

The Throughput Potential of Cognitive Radio: A Theoretical Perspective

(Invited Paper)

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

Abstract—Cognitive radios are promising solutions to the problem of overcrowded and inefficient licensed spectrum. In this work we explore the throughput potential of cognitive communication. We summarize different cognitive radio techniques that underlay, overlay and interweave the transmissions of the cognitive user with those of the licensed users. Recently proposed models for cognitive radios based on the overlay technique are described. For the interweave technique, we present a ‘two switch’ cognitive radio model and develop inner and outer bounds on the secondary radio capacity. Using the two switch model, we investigate the inherent tradeoff between the sensitivity of primary detection and the cognitive link capacity. With numerical results, we compare the throughputs achieved by the secondary user in the different models.

I. INTRODUCTION

The widespread acceptance of diverse wireless technologies has triggered a huge demand for bandwidth that is expected to grow well into the future. The traditional approach used to ensure co-existence of multiple wireless systems is to split the available bandwidth into frequency bands and auction/allocate them to different licensed (primary) users. This kind of spectrum licensing has created a very crowded spectrum as the FCC’s frequency allocation chart shows [1], with almost all frequency bands already assigned to different primary users for specific purposes. A natural question is to explore if there is any room in the spectrum to accommodate secondary (unlicensed) wireless devices without interfering with the communications of the primary (licensed) users of the spectrum. In a very broad sense, the term ‘cognitive radio’ can be used to refer to various solutions to this problem that seek to *overlay, underlay or interweave* the secondary user’s signals with the primary users’ signals in such a way that the primary users of the spectrum are as unaffected as possible.

In the ‘underlay’ technique, simultaneous primary and secondary transmissions are allowed as in Ultrawideband (UWB) systems. The secondary radio spreads its signal over a bandwidth large enough to ensure that the amount of interference caused to the primary users is within tolerable limits. Due to the interference constraints associated with underlay systems, the underlay technique is only useful for short range communications.

The ‘overlay’ technique also allows concurrent primary and secondary transmissions. In this technique, primary message knowledge at the secondary transmitter is used to perform dirty paper coding in order to mitigate the interference seen by the secondary receiver. The secondary transmitter can also employ this side information to relay the primary signal with a power

large enough to ensure that the SNR at the licensed receiver remains unaffected.

The ‘interweave’ technique is based on the idea of *opportunistic communication* [2]. Recent studies conducted by the FCC [3] and industry [4] show that in spite of the spectrum being overcrowded, a major part of the spectrum is typically underutilized. In other words, there exist frequency voids (referred to as *spectrum holes*) that are not in use by the primary owners and consequently can be used for secondary communication. These spectrum holes change with time, location and geographic location. The secondary radio in this technique, therefore, is an intelligent wireless communication system that periodically monitors the radio spectrum, detects the presence/absence of primary users in the different frequency bands and then opportunistically interweaves the secondary signal through the gaps that arise in frequency and time. Spectrum utilization is thus improved by frequency re-use over the spectrum holes. In this technique, accurate detection of the presence of primary systems, especially in low SNR scenarios, is critical to cognitive radio operation: some interesting results in this area can be found in [5].

The underlay technique is usually associated with UWB and spread spectrum technologies. While cognitive radio is most commonly identified with the interweave technique [2], [6], recent literature [7]–[9] considers cognitive communication using the overlay approach. In this work, we are interested in the throughput potential of cognitive radio technology as revealed by the recent theoretical studies in [7]–[10]. We begin with a discussion of the overlay models presented in [7], [9].

II. OVERLAY MODELS

The overlay technique permits the secondary system to transmit simultaneously with the primary user. Consider the communication scenario shown in Figure 1(a), where the primary transmitter (PT) and secondary transmitter (ST) wish to communicate over the same frequency band with the primary receiver (PR) and the secondary receiver (SR), respectively. All the channel gains are known to both the transmitters and both the receivers. The defining assumption made in the overlay models [7], [9] is that the secondary transmitter has non-causal knowledge of the primary user’s transmissions, i.e., the primary message W_1 is known a priori to the secondary transmitter. In such a scenario, there are two interesting strategies the secondary transmitter can pursue [7]–[9]. We discuss the ideas behind both these approaches.

- **Selfish approach:** This is a greedy approach - the secondary transmitter uses *all* the available power to send its

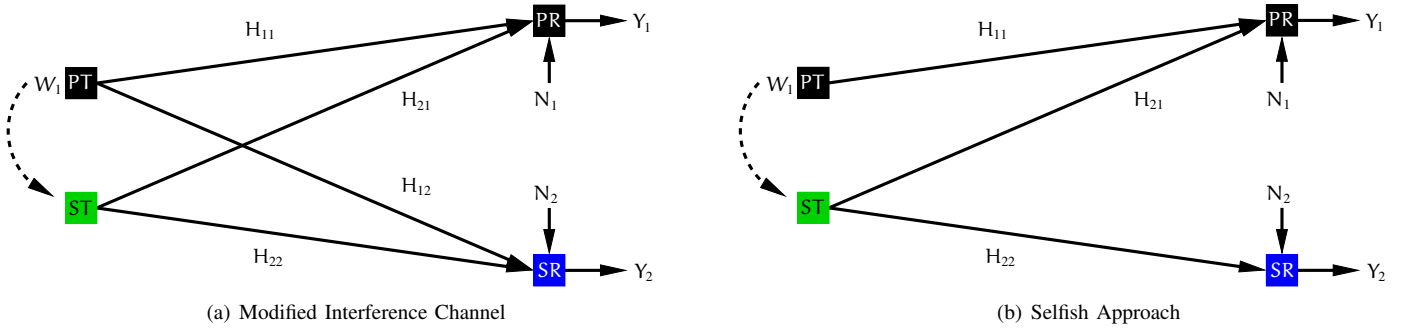


Fig. 1: **PT** and **PR** represent the primary transmitter and receiver. **ST** and **SR** represent the secondary counterparts.

own message to the secondary receiver. The primary message knowledge at the secondary transmitter is used to effectively null the interference at the secondary receiver by using dirty paper coding [7]. Therefore the secondary user maximizes its own throughput without any concern about the interference caused to the primary receiver, as shown by the equivalent system model in Figure 1(b). While the selfish approach violates the cognitive radio principle of protecting the primary users, it provides a theoretical upperbound on the maximum throughput achievable by the secondary users.

- **Selfless approach:** In this approach, the secondary transmitter uses a part of its power to *relay* the primary user's message to the primary receiver. The remaining power is used to transmit the secondary user's message. The power split is chosen such that the increase in the primary user's SNR due to the relaying is exactly balanced by the decrease in its SNR due to interference caused by secondary transmissions, i.e., the SNR at the primary receiver remains the same with or without the secondary user [9]. The primary receiver is therefore virtually unaware of the existence of the secondary user. Further, the secondary transmitter uses dirty paper coding on its own message to eliminate interference at the secondary receiver. Figure 2 shows the equivalent system model. The capacity of the secondary user in the low interference gain regime ($|H_{21}| \leq |H_{22}|$) is characterized in [9].

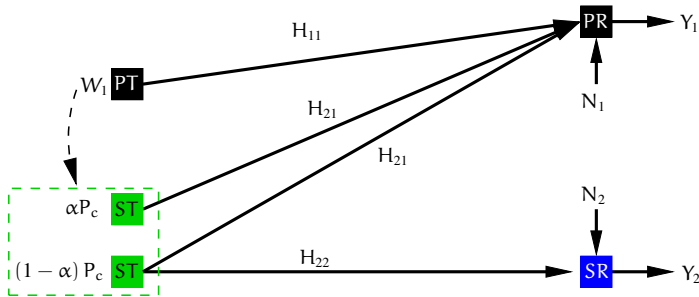


Fig. 2: Overlay Model, Selfless Approach.

The utility of overlay models lies in the fact that they characterize the ultimate limits of cognitive radio when the secondary user has access to side information and sophisticated coding techniques.

III. THE INTERWEAVE MODEL

Non-causal knowledge of the interference is difficult to obtain when the transmitters are not in close proximity of each other or do not share codebooks. In such scenarios, the overlay techniques are invariably associated with interference at the primary receiver, which is not desired. The interweave technique, on the other hand, completely *avoids this interference* by allowing the secondary user to transmit only over spectral segments unoccupied by the primary radios. In this section, the *two switch* interweave model we propose in [10] is described.

A. Two Switch Model

A mathematical model for cognitive radio links can be obtained from a conceptual understanding of the interweave technique. Consider a secondary transmitter (ST) and a secondary receiver (SR) in the presence of primary users (PU) A, B and C located as shown in Figure 3(a). It is assumed that the secondary transmitter and receiver can detect primary transmissions perfectly within their respective *sensing regions* represented by the dotted regions in Figure 3(a). The cognitive transmitter ST can therefore only sense whether or not primary users A or B are active, i.e., ST detects spectral holes when both A and B are inactive. Similarly, the cognitive receiver SR detects spectral holes when both B and C are inactive. Therefore, the spectral holes (communication opportunities) detected at the secondary transmitter and receiver are not identical.

The conceptual model of Figure 3(a) reveals two fundamental properties of the underlying spectral environment:

- **Distributed:** As seen from Figure 3(a), the primary user activity detected in the vicinity of the cognitive transmitter differs from that detected around the cognitive receiver. Further, the secondary transmitter ST does not automatically have full knowledge of the primary user activity in the vicinity of the receiver SR and vice versa. The larger the physical separation between the secondary transmitter and receiver, the lesser the overlap in their respective sensing regions, the more *distributed* the spectral environment, and consequently the higher the uncertainties at the transmitter and receiver.
- **Dynamic:** The primary users' activity is also *dynamic* - over time, different primary users can become active/inactive in different segments of the spectrum. Therefore the primary user activity sensed at the secondary

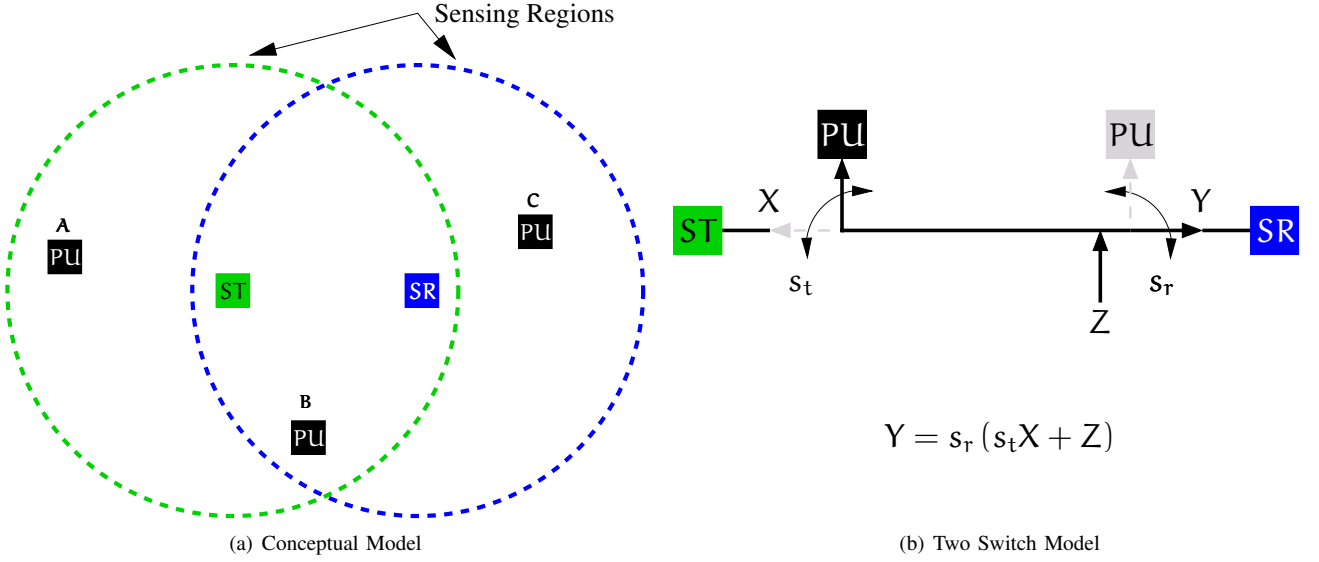


Fig. 3: The different perspectives on local spectral activity at cognitive radio transmitter **ST** and receiver **SR** are depicted in 3(a). Nodes marked A, B and C represent the primary users (**PU**) of the spectrum. The dotted circles represent the corresponding sensing regions. Figure 3(b) represents the corresponding two switch model where the primary user occupancy processes are captured in the switch states s_t and s_r .

transmitter and receiver change with time. This increases the uncertainty at either end of the link about the communication opportunities sensed at the other end. As the primary users become more dynamic, the spectral activity changes faster and is consequently less predictable.

The conceptual model of Figure 3(a) can be reduced to a *two switch* mathematical model shown in Figure 3(b). The communication opportunities sensed at the secondary transmitter are modeled using a two-state switch $s_t \in \{0, 1\}$. The transmitter switch state $s_t = 0$, i.e., the transmitter switch is open, whenever cognitive transmitter perceives that a primary user is active in its sensing region. Transmission can take place only when $s_t = 1$, i.e., when the switch s_t is closed. Similarly the spectral activity sensed at the receiver is captured in the switch s_r . The switch s_r is closed ($s_r = 1$) when the receiver SR detects no primary user activity in its sensing region. The receiver discards the channel output ($s_r = 0$) when it is not believed to be a communication opportunity (when a primary user is present in its sensing region).

The *correlation* between the transmitter state s_t and the receiver state s_r is a measure of the distributed nature of the system - if the transmitter and receiver are far apart, the more distributed the primary activity and therefore the lower the correlation. The dynamic nature of the primary user activity is reflected in the rate at which the switches change state.

B. Capacity of the Two Switch Model

The relationship between the input signal X at the secondary transmitter and the signal output Y at the secondary receiver is described in the following equation:

$$Y = s_r (s_t X + Z), \quad (1)$$

where N is the additive white Gaussian noise at the secondary receiver. We consider an average power constraint of P at the transmitter. A block static model with a coherence interval of

T_c is assumed for the primary user activity, i.e., the switches at the secondary transmitter and receiver retain their state for a period of T_c channel uses (one block) after which they change to an i.i.d state.

Notice that the knowledge of both the switch states s_t and s_r completely characterizes the underlying channel. However, s_t is known only to the secondary transmitter and s_r only to the secondary receiver, i.e., the secondary transmitter and receiver only have *partial* channel knowledge. Cognitive radio therefore corresponds to communication with *partial side information*.

Capacity expressions with partial side information involve an input distribution maximization that is difficult to solve [10]. We instead provide tight upper and lower bounds on the capacity of the two switch cognitive radio channel.

1) *Capacity upperbound*: An upperbound on the capacity can be obtained by assuming additional side information at the receiver provided by a hypothetical genie. Suppose we assume full side information at the receiver, i.e., that the receiver has knowledge of the switch states of *both* the transmitter s_t and the receiver s_r . The transmitter is still assumed to know only s_t . It can be easily shown that Gaussian inputs are optimal in this case and the capacity is given by:

$$C_{s_t, \star}(P) = \text{Prob}[s_t = s_r = 1] \log \left(1 + \frac{P}{\text{Prob}[s_t = 1]} \right) \quad (2)$$

2) *Capacity lowerbound*: The results of [11] show that, interestingly, a genie argument can also be used to obtain lowerbounds on the capacity. Consider the cognitive transmitter receiver pair of Figure 3(b). Suppose a genie provides some amount of side information to the receiver every channel use in a genie variable G . The genie bound result [11] proves that the improvement in capacity due to the genie information G provided to the receiver cannot exceed the entropy rate of the

genie information itself, i.e.,

$$C_{s_t, (s_r, G)} - C_{s_t, s_r} \leq \mathcal{H}(G | s_r), \quad (3)$$

where $\mathcal{H}(G | s_r)$ is the entropy rate of the genie information G given the receiver state s_r . The genie bound can be used to provide bounds on the capacity of the cognitive link of Section III-A. Suppose the genie provides the receiver with the transmitter state s_t once every T_c channel uses in the genie variable G . Since the receiver has knowledge of both the transmitter state and receiver state, we have $C_{s_t, (s_r, G)} = C_{s_t, *}$. The amount of genie information provided to the receiver is no more than 1 bit every T_c channel uses. Consequently we have

$$\begin{aligned} C_{s_t, s_r}(P) &\geq C_{s_t, *}(P) - \mathcal{H}(G | s_r) \\ &= C_{s_t, *}(P) - \mathcal{H}(s_t | s_r) \\ &\geq C_{s_t, *}(P) - \frac{1}{T_c} \end{aligned} \quad (4)$$

An achievable bound on the capacity based on Gaussian inputs can also be obtained [10] and is found to be fairly tight with the upperbound of Section III-B.1 even in highly dynamic scenarios ($T_c = 3$). As T_c increases, equation (4) establishes that the genie lower bound quickly approaches the capacity with full information at the receiver. In the sequel, we will therefore assume that the capacity of the two switch channel model is given by equation (2).

C. Sensitivity of Primary User Detection

To explore the tradeoff between the sensitivity of primary user detection and the capacity of the cognitive radio link, we consider a secondary transmitter and secondary receiver separated by a distance d as shown in Figure 4. The locations of the primary users in the system are captured by a Poisson point process with a density of λ primary nodes per unit area, i.e., the probability of finding k primary in an area $\mathbb{A} \subset \mathbb{R}^2$ is given by

$$\text{Prob}[k \text{ nodes in } \mathbb{A}] = \text{Prob}[N(\mathbb{A}) = k] = \frac{e^{-\lambda \mathbb{A}} (\lambda \mathbb{A})^k}{k!}. \quad (5)$$

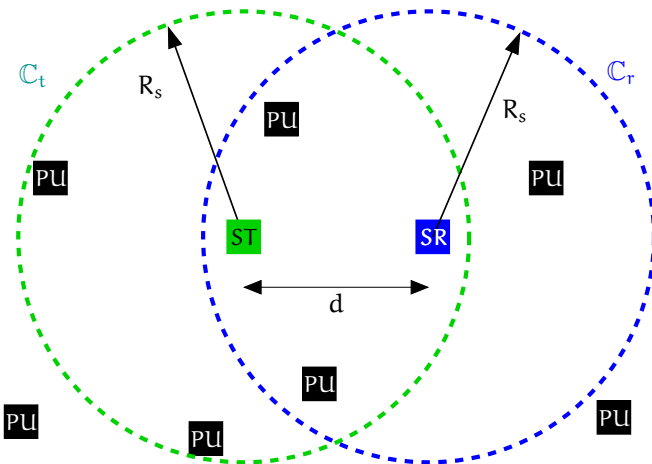


Fig. 4: Sensing regions of radius R_s around the secondary transmitter and receiver.

We assume two-way communication between the primary nodes, i.e. that every primary node functions as both a

transmitter and receiver. The sensing regions at the secondary transmitter and receiver (denoted by \mathbb{C}_t and \mathbb{C}_r in Figure 4) are assumed to be circles of radius R_s centered around ST and SR respectively. We assume perfect detection of primary users within the sensing regions. The radius R_s is a measure of the sensitivity of primary user detection and is decided by the amount of interference tolerable at the primary nodes.

The probabilities required in the calculation of the capacity (equation (2)) are determined as follows:

$$\begin{aligned} \text{Prob}[s_t = 1] &= \text{Prob}[\text{No PUs within ST's sensing region}] \\ &= \text{Prob}[N(\mathbb{C}_t) = 0] \\ &= \text{Prob}[N(\pi R_s^2) = 0] \\ &= e^{-\lambda \pi R_s^2}. \end{aligned} \quad (6)$$

Similarly, we have

$$\begin{aligned} \text{Prob}[s_t = s_r = 1] &= \text{Prob}[\text{No PUs within ST's and SR's sensing regions}] \\ &= \text{Prob}[N(\mathbb{C}_t \cup \mathbb{C}_r) = 0] \\ &= e^{-\lambda \left(2R_s^2 \left(\pi - \cos^{-1} \left(\frac{d}{2R_s} \right) \right) + d R_s \sqrt{1 - \frac{d^2}{4R_s^2}} \right)} \end{aligned} \quad (7)$$

Substituting equations (6) and (7) into equation (2), the capacity of the secondary link is given by

$$C = e^{-\lambda \left(2R_s^2 \left(\pi - \cos^{-1} \left(\frac{d}{2R_s} \right) \right) + d R_s \sqrt{1 - \frac{d^2}{4R_s^2}} \right)} \log \left(1 + P e^{\lambda \pi R_s^2} \right), \quad (8)$$

where P is the power constraint at the secondary transmitter.

Figure 5 plots the secondary user throughput against the radius of the sensing regions R_s for different primary user densities λ . As R_s increases, the sensitivity of detection increases, the average number of communication opportunities decreases resulting in a lower throughput as expected. The same is true as λ increases. Similar behavior is also observed

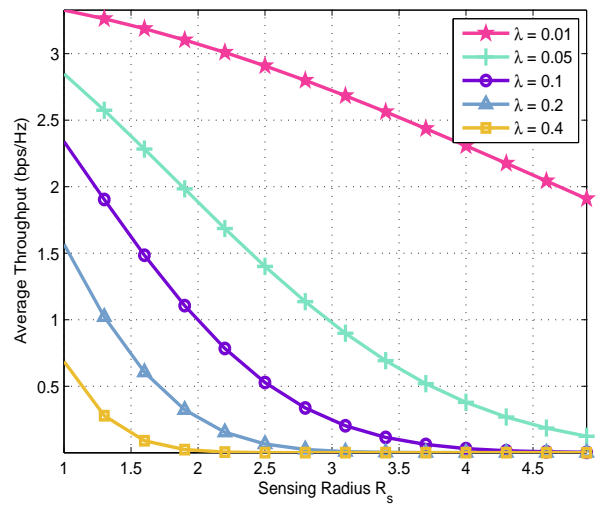


Fig. 5: Throughput vs. sensing radius for different values of λ . We assume $P = 1$ and $d = 1$.

even in cases where the primary user detection is not perfect.

IV. OVERLAY VS. INTERWEAVE: A QUANTITATIVE COMPARISON

In this section, we present some numerical results comparing the theoretical performance limits of the overlay and interweave cognitive models discussed previously. Consider a communication scenario with the primary and secondary transmitter-receiver pairs located as shown in Figure 6. For every link in Figure 6, we assume path loss of the form d^{-4} and unit variance AWGN noise. The channel gains are assumed to be known to all the nodes at all instants. The primary user activity follows an i.i.d Bernoulli process with an average on-time of 40%. We consider a short term power constraint of $P_p = 10$ at the primary transmitter and P_s at the secondary transmitter. For the sake of simplicity, primary user detection is assumed to be perfect (error free).

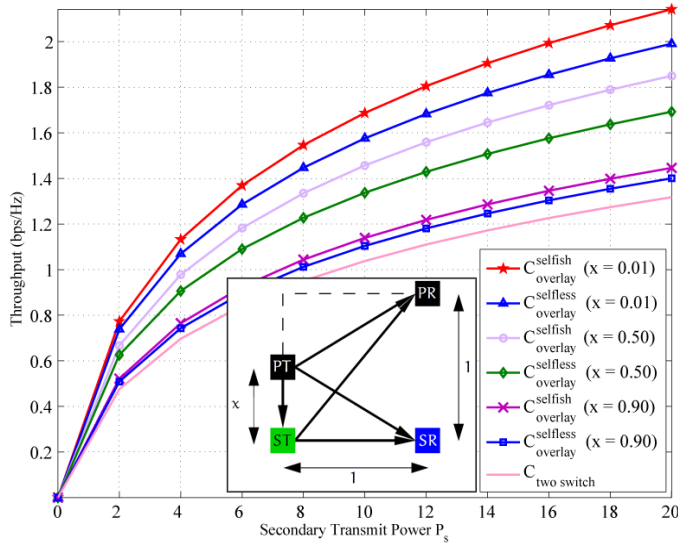


Fig. 6: Throughput comparison of the overlay and interweave models. The inset shows the communication model considered.

Recall that the overlay technique requires non-causal knowledge of the primary message, which is obtained as follows. We focus on the case where the primary and secondary transmitters are located in close proximity to one another ($x \leq 1$). The capacity, C_{ps} of the PT–ST link is then higher than the capacity, C_{pp} , of the PT–PR link. The secondary transmitter can therefore decode the primary message in a fraction $\nu = \frac{C_{pp}}{C_{ps}}$ of the time it takes the primary receiver to decode the same message. Therefore for a fraction $(1 - \nu)$ of the time, the primary user’s transmitted signal is non-causally known to the secondary transmitter. The cost of acquiring this non-causal interference knowledge is the fraction of time ν that must be spent listening to the primary user’s transmissions. The secondary user’s throughput in the overlay models is therefore scaled by the fraction $(1 - \nu)$ to account for this overhead.

The throughput performance of the secondary user is compared for three cognitive radio models - the two overlay approaches of Section II and the interweave approach of Section III. Since the transmitter in the two switch interweave model does not transmit when the primary user is active, the achievable throughput $C_{two\ switch}$ is independent of x . The potential throughput improvements from interference knowledge and dirty paper coding techniques are captured

by the $C_{selfless_overlay}$ and $C_{selfish_overlay}$ curves, which depend on the distance x between the primary and secondary transmitters. As $x \rightarrow 1$, $C_{pp}(x) \rightarrow C_{ps}$ and $1 - \nu(x) \rightarrow 0$. The fraction of overlay transmission time therefore decreases to 0 and both $C_{selfless_overlay}$ and $C_{selfish_overlay}$ approach $C_{two\ switch}$. However, when the primary and secondary transmitters are located close to each other ($x \approx 0$), the secondary transmitter is able to obtain the primary message sooner and therefore $C_{selfless_overlay}$ and $C_{selfish_overlay}$ increase. Since all the available power in the selfish approach is used for secondary transmissions, the $C_{selfish_overlay}$ curves represent an upperbound on the secondary user’s capacity. The throughput improvement of the overlay scheme over interweave techniques quickly disappears as x increases.

V. CONCLUSION

We provide an overview of different techniques to cognitive radio that underlay, overlay and interweave secondary transmissions with the primary users’ signals. Models for cognitive radio links based on these techniques are studied. Numerical results comparing the throughputs of the different cognitive radio models show that the overlay technique can increase the throughput of secondary communications significantly over the interweave technique. This improvement, however, is critically dependent on the availability of interference knowledge at the secondary transmitter and quickly disappears as the distance between the primary and secondary transmitters increases.

REFERENCES

- [1] National Telecommunications and Information Administration (NTIA), “FCC Frequency Allocation Chart,” 2003. Download available at www.ntia.doc.gov/osmhome/allocchr.pdf.
- [2] Joseph Mitola, “Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio,” *PhD Dissertation, KTH, Stockholm, Sweden*, December 2000.
- [3] Federal Communications Commission Spectrum Policy Task Force, “Report of the Spectrum Efficiency Working Group,” *Technical Report 02-135*, no. November, 2002. Download available at http://www.fcc.gov/sptf/files/SEWGFFinalReport_1.pdf.
- [4] Shared Spectrum Company, “Comprehensive Spectrum occupancy measurements over six different locations,” August 2005. Download available at http://www.sharespectrum.com/?section=nsf_summary.
- [5] Anant Sahai, Nigel Hoven, Shrishar Mubaraq Mishra, and Rahul Tandra, “Fundamental Tradeoffs in Robust Spectrum Sensing for Opportunistic Frequency Reuse,” *Technical Report*, March 2006. Available online at <http://www.eecs.berkeley.edu/~sahai/Papers/CognitiveTechReport06.pdf>.
- [6] Simon Haykin, “Cognitive Radio: Brain-Empowered Wireless Communications,” *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 201 – 220, February 2005.
- [7] Natasha Devroye, Patrick Mitran and Vahid Tarokh, “Achievable Rates in Cognitive Radio Channels,” *IEEE Transactions on Information Theory*, vol. 52, pp. 1813–1827, May 2006.
- [8] Natasha Devroye, Patrick Mitran and Vahid Tarokh, “Limits on Communications in a Cognitive Radio Channel,” *IEEE Communications Magazine*, vol. 44, pp. 44 – 49, June 2006.
- [9] Aleksandar Jovičić and Pramod Viswanath, “Cognitive Radio: An Information-Theoretic Perspective,” *Submitted to the IEEE Transactions on Information Theory*, April 2006. Available online at <http://www.ifp.uiuc.edu/pramodv/pubs/JV06.pdf>.
- [10] Syed Ali Jafar and Sudhir Srinivasa, “Capacity Limits of Cognitive Radio with Distributed and Dynamic Spectral Activity,” *To appear in the IEEE Journal on Selected Areas in Communications*, First Quarter 2007.
- [11] Syed Ali Jafar, “Capacity with Causal and Non-Causal Side Information - A Unified View,” *Submitted to the IEEE Transactions on Information Theory*, November 2005.