Abstract—In this paper, we study the impact of a full-duplex secondary node on a cognitive cooperative network with Multipacket Reception (MPR) capabilities at the receivers. Motivated by recent schemes that make full-duplex communication feasible, we study a model with one primary and one secondary transmitter-receiver pair, where the secondary transmitter is able to relay primary unsuccessful packets. Cooperation between primary and secondary users has been previously shown to be beneficial for the primary and the secondary users in terms of stable throughput. Our model assumes an imperfect full-duplex secondary node that can transmit and receive simultaneously, cancelling self-interference to a certain extent. Furthermore, we assume that the secondary transmitter chooses between cooperating with the primary user and transmitting secondary packets probabilistically according to some optimized probabilities that depend on both the channels in the network and the state of the primary user. We determine these probabilities by formulating a constrained optimization problem with the secondary throughput as the objective function and the stability of the primary queues as constraints. Using the dominant system approach, we show that the optimization problem has a quasi-concave structure, to which the optimal solution can be easily found. Using numerical results, we characterize the cases where the full-duplex capability is beneficial to the system, namely, we show that the full-duplex secondary node greatly increases both the secondary throughput and the primary maximum stable throughput in channels with receivers that have strong MPR capability.

I. INTRODUCTION

Cognitive radio technology has the potential to alleviate the scarcity of the spectrum by allowing unlicensed (secondary) nodes access the licensed but under-utilized radio spectrum as long as the interference to the licensed (primary) nodes is completely avoided or limited. Initial works on cognitive radio networks consider only opportunistic spectrum access whereby the secondary users sense the channel for primary activity and access the channel only when the primary user is inactive. However, in a more recent approach proposed by [1], the secondary users cooperate with primary users to relay the primary packets to the destination whenever they are able to decode the primary users’ packets that are not decoded by the primary destination. Meanwhile, the secondary nodes are allowed to transmit their own packets whenever the primary users are idle. It was shown that this cooperative approach can increase both the primary and secondary user throughputs.

This approach was extended for MultiPacket Reception (MPR) in [2], wherein the secondary user is able to access the channel simultaneously with the primary user at the cost of mutual interference at the primary and the secondary destinations. In [3], [4] a cognitive spectrum sharing network is considered. A Markov decision process based framework was used to determine the optimal cooperation policy of the secondary user that maximizes the secondary throughput. In this framework, a secondary user makes the cooperation decisions dynamically depending on the queue state of the primary user. It is assumed that each time slot is divided into two portions, one for primary transmissions and one for possible secondary relaying. This setting may become wasteful when the secondary user decides not to cooperate.

Most of the work on cooperative cognitive networks assume that the nodes are half-duplex, i.e., when the secondary user accesses the channel simultaneously with the primary user, it cannot decode the packets of the primary user. Recent works in [5] and [6] propose practical schemes to enable full-duplex communications. Full-duplex communications was deemed infeasible in the past due to the self-interference, i.e., the interference caused by the transmitted signal on a received signal at the same node. However, [5] and [6] demonstrated that self-interference can be significantly suppressed allowing nodes to transmit and receive simultaneously.

Cooperation between primary and secondary user(s) has been proposed as one way to exploit possible full-duplex capability at the secondary user(s). The full-duplex capability enables the secondary user to simultaneously transmit and listen for primary transmission, cooperating with the primary user whenever needed. In [7], an imperfect full-duplex multi-antenna secondary user is able to relay the primary user signal. In return, the secondary user is able to overlay its own signal on top of the primary signal. The beamforming vector is chosen in a way that maximizes the secondary rate under a primary target rate condition. It was shown that the full-duplex operation of the secondary user increases both the secondary and primary maximum achievable rates. In contrast to our work, the work in [7] investigates the performance of a full-duplex secondary node in a cognitive channel using information theoretic metrics, namely, the rate region. In our work, we investigate how the full duplex secondary node...
affects the system using network layer metrics, namely the stable throughput. In [8], a multiaccess channel with a full-duplex relay is studied. It was shown that a full-duplex relay can help improve the system’s throughput and delay. However, this work assumes that the relay has no data of its own and the only constraint on the system is the stability of the users queue.

Our main contribution in this paper is to study the effect of an imperfect full-duplex secondary node, which is able to decode and retransmit the primary unsuccessful packets on a cognitive interference channel to the receivers with MPR capabilities from a queuing theoretical perspective. The full-duplex capability enables the secondary transmitter to decode the primary transmission while simultaneously accessing the channel with the primary user. The simultaneous transmission of primary and secondary users increases the mutual interference, and decreases the probability of successful decoding at the receivers. On the other hand, if the interference is not high, the simultaneous primary and secondary transmissions may increase the throughput of the secondary node. We show that by utilizing the full-duplex capability of the secondary transmitter and by optimally scheduling the transmissions of the secondary user, significant gains can be achieved depending on the topology of the network. We characterize the cases where the full-duplex outperforms the half-duplex operation.

The rest of this paper is organized as follows. In Section II, we present the system model. Afterwards, we present the problem formulation in Section III. In Section IV, we present the solution approach. In Section V, we provide numerical results. Finally, conclusions are drawn in Section VI.

II. SYSTEM MODEL

We study a two user cognitive radio interference channel consisting of a primary and a secondary transmitter-receiver pair as shown in Fig. 1. The secondary transmitter is able to perfectly determine the state of the primary transmitter being “busy” or “idle”. Time is slotted, where one packet duration is equal to one slot duration. The packet arrivals at the primary transmitter follow a stationary Bernoulli process with mean $\lambda_p$ packets per slot. The secondary transmitter is assumed to have infinite number of secondary packets backlogged. We assume that the primary destination has MultiPacket Reception (MPR) capability, i.e., the primary destination can decode multiple packets simultaneously under certain conditions that shall be stated later. Throughout this work, we refer to the primary source and destination as $S_p$ and $D_p$, respectively, and to the secondary source and destination as $S_s$ and $D_s$ respectively.

A. Physical Layer Model

The link between any pair of nodes $(i,j)$ is subject to independent stationary Rayleigh flat-fading, where the channel gain between nodes $i$ and $j$ is $h_{ij}$ with $\mathbb{E}[|h_{ij}|^2] = \sigma_{ij}^2$. Channel gains are independent over time and they are mutually independent among links. All nodes are subject to independent additive white complex Gaussian noise with zero mean and variance $N_0$. The primary and secondary sources transmit with fixed power, $P_p$ and $P_s$, respectively.

The secondary transmitter $S_s$ has the complete knowledge of the primary user codebook, and it is able to decode the primary transmissions. However, neither the primary nor the secondary destination, i.e., $D_p$ or $D_s$, has the knowledge of each other’s codebook. Hence, $D_s$ is unable to decode the primary packets and treats them as noise. Similarly, $D_p$ is unable to decode the secondary packets, and treats them as noise.

A receiving node is able to decode a transmitted packet correctly, if the received instantaneous signal-to-interference-plus-noise ratio (SINR) is larger than a certain threshold $\beta$ (SNR in case of no interference). Otherwise, the packet is assumed to be lost, and it is retransmitted. We define $\text{SINR}_{ij}$ as the instantaneous SINR between nodes $i$ and $j$. The channel operates in the following three modes depending on which of the nodes transmit as illustrated in Fig. 2(a)-(c).

(a) Single User Channel: If either the primary or the secondary node transmits alone, the probability of decoding the transmission of node $S_i$ successfully at node $D_i$, where $i \in \{p, s\}$, is given by $P_{S_{is}}^{\text{single}}$.

$$P_{S_{is}}^{\text{single}} = P \{ \text{SINR}_{Si, Di} \geq \beta \} = P \left\{ \frac{|P_i | h_{Si, Di} |^2}{N_0} \geq \beta \right\} = \exp \left( - \frac{\beta N_0}{P_i \sigma_{Si, Di}^2} \right), i \in \{p, s\}$$

(b) Multiaccess (MAC) Channel: Both $S_p$ and $S_s$ transmits primary packets. Since $D_p$ has multipacket reception (MPR) capability, $D_p$ attempts decoding both transmissions by performing Successive Interference Cancellation (SIC) [9]. The decoding order is adaptive, so the primary receiver $D_p$ attempts decoding the signal with the highest instantaneous SINR treating the other signal as noise. If successful, the primary receiver then subtracts the decoded signal from the received signal and decodes the signal with lower SINR. Let $P_{S_{is}}^{\text{MAC}}$ denote the probability of successful decoding of the signal transmitted by node $S_i$ at $D_p$, where $i \in \{p, s\}$ and $j \in \{p, s\}, j \neq i$.

$$P_{S_{is}}^{\text{MAC}} = P \left\{ \frac{|P_i | h_{Si, Dp} |^2}{N_0 + |P_j | h_{Sj, Dp} |^2} \geq \beta \right\}$$
where the first term in the summation is the probability of decoding the packet transmitted by \( S_i \) first, and the second term is the probability of decoding the packet transmitted by \( S_f \) first. We do not allow keeping multiple replicas of the same packet in the network. Hence, during a primary transmission, if \( S_s \) simultaneously transmits with \( S_p \), then the primary packet from \( S_s \) is different from the packet transmitted by \( S_p \).

(c) Interference Channel: \( S_p \) transmits a primary packet, and \( S_s \) transmits a secondary packet simultaneously. The channel becomes an interference channel, where the primary receiver \( D_p \) (secondary receiver \( D_s \)) treats the secondary (primary) transmission as noise. Let \( P_{S_i}^{\text{inf}} \) be the probability of decoding the transmission from \( S_i \) at \( D_i \) successfully, where \( i = \{ p, s \} \).

\[
P_{S_i}^{\text{inf}} = \mathbb{P} \left\{ \frac{P_i |h_s D_i|^2}{N_0 + P_i |h_s D_i|^2} \geq \beta \right\}.
\]

Full-Duplex operation: The secondary transmitter \( S_s \) can leverage the full duplex capability by having a non-zero probability of decoding the primary packet, while transmitting either a primary or a secondary packet. Let \( P_{S_p}^{\text{dup}} \) be the probability that \( S_s \) decodes a primary packet in the full duplex mode.

\[
P_{S_p}^{\text{dup}} = \mathbb{P} \left\{ \frac{P_p |h_s s_p|^2}{N_0 + P_p |h_s s_p|^2} \geq \beta \right\}
\]

where we model the effectiveness of self-interference cancellation techniques by a scalar gain \( g \in [0, 1] \), following other works in the literature, e.g., [8] and [10]. For instance, if \( g = 1 \), no self-interference cancellation is adopted, while if \( g = 0 \) the node cancels self-interference perfectly. The details of the methods used for self-interference cancellation are beyond the scope of this paper. More details on techniques used can be found in [11] and references therein.

B. MAC Layer Model

Let \( Q_p \) denote the primary user queue. The secondary transmitter has two queues; \( Q_p \) for storing the primary packets that were not decoded successfully by \( D_p \) but were successfully decoded by the secondary user, \( S_s \), and \( Q_s \) which is used for storing the secondary user’s packets. In the subsequent analysis, we assume that \( Q_s \) is infinitely backlogged. At the beginning of each slot, the secondary transmitter senses the channel perfectly for primary transmissions. Let \( p_s^{\text{ps}} \) and \( p_s^{\text{ps}} \) be the probability that secondary user transmits a packet from \( Q_s \) and \( Q_{ps} \) respectively when the primary node is idle. Due to its full-duplex capability the secondary user may also transmit when primary user is busy. Let \( p_s^{\text{ps}} \) and \( p_s^{\text{ps}} \) be the probability that secondary user transmits a packet from \( Q_s \) and \( Q_{ps} \) respectively when the primary node is busy.

At the end of each slot, an error-free ACK/NACK packet is sent by \( D_s, D_p \) and \( S_s \). We assume that the ACK/NACK packets are perfect and available at all nodes. At the beginning of each slot, the secondary transmitter senses the channel perfectly for primary transmissions. Two cases may arise depending on the outcome of this sensing:

- **Primary is Busy**: Secondary node \( S_s \) attempts to decode the primary user packet. If successful, \( S_s \) adds the primary packet to \( Q_p \) and sends ACK to \( S_p \) only if \( D_p \) sends a NACK. Otherwise \( S_s \) takes no action. \( Q_p \) drops the packet if it receives an ACK from \( D_p \) or \( S_s \). While attempting to decode the primary transmission, \( S_s \)

### Table 1: List of Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Probability that transmission from ...</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_{\text{single}}^{\text{S_p D_p}} )</td>
<td>( S_p ) is successfully decoded at ( D_p ) when ( S_p ) transmits alone</td>
</tr>
<tr>
<td>( P_{\text{single}}^{\text{S_p S_s}} )</td>
<td>( S_p ) is successfully decoded at ( S_s ) when ( S_p ) transmits alone</td>
</tr>
<tr>
<td>( P_{\text{single}}^{\text{S_s D_s}} )</td>
<td>( S_s ) is successfully decoded at ( D_s ) when ( S_s ) transmits alone (secondary packet transmitted)</td>
</tr>
<tr>
<td>( P_{\text{single}}^{\text{S_s D_p}} )</td>
<td>( S_s ) is successfully decoded at ( D_p ) when ( S_s ) transmits alone (primary packet transmitted)</td>
</tr>
<tr>
<td>( P_{\text{SM}}^{\text{S_p}} )</td>
<td>( S_p ) is successfully decoded at ( D_p ) when both ( S_p ) and ( S_s ) transmits a primary packet</td>
</tr>
<tr>
<td>( P_{\text{SM}}^{\text{S_s}} )</td>
<td>( S_s ) is successfully decoded at ( D_p ) when both ( S_p ) and ( S_s ) transmits a primary packet</td>
</tr>
<tr>
<td>( P_{\text{SM}}^{\text{S_p S_s}} )</td>
<td>( S_p ) is successfully decoded at ( D_p ) when ( S_p ) transmits a primary packet and ( S_s ) transmits a secondary packet</td>
</tr>
<tr>
<td>( P_{\text{SM}}^{\text{S_s S_p}} )</td>
<td>( S_s ) is successfully decoded at ( D_p ) when ( S_s ) transmits a primary packet and ( S_p ) transmits a secondary packet</td>
</tr>
</tbody>
</table>
may transmit from $Q_{ps}$ or $Q_s$ with probability $p^B_{ps}$ or $p^B_s$, respectively, where $p^B_{ps} + p^B_s \leq 1$.

- **Primary is Idle**: Secondary node $S_s$ transmits from $Q_{ps}$ or $Q_s$ with probability $p^I_{ps}$ or $p^I_s$ respectively such that $p^I_{ps} + p^I_s = 1$, i.e., the secondary user always transmits when the primary user is idle.

### C. Queue Evolution

In the subsequent analysis we assume for mathematical tractability that the system is non-work conserving, i.e., during an idle (or busy) slot, $S_s$ serves $Q_{ps}$, with probability $p^I_{ps}$ (or $p^I_s$) even if $Q_{ps}$ is empty. A list of the parameters used are given in Table I. Let $Q_p(n)$ and $Q_s(n)$ be the length of the primary transmitter queue and the secondary relaying queue at the beginning of slot $n$, respectively. Let $X_p(n)$ and $Y_p(n)$ be the arrival and service processes at the primary queue, respectively. All processes are assumed to be stationary with $E(X_p(n)) = \lambda_p$ and $E(Y_p(n)) = \mu_p$. Similarly, $E(X_s(n)) = \lambda_{ps}$, and $E(Y_s(n)) = \mu_{ps}$. The evolution of $Q_p$ and $Q_s$ is

\[
Q_p(n + 1) = (Q_p(n) - Y_p(n))^+ + X_p(n) \quad (5)
\]

\[
Q_s(n + 1) = (Q_s(n) - Y_s(n))^+ + X_s(n) \quad (6)
\]

where $x^+ = \max(x, 0)$. The primary node service rate $\mu_p$ depends on $Q_{ps}$ and the action taken by the secondary node, and is given as

\[
\mu_p = \mathbb{P}\{Q_{ps} > 0\} p^B_{ps} (p_{inf} + p_{dup}^{ps} - p_{inf}^{ps} p_{dup}^{ps}) + p^I_{ps} (p_{MAC}^{ps} + p_{dup}^{ps} - p_{MAC}^{ps} p_{dup}^{ps}) + (1 - p^B_{ps} - p^I_{ps}) (p_{inf}^{ps} + p_{macro}^{ps} - p_{inf}^{ps} p_{macro}^{ps}) + \mathbb{P}\{Q_{ps} = 0\} (p^B_{ps} (p_{inf}^{ps} + p_{dup}^{ps} - p_{inf}^{ps} p_{dup}^{ps}) + (1 - p^B_{ps} - p^I_{ps}) (p_{inf}^{ps} + p_{macro}^{ps} - p_{inf}^{ps} p_{macro}^{ps})). \quad (7)
\]

Note that $\mu_p$ represents the rate of packets transmitted by $Q_{ps}$ that are either successfully received by $D_p$ or by $S_s$ in case $D_p$ fails to decode. The primary service rate (7) follows from conditioning on the state of $Q_{ps}$, and the action of $S_s$:

- **$Q_{ps}$ is non-empty. ($\mathbb{P}\{Q_{ps} > 0\}$)**
  - $S_s$ serves $Q_s$, and the mode of operation is that of the interference channel. The service rate is the probability that the packet is decoded either by $D_p$ or by $S_s$, using the full duplex capability.
  - $S_s$ serves $Q_{ps}$, and the mode of operation is that of the MAC channel. The service rate is the probability that the packet is decoded either by $D_p$ or by $S_s$ using the full duplex capability.
  - $S_s$ remains idle, and the mode of operation is that of the single user channel. The service rate is the probability that the packet is decoded either by $D_p$ or by $S_s$ using the full duplex capability.

- **$Q_{ps}$ is empty. ($\mathbb{P}\{Q_{ps} = 0\}$)**
  - $S_s$ serves $Q_s$, and the mode of operation is that of the interference channel. The service rate is the probability that the packet is decoded either by $D_p$ or by $S_s$, using the full duplex capability.

## D. Dominant System Approach

Note that $Q_p$ and $Q_{ps}$ are interacting queues, i.e., the service rate of each queue depends on the steady state distribution of another queue. The *dominant system approach* was proposed in [12] and [13] to obtain the sufficient conditions for stability. For a system with two interacting queues, a dominant system is the one where one of the queues sends dummy packets whenever the channel is empty, and, thus, constantly interferes with the transmission of the other queue at every slot. It is clear that the queues of the dominant system can never be shorter than those of the original system, so the stability of the dominant system implies the stability of the original system.

In our system, the service rate of $Q_p$ in the dominant system is derived by modifying (7) by letting $\mathbb{P}\{Q_{ps} = 0\} = 0$, i.e., $Q_{ps}$ is always backlogged.

\[
\begin{align*}
\mu_p &= p^B_{ps} (p_{inf} + p_{macro}^{ps} - p_{inf}^{ps} p_{macro}^{ps}) + p^I_{ps} (p_{MAC}^{ps} + p_{macro}^{ps} - p_{MAC}^{ps} p_{macro}^{ps}) + (1 - p^B_{ps} - p^I_{ps}) (p_{inf}^{ps} + p_{macro}^{ps} - p_{inf}^{ps} p_{macro}^{ps}) = \frac{\lambda_p}{\mu_p} p^I_{ps} p_{MAC}^{ps} + \lambda_{ps} = \frac{\lambda_p}{\mu_p} p^I_{ps} p_{MAC}^{ps} + \lambda_{ps}, \quad (11)
\end{align*}
\]

Similarly, the arrival rate $\lambda_{ps}$ is the rate of primary packets that $D_p$ failed to decode but successfully decoded by $S_s$ in the dominant system, and can be obtained as

\[
\begin{align*}
\lambda_{ps} &= \mathbb{P}\{Q_p > 0\} (p^B_{ps} p_{MAC}^{ps} + p^B_{ps} p_{MAC}^{ps} + (1 - p^B_{ps} - p^I_{ps}) p_{inf}^{ps} p_{macro}^{ps}) = \frac{\lambda_p}{\mu_p} p^I_{ps} p_{MAC}^{ps} + \lambda_{ps} = \frac{\lambda_p}{\mu_p} p^I_{ps} p_{MAC}^{ps} + \lambda_{ps}, \quad (12)
\end{align*}
\]

where $x = (1 - x)$ represents the probability of the complement of an event occurring with probability $x$. 

}\]

478
\[ \gamma(p^B_p, p^B_{bs}) = \frac{\lambda_p p^B_{bs} P^{\text{MAC}}_{S_d} \mu_p}{(\mu_p - \lambda_p) P^{\text{MAC}}_{S_d} D_p} \left( \frac{1 - p^B_p - p^B_{bs}}{\lambda_p} \right) \]  
\[ = \frac{\lambda_p p^B_{bs} P^{\text{MAC}}_{S_d} \mu_p}{(\mu_p - \lambda_p) P^{\text{MAC}}_{S_d} D_p} \]  
\[ = \frac{\lambda_p p^B_{bs} P^{\text{MAC}}_{S_d} \mu_p}{(\mu_p - \lambda_p) P^{\text{MAC}}_{S_d} D_p} \]  

III. Problem Formulation

We aim to maximize the secondary throughput \( \mu_s \) subject to the stability of the primary system. We assume that \( Q_s \) is always backlogged. We apply the stability conditions on the dominant system. Hence, the problem we formulate and the solution we obtain are those of the dominant system defined in the previous section. However, we show that at the optimal operating point, the dominant system is indistinguishable from the original system. Hence, maximizing the secondary throughput subject to the stability conditions of the dominant system is equivalent to maximizing the secondary throughput in the original system.

By using Little’s law, we can rewrite (9) as

\[ \mu_s = \frac{\lambda_p p^B_{bs} P^{\text{inf}}_{S_d} \mu_p}{(\mu_p - \lambda_p) p^B_{bs} P^{\text{inf}}_{S_d} D_p} \]  
\[ \mu_s = \frac{\lambda_p p^B_{bs} P^{\text{inf}}_{S_d} \mu_p}{(\mu_p - \lambda_p) p^B_{bs} P^{\text{inf}}_{S_d} D_p} \]  

where \( \mu_s \) is as given in (10). We use the definition of stability given in [15], i.e., the system is stable if there exists a unique stationary distribution for each queue, and the queue lengths do not grow to infinity with time. According to Looynes criteria, the arrival rate should be less than the service rate for each queue in order for the queues to be stable [16].

Our optimization problem is formulated as follows:

\[ \text{P1: max } \mu_s \]
subject to
\[ \lambda_p < \mu_p, \]
\[ \lambda_{ps} < \mu_{ps}, \]
\[ p^B_{bs} + p^B_{ps} \leq 1, \]
\[ p^B_{ps} + p^B_{ps} \leq 1, \]
\[ p^B_{ps}, p^B_{ps} \geq 0, \]

The decision variables are the access probabilities \((p^B_{bs}, p^B_{ps})\). Constraint (15) ensures the stability of \( Q_p \), and (16) ensures the stability of \( Q_{ps} \). (17) ensures that the secondary user either relays a primary packet from \( Q_{ps} \), transmits a secondary packet from \( Q_s \), or abstains from transmitting to limit the interference to the transmission from \( S_p \). The equality in (18) enforces \( S_s \) to transmit either from \( Q_{ps} \) or \( Q_s \) during a primary idle slot.

The objective function \( \mu_s \), as given in (14), is not concave, since \( \mu_p \), which is a linear function of \( p^B \) and \( p^B_{ps} \), is multiplied by \( p^I \). Furthermore, the feasible region is not convex due to the non-convexity of the constraint in (16), since the expression for \( \mu_{ps} \) in (11) also has \( p^I \) multiplied by \( \mu_p \). Nevertheless, the structure of the problem enables us to determine the optimal solution efficiently as presented next.

IV. Solution Approach

Note that the optimal value of \( p^I_s \) is a function of both \( p^B_p \) and \( p^B_{ps} \), i.e., for fixed values of \( p^B_p \) and \( p^B_{ps} \), a closed form can be found for optimal \( p^I_s \) by exploiting the structure of the problem. We divide the problem into two subproblems with non-overlapping feasible regions. Both subproblems turn out to be linear fractional optimization problems. It is well known that linear fractional problems are both quasi-convex and quasi-concave, and they can be solved efficiently by using the bisection algorithm [17]. The two sub-problems can be solved simultaneously, and the solution that achieves the highest objective function is solution to P1.

We first reduce our decision variables, by observing that

\[ \mu_s = 1 - p^B_s \]

\[ \text{P2: max } \mu_s \]
subject to
\[ \lambda_p < \mu_p, \]
\[ \lambda_{ps} < \mu_{ps}, \]
\[ p^B_{bs} + p^B_{ps} \leq 1, \]
\[ p^B_{ps} + p^B_{ps} \leq 1, \]
\[ \lambda_{ps} - \mu_{ps} \geq 0, \]
\[ p^B_{ps}, p^B_{ps} \geq 0, \]

Lemma 1. The optimal value of \( p^I_s \) solving the optimization problem P2 is given by

\[ p^I_s^* = \min \{ \gamma(p^B_p, p^B_{bs}), 1 \}, \]

where \( \gamma(p^B_p, p^B_{bs}) \) is as defined in (26).

Proof. The secondary service rate \( \mu_s \) as given in (14) is a monotonically increasing function with respect to decision variable \( p^I_s \). Thus, the optimal \( p^I_s^* \) corresponds to the maximum \( \mu_s \) given any \((p^B_p, p^B_{ps})\) in the feasible region. By manipulating the inequality in (22), an upper bound \( \gamma(p^B_p, p^B_{bs}) \) (defined in (26)) on \( p^I_s^* \) is found as a function of the decision variables \((p^B_p, p^B_{ps})\), thus,

\[ p^I_s < \gamma(p^B_p, p^B_{bs}). \]

Since \( p^I_s^* \) is a probability, another natural upper bound on \( p^I_s \) is 1. Since each of these constraints gives an upper bound on \( p^I_s \), we take the minimum of the two upper bounds to obtain the maximum (optimal) value of \( p^I_s^* \).

Since there are two possible upper bounds given in (25), the maximum \( p^I_s^* \) takes one of those two values depending on the value of the decision variables \((p^B_p, p^B_{bs})\). Finally, we solve our optimization problem P2 by dividing it into two subproblems. In the first subproblem, we add a new constraint \( \gamma(p^B_p, p^B_{bs}) \leq 1 \), and make the substitution \( p^I_s^* = \gamma(p^B_p, p^B_{bs}) \) to get a new objective function \( \mu_s^* \) as given in (27). In the
second subproblem, we add the complementary constraint \( \gamma(p_p^B, p_{ps}^B) > 1 \), and make the substitution \( p^*_s = 1 \) to get a new objective function \( \mu^2_s \), as given in (29).

\[
\mu^2_s = \frac{\lambda_p p_p^B p_{ps}^B}{\mu_p} t^\text{inf} S_t + (1 - \frac{\lambda_p p_p^B}{\mu_p}) P^\text{single} S_t, D_s
\]

Note that solving each subproblem is equivalent to solving the problem P2 (equivalently P1) over two non-overlapping feasible regions, whose union gives the original problem’s feasible region. The first subproblem P3 is formulated as

**P3:** \[
\max_{p^B_p, p^B_{ps}} \mu^2_s(p_p^B, p_{ps}^B)
\]

subject to

1. \( \lambda_p < \mu_p \)
2. \( \gamma(p_p^B, p_{ps}^B) \leq 1 \)
3. \( p_{ps}^B + p_p^B \leq 1 \)
4. \( p_{ps}^B, p_p^B \geq 0 \)

The objective function of P3 is given in (27), and it is a linear fractional function with respect to decision variables \((p_p^B, p_{ps}^B)\). Note that (32) is also a linear fractional function of \((p_p^B, p_{ps}^B)\). Likewise, (31), (33) and (34) are all linear functions of \((p_p^B, p_{ps}^B)\). Linear fractional optimization problems are quasi-concave [17], and they can be solved efficiently by using the bisection algorithm. The bisection algorithm works by formulating a sequence of feasibility problems, where each problem is a linear program.

Next, we present the solution approach for subproblem P3. In order to implement the bisection algorithm, we first find a feasible lower bound, \( t_1 \), and an infeasible upper bound, \( t_u \), for the optimization problem. We let \( t_1 = 0 \) packets/slot and \( t_u = 1 \) packets/slot, since the secondary throughput \( \mu_s \) is guaranteed to fall between these two values. We re-write P3 in hypograph form [17] as

\[
\max_{p_p^B, p_{ps}^B} t
\]

subject to

1. \( \mu_p t^\text{inf} (p_p^B, p_{ps}^B) \geq t \mu_p \)
2. \( \lambda_p \leq \mu_p \)
3. \( (\mu_p - \lambda_p) P^\text{single} S_t, D_p \gamma(p_p^B, p_{ps}^B) < (\mu_p - \lambda_p) P^\text{single} S_t, D_p \)
4. \( p_{ps}^B + p_p^B \leq 1 \)
5. \( p_{ps}^B, p_p^B \geq 0 \)

where \( t \) is the new scalar objective. The inequality in (35) is the new constraint function \( \mu^2_s(p_p^B, p_{ps}^B) - t \) added to the hypograph form problem [17]. Inequality (36) guarantees the stability of \( Q_{ps} \). Inequality (37) is obtained by manipulating inequality (32) in P2 to make it linear. Given the value of \( t \), the feasibility problem is a linear program. Since the optimal value of the objective function \( \mu^2_s \) is between \( t_1 \) and \( t_u \), we solve the feasibility problem iteratively at \( t = \frac{t_1 + t_u}{2} \). If the problem is feasible we update \( t_1 = t \) and if the problem is infeasible we update \( t_u = t \). We repeat this procedure until interval \([t_1, t_u]\) falls below a predetermined threshold, and the optimal value is then given by \( t_1 \).

The second subproblem P4 is formulated as:

**P4:** \[
\max_{p^B_p, p_{ps}^B} \mu^2_s(p_p^B, p_{ps}^B)
\]

subject to

1. \( \lambda_p < \mu_p \)
2. \( \gamma(p_p^B, p_{ps}^B) > 1 \)
3. \( p_{ps}^B + p_p^B \leq 1 \)
4. \( p_{ps}^B, p_p^B \geq 0 \)

The objective function of P4 as given in (29) is also linear-fractional in \((p_p^B, p_{ps}^B)\). P4 can be solved in a similar fashion as in P3 by transforming it into a set of feasibility problems, and finding the optimal solution via the bisection algorithm. After the solutions of two subproblems are found, the optimal values \( \mu^*_s \) and \( \mu^2_s \) are calculated. Summarizing the solution of P2 in Theorem 1.

**Theorem 1.** The optimal values of the decision variables \((p_p^B, p_{ps}^B, p_s^*, p_{ps}^*)\) for P2 and the corresponding optimal \( \mu^*_s \) are given as

\[
(p_p^B, p_{ps}^B) = \arg \max p_p^B, p_{ps}^B \mu^*_s(p_p^B, p_{ps}^B) \]

\[
p_p^* = 1 - p_s^*
\]

\[
p_s^* = \begin{cases} \gamma(p_p^B, p_{ps}^B) & \text{if } \mu^*_s \geq \mu^2_s \\ 1 & \text{if } \mu^*_s < \mu^2_s \\ \end{cases}
\]

\[
\mu^*_s = \max(\mu^*_s(p_p^B, p_{ps}^B), \mu^2_s(p_p^B, p_{ps}^B))
\]

Note that we solved the optimization problem for the dominant system, wherein \( Q_{ps} \) sends dummy packets whenever it is empty. We use the “indistinguishability” argument in [12], which was also used in a similar setting as ours in [1] to prove that at the optimal operating point \((p_p^B, p_{ps}^B, p_s^* , p_{ps}^*)\), the behaviour of the original system is identical to the dominant system. This means that our solution gives the optimal access probabilities, and the maximum secondary throughput for the original system with interacting queues.

To use the indistinguishability argument, we begin by inspecting how the value of the objective function \( \mu^*_s \) changes with increasing value of the decision variables \( p_{ps}^B \) and \( p_p^B \). When \( p_{ps}^B \) increases, \( p_s^* \) decreases, and \( \mu^*_s \) also decreases since it is monotonically increasing with \( p_s^* \). When \( p_p^B \) increases, \( p_p^B \) decreases if the inequality \( p_p^B + p_{ps}^B \leq 1 \) is satisfied with equality. Furthermore, \( \mu^*_s \) decreases, since \( Q_p \) now sees interference in more number of slots. This also decreases \( \mu^*_s \) due to a decrease in the number of primary idle slots. In short, the objective function \( \mu^*_s \) decreases, whenever \( p_{ps}^B \) or \( p_p^B \) increases. The decision variables \( p_{ps}^B \) and \( p_p^B \) control the service rate of \( Q_{ps} \), so increasing \( p_{ps}^B \) and \( p_p^B \) means more transmission attempts from \( Q_{ps} \) increasing the service rate \( \mu_{ps} \). We conclude that if \( Q_{ps} \) increases its transmission attempts in idle and/or busy slots, the secondary throughput

---

1. If the problem for \( t_1 = 0 \) is infeasible, then the feasible set is empty and no value for the decision variables make \( Q_p \) and/or \( Q_{ps} \) stable.
powers of the primary and secondary transmitters are equal to the channel variances between all nodes is $r_{ij}$. We use the setting depicted in Fig. 1.

If we decrease the transmission attempts of $\mu_s$ too much, $\mu_{ps}$ becomes less than $\lambda_{ps}$, making $Q_{ps}$ unstable. Thus, optimally the transmission attempts of $Q_{ps}$ should be decreased as much as possible until $\lambda_{ps} \approx \mu_{ps}$, as making $\mu_{ps}$ larger than $\lambda_{ps}$ is suboptimal, and decreasing $\mu_{ps}$ to be less than $\lambda_{ps}$ causes $Q_{ps}$ to become unstable. Hence, at the optimal operating point $Q_{ps}$ is always backlogged and the probability of transmitting a dummy packet approaches zero. Consequently, the dominant system is indistinguishable from the original system.

V. NUMERICAL RESULTS

We performed numerical experiments to illustrate the effect of a full duplex secondary node with self-interference cancellation capability. We use the setting depicted in Fig. 1. Let $r_{ij}$ be the distance between nodes $i$ and $j$, and we set $r_{S_p}D_p = 50m$, $r_{S_p}S_s = 30m$, $r_{S_p}D_s = 20m$, $r_{S_s}D_s = 55m$ and $r_{S_s}D_s = 20m$. We use a path loss exponent $\eta = 4$, and the channel variances between all nodes is $\sigma_{ij}^2 = 3$ dB. The powers of the primary and secondary transmitters are equal to $P_{S_p} = P_{S_s} = 8$ dB. Obviously $S_s$ has better channel to both $D_s$ and $D_p$ than $S_p$, which justifies the secondary transmitter relaying the primary unsuccessful packets. Furthermore, the system parameters are chosen such that $D_p$ has strong MPR capability, specifically $f_{S_p}^{\text{MAC}} = 0.42$ and $f_{S_s}^{\text{MAC}} = 0.79$ which means that with high probability, $D_p$ can decode simultaneous transmissions from $S_s$ and $S_p$ in the MAC channel configuration. Meanwhile, we set $P_{S_p}^{\text{inf}} = 0.002$, and $P_{S_s}^{\text{inf}} = 0.78$ which means that in the interference channel mode of operation, only the secondary transmission has a high probability of successful decoding.

In Fig. 3, we plot the secondary throughputs $\mu_s$ against the primary arrival rate $\lambda_p$ for $g = 1$, which represents no self-interference cancellation capability, i.e., half duplex operation, and for $g = 10^{-7}$, which represents a high self-interference cancellation capability making full duplex operation feasible. The probability $P_{S_p}^{\text{inf}}$ corresponding to $g = 10^{-7}$ is equal to 0.625. Results show that the full duplex operation gives a higher secondary throughput than the half duplex operation for all $\lambda_p$, and increases the primary maximum stable throughput over the half duplex system.

Fig. 4 shows the state dependent access probabilities for both full- and half-duplex systems. We note that for low arrival rates (e.g., $\lambda_p < 0.7$), $p_s^D + p_p^D = 1$ for the full-duplex system. Hence, $S_s$ accesses the channel at every slot at the

Fig. 3: Secondary Throughput in strong MPR case

Fig. 5: Secondary Throughput in weak MPR case

Fig. 4: Access probabilities in the Full-duplex and the Half-duplex systems

Fig. 6: Secondary Throughput in weak MPR case

2Note that the value chosen for $g$ is close to the practical values reported in [6].
cost of relaying the primary user packets in both the busy and idle slots. On the other hand, at higher arrival rates (e.g., \( \lambda_p > 0.7 \)), \( S_s \) remains idle during the busy slots to limit the interference on \( S_p \) in order to maintain the stability of \( Q_p \).

On the other hand, in the half-duplex system \( P_{i}^{peak} = 0 \) for all \( \lambda_p \), which means that \( S_s \) cannot use any of the busy slots for secondary transmissions due to high interference, and the half duplex constraint.

In Fig. 5, we plot the secondary throughput for a symmetric system, i.e., \( r_{ij} = 40m \) and \( \sigma_{ij}^2 = 3dB \) for every pair of nodes \( i \) and \( j \). This creates a very poor MPR capability at \( D_p \), i.e., \( P_{Sp}^{MAC} = P_{Ds}^{MAC} = 0.08 \). Fig. 5 shows that the full duplex system performs slightly worse than the half duplex system. This is because the full-duplex system loses the advantage of simultaneous transmissions when the MPR capability is weak, while still having to relay extra packets decoded in the full duplex mode (\( P_{Sp}^{Sup} = 0.72 \)).

Fig. 6 depicts the secondary throughput for a system, where the primary user has to depend heavily on the secondary transmitter relaying its packets. The distances are \( r_{SpDp} = 70m \), \( r_{SpDs} = 50m \), \( r_{SpDs} = 5m \), \( r_{SpDs} = 5m \) and \( r_{SpDs} = 80m \). The transmission powers are chosen as \( P_{Sp} = 8dB \) and \( P_{Ds} = 2dB \). The secondary transmitter is a low power node that is much closer to the primary and secondary receivers than the primary transmitter. This resembles a scenario where the primary transmitter is a high power macro base station and the secondary transmitter is an indoor low power femtocell.

The parameters chosen here make the primary transmitter-receiver channel very unreliable, in particular \( P_{Sp}^{single} = 0.04 \). \( S_p \) has a much better channel to \( S_s \), \( P_{Sp}^{single} = 0.52 \). \( S_s \) has a reliable channel to both \( D_p \) and \( D_s \), \( P_{Sp}^{single} = 0.97 \). Furthermore, \( S_s \) can still transmit reliably when \( S_p \) is transmitting simultaneously (\( P_{Sp}^{MAC} = 0.88 \), \( P_{Sp}^{MAC} = 0.78 \)). The full-duplex capability and the low power of \( S_s \) make \( P_{Sp}^{Sup} = 0.44 \) which is fairly close to \( P_{Sp}^{single} = 0.52 \). Fig. 6 shows that the full-duplex system performs significantly better than the half-duplex system. For \( \lambda_p = 0.2 \), the full-duplex system has a gain of 75% over the half-duplex system in terms of the secondary throughput. For \( \lambda_p = 0.3 \), the secondary throughput of the full-duplex system is more than 4.5 times higher than that of the half-duplex system. We note that the benefits of full duplex relaying are more pronounced in systems that have the primary transmitter depending heavily on secondary relaying such as the cellular networks with macro and femto base stations.

**VI. Conclusions**

In this paper, we have studied the effect of a full duplex secondary transmitter, capable of relaying unsuccessful primary packets, in a cognitive cooperative network. We introduced a channel access protocol for the secondary transmitter that takes into account the sensing outcome, the transmission’s probability of success and the secondary transmitter’s full-duplex capability. The problem of finding optimal channel access probabilities of the secondary user is formulated as a constrained optimization problem. The problem is transformed to a linear fractional problem that is known to be quasi-concave. We compared the performance of the developed protocol for both the half-duplex and the practical full-duplex scenarios. The full-duplex capability significantly increases both the secondary throughput and the primary maximum stable throughput in networks that have strong MPR capability, as the secondary transmitter is able to transmit primary or secondary packets while attempting to decode primary unsuccessful packets. Clearly, this increases throughput compared to the half-duplex system where each slot is used to either transmit or receive. However, the full-duplex capability does not improve the performance of systems that have weak/no MPR capability. This is because simultaneous transmissions of primary and secondary packets in that case are not beneficial, as at least one of them is guaranteed to fail.

**References**


