

On the Capacity Region of the Two-User Interference Channel With a Cognitive Relay

Alex Dytso, Stefano Rini, Natasha Devroye, and Daniela Tuninetti

Abstract—This paper considers a variation of the classical two-user interference channel where the communication of two interfering source-destination pairs is aided by an additional node that has *a priori* knowledge of the messages to be transmitted, which is referred to as the *cognitive relay*. For this interference channel with a cognitive relay (ICCR), novel outer bounds and capacity region characterizations are derived. In particular, for the class of injective semi-deterministic ICCRs, a sum-rate upper bound is derived for the general memoryless ICCR and further tightened for the linear deterministic approximation (LDA) of the Gaussian noise channel at high SNR, which disregards the noise and focuses on the interaction among the users' signals. The capacity region of the symmetric LDA is completely characterized except for the regime of moderately weak interference and weak links from the CR to the destinations. The insights gained from the analysis of the LDA are then translated back to the symmetric Gaussian noise channel (GICCR). For the symmetric GICCR, an approximate characterization (to within a constant gap) of the capacity region is provided for a parameter regime where capacity was previously unknown. The approximately optimal scheme suggests that message cognition at a relay is beneficial for interference management as it enables simultaneous *over the air neutralization* of the interference at both destinations.

Index Terms—Cognitive relay, interference channel, interference neutralization, capacity region, constant gap.

I. INTRODUCTION

IN the last two decades the wireless industry has grown at such a rapid rate that it has started to exhaust much of the precious frequency spectrum [3]. As a response, new technologies have emerged to improve spectrum management. Pico and femto cells technologies [4], [5], for example, use many small base-stations with relatively small coverage areas (as opposed to and in addition to standard macro base-stations with larger

Manuscript received December 3, 2013; revised June 12, 2014; accepted August 21, 2014. Date of publication September 8, 2014; date of current version December 8, 2014. Alex Dytso, Natasha Devroye, and Daniela Tuninetti work was partially funded by NSF under award number 1017436; the contents of this article are solely the responsibility of the author and do not necessarily represent the official views of the NSF. Stefano Rini work was partially funded by National Science Council under MOST 103-2218-E-009-014-MY2. The results in this paper were presented in part at the IEEE Information Theory Workshop, Dublin, Ireland, September 2010 and the IEEE International Conference on Communications, Ottawa, Canada, June 2012. The associate editor coordinating the review of this paper and approving it for publication was M. Bennis.

A. Dytso, N. Devroye, and D. Tuninetti are with the Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, IL 60607 USA (e-mail: odytso2@uic.edu; devroye@uic.edu; danielat@uic.edu).

S. Rini was with Stanford University, Stanford, CA 94305 USA. He is now with the Department of Electrical and Computer Engineering National Chiao Tung University (NCTU), Hsinchu 300, Taiwan (e-mail: stefano@nctu.edu.tw).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TWC.2014.2355826

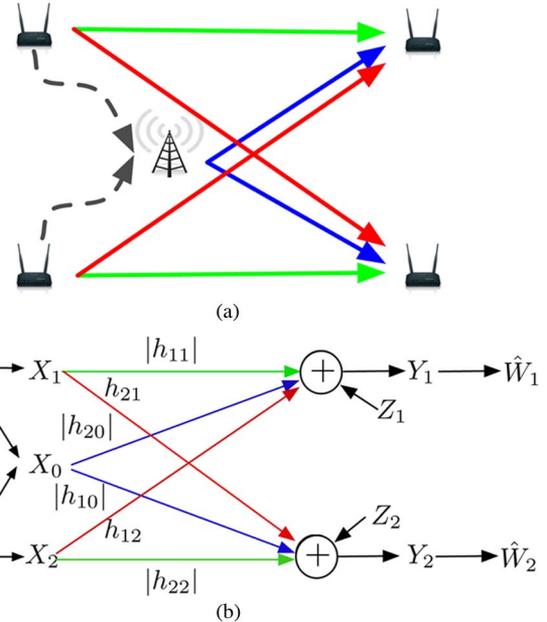


Fig. 1. Network model under investigation. (a) Network aided by a small cell. (b) The two-user GICCR.

coverage areas) to achieve higher throughputs through aggressive spatial reuse. When the small cells have knowledge of the messages to be transmitted by the macro base-stations, they may act as relays to help other devices on the network, as shown in Fig. 1(a), by providing an additional communication path for a message to the desired destination, and by allowing the small cell to better manage/combat the interference. In this work we seek to obtain insights into the performance of such small cell inspired systems. We take an information theoretic approach to the study of such architectures to obtain technology-independent characterizations on the possible performance of the system, measured here in terms of capacity regions.

We study the interference channel with a cognitive relay (ICCR) shown in Fig. 1(b). In this simple model, the ICCR has two independent sources (macro base stations) that send information to their respective destinations by sharing the same channel, i.e., interfering with each other. In addition, a relay (small cell base station) that is non-causally aware of both messages before transmission starts, aids the two sources. Since the relay knows both messages, we term it the *Cognitive Relay* (CR) following [6]. Non-causal message knowledge at the relay may be possible when the relay backhauls to the other transmitting nodes. Alternatively, if no backhauls are possible, the relay may learn the messages of the other transmitters over the air in a causal fashion—in this case the model studied may provide a useful outer bound to the performance of any causal

system. Non-causal message knowledge could also be the result of a failed transmission in systems employing retransmission protocols. Besides its practical motivations, the ICCR is also independently interesting from a theoretical perspective as it generalizes several channel models: the Interference Channel (IC) [7], when the relay is not present, the Broadcast Channel (BC) [8], when both transmitters are not present, and the Cognitive Interference Channel (CIC), when one source is not present [6].

Past Work: In this work we focus on the case where the relay has *non-causal message knowledge* and is *in-band*, that is, the CR shares the same channel as the two source-destination pairs. We note however that significant work exists on models with causal cognition at the relay (where the relay is in-band [9], or out-of-band [10], or out-of-band with noiseless rate-limited links from the CR to the destinations [11], and others variations such as those investigated in [12], [13]). If only one message is available non-causally we obtain a MIMO CIC studied in [14]. Finally, if only portions of the messages are available, the techniques developed in this paper would apply to the portion of the messages known at the relay, but decoding rates would differ (as for example, partial message knowledge does not allow for complete neutralization of the interference) and might more resemble that of an IC.

To the best of our knowledge, the ICCR was first considered in [15], where an achievable rate region was proposed. This rate region was improved upon in Gaussian noise in [16], and again for a general discrete memoryless channel in [17]. The authors of [16] first proposed a sum-rate outer bound for the Gaussian channel. In our conference work [1], we derived the first outer bound for a general memoryless channel, which we further tightened for a class of semi-deterministic channels subject to injectivity conditions in the spirit of [18]. The tightened outer bound was shown in [1] to be capacity for the class of linear deterministic approximation (LDA) of the Gaussian noise channel at high SNR (first introduced in [19]) in the absence of interfering links and in several other special cases. In [20], [21] general inner and outer bounds were obtained, and shown to match for a class of ICCRs with “very strong interference at one destination.” In our conference work [2], the capacity of the symmetric ICCR LDA was shown for almost all channel parameters with a tighter outer bound than that in [1]. The insight from the capacity achieving schemes were used to show capacity to within 3 bits/s/Hz in the Gaussian ICCR (GICCR) without interfering links in [22], which was recently improved upon by [23], where capacity was shown for this channel through the derivation of a new outer bound tailored to the channel without interfering links.

Contributions: The general ICCR is a complex channel model that generalizes many classical channel models for which capacity is open, including the IC and the BC. As such, deriving its capacity region is a challenging and ambitious task. We approach this task by first focusing our attention on the LDA, which highlights the interplay between users’ signals by eliminating the randomness of the noise [19]. The LDA models the Gaussian channel at high SNR, and as such, schemes developed for the LDA can often be translated into “good” achievable strategies for the Gaussian noise channel at any finite SNR, that is, although not optimal in the sense of exactly achieving an

outer bound, they lie within a bounded distance of the outer bounds regardless of the channel parameters. This approach has allowed for progress on long standing open problems; for example, the capacity of the IC [24] and of the CIC [25] are known to within 1 bit/s/Hz. In this work, we first analyze the symmetric LDA by determining its capacity region in almost all parameter regimes (roughly speaking, the case of weak links from the CR to the destinations and of moderately weak interference at a destination from the non-intended source is excluded). We present new achievable techniques that are sum-capacity optimal for the LDA model that were not presented in our conference work [2]. Then, with the insight gained from the study of the LDA, we move back to the symmetric Gaussian noise channel and show capacity to within a constant gap in several parameter regimes that mimic the capacity results for the LDA, which has not appeared in any conference version of our work.

Our central contributions are: (1) Deriving novel (non cut-set) outer bounds for the class of injective semi-deterministic ICCRs; (2) Further tailoring and tightening of the outer bounds for the LDA and GICCR; (3) Deriving optimal achievability schemes in almost all parameter regimes for the symmetric LDA and providing insight into what might be missing in the parameter regime for which we do not have capacity; (4) Deriving the capacity to within a constant gap for the symmetric Gaussian channel in regimes where it was open; (5) Numerically comparing the proposed inner and outer bounds with other achievability schemes.

We note that the central contribution of this work lies in considering a *general* ICCR for the outer bound, rather than models where the assumptions of strong interference [21] or the absence of interfering links [23] significantly simplify the problem. For sake of space, and to convey the key contributions of this work, we focus here only on symmetric channels for achievability results, that is, ICCRs in which the capacity is the same when the role of the sources is swapped. This is done so as to reduce the number of parameters and obtain insightful analytically tractable results. Nevertheless, our outer bounds and achievable scheme apply to general non-symmetric LDAs and GICCRs. We expect that the extension of the presented analysis to the fully general ICCR may follow the same approach used here for the symmetric case, albeit with more involved analytical computations.

Paper Organization: We introduce the channel model in Section II. In Section III we present our novel outer bounds. In Section IV we determine the capacity region for the LDA in almost all parameter regimes. In Section V we derive the capacity to within a constant gap for some parameter regimes of the Gaussian channel which were open, and numerically compare the inner and outer bounds with other relaying schemes. Section VI concludes the paper. Some proofs are found in the Appendix.

Notation: We use the notation convention of [26]: $[n_1 : n_2]$ is the set of integers from n_1 to $n_2 \geq n_1$; $[x]^+ := \max\{0, x\}$ for $x \in \mathbb{R}$; x^n denotes a vector of length n with components (x_1, \dots, x_n) ; lower case x is an outcome of random variable in upper case X which lies in calligraphic case alphabet \mathcal{X} ; $\mathcal{N}(\mu, \sigma^2)$ denotes a proper-complex Gaussian random variable with mean $\mu \in \mathbb{C}$ and variance $\sigma^2 \in \mathbb{R}_+$; $\delta(\cdot)$ denotes the Dirac delta function.

II. CHANNEL MODELS

A. The General Memoryless ICCR

The general two-user memoryless ICCR is characterized by three input alphabets $(\mathcal{X}_0, \mathcal{X}_1, \mathcal{X}_2)$, two output alphabets $(\mathcal{Y}_1, \mathcal{Y}_2)$, and a memoryless transition probability $\mathbb{P}_{Y_1, Y_2 | X_0, X_1, X_2}$. Source i , $i \in [1 : 2]$, encodes the message W_i , assumed independent of everything else and uniformly distributed on $[1 : 2^{nR_i}]$, into a codeword $X_i^n \in \mathcal{X}_i^n$, where $n \in \mathbb{N}$ denotes the codeword length and R_i the rate in bits per channel use. Message W_i is intended for receiver i , $i \in [1 : 2]$, which forms the estimate \widehat{W}_i from channel output $Y_i^n \in \mathcal{Y}_i^n$. The two sources are aided by a cognitive relay that has knowledge of and encodes W_1 and W_2 into the codeword $X_0^n \in \mathcal{X}_0^n$. A non-negative rate pair (R_1, R_2) is said to be achievable if there exists a sequence of encoding functions $X_1^n(W_1), X_2^n(W_2), X_0^n(W_1, W_2)$, and decoding functions $\widehat{W}_1(Y_1^n), \widehat{W}_2(Y_2^n)$, such that the maximum probability of error satisfies $\max_{i \in [1:2]} \mathbb{P}[\widehat{W}_i \neq W_i] \rightarrow 0$ as $n \rightarrow +\infty$. The capacity region is the convex closure of all achievable rate pairs (R_1, R_2) [26].

Since the destinations do not cooperate, the channel capacity only depends on the conditional marginal distributions $\mathbb{P}_{Y_\ell | X_0, X_1, X_2}(y_\ell | x_0, x_1, x_2)$, $\ell \in [1 : 2]$. In other words, all ICCRs that share the same conditional marginal distributions have the same capacity region, as for the BC [26, Lemma 5.1]. Note that the ICCR contains three important channels as special cases: (a) the IC if $X_0 = \emptyset$, (b) the BC if $X_1 = X_2 = \emptyset$, (c) the CIC if either $X_1 = \emptyset$ or $X_2 = \emptyset$.

B. The Injective Semi-Deterministic ICCR

The injective semi-deterministic ICCR was introduced in [18] and corresponds to the special case when the transition probability satisfies

$$\begin{aligned} & \mathbb{P}_{Y_1, Y_2 | X_0, X_1, X_2}(y_1, y_2 | x_0, x_1, x_2) \\ &= \sum_{v_1, v_2} \mathbb{P}_{V_1 | X_1}(v_1 | x_1) \mathbb{P}_{V_2 | X_2}(v_2 | x_2) \\ & \quad \cdot \delta(y_1 - f_1(x_1, x_0, v_2)) \delta(y_2 - f_2(x_2, x_0, v_1)), \end{aligned} \quad (1)$$

for some memoryless transition probabilities $\mathbb{P}_{V_1 | X_1}$ and $\mathbb{P}_{V_2 | X_2}$, and some deterministic functions f_1 and f_2 that are injective when (X_1, X_0) and (X_2, X_0) , respectively, are held fixed, which implies that for all $\mathbb{P}_{X_0, X_1, X_2}$ one has $H(Y_1 | X_1, X_0) = H(V_2 | X_1, X_0) = H(V_2 | X_0)$, and similarly for the other source. Injective semi-deterministic channels are important because approximate capacity results for this class of channels are available while those of their more general counterpart are still open. For example, the injective deterministic IC (where $\mathbb{P}_{V_1 | X_1}$ and $\mathbb{P}_{V_2 | X_2}$ are noiseless) was completely solved in [27] and the injective semi-deterministic IC capacity was characterized to within a constant gap in [18]. Intuitively, it is easier to characterize the capacity of an injective semi-deterministic IC, compared to the general IC, as one knows exactly what the interference signals are through the random variables V_1 and V_2 [18]. We also note that the important

Gaussian channel is a special case of the injective semi-deterministic model. For continuous alphabets, the summations in (1) must be replaced with integrals.

C. The GICCR

The complex-valued single-antenna power-constrained GICCR in a standard form [21] is shown in Fig. 1(b) and is described by the input-output relationships

$$Y_1 = |h_{11}|X_1 + |h_{10}|X_0 + h_{12}X_2 + Z_1, \quad (2a)$$

$$Y_2 = h_{21}X_1 + |h_{20}|X_0 + |h_{22}|X_2 + Z_2, \quad (2b)$$

where, without loss of generality, the input $X_i \in \mathbb{C}$ is subject to power constraint $\mathbb{E}[|X_i|^2] \leq 1$, $i \in [0 : 2]$, the noise $Z_j \sim \mathcal{N}(0, 1)$, $j \in [1 : 2]$, and the channel gains h_{ij} , $i \in [1 : 2]$, $j \in [0 : 2]$, are complex-valued, fixed and known to all nodes. Without loss of generality, some channel gains can be taken to be real-valued and non-negative [21, Appendix M]. The Gaussian GICCR is a special case of the injective semi-deterministic ICCR in (1), where $V_1 := h_{21}X_1 + Z_2$, $V_2 := h_{12}X_2 + Z_1$, and f_1 and f_2 are complex-valued linear combinations.

The capacity region of the GICCR is open. Progress towards understanding its fundamental limits is possible by providing an approximate characterization of its capacity as pioneered in [19]. The capacity is said to be known *to within GAP bits* if one can show an inner bound region \mathcal{I} and an outer bound region \mathcal{O} such that $(R_1, R_2) \in \mathcal{O} \implies ([R_1 - \text{GAP}]^+, [R_2 - \text{GAP}]^+) \in \mathcal{I}$. The GAP upper bounds the worst-case distance between the inner and outer bounds [24].

D. The LDA

The *linear deterministic approximation* (LDA) of the GICCR is a model that captures the behavior of the GICCR in (2) at high SNR. The LDA is a fully deterministic model described by [19]

$$Y_1 = \mathbf{S}^{m-n_{11}} X_1 \oplus \mathbf{S}^{m-n_{10}} X_0 \oplus \mathbf{S}^{m-n_{12}} X_2, \quad (3a)$$

$$Y_2 = \mathbf{S}^{m-n_{21}} X_1 \oplus \mathbf{S}^{m-n_{20}} X_0 \oplus \mathbf{S}^{m-n_{22}} X_2, \quad (3b)$$

where \mathbf{S} is the binary down-shift matrix of dimension $m := \max\{n_{11}, n_{12}, n_{10}, n_{21}, n_{22}, n_{20}\}$, for $\{n_{ij} \in \mathbb{N}, i \in [1 : 2], j \in [0 : 2]\}$, all inputs and outputs are binary column vectors of dimension m , and where the symbol \oplus denotes the component-wise modulo-2 addition of the binary vectors. The LDA is a special case of injective semi-deterministic ICCR in (1) where $V_1 := \mathbf{S}^{m-n_{21}} X_1$, $V_2 := \mathbf{S}^{m-n_{12}} X_2$, and f_1 and f_2 are modulo-2 additions. The LDA in (3) may be related to the GICCR in (2) by taking $n_{ij} = \lceil \log(1 + |h_{ij}|^2) \rceil$ [24]. The capacity of the LDA often gives insight into strategies that are optimal to within a constant gap for the GICCR [28].

III. OUTER BOUNDS

We start off stating a known outer bound for the general memoryless ICCR, and then deriving new outer bounds for the injective semi-deterministic ICCR, the LDA and the GICCR.

Theorem 1. (Outer bound to the capacity of the general memoryless ICCR [21, Thm. III.1]): If (R_1, R_2) lies in the capacity region of the general memoryless ICCR, then

$$R_1 \leq I(Y_1; X_1, X_0|Q, X_2), \quad (4a)$$

$$R_2 \leq I(Y_2; X_2, X_0|Q, X_1), \quad (4b)$$

$$R_1 + R_2 \leq I(Y_2; X_1, X_2, X_0|Q) + I(Y_1; X_1, X_0|Q, \bar{Y}_2, X_2), \quad (4c)$$

$$R_1 + R_2 \leq I(Y_1; X_1, X_2, X_0|Q) + I(Y_2; X_2, X_0|Q, \bar{Y}_1, X_1), \quad (4d)$$

for some input distribution that factors as $\mathbb{P}_{Q, X_1, X_2, X_0} = \mathbb{P}_Q \mathbb{P}_{X_1|Q} \mathbb{P}_{X_2|Q} \mathbb{P}_{X_0|X_1, X_2, Q}$, where \bar{Y}_1 and \bar{Y}_2 have the same conditional marginal distributions of the channel outputs Y_1 and Y_2 given the inputs (X_1, X_2, X_0) , respectively, but are otherwise arbitrarily jointly distributed. \square

The region in Theorem 1 is not the tightest in general. For example, [21] reports other outer bounds that can actually be used to prove capacity in some regimes. However, these other outer bounds depend on auxiliary random variables for which no cardinality bound is known on the corresponding alphabets. The advantage of Theorem 1 is that it only contains random variables defined in the channel model (with the exception of the time-sharing random variable Q) and it is therefore in principle computable. Note that the correlation among \bar{Y}_1 and \bar{Y}_2 in Theorem 1 may be chosen to tighten the bound since the capacity region of the ICCR is only a function of the output conditional marginal distributions, as for the BC [26, Lemma 5.1] and the CIC [21].

Theorem 1 reduces to: (a) the capacity region of a deterministic BC when $X_1 = X_2 = \emptyset$ [26], and (b) the capacity region of a deterministic CIC when either $X_2 = \emptyset$ or $X_1 = \emptyset$ [29]. However, it does not reduce to the capacity region of the class of fully deterministic IC when $X_0 = \emptyset$ [27]. Hence, in the following we develop additional rate bounds that reduce to the bounds for the injective semi-deterministic IC developed in [18], which includes the fully deterministic IC studied in [27], when $X_0 = \emptyset$.

A. Novel Outer Bounds for the Injective Semi-Deterministic ICCR

The outer bound of Theorem 1 may be tightened for the injective semi-deterministic ICCR defined in (1) as follows, whose proof can be found in the Appendix A.

Theorem 2: If (R_1, R_2) lies in the capacity region of the injective semi-deterministic ICCR, then in addition to the bounds in (4), the following must hold

$$R_1 + R_2 \leq H(Y_1|\tilde{V}_1, Q) - H(\tilde{V}_2|X_2) + H(Y_2|\tilde{V}_2, Q) - H(\tilde{V}_1|X_1) + \text{MLP}_1, \quad (5a)$$

$$2R_1 + R_2 \leq -H(\tilde{V}_1|X_1) - 2H(V_2|X_2) + H(Y_1|Q) + H(Y_1|\tilde{V}_1, X_2, Q) + H(Y_2|\tilde{V}_2, Q) + \text{MLP}_1, \quad (5b)$$

$$R_1 + 2R_2 \leq -H(\tilde{V}_2|X_2) - 2H(V_1|X_1) + H(Y_2|Q) + H(Y_2|\tilde{V}_2, X_1, Q) + H(Y_1|\tilde{V}_1, Q) + \text{MLP}_1, \quad (5c)$$

where the multi-letter portion (MLP) MLP_1 is given by

$$\text{MLP}_1 := \sup_{n \in \mathbb{N}} \frac{1}{n} (I(V_1^n; X_0^n|W_2) + I(V_2^n; X_0^n|W_1)), \quad (5d)$$

and where the random variables \tilde{V}_1 and \tilde{V}_2 are conditionally independent copies of V_1 and V_2 , respectively, that is, they are jointly distributed with (Q, X_1, X_2, X_0) as

$$\mathbb{P}_{\tilde{V}_1, \tilde{V}_2|Q, X_1, X_2, X_0}(v_1, v_2|q, x_1, x_2, x_0) = \mathbb{P}_{V_1|X_1}(v_1|x_1) \mathbb{P}_{V_2|X_2}(v_2|x_2), \quad (5e)$$

where $\mathbb{P}_{V_1|X_1}$ and $\mathbb{P}_{V_2|X_2}$ are part of the channel model definition in (1). \square

The auxiliary random variables \tilde{V}_1 and \tilde{V}_2 are provided as “genie side information” at receivers 1 and 2, respectively, as a mathematical tool to enable the derivation of “single letter” outer bounds; they are identical to those used in [18], and thus with this choice Theorem 2 reduces to [18, Theorem 1], which is tight for the LDA [28] and is optimal to within 1 bit for the Gaussian IC [24], when $X_0 = \emptyset$. Theorem 2 is however not in the desirable “single-letter” format, as it contains the MLP in (5d). We discuss how to “single-letterize” the MLP in (5d) for the LDA and the GICCR in the rest of the section.

Note that the step of tightening the bound used in the proof of Theorem 2 (i.e., conditioning on the interference function V_j rather than on the interfering codeword $X_j, j \in [1, 2]$) highlights a stumbling block in deriving outer bounds for general IC and BC: in general we do not know the exact form of the interfering signal(s) at a receiver for any possible input distribution. Assuming that the channel is deterministic and in a certain way invertible, allows one to exactly determine the interference. Notice that “conditioning” on the interference functions V_1 or V_2 may be interpreted as if the interference has been removed without necessarily decoding the corresponding messages. On the other hand, conditioning on X_1 or X_2 may be interpreted as if the message carried by X_1 or X_2 were known through decoding.

B. Outer Bounds for the LDA

For a discrete-valued channel (for which the entropy is non-negative), one may turn the MLP in (5d) into a single-letter expression as

$$\begin{aligned} \text{MLP}_1 &\leq \sup_{n \in \mathbb{N}} \frac{1}{n} (\min \{H(V_2^n), H(X_0^n)\} \\ &\quad + \min \{H(V_1^n), H(X_0^n)\}) \\ &\leq \min \{H(V_2), H(X_0)\} + \min \{H(V_1), H(X_0)\}. \end{aligned} \quad (6)$$

For the LDA we next provide a tighter bound than that in (6). The LDA belongs to a special class of injective deterministic channels whose outputs are described by

$$Y_1 = f_1(X_1, q_1(X_0), V_2), \quad V_2 = g_2(X_2), \quad (7a)$$

$$Y_2 = f_2(X_1, q_2(X_0), V_1), \quad V_1 = g_1(X_1), \quad (7b)$$

where $q_1, q_2, g_1, g_2, f_1, f_2$ are deterministic functions. The difference between (7) and (1) is that in the former the output at receiver $i \in [1 : 2]$ depends on a function $q_i(X_0)$ rather than on X_0 ; this distinction is important when the function $q_i(\cdot)$ is not a

bijection, as it may be the case in the LDA. We further require the function f_1 to be injective when its first two arguments are known, that is, $H(Y_1|q_1(X_0), X_1) = H(V_2|q_1(X_0))$, and analogously for f_2 .

For the LDA, Theorem 2 may be tightened as follows, whose proof may be found in Appendix B.

Theorem 3: For the LDA the term MLP_1 in (5d) can be tightened by using instead

$$MLP_2 := \min\{n_{20}, n_{21}\} + \min\{n_{10}, n_{12}\}, \quad (8)$$

and the resulting region in (5) is exhausted by considering i.i.d. Bernoulli(1/2) inputs. \square

C. Outer Bounds for the GICCR

The proof of Theorem 3 inspired the following bound on the MLP for the Gaussian noise channel. For the GICCR, Theorem 2 may be tightened as follows, whose proof may be found in Appendix C.

Theorem 4: For the GICCR the term MLP_1 in (5d) can be tightened by using instead

$$MLP_2 := \log(1 + \min\{|h_{12}|^2, |h_{10}|^2\}) + \log(1 + \min\{|h_{21}|^2, |h_{20}|^2\}) + 2\log(2), \quad (9)$$

and the resulting region in (5) is exhausted by considering jointly Gaussian inputs. \square

IV. CAPACITY FOR THE SYMMETRIC LDA

In this section we propose achievable schemes that match the outer bound in Theorem 3 for almost all channel parameters, where channel gains and the rates are parametrized as

$$n_{11} = n_{22} = n_S > 0, \quad (10a)$$

$$n_{12} = n_{21} = n_I = \alpha n_S, \quad \alpha \geq 0, \quad (10b)$$

$$n_{10} = n_{20} = n_C = \beta n_S, \quad \beta \geq 0, \quad (10c)$$

$$R_i = r_i n_S, \quad r_i \geq 0, \quad i \in [1:2]. \quad (10d)$$

The focus on the symmetric case is not for lack of generality of the developed theory but for simplicity of exposition (the symmetric model is specified by three parameters rather than six).

Under the symmetric condition in (10), the outer bound in Theorem 3 simplifies to

$$r_1 \leq \max\{1, \beta\}, \quad r_2 \leq \max\{1, \beta\}, \quad (11a)$$

$$r_1 + r_2 \leq [1 - \max\{\alpha, \beta\}]^+ + \beta + \max\{1, \alpha\}, \quad (11b)$$

$$r_1 + r_2 \leq \max\{1, \beta\}, \quad \text{apply for } \alpha = 1 \text{ only}, \quad (11c)$$

$$r_1 + r_2 \leq 2 \max\{1 - \alpha, \alpha, \beta\} + 2 \min\{\alpha, \beta\}, \quad (11d)$$

$$2r_1 + r_2 \leq \max\{1, \beta, \alpha\} + \max\{1 - \alpha, \alpha, \beta\} + \max\{1 - \alpha, \beta\} + 2 \min\{\alpha, \beta\}, \quad (11e)$$

$$r_1 + 2r_2 \leq \max\{1, \beta, \alpha\} + \max\{1 - \alpha, \alpha, \beta\} + \max\{1 - \alpha, \beta\} + 2 \min\{\alpha, \beta\}. \quad (11f)$$

The outer bound in (11) naturally leads to the division of the channel parameter space (α, β) in (10) into the six regimes illustrated in Fig. 2 based on the different values of the max/min terms in (11).

Theorem 5: For the symmetric LDA, capacity is known for the following regimes (see Fig. 2): Regimes I to V ($1 \leq$

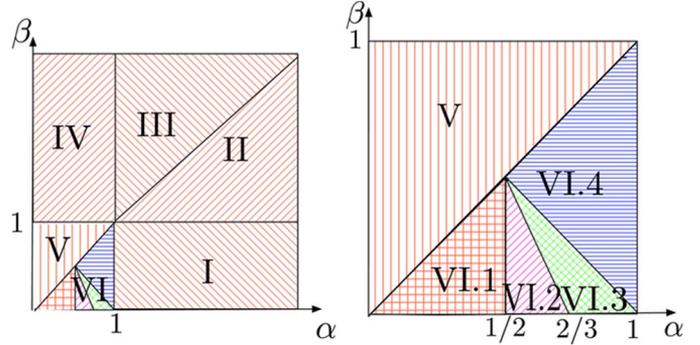


Fig. 2. Parameter regimes for the LDA and the GICCR at high SNR. Regimes I to VI.1: capacity is known for the LDA and to within a constant gap for the GICCR. Regimes VI.2 to VI.4: capacity is open for both models (but sum-capacity is known in some cases for the LDA).

$\max\{\alpha, \beta\}$) and Regime VI.1 ($\beta \leq \alpha \leq \frac{1}{2}$). For the remaining regimes the sum-capacity is known for $4\alpha - 3 \leq \beta \leq 3\alpha - 2$, $\frac{2}{3} \leq \alpha \leq 1$, which in Regime VI.4 implies that the whole capacity region is known. \square

Proof: We now prove Theorem 5 for different cases and regimes.

a) *Case $n_S > 0$ and $\alpha = 1$ (line $\alpha = 1$ in Fig. 2):* The outer bound in (11) when $\alpha = 1$ is simply the triangle formed by $r_1 + r_2 \leq \max\{1, \beta\}$, from (11c) only, which is trivially achieved by time division between the cases where one source is silent and the cognitive relay fully helps the other source. In particular, to prove capacity, it suffices to show the achievability of $(r_1, r_2) = (\max\{1, \beta\}, 0)$, which can be attained as follows. Case 1) If $1 \geq \beta$: $X_2 = X_0 = 0$, i.e., the information bits for destination 1 are carried by X_1 . The achievable rate is $r_1 = 1$, $r_2 = 0$. Case 2) If $1 < \beta$: $X_2 = X_1 = 0$, i.e., the information bits for destination 1 are carried by X_0 . The achievable rate is $r_1 = \beta$, $r_2 = 0$. The other corner point $(r_1, r_2) = (0, \max\{1, \beta\})$ is achieved by swapping the role of the users. By time-sharing, the whole dominant face of the outer bound region is achievable, thus proving capacity.

Remark 1: The points $(r_1, r_2) = (\max\{1, \beta\}, 0)$ and $(r_1, r_2) = (0, \max\{1, \beta\})$ are always corner points of the capacity region, but are not the dominant ones in general.

b) *Case $n_S > 0$, $\alpha \neq 1$ and $\max\{\alpha, \beta\} > 1$ (Regimes I to IV in Fig. 2):* When (11e), (11f), and (11d) are redundant, that is, for $\max\{\alpha, \beta\} > 1$ (all but Regimes V and VI in Fig. 2), the region in (11) simplifies to the pentagon region

$$r_1 \leq \max\{1, \beta\}, \quad r_2 \leq \max\{1, \beta\}, \quad (12a)$$

$$r_1 + r_2 \leq \beta + \max\{1, \alpha\}. \quad (12b)$$

We show achievability with two different strategies.

b.1) *Regimes II, III, and IV in Fig. 2:* For $\beta \geq 1$, to prove capacity, it suffices to show the achievability of the corner point $(r_1, r_2) = (\beta, \min\{\beta, \max\{1, \alpha\}\})$. The other corner point $(r_1, r_2) = (\min\{\beta, \max\{1, \alpha\}\}, \beta)$ is achieved by swapping the role of the users. By time-sharing between the corner points, the whole dominant face of the outer bound region is achievable, thus proving capacity. Let U_0, U_{1p}, U_{2p} be independent vectors. Consider the following strategy

$$X_1 = \mathbf{S}^{m-n_C} U_{1p}, \quad X_2 = \mathbf{S}^{m-n_C} U_{2p}, \quad (13a)$$

$$X_0 = \mathbf{S}^{m-n_1} (U_{1p} + U_{2p}) + U_0, \quad (13b)$$

where X_0 is so as to *neutralize over the air* the interference at the destinations. The received signal at destination 1 is

$$\begin{aligned} Y_1 &= \underbrace{(\mathbf{S}^{m-n_S} \mathbf{S}^{m-n_C} + \mathbf{S}^{m-n_C} \mathbf{S}^{m-n_1})}_{\neq 0 \text{ if } n_S \neq n_1 \iff \alpha \neq 1} U_{1p} \\ &\quad + \underbrace{(\mathbf{S}^{m-n_C} \mathbf{S}^{m-n_1} + \mathbf{S}^{m-n_1} \mathbf{S}^{m-n_C})}_{=0, \text{ interference neutralized}} U_{2p} \\ &\quad + \mathbf{S}^{m-n_C} U_0 \\ &= \mathbf{S}^{m-n_C} ((\mathbf{S}^{m-n_S} + \mathbf{S}^{m-n_1}) U_{1p} + U_0), \end{aligned} \quad (14)$$

and similarly for the received signal at destination 2. Let the top $m - \max\{n_S, n_1\}$ bits of U_0 , which are received clean on top of the bits of $(\mathbf{S}^{m-n_S} + \mathbf{S}^{m-n_1}) U_{ip}$ at each destination $i \in [1 : 2]$, be i.i.d. Bernoulli(1/2) bits dedicated to user 1 and the rest of U_0 be set to zero. U_{1p} and U_{2p} are i.i.d. Bernoulli(1/2) bits. Hence, $R_1 = n_C$ and $R_2 = \min\{n_C, \max\{n_S, n_1\}\}$ (note that the rates cannot be larger than n_C because of the multiplication by \mathbf{S}^{m-n_C} of the signals at each receiver (14)). By normalizing the rates by n_S the claim follows.

Notice that by setting $U_0 = 0$ and by using the “neutralize over the air” technique in (13), it is always possible to achieve the normalized private rates $r_1 = r_2 = \min\{\beta, \max\{1, \alpha\}\}$, where we use the qualifier “private” to follow the nomenclature convention for the classical IC: a message that is decoded only at an intended destination is referred to as a “private message.” A message also decoded at a non intended destination is referred to as a “common message.” In Regimes II to IV, a “common message” for user 1 is sent by the CR through the top bits of U_0 whenever $\beta > \max\{1, \alpha\}$.

Remark 2: The achievability in this case can also be used to show the achievability of the region in (15) when $n_S = 0$, in which case the region in (11) simplifies to (here we do not normalize by the strength of the direct link as this link does not exist)

$$R_1 \leq n_C, \quad R_2 \leq n_C, \quad R_1 + R_2 \leq n_C + n_1. \quad (15)$$

Clearly this is a special case of $1 < \max\{\alpha, \beta\}$ since $0 = n_S \leq \max\{n_1, n_C\}$.

b.2) *Regime I in Fig. 2:* For $\beta < 1$ (and as a consequence of $\max\{\alpha, \beta\} > 1$ we must have $\alpha > 1$), to prove capacity, it suffices to show the achievability of $(r_1, r_2) = (1, \min\{1, \beta + \alpha - 1\})$. Here we build on the observation made for the achievable scheme in Regimes II to IV and develop a scheme that in addition to the “private rates” $r_{1p} = r_{2p} = \min\{\beta, \max\{1, \alpha\}\} = \beta$ also conveys common rates $r_{1c} = 1 - \beta$ and $r_{2c} = \min\{\alpha - 1, 1 - \beta\}$. In this regime some interfering bits can be decoded because the interference is strong ($\alpha > 1$) at the non-intended destination. As opposed to Regimes II to IV where the “common bits” were carried by the CR through U_0 , here they will be carried by X_1 and X_2 , i.e., cooperation through the CR in this regime is too weak and it is better used to neutralize the interference rather than to deliver common bits. Let $U_{1c}, U_{1p}, U_{2c}, U_{2p}$ be independent vectors. Consider

$$X_1 = \mathbf{S}^{m-n_C} U_{1p} + U_{1c}, \quad X_2 = \mathbf{S}^{m-n_C} U_{2p} + U_{2c}, \quad (16a)$$

$$X_0 = \mathbf{S}^{m-n_1} (U_{1p} + U_{2p}). \quad (16b)$$

The received signal at destination 1 is

$$Y_1 = \mathbf{S}^{m-n_S} U_{1c} + \mathbf{S}^{m-n_1} U_{2c} + \mathbf{S}^{m-n_C} (\mathbf{S}^{m-n_S} + \mathbf{S}^{m-n_1}) U_{1p}, \quad (17)$$

and similarly for destination 2. Clearly, if only the top $n_S(1 - \beta)$ bits of U_{1c} are non-zero and the top $n_S \min\{\alpha - 1, 1 - \beta\}$ bits of U_{2c} are non-zero, then destination 1 can decode U_{2c}, U_{1c}, U_{1p} in this order and destination 2 can decode U_{1c}, U_{2c}, U_{1p} in this order, thus achieving the desired rates.

Remark 3: Interestingly, the region in (12) is equivalent to the capacity region under “strong interference at both receivers” in [21, Theorem V.2], defined as the channel parameters for which

$$I(Y_2; X_2, X_c | X_1) \leq I(Y_1; X_2, X_0 | X_1), \quad (18a)$$

$$I(Y_1; X_1, X_c | X_2) \leq I(Y_2; X_1, X_0 | X_2), \quad (18b)$$

hold for all distributions that factor as $\mathbb{P}_{X_1, X_2, X_0} = \mathbb{P}_{X_1} \mathbb{P}_{X_2} \mathbb{P}_{X_0 | X_1, X_2}$. Evaluation of the condition of “strong interference at both receivers” in (18) is difficult because all possible input distributions must be tested—or an argument must be found that allows restriction to a specific subset of input distributions without loss of generality. For the LDA, it was not clear that i.i.d. Bernoulli(1/2) input bits at all nodes would exhaust all possible input distributions, as this does not capture the possible correlation between X_0 and (X_1, X_2) . It is interesting to notice that, with i.i.d. Bernoulli(1/2) input bits at all terminals, that the condition of “strong interference at both receivers” in (18) gives $\max\{\alpha, \beta\} \geq \max\{1, \beta\}$, or equivalently, $\max\{\alpha, \beta\} \geq 1$.

c) *Case $n_S > 0$, $\alpha \neq 1$ and $\max\{\alpha, \beta\} \leq 1$: sub-case $0 \leq \alpha \leq \beta \leq 1$ (Regime V in Fig. 2):* In Regime V the outer bound region is a square and has only one dominant corner point given by $r_1 = r_2 = 1$. Let U_{1p}, U_{2p} be independent vectors and set

$$X_1 = U_{1p}, \quad X_2 = U_{2p}, \quad X_0 = \mathbf{S}^{n_C - n_1} (U_{1p} + U_{2p}), \quad (19)$$

so as to neutralize the interference at the destinations (note the different shifts of the “private codewords” as compared to the scheme for Regimes I to IV in Fig. 2). In this regime $m = \max\{n_S, n_C, n_1\} = n_S$. The received signal at destination 1 is

$$Y_1 = (\mathbf{S}^{m-n_S} + \mathbf{S}^{m-n_C} \mathbf{S}^{n_C - n_1}) U_{1p}, \quad (20)$$

and similarly for the received signal at destination 2. Hence $R_1 = R_2 = n_S \max\{1, \alpha\} = n_S \cdot 1$. By normalizing the rates by n_S the claim follows.

Remark 4: Notice that X_0 in (19) is obtained by downshifting $U_{1p} + U_{2p}$ by $n_C - n_1$ positions, or in other words, the top $n_C - n_1$ bits of X_0 are zero. This strategy is slightly counter-intuitive as the cognitive relay, with knowledge of all messages, should be able to use all its bits without harm. However, including bits here would not improve rates as the direct link is already able to convey these bits directly, and the cognitive relay is only really needed to simultaneously cancel the interference at both receivers. The desired signal can be obtained by multiplying the received signal by the inverse of $\mathbf{S}^{m-n_S} + \mathbf{S}^{m-n_1}$, which is well defined as long as $n_S \neq n_1 \iff \alpha \neq 1$.

d) *Case $n_S > 0$, $\alpha \neq 1$ and $\max\{\alpha, \beta\} \leq 1$: sub-case $0 \leq \beta \leq \alpha < 1$ (Regime VI in Fig. 2):* In Region VI in Fig. 2, the

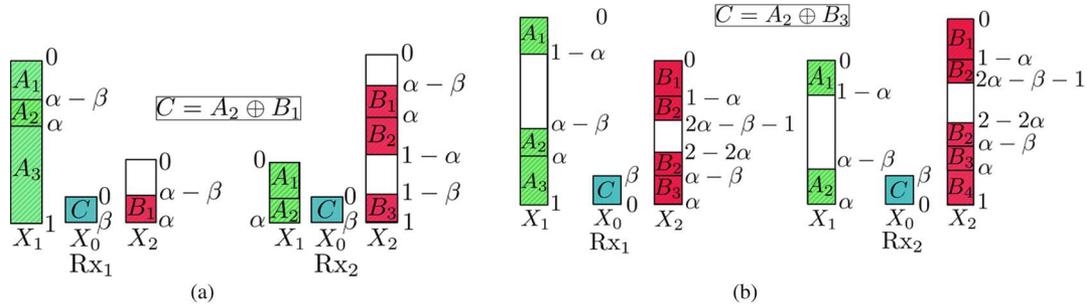


Fig. 3. Achievability strategies for portions of Regime VI of Fig. 2. (a) Optimal strategy for Regime VI.1. (b) Sum-capacity optimal strategy for parts of Regimes VI.3 and VI.4.

region in (11) simplifies to

$$r_1 \leq 1, \quad r_2 \leq 1, \quad (21a)$$

$$r_1 + r_2 \leq 2 - \alpha + \beta, \quad (21b)$$

$$r_1 + r_2 \leq 2 \max\{1 - \alpha, \alpha\} + 2\beta, \quad (21c)$$

$$2r_1 + r_2 \leq 1 + \max\{1 - \alpha, \alpha\} + \max\{1 - \alpha, \beta\} + 2\beta, \quad (21d)$$

$$r_1 + 2r_2 \leq 1 + \max\{1 - \alpha, \alpha\} + \max\{1 - \alpha, \beta\} + 2\beta. \quad (21e)$$

Due to the complexity of the outer bound region in (21), Regime VI is further divided into four sub-regimes, which also correspond to a generalization of the division of the W-curve in [24] as β is relatively small in this regime. The boundary between Regimes VI.1 and VI.2 occurs at $2\alpha = 1$, that between Regimes VI.2 and VI.3 at $\beta + 3\alpha = 2$, and that between Regimes VI.3 and VI.4 at $\beta + \alpha = 1$. These boundaries reduce to those of the W-curve in weak interference for $\beta = 0$. So far we were unable to show capacity for the whole Regime VI. We propose next a capacity achieving scheme Regime VI.1 and discuss strategies for the remaining cases.

In Regime VI.1 ($0 \leq \beta \leq \alpha \leq 1/2$) capacity can be proved by showing the achievability of the corner point $(r_1, r_2) = (1, 1 - 2\alpha + 2\beta)$ from (21), because in this regime the bounds in (21b), (21d), and (21e) are redundant. We will demonstrate our achievable scheme by using the graphical representation proposed in [28]. Fig. 3(a) shows such a strategy. The blocks represent the signals arriving at each destination, where block lengths has been normalized by n_S . Due to the channel down-shift operation, the desired signal at a destination has normalized length of 1, the signal from a cognitive relay has normalized length β , and the interfering signal has normalized length α . Bits intended for destination 1 are denoted by A_i , and those destined to destination 2 by B_i , $i \in [1 : 3]$. In our example, the relay sends $C := A_2 \oplus B_1$, where A_2 and B_1 play the role of U_{1p} and U_{2p} , respectively, in the previous regimes, i.e., they are “private bits” whose effect is “neutralized over the air” by the relay. Source 1 sends the block of bits indicated as A_1 “on top” of A_2 ; here A_1 plays role of U_{1c} in the previous regimes, i.e., they are “common bits” decodes at both destination; as for the classical IC, bits A_2 can be decoded at destination 2 if they are received interference-free at destination 2 [28], which is possible thanks to the fact that a portion of the signal sent by source 2 contains zeros (in between blocks B_2 and B_3). Blocks A_3 , B_2 , and B_3 are “private bits” too. However these

bits do not require “neutralization” by the relay as they appear “below the noise floor” at the non-intended receiver (similarly to the classical IC [28], these bits are actually not received at the non-intended destination). Notice that the top portion of the signal sent by source 2 is also populated by zeros (above block B_1); this is needed to allow destination 1 to decode A_3 . Destination/Rx1 decodes A_1, A_2, A_3 in this order, as does not suffers any interference from user 2, and achieves normalized rate $r_1 = 1$. Destination/Rx2 decodes B_1, B_2, A_1, B_3 in this order, and achieves normalized rate $r_2 = 1 - 2\alpha + 2\beta$.

e) On capacity and sum-capacity for parts of Regimes VI.3 and VI.4 in Fig. 2: Fig. 3(b) shows an achievable scheme for the case $4\alpha - 3 \leq \beta \leq 3\alpha - 2$, $2/3 \leq \alpha \leq 1$, where the restriction of the possible values of (α, β) is due to the fact that certain pieces of X_2 must have non-negative length. The corner point we aim to achieve is $(r_1, r_2) = (1, 1 - \alpha + \beta)$. Because the outer bound region in Regime VI.4, described by

$$r_1 \leq 1, \quad r_2 \leq 1, \quad r_1 + r_2 \leq 2 - \alpha + \beta, \quad (22a)$$

has only two corner points, achieving one of them implies the achievability of the entire capacity region by a time sharing argument. In contrast, Regime VI.3 described by

$$r_1 \leq 1, \quad r_2 \leq 1, \quad (23a)$$

$$r_1 + r_2 \leq 2 - \alpha + \beta, \quad (23b)$$

$$2r_1 + r_2 \leq 1 + \alpha + \max\{1 - \alpha, \beta\} + 2\beta, \quad (23c)$$

$$r_1 + 2r_2 \leq 1 + \alpha + \max\{1 - \alpha, \beta\} + 2\beta, \quad (23d)$$

has four corner points and thus achieving $(r_1, r_2) = (1, 1 - \alpha + \beta)$ does not suffice to show capacity. The other dominant corner point in Regime VI.3 is determined by the intersection of the r_1 -bound in (23a) with the $(2r_1 + r_2)$ -bound (23c). Thus, the strategy in Fig. 3(b) is only sum-rate optimal and works in the following way. Destination/Rx1 decodes the desired vector A_1 and the undesired vectors B_1 and B_2 . Now, since B_2 has been decoded, it can be subtracted at the points where its repetition interference with A_2 and A_3 . Thus, A_2 and A_3 are decoded too (note that the effect of B_3 has been “neutralized” by the relay and the block A_2 is decoded before decoding A_3 so its effect can be removed from A_3). Destination/Rx2 first decodes B_1 . Next, because vectors B_2, B_3 , and B_4 do not experience any interference, they can be decoded as well. Finally, since B_2 has been already decoded, the portion where A_1 interferes with B_2 can be ignored. This achieves $r_1 + r_2 = 2 - \alpha + \beta$. ■

It would be interesting to know what could be missing for Regimes VI.2 to VI.4. The capacity region for Regimes VI.2 to VI.4 remains unknown. These regimes are related to the most involved region of the W-curve in [24] for the IC in moderately weak interference (i.e., for $\alpha \in [1/2, 1]$) and as such it is not surprising that these are also the most difficult cases for the ICCR. At this point we conjecture that the way we have bounded the multi-letter portion (MLP) in (6) is too loose. We note that the entropy of a discrete random variable is non-negative and is not decreased by removing conditioning. Possibly the bound in (6) does not accurately capture the correlation between X_0 and (X_1, X_2) . Essentially the bound in (6), which for the symmetric LDA is given in (8), appears to assert that X_0 can be simultaneously maximally correlated with both X_1 and X_2 . However, if X_0 is maximally correlated with X_1 , i.e., $X_0 = X_1$, then it is independent of X_2 (recall that X_1 and X_2 are independent because carry independent messages); in this case the MLP expression would be $\min\{\alpha, \beta\}$ rather than $2 \min\{\alpha, \beta\}$. Tightening the bounds in (11d), (11e), and (11f) so as to capture the correlation among transmitted signals, and/or to derive another bound of the form $2R_1 + R_2$ or $R_1 + 2R_2$ (such a bound was needed for the IC with rate-limited receiver cooperation [30]) is the subject of ongoing investigation.

V. APPROXIMATE CAPACITY FOR THE SYMMETRIC GICCR

We now concentrate our attention on the symmetric GICCR. We will use the insights gained from the symmetric LDA to prove a constant gap result in those regimes where capacity is not known [21]. The symmetric GICCR is parameterize as

$$|h_{11}|^2 = |h_{22}|^2 = |h_S|^2 := \text{SNR}^1, \quad (24a)$$

$$|h_{12}|^2 = |h_{21}|^2 = |h_I|^2 := \text{SNR}^\alpha, \quad \alpha \geq 0, \quad (24b)$$

$$|h_{20}|^2 = |h_{10}|^2 = |h_C|^2 := \text{SNR}^\beta, \quad \beta \geq 0 \quad (24c)$$

where here α and β have meaning similar to the parameters used in the LDA model in (10), in particular α is the ratio of the received power on the interference link expressed in dB over the received power on the direct link expressed in dB, and β is the ratio of the received power on the relay-destination link expressed in dB over the received power on the direct link expressed in dB. The normalization of the SNR-exponent of the direct link to 1 is without loss of generality and parallels the normalization by n_S in the LDA. The following results parallel Theorem 5.

A. Capacity in Regimes I to IV in Fig. 2

Recently, the capacity of (18) was characterized in the “strong interference at both receivers” [21], which in the symmetric GICCR reduces to [21, eq. (27)]¹

$$||h_S| + |h_C||^2 \leq ||h_I|e^{+j\theta} + |h_C||^2, \quad \theta \in \{\angle h_{12}, \angle h_{21}\}, \quad (25)$$

¹The detailed proof is rather involved and uses the so called extremal inequality [31]. The main difficulty arises from the fact that X_1 and X_2 are correlated with X_0 and a more elaborate argument to show that Gaussians are optimal is needed. For interested readers the proof may be found in [21, Theorem VI.1].

where $\angle h_{21}, \angle h_{12}$ are the phases of the cross-link channel gains (the ones that could not be taken to be real-valued and non-negative without loss of generality in (2)). Using (24) and taking $\text{SNR} \rightarrow \infty$ the condition in (25), by assuming that $|h_I|e^{+j\angle h_{ij}} + |h_C| \neq 0$, reduces to

$$\max\{\text{SNR}, \text{SNR}^\beta\} \leq \max\{\text{SNR}^\alpha, \text{SNR}^\beta\} \iff 1 \leq \max\{\alpha, \beta\}. \quad (26)$$

The high-SNR regime of (26) coincides with Regimes I to IV in Fig. 2 for the LDA (see also Remark 3). In [21] it was shown that joint decoding of all messages at both receivers is optimal or capacity achieving when the “strong interference at both receivers” condition is satisfied. We therefore concentrate here on mimicking, in the Gaussian case, those regimes for which we could prove capacity in the LDA, namely Regimes V and VI.1.

B. Capacity to Within a Constant Gap in Regime V in Fig. 2

Regime V in the LDA is characterized by $\alpha \leq \beta \leq 1$, which we try to match with something of the form $|h_I|^2 \leq |h_C|^2 \leq |h_S|^2$ for the GICCR. We now build on the intuition developed for the LDA and propose a simple scheme that is optimal to within an additive gap.

Theorem 6: For the symmetric GICCR, the capacity outer bound in Theorem 1 is achievable to within $\log_2(4/(1 - (1/\sqrt{2}))^2) \approx 5.5$ bits per user for $2|h_I|^2 \leq |h_C|^2 \leq |h_S|^2$. \square

Proof: In Regime V for the LDA, the cognitive relay simultaneously neutralizes over the air the interference at both receivers. This mode of operation is reminiscent of zero forcing. We therefore propose: let U_{1p} and U_{2p} be two independent Gaussian random variables with zero mean and unit variance and define for some (ρ_1, ρ_2) such that $|\rho_1|^2 + |\rho_2|^2 \leq 1$

$$X_1 = U_{1p}, \quad X_2 = U_{2p}, \quad X_0 = \rho_1 U_{1p} + \rho_2 U_{2p}. \quad (27)$$

Next we choose ρ_1 and ρ_2 so as to simultaneously neutralize the contribution of U_{2p} at destination 1 and of U_{1p} at destination 2. This is possible if

$$\rho_1 = -\frac{|h_I|e^{+j\angle h_{21}}}{|h_C|}, \quad \rho_2 = -\frac{|h_I|e^{+j\angle h_{12}}}{|h_C|}, \quad (28)$$

which requires $2|h_I|^2 \leq |h_C|^2$. With this assignment the channel outputs become

$$Y_1 = (|h_S| - |h_I|e^{+j\angle h_{21}}) U_{1p} + Z_1, \quad (29a)$$

$$Y_2 = (|h_S| - |h_I|e^{+j\angle h_{12}}) U_{2p} + Z_2, \quad (29b)$$

and thus the following rates are achievable:

$$R_1 \leq \log \left(1 + ||h_S| - |h_I|e^{+j\angle h_{21}}|^2 \right), \quad (30a)$$

$$R_2 \leq \log \left(1 + ||h_S| - |h_I|e^{+j\angle h_{12}}|^2 \right). \quad (30b)$$

From the outer bound we have

$$\begin{aligned} R_1 &\leq I(Y_1; X_1, X_0|Q, X_2) \leq \log \left(1 + (|h_S| + |h_C|)^2 \right) \\ &\leq \log \left(1 + 4 \max \{ |h_S|^2, |h_C|^2 \} \right), \end{aligned} \quad (31)$$

and similarly for R_2 . Next, the argument of the log-function in (30) can be lower bounded by $(|h_S| - |h_I|)^2$. Imposing $|h_C|^2 \leq |h_S|^2$, to mimic Regime V of the LDA, and $2|h_I|^2 \leq |h_C|^2$, implies $|h_I|^2 \leq |h_S|^2/2$, so that $(|h_S| - |h_I|)^2 \geq (1 - (1/\sqrt{2}))^2 |h_S|^2$. Finally, by taking the difference between the upper bound in (31) and the lower bound in (30) we arrive at the claimed gap result. ■

It is pleasing to see that a simple interference management technique reminiscent of zero-forcing is optimal to within a constant gap for the GICCR. Notice that in this regime the channel behaves effectively as two non-interfering point-to-point links.

C. Capacity to Within a Constant Gap in Regime VI.1 in Fig. 2

Regime VI.1 for the LDA is characterized by $\beta \leq \alpha \leq 1/2$, which we try to match with something of the form $|h_C|^2 \leq |h_I|^2 \leq \sqrt{|h_S|^2}$ for the GICCR. Next, we build on the intuition developed in the LDA and propose a scheme that is optimal to within an additive gap.

Theorem 7: For the symmetric GICCR, the capacity outer bound in Theorem 4 is achievable to within 11.7 bits per user if the channel gains satisfy the following three conditions: (c1) $|h_C|^2 \leq |h_I|^2(|h_I|^2/(1 + |h_I|^2))$, (c2) $|h_C|^2 \leq (1/2)((1 + |h_S|)/(1 + |h_I|^2))$, (c3) $|h_I|^2(1 + |h_I|^2) \leq |h_S|^2$, (c4) $1 \leq \min\{|h_S|^2, |h_I|^2\}$, (c5) $(|h_S|^2/(1 + |h_I|^2)) \geq 9$. □

Proof: The conditions (c1)–(c3) at high SNR are equivalent to $\beta \leq \alpha \leq 1/2$; conditions (c4)–(c5) are convenient for gap computation. In Regime VI.1 for the LDA, the CR simultaneously neutralizes interference at destination 1 and part of the interference at destination 2, see Fig. 3(a). We therefore propose the following choice of inputs: for $X_{A_1}, X_{A_2}, X_{A_3}, X_{B_1}, X_{B_2}, X_{B_3}$ i.i.d. Gaussian random variables with zero mean and unit variance, let

$$\begin{aligned} X_1 &= a_1 X_{A_1} + a_2 X_{A_2} + a_3 X_{A_3}, \\ X_2 &= b_1 X_{B_1} + b_2 X_{B_2} + b_3 X_{B_3}, \\ X_0 &= \ell_1 X_{A_2} + \ell_2 X_{B_1}, \\ \ell_1 &= \frac{1}{\sqrt{2}} e^{j\angle \ell_1}, \quad \ell_2 = \frac{1}{\sqrt{2}} e^{j\angle \ell_2}, \\ a_1 &= \frac{|h_I|^2}{1 + |h_I|^2} - \frac{|h_C|^2}{2|h_I|^2}, \quad b_1 = \frac{-|h_C|e^{j\angle \ell_2}}{\sqrt{2}|h_I|e^{j\angle h_{21}}}, \\ a_2 &= \frac{-|h_C|e^{j\angle \ell_1}}{\sqrt{2}|h_I|e^{j\angle h_{12}}}, \quad |b_2|^2 = \frac{1}{1 + |h_I|^2} - \frac{|h_C|^2}{1 + |h_S|^2}, \\ |a_3|^2 &= \frac{1}{1 + |h_I|^2}, \quad |b_3|^2 = \frac{|h_C|^2}{1 + |h_S|^2}, \end{aligned} \quad (32)$$

Under the channel conditions $|h_C|^2 \leq 2|h_I|^2(|h_I|^2/(1 + |h_I|^2))$ so that $|a_1|^2 \geq 0$, and $|h_C|^2 \leq ((1 + |h_S|)/(1 + |h_I|^2))$ so that $|b_2|^2 \geq 0$, the transmitter power constraints are satisfied; these conditions are true by (c1) and (c2), respectively. Note that transmitter 2 does not fully utilize its power. With

this choice of coefficients/power allocation, the channel outputs become

$$\begin{aligned} Y_1 &= |h_S| (a_1 X_{A_1} + a_3 X_{A_3}) \\ &\quad + \frac{|h_C|}{\sqrt{2}} \left(\frac{|h_S|}{|h_I|} e^{-j\angle h_{21}} - 1 \right) X_{A_2} \\ &\quad + |h_I| (b_2 X_{B_2} + b_3 X_{B_3}) e^{+j\angle h_{12}} + Z_1, \\ Y_2 &= |h_I| (a_1 X_{A_1} + a_3 X_{A_3}) e^{+j\angle h_{21}} \\ &\quad + \frac{|h_C|}{\sqrt{2}} \left(\frac{|h_S|}{|h_I|} e^{-j\angle h_{12}} - 1 \right) X_{B_1} \\ &\quad + |h_S| (b_2 X_{B_2} + b_3 X_{B_3}) + Z_2, \end{aligned}$$

since X_{B_1} has been zero forced at Y_1 , and X_{A_2} at Y_2 , similarly to the scheme in Fig. 3(a) for the LDA. By mimicking the corresponding scheme for the LDA, destination 1 successively decodes $X_{A_1}, X_{A_2}, X_{A_3}$ in this order, and destination 2 successively decodes $X_{B_1}, X_{B_2}, X_{A_1}, X_{B_3}$ in this order; with this decoding procedure the following rates are achievable (see Appendix E)

$$\begin{aligned} R_{A_1} &= \log \left(1 + \frac{|h_I|^2}{4(3 + |h_C|^2)} \right), \\ R_{A_2} &= \log \left(1 + \frac{|h_C|^2}{10} \right), \\ R_{A_3} &= \log \left(1 + \frac{|h_S|^2}{1 + 2|h_I|^2} \right), \\ R_{B_1} &= \log \left(1 + \frac{|h_C|^2}{10} \right), \\ R_{B_2} &= \log \left(1 + \frac{|h_S|^2}{4(1 + |h_I|^2)^2} \right), \\ R_{B_3} &= \log \left(1 + \frac{|h_C|^2}{4} \right). \end{aligned}$$

We next compare this lower bound with the outer bound obtained by intersecting the sum-rate upper bound in (5a) with the MLP tightened as in Theorem 4 (see Appendix D) and the single-rate upper bound in (4a) (see (31)), that is, the corner point outer bound with coordinates

$$R_1 = \log(1 + 4|h_S|^2), \quad (33a)$$

$$\begin{aligned} R_2 &= 2 \log \left(\left(1 + |h_I|^2 + \frac{|h_S|^2}{1 + |h_I|^2} \right) \right. \\ &\quad \cdot (1 + |h_C|^2) 2(1 + 1/\sqrt{2})^2 \left. \right) \\ &\quad - \log(1 + 4|h_S|^2). \end{aligned} \quad (33b)$$

In Appendix E we show that the gap between the inner and outer bound is at most 11.7 bits per user. By swapping the role of the users, the other sum-capacity achieving corner point of the capacity region outer bound can be attained to within the same gap.

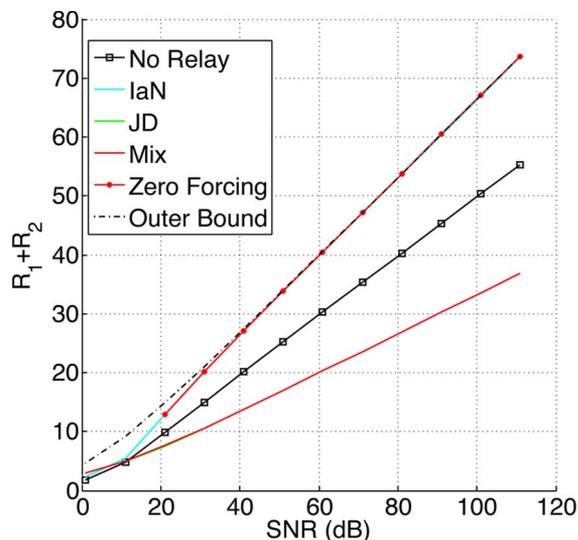


Fig. 4. Numerical comparison of various strategies for the GICCR for $\alpha = 0.5$, $\beta = 0.7$. No relay: the relay is not used and the rates are given by the optimal interference channel strategy. For the other curves the relay uses a linear strategy and the receivers apply JD (jointly decode both messages), or IaN (treat interference as noise), or Mix (one receiver decodes both messages and the other only its intended one), or ZF (relay performs zero forcing as in (30), which is special case of IaN). The outer bound is from (31).

By setting $R_2 = 0$ and not using the CR we can achieve $R_1 = \log(1 + |h_S|^2)$, which is at most 2 bits away from the corner point where R_1 is upper bounded by (33a) and $R_2 = 0$. The same reasoning holds with the role of the users swapped. This shows that all corner points of the outer bound region can be achieved to within 11.7 bits per user. Therefore, by time sharing, the whole capacity region outer bound can be achieved to within a constant gap. This concludes the proof. ■

The gap in this regime is fairly large; we believe that this is due to the crude lower bounding steps for the achievable rates and to the simplicity of the proposed interference zero-forcing scheme. Numerical evaluations show that the actual gap when optimizing the power splits in the proposed scheme is actually lower.

D. Numerical Comparisons

We conclude this section with some numerical examples. Figs. 4 and 5 compare the performance of different achievable strategies as a function of the SNR (in dB) in Regime V, where the new constant gap result is obtained from Theorem 6. We note that the purpose of this paper is to provide *simple* achievable schemes for the Gaussian channel that are provably optimal to within a constant gap, rather than focussing on finding the parameters that optimize the largest known (quite involved) achievable rate region for the ICCR derived in [21, Theorem IV.1]. To this end, we compare several simple achievability schemes, including the constant gap to capacity scheme in (30) and the outer bound in (4).

In Fig. 4, we increase the SNR with fixed $\alpha = 0.5$ and $\beta = 0.7$ in (24) and compare the following strategies. In the first strategy the relay stays silent and we use a well-known achievability strategy for the Gaussian IC (a version of the Han and Kobayashi strategy [24]). For the second achievabil-

ity scheme, the relay is used and performs the simple linear combination scheme (rather than more complex schemes such as dirty paper coding) $X_0 = a_1 X_1 + a_2 X_2 : |a_1|^2 + |a_2|^2 \leq 1$, where we optimize over a_1 and a_2 . We consider the following strategies at the receivers:

- 1) JD (Joint Decoding): both transmitters use common messages only, which are decoded at both destinations—the region thus looks like a compound multiple access channel with each message amplified at the receiver due to the relay’s transmission.
- 2) IaN (Interference as Noise): destinations treat non-desired interference as noise. All messages are therefore private.
- 3) Mix: one of the transmitters uses a common message and the other uses a private message; the common message is decoded at both receivers and the private is decoded at the appropriate receiver only and treated as noise at the other.
- 4) ZF (Zero Forcing): use $a_1 = -(|h_1|e^{+j\angle h_{21}}/|h_C|)$ and $a_2 = -(|h_1|e^{+j\angle h_{12}}/|h_C|)$ as in Theorem 6 (when possible).

Finally, the sum-rate outer bound from (31) is plotted (i.e., in this case the whole capacity region is a square).

From Fig. 4 we see that as the SNR increases, the IaN and ZF schemes (ZF is actually one very specific choice of the IaN scheme where a_1, a_2 are specified explicitly) essentially overlap with the outer bound, which verifies the constant gap to capacity claim numerically. This scheme significantly outperforms (diverging slopes means the gap can be arbitrarily large) not using a relay at all, even with an optimizing transmission strategy, or using a JD or Mix strategy where the relay uses a simple linear combination scheme. Fig. 5 shows the actual regions, rather than sum-rates, for the same settings and conventions as in Fig. 4 for two different SNRs. We note that our goal is not to derive the best achievability scheme at any SNR, but rather to derive a simple, constant gap to capacity scheme and compare it to other, simple schemes.

VI. CONCLUSION

We considered an interference channel in which a cognitive relay aids in the transmission of the two independent messages. We obtained the capacity region in almost all regimes for the symmetric LDA and translated these insights into a constant gap to capacity result for the corresponding Gaussian model. The capacity achieving schemes for the symmetric LDA use a variety of techniques at the cognitive relay, which both aid in the transmission of the messages to the receivers, and simultaneously neutralize interference at the two receivers. Given the generality of this challenging channel model, it is not surprising that a number of open questions remain: capacity is missing in a parameter regime of the symmetric LDA which has typically been the most challenging one for the interference channel as well (the moderately weak interference regime). Constant gap to capacity results for the corresponding regime in the Gaussian channel are also missing and are an interesting topic for further investigation.

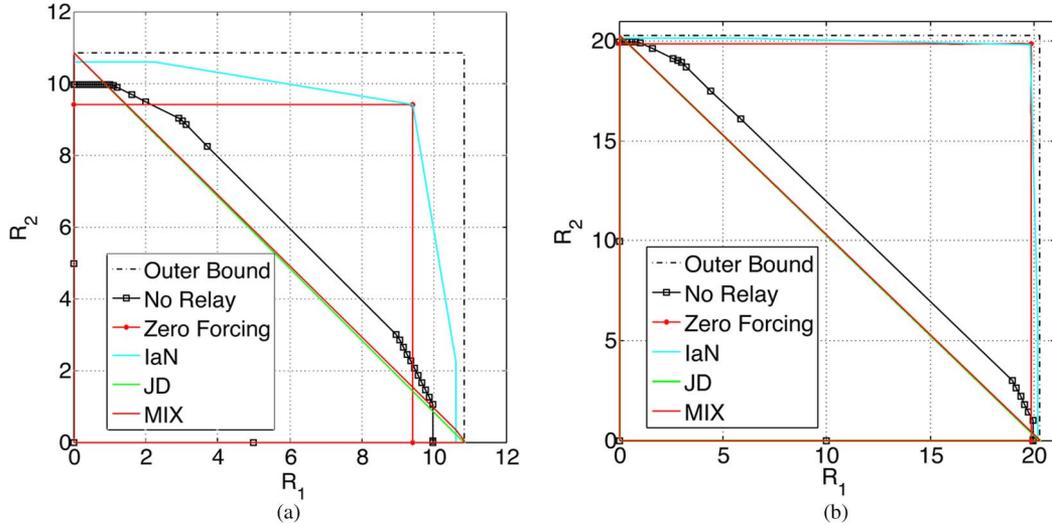


Fig. 5. Achievable and outer bound regions for the GICCR with $\alpha = 0.5$, $\beta = 0.7$. Same settings and conventions as in Fig. 4. (a) SNR = 30 dB. (b) SNR = 60 dB.

APPENDIX

A. Proof of Theorem 2

Given the random variables $(Q, X_1, X_2, X_0, V_1, V_2, Y_1, Y_2)$ with joint distribution

$$\begin{aligned} & \mathbb{P}_{Q, X_1, X_2, X_0, V_1, V_2, Y_1, Y_2} \\ &= \mathbb{P}_Q \mathbb{P}_{X_1|Q} \mathbb{P}_{X_2|Q} \mathbb{P}_{X_0|Q, X_1, X_2} \mathbb{P}_{V_1|X_1} \mathbb{P}_{V_2|X_2} \\ & \quad \cdot \delta(Y_1 - f_1(X_1, X_0, V_2)) \delta(Y_2 - f_2(X_2, X_0, V_1)) \end{aligned}$$

let \tilde{V}_1 and \tilde{V}_2 be conditionally independent copies of V_1 and V_2 , that is, distributed jointly with (Q, X_1, X_2, X_0) as

$$\begin{aligned} & \mathbb{P}_{\tilde{V}_1, \tilde{V}_2|Q, X_1, X_2, X_0}(v_1, v_2|q, x_1, x_2, x_0) \\ &= \mathbb{P}_{V_1|X_1}(v_1|x_1) \mathbb{P}_{V_2|X_2}(v_2|x_2). \end{aligned}$$

Similar arguments to those in [18] yield:

$$\begin{aligned} & n(R_1 + R_2 - 2\epsilon_n) \\ & \leq I(W_1; Y_1^n, \tilde{V}_1^n) + I(W_2; Y_2^n, \tilde{V}_2^n) \end{aligned} \quad (34a)$$

$$= H(\tilde{V}_1^n) - H(\tilde{V}_1^n|W_1, X_1^n) \quad (34b)$$

$$\begin{aligned} & + H(Y_1^n|\tilde{V}_1^n) - H(Y_1^n|\tilde{V}_1^n, W_1, X_1^n) \\ & + H(\tilde{V}_2^n) - H(\tilde{V}_2^n|W_2, X_2^n) \end{aligned} \quad (34c)$$

$$+ H(Y_2^n|\tilde{V}_2^n) - H(Y_2^n|\tilde{V}_2^n, W_2, X_2^n) \quad (34d)$$

$$\stackrel{(a)}{\leq} H(\tilde{V}_1^n) - H(\tilde{V}_1^n|X_1^n) + H(Y_1^n|\tilde{V}_1^n) - H(Y_1^n|\tilde{V}_1^n, W_1, X_1^n, X_0^n) \quad (34e)$$

$$\begin{aligned} & + H(\tilde{V}_2^n) - H(\tilde{V}_2^n|X_2^n) + H(Y_2^n|\tilde{V}_2^n) \\ & - H(Y_2^n|\tilde{V}_2^n, W_2, X_2^n, X_0^n) \end{aligned} \quad (34f)$$

$$\begin{aligned} & \stackrel{(b)}{=} H(Y_1^n|\tilde{V}_1^n) + H(Y_2^n|\tilde{V}_2^n) \\ & - H(\tilde{V}_1^n|X_1^n) - H(\tilde{V}_2^n|X_2^n) \\ & + H(\tilde{V}_1^n) - H(V_1^n|W_2, \tilde{V}_2^n, X_2^n, X_0^n) \\ & + H(\tilde{V}_2^n) - H(V_2^n|W_1, \tilde{V}_1^n, X_1^n, X_0^n) \end{aligned} \quad (34g)$$

$$\stackrel{(c)}{=} H(Y_1^n|\tilde{V}_1^n) + H(Y_2^n|\tilde{V}_2^n) \quad (34h)$$

$$\begin{aligned} & - H(\tilde{V}_1^n|X_1^n) - H(\tilde{V}_2^n|X_2^n) \\ & + H(\tilde{V}_1^n) - H(V_1^n|W_2, X_0^n) \\ & + H(\tilde{V}_2^n) - H(V_2^n|W_1, X_0^n) \end{aligned} \quad (34i)$$

$$\begin{aligned} & = H(Y_1^n|\tilde{V}_1^n) + H(Y_2^n|\tilde{V}_2^n) \\ & - H(\tilde{V}_1^n|X_1^n) \\ & - H(\tilde{V}_2^n|X_2^n) \\ & + I(V_1^n; X_0^n|W_2) + I(V_2^n; X_0^n|W_1) \end{aligned} \quad (34j)$$

where the inequality in “(a)” follows from further conditioning on X_0 (and because condition on X_i^n we have that V_i^n is independent of everything else, so that in particular we can drop the message W_i from the conditioning, $i = 1, 2$), the equality in “(b)” follows from the assumed determinism, the equality in “(c)” follows from the since V_1^n is independent of (\tilde{V}_2^n, X_2^n) so it can be dropped from the conditioning (however X_0^n depends on (W_1, W_2) so we must keep the messages in the conditioning) and similarly for user 2.

B. Proof of Theorem 3

For the channels in (7), instead of conditioning on X_0 in the step marked by (a) in Appendix A, we condition on the $q_i(X_0)$, $i \in [1, 2]$, to obtain the tighter bound

$$\begin{aligned} & H(\tilde{V}_2^n) - H(Y_1^n|\tilde{V}_1^n, W_1, X_1^n) \\ & \stackrel{(a')}{\leq} H(\tilde{V}_2^n) - H(Y_1^n|\tilde{V}_1^n, W_1, X_1^n, q_1(X_0^n)) \\ & = H(V_2^n) - H(V_2^n|\tilde{V}_1^n, W_1, X_1^n, q_1(X_0^n)) \\ & = H(V_2^n) - H(V_2^n|W_1, q_1(X_0^n)) \\ & = I(V_2^n; W_1, q_1(X_0^n)) \\ & = I(V_2^n; q_1(X_0^n)|W_1) \\ & \leq \min\{H(V_2^n), H(q_1(X_0^n))\} \\ & \leq n \min\{H(V_2|Q), H(q_1(X_0)|Q)\}, \end{aligned}$$

and similarly for the other users. The fact that the resulting region is exhausted by i.i.d. Bernoulli(1/2) bits for the input vectors follows by arguments similar to [19].

C. Proof of Theorem 4

Inspired by the proof of Theorem 3—where the term $H(Y_i^n | \tilde{V}_i^n, W_i, X_i^n)$, $i \in [1:2]$, was further conditioned on $q_i(X_0^n)$ rather than on X_0^n (i.e., compare step marked by (a) in Appendix A with step marked by (a') in Appendix B)—we mimic here the function $q_i(X_0^n)$ for the LDA with $|h_{i0}|X_0^n - Z_0^n$ for the GICCR, where Z_0 i.i.d. $\mathcal{N}(0, 1)$ independent of $(Z_1, Z_2, \tilde{Z}_1, \tilde{Z}_2, W_1, W_2)$. Recall that

$$\begin{aligned} V_2 &= h_{12}X_2 + Z_1 \sim \tilde{V}_2 = h_{12}X_2 + \tilde{Z}_1 \\ &: Z_2 \text{ independent of } \tilde{Z}_1 \sim Z_1, \\ V_1 &= h_{21}X_1 + Z_2 \sim \tilde{V}_1 = h_{21}X_1 + \tilde{Z}_2 \\ &: Z_1 \text{ independent of } \tilde{Z}_2 \sim Z_2. \end{aligned}$$

Then, we replace the step marked with (a) in Appendix A with

$$\begin{aligned} &h(\tilde{V}_2^n) - h(Y_1^n | \tilde{V}_1^n, W_1, X_1^n) \\ &\stackrel{(a')}{\leq} h(h_{12}X_2^n + \tilde{Z}_1^n) \\ &\quad - h(|h_{11}|X_1^n + |h_{10}|X_0^n + h_{12}X_2^n + Z_1^n | h_{21}X_1^n + \tilde{Z}_2^n, \\ &\quad\quad W_1, X_1^n, |h_{10}|X_0^n - Z_0^n) \\ &= h(h_{12}X_2^n + Z_1^n) \\ &\quad - h(h_{12}X_2^n + Z_1^n + Z_0^n | W_1, |h_{10}|X_0^n - Z_0^n) \\ &= -I(h_{12}X_2^n + Z_1^n + Z_0^n; Z_0^n) \\ &\quad + I(h_{12}X_2^n + Z_1^n + Z_0^n; |h_{10}|X_0^n - Z_0^n | W_1) \\ &\leq -0 + h(|h_{10}|X_0^n - Z_0^n) \\ &\quad - h(|h_{10}|X_0^n - Z_0^n | W_1, h_{12}X_2^n + Z_1^n + Z_0^n, W_2) \\ &= I(|h_{10}|X_0^n - Z_0^n; X_0^n) + I(Z_0^n; Z_1^n + Z_0^n) \\ &\leq n \log(1 + |h_{10}|^2) + n \log(2). \end{aligned}$$

We can also trivially upper bound MLP_1 in (5d) as

$$\begin{aligned} &h(\tilde{V}_2^n) - h(Y_1^n | \tilde{V}_1^n, W_1, X_1^n) \leq I(V_2^n; X_0^n | W_1) \\ &\leq h(h_{12}X_2^n + Z_1^n) \\ &\quad - h(h_{12}X_2^n + Z_1^n | X_0^n, X_1^n, W_1, W_2) \\ &= I(h_{12}X_2^n + Z_1^n; X_0^n) \leq n \log(1 + |h_{12}|^2). \end{aligned}$$

Therefore, we conclude that

$$\frac{h(\tilde{V}_2^n) - h(Y_1^n | \tilde{V}_1^n, W_1, X_1^n)}{n} \leq \log(1 + \min\{|h_{12}|^2, |h_{10}|^2\}) + \log(2).$$

By repeating the same reasoning for the other receiver, we conclude that for the GICCR Theorem 2 holds with MLP_1 in (5d) replaced by

$$\begin{aligned} \text{MLP}_2 &:= \log(1 + \min\{|h_{12}|^2, |h_{10}|^2\}) \\ &\quad + \log(1 + \min\{|h_{21}|^2, |h_{20}|^2\}) + 2 \log(2). \end{aligned}$$

The resulting region is exhausted by jointly Gaussian inputs by arguments similar to [21].

D. Evaluation of the Sum-Rate Upper Bound in (5a) for the GICCR

By Theorem 4 we can restrict attention to jointly Gaussian inputs. Let parameterize the possible jointly Gaussian inputs as

$$\begin{bmatrix} X_1 \\ X_2 \\ X_0 \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} 1 & 0 & r_1^* \\ 0 & 1 & r_2^* \\ r_1 & r_2 & 1 \end{bmatrix}\right) : |r_1|^2 + |r_2|^2 \leq 1,$$

that is, $X_0 = r_1X_1 + r_2X_2 + X'_0$ with $X'_0 \sim \mathcal{N}(0, 1 - |r_1|^2 - |r_2|^2)$ and independent of everything else. In (5a), consider the term

$$\begin{aligned} &h(Y_1 | \tilde{V}_1, Q) - h(\tilde{V}_2 | X_2) \\ &\leq h(|h_{11}|X_1 + |h_{10}|X_0 + h_{12}X_2 + Z_1 | h_{21}X_1 + \tilde{Z}_2) \\ &\quad - h(\tilde{Z}_1) \\ &= h(a_1X_1 + |h_{10}|X'_0 + a_2X_2 + Z_1 | h_{21}X_1 \\ &\quad\quad + \tilde{Z}_2) \Big|_{a_1:=|h_{11}|+r_1|h_{10}|, a_2:=h_{12}+r_2|h_{10}|} - h(Z_1) \\ &= \log\left(\frac{|a_1|^2}{1 + |h_{21}|^2} + |h_{10}|^2(1 - |r_1|^2 - |r_2|^2) \right. \\ &\quad\quad \left. + |a_2|^2 + 1\right) \Big|_{a_1:=|h_{11}|+r_1|h_{10}|, a_2:=h_{12}+r_2|h_{10}|} \\ &\leq \log\left(\frac{(|h_{11}| + |r_1||h_{10}|)^2}{1 + |h_{21}|^2} + |h_{10}|^2(1 - |r_1|^2 - |r_2|^2) \right. \\ &\quad\quad \left. + (|h_{12}| + |r_2||h_{10}|)^2 + 1\right) \\ &= \log\left(\frac{(|h_{11}| + |r_1||h_{10}|)^2}{1 + |h_{21}|^2} + |h_{10}|^2(1 - |r_1|^2) \right. \\ &\quad\quad \left. + 2|r_2||h_{10}||h_{12}| + |h_{12}|^2 + 1\right) \end{aligned}$$

where clearly the last expression, for any r_1 such that $1 - |r_1|^2 \geq 0$, is maximized by $|r_2| = \sqrt{1 - |r_1|^2}$ (recall that the bound must be optimized over $|r_1|^2 + |r_2|^2 \leq 1$); this implies that for some $|r_1|^2 + |r_2|^2 = 1$

$$\begin{aligned} &h(Y_1 | \tilde{V}_1, Q) - h(\tilde{V}_2 | X_2) \\ &\leq \log\left(\frac{(|h_{11}| + |r_1||h_{10}|)^2}{1 + |h_{21}|^2} + (|h_{12}| + |r_2||h_{10}|)^2 + 1\right). \quad (35) \end{aligned}$$

By a similar reasoning for the other receiver, we have that for some $|r_1|^2 + |r_2|^2 = 1$

$$\begin{aligned} &h(Y_2 | \tilde{V}_2, Q) - h(\tilde{V}_1 | X_1) \\ &\leq \log((|h_{21}| + |r_1||h_{20}|)^2 + \frac{(|h_{22}| + |r_2||h_{20}|)^2}{1 + |h_{12}|^2} + 1) \quad (36) \end{aligned}$$

Finally, by summing (35) and (36), the sum-rate upper bound from Theorem 2 with the MLP from Theorem 4 reads

$$\begin{aligned}
R_1 + R_2 &\leq \max_{|r_1|^2 + |r_2|^2 = 1} \log \left(\frac{(|h_{11}| + |r_1||h_{10}|)^2}{1 + |h_{21}|^2} + (|h_{12}| + |r_2||h_{10}|)^2 + 1 \right) \\
&\quad + \log \left((|h_{21}| + |r_1||h_{20}|)^2 + \frac{(|h_{22}| + |r_2||h_{20}|)^2}{1 + |h_{12}|^2} + 1 \right) \\
&\quad + \log(1 + \min\{|h_{12}|^2, |h_{10}|^2\}) \\
&\quad + \log(1 + \min\{|h_{21}|^2, |h_{20}|^2\}) + 2 \log(2). \tag{37}
\end{aligned}$$

In the symmetric case in (24), by the symmetry of the problem, it is easy to see that the maximizing (r_1, r_2) is such that $|r_1|^2 = |r_2|^2 = 1/2$; hence the sum-rate upper bound in (37) reads

$$\begin{aligned}
R_1 + R_2 &\leq 2 \log \left(\frac{\max\{|h_S|^2, |h_C|^2\}}{1 + |h_I|^2} + \max\{|h_I|^2, |h_C|^2\} + 1 \right) \\
&\quad + 2 \log(1 + \min\{|h_I|^2, |h_C|^2\}) + 2 \log(2(1 + 1/\sqrt{2})^2) \tag{38}
\end{aligned}$$

where $2 \log(2(1 + 1/\sqrt{2})^2) \leq 2 \log(6) < 5.17$ bits.

E. Lower Bounds on the Achievable Rates for the Scheme in Section V-C

The achievable scheme in Section V-C attains the following rates (where the further lower bounds follow from straightforward but tedious algebraic manipulations by using the conditions (c1)–(c5) of Theorem 7), see (39a)–(39g) (shown at the bottom of the page) and, because X_{A_1} is a “common message” decoded at both destinations, we finally choose

$$R_{A_1} = \min\{\text{eq. (39a), eq. (39f)}\} \geq \log \left(1 + \frac{|h_I|^2}{4(3 + |h_C|^2)} \right). \tag{39h}$$

As outer bound consider the corner point obtained by intersecting the sum-rate upper bound in (38) with the single-rate upper bound in (4a) (see (31)) whose coordinates are given in (33) (note that in this regime the channel gains satisfy $|h_C|^2 \leq |h_I|^2 \leq |h_S|^2$). We next compare the lower bound in (39) with the corner point outer bound in (33). It can be easily seen that the gap for R_1 is, for $n_{A_2} \geq 3, n_{B_1} \geq 4$,

$$\text{gap}_{R_1} \leq \log(4 \cdot n_{A_2} \cdot 8)$$

$$\text{gap}_{R_2} \leq \log(4(1 + 1/\sqrt{2})^4 \cdot 5 \cdot n_{B_1} \cdot 4).$$

$$\begin{aligned}
R_{A_1} &\leq \log \left(1 + \frac{|h_S|^2 \left(\frac{|h_I|^2}{1 + |h_I|^2} - \frac{|h_C|^2}{2|h_I|^2} \right)}{1 + \frac{|h_S|^2}{1 + |h_I|^2} + \frac{|h_I|^2}{1 + |h_I|^2} + \frac{|h_C|^2}{2} \left| \frac{|h_S|}{|h_I|} e^{-j\angle h_{21}} - 1 \right|^2} \right) \\
&\quad : \frac{|h_S|^2 \left(\frac{|h_I|^2}{1 + |h_I|^2} - \frac{|h_C|^2}{2|h_I|^2} \right)}{1 + \frac{|h_S|^2}{1 + |h_I|^2} + \frac{|h_I|^2}{1 + |h_I|^2} + \frac{|h_C|^2}{2} \left| \frac{|h_S|}{|h_I|} e^{-j\angle h_{21}} - 1 \right|^2} \geq \frac{|h_I|^2}{4} \frac{1}{3 + |h_C|^2}, \tag{39a}
\end{aligned}$$

$$\begin{aligned}
R_{A_2} &= \log \left(1 + \frac{|h_C|^2}{2} \frac{\left| \frac{|h_S|}{|h_I|} e^{-j\angle h_{21}} - 1 \right|^2}{1 + \frac{|h_S|^2}{1 + |h_I|^2} + \frac{|h_I|^2}{1 + |h_I|^2}} \right) \geq \log \left(1 + |h_C|^2 \frac{(\sqrt{t} - 1)^2}{2(2 + t)} \right) \Big|_{t := \frac{|h_S|^2}{1 + |h_I|^2}} \\
&\geq \log \left(1 + \frac{|h_C|^2}{n_{A_2}} \right) \text{ for } \sqrt{t} \geq \frac{1 + \sqrt{4/n_{A_2}(3/2 - 2/n_{A_2})}}{1 - 2/n_{A_2}}, \tag{39b}
\end{aligned}$$

$$R_{A_3} = \log \left(1 + \frac{\frac{|h_S|^2}{1 + |h_I|^2}}{1 + \frac{|h_I|^2}{1 + |h_I|^2}} \right) = \log \left(1 + \frac{|h_S|^2}{1 + 2|h_I|^2} \right), \tag{39c}$$

$$\begin{aligned}
R_{B_1} &= \log \left(1 + \frac{|h_C|^2}{2} \frac{\left| \frac{|h_S|}{|h_I|} e^{-j\angle h_{12}} - 1 \right|^2}{1 + |h_I|^2 \left(\frac{|h_I|^2}{1 + |h_I|^2} - \frac{|h_C|^2}{2|h_I|^2} \right) + \frac{|h_I|^2}{1 + |h_I|^2} + \frac{|h_S|^2}{1 + |h_I|^2}} \right) \geq \log \left(1 + |h_C|^2 \frac{(\sqrt{t} - 1)^2}{4(1 + t)} \right) \\
&\geq \log \left(1 + \frac{|h_C|^2}{n_{B_1}} \right) \text{ for } \sqrt{t} \geq \frac{1 + \sqrt{8/n_{B_1}(1 - 2/n_{B_1})}}{1 - 4/n_{B_1}}, \tag{39d}
\end{aligned}$$

$$R_{B_2} = \log \left(1 + \frac{|h_S|^2 \left(\frac{1}{1 + |h_I|^2} - \frac{|h_C|^2}{1 + |h_S|^2} \right)}{1 + |h_I|^2 \left(\frac{|h_I|^2}{1 + |h_I|^2} - \frac{|h_C|^2}{2|h_I|^2} \right) + \frac{|h_I|^2}{1 + |h_I|^2} + \frac{|h_S|^2 |h_C|^2}{1 + |h_S|^2}} \right) \geq \log \left(1 + \frac{|h_S|^2}{4(1 + |h_I|^2)^2} \right), \tag{39e}$$

$$R_{A_1} \leq \log \left(1 + \frac{|h_I|^2 \left(\frac{|h_I|^2}{1 + |h_I|^2} - \frac{|h_C|^2}{2|h_I|^2} \right)}{1 + \frac{|h_I|^2}{1 + |h_I|^2} + \frac{|h_S|^2 |h_C|^2}{1 + |h_S|^2}} \right) : \frac{|h_I|^2 \left(\frac{|h_I|^2}{1 + |h_I|^2} - \frac{|h_C|^2}{2|h_I|^2} \right)}{1 + \frac{|h_I|^2}{1 + |h_I|^2} + \frac{|h_S|^2 |h_C|^2}{1 + |h_S|^2}} \geq \frac{|h_I|^2}{4(2 + |h_C|^2)}, \tag{39f}$$

$$R_{B_3} = \log \left(1 + \frac{|h_C|^2 \frac{|h_S|^2}{1 + |h_S|^2}}{1 + \frac{|h_I|^2}{1 + |h_I|^2}} \right) \geq \log \left(1 + \frac{|h_C|^2}{4} \right) \tag{39g}$$

$$\begin{aligned}
\text{gap}_{R_1} &\leq \text{eq. (33a)} - \text{eq. (39h)} - \text{eq. (39b)} - \text{eq. (39c)} \\
&= \log \frac{(1 + 4|h_S|^2)(3 + |h_C|^2)(1 + 2|h_I|^2)}{(1 + |h_S|^2 + 2|h_I|^2)(1 + |h_C|^2/n_{A_2})(3 + |h_C|^2 + |h_I|^2/4)} \leq \log(4 \cdot n_{A_2} \cdot 8), \\
\text{gap}_{R_2} &\leq \text{eq. (33b)} - \text{eq. (39d)} - \text{eq. (39e)} - \text{eq. (39g)} \\
&= \log \frac{4(1 + 1/\sqrt{2})^4 \left(1 + |h_I|^2 + \frac{|h_S|^2}{1 + |h_I|^2}\right)^2 (1 + |h_C|^2)^2}{(1 + 4|h_S|^2) \left(1 + \frac{|h_S|^2}{4(1 + |h_I|^2)^2}\right) (1 + |h_C|^2/n_{B_1})(1 + |h_C|^2/4)} \\
&\leq \log \frac{4(1 + 1/\sqrt{2})^4 \cdot (1 + 2t)^2 \cdot n_{B_1} \cdot 4}{1 + t^2} \Bigg|_{t := \frac{|h_S|^2}{1 + |h_I|^2} \geq t_0, \sqrt{t_0} := \max \left\{ \frac{1 + \sqrt{4/n_{A_2}(3/2 - 2/n_{A_2})}}{1 - 2/n_{A_2}}, \frac{1 + \sqrt{8/n_{B_1}(1 - 2/n_{B_1})}}{1 - 4/n_{B_1}} \right\}} \\
&\leq \log \left(4(1 + 1/\sqrt{2})^4 \cdot 5 \cdot n_{B_1} \cdot 4 \right)
\end{aligned} \tag{41}$$

see (41) for the proof, shown at the top of the page. Thus the proposed scheme achieves a corner point of the capacity region outer bound to within at most $\text{gap} = \max\{\text{eq. (40), eq. (41)}\}$. For example, for $n_{A_2} = n_{B_1} = 10$, we have $t_0 = 9$ and $\text{gap} = 11.7$ bits per user. The gap can be reduced by increasing the value of t_0 .

REFERENCES

- [1] S. Rini, D. Tuninetti, and N. Devroye, "Outer bounds for the interference channel with a cognitive relay," in *Proc. IEEE Inf. Theory Workshop*, Dublin, Ireland, Sep. 2010, pp. 1–5.
- [2] A. Dytso, N. Devroye, and D. Tuninetti, "On the capacity of the symmetric interference channel with a cognitive relay at high SNR," in *Proc. IEEE Int. Conf. Commun.*, 2012, pp. 2350–2354.
- [3] R. Merritt, FCC Gives More Details on Spectrum Plan, Oct. 2010. [Online]. Available: <http://www.eetimes.com/electronics-news/4209892/FCC-gives-more-details-on-spectrum-plan>
- [4] J. Cox, LTE Performance Will Hinge on Picocell Backhaul, Mar. 2011. [Online]. Available: <http://www.networkworld.com/news/2011/032211-ctia-lte-picocell-backhaul.html>
- [5] R. Kumar, A Picocell Primer, Jan. 2006. [Online]. Available: <http://eetimes.com/design/audio-design/4012603/A-picocell-primer>
- [6] 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.
- [7] E. van der Meulen, "A survey of multi-way channels in information theory: 1961–1976," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 1, pp. 1–37, Jan. 1977.
- [8] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 1, pp. 2–14, Jan. 1972.
- [9] O. Sahin and E. Erkip, "Achievable rates for the Gaussian interference relay channel," in *Proc. IEEE GLOBECOM*, Washington, DC, USA, Nov. 2007, pp. 1627–1631.
- [10] O. Sahin, E. Erkip, and O. Simeone, "Interference channel with a relay: Models, relaying strategies, bounds," in *Proc. Workshop Inf. Theory Appl.*, La Jolla, CA, USA, 2009, pp. 90–95.
- [11] P. Razaghi, S. Hong, L. Zhou, W. Yu, and G. Caire, "Two birds and one stone: Gaussian interference channel with a shared out-of-band relay," *Arxiv preprint*, vol. arXiv:1104.0430, 2011. [Online]. Available: <http://arxiv.org/pdf/1104.0430.pdf>
- [12] Y. Tian and A. Yener, "The Gaussian interference relay channel: Improved achievable rates and sum rate upperbounds using a potent relay," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2865–2879, May 2011.
- [13] Y. Tian and A. Yener, "Symmetric capacity of the Gaussian interference channel with an out-of-band relay to within 1.15 bits," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5151–5171, Aug. 2012.
- [14] S. Rini and A. Goldsmith, "On the capacity of the MIMO cognitive interference channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013, pp. 2691–2695.
- [15] O. Sahin and E. Erkip, "On achievable rates for interference relay channel with interference cancellation," in *Proc. Annu. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2007, pp. 1–5.
- [16] S. Sridharan, S. Vishwanath, S. Jafar, and S. Shamai, "On the capacity of cognitive relay assisted Gaussian interference channel," in *Proc. IEEE Int. Symp. Inf. Theory*, 2008, pp. 549–553.
- [17] J. Jiang, I. Maric, A. Goldsmith, and S. Cui, "Achievable rate regions for broadcast channels with cognitive radios," in *Proc. IEEE Inf. Theory Workshop*, Oct. 2009, pp. 500–504.
- [18] E. Telatar and D. Tse, "Bounds on the capacity region of a class of interference channels," in *Proc. IEEE Int. Symp. Inf. Theory*, 2008, pp. 2871–2874.
- [19] A. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.
- [20] S. Rini, D. Tuninetti, and N. Devroye, "The capacity of the interference channel with a cognitive relay in strong interference," in *Proc. IEEE Int. Symp. Inf. Theory*, St. Petersburg, Russia, Aug. 2011, pp. 2632–2636.
- [21] S. Rini, D. Tuninetti, N. Devroye, and A. Goldsmith, "On the capacity of the interference channel with a cognitive relay," *IEEE Trans. Inf. Theory*, vol. 60, no. 4, pp. 2148–2179, Apr. 2014.
- [22] S. Rini, D. Tuninetti, and N. Devroye, "Capacity to within 3 bits for a class of Gaussian interference channels with a cognitive relay," in *Proc. IEEE Int. Symp. Inf. Theory*, St. Petersburg, Russia, Aug. 2011, pp. 2627–2631.
- [23] H. Charmchi, G. A. Hodtani, and M. Nasiri-Kenari, "A new outer bound for a class of interference channels with a cognitive relay and a certain capacity result," *IEEE Commun. Lett.*, vol. 17, no. 2, pp. 241–244, Feb. 2013.
- [24] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.
- [25] S. Rini, D. Tuninetti, and N. Devroye, "On the capacity of the gaussian cognitive interference channel: New inner and outer bounds and capacity to within 1 bit," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 820–848, Feb. 2012.
- [26] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
- [27] A. El Gamal and M. Costa, "The capacity region of a class of deterministic interference channels," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 2, pp. 343–346, Mar. 1982.
- [28] G. Bresler and D. Tse, "The two-user Gaussian interference channel: A deterministic view," *Eur. Trans. Telecommun.*, vol. 19, no. 4, pp. 333–354, Apr. 2008.
- [29] S. Rini, D. Tuninetti, and N. Devroye, "New inner and outer bounds for the discrete memoryless cognitive interference channel and some new capacity results," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4087–4109, Jul. 2011.
- [30] I.-H. Wang and D. Tse, "Interference mitigation through limited transmitter cooperation," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2941–2965, May 2011.
- [31] T. Liu and P. Viswanath, "An extremal inequality motivated by multiterminal information-theoretic problems," *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1839–1851, May 2007.



Alex Dytso received the B.S. degree from the University of Illinois at Chicago in 2011, where he received the International Engineering Consortium's William L. Everitt Student Award of Excellence for outstanding seniors who have demonstrated an interest in the communications field, and where he is currently working toward the Ph.D. degree. His current research topic focuses on multi-user information theory and its applications to wireless networks.



Natasha Devroye received the B.Eng. (honors) in electrical engineering from McGill University in 2001, the M.Sc. degree from Harvard University in 2003, and the Ph.D. degree in engineering sciences from the School of Engineering and Applied Sciences, Harvard University, in 2007. She has been an Assistant Professor in the Department of Electrical and Computer Engineering, University of Illinois at Chicago (UIC), since January 2009. From July 2007 until July 2008, she was a Lecturer at Harvard University. She was a recipient of an NSF CAREER award in 2011, is an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and the IEEE JOURNAL OF SELECTED AREAS OF COMMUNICATIONS (Cognitive Radio Series) and was named UIC's Researcher of the Year in the "Rising Star" category in 2012. Her research focuses on multi-user information theory and applications to cognitive and software-defined radio, radar, and relay and two-way communication networks.



Stefano Rini received the B.A. degree in computer science from the Politecnico di Milano, Italy, in 2005, the M.S. degree in both electrical and computer engineering and statistics from the University of Illinois at Chicago (UIC) in 2009. He received the doctoral degree in electrical and computer engineering at the University of Illinois in Chicago (UIC) in 2011, with adviser Professor Tuninetti and co-adviser Professor Devroye, with a thesis titled "Cognition and cooperation in wireless networks: an information theoretic perspective." In 2012, he

was a Postdoctoral Fellow at the Institute for Communications Engineering, Technical University of Munich (TUM), Germany, with Professor Kramer. In 2013, he was a Postdoctoral Fellow with the Department of Electrical Engineering, Stanford University, with Professor Goldsmith. He is currently an Assistant Professor with the Department of Communication Engineering, National Chiao-Tung University (NCTU), Taiwan. He is an active researcher in multi-user information theory, communication theory, coding theory and computational neuroscience. He received the Best Paper Award at the 2012 APSIPA Annual Summit and Conference and was recognized as an Exemplary Reviewer in 2012 by IEEE COMMUNICATION LETTERS.



Daniela Tuninetti received the M.S. degree in telecommunication engineering from Politecnico di Torino, Turin, Italy, in 1998 and the Ph.D. degree in electrical engineering from ENST/Telecom ParisTech, Paris, France, (with work done at the Eurecom Institute in Sophia Antipolis, France) in 2002. From 2002 to 2004, she was a Postdoctoral Research Associate at the School of Communication and Computer Science, EPFL/Swiss Federal Institute of Technology, Lausanne, Switzerland. Since January 2005, she is with the Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, IL, USA, where she currently is an Associate Professor. She was the Editor-in-Chief of the IEEE Information Theory Society Newsletter from 2006 to 2008, an Associate Editor for the IEEE COMMUNICATION LETTERS from 2006 to 2009, and an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2009 to 2014. She currently is an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY. She regularly serves on the Technical Program Committees of IEEE workshops and conferences and she was the Communication Theory Symposium Co-chair of the 2010 IEEE International Conference on Communications (ICC 2010). She received the Best Student Paper Award at the European Wireless Conference in 2002 and was the recipient of an NSF CAREER award in 2007. Her research interests are in the ultimate performance limits of wireless interference networks, with special emphasis on cognition and user cooperation.