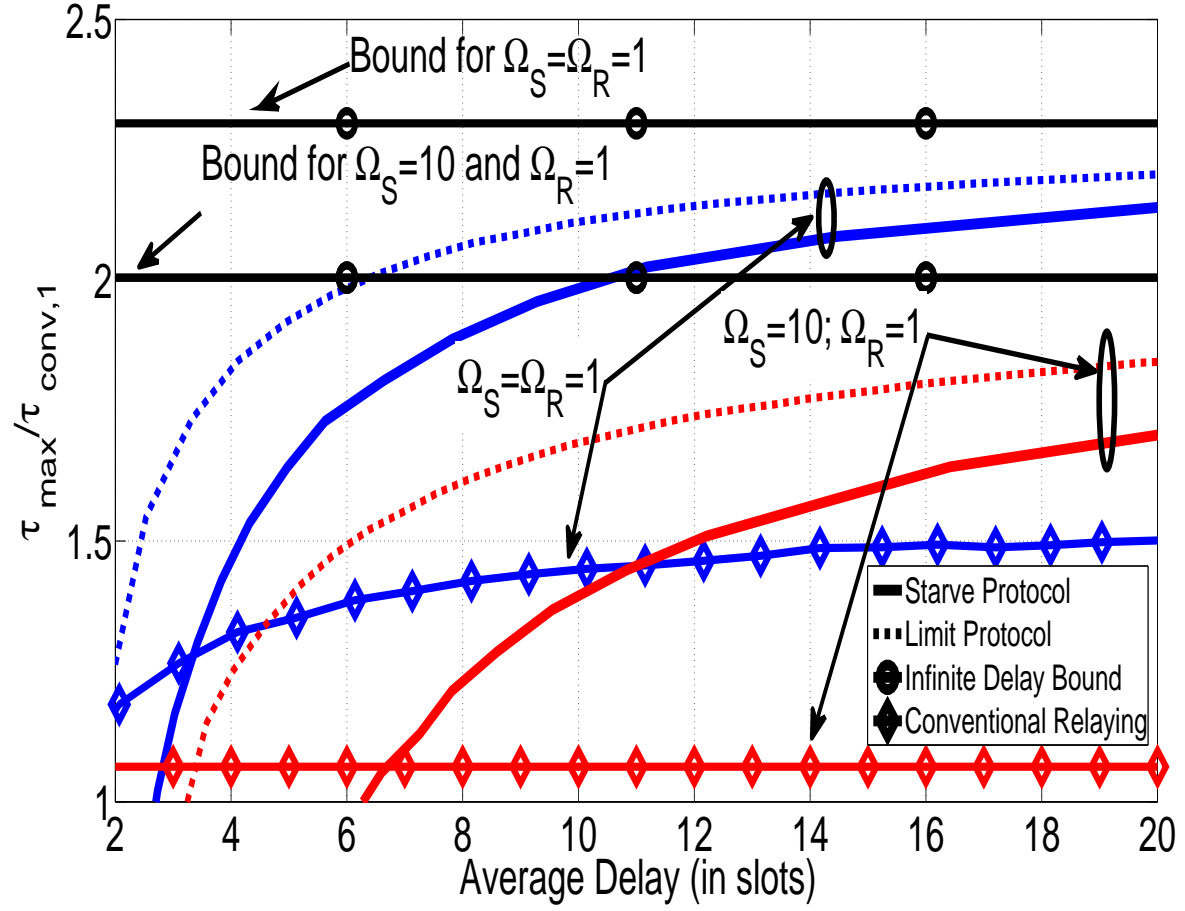
Fig. 9.   Throughput ratio $\tau_{\mathrm{max}}/\tau_{\mathrm{conv},1}$ vs. average delay for the proposed buffer-aided relaying schemes with starved buffer and limited buffer and conventional relaying with buffer [13].