Achievable Rates in Cognitive Radio Channels

Natasha Devroye, Patrick Mitran, and Vahid Tarokh

Abstract

Cognitive radio promises a low cost, highly flexible alternative to the classic single frequency band, single protocol wireless device. By sensing and adapting to its environment, such a device is able to fill voids in the wireless spectrum and dramatically increase spectral efficiency. In this paper, the *cognitive radio channel* is defined as an n-transmitter, m-receiver interference channel in which sender i obtains the messages senders 1 through $i-1$ plan to transmit. The two sender, two receiver case is considered. In this scenario, one user, a cognitive radio, obtains (genie assisted, or causally) knowledge of the data to be transmitted by the other user. The cognitive radio may then simultaneously transmit over the same channel, as opposed to waiting for an idle channel as in a traditional cognitive radio channel protocol. Dirty-paper coding and ideas from achievable region constructions for the interference channel are used, and an achievable region for the *cognitive radio channel* is computed. It is shown that in the Gaussian case, the described achievable region approaches the upper bounds provided by the 2×2 Gaussian MIMO broadcast channel, and an interference-free channel. Results are extended to the case in which the message is causally obtained.

Index Terms

Cognitive radio channel, interference channel, dirty-paper coding, wireless communication.

N. Devroye, P. Mitran, and V. Tarokh are with the Division of Engineering and Applied Sciences, Harvard University. (e-mail: {ndevroye, mitran, vahid}@deas.harvard.edu)

This material is based upon research supported by the National Science Foundation under the Alan T. Waterman Award, Grant No. CCR-0139398.

I. MOTIVATION

Recently, there has been an explosion of interest in cognitive and software radios, as is evidenced by FCC proceedings [7], [8], talks [22], and papers [17], [21]. *Software Defined Radios* (SDR) [16] are devices used to communicate over the wireless medium equipped with either a general purpose processor or programmable silicon as hardware base, and enhanced by a flexible software architecture. They are low-cost, can be rapidly upgraded, and may adapt to the environment in real-time. Such devices are able to operate in many frequency bands under multiple transmission protocols and employ a variety of modulation and coding schemes. Taking this one step further, Mitola [17] coined the term *cognitive radio* for software defined radios capable of sensing their environment and making decisions instantaneously, without any user intervention. This allows them to change their modulation schemes or protocols so as to better communicate with the sensed environment.

Apart from their low cost and flexibility, another benefit of SDR technology is spectral efficiency. Currently, FCC measurements [9], indicate that at any time roughly 10% of the unlicensed frequency spectrum is actively in use (leaving 90% unused). If a wireless device such as a cognitive radio is able to sense an idle channel at a particular frequency band (or time), then it can shift to that frequency band (or time slot) to transmit its own information, dramatically increasing spectral (or temporal) efficiency.

In current cognitive radio protocol proposals, the device listens to the wireless channel and determines, either in time or frequency, which part of the spectrum is unused [13]. It then adapts its signal to fill this void in the spectrum space. Thus, a device transmits over a certain time or frequency only when no other user does. In this paper, the cognitive radio behavior is generalized to allow two users to simultaneously transmit over the same time or frequency. Under our scheme, a cognitive radio will listen to the channel and, if sensed idle, proceed with the traditional cognitive radio channel model, that is, transmit during the voids. 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.

Although cognitive radios have spurred great interest and excitement in industry, many of the fundamental theoretical questions on the limits of such technology remain unanswered. In this paper, we propose a general model of a cognitive radio channel and study its theoretic limits. Specifically, we will prove achievability, in the information theoretic sense, of a certain set of rates at which two senders (cognitive radios, denoted as $\mathcal{S}_1$ and $\mathcal{S}_2$) can transmit simultaneously over a common channel to two independent receivers $\mathcal{R}_1,\ \mathcal{R}_2$ when $\mathcal{S}_2$ is aware of the message to be sent by $\mathcal{S}_1$. Our methods borrow ideas from Costa's dirty paper coding [3], the interference channel [2], the Gaussian MIMO broadcast channel [28], and the achievable region of the interference channel described by Han and Kobayashi [12]. The results are also related, conceptually, to other communication systems in which user cooperation is employed in order to enhance the capacity. These schemes can be traced back to telegraphy, and have recently been considered in the *collaborative communications* of [23], the spatial diversity enhancing schemes obtained through user cooperation described in [26], [27], and many others such as [14], [15], [19], [20].

The key idea behind achieving high data rates in an environment where two senders share a common channel is interference cancellation or mitigation. *Dirty paper coding* is the term first used by Costa [3] to describe a technique which completely mitigates *a-priori* known interference over an additive white Gaussian noise channel. In the case that both the noise and interference are Gaussian, he demonstrated that there is no loss in capacity regardless of the interference power. In order to prove this, Costa started with the well-known formula obtained by Gel'fand and Pinsker [11] for the capacity of a channel with non-causal knowledge of the interference at the transmitter only, given by

$$C = \max_{p(u,x|s)} \left[I(U;Y) - I(U;S) \right],$$

where X is the input to the channel, Y is the output, S is the interference, and U is an auxiliary random variable chosen to make the channel $U \to Y$ appear causal. The channel model and variables are shown in Fig. 1 for additive interference and noise. In the Gaussian noise and interference case, Costa achieves the capacity of an interference-free channel by assuming the input X to the channel is Gaussian, and then considering an auxiliary variable U of the form $U = X + \alpha S$ for some parameter α whose optimal value is equal to the ratio of the signal power to the signal plus noise power. Since the rate thus obtained is equal to the capacity of an interference-free channel, which provides an upper bound, optimality is achieved by the assumed Gaussian input X. We will make use of the coding techniques of Costa [3], Gel'fand and Pinsker [11], as well as Cover and Chiang [4] in our main results in Sections II and IV.

Our methods are also closely related to the interference channel, which is briefly described next. Consider a discrete memoryless *interference channel* [2], with random variables $X_1 \in \mathcal{X}_1,\ X_2 \in \mathcal{X}_2$ as inputs to the channel characterized by the conditional probabilities $p(y_1|x_1,x_2),\ p(y_2|x_1,x_2)$ with resulting channel output random variables $Y_1 \in \mathcal{Y}_1, Y_2 \in \mathcal{Y}_2$. The interference channel

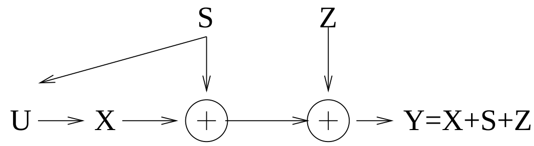


Fig. 1. Dirty paper coding channel with input X, auxiliary random variable U, interference S known non-causally to the transmitter, additive noise Z and output Y.

corresponds to two independent senders $\mathcal{S}_1, \mathcal{S}_2$, with independent non-cooperating receivers $\mathcal{R}_1, \mathcal{R}_2$, transmitting over the same channel, and thus interfering with each other.

The additive interference channel is shown in Fig. 2 below. There, in addition to the additive interference from the other sender, each output is affected by independent additive noise Z_1, Z_2. The parameters a_{12}, a_{21} capture the effects of the interference. The channel outputs are:

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 1 & a_{21} \\ a_{12} & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} + \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \tag{1}$$

The interference channel capacity, in the most general case, is still an open problem. In the

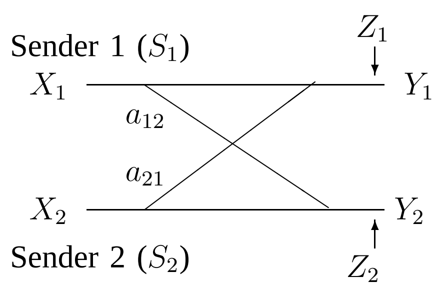


Fig. 2. The interference channel with inputs X_1, X_2, outputs Y_1, Y_2, additive noise Z_1, Z_2 and interference coefficients a_{12}, a_{21}.

case of strong interference, as defined in [25], and very-strong interference, as defined in [2], the capacity is known. Achievable regions of the interference channel have been calculated in [12], and recently in [24]. We will make use of techniques as in [12] to provide an achievable region for the cognitive radio channel, as defined next.

The main idea of this paper is to define and prove achievability of a region of rate pairs for a *cognitive radio channel*. Define a *cognitive radio channel* to be an interference channel in which $\mathcal{S}_2$ has knowledge of the message to be transmitted by $\mathcal{S}_1$. This is either obtained causally, or could possibly be given to the sender non-causally by a "genie". The main theorems (1,2 and 3) will be proved for the non-causal case, or the *genie-aided cognitive radio channel.* In the proof of achievability, $\mathcal{S}_2$ treats the message of $\mathcal{S}_1$ as interference and tries to compensate for it using a dirty-paper coding technique. This results in an achievable region for the rate pair that enlarges the region in [12], and reduces to that region in the case where no interference mitigation is performed. Simulations in the Gaussian noise case show the rate region described in this paper comes close to both the upper bound provided by a 2×2 Multiple Input Multiple Output (MIMO) Gaussian broadcast channel, and to another upper bound provided by an interference-free channel. Simulations also suggest that the larger the power mismatch between the two senders, the better this scheme performs. This is relevant in rich scattering environments, where fading commonly causes power mismatches between the two senders and the signals at the receivers. The variations in path loss and shadowing effects may further the mismatches and thus aid our scheme.

The paper is structured as follows: Section II defines the *genie-aided cognitive radio channel* as an interference channel in which one sender is non-causally given the other sender's message. Section II also proves the main result: achievability of a certain rate region. The technique employed merges the results of Costa [3] on dirty paper coding and the achievable region for the interference channel described by Han and Kobayashi [12]. The significance of our result is shown in Section III, where numerical methods are used to compute an achievable region in the additive white Gaussian noise case. Here, it is clear that our region not only extends that of [12], but that in the case of large power mismatches between the two senders, as would be expected in a rich fading environment, the achievable region described here approaches the upper bounds given by the 2×2 Gaussian MIMO broadcast channel [28], and an interference-free channel. Section IV extends the genie-aided cognitive radio channel model of Section II to a more realistic scenario in which all signals are obtained causally. In Section V, we summarize the main contributions of this paper: the definition of a *cognitive radio channel*, and the proof and significance of a certain achievable rate region for this channel.

II. GENIE-AIDED COGNITIVE RADIO CHANNEL DEFINITION

Define a *genie-aided cognitive radio channel* C_{COG} to be an interference channel in which $\mathcal{S}_2$ is given, in a non-causal manner (i.e., by a genie), the message x_1^n which $\mathcal{S}_1$ will transmit, as illustrated in Fig. 3 below. This non-causal constraint will be relaxed in Section IV, and a *cognitive radio channel* describes the case where the message is causally obtained. $\mathcal{S}_2$ can then exploit the knowledge of $\mathcal{S}_1$'s message, and potentially improve the transmission rate. It can do so using a dirty paper coding technique. In the following, an achievable rate region for such a cognitive radio channel is constructed in a way which combines the results of Gel'fand and Pinsker [11] on the capacity of channels with known interference, Costa's [3] dirty paper coding, and the largest known achievable region of the interference channel as described by Han and Kobayashi [12], in which senders are completely independent. Intuitively, the achievable region in [12] should lie entirely within our achievable region, since our senders are permitted to at least partially cooperate. They could choose not to cooperate at all and in that case reduce to the scenario in [12]. It is unclear whether this achievable region lies within the current outer bounds of the interference channel [18], or not. However, an upper bound for our region in the Gaussian case is provided by the 2×2 MIMO broadcast channel whose capacity, in the Gaussian case, has recently been calculated in [28]. In [28], dirty paper coding techniques are shown to be optimal for non-degraded vector broadcast channels. Our channel model resembles that of [28], with one important difference. In the scheme of [28] it is presumed that both senders can cooperate in order to precode the transmitted signal. In our scheme, the relation between the two senders is asymmetric. We believe this is a reasonable model for the target application of a cognitive radio channel in which one sender is transmitting and a second sender obtains the first sender's message before transmitting its own message. The rate of $\mathcal{S}_2$ is also bounded by the rate achievable in an interference-free channel, with $a_{12} = 0$. For some rate pairs, this interference-free channel provides a tighter bound than the 2×2 MIMO broadcast channel, and vice versa.

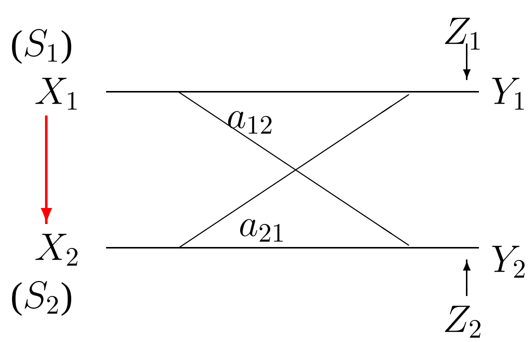


Fig. 3. The genie-aided cognitive radio channel with inputs X_1, X_2, outputs Y_1, Y_2, additive noise Z_1, Z_2 and interference coefficients a_{12}, a_{21}. S_1's input X_1 is given to S_2, but not the other way around.

An (n, K_1, K_2, λ) *code* for the *genie-aided cognitive radio channel* consists of K_1 codewords $x_1^n(i) \in \mathcal{X}_1^n$ for $\mathcal{S}_1$ and $K_1 \cdot K_2$ codewords $x_2^n(i,j) \in \mathcal{X}_2^n$ for $\mathcal{S}_2$ which together form the *codebook*, revealed to both senders and receivers such that the average error probabilities under some decoding scheme are less than λ.

A rate pair (R_1, R_2) is said to be *achievable* for the genie-aided cognitive radio channel if there exists a sequence of $(n, 2^{nR_1}, 2^{nR_2}, \epsilon_n)$ codes such that $\epsilon_n \to 0$ as $n \to \infty$. An *achievable region* is a closed subset of the positive quadrant of $\mathbb{R}^2$ of achievable rate pairs.

The interference channel capacity, in the most general case, is still an open problem. This is the case for the *genie-aided cognitive radio channel* as well. In [12], an achievable region of the interference channel is found by first considering a modified interference channel and then establishing a correspondence between the achievable rates of the modified and the original channel models. A similar modification is made in the next subsection.

A. The Modified Genie-aided Cognitive Channel C^m_{COG}

As in [12], we introduce a modified genie-aided cognitive radio channel, C^m_{COG}, (m for modified) and demonstrate an achievable region for C^m_{COG}. Then, a relation between an achievable rate for C^m_{COG} and an achievable rate for C_{COG} is used to establish an achievable region for C_{COG}. Define the modified genie-aided cognitive radio channel C^m_{COG} as in Fig. 4 below.

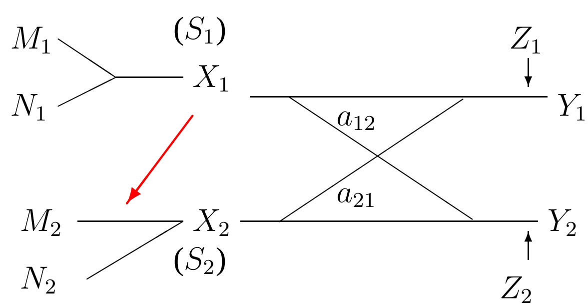


Fig. 4. The modified cognitive radio channel with auxiliary random variables M_1, M_2, N_1, N_2, inputs X_1, X_2, additive noise Z_1, Z_2, outputs Y_1, Y_2 and interference coefficients a_{12}, a_{21}.

Let $X_1 \in \mathcal{X}_1$ and $X_2 \in \mathcal{X}_2$ be the random variable inputs to the channel. Let $Y_1 \in \mathcal{Y}_1$ and $Y_2 \in \mathcal{Y}_2$ be the random variable outputs of the channel. The conditional probabilities of the discrete memoryless C^m_{COG} are the same as those of the discrete memoryless C_{COG} and are fully described by $p(y_1|x_1, x_2)$ and $p(y_2|x_1, x_2)$ for all values $x_1 \in \mathcal{X}_1,\ x_2 \in \mathcal{X}_2,\ y_1 \in \mathcal{Y}_1$ and $y_2 \in \mathcal{Y}_2$.

The modified genie-aided cognitive radio channel introduces two pairs of auxiliary random variables: (M_1, N_1) and (M_2, N_2). The random variables $M_1 \in \mathcal{M}_1$ and $M_2 \in \mathcal{M}_2$ represent, as in [12], the private information to be sent from $\mathcal{S}_1$ and $\mathcal{S}_2$ to $\mathcal{R}_1$ and $\mathcal{R}_2$ respectively. In contrast, the random variables $N_1 \in \mathcal{N}_1$ and $N_2 \in \mathcal{N}_2$ represent the public information to be sent from $\mathcal{S}_1$ and $\mathcal{S}_2$ to both $\mathcal{R}_1$ and $\mathcal{R}_2$. The function of these M_1, N_1, M_2, N_2 is as in [12]: to decompose or define *explicitly* the information to be transmitted between various input and output pairs.

In this work, M_2 and N_2 also serve a dual purpose: these auxiliary random variables are analogous to the auxiliary random variables of Gel'fand and Pinsker [11] or Cover and Chiang [4]. They serve as fictitious inputs to the channel, so that after $\mathcal{S}_2$ is informed of the message of $\mathcal{S}_1$ non-causally (or equivalently, is given x_1^n), the channel still looks or behaves like a discrete memoryless channel (DMC) from (M_1, N_1, M_2, N_2) to (Y_1, Y_2). As in [4], [11], there is a penalty in using this approach which will be reflected by a reduction in achievable rates (compared to the ficticious DMC from (M_1, N_1, M_2, N_2) to (Y_1, Y_2)) for the links which use the non-causal information.

Similar to the definition of a *code* in the cognitive radio channel case, define an $(n, K_{11}, K_{12}, K_{21}, K_{22}, \lambda)$ code for the modified genie-aided cognitive radio channel as a set of $K_{11} \cdot K_{12}$ codewords $x_1^n(i,j) \in \mathcal{X}_1^n$ for $\mathcal{S}_1$ and $K_{11} \cdot K_{12} \cdot K_{21} \cdot K_{22}$ codewords $x_2^n(i,j,k,l) \in \mathcal{X}_2^n$ such that the average probability of decoding error is less than λ. Call a quadruple $(R_{11}, R_{12}, R_{21}, R_{22})$ achievable if there exists a sequence of $(n, 2^{nR_{11}}, 2^{nR_{12}}, 2^{nR_{21}}, 2^{nR_{22}}, \epsilon_n)$ codes such that $\epsilon_n \to 0$ as $n \to \infty$. An achievable region of a modified genie-aided cognitive radio channel is the closure of a subset of the positive region of $\mathbb{R}^4$ of achievable rate quadruples.

As mentioned in [12], the introduction of a time-sharing random variable W is thought to strictly extend the achievable region obtained using a convex hull operation. Thus, let $W \in \mathcal{W}$ be a time-sharing random variable whose n-sequences $w^n \triangleq (w^{(1)}, w^{(2)}, \ldots, w^{(n)})$ are generated independently of the messages, according to $\prod_{t=1}^n p(w^{(t)})$. The n-sequence w^n is given to both senders and both receivers. The paper's main theorems (1,2 and 3) are proved next.

Theorem 1: Let $Z \triangleq (Y_1, Y_2, X_1, X_2, M_1, N_1, M_2, N_2, W)$, and let $\mathcal{P}$ be the set of distributions on Z that can be decomposed into the form

$$p(w)p(m_1|w)p(n_1|w)p(x_1|m_1, n_1, w)p(m_2|x_1, w)p(n_2|x_1, w) \times p(x_2|m_2, n_2, w)p(y_1|x_1, x_2)p(y_2|x_1, x_2). \quad (2)$$

For any $Z \in \mathcal{P}$, let $S(Z)$ be the set of all quadruples $(R_{11}, R_{12}, R_{21}, R_{22})$ of non-negative real numbers such that there exist non-negative real (L_{21}, L_{22}) satisfying:

$$R_{21} \leq L_{21} - I(N_2; X_1|W) \quad (3)$$

$$R_{22} \leq L_{22} - I(M_2; X_1|W) \tag{4}$$

$$R_{11} \leq I(Y_1, N_1, N_2; M_1|W) \tag{5}$$

$$R_{12} \leq I(Y_1, M_1, N_2; N_1|W) \tag{6}$$

$$L_{21} \leq I(Y_1, M_1, N_1; N_2|W) \tag{7}$$

$$R_{11} + R_{12} \leq I(Y_1, N_2; M_1, N_1|W) \tag{8}$$

$$R_{11} + L_{21} \leq I(Y_1, N_1; M_1, N_2|W) \tag{9}$$

$$R_{12} + L_{21} \leq I(Y_1, M_1; N_1, N_2|W) \tag{10}$$

$$R_{11} + R_{12} + L_{21} \leq I(Y_1; M_1, N_1, N_2|W) \tag{11}$$

$$L_{22} \leq I(Y_2, N_1, N_2; M_2|W) \tag{12}$$

$$R_{12} \leq I(Y_2, N_2, M_2; N_1|W) \tag{13}$$

$$L_{21} \leq I(Y_2, N_1, M_2; N_2|W) \tag{14}$$

$$L_{22} + L_{21} \leq I(Y_2, N_1; M_2, N_2|W) \tag{15}$$

$$L_{22} + R_{12} \leq I(Y_2, N_2; M_2, N_1|W) \tag{16}$$

$$R_{12} + L_{21} \leq I(Y_2, M_2; N_1, N_2|W) \tag{17}$$

$$L_{22} + R_{21} + L_{12} \leq I(Y_2; M_2, N_1, N_2|W). \tag{18}$$

Let S be the closure of $\cup_{Z \in \mathcal{P}} S(Z)$. Then any element of S is achievable for the modified genie-aided cognitive radio channel C^m_{COG}. $\square$

Proof: It is sufficient to show the achievability of the interior elements of $S(Z)$ for each $Z \in \mathcal{P}$. So, fix $Z = (Y_1, Y_2, X_1, X_2, M_1, N_1, M_2, N_2, W)$ and take any $(R_{11}, R_{12}, R_{21}, R_{22})$ and (L_{21}, L_{22}) satisfying the constraints of the theorem. The standard notation and notions of strong ϵ-typicality, strong joint typicality, and strongly typical sets of [5] will be used.

Codebook generation: Let some distribution on Z of the form (2) be given. For any $\epsilon > 0$ it is sufficient to prove that there exists a large enough block length n to ensure that the probability of error is less than ϵ. To generate the codebook, first let $w^n \stackrel{\triangle}{=} (w^{(1)}, w^{(2)}, \ldots, w^{(n)})$ be a sequence in $\mathcal{W}^n$ chosen randomly according to $\prod_{t=1}^n p(w^{(t)})$ and known to $\mathcal{S}_1, \mathcal{S}_2, \mathcal{R}_1$ and $\mathcal{R}_2$. Next, note that $p(m_2|w) = \sum_{x_1 \in \mathcal{X}_1} p(m_2|x_1, w)p(x_1)$, and $p(n_2|w) = \sum_{x_1 \in \mathcal{X}_1} p(n_2|x_1, w)p(x_1)$. We will generate the codebook according to the distribution

$$\begin{aligned} & p(w)p(m_1|w)p(n_1|w)p(x_1|m_1, n_1, w) \\ & \quad p(m_2|w)p(n_2|w)\ p(x_2|m_2, n_2, w)p(y_1|x_1, x_2)p(y_2|x_1, x_2). \end{aligned} \tag{19}$$

Then,

1. Generate $2^{n(R_{11}-6\epsilon)}$ n-sequences $m_1(i)$ i.i.d. according to $\prod_{t=1}^{n} p(m_1^{(t)}|w^{(t)})$
2. Generate $2^{n(R_{12}-6\epsilon)}$ n-sequences $n_1(j)$ i.i.d. according to $\prod_{t=1}^{n} p(n_1^{(t)}|w^{(t)})$
3. Generate $2^{n(L_{21}-6\epsilon)}$ n-sequences $n_2(l)$ i.i.d. according to $\prod_{t=1}^{n} p(n_2^{(t)}|w^{(t)})$

   $\rightarrow$ throw into $2^{n(R_{21}-6\epsilon)}$ bins uniformly
4. Generate $2^{n(L_{22}-6\epsilon)}$ n-sequences $m_2(k)$ i.i.d. according to $\prod_{t=1}^{n} p(m_2^{(t)}|w^{(t)})$

   $\rightarrow$ throw into $2^{n(R_{22}-6\epsilon)}$ bins uniformly

Define the message index spaces $\mathcal{S}_{11} \triangleq \{1, 2, \dots, 2^{n(R_{11}-6\epsilon)}\}$, $\mathcal{S}_{12} \triangleq \{1, 2, \dots, 2^{n(R_{12}-6\epsilon)}\}$, $\mathcal{S}_{21} \triangleq \{1, 2, \dots, 2^{n(R_{21}-6\epsilon)}\}$ and $\mathcal{S}_{22} \triangleq \{1, 2, \dots, 2^{n(R_{22}-6\epsilon)}\}$. The aim is to send a four dimensional message $s \triangleq (s_{11}, s_{12}, s_{21}, s_{22}) \in \mathcal{S} \triangleq \mathcal{S}_{11} \times \mathcal{S}_{12} \times \mathcal{S}_{21} \times \mathcal{S}_{22}$ whose first two components are message indices, and last two components are bin indices. Note that if such a message can be sent with arbitrarily small probability of error, then the rates achieved will be $(R_{11}, R_{12}, R_{21}, R_{22})$ for the respective sender, receiver pairs.

Recall that the messages actually sent over the genie-aided cognitive radio channel are elements of $\mathcal{X}_1^n$, $\mathcal{X}_2^n$. The message indices are mapped into the signal space as follows:

1) To send s_{11} and s_{12} look up the sequences $m_1^n(s_{11})$ and $n_1^n(s_{12})$.

2) Send $x_1^n = f^n(m_1^n(s_{11}), n_1^n(s_{12})|w^n)$.

It is assumed that x_1^n is a deterministic function of m_1^n and n_1^n defined as $f^n(m_1^n, n_1^n|w^n) \triangleq (f(m_1^{(1)}, n_1^{(1)}|w^{(1)}), \cdots, f(m_1^{(n)}, n_1^{(n)}|w^{(n)}))$, for some function $f(\cdot, \cdot|w)$ for each $w \in \mathcal{W}$. In order for $\mathcal{S}_2$ to send the two messages s_{21} and s_{22} (recall that these are bin indices), its encoder is given the message x_1^n (equivalently $m_1^n(s_{11})$ and $n_1^n(s_{12})$) to be transmitted by $\mathcal{S}_1$. It will use this to generate x_2^n from s_{21} and s_{22} as follows:

1) To send s_{21} and s_{22} look in bin s_{21} and s_{22} for sequences n_2^n and m_2^n such that $(n_2^n, m_1^n, n_1^n, x_1^n)$ and $(m_2^n, m_1^n, n_1^n, x_1^n)$ are *jointly typical* quadruples respectively according to the joint distribution in (2).

2) Generate x_2^n i.i.d. according to $\prod_{t=1}^{n} p(x_2^{(t)}|m_2^{(t)}, n_2^{(t)}, w^{(t)}),$ and send this x_2^n.

Notice that for $\mathcal{S}_2$, x_2^n is not necessarily a deterministic function of m_2^n, n_2^n, x_1^n and w^n. At $\mathcal{S}_1$, the sent signal x_1^n was a deterministic function of the messages m_1^n and n_1^n so that $\mathcal{S}_2$, by decoding m_1^n and n_1^n knows the exact n-sequence that was sent, or the exact interference it will encounter. If $\mathcal{S}_2$ is given the x_1^n explicitly, then x_1^n need not be a deterministic function of m_1^n and n_1^n. Also note that $(M_1, N_1) \to X_1 \to (M_2, N_2, X_2)$ forms a Markov chain. This will be made full use of at the decoding stage.

Decoding: $\mathcal{R}_1$ and $\mathcal{R}_2$ decode independently, based on strong joint typicality. The inputs x_1^n, x_2^n to the genie-aided cognitive radio channel are received at $\mathcal{R}_1$, $\mathcal{R}_2$ as y_1^n, y_2^n according to the conditional distributions $p^n(y_1^n|x_1^n, x_2^n) = \prod_{t=1}^n p(y_1^{(t)}|x_1^{(t)}, x_2^{(t)})$ and $p^n(y_2^n|x_1^n, x_2^n) = \prod_{t=1}^n p(y_2^{(t)}|x_1^{(t)}, x_2^{(t)})$. $\mathcal{R}_1$ aims to recover (s_{11}, s_{12}, s_{21}) ($\mathcal{R}_2$ aims to recover (s_{12}, s_{21}, s_{22})) based on y_1^n (y_2^n resp.) and w^n. Thus, the decoders at $\mathcal{R}_1$ ($\mathcal{R}_2$ resp.) are functions

$$\psi_1 : \mathcal{Y}_1^n \times \mathcal{W}^n \to \mathcal{S}_{11} \times \mathcal{S}_{12} \times \mathcal{S}_{21}, \quad \psi_1(y_1^n, w^n) = (\psi_1^{11}(y_1^n, w^n), \psi_1^{12}(y_1^n, w^n), \psi_1^{21}(y_1^n, w^n))$$

$$\psi_2 : \mathcal{Y}_2^n \times \mathcal{W}^n \to \mathcal{S}_{12} \times \mathcal{S}_{21} \times \mathcal{S}_{22}, \quad \psi_2(y_2^n, w^n) = (\psi_2^{12}(y_2^n, w^n), \psi_2^{21}(y_2^n, w^n), \psi_2^{22}(y_2^n, w^n)).$$

When $\mathcal{R}_1$ ($\mathcal{R}_2$ resp.) receives the n-sequence y_1^n (y_2^n resp.) and w^n, it looks at the set of all input sequences m_1^n, n_1^n, n_2^n (m_2^n, n_2^n, n_1^n resp.) that are *jointly typical*, according to the distribution (19) with the received y_1^n (y_2^n resp.) and w^n. Thus, $\mathcal{R}_1$ ($\mathcal{R}_2$) forms the set, for the given $w^n \in \mathcal{W}^n$

$$S_1(y_1^n, w^n) \stackrel{\triangle}{=} \{(m_1^n, n_1^n, n_2^n) : (y_1^n, m_1^n, n_1^n, n_2^n, w^n) \in A_\epsilon^n(Y_1, M_1, N_1, N_2|W)\}$$

$$S_2(y_2^n, w^n) \stackrel{\triangle}{=} \{(m_2^n, n_1^n, n_2^n) : (y_2^n, m_2^n, n_1^n, n_2^n, w^n) \in A_\epsilon^n(Y_2, M_2, N_1, N_2|W)\}.$$

Since $\mathcal{R}_1$, $\mathcal{R}_2$ will be decoding message and bin indices, let $B(m_1^n)$ and $B(m_2^n)$ be the message indices of the n-sequences $m_1^n \in \mathcal{S}_{11}$, $n_1^n \in \mathcal{S}_{12}$ respectively, while $B(m_2^n)$ and $B(n_2^n)$ are bin indices of the n-sequences $m_2^n \in \mathcal{S}_{22}$, $n_2^n \in \mathcal{S}_{21}$ respectively. Then the decoding function $\psi_1(\cdot, \cdot)$ is as follows:

If all $(m_1^n, \cdot, \cdot) \in S_1(y_1^n, w^n)$ have the same message index, then we let $\psi_1^{11}(y_1^n, w^n) = B(m_1^n)$.

If all $(\cdot, n_1^n, \cdot) \in S_1(y_1^n, w^n)$ have the same message index, then we let $\psi_1^{12}(y_1^n, w^n) = B(n_1^n)$.

If all $(\cdot,\cdot,n_2^n) \in S_1(y_1^n, w^n)$ have the same bin index, then we let $\psi_1^{21}(y_1^n, w^n) = B(n_2^n)$.

Otherwise, an error is declared. $\psi_2(\cdot)$ if defined analogously.

We defer the probability of error analysis to the Appendix. The analysis shows that if $(R_{11}, R_{12}, R_{21}, R_{22})$ and (L_{21}, L_{22}) are as in the statement of the theorem, then reliable communication is possible. ■

Direct application of Lemma 2.1 in [12] to the modified genie-aided cognitive radio channel demonstrates that if the rate quadruple $(R_{11}, R_{12}, R_{21}, R_{22})$ is achievable for the modified genie-aided cognitive radio channel, then the rate pair $(R_{11} + R_{12}, R_{21} + R_{22})$ is achievable for the genie-aided cognitive radio channel.

Another important rate pair for the genie-aided cognitive radio channel is achievable: that in which $\mathcal{S}_2$ transmits no information of its own to $\mathcal{R}_2$, and simply aids $\mathcal{S}_1$ in sending its message to $\mathcal{R}_1$. When this is the case, the rate pair $(R_1^*, 0)$ is achievable, where R_1^* is the capacity of the vector rate channel $(\mathcal{S}_1, \mathcal{S}_2) \to \mathcal{R}_1$.

Theorem 2: Consider the vector channel from $\mathcal{S}_1$ (input X_1), $\mathcal{S}_2$ (input X_2) $\to \mathcal{R}_1$ (output Y_1) described by the conditional probability density $p(y_1|x_1, x_2)$ for all $y_1 \in \mathcal{Y}_1$, $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$, and define

$$R_1^* \triangleq \max_{p(x_1,x_2)} I(X_1, X_2; Y_1). \tag{20}$$

Then the rate pair $(R_1^*, 0)$ is achievable. □

Note however, that the analogous rate pair $(0, R_1^*)$ is not achievable, since that would involve $\mathcal{S}_1$ serving aiding $\mathcal{S}_2$ in sending its message, which cannot happen under our assumptions; $\mathcal{S}_2$ knows $\mathcal{S}_1$'s message, but not vice versa.

Theorem 3: The convex hull of the points of Theorem 1 and Theorem 2 is achievable. □

Proof: Follows by standard time-sharing techniques and the fact that the achievable region is the closure of achievable rates. ■

Next, an achievable region is demonstrated in the Gaussian case.

III. THE GAUSSIAN COGNITIVE RADIO CHANNEL

Consider the genie-aided cognitive radio channel, depicted in Fig. 5 with independent additive noise $Z_1 \sim \mathcal{N}(0, Q_1)$ and $Z_2 \sim \mathcal{N}(0, Q_2)$. In order to determine an achievable region for the modified Gaussian genie-aided cognitive radio channel, specific forms of the random variables described in Theorem 1 are assumed. As in [3], [10], [12], Theorem 1 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). Let W, the time-sharing random variable be constant. Consider the case where, for certain $\alpha, \beta \in \mathbb{R}$ and $\lambda, \overline{\lambda}, \gamma, \overline{\gamma} \in [0,1]$, with

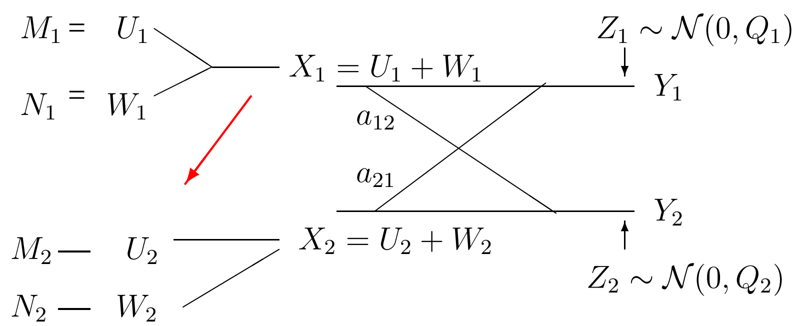


Fig. 5. The modified Gaussian genie-aided cognitive radio channel with inputs X_1, X_2, auxiliary random variables $U_1, W_1, U_2, W_2, M_1, N_1, M_2, N_2$, outputs Y_1, Y_2, additive Gaussian noise Z_1, Z_2 and interference coefficients a_{12}, a_{21}.

$\lambda + \overline{\lambda} = 1,\ \gamma + \overline{\gamma} = 1$, and additional independent auxiliary random variables U_1, W_1, U_2, W_2 as in Fig. 5, the following hold:

$U_1 = M_1$ distributed according to $\mathcal{N}(0, \lambda P_1)$

$W_1 = N_1$ distributed according to $\mathcal{N}(0, \overline{\lambda} P_1)$

$X_1 = U_1 + W_1 = M_1 + N_1$ distributed according to $\mathcal{N}(0, P_1)$

$M_2 = U_2 + \alpha X_1$ where U_2 is distributed according to $\mathcal{N}(0, \gamma P_2)$

$N_2 = W_2 + \beta X_1$ where W_2 is distributed according to $\mathcal{N}(0, \overline{\gamma} P_2)$

$X_2 = U_2 + W_2$ distributed according to $\mathcal{N}(0, P_2)$.

In this model, the received signals are given by

$$\begin{aligned} Y_1 &= X_1 + a_{21} X_2 + Z_1 \\ &= U_1 + W_1 + a_{21}(U_2 + W_2) + Z_1 \\ Y_2 &= a_{12} X_1 + X_2 + Z_2 \\ &= a_{12}(U_1 + W_1) + U_2 + W_2 + Z_2. \end{aligned}$$

Notice that although U_1, W_1, U_2, W_2 are independent, M_1, N_1, M_2, N_2 are not necessarily so. Bounds on the rates $R_{11}, R_{12}, R_{21}, R_{22}$ can be calculated as functions of the free parameters

$\alpha, \beta, \lambda, \gamma$, as well as $a_{12}, a_{21}, Q_1, Q_2, P_1, P_2$. First, we calculate the covariance matrix between all variables:

$$cov(Y_1, Y_2, M_1, N_1, M_2, N_2, X_1) = E[\Theta^T \Theta]$$

$$= \begin{pmatrix} P_1 + a_{21}^2 P_2 + Q_1 & a_{12}P_1 + a_{21}P_2 & \lambda P_1 & \overline{\lambda} P_1 & \alpha P_1 + a_{21}\gamma P_2 & \beta P_1 + a_{21}\overline{\gamma} P_2 & P_1 \\ a_{12}P_1 + a_{21}P_2 & a_{12}^2 P_1 + P_2 + Q_2 & a_{12}\lambda P_1 & a_{12}\overline{\lambda} P_1 & a_{12}\alpha P_1 + \gamma P_2 & a_{12}\beta P_1 + \overline{\gamma} P_2 & a_{12}P_1 \\ \lambda P_1 & a_{12}\lambda P_1 & \lambda P_1 & 0 & \alpha\lambda P_1 & \beta\lambda P_1 & \lambda P_1 \\ \overline{\lambda} P_1 & a_{12}\overline{\lambda} P_1 & 0 & \overline{\lambda} P_1 & \alpha\overline{\lambda} P_1 & \beta\overline{\lambda} P_1 & \overline{\lambda} P_1 \\ \alpha P_1 + a_{21}\gamma P_2 & a_{12}\alpha P_1 + \gamma P_2 & \alpha\lambda P_1 & \alpha\overline{\lambda} P_1 & \gamma P_2 + \alpha^2 P_1 & \alpha\beta P_1 & \alpha P_1 \\ \beta P_1 + a_{21}\overline{\gamma} P_2 & a_{12}\beta P_1 + \overline{\gamma} P_2 & \beta\lambda P_1 & \beta\overline{\lambda} P_1 & \alpha\beta P_1 & \overline{\gamma} P_2 + \beta^2 P_1 & \beta P_1 \\ P_1 & a_{12}P_1 & \lambda P_1 & \overline{\lambda} P_1 & \alpha P_1 & \beta P_1 & P_1 \end{pmatrix} \tag{21}$$

where $\Theta \stackrel{\triangle}{=} (Y_1\, Y_2\, M_1\, N_1\, M_2\, N_2\, X_1)$. The values for λ and γ are repeatedly randomly selected from the interval $[0, 1]$. The values of α and β are also repeatedly generated according to $\mathcal{N}(0, 1)$. There exist bounds on the admissible values of α and β in order to keep all upper bounds on the rates R_{11}, R_{12}, R_{21} and R_{22} positive. However, these are not explicitly considered, and whenever α, β values cause any bound to be negative, those particular values of α and β are rejected. For each 4-tuple $\lambda, \gamma, \alpha, \beta$, the above covariance matrix (21) yields all the information necessary to calculate the 14 bounds on $(R_{11}, R_{12}, R_{21}, R_{22})$ of Theorem 1. Each mutual information bound can be expanded in terms of entropies, which can then be evaluated by taking the determinant of appropriate sub-matrices of (21). The achievable regions thus obtained for the Gaussian genie-aided cognitive radio channel are plotted in Fig. 6. The innermost region (black) corresponds to the achievable region of [12], and is obtained by setting $\alpha = \beta = 0$. As expected, because of the extra information at the encoder and the partial use of a dirty-paper coding technique, our achievable region, the second to smallest region (cyan), extends that of [12]. Simulations were carried out until further simulations extended the regions negligibly. An upper bound on our achievable rate region is provided by the 2×2 Gaussian MIMO broadcast channel, whose capacity was recently computed in [28]. The largest in Fig. 6 is the intersection of the 2×2 Gaussian MIMO broadcast channel capacity region with the bound on $\mathcal{S}_2$'s rate $R_2 \leq \frac{1}{2}\log(1 + \text{SNR})$ provided by the interference-free channel in which $a_{12} = 0$. The Gaussian MIMO broadcast channel capacity region is computed using a covariance matrix constraint on the inputs $x = (x_1, x_2)^T$ of the form

$E[xx^T] \preceq S$, where S is of the form

$$S = \begin{pmatrix} P_1 & c \\ c & P_2 \end{pmatrix},$$

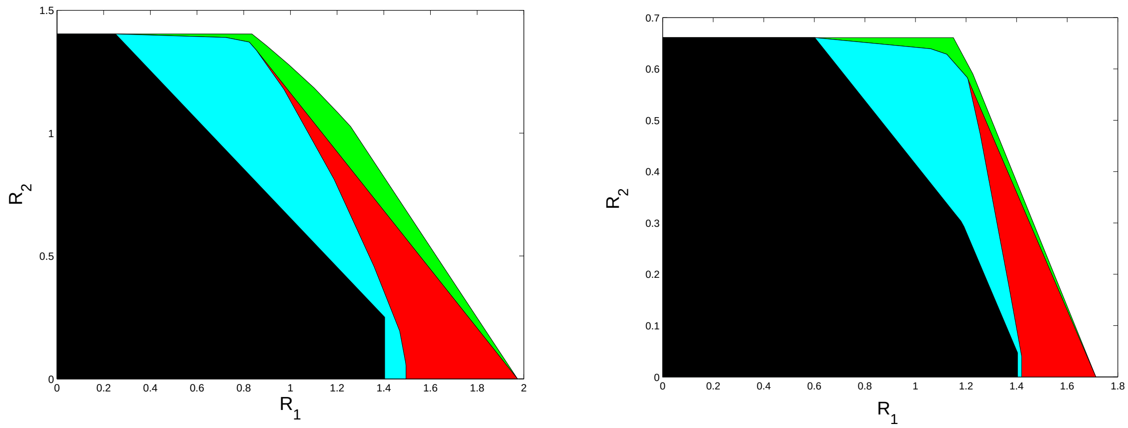
for some $-\sqrt{P_1 P_2} \leq c \leq \sqrt{P_1 P_2}$ (which ensures S is positive semi-definite). For each such S, and positive semi-definite matrices B and D, where $B + D \preceq S$, both rate pairs

$$R_1 \leq \frac{1}{2} \log \left(\frac{\det(H_1 B H_1^T + N_1)}{\det(N_1)} \right), \; R_2 \leq \frac{1}{2} \log \left(\frac{\det(H_2 (B+D) H_2^T + N_2)}{\det(H_2 B H_2^T + N_2)} \right)$$

and

$$R_1 \leq \frac{1}{2} \log \left(\frac{\det(H_1 (B+D) H_1^T + N_1)}{\det(H_1 D H_1^T + N_1)} \right), \; R_2 \leq \frac{1}{2} \log \left(\frac{\det(H_2 D H_2^T + N_2)}{\det(N_2)} \right)$$

are achievable, where $H_1 = (1, a_{21})$ and $H_2 = (a_{12}, 1)$. The convex hull of the union of these pairs over all possible S, B and D matrices will yield the capacity region of the 2×2 Gaussian MIMO channel with channel described by H_1 and H_2, and input covariance constraint matrix S. The 2×2 Gaussian MIMO broadcast channel is a channel in which two transmitters can cooperate in order to send messages to two independent, non-cooperating receivers. In the MIMO channel, both $\mathcal{S}_1$ and $\mathcal{S}_2$ know each others' messages, whereas in the genie-aided cognitive radio channel, $\mathcal{S}_2$ knows $\mathcal{S}_1$'s message, but not vice versa. There is a lack of symmetry, and this is apparent in the simulation plots, where it can be seen that the dirty paper coding technique aims to eliminate the interference from $\mathcal{S}_1$, and thus boosts the rate of $\mathcal{S}_2$ more than that of $\mathcal{S}_1$. Although $\mathcal{S}_1$ also sees rate increases, it is unclear whether the dirty paper coding performed by $\mathcal{S}_2$ is optimal for $\mathcal{S}_1$'s rate. An upper bound on the rate of $\mathcal{S}_2$ is provided by the interference-free channel in which $a_{12} = 0$. Thus, $R_2^{max} \leq \frac{1}{2} \log(1 + P_2/Q_2)$. For small R_1 this provides a tighter bound than the MIMO channel outer bound. *However, R_1 cannot be similarly bounded, as $\mathcal{S}_2$, which knows $\mathcal{S}_1$'s message could aid $\mathcal{S}_1$ in sending it, thus boosting $\mathcal{S}_1$'s rate above the interference-free channel case of* $a_{21} = 0$. In fact, the point $(R_1^*, 0)$ is achievable, where $R_1^* = \frac{1}{2} \log(1 + (\sqrt{P_1} + a_{21}\sqrt{P_2})^2/Q_1)$ is also achievable. Note that in the region of [28], the two transmit antennas can fully cooperate subject to the sum power constraint $P_1 + P_2$. In contrast, our simulations restrict the power of $\mathcal{S}_1$ to P_1, and the power of $\mathcal{S}_2$ to P_2. The result of [28] thus provides a bound. However, in the simulations, in order to mimic the individual power constraints on the two users for the MIMO case, the input covariance matrix was constrained to have diagonal elements P_1 and P_2.



(a) (b)

Fig. 6. The innermost polyhedron (black) is the achievable region of [12]. The next to smallest (cyan) is the achievable region for the genie-aided cognitive radio channel in Theorem 1. The second to largest region (red) is the achievable region of the cognitive radio channel (Theorem 3). The largest region (green) is the intersection of the capacity region of the 2×2 MIMO broadcast channel with the outer bound on R_2 of an interference-free Gaussian channel of capacity $1/2\log(1+P_2/Q_2)$. In (a) $Q_1 = Q_2 = 1$, $a_{12} = a_{21} = 0.55$, $P_1 = P_2 = 6$, in (b) $Q_1 = Q_2 = 1$, $a_{12} = a_{21} = 0.55$, $P_1 = 6$, $P_2 = 1.5$. *Note that since $\mathcal{S}_2$ knows $\mathcal{S}_1$'s message, it could aid $\mathcal{S}_1$ in sending it and boost R_1 above the interference-free channel case of $a_{21} = 0$, up to the vector channel rate of R_1^*.*

IV. COGNITIVE RADIO CHANNEL: THE CAUSAL CASE

In practice, the message x_1^n that $\mathcal{S}_1$ wants to transmit cannot be non-causally given to $\mathcal{S}_2$. The transmitter $\mathcal{S}_2$ must obtain the message in real time, and one possible way to do so is by exploiting proximity to $\mathcal{S}_1$. As in [23], this proximity is modeled by a reduction G in path loss, or equivalently, an increase in capacity between $\mathcal{S}_1$ and $\mathcal{S}_2$, relative to the channels between the senders and the receivers. If, for example, the channel between $\mathcal{S}_1$ and $\mathcal{S}_2$ is an Additive White Gaussian Noise channel, then the capacity increase, by a factor G, would be $C = \frac{1}{2}\log\left(1 + G\cdot \text{SNR}\right)$ and if $G \geq 1$ there is a geometric capacity gain. Alternatively, if $\mathcal{S}_1$ and $\mathcal{S}_2$ are base-stations, then it may be possible for S_2 to obtain S_1's message through a high bandwidth wired connection (if one exists) in real time. In the Gaussian cognitive radio channel model, all receivers know the channel between themselves and the relevant sender(s). In addition, both senders and receivers know the interference channel parameters a_{12} and a_{21}. We propose four protocols, in Lemmas 1, 2, 3 and 4, and derive the corresponding achievable regions shown in Fig. 8 (Protocol 1), and Fig. 9 (Protocol 2) which allow S_2 to obtain S_1's message in a causal manner. Protocol 3 is the achievable region of the interference channel shown in black in Fig. 6, while Protocol 4 produces a set of points of the form $(R, 0)$, and are not shown explicitly, but form a part of the convex hull of the 4 achievable regions shown in Fig. 10 (which is again achievable), and forms an inner bound on the causal achievable region. For Lemma 1 and 4, we assume that $\mathcal{S}_1$ knows the channel between itself and $\mathcal{S}_2$.

Lemma 1: Let Protocol 1 be a two phase protocol, for which phase 1 consists of a Gaussian broadcast channel between $\mathcal{S}_1 \rightarrow \mathcal{S}_2$ and $\mathcal{S}_1 \rightarrow \mathcal{R}_1$. During phase 1, S_2 is in "listening" mode, while S_1 transmits some portion of the M_1 message, $\mu n R_{11}$ bits of the total nR_1 bits to Y_1, and all of M_1 (nR_{11} bits) and N_1 (nR_{12} bits) to S_2. Let the 4-tuple $(R_{11}, R_{12}, R_{21}, R_{22})$ be an achievable rate pair of the modified genie-aided cognitive radio channel. Phase 2 transmission follows the Gaussian modified genie-aided cognitive radio channel scheme. Define $R_s(\alpha) = 1/2\ \log(1 + \frac{G\alpha P_1}{Q})$, and $R_y(\alpha) = 1/2\ \log(1 + \frac{(1-\alpha)P_1}{\alpha P_1 + Q_1})$, where Q is the additive Gaussian noise power at S_2, and G is the gain factor. Let $\hat{\alpha} \in [0, 1]$ be such that

$$\frac{n\mu R_{11}}{R_y(\hat{\alpha})} = \frac{n(R_{11} + R_{12})}{R_y(\hat{\alpha}) + R_s(\hat{\alpha})}. \tag{22}$$

Define $f \stackrel{\triangle}{=} \frac{\mu R_{11}}{R_y(\hat{\alpha})}$. Then if $\frac{1}{1-f}((1-\mu)R_{11}, R_{12}, R_{21}, R_{22})$ is achievable for the modified genie-aided cognitive radio channel, then the rate pair $(R_1 = R_{11} + R_{12}, R_2 = R_{21} + R_{22})$ is achievable for the causal case. □

Proof: In phase 1, consider the Gaussian broadcast channel between $\mathcal{S}_1 \rightarrow \mathcal{S}_2$ and $\mathcal{S}_1 \rightarrow \mathcal{R}_1$, and let R_s denote the rate between $\mathcal{S}_1$ and $\mathcal{S}_2$, and R_y denote the rate between $\mathcal{S}_1$ and $\mathcal{R}_1$ (output Y_1). Let the noise at $\mathcal{S}_2$ be additive Gaussian noise of power Q, and the gain factor between $\mathcal{S}_1$ and $\mathcal{S}_2$ be G. Then the following broadcast rates [5] are achievable for $0 \leq \alpha \leq 1$,

$$R_s(\alpha) < \frac{1}{2}\log\left(1 + \frac{G\alpha P_1}{Q}\right)$$
$$R_y(\alpha) < \frac{1}{2}\log\left(1 + \frac{(1-\alpha)P_1}{\alpha P_1 + Q_1}\right)$$

Now, consider trying to achieve a rate of $R_1 = R_{11} + R_{12}$ for the causal Gaussian cognitive radio channel, where R_{11}, R_{12} are achievable rates for the modified Gaussian cognitive radio channel. For a given $\mu \in [0, 1]$, we try to find $\hat{\alpha}$ such that the messages at Y_1 and $\mathcal{S}_2$ are fully received *simultaneously*. Thus, we try to find α such that, $\frac{n\mu R_{11}}{R_y(\hat{\alpha})} = \frac{n(R_{11}+R_{12})}{R_y(\hat{\alpha})+R_s(\hat{\alpha})}$. We let $f \stackrel{\triangle}{=} \frac{\mu R_{11}}{R_y(\hat{\alpha})} = \frac{R_{11}+R_{12}}{R_y(\hat{\alpha})+R_s(\hat{\alpha})}$. This is the fraction of the transmission duration $\mathcal{S}_1$ spends in the broadcast channel phase. During this phase, Y_1 has obtained μR_{11}. Thus, in order to send the overall rates R_{11}, R_{12}, during phase 2, of duration $(1-f)$ of the total transmission length, the rate $(\frac{1}{1-f}((1-\mu)R_{11}, R_{12}, R_{21}, R_{22})$ must be achievable for the modified Gaussian cognitive radio channel. If this is the case, then the rate $(R_{11}, R_{12}, R_{21}, R_{22})$ has been achieved for the overall causal modified cognitive radio channel, leading to a rate of $(R_{11}+R_{12}, R_{21}+R_{22})$ for the causal cognitive radio channel. ■

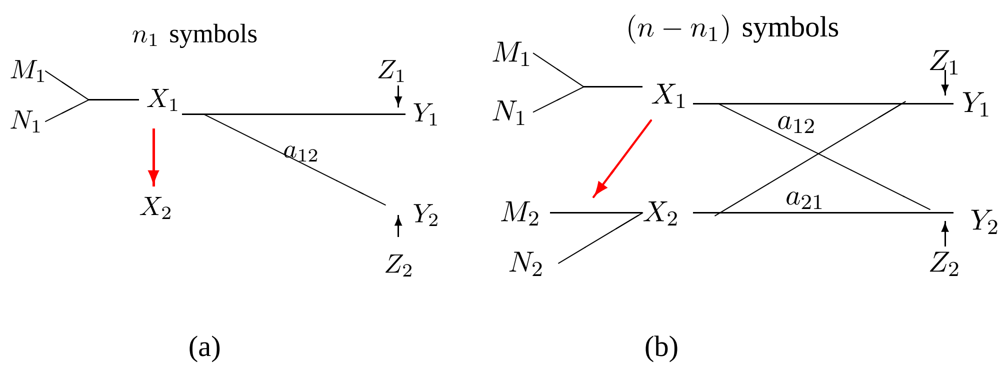


Fig. 7. (a) illustrates the listening phase of the cognitive radio channel and (b) illustrates the cognitive radio channel phase.

Next consider Protocol 2 which consists of two phases, and in which $\mathcal{S}_1$ transmits using the distribution for C_{cog}^m during both phases. The two phases can still be viewed as in Fig. 7, however the underlying distributions differ from those of Protocol 1. In phase 1, $\mathcal{S}_1$ transmits to Y_1 and Y_2, while $\mathcal{S}_2$ is in "listening" mode, and refrains from transmission until it has completely overheard and decoded the message of $\mathcal{S}_1$. At this point, the scheme enters phase 2, in which $\mathcal{S}_2$ starts transmission as well, according to the cognitive radio channel of Section II. In phase 2, $\mathcal{S}_1$ continues transmitting according to the same distribution, however a reduced rate will be necessary, due to the added interference from the now-transmitting $\mathcal{S}_2$.

In order to determine the rate pairs (R_1, R_2) achievable in this causal scheme, let nR_1 be the total number of bits to be transmitted by $\mathcal{S}_1$. Define $f = \frac{n_1}{n}$, where $0 \leq n_1 \leq n$ is the number of symbols for phase 1, and $(n-n_1)$ is the number of symbols for phase 2. Then,

Lemma 2: Let $R_{11}^{'}$ be the rate achieved by $\mathcal{S}_1$ to $\mathcal{R}_1$ and R_{12}' be the rate achieved by $\mathcal{S}_1$ to $(\mathcal{R}_1, \mathcal{R}_2)$ in phase 1, and $R_{11}^{''}, R_{12}^{''}, R_{21}^{''}, R_{22}^{''}$ be the rates achieved by $\mathcal{S}_1, \mathcal{S}_2$ respectively during the cognitive phase 2. Then if

$$\max\left(\frac{R_{11}^{''}}{R_{11}^{''}+I(M_1;S_2|N_1)-R_{11}^{'}}, \frac{R_{12}^{''}}{R_{12}^{''}+I(N_1;S_2|M_1)-R_{12}^{'}}, \frac{R_{11}^{''}+R_{12}^{''}}{R_{11}^{''}+R_{12}^{''}+I(M_1,N_1;S_2)-R_{11}^{'}-R_{12}^{'}}\right) \leq f \leq 1,$$

the rate pair $(R_1 = f(R_{11}^{'}+R_{12}^{'}) + (1-f)(R_{11}^{''}+R_{12}^{''}), R_2 = (1-f)(R_{21}^{''}+R_{22}^{''}))$ is achieved by Protocol 2. □

Proof: Let $n_1 = fn$. Phase 1 and phase 2 have durations n_1 and $(n - n_1)$ symbols respectively. By definition of $R_{11}', R_{12}', R_{11}'', R_{12}''$, the overall data transmitted by $\mathcal{S}_1$ during the n symbols is $nR_1 = n_1(R_{11}' + R_{12}') + (n - n_1)(R_{11}'' + R_{12}'')$. However, in order for $\mathcal{S}_2$ to reliably obtain the message of $\mathcal{S}_1$ in the first n_1 symbols over the channel between M_1, N_1 and S_2, using the distribution employed in Section II, the Multiple Access Channel (MAC) constraints must be satisfied. This requires choosing n_1 large enough to simultaneously satisfy all 3 constraints (we abuse notation and let S_2 denote the received signal at $\mathcal{S}_2$)

$$n_1 I(M_1; S_2|N_1) \geq nR_{11} = n_1 R_{11}' + (n - n_1)R_{11}''$$

$$n_1 I(N_1; S_2|M_1) \geq nR_{12} = n_1 R_{12}' + (n - n_1)R_{12}''$$

$$n_1 I(M_1, N_1; S_2) \geq nR_1 = n_1(R_{11}' + R_{12}') + (n - n_1)(R_{11}'' + R_{12}'')$$

Note that these mutual informations are evaluated according to the distribution for S_1 given in Section II. This leads to the requirement of

$\max\left(\frac{R_{11}''}{R_{11}''+I(M_1;S_2|N_1)-R_{11}'}, \frac{R_{12}''}{R_{12}''+I(N_1;S_2|M_1)-R_{12}'}, \frac{R_{11}''+R_{12}''}{R_{11}''+R_{12}''+I(M_1,N_1;S_2)-R_{11}'-R_{12}'}\right) \leq f = \frac{n_1}{n} \leq 1.$

During phase 2, of length $(n - n_1)$ symbols, the rates $R_{11}'' + R_{12}''$ ans $R_{21}'' + R_{22}''$ are achievable for a fraction $(1 - f)$ of the total transmission length. Thus, weighting the two portions yields the achievable rate pair $(R_1 = f(R_{11}' + R_{12}') + (1 - f)(R_{11}'' + R_{12}''), R_2 = (1 - f)(R_{21}'' + R_{22}''))$. ■

Finally, yet another protocol is achievable: let Protocol 3 denote a scheme in which $\mathcal{S}_2$ starts transmission immediately and does not obtain the message of $\mathcal{S}_1$, as in the interference channel. Any point achievable for the interference channel is achievable here. Lemma 3 formally states this:

Lemma 3: The interference channel rates of [12] can be causally achieved. □

This final lemma describes a causal scheme, Protocol 4, in which to achieve a rate pair of the form $(R_1^*, 0)$ where $\mathcal{S}_2$ sends no information of its own and simply aids $\mathcal{S}_1$ in sending $\mathcal{S}_1's$ message.

Lemma 4: Let Protocol 4 be a two phase protocol, for which phase 1 consists of a Gaussian broadcast channel between $\mathcal{S}_1$ and both $\mathcal{S}_2$ and $\mathcal{R}_1$. For any $\alpha \in [0, 1]$, let $R_1(\alpha)$ and $R_2(\alpha)$ denote the broadcast rates [5] between $\mathcal{S}_1 \to \mathcal{R}_1$ and $\mathcal{S}_1 \to \mathcal{S}_2$ respectively. Let the additive Gaussian noise at sender $\mathcal{S}_2$ be of power Q, and the gain factor between $\mathcal{S}_1$ and $\mathcal{S}_2$ be G. Let $R_1^* = \frac{1}{2}\log(1 + \frac{(\sqrt{P_1}+a_{21}\sqrt{P_2})^2}{Q_1})$, the rate achievable in phase 2 during which $\mathcal{S}_2$ collaborates to transmit $\mathcal{S}_1$'s message according to the optimal distribution for the vector channel between

$\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{R}_1$ (Theorem 2). Let $f = \frac{1}{1+R_2(\alpha)/R_1^*}$. Then the rate pair

$$(f(R_1(\alpha) + R_2(\alpha)), 0) \tag{23}$$

is achievable in a causal fashion. □

Proof: In phase 1, consider the Gaussian broadcast channel between $\mathcal{S}_1$ and both $\mathcal{S}_2$ and $\mathcal{R}_1$, and let R_1 denote the rate between $\mathcal{S}_1$ and $\mathcal{R}_1$, and R_2 denote the rate between $\mathcal{S}_1$ and $\mathcal{S}_2$. Let the noise at $\mathcal{S}_2$ be additive Gaussian noise of power Q, and the gain factor between $\mathcal{S}_1$ and $\mathcal{S}_2$ be G. Then the following broadcast rates [5] are achievable for any given $0 \leq \alpha \leq 1$,

$$R_1(\alpha) < \frac{1}{2}\log\left(1 + \frac{(1-\alpha)P_1}{\alpha P_1 + Q_1}\right)$$
$$R_2(\alpha) < \frac{1}{2}\log\left(1 + \frac{G\alpha P_1}{Q}\right)$$

Let phase 1 be of duration n_1 symbols, and phase 2 be of duration n_2 symbols. During phase 1, $\mathcal{R}_1$ receives $R_1(\alpha)n_1$ bits, while $\mathcal{S}_2$ receives $(R_1(\alpha)+R_2(\alpha))n_1$ bits. We also require that $\mathcal{S}_2$ receives the total number of bits to be sent to $\mathcal{R}_1$ during the first n_1 symbols. Thus, if the overall rate (from $\mathcal{S}_1$ to $\mathcal{R}_1$) achieved is denoted by R, then $n_1(R_1(\alpha)+R_2(\alpha)) = (n_1+n_2)R$. During phase 2, both senders form a vector channel in order to send the remaining $n_1R_2(\alpha)$ bits. They do so at the maximal rate possible for this vector channel, given by $R_1^* = \max_{p(x_1,x_2)} I(X_1, X_2; Y_1) = \frac{1}{2}\log(1 + \frac{(\sqrt{P_1}+a_{21}\sqrt{P_2})^2}{Q_1})$. Thus, equating the number of bits sent during phase 2 we obtain $n_2R_1^* = n_1R_2(\alpha)$. Defining $f = \frac{n_1}{n_1+n_2}$ to be the fraction of the total transmission duration spent in phase 1, we have $f = \frac{1}{1+R_2(\alpha)/R_1^*}$ and $R = f(R_1(\alpha) + R_2(\alpha))$. ■

These 4 Lemmas can be combined to form an overall causal achievable region.

Theorem 4: The convex hull of the regions achieved in Lemmas 1, 2, 3 and 4 under Protocols 1, 2, 3 and 4 respectively is also achievable. □

In order to demonstrate the effect of causality on the achievable region in the Gaussian noise case, for Protocol 1, consider Fig. 8. For values of the gain factor $G = 1$ and $G = 10$, a finite set of μ taken from $[0, 1]$, and for a certain genie-aided achievable rate-tuple $(R_{11}, R_{12}, R_{21}, R_{22})$, we solve for $\hat{\alpha}$ such that (22) is satisfied. If such an $\hat{\alpha}$ exists, form f and verify whether $\frac{1}{1-f}((1-\mu)R_{11}, R_{12}, R_{21}, R_{22})$ lies in the achievable region of the modified genie-aided Gaussian cognitive radio channel. If so, $(R_{11}, R_{12}, R_{21}, R_{22})$ is achievable in the causal case. Fig. 8 demonstrates the regions attained by Protocol 1 for $G = 1$ (innermost, blue) and $G = 10$ (middle, yellow) as compared to the overall achievable region of the genie-aided cognitive radio channel of Theorem 3 (cyan). For Protocol 2, the regions of Fig. 9 are achievable for $G = 1$ (innermost, blue)

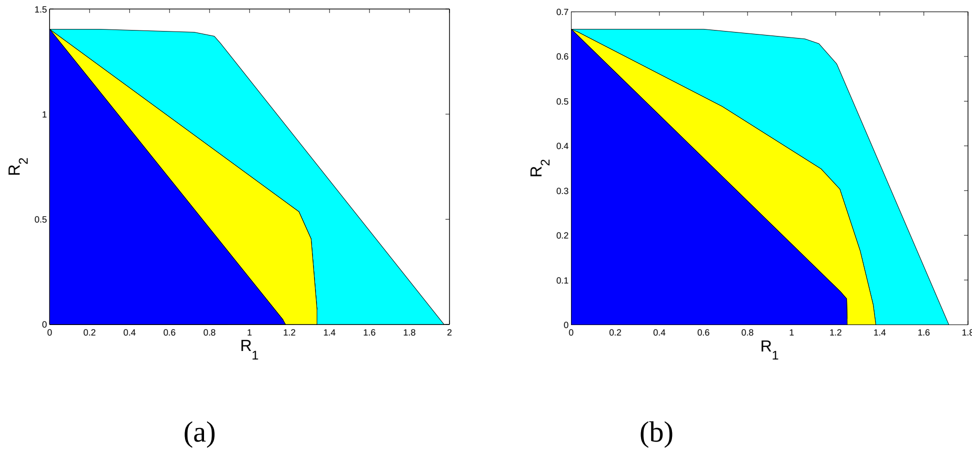
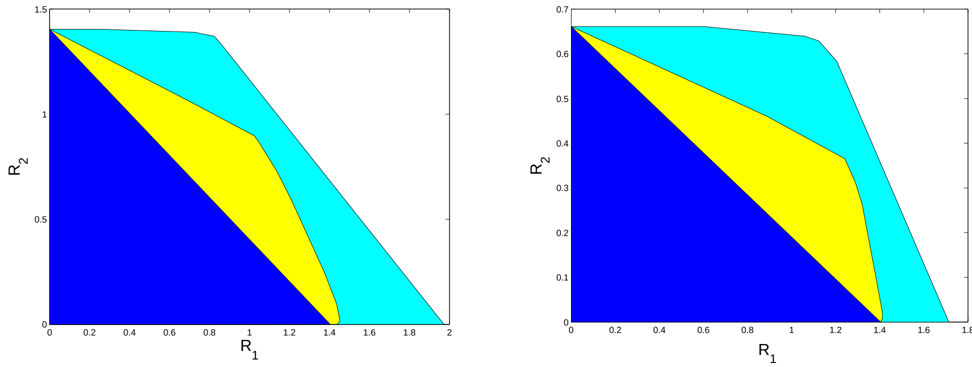


Fig. 8. The outermost (cyan) curve is the overall genie-aided achievable region of the genie-aided cognitive radio channel as in Theorem 3. Both plots demonstrate the various regions causally attained using Protocol 1 for gain factor values $G = 1$ (innermost, blue) and 10 (middle, yellow). Both figures illustrate the regions with parameters $Q_1 = Q_2 = 1$ and $a_{12} = a_{21} = 0.55$, and in (a) $P_1 = P_2 = 6$, in (b) $P_1 = 6,\ P_2 = 1.5$

and $G = 10$ (middle, yellow), and are compared to the genie-aided achievable region of the cognitive radio channel of Theorem 3 (cyan). In order to calculate these regions, we use the same assumptions on the forms of the relevant random variables as in Section III. To calculate $R_{11}^{'}$ and $R_{12}^{'}$ one can use the equations of Theorem 1, ignoring all of $\mathcal{S}_2$'s signals, as it is not transmitting anything during phase 1. That is, (R'_{11}, R'_{12}) satisfy:

$$R'_{11} \leq I(Y_1, N_1; M_1)$$
$$R'_{12} \leq I(Y_1, M_1; N_1)$$
$$R'_{11} + R'_{12} \leq I(Y_1; M_1, N_1)$$
$$R'_{12} \leq I(Y_2; N_1).$$

These mutual information terms are evaluated using the assumed Gaussian forms on the random variables of Section III. Finally, $R_1^{''}$ and $R_2^{''}$ are exactly the rates calculated in Section III. However, these rates are only achieved for a fraction of the total n symbols. Carrying out the simulation yields Fig. 9. Protocol 3 yields the same region as the interference channel, as computed in [12], and is plotted here as the innermost (black) region of Fig. 6. Protocol 4 yields points of the form $(R, 0)$ for each selected value of α, the power tradeoff parameter for the broadcast channel, and for each gain factor value G. For $G = 1$, the maximal Protocol 4 point was 1.4037 bits/second



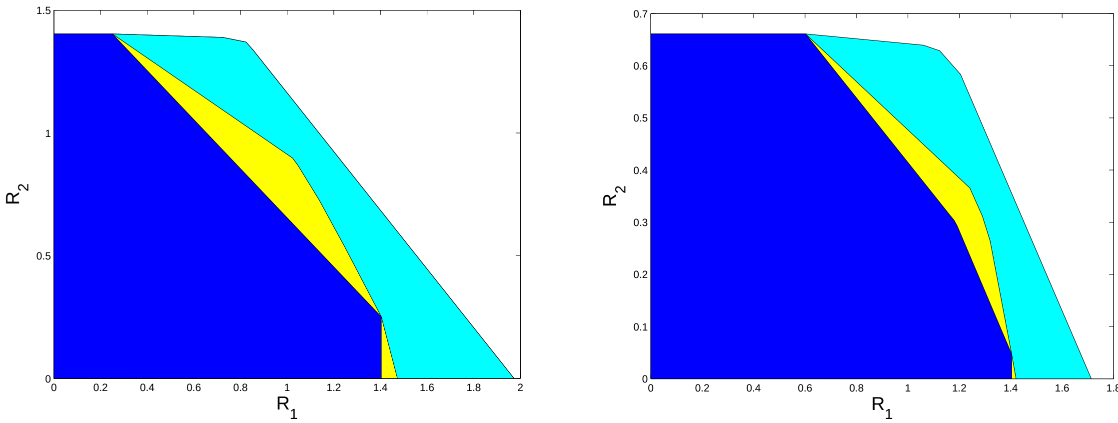
(a) (b)

Fig. 9. The outermost (cyan) curve is the genie-aided cognitive channel achievable region of Theorem 3. Both plots demonstrate the various regions causally attained using Protocol 2 for values of $G = 1$ (innermost, blue), $G = 10$ (middle, yellow), and the genie-aided achievable region (outermost, cyan). Both figures illustrate the regions for parameters $Q_1 = Q_2 = 1$ and $a_{12} = a_{21} = 0.55$, and in (a) $P_1 = P_2 = 6$, in (b) $P_1 = 6,\ P_2 = 1.5$.

and for $G = 10$, the maximal point achieved by Protocol 4 was 1.4730 bits/second for $P_2 = 6$ and 1.4026 bits/second for $P_2 = 1.5$. The overall causal achievable region is then the convex hull of the regions achieved under Protocols 1, 2, 3 and 4. This region is shown in Fig. 10

V. CONCLUSION

Although interest in cognitive radio technology has exploded recently, theoretical knowledge concerning its limits is still being acquired. In this paper, we contribute to this emerging field by defining and proving an achievable region for a more flexible and potentially more efficient transmission model for cognitive radio channels. In contrast to the traditional cognitive radio channel model or protocol in which a sender fills voids in time/spectrum (i.e., wait for silence or unused frequencies), a second sender may transmit with an existing sender at the same time or in the same frequency band. Thus the generalized *cognitive radio channel* is modeled as an interference channel in which two senders (more generally m) communicate over a common medium to two independent, non-cooperating receivers (more generally n), and the k-th sender is given, or causally obtains the messages of the $k-1$ preceding senders. We computed an achievable region for the *genie-aided cognitive radio channel* in which one sender is non-causally given the other's message. We then removed the non-causal constraint, and four protocols which allow $\mathcal{S}_2$ to causally obtain $\mathcal{S}_1$'s message were proposed. Three of the four protocols use a 2 phase technique.



(a) (b)

Fig. 10. The outermost region (cyan) is the achievable region of the genie-aided cognitive radio channel. Both plots demonstrate the various regions attained using a convex combination of the Protocol 1, 2, 3 and 4 for values of $G = 1$ (innermost, blue) and $G = 10$ (middle, yellow) and parameters $Q_1 = Q_2 = 1$, $a_{12} = a_{21} = 0.55$, and in (a) $P_1 = P_2 = 6$, in (b) $P_1 = 6$, $P_2 = 1.5$. *Note that since $\mathcal{S}_2$ knows $\mathcal{S}_1$'s message, it could aid $\mathcal{S}_1$ in sending it, and boost R_1 above the interference-free channel case of $a_{21} = 0$.*

During the first phase $\mathcal{S}_2$ obtains $\mathcal{S}_1$'s message through either a degraded broadcast, or a MAC channel between itself and $\mathcal{S}_1$, and during the second phase the genie-aided rates are achievable. In this genie-aided scheme, the sender with the non-causal interference knowledge uses dirty paper coding, as in [3], to cancel the interference from $\mathcal{S}_1$ to $\mathcal{S}_2$. Dirty paper coding is performed on top of the information-separating technique first proposed by Han and Kobayashi in [12], which yields, in most cases [18], the largest to date known achievable region for the interference channel. Simulations in a Gaussian noise case show that the region achieved approaches the 2×2 MIMO channel upper bound, as well as the ideal upper bound on R_2 provided by an interference-free channel. We described a coding technique and provided theoretical answers to some of the questions in the emerging field of cognitive radios.

APPENDIX

Probability of error analysis:

Consider P_e, the sum of the average probability of errors of the two senders. The average is taken over all random codes generated as described in Section II. It is assumed that all messages $s \in \mathcal{S}$ are equi-probable. Without loss of generality it is assumed that $s = (1, 1, 1, 1)$ is sent with dither w^n. Notice that the first two components s_{11} and s_{12} are message indices, whereas the last two components s_{21} and s_{22} are bin indices. Then P_e may be bounded by, for each dither sequence w^n,

$$P_e \leq \Pr\{\psi_1(y_1^n, w^n) \neq (1,1,1)|s = (1,1,1,1)\} + \Pr\{\psi_2(y_2^n, w^n) \neq (1,1,1)|s = (1,1,1,1)\}.$$

Although the decoding at $\mathcal{R}_1$ and $\mathcal{R}_2$ is described in terms of $\psi_1(\cdot,\cdot)$ and $\psi_2(\cdot,\cdot)$, it is more convenient to write certain probabilities of error events directly in terms of n-sequences $m_1^n, n_1^n, m_2^n, n_2^n$. One type of decoding error occurs when a decoded message/bin index does not equal the sent message/bin index. Recall that in order to send message indices s_{11} and s_{12} the n-sequences $m_1^n(s_{11})$ and $n_1^n(s_{12})$ are selected and used to compute x_1^n. Also, the bin indices s_{21} and s_{22} are used to find n-sequences $n_2^n(s_{11}, s_{12}, s_{21}, l)$ (n_2^n is in bin s_{21} and is sequence number l in that bin, chosen so that it and $m_1^n(s_{11}), n_1^n(s_{12})$ are jointly typical) and $m_2^n(s_{11}, s_{12}, s_{22}, k)$ (m_2^n is in bin s_{22} and is sequence number k in that bin, chosen so that it and $m_1^n(s_{11}), n_1^n(s_{12})$ are jointly typical) respectively, which are used to obtain x_2^n. Without loss of generality, assume the selected n-sequences are $m_1^n(1), n_1^n(1), m_2^n(1,1,1,\hat{k}), n_2^n(1,1,1,\hat{l})$. A sufficient condition for correct decoding of the message $s = (1,1,1,1)$ is that $S_1(y_1^n, w^n)$ and $S_2(y_2^n, w^n)$ each contain exactly one tuple $(m_1^n(1), n_1^n(1), n_2^n(1,1,1,\hat{l}), w^n)$ and $(n_1^n(1), m_2^n(1,1,1,\hat{k}), n_2^n(1,1,1,\hat{l}), w^n)$ respectively. Then the probabilities of error can be upper bounded as

$$\begin{aligned}
&\Pr\{\psi_1^n(y_1^n, w^n) \neq (1,1,1)|s = (1,1,1,1)\} \\
&\quad \leq \Pr\{(m_1^n(1), n_1^n(1), n_2^n(1,1,1,\hat{l}), w^n) \text{ is not the only element in } S_1(y_1^n, w^n)|s = (1,1,1,1)\} \\
&\Pr\{\psi_2^n(y_2^n, w^n) \neq (1,1,1)|s = (1,1,1,1)\} \\
&\quad \leq \Pr\{(n_1^n(1), m_2^n(1,1,1,\hat{k}), n_2^n(1,1,1,\hat{l}), w^n) \text{ is not the only element in } S_2(y_2^n, w^n)|s = (1,1,1,1)\}.
\end{aligned}$$

The indices i, j, k, l are associated with the random variables M_1, N_1, M_2, N_2 respectively. Define the error events:

$$\begin{aligned}
E_0^1 &= \{\nexists \hat{l} \text{ s.t. } (m_1^n(1), n_1^n(1), n_2^n(1,1,1,\hat{l}), w^n) \notin A_\epsilon^n(M_1, N_1, N_2|W)), 1 \leq \hat{l} \leq 2^{n(L_{21}-R_{21})}\} \\
E_1^1 &= \{(y_1^n, m_1^n(1), n_1^n(1), n_2^n(1,1,1,\hat{l}), w^n) \notin A_\epsilon^n(Y_1, M_1, N_1, N_2|W)\} \\
E_{ijs_{21}l}^1 &= \{(y_1^n, m_1^n(i), n_1^n(j), n_2^n(1,1,s_{21},l), w^n) \in A_\epsilon^n(Y_1, M_1, N_1, N_2|W)\} \\
E_0^2 &= \{\nexists \hat{k} \text{ s.t. } (m_1^n(1), n_1^n(1), m_2^n(1,1,1,\hat{k}), n_2^n(1,1,1,\hat{l}), w^n) \notin A_\epsilon^n(M_1, N_1, M_2, N_2|W), \\
&\qquad 1 \leq \hat{k} \leq 2^{n(L_{22}-R_{22})}\}
\end{aligned}$$

$$E_1^2 = \{(y_2^n, n_1^n(1), m_2^n(1,1,1,\hat{k}), n_2^n(1,1,1,\hat{l}), w^n) \notin A_\epsilon^n(Y_2, M_2, N_1, N_2|W)\}$$

$$E_{js_{22}ks_{21}l}^2 = \{(y_2^n, n_1^n(j), m_2^n(1,1,s_{22},k), n_2^n(1,1,s_{21},l), w^n) \in A_\epsilon^n(Y_2, N_1, M_2, N_2|W)\}$$

Then, if $\overline{X}$ denotes the complement of event X,

$$\begin{aligned} P_e &\le \Pr\{E_1^1 \cup \bigcup_{ijs_{21}l \neq 111\hat{l}} E_{ijs_{21}l}^1 \cap \overline{E_1^1}\} \\ &\quad + \Pr\{E_1^2 \cup \bigcup_{js_{22}ks_{21}l \neq 11\hat{k}1\hat{l}} E_{js_{22}ks_{21}l}^2 \cap \overline{E_1^2}\} \\ &\le \Pr\{E_0^1\} + \Pr\{E_1^1|\overline{E_0^1}\}\Pr\{\overline{E_0^1}\} + \sum_{ijs_{21}l \neq 111\hat{l}} \Pr\{E_{ijs_{21}l}^1|\overline{E_1^1}\}\Pr\{\overline{E_1^1}\} \\ &\quad + \Pr\{E_0^2\} + \Pr\{E_1^2|\overline{E_0^2}\}\Pr\{\overline{E_0^2}\} + \sum_{js_{22}ks_{21}l \neq 11\hat{k}1\hat{l}} \Pr\{E_{js_{22}ks_{21}l}^2|\overline{E_1^2}\}\Pr\{\overline{E_1^2}\}. \end{aligned}$$

We examine each error event separately:

$$\begin{aligned} \Pr\{E_0^1\} &\le \Pr\{(m_1^n(1), n_1^n(1), w^n) \notin A_\epsilon^n(M_1, N_1|W)\} \\ &\quad + \prod_{1 \le l \le 2^{n(L_{21}-R_{21})}} \Pr\{(m_1^n(1), n_1^n(1), n_2^n(1,1,1,l), w^n) \notin A_\epsilon^n(M_1, N_1, N_2|W)|(m_1^n, n_1^n, w^n) \in A_\epsilon^n(M_1, N_1|W)\} \\ &= \Pr\{(m_1^n(1), n_1^n(1), w^n) \notin A_\epsilon^n(M_1, N_1|W)\} \qquad (24) \\ &\quad + \prod_{1 \le l \le 2^{n(L_{21}-R_{21})}} \Pr\{(m_1^n(1), n_1^n(1), n_2^n(1,1,1,l), x_1^n, w^n) \notin A_\epsilon^n(M_1, N_1, N_2|W)|(m_1^n, n_1^n, x_1^n, w^n) \in A_\epsilon^n(M_1, N_1, X_1|W)\} \\ &\le \epsilon + (1 - 2^{-n(I(N_2;X_1,M_1,N_1|W)+3\epsilon)})^{2^{n(L_{21}-R_{21})}} \qquad (25) \\ &\le \epsilon + e^{-2^{-n(I(N_2;X_1,M_1,N_1|W)+3\epsilon-L_{21}+R_{21})}} \qquad (26) \end{aligned}$$

The equality (24) follows from $x_1^n = f^n(m_1^n, n_1^n)$ and (25), (26) follow the form of [5], p. 356. Thus, $\Pr\{E_0^1\}$ will decay to 0 as $n \to \infty$ as long as $L_{21} - R_{21} > I(N_2; X_1, M_1, N_1|W) + 3\epsilon$. Since $(M_1, N_1) \to X_1 \to N_2$ is a Markov chain, $I(N_2; X_1, M_1, N_1|W) = I(N_2; X_1|W)$, and so (3) is an equivalent condition.

For S_2, we wish to show that if equations (3) and (4) hold, then $\Pr\{E_0^2\} \to 0$ as $n \to \infty$. Equation (3) ensures that there are enough sequences n_2^n in each bin to find a jointly typical one with x_1^n, and likewise Equation (4) ensures that a sufficient number of m_2^n are generated so that in each bin an m_2^n can be found that is jointly typical with x_1^n. To support the latter claim, consider:

$$\begin{aligned} &\prod_{1 \le k \le 2^{n(L_{22}-R_{22})}} \Pr\{(m_1^n(1), n_1^n(1), m_2^n(1,1,1,k), w^n) \notin A_\epsilon^n(M_1, N_1, M_2|W)|(m_1^n, n_1^n, w^n) \in A_\epsilon^n(M_1, N_1|W)\} \\ &\qquad \le (1 - 2^{-n(I(M_2;X_1,M_1,N_1|W)+3\epsilon)})^{2^{n(L_{22}-R_{22})}} \\ &\qquad \le e^{-2^{-n(I(M_2;X_1,M_1,N_1|W)+3\epsilon-L_{22}-R+22)}} \end{aligned}$$

This will decay as $n \to \infty$ provided $L_{22} - R_{22} > I(M_2; X_1, M_1, N_1|W) + 3\epsilon = I(M_2; X_1|W) + 3\epsilon$ (equation (4)). Similarly, if $L_{21} - R_{21} > I(N_2; X_1, M_1, N_1|W) + 3\epsilon = I(N_2; X_1|W) + 3\epsilon$ (equation (3)), then the analogous equations for n_2^n will $\to 0$ as $n \to \infty$. Thus, we have shown that provided equations (3) and (4) hold, indices $\hat{k}$ and $\hat{l}$ may be found separately such that (m_1^n, n_1^n, m_2^n) and (m_1^n, n_1^n, n_2^n) are strong jointly typical triples. We now wish to show that the 4-tuple $(m_1^n, n_1^n, m_2^n, n_2^n)$ is also strongly jointly typical. This is not the case in general, but when the distribution $p(m_2, n_2|m_1, n_1)$ is of the form $p(m_2|m_1, n_1)p(n_2|m_1, n_1)$ then with high probability this will be the case, as proven in Lemma 5 at the end of this section.

Next, consider the error events E_1^1, E_1^2. Since $(M_1, N_1, M_2, N_2) \to (X_1, X_2) \to (Y_1, Y_2)$ is a Markov chain, by the Markov Lemma (which holds for strongly typical sequences [5], [6], [1]), the probability that $(y_1^n, m_1^n, n_1^n, n_2^n, w^n)$ are not jointly typical goes to 0 as $n \to \infty$, and likewise for $(y_2^n, m_2^n, n_1^n, n_2^n, w^n)$.

Finally, consider all the possible joint decoding errors, given that the channel inputs and outputs are jointly typical. We suppose indices $\hat{l}$ and $\hat{k}$ have been chosen and that

$$T_1 = \{(y_1^n, m_1^n(1), n_1^n(1), n_2^n(1,1,1,\hat{l}), w^n) \in A_\epsilon^n(Y_1, M_1, N_1, N_2|W)\}$$

$$T_2 = \{(y_2^n, n_1^n(1), m_2^n(1,1,1,\hat{k}), n_2^n(1,1,1,\hat{l}), w^n) \in A_\epsilon^n(Y_2, N_1, M_2, N_2|W)\}.$$

Then, for any $\bar{l} \neq \hat{l}$ and $\bar{k} \neq \hat{k}$, since $\Pr\{E_{111\bar{l}}^1|T_1\} = \Pr\{E_{1121}^1|T_1\} = \Pr\{E_{112\bar{l}}^1|T_1\}$ and analogously for $\Pr\{E_{1s_{22}ks_{21}l}^2|T_2\}$ at $\mathcal{R}_2$,

$$\begin{aligned}
&\sum_{ijs_{21}l\neq 111\hat{l}} \Pr\{E_{ijs_{21}l}^1|T_1\} \\
&\quad = \sum_{ijs_{21}l\neq 111\hat{l}} \Pr\{(y_1^n, m_1^n(i), n_1^n(j), n_2^n(1,1,s_{21},l), w^n) \in A_\epsilon^n(Y_1, M_1, N_1, N_2|W)|T_1\} \\
&\quad \leq (2^{n(R_{11}-6\epsilon)} - 1) \cdot \Pr\{E_{211\hat{l}}^1|T_1\} \\
&\quad\quad + (2^{n(R_{12}-6\epsilon)} - 1) \cdot \Pr\{E_{121\hat{l}}^1|T_1\} \\
&\quad\quad + (2^{n(L_{21}-6\epsilon)} - 1) \cdot \Pr\{E_{112\bar{l}}^1|T_1\} \\
&\quad\quad + (2^{n(R_{11}-6\epsilon)} - 1)(2^{n(R_{12}-6\epsilon)} - 1) \cdot \Pr\{E_{221\hat{l}}^1|T_1\} \\
&\quad\quad + (2^{n(R_{11}-6\epsilon)} - 1)(2^{n(L_{21}-6\epsilon)} - 1) \cdot \Pr\{E_{212\bar{l}}^1|T_1\} \\
&\quad\quad + (2^{n(R_{12}-6\epsilon)} - 1)(2^{n(L_{21}-6\epsilon)} - 1) \cdot \Pr\{E_{122\bar{l}}^1|T_1\} \\
&\quad\quad + (2^{n(R_{11}-6\epsilon)} - 1)(2^{n(R_{12}-6\epsilon)} - 1)(2^{n(L_{21}-6\epsilon)} - 1) \cdot \Pr\{E_{222\bar{l}}^1|T_1\}
\end{aligned}$$

$$
\begin{aligned}
&\le 2^{n(R_{11}-I(M_1;Y_1,N_1,N_2|W)-6\epsilon+2\epsilon)} \\
&\quad + 2^{n(R_{12}-I(N_1;Y_1,M_1,N_2|W)-6\epsilon+2\epsilon)} \\
&\quad + 2^{n(L_{21}-I(N_2;Y_1,M_1,N_1|W)-6\epsilon+2\epsilon)} \\
&\quad + 2^{n(R_{11}+R_{12}-I(M_1,N_1;Y_1,N_2|W)-6\epsilon+2\epsilon)} \\
&\quad + 2^{n(R_{11}+L_{21}-I(M_1,N_2;Y_1,N_1|W)-6\epsilon+2\epsilon)} \\
&\quad + 2^{n(R_{12}+L_{21}-I(N_1,N_2;Y_1,M_1|W)-6\epsilon+2\epsilon)} \\
&\quad + 2^{n(R_{11}+R_{12}+L_{21}-I(M_1,N_1,N_2;Y_1|W)-6\epsilon+2\epsilon)}
\end{aligned}
$$

$$
\begin{aligned}
\sum_{js_{22}ks_{21}l\neq 11\hat{k}1\hat{l}} \Pr\{E^2_{js_{22}ks_{21}l}|T_2\}
&= \sum_{js_{22}ks_{21}l\neq 11\hat{k}1\hat{l}} \Pr\{(y_2^n, n_1(j), m_2^n(1,1,s_{22},k), n_2^n(1,1,s_{21},l), w^n) \in A_\epsilon^n(Y_2, M_2, N_1, N_2|W)|T_2\} \\
&\le (2^{n(R_{12}-6\epsilon)}-1)\cdot \Pr\{E^2_{21\hat{k}1\hat{l}}|T_2\} \\
&\quad + (2^{n(L_{22}-6\epsilon)}-1)\cdot \Pr\{E^2_{12\bar{k}1\hat{l}}|T_2\} \\
&\quad + (2^{n(L_{21}-6\epsilon)}-1)\cdot \Pr\{E^2_{11\hat{k}2\bar{l}}|T_2\} \\
&\quad + (2^{n(L_{22}-6\epsilon)}-1)(2^{n(R_{12}-6\epsilon)}-1)\cdot \Pr\{E^2_{22\bar{k}1\hat{l}}|T_2\} \\
&\quad + (2^{n(L_{22}-6\epsilon)}-1)(2^{n(L_{21}-6\epsilon)}-1)\cdot \Pr\{E^2_{12\bar{k}2\bar{l}}|T_2\} \\
&\quad + (2^{n(R_{12}-6\epsilon)}-1)(2^{n(L_{21}-6\epsilon)}-1)\cdot \Pr\{E^2_{21\hat{k}2\bar{l}}|T_2\} \\
&\quad + (2^{n(L_{22}-6\epsilon)}-1)(2^{n(R_{12}-6\epsilon)}-1)(2^{n(L_{21}-6\epsilon)}-1)\cdot \Pr\{E^2_{22\bar{k}2\bar{l}}|T_2\} \\
&\le 2^{n(R_{12}-I(N_1;Y_2,M_2,N_2|W)-6\epsilon+2\epsilon)} \\
&\quad + 2^{n(L_{22}-I(M_2;Y_2,N_1,N_2|W)-6\epsilon+2\epsilon)} \\
&\quad + 2^{n(L_{21}-I(N_2;Y_2,M_2,N_1|W)-6\epsilon+2\epsilon)} \\
&\quad + 2^{n(L_{22}+R_{12}-I(M_2,N_1;Y_2,N_2|W)-6\epsilon+2\epsilon)} \\
&\quad + 2^{n(L_{22}+L_{21}-I(M_2,N_2;Y_2,N_1|W)-6\epsilon+2\epsilon)} \\
&\quad + 2^{n(R_{12}+L_{21}-I(N_1,N_2;Y_2,M_2|W)-6\epsilon+2\epsilon)} \\
&\quad + 2^{n(L_{22}+R_{12}+L_{21}-I(M_2,N_1,N_2;Y_2|W)-6\epsilon+2\epsilon)}
\end{aligned}
$$

If $(R_{11}, R_{12}, R_{21}, R_{22})$ and (L_{21}, L_{22}) are as in the theorem statement, these quantities will tend to zero as the block length $n \to \infty$.

Lemma 5: Let sequences x_1^n, y_1^n and z_1^n be generated independently with each letter distributed i.i.d. according to $p(x)$, $p(y)$ and $p(z)$. If x_1^n and y_1^n are strongly jointly typical according to $p(x)q(y|x)$ ($q(y|x)$ not necessarily equal to $p(y)$) and x_1^n and z_1^n are jointly typical according to $p(x)q(z|x)$ ($q(z|x)$ not necessarily equal to $p(z)$) then with probability $\rightarrow 1$ as $n \rightarrow \infty$, (x_1^n, y_1^n, z_1^n) are jointly typical according to $p(x)q(y|x)q(z|x)$. □

Proof: For each $x' \in \mathcal{X}$, consider the subsequences $(x_{i_1}, x_{i_2}, \ldots, x_{i_M})$ of x_1^n such that $x_{i_1} = x_{i_2} = \cdots = x_{i_M} = x'$, where M denotes the number of occurrences of the letter x' in x_1^n. Then the subsequences $(y_{i_1}, y_{i_2}, \ldots, y_{i_M})$ of y_1^n and $(z_{i_1}, z_{i_2}, \ldots, z_{i_M})$ of z_1^n have distribution near (in the strongly typical sense) $q(y|x = x')$ and $q(z|x = x')$ respectively. By the independence of the choice of these sequences, the joint distribution is near (in the strongly typical sense) $q(y|x = x')q(z|x = x')$ with probability $1 - \epsilon_{x'}$, with $\epsilon_{x'} \rightarrow 0$ as $n \rightarrow \infty$. Since the alphabet is finite, $\prod_{x' \in \mathcal{X}}(1 - \epsilon_{x'}) \rightarrow 1$ as $n \rightarrow \infty$. ■

REFERENCES

[1] T. Berger, "Multiterminal source coding," in *The Information Theory Apporach to Communications*, G. Longo, Ed. New York: Springer-Verlag, 1977.

[2] A. Carleial, "Interference channels," *IEEE Trans. Inform. Theory*, vol. IT-24, no. 1, pp. 60–70, Jan. 1978.

[3] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 439–441, May 1983.

[4] T. Cover and M. Chiang, "Duality between channel capacity and rate distortion," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, 2002.

[5] T. Cover and J. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, 1991.

[6] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic Press, 1981.

[7] FCC. [Online]. Available: http://www.fcc.gov/oet/cognitiveradio/

[8] ——, "FCC ET docket no. 03-108: Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies," FCC, Tech. Rep., 2003.

[9] F. C. C. S. P. T. Force, "Fcc report of the spectrum efficiency working group," FCC, Tech. Rep., 2002.

[10] R. G. Gallagher, *Information Theory and Reliable Communication*. New York: Wiley, 1968, ch. 7.

[11] S. Gel'fand and M. Pinsker, "Coding for channels with random parameters," *Probl. Contr. and Inform. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[12] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inform. Theory*, vol. IT-27, no. 1, pp. 49–60, 1981.

[13] W. D. Horne, "Adaptive spectrum access: Using the full spectrum space." [Online]. Available: http://tprc.org/papers/2003/225/Adaptive_Spectrum_Horne.pdf

[14] T. Hunter, A. Hedayat, M. Janani, and A. Nostratinia, "Coded cooperation with space-time transmission and iterative decoding," in *WNCG Wireless Networking Symposium*, Oct. 2003.

[15] T. Hunter and A. Nostratinia, "Coded cooperation under slow fading, fast fading, and power control," in *Asilomar Conference on Signals, Systems, and Computers*, Nov. 2002.

[16] J.Mitola, "The software radio architecture," *IEEE Commun. Mag.*, vol. 33, no. 5, pp. 26–38, May 1995.

[17] ——, "Cognitive radio," Ph.D. dissertation, Royal Institute of Technology (KTH), 2000.

[18] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 3, Mar. 2004.

[19] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, 2004. [Online]. Available: http://www.nd.edu/ jnl/pubs/it2002.pdf

[20] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[21] J. Mitola, "Cognitive radio for flexible mobile multimedia communications," in *IEEE Mobile Multimedia Conference*, 1999.

[22] ——, "Future of signal processing - cognitive radio," in *IEEE ICASSP*, May 1999, keynote address.

[23] P. Mitran, H. Ochiai, and V. Tarokh, "Space-time diversity enhancements using collaborative communication," Submitted to *IEEE Trans. Inform. Theory*, June 2004.

[24] I. Sason, "On achievable rate regions for the Gaussian interference channel," *IEEE Trans. Inform. Theory*, June 2004.

[25] H. Sato, "The capacity of Gaussian interference channel under strong interference," *IEEE Trans. Inform. Theory*, vol. IT-27, no. 6, Nov. 1981.

[26] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity–part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[27] ——, "User cooperation diversity–part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[28] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," Submitted to *IEEE Trans. Inform. Theory*, July 2004.