Effective Capacity of Buffer-Aided Relay Systems with Selection Relaying

Deli Qiao
School of Information Science and Technology
East China Normal University, Shanghai, China, 200241
E-mail: dlqiao@ce.ecnu.edu.cn

Abstract—In this work, the achievable rate of three-node relay systems with selection relaying under statistical delay constraints, imposed on the limitations of the maximum delay violation probabilities, is investigated. It is assumed that there are queues of infinite size at both the source and relay node, and the source can select the relay or destination for data transmission. Given a selection relaying policy, the effective bandwidth of the arrival processes of the queue at the relay is derived. Then, the maximum constant arrival rate can be identified as the maximum effective capacity as a function of the statistical end-to-end queueing delay constraints and signal-to-noise ratios (SNR) at the source and relay, the fading distributions of the links, and the relay policy. Subsequently, a relay policy that incorporates the statistical delay constraints is proposed. It is shown that the proposed relay policy can achieve better performance than the one selecting the relay when source-relay link is stronger. Moreover, it is demonstrated that buffering relay model can still help improve the throughput of relay systems in the presence of statistical delay constraints and source-destination link.

I. INTRODUCTION

Relay channels can help improve the system coverage and throughput, and hence information-theoretic analysis of relay channels has been the research forefront for decades (see, e.g., [1]-[3]). Among the relaying protocols, selection relaying schemes are attractive due to their potential to improve bandwidth utilization and cooperation diversity [3]. While providing powerful results, information-theoretic studies generally assume no buffer at the relay. Recently, it has been shown that the achievable throughput can be further improved with the introduction of buffering relay model [4]. This is generally due to the information storage at the relay such that the shortcomings of existing relaying schemes caused by bad channel conditions can be overcome. The analysis of the buffer-aided relay systems has attracted much attention recently (see, e.g., [5], [6] and references therein). For instance, in [6], the authors proposed a max-max relay selection protocol, which chooses the relay with the strongest channel gain. They found that this policy can achieve better performance than systems without buffering relay [3]. However, the works on buffer-aided relaying systems rarely consider the buffer at the source. In [7], the authors investigated the buffer-aided relay systems with buffer at the source and energy-harvesting capability at each node. However, the analysis is based on the average arrival and service rate, and only stability regions of different strategies are considered.

In this paper, we consider the buffer-aided relay system under statistical delay constraints, imposed on the limitations of the maximum delay violation probabilities. We assume that there are buffers of infinite size at both the source and relay node, each subject to the statistical queueing constraints imposed on the limitations of buffer overflow probabilities. To handle the queueing dynamics of the relay networks, we employ the concept of effective bandwidth, which defines the bandwidth usage of given processes [15]. More recently, Wu and Negi in [8] defined the dual concept of effective capacity, which provides the maximum constant arrival rate that can be supported by a given departure process while satisfying statistical delay constraints. The analysis and application of effective capacity in various settings have attracted much interest recently (see e.g. [9]-[14] and references therein).

In this work, we consider the selection relaying strategy. The source selects the relay or destination for data transmission based on the channel state information (CSI) available. We assume that the relay policy is informed to the destination via an acknowledge (ACK) signal such that the destination can perform successive decoding of the received signals when destination is selected for data transmission. Given the relay protocol and statistical queueing constraints of the queues, we characterize the effective bandwidth of the arrival processes to the relay. Then, based on the established statistical delay tradeoff, we find the maximum effective capacity under the statistical delay constraints. Also, we propose a relay scheme taking into account the statistical delay constraints. Through numerical evaluation, this study reveals the benefits of buffering relay model in the presence of delay constraints and source-destination link.

The rest of this paper is organized as follows. Section II
discusses the necessary preliminaries on the system model, the statistical delay constraints and effective capacity. In Section III, we present our main results on the effective capacity and relay policy, with numerical results given in Section IV. Finally, Section V concludes this paper.

II. PRELIMINARIES

A. System Model

The three-node relay communication link is depicted in Fig. 1. The source can select the relay or destination for data transmission. In this model, there are buffers of infinite size at both the source and relay node. In this work, we assume full-duplex relay such that transmission and reception can be performed simultaneously.

The discrete-time input and output relationships in the $i$th symbol duration are given by

$$Y_i = g_{sr}[i]X_s + n_r$$
$$Y_d = g_{sd}[i]X_s + g_{rd}[i]X_r + n_d$$

where $X_j$ for $j \in \{s, r\}$ denotes the inputs for the links $S \rightarrow R$ and $R \rightarrow D$, respectively. The inputs are subject to individual average energy constraints $E[|X_j|^2] \leq P_j/B, j \in \{s, r\}$ where $B$ is the bandwidth. $Y_r, Y_d$ represent the received signal at $R$ and $D$, respectively. We assume that the fading coefficients $g_{sd}, g_{sr}, g_{rd}$ are jointly stationary and ergodic discrete-time processes, and we denote the magnitude-square of the fading coefficients by $z_{sd}[i] = |g_{sd}[i]|^2$, $z_{sr}[i] = |g_{sr}[i]|^2$, and $z_{rd}[i] = |g_{rd}[i]|^2$. Denote $z = (z_{sd}, z_{sr}, z_{rd})$. Assuming that there are $B$ complex symbols per second, we can easily see that the symbol energy constraint of $P_j/B$ implies that the channel input has a power constraint of $P_j$. Above, in the channel input-output relationships, the noise component $n_j[i]$ is a zero-mean, circularly symmetric, complex Gaussian random variable with variance $E[n_j[i]^2] = N_0$ for $j \in \{r, d\}$. The additive Gaussian noise samples $\{n_j[i]\}$ are assumed to form an independent and identically distributed (i.i.d.) sequence. We denote the signal-to-noise ratio at source as $SNR_s = P_s/N_0B$, and at relay as $SNR_r = P_r/N_0B$.

B. Statistical Delay Constraints

With the above mentioned settings, we first need the following result from [15].

**Theorem 1 ([15]):** Suppose that the queue is stable and that both the arrival process $a[n], n = 1, 2, \ldots$ and service process $c[n], n = 1, 2, \ldots$ satisfy the Gärtner-Ellis limit, i.e., for all $\theta \geq 0$, there exists a differentiable logarithmic moment generating function (LMGF) $\Lambda_A(\theta)$ such that

$$\lim_{n \to \infty} \frac{\log E[e^{\theta \sum_{n=1}^{\infty} a[n]}]}{n} = \Lambda_A(\theta),$$

and a differentiable LMGF $\Lambda_C(\theta)$ such that

$$\lim_{n \to \infty} \frac{\log E[e^{\theta \sum_{n=1}^{\infty} c[n]}]}{n} = \Lambda_C(\theta).$$

If there exists a unique $\theta^* > 0$ such that

$$\Lambda_A(\theta^*) + \Lambda_C(-\theta^*) = 0,$$

then

$$\lim_{Q_{\max} \to \infty} \frac{\log Pr\{Q > Q_{\max}\}}{Q_{\max}} = -\theta^*.$$  

where $Q$ is the stationary queue length.

Assume that the data queues are saturated, and hence they always attempt to transmit [7]. Then the delay violation probability can be written equivalently as [8], [11]

$$Pr\{D > D_{\max}\} = e^{-J(\theta)D_{\max}}$$

where we defined $f(x) = e^{cx}$ when $\lim_{x \to \infty} \frac{\log f(x)}{x} = c$, and

$$J(\theta) = \theta R = -\Lambda_C(-\theta)$$

is the statistical delay exponent associated with the queue, with $\Lambda_C(\theta)$ the LMGF of the service rate. Now, we can express the probability density function of random variable $D$ as

$$p_D(x) = \frac{d}{dx} (1 - Pr\{D > x\}) = J(\theta)e^{-J(\theta)x}.$$ 

Consider two concatenated queues with statistical queueing constraints specified by $\theta_1$ and $\theta_2$, for queue 1 and queue 2, respectively. Given the queuing constraints specified by $\theta_1$ and $\theta_2$ with (6) satisfied for each queue, we define

$$J_1(\theta_1) = -\Lambda_{C,1}(-\theta_1),$$

and

$$J_2(\theta_2) = -\Lambda_{C,2}(-\theta_2),$$

where $\Lambda_{C,1}(\theta_1)$ and $\Lambda_{C,2}(\theta_2)$ are the LMGF functions of the service rate of queue 1, 2, respectively. For data going through both queues, the delay violation probability can be characterized as

$$Pr\{D_1 + D_2 > D_{\max}\} = 1 - \int_0^{D_{\max}-D_1} \frac{dD_1}{dD_{\max}} \int_0^{D_{\max}-D_1} \frac{dD_2}{dD_{\max}} dD_1$$

where $D_{\max} = J_1(\theta_1)D_{\max}$. Denote $J_1(\theta_1) \neq J_2(\theta_2)$,

$$J_1(\theta_1) = J_2(\theta_2).$$

Therefore, we need to guarantee that

$$Pr\{D_1 + D_2 > D_{\max}\} \leq \epsilon.$$ 

In this way, we can guarantee that the data transmissions through the relay, i.e., information flow over two queues, satisfy the statistical delay constraints. Then, the delay constraints of the whole system can be satisfied. Note that $(\epsilon, D_{\max})$ characterizes the statistical delay constraints with maximum delay violation probability $\epsilon$ and maximum delay $D_{\max}$.

To facilitate the following analysis, we need the following tradeoff between $J_1(\theta_1)$ and $J_2(\theta_2)$.

**Lemma 1 ([14]):** Consider the following function

$$\vartheta(J_1(\theta_1), J_2(\theta_2)) = \frac{J_2(\theta_2)e^{-J_1(\theta_1)D_{\max}} - J_1(\theta_1)e^{-J_2(\theta_2)D_{\max}}}{J_2(\theta_2) - J_1(\theta_1)}$$

for $0 \leq \epsilon \leq 1$,

$$J_1(\theta_1) = J_2(\theta_2) = \Phi(J_1(\theta_1))$$

where $J_0 = \frac{-\log(\epsilon)}{D_{\max}}$ is defined as the statistical delay exponent associated with $(\epsilon, D_{\max})$. Denoting $J_2(\theta_2) = \Phi(J_1(\theta_1))$ as a
function of $J_1(\theta_1)$, we have

a) $\Phi$ is continuous. For $J_1(\theta_1) = J_{th}(\epsilon)$, we have

$$\Phi(J_1(\theta_1)) = J_{th}(\epsilon)$$

(14)

where

$$J_{th}(\epsilon) = -\frac{1}{D_{\text{max}}} \left(1 + W_{-1}\left(-\frac{\epsilon}{c}\right)\right)$$

(15)

where $W_{-1}(\cdot)$ is the Lambert W function, which is the inverse function of $y = xe^x$ in the range $(-\infty, -1]$.

b) $\Phi$ is strictly decreasing in $J_1(\theta_1)$.

c) $\Phi$ is convex in $J_1(\theta_1)$.

d) $J_1(\theta_1) \in [0, \infty)$, and $J_2(\theta_2) = \Phi(J_1(\theta_1)) \in [0, \infty)$.

Remark 1: This lemma indicates that to have the end-to-end delay constraints satisfied, we must increase $J_1(\theta_1)$ if $J_2(\theta_2)$ is decreased; vice versa.

C. Effective Capacity

Denote the queue at source $S$ as queue 1, and the queue at relay $R$ as queue 2. Denote $\Omega$ as the set of pairs $(\theta_1, \theta_2)$ such that (12) can be satisfied. Assume $\theta_1 > 0$ and $\theta_2 > 0$ at the source and the relay node with $(\theta_1, \theta_2) \in \Omega$. Assume that the constant arrival rate at the source is $R \geq 0$, and the channels operate at their capacities. Then, the effective capacity is defined as the maximum constant arrival rate such that the queueing constraints of the source and relay node can be satisfied. More specifically, to satisfy the queueing constraint at the source, we should have

$$\tilde{\theta} \geq \theta_1$$

(16)

where $\tilde{\theta}$ is the solution to

$$R = -\frac{\Lambda_{C,1}(\tilde{\theta})}{\tilde{\theta}}$$

(17)

and $\Lambda_{C,1}(\theta)$ is the LGMF of the service rate for queue 1, i.e., queue at the source.

Also, in order to satisfy the queueing constraint of the relay node $R$, we must have

$$\tilde{\theta} \geq \theta_2$$

(18)

where $\tilde{\theta}$ is the solution to

$$\Lambda_{A,2}(\tilde{\theta}) + \Lambda_{C,2}(\tilde{\theta}) = 0.$$  

(19)

where $\Lambda_{A,2}(\tilde{\theta})$ is the LGMF of the arrival process to queue 2, $\Lambda_{C,2}(\tilde{\theta})$ is the LGMF of the service process of queue 2, i.e., queue at the relay.

Note that we can derive the effective capacity $R_E(\theta_1, \theta_2)$ with $(\theta_1, \theta_2)$ following the method provided in [13, Theorem 2]. After these characterizations, effective capacity of the buffer-aided relay system under statistical delay constraints $(\epsilon, D_{\text{max}})$ can be formulated as follows.

Definition 1: The effective capacity of the buffer-aided relay system with statistical delay constraints specified by

$$(\epsilon, D_{\text{max}})$$

is given by

$$R_E(\epsilon, D_{\text{max}}) = \sup_{(\theta_1, \theta_2) \in \Omega} R_E(\theta_1, \theta_2)$$

(20)

III. EFFECTIVE CAPACITY WITH SELECTION RELAYING PROTOCOL IN BLOCK-FADING CHANNEL

In the following, we first discuss the transmission strategy in detail to obtain the associated channel rate. Then, assuming that $(\theta_1, \theta_2)$ are given, we obtain the effective bandwidth of the arrival processes of the queue at the relay given selection relaying policy. Next, we derive the effective capacity for arbitrary relay policy in a general form. Afterwards, we propose a relay policy taking into account the statistical delay constraints.

A. Transmission Strategy and Channel Rate

We assume that perfect CSI of all links is available at $S$ and $R$, while only the CSI of the links $S - D$ and $R - D$ is available at $D$. The transmission power levels at the source and relay are fixed and hence no power control is employed. We further assume that the channel capacity for each link can be achieved, i.e., the service processes are equal to the instantaneous Shannon capacities of the links. We consider a block fading scenario in which the fading stays constant for a block of $T$ seconds and change independently from one block to another.

We consider selection relaying protocols, i.e., the source sends information to the relay when the channel condition at the relay is larger than certain threshold [2]. Denote $Z$ as the region such that when $z \in Z$, the source $S$ selects the relay $R$ for data transmission. The relay strategy is forwarded by the source to the destination through an acknowledge (ACK) signal such that the destination can perform successive decoding of the received signals when $z \in Z^c$.

When $z \in Z$, we have a two-hop channel while the transmitted signal of $S - R$ link forms interference to the $R - D$ link. When $z \in Z^c$, we have a two-user multiple access channel, and the destination can perform successive decoding of the received signal from the source and relay node. Define $Z_0$ as the region depending on $z_{sd}$ and $z_{rd}$ such that when $z \in Z_0$, the destination $D$ decodes the received signal in the order of $(R, S)$, i.e., the sent signal from the source sees no interference. On the other hand, when $z \in Z^c \cap Z_0^c$, the decoding order at $D$ is $(S, R)$, i.e., the sent signal from the relay sees no interference.

To summarize, we have the service rates of the queues at the source and relay node as [16]

$$C_s = \begin{cases} 
TB \log_2 (1 + \frac{\text{SNR}_{z_{sd}}}{\epsilon}) & z \in Z \\
TB \log_2 (1 + \frac{\text{SNR}_{z_{rd}}}{\epsilon}) & z \in Z^c \cap Z_0 \\
TB \log_2 (1 + \frac{\text{SNR}_{z_{rd}}}{\epsilon}) & z \in Z^c \cap Z_0^c 
\end{cases}$$

(21)

and

$$C_r = \begin{cases} 
TB \log_2 (1 + \frac{\text{SNR}_{z_{rd}}}{\epsilon}) & z \in Z \cup \{Z^c \cap Z_0\} \\
TB \log_2 (1 + \frac{\text{SNR}_{z_{sd}}}{\epsilon}) & z \in Z^c \cap Z_0^c
\end{cases}$$

(22)
respectively. Above, the rates are in the units of bits per block or equivalently bits per $T$ seconds. These can be regarded as the service processes of the queues at the source and relay, respectively.

To ensure the stability of the queues, we need to enforce the following condition [15]

\[ \mathbb{E}_{z \in Z} \{ TB \log_2 (1 + \text{SNR}_{z_{sr}}) \} < \mathbb{E}_z \{ C_r \}. \quad (23) \]

That is, the average arrival rate of the queue at the relay should be less than the average service rate. We assume that the above condition is satisfied for the relay policies in consideration.

B. Effective Bandwidth of the S-R Link

Note that $\Lambda_{C,1}(\theta)$ and $\Lambda_{C,2}(\theta)$ can be easily derived with the relay policy $Z$ after we obtain the service rate of queue 1 and 2 in (21) and (22), respectively. Then, from the discussions in Section II-C, we must obtain the effective bandwidth of the arrival processes of the queue at the relay to derive the effective capacity.

This can be achieved by borrowing the idea of internetwork [15]. Note that queue 1 is the predecessor of queue 2, or equivalently, queue 2 is the successor of queue 1. Now the three-node relay network can be viewed as an internetwork. The routing depends on the relay protocol designed. Following the definitions in [15, Section 9.4], we define the routing variable

\[ p_1[i] = \begin{cases} 1, & z \in Z, \\ 0, & z \in Z^c. \end{cases} \quad (24) \]

That is, if the departure from queue 1 at $i$th symbol is routed to the successor, i.e., queue 2, and $p_1[i] = 0$ if the departure from queue 1 at $i$th symbol leaves the internetwork, i.e., goes to the destination node. Now, we have the log-moment generating function for the routing process as

\[ \Lambda_{p_1}(\theta, Z) = \lim_{n \to \infty} \log \mathbb{E} \left\{ e^{\theta \sum_{i=1}^n p_1[i]} \right\}. \quad (25) \]

Recalling the definitions in Section II-C, we have the following result.

Proposition 1: Given $(\theta_1, \theta_2)$ and the routing process specified by $Z$, the LMGF of the departure process from the source to the relay, or equivalently the arrival process to the relay node, is given by

\[ \Lambda_{A,2}(\theta) = \begin{cases} R\Lambda_{p_1}(\theta, Z), & 0 \leq \Lambda_{p_1}(\theta, Z) \leq \bar{\theta}, \\ R\bar{\theta} + \Lambda_{C,1}(\Lambda_{p_1}(\theta, Z) - \bar{\theta}), & \Lambda_{p_1}(\theta, Z) > \bar{\theta}. \end{cases} \]

where $R$ is the constant arrival rate to queue 1, and $\bar{\theta}$ is defined in (17).

C. Effective Capacity

Under the block fading assumption, the logarithmic moment generating functions for the service processes of queues at the source $S$ and the relay $R$ as functions of $\theta$ are given by [9]

\[ \Lambda_{C,1}(\theta) = \log \mathbb{E} \left\{ e^{\theta C_r} \right\} \quad \text{and} \quad \Lambda_{C,2}(\theta) = \log \mathbb{E} \left\{ e^{\theta C_r} \right\} \quad (26) \]

where $C_r$ and $C_s$ are given by (21) and (22), respectively. Combining (26) with (10) gives us $J_1(\theta_1)$ and $J_2(\theta_2)$. Now, the LMGF for the routing process (25) can be written as

\[ \Lambda_{p_1}(\theta, Z) = \log \left( \Pr\{z \in Z^c\} + e^\theta \Pr\{z \in Z\} \right). \quad (27) \]

The LMGF for the arrival process of the queue at the relay is

\[ \Lambda_{A,2}(\theta) = \begin{cases} R\Lambda_{p_1}(\theta, Z), & 0 \leq \Lambda_{p_1}(\theta) \leq \bar{\theta}, \\ R\bar{\theta} + \log \mathbb{E} \left\{ e^{(\Lambda_{p_1}(\theta,z)-\bar{\theta})C_s} \right\}, & \Lambda_{p_1}(\theta, Z) > \bar{\theta}. \end{cases} \quad (28) \]

For the following analysis, we need to characterize the relationship between $\Lambda_{p_1}(\theta, Z)$ and $\theta$. We have the following result.

Lemma 2: Consider the function

\[ \Lambda_{p_1}(\theta, Z) = \log \left( \Pr\{z \in Z^c\} + e^\theta \Pr\{z \in Z\} \right). \quad (29) \]

This function has the following properties:

a) $\Lambda_{p_1}(0, Z) = 0$, and $\Lambda_{p_1}(\theta, Z) \leq \theta$.

b) $\Lambda_{p_1}(\theta, Z)$ is increasing in $\theta$.

c) $\Lambda_{p_1}(\theta, Z)$ is a convex function of $\theta$.

Remark 2: Therefore, when $\theta = 0$, we have $\Lambda_{p_1}(0, Z) = 0$. As $\theta$ increases, $\Lambda_{p_1}(\theta, Z)$ increases. Also, $\Lambda_{p_1}(\theta, Z)$ increases at least linearly with $\theta$.

With the selection relaying strategy $Z$, we can establish an upperbound on the arrival rates supported by the relay system with specific $(\theta_1, \theta_2)$.

Proposition 2: The constant arrival rates, which can be supported by the buffering relay system with statistical queueing constraints specified by $(\theta_1, \theta_2)$ at the source and relay, respectively, are upperbounded by

\[ R \leq \min \left\{ -\frac{1}{\theta_1} \log \mathbb{E} \left\{ e^{-\theta_1 C_s} \right\}, -\frac{1}{\Lambda_{p_1}(\theta_2)} \log \mathbb{E} \left\{ e^{-\theta_2 C_r} \right\} \right\} \quad (30) \]

\[ = \min \left\{ \frac{J_1(\theta_1)}{\theta_1}, \frac{J_2(\theta_2)}{\Lambda_{p_1}(\theta_2, Z)} \right\} \quad (31) \]

Proof: The idea of this proof is similar to [13, Proposition 1]. The difficulty comes from proving the second term in the upperbound. Since there is direct link between $S$ and $D$, we should assume that the channel of links $S - R$ and $S - D$ are deterministic, and the relay strategy selects the relay randomly with probability $\Pr\{z \in Z\}$ for transmission. In this way, $\Lambda_{p_1}(\theta, Z)$ is still given by (27). Then, we can argue that the rate is upperbounded by $\frac{J_2(\theta_2)}{\Lambda_{p_1}(\theta_2, Z)}$. Details are omitted here due to space limit.

Considering Lemma 2, we can show that $\frac{J_2(\theta_2)}{\Lambda_{p_1}(\theta_2, Z)}$ is still an increasing and convex function of $\theta_2$ similar to [13, Lemma 1], which is fundamental to this article. Define

\[ \Omega_e = \{ (\theta_1, \theta_2) : J_1(\theta_1) \text{ and } J_2(\theta_2) \text{ are solutions to } (13) \}. \]

Similar to the discussions in [14], we can iterate over $(J_1(\theta_1), J_2(\theta_2))$ satisfying (13) and obtain the following result. Note that $z_{ij,\text{max}}$ and $z_{ij,\text{min}}$ represent the largest or smallest channel gain of link $i - j$, respectively.
The effective capacity of the buffer-aided relay systems with selection relaying strategy given by \( X \) subject to statistical delay constraints specified by \((e, D_{\text{max}})\) is given by the following:

**Case I:** If \( \theta_{1,th} = \Lambda_{p_1}(\theta_{2,th}, Z) \),
\[
R_e(e, D_{\text{max}}) = \frac{J_{th}(e)}{\theta_{1,th}},
\]
where \((\theta_{1,th}, \theta_{2,th})\) is the unique solution pair to \( J_1(\theta_1) = J_{th}(e) \), and \( J_2(\theta_2) = J_{th}(e) \).

**Case II:** If \( \theta_{1,th} > \Lambda_{p_1}(\theta_{2,th}, Z) \),
\[
R_e(e, D_{\text{max}}) = \begin{cases} 
\frac{J_{th}(e)}{\theta_{1,th}} & \text{if } \theta_{1,th} = \Lambda_{p_1}(\theta_{2,th}, Z) \\
TB \log_2 \left( 1 + \frac{\text{SNR}_{sr,\text{max}}}{1 + \text{SNR}_{zd,\text{max}}} \right), & \text{if } \theta_{1,th} > \Lambda_{p_1}(\theta_{2,th}, Z), \\
\frac{J_{th}(e)}{\theta_{1,th}} & \text{otherwise,}
\end{cases}
\]
where \( \theta_{1,0} \) is the solution to \( J_1(\theta_1) = J_0 \), and \( \tilde{\theta}_1 \) is the smallest value of \( \theta_1 \) with \((\theta_1, \theta_2) \in \Omega_e \) satisfying
\[
-\frac{1}{\theta_1} \log \mathbb{E}_{z_1} \left\{ e^{-\theta_1 C_{sr}} \right\} = -\frac{1}{\theta_1} \left( \log \mathbb{E}_{z_2} \left\{ e^{-\theta_1 C_{sr}} \right\} + \log \mathbb{E}_{z_1} \left\{ e^{(\Lambda_{p_1}(\theta_{2,0}, Z) - \theta_1) C_{sr}} \right\} \right).
\]
Moreover, if \( \frac{dJ_2(\theta)}{d\theta} \bigg|_{\theta=\tilde{\theta}_1} \leq \frac{dJ_1(\theta)}{d\theta} \bigg|_{\theta=\theta_1} \), where \( \tilde{\theta}_1 \) is given by \((\theta_1, \theta_2) \in \Omega_e \) with \( \theta_1 = \Lambda_{p_1}(\theta_{2,0}, Z) \), then the solution to (34) with \((\theta_1, \theta_2) \in \Omega_e \) is unique.

**Case III:** If \( \theta_{1,th} < \Lambda_{p_1}(\theta_{2,th}, Z) \),
\[
R_e(e, D_{\text{max}}) = \begin{cases} 
\frac{J_{th}(e)}{\theta_{1,th}} & \text{if } \theta_{1,0} = \Lambda_{p_1}(\theta_{2,0}, Z) \\
TB \log_2 \left( 1 + \frac{\text{SNR}_{zd,\text{min}}}{1 + \text{SNR}_{sr,\text{max}}} \right), & \text{if } \theta_{1,0} < \Lambda_{p_1}(\theta_{2,0}, Z) \\
\frac{J_{th}(e)}{\theta_{1,th}} & \text{otherwise,}
\end{cases}
\]
where \( \theta_{2,0} \) is the solution to \( J_2(\theta_2) = J_0 \), and \((\tilde{\theta}_1, \tilde{\theta}_2) \) is the unique solution to
\[
J_1(\theta_1) = J_2(\tilde{\theta}_2) = \frac{J_{th}(e)}{\Lambda_{p_1}(\theta_{2,0}, Z)}
\]
with \((\theta_1, \tilde{\theta}_2) \in \Omega_e \).

**Proof:** The idea of this proof follows that in [14, Appendix C], except that the effective bandwidth of the arrival process to the queue at the relay is now given by (28), and the upperbound on the effective capacity is given by (30) instead. Details are omitted due to space limit. \( \square \)

**Remark 3:** The above theorem covers all the possibilities that symmetric or asymmetric delay constraints on the queues at the source and relay node can be optimal for achieving the maximum effective capacity of the relay system.

**D. Selection Relaying Protocols**

Considering the expression of the effective capacity and the associated conditions in Theorem 2, we note that finding the optimal relaying protocol in closed-form analytical expressions seems intractable for a general scenario. With this in mind, we consider a simplified case in which the relaying protocol is decided by a function of \( z_{sr} \) and \( z_{zd} \), and is denoted as \( z_{sr} = g(z_{zd}) \). The channel state region \( Z \) is given by
\[
Z = \{ z : z_{sr} > g(z_{zd}) \}.
\]
Also, assume that the decoding strategy at the destination is given by \( z_{rd} = f(z_{zd}) \), such that
\[
Z_0 = \{ z : z_{rd} > f(z_{zd}) \}.
\]

1) **Max Channel Gain:** A typical relay policy with buffer-aided relay system is to select the link with stronger channel condition [6]. Then, we have
\[
g(z_{zd}) = z_{zd}.
\]

With this relaying scheme, we can obtain the effective capacity according to Theorem 2.

2) **Max Delay Exponent:** In this part, we propose a relay policy that takes into account the statistical delay constraints as well. Assume that the optimal statistical queueing constraints \((\theta_1, \theta_2) \in \Omega_e \) that maximize the effective capacity are given. With this parameter set \((\theta_1, \theta_2) \), we consider the relay strategy that maximizes the statistical delay exponent \( J_1(\theta_1) \) at the source, in which case the effective capacity can be potentially improved. Assume that the channel of the link \( R - D \) is independent of the links \( S - R \) and \( S - D \). Combining (10), (21), (26), (38) and (39), we can express the statistical delay exponent \( J_1(\theta_1) \) at the source as (41) on top of next page. Then, the associated relaying strategy should be the solution to the following optimization problem
\[
\max_{g} J_1(\theta_1, g(z_{zd})).
\]

We can obtain the relaying strategy specified as below.

**Theorem 3:** Assume \((\theta_1, \theta_2) \in \Omega_e \). The relay strategy as a function of \((z_{zd}, z_{sr})\) that maximizes the statistical delay exponent at the source is given by
\[
g(z_{zd}, \theta_1) = \frac{1}{\text{SNR}} \left( e^{-\frac{1}{\beta_1} \log \mathbb{E}_{z_{zd}} \left\{ e^{-\theta_1 C_{zd,MAC}} \right\} } - 1 \right)
\]
where \( \beta_1 = \frac{\theta_1 TB}{\log 2} \), and
\[
C_{zd,MAC} = \begin{cases} 
TB \log_2 \left( 1 + \frac{\text{SNR}_{zd}}{1 + \text{SNR}_{sr}} \right), & \text{if } z_{rd} \geq f(z_{zd}), \\
TB \log_2 \left( 1 + \frac{\text{SNR}_{zd}}{1 + \text{SNR}_{sr}} \right), & \text{if } z_{rd} < f(z_{zd}).
\end{cases}
\]

**Proposition 3:** \( J_1(\theta_1, g(z_{zd}, \theta_1)) \) is increasing in \( \theta_1 \). Therefore, given \( J_1(\theta_1, g(z_{zd}, \theta_1)) \), achieving \( J_1(\theta_1, g(z_{zd}, \theta_1)) = J_1 \) can be done by choosing \( \theta_1 \) such that \( g(z_{zd}, \theta_1) \) is increasing in \( \theta_1 \). Then, we can still apply the method in Theorem 2 to obtain the maximum effective capacity associated with the proposed relay policy.
\[ J_1(\theta_1, g(z_{sd})) = -\log \left( \int_0^\infty \int_0^\infty (1 + \text{SNR}_s) e^{-\beta_1} p(z_{sd}, z_{sr}) dz_{sr} dz_{sd} + \int_0^\infty \int_0^\infty (1 + \text{SNR}_s) e^{-\beta_1} p(z_{sd}, z_{sr}) p(z_{rd}) dz_{rd} dz_{sr} dz_{sd} \right) \]

\[ + \int_0^\infty \int_0^\infty \int_0^\infty (1 + \text{SNR}_s) e^{-\beta_1} p(z_{sd}, z_{sr}) p(z_{rd}) dz_{rd} dz_{sr} dz_{sd} \]

Remark 4: As \( \epsilon \to 1 \), i.e., \( \theta_1 \to 0 \) and \( \theta_2 \to 0 \), we have

\[ g(z_{sd}, 0) = \frac{1}{\text{SNR}} \left( e^{\text{SNR}_r \{C_r_{\text{MAC}} \}} - 1 \right). \] (45)

This is equivalent to selecting the relay when the instantaneous channel rate of \( S - R \) is larger than the average channel rate of \( S - D \) given \( z_{sd} \). Note that \( g(z_{sd}, 0) < z_{sd} \), and hence the source may select the relay for transmission even when the link \( S - D \) is better due to the potential interference from the relay. In this case, according to Theorem 2, we have the achievable throughput given by

\[ \min \left\{ \mathbb{E}\{C_s\}, \frac{\mathbb{E}\{C_r\}}{Pr\{z_{sr} > g(z_{sd}, 0)\}} \right\} \] (46)

where the first term is the average service rate of the queue at the source, while the second term is maximum supported rate considering the buffer at the relay.

Remark 5: If \( f(z_{sd}) = 0 \), i.e., the destination always decodes the signal from \( S \) last, we know that \( g(z_{sd}, 0) = z_{sd} \), i.e., the proposed relay policy reduces to the max channel gain selection.

3) No-Buffer Relay: For comparison, we also consider the performance of the relay system without buffer at the relay. We assume buffer at the source node only. Information-theoretical analysis have shown that the maximum rate for the three-node relay network with DF is given by [2]

\[ C = \frac{TB}{2} \min\{\log_2(1 + 2\text{SNR}_{sr}), \log_2(1 + 2\text{SNR}_{zd} + 2\text{SNR}_{zd})\}. \]

The effective capacity associated with arbitrary queueing constraints \( \theta \) is

\[ R(\theta) = -\frac{1}{\theta} \log \mathbb{E}\{e^{-\theta C}\}. \] (47)

Then, to guarantee the statistical delay constraints \( (\epsilon, D_{\text{max}}) \), we should have from (7)

\[ \theta R(\theta) \geq -\frac{\log \epsilon}{D_{\text{max}}}. \] (48)

Since \( R(\theta) \) is decreasing in \( \theta \), there must be one smallest \( \theta_{\text{min}} \) such that the above inequality holds with equality. The maximum achievable throughput is then given by \( R(\theta_{\text{min}}) \).

IV. NUMERICAL RESULTS

We consider the relay model depicted in Fig. 2. The source, relay, and destination nodes are located on a straight line. The distance between the source and the destination is normalized to 1. Let the distance between the source and the relay node be \( d \in (0, 1) \). Then, the distance between the relay and the destination is \( 1 - d \). We assume the fading distributions for \( S - D, S - R \) and \( R - D \) links follow independent Rayleigh fading with means \( \mathbb{E}\{z_{sr}\} = 1, \mathbb{E}\{z_{rd}\} = 1/d^\alpha \) and \( \mathbb{E}\{z_{zd}\} = 1/(1 - d)^\alpha \), respectively, where we assume that the path loss \( \alpha = 4 \). We assume \( \text{SNR}_r = 0 \) dB, \( T = 1 \) ms, \( B = 180 \) kHz, \( d = 0.5 \), and \( f(z_{sd}) = z_{sd} \) in the following numerical results.

In Fig. 3, we plot the effective capacities for the different relay policies as a function of \( \text{SNR}_r \). The relay node. We see from the figure that buffering relay can still help improve the system throughput under statistical delay constraints. In addition, the proposed relay policy always achieve better performance than the policy that selects relay when its channel condition is better. As \( \text{SNR}_r \) increases, the achievable throughput without buffer at relay approaches some finite value, since the service rate of the queue at the source is limited by \( \frac{TB}{2} \log_2(1 + 2\text{SNR}_{sr}) \).

In Fig. 4, we plot the effective capacity as \( \epsilon \) varies for \( \text{SNR}_r = \{5, 10, 15\} \) dB. It is interesting that buffering relay helps improve the throughput for a wide range of statistical delay constraints.

V. CONCLUSION

In this paper, we have analyzed the maximum arrival rates that can be supported by a full-duplex buffer-aided relaying system under statistical delay constraints. We have assumed that both the source and the relay have perfect CSI of the all
Fig. 4. Effective capacity as a function of $\epsilon$. SNR$_1 = 0$ dB, $D_{\text{max}} = 1$ sec.

links, while the destination only has the CSI of the links to itself. We have assumed that the source either transmit data to the destination or the relay based on selection relaying strategy, which is informed to the destination by the source through an ACK signal such that the destination can perform simultaneous decoding of the received signal. Given arbitrary selection relay policy, we have analytically expressed the effective bandwidth of the departure process from the queue at the source. We have obtained the maximum effective capacity under statistical end-to-end delay constraints. In the subsequent analysis, we have proposed a relay policy that takes into account the statistical delay constraints as well. Through numerical results, we have shown that the delay constraint based relay scheme can achieve better performance than max channel gain selection policy. Also, we have found that buffering relay can still help improve the throughput in the presence of statistical delay constraints and direct transmission link.

REFERENCES