On the Capacity of the Cognitive Tracking Channel

Sudhir Srinivasa and Syed Ali Jafar
Electrical Engineering and Computer Science
University of California Irvine, Irvine, CA 92697-2625
Email: sudhirs@uci.edu, syed@uci.edu

Nihar Jindal
Department of Electrical and Computer Engineering
University of Minnesota, Minneapolis, MN 55455
Email: nihar@umn.edu

Abstract—We explore the capacity of opportunistic secondary(cognitive) communication over a spectral pool of two independent channels. Due to the distributed nature of the primary user’s spectral activity, the cognitive receiver does not have full knowledge of the channel used at the transmitter for secondary communication. Tracking the transmitter state at the receiver is therefore a primary issue in such channels. The problem is further complicated as the channel availability changes with time. The tracking uncertainty also makes decoding at the receiver non-trivial. Using genie based outer bounds and training based lower bounds, we estimate the capacity of the secondary link. The capacity analysis shows that the benefits of spectral pooling are lost in dynamic spectral environments.

I. INTRODUCTION

Cognitive radios [1]–[4] are encouraging solutions to improve the utilization of the radio spectrum. The main idea in cognitive radio is to periodically monitor the radio spectrum, intelligently detect occupancy in the different parts (channels) of the spectrum and then opportunistically communicate over unused channels (spectrum holes) with minimal interference to the active licensed (primary) users. Cognitive radio, however, is still an emerging technology and faces a number of challenges in how the radio learns and adapts to the local spectral activity at each end of the cognitive link. Many of these challenges can be traced down to the distributed and dynamic nature of the underlying spectral environment [5].

Fig. 1: Different Perspectives on Local Spectral Activity at Cognitive Radio Transmitter $T$ and Receiver $R$

Consider the conceptual depiction of a cognitive radio link shown in Figure 1. The white nodes marked $T$ and $R$ represent the cognitive radio transmitter and receiver respectively while the black nodes marked $A$, $B$ and $C$ are primary users occupying different parts of the spectrum. The dotted regions around the cognitive radio transmitter and receiver are the boundaries of the respective sensing regions - spectral activity can only be sensed within these regions. We assume that the secondary transmitter $T$ can transmit on a channel if no primary user within its sensing region is active on that channel. If more than one channel is available, $T$ chooses one of the available channels. Similarly, the secondary receiver can choose to listen to any channel but it cannot receive signals from multiple channels at the same time. These are typical constraints for a tunable but narrowband transmitter and receiver.

Distributed Spectral Environment: In order to avoid interference to the primary users, the secondary transmitter $T$ uses a channel only when primary users $A$ and $B$ are sensed to be inactive. Due to the physical separation of the secondary transmitter and receiver, their sensing regions are different. Therefore, the secondary receiver $R$ does not automatically have full knowledge of the channel selected by the secondary transmitter $T$. The distributed nature of the spectral environment therefore leads to the uncertainty at the receiver about the communication opportunities sensed at the transmitter $T$.

Dynamic Spectral Environment: In a static environment, where the primary users are always active or inactive, the uncertainty at the receiver can be resolved by an initial search phase where the receiver $R$ identifies the channel used by the secondary transmitter $T$ by scanning all channels, one by one, for a pre-determined training sequence sent by the transmitter. The overhead of such a training scheme would be negligible compared to the asymptotic lengths of capacity achieving codes. However, the problem becomes especially challenging when the primary users spectral activity is also time varying. In a dynamic environment, the secondary transmitter $T$ must respond to the primary users within its sensing region by dynamically switching out of the channels where the primary users become active and into the channels where the primary users become inactive. The dynamic nature of the spectral environment therefore increases the overhead of any handshake protocol between the transmitter and receiver.

Overall, opportunistic communication in a distributed and dynamic spectral environment presents the following challenges from the receiver’s perspective.

1) Tracking problem (choosing the right channel): Based on the past received signals, the receiver must choose the best channel to listen to so that the probability that $T$ and $R$ are on the same channel is maximized. From an information theoretic perspective this is a significant departure from the conventional communication model. Traditionally, real time is not important at the receiver because the receiver waits till the end of transmission to decode the message. However, in the tracking problem, the receiver must make real time choices to stay matched with the transmitter.

2) Decoding problem (decoding with tracking uncertainty): At the end of transmission, the decoding problem is similar to the scenario with partial channel state information at the receiver. The partial state information
in this case refers to the receiver’s estimate of whether each observed symbol corresponds to a matched scenario (i.e., affected by the secondary transmitted symbol) or a mismatch (i.e., independent of the secondary transmitted symbol).

An obvious way to eliminate the overhead of the handshake is to have an ‘offline’ scheme where a pre-determined sequence of channels is agreed upon by the secondary transmitter and receiver. During time slot \( n \), the secondary transmitter monitors the \( n^{th} \) frequency band in the sequence, \( f_n \), and uses it for communication if it is primary user free. The receiver is always matched with the transmitter because it monitors the same frequency segment \( f_n \) for secondary transmissions.

The drawback of the offline scheme is that it does not take advantage of the real-time ‘online’ information obtained about the channel. For example, consider a scenario where the primary users’ transmissions are long and infrequent, i.e., the primary users’ occupancy is characterized by extended periods of activity followed by long periods where the primary user is inactive. Channels (frequency bands) that are free in a particular time slot are therefore more likely to be unoccupied even in the next time slot. It makes sense therefore to keep using a channel that is presently available until it is no longer available. However, the offline scheme does not take advantage of knowledge of the present availability of the channel. In this light, we note that it may be possible to use availability information of the past to predict future availabilities, identify more communication opportunities and consequently perform better.

In order to evaluate the cost of the overhead information and to determine the benefits of these overheads to the cognitive user, a capacity perspective would be immensely helpful. Our goal in this paper is to characterize the fundamental limitations on the capacity of the cognitive user in the distributed and dynamic environment described above. We begin with a formal system model in Section II.

II. System and Channel Model

Our system model is defined on a spectrum pool consisting of \( L \) channels assigned for use to different primary users. For simplicity of exposition, we only deal with \( L = 2 \) channels in this work - analysis of cases where \( L \geq 2 \) will be postponed to [6]. We consider a secondary (cognitive) transmitter-receiver pair trying to communicate with one another using the two channels as shown in Figure 2 (A).

1) Channel Availability Model: The primary user occupancies on the two channels at time \( n \) are collected in the binary random processes \( S_{PU}^1(n) \in \{0, 1\} \) and \( S_{PU}^2(n) \in \{0, 1\} \). A value of \( S_{PU}^1(n) = 0 \) indicates that a primary user is actively transmitting on channel \( l \) at time \( n \). Similarly, a value of \( S_{PU}^1(n) = 1 \) implies that channel \( l \) is free for secondary transmissions. We model the occupancy processes \( S_{PU}^1(n) \) and \( S_{PU}^2(n) \) with independent and identical Markov chains as shown in Figure 2 (B).

2) Cognitive Transmitter:: The secondary transmitter monitors the two channels every time slot to determine whether or not it is in use by the primary radios. When one or more of the two channels is/are deemed temporally unoccupied, it chooses one of the available channels for secondary transmissions. If neither of the two channels is detected to be free, the secondary transmitter does not transmit and goes to the idle state. Therefore in any time slot the cognitive transmitter is in one of three states - it is idle (no transmission), transmits on channel 1 or transmits on channel 2. We capture this state information in the variable \( S_T(n) \in \{0, 1, 2\} \).

3) Input Alphabet and Probability of error: The input alphabet used by the cognitive transmitter is assumed to be a \( Q \)-ary constellation with equiprobable symbols. We further assume that the two channels are \( Q \)-ary symmetric with a symbol error probability of \( \epsilon \), i.e., on either channel the transmitted symbol is received as itself with a probability \((1 - \epsilon)\) and is received as any of the other \((Q - 1)\) symbols with a probability \(\frac{\epsilon}{Q - 1}\). The discrete symmetric channel can be thought of as arising due to hard decision decoding at the secondary receiver. \(\epsilon\) is a function of the constellation size \(Q\) and the constellation power used at the secondary/primary transmitters.

4) Cognitive Receiver: The secondary receiver, at any given time, is only able to scan a single channel for secondary transmissions. It picks the channel which it considers will be used by the secondary transmitter in the next time slot.
and then listens to the corresponding frequency for secondary transmissions. We represent the receiver state at time \( n \), i.e., the channel the receiver chooses to monitor during time slot \( n \), by the variable \( S_R(n) \in \{1, 2\} \).

For simplicity of analysis, we make the optimistic assumption that during each time slot, the secondary transmitter and receiver have accurate information of the presence/absence of primary users in both the channels in their local vicinities, i.e., no errors are made in the detection of spectral holes. We analyze and discuss the effect of missed detection and false alarm [7] on the capacity in [6].

We denote the symbol transmitted from the cognitive transmitter (if any) at time \( n \) by \( X(n) \) and the corresponding signal received by \( Y(n) \). When the transmitter and receiver states are matched \( (S_T(n) = S_R(n)) \), \( Y(n) \) and \( X(n) \) are related through the probability of error distribution dictated by the \( Q \)-ary symmetric channel (QSC) model. On the other hand, when \( S_T(n) \neq S_R(n) \) (includes \( S_T(n) = 0 \), the cognitive receiver only sees random signals (RS in Figure 2 (A)).

### III. CAPACITY PERSPECTIVE

In a single user point-to-point scenario with different causal side information at the transmitter and the receiver, the system capacity is the solution to the following optimization problem [8]:

\[
C = \max_{p_u(U), x=f(U; s_T), y=y(U; S_T, S_R)} I(U; Y, S_R),
\]

where \( X \) and \( Y \) represent the input and the output, \( S_T \) and \( S_R \) denote the side information at the transmitter and the receiver and \( U \) is an auxiliary random variable independent of \( S_T \). The basic premise of the capacity definition of equation (1) is that the channel definition \( \text{Prob}[Y, S_R|X, S_T] \) is known.

One might be tempted to use equation (1) to determine the capacity of the cognitive link in the system model of Section II. However, it should be noted that the channel definition \( \text{Prob}[Y(n), S_R(n)|X(n), S_T(n)] \) in the model is not complete because it critically depends on the following:

- **Transmitter strategy**: The strategy used at the transmitter to choose the channel for transmission when more than one spectral hole is detected.
- **Receiver strategy**: The strategy used at the receiver to track the transmitter state \( S_T(n) \).

Consequently equation (1) cannot be used directly to compute capacity.

In this paper, we characterize the capacity assuming specific strategies at the transmitter. Note, however, that since the receiver strategy is not known, to apply equation (1) requires a maximization of the mutual information over all possible receiver strategies. We approach the problem through the notion of matching probability \( \alpha \) which, as we show, can be used to bound the system capacity given only the transmitter strategy.

We define the matching probability \( \alpha \) as the average fraction of time the secondary transmitter and receiver are, as the name implies, matched to the same channel, i.e.,

\[
\alpha = \lim_{N \to \infty} \frac{\sum_{n=1}^{N} \mathbb{I}[S_T(n) = S_R(n)]}{N},
\]

where \( N \) is the codeword duration and \( \mathbb{I}[\cdot] \) denotes the indicator function. If both the transmitter and the receiver strategies are known, the computation of the matching probability \( \alpha \) is straightforward. However, as we pointed out earlier, we only assume that the transmitter strategy is given. Using the concept of matching probability, we now discuss how upper and lower bounds on the system capacity (for the given transmit strategy) can be derived.

#### A. Upper Bounds

For every set of transmitter and receiver strategies, there exists a corresponding matching probability \( \alpha \). It is easy to see that the capacity (for the given transmitter receiver strategies) cannot exceed \( \alpha C_{QSC}(Q) \), where \( C_{QSC}(Q) \) is the capacity of the \( Q \)-ary symmetric channel. Suppose we fix the transmitter strategy and determine an upper bound \( \beta \) on \( \alpha \) such that \( \alpha \leq \beta \) for all possible receiver strategies. Let \( C_{S_T, S_R} \) be the capacity of the system of Section II for a given transmitter strategy. Clearly \( \beta \) can be used to upperbound this capacity because

\[
\alpha \leq \beta \Rightarrow C_{S_T, S_R} \leq \alpha C_{QSC}(Q) \leq \beta C_{QSC}(Q).
\]

Using this idea, we derive different upperbounds on the capacity for specific transmit strategies in Section IV-C.

#### B. Lower Bounds

Any achievable scheme provides a lower bound on the capacity. A trivial lower bound would be the offline scheme discussed in Section I. Since the transmitter and the receiver are always matched, the receiver only needs to determine if the transmitter is active or idle. Using the genie bound [9], where the genie provides this information to the receiver, a lowerbound on the capacity of the cognitive link easily follows:

\[
C_{S_T, S_R} \geq \min \left\{ \text{Prob}(S^1_{PU}(n) = 1) C_{QSC}(Q) - \mathcal{H}(G) \right\},
\]

where \( \text{Prob}(S^1_{PU}(n) = 1) = \text{Prob}(S^2_{PU}(n) = 1) \) is the steady state probability of the channel being free of primary users and \( \mathcal{H}(G) \) is the entropy measure of the genie information provided to the receiver. This is also essentially the i.i.d scheme discussed in [5]. More sophisticated lower bounds can also be derived with achievable transmitter and receiver strategies. As an example, consider a scenario where the transmitter sends training symbols between the data symbols to help the receiver track the transmitter state. Capacity expressions similar to equation (4) can be derived even in this case. To summarize, given an achievable scheme the corresponding matching probability and the genie bound [9] can be used to derive a lower bound on the capacity.

In this light, the problem of bounding the capacity reduces to finding good upper bounds for the matching probability and tight achievable capacity inner bounds.
IV. BOUNDS ON $\alpha$ FOR THE PREFERENCE STRATEGY

In this section we consider the ‘preference strategy’ at the transmitter and present some bounds on the matching probability $\alpha$. While the upperbound and lowerbound calculations are specific to the preference scheme, the technique employed is general and can be followed to obtain similar bounds for any transmitter strategy.

A. Preference Transmitter Strategy

In this strategy, the secondary transmitter prefers one of the two channels over the other. Without any loss of generality, we assume that the preferred channel is channel $l = 1$. Under this policy, if both the channels are unoccupied, channel 1 is used. Table 1 lists the transmitter states depending on the primary user occupancies in the two channels. The Markov chain for the transmitter state $S_T(n)$ can be easily derived from the Markov chains of the underlying channels $S_{PU}^1(n)$ and $S_{PU}^2(n)$ and is shown in Figure 2 (C).

<table>
<thead>
<tr>
<th>$S_{PU}^1 (n)$</th>
<th>$S_{PU}^2 (n)$</th>
<th>$S_T (n)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

TABLE 1: State choices for the preference policy.

B. Upper Bounds on $\alpha$

1) Delayed Side Information at Receiver ($\alpha_{\text{delayed}}$): Delayed side information at the receiver refers to the case where the receiver has knowledge of the transmitter state in the previous time slot, i.e., at time $n$, the receiver chooses the channel to monitor in the next time slot, $S_R(n+1)$, using the knowledge of $S_T(n)$. The decision taken by the receiver is based on the maximum likelihood rule, i.e.,

$$k^* = \arg \max_{k \in \{1, 2\}} \text{Prob} [S_T(n) = k | S_T(n-1)]$$  \hspace{1cm} (5)

The receiver decision rule therefore reduces to picking the non-zero state that has the highest transition probability from $S_T(n)$. Since the receiver has additional information about the previous state at the transmitter, it is easy to see that $\alpha_{\text{delayed}} \geq \alpha$.

When $a = b$, the preference policy yields steady state probabilities independent of $a$, given by $\{\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\}$ for the transmitter states $\{0, 1, 2\}$ respectively. In such a scenario it can be shown [6] that

$$\alpha_{\text{delayed}}(1) = \left\{ \begin{array}{ll}
\frac{3-\sqrt{3}}{4} & a \leq \frac{3-\sqrt{3}}{2} \\
\frac{1}{2} & \frac{3-\sqrt{3}}{2} \leq a \leq \frac{2}{3} \\
\frac{2a}{3} & a \geq \frac{2}{3} \end{array} \right. \hspace{1cm} (6)$$

2) All $Y$ Information ($\alpha_Y$): Consider a scenario where the secondary receiver is able to monitor the received signals on both the channels. At the end of the time slot, however, it is required to output the estimate of the transmitter state in the next time slot, i.e., at the end of time slot $n$, the receiver has knowledge of $Y(n) = \{Y(0), Y(1), \cdots, Y(n)\}$ where $Y(i) = [Y_1(i), Y_2(i)]$ is the received vector at time $i$. Based

on $Y(n)$ it has to determine $S_R(n+1)$. The optimal receiver strategy [6] in this case is to estimate the subsequent state $S_R(n+1)$ to be

$$k^* = \arg \max_{k \in \{1, 2\}} \text{Prob} [S_T(n+1) = k | Y(n)]$$  \hspace{1cm} (7)

The corresponding matching probability $\alpha_Y$ is therefore an upperbound to $\alpha$.

C. Achievable $\alpha$ Bounds

1) Offline Bound: For the offline scheme discussed earlier, the transmitter and the receiver are always matched. The best possible matching probability that can be achieved in this case is equal to the steady state probability of any of the identical channels being free.

For the $(a = b)$ case, $\alpha_{\text{offline}} = 0.5$ [6]. Since $\alpha_{\text{delayed}}$ is an upperbound on $\alpha$, equation (6) shows that for all $\frac{3-\sqrt{3}}{2} \leq a \leq \frac{2}{3}$, the offline scheme is optimal in terms of achieving the best possible matching probability. The corresponding capacity bounds, however, differ in the genie information $H(G)$.

2) Training Bound ($\alpha_{\text{training}}$): Consider the case where the secondary transmitter sends a known training symbol to the receiver once every $N$ time slots (symbol periods). The training symbol is sent to help the secondary receiver track the state information of the secondary transmitter more reliably. Without any loss of generality, we assume that the training symbol sent is the first constellation symbol $q_0$, i.e., $X(kN) = q_0 \forall k \in \{1, 2, \cdots\}$. Based on its state at $n = kN$, $S_R(kN)$ and the received signal $Y(kN)$ the receiver decides either to continue monitoring the same channel or switch to the other channel. Between the training symbol time slots, the receiver is not allowed to change states and monitors the same channel chosen in the previous training symbol time slot. The resulting matching probability (not counting training symbols) can be used to lower bound the capacity.

Note that the matching probability $\alpha_{\text{training}}(N)$ is a function of the training delay $N$. If $N$ is too small, due to the training overhead, $\alpha_{\text{training}}(N)$ is small. On the other hand when $N$ is too large, since the receiver is not allowed to switch states between training symbols, the receiver cannot follow the transmitter state accurately. Consequently the resulting $\alpha_{\text{training}}(N)$ is low. There is therefore a tradeoff between the training delay and the matching probability. We define $\alpha_{\text{training}}^*$ as the maximum possible achievable matching probability and the corresponding optimal delay by $N^*$.

V. SIMULATION RESULTS

For our simulations we assume symmetric Markov chains for the primary user occupancy processes, i.e., $a = b$. We present numerical results which show how the matching probability changes with the loop probability $\bar{a} = (1 - a)$ of the primary user occupancy process. $\bar{a}$ is a measure of the memory present in the primary activity process - higher the $\bar{a}$, higher the memory and slower the change in the primary user state $S_{PU}^i(n)$.

Figure 3 plots the upper and lower bounds on the matching probability as a function of $\bar{a}$ for $Q = 8$ and $\epsilon = 0.01$. We note here that the offline lower bound $\alpha_{\text{offline}}$ and
The corresponding plots for \( Q = 2 \) and \( \epsilon = 0.2 \) are shown in Figure 4. Since the probability of error \( \epsilon = 0.2 \) is higher, the separation between \( \alpha_{\text{delayed}} \) and \( \alpha_Y \) increases. Consider the upperbound plot \( \alpha_Y \) and the achievable lowerbound \( \alpha_{\text{offline}} \). The small gap indicates that for scenarios with high error probabilities \( \epsilon \), the offline scheme is fairly robust since if offers nearly the same matching probability as that possible in the \( \alpha_Y \) upperbound (Equation (6) already shows that the offline scheme is \( \alpha \)-optimal for \( a \in \{0.3821, 0.666\} \)). The matching probability achieved by the training scheme suffers due to the fairly high error probability involved.

Capacity plots can be directly obtained from the plots for the matching probability and the genie bound [9]. A detailed capacity analysis for the preference transmitter strategy and other strategies, numerical capacity results for scenarios where \( a \neq b \) and a study of the influence of the constellation power (which affects the resulting error probability) will be presented in [6].

Fig. 4: Upper and achievable bounds on \( \alpha \) with increasing \( \bar{a} \) for \( Q = 2 \) and \( \epsilon = 0.2 \).

VI. CONCLUSIONS

We explore the capacity of a cognitive radio system in a spectrum pool consisting of two channels with independent and identically distributed occupancy processes. The distributed channel information at either end of the cognitive link can be modeled with multi-state switches at the transmitter and receiver. The formulation of the problem does not allow a direct computation of the capacity using the conventional mutual information maximization. By estimating the probability of the receiver and transmitter being matched to the same state, we derive both upper and lower bounds on the capacity. The bounds can be used to explore the benefits and costs associated with the forward and feedback overheads. The capacity analysis shows that the benefits of spectral pooling are lost in dynamic spectral environments.

REFERENCES