

Maximum Throughput of a Secondary User Cooperating with an Energy-Aware Primary User

Ahmed El Shafie[†], Ahmed Sultan^{*}, Tamer Khattab^{*}

[†]Wireless Intelligent Networks Center (WINC), Nile University, Giza, Egypt.

^{*}Department of Electrical Engineering, Alexandria University, Alexandria, Egypt.

^{*}Electrical Engineering, Qatar University, Doha, Qatar.

Abstract—This paper proposes a cooperation protocol between a secondary user (SU) and a primary user (PU) which dedicates a free frequency subband for the SU if cooperation results in energy saving. Time is slotted and users are equipped with buffers. Under the proposed protocol, the PU releases portion of its bandwidth for secondary transmission. Moreover, it assigns a portion of the time slot duration for the SU to relay primary packets and achieve a higher successful packet reception probability at the primary receiver. We assume that the PU has three states: idle, forward, and retransmission states. At each of these states, the SU accesses the channel with adaptive transmission parameters. The PU cooperates with the SU if and only if the achievable average number of transmitted primary packets per joule is higher than the number of transmitted packets per joule when it operates alone. The numerical results show the beneficial gains of the proposed cooperative cognitive protocol.

Index Terms—Cognitive radio, energy-aware, closure, Markov chain, stability analysis.

I. INTRODUCTION

Cognitive usage of electromagnetic spectrum has been identified as a useful means for efficient utilization of the scarce radio resources. Under such paradigm, the spectrum owned by the primary/licensed system is opportunistically accessed by the secondary/unlicensed system using cognitive radio technology.

Energy efficient communications has received significant attention recently [1]–[5]. The problem of energy-efficient design was investigated from physical layer standpoint [2], medium-access-control (MAC) layer standpoint [3], and cross-layer standpoint [4]. Cooperative communication is a method for achieving energy-efficient data transmission. Cooperative terminals act as spatially distributed antennas to provide alternative paths of the signal to the destinations. This may lead to significant reduction in transmit power and a better use of communication resources.

Combining cognitive radio technology and cooperative communications has been investigated in many papers, e.g. [6]–[9]. In [6] the secondary transmitter is used as a relay for the undelivered packets of the primary transmitter. The secondary user (SU) optimizes its power to expand the queueing stability region of the network. In [7], the authors assumed that the cognitive transmitter is allowed to use the channel whenever the primary queue is empty. The SU cannot access the channel

with its own data packets unless the relaying queue becomes empty. The secondary relays a certain fraction of the primary undelivered packets to minimize the secondary packets average delay subject to a power budget for the relayed primary packets. The maximum stable throughput of a secondary terminal sharing the channel with a primary terminal under multi-packet reception channel model was characterized in [8]. Based on the proposed scheme, the SU transmits its packets with some access probability based on the primary state. In particular, when the primary user (PU) is inactive, the SU immediately accesses the channel. When the PU is active, the SU probabilistically accesses the channel. A new cooperation protocol was proposed in [9]. At each time slot, the PU releases portion of its bandwidth for the cognitive radio (CR) user and half of its time slot duration. During the first half of the time slot, the CR user receives the PU data. Afterwards, it amplifies-and-forwards the received packet during the second half of the time slot. The primary and secondary transmitters are assumed to be bufferless and to have a complete knowledge of the instantaneous transmit channel state information (CSI).

In this paper, we consider a cognitive network with an energy-aware primary terminal and a secondary terminal. We propose the following cooperation protocol between the PU and the SU. The PU releases portion of its bandwidth for secondary transmission. Moreover, it assigns a portion of the time slot duration for the SU to relay primary packets and achieve a higher successful packet reception probability at the primary receiver. The SU spends portion of its energy/power for aiding the PU to achieve certain quality of service requirements characterized by primary queue stability and higher number of transmitted primary packets per unit energy. The SU may leverage the primary feedback signal broadcasted in a time slot to ascertain the primary state in the following time slot. The secondary transmission time and bandwidth in a time slot change according to the state of the PU. We do not assume the availability of the transmit CSI at the transmitters.

The contributions of this paper can be summarized as follows.

- We propose a new cooperative cognitive relaying protocol for buffered primary and secondary terminals. The proposed protocol allows bandwidth and time sharing between the primary and the secondary users. Furthermore, the proposed protocol leverages the primary feedback channel.
- We consider an energy-aware PU and investigate the

This paper was made possible by a NPRP grant 09-1168-2-455 from the Qatar National Research Fund (a member of The Qatar Foundation). The statements made herein are solely the responsibility of the authors.

average number of transmitted primary packets per unit energy.

- In contrast to most of the existing works, we assume two quality of service requirements for the PU. Specifically, we put a constraint on the required number of transmitted primary packets per time slot, and another constraint on the primary queue stability. Violating any of these constraints obviates cooperation between the users.
- We derive closed-form expressions for the service rates of queues and the average number of transmitted primary packets per joule without and with cooperation.
- We characterize the maximum secondary throughput of the proposed protocol under the aforementioned constraints.

The remainder of this paper is organized as follows. In the next section, we describe the system model adopted in this paper. The description of the proposed cooperative protocol is found in Section III. In Section IV, we discuss the stable-throughput region of the proposed protocol. Numerical results and conclusions are provided in Section V.

II. SYSTEM MODEL

We consider a cognitive network with one PU and one SU. This can be seen as a part of a larger network with multiple primary bands operating under frequency division multiple-access. Each band is composed of one PU and one SU. For simplicity of presentation, we provide the analysis of one of the available primary bands. The SU is assumed to be equipped with \mathcal{M} antennas. The PU is assumed to be an energy-aware terminal, which operates in a time-slotted fashion with slot duration of T seconds and a total bandwidth of W Hz.

A. MAC Layer

Each user is equipped with an infinite capacity buffer for storing its incoming traffic denoted by Q_n , where n is ‘p’ for primary and ‘s’ for secondary. The SU has an additional finite capacity queue for storing the relaying packet, denoted by Q_{ps} . Under the proposed protocol, the relaying queue maintains one packet at most, as explained later. We consider time-slotted transmissions where all packets have the same size of b bits. The arrivals at Q_n , $n \in \{p, s\}$, are assumed to be independent and identically distributed Bernoulli random variables from slot to slot with mean arrival rate $\lambda_n \in [0, 1]$ packets per time slot. The arrivals are also mutually independent from terminal to terminal. For similar network model and queue assumptions, the reader is referred to [6], [7].

The retransmission mechanism is based on the feedback acknowledgement/negative-acknowledgement (ACK/NACK) messages. More specifically, at the end of the time slot, each destination sends a feedback message to inform the respective transmitter about the decodability status of the transmitted packet. If a packet is received correctly at the respective destination, an ACK is fed back to the respective transmitter. On the other hand, if a packet is received erroneously at the respective receiver, a NACK message is fed back to the respective transmitter. We assume that all nodes in the system can hear the feedback ACK/NACK messages. Therefore, the primary feedback messages are overheard by the SU. The

overhead for transmitting the ACK and NACK messages is assumed to be very small compared to packet sizes. Furthermore, the errors in packet acknowledgement feedback are negligible [10], which is reasonable for short length ACK/NACK packets as low rate strong codes can be employed in the feedback channel.

The PU is assumed to have three states: idle, forward, and retransmission. The PU is said to be ‘idle’, when its queue is empty. In the ‘forward’ states, the PU sends the packet at the head of its queue, *for the first time*, while in the ‘retransmission’ states, the PU transmits packets that have been erroneously decoded at the primary receiver. The primary feedback is either “ACK”, if the primary destination decodes the primary packet correctly; “NACK”, if the primary decoder fails in decoding the packet; or “nothing”, if there is no primary transmission.

Without cooperation, the SU does not gain any access to the primary spectrum. This assumption is motivated by the fact that the primary system owns the spectrum. For the SU to gain access to the spectrum, it should provide an economic incentive or aid in enhancing the performance of the PU [9], [11], [12]. In this work, we consider performance enhancement incentives.

The PU cooperates with the SU if and only if cooperation yields more average number of transmitted packets per joule than when it operates alone. If cooperation is beneficial for the PU, it releases portion of its time slot duration and bandwidth for the SU. The portions of bandwidth and time slot released by the PU differ from state to state. In particular, if the primary queue is empty, the PU remains idle and the SU transmits its own packet over the whole channel bandwidth, W . Thus, the released bandwidth for the SU in idle states is W Hz. If the primary queue is in a forward state, the released bandwidth and time slot duration are $W_s = W - W_p$ Hz and $T_{s,F} = T - T_{p,F}$ seconds, respectively. If the primary queue is in a retransmission state, the released bandwidth and time slot duration are $W_s = W - W_p$ Hz and $T_{s,R} = T - T_{p,R}$ seconds, respectively. In both cases, the SU uses the released bandwidth for its own data packets transmission. In addition, the SU assists the primary system in delivering its packet to the primary receiver using the assigned time and the remainder of the bandwidth, W_p , which is used for both primary transmission and secondary transmission of the primary packet. The immediate benefits of the proposed cooperation protocol for the PU are transmission time reduction and transmission energy savings per time slot.

B. Physical (PHY) Layer

Wireless links exhibit fading and are corrupted by additive white Gaussian noise (AWGN). The fading is assumed to be stationary, with frequency non-selective Rayleigh block fading. This means that the fading coefficient of a certain link remains constant during one slot. We do not assume the availability of the CSI at the transmitting terminals. The channel is assumed to be known perfectly only at the receivers. The AWGN at each receiving node is assumed to have zero mean and power spectral density \mathcal{N}_0 Watts per unit frequency.

Denote the primary transmitter as ‘p’, the primary destination as ‘pd’, the ν th antenna of the secondary transmitter

as ‘ s_ν ’, $\nu \in \{1, 2, \dots, \mathcal{M}\}$, and the secondary destination as ‘sd’. Let H denote the state of the primary terminal, where H is ‘o’ for idle state, ‘F’ for forward state, and finally, ‘R’ for retransmission state. The probability of channel outage of the link between node j and node ρ (link $j \rightarrow \rho$), for a given state $H \in \{o, F, R\}$, is given by $\mathbb{P}_{j,\rho}^H$. This outage probability is a function of the number of bits in a data packet, the slot duration, the transmission bandwidth, the transmit powers, number of receive and transmit antennas, the average channel gains, and the state of the PU as detailed in Appendix A.

We assume one secondary antenna is used for transmission, and *all* antennas are used for channel sensing and data reception. Intuitively, increasing the expected link gain, $\sigma_{j,\rho}$ for the link between j and ρ , decreases the outage probability of that link and, consequently, enhances queues’ service rates. This is clear from the outage probability expression (19) in Appendix A. Based on this, the SU transmits to the primary or secondary receivers with the antennas having the highest expected link gain to the respective receivers. Let us assume, without loss of generality, that the link between the i th secondary antenna and the secondary destination has the highest expected link gain among all the links between the SU’s antennas and the secondary destination, and the link between the m th secondary antenna and the primary destination has the highest expected link gain among all the links between the SU’s antennas and the primary destination. The SU transmits a relaying packet using the m th antenna, and transmits its own packet using the i th antenna. Note that there possibly concurrent transmissions occur over orthogonal frequency subbands.

For proper operation of the proposed protocol, we assume that the minimum $W_p T_{p,F}$ is \mathcal{E} . This value is required for almost unity successful decoding of the primary packet at the SU terminal under the availability of \mathcal{M} antennas and for certain channels parameters. In particular, \mathcal{E} is designed such that the required outage probability of the link $p \rightarrow s$ is at most \mathcal{Q} , where $\mathcal{Q} \ll 1$. Using the expression of probability of channel outage of link $p \rightarrow s$ in Appendix A, we can find an expression for \mathcal{E} . Assume that P^j is the transmit power per unit frequency used by node j . Letting $\sigma_{p,s_\nu} = \sigma_{p,s} \forall \nu$ and noting that $\gamma_{p,s_\nu} = P^p / \mathcal{N}_o = \gamma_{p,s}$ (equal for all receiving antennas), the outage probability (22) in Appendix A can be upper bounded as follows:

$$\mathbb{P}_{p,s}^F \leq \left[1 - \exp\left(-\frac{2^{\frac{b}{W_p T_{p,F}}} - 1}{\sigma_{p,s} \gamma_{p,s}}\right) \right]^{\mathcal{M}} \leq \mathcal{Q} \quad (1)$$

where the first inequality holds strictly to equality when the SU decodes separately the received copies of the primary packet at each antenna. After some mathematical manipulations, we get

$$W_p T_{p,F} \geq \frac{b}{\log_2 \left[1 - \sigma_{p,s} \gamma_{p,s} \ln(1 - \mathcal{Q}^{\frac{1}{\mathcal{M}}}) \right]} = \mathcal{E} \quad (2)$$

From the expression of \mathcal{E} , increasing the number of antennas reduces the required \mathcal{E} . Furthermore, as the received SNR, $\sigma_{p,s} \gamma_{p,s}$, increases, the signal quality increases, and \mathcal{E} needed for a negligible decoding failure of the primary packet at the

SU decreases as well. Finally, increasing the packet size, b , increases the required \mathcal{E} .

III. PROPOSED COOPERATIVE COGNITIVE PROTOCOL

One of the most important functions for the SU is to ascertain the primary state at each time slot. The SU can discern the state of the PU via observing the primary feedback channel and/or channel sensing. In particular, if there is no primary feedback at the end of the previous time slot due to PU’s inactivity, or if the SU hears an ACK, the SU senses the channel to discern the current state of activity of the PU. If the previous primary feedback is a NACK feedback, the SU does not sense the channel because it knows with certainty that the PU will be active in the current time slot. It is worth pointing out here that since the SU’s operation is based on the channel sensing outcomes, the time assigned for channel sensing, τ , is less than or equal to the primary transmission $T_{p,F}$. That is, the domain of $T_{p,F}$ is $[\tau, T]$.

Let us assume without loss of generality that the primary feedback received at the end of the previous time slot was not a NACK. In the current time slot, the SU must sense the channel for τ seconds to discern the state of activity of the PU. When the primary queue is empty, the SU transmits the packet at the head of its own traffic queue, Q_s , with transmission rate $b/(T - \tau)$ and transmission bandwidth W Hz. When the primary queue is nonempty, i.e., the PU is in the forward state, it releases $W_s = W - W_p$ for the SU to be used for its own data transmission. Hence, the SU’s transmission rate is still $b/(T - \tau)$, but the transmission bandwidth becomes $W_s \leq W$ Hz. The PU and the SU use the remainder of the bandwidth, $W_p = W - W_s$, for the transmission of the primary packet over two non-overlapped time intervals. In particular, the PU sends its data packet over the time interval $[0, T_{p,F}]$, at the same time, the SU attempts to decode the packet. If the SU is able to decode the primary packet, the packet is added to the relaying queue, Q_{ps} . As mentioned earlier, we assume here that the SU always correctly decodes primary transmission by exploiting the spatial diversity provided by its multiple antennas. Over the remainder of the time slot, i.e., over the time interval $[T_{p,F}, T]$, the SU retransmits the primary packet to the primary receiver. At the end of the time slot, the primary destination decodes separately the received primary packet versions. If the primary destination decodes the primary packet successfully, the primary destination sends an ACK to inform the transmitters about the successful decoding of the primary packet. The primary packet is then dropped from both the primary and the relaying queues.

If the primary destination fails in decoding the primary packet at the first transmission, a NACK is sent by the primary destination to inform both the PU and the SU about the decoding failure of the primary packet. The retransmission of the primary packet starts at the following time slots until an ACK is issued by the primary destination. Since the SU does not sense the channel, and it knows with certainty that the PU is active, the SU transmits its own data packet over whole slot duration. Moreover, since the SU already has the primary packet in its buffer from the previous time slot, it does not have to receive/decode the primary packet again.

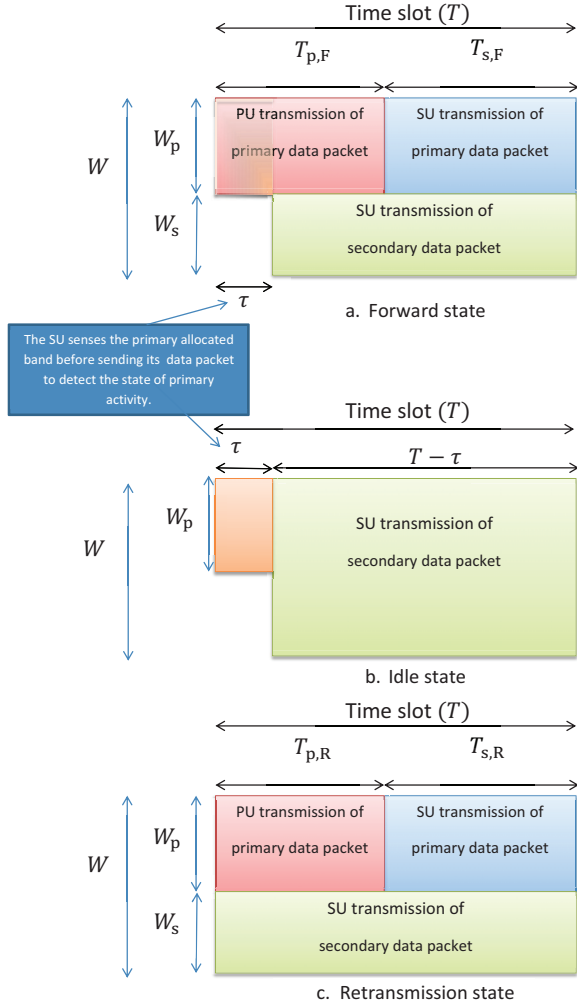


Fig. 1. System operation and resource sharing at each of the primary states. In the figure, $T_{s,F} = T - T_{p,F}$ and $T_{s,R} = T - T_{p,R}$.

During the retransmission states, the PU and the SU split the time slot into two portions as in the forward states such that the primary transmission takes place over $[0, T_{p,R}]$, while the secondary transmission of the primary packet takes place over $[T_{p,R}, T]$. The system operation at each of the primary states is shown in Fig. 1. Note that since there is no channel sensing in retransmission states, the primary transmission duration can take any value from 0 to T . Hence, the domain of $T_{p,R}$ is $[0, T]$.

It is worth noting that the proposed protocol requires time-synchronization of the cognitive radio system to the primary system. That can be ensured via a beacon channel [13].

We assume that the misdetection and false-alarm probabilities are negligible (the same assumption is found in [7], [13]). This means that the sensing duration τ is large enough to gather enough statistically independent samples using the SU's multiple antennas such that the error probabilities become almost equal to zero.

IV. STABILITY ANALYSIS

We adopt a late arrival model where new arrivals will not be served at the arriving slot even if the queue is empty

[14]. Denote by A_ζ^t the number of arrivals to queue Q_ζ , $\zeta \in \{p, s, ps\}$, at time slot t , and D_ζ^t the number of departures from queue Q_ζ at time slot t . The queue length evolves according to the following form:

$$Q_\zeta^{t+1} = (Q_\zeta^t - D_\zeta^t)^+ + A_\zeta^t \quad (3)$$

where $(z)^+$ denotes $\max(z, 0)$. A fundamental performance measure of a communication network is the stability of its queues. Stability can be defined as follows [14].

Definition: Queue Q_ζ^t , $\zeta \in \{p, s, ps\}$, is stable, if

$$\lim_{y \rightarrow \infty} \lim_{t \rightarrow \infty} \Pr\{Q_\zeta^t < y\} = 1 \quad (4)$$

If the arrival and service processes are strictly stationary, then we can apply Loynes theorem to check for stability conditions [15]. This theorem states that if the arrival process and the service process of a queue are strictly stationary processes, and the average service rate is greater than the average arrival rate of the queue, then the queue is stable. If the mean service rate is lower than the mean arrival rate, then the queue is unstable [14].

A. Non-cooperative Users

As mentioned earlier, without cooperation, the SU cannot access the channel and the PU transmits a lone at the beginning of the time slot and over the whole slot duration if its queue is nonempty. The primary queue mean service rate is equal to the complement of the channel outage between the PU and its respective receiver. Since the PU is an energy-aware terminal, at any given primary arrival rate, it maximizes its number of transmitted packets/joule. The PU transmits b bits over T seconds with transmission bandwidth $\mathbb{W} < W$, where \mathbb{W} is assumed to be an optimizable parameter. The probability of successful primary transmission (i.e., the mean service rate) is given by

$$\mu_{p,nc} = \overline{\mathbb{P}}_{p,pd} = \exp\left(-\frac{2^{\frac{b}{\mathbb{W}T}} - 1}{\gamma_{p,pd}\sigma_{p,pd}}\right) \quad (5)$$

where $\overline{\mathcal{X}} = 1 - \mathcal{X}$. The probability of successful primary transmission is maximized when \mathbb{W} is equal to W ; hence, the maximum primary stable throughput is $\mu_{p,nc}^{\max} = \exp\left(-\frac{2^{\frac{b}{WT}} - 1}{\gamma_{p,pd}\sigma_{p,pd}}\right)$.

When the PU's queue is nonempty, the packet at the head of the queue is successfully received at the primary destination with probability $\mu_{p,nc}$. Hence, the average number of transmitted packets per slot is $\mu_{p,nc}\Pr\{Q_p \neq 0\}$ packets/slot. Since the PU transmits with power P^p Watts/Hz over bandwidth \mathbb{W} and with transmission time T seconds, the average energy per slot is $P^p\mathbb{W}T$ joule/slot. Hence, the average number of transmitted primary packets per joule, $B_{p,nc}$, is given by

$$B_{p,nc} = \frac{1}{P^pT\mathbb{W}}\mu_{p,nc}\Pr\{Q_p \neq 0\} \quad (6)$$

If the primary queue is stable, i.e., $\mu_{p,nc} > \lambda_p$, the probability of the primary queue being nonempty is given by $\lambda_p/\mu_{p,nc}$. Substituting with $\Pr\{Q_p \neq 0\} = \lambda_p/\mu_{p,nc}$ into (6), the number

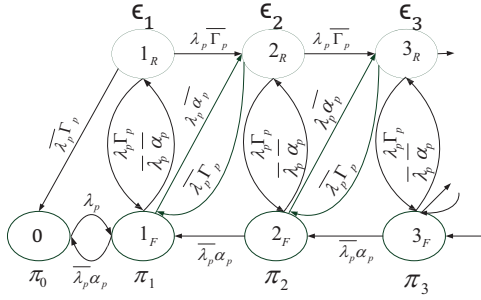


Fig. 2. Markov chain of the PU for the proposed protocol. State self-transitions are not depicted for visual clarity. Note that $\bar{\mathcal{X}} = 1 - \mathcal{X}$.

of transmitted primary packets per joule when the PU operates lonely is given by

$$B_{p,nc} = \frac{\lambda_p}{P_p T \mathbb{W}} \quad (7)$$

The maximum average number of transmitted primary packets per joule is obtained via solving the following optimization problem:

$$\max_{\mathbb{W}} B_{p,nc} = \frac{\lambda_p}{P_p T \mathbb{W}}, \text{ s.t. } \lambda_p \leq \mu_{p,nc}, \mathbb{W} \in [0, W] \quad (8)$$

The optimization problem can be converted to the following linear program:

$$\min_{\mathbb{W} \in [0, W]} \mathbb{W}, \text{ s.t. } \mathbb{W} \geq \frac{b}{T \log_2 (1 - \gamma_{p,pd} \sigma_{p,pd} \ln(\lambda_p))} = \mathbb{W}_{\min} \quad (9)$$

It is straightforward to show that the optimal primary transmission bandwidth and the maximum average number of transmitted primary packets per joule are given by

$$\mathbb{W}^* = \mathbb{W}_{\min}, B_{p,nc}^{\max} = \frac{\lambda_p}{P_p T \mathbb{W}^*} \quad (10)$$

with $\lambda_p \leq \mu_{p,nc}^{\max}$. From the optimal solution, we note that \mathbb{W}_{\min} increases with λ_p . This is expected as the primary queue starts to saturate and the needed bandwidth expands to increase the successful probability of the primary transmissions, which in turn maintain the stability of the primary queue. Furthermore, for a given λ_p , the bandwidth used for transmission decreases with $\gamma_{p,pd} \sigma_{p,pd}$ and T/b . This is because μ_p must be at least equal to λ_p for maintaining the primary queue stability.

B. Cooperative Users

1) *Markov Chain Analysis of The PU's Queue*: The Markov chain representing the PU's queue when the PU cooperates with the SU is shown in Fig. 2. The probability of the primary queue having k packets and transmitting for the first time is π_k , where F in Fig. 2 denotes first transmission. The probability of the primary queue having k packets and retransmitting is ϵ_k , where R in Fig. 2 denotes retransmission. Let α_p denote the probability of primary successful transmission in case of first transmission, and Γ_p denote the probability of successful transmission in case of retransmission. Recall that the m th antenna is the antenna element within secondary array that has the maximum average link gain to the primary destination.

TABLE I
STATE PROBABILITIES FOR THE PU'S MARKOV CHAIN.

Parameter	Value	Parameter	Value
η	$\lambda_p \alpha_p + \bar{\lambda}_p \Gamma_p$	π_0	$\frac{\eta - \lambda_p}{\Gamma_p}$
		ψ	$\bar{\lambda}_p \eta$
π_1	$\pi_0 \frac{\lambda_p}{\psi} (\lambda_p + \bar{\lambda}_p \Gamma_p)$	ϵ_1	$\pi_0 \frac{\lambda_p \alpha_p}{\eta}$
$\pi_k, k \geq 2$	$\pi_0 \frac{\lambda_p \alpha_p}{\eta^2} \left[\frac{\lambda_p \eta}{\psi} \right]^k$	$\epsilon_k, k \geq 2$	$\pi_0 \frac{\lambda_p \alpha_p}{\eta^2} \left[\frac{\lambda_p \eta}{\psi} \right]^k$
$\sum_{k=1}^{\infty} \pi_k$	$\pi_0 \frac{\lambda_p \Gamma_p}{\eta - \lambda_p} = \lambda_p$	$\sum_{k=1}^{\infty} \epsilon_k$	$\pi_0 \frac{\lambda_p}{\eta - \lambda_p} \bar{\alpha}_p = \frac{\lambda_p}{\Gamma_p} \bar{\alpha}_p$

Since the primary receiver decodes separately the received copies of the primary packet from each user, the probability of primary packet correct reception in forward or retransmission states is equal to one minus the probability of both $p \rightarrow pd$ and $s_m \rightarrow pd$ being in outage simultaneously. Hence, α_p and Γ_p are given by

$$\alpha_p = 1 - \mathbb{P}_{p,pd}^F \mathbb{P}_{s_m,pd}^F, \Gamma_p = 1 - \mathbb{P}_{p,pd}^R \mathbb{P}_{s_m,pd}^R \quad (11)$$

where $\mathbb{P}_{p,pd}^F$, $\mathbb{P}_{s_m,pd}^F$, $\mathbb{P}_{p,pd}^R$, and $\mathbb{P}_{s_m,pd}^R$ follow (19) with $H \in \{F, R\}$, $r_{s_m}^H = b/T_{s,H}$, $r_p^H = b/T_{p,H}$ and $W_{s_m,pd} = W_p$.

Solving the state balance equations, we can obtain the state probabilities which are provided in Table I. The detailed solution is omitted here due to space limitations. The stability of the PU's queue is attained when the probability of its queue being empty is greater than zero. This can be guaranteed when $\lambda_p < \eta$, where η is defined in Table I.

2) *Secondary Throughput and Average Number of transmitted Primary Packets per joule*: The packet at the head of Q_s is served if, for a given primary state, the channel between the i th antenna¹ and the secondary receiver is not in outage. The average service rate of the secondary queue, denoted by μ_s , is then given by

$$\begin{aligned} \mu_s &= \pi_0 \overline{\mathbb{P}_{s_i,sd}^0} + \left(\sum_{k \geq 1} \pi_k \right) \overline{\mathbb{P}_{s_i,sd}^F} + \left(\sum_{k \geq 1} \epsilon_k \right) \overline{\mathbb{P}_{s_i,sd}^R} \\ &= \pi_0 \overline{\mathbb{P}_{s_i,sd}^0} + \lambda_p \overline{\mathbb{P}_{s_i,sd}^F} + \frac{\lambda_p \bar{\alpha}_p}{\Gamma_p} \overline{\mathbb{P}_{s_i,sd}^R} \end{aligned} \quad (12)$$

where the probabilities π_0 , $\sum_{k \geq 1} \pi_k$ and $\sum_{k \geq 1} \epsilon_k$ are given in Table I.

Under the proposed protocol, the average number of transmitted primary packets per joule, $B_{p,c}$, is given by

$$B_{p,c} = \begin{cases} \frac{\alpha_p (\sum_{k \geq 1} \pi_k)}{P_p W_p T_{p,F}} + \frac{\Gamma_p (\sum_{k \geq 1} \epsilon_k)}{P_p W_p T_{p,R}} = \frac{\alpha_p}{W_p P_p} + \frac{\bar{\alpha}_p}{T_{p,R}} & \text{if } T_{p,R} > 0; \\ \frac{\alpha_p (\sum_{k \geq 1} \pi_k)}{P_p W_p T_{p,F}} = \frac{\alpha_p \lambda_p}{P_p W_p T_{p,F}} & \text{if } T_{p,R} = 0. \end{cases} \quad (13)$$

The expression in (13) is explained as follows. First, we note that $T_{p,F}$ is at least equal to $\min\{\tau, \mathcal{E}/W_p\} > 0$, whereas $T_{p,R}$ can be assigned any value between 0 and T . If $T_{p,R} = 0$, this means that the PU does not spend any energy for packets retransmissions, and the responsibility of delivering

¹Recall that the i th antenna is the antenna element within secondary array that has the maximum average link gain to the secondary destination.

the primary packets at retransmission states is entirely given to the SU. Consequently, the value of $T_{p,R}$ partitions $B_{p,c}$ into two sets. Given that the probability of the PU being in a forward state is $\sum_{k \geq 1} \pi_k$, the packet at the head of the primary queue is received successfully at the primary destination with probability α_p . Hence, the number of successfully transmitted primary packets per slot is $(\sum_{k \geq 1} \pi_k) \alpha_p$ packets/slot. Since the PU transmits with power P^p Watts/Hz over bandwidth W_p Hz and with transmission time $T_{p,F}$ seconds, the average primary transmit energy per slot is $P^p W_p T_{p,F}$ joule/slot.

In a similar fashion, if the PU is in a retransmission state, given that the probability of the PU being in a retransmission state is $\sum_{k \geq 1} \epsilon_k$, the average number of successfully transmitted primary packets per slot is $(\sum_{k \geq 1} \epsilon_k) \Gamma_p$ packets/slot, and the average primary transmit energy per slot is $P^p W_p T_{p,R}$ joule/slot. If $T_{p,R} = 0$, the PU does not spend any energy for its packet delivery during retransmission states and the average primary transmit energy per slot is zero; hence, $B_{p,c}$ is given by the value explained in the forward states. Specifically, $B_{p,c} = \frac{\alpha_p (\sum_{k \geq 1} \pi_k)}{W_p P^p T_{p,F}}$ packets/joule for $T_{p,R} = 0$.

3) *Problem Formulation:* For cooperation to be established, $B_{p,c}$ must be strictly greater than $B_{p,nc}^{\max}$, i.e., $B_{p,c} > B_{p,nc}^{\max}$. We characterize the maximum secondary throughput under such quality of service condition. The maximum secondary throughput of the proposed protocol is characterized by the closure of the rate pairs (λ_p, μ_s) . To obtain this closure, we maximize the mean service rate of the SU under constraints on the stability of system's queues², $B_{p,c} > B_{p,nc}^{\max}$, and $W_p T_{p,F} \geq \mathcal{E}$. The optimization problem is stated as:

$$\begin{aligned} & \max_{W_p, T_{p,F}, T_{p,R}} \mu_s, \\ & \text{s.t. } \lambda_p \leq \eta, B_{p,c} > B_{p,nc}^{\max}, W_p T_{p,F} \geq \mathcal{E} \\ & W_p \in [0, W], T_{p,F} \in [\tau, T], T_{p,R} \in [0, T] \end{aligned} \quad (14)$$

The optimization problem (14) is solved numerically.³ In particular, we make a grid search over the optimization parameters. It should be pointed out here that the impact of the number of antennas, \mathcal{M} , affects the feasible set of (14) as it changes the value of \mathcal{E} . As is shown in Appendix B, increasing \mathcal{M} decreases \mathcal{E} which in turn may reduce the required value of bandwidth assigned for primary packet transmission. Consequently, the bandwidth assigned for the SU when the PU is active may increase, which boosts the secondary throughput.

If the distance between the PU and its respective receiver is long or the direct link is in deep shadowing due to surrounding physical obstacles, the primary direct link will be disconnected [16], [17]. This means that $\sigma_{p,pd} \approx 0$. In this case, the time assigned to the PU for transmission must be minimized. This is because there is no beneficial gain for the PU to use most of the time slot for delivering the packet to its receiver. Instead assigning most of the time slot for the SU to relay

²The stability of the relaying queue is obvious as it contains at most one packet.

³The optimal parameters obtained via solving the optimization problem (14) are announced to both users so that W_p , $T_{p,F}$ and $T_{p,R}$ are known at the PU and the SU before actual operation.

that packet may be more beneficial if it has relatively better channel quality to the primary destination. Thus, the optimal parameters are given by

$$T_{p,F} = \max\left\{\frac{\mathcal{E}}{W_p}, \tau\right\}, T_{p,R} = 0 \quad (15)$$

The problem reduces to an optimization problem with single optimization parameter, W_p , and with one constraint, $\lambda_p \leq \eta$. In this case, the successful probabilities of the primary packets at forward and retransmission states are

$$\alpha_p = \exp\left(-\frac{2 \frac{b}{W_p (T - \max\{\frac{\mathcal{E}}{W_p}, \tau\})} - 1}{\gamma_{s,pd} \sigma_{s,pd}}\right), \Gamma_p = \exp\left(-\frac{2 \frac{b}{W_p T} - 1}{\gamma_{s,pd} \sigma_{s,pd}}\right) \quad (16)$$

The feasible set of the optimization problem is given by

$$\lambda_p \leq \eta \iff \frac{\lambda_p}{\lambda_p} \leq \frac{\Gamma_p}{1 - \alpha_p} \quad (17)$$

The optimization problem becomes:

$$\begin{aligned} & \max_{W_p \in [0, W]} \mu_s, \\ & \text{s.t. } \frac{\lambda_p}{\lambda_p} \leq \frac{\Gamma_p}{1 - \alpha_p} \end{aligned} \quad (18)$$

This optimization problem is solved via a simple grid search over W_p .

V. NUMERICAL RESULTS AND CONCLUSIONS

In this section, we provide numerical results to illustrate the gains of the proposed protocol. For sake of simplicity, we denote $\sigma_{s_i, sd} = \sigma_{s_i, sd}$ and $\sigma_{s_m, sd} = \sigma_{s, pd}$. The common parameters used to generate the figures are: $\mathcal{Q} = 10^{-8}$, $b = 2000$ bits, $W = 10$ MHz, $T = 4 \times 10^{-4}$ seconds, $\mathcal{N}_o = 10^{-11}$ Watts/Hz, $P^p = 10^{-10}$ Watts/Hz, $\sigma_{s, sd} = 0.1$, $\sigma_{p, pd} = 0.2$, $\sigma_{s, pd} = 0.5$, $\sigma_{p, s\nu} = \sigma_{p, s} = 1$ for all ν , and $\tau = 0.2T$.

In Fig. 3, we present the average number of transmitted primary packets per joule versus λ_p for the proposed protocol. The non-cooperation case is also plotted for comparison. The figure reveals that the average number of transmitted primary packets per joule of the proposed protocol is higher than the maximum average number of transmitted primary packets per joule of the non-cooperative case over most of the λ_p range. More specifically, over $\lambda_p < 0.475$ packets/slot, the proposed protocol outperforms the non-cooperative case. For $\lambda_p \geq 0.475$ packets/slot, $B_{p,c} \leq B_{p,nc}^{\max}$; hence, the cooperation becomes non-beneficial for the PU and it ceases to use the SU for relaying. In this case, the optimization problem of the proposed protocol becomes infeasible. Hence, the SU does not gain any access to the channel. The parameters used to generate this figure are the common parameters, $\mathcal{M} = 6$, and $P^s = 5 \times 10^{-11}$ Watts/Hz.

Fig. 4 demonstrates the impact of the number of secondary antennas on the maximum secondary stable throughput. Under the used parameters, for $\mathcal{M} < 6$, the SU cannot access the channel because the condition $W_p T_{p,F} \geq \mathcal{E}$, which guarantees unity probability of primary packet decoding at the SU, is not satisfied for the feasible range of W_p and $T_{p,F}$. The figure also demonstrates the gains of increasing the number of secondary antennas which boosts the secondary throughput.

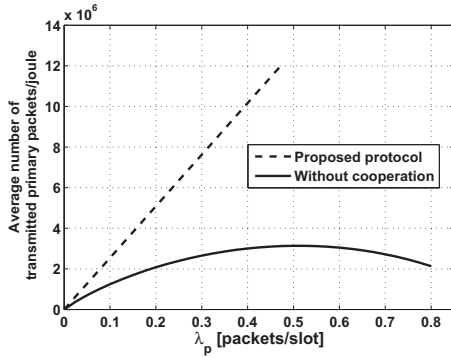


Fig. 3. The primary average packets per joule versus λ_p .

The secondary throughput increases with \mathcal{M} because as the number of antennas increases, the assigned bandwidth for the PU may be decreased due to the fact that $\mathcal{E} \leq W_p T_{p,F}$ is monotonically decreasing with \mathcal{M} . Hence, the assigned bandwidth for the SU when the PU is active, $W - W_p$, may increase, which in turn increases the probability of secondary packet successful decoding and the secondary throughput. The figure is generated using the common parameters, $\mathcal{M} = 7$ and $P^s = 10^{-10}$ Watts/Hz. Fig. 5 shows the number of primary transmitted packets per joule with and without cooperation. The figure is generated with the same parameters used to generate Fig. 4 and with $\mathcal{M} = 7$ antennas. From the figure, the gain of the proposed protocol over the noncooperation case is obvious. When $\lambda_p = 0.7$ packets/slot, the gain of the proposed protocol over the noncooperation case is almost 765%.

Fig. 6 shows the maximum secondary throughput of the proposed system when the highest mean among the channels between the SU's antennas and the primary destination, $\sigma_{s_m,p,d}$, varies. It can be noticed that the feasible range of the primary arrivals increases with $\sigma_{s_m,p,d}$. Fig. 7 presents the maximum secondary throughput of the proposed system when the highest mean among the channels between the SU's antennas and its respective destination, $\sigma_{s_i,s,d}$, varies. It can be noticed that the feasible range of the secondary arrivals increases with $\sigma_{s_i,p,d}$. From Figs. 6 and 7 we can conclude that for a given λ_p , increasing $\sigma_{s_i,s,d}$ or $\sigma_{s_m,p,d}$ boosts the secondary throughput. Ditto for a given λ_s , increasing $\sigma_{s_i,s,d}$ or $\sigma_{s_m,p,d}$ boosts the primary throughput. The parameters used to generate each of the figures are the common parameters, $\mathcal{M} = 7$, $P^s = 10^{-10}$ Watts/Hz and the parameters in the legend of each figure. Fig. 8 shows the maximum secondary stable throughput of the proposed system for different $P_T = WP^s$. Increasing P^s boosts the maximum secondary stable throughput as well as the feasible primary mean arrival rate. This is because increasing the transmit power decreases the outage probability of the link. This can be noticed from the outage probability formula. The figure is generated using the common parameters and the parameters in the figure's legend.

APPENDIX A

A. Channels Outages

Let $g_{j,\rho}^t$ denote the channel gain between transmitting node j and receiving node ρ at instant t , where $j \in \{s_1, s_2, \dots, s_{\mathcal{M}}, p\}$, $\rho \in \{s_1, s_2, \dots, s_{\mathcal{M}}, sd, pd\}$ and $j \neq \rho$,

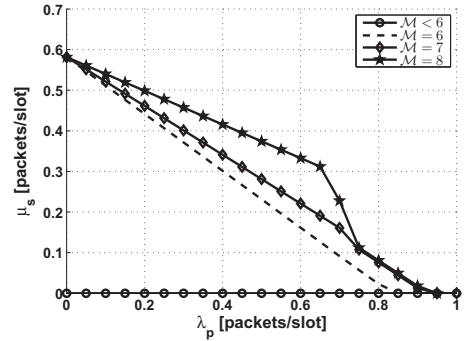


Fig. 4. The impact of the number of antennas on the secondary throughput.

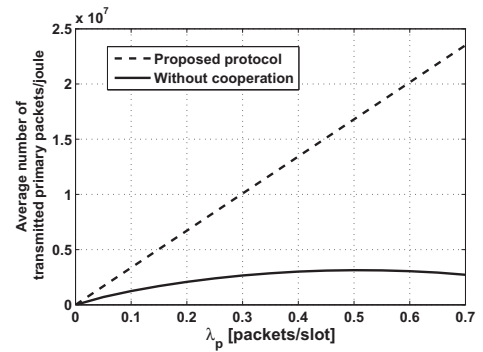


Fig. 5. The number of transmitted primary packets per joule.

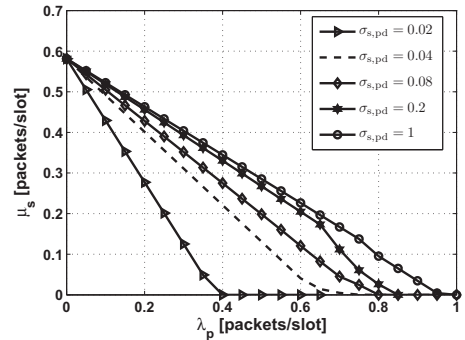


Fig. 6. Maximum mean secondary service rate for the proposed system when the expected channel gain of the highest channel gain of the link connecting the m th antenna and the primary destination varies.

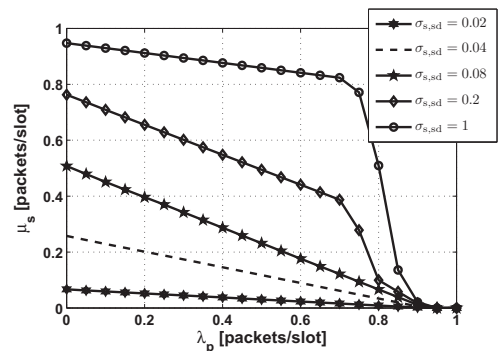


Fig. 7. Maximum mean secondary service rate for the proposed system when the expected channel gain of the highest channel gain of the link connecting the i th antenna and the secondary destination varies.

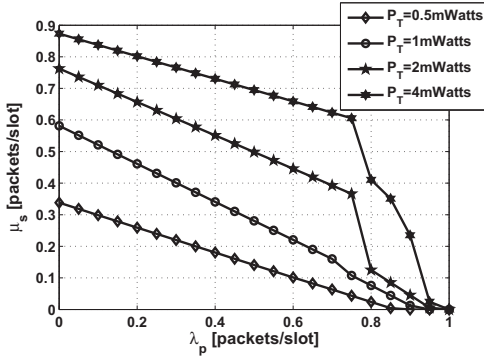


Fig. 8. Maximum secondary service rate of the proposed protocol when secondary transmit power, $P_T = P^s W$, varies.

which is exponentially distributed in case of Rayleigh fading channel with mean $\sigma_{j,\rho}$. Hereinafter, the time notation is omitted from all symbols for simplicity. Channel gains are independent from slot to slot and link to link. Outage occurs when the transmission rate exceeds the channel capacity. Since the transmission rates of all nodes change from state to state, the outage probability should be parameterized by the state of the PU. The probability of channel outage of the link between node j and node ρ , for a given state $H \in \{\circ, F, R\}$, is given by [10]

$$\mathbb{P}_{j,\rho}^H = \Pr\left\{r_j^H > W_{j,\rho}^H \log_2(1 + \beta_{j,\rho})\right\} = 1 - \exp\left(-\frac{2^{\frac{r_j^H}{W_{j,\rho}^H}} - 1}{\gamma_{j,\rho} \sigma_{j,\rho}}\right) \quad (19)$$

where r_j^H is the transmission rate of transmitter j at state H , $W_{j,\rho}^H$ is the transmission bandwidth used for the communication between node j and node ρ at primary state H , $\beta_{j,\rho} = \frac{P^j g_{j,\rho}}{N_o}$ is the instantaneous signal-to-noise ratio (SNR), $\gamma_{j,\rho} = P^j / N_o$ is the received SNR at node ρ when the channel gain is equal to unity, and P^j is the transmit power per unit frequency used by node j . Note that $W_{p,pd}^H = W_{s\nu,pd}^H = W_p$, where $H \in \{F, R\}$, $W_{s\nu,sd}^F = W_{s\nu,sd}^R = W_s$ and $W_{s\nu,sd}^\circ = W$. It should be mentioned that the transmit power by any of the SU's antennas is P^s . Hence, the received SNR at destination ρ , when the channel gain is unity, is $\gamma_{s\nu,\rho} = \gamma_{s,\rho} = P^s / N_o$. We note that the outage probability (19) decreases with $\sigma_{j,\rho}$.

B. Secondary Transmission Rate

The transmission rate of the secondary transmitter differs from state to state. If the PU is idle or sending a packet for the first time (forward state), the secondary rate is given by

$$r_{s_i}^F = r_{s_i}^\circ = \frac{b}{T - \tau} \quad (20)$$

In case of retransmission of a primary packet, the SU does not sense the channel because it knows with certainty that the PU is active. Hence, the secondary transmission rate is given by

$$r_{s_i}^R = \frac{b}{T} \quad (21)$$

C. Primary Packet Decoding at the SU

In forward states, the PU sends a packet of size b bits over $[0, T_{p,F}]$; hence, the primary rate is $r_p^F = b / T_{p,F}$. We assume

that the SU combines what it gets from each antenna⁴; hence, the probability of primary packet decoding failure by the SU is given by

$$\mathbb{P}_{p,s}^F = \Pr\left\{\sum_{\nu=1}^{\mathcal{M}} g_{p,s\nu} \gamma_{p,s\nu} < 2^{\frac{b}{T_{p,F} W_p}} - 1\right\} \quad (22)$$

where $\gamma_{p,s\nu} = P^p / N_o$ and P^p is the primary transmit power. The closed-form of the probability in (22) can be found in [18].⁵

REFERENCES

- [1] G. Y. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: tutorial, survey, and open issues," *IEEE Wireless Communications*, vol. 18, no. 6, pp. 28–35, 2011.
- [2] E. V. Belmega, S. Lasaulce, and M. Debbah, "A survey on energy-efficient communications," in *IEEE 21st International Symposium on PIMRC Workshops*, 2010, pp. 289–294.
- [3] L. Wang and Y. Xiao, "A survey of energy-efficient scheduling mechanisms in sensor networks," *Mobile Networks and Applications*, vol. 11, no. 5, pp. 723–740, 2006.
- [4] G. Miao, N. Himayat, Y. G. Li, and A. Swami, "Cross-layer optimization for energy-efficient wireless communications: a survey," *Wireless Communications and Mobile Computing*, vol. 9, no. 4, pp. 529–542, Apr. 2009.
- [5] R. Devarajan, A. PUNCHIHewa, and V. K. Bhargava, "Energy-aware power allocation in cooperative communication systems with imperfect CSI," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 1633–1639, May 2013.
- [6] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Stable throughput of cognitive radios with and without relaying capability," *IEEE Trans. Commun.*, vol. 55, no. 12, pp. 2351–2360, Dec. 2007.
- [7] M. Elsaadany, M. Abdallah, T. Khattab, M. Khairy, and M. Hasna, "Cognitive relaying in wireless sensor networks: Performance analysis and optimization," in *Proc. IEEE GLOBECOM*, Dec. 2010, pp. 1–6.
- [8] S. Kompella, G. Nguyen, J. Wieselthier, and A. Ephremides, "Stable throughput tradeoffs in cognitive shared channels with cooperative relaying," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 1961–1969.
- [9] W. Su, J. D. Matyas, and S. Batalama, "Active cooperation between primary users and cognitive radio users in heterogeneous ad-hoc networks," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1796–1805, Apr. 2012.
- [10] B. Rong and A. Ephremides, "Cooperative access in wireless networks: stable throughput and delay," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 5890–5907, Sept. 2012.
- [11] H.-B. Chang and K.-C. Chen, "Cooperative spectrum sharing economy for heterogeneous wireless networks," in *IEEE GLOBECOM Workshops (GC Wkshps)*, 2011, pp. 458–463.
- [12] H. Roh, C. Jung, W. Lee, and D.-Z. Du, "A stackelberg game for cooperative cognitive radio network with active sus," in *ICNC*, Jan. 2013.
- [13] I. Krikidis, J. Laneman, J. Thompson, and S. McLaughlin, "Protocol design and throughput analysis for multi-user cognitive cooperative systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4740–4751, Sept. 2009.
- [14] A. Sadek, K. Liu, and A. Ephremides, "Cognitive multiple access via cooperation: protocol design and performance analysis," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3677–3696, Oct. 2007.
- [15] R. Loynes, "The stability of a queue with non-independent inter-arrival and service times," in *Proc. Cambridge Philos. Soc.*, vol. 58, no. 3. Cambridge University Press, 1962, pp. 497–520.
- [16] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex mimo relays," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5983–5993, Mar. 2011.
- [17] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 659–672, Aug. 2006.
- [18] E. M. Scheuer, "Reliability of an m out of n system when component failure induces higher failure rates in survivors," *IEEE Trans. Rel.*, vol. 37, no. 1, pp. 73–74, 1988.

⁴Recall that the receiving nodes have the CSI.

⁵This probability, $\mathbb{P}_{p,s}^F$, can be upper bounded by the probability of having \mathcal{M} antennas each of which is independently decodes the primary packet.