Breaking Spectrum Gridlock with Cognitive Radios: An Information Theoretic Perspective[‡]

Andrea Goldsmith*, Syed Ali Jafar†, Ivana Marić* and Sudhir Srinivasa†
Email: *andrea@systems.stanford.edu, syed@uci.edu, ivanam@wsl.stanford.edu, sudhirs@uci.edu*

Abstract—**Cognitive radios hold tremendous promise for increasing spectral efficiency in wireless systems. This paper surveys the fundamental capacity limits and associated transmission techniques for different wireless network design paradigms based on this promising technology. These paradigms are unified by the definition of a cognitive radio as an intelligent wireless communication device that exploits side information about its environment to improve spectrum utilization. This side information typically comprises knowledge about the activity, channels, codebooks and/or messages of other nodes with which the cognitive node shares the spectrum. Based on the nature of the available side information as well as a priori rules about spectrum usage, cognitive radio systems seek to underlay, overlay or interweave the cognitive radios' signals with the transmissions of noncognitive nodes. We provide a comprehensive summary of the known capacity characterizations in terms of upper and lower bounds for each of these three approaches. The increase in system degrees of freedom obtained through cognitive radios is also illuminated. This information theoretic survey provides guidelines for the spectral efficiency gains possible through cognitive radios, as well as practical design ideas to mitigate the coexistence challenges in today's crowded spectrum.**

I. INTRODUCTION

Radio is a broadcast medium, and thus all users coexisting in the same frequency band interfere with each other. As the number of wireless systems and services has grown exponentially over the last two decades, the availability of prime wireless spectrum has become severely limited. This is evident by a glance at the NTIA's frequency allocation chart [1], which reveals that almost all frequency bands have been assigned, and there is little new bandwidth available for emerging wireless products and services. Out of this spectrum shortage was born the idea for cognitive radios. These devices utilize advanced radio and signal processing technology along with novel spectrum allocation policies to support new wireless users operating in the existing crowded spectrum, without degrading the performance of entrenched users. If successful, this technology could revolutionize the way spectrum is allocated worldwide as well as provide sufficient bandwidth to support the demand for higher quality and higher data rate wireless products and services well into the future.

A cognitive radio must collect and process information about coexisting users within its spectrum, which requires advanced sensing and signal processing capabilities. Technological advances to support these capabilities are either here today or on the horizon, and thus do not form a major barrier to success. The larger barrier is the requirement for significant changes in the way wireless spectrum is currently allocated to enable cognitive techniques. Unfortunately, the regulatory bodies governing spectrum allocation have not shown much appetite for change since their inception in the early 1900s. In particular, until recently spectrum regulatory bodies such as the Federal Communications Commission (FCC) in the US or the European Telecommunications Standards Institute (ETSI) in Europe always allocated spectral frequency blocks for specific uses, and assigned licenses for these blocks to specific groups or companies. This *divide and set aside* approach involves (a) dividing the spectrum into distinct bands, each defined over a range of frequencies; (b) assigning specific communication uses to specific bands, and (c) determining a licensee for each band, who is generally granted exclusive use of the band. Examples of licensed frequency bands today are the radio and television bands, cellular and satellite bands, and the air traffic control bands. The main advantage of the licensing approach is that the licensee completely controls its assigned spectrum, and can thus unilaterally manage interference between its users and hence their quality-of-service (QoS). Until the mid 1990s licenses were generally granted free of charge based on an application process. Today most licenses that are not for public safety or military use are granted to the highest bidder in a spectral auction.

In addition to the licensed spectrum, in recent years spectrum has been set aside in specific frequency bands that can be used without a license by radios following a specific set of etiquette rules, such as a maximum power per Hertz or a shared channel access mechanism. The purpose of these *unlicensed bands* is to encourage innovation without the high cost to entry associated with purchasing licensed spectrum through auctions. The unlicensed bands have proven a great vehicle for innovation, and the 2.4 GHz unlicensed band currently hosts systems such as Bluetooth, 802.11b/g/n Wifi, and cordless phones. Unfortunately, the unlicensed bands can be killed by their own success, since the more devices that occupy these bands, the more interference they cause to each other.

Spectrum allocation is not just limited to licensed and unlicensed paradigms. The licensed or unlicensed bands may accommodate many additional wireless devices if these devices can exploit advanced technology to only minimally disrupt the

---

[‡] The ordering of the authors in this paper is alphabetical.

[†] Sudhir Srinivasa and Prof. Syed Ali Jafar are with the Electrical Engineering and Computer Science Department, University of California Irvine, Irvine, CA 92697-2625. Email: sudhirs@uci.edu, syed@uci.edu

[*] Ivana Marić and Prof. Andrea Goldsmith are with the *Wireless Systems Lab, Stanford University, Stanford CA, 94305. Email: ivanam@wsl.stanford.edu, andrea@systems.stanford.edu

communications of coexisting noncognitive devices. *Cognitive Radio* originated in the form of various solutions to this problem that allow cognitive communication with minimal impact on noncognitive users. Since its introduction in [2], the idea of cognitive radio has evoked much enthusiasm, including within the FCC, which tapped a spectrum policy task force to provide new policy recommendations that support cognitive radio innovations [3]. The enthusiasm behind the initial ideas evolved in various directions, leading to a variety of different visions. However, behind these diverse cognitive radio interpretations lies a common defining feature: awareness of its environment [4]. In the terminology of information theory, it is the availability and utilization of *network side information* that defines a cognitive radio, which we formalize as follows:

A cognitive radio is a wireless communication system that intelligently utilizes any available side information about the (a) activity, (b) channel conditions, (c) codebooks or (d) messages of other nodes with which it shares the spectrum.

Based on the type of available network side information along with the regulatory constraints, cognitive radio systems seek to *underlay*, *overlay*, or *interweave* their signals with those of existing users without significantly impacting their communication [5]. In the next section we describe in more detail these three different paradigms for cognitive radio systems. The rest of the paper outlines capacity results and coding strategies for these systems as well as for the interference channel, which forms the building block underlying these systems. In cases where the exact capacity is unknown, we characterize the capacity through upper and lower bounds.

Obtaining the capacity region of a wireless network is an open problem for most networks of interest. Moreover, expressions for these capacity regions or region bounds are often cumbersome and yield little insight. More insight may be obtained by characterizing wireless network capacity in terms of the network *degrees of freedom*. The degrees of freedom (also called the capacity pre-log or the multiplexing gain) of a wireless network provides an approximation to the network sum capacity, i.e. the maximum sum of rates that all users can achieve simultaneously. Specifically, the sum capacity $C_{\sum}(\mathrm{SNR}) = d\log(\mathrm{SNR}) + o(\log(\mathrm{SNR}))$, where d is the network degrees of freedom and $o(\log(\mathrm{SNR}))$ represents the approximation error term, which becomes negligible compared to $\log(\mathrm{SNR})$ as SNR increases. For simplicity, we will use the abbreviated notation $x \doteq y$ to indicate that x and y are equal to within $o(\log(\mathrm{SNR}))$. With this notation, $C_{\Sigma}(\mathrm{SNR}) \doteq d\log(\mathrm{SNR})$, hence a network's degrees of freedom yields its approximate sum rate capacity. This degrees of freedom perspective is especially useful for insight into the fundamental rate limits of wireless networks when exact expressions or bounds for its capacity region are unknown. Thus, we will characterize the network degrees of freedom for many of the cognitive systems we consider.

II. COGNITIVE RADIO NETWORK PARADIGMS

There are three main cognitive radio network paradigms: underlay, overlay, and interweave. The underlay paradigm allows cognitive users to operate if the interference caused to noncognitive users is below a given threshold. In overlay systems the cognitive radios use sophisticated signal processing and coding to maintain or improve the communication of noncognitive radios while also obtaining some additional bandwidth for their own communication. In interweave systems the cognitive radios opportunistically exploit spectral holes to communicate without disrupting other transmissions. We now describe each of these three paradigms in more detail, including the associated regulatory policy as well as underlying assumptions about what network side information is available, and the practicality of obtaining this information.

Underlay Paradigm: The underlay paradigm encompasses techniques that allow communication by the cognitive radio assuming it has knowledge of the interference caused by its transmitter to the receivers of all noncognitive users. In this setting the cognitive radio is often called a *secondary user* which cannot significantly interfere with the communication of existing (typically licensed) users, who are referred to as *primary users*. Specifically, the underlay paradigm mandates that concurrent noncognitive and cognitive transmissions may occur only if the interference generated by the cognitive devices at the noncognitive receivers is below some acceptable threshold. The interference constraint for the noncognitive users may be met by using multiple antennas to guide the cognitive signals away from the noncognitive receivers, or by using a wide bandwidth over which the cognitive signal can be spread below the noise floor, then despread at the cognitive receiver. The latter technique is the basis of both spread spectrum and ultrawideband (UWB) communications. The interference caused by a cognitive transmitter to a noncognitive receiver can be approximated via reciprocity if the cognitive transmitter can overhear a transmission from the cognitive receiver's location. Alternatively, the cognitive transmitter can be very conservative in its output power to ensure that its signal remains below the prescribed interference threshold. In this case, since the interference constraints in underlay systems are typically quite restrictive, this limits the cognitive users to short range communications. While the underlay paradigm is most common in the licensed spectrum (e.g. UWB underlays many licensed spectral bands), it can also be used in unlicensed bands to provide different classes of service to different users.

Overlay Paradigm: The enabling premise for overlay systems is that the cognitive transmitter has knowledge of the noncognitive users' codebooks and possibly its messages as well. The codebook information could be obtained, for example, if the noncognitive users follow a uniform standard for communication based on a publicized codebook. Alternatively, they could broadcast their codebooks periodically. A noncognitive user message might be obtained by decoding the message at the cognitive receiver. However, the overlay model assumes the noncognitive message is known at the cognitive transmitter when the noncognitive user begins its transmission. While this is impractical for an initial transmission, the assumption holds for a message retransmission where the cognitive user hears the first transmission and decodes it, while the intended receiver cannot decode the initial transmission due to fading or interference. Alternatively, the noncognitive user may send its message to the cognitive user (assumed to be close by) prior to its transmission. Knowledge of a noncognitive user's message and/or codebook can be exploited in a variety of ways to either cancel or mitigate the interference seen at the cognitive and noncognitive receivers. On the one hand, this information can be used to completely cancel the interference due to the noncognitive signals at the cognitive receiver by sophisticated techniques like dirty paper coding. On the other hand, the cognitive users can utilize this knowledge and assign part of their power for their own communication and the remainder of the power to assist (relay) the noncognitive transmissions. By careful choice of the power split, the increase in the noncognitive user's signal-to-noise power ratio (SNR) due to the assistance from cognitive relaying can be exactly offset by the decrease in the noncognitive user's SNR due to the interference caused by the remainder of the cognitive user's transmit power used for its own communication. This guarantees that the noncognitive user's rate remains unchanged while the cognitive user allocates part of its power for its own transmissions. Note that the overlay paradigm can be applied to either licensed or unlicensed band communications. In licensed bands, cognitive users would be allowed to share the band with the licensed users since they would not interfere with, and might even improve, their communications. In unlicensed bands cognitive users would enable a higher spectral efficiency by exploiting message and codebook knowledge to reduce interference.

Interweave Paradigm: The 'interweave' paradigm is based on the idea of *opportunistic communication*, and was the original motivation for cognitive radio [2]. The idea came about after studies conducted by the FCC [3] and industry [6] showed that a major part of the spectrum is not utilized most of the time. In other words, there exist temporary space-time-frequency voids, referred to as *spectrum holes*, that are not in constant use in both the licensed and unlicensed bands. These gaps change with time and geographic location, and can be exploited by cognitive users for their communication. Thus, the utilization of spectrum is improved by opportunistic frequency reuse over the spectrum holes. The interweave technique requires knowledge of the activity information of the noncognitive (licensed or unlicensed) users in the spectrum. One could also consider that all the users in a given band are cognitive, but existing users become primary users, and new users become secondary users that cannot interfere with communications already taking place between existing users. To summarize, an interweave cognitive radio is an intelligent wireless communication system that periodically monitors the radio spectrum, intelligently detects occupancy in the different parts of the spectrum and then opportunistically communicates over spectrum holes with minimal interference to the active users. For a fascinating motivation and discussion of the signal processing challenges faced in interweave cognitive radio, we refer the reader to [4].

Table I summarizes the differences between the underlay, overlay and interweave cognitive radio approaches. While underlay and overlay techniques permit concurrent cognitive and noncognitive communication, avoiding simultaneous transmissions with noncognitive or existing users is the main goal in the interweave technique. We also point out that the cognitive radio approaches require different *amounts* of side information: underlay systems require knowledge of the interference caused by the cognitive transmitter to the noncognitive receiver(s), interweave systems require considerable side information about the noncognitive or existing user activity (which can be obtained from robust primary user sensing) and overlay systems require a large amount of side information (non-causal knowledge of the noncognitive user's codebook and possibly its message). Apart from device level power limits, the cognitive user's transmit power in the underlay and interweave approaches is decided by the interference constraint and range of sensing, respectively. While underlay, overlay and interweave are three distinct approaches to cognitive radio, hybrid schemes can also be constructed that combine the advantages of different approaches. For example, the overlay and interweave approaches are combined in [7].

Before launching into capacity results for these three cognitive radio networks, we will first review capacity results for the interference channel. Since cognitive radio networks are based on the notion of minimal interference, the interference channel provides a fundamental building block to both the capacity as well as encoding and decoding strategies for these networks.

III. INTERFERENCE CHANNELS: AN OVERVIEW

The interference channel model [8], [9] captures scenarios in which multiple terminal pairs wish to communicate simultaneously in the presence of mutual interference. The users are not assumed to be cognitive - they do not monitor the activity or decode messages of other users. However, it is commonly assumed that all terminals know the channel gains and the codebooks of all the encoders. The communication problem is to determine the highest rates that can simultaneously be achieved with arbitrarily small error probability at the desired receivers, i.e., to determine the *capacity region*. This performance can serve as a benchmark to evaluate the gains of cognition. Even for the smallest interference network consisting of two transmitter-receiver pairs, this problem has remained unsolved for more than thirty years, emphasizing that one of the fundamental problems in networks - coping with and exploiting interference - is not yet entirely understood. Still, there has been a lot of progress in understanding communications in interference channels. We review some of these results next, focusing on the two-transmitter, two-receiver scenario, shown in Fig. 1.

We assume that each encoder t for $t = 1, 2$ wishes to send one of M_t messages, denoted W_t, to its receiver at rate R_t. To do so, an encoder makes a codeword of length n and transmits at rate $R_t = \log M_t / n$. In an interference channel, transmission at each user is affected by a random perturbation of the

Underlay	**Overlay**	**Interweave**
Channel Side Information: Cognitive (secondary) transmitter knows the channel strengths to noncognitive (primary) receiver(s).	Message Side Information: Cognitive nodes know channel gains, codebooks and possibly the messages of the noncognitive users.	Activity Side Information: Cognitive user knows the spectral holes in space, time, or frequency when the noncognitive user is not using these holes.
Cognitive user can transmit simultaneously with noncognitive user as long as interference caused is below an acceptable limit.	Cognitive user can transmit simultaneously with noncognitive user; the interference to noncognitive user can be offset by using part of the cognitive user's power to relay the noncognitive user's message.	Cognitive user transmits simultaneously with a noncognitive user only in the event of a false spectral hole detection.
Cognitive user's transmit power is limited by the interference constraint.	Cognitive user can transmit at any power, the interference to noncognitive users can be offset by relaying the noncognitive user's message.	Cognitive user's transmit power is limited by the range of its spectral hole sensing.

TABLE I: Comparison of underlay, overlay and interweave cognitive radio techniques.

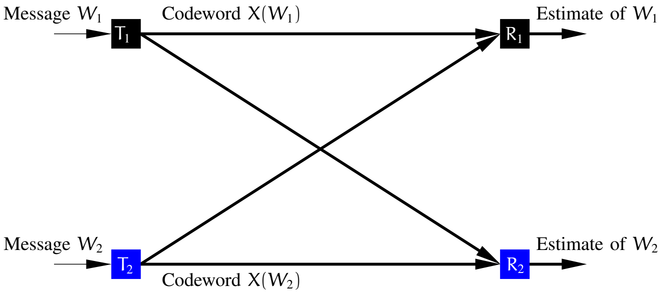


Fig. 1: The Interference Channel.

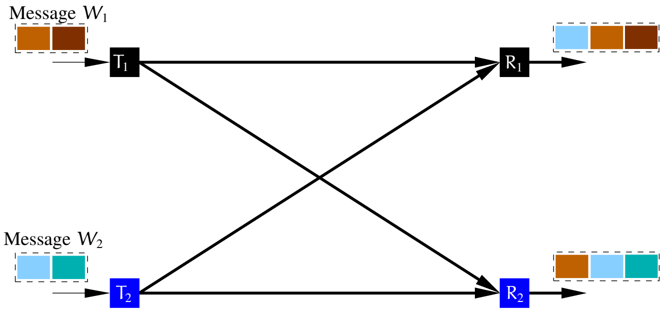


Fig. 2: Rate-splitting. Each encoder splits its message into two sub-messages. User 1 and 2 sub-messages are represented with shades of brown and blue, respectively. A decoder decodes one sub-message of the other user.

channel and the transmission of the other user. Information-theoretic analysis towards obtaining the capacity region of the interference channel (or any channel) is typically performed in two steps:

1) Propose a specific encoding and decoding scheme, and evaluate its achievable rate region.
2) Determine an outer bound to the rate region that cannot be exceeded by any encoding scheme.

If the two bounds meet, then the capacity region is known and the proposed encoding scheme is capacity-achieving. Depending on the level of interference at the receivers, different regimes can be distinguished. The capacity region is known when the interference is *strong*. In this regime, the received interfering signal component carrying the unwanted message is strong enough so that this message can be decoded. The interfering component can then be canceled from the received signal, allowing for interference-free decoding of the desired information. This strategy leads to the capacity in strong interference [10].

However, in general the interference is not strong enough to allow for decoding of the unwanted message without reducing its rate. In this case, *rate-splitting* [9] can be used at the encoders to allow the receivers to decode a *part* of the unwanted message. Rate-splitting achieves the best rates known today [11]. In this encoding scheme, each encoder divides its message into two sub-messages and encodes them separately. A receiver decodes one sub-message of the other user and cancels a part of the interference. This will increase the rate for his communication, but will lower the rate for the other communicating pair due to the additional decoding requirement. Hence, there is a tradeoff between sending a message only to the desired receiver and allowing partial decoding at the other one. The rate-splitting concept is illustrated in Fig. 2.

These results have been specialized to the interference channel with additive white Gaussian noise. Two outer bounds for the Gaussian interference channel have also been developed in [12]. An insightful overview of the known results for Gaussian interference channels can be found in [13]. In general, there is a gap between the rates achieved with rate-splitting and the outer bounds. For the Gaussian interference channel in the high SNR regime, this gap has recently been tightened in [14]. A new outer bound that gives the sum-capacity in the regime in which treating interference as noise is optimal, i.e. for *noisy-interference*, was developed in [15].

Degrees of Freedom: We now consider the degrees of freedom in a K user interference channel. If there was only one user, he would achieve a capacity on the order of $\log(\mathrm{SNR})$. Thus, this user can increase his rate by increasing his power. However, in the presence of multiple users, if all users try to increase their rates by increasing their transmit power, the users interfere and both the signal and the interference powers increase. Consequently, the signal-to-interference-plus-noise power ratio (SINR) and therefore the users' rates saturate at a constant value. The system is then said to be interference-limited. To eliminate the interference, the users can be orthogonalized along space, frequency, or time dimensions of the channel. In this case user k achieves a rate proportional to $d_k \log(\mathrm{SNR})$, where d_k is the fraction of the channel degrees of freedom (typically in terms of bandwidth or timeslots or spatial streams) allocated to the k^{th} user. The total degrees of freedom of the interference network $d = \sum_{k=1}^{K} d_k$. Thus, the degrees of freedom allocated to a user can be interpreted as the fraction of the channel degrees of freedom that the user is allocated relative to the interference-free scenario when the user has all the channel degrees of freedom to himself. The traditional approach to orthogonal spectrum allocation is the "cake-cutting" approach where each user gets a fraction of the channel bandwidth and the sum of these fractions is equal to one. For the two user interference channel the cake-cutting interpretation is confirmed by theory. Specifically, it is known that the two user interference channel has only one degree of freedom [16]. The sum capacity of the two user interference channel $\log(\mathrm{SNR})$ and therefore, its capacity per user, is only a half of what the user would achieve in the absence of interference. Thus, each user gets half the cake (half the channel).

The cake-cutting interpretation is also consistent with the degrees of freedom results of [17], which suggest that each user gets $1/K$ degrees of freedom when only local channel knowledge is available at each user. In other words, the capacity per user for the K user interference channel is on the order of $C_k \doteq \frac{1}{K} C$, where C is the single user capacity in the absence of interference. In terms of the cake-cutting analogy, each user gets a fraction $1/K$ of the cake, which seems like the obvious division of channel resources.

Somewhat surprisingly, recent results have shown that for interference channels with time varying (or frequency selective) channel coefficients, if global channel knowledge is available, then each user can simultaneously achieve rates of the order of $\frac{1}{2}\log(\mathrm{SNR})$ [18]. The result is counterintuitive since, in terms of the cake-cutting analogy, it implies that everyone gets half the cake even when the number of users is more than two. This result does not contradict [17] since it assumes global, rather than local, channel knowledge. Since global channel knowledge is the underlying assumption of underlay cognitive radio networks, we will discuss this surprising degrees of freedom result in the context of the next section.

IV. UNDERLAY COGNITIVE RADIO

The underlay approach to cognitive radio allows simultaneous cognitive and noncognitive communication under the constraint that the interference caused to the noncognitive user by cognitive user does not degrade its communication. In this section, we overview some of the information theoretic results concerning underlay cognitive communication.

Underlay cognitive radio can be modeled as cognitive communication with certain constraints placed on the signal power received (interference caused) at a noncognitive receiver. The capacity for this underlay system assuming different channel models (SISO AWGN and fading channels, Gaussian multiple access channels) and an average receive power constraint at the noncognitive receiver can be characterized by translating this receive power constraint into a transmit power constraint at the cognitive transmitter [19], [20]. For example, in an AWGN scenario, an average interference (receive power) constraint at the primary receiver is equivalent to a corresponding average transmit power constraint at the secondary transmitter. Gaussian codebooks are optimal in this case and the well known logarithmic relationship between the cognitive user's capacity and its SNR holds (with the noncognitive user's signal being treated as Gaussian noise). In a Gaussian MAC with independent messages at each user, a receive power constraint at the noncognitive receiver reduces to a constraint on the *sum* of transmit powers of the cognitive users. The resulting capacity region will be the union of the capacity regions of different multiple access channels for which the sum of the transmit powers of the different users is a constant. Consider the extension of this model with fading channels between the cognitive transmitter and cognitive receiver, and between the cognitive transmitter and noncognitive receiver (interfering link). With full knowledge of both the channel gains at the cognitive transmitter, Gaussian codebooks achieve capacity [20]. The power adaptation that maximizes the capacity is similar to waterfilling with a non-constant water level that depends on the channel between the cognitive transmitter and noncognitive receiver. When a *peak* interference power constraint at the noncognitive receiver exists, the cognitive transmitter cannot transmit above a certain power level, depending on the interfering link channel. Thus the peak receive power constraint reduces to a peak cognitive transmit power constraint, and the capacity characterization is then straightforward.

In underlay cognitive communication with multiple cognitive and noncognitive users, the cognitive user's sum rate optimization (with interference from other users regarded as noise) can be formulated as a general multiuser communication problem with transmit power constraints (or quality of service (QoS) constraints) at the cognitive transmitters and interference constraints at the noncognitive receivers. We first consider the cognitive sum rate maximization with peak transmit power and peak interference constraints in a scenario where all the secondary transmitters and receivers have full channel knowledge. In the power control strategy that maximizes the secondary sum rate, it is seen that at most one user transmits with a power between zero and its peak transmit power - all the other users either transmit with zero or their peak powers [21]. For a similar model with minimum SINR constraints at the cognitive users and interference constraints at a single noncognitive user, conditions under which a feasible power allocation policy at the cognitive users exists can be derived [22], [23]. In the case when not all of the cognitive users can be supported with their SINR requirements, the different cognitive users can be assigned different spectrum access priorities. While this modified problem is not analytically tractable, some properties of the power control optimization problem can be derived by defining the problem in terms of a spectrum sharing game [23]. Game theoretic concepts have also been applied to investigate power control and spectrum sharing in underlay systems [24]–[29]. The idea is to model the cognitive radios as the different players; the controllable communication parameters (transmit power [25]–[27], the signaling waveform [25], or the channel

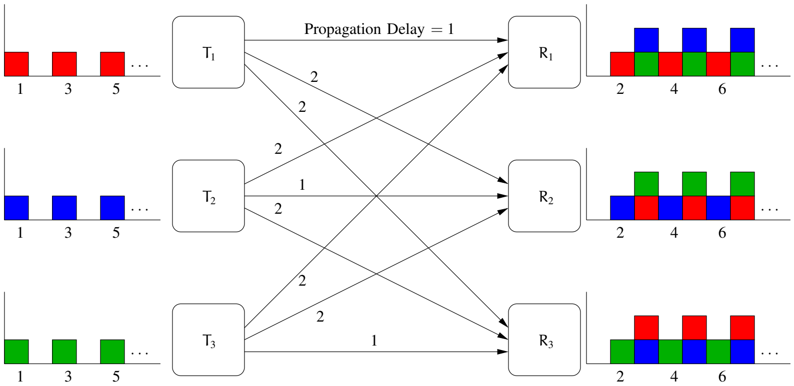


Fig. 3: Interference alignment example where everyone is able to access the channel half the time with no interference (Everyone gets half the cake).

to be used for transmission [27]) as the actions that can be performed by each of the players; and the SINRs (or user rates) of the cognitive radios as the players' utility functions in the game. The outcome of the game is analyzed under the assumption that each player, while influenced by the other players' decisions, acts in its own self interest in such a way as to maximize its utility. The goal is to determine if there exists a convergence point from which any deviation by any player decreases the player's utility (a *Nash equilibrium*). Based on the utility functions of the cognitive radios, different kinds of game models exist [24], [26].

Underlay cognitive systems may also exploit multiple transmit/receive antennas at the cognitive and noncognitive transmitters and receivers. Consider a network with multiple transmit antennas at the cognitive transmitter and single antennas at the noncognitive transmitter and all receivers. Assume that each of the nodes has global channel knowledge. The noncognitive receiver is, as usual, to be protected by an interference constraint. Notice that the channels from the cognitive transmitter to the noncognitive and cognitive receivers are MISO channels. With an interference constraint at the noncognitive receiver, the cognitive transmitter can direct any amount of power in a direction perpendicular to the cognitive transmitter-noncognitive receiver channel *without* interfering with the noncognitive communication. Then as long as the cognitive receiver lies within this perpendicular direction, the cognitive user's capacity is limited by its maximum transmit power, i.e. the noncognitive system imposes no constraint. Conversely, suppose we only constrain the cognitive transmit power and not the interference caused at the noncognitive receiver. Then it is well known that the cognitive user's capacity is maximized when its transmitter beamforms along the direction of its channel [30]. In general the cognitive transmitter will need to direct its available power along one (or more) directions such that the total projection along the channel between the cognitive transmitter and the noncognitive receiver satisfies the interference constraint. Interestingly, it turns out that unit rank beamforming remains the optimal strategy at the cognitive transmitter [31]. This is true even with multiple noncognitive receivers with different interference constraints.

So far we have only constrained received power in terms of an interference constraint on cognitive transmitters. However, different interference waveforms can cause vastly different performance degradation on a particular desired signal, even if received at the same power. How and what kind of constraints should be imposed on the interference generated by cognitive transmissions has a variety of regulatory ramifications. In 2002 the FCC proposed *interference temperature* as the appropriate metric [3], which is a measure of the RF power at an antenna generated by other transmitters and noise sources [32]. This metric was somewhat controversial in terms of how it could be measured at the cognitive transmitter, and whether it would provide sufficient protection for licensed users with a cognitive underlay. We refer the reader to [33] for an interesting discussion on general considerations for specifying an interference constraint in underlay networks.

Degrees of Freedom: Degrees of freedom results for underlay networks reveal interesting insights into the optimal ways to deal with interference in wireless networks. It is shown in [18] that in the K user interference channel with time varying (frequency selective) channel coefficients and global channel knowledge, as the SNR increases every user is able to simultaneously achieve nearly one half of the capacity that he could achieve in the absence of all interference. Thus, everyone gets half the cake.

The key to this counter-intuitive result is the concept of interference alignment [18], [34], [35]. Wireless networks are inherently asynchronous. Unlike centralized (multiple access) networks where signals may be aligned according to the common frame of reference provided by the single receiver, a wireless network has a different frame of reference associated with each receiver. For example, there is a different propagation delay between each transmitter-receiver pair that makes it impossible to align all signals simultaneously at all receivers. The asynchronicity of the network is evident in time, frequency, space, and code dimensions. This "relativity of alignment" is the enabling premise for the novel idea of interference alignment. Since each receiver sees a different picture, it is possible to design signals intelligently in a way that each receiver, from its own perspective, appears privileged relative to other receivers. Interference alignment refers to schemes that design signals so that they cast overlapping shadows at the receivers where they constitute interference and remain distinguishable at the receivers where they are desired.

The canonical example of interference alignment (from [18]) is illustrated in Figure 3. This example is interesting because it shows how the asynchronous nature of the network allows each user to access half the channel (in the time dimension) without any interference. So all users get half the cake. In this example, the propagation delays between a transmitter and receiver are given by the numbers alongside the corresponding links in Figure 3. Notice that the propagation delays are equal to one time slot (symbol duration) for all desired signal paths and two time slots for all paths that carry interference signals. Suppose all transmitters transmit simultaneously only during odd time slots and remain quiet during the even time slots. From Figure 3, we can see that symbols sent from the interfering transmitters are received simultaneously in the odd time slots while the desired signals are received with no interference in the even time slots. Every user is therefore able to access the channel one-half of the time *without* interference from other users. While the example in Figure 3 uses a channel with artificially selected propagation delays to illustrate the idea of interference alignment, the result extends to networks with random time varying channel coefficients even without propagation delays. There is, actually, a capacity penalty with random coefficients since the most suitable channel coefficients for interference alignment cannot be hand-picked. However, the penalty is $o(\log(\mathrm{SNR}))$, so it becomes a negligible fraction of the users' rates as SNR increases [18]. Interestingly, the outer bound on sum capacity is given by $C_\Sigma = KC_k \doteq \frac{K}{2}\log(\mathrm{SNR})$, so that interference alignment is the optimal scheme in terms of maximizing its degrees of freedom.

V. OVERLAY COGNITIVE RADIO

Overlay cognitive radio networks allow concurrent cognitive and noncognitive transmissions, but in contrast to underlay networks, the cognitive transmitter may now facilitate the transmission of the noncognitive user. The smallest overlay cognitive radio network is a two-user (cognitive and noncognitive) interference channel where the cognitive transmitter has non-causal knowledge of the noncognitive user's message, as shown in Fig. 4. This overlay cognitive radio system is also referred to in the information theory literature as an interference channel with asymmetric message knowledge, degraded message sets or one cooperating encoder. In this section, we describe capacity results for this overlay cognitive radio channel model. The case in which the cognitive transmitter learns only a part of the noncognitive user's message is analyzed in [36].

As in Sec. III, we assume that all the codebooks and channel gains are known to the two encoders. This means, for example, that the cognitive user (User 2), knowing the noncognitive user's message W_1, also knows the codeword of the noncognitive user (User 1).

Knowledge of the noncognitive user's message allows the cognitive transmitter to apply several encoding schemes that will improve both its own rates as well as the rates of the noncognitive user [37]–[40]. For example, encoding can be done to achieve a non-zero rate for the noncognitive user such that the noncognitive user's transmission causes no interference to the cognitive receiver [39]. This is but one encoding strategy, and in the next section we describe various encoding schemes that the cognitive transmitter may employ to exploit its knowledge of the noncognitive user's message. As described in Section II, the assumption that the cognitive transmitter knows the message of the noncognitive transmitter at the start of their transmissions is somewhat problematic for practical systems. However, this assumption is very reasonable if the two transmitters are close together, the noncognitive message is being retransmitted after an initial failure, and the cognitive user was able to successfully decode the first transmission. This assumption is also applicable when the noncognitive transmitter sends its message to the cognitive transmitter in advance, which might be done in a separate frequency band. If the two transmitters are close together, little power and bandwidth is needed for this separate message transmission. Both of these scenarios might lead to partial message decoding of the noncognitive message instead of full decoding, which also fits within the overlay network paradigm.

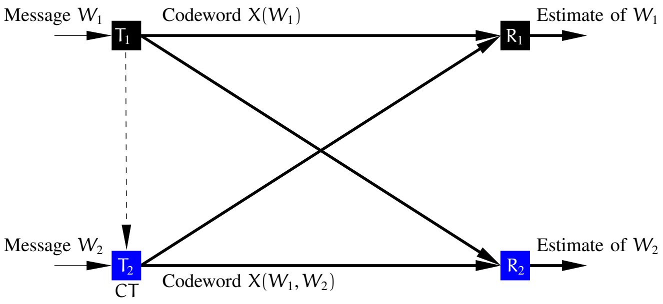


Fig. 4: Interference channel with one cognitive encoder (CT).

A. *Overlay Encoding Techniques*

Overlay encoding techniques have mostly been investigated for the interference channel with one cognitive encoder (Fig. 4), which comprises the most basic overlay network. Note that when neither user is cognitive, i.e. neither user knows the message of the other, this model reduces to the basic interference channel of Sec. III. Another special case of an overlay network occurs when the noncognitive encoder does not transmit. Since the cognitive transmitter knows the messages intended for both receivers, the model reduces to a broadcast channel, whose capacity was first analyzed in [41]. We see from these special cases that the overlay cognitive radio network has elements of both interference channels and broadcast channels. Thus, encoding techniques developed for either of these channel models, or their combinations, may be capacity achieving under certain channel conditions. We now review the various encoding techniques that have been proposed for overlay cognitive networks, which are mostly derived from encoding strategies for the interference channel.

Rate-splitting is the best known encoding technique for interference channels, so it is natural to consider it for overlay encoding as well. This technique was first applied to overlay networks in [37], [42], [43]. However, additional techniques beyond rate-splitting can be used to best exploit the cognitive transmitter's message knowledge. In particular, to improve the rate for the noncognitive communicating pair, the cognitive encoder can *cooperate* by encoding the noncognitive user's message in order to help convey it to the noncognitive decoder. In this way, the cognitive encoder dedicates a fraction of its resources (power) to send the noncognitive user's message W_1 and increase its rate R_2. On the other hand, any signal conveying W_1 is interference to the cognitive encoder's receiver. This interference is known at the cognitive transmitter since it consists of codewords used for W_1. In this setting, the precoding technique referred to as the Gel'fand-Pinsker (GP) binning [44] and, specifically, dirty-paper coding (DPC) [45] in Gaussian channels, can be applied. These techniques allow the cognitive encoder to precode its message at a rate associated with interference-free communication. In fact, as will be explained later in this section, GP binning is crucial for the overlay cognitive radio channel: together with cooperation, it leads to capacity in certain scenarios, [38], [39], [46].

It is not surprising that DPC brings gains in the Gaussian cognitive radio channel. As previously pointed out, if the noncognitive encoder is silent, the model reduces to a broadcast channel from the cognitive encoder to the two receivers, for which dirty-paper coding is the optimal strategy [47], [48]. In general, however, there are two differences at the cognitive encoder from the classical GP setting. First, the interference carries useful information for the noncognitive receiver. Second, the interference is a *codebook* of some rate and can thus have lower entropy than in the GP setting. As shown in [43], the latter can be exploited to achieve a higher rate.

Therefore, although the encoding techniques for overlay cognitive radio certainly borrows from existing strategies for the classical interference channel, a number of additional techniques are needed to fully exploit the knowledge of the noncognitive user's message. The three overlay network encoding strategies that have been investigated in the literature for the network of Fig. 4 are as follows:

- Rate-splitting: This technique improves rates by enabling (partial) interference cancellation at the decoders.
- GP binning and binning against a codebook: the cognitive encoder improves its rate by precoding against interference.
- Cooperation: The cognitive encoder increases the rate of the noncognitive user by (partially) relaying its message to its decoder.

B. Capacity Results

Determining the capacity region for the overlay cognitive network, even for the simple model of Fig. 4, remains an open problem in most cases. However, there are some regimes for which a subset of the three encoding techniques described in the previous section achieve capacity for this model. These regimes, and their capacity-achieving encoding strategies, are enumerated below.

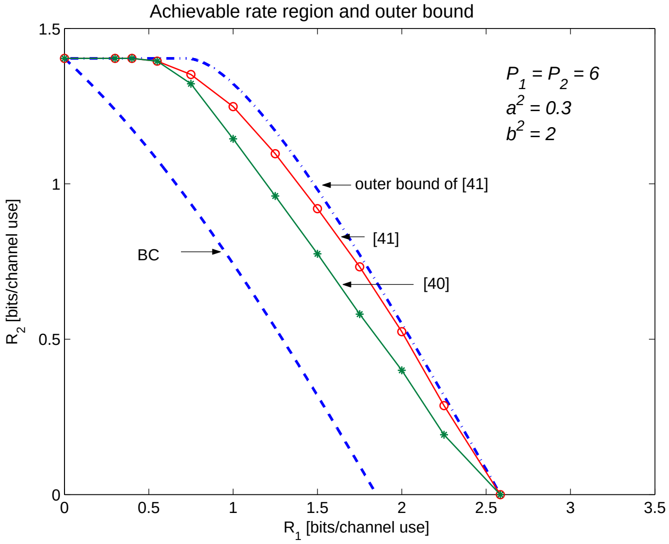


Fig. 5: Achievable rate region (R_1, R_2) in the Gaussian channel. R_1 is the noncognitive user rates and R_2 the cognitive user rates. Star-labeled and circle-labeled curves respectively present performance achieved with the schemes proposed in [42] and [43]. Also shown is an outer bound of [43] and the BC performance.

1) **Strong Interference:** As in the interference channel model, if the interference to both decoders in the cognitive radio network of Fig. 4) is strong, both decoders can decode both messages and cancel interference as if it was not present in the network. Thus, there is no need for rate-splitting or binning, and cooperation achieves capacity [40].
2) **Weak Interference:** If the interference to the noncognitive decoder is weak, this decoder does not need to cancel interference. Interference at the cognitive decoder can be eliminated by binning. Therefore, there is no need for rate-splitting. DPC and cooperation achieve capacity for the Gaussian cognitive channel model [38], [39], [46]. The general case remains an open problem.
3) **Common Information:** If the cognitive decoder decodes both messages there is no need for binning, as there is no interference at that decoder. Thus, rate-splitting and superposition coding achieve capacity [49], [50].

For most regimes capacity of the simple interference channel with one cognitive encoder is still unknown, and achievable rates are based on encoding strategies that combine the above techniques [37], [42], [43]. The relative performance of these various encoding schemes depends on the channel conditions and topology. For the Gaussian channel, a comparison of the achievable rate regions for the encoding schemes proposed in [42] (star curve) and in [43, Thm. 1] (circle curve) is shown in Fig. 5. These achievable rates are also compared to the capacity outer bound of [43], shown with dashed-dotted curve. As discussed above, when the noncognitive user does not transmit, the channel reduces to a broadcast channel (BC), whose capacity region is shown with a dashed curve. For this comparison we assume both users transmit with the same

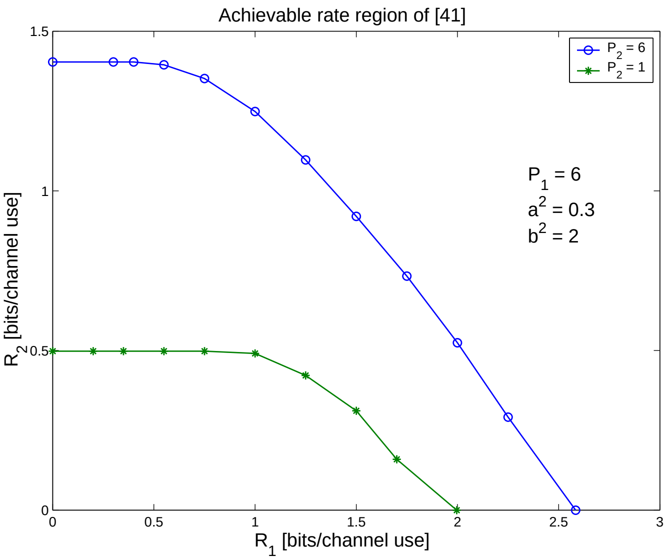


Fig. 6: Achievable rate region (R_1, R_2) in the Gaussian channel for different values of the cognitive user's (User 2's) power.

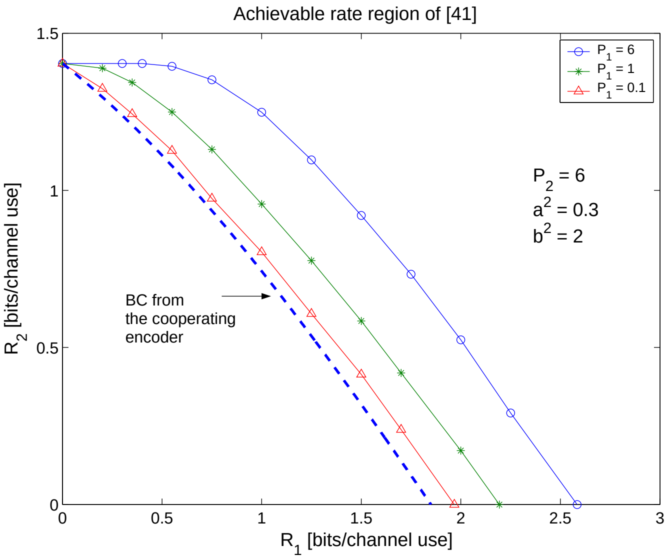


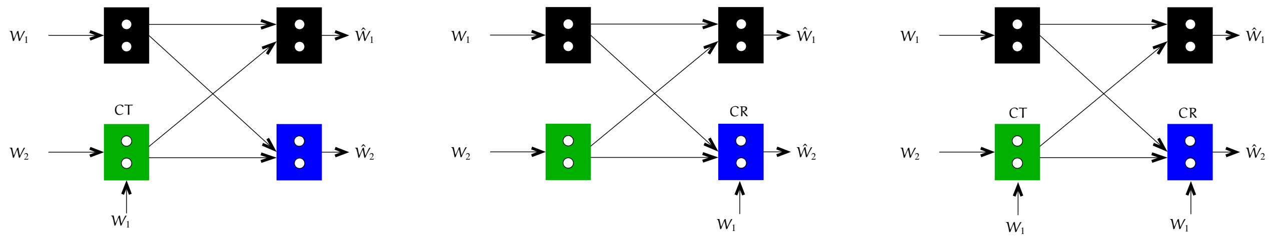
Fig. 7: Achievable rate region (R_1, R_2) in the Gaussian channel for different values of the noncognitive user's (User 1's) power.

power $(P_1 = P_2 = 6)$. Figures 6 and 7 show the impact of changing the power at the cognitive and noncognitive transmitter, respectively. We observe from these figures that reducing power at the cognitive transmitter has a more drastic impact, which is expected since this transmitter knows both messages and can therefore improve rates to both receivers. We also see in Fig. 7 that as the power of the primary transmitter is reduced, the rate region approaches that of the BC (dashed line). This is expected since, as discussed above, the model reduces to a BC when the noncognitive user does not transmit. The rate gains of having one cognitive encoder versus the two noncognitive encoders in the traditional interference channel were evaluated in [37]. The encoding schemes we have described are known to be capacity-achieving under certain assumptions about the channel or specific encoding/decoding constraints. Finding additional regimes where these and other encoding schemes achieve capacity is the topic of much ongoing investigation. The impact of feedback and common information has also been investigated in [51], [52].

In general, characterizing the capacity region of a network with cognitive users that have knowledge of other users' messages, of which the overlay cognitive radio network is but one example, offers both insight and motivation for practical designs based on this new paradigm. The capacity and achievable rate results indicate the best known encoding strategies for such networks, and quantify the rate gains possible due to cognition. Different assumptions and constraints about message knowledge will affect these conclusions, and many of these issues have yet to be investigated. An important question for a large network with many cognitive and noncognitive users is the best protocol for coexistence among the cognitive users. Intuitively some form of cooperation between these users will be required, but the best form of cooperation is unclear. Another interesting question is whether cognition is more beneficial at the transmitter or receiver. In particular, consider the interference channel where, instead of noncognitive message knowledge of W_1 at the cognitive encoder, it is known at the cognitive decoder. This problem has yet to be investigated, but the problem changes significantly since only rate-splitting encoding can be applied, and the cognitive decoder has no interference since it can be subtracted out.

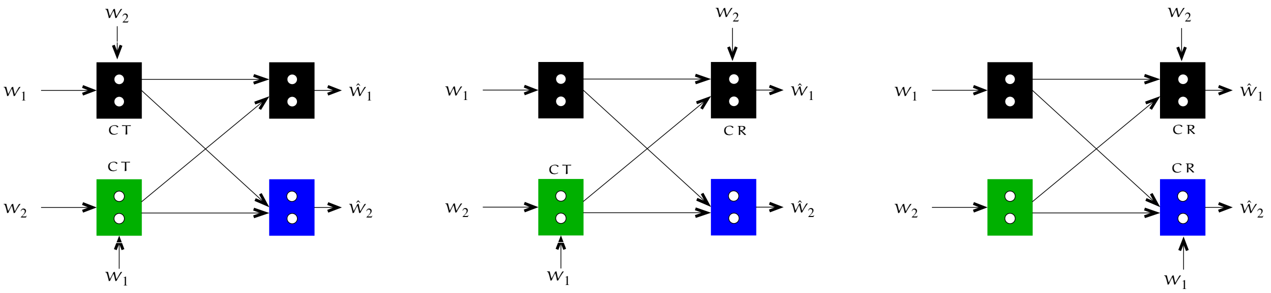
Recent work has also investigated overlay cognitive networks with perfect message knowledge at the cognitive encoder but only partial channel knowledge. Lack of channel knowledge changes the problem considerably. In particular, consider the case where the *phase* of the channel between the noncognitive transmitter and the cognitive receiver is unknown at the cognitive transmitter. The main challenge in this scenario is that Costa's [45] DPC technique cannot be applied directly. Even when the phase uncertainty at the cognitive transmitter is limited to distinguishing between one of two possible phases, the cognitive user's capacity is reduced considerably by this uncertainty. At moderate SINRs, the capacity with this phase uncertainty is almost equal to that obtained by treating the interference as noise [53]. Learning the phase information at the cognitive transmitter, therefore, can yield substantial throughput benefits.

Degrees of Freedom: We conclude this section with results on the degrees of freedom for overlay cognitive communication. The number of degrees of freedom for the two-user overlay cognitive radio channel lies between one (interference channel with no cooperation) and two (2×2 MIMO channel obtained by full cooperation between transmitters and full cooperation between receivers). Fig. 8 shows scenarios where one user has a cognitive transmitter (CT), a cognitive receiver (CR), or both a cognitive transmitter and a cognitive receiver. In these scenarios the number of degrees of freedom for the overlay cognitive radio channel is equal to one [35], [54]. Thus, the sum rate capacity for all three scenarios is $C_\Sigma \doteq \log(\mathrm{SNR})$. The scenario where both transmitters are cognitive, both receivers are cognitive or one users' transmitter is cognitive and the other users' receiver is cognitive is



(a) (b) (c)

Fig. 8: The degrees of freedom if one user has a cognitive transmitter, a cognitive receiver, or both a cognitive transmitter and a cognitive receiver is one; the sum rate capacity is $C_\Sigma \doteq \log(\mathrm{SNR})$.



(a) (b) (c)

Fig. 9: The degrees of freedom if both users have cognitive transmitters, both users have cognitive receives, or one user has a cognitive transmitter and the other user has a cognitive receiver is increased to 2; the sum rate capacity is $C_\Sigma \doteq 2\log(\mathrm{SNR})$.

shown in Fig. 9. In these scenarios the network is able to achieve 2 degrees of freedom [35]. Note that these capacity results assume a genie provides a users' message to another (cognitive) user. The degrees of freedom benefits of this cognition disappear if the cost of the genie is factored in. In other words, it is shown in [16] that if message sharing takes place only through physical channels (no genies involved) then cognitive cooperation does not provide any increase in the network degrees of freedom. This sobering result provides pause for the potential gains of cognition in practical operating environments where message side information is not easily obtained.

The encoding strategies for cognitive overlays involve one-sided cooperation of the cognitive transmitter to help the noncognitive receiver relay its message. This notion of node cooperation can be generalized to a wide range of cooperative strategies in wireless networks. In the next section we discuss in more detail various cooperation methods for wireless networks, their capacity results, and the connection between cognition and cooperation.

C. Cognition and Cooperation

Cooperative techniques allow nodes to relay each other's information to improve network capacity. For example, classic multi-hop relaying in which data from a source to a destination is relayed by an intermediate node is a simple and common method of cooperation in many wireless networks. This form of cooperation requires that the relay node has some information about the source message in order to forward it. This information might be its noisy observation of the signal transmitted by the source, the decoded message from this noisy observation, or a part of the decoded message. In cooperative communication, the relay can obtain this information by assistance of the source node, for example, through block-Markov encoding [55] or simply by listening to the channel. In a cognitive network, a cognitive node can acquire such information in a similar manner, and therefore use the same cooperative encoding approaches.

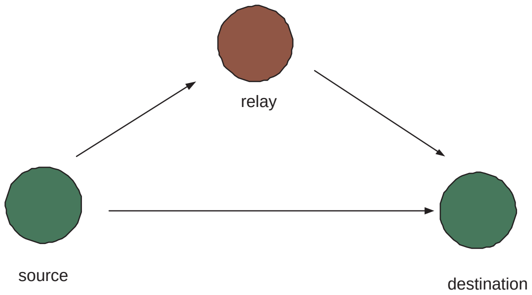


Fig. 10: Relay channel.

The simplest communication scenario with cooperation is the *relay channel* [56], shown in Fig. 10. In this setting, a single communicating pair is helped by a relay node. The relay has no data of its own to send and uses its resources solely to help the communicating pair. In cooperation with the source, the relay can *decode-and-forward (DF)* - the source transmits at the rate such that the relay can decode the message, re-encode it, and transmit it together with the source. Alternatively, the relay can employ *compress-and-forward (CF)* or *amplify-and-forward* to respectively quantize or amplify the received signal and send it to the destination.

These strategies can also be applied to the multiple-relay network where a group of intermediate relays helps forward messages between a given source-destination pair [57]. The above strategies were shown to be optimal in special cases. In general, however, the capacity of the relay channel, this simplest of cooperation networks, has been an open problem since its introduction in the 1970s.

In networks, the presence of multiple communicating pairs adds the additional impairment of interference. We have seen that cognition can be used to avoid or exploit this interference. In interweave communications, cognitive transmitters avoid interference by detecting existing users and transmitting in unused bands. In underlay networks cognitive users also avoid interference by maintaining it below a prescribed threshold. However, as shown by the cognitive radio channel paradigm, avoiding interference is in general suboptimal, since interference can be exploited to help other communicating nodes with their transmissions.

Our model for a cognitive network assumed that the noncognitive users were oblivious to the cognitive users, and in particular did not cooperate with them to improve their communications. We can relax this assumption and allow various forms of cooperation between the cognitive and noncognitive users. We now delineate these various cooperation possibilities.

1) **Oblivious noncognitive users:** This form of cooperation is the overlay network described above; noncognitive users are oblivious to the cognitive users' presence. The cognitive transmitter decodes messages from its received signal without the assistance of the noncognitive transmitter. The noncognitive receiver's decoding is not affected by the presence of the cognitive user. This form of cooperation might motivate licensed users to share their dedicates spectrum with cognitive unlicensed users in exchange for improving the licensed users' performance via their cognitive capabilities.
2) **Aware noncognitive users:** In this setting noncognitive users are aware of the cognitive users' presence. They can use this awareness to improve their own communication. For example, they might adapt their decoding rules to exploit the signals received from the cognitive users instead of treating them as interference. In particular, if the interfering component of a cognitive message is strong, the noncognitive user can decode this message and cancel the interference. In the cognitive radio channel, this interference cancellation is required in the regime of strong interference [39], [40]. Another possibility is for a multiantenna noncognitive receiver to reduce or increase cognitive user interference via beamsteering, to place it in a weak or strong interference regime, where capacity-achieving strategies may be known.
3) **Cooperative noncognitive users:** In this setting noncognitive users may assist in delivering data to cognitive users. For example, a primary encoder may know channel conditions at a cognitive user in its vicinity. Based on this, it can transmit at a rate which ensures that the cognitive user can decode its transmitted message. Then, classic decode-and-forward, [10], where the source and the relay cooperatively deliver message to the destination, can be employed. For the cognitive radio channel, this was discussed in [37, Sec. IV].
4) **Full cooperation and cognition:** In this setting all users are cognitive and cooperative. This is the most general case in which any node can sense the environment, cooperate based on the obtained information, and forward messages. This allows for the maximum benefits possible via cooperative communications.

In a network with many nodes, it is more likely that the cognitive users will obtain information about the nodes in their vicinity. Moreover, it is easier for nodes that are in the same vicinity to cooperate, since cooperating transmitters that are close to each other can exchange messages without transmitting significant power and without creating much interference to the rest of the network. Once these neighboring nodes know each other messages, either through cognition or cooperation, they can jointly transmit these messages using encoding strategies for multiple-antenna transmitters, with each antenna associated with a different neighboring node. Similarly, a group of neighboring receivers can exchange their estimates of observed signals before decoding, thus mimicking a multiple-antenna receiver. In effect, the nodes create a *virtual MIMO* (multi-input, multi-output) transmitter/receiver. Indeed, this strategy of virtual MIMO for node *clustering* leads to capacity for several scenarios. Specifically, is has been shown that if the nodes form a cluster with the source and a destination, DF and CF respectively achieve capacity [57]; cooperation based on CF achieves high gains [58]; and sources can use dirty paper coding for improved performance [59]. In large networks, hierarchical clustering can achieve a linear scaling of the total capacity [60], which is significantly better than scaling under a classic multi-hop strategy [61]. These approaches can also be adapted to the scenario where cognition enable users to obtain necessary information for relaying. A more detailed discussion of connections between cognition and cooperation can be found in [62]. Clustering in the cognitive multi-access network has also been evaluated in [63].

While scaling laws offer significant insights into the capacity limits of large networks and the associated transmission strategies, these results are only applicable to asymptotically large networks. Transmission strategies for small networks based on information-theoretic analysis are generally quite involved and impractical for real networks. However, it is clear that networks with cognitive users can employ cooperative strategies and benefit from cooperation. Conversely, some level of cognition in the network is necessary in order to realize cooperation. Furthermore, cognition at the nodes can *simplify* cooperative strategies because less assistance from source nodes will be required to deliver necessary information to the relays, and relays can obtain some of this information through cognition. Overall, capacity results to date indicate significant performance gains from cooperation as well as cognition, and the challenge is to find practical techniques to exploit these ideas in real systems.

VI. INTERWEAVE COGNITIVE RADIO

The overlay approaches described previously requires a proiri knowledge of the noncognitive message at the cognitive transmitter, which is generally very difficult to obtain in practice unless the transmitters are in close proximity. Specifically, when the cognitive and noncognitive transmitters are close to each other, the SNR of the noncognitive signal at the cognitive transmitter is higher than that at the noncognitive receiver. The cognitive transmitter might therefore be able to decode the noncognitive message in a fraction of the time it takes the primary receiver to do so. However, whether or not this is possible will depend on the transmission strategy of the noncognitive encoder; a higher SNR does not in general allow faster decoding of the message. After decoding the primary message, the cognitive transmitter can then exploit the message and codebook knowledge via techniques described in the previous section [64]. Clearly the overlay approach requires many underlying assumptions about the network which are not true in general scenarios.

In general scenarios, concurrent cognitive and noncognitive user operation is invariably associated with interference at the noncognitive receiver, which is not desired. The solution then is to try to completely *avoid* this interference by allowing the cognitive user to transmit *only over spectrum gaps* that arise in time and in frequency. This is the central idea behind the *interweave approach*, originally outlined in [2].

A. Noncognitive User Detection

We observe that the underlay and overlay cognitive radio models involve simultaneous noncognitive and cognitive transmissions. Protection to the noncognitive users can be guaranteed by limiting some measure of interference caused to the noncognitive users from cognitive communication. Sensing for noncognitive users is not necessary and is therefore not incorporated into the underlying models. However, in order for an interweave cognitive radio to be able to efficiently communicate through the spectral holes without causing any interference to the active noncognitive users, it requires occupancy knowledge of the noncognitive users' in the different frequency bands. Accurate sensing of the presence of noncognitive systems over a wide bandwidth is therefore crucial to interweave cognitive radio operation.

The task of noncognitive user detection is rendered especially difficult due to signal degradations caused by fading and shadowing effects. Further, device level non-linearities and interference from other unlicensed radios result in an uncertainty in the noise power seen at the cognitive receiver which imposes additional limitations on sensing. In addition, due to the dynamic nature of the noncognitive user activity, spectrum sensing needs to be periodically performed to update the occupancy information. The plurality of problems encountered in sensing has attracted a lot of research activity [65]–[69] (and references therein).

Noncognitive user sensing in practical environments is usually governed by limits on the probabilities of missed detection and false alarms, specified together in a receiver operating characteristic (ROC) constraint. In the absence of noise uncertainty (uncertainty in the variance of the noise), detection at low noncognitive SNRs directly translates to longer observation times at the detector. Analysis of detection times in scenarios with noise uncertainty reveals that when the SNR is below a certain threshold, detection may not be possible, even with infinite sensing times [65]. This threshold is called the 'SNR wall' [65] and noncognitive user detection at SNRs below this wall is impossible. While the existence of the SNR wall was initially observed in moment detectors, such SNR limits are also seen in cyclostationary feature detectors [66], [70].

One solution to combat fading and noise uncertainty is, fundamentally, to take a collaborative approach to sensing [65], [67]–[69], [71]. Multiple cognitive users have to independently monitor noncognitive user activity and then exchange spectrum availability estimates to infer the presence of the noncognitive users. However, there has been limited research into designing suitable protocols for information exchange between the sensing nodes in the dynamic cognitive radio environment.

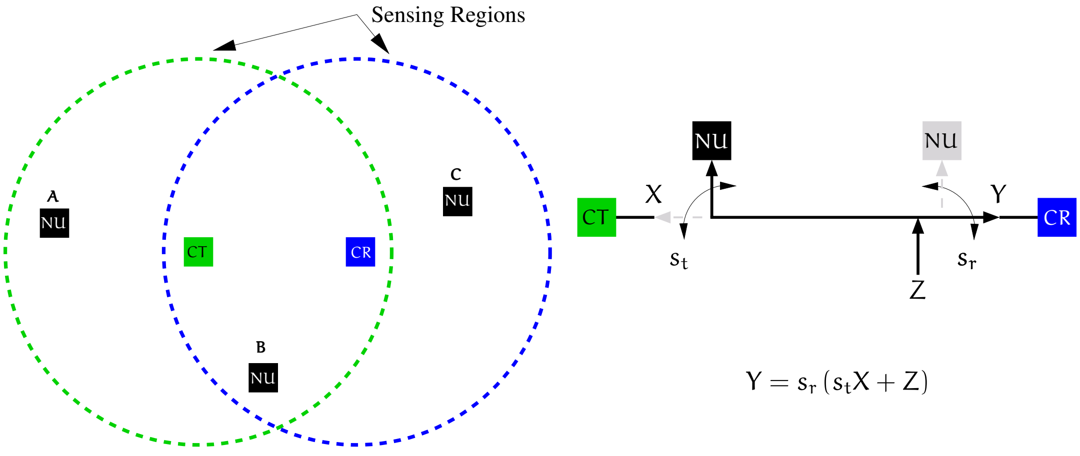
B. Cognitive Radio Link Modeling and Capacity

We now overview some of the recent theoretical results pertaining to the fundamental capacity limits of interweave cognitive radio systems.

1) Single cognitive user: In this section, we discuss some capacity limits of an isolated single cognitive link in the presence of one or more noncognitive users. When the cognitive transmitter (CT) and receiver (CR) are physically separated by a sufficiently large distance, the local noncognitive user activity at the transmitter and receiver will be different [72], [73]. In such scenarios, noncognitive user sensing will have to be performed *both* at CT and CR. In the absence of any information exchange between them, there is an uncertainty at the CT if there is a spectrum hole available near CR and vice versa. The capacity of the cognitive link can be calculated by modeling the uncertainty in terms of partial spectral activity knowledge at the transmitter and receiver. We discuss this approach and some of the associated results in detail in the following.

Consider a cognitive transmitter and a cognitive receiver operating in the presence of noncognitive users (NU) denoted by A, B and C, as shown in Figure 11(a). For the sake of simplicity, assume that all the users are operating over the same frequency band. The dotted regions around the cognitive transmitter and receiver represent their respective *sensing regions* - noncognitive transmissions can only be detected within these regions. Cognitive transmitter CT can therefore only sense whether or not noncognitive users A or B are active, i.e., CT detects spectral holes when both A and B are inactive. Similarly, the cognitive receiver CR can only sense whether or not noncognitive users B or C are active, i.e., CR detects spectral holes when both B and C are inactive. As a consequence, the spectral holes (communication opportunities) detected at the cognitive transmitter and receiver are not identical.

We can describe fundamental characteristics of the underlying spectral environment as being *dynamic* and *distributed* [72], [73]. The term 'distributed' is used to emphasize that the noncognitive user activity detected in the vicinity of the



(a) Conceptual Model

(b) Two Switch Model

Fig. 11: The different perspectives on local spectral activity at the cognitive radio transmitter CT and receiver CR are depicted in 11(a). Nodes marked A, B and C represent the noncognitive users (NU) of the spectrum. The dotted circles represent the corresponding sensing regions. Figure 11(b) represents the corresponding two switch model where the noncognitive user occupancy processes are captured in the switch states s_t and s_r. In the scenario considered, we have $s_t = 1$ and $s_r = 0$.

cognitive transmitter differs from that detected around the cognitive receiver. The cognitive transmitter CT does not automatically have full knowledge of the noncognitive user activity in the vicinity of the receiver CR. Similarly there is an uncertainty at the cognitive receiver about the noncognitive user activity at the transmitter. The larger the separation between the cognitive transmitter and receiver, the less the overlap in their respective sensing regions. This leads to a more *distributed* spectral environment, and consequently a higher sensing uncertainty at the transmitter and receiver. The noncognitive users' activity is dynamic - over time, different noncognitive users can become active/inactive in different segments of the spectrum. Therefore the noncognitive user activity sensed at the cognitive transmitter and receiver will change with time. This increases the uncertainty at either end of the link about the communication opportunities sensed at the other end. As the noncognitive users become more dynamic, the spectral activity changes faster and is consequently less predictable.

The simple conceptual model of Figure 11(a) can be reduced to the *two switch* mathematical model shown in Figure 11(b). The spectrum holes sensed at the cognitive 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 the cognitive transmitter perceives that a noncognitive user is active in its sensing region. The transmitter switch state $s_t = 1$ when no noncognitive user is detected at the cognitive transmitter. Similarly, the receiver switch state s_r is equal to 0 or 1 depending on whether or not an active noncognitive user is detected in the sensing region of the cognitive receiver.

The switch state s_t is known only to the transmitter while the switch state s_r is known only to the receiver. 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 noncognitive activity and therefore the lower the correlation. The dynamic nature of the noncognitive user activity is reflected in the rate at which the switches change state.

The relationship between the input signal X at the cognitive transmitter and the signal output Y at the cognitive receiver is described in Figure 11(b) (an AWGN channel is considered for the sake of simplicity). Notice that the knowledge of both the switch states s_t and s_r completely characterizes the communication channel. However, s_t is known only to the cognitive transmitter and s_r only to the cognitive receiver, i.e., the cognitive transmitter and receiver only have *partial* channel knowledge. Opportunistic cognitive radio therefore corresponds to communication with *partial side information*.

The switch state s_t (or s_r) represents whether or not the cognitive transmitter (or cognitive receiver) perceives a possible communication opportunity based on its sensing of local spectral activity. In practice, local sensing at the transmitter and receiver is usually not perfect and depend on the ROC of the spectrum sensor. It should be noted that the two switch model is also applicable to imperfect sensing scenarios. In such cases, the probabilities of missed detection and false alarm can be directly mapped to the probability of the switches s_t and s_r being in state 0 or 1.

Using information theoretic formulations for memoryless channels with causal/non-causal partial channel knowledge at the transmitter and receiver(see [74]), one can characterize the capacity of the two switch cognitive radio link with causal and non-causal knowledge of the communication opportunities at the cognitive transmitter and receiver [72]. For the case of perfect sensing, inner bounds based on Gaussian inputs at the cognitive transmitter can be easily constructed. Outer bounds based on *genie* information, wherein additional side information is provided to the cognitive transmitter or receiver, can also be obtained. As an example, suppose a genie provides the cognitive transmitter state to the cognitive receiver. Then, the capacity with this additional knowledge at the receiver is higher than that of the original link and consequently serves as an upper bound. The inner and outer bounds are also invaluable to characterize the benefits of feeding forward the transmitter state to the receiver; and feeding back the receiver state to the transmitter. In an overpopulated system (i.e., high noncognitive user activity), transmitter side information is found to be more valuable (i.e., feeding back the spectrum hole availability at the cognitive receiver to the cognitive transmitter results in a higher capacity than feeding forward the cognitive transmitter's spectrum hole availability to the cognitive receiver) [74]. However, the opposite is found to be true in underpopulated systems. An interesting observation from this study is that when the noncognitive user activity is moderately dynamic (switch states remain the same for three or more channel uses), the advantage of receiver side information is less pronounced.

One can also look at the throughputs achievable in cognitive radio systems from a queue stability perspective [75], [76]. Consider a certain cognitive user operating in the presence of a single noncognitive radio. Assume that the noncognitive user's packet arrival rate is held fixed to guarantee a certain average throughput. With given sensing limits at the cognitive user, the maximum arrival rate of the cognitive packets which guarantees noncognitive and cognitive queue stability can be characterized [75]. Similarly, with multiple cognitive users sharing the channel resource with a single noncognitive radio, the queue stability region of the cognitive users can be investigated under a maximum packet delay constraint at the noncognitive user.

There has also been work exploring soft sensing based power control at the cognitive transmitter to maximize the cognitive capacity [77]. Some cognitive radio models implement the sensing as a binary hypothesis test, i.e., the spectrum sensor outputs a single bit indicating the presence/absence of the noncognitive user. Any power adaptation at the cognitive transmitter depends on the sensing only through this single bit. There is a loss of information in translating the analog noncognitive signals to a bit decision, and the soft information from sensing is not made use of. However, considerable throughput benefits can be obtained from power control exploiting this soft information [77].

2) Opportunistic Channel Selection: The two switch cognitive radio model hides the details of opportunistic channel (frequency) selection. The choice of the frequency band(s) that can be used for communication is dictated by the type of cognitive transmitter and receiver - narrowband or wideband. In other words, channel selection depends on the amount of noncognitive occupancy information available at the cognitive transmitter and receiver.

- **Narrowband Techniques**: With a narrowband transmitter and receiver, the frequency band to be used for transmission can be predetermined, or dynamically chosen based on the noncognitive user occupancy. This gives rise to two different frequency selection techniques:
  - *Frequency Hopping*: In the frequency hopping scheme, the cognitive transmitter and receiver simultaneously hop across multiple frequencies according to a predetermined hopping sequence. Thus, the cognitive transmitter and receiver are *always* matched to the same frequency band. Frequency hopping is a very simple scheme - it does not exploit the knowledge of the past and present channel availabilities. This method is not really opportunistic or cognitive, and causes interference when the hopping signal lands in a band occupied by other users, which can significantly degrade their performance. Frequency hopping is the basic premise behind several commercial and military systems. The interference caused by frequency hopping is well documented in the coexistence challenges between Bluetooth and 802.11 b/g/n systems in the 2.4 GHz unlicensed band.
  - *Frequency Tracking*: In the frequency tracking scheme, the cognitive transmitter, based on a given strategy, chooses *one* (if any) of the free frequency bands for transmission. The cognitive receiver, based on past received signals, chooses the best channel to listen to so that the probability that the transmitter and receiver are matched to the same channel is maximized. Notice that frequency tracking is a significant departure from the conventional communication model. Traditionally, the receiver has knowledge of the frequency band used for transmission and can wait 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. After each transmission, the receiver must determine whether each observed symbol corresponds to a matched scenario (i.e. the cognitive transmitted symbol is received) or a mismatch (i.e. received signal is independent of the cognitive transmitted symbol).
- **Wideband Technique** - *Frequency Coding*: With a wideband cognitive system, the transmitter and receiver can scan the spectral activity in all the frequency bands and communicate opportunistically through a codeword spanning multiple frequency slots that are presumed to be idle. Unlike the narrowband schemes, such a *frequency coding* scheme requires the channel availabilities in *all* the different frequency bands *before* every transmission. Such wideband techniques pose difficult design challenges in developing a wideband receiver front end with good sensitivity.

In highly dynamic noncognitive activity environments, frequency hopping outperforms frequency tracking [72], [73]. On the other hand, when the noncognitive user switches ON/OFF much slowly, frequency tracking utilizes the memory in the noncognitive occupancy process and provides higher throughputs than frequency hopping. This is intuitive since as the noncognitive users become more dynamic, the cost of tracking the channel is higher than the benefits of being matched. Further, the throughput benefits of frequency coding over frequency hopping are found to be very small [72].

Degrees of Freedom: The interweave approach mandates cognitive transmissions without noncognitive interference. With perfect sensing, since the cognitive users are able to avoid transmissions with the noncognitive users, the number of degrees of freedom that can be achieved is equal to one. The degrees of freedom therefore depends on the fraction of the time the channel is available for cognitive transmissions (noncognitive user duty cycle). With imperfect sensing, the number of degrees of freedom will depend not only on the duty cycle of the noncognitive users but also on the probabilities of false alarm and missed detection.

C. Multiple cognitive users

A major issue in a multiple cognitive user environment is dynamic spectrum access and sharing of the available spectrum holes, a topic that has generated a lot of research interest in the recent past [78]–[83]. This problem is similar to that of multiple access in multichannel wireless networks; in both these cases multiple independent transmitters need to access a set of shared channel resources. Many access protocols for cognitive networks have therefore been derived from conventional MAC protocols like ALOHA and CSMA. Cognitive radio operation in practical multiuser environments is governed by interference tolerance and sensing limits at the noncognitive and cognitive users. The interference limits at the noncognitive and cognitive users indicate the amount of protection needed at each (noncognitive or cognitive) user from the multiuser interference to maintain a certain rate. In other words, the interference limit is a measure of how *tolerant* the users are to multiuser interference. On the other hand, the sensing limits (minimum SNR needed for detection) at the cognitive users reflect the amount of protection that each cognitive user is individually able to provide to the noncognitive users. Put differently, the sensing limit is a measure of how *aggressively* the cognitive users transmit their signals. In these scenarios, the key is to strike a balance between the two conflicting goals - minimizing the interference to the noncognitive users, and maximizing the performance of the entire system. One of the ways this issue can be handled is by limiting the number of cognitive users. Therefore, the natural question that arises is: *What is the optimal number of cognitive users (opportunistic access) relative to the number of noncognitive users (licensed access) that maximizes the sum throughput in the system?* This is reminiscent of the familiar debate of licensing versus autonomy, a tradeoff that is fundamental to many areas of systems and control theory, and has provoked spirited debate in regulatory bodies about whether unused spectrum is better allocated to licensed or unlicensed users. The generality of this tradeoff is evident through an analogy with traffic control: Too much regulation, i.e., too few cognitive users (traffic lights at every intersection) and the system is inefficient due to unoccupied spectral holes. On the other hand, too much autonomy/opportunistic behavior, i.e., too many cognitive users, (no traffic lights) and the system becomes self-disruptive due to collisions between the cognitive users. It is found that when the noncognitive and cognitive users have identical packet arrival rates, with perfect sensing at the cognitive users and zero interference tolerance at each of the users, the optimal fraction of noncognitive users is equal to the duty cycle [84].

There has also been some recent research interest focused on scaling laws in cognitive radio networks, i.e., how the sum throughput in a random network with multiple noncognitive and cognitive transmitter-receiver pairs scales with the number of cognitive users. It is well known from seminal work in homogeneous non-cooperative adhoc networks [85] that the total network throughput scales as $\sqrt{n}$, where n is the number of transmitter-receiver pairs, i.e, the per user capacity scales as $\frac{1}{\sqrt{n}}$. Consider a similar setup in heterogeneous wireless networks with multiple independent noncognitive and cognitive users. To protect the noncognitive users from cognitive user interference, non-overlapping noncognitive exclusive regions (PER) are defined, one around every noncognitive transmitter within which cognitive transmitters are not allowed. Further, the cognitive transmitter and receiver are constrained to be located within a certain distance of each other. This model differs from the adhoc network model of [85] in that only single hop communication between the cognitive transmitter receivers is considered. The sum throughput is found to scale *linearly* in the number of cognitive links [86]. This result is shown to hold true regardless of whether the cognitive transmitters use constant power or scale the power depending on the distance from the noncognitive transmitters.

VII. SUMMARY AND FUTURE DIRECTIONS

Cognitive radios hold tremendous promise to unlock spectral gridlock through advanced radio designs, powerful encoding and signal processing techniques, and novel coexistence protocols. There are multiple paradigms associated with cognitive radios, the most common of which are underlay, overlay, and interweave networks. These paradigms are unified in their use of cognitive radios to sense their environment and exploit the network side information obtained from this sensing to improve spectral efficiency and performance for all users. The paradigms differ in the nature of the side information the cognitive radios can obtain, as well as the coexistence protocols imposed on the cognitive users. In particular, underlay networks impose strict constraints on the interference a cognitive user may cause to other users. Interweave networks require cognitive users to communicate using spectral holes in space, time, and frequency that are not occupied by other users, so ideally they cause no interference. In contrast, overlay networks seek to exploit interference through sophisticated coding strategies at the cognitive transmitters that facilitate communications for other users.

While we have described known capacity results and bounds for cognitive networks along with their associated encoding and decoding strategies, we have not included capacity formulas for these results. These formulas are generally quite cumbersome, consisting of an implicit characterization via multiple rate equations that typically yield little insight. Instead of focusing on exact capacity results, we have instead opted to illuminate the degrees of freedom in cognitive networks as a metric for their sum capacity. We have seen that in a system with two transmitter-receiver pairs and different assumptions about cognition, the network degrees of freedom ranges from one to two. Specifically, if only one transmitter, one receiver, or one transmitter-receiver pair is cognitive, there is only one degree of freedom. If both transmitters or both receivers are cognitive, or one transmitter is cognitive and the receiver of the other transmitter is cognitive, then there are two degrees of freedom in the network. A surprising and promising result for large networks is the degrees of freedom possible via interference alignment. Specifically, in an interference channel with K users, assuming global knowledge of the time-varying channel coefficients, the network degrees of freedom is $K/2$, i.e. it grows with the number of users. This growth does not require any side information about user messages. However, obtaining global channel knowledge at all nodes may require significant overhead, especially in a highly dynamic propagation environment.

The side information assumptions in cognitive radio network models are not always practical for typical operating scenarios. Underlay networks assume that a cognitive transmitter can determine the interference it causes to a receiver in another location. Given the random and sometimes erratic characteristics of many wireless channels, it seems unlikely this side information could be obtained except by exploiting reciprocity in an overheard transmission. The noncognitive receiver might also send the signal strength it receives from the cognitive transmitter via a feedback path. Without knowledge of the interference strength, the cognitive radio must generally transmit at a very low power or power per Hertz, which can significantly hamper its capabilities. For overlay networks the ability of a cognitive transmitter to learn the message of a noncognitive transmitter is generally only practical in a retransmission scenario or when there is cooperation between the transmitters. Interweave communications requires fast detection and rapid frequency changes or wideband receiver front ends to determine spectral holes in space, time, and frequency. Such holes must be common to the transmitter and receiver, and a protocol established for them to coordinate on which hole to use. In all cognitive radio systems, sensing will consume battery power, and the capacity benefits of such sensing must be weighed against the desire for low-power devices. Moreover, while technology advances will address some of the practicality issues, some of these side information assumptions will never be realizable in certain networks. For such networks, new capacity results are needed based on more realistic assumptions about what network side information can be obtained by cognitive users.

Although cognitive radio ideas have permeated much research throughout this decade, there are still many open questions as well as new directions to explore. Few works have investigated protocols for cooperation between and among cognitive and noncognitive users, especially in networks with more than a few users. New paradigms for cognitive radio networks may also be developed that combine notions of underlay, overlay, and interweave, or come up with completely new ways to exploit network side information for capacity gains. Radio advances such as multiple antennas or multiuser detection may change the characteristics of network interference, paving the way for new cognitive strategies to exploit or avoid it. Finally, software radios that can adapt their waveforms and protocols on the fly may provide new breakthroughs for cognitive network design.

Research breakthroughs will not be sufficient to enable widespread use of cognitive networks. Regulatory bodies must fundamentally change their philosophy of spectrum allocation. Licensed users must give up the assumed right to sole use of their spectrum. A cognitive radio may be larger, cost more money, and consume more power than a noncognitive one. This confluence of technical, political, and commercial challenges ensures that cognitive radios will be a hot topic in research, politics, and commercial development for many years to come.

References

[1] National Telecommunications and Information Administration (NTIA), “FCC Frequency Allocation Chart,” 2003. Download available at www.ntia.doc.gov/osmhome/allochrt.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/SEWGFinalReport_1.pdf.

[4] Simon Haykin, “Cognitive Radio: Brain-Empowered Wireless Communications,” *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 201 – 220, February 2005.

[5] Paul J. Kolodzy, “Cognitive Radio Fundamentals,” *SDR Forum, Singapore*, April 2005.

[6] Shared Spectrum Company, “Comprehensive Spectrum occupancy measurements over six different locations,” August 2005. Download available at http://www.sharedspectrum.com/?section=nsf_summary.

[7] Zhiqiang Wu and Bala Natarajan, “Interference Tolerant Agile Cognitive Radio: Maximize Channel Capacity of Cognitive Radio,” *Consumer Communications and Networking Conference*, pp. 1027 – 1031, January 2007.

[8] H. Sato, “Two user communication channels,” *IEEE Transactions on Information Theory*, vol. 23, pp. 295–304, May 1977.

[9] A. B. Carleial, “Interference channels,” *IEEE Transactions on Information Theory*, vol. 24, pp. 60–70, Jan. 1978.

[10] M. H. M. Costa and Abbas El Gamal, “The capacity region of the discrete memoryless interference channel with strong interference,” *IEEE Transactions on Information Theory*, vol. 33, pp. 710–711, Sept. 1987.

[11] T. Han and K. Kobayashi, “A new achievable rate region for the interference channel,” *IEEE Transactions on Information Theory*, vol. 27, pp. 49–60, Jan. 1981.

[12] G. Kramer, “Outer bounds on the capacity of Gaussian interference channels,” *IEEE Transactions on Information Theory*, vol. 50, pp. 581–586, Mar. 2004.

[13] G. Kramer, “Review of rate regions for interference channels,” in *Proc. 2006 Int. Zurich Seminar on Communications, Zurich, Switzerland*, pp. 162–165, Feb. 2006.

[14] R. Etkin, D. Tse, and H. Wang, “Gaussian interference channel capacity to within one bit,,” *Submitted to the IEEE Transactions on Information Theory*, 2007.

[15] Xiaohu Shang, Gerhard Kramer and Biao Chen, “A New Outer Bound and the Noisy-Interference Sum-Rate Capacity for Gaussian Interference Channels,” *Submitted to the IEEE Transactions on Information Theory*, 2007.

[16] A Host-Madsen and Aria Nosratinia, “The Multiplexing Gain of Wireless Networks,” in *IEEE International Symposium on Information Theory*, pp. 2065 – 2069, September 2005.

[17] Xiaohu Shang, Biao Chen and Michael J. Gans, “On Achievable Sum Rate for Vector Gaussian Interference Channels,” *IEEE International Symposium on Information Theory*, pp. 2215 – 2219, 2006.

[18] Viveck R. Cadambe, Syed A. Jafar, “Interference Alignment and the Degrees of Freedom for the K User Interference Channel,” *Preprint available at http://newport.eecs.uci.edu/ syed/papers/*, 2007.

[19] Michael C. Gastpar, "On Capacity under Receive and Spatial Spectrum-Sharing Constraints," *IEEE Transactions on Information Theory*, vol. 53, pp. 471 – 487, February 2007.

[20] Amir Ghasemi and Elvino S. Sousa, "Capacity of Fading Channels Under Spectrum-Sharing Constraints," *IEEE International Conference on Communications*, June 2006.

[21] Wei Wang, Tao Peng and Wenbo Wang, "Optimal Power Control Under Interference Temperature Constraints in Cognitive Radio Network," *IEEE Wireless Communications and Networking Conference*, pp. 116–120, March 2007.

[22] Yiping Xing and R. Chandramouli, "QoS Constrained Secondary Spectrum Sharing," *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 658 – 661, November 2005.

[23] Yiping Xing, Chetan Nanjunda Mathur, M. A. Haleem, R. Chandramouli and K. P. Subbalakshmi, "Dynamic Spectrum Access with QoS and Interference Temperature Constraints," *IEEE Transactions on Mobile Computing*, vol. 6, pp. 423 – 433, April 2007.

[24] James O. Neel, Jeffrey H. Reed and Robert P. Gilles, "Convergence of Cognitive Radio Networks," *IEEE Wireless Communications and Networking Conference*, vol. 4, pp. 2250 – 2255, 2004.

[25] James O. Neel, R. Michael Buehrer, Jeffrey H. Reed, Robert P. Gilles, "Game Theoretic Analysis of a Network of Cognitive Radios," *Midwest Symposium on Circuits and Systems*, vol. 3, pp. 409 – 412, August 2002.

[26] J. Neel, Rekha Menon, Jeffrey H. Reed, and A. MacKenzie, "Using Game Theory to Analyze Physical Layer Cognitive Radio Algorithms," 2005.

[27] N. Nie, C. Comaniciu and P. Agrawal, "A Game Theoretic Approach to Interference Management in Cognitive Networks," *Wireless Communications (The IMA Volumes of Mathematics and its Applications)*, November 2006.

[28] Dusit Niyato and Ekram Hossain, "A Game-Theoretic Approach to Competitive Spectrum Sharing in Cognitive Radio Networks," *IEEE Wireless Communications and Networking Conference*, pp. 16–20, March 2007.

[29] Xia Wang and Qi Zhu, "Power Control for Cognitive Radio Base on Game Theory," *International Conference on Wireless Communications, Networking and Mobile Computing*, pp. 1256 – 1259, September 2007.

[30] I. Emre Telatar, "Capacity of Multi-antenna Gaussian Channels," *European Transactions on Telecommunications*, vol. 10, pp. 585 – 596, November 1999.

[31] Rui Zhang and Ying-Chang Liang, "Exploiting Multi-Antennas for Opportunistic Spectrum Sharing in Cognitive Radio Networks," *Preprint available at http://arxiv.org/abs/0711.4414v1*, November 2007.

[32] Paul J. Kolodzy, "Interference Temperature: A Metric for Dynamic Spectrum Utilization," *International Journal of Network Management*, pp. 103 – 113, March 2006.

[33] Qing Zhao and Brian M. Sadler, "A Survey of Dynamic Spectrum Access," *IEEE Signal Processing Magazine*, pp. 79 – 89, May 2007.

[34] Mohammad Ali Maddah-Ali, Abolfazl S. Motahari and Amir K. Khandani, "Signaling over MIMO Multi-Base Systems: Combination of Multi-Access and Broadcast Schemes," *IEEE International Symposium on Information Theory*, pp. 2104 – 2108, 2006.

[35] Syed Ali Jafar and Shlomo Shamai, "Degrees of Freedom Region for the MIMO X Channel," *To appear in the IEEE Transactions on Information Theory*, 2007.

[36] Ivana Marić, Andrea J. Goldsmith, Gerhard Kramer and Shlomo Shamai, "On the Capacity of Interference Channels with a Partially-Cognitive Transmitter," *IEEE International Symposium on Information Theory*, June 2007.

[37] 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.

[38] Wei Wu, Sriram Vishwanath, and A. Arapostathis, "On the Capacity of the Interference Channel with Degraded Message Sets," *Submitted to the IEEE Transactions on Information Theory*, 2006.

[39] 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.

[40] Ivana Marić, Roy D. Yates and Gerhard Kramer, "Capacity of Interference Channels with Partial Transmitter Cooperation," *IEEE Transactions on Information Theory*, vol. 53, pp. 3536 – 3548, October 2007.

[41] T. Cover, "Broadcast channels," *IEEE Transactions on Information Theory*, vol. 18, pp. 2–14, 1972.

[42] J. Jiang and Y. Xin, "On the achievable rate regions for interference channels with degraded message sets," *Submitted to the IEEE Transactions on Information Theory*, Apr. 2007.

[43] Ivana Marić, Andrea Goldsmith, Gerhard Kramer and S. Shamai(Shitz), "On the capacity of interference channels with a cognitive transmitter," *European Transactions on Telecommunications*, 2007.

[44] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Problemy Peredachi Informatsii*, vol. 9, no. 1, pp. 19–31, 1980.

[45] M. H. M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, pp. 439–441, May 1983.

[46] S. Sridharan and S. Vishwanath, "On the capacity of a class of MIMO cognitive radios," in *IEEE Information Theory Workshop (ITW 2007), Lake Tahoe, CA*, Sept. 2007.

[47] H. Weingarten, Y. Steinberg, and S. Shamai(Shitz), "The capacity region of the Gaussian multiple input multiple output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, pp. 3936–3964, Sept. 2006.

[48] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates and sum rate capacity of Gaussian MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 49, pp. 2658–2668, Oct. 2003.

[49] J. Jiang, Y. Xin, and H. Garg, "Interference channels with common information," *Submitted to the IEEE Transactions on Information Theory*, 2007.

[50] Y. Liang, A. Somekh-Baruch, V. Poor, S. Shamai(Shitz), and S. Verdú, "Cognitive interference channels with confidential messages," in *45th Annual Allerton Conference on Communication, Control and Computing, Allerton House, Monticello, IL*, Sept. 2007.

[51] D. Tuninetti, "The interference channels with generalized feedback," in *IEEE Proc. Int. Symp. Inf. Th., Nice, France*, June 2007.

[52] Y. Cao, B. Chen, and J. Zhang, "A new achievable rate region for interference channels with common information," in *Proc. IEEE Wireless Comm. and Networking Conf., Hong Kong*, Mar. 2007.

[53] Pulkit Grover and Anant Sahai, "What is needed to exploit knowledge of primary transmissions?," *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, 2007.

[54] Natasha Devroye and Masood Sharif, "The Multiplexing Gain of MIMO X channels with Partial Transmit Side Information," *IEEE International Symposium on Information Theory*, June 2007.

[55] Thomas M. Cover and Abbas El Gamal, "Capacity Theorems for the Relay Channel," *IEEE Transactions on Information Theory*, vol. 25, pp. 572–584, September 1979.

[56] E. C. V. D. Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120–154, June 1971.

[57] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Transactions on Information Theory*, vol. 51, pp. 3037–3063, Sept. 2005.

[58] M. Katz and S. Shamai(Shitz), "Relaying protocols for two co-located users," in *Proc. IEEE Int. Symp. Inf. Theory*, pp. 936–940, Sept. 2005.

[59] N. Jindal, U. Mitra, and A. Goldsmith, "Capacity of ad-hoc networks with node cooperation," in *Proc. IEEE Int. Symp. Inf. Theory*, p. 271, June 2004.

[60] A. Ozgur, O. Leveque, and D. Tse, "Hierarchial cooperation achieves optimal capacity scaling in ad hoc networks," *Submitted to the IEEE Transactions on Information Theory*, 2006.

[61] Piyush Gupta and P.R. Kumar, "The Capacity of Wireless Networks," *IEEE Transactions on Information Theory*, vol. 42, pp. 388 – 404, March 2000.

[62] Natasha Devroye and Patrick Mitran and Vahid Tarokh, "Cognitive Decomposition of Wireless Networks," *Cognitive Radio Oriented Wireless Networks and Communications*, June 2006.

[63] Natasha Devroye, Patrick Mitran and Vahid Tarokh, "Cognitive Multiple Access Networks," *IEEE International Symposium on Information Theory*, pp. 57–61, September 2005.

[64] Onur Ozan Koyluoglu and Hesham El Gamal, "On Power Control and Frequency Re-Use in the Two-User Cognitive Channel," *Submitted to the IEEE Transactions on Wireless Communications*, October 2007.

[65] Anant Sahai, Nigel Hoven, Shridhar 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.

[66] Rahul Tandra and Anant Sahai, "SNR Walls for signal detection," *IEEE Journal on Selected Topics in Signal Processing*, February 2008.

[67] Shridhar Mubaraq Mishra, Anant Sahai and Robert W. Brodersen, "Cooperative Sensing among Cognitive Radios," *IEEE International Conference on Communications*, vol. 4, pp. 1658–1663, June 2006.

[68] Amir Ghasemi and Elvino S. Sousa, "Collaborative Spectrum Sensing for Opportunistic Access in Fading Environments," *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 131 – 136, November 2005.

[69] Ghurumuruhan Ganesan and Ye Li, "Agility Improvement through Cooperative Diversity in Cognitive Radio," *IEEE Global Communications Conference*, vol. 5, November - December 2005.

[70] Rahul Tandra and Anant Sahai, "SNR Walls for Feature Detectors," *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 559 – 570, April 2007.

[71] Danijela Cabric, Shridhar Mubaraq Mishra, Robert W. Brodersen, "Implementation Issues in Spectrum Sensing for Cognitive Radios," *Asilomar Conference on Signals, Systems and Computers*, November 2004.

[72] Syed Ali Jafar and Sudhir Srinivasa, "Capacity Limits of Cognitive Radio with Distributed and Dynamic Spectral Activity," *IEEE Journal on Selected Areas in Communications*, vol. 25, pp. 529–537, April 2007.

[73] Sudhir Srinivasa, Syed Ali Jafar and Nihar Jindal, "On the Capacity of the Cognitive Tracking Channel," *IEEE International Symposium on Information Theory*, January 2006.

[74] Syed Ali Jafar, "Capacity with Causal and Non-Causal Side Information - A Unified View," *IEEE Transactions on Information Theory*, vol. 52, pp. 5468–5475, November 2006.

[75] Osvaldo Simeone, Umberto Spagnolini and Yeheskel Bar-Ness, "Stable Throughput of Cognitive Radios with and without Relaying Capability," *To appear in the IEEE Transactions on Communications*, 2007.

[76] J. Gambini, Osvaldo Simeone, Umberto Spagnolini, Yeheskel Bar-Ness and Yungsoo Kim, "Stability Analysis of a Cognitive Multiple Access Channel with Primary QoS Constraints," *Asilomar Conference on Signals, Systems and Computers*, 2007.

[77] Sudhir Srinivasa and Syed Ali Jafar, "Soft Sensing and Optimal Power Control for Cognitive Radio," *IEEE Global Communications Conference*, November 2007.

[78] Qing Zhao, Lang Tong and Anantharam Swami, "Decentralized Cognitive MAC for Dynamic Spectrum Access," *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 224–232, November 2005.

[79] Qing Zhao, Lang Tong and Anantharam Swami, "Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework," *IEEE Journal on Selected Areas in Communications*, pp. 224–232, April 2007.

[80] Alex Chia-Chun Hsu, David S. L. Wei and C.-C. Jay Kuo, "A Cognitive MAC Protocol Using Statistical Channel Allocation for Wireless Ad-Hoc Networks," *IEEE Wireless Communications and Networking Conference*, pp. 105–110, March 2007.

[81] Jianwei Huang, Randall A. Berry and Michael L. Honig, "Spectrum Sharing with Distributed Interference Compensation," *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 88 – 93, November 2005.

[82] Chandrasekharan Raman, and Roy D. Yates and Narayan B. Mandayam, "Scheduling Variable Rate Links via a Spectrum Server," *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 110 – 118, November 2005.

[83] Srivatsan Sankaranarayanan, Panagiotis Papadimitratos, A.M. Bradley and S. Hershey, "A Bandwidth Sharing Approach to Improve Licensed Spectrum Utilization," *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 279 – 288, November 2005.

[84] Sudhir Srinivasa and Syed Ali Jafar, "Cognitive Radio Networks: How much Spectrum Sharing is Optimal?," *Submitted to the IEEE Transactions on Wireless Communications*, November 2007.

[85] Piyush Gupta and P. R. Kumar, "Towards an Information Theory of Large Networks: An Achievable Rate Region," *IEEE Transactions on Information Theory*, vol. 49, pp. 1877 – 1894, August 2003.

[86] Mai Vu, Natasha Devroye, Masoud Sharif and Vahid Tarokh, "Scaling laws of Cognitive Networks," *IEEE Journal on Selected Topics in Signal Processing*, June 2007.