On the Capacity of Interference Channels with One Cooperating Transmitter

Ivana Marić, Andrea Goldsmith, Gerhard Kramer and Shlomo Shamai (Shitz)

***Abstract*— Inner and outer bounds are established on the capacity region of two-sender, two-receiver interference channels where one transmitter knows both messages. The transmitter with this extra message knowledge is referred to as being cognitive. The inner bound is based on strategies that generalize prior work, and includes rate-splitting, Gel'fand-Pinsker coding and cooperative transmission. A general outer bound is based on the Nair-El Gamal outer bound for broadcast channels. A simpler bound is presented for the case in which one of the decoders can decode both messages. The bounds are evaluated and compared for Gaussian channels.**

***Index Terms*— Capacity, cognitive radio, cooperation, interference channels.**

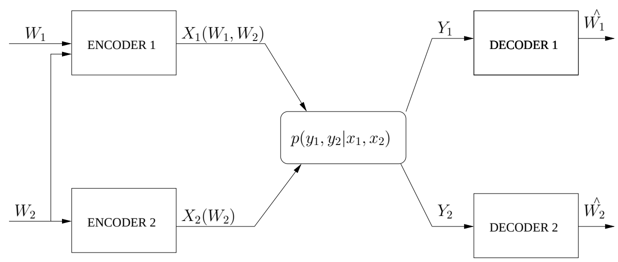

Fig. 1. Interference channel with cooperating encoder.

I. Introduction and Related Work

Two-sender, two-receiver channel models allow for various forms of transmitter cooperation. When senders are unaware of each other's messages, we have the interference channel [1], [2]. In wireless networks, however, the broadcast nature of the wireless medium allows nodes to overhear transmissions and possibly decode parts of other users' messages. An encoder that has such knowledge can use it to improve its own rate as well as the other user's rate. The level of cooperation and the resulting performance improvement will depend on the amount of information the encoders share, as demonstrated in [3].

Channel models with cooperating nodes are of interest also for networks with cognitive users. Cognitive radio [4] technology is aimed at developing smart radios that are both aware of and adaptive to the environment. Such radios can efficiently sense the spectrum, decode information from detected signals and use that knowledge to improve the system performance. This technology motivates information-theoretic models that try to capture the cognitive radio characteristics. In that vein, this paper considers a two-sender, two-receiver channel model in which, somewhat idealistically, we assume that cognitive capabilities allow one user to know the full message of the other encoder, as shown in Fig. 1. The capacity region for this channel is unknown in general, although it has been determined for special cases. No existing coding scheme is known to be uniformly better than other known techniques for all channel characteristics and topologies.

The interference channel with one cooperating encoder was dubbed the *cognitive radio channel* and achievable rates were presented in [3], [5]. The capacity region for the strong interference regime in which both receivers can decode both messages was determined in [6]. The capacity region for the Gaussian case of weak interference was determined in [7] and [8]. The results of [7], [8] were extended to the Gaussian MIMO cognitive radio network and shown to achieve the sum-capacity in [9]. The conclusions of [9] apply to the single-antenna cognitive radio channel as well. A general encoding scheme was recently proposed in [10]. Related work can also be found in [11], [12]. Scaling laws for cognitive networks were analyzed in [13].

In this paper, we present a scheme that generalizes those in [6]-[8]. The scheme is similar to the one in [10]: as in [10] and [3], an encoder uses *rate-splitting* [2] to enable the other receiver to decode part of the interference; the cognitive transmitter cooperates in sending the other user's message to its intended receivers and uses Gel'fand-Pinsker (GP) binning [14] to reduce interference to its own receiver. The key difference of our contribution to the prior work is in the way the binning is performed.

An overview of the encoding scheme is given in the next section. The channel model is described in Section III. Details of the encoding scheme are developed in Section IV. In Section VI, the encoding scheme is adapted for Gaussian channels and its performance is compared to performance of other coding schemes. Our results demonstrate improvements compared to the general scheme of [10]. In Section V, we present two outer bounds for the interference channel with one cooperating encoder. The first bound is based on [15] and the only difference is in the input distribution over which the optimization is performed. We then present an outer bound for the strong interference case that is of the same form as the

I. Marić and A. Goldsmith are with Department of Electrical Engineering, Stanford University, Stanford, CA (email: {ivanam,andrea}@wsl.stanford.edu).

G. Kramer is with Bell Labs, Alcatel-Lucent, Murray Hill, NJ (email: gkr@bell-labs.com).

S. Shamai (Shitz) is with Technion, Haifa, Israel (email: sshlomo@ee.technion.ac.il).

[1]The work by I. Marić and A. Goldsmith was supported in part from the DARPA ITMANET program under grant 1105741-1-TFIND, Stanford's Clean Slate Design for the Internet Program and the ARO under MURI award W911NF-05-1-0246. The work of G. Kramer was partially supported by the Board of Trustees of the University of Illinois Subaward No. 04-217 under NSF Grant No. CCR-0325673 and by ARO Grant W911NF-06-1-0182. The work of S. Shamai was supported by the EU 7th framework program via NEWCOM++, and by the ISRC Consortium.

one in [6, Sec.V] and compare it to the achievable rate region for Gaussian channels.

II. Overview of the Encoding Strategy

The channel model in Fig. 1 has elements of both the interference channel (IC) and the broadcast channel (BC). Encoding techniques developed for either of these channel models are therefore useful for our model. If the message W_2 of encoder 2 was not known at the cognitive encoder, the channel would reduce to the IC. The best achievable rate region for the IC is achieved by rate-splitting [2], [16]: each encoder divides its message into two parts and encodes each of them with a separate codebook. This allows receivers to decode one of the other user's messages and cancel a part of the interference that it would otherwise create. Rate-splitting was applied to the cognitive radio channel model in [3], [10]. In the encoding scheme presented in this paper, rate-splitting is performed at the cognitive encoder.

The cognitive encoder can employ a number of techniques in addition to rate-splitting. For example, to improve the rate for the noncognitive communicating pair, the cognitive encoder can *cooperate* by encoding W_2 to help convey it to the other decoder. On the other hand, any signal carrying information about W_2 creates interference to the cognitive encoder's receiver. This interference is known at the cognitive transmitter and therefore techniques for precoding against the interference, e.g. GP binning [14] or dirty-paper coding (DPC) [17], can be employed. In fact, GP binning is crucial for the cognitive radio channel: together with cooperation, it achieves capacity in certain scenarios [7], [8], [9]. It is not surprising that DPC brings gains in the Gaussian cognitive radio channel: if the non-cognitive encoder is silent, we have a broadcast channel from the cognitive encoder to the two receivers, for which DPC is the optimal strategy [18], [19].

In general, however, there are two differences at the cognitive encoder from the classical GP setting. First, the interference carries useful information for receiver 2. Second, the interference is a *codebook* of some rate and can thus have lower entropy than in the GP setting. As we will see in Sec. II-B, the latter can be exploited to achieve a higher rate.

We note that due to rate-splitting, there is a common part of W_1 decoded at the both receivers and precoded against interference. Since the signal carrying this common message experiences *different* interference at the two receivers, we use the ideas of [20] and [21] that respectively extend [14] and [17] to channels with different states known non-causally to the encoder. For Gaussian channels, DPC is generalized to *carbon-copying* onto dirty paper [21] to adjust to the interference experienced at both receivers.

A. Summary of Techniques and Special Cases

Based on the above discussion, a number of techniques may be applied to exploit the additional knowledge of the cognitive encoder:

- Rate-splitting at encoder 1: to improve R_2 through interference cancelation at decoder 2.

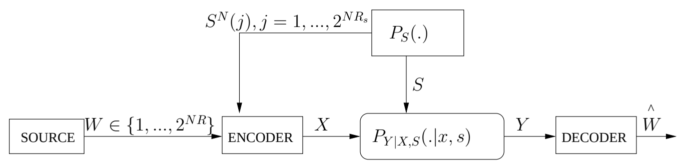

Fig. 2. Communication from the cognitive transmitter to its receiver.

- GP binning and binning against a codebook: to improve R_1 by precoding against interference. This approach also allows decoder 1 to decode message W_2 (or part of it) when R_2 is small, as will be shown in Sec. II-B.
- Carbon-copying onto dirty paper: to improve the rate of the common message sent by the cognitive encoder.
- Cooperation: Encoder 1 contributes to R_2 by encoding W_2.

A general encoding scheme that combines these techniques is described in Section IV. While this general encoding scheme may not achieve capacity in all scenarios, there are a number of special cases in the scheme for which a subset of techniques suffice to achieve capacity, as we now describe:

1) Strong interference: both decoders can decode both messages with no rate penalty, so there is no need for either rate-splitting or binning. Superposition coding achieves capacity [6].
2) Cognitive decoder has to decode both messages: again, there is no need for binning. Rate-splitting and superposition coding achieve capacity [22], [23].
3) Weak interference at receiver 2: there is no need for a common part of message W_1 and hence for rate-splitting. DPC and cooperation achieve capacity for Gaussian channels [7], [8], [9].

B. Rate Improvement due to Binning Against a Codebook

For the communication between the cognitive transmitter and its receiver, a codebook carrying W_2 creates interference. The situation is depicted in Fig. 2, where S plays the role of the codebook of rate R_s interfering with the communication of message W at rate R. While in the GP problem the interference S is generated by a discrete memoryless source, the interference in the cognitive setting is a *codebook* of some rate R_s. The next lemma reflects the fact that when R_s is small, this can be exploited for potential rate gains.

Lemma 1: For the communication situation of Fig. 2, the rate

$$\begin{aligned} R \le \max_{P_{U|S}, f(\cdot)} \min\{&I(X;Y|S), \\ &\max\{I(U,S;Y) - R_s, I(U;Y) - I(U;S)\}\} \end{aligned} \tag{1}$$

is achievable. For $I(S;U,Y) \le R_s \le H(S)$, binning achieves the GP rate given by the second term in (1). For $R_s \le I(S;U,Y)$, superposition coding achieves the rate given by the first term in (1). The two cases are shown in Fig. 3.

Proof: See Appendix B. ■

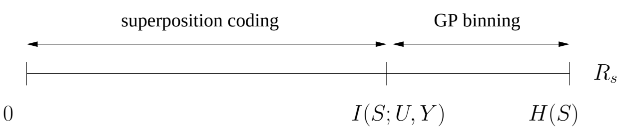

Fig. 3. Binning against a codebook.

Remark 1: Rate (1) can be written as

$$R \leq \max_{P_{U|S}, f(\cdot)} \{I(X,S;Y) \\ - \max\{I(S;Y), \min\{R_s, I(U,Y;S)\}\}. \quad (2)$$

From (1) and (2), we observe that $I(S;U,Y) \leq R_s \leq H(S)$ corresponds to the classical GP setting. Potential rate improvement comes for $R_s \leq I(S;U,Y)$. Interestingly, in this case the receiver decodes both indexes (w,j), thus learning both its message and the interference. A related setting in which both data and the channel state information are communicated to the receiver was analyzed in [24], [25].

In the cognitive setting of Fig. 1, index j carries information about W_2. The implication is that, when R_s is small, receiver 1 will decode a part (or the whole) of W_2 without having encoder 2 rate-split to send common information in the sense of [2], [16].

Recall that, due to rate-splitting, encoder 1 uses two codebooks to send a common and a private index. We denote these respective codebooks as (U_{1c}^N, U_{1a}^N). We can distinguish four cases depending on whether the two codebooks are generated through binning or superposition coding with respect to X_2^N:

1) Binning: both (U_{1c}^N, U_{1a}^N) are binned against the codebook X_2^N of the non-cognitive encoder.
2) Superposition coding: codebooks are superimposed on X_2^N.
3) Binning then superposition coding: U_{1c}^N is binned against X_2^N, and U_{1a}^N is superimposed on (X_2^N, U_{1c}^N).
4) Superposition coding then binning: U_{1c}^N is superimposed on X_2^N; U_{1a}^N is superimposed on U_{1c}^N and binned against X_2^N.

In the last two cases, decoder 1 can decode W_2 due to superposition coding of U_{1a}^N or U_{1c}^N on X_2^N, as shown in Lemma 1. The setting thus corresponds to the *cognitive radio with degraded message sets*. For this channel model, superposition coding achieves capacity [22], [23]. The last two cases can therefore bring no rate improvement. The achievable rate region is the union of two rate regions achieved by binning or superposition coding. We will derive these regions after formally defining the problem in the next section. We remark that in the above encoding scheme, codebook U_{1a}^N is always superimposed on U_{1c}^N. The other encoding choice would be to use binning for U_{1a}^N against the codebook U_{1c}^N.

Finally, we note that encoder 2 also uses rate-splitting and forms two codebooks (X_{2a}^N, X_{2b}^N) using superposition coding. Encoder 1 bins against both codebooks and does not decode a part of W_2. Following Lemma 1, the respective rates R_{2a} and R_{2b} could be chosen such that (U_{1a}^N, U_{1c}^N) are binned against one of the two codebooks, but superimposed on the other. That would facilitate decoding a part of W_2 at receiver 1.

III. Channel Model

Consider a channel with finite input alphabets $\mathcal{X}_1, \mathcal{X}_2$, finite output alphabets $\mathcal{Y}_1, \mathcal{Y}_2$, and a conditional probability distribution $p(y_1, y_2|x_1, x_2)$, where $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$ are channel inputs and $(y_1, y_2) \in \mathcal{Y}_1 \times \mathcal{Y}_2$ are channel outputs. Each encoder t, $t = 1, 2$, wishes to send a message $W_t \in \{1, \ldots, M_t\}$ to decoder t in N channel uses. Message W_2 is also known at encoder 1 (see Fig. 1). The channel is memoryless and time-invariant in the sense that

$$p(y_{1,n}, y_{2,n}|x_1^n, x_2^n, y_1^{n-1}, y_2^{n-1}, \bar{w}) \\ = p_{Y_1,Y_2|X_1,X_2}(y_{1,n}, y_{2,n}|x_{1,n}, x_{2,n}) \quad (3)$$

for all n, where X_1, X_2 and Y_1, Y_2 are random variables representing the respective inputs and outputs, $\bar{w} = [w_1, w_2]$ denotes the messages to be sent, and $x_t^n = \left[x_{t,1}, \quad \ldots, \quad x_{t,n}\right]$. We will follow the convention of dropping subscripts of probability distributions if the arguments of the distributions are lower case versions of the corresponding random variables.

An (M_1, M_2, N, P_e) code has two encoding functions

$$X_1^N = f_1(W_1, W_2) \quad (4)$$
$$X_2^N = f_2(W_2) \quad (5)$$

two decoding functions

$$\hat{W}_t = g_t(Y_t^N) \quad t = 1, 2 \quad (6)$$

and an error probability

$$P_e = \max\{P_{e,1}, P_{e,2}\} \quad (7)$$

where, for $t = 1, 2$, we have

$$P_{e,t} = \sum_{(w_1,w_2)} \frac{1}{M_1 M_2} P\left[g_t(Y_t^N) \neq w_t | (w_1, w_2) \text{ sent}\right]. \quad (8)$$

A rate pair (R_1, R_2) is achievable if, for any $\epsilon > 0$, there is an (M_1, M_2, N, P_e) code such that

$$M_t \geq 2^{NR_t}, \quad t = 1, 2, \text{ and } P_e \leq \epsilon.$$

The capacity region of the interference channel with a cooperating encoder is the closure of the set of all achievable rate pairs (R_1, R_2).

IV. Achievable Rate Region

To obtain an inner bound, we employ rate-splitting. We let

$$R_1 = R_{1a} + R_c \quad (9)$$
$$R_2 = R_{2a} + R_{2b} \quad (10)$$

for nonnegative $R_{1a}, R_c, R_{2a}, R_{2b}$, which we now specify.

In the encoding scheme, encoder 2 uses superposition coding with two codebooks X_{2a}^N, X_{2b}^N. Encoder 1 repeats the steps of encoder 2 and adds binning: it encodes the split message W_1 with two codebooks which are Gel'fand-Pinsker precoded against X_{2a}^N, X_{2b}^N. In particular:

1) Binning against X_{2a}^N, X_{2b}^N is used to create a codebook U_{1c}^N of common rate R_c.

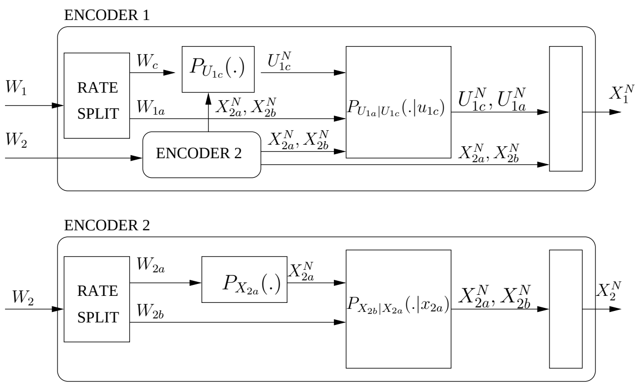

Fig. 4. Encoding structure.

2) Binning against X_{2a}^N, X_{2b}^N conditioned on U_{1c} is used to create a codebook U_{1a}^N with private rate R_{1a}.

The encoding structure is shown in Fig. 4.

We have the following result.

Theorem 1: (joint decoding) Rates (9)-(10) are achievable if

$$R_{1a} \le I(U_{1a};Y_1|U_{1c},Q) - I(U_{1a};X_{2a},X_{2b}|U_{1c},Q) \quad (11)$$
$$R_1 \le I(U_{1c},U_{1a};Y_1|Q) - I(U_{1c},U_{1a};X_{2a},X_{2b}|Q) \quad (12)$$
$$R_2 \le I(X_2;Y_2,U_{1c}|Q) \quad (13)$$
$$R_2 + R_c \le I(X_2,U_{1c};Y_2|Q) \quad (14)$$
$$R_{2b} \le I(X_{2b};Y_2,U_{1c}|X_{2a},Q) \quad (15)$$
$$R_{2b} + R_c \le I(X_{2b},U_{1c};Y_2|X_{2a},Q) \quad (16)$$

for some joint distribution that factors as

$$p(q)p(x_{2a},x_{2b},u_{1c},u_{1a},x_1,x_2|q)p(y_1,y_2|x_1,x_2) \quad (17)$$

and for which the right-hand sides of (11)-(12) are nonnegative. Q is a time-sharing random variable.

Proof: See Appendix A. ■

Theorem 2: (sequential decoding) Rates (9)-(10) are achievable if

$$R_{1a} \le I(U_{1a};Y_1|U_{1c},Q) - I(U_{1a};X_2|U_{1c},Q) \quad (18)$$
$$R_c \le \min\{I(U_{1c};Y_1|Q), I(U_{1c};Y_2,X_{2a}|Q)\} - I(U_{1c};X_2|Q) \quad (19)$$
$$R_{2a} \le I(X_{2a};Y_2|Q) \quad (20)$$
$$R_{2b} \le I(X_{2b};Y_2,U_{1c}|X_{2a},Q) \quad (21)$$

for some joint distribution that factors as

$$p(q)p(x_{2a},x_{2b},u_{1c},u_{1a},x_1,x_2|q)p(y_1,y_2|x_1,x_2)$$

and for which the right-hand sides of (18)-(19) are nonnegative.

Proof: The proof follows similar steps as the proof of Thm. 1 and is omitted. Details can be found in [26]. ■

Remark 2: The rates of Thm. 1 include the rates of Thm. 2.

Remark 3: Thm. 1 includes the rates of the following schemes:

- The scheme of [7, Thm 3.1]: set $X_{2a} = \emptyset, U_{1c} = \emptyset, X_{2b} = (X_2, U)$ and $U_{1a} = V$ so that (11)-(17) become
$$R_2 \le I(X_2,U;Y_2) \quad (22)$$
$$R_1 \le I(V;Y_1) - I(V;X_2,U) \quad (23)$$
for $p(u,x_2)p(v|u,x_2)p(x_1|v)$.
- The scheme of [8, Lemma 4.2]: set $X_{2a} = \emptyset$, $X_{2b} = X_2$, $U_{1a} = \emptyset$, and $R_1 = R_c$, $R_2 = R_{2b}$ so that (11)-(17) become
$$R_2 \le I(X_2;Y_2|U_{1c})$$
$$R_1 \le \min\{I(U_{1c};Y_1), I(U_{1c};Y_2)\} \quad (24)$$
for $p(x_2)p(u_{1c})$. The strategy in [8] considers the case when
$$I(U_{1c};Y_1) \le I(U_{1c};Y_2). \quad (25)$$
.
- Carbon-copy on dirty paper [21]: set $X_{2a} = \emptyset, U_{1a} = \emptyset$.
- For $X_{2a} = \emptyset$, our scheme closely resembles the scheme in [10]. One difference in our scheme is that the two binning steps are not done independently which brings potential improvements. Another difference is in the evaluation of error events.

It is also interesting to compare our scheme to the encoding scheme in [3]. The latter combines rate-splitting at both users, with two-step binning at the cognitive user. Each user sends a private index decoded by its receiver and a common index decoded by both. Again, one difference in our scheme is that two binning steps are not independent. Another is that in our scheme the cognitive encoder cooperates by encoding W_2.

We next exploit Lemma 1.

A. An Achievable Rate Region with Superposition Coding

Consider a joint distribution (17) and rate R_2 that satisfies

$$R_2 \le I(X_2;U_{1c},Y_1) \quad (26)$$
$$R_2 \le I(X_2;U_{1a},Y_1|U_{1c}). \quad (27)$$

From Lemma 1, we know that under conditions (26) and (27), superposition of U_{1c}^N and U_{1a}^N with X_2^N should be used instead of binning. The encoding scheme of the cognitive encoder reduces to rate-splitting and superposition coding. The scheme and the obtained rates reduce to that of [10, Thm. 5] derived for the cognitive radio with degraded message sets, in which the cognitive decoder needs to decode both messages. No rate-splitting at encoder 2 is needed. The resulting achievable rates (R_1, R_2) satisfy

$$R_{1a} \le I(X_1;Y_1|X_2,U_{1c})$$
$$R_1 \le I(X_1;Y_1|X_2)$$
$$R_1 + R_2 \le I(X_1,X_2;Y_1)$$
$$R_c + R_2 \le I(U_{1c},X_2;Y_2) \quad (28)$$

for some joint input distribution $p(x_2,u_{1c},x_1)$. After Fourier-Motzkin elimination [27], we obtain the following region.

Theorem 3: [22]. Achievable rates (R_1, R_2) satisfy

$$\begin{aligned} R_1 &\le I(X_1;Y_1|X_2) \\ R_2 &\le I(U_{1c},X_2;Y_2) \\ R_1+R_2 &\le I(X_1;Y_1|X_2,U_{1c}) + I(U_{1c},X_2;Y_2) \\ R_1+R_2 &\le I(X_1,X_2;Y_1) \end{aligned} \tag{29}$$

for some joint input distribution $p(x_2, u_{1c}, x_1)$.

Remark 4: The above region is the capacity region for the cognitive radio with degraded message sets: the converse follows from [23] where a more general case of confidential messages is analyzed.

V. Outer Bounds

We next derive two capacity outer bounds.

Theorem 4: The union of the set of rate pairs (R_1, R_2) satisfying

$$R_1 \le I(V,U_1;Y_1) \tag{30}$$
$$R_2 \le I(V,U_2;Y_2) \tag{31}$$
$$R_1+R_2 \le \min\{I(V,U_1;Y_1)+I(U_2;Y_2|U_1,V), \tag{32}$$
$$I(U_1;Y_1|U_2,V)+I(V,U_2;Y_2)\} \tag{33}$$

for input distributions $p(v, u_1, u_2, x_1, x_2)$ that factor as

$$p(u_1)p(u_2)p(v|u_1,u_2)p(x_2|u_2)p(x_1|u_1,u_2) \tag{34}$$

is an outer bound to the capacity region.

Proof: See Appendix C. ■

Remark 5: We observe that (30)-(33) is of the same form as the outer bound for the broadcast channel in [15, Sect. 3]. The difference is the factorization of the input distribution.

Remark 6: One can restrict attention to distributions (34) where X_2 is a function of U_2 and X_1 is a function of (U_1, U_2). The bounds (30)-(33) can then be written as

$$R_1 \le I(V,U_1;Y_1) \tag{35}$$
$$R_2 \le I(V,U_2,X_2;Y_2) \tag{36}$$
$$R_1+R_2 \le \min\{I(V,U_1;Y_1)+I(X_1,X_2;Y_2|U_1,V), \tag{37}$$
$$I(X_1;Y_1|X_2,U_2,V)+I(V,U_2,X_2;Y_2)\} \tag{38}$$

From (36) and (38), we obtain the outer bound of [7, Thm. 3.2]:

$$R_1 \le I(X_1;Y_1|X_2) \tag{39}$$
$$R_2 \le I(U,X_2;Y_2) \tag{40}$$
$$R_1+R_2 \le I(X_1;Y_1|X_2,U)+I(U,X_2;Y_2) \tag{41}$$

where we use the notation $U = [U_2, V]$ and also add (39) as it follows by standard methods. The probability distribution factors as

$$p(u,x_1,x_2)p(y_1,y_2|x_1,x_2). \tag{42}$$

Interestingly, (39)-(42) was shown to be tight under weak interference [7, Def. 2.3] and in particular for Gaussian channels with weak interference [7], [8].

The following theorem gives a simple capacity outer bound in strong interference.

Theorem 5: For an interference channel with one cooperating encoder satisfying

$$I(X_1;Y_1|X_2) \le I(X_1;Y_2|X_2) \tag{43}$$

for all input distributions $p(x_1, x_2)$, the set of rate pairs (R_1, R_2) satisfying

$$R_1 \le I(X_1;Y_1|X_2) \tag{44}$$
$$R_1+R_2 \le I(X_1,X_2;Y_2) \tag{45}$$

for all input distributions $p(x_1, x_2)$ is an outer bound to the capacity region.

Proof: See Appendix D. ■

Remark 7: The bound (45) reflects the fact that, because decoder 2 experiences strong interference, as given by (43), it can decode W_1 with no rate penalty.

This bound is evaluated for Gaussian channels in (55).

VI. Gaussian Channel

To illustrate our results more concretely, consider the Gaussian interference channel described by

$$Y_1 = X_1 + aX_2 + Z_1 \tag{46}$$
$$Y_2 = bX_1 + X_2 + Z_2 \tag{47}$$

where $Z_t \sim \mathcal{N}[0,1]$, $E[X_t^2] \le P_t$, $t = 1, 2$ and $\mathcal{N}[0, \sigma^2]$ denotes the normal distribution with zero mean and variance σ^2.

A. Previous Work: Capacity Results

The capacity region for Gaussian cognitive radio channels (46)-(47) was determined for the case of weak interference, i.e., $b \le 1$, in [7], [8]. The optimum coding strategy at the cognitive encoder consists of encoding message W_1 via DPC while treating X_2^N as interference, and superposition coding to help convey W_2 to receiver 2. Receiver 1 does not suffer interference due to DPC. Because the interference is weak, receiver 2 does not attempt to decode the unwanted message.

For $b \ge 1$, the interference at receiver 2 is stronger than in the previous case, and it is plausible to expect that decoding W_1 (or a part of it) may be beneficial. In fact, we observe from (46)-(47) that *after* decoding W_2, decoder 2 has a better observation of X_1 than receiver 1, since $b \ge 1$. Therefore, decoder 2 can decode W_1. However, the conditions under which decoder 2 can decode W_1 also depend on the encoding/decoding approach. In particular, the capacity is known for the strong interference regime in which both decoders can decode both messages. The considered channel becomes a *compound* multiaccess (MAC) channel consisting of two MAC channels, one to each receiver. The strong interference conditions under which this is optimal were determined in [6] leading to the capacity region. These conditions depend on P_1 and P_2. For $P_1 = P_2$ they simplify to $b > 1$ and $a > b$. Encoder 1 uses a superposition code

$$X_1 = X_{1c} + \sqrt{\frac{\bar{\alpha}P_1}{P_2}}X_2 \tag{48}$$

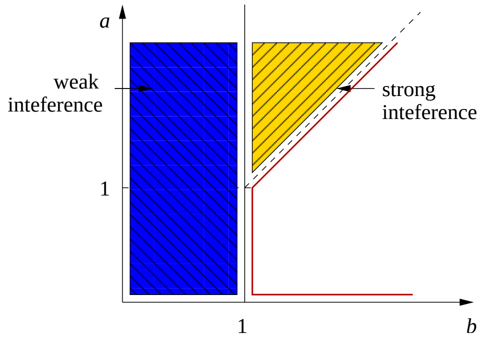

Fig. 5. Values of (a,b) for which the capacity region of the Gaussian channel is known. For $b<1$ the channel has weak interference and the capacity was determined in [7], [8]. Strong interference conditions for case $P_1=P_2$ are $b>1$ and $a>b$. The capacity was determined in [6].

where $X_2 \sim \mathcal{N}[0,P_2]$, $X_{1c} \sim \mathcal{N}[0,\alpha P_1]$ and $0 \leq \alpha \leq 1$. Fig. 5 shows the range of channel gains a and b for which the above capacity results apply. In Fig. 5, we choose $P_1 = P_2$.

When the channel conditions do not allow decoding of both messages at decoder 1, they may still enable decoder 2 to decode W_1. [8, Lemma. 4.2] considers such a case, assuming superposition coding (48). Decoder 2 sequentially decodes: it first decodes W_1, subtracts the part of the received signal that carries it, and decodes W_2 in the interference-free channel. The obtained rates are given by (24). The conditions under which this encoding/decoding procedure is optimal are given by (25).

When these conditions are not satisfied, there may be other interesting scenarios for which the presented techniques lead to capacity results. For example, consider the case when $a=0$ in (46). Then (46)-(47) describe a Z-channel where receiver 1 does not suffer interference and the strong interference conditions of [6] are not satisfied. Suppose that encoder 1 is non-cooperating, aimed to achieve the largest possible R_1. It follows that $\alpha = 0$ in (48). Consider the case for which capacity is unknown, i.e., conditions (25) are not satisfied:

$$\frac{1}{2}\log(1+P_1) \geq \frac{1}{2}\log\left(1+\frac{b^2P_1}{1+P_2}\right). \tag{49}$$

For this case, the achievable rates evaluate to

$$\begin{aligned} R_1 &\leq \frac{1}{2}\log(1+P_1) \\ R_1 + R_2 &\leq \frac{1}{2}\log(1+b^2P_1+P_2). \end{aligned} \tag{50}$$

For Gaussian channels, the outer bound (44)-(45) is given by (55). Due to the above assumption of a non-cooperative cognitive encoder, X_1 and X_2 are independent, and $\rho = 0$ in (55). Therefore, the outer bound (55) coincides with the achievable rates (50), yielding capacity.

As the above discussion illustrates, there are still regimes for which the capacity of Gaussian cognitive channels is unknown. This motivates evaluating the general strategy proposed in Thm. 1. We present such an evaluation next.

B. Numerical Results

We evaluate the rates of Thm. 1 for the special case $X_{2a} = \emptyset$ and $Q = \emptyset$. The rates of Thm.1 reduce to

$$\begin{aligned} R_{1a} &\leq I(U_{1a};Y_1|U_{1c}) - I(U_{1a};X_2|U_{1c}) \\ R_1 &\leq I(U_{1c},U_{1a};Y_1) - I(U_{1c},U_{1a};X_2) \\ R_2 &\leq I(X_2;Y_2,U_{1c}) \\ R_2 + R_c &\leq I(X_2,U_{1c};Y_2). \end{aligned} \tag{51}$$

To simplify (51), we express the conditional entropies in terms of joint entropies, recall that $R_1 = R_c + R_{1a}$, and apply Fourier-Motzkin elimination to obtain

$$\begin{aligned} R_1 &\leq I(U_{1c},U_{1a};Y_1) - I(U_{1c},U_{1a};X_2) \\ R_2 &\leq I(X_2;Y_2,U_{1c}) \\ R_2 &\leq I(X_2,U_{1c};Y_2) \\ R_1 + R_2 &\leq I(X_2,U_{1c};Y_2) + I(U_{1a};Y_1,U_{1c}) \\ &\quad - I(U_{1a};X_2,U_{1c}). \end{aligned} \tag{52}$$

It is also interesting to evaluate the rates of Thm. 2 achieved with sequential decoding for $X_{2a} = \emptyset, Q = \emptyset$. This evaluation results in

$$\begin{aligned} R_{1a} &\leq I(U_{1a};Y_1|U_{1c}) - I(U_{1a};X_2|U_{1c}) \\ R_c &\leq \min\{I(U_{1c};Y_1), I(U_{1c};Y_2)\} - I(U_{1c};X_2) \\ R_2 &\leq I(X_2;Y_2,U_{1c}). \end{aligned} \tag{53}$$

Remark 8: When $I(U_{1c};Y_1) \leq I(U_{1c};Y_2)$, decoder 2 can decode W_1. Thus, there is no need to rate split at encoder 1 and we choose $U_{1a} = \emptyset$. It follows from (52) and (53) that for this case the same rates can be achieved by sequential decoding or by joint decoding.

Remark 9: We observe from (53) that R_c, being a common rate, is bounded by the worst channel, as reflected by the $\min\{I(U_{1c};Y_1), I(U_{1c};Y_2)\}$ term. If $I(U_{1c};Y_1) > I(U_{1c};Y_2)$, transmitting X_{2a} will allow decoder 2 to decode part of W_2 before decoding W_c. It will also serve as an observation when decoding W_c as suggested by the expression $I(U_{1c};Y_2,X_{2a})$ in (19). This will improve the common rate R_c.

We evaluate region (52) for

$$X_2 \sim \mathcal{N}[0,P_2],\ X_{1c} \sim \mathcal{N}[0,\alpha\beta P_1],\ X_{1a} \sim \mathcal{N}[0,\alpha\bar{\beta}P_1]$$

$$\begin{aligned} U_{1c} &= X_{1c} + \lambda_1 X_2 \\ U_{1a} &= X_{1a} + \lambda_2 X_2 \\ X_1 &= X_{1c} + X_{1a} + \sqrt{\frac{\bar{\alpha}P_1}{P_2}}X_2 \end{aligned} \tag{54}$$

where $0 \leq \alpha,\beta \leq 1$ and $0 \leq \lambda_1,\lambda_2$. Parameters α and β determine the amount of power that the cognitive user dedicates for cooperation and for sending the common message, respectively.

We compare the region (52) to the outer bound of Thm. 5 which for Gaussian channels is given by the following corollary:

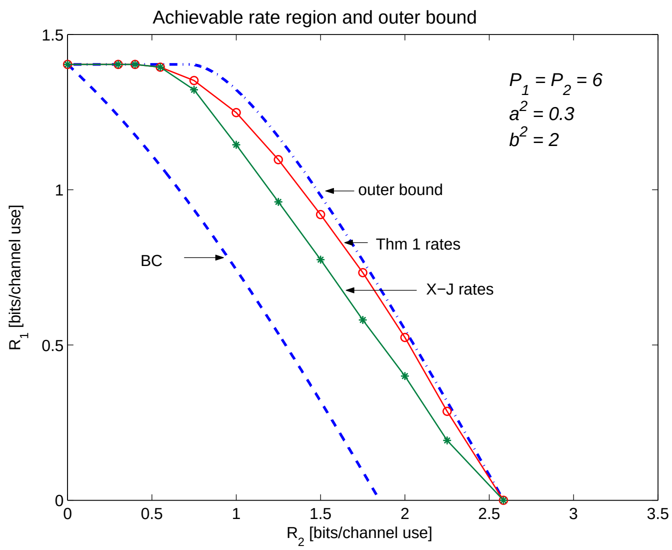

Fig. 6. Achievable rates of Thm. 1 and [10, Thm.5] and outer bound of Cor. 1. Also shown is the capacity region of a BC from the cooperative encoder, i.e. the case $P_2 = 0$.

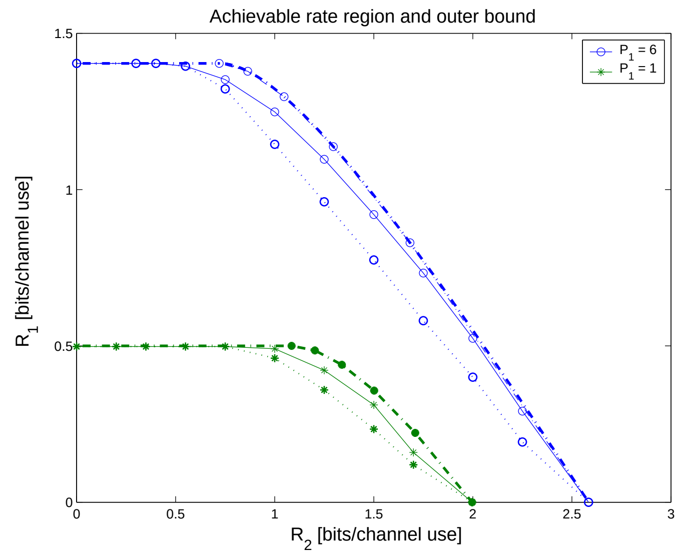

Fig. 8. Impact of reduced power of the cognitive transmitter. Rates achieved with Thm. 1 are shown in solid lines and rates of [10] are shown with dotted lines. Dash-dotted lines show the outer bound (55).

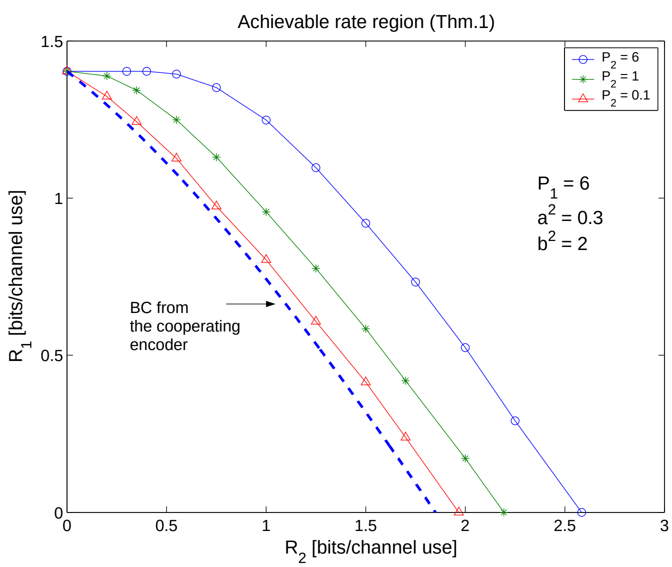

Fig. 7. Achievable rates for different values of P_2.

Corollary 1: When $b \geq 1$, any achievable rate pair (R_1, R_2) satisfies

$$\begin{aligned} R_1 &\leq \frac{1}{2}\log(1 + (1-\rho^2)P_1) \\ R_1 + R_2 &\leq \frac{1}{2}\log(1 + b^2 P_1 + P_2 + 2\rho\sqrt{b^2 P_1 P_2}) \end{aligned} \quad (55)$$

for some ρ, $0 \leq \rho \leq 1$.

Fig. 6 shows the achievable rate region (52) and the outer bound (55) for channel gain values $a^2 = 0.3$, $b^2 = 2$ and equal powers $P_1 = P_2 = 6$. Thm. 1 gives larger rates than those of [10, Thm. 5].

When encoder 2 does not transmit (i.e. $P_2 = 0$), the channel reduces to the BC in which there is only the cooperating encoder communicating to two receivers. Unlike the BC channel rate region, the region for the IC with one cooperating encoder is flat for small values of R_2, reflecting that a cognitive transmitter does not need to cooperate in this regime. It can instead use its full power to precode and transmit W_1 at the single-user rate as if the second user was not present. On the other hand, at $R_1 = 0$ the cooperating encoder fully helps encoder 2, i.e. $\alpha = 0$ and user 2 benefits from the coherent combining gain as indicated by the rate expression

$$R_{2,max} = \frac{1}{2}\log\left(1 + \left(1 + b\sqrt{\frac{P_1}{P_2}}\right)^2 P_2\right).$$

The achievable rates are very close to the outer bound, especially for large values of R_2, in the regime where the cognitive encoder dedicates more of its power to cooperate.

Fig. 7 shows achievable rates for different values of P_2 and fixed P_1. As P_2 decreases, the performance approaches the rate achieved in the BC with only the cooperating encoder transmitting to the two receivers. Since in the BC encoder 2 is not present, the rate region does not depend on P_2 and is given by the dashed line. Fig. 8 shows the effect of reducing power at the cognitive encoder, keeping P_2 constant. This has a strong impact, drastically reducing R_1.

For the Gaussian channel, the rates achieved with sequential encoding (53) can be evaluated for the choice of random variables X_2, X_1, U_{1a}, X_{1a} as in (54). U_{1c} carries a common message and is precoded against interference. Since the two channels from encoder 1 to the two receivers experience different interference, the carbon-copy method of [21] can be used. More details on this approach are presented in [26]. Fig. 9 shows the performance of the two decoding schemes which can differ significantly.

VII. Conclusions and Future Work

We have developed an encoding strategy for the IC with one cooperating encoder that generalizes previously proposed

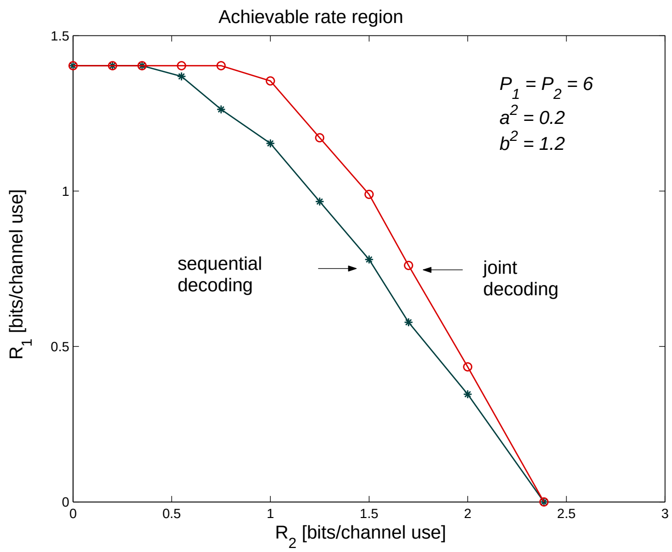

Fig. 9. Comparison of achievable rates with joint and sequential decoding.

strategies. We evaluated its performance and compared it to the performance of other schemes, focusing on the Gaussian channel. A comparison with [3] would be an interesting next step. As explained in Sec. IV, the latter combines rate-splitting at both users, with two-step binning at the cognitive user. One of the differences is that, in our scheme the cognitive encoder cooperates by encoding W_2. However, it is unclear whether our strategy generalizes the scheme in [3], or whether a combination of the two techniques would achieve higher rates. We also compared the proposed scheme to the outer bound that we developed for the strong interference regime. We further developed a new outer bound that extends the Nair-El Gamal broadcast outer bound. Evaluating this bound for specific channels such as Gaussian channels may give capacity results for special cases.

The cognitive radio channel shares some characteristics of both ICs and BCs. Combining encoding strategies developed for either of the two channel models therefore seems a natural approach. However, the optimality of a particular encoding scheme seems to be in part dictated by the channel conditions: for the Gaussian channel in which decoder 2 experiences weak interference, dirty-paper coding achieves capacity. On the other hand, strong interference conditions may allow the cognitive receiver to decode the message not intended for him and therefore DPC against that message is not needed; superposition coding and rate-splitting achieve capacity. An even simpler scheme suffices when both receivers experience strong interference and can both decode the two messages. Neither DPC nor rate-splitting is needed; superposition coding achieves capacity. The encoding scheme presented in this paper is a combination of rate-splitting, GP binning and superposition coding. We believe that this general encoding scheme may achieve capacity for certain special cases related to the channel or specific encoding/decoding constraints. Finding such special cases is a topic of ongoing investigation

APPENDIX A: PROOF OF THEOREM 1

Proof: *(Theorem 1)* **Code construction:** Ignore Q. Choose a distribution $p(x_{2a}, x_{2b}, u_{1c}, u_{1a}, x_1, x_2)$.

- Split the rates as in (9)-(10).
- Generate $2^{NR_{2a}}$ codewords $x_{2a}^N(w_{2a})$, $w_{2a} = 1, \ldots, 2^{NR_{2a}}$, by choosing $x_{2a,n}(w_{2a})$ independently according to $P_{X_{2a}}(\cdot)$.
- For each w_{2a}: Generate $2^{NR_{2b}}$ codewords $x_{2b}^N(w_{2a}, w_{2b})$ using $\prod_{n=1}^N P_{X_{2b}|X_{2a}}(\cdot|x_{2a,n}(w_{2a}))$, $w_{2b} = 1, \ldots, 2^{NR_{2b}}$.
- For each pair (w_{2a}, w_{2b}) : Generate $x_2^N(w_{2a}, w_{2b})$ where x_2 is a deterministic function of (x_{2a}, x_{2b}).
- Generate $2^{N(R_{1c}+R_{1c'})}$ codewords $u_{1c}^N(w_c, b_c)$, $w_c = 1, \ldots, 2^{NR_{1c}}$, $b_c = 1, \ldots, 2^{NR_{1c'}}$ using $P_{U_{1c}}(\cdot)$.
- For each $u_{1c}^N(w_c, b_c)$: Generate $2^{N(R_{1a}+R'_{1a})}$ codewords $u_{1a}^N(w_c, b_c, w_{1a}, b_{1a})$, $w_{1a} = 1, \ldots, 2^{NR_{1a}}$, $b_{1a} = 1, \ldots, 2^{NR'_{1a}}$ using $\prod_{n=1}^N P_{U_{1a}|U_{1c}}(\cdot|u_{1c,n}(w_c, b_c))$.
- For (w_1, w_2) : Generate $x_1^N(w_{2a}, w_{2b}, w_c, b_c, w_{1a}, b_{1a})$ where x_1 is a deterministic function of $(x_{2a}, x_{2b}, u_{1c}, u_{1a}, x_2)$.

Following the proof in [28, Appendix D], it can be shown that it is sufficient to choose respective x_2 and x_1 to be deterministic functions of (x_{2a}, x_{2b}) and $(x_{2a}, x_{2b}, u_{1c}, u_{1a}, x_2)$.

Encoders: Encoder 1:

1) Split the NR_1 bits w_1 into NR_{1a} bits w_{1a} and NR_c bits w_c. Similarly, split the NR_2 bits w_2 into NR_{2a} bits w_{2a} and NR_{2b} bits w_{2b}. We write this as
$$w_1 = (w_{1a}, w_c), \qquad w_2 = (w_{2a}, w_{2b}).$$
2) Try to find a bin index b_c so that $(u_{1c}^N(w_c, b_c), x_{2a}^N(w_{2a}), x_{2b}^N(w_{2a}, w_{2b})) \in T_\epsilon(P_{U_{1c}X_{2a}X_{2b}})$ where $T_\epsilon(P_{XY})$ denotes the jointly ϵ-typical set with respect to P_{XY}, see [29, Sect.8.6]. If no such b_c is found, choose $b_c = 1$.
3) For each (w_c, b_c): Try to find a bin index b_{1a} such that $(u_{1a}^N(w_c, b_c, w_{1a}, b_{1a}), u_{1c}^N(w_c, b_c)), x_{2a}^N(w_{2a}), x_{2b}^N(w_{2a}, w_{2b}) \in T_\epsilon(P_{U_{1a}U_{1c}X_{2a}X_{2b}})$. If one cannot, choose $b_{1a} = 1$.
4) Transmit $x_1^N(w_{2a}, w_{2b}, w_c, b_c, w_{1a}, b_{1a})$.

Encoder 2: Transmit $x_2^N(w_{2a}, w_{2b})$.

Decoders: Decoder 1: Given y_1^N, choose $(\hat{w}_c, \hat{b}_c, \hat{w}_{1a}, \hat{b}_{1a})$ if $(u_{1c}^N(\hat{w}_c, \hat{b}_c), u_{1a}^N(\hat{w}_c, \hat{b}_c, \hat{w}_{1a}, \hat{b}_{1a}), y_1^N) \in T_\epsilon(P_{U_{1c}U_{1a}Y_1})$. If there is more than one such a quadruple, choose one. If there is no such quadruple, choose $(1, 1, 1, 1)$

Decoder 2: Given y_2^N, choose $(\hat{w}'_{2a}, \hat{w}'_c, \hat{b}'_c, \hat{w}'_{2b})$ if $(x_{2a}^N(\hat{w}'_{2a}), u_{1c}^N(\hat{w}'_c, \hat{b}'_c), x_{2b}^N(\hat{w}'_{2a}, \hat{w}'_{2b}), y_2^N) \in T_\epsilon(P_{X_{2a}U_{1c}X_{2b}Y_2})$. If there is more than one such a quadruple, choose one. If there is no such quadruple, choose $(1, 1, 1, 1)$.

Analysis: Suppose $(w_{1a}, w_c, w_{2a}, w_{2b}) = (1, 1, 1, 1)$ was sent. An encoder error occurs if

1) Encoder 1 cannot find a bin index b_c such that $(u_{1c}^N(1, b_c), x_{2a}^N(1), x_{2b}^N(1, 1)) \in T_\epsilon(P_{U_{1c}X_{2a}X_{2b}})$ which

	Error event	Arbitrarily small positive error probability if
E_1	$(\hat{w}_c \neq 1, \hat{w}_{1a} = 1)$	$R_c + R_c' \leq I(U_{1c}, U_{1a}; Y_1)$
E_2	$(\hat{w}_c = 1, \hat{w}_{1a} \neq 1)$	$R_{1a} + R_{1a}' \leq I(U_{1a}; Y_1 \vert U_{1c})$
E_3	$(\hat{w}_c \neq 1, \hat{w}_{1a} \neq 1)$	$R_c + R_c' + R_{1a} + R_{1a}' \leq I(U_{1c}, U_{1a}; Y_1)$
E_1'	$(\hat{w}_{2a}' \neq 1, \hat{w}_{2b}' = 1, \hat{w}_c' = 1)$	$R_{2a} \leq I(X_{2a}, X_{2b}; Y_2, U_{1c})$
E_2'	$(\hat{w}_{2a}' \neq 1, \hat{w}_{2b}' \neq 1, \hat{w}_c' = 1)$	$R_{2a} + R_{2b} \leq I(X_{2a}, X_{2b}; Y_2, U_{1c})$
E_3'	$(\hat{w}_{2a}' \neq 1, \hat{w}_{2b}' = 1, \hat{w}_c' \neq 1)$	$R_{2a} + R_c + R_c' \leq I(X_{2a}, X_{2b}, U_{1c}; Y_2) + I(U_{1c}; X_{2a}, X_{2b})$
E_4'	$(\hat{w}_{2a}' \neq 1, \hat{w}_{2b}' \neq 1, \hat{w}_c' \neq 1)$	$R_{2a} + R_{2b} + R_c + R_c' \leq I(X_{2a}, X_{2b}, U_{1c}; Y_2) + I(U_{1c}; X_{2a}, X_{2b})$
E_5'	$(\hat{w}_{2a}' = 1, \hat{w}_{2b}' \neq 1, \hat{w}_c' = 1)$	$R_{2b} \leq I(X_{2b}; Y_2, U_{1c} \vert X_{2a})$
E_6'	$(\hat{w}_{2a}' = 1, \hat{w}_{2b}' \neq 1, \hat{w}_c' \neq 1)$	$R_{2b} + R_c + R_c' \leq I(X_{2b}, U_{1c}; Y_2 \vert X_{2a}) + I(U_{1c}; X_{2a}, X_{2b})$

TABLE I

ERROR EVENTS IN JOINT DECODING AND CORRESPONDING RATE BOUNDS.

happens with probability

$$\begin{aligned}
P_{e,enc1}^{(1)} &= P[\bigcap_{b_c=1}^{2^{NR_c'}} (U_{1c}^N(1,b_c), x_{2a}^N(1), x_{2b}^N(1,1)) \\
&\qquad \notin T_\epsilon(P_{U_{1c}X_{2a}X_{2b}})] \\
&= \left(1 - P\left[(U_{1c}^N(1,b_c), x_{2a}^N(1), x_{2b}^N(1,1)) \right.\right. \\
&\qquad \left.\left. \in T_\epsilon(P_{U_{1c}X_{2a}X_{2b}})\right]\right)^{2^{NR_c'}} \\
&\leq \left(1 - (1-\epsilon)2^{-N[I(U_{1c};X_{2a},X_{2b})+\delta]}\right)^{2^{NR_c'}} \\
&\leq \exp\left(-(1-\epsilon)2^{N[R_c' - I(U_{1c};X_{2a},X_{2b})-\delta]}\right) \qquad (56)
\end{aligned}$$

where the first inequality follows from [29, Thm.8.6.1] and the second from $(1-x)^m \leq e^{-mx}$. From (56), the error probability $P_{e,enc1}^{(1)}$ can be made arbitrarily small if

$$R_c' > I(U_{1c}; X_{2a}, X_{2b}) + \delta. \qquad (57)$$

2) After it has determined bin index b_c, say $b_c = 1$, encoder 1 cannot find a bin index b_{1a} such that $(u_{1a}^N(1,1,1,b_{1a}), u_{1c}^N(1,1), x_{2a}^N(1), x_{2b}^N(1,1)) \in T_\epsilon(P_{U_{1a}U_{1c}X_{2a}X_{2b}})$ which happens with probability

$$\begin{aligned}
P_{e,enc1}^{(2)} &= P[\bigcap_{b_{1a}=1}^{2^{NR_{1a}'}} (U_{1a}^N(1,1,1,b_{1a}), u_{1c}^N(1,1), x_{2a}^N(1), \\
&\qquad x_{2b}^N(1,1)) \notin T_\epsilon(P_{U_{1a}U_{1c}X_{2a}X_{2b}})] \\
&= \left(1 - P\left[(U_{1a}^N(1,1,1,b_{1a}), u_{1c}^N(1,1), x_{2a}^N(1), \right.\right. \\
&\qquad \left.\left. x_{2b}^N(1,1)) \in T_\epsilon(P_{U_{1a}U_{1c}X_{2a}X_{2b}})\right]\right)^{2^{NR_{1a}'}}. \qquad (58)
\end{aligned}$$

We have

$$\begin{aligned}
&P[(U_{1a}^N(1,1,1,b_{1a}), u_{1c}^N(1,1), x_{2a}^N(1), \\
&x_{2b}^N(1,1) \in T_\epsilon(P_{U_{1a}U_{1c}X_{2a}X_{2b}})] \\
&= \sum_{u_{1a}^N \in T_\epsilon(P_{U_{1a}U_{1c}X_{2a}X_{2b}} | u_{1c}^N x_{2a}^N x_{2b}^N)} P[u_{1a}^N | u_{1c}^N] \\
&\geq (1-\epsilon)2^{-N(H(U_{1a}|U_{1c}X_{2a}X_{2b}) - H(U_{1a}|U_{1c})+\delta)} \\
&= (1-\epsilon)2^{-N(I(U_{1a};X_{2a}X_{2b}|U_{1c})+\delta)} \qquad (59)
\end{aligned}$$

where, as in [30, Handout 2.]), we denote

$$T_\epsilon(P_{XY}|x^N) = \{y^N : (x^N, y^N) \in T_\epsilon(P_{XY})\}.$$

The first inequality follows from the fact that U_{1a}^N was generated according to $P_{U_{1a}|U_{1c}}$ (also directly from [30, Handout 3.]). Employing (59), we can bound (58) as

$$\begin{aligned}
P_{e,enc1}^{(2)} &\leq \left(1 - (1-\epsilon)2^{-N[I(U_{1a};X_{2a},X_{2b}|U_{1c})+\delta]}\right)^{2^{NR_{1a}'}} \\
&\leq \exp\left(-(1-\epsilon)2^{N[R_{1a}' - I(U_{1a};X_{2a},X_{2b}|U_{1c})-\delta]}\right).
\end{aligned} \qquad (60)$$

We need

$$R_{1a}' > I(U_{1a}; X_{2a}, X_{2b} | U_{1c}) + \delta. \qquad (61)$$

Decoder errors: The possible error events at the decoders are shown in the first column of Table I. We next derive the corresponding rate bounds given in the second column of the same table, which guarantee that the error probability of each event can be made small as N gets large. Bounds for E_1, E_1', E_3' are loose. The rest of the rate expressions in Table I yield (11)-(16).

Consider the probability of event E_1 :

$$\begin{aligned}
P[\hat{W}_c \neq 1, \hat{W}_{1a} = 1] &= \sum_{w_c=2}^{2^{NR_c}} \sum_{b_c=1}^{2^{NR_c'}} P[\left(U_{1c}^N(w_c, b_c), \right. \\
&\quad \left. U_{1a}^N(w_c, b_c, 1, 1), Y_1^N\right) \in T_\epsilon(P_{U_{1c}U_{1a}Y_1})] \\
&\leq 2^{-N[I(U_{1c},U_{1a};Y_1) - (R_c + R_c') - \delta]}
\end{aligned} \qquad (62)$$

by [29, Thm.8.6.1] and [30, Handout 1, Thm. 2]. From (62), achieving arbitrarily small error probability of E_1 requires

$$R_c + R_c' < I(U_{1c}, U_{1a}; Y_1). \qquad (63)$$

Similarly, the probability of E_2 is

$$\begin{aligned}
P[\hat{W}_c = 1, \hat{W}_{1a} \neq 1] &= \sum_{w_{1a}=2}^{2^{NR_{1a}}} \sum_{b_{1a}=1}^{2^{NR_{1a}'}} P[\left(U_{1c}^N(1,1), \right. \\
&\quad \left. U_{1a}^N(1,1,w_{1a},b_{1a}), Y_1^N\right) \in T_\epsilon(P_{U_{1c}U_{1a}Y_1})] \\
&\leq 2^{-N[I(U_{1a};Y_1|U_{1c}) - (R_{1a} + R_{1a}') - \delta]}
\end{aligned} \qquad (64)$$

where the inequality follows by [30, Thm., Handout 3]. We need

$$R_{1a} + R_{1a}' < I(U_{1a}; Y_1 | U_{1c}). \qquad (65)$$

The probability of E_3 is, similarly as in (62),

$$P[\hat{W}_c \neq 1, \hat{W}_{1a} \neq 1] = \sum_{w_c=2}^{2^{NR_c}} \sum_{b_c=1}^{2^{NR'_c}} \sum_{w_{1a}=2}^{2^{NR_{1a}}} \sum_{b_{1a}=1}^{2^{NR'_{1a}}} P[\big(U_{1c}^N(w_c, b_c), U_{1a}^N(w_c, b_c, w_{1a}, b_{1a}), Y_1^N\big) \in T_\epsilon(P_{U_{1c}U_{1a}Y_1})] \leq 2^{-N[I(U_{1c}, U_{1a}; Y_1) - (R_c + R'_c + R_{1a} + R'_{1a}) - \delta]} \quad (66)$$

and thus requires

$$R_c + R'_c + R_{1a} + R'_{1a} < I(U_{1c}, U_{1a}; Y_1). \quad (67)$$

We next consider the error events at decoder 2. For E'_1, we have

$$P[\hat{W}'_{2a} \neq 1, \hat{W}'_{2b} = 1, \hat{W}'_c = 1] = \sum_{w_{2a}=2}^{2^{NR_{2a}}} P[\big(U_{1c}^N(1,1), X_{2a}^N(w_{2a}), X_{2b}^N(w_{2a}, 1), Y_2^N\big) \in T_\epsilon(P_{U_{1c}X_{2a}X_{2b}Y_2})]. \quad (68)$$

We further have

$$\begin{aligned} P[(U_{1c}^N(1,1), &X_{2a}^N(w_{2a}), X_{2b}^N(w_{2a},1), Y_2^N) \\ &\in T_\epsilon(P_{U_{1c}X_{2a}X_{2b}Y_2})] \\ &= \sum_{(u_{1c}^N, x_{2a}^N, x_{2b}^N, y_2^N) \in T_\epsilon} P[x_{2a}^N, x_{2b}^N] P[y_2^N u_{1c}^N] \\ &\leq 2^{-N[I(X_{2a}, X_{2b}; Y_2, U_{1c}) - \delta]}. \end{aligned} \quad (69)$$

From (68) and (69), we require

$$R_{2a} < I(X_{2a}, X_{2b}; Y_2, U_c). \quad (70)$$

The probability of event E'_2 is

$$P[\hat{W}'_{2a} \neq 1, \hat{W}'_{2b} \neq 1, \hat{W}'_c = 1] = \sum_{w_{2a}=2}^{2^{NR_{2a}}} \sum_{w_{2b}=2}^{2^{NR_{2b}}} P[\big(U_{1c}^N(1,1), X_{2a}^N(w_{2a}), X_{2b}^N(w_{2a}, w_{2b}), Y_2^N\big) \in T_\epsilon(P_{U_{1c}X_{2a}X_{2b}Y_2})]. \quad (71)$$

Following the same steps as in (69) and using (71) it can be shown that arbitrarily small $P[E'_2]$ requires

$$R_{2a} + R_{2b} < I(X_{2a}, X_{2b}; Y_2, U_{1c}). \quad (72)$$

We next consider E'_3:

$$\begin{aligned} &P[\hat{W}'_{2a} \neq 1, \hat{W}'_{2b} = 1, \hat{W}'_c \neq 1] \\ &= \sum_{w_{2a}=2}^{2^{NR_{2a}}} \sum_{w_c=2}^{2^{NR_c}} \sum_{b_c=1}^{2^{NR'_c}} P[\big(U_{1c}^N(w_c, b_c), X_{2a}^N(w_{2a}), \quad (73) \\ &X_{2b}^N(w_{2a}, 1), Y_2^N\big) \in T_\epsilon(P_{U_{1c}X_{2a}X_{2b}Y_2})]. \quad (74) \end{aligned}$$

We have

$$\begin{aligned} P[(U_{1c}^N(w_c, b_c), &X_{2a}^N(w_{2a}), X_{2b}^N(w_{2a},1), Y_2^N) \\ &\in T_\epsilon(P_{U_{1c}X_{2a}X_{2b}Y_2})] \\ &= \sum_{(u_{1c}^N, x_{2a}^N, x_{2b}^N, y_2^N) \in T_\epsilon} P[x_{2a}^N, x_{2b}^N] P[u_{1c}^N] P[y_2^N] \\ &\leq 2^{-N[I(X_{2a}, X_{2b}, U_{1c}; Y_2) + I(U_{1c}; X_{2a}, X_{2b}) - \delta]}. \end{aligned} \quad (75)$$

From (74) and (75) it follows that

$$R_{2a} + R_c + R'_c < I(X_{2a}, X_{2b}, U_{1c}; Y_2) + I(U_{1c}; X_{2a}, X_{2b}). \quad (76)$$

For E'_4 we use the same approach as in (74) and reuse (75) to obtain

$$R_{2a} + R_{2b} + R_c + R'_c < I(X_{2a}, X_{2b}, U_{1c}; Y_2) + I(U_{1c}; X_{2a}, X_{2b}). \quad (77)$$

We continue by considering error event E'_5:

$$P[\hat{W}'_{2a} = 1, \hat{W}'_{2b} \neq 1, \hat{W}'_c = 1] = \sum_{w_{2b}=2}^{2^{NR_{2b}}} P[\big(U_{1c}^N(1,1), X_{2a}^N(1), X_{2b}^N(1, w_{2b}), Y_2^N\big) \in T_\epsilon(P_{U_{1c}X_{2a}X_{2b}Y_2})]. \quad (78)$$

We have

$$\begin{aligned} P[(U_{1c}^N(1,1), &X_{2a}^N(1), X_{2b}^N(1, w_{2b}), Y_2^N) \in T_\epsilon(P_{U_{1c}X_{2a}, X_{2b}, Y_2})] \\ &= \sum_{(u_{1c}^N, x_{2a}^N, x_{2b}^N, y_2^N) \in T_\epsilon} P[x_{2a}^N, x_{2b}^N] P[u_{1c}^N, y_2^N | x_{2a}^N] \\ &\leq 2^{-N[I(X_{2b}, U_{1c}; Y_2 | X_{2a}) - \delta]}. \end{aligned} \quad (79)$$

From (78) and (79) it follows that

$$R_{2b} < I(X_{2b}, U_{1c}; Y_2 | X_{2a}). \quad (80)$$

For the error event E'_6 we have

$$P[\hat{W}'_{2a} = 1, \hat{W}'_{2b} \neq 1, \hat{W}'_c \neq 1] = \sum_{w_{2b}=2}^{2^{NR_{2b}}} \sum_{w_c=2}^{2^{NR_c}} \sum_{b_c=1}^{2^{NR'_c}} P[\big(U_{1c}^N(w_c, b_c), X_{2a}^N(1), X_{2b}^N(1, w_{2b}), Y_2^N\big) \in T_\epsilon(P_{U_{1c}X_{2a}X_{2b}Y_2})]. \quad (81)$$

Again

$$\begin{aligned} P[(U_{1c}^N(w_c, b_c), &X_{2a}^N(1), X_{2b}^N(1, w_{2b}), Y_2^N) \\ &\in T_\epsilon(P_{U_{1c}X_{2a}X_{2b}Y_2})] \\ &= \sum_{(u_{1c}^N, x_{2a}^N, x_{2b}^N, y_2^N) \in T_\epsilon} P[x_{2a}^N, x_{2b}^N] P[u_{1c}^N] P[y_2^N | x_{2a}^N] \\ &\leq 2^{-N[I(X_{2b}, U_{1c}; Y_2 | X_{2a}) + I(U_{1c}; X_{2a}, X_{2b}) - \delta]}. \end{aligned} \quad (82)$$

From (81) and (82) it follows that

$$R_{2b} + R_c + R'_c < I(X_{2b}, U_{1c}; Y_2 | X_{2a}) + I(U_{1c}; X_{2a}, X_{2b}). \quad (83)$$

■

Appendix B: Proof of Lemma 1

Proof: (Lemma 1) For $I(S; Y, U) \leq R_s \leq H(S)$ use GP coding [14]. The achieved rate is

$$R \leq I(U; Y) - I(U; S). \quad (84)$$

Note that $I(U; Y) - I(U; S) \leq I(X; Y | S)$.

For $R_s < I(S; Y, U)$ proceed as follows.

Code construction: For every codeword $s^N(j), j = 1, \ldots, 2^{NR_s}$ generate 2^{NR} codewords $u^N(w, j)$, $w = 1, \ldots, 2^{NR}$ using $\prod_{n=1}^N P_{U|S}(\cdot | s_n(j))$.

Encoder: Given w and $s^N(j)$, choose $u^N(w, j)$ and transmit $x^N = f^N(u^N(w, j), s^N(j))$.

Decoder: Given y^N, try to find (w,j) such that $(u^N(w,j), s^N(j), y^N) \in T_\epsilon(P_{SUY})$.

Analysis: Suppose $w=1, j=1$ was sent. Error $\{\hat{w} \neq 1\}$ occurs if $\{\hat{w} \neq 1, \hat{j} \neq 1\}$ or $\{\hat{w} \neq 1, \hat{j} = 1\}$. The probability of error is

$$\begin{aligned} P_e = & \sum_{j=2}^{2^{NR_s}} \sum_{w=1}^{2^{NR}} P[(U^N(w,j), S^N(j), Y^N) \in T_\epsilon(P_{SUY})] \\ & + \sum_{w=2}^{2^{NR}} P[(U^N(w,1), s^N(1), Y^N) \in T_\epsilon(P_{SUY})] \\ & \leq 2^{N(R+R_s-I(US;Y)+\delta_1)} + 2^{N(R-I(U;Y|S)+\delta_2)} \end{aligned} \tag{85}$$

where $\delta_1, \delta_2 \to 0$ as $N \to \infty$. From (85) and $I(U;Y|S) = I(X;Y|S)$ it follows that

$$R \leq \min\{I(U,S;Y) - R_s, I(X;Y|S)\}. \tag{86}$$

Note that we could have chosen $u^N = x^N$ in the superposition coding above, so that (86) is

$$R \leq \min\{I(X,S;Y) - R_s, I(X;Y|S)\}. \tag{87}$$

■

Appendix C: Proof of Theorem 4

Proof: (Theorem 4) Consider a code (M_1, M_2, N, P_e) for the interference channel with one cooperating encoder. We first consider the bound (33). Fano's inequality implies that for reliable communication we require

$$\begin{aligned} & N(R_1+R_2) \\ & \leq I(W_1;Y_1^N) + I(W_2;Y_2^N) \\ & \leq^{(a)} I(W_1;Y_1^N|W_2) + I(W_2;Y_2^N) \\ & = \sum_{i=1}^N I(W_1;Y_1^i|W_2,Y_{2,i+1}^N) - I(W_1;Y_1^{i-1}|W_2,Y_{2,i}^N) \\ & \qquad + I(W_2;Y_{2,i}|Y_{2,i+1}^N) \\ & = \sum_{i=1}^N I(W_1;Y_1^i|W_2,Y_{2,i+1}^N) - [I(W_1,Y_{2,i};Y_1^{i-1}|W_2,Y_{2,i+1}^N) \\ & \qquad - I(Y_{2,i};Y_1^{i-1}|W_2,Y_{2,i+1}^N)] + I(W_2;Y_{2,i}|Y_{2,i+1}^N) \\ & =^{(b)} \sum_{i=1}^N I(W_1;Y_{1,i}|W_2,V_i) - I(Y_{2,i};Y_1^{i-1}|W_1,W_2,Y_{2,i+1}^N) \\ & \qquad + I(W_2,Y_1^{i-1};Y_{2,i}|Y_{2,i+1}^N) \\ & \leq \sum_{i=1}^N I(W_1;Y_{1,i}|W_2,V_i) + I(W_2,V_i;Y_{2,i}) \end{aligned} \tag{88}$$

where (a) follows from the independence of W_1, W_2; in (b), we let $Y_{t,i}^j = (Y_{t,i}, \ldots, Y_{t,j})$ and $V_i = [Y_1^{i-1}, Y_{2,i+1}^N]$.

We next consider the bound (31). Fano's inequality implies that reliable communication requires

$$\begin{aligned} NR_2 & \leq I(W_2;Y_2^N) \\ & = \sum_{i=1}^N I(W_2;Y_{2,i}|Y_{2,i+1}^N) \\ & \leq \sum_{i=1}^N I(W_2,Y_1^{i-1},Y_{2,i+1}^N;Y_{2,i}) \\ & = \sum_{i=1}^N I(W_2,V_i;Y_{2,i}). \end{aligned} \tag{89}$$

Note that for (88)-(89) we have used only the independence of W_1 and W_2, and the non-negativity of mutual information. The bounds (30) and (32) thus follow by symmetry.

We introduce random variables $U_{1,i} = W_1$ and $U_{2,i} = W_2$ for all i, to get the bounds in the form (30)-(33). Observe that $U_{1,i}$ and $U_{2,i}$ are independent. Furthermore, due to unidirectional cooperation, the joint probability distribution factors as in (34). ■

Appendix D: Proof of Theorem 5

Proof: (Theorem 5) The bound (44) follows by standard methods. To prove (45), consider (33) and

$$\begin{aligned} I(U_1;Y_1|U_2,V) & \leq I(U_1;Y_1,X_2|U_2,V) \\ & = I(U_1;Y_1|U_2,V,X_2) \\ & \leq I(U_1,X_1;Y_1|U_2,V,X_2) \\ & = I(X_1;Y_1|U_2,V,X_2) \\ & \leq I(X_1;Y_2|U_2,V,X_2) \end{aligned} \tag{90}$$

where the second step follows by the Markov chain (34), and the last step follows by (43). We similarly have

$$I(V,U_2;Y_2) \leq I(U_2,V,X_2;Y_2). \tag{91}$$

Combining inequalities (33), (90) and (91) gives (45). ■

References

[1] H. Sato, "Two user communication channels," *IEEE Trans. Inf. Theory*, vol. 23, no. 3, pp. 295–304, May 1977.

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

[3] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

[4] J. Mitola, *Cognitive Radio Architecture*. John Wiley Sons, Inc., 1991.

[5] N. Devroye, P. Mitran, and V. Tarokh, "Limits on communications in a cognitive radio channel," in *IEEE Comm. Magazine*, vol. 44, no. 6, June 2006, pp. 44–49.

[6] I. Marić, R. D. Yates, and G. Kramer, "Capacity of interference channels with partial transmitter cooperation," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3536–3548, Oct. 2007.

[7] W. Wu, S. Vishwanath, and A. Arapostathis, "Capacity of a class of cognitive radio channels: Interference channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4391–4399, Nov. 2007.

[8] A. Jovičić and P. Viswanath, "Cognitive radio: An information-theoretic perspective," *IEEE Trans. Inf. Theory, submitted, http://www.arxiv.org/pdf/cs.IT/0604107.pdf*, 2006.

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

[10] J. Jiang and Y. Xin, “On the achievable rate regions for interference channels with degraded message sets,” *IEEE Trans. Inf. Theory, submitted*, Apr. 2007.

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

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

[13] N. Devroye, M. Vu, and V. Tarokh, “Achievable rates and scaling laws for cognitive radio channels,” *to appear in EURASIP J. on Wireless Communications and Networking*, Feb. 2008.

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

[15] C. Nair and A. E. Gamal, “An outer bound to the capacity region of the broadcast channel,” *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 350–355, Jan. 2007.

[16] T. Han and K. Kobayashi, “A new achievable rate region for the interference channel,” *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.

[17] M. H. M. Costa, “Writing on dirty paper,” *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[18] H. Weingarten, Y. Steinberg, and S. Shamai(Shitz), “The capacity region of the Gaussian multiple input multiple output broadcast channel,” *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sept. 2006.

[19] S. Vishwanath, N. Jindal, and A. Goldsmith, “Duality, achievable rates and sum rate capacity of Gaussian MIMO broadcast channels,” *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[20] P. Mitran, N. Devroye, and V. Tarokh, “On compound channels with side-information at the transmitter,” *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1745–1755, Apr. 2006.

[21] A. Khisti, U. Erez, A. Lapidoth, and G. W. Wornell, “Carbon copying onto dirty paper,” *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1814–1827, May 2007.

[22] J. Jiang, Y. Xin, and H. Garg, “Interference channels with common information,” *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 171–187, Jan. 2008.

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

[24] T. Cover, Y. Kim, and A. Sutivong, “Channel capacity and state uncertainty reduction for state-dependent channels,” in *Proc. IEEE Int. Symp. Inf. Theory, Nice, France*, June 2007.

[25] ——, “Simultaneous communication of data and state,” *http://arxiv.org/PScache/cs/pdf/0703/0703005v1.pdf*, 2007.

[26] I. Marić, A. Goldsmith, G. Kramer, and S. Shamai(Shitz), “On the capacity of interference channels with a cognitive transmitter,” in *Information Theory and Applications (ITA), UCSD, La Jolla, CA http://ita.ucsd.edu/workshop/07/files/paper/paper431.pdf*, Jan. 2007.

[27] S. Lall, *Advanced Topics in Computation for Control*. Lecture notes, Stanford University, 2004.

[28] F. M. J. Willems, “The discrete memoryless multiple access channel with cribbing encoders,” *IEEE Trans. Inf. Theory*, vol. 31, no. 3, pp. 313–327, May 1985.

[29] T. Cover and J. Thomas, *Elements of Information Theory*. John Wiley Sons, Inc., 1991.

[30] G. Kramer, *Multiple User Information Theory, lecture notes*. unpublished, 2006.

PLACE PHOTO HERE

Ivana Marić received her B.S. degree from the University of Novi Sad, Yugoslavia. She finished her M.S and Ph.D. in the Wireless Network Information Laboratory (WINLAB), Rutgers University in 2000 and 2006, respectively. She was a summer intern at AT&T Research Labs in 1998. She is currently a postdoctoral scholar at Stanford University. Her research focuses on network information theory and wireless communications.

PLACE PHOTO HERE

Andrea Goldsmith is a professor of Electrical Engineering at Stanford University, and was previously an assistant professor of Electrical Engineering at Caltech. She has also held industry positions at Maxim Technologies and at AT&T Bell Laboratories. She is currently on leave from Stanford as co-founder and CTO of Quantenna Communications, Inc. (www.quantenna.com). Her research includes work on capacity of wireless channels and networks, wireless communication and information theory, MIMO and energy-constrained wireless systems, wireless communications for distributed control, and cross-layer design of wireless networks. She is author of the book ”Wireless Communications” and co-author of the book ”MIMO Wireless Communications”, both published by Cambridge University Press. She received the B.S., M.S. and Ph.D. degrees in Electrical Engineering from U.C. Berkeley.

Dr. Goldsmith is a Fellow of the IEEE and of Stanford. She has received serveral awards for her research, including the National Academy of Engineering Gilbreth Lectureship, the Alfred P. Sloan Fellowship, the Stanford Terman Fellowship, the National Science Foundation CAREER CAREER Development Award, and the Office of Naval Research Young Investigator Award. She was also a co-recipient of the 2005 IEEE Communications Society and Information Theory Society joint paper award. She currently serves as associate editor for the IEEE Transactions on Information Theory and as editor for the Journal on Foundations and Trends in Communications and Information Theory and in Networks. She She was previously an editor for the IEEE Transactions on Communications and for the IEEE Wireless Communications Magazine, and has served as guest editor for several IEEE journal and magazine special issues. Dr. Goldsmith is active in committees and conference organization for the IEEE Information Theory and Communications Societies and is an elected member of the Board of Governors for both societies. She is a distinguished lecturer for the IEEE Communications Society, the second vice-president and student committee founder of the IEEE Information Theory Society, and the technical program co-chair for the 2007 IEEE International Symposium on Information Theory.

PLACE PHOTO HERE

Gerhard Kramer received the B.Sc. and M.Sc. degrees in electrical engineering from the University of Manitoba, Winnipeg, MB, Canada, in 1991 and 1992, respectively, and the Dr. sc. techn. (Doktor der Technischen Wissenschaften) degree from the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, in 1998.

From July 1998 to March 2000, he was with Endora Tech AG, Basel, Switzerland, as a communications engineering consultant. Since May 2000 he has been with Bell Laboratories, Murray Hill, NJ, USA. His research has been focused on information theory, communications theory, iterative decoding, and source coding.

Dr. Kramer is currently serving as an Associate Editor for Shannon Theory for the IEEE Transactions on Information Theory. He served as Guest Editor-in-Chief for an IEEE Transactions on Information Theory Special Issue on Relaying and Cooperation in Communication Networks in 2006-2007, and as Publications Editor for the same Transactions during 2004-2005. He is serving as Co-Chair of the Technical Program Committee for the 2008 IEEE International Symposium on Information Theory. Dr. Kramer is a co-recipient of the IEEE Communications Society 2005 Stephen O. Rice Prize paper award, a Bell Labs President’s Gold Award in 2003, and a recipient of an ETH Medal in 1998.

PLACE
PHOTO
HERE

Shlomo Shamai (Shitz) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology, in 1975, 1981 and 1986 respectively.

During 1975-1985 he was with the Communications Research Labs in the capacity of a Senior Research Engineer. Since 1986 he is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, where he is now the William Fondiller Professor of Telecommunications. His research interests encompasses a wide spectrum of topics in information theory and statistical communications.

Dr. Shamai (Shitz) is an IEEE Fellow and a member of the Union Radio Scientifique Internationale (URSI). He is the recipient of the 1999 van der Pol Gold Medal of URSI, and a co-recipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003, and the 2004 joint IT/COM societies paper award, and the 2007 IEEE Information Theory Society Paper Award. He is also the recipient of 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for the Shannon Theory of the IEEE Transactions on Information Theory, and also serves on the Board of Governors of the Information Theory Society.