Information Theoretic Analysis of Cognitive Radio Systems

Natasha Devroye[1], Patrick Mitran[1], Masoud Sharif[2], Saeed Ghassemzadeh[3], and Vahid Tarokh[1]

[1] Division of Engineering and Applied Sciences, Harvard University
ndevroye, mitran, vahid @deas.harvard.edu
[2] Department of Electrical and Computer Engineering, Boston University
sharif@bu.edu
[3] Communications Technology Research Department, AT&T Labs-Research
saeedg@research.att.com

1 Introduction

Cognitive radios have recently emerged as a prime candidates for exploiting the increasingly flexible licensing of wireless spectrum. Regulatory bodies have come to realize that most of the time, large portions of certain licensed frequency bands remain empty [17]. To remedy this, legislators are easing the way frequency bands are licensed and used. In particular, new regulations would allow for devices which are able to sense and adapt to their spectral environment, such as cognitive radios, to become *secondary users.* [4] Such users are wireless devices that opportunistically employ the spectrum already licensed to *primary users*. Primary users generally associate with the primary spectral license holder, and thus have a higher priority right to the spectrum.

The intuitive goal behind secondary spectrum licensing is to increase the spectral efficiency of the network, while, depending on the type of licensing, not affecting higher priority users. The exact regulations governing secondary spectrum licensing are still being formulated [18], but it is clear that networks consisting of heterogeneous devices, both in terms of physical capabilities and in the right to the spectrum, will emerge.

Among the many questions that remain to be answered about cognitive networks, is that of the *fundamental limits* of possible communication. Although this may be defined in various ways, information theory is an ideal tool and approach from which to explore the underlying, implementation-independent limits of such heterogeneous networks. In this chapter, we will

[4] In this chapter, we will use the terms secondary user and cognitive user interchangeably. Cognitive radio will be clearly defined in section 1.2, and can be thought of as "smart" radios which are able to adapt to their environment for now.

outline the current state of the art in information theoretic analysis of cognitive systems.

1.1 Secondary Spectrum Licensing

The emergence of the FCC's Secondary Markets Initiative (SMI, [18]) was brought on by both the obvious desire for spectral efficiency, as well as empirical measurements showing that most of the time certain licensed frequency bands remain unused. The goal of the SMI is to remove unnecessary regulatory barriers to new secondary market oriented policies such as:

- **Spectrum leasing,** which allows non-licensed users to lease any part, or all of the spectrum from the licensed user.
- **Dynamic spectrum leasing,** which is a temporary and opportunistic usage of spectrum rather than a longer-term sub-lease.
- **Private commons,** whereby a licensee could allow non-licensed users access to his/her spectrum without a contract, optionally with an access fee.
- **Interruptible spectrum leasing,** which would be suitable for a lessor that wants a high level of assurance that any spectrum temporarily in use, or leased, to an incumbent cognitive radio could be efficiently reclaimed if needed. A prime example would be the leasing of the generally unoccupied spectrum alloted to the US government or local enforcement agencies, which in times of emergency could be quickly reclaimed.

Of interest in this chapter is *dynamic spectrum leasing*, in which some wireless devices opportunistically employ the spectrum rather than opt for a longer term sub-lease. In order to exploit the spectrum, we require a device which is able to sense the communication opportunities, and then take actions based on the sensed information. In this chapter, such actions will include transmitting (or refraining from transmitting) and adapting their modulation and/or coding strategies so as to 'better" employ the sensed spectral environment. Cognitive radios are prime candidates for such actions.

1.2 Cognitive Radios and Behavior

Over the past few years, the incorporation of software into radio systems has become increasingly common. This has allowed for faster upgrades, and has given these wireless communication devices the ability to transmit and receive using a variety of protocols and modulation schemes (enabled by reconfigurable software rather than hardware). Furthermore, as their name suggests, such radios can even become "cognitive", and, as dictated by the software, adapt their behavior to their wireless surroundings without user intervention. According to the FCC software defined radios (SDR) encompasses any "radio that includes a transmitter in which operating parameters such as frequency range, modulation type or maximum output power can be altered by software without making any changes to hardware components that affect the radio frequency emissions." Mitola [37] took the definition of an SDR one step further, and envisioned a radio which could make decisions as to the network, modulation and/or coding parameters *based on its surroundings,* and called such a "smart" radio a *cognitive radio.* Such radios could even make decisions based on the availability of nearby collaborative nodes, or on the regulations dictated by their current location and spectral conditions.

The spectral conditions sensed by the cognitive radio may be utilized in any number of ways. In this chapter, we consider and survey the information theoretic results on three main categories of cognitive behavior:

1. **Interference mitigating cognitive behavior:** This behavior allows *two users to simultaneously transmit* over the same time or frequency band(s). Under this scheme, a cognitive radio will listen to the channel and, if sensed idle, could transmit during the void, not worrying about interference to the primary user (who is not transmitting). On the other hand, if another sender is sensed, the radio may decide to proceed with simultaneous transmission. The cognitive radio need not wait for an idle channel to start transmission. There will be interference between the primary and secondary users, but as we will show, this could potentially be mitigated. Here, the sensed information is fully utilized as side information, which will be the main aid in interference mitigation.
2. **Collaborative behavior (interference-free cognitive behavior):** When cognitive devices exist in a network but have no information of their own to transmit, they could potentially act as relays, and *collaborate* with the primary users. Rather than cause interference to the primary link, they boost it. Neglecting any other possibly active cognitive clusters [15], this system is interference-free. Incentives for cognitive radios to collaborate with primary users is beyond the scope of this chapter, but must also be considered. Here the sensing capability of the cognitive radio is used to obtain the message of the primary user, in order to relay it.
3. **Interference avoiding cognitive behavior:** In current FCC proposals on opportunistic channel usage, the cognitive radio listens to the wireless channel and determines, either in time or frequency, which part of the spectrum is unused [17]. It then adapts its signal to fill this void in the spectral domain, by either transmitting at a different time, or in a different band. A device transmits over a certain time and/or frequency band *only when no other user does*, thus avoiding interference, rather than mitigating it. Such behavior employs the sensing capability to determine a suitable moment, protocol, and band to transmit in.

1.3 Chapter Outline

The chapter is structured as follows. In section 2 we look at interference-mitigating cognitive behavior, where a prime example is the cognitive radio channel. We outline strategies and their resulting achievable rate regions (for general discrete memoryless cognitive radio channels [14, 12]) and capacity regions (for Gaussian cognitive radio channels [31]). We also demonstrate applicable and related results on interference channels with degraded message sets [47] and interference channels with unidirectional cooperation [27]. In section 3 we demonstrate that the multiplexing gain of the cognitive radio channel is 1. This somewhat pessimistic result motivates the definition of the cognitive X-channel in section 4. We study an achievable rate region for this channel before demonstrating that it achieves a multiplexing gain of 2 in section 5. In section 6 the limits of collaborative communications [38] are examined. There, the cognitive radio serves as a relay, and many of the previous idealistic assumptions often encountered in the relay channel literature are removed in establishing achievable rate regions. In section 7 we take a look at the capacity limits of interference-avoiding cognitive behavior. The problem of tracking and matching the cognitive transmitter and receiver channels in a distributed and dynamic spectral environment is posed, and capacity inner and outer bounds are examined.

2 Interference-mitigating Cognitive Behavior: the Cognitive Radio Channel

We start our discussion by looking at the simplest possible scenario in which a cognitive radio could be employed. We assume there exists a primary transmitter and receiver pair ($\mathcal{S}_1 \to \mathcal{R}_1$), as well as the cognitive secondary transmitter and receiver pair ($\mathcal{S}_2 \to \mathcal{R}_2$). As shown in Fig.1.1, there are three possibilities for transmitter cooperation in these two point-to-point channels. We have chosen to focus on transmitter cooperation because such cooperation is often more insightful and general than receiver-side cooperation [29, 11]. We thus assume that each receiver decodes independently. Transmitter cooperation in this figure is denoted by a directed double line. These three channels are simple examples of the cognitive decomposition of wireless networks seen in [15]. The three possible types of transmitter cooperation in this simplified scenario are:

1. **Competitive behavior:** The two transmitters transmit independent messages. There is no cooperation in sending the messages, and thus the two users *compete* for the channel. This is the same channel as the 2 sender, 2 receiver interference channel [5].
2. **Cognitive behavior:** Asymmetric cooperation is possible between the transmitters. This asymmetric cooperation is a result of $\mathcal{S}_2$ knowing $\mathcal{S}_1$'s message, but not vice-versa. As a first step, we idealize the concept of message knowledge: whenever the cognitive node $\mathcal{S}_2$ is able to hear and decode the message of the primary node $\mathcal{S}_1$, we assume it has full *a-priori* knowledge. We call this the *genie assumption*, as these messages could

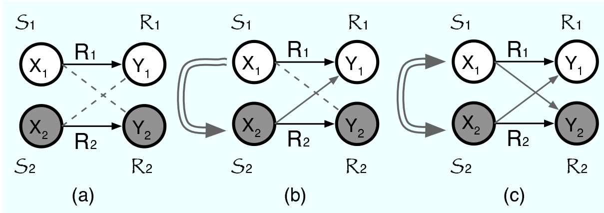

Fig. 1.1. (a) Competitive behavior, the interference channel. The transmitters may not cooperate. (b) Cognitive behavior, the cognitive radio channel. Asymmetric transmitter cooperation. (c) Cooperative behavior, the two antenna broadcast channel. The transmitters, but not the receivers, may fully and symmetrically cooperate.

have been given to the appropriate transmitters by a genie. The one way double arrow indicates that $\mathcal{S}_2$ knows $\mathcal{S}_1$'s message but not vice versa. This is the simplest form of asymmetric non-causal cooperation at the transmitters. We use the term cognitive behavior to emphasize the need for $\mathcal{S}_2$ to be a "smart" device capable of altering its transmission strategy according to the message of the primary user. We can motivate considering asymmetric side information in practice in three ways:

- Depending on the device capabilities, as well as the geometry and channel gains between the various nodes, certain cognitive nodes may be able to hear and/or obtain the messages to be transmitted by other nodes. These messages would need to be obtained in real time, and could exploit the geometric gains between cooperating transmitters relative to receivers in, for example, a 2 phase protocol [14].
- In an Automatic Repeat reQuest (ARQ) system, a cognitive transmitter, under suitable channel conditions (if it has a better channel to the primary transmitting node than the primary receiver), could decode the primary user's transmitted message during an initial transmission attempt. In the event that the primary receiver was not able to correctly decode the message, and it must be re-transmitted, the cognitive user would already have the to-be-transmitted message, or asymmetric side information, at no extra cost (in terms of overhead in obtaining the message).
- The authors in [47] consider a network of wireless sensors in which a sensor $\mathcal{S}_2$ has a better sensing capability than another sensor $\mathcal{S}_1$ and thus is able to sense two events, while $\mathcal{S}_1$ is only able to sense one. Thus, when they wish to transmit, they must do so under an asymmetric side-information assumption: sensor $\mathcal{S}_2$ has two messages, and the other has just one.

3. **Cooperative behavior:** The two transmitters know each others' messages (two way double arrows) and can thus fully and symmetrically cooperate in their transmission. The channel pictured in Fig.1.1 (c) may be

thought of as a two antenna sender, two single antenna receivers broadcast channel [44].

Many of the classical, well known information theoretic channels fall into the categories of competitive and cooperative behavior. For more details, we refer the interested reader to the cognitive network decomposition theorem of [15] and [12]. We now turn to the much less studied behavior which spans and in a sense interpolates between the symmetric cooperative and competitive behaviors. We call this behavior asymmetric cognitive behavior. In this section we will consider one example of cognitive behavior: a two sender, two receiver (with two independent messages) interference channel with asymmetric and a priori message knowledge at one of the transmitters, as shown in Fig.1.1(b). Certain asymmetric (in transmitter cooperation) channels have been considered in the literature: for example in [43], the capacity region of a multiple access channel with asymmetric cooperation between the two transmitters is computed. The authors in [27] consider a channel which could involve asymmetric transmitter cooperation, and explore the conditions under which the capacity of this channel coincides with the capacity of the channel in which *both messages are decoded at both receivers.* In [13, 12] the authors introduced the *cognitive radio channel*, which captures the most basic form of asymmetric transmitter cooperation for the interference channel. We now study the information theoretic limits of interference channels with asymmetric transmitter cooperation, or *cognitive radio channels.*

Our survey on the work on the 2 sender, 2 receiver channel with asymmetric cooperation at the transmitters will proceed as follows. First we will define and demonstrate an achievable rate region for the case of two independent messages for the discrete memoryless cognitive radio channel. This will be followed by the results of [31], who, under certain channel conditions, find the capacity region of the Gaussian interference channel with degraded message sets, a formulation equivalent to the Gaussian cognitive radio channel. We then consider the work of [47] on the general discrete memoryless interference channel with degraded message sets. In particular, they look for conditions under which the derived achievable rate regions are tight. In the Gaussian noise case, their result explicitly equals that of [31]. We then look at the work of [27] on the interference channel with unidirectional cooperation, where the capacity region of the cognitive radio channel when both messages are to be decoded at both receivers, under certain *strong interference* conditions, is derived. We proceed to explore the multiplexing gain of the Gaussian cognitive radio channel, which turns out to be 1. Motivated by this result, we define and derive an achievable rate region for the Gaussian X-channel with partial asymmetric (or cognitive) side information at the transmitter. In this case, the multiplexing gain turns out to be 2.

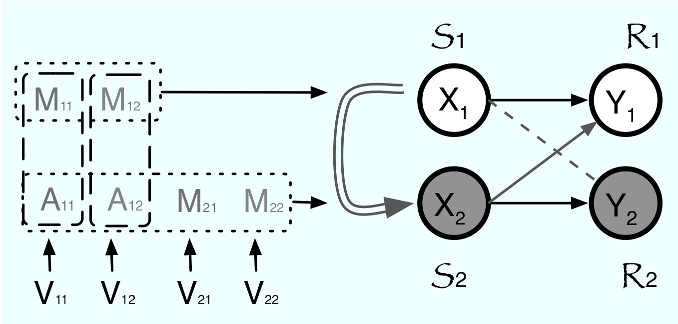

Fig. 1.2. The modified cognitive radio channel with auxiliary random variables M_{11}, M_{12} and M_{21}, M_{22}, inputs X_1 and X_2, and outputs Y_1 and Y_2. The auxiliary random variable A_{11}, A_{12} associated with $\mathcal{S}_2$, aids in the transmission of M_{11} and M_{12} respectively. The vectors V_{11}, V_{12}, V_{21} and V_{22} denote the effective random variables encoding the transmission of the private and public messages.

2.1 Cognitive radio channel: an achievable rate region

We define a 2×2 *genie-aided cognitive radio channel* C_{COG}, as in Fig.1.2, to be two point-to-point channels $\mathcal{S}_1 \to \mathcal{R}_1$ and $\mathcal{S}_2 \to \mathcal{R}_2$ in which the sender $\mathcal{S}_2$ is given, in a non-causal manner (i.e., by a genie), the message X_1 which the sender $\mathcal{S}_1$ will transmit. Let X_1 and X_2 be the random variable inputs to the channel, and let Y_1 and Y_2 be the random variable outputs of the channel. The conditional probabilities of the discrete memoryless C_{COG} are fully described by $P(y_1|x_1, x_2)$ and $P(y_2|x_1, x_2)$.

In [23], an achievable region for the interference channel is found by first considering a modified problem and then establishing a correspondence between the achievable rates of the modified and the original channel models. We proceed in the same fashion.

The channel C^m_{COG}, defined as in Fig.1.2 introduces many new auxiliary random variables, whose purposes can be made intuitively clear by relating them to auxiliary random variables in previously studied channels. They are defined and described in Table 1. Standard definitions of achievable rates and regions are employed [8, 13] and omitted for brevity. Then an achievable region for the 2×2 cognitive radio channel is given by:

Theorem 1. *Let* $Z \triangleq (Y_1, Y_2, X_1, X_2, V_{11}, V_{12}, V_{21}, V_{22}, W)$, *be as shown in Fig.1.2. Let* $\mathcal{P}$ *be the set of distributions on* Z *that can be decomposed into the form*

Table 1. Description of random variables and rates in Theorem 1.

(Random) variable names	(Random) variable descriptions
M_{11}, M_{22}	Private information from $\mathcal{S}_1 \to \mathcal{R}_1$ and $\mathcal{S}_2 \to \mathcal{R}_2$ resp.
M_{12}, M_{21}	Public information from $\mathcal{S}_1 \to (\mathcal{R}_1, \mathcal{R}_2)$ and $\mathcal{S}_2 \to (\mathcal{R}_1, \mathcal{R}_2)$ resp.
R_{11}, R_{22}	Rate between $\mathcal{S}_1 \to \mathcal{R}_1$ and $\mathcal{S}_2 \to \mathcal{R}_2$ resp.
R_{12}, R_{21}	Rate between $\mathcal{S}_1 \to (\mathcal{R}_1, \mathcal{R}_2)$ and $\mathcal{S}_2 \to (\mathcal{R}_1, \mathcal{R}_2)$ resp.
A_{11}, A_{12}	Variables at $\mathcal{S}_2$ that aid in transmitting M_{11}, M_{12} resp.
$V_{11} = (M_{11}, A_{11}), V_{12} = (M_{12}, A_{12})$	Vector helping transmit the private/public (resp.) information of $\mathcal{S}_1$
$V_{21} = M_{21}, V_{22} = M_{22}$	Public and private message of $\mathcal{S}_2$.
	Also the auxiliary random variables for Gel'fand-Pinsker coding
W	Time-sharing random variable, independent of messages

$$\begin{aligned}
&P(w) \times [P(m_{11}|w)P(m_{12}|w)P(x_1|m_{11}, m_{12}, w)] \\
&\times [P(a_{11}|m_{11}, w)P(a_{12}|m_{12}, w)] \\
&\times [P(m_{21}|v_{11}, v_{12}, w)P(m_{22}|v_{11}, v_{12}, w)] \\
&\times [P(x_2|m_{21}, m_{22}, a_{11}, a_{12}, w)]\, P(y_1|x_1, x_2)P(y_2|x_1, x_2), \qquad (1)
\end{aligned}$$

where $P(y_1|x_1, x_2)$ and $P(y_2|x_1, x_2)$ are fixed by the channel. Let $T_1 \triangleq \{11, 12, 21\}$ and $T_2 \triangleq \{12, 21, 22\}$. For any $Z \in \mathcal{P}$, let $S(Z)$ be the set of all rate tuples $(R_{11}, R_{12}, R_{21}, R_{22})$ (as defined in Table 1) of non-negative real numbers such that there exist non-negative reals $L_{11}, L_{12}, L_{21}, L_{22}$ satisfying:

$$\bigcap_{T \subset \{11,12\}} \left(\sum_{t \in T} R_t \right) \leq I(X_1; \mathbf{M}_T | \mathbf{M}_{\overline{T}}) \qquad (2)$$

$$R_{11} = L_{11} \qquad (3)$$

$$R_{12} = L_{12} \qquad (4)$$

$$R_{21} \leq L_{21} - I(V_{21}; V_{11}, V_{12}) \qquad (5)$$

$$R_{22} \leq L_{22} - I(V_{22}; V_{11}, V_{12}) \qquad (6)$$

$$\bigcap_{T \subset T_1} \left(\sum_{t_1 \in T} L_{t_1} \right) \leq I(Y_1, \mathbf{V}_{\overline{T}}; \mathbf{V}_T | W) + f(\mathbf{V}_T | W) \qquad (7)$$

$$\bigcap_{T \subset T_2} \left(\sum_{t_2 \in T} L_{t_2} \right) \leq I(Y_2, \mathbf{V}_{\overline{T}}; \mathbf{V}_T | W) + f(\mathbf{V}_T | W), \qquad (8)$$

where $f(\mathbf{v}_T)$ denotes the divergence between the joint distribution of the random variables $\mathbf{V}_T$ in (1) and their product distribution (where all components are independent). $\overline{T}$ denotes the complement of the subset T with respect to T_1 in (7), with respect to T_2 in (8), and $\mathbf{V}_T$ denotes the vector of V_i such that $i \in T$. Let S be the closure of $\cup_{Z \in \mathcal{P}} S(Z)$. Then any pair $(R_{11}+R_{12}, R_{21}+R_{22})$ for which $(R_{11}, R_{12}, R_{21}, R_{22}) \in S$ is achievable for C_{COG}.

Proof outline: The main intuition is as follows: the equations in (2) ensure that when $\mathcal{S}_2$ is presented with X_1 by the genie, the auxiliary variables M_{11} and M_{12} can be recovered. Eqs. (7) and (8) correspond to the equations for two overlapping MAC channels seen between the effective random variables $\mathbf{V}_{T_1} \to \mathcal{R}_1$, and $\mathbf{V}_{T_2} \to \mathcal{R}_2$. Eqs. (5) and (6) are necessary for the Gel'fand-Pinsker [21] coding scheme to work ($I(V_{21}; V_{11}, V_{12})$ and $I(V_{22}; V_{11}, V_{12})$ are the penalties for using non-causal side information. The $f(\mathbf{V}_T)$ terms correspond to the highly unlikely events of certain variables being correctly decoded despite others being in error. Intuitively, the sender $\mathcal{S}_2$ could aid in transmitting the message of $\mathcal{S}_1$ (the A_{11}, A_{12} random variables) or it could dirty paper code against the interference it will see (the M_{21}, M_{22} variables). The theorem smoothly interpolates between these two options. Details may be found in [14, 12].

2.2 Achievable rates for the Gaussian Cognitive Radio Channel

The previous section proposed inner and outer bounds on the capacity of the cognitive radio channel for discrete memoryless channels. Although the regions can be succinctly expressed, as done in Theorem 1, because this expression involves evaluation of the mutual information terms over all distributions of the specified form, it is unclear what these regions look like in general (and numerically intractable to try all possible input distributions). When the channel is affected by additive white Gaussian noise, as is often done in the literature, one can assume the input distributions to be of a certain form, and thus obtain a possible achievable rate region (not necessarily the largest one). In this section, we use this approach to arrive at the inner and outer bound regions shown in Fig.1.3

We consider the 2×2 Gaussian cognitive radio channel described by the input, noise and output relations:

$$Y_1 = X_1 + a_{21} X_2 + Z_1 \tag{9}$$

$$Y_2 = a_{12} X_1 + X_2 + Z_2 \tag{10}$$

where a_{12}, a_{21} are the crossover (channel) coefficients, $Z_1 \sim \mathcal{N}(0, Q_1)$ and $Z_2 \sim \mathcal{N}(0, Q_2)$ are independent AWGN terms, X_1 and X_2 are channel inputs constrained to to average powers P_1 and P_2 respectively, and $\mathcal{S}_2$ is given X_1 non-causally. Thus the Gaussian cognitive radio channel is simply the cognitive radio channel, where we have specified the conditional distributions which describe the channel, $p(y_1, y_2|x_1, x_2)$ to be of the above (9), (10) form. In order to determine an achievable region for the modified Gaussian cognitive radio channel, specific forms of the random variables described in Theorem 2 are assumed, and are analogous to the assumptions found in [12].

The resulting achievable region, in the presence of additive white Gaussian noise for the case of identical transmitter powers ($P_1 = P_2$) and identical receiver noise powers ($Q_1 = Q_2$), is presented in Figure 1.3. The ratio of transmit power to receiver noise power is 7.78 dB. The cross-over coefficients in the interference channel are $a_{12} = a_{21} = 0.55$, while the direct coefficients are 1.

In the figure, we see 4 regions. The time-sharing region (1) displays the result of pure time sharing of the wireless channel between users X_1 and X_2. Points in this region are obtained by letting X_1 transmit for a fraction of the time, during which X_2 refrains, and vice versa. The interference channel region (2) corresponds to the best known achievable region [23] of the classical information theoretic interference channel. In this region, both senders encode independently, and there is no *a-priori* message knowledge by either transmitter of the other's message. The cognitive channel region (3) is the achievable region described here and in [12]. In this case X_2 received the message of X_1 non-causally from a genie, and X_2 uses a coding scheme which combines interference mitigation with relaying the message of X_1. We see that both users – not only the incumbent X_2 which has the extra message knowledge – benefit from using this scheme. This is as expected: the selfish strategy boosts R_2 rates, while the selfless one boosts R_1 rates, and so gracefully combining the two will yield benefits to both users. Thus, the presence of the incumbent cognitive radio X_2 can be beneficial to X_1, a point which is of practical significance. This could provide yet another incentive for the introduction of such schemes.

The modified MIMO bound region (4) is an outer bound on the capacity of this channel: the 2 antenna Gaussian broadcast channel capacity region [44], where we have restricted the form of the transmit covariance matrix to be of the form $\begin{pmatrix} P_1 & c \\ c & P_2 \end{pmatrix}$, to more closely resemble our constraints, intersected with the capacity bound on $R_2 \leq I(Y_2; X_2|X_1)$ for the channel for $X_2 \to Y_2$ in the absence of interference from X_1. Let $H_1 = [1\ a_{21}]$ and $H_2 = [a_{12}\ 1]$. Then modified MIMO bound region is explicitly given by the set:

$$\text{Convex hull } \{(R_1, R_2) :$$

$$\begin{array}{l} R_1 \leq \frac{1}{2}\log_2\left(\frac{H_1(B_1+B_2)H_1^T+Q_1}{H_1B_2H_1^T+Q_1}\right) \\ R_2 \leq \frac{1}{2}\log_2\left(\frac{H_2B_2H_2^T+Q_2}{Q_2}\right) \\ R_2 \leq \frac{1}{2}\log_2\left(1+\frac{P_2}{Q_2}\right) \end{array} \bigcup \begin{array}{l} R_1 \leq \frac{1}{2}\log_2\left(\frac{H_1B_1H_1^t+Q_1}{Q_1}\right) \\ R_2 \leq \frac{1}{2}\log_2\left(\frac{H_2(B_1+B_2)H_2^T+Q_2}{H_2B_1H_2^T+Q_2}\right) \\ R_2 \leq \frac{1}{2}\log_2\left(1+\frac{P_2}{Q_2}\right) \end{array}$$

$$\text{for any 2x2 matrices } B_1, B_2 \text{ such that}$$

$$\begin{array}{c} B_1 \succeq 0, \quad B_2 \succeq 0 \\ B_1 + B_2 \preceq \begin{pmatrix} P_1 & c \\ c & P_2 \end{pmatrix} \\ c^2 \leq P_1P_2\} \end{array}$$

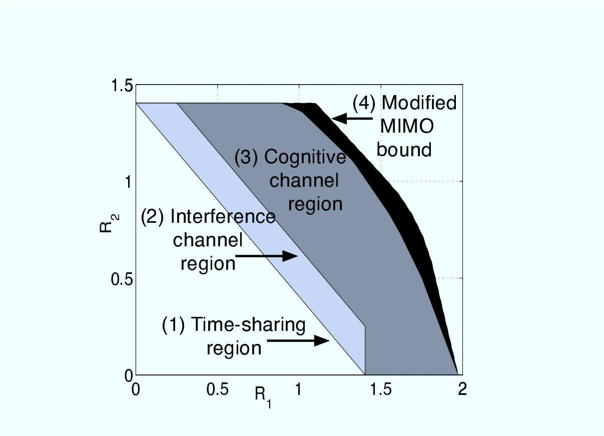

Fig. 1.3. Rate regions (R_1, R_2) for 2×2 wireless channels.

Here $X \succeq 0$ denotes that the matrix X is positive semi-definite.

2.3 Further results on the cognitive radio channel

Following the introduction of the *cognitive radio channel* [13], [31] and [47] considered the Gaussian cognitive radio channel, albeit under different names, and subsequently obtained its capacity in weak interference. The authors in [27] consider a channel which could involve asymmetric transmitter cooperation, and explore the conditions under which the capacity of this channel coincides with the capacity of the channel in which both messages are decoded at both receivers. We briefly review the results of these three works.

The authors of [31] consider a two sender two receiver channel which consists of a primary user and a secondary, or cognitive user. Like in the cognitive radio channel, each has its own independent message to send, and the cognitive user is assumed to know, a priori, the message of the primary user. They term their channel the *interference channel with degraded message sets (IC-DMS).* This work is particularly interested in determining the maximal rate at which the secondary cognitive user may transmit such that the primary user's rate remains unchanged (that is, the primary user's rate continues to be the same as if there were no interference), in the Gaussian noise channel. This would correspond to a single point in the capacity region of the channel in general. They furthermore require the primary receiver to employ a single-user decoder, which would be the case if no cognitive user were present. In essence, these two conditions, which they term *co-existence conditions,* require the cognitive user to remain transparent to the primary user. In fact, the only difference between the IC-DMS and the cognitive radio channel is that the IC-DMS, and all the results pertaining to it, are only valid in the Gaussian noise case. In addition, the *co-existence conditions* are not explicitly required in cognitive radio channel. In [31], these co-existence conditions are also relaxed (allowing for joint codebook design between primary and secondary users), and the authors show that the capacity achieving coding/decoding scheme in fact satisfy these *co-existence conditions*, that is, that the primary user decoder behaves as a single user decoder.

Let R_1 and R_2 denote the rates achieved by the primary and cognitive users, respectively. The main results of [31] stated in their Theorems 3.1 and 4.1 are summarized in the following single theorem. Here the primary user is expected power limited to P_1, the secondary user is expected power limited to P_2, and the noises at the two receivers are Gaussian of zero mean and variance N_1 and N_2 respectively. The conditions, and notation, are the same as in the Gaussian cognitive radio channel of section 2, save the *co-existence conditions.*

Theorem 2. *The capacity region of the IC-DMS defined in* (9), (10) *is given by the union, over all* $\alpha \in [0,1]$, *of the rate regions*

$$\begin{array}{l} 0 \leq R_1 \leq \frac{1}{2}\log_2\left(1 + \frac{(\sqrt{P_1}+a_{21}\sqrt{\alpha P_2})^2}{1+a_{21}^2(1-\alpha)P_2}\right) \\ 0 \leq R_2 \leq \frac{1}{2}\log_2\left(1+(1-\alpha)P_2\right) \end{array}$$

In particular, the maximal rate R_2 (or capacity) at which a cognitive user may transmit such that the primary user's rate R_1 remains as in the interference-free regime ($R_1 = \frac{1}{2}\log_2(1+P_1/N)$) is given by

$$\begin{array}{l} R_1 = \frac{1}{2}\log_2\left(1+\frac{P_1}{N}\right) \\ R_2 = \frac{1}{2}\log_2\left(1+(1-a^*)P_2\right). \end{array}$$

as long as $a_{21} < 1$,and a^* is

$$a^* = \left(\frac{\sqrt{P_1}\left(\sqrt{1+a_{21}^2P_2(1+P_1)}-1\right)}{a_{21}\sqrt{P_2}(1+P_1)}\right)^{\frac{1}{2}}.$$

Both these results are obtained using a Gaussian encoder at both the primary and cognitive transmitters. For more precise definitions of achievability in this channel, we refer to [31]. We paraphrase their achievability results here. The primary user generates its 2^{nR_1} codewords, X_1^n (block length n), by drawing the coordinates i.i.d. according to $\mathcal{N}(0,P_1)$, where we recall P_1 is the expected noise power constraint. Then, since the cognitive radio knows the message the primary user, it can form the primary user's encoding X_1^n, and performs superposition coding as:

$$X_2^n = \hat{X}_2^n + \sqrt{\frac{\alpha P_2}{P_1}}X_1^n,$$

where $\alpha \in [0,1]$. The codeword $\hat{X}_2^n$ encodes one of the 2^{nR_2} messages, and is generated by performing Costa precoding [6] (dirty-paper coding). Costa showed that to optimize the rate achieved by this dirty-paper coding, one selects $\hat{X}_2^n$ statistically independently from X_1^n, and thus i.i.d. Gaussian. Encoding is done using a standard information theoretic binning technique, which treats the message X_1^n as non-causally known interference. In order to satisfy the average power constraint of P_2 on the components of X_2^n, $\hat{X}_2^n$ must be $\mathcal{N}(0,(1-\alpha)P_2)$. A converse, resulting in the capacity region of the cognitive radio channel under weak interference, is given in [31] and is based on the conditional entropy power inequality, and results from [44].

Whereas the paper [31] considers only the Gaussian IC-DMS with specific *co-existence conditions*, the work [47] considers the discrete memoryless IC-DMS (not necessarily Gaussian), and looks at the Gaussian IC-DMS as a special case. The authors in this work are motivated by a sensor network in which one sensor has better sensing capabilities than another. The one with the better channel is thus able to detect two sensed events, while another is only able to detect one. This problem then reduces to the interference channel with degraded message sets (where the message of one user is a subset of the other user's message). The authors define three types of *weak interference* (as opposed to the *very strong* and *strong* interference typically seen in the interference channel literature [5]), an achievable rate region, outer bounds, and conditions under which these outer bounds are tight. They then look at a Gaussian noise example in which their region is tight, and for which the result is as described in the capacity region of [31]. We summarize some of their main results in the single following theorem. It provides an inner and an outer bound on the IC-DMS, which turns out to be the capacity region for the types of interference specified.

Theorem 3. Inner bound: *Let $\mathcal{R}_{in}$ be the set of all rate pairs (R_1, R_2) (same as in the cognitive radio channel) such that*

$$\begin{aligned} R_1 &\le I(V, X_1; Y_1) \\ R_2 &\le I(U; Y_2) - I(U; V, X_1) \end{aligned}$$

for the probability distribution $p(x_1, x_2, u, v, y_1, y_2)$ that factors as

$$p(v, x_1)p(u|v, x_1)p(x_1|u)p(y_1, y_2|x_1, x_2).$$

Then $\mathcal{R}_{in}$ is an achievable rate region for the IC-DMS where transmitter $\mathcal{S}_2$ knows both messages and transmitter $\mathcal{S}_1$ only knows one.
Outer bound: *Define $\mathcal{R}_o$ to be the set of all rate pairs (R_1, R_2) such that*

$$\begin{aligned} R_1 &\le I(V, X_1; Y_1) \\ R_2 &\le I(X_1; Y_2|X_1) \\ R_1 + R_2 &\le I(V, X_1; Y_1) + I(X_2; Y_2|V, X_1), \end{aligned}$$

for the probability distribution $p(x_1, x_2, v, y_1, y_2)$ *that factors as*

$$p(v, x_1)p(x_2|v)p(y_1, y_2|x_1, x_2).$$

Then R_o *is an outer bound for the capacity of the IC-DMS.*
Capacity conditions: *If there exists a probability transition matrix* $q_1(y_2|x_2, y_1)$ *such that*

$$p(y_2|x_1, x_2) = \sum_{y_1} p(y_1|x_1, x_2)q_1(y_2|x_2, y_1),$$

or if there exists a probability transition matrix $q_2(y_1|x_1, y_2)$ *such that*

$$p(y_1|x_1, x_2) = \sum_{y_2} p(y_2|x_1, x_2)q_2(y_1|x_1, y_2),$$

then the set of all rate pairs (R_1, R_2) *such that*

$$R_1 \leq I(V, X_1; Y_1) \tag{11}$$
$$R_2 \leq I(X_2; Y_2|V, X_1) \tag{12}$$

for the probability distribution $p(x_1, x_2, y_1, y_2)$ *that factors as*

$$p(v, x_1)p(x_2|v)p(y_1, y_2|x_1, x_2),$$

is the capacity region of the IC-DMS.

Since the channel of [47] is the same as the cognitive radio channel [14, 12], direct comparisons between their resepective bounds may be made. Whereas the outer bounds are equivalent, due to the fact that the inner bounds for the discrete memoryless channel involve non-trivial unions over all distributions of a certain form, it is unclear a priori which region will be larger. However, the authors demonstrate that all *Gaussian* weak interference channels satisfy the *capacity conditions* of the theorem, and thus the region of (11)-(12) is the capacity region. This capacity region in the Gaussian noise case is shown to be explicitly equal to that of [31], and, numerically, to that of of [12], specialized to the Gaussian noise case.

Finally, the work [27] considers again the cognitive radio channel, referred to as the *interference channel with unidirectional cooperation*. There, one set of conditions for which the capacity region of the channel coincides with that of the channel in which both messages are required at both receivers is derived. Notice that in the cognitive radio channel this added condition, of being able to decode both messages at both receivers, is not assumed. This is related to the work [26] on the compound multiple access channel with common information, in which the capacity region for another set of *strong interference*-type conditions is computed. Notice that whereas [47] considers *weak interference* conditions, [27] considers *strong interference conditions.* Their results on the cognitive radio channel capacity read as follows:

Theorem 4. *For an interference channel with unidirectional cooperation satisfying*

$$I(X_2; Y_2|X_1) \leq I(X_2; Y_1|X_1)$$
$$I(X_1, X_2; Y_1) \leq I(X_1, X_2; Y_2)$$

for all joint distributions on X_1 *and* X_2*, the capacity region* $\mathcal{C}$ *is given by*

$$\begin{aligned} \mathcal{C} = \bigcup \{(R_1, R_2) :& \\ R_2 &\leq I(X_2; Y_2|X_1) \\ R_1 + R_2 &\leq I(X_1, X_2; Y_1)\}, \end{aligned}$$

where the union is over joint distributions $p(x_1, x_2, y_1, y_2)$.

2.4 Cognitive Radio Channel Conclusions

As we have seen, various authors have studied the fundamental information theoretic limits of *cognitive behavior*, albeit sometimes under different names, with the common idea of partial asymmetric side information at one transmitter. In addition, in Gaussian noise, it can be seen that cognitive behavior allows for a secondary user to transmit at a non-zero rate while the primary user remains unaffected. Alternatively, tradeoffs between the primary and secondary users' rates can also be analyzed. The capacity regions are known under certain conditions, but as is the case for the interference channel, the capacity region of the most general discrete memoryless cognitive radio channel remains an open problem.

3 The Multiplexing Gain of Cognitive Radio Channels

The previous section showed that when two interfering point-to-point links act in a *cognitive* fashion, or employ asymmetric non-causal side information, interference may be at least partially mitigated, allowing for higher spectral efficiency. That is, it is possible for the cognitive user to communicate at a non-zero rate while the primary user suffers no loss in rate. Thus, at medium SNR levels, there is an advantage to cognitive transmission. One immediate question that arises is how cognitive transmission performs in the high SNR regime. The *multiplexing gain* is defined as the limit of the ratio of the maximal achieved rate to the log(SNR) as the SNR tends to infinity. That is,

$$\text{multiplexing gain} = \lim_{\max\,\text{SNR}\to\infty} \frac{R(\text{SNR})}{\log(\text{SNR})}.$$

The multiplexing gain of various multiple input multiple output (MIMO) systems has been extensively studied in the literature [30]. For the single user point-to-point MIMO channel with M_T transmit and N_R receive antennas, the maximum multiplexing gain is known to be $\min(M_T, N_R)$ [**?**, 42]. For the two user MIMO multiple access channel with N_R receive antennas and M_{T_1}, M_{T_2} transmit antennas at the two transmitters, the maximal multiplexing gain is $\min(M_{T_1} + M_{T_2}, N_R)$. For its counterpart, the two user MIMO broadcast channel with M_T transmit antennas and N_{R_1}, N_{R_2} receive antennas at the two transmitters respectively, the maximum multiplexing gain is $\min(M_T, N_{R_1} + N_{R_2})$. These results, as outlined in [30] demonstrate that when joint signal processing is available at either the transmit or receive sides (as is the case in the MAC and BC channels), then the multiplexing gain is significant. However, when joint processing is neither possible at the transmit nor receive side, as is the case for the interference channel, then the multiplexing gain is severely limited. Results for the maximal multiplexing gain when cooperation is permitted at the transmitter or receiver side through noisy communication channels can be found in [25, 24]. In the cognitive radio channel, a form of partial joint processing is possible at the transmitter. It is thus unclear whether this channel will behave more like the cooperative MAC and BC channels, or whether it will suffer from interference at high SNR as in the interference channel. We thus outline results on the multiplexing gain in this scenario, under additive white Gaussian noise [16].

We expect the multiplexing gain (which intuitively corresponds to the number of information streams one can push through a channel) to lie somewhere between 1 and 2, as we have two independent messages, and single antennas at all nodes. One can show that the sum-rate of the Gaussian cognitive radio channel, with two independent messages $\mathcal{S}_1 \to \mathcal{R}_1$ and $\mathcal{S}_2 \to \mathcal{R}_2$, as shown in Fig.1.4(a) scales at best like $\log P$ (not $2 \log P$). In other words, although partial side information may help the interference channel in a medium SNR-regime [14, 31], at high SNR, one cannot improve the scaling law of the sum-rate.

Theorem 5. *Consider a Gaussian interference channel defined in* (9)*,* (10)*, and where additionally $\mathcal{S}_2$ has non-causal knowledge of the message of $\mathcal{S}_1$. Then the sum-rate capacity of this channel satisfies*

$$\lim_{P \to \infty} \frac{\max_{(R_1, R_2) \in \mathcal{C}} R_1 + R_2}{\log P} = 1, \tag{13}$$

where R_i corresponds to the rates from the i-th source to the i-th receiver, P is the expected transmit power constraint at each transmitter and $\mathcal{C}$ is the capacity region of the channel.

Proof. The $a_{21} \leq 1$ condition ensures that we are operating in the *weak interference* regime. Consider the capacity region denoted by $\mathcal{C}$ claimed in eqns. (24) and (25) of [31]. Notice that $a_{21} \leq 1$ corresponds to $a \leq 1$, $P_1 = P_p$, $P_2 = P_c$, and $R_1 = R_p, R_2 = R_c$ in the notation of [31]. If $P_p = P_c = P$, then it follows that

$$\lim_{P\to\infty} \frac{\max_{(R_1,R_2)\in\mathcal{C}} R_1 + R_2}{\log P} \tag{14}$$

$$= \lim_{P\to\infty} \frac{\max_\alpha \frac{1}{2}\log\left(\left(1 + \frac{(\sqrt{P}+a\sqrt{\alpha P})^2}{1+a^2(1-\alpha)P}\right)\cdot(1+(1-\alpha)P)\right)}{\log P} \tag{15}$$

$$= 1 \tag{16}$$

For the case when $a_{21} > 1$ the sum-rate again scales like $\log(P)$, which can be seen by using Theorem 2.

4 The X-Channel with Asymmetric Side Information

Section 2 showed that when two non-overlapping single sender, single receiver channels act in a *cognitive* fashion, or employ asymmetric non-causal side information, interference may be at least partially mitigated, allowing for higher spectral efficiency. In this scenario, the two senders and the two receivers were independent. However, at high SNR, the multiplexing gain was limited to 1. This is in fact equal to that of a channel with no cognition. We ask ourselves if there are other *cognitive* channels in which partial asymmetric message knowledge does provide a multiplexing gain greater than 1. The answer, as we will see in the next section, is yes. The channel for which the multiplexing gain using partial asymmetric side information is the cognitive *X-channel*, which we define next. This channel is equivalent to the cognitive version of the X-channel, defined in [30, 36], where the degrees of freedom, or multiplexing gain, is considered in the multiple antenna, non cognitive case. We will ultimately be interested in the multiplexing gain for Gaussian noise channels, and so introduce the Gaussian cognitive radio channel, and the Gaussian cognitive X-channel.

Repeating for clarity, in the cognitive radio channel, defined in Section 2.1 and shown in Fig. 1.4(a), there are two messages, one from $(\mathcal{S}_1 \to \mathcal{R}_1)$, and the other from $(\mathcal{S}_2 \to \mathcal{R}_2)$. There is no cross-over information from $(\mathcal{S}_1 \to \mathcal{R}_2)$ or $(\mathcal{S}_2 \to \mathcal{R}_1)$. Here $\mathcal{S}_2$ knows the message X_1, as seen by the directed double arrow in Fig. 1.4(a). The multiplexing gain of this channel is 1. Consider now the same 2 sender, 2 receiver Gaussian noise channel as Fig. 1.4(a) except that here we *do* have cross-over information. That is, each sender has an independent message destined to each receiver, for a total of four messages, as shown in Fig. 1.4(b). $\mathcal{S}_1$ wishes to send message $s_{11} \in \{1, 2, \cdots 2^{nR_{11}}\}$, encoded as $A_1 \in \mathcal{A}_1$ to $\mathcal{R}_1$ (at rate R_{11}) and $s_{12} \in \{1, 2, \cdots, 2^{nR_{12}}\}$, encoded as $A_2 \in \mathcal{A}_2$ to $\mathcal{R}_2$ (at rate R_{12}) in n channel uses. Similarly, $\mathcal{S}_2$ wishes to send message $s_{21} \in \{1, 2, \cdots 2^{nR_{21}}\}$, encoded as $B_1 \in \mathcal{B}_1$ to $\mathcal{R}_1$ (at rate R_{21}) and $s_{22} \in \{1, 2, \cdots, 2^{nR_{22}}\}$ encoded as $B_2 \in \mathcal{B}_2$ to $\mathcal{R}_2$ (at rate R_{22}) in n channel uses. The double arrow from X_1 to X_2 denotes *partial side information*, specifically, that the encoding A_1 is known fully, non-causally (or *a-priori*) to the second

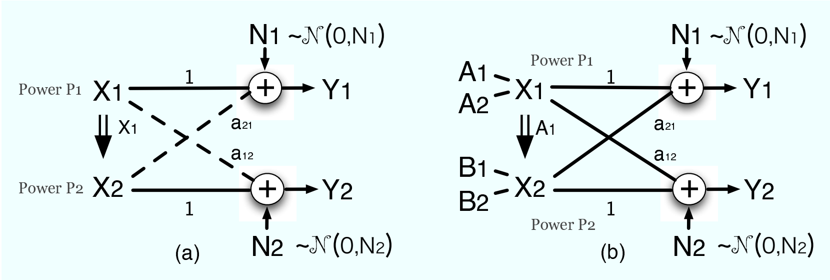

Fig. 1.4. Both channels are additive Gaussian noise interference channels with cross-over parameters α_{12}, α_{21}, transmitted encodings X_1, X_2 with expected transmit power limitations P_1 and P_2, and received signals Y_1 and Y_2. (a) Cognition in the interference channel: there are two information streams $(X_1 \to Y_1)$ and $(X_2 \to Y_2)$, and X_1 is the asymmetric side information known at X_2. (b) Cognition in the X-channel: there are four message streams $(A_1 \to Y_1)$, $(A_2 \to Y_2)$, $(B_1 \to Y_1)$ and $(B_2 \to Y_2)$. A_1 is the partial and asymmetric message knowledge at X_2.

transmitter. Notice also that only one of $\mathcal{S}_1$'s messages is known to $\mathcal{S}_2$, that is, only *partial* knowledge is used in the following. We could alternatively have allowed A_2 to be known at the second transmitter. This would lead to analogous results when indices are permuted. The channel is still an additive Gaussian noise channel with independent noise at the receivers, so the received signals are:

$$Y_1 \;\; = A_1 + A_2 + a_{21}(B_1 + B_2) + N_1 \tag{17}$$

$$Y_2 = a_{12}(A_1 + A_2) + (B_1 + B_2) + N_2. \tag{18}$$

Standard definitions of achievable rates and regions are employed [8, 13] or chapters 8 and 14 of [8]. Although our achievable rate region will be defined for finite alphabet sets, in order to determine an achievable region for the Gaussian noise channel, specific forms of the random variables described in Theorem 6 are assumed. As in [6, 19, 23], Theorem 6 can readily be extended to memoryless channels with discrete time and continuous alphabets by finely quantizing the input, output, and interference variables (Gaussian in this case).

We now outline an achievable region for this Gaussian noise channel. The capacity region of the Gaussian MIMO broadcast channel [44] is achieved by Costa's dirty-paper coding techniques [6]. In the X-channel, at $\mathcal{S}_1$, the encodings A_1 and A_2 may be jointly generated, for example using a dirty-paper like coding scheme. That is, one message may treat the other as non-causally known interference and code so as to mitigate it. At $\mathcal{S}_2$, not only may the encodings B_1 and B_2 be jointly designed, but they may additionally use A_1 as *a-priori* known interference. Thus, transmitter 2 could encode B_2 so as to potentially mitigate the interference Y_2 will experience from A_1 as well as B_1.

We demonstrate an achievable region for the discrete, finite alphabet case in Theorem 6 and look at the achieved rate scalings in the Gaussian noise case, assuming specific forms for all involved variables in Theorem 6. Let R_{11} be the rate from $A_1 \to Y_1$, R_{12} from $A_2 \to Y_2$, R_{21} from $B_1 \to Y_1$ and R_{22} from $B_2 \to Y_2$.

Theorem 6. *Let* $Z \triangleq (Y_1,Y_2,X_1,X_2,A_1,A_2,B_1,B_2)$*, and let* $\mathcal{P}$ *be the set of distributions on* Z *that can be decomposed into the form*

$$\begin{array}{c} p(a_1|a_2)p(a_2)p(b_1)p(b_2|a_1,b_1) \\ p(x_1|a_1,a_2)p(x_2|a_1,b_1,b_2) \\ p(y_1|x_1,x_2)p(y_2|x_1,x_2), \end{array} \tag{19}$$

where we additionally require $p(a_2,b_2) = p(a_2)p(b_2)$*. For any* $Z \in \mathcal{P}$*, let* $S(Z)$ *be the set of all tuples* $(R_{11},R_{12},R_{21},R_{22})$ *of non-negative real numbers such that:*

$$\left.\begin{aligned} R_{11} &\le I(A_1;Y_1|B_1) - I(A_1;A_2) \\ R_{21} &\le I(B_1;Y_1|A_1) \\ R_{11}+R_{21} &\le I(A_1,B_1;Y_1) - I(A_1;A_2) \end{aligned}\right\} \begin{array}{c} MAC \\ (A_1,B_1) \\ \searrow\swarrow \\ Y_1 \end{array}$$

$$\left.\begin{aligned} R_{12} &\le I(A_2;Y_2|B_2) \\ R_{22} &\le I(B_2;Y_2|A_2) - I(B_2;A_1,B_1) \\ R_{12}+R_{22} &\le I(A_2,B_2;Y_2) - I(B_2;A_1,B_1) \end{aligned}\right\} \begin{array}{c} MAC \\ (A_2,B_2) \\ \searrow\swarrow \\ Y_2 \end{array}$$

Let S *be the closure of* $\cup_{Z\in\mathcal{P}} S(Z)$*. Then any element of* S *is achievable.*

Proof. The codebook generation, encoding, decoding schemes and formal probability of error analysis are deferred to the manuscript in preparation [16]. Heuristically, notice that the channel from $(A_1,B_1) \to Y_1$ is a multiple access channel with encoders that are possibly correlated [7, 45] and employ dirty paper coding [21, 6]. However, by (19) we see that A_1 and B_1 are in fact independent, and thus the regular MAC equations hold. A_1 does use a binning scheme with respect to A_2, but this does not alter the $(A_1,B_1) \to Y_1$ MAC equations other than reduce the rate R_{11} by $I(A_1;A_2)$ (like in Gel'fand-Pinsker [21] coding). Similarly, for the MAC $(A_2,B_2) \to Y_2$ the encodings A_2 and B_2 are independent (this is true in particular in the Gaussian case of interest in the next subsection, and so we simplify our theorem by ensuring the condition $p(a_2,b_2) = p(a_2)p(b_2)$) so that the regular MAC equations also hold here. Again, there is a penalty of $I(B_2;A_1,B_1)$ for the rate R_{22} incurred in order to guarantee finding an n-sequence b_2 in the desired bin that is jointly typical with any *given* a_1, b_1 pair.

5 Multiplexing Gains in Overlapping Cognitive Broadcast Channels

The multiplexing gain of the Gaussian cognitive radio channel was shown to be 1. We now proceed to examine the multiplexing gain of the cognitive Gaussian X-channel. We wish to see how the achievable rate tuple varies as a function of the transmit powers, or equivalently, of the SNRs when the white Gaussian noise variance is held fixed. To do so, the achievable rate region is evaluated in the proof of the following corollary, which emphasizes that the sum-rate of two the X-channel with *partial non-causal side information* has a multiplexing gain of 2.

Corollary 1. *Consider the Gaussian X-channel with asymmetric side information described in Theorem 6. Then*

$$\lim_{P\to\infty} \frac{\max_{(R_{11},R_{12},R_{21},R_{22})\in\mathcal{C}_{OBC}} R_{11}+R_{12}+R_{21}+R_{22}}{\log P} = 2, \tag{20}$$

where $\mathcal{C}_{OBC}$ *is the capacity region of the cognitive X-channel.*

Proof. First, note that the multiplexing gain of a single sender, 2 receiver broadcast channel is 2, and as this channel's capacity region provides an upper bound to our channel's region, we cannot have a multiplexing gain larger than 2. We will in fact prove that 2 is achievable using the scheme of Theorem 6. To prove this result, we specify forms for the variables, and then optimize the dirty paper coding parameters, similar to Costa's technique [6]. The Gaussian distributions we assume on all variables are of the form

$$\begin{array}{ll}
A_1 = U_1 + \gamma_1 U_2 & U_1 \sim \mathcal{N}(0, P_{11}) \\
A_2 = U_2, & U_2 \sim \mathcal{N}(0, P_{12}) \\
B_1 = V_1, & V_1 \sim \mathcal{N}(0, P_{21}) \\
B_2 = V_2 + \gamma_2(V_1 + a_{12}U_1) & V_2 \sim \mathcal{N}(0, P_{22}) \\
 & \\
X_1 = U_1 + U_2 \quad \sim \mathcal{N}(0, P_1) & P_1 = P_{11} + P_{12} \\
X_2 = V_1 + V_2 \quad \sim \mathcal{N}(0, P_2) & P_2 = P_{21} + P_{22} \\
Y_1 = U_1 + U_2 + a_{21}(V_1 + V_2) + N_1 & N_1 \sim \mathcal{N}(0, N_1) \\
Y_2 = a_{12}(U_1 + U_2) + (V_1 + V_2) + N_2 & N_2 \sim \mathcal{N}(0, N_2).
\end{array}$$

Here the variables U_1, U_2, V_1, V_2 are all independent, encoding the four messages to be transmitted. Notice that here $p(a_1, b_1) = p(a_1)p(b_1)$ and $p(a_2, b_2) = p(a_2)p(b_2)$ as needed in Theorem 6. The sum rates $R_1 = R_{11} + R_{21}$ and $R_2 = R_{12} + R_{22}$ to each receiver can be calculated separately. Each can be maximized with respect to the relevant dirty-paper coding parameter (γ_1 for $\mathcal{S}_1$, and γ_2 for $\mathcal{S}_2$). The bounds of Theorem 6 may be evaluated by combining the appropriate determinants of sub-matrices of the overall covariance matrix $E[\Theta\Theta^T]$ where $\Theta \stackrel{\triangle}{=} (A_1, B_1, A_2, B_2, Y_1, Y_2)$. The details may be found in [16]. The main idea is that when the dirty paper coding parameters are properly chosen, and when we let the powers $P_{11} = P_{12} = P_{21} = P$ scale like

$P \to \infty$ while keeping P_{22} fixed, then the multiplexing gain of 2 is achieved. Keeping P_{22} fixed is crucial for achieving the $\log P$ scaling in R_1. Intuitively, this is because of asymmetric message knowledge; the interference the second cognitive transmitter causes the first is not mitigated. Keeping P_{22} constant still allows the second transmitter to dirty paper code, or mitigate the interference caused by A_1 and B_1 to the second receiver's signal Y_2, while causing asymptotically (as $P_{11}, P_{12}, P_{21} \to \infty$) negligible interference to Y_1. This is a remarkable fact: only partial side information is needed to attain the full multiplexing gain of a broadcast channel with a two antenna transmitter.

6 Collaborative Communications

We now consider another example of cognitive behavior where rather than having two independent messages to be transmitted, there is only one message to be sent from a given source to a given destination, possibly with the help of a relay. This relay help can be considered as asymmetric transmitter cooperation, or cognitive behavior. We first survey some relay channel results before moving onto the case considered in [38], which has removed many of the classical, and somewhat unrealistic constraints,

6.1 The relay channel

The relay channel, which in its simplest and most classical form is a three-terminal channel with one source, one relay (without its own information to transmit) and one destination, is another example of *cognitive behavior*. Relay channels were introduced by van der Meulen [43], and various variations of the problem were later studied by others [1], [9]. The current state of the art is well summarized in [33].

The classical relay channel is shown in Fig.1.5. It consists of a source, with information, a relay, with no independent information of its own, and destination. Here, as in the cognitive radio channel, full channel-state information is assumed at all terminals. The paper [9] introduced two fundamental coding schemes for the relay channel often called Decode-and-Forward (DF, Theorem 1) and Compress-and-Forward (CF, Theorem 6). This formulation may be extended to multiple relays, as done in [1], [22] and improved in [49], [48]. We defer to the very informative and insightful [33] for further information on relay channels.

Three major issues are ignored in the classical relay channel framework: the half-duplex constraint of most practical wireless systems, the compound nature, and the non-degraded nature of most wireless channels. To elaborate,

1. The first constraint often ignored in the classical relay-channel framework is the *duplex* constraint. Most of the results on relay channel assume full-duplex relays, that is, relays which may receive and transmit simultaneously. In realistic wireless channel, this assumption begins to break

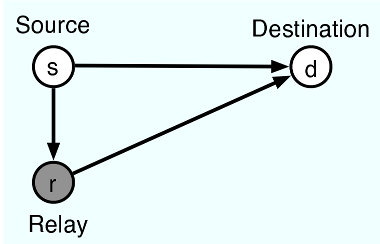

Fig. 1.5. The classical relay channel has a source, with information, a relay, with no information, and a destination. The relay aids the source in transmitting its message to the destination.

down, since the intensity of the near-field of the transmitted signal is much higher than that of the far field of the received signal. In essence, a full-duplex relay would, in practice, interfere with itself. Cognitive relay schemes which operate under a *half-duplex* constraint, that is, where a node cannot simultaneously transmit and receive data, must be considered.

Although the capacity of a half duplex relay channel has yet to be found, there has been a large body of work to understand optimal schemes in the asymptotic regimes of low and high signal to noise ratio (SNR) in slowly fading wireless channels [20, 4, 3, 2]. In large SNR, the outage capacity of such a channel has been analyzed in [4, 34]. Interestingly, it is proved that for small multiplexing gains, the diversity gain achieved by the relay channel matches the maximum diversity gain achieved by max-flow min-cut bound in Rayleigh fading channels [4]. In other words, for small multiplexing gains r, i.e., $r \leq \frac{1}{2}$, the relay channel can provide the same diversity gain as that of a system with two transmit antennas and a receiver with a single antenna. This result is achieved by a variation of decode and forward (DF) scheme in which the relay starts forwarding the message as soon as it can decode the message.

As for the low SNR regime, it has been recently shown that the decode and forward scheme is strictly suboptimal in terms of outage capacity [3]. It is further proved that a bursty variant of the Amplify and Foward cooperation scheme in which the source broadcasts with a larger power $\frac{P}{\alpha}$ for a short fraction α of the transmission time and then remains silent for the rest of the time [20, 3], is outage optimal for Rayleigh fading channels. Intuitively, sending bursty signals with high power significantly improves the quality of the received signal at the relay. This scheme turns out to be optimal not only for Rayleigh fading channels, but also for a wide class of channel distributions, namely the distributions that are analytic in the neighborhood of zero [2].

2. The second assumption often made in the context of wireless communications is the quasi-static fading model. That is, traditionally many authors

assume that the fading coefficients remain fixed for the entire duration of the transmission frame. In an information theoretic framework, where block lengths tend to infinity, all realizations of a channel are thus *not* experienced in a frame, and ergodic capacity results seem limited in their applicability. This, in addition to the fact that the channel state is often not known to the transmitters but only to receivers motivates the study of more realistic *compound* channels [46, 10].

3. Finally, while the degraded relay channel has been completely solved [9, 39], in wireless systems most noise is due to thermal noise in the receiver frontend. While it may be reasonable to assume that the relay has a better signal to noise ratio (SNR) than the ultimate receiver, it is unrealistic to assume that the receiver is a degraded version of the relay.

These three drawbacks of traditional approaches to the relay channel motivate the study of non-degraded compound relay channels which satisfy the half-duplex constraint. In [38] the authors investigate a bandwidth efficient decode and forward approach that does not employ predetermined phase durations or orthogonal sub-channels to resolve the half-duplex constraint: each relay determines based on its own receive channel when to listen and when to transmit. Furthermore, the transmitters are not aware of the channel and no assumption of degradedness are made: the noise at the relays is independent of that at the destination. Also, as opposed to previous relay and collaborative literature, the results still hold under a bounded asynchronous model. Finally, in the case of multiple relays assisting the source, their approach permits one relay to assist another in receiving the message, a feature not present in much of the early work on communications over compound channels. However, more recent work along this line may be found in [32, 4].

6.2 Collaborative Communications

We now present a brief summary of this important and alternate view of the compound relay channel [38], which is a prime example of cognitive behavior in a network where the cognitive nodes do not have information of their own to send. The authors of [38] use the term *collaborative communications* to describe their category of work. This falls into the category of cognitive behavior in the setting considered here.

Spatial diversity is the term often used to capture the potential gain (reliability in this case) of independent paths between sources and destinations which result from spatial separation of nodes or antennas. Of primary interest then is to determine if one can achieve the genie bound on diversity: the diversity gain that would be achieved if all the transmit antennas of the source and relay nodes were in fact connected to a single node (in [35, 34] this is referred to as the transmit diversity bound). For example, consider the three transmit collaborators and one receiver node scenario (each equipped with a single antenna) as illustrated in Fig.1.6. If all the collaborators were aware

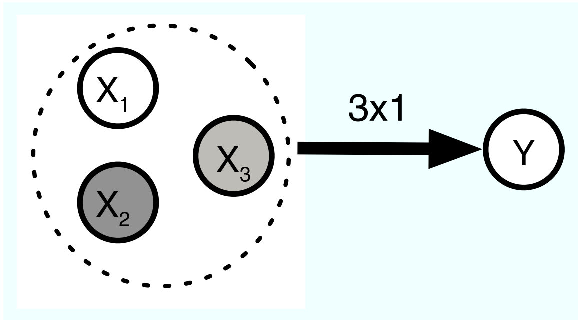

Fig. 1.6. Is an ideal 3×1 space-time gain achievable with three separate transmit nodes and one receive node.

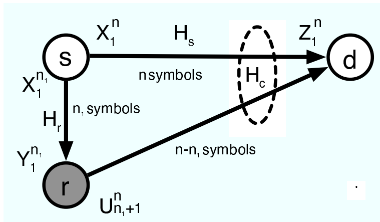

Fig. 1.7. The collaborative communications problem for two transmit collaborators and one receiver.

of the message *a priori*, one could in principle achieve the ideal performance of a 3×1 space-time system between the transmit cluster and the receiver node. However, only the source node in the transmit cluster is aware of the message *a priori*. The other two nodes in the cluster must serve as relays and are not aware of the message *a priori*. There will be a loss in performance (as measured by the probability of outage) compared to the idealized 3×1 space-time system. In particular, the authors in [38] are interested in determining sufficient conditions on the geometry and signal path loss of the transmitting cluster for which performance close to the genie bound can be guaranteed.

To determine an upperbound on this loss, the authors [38] derive a novel approach to the compound relay channel. This approach is best summarized as follows. In a traditional compound channel, a set of possible channel realizations are given and one seeks to prove the existence of a code (with maximal rate) which is simultaneously good on all channel realizations. In [38], the problem is framed in the opposite direction. They fix a rate and ask how large the set of compound channels can be made while guaranteeing that the code is still good.

Consider three nodes denoted as source (s), relay (r) and destination (d) as illustrated in Fig.1.7 and each equipped with N_s, N_r and N_d antennas respectively (the results readily generalize to multiple relay nodes).

It is assumed that while listening to the channel, the relay may not transmit, satisfying the half-duplex constraint. Hence, the communications protocol proposed is as follows. The source node wishes to transmit one of 2^{nR} messages to the destination employing n channel uses. While not transmitting, the relay node listens. Due to the relay node's proximity to the source, after n_1 samples from the channel (a number which the relay determines on its own and for which the source has no knowledge), it may correctly decode the message. After decoding the message, it then proceeds to transmit for the remaining $n - n_1$ transmissions in an effort to improve the reception of the message at the destination. The destination is assumed to be made aware of n_1 before attempting to decode the message. This may be achieved by an explicit low-rate transmission from the relay to the destination. Alternatively, if the value of n_1 is constrained to some integer multiple of a fundamental period n_0 (say $n_0 \sim \sqrt{n}$), then the destination may estimate n_1 accurately using power detection methods. Denote the first phase of the n_1 transmissions as the *listening* phase and the last $n - n_1$ transmissions as the *collaboration* phase.

All channels are modeled as additive white Gaussian noise (AWGN) with quasi-static fading. In particular, X and U are column vectors representing the transmission from the source and relay nodes respectively and denote by Y and Z the received messages at the relay and destination respectively. Then during the listening phase,

$$Z = H_s X + N_Z \tag{21}$$

$$Y = H_r X + N_Y, \tag{22}$$

where the N_Z and N_Y are column vectors of statistically independent complex AWGN with variance $1/2$ per row per dimension, H_s is the the fading matrix between the source and destination nodes and likewise, H_r is the fading matrix between the source and relay nodes. During the collaboration phase,

$$Z = H_c[X^T, U^T]^T + N_Z, \tag{23}$$

where H_c is a channel matrix that contains H_s as a submatrix (see Fig.1.7).

It is further assumed that the source has no knowledge of the H_r and H_c matrices (and hence the H_s matrix too). Similarly, the relay has no knowledge of H_c but is assumed to know H_r. Finally, the destination knows H_c.

Without loss of generality, we will assume that all transmit antennas have unit average power during their respective transmission phases. Likewise, the receive antennas have unit power Gaussian noise. If this is not the case, the respective H matrices may be appropriately scaled row-wise and column-wise.

Under the above unit transmit power per transmit antenna and unit noise power per receive antenna constraint, it is well known that a Multiple Input Multiple Output (MIMO) system with Gaussian codebook and with rate R bits/channel use can reliably communicate over any channel with transfer matrix H such that $R < \log_2 \det(I + HH^\dagger) \stackrel{\triangle}{=} C(H)$ [5][42, **?**], where I denotes the identity matrix and $H^\dagger$ is the conjugate transpose of H.

Intuition for the above problem then suggests the following. During the listening phase, the relay knowing H_r listens for an amount of time n_1 such that $nR < n_1 C(H_r)$. During this time, the relay receives at least nR bits of information and may reliably decode the message. The destination, on the other hand, receives information at the rate of $C(H_s)$ bits/channel use during the listening phase and at the rate of $C(H_c)$ bits/channel use during the collaborative phase. It may reliably decode the message provided that $nR < n_1 C(H_s) + (n - n_1)C(H_c)$. In the limit as $n \to \infty$, the ratio n_1/n approaches a fraction f and one may conjecture that there exists a "good" code of rate R for the set of channels (H_r, H_c) which satisfy

$$R \leq fC(H_s) + (1 - f)C(H_c) \tag{24}$$
$$R \leq fC(H_r), \tag{25}$$

for some $f \in [0, 1]$. Note that if the channel between the source and the relay is particularly poor, one may fall back on the traditional point-to-point communications paradigm and add the following region to that given in (24) – (25)

$$R \leq C(H_s). \tag{26}$$

The above intuition is not a proof of achievability but it does provide an upper bound on the performance of the protocol. The essential difficulty in proving that there exists a code which is "good" for any such pair of channels (H_r, H_c) is two-fold. The problem considered is a *relay channel* which is also a *compound channel*: the authors seek to prove the existence of a code which performs well over an entire set of channels (unknown to the transmitters). The key will be to show the existence of a code that may essentially be refined. Regardless of the actual value of n_1, there exists a codebook for the source which, starting at time $n_1 + 1$, may be layered with the transmission of the relay and perform just as well as if the value of n_1 had been known to the source. For a formal statement and proof of these results, we defer to [38].

The authors simulated the outage probability of their scheme under a quasi-static Rayleigh fading assumption. These numerical and simulation results showed that if the intra-cluster communication has a 10 dB path loss advantage over the receiver at the destination node, in most cases there is essentially no penalty for the intra-cluster communication. Physically, in a two collaborator scenario, this corresponds to a transmit cluster whose radius is 1/3 the distance between the source and destination nodes. By comparison, for a time-division scheme (first the source sends to the relay for a *half* of

[5] Here, $C(H)$ does not, in general, designate the capacity of each link as is witnessed by the fact that only for a special subset of matrices is capacity achieved by placing an equal transmit power on each antenna.

the time rather than the adjustable fraction f allowed by the authors, then the relay and source send to the destination for the remaining half) with a 5 dB geometric penalty, the allowable cluster size is at most 0.178 times the distance between the source and the destination. This work demonstrates the power of this flexible technique with more realistic assumptions on the wireless channel.

7 Interference avoiding cognitive behavior

Up to now the schemes for channels employing cognitive radios have either involved simultaneous transmission, over the same time and frequency, of the primary and secondary users' data (using an interference-mitigating technique), or have not caused any interference at all (collaborative communications). The primary user's message was used as side-information at the secondary transmitter in order to mitigate interference effects. Another way cognitive radios may improve spectral efficiency is by sensing and filling in *spectral gaps.* This can be seen as interference-avoiding cognitive behavior. Suppose the wireless spectrum is populated by some primary users, transmitting on any number of bands. At any point in time, a number of frequency bands will be occupied by primary users, leaving the remainder unoccupied. If a cognitive radio can sense these spectral nulls, it can opportunistically transmit during these times at these frequencies. The work in [28] and [41] addresses issues involved in the opportunistic sensing of and communication over spectral holes. We outline some of these results next.

The authors in [28, 41] are interested in deriving capacity inner and outer bounds for a cognitive transmitter-receiver pair acting as secondary users in a network of primary users. The capacity is limited by the *distributed* and *dynamic* nature [28] of the spectral activity which these cognitive radios wish to exploit. To illustrate these points, consider a cognitive transmitter (T) and receiver (R) pair denoted by the grey circles in Fig.1.8. Each of these is able to sense transmissions within a certain circular radius around themselves, denoted by the dotted circles. Thus, each transmitter and each receiver has a different *local* view of the spectrum utilization. The white circles indicate the primary users (PU), which may or may not be transmitting at a particular point in time. The authors use the term *distributed* to denote the different views of local spectral activity at the cognitive transmitter T and receiver R. In addition to the spectrum availability being location-dependent, it will also vary with time, depending on the data that must be sent at different moments. The authors use the term *dynamic* to indicate the temporal variation of the spectral activity of the primary users.

Communication by the cognitive transmitter-receiver pair takes place as follows. The transmitter senses the channel and detects the presence of primary users. If primary users are detected, the secondary user refrains from transmission. If not, the cognitive user may opportunistically transmit to the

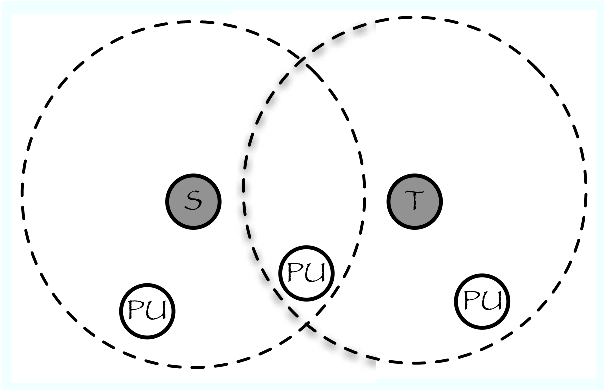

Fig. 1.8. The grey cognitive transmitter (T) receiver (R) pair each have a radius in which they can sense the transmissions of primary users (PU). This leads to different views of local spectral activity, or a *distrbiuted* view on the spectral activity. The PU may change their transmissions over time, leading to *dynamic* spectral activity.

receiver. The cognitive receiver may similarly sense the presence of primary users. If none are present, it may opportunistically receive from the secondary transmitter. If primary users are present, in a simplified model, these will cause interference at the receiver, thus making the reception of a cognitive transmission impossible. Cognitive transmission may thus take place when both the cognitive transmitter and the cognitive receiver sense a spectral hole. The communication opportunities detected at the transmitter T and the receiver R are in general correlated but not identical. The authors in [28] wish to quantify the effect of this *distributed* nature of the spectral environment. To do so they model the channel as a *switched* channel, shown in Fig.1.9. The input X is related to the output Y (all of the cognitive link) as

$$Y = (XS_T + N)S_R, \tag{27}$$

where N is the additive white Gaussian noise, and $S_T, S_R \in \{0, 1\}$ are binary random variables modeled as switches that represent communication opportunities sensed at the transmitter and the receiver respectively. An S_T or S_R value of 0 indicates that communication is not possible at that end of the cognitive link. The authors proceed to model and analyze this switched model using causal and non-causal side information tools [29]. The capacity of the channel depends on whether the transmitter, the receiver, or both, know the states of the switches S_T and S_R. Knowing whether the switch is open at the transmitter allows it to transmit or remain idle. This *side information* allows the secondary link to transmit more efficiently. Intuitively, if the transmitter lacks this side information (on whether the channel is unoccupied or not), power will be lost in failed transmissions, which are caused by collisions with primary user messages. Similarly, power will also be more efficiently used if the transmitter is aware of the receiver's switch state S_R, as it will refrain from transmission if $S_R = 0$. However, the *distributed* nature of the channel will cause a loss in the capacity of such systems, as analyzed in [28]. The effect of the *dynamic*, or temporal variation in the spectral activity is also considered.

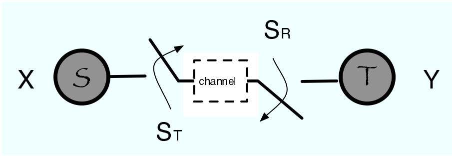

Fig. 1.9. The two switch channel model representing the distributed and dynamic nature of the cognitive channel spectral activity. For successful transmission of the encoded message X to the received message Y in the secondary link $\mathcal{S} \rightarrow \mathcal{T}$, both switches S_T and S_R must be closed, and have a value of 1 in (27).

In [28], the capacity limits of a secondary cognitive radio link is explored in terms of how well the spectral holes at the transmitter and the receiver are *matched*, that is, as a function of the state switches S_T and S_R and how well they are known to the cognitive transmitter and receiver. In the work [41], a similar switching framework is used to analyze the effect of spectral hole *tracking*. That is, once the detection of spectral holes is complete, the secondary cognitive user selects one of the locally free spectral segments for opportunistic transmission. The cognitive receiver must also select one of the locally free spectral segments to monitor in order to detect and decode this cognitive message. For communication to be successful, the transmitter and receiver must select the same spectral hole, which must also be empty (of primary users) at both ends. To coordinate the selection of opportunistic spectral holes, protocols resulting in transmission overhead could be used. The purpose of [41] is to determine the cost and benefits, in terms of capacity, of these overheads to the cognitive user.

Their model is depicted in Fig.1.10 for the case of two spectral channels. Here, the primary user occupancy on the two channels are modeled as binary random processes $S_{PU}^1,\ S_{PU}^2 \in \{0,1\}$. A value of 0 indicates that a primary user is transmitting on the channel indicated by the superscript, while a 1 indicates that channel is free for the secondary user. These processes are modeled as independent identical Markov chains. The cognitive user may be in one of three states, as indicated by $S_T \in \{0,1,2\}$. If $S_T = 0$ then the cognitive transmitter is idle, if is it 1 or 2, it means the cognitive user is transmitting on channel 1 or 2 respectively. The cognitive receiver monitors the channel indicated by $S_R \in \{1,2\}$. When the cognitive transmitter and receiver states are matched, that is, $S_T = S_R$, then the input and output are related through the channel model (in [41] this is a Q-ary symmetric channel), and when they are not matched the cognitive receiver sees random signals. Thus, it is of interest to calculate the channel capacity assuming that the transmitter knows only S_T and the receiver knows only S_R. They can of course exchange this information, but this would cause a loss in capacity. The goal of [41] is to evaluate this loss.

Capacity inner and outer bounds of this *cognitive tracking channel* are determined and simulated. The inner bounds consist of suggesting particular spectral hole selection strategies at the transmitter and receiver, and seeing what fraction of the time these match up (or track each other). Outer bounds are constructed using a *genie* that gives the transmitter and receiver various amounts of side information, which can only improve what can be achieved in reality. For details, we refer to [41, 40].

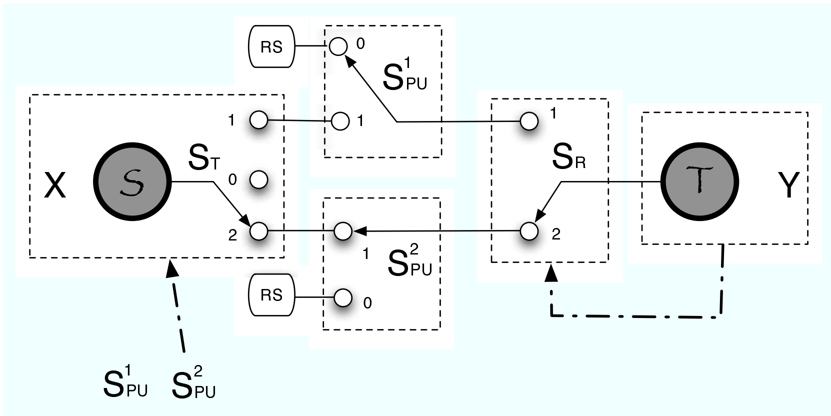

Fig. 1.10. The tracking model of [41]. Secondary transmitter $\mathcal{S}$ wishes to communicate with the secondary receiver $\mathcal{T}$ on one of two channels. The primary user occupancy on the two channels are modeled as binary random processes S^1_{PU}, $S^2_{PU} \in \{0,1\}$. The cognitive user may be in one of three states indicated by S_T, and the cognitive receiver may listen to one of the two channel, as indicated by S_R. For successful communication, S_T must equal S_R (they must be matched).

Conclusion

Due to their ability to adapt to their spectral environment, cognitive radios allow for much more flexible and potentially more spectrally efficient wireless networks. Heterogeneous networks consisting of both cognitive and non-cognitive devices will soon be a reality. In order to exploit the full capabilities of cognitive radios, many questions must be addressed. One of the foremost, from a physical layer, communications perspective, is that of the fundamental limits of the communication possible over a network when using cognitive devices. In order to effectively study this question, researchers have looked at simplified versions of the problem which capture the essence of the communication characteristics particular to such devices. For example, cognitive devices allow for asymmetric side information between transmitting nodes. Information theoretic limits of cognitive channels have been studied in, among others, [13, 12, 14, 27, 31, 47, 41, 28]. In this chapter we summarized some of the most important results in these works. They all had the property that the primary and secondary users had independent information to transmit, and did so by either mitigating the interference using non-causal side information at the cognitive transmitter, or by filling in spectral gaps. Alternatively, when cognitive radios do not have any information of their own to transmit, they can act as relays, a form of asymmetric behavior. As an example, we outlined the work of [38], where some of the idealistic assumptions of relay channels are removed. The benefits and feasibility of cognitive behavior are intimately linked to the topology of the network: poor primary to secondary user wireless links will make the partial asymmetric side information inherent in cognitive behavior to become very costly to obtain. The value of side-information in wireless networks, in terms of diversity, multiplexing, or delays gains, is another fundamental, and not yet fully understood research problem. In summary, research has thus far looked at simplified scenarios in which cognitive radios may be used. Even there many open problems remain. However, the true question that must be answered in order to understand the limits of communication using cognitive radios, is how their capabilities may be harnessed in order to optimize some network communication utility function. We hope that the research outlined in this chapter serves as a first step to this ultimate goal.

References

1. M. Aref, "Information flow in relay networks," Stanford University, Tech. Rep., 1980.
2. G. Atia, M. Sharif, and V. Saligrama, "On optimal outage in relay channels with general fading distributions," in *Proc. of Allerton Conference on Communications, Control and Computing*, Oct. 2006.
3. A. S. Avestimehr and D. N. Tse, "Outage-optimal relaying in the low SNR regime," in *Proc. IEEE Int. Symp. Inf. Theory*, Sept. 2005.
4. K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inf. Theory*, Dec. 2005.
5. A. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 1, pp. 60–70, Jan. 1978.
6. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT-29, pp. 439–441, May 1983.
7. T. Cover, A. E. Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inf. Theory*, vol. IT-26, no. 6, pp. 648–657, Nov. 1980.

8. T. Cover and J. Thomas, *Elements of Information Theory.* New York: John Wiley & Sons, 1991.
9. T. M. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sept. 1979.
10. I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* New York: Academic Press, 1981.
11. C.T.K.Ng and A. Goldsmith, "Capacity gain from transmitter and receiver cooperation," in *Proc. IEEE Int. Symp. Inf. Theory*, Sept. 2005.
12. N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive networks," in *2005 IEEE International Symposium on Information Theory*, Sept. 2005.
13. ——, "Achievable rates in cognitive radio channels," in *39th Annual Conf. on Information Sciences and Systems (CISS)*, Mar. 2005.
14. ——, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.
15. ——, "Cognitive decomposition of wireless networks," in *Proceedings of CROWNCOM*, Mar. 2006.
16. N. Devroye and M. Sharif, "The value of partial side information in interfering channels," 2006, in preparation.
17. FCC. [Online]. Available: http://www.fcc.gov/oet/cognitiveradio/
18. FCC, "Secondary markets initiative." [Online]. Available: http://wireless.fcc.gov/licensing/secondarymarkets/
19. R. G. Gallagher, *Information Theory and Reliable Communication.* New York: Wiley, 1968, ch. 7.
20. A. E. Gamal, M. Mohseni, and S. Zahedi, "On reliable communication over additive white gaussian noise relay channels," *IEEE Trans. Inf. Theory*, 2006.
21. S. Gel'fand and M. Pinsker, "Coding for channels with random parameters," *Probl. Contr. and Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.
22. P. Gupta and P. R. Kumar, "Towards an information theory of large networks: an achievable rate region," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1877–1894, Aug. 2003.
23. T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 1, pp. 49–60, 1981.
24. A. Host-Madsen, "The multiplexing gain of wireless networks," in *Proc. of ISIT*, Sept. 2005.
25. ——, "Capacity bounds for cooperative diversity," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1522–1544, Apr. 2006.
26. I.Maric, R. Yates, and G. Kramer, "The strong interference channel with common information," in *Proc. of Allerton Conference on Communications, Control and Computing*, Sept. 2005.
27. ——, "The strong interference channel with unidirectional cooperation," in *Information Theory and Applications ITA Inaugural Workshop*, Feb. 2006.
28. S. Jafar and S. Srinivasa, "Capacity limits of cognitive radio with distributed dynamic spectral activity," in *Proc. of ICC*, June 2006.
29. S. Jafar, "Capacity with causal and non-causal side information – a unified view," Submitted to *IEEE Trans. Inf. Theory*, Oct. 2005.
30. ——, "Degrees of freedom on the MIMO X channel - optimality of zero forcing and the MMK scheme," Submitted to *IEEE Trans. Inf. Theory*, Sept. 2006.
31. A. Jovicic and P. Viswanath, "Cognitive radio: An information-theoretic perspective," Submitted to *IEEE Trans. Inf. Theory*, 2006.

32. M. Katz and S. Shamai, "Communicating to co-located ad-hoc receiving nodes in a fading environment," in *Proc. IEEE Int. Symp. Inf. Theory*, Chicago, IL, July 2004, p. 115.
33. G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, p. 30373063, Sept. 2005.
34. J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, 2004.
35. J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.
36. M. Maddah-Ali, A. Motahari, and A.Khandani, "Combination of multi-access and broadcast schemes," in *Proc. IEEE Int. Symp. Inf. Theory*, Seattle, WA, July 2006, pp. 2104–2108.
37. J. Mitola, "Cognitive radio," Ph.D. dissertation, Royal Institute of Technology (KTH), 2000.
38. P. Mitran, H. Ochiai, and V. Tarokh, "Space-time diversity enhancements using collaborative communication," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2041–2057, June 2005.
39. A. Reznik, S. Kulkarni, and S. Verdú, "Capacity and optimal resource allocation in the degraded Gaussian relay channel with multiple relays," in *Proc. Allerton Conf. Commun., Control and Comp.*, Monticello, IL, Oct. 2002.
40. S. Srinivasa and S. Jafar, "On the capacity of the cognitive tracking channel," *Journal paper in preparation*, 2006.
41. S. Srinivasa, S. Jafar, and N. Jindal, "On the capacity of the cognitive tracking channel," in *Proc. of ISIT*, July 2006.
42. I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecomm.*, vol. 10, no. 6, pp. 585–595, 1999.
43. E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120–154, 1971.
44. H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sept. 2006.
45. F. Willems and E. van der Meulen, "The discrete memoryless multiple-access channel with cribbing encoders," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 3, pp. 313–327, Nov. 1985.
46. J. Wolfowitz, *Coding Theorems of Information Theory.* New York: Springer-Verlag, 1978.
47. W. Wu, S. Vishwanath, and A. Arapostathis, "On the capacity of the interference channel with degraded message sets," Submitted to *IEEE Trans. Inf. Theory*, June 2006.
48. L.-L. Xie and P. R. Kumar, "An achievable rate for the multiple level relay channel," Submitted to *IEEE Trans. Inf. Theory*, Nov. 2003.
49. ——, "A network information theory for wireless communication: scaling laws and optimal operation," *IEEE Trans. Inf. Theory*, vol. 50, pp. 748–767, May 2004.