Cross-Layer Analysis of Cognitive Radio Relay Networks under Quality of Service Constraints

Invited Paper

Leila Musavian and Sonia Aissa
INRS-EMT
University of Quebec
Montreal, QC, Canada

Abstract—In this paper, we investigate the performance gains of cognitive radio relay networks under delay quality of service (QoS) limitations at the secondary users, and spectrum-sharing restrictions imposed by the primary users of the channel. In particular, we assume that the primary user allows secondary users to gain access to its allocated spectrum band as long as a certain threshold on its corresponding outage probability is satisfied. Using this constraint, we find the maximum limit on the interference-power inflicted on the primary receiver that should not be exceeded by the transmission of the secondary users. In addition, we assume that the secondary transmitter benefits from an intermediate node, chosen from \( K \) terminals, to relay its signal to the destination. Considering that the transmission of the secondary user is subject to satisfying a statistical delay QoS constraint, we obtain the maximum arrival-rate supported by the secondary user’s relaying link. In this respect, we derive closed-form expressions for the effective capacity of the channel in Rayleigh block-fading environment. Numerical simulations are provided to endorse our theoretical results.

I. INTRODUCTION

During recent years, demand for radio spectrum has considerably increased. These frequencies are controlled and allocated by national regulation bodies, e.g., the federal communications commission (FCC) [1], whose recent measurements have shown that the spectrum is not optimally utilized [2]. For instance, cellular networks are overloaded while amateur radios and paging are underutilized. The apparent scarcity of spectrum is in fact due to an inefficiency in its utilization rather than a shortage of this precious resource. As a result, we have recently witnessed increasing interest in developing more flexible spectrum usage models. In particular, new ideas emerged to allow non-licensed users, referred to as secondary users, to be able to utilize the frequency bands that are allocated to primary users, as long as the communication of the former does not interfere or disturb the latter.

Cognitive radio, coined by J. Mitola [3], is a technology that improves the utilization of the spectrum by intelligently sharing its resources. One of the major challenges of CR systems is indeed to support quality of service (QoS) requirements for different applications of the secondary user network [4]. For instance, in systems that carry delay-sensitive applications, we need to ensure that the delay adheres to the system requirements. However, since the capacity varies as a function of the channel quality, satisfying deterministic delay QoS constraints is very challenging, and even impossible in some cases [5]. Therefore, statistical delay QoS constraints, where delay is required to be lower than a specific threshold only for a certain percentage of time are considered in various applications [6]. Recently, the concept of effective capacity has been introduced as a link-layer model for supporting QoS requirements of the system [5]. Effective capacity is the dual of effective bandwidth [6], and can be interpreted as the maximum constant arrival-rate that can be supported by the channel while satisfying the delay QoS requirement [5]. In this regard, an optimum power and rate allocation strategy that achieves the maximum on the effective capacity in fading channels has been obtained in [7].

On the other hand, relaying has emerged as a powerful approach to improve the reliability and the throughput of wireless networks. The capacity of relay channels for different techniques, e.g., Decode-and-forward (DF) relaying, was obtained in [8]. Recently, several studies have been conducted to determine capacities, bandwidth and power allocation for cooperative and relay channels, e.g., [9].

In this paper, we consider a relay channel where \( K \) relay nodes are present in order to help the secondary transmitter’s communication with its receiver. The transmission parameters of the secondary transmitter and the relay nodes are assumed to be limited such that the primary user is supported with a minimum-rate for a certain percentage of time. Translating this limitation into an interference-power constraint, we obtain the maximum throughput of the secondary user’s channel under delay QoS constraint by obtaining the effective capacity of the channel. We refer to the earlier work of Gastpar who presented capacity investigations of additive white Gaussian noise (AWGN) spectrum-sharing channels under interference-power constraint rather than transmit power constraint [10]. Later, ergodic and outage capacity metrics of a system with constraints on the received-power at the primary’s receiver in fading environment were derived in [11], [12]. A QoS-driven power and rate allocation scheme under spectrum-sharing constraint was proposed in [13], wherein the effective capacity of a point-to-point channel in Rayleigh fading environment was found. In this paper, we obtain the effective capacity of cognitive radio relay networks and investigate the capacity gains that can be achieved through opportunistic relay selection in spectrum-sharing channels.

This work was supported in part by a Discovery Grant (DG) from the Natural Sciences and Engineering Research Council (NSERC) of Canada.
II. SYSTEM MODEL

We consider a spectrum-sharing network, where in the secondary user communication system, the upper layer packets are organized into frames with the same time duration, $T_f$, at the data-link layer. These frames are stored in the transmit buffer and then divided into bit streams that will be transmitted through the channel.

The secondary transmitter communicates with its receiver through an intermediate (relay) node. We assume that multiple relay nodes are available and that the relay node that provides the highest achievable rate is used for the communication between the secondary transmitter and receiver. The relaying is based on the DF technique. Hence, each relay node listens to the signal transmitted from the secondary transmitter in the first time slot and then decodes the received signal and relays it to the secondary’s receiver in the second time slot. The system model is shown in Fig. 1.\textsuperscript{1}

Full channel state information (CSI) is assumed to be available at the transmitter and receiver of the secondary user. Each relay node, on the other hand, is provided with perfect knowledge about the channel gain between its receiver and the secondary’s transmitter, $h_{SR}$, the one between its transmitter and the secondary’s receiver, $h_{R,D}$, and the one between its transmitter and the primary’s receiver, $h_{R,P}$. Information about the latter can be carried out by a band manager that intervenes between the primary and secondary users [14], or can be directly fed back from the primary’s receiver to the secondary user, as proposed in [15] where algorithms that allow the primary and secondary users to collaborate and exchange CSI are proposed. The secondary transmitter analyzes the CSI in order to choose the relay node to be active in the next time slot. The index of the active relay node can then be broadcast to the relay nodes through a parallel channel.

We define the channel gain between the secondary’s transmitter and the primary’s receiver by $h_{SP}[n]$, and assume the channel gains to be independent and identically distributed (i.i.d.) according to Gamma distribution with unit variance. Channel gains are stationary and ergodic random processes. The transmitters are assumed to use Gaussian codebooks, and the noise power spectral density and received signal bandwidth are denoted by $N_0$ and $B$, respectively.

We further assume that the transmission technique of the secondary user must satisfy a statistical delay QoS constraint. It is shown that the probability for the queue length of the transmit buffer exceeding a certain threshold, $x$, decays exponentially as a function of $x$ [6] and [7]. We now define $\theta$ as a delay QoS exponent such that

$$\theta = -\lim_{x \to \infty} \frac{\ln \{Pr \{q(\infty) > x\}\}}{x},$$

(1)

with $q(n)$ indicating the transmit buffer length at time $n$ and $Pr \{a \leq b\}$ denoting the probability that the inequality $a \leq b$ holds true. Note that $\theta \to 0$ corresponds to a system with no delay constraint, while $\theta \to \infty$ implies a strict delay constraint. Considering $\theta$ as the delay QoS exponent in our system, we obtain the secondary user’s maximum supported arrival-rate given that the QoS constraint is satisfied.

III. INTERFERENCE-POWER CONSTRAINT

We recall that the transmission power of the secondary transmitter and relay transmitters are limited such that the primary user is provided with a minimum-rate $R_{\min}$ for a certain percentage of time $(1 - P_{P_{\text{out}}})$.

Since the channel gains are i.i.d., we study the interference-power constraint for the secondary transmitter and the same approach applies to the relay transmitters. We start by formulating the interference constraint according to

$$Pr \{R_p \leq R_{\min}\} \leq P_{P_{\text{out}}},$$

(2)

where $R_p$ indicates the rate of the primary user link. The transmission power of the primary user is assumed to be constrained to an average level, i.e.,\textsuperscript{2}

$$E \{P_T(h_P)\} \leq \bar{P},$$

(3)

where $E$ defines the expectation over the probability density function (PDF) of $h_P$ and $P_T(h_P)$ is the input transmit power of the primary user as a function of $h_P$. The transmission scheme of the primary user is chosen without considering the presence of the secondary user network. The primary transmitter employs the adaptive power and rate allocation scheme (opera) [16]. Hence, the power allocation of the primary can be found as

$$P_T(h_P) = \left[\mu - \frac{N_0 B}{h_P}\right]^+, \quad (4)$$

where $[x]^+$ indicates $\max\{x,0\}$ and the cutoff threshold $\mu$ is found such that the power constraint (3) is satisfied with equality.

The left-hand-side of constraint (2) can now be expanded according to

$$Pr \left\{ \ln \left(1 + \frac{P_T(h_P)h_P}{P(\theta,\mu,h_{SR},h_{SP})h_{SR} + N_0 B}\right) \leq R_{\min}, \quad h_P \geq \frac{N_0 B}{\mu} \right\} + Pr \left\{ h_P < \frac{N_0 B}{\mu} \right\} \leq P_{P_{\text{out}}},$$

(5)

where $P(\theta,\mu,h_{SR},h_{SP})$ denotes the transmit power of the secondary user as a function of $\theta$, $h_{SR}$ and $h_{SP}$. The interference

\textsuperscript{1}Note that Fig. 1 shows the channel power gains and not the channel coefficients.

\textsuperscript{2}Hereafter, we omit the time index $n$ whenever it is clear from the context.
constraint can further be expanded as
\[
\Pr \left\{ k_2 \leq h_p \leq k_1 \left( P(\theta, h_{SR}, h_{SP})h_{SP} + N_0B \right) + k_2 \right\} \\
+ \left( 1 - e^{-k_2} \right) \leq P_{out}^p,
\]
where \( k_1 = \frac{e^{R_{min}} - 1}{\mu} \) and \( k_2 = \frac{N_0B}{\mu} \). The probability function in (6) can be expanded to
\[
P_{out}^p - \left( 1 - e^{-k_2} \right) \\
\geq \int_0^\infty \int_0^\infty k_1 \left( P(\theta, h_{SR}, h_{SP})h_{SP} + N_0B \right) + k_2 \\
\times f_{h_p}(h_p) f_{h_{SR}}(h_{SR}) dh_p dh_{SR} \\
= \int_0^\infty \int_0^\infty e^{-k_2} - e^{-k_1 \left( P(\theta, h_{SR}, h_{SP})h_{SP} + N_0B \right) - k_2} \\
\times f_{h_p}(h_p) f_{h_{SR}}(h_{SR}) dh_p dh_{SR},
\]
where \( f_x(x) \) indicates the PDF of the random variable \( x \). Satisfying constraint (7) guarantees that the achievable-rate of the primary user is bigger than \( R_{min} \) for at least \( 1 - P_{out}^p \) percentage of time. We now simplify (7) assuming that the amount of the interference-power \( P(\theta, h_{SR}, h_{SP})h_{SP} \) should satisfy the inequality (7) at all times. Hence, we have
\[
P(\theta, h_{SR}, h_{SP})h_{SP} \leq I_{th},
\]
\[
P(\theta, h_{R,D}, h_{R,P})h_{R,P} \leq I_{th}, \quad i = 1, \ldots, K,
\]
where
\[
I_{th} = -\frac{\ln \left( 1 - P_{out}^p \right) + k_2}{k_1} - N_0B,
\]
which, hereafter, is referred to as interference-limit. It is worth noting that when \( I_{th} \leq 0 \), no feasible power allocation satisfying (8) and (9) exists, hence, the capacity lower bound is zero. In the following, we assume \( I_{th} > 0 \).

IV. EFFECTIVE CAPACITY

As stated earlier, effective capacity was originally defined in [5] as the dual concept of effective bandwidth. In this section, we first introduce this concept and then find the maximum effective capacity of a Rayleigh fading relay channel under spectrum-sharing constraints ((8) and (9)). Defining \( \{R[n], n = 1, 2, \ldots\} \) as the stochastic service process which is assumed to be stationary and ergodic and assuming that the function
\[
\Lambda(-\theta) = \lim_{N \to \infty} \frac{1}{N} \ln \left( \mathcal{E} \left\{ e^{-\theta \sum_{n=1}^N R[n]} \right\} \right), \tag{11}
\]
exists, the effective capacity is outlined as [5]
\[
E_c(\theta) = -\frac{\Lambda(-\theta)}{\theta} \\
= -\lim_{N \to \infty} \frac{1}{N\theta} \ln \left( \mathcal{E} \left\{ e^{-\theta \sum_{n=1}^N R[n]} \right\} \right). \tag{12}
\]
It is worth noting that the effective capacity, given in (12), shows the maximum arrival-rate that can be supported by the channel under the constraint of QoS exponent \( \theta \), interpreted as the delay constraint. Moreover, in block-fading channels, where the sequence \( R[n], n = 1, 2, \ldots \), is uncorrelated, the effective capacity can be simplified to
\[
E_c(\theta) = -\frac{1}{\theta} \ln \left( \mathcal{E} \left\{ e^{-\theta R[n]} \right\} \right). \tag{13}
\]

Having introduced the effective capacity formulations, in the following subsections, we obtain closed-form expressions for the effective capacity of secondary relaying networks in two cases: (i) when a single relay node is available, and (ii) when multiple relay nodes are implemented in the system.

A. Single-Relay Node

As stated earlier, in the first time slot the secondary’s transmitter uses the spectrum band to broadcast its signal to the relay nodes. Its transmission power, however, is limited to comply with the interference-power constraint (8). Furthermore, assuming that a single relay node \( R_i \) is available to relay the signal to the secondary’s transmitter, its power should also be limited so as to adhere to the interference requirement of the primary user. Hence, one can show that
\[
P(\theta, h_{SR}, h_{SP}) = \frac{I_{th}}{h_{SP}} \quad \text{and} \quad P(\theta, h_{R,D}, h_{R,P}) = \frac{I_{th}}{h_{R,P}},
\]
The data rate of the relay channel can now be found as
\[
R[n] = T_f B \min \left\{ \frac{1}{2} \ln \left( 1 + \frac{h_{SR}[n]}{h_{SP}[n] N_0B} \right), \frac{1}{2} \ln \left( 1 + \frac{h_{R,D}[n]}{h_{R,P}[n] N_0B} \right) \right\}. \tag{14}
\]
We now obtain closed-form expression for the effective capacity of the secondary relaying channel in Rayleigh flat-fading environment. We proceed by defining a new parameter
\[
z_i = \min \left\{ h_{SR}, \frac{h_{R,D}[n]}{h_{SP}} \right\}
\]
whose cumulative distribution function (CDF) can be found as follows:
\[
F_{z_i}(z_i) = \Pr \left( \min \left\{ h_{SR}, \frac{h_{R,D}[n]}{h_{SP}} \right\} \leq z_i \right) \tag{15}
\]
\[
= 1 - \frac{1}{(1 + z_i)^2}, \tag{16}
\]
where to derive (16) from (15) we use the fact that the distribution of the ratio between two Gamma distributed random variables, with parameters \( \alpha_1 \) and \( \alpha_2 \), is a beta prime distribution, with parameters \( \alpha_1 \) and \( \alpha_2 \) [17]. Specifically, the distribution of the ratio of two independent Rayleigh distributed random variables, e.g. \( v = \frac{h_{SR}}{h_{SP}} \), can be found according to \( f_v(v) = \frac{1}{(v+1)^2} \). Note that \( D \) in \( h_{R,D}[n] \) refers to the destination node, i.e, the secondary receiver. We proceed by using (16) to obtain the PDF of the random variable \( z_i \) as
\[
f_{z_i}(z_i) = \frac{2}{(1 + z_i)^3}. \]
The effective capacity of the secondary user channel can now be found as
\[
E_c(\theta) = -\frac{1}{\theta} \ln \left( \mathcal{E} \left\{ e^{-\theta z_i I_{th}} \frac{1}{N_0B} \ln \left( 1 + \frac{z_i I_{th}}{N_0B} \right) \right\} \right) \\
= -\frac{1}{\theta} \ln \left( \int_0^\infty \left( 1 + \frac{z_i I_{th}}{N_0B} \right)^{-\alpha} \frac{2}{(1 + z_i)^3} dz_i \right),
\]
where \( \alpha = \frac{\theta T_f B}{2} \). A closed-form solution for \( E_c(\theta) \) when \( I_{th} \geq \frac{N_0 B}{2} \) can then be obtained according to

\[
E_c(\theta) = -\frac{1}{\theta} \ln \left( \frac{2I_{th}^\alpha (N_0 B)^\alpha}{\alpha + 2} \times {}_2F_1 \left( \alpha, \alpha + 2, \alpha + 3; \frac{I_{th} - N_0 B}{I_{th}} \right) \right),
\]

where \( {}_2F_1(a, b; c; z) \) denotes the Gauss’s hypergeometric function [18]. The proof on how (17) is obtained is provided in Appendix A.

### B. Multiple-Relay Nodes

In this case, we assume that multiple intermediate nodes are available to relay the data from the secondary transmitter to the secondary receiver. The relay node that provides the highest rate is chosen to access the spectrum band in the second time-slot. As such, the rate of the system under consideration can be obtained according to

\[
R[n] = \max \{ R_i[n] \}, \quad i = 1, \ldots, K,
\]

where \( R_i[n] \) can be found from (14). In the following, we derive a closed-form expression for the effective capacity of the secondary relay channel in the mentioned configuration.

We define the random variable \( z \triangleq \max \{ z_1, \ldots, z_K \} \) and obtain its CDF as follows:

\[
F_z(z) = \Pr (\max \{ z_1, \ldots, z_K \} \leq z) = \prod_{i=1, \ldots, K} \Pr (z_i \leq z) = \sum_{k=0}^{K} \binom{K}{k} (-1)^k \frac{(1 + z)^{2k}}{2k},
\]

from which the PDF of \( z \) can be obtained as

\[
f_z(z) = \sum_{k=1}^{K} (-1)^{k+1} \binom{K}{k} \frac{2k}{(1 + z)^{2k+1}}.
\]

We now expand the integration in the effective capacity formula according to

\[
E_c(\theta) = -\frac{1}{\theta} \ln \left( \int_0^{\infty} \left( 1 + \frac{z}{N_0 B} \right)^{-\alpha} \times \sum_{k=1}^{K} (-1)^{k+1} \binom{K}{k} \frac{2k}{(1 + z)^{2k+1}} dz \right),
\]

whose closed-form expression can be obtained by following a similar approach as proposed in Appendix A, according to

\[
E_c(\theta) = -\frac{1}{\theta} \ln \left( \frac{I_{th}}{N_0 B} \sum_{k=1}^{K} (-1)^{k+1} \binom{K}{k} \frac{2k}{\alpha + 2k} \times {}_2F_1 \left( \alpha, \alpha + 2k, \alpha + 2k + 1; \frac{I_{th} - N_0 B}{I_{th}} \right) \right),
\]

### V. Numerical Results

In this section, we numerically illustrate the effective capacity of a relaying channel under interference-power constraints

at the primary’s receiver in Rayleigh fading environment. Hereafter, we assume \( N_0 B = 1 \) and \( T_f B = 1 \).

In Fig. 2, the normalized effective capacity for a Rayleigh block-fading relay channel under interference-power constraints, (8) and (9), is plotted versus delay QoS exponents, \( \theta \), for various interference-limit values and considering the single relay case or multiple relays (\( K = 2 \)). The effective capacity is shown in nats/s/Hz. In this figure, the steady and dashed lines refer to the single relay and multiple relay channels, respectively. We observe that the capacity increases as \( \theta \) decreases; the gain in the effective capacity, however, decreases for lower values of \( \theta \). The figure shows that in cases with loose QoS restrictions, i.e., lower values of \( \theta \), the capacity benefits significantly from implementing multiple relay nodes in the system whereas in systems with higher values of \( \theta \), e.g., \( \theta = 100 \) 1/bit, the capacity does not benefit much from the existence of multiple relay nodes.

We further examine the capacity gain achieved by implementing multiple relay nodes in the cognitive radio relay network. Fig. 3 shows plots for the normalized effective capacity of the secondary relay channel versus the number of relay nodes for various value of the QoS exponent, \( \theta \). The figure shows significant capacity gain as a result of increasing
the number of relay nodes when \( \theta = 0.1 \) bit. The slope in the effective capacity gain decreases as \( K \) increases, indicating hence that the gain in the effective capacity as a result of implementing more relay nodes is bigger for smaller values of \( K \).

\[ \text{VI. CONCLUSION} \]

In this paper, we considered a spectrum-sharing relay network in which a secondary transmitter communicates with its receiver through a relay node under a joint delay QoS constraint required for the secondary successful transmission and a constraint on the outage probability of the primary user. In particular, the primary user requires its transmission rate to be bigger than a minimum required threshold for at least a certain percentage of time. We obtained a maximum interference-power constraint which, when guaranteed by the secondary user, allows the primary user to be supported with its required rate. On the secondary network side, we considered a delay QoS constraint such that the delay should be limited by a predefined value for a certain percentage of time. The secondary user benefits from an intermediate relay node, chosen from \( K \) possible terminals. Considering the interference power limitations, we obtained the effective capacity of the secondary relay network when a single relay node is available and also when multiple relay nodes are implemented. We found closed-form expressions for the effective capacity of these systems in Rayleigh fading channels. Numerical results were also provided and showed that in systems with loose QoS constraints the capacity benefits tremendously by increasing the number of relaying terminals in the system.

\[ \text{APPENDIX A} \]

The integration in the effective capacity formula (17) can be expanded according to

\[ E_c(\theta) = \frac{1}{\theta} \ln \left( \int_0^\infty \left( 1 + \frac{z_i I_{th}}{N_0 B} \right)^{\alpha - 1} \frac{2}{(1 + z_i)^3} dz_i \right) \]

where \( I \) can be simplified by using the change of variable \( x = 1 + z_i \) according to

\[ I = \int_0^\infty \left( 1 - \frac{I_{th}}{N_0 B} + \frac{x I_{th}}{N_0 B} \right)^{-2} \frac{1}{x^3} dx. \]  

We further apply the change of variable \( y = \frac{1}{x} \) to (21), thus yielding

\[ I = 2 \left( \frac{I_{th}}{N_0 B} \right)^{-\alpha} \int_0^1 \left( 1 - \frac{I_{th} - N_0 B}{I_{th} y} \right)^{-\alpha} y^{\alpha+1} dy. \]

Then, making use of the following expression [18],

\[ _2F_1(a, b; c; u) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \times \int_0^1 t^{b-1}(1-t)^{-b+c-1}(1-t u)^{-a} dt, \]

for \( \text{Re}(c) > \text{Re}(b) > 0 \) and \( |u| < 1 \), where \( \Gamma(x) = \int_0^\infty t^{x-1}e^{-t} dt \) indicates the Gamma function, namely, by inserting (23) into (22) when setting \( b = \alpha + 2, c = \alpha + 3, a = \alpha \) and \( u = \frac{I_{th} - N_0 B}{I_{th}} \), we get

\[ I = 2 \left( \frac{I_{th}}{N_0 B} \right)^{-\alpha} \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + 3)} \times _2F_1 \left( \alpha, \alpha + 2, \alpha + 3; \frac{I_{th} - N_0 B}{I_{th}} \right) \]

\[ = 2F_1(\alpha, \alpha + 2, \alpha + 3; \frac{I_{th} - N_0 B}{I_{th}}), \]

where to derive (25) from (24), we use the equality \( (1 + z) = z \Gamma(z) \). This concludes the proof for (17).

Note that the condition \( |u| < 1 \) implies that \( I_{th} \geq \frac{N_0 B}{2} \). For \( N_0 B = 1 \), this limiting value corresponds to \( I_{th} \geq -3 \) dB.

\[ \text{REFERENCES} \]


