

Inner and Outer Bounds for the Gaussian Cognitive Interference Channel and New Capacity Results

Stefano Rini, Daniela Tuninetti, and Natasha Devroye

Abstract—The capacity of the Gaussian cognitive interference channel, a variation of the classical two-user interference channel where one of the transmitters (referred to as *cognitive*) has knowledge of both messages, is known in several parameter regimes but remains unknown in general. This paper provides a comparative overview of this channel model as it proceeds through the following contributions. First, several outer bounds are presented: a) a new outer bound based on the idea of a broadcast channel with degraded message sets, and b) an outer bound obtained by transforming the channel into channels with known capacity. Next, a compact Fourier–Motzkin eliminated version of the largest known inner bound derived for the discrete memoryless cognitive interference channel is presented and specialized to the Gaussian noise case, where several simplified schemes with jointly Gaussian input are evaluated in closed form and later used to prove a number of results. These include a new set of capacity results for: a) the “primary decodes cognitive” regime, a subset of the “strong interference” regime that is not included in the “very strong interference” regime for which capacity was known, and b) the “S-channel in strong interference” in which the primary transmitter does not interfere with the cognitive receiver and the primary receiver experiences strong interference. Next, for a general Gaussian channel the capacity is determined to within one bit/s/Hz and to within a factor two regardless of the channel parameters, thus establishing rate performance guarantees at high and low SNR, respectively. The paper concludes with numerical evaluations and comparisons of the various simplified achievable rate regions and outer bounds in parameter regimes where capacity is unknown, leading to further insight on the capacity region.

Index Terms—Broadcast channel with degraded message sets, capacity in the primary decodes cognitive regime, capacity for the Z-channel in strong interference, capacity to within one bit, capacity to within a factor of two, cognitive interference channel, inner bound, outer bound.

I. INTRODUCTION

A well studied channel model inspired by the newfound abilities of cognitive radio technology and its potential impact on spectral efficiency in wireless networks is the *cognitive radio channel* [4]. The cognitive radio channel is also referred to as the *interference channel with unidirectional*

cooperation [5], the *interference channel with degraded message sets* [6], or the *cognitive interference channel* [7]. This channel consists of a two-user interference channel [8] where one transmitter-receiver pair is referred to as the *primary* user and the other as the *cognitive* user. The primary transmitter has knowledge of one of the two independent messages to be sent, while the cognitive transmitter has full, non-causal knowledge of both messages, thus idealizing the cognitive user’s ability to detect transmissions taking place in the network.¹ Since the cognitive transmitter can “broadcast” information to both receivers, the capacity of the cognitive interference channel contains features of both the interference and the broadcast channel [10].

A. Past Work

Capacity results. The cognitive interference channel was first posed in an information theoretic framework in [4], where an achievable rate region (for general discrete memoryless channels) and a outer bound based on a broadcast channel idea (in Gaussian noise) were proposed. The first capacity results were determined in [11], [6] for a class of channels with “weak interference” at the primary receiver. In this regime, capacity is achieved by having the cognitive transmitter pre-code against the interference created at its receiver, while the primary receiver treats the interference from the cognitive transmitter as noise. Capacity is also known in the “very strong interference” regime [12]. In this regime, capacity is achieved by having both receivers decode both messages as in a compound multiple access channel. In [13, Th. 7.1], we showed that the outer bound of [6] is achievable in the “better cognitive decoding” regime, which includes both the “very weak interference” and the “very strong interference” regimes.

Outer bounds. An outer bound for a general cognitive interference channel was derived in [14, Th. 4] using a technique developed for the general broadcast channel in [15]. Both the “weak interference” outer bound of [6] and the “strong interference” outer bound of [12] may be derived by loosening [14, Th. 4]. Although the outer bound in [14, Th. 4] is the tightest known, it is difficult to evaluate because it contains three auxiliary random variables for which no cardinality bounds are given on the corresponding alphabets. Moreover, for the Gaussian channel, the “Gaussian maximizes entropy” property

Manuscript received October 27, 2010; revised September 07, 2011; accepted September 08, 2011. Date of publication January 9, 2012; date of current version February 08, 2012. The material in this paper was presented in part at the International Zurich Seminar on Communications, Zurich, Switzerland, March 2010, the Allerton Conference on Communication, Control, and Computing, September 2010, and at the 2011 IEEE International Conference on Communications.

The authors are with the Department of Electrical and Computer Engineering University of Illinois-Chicago, Chicago IL 60607 USA (e-mail: srini2@uic.edu; danielat@uic.edu; devroye@uic.edu).

Communicated by S. A. Jafar, Associate Editor for Communications.
Digital Object Identifier 10.1109/TIT.2011.2168382

¹Although the assumption of full, non-causal knowledge of the primary user’s message at the cognitive transmitter might not be practical, the simplicity of the resulting model leads to closed form results and provides powerful insight on the role of unilateral cooperation among the users. The more practical scenario of causal unilateral cooperation may be studied in the framework of the *interference channel with generalized feedback* (see [9] and references therein), but is outside the scope of this work.

[16] alone does not suffice to show that jointly Gaussian inputs exhaust the outer bound. For these reasons, in [13, Th. 4.1] we proposed an outer bound that exploits the fact that the capacity region only depends on the conditional marginal distributions of the outputs given the inputs (as for broadcast channels [17]) since the receivers do not cooperate. The resulting outer bound does not include auxiliary random variables and every mutual information term involves all the inputs (like in the cut-set bound [16]) and thus may be evaluated for general channels including the Gaussian channel. The bound in [13, Th. 4.1] was shown to be tight for a class of semi-deterministic cognitive interference channels with a noiseless output at the primary receiver.

Achievable rate regions. Different achievable schemes have been proposed for the cognitive interference channel which include features originally devised for the interference channel and for the broadcast channel, such as rate splitting, superposition coding, binning and simultaneous decoding. The scheme of [18] generalized the “weak interference” capacity achieving scheme of [6] by making part of the cognitive message common. The same rate splitting idea was used in [14] along with a more elaborated binning operation. The region in [19] introduced a binning scheme inspired by Marton’s achievable rate region for a general broadcast channel [20]. This feature was further generalized in [21] and in [1] where more refined binning and superposition steps were added in the cognitive encoding process. Given the different encoding choices, a comparison of the different achievable schemes is often not straightforward. In particular, despite possible simplifications of the original scheme in [4] as described in [22], no region was shown to conclusively encompass [4], or the larger region of [23], until recently. A comparison of all the transmission schemes proposed in the literature was presented in [13], in which we showed that our region in [13, Th. 5.1] is provably the largest known achievable rate region to date.

Capacity to within a constant gap. While the capacity region remains unknown for a general channel, in [7] we demonstrated achievable rate regions which lie within 1.87 bits/s/Hz for any real-valued Gaussian cognitive interference channel. We derived this constant gap result by using insights from the high SNR deterministic approximation of the Gaussian cognitive interference channel [24], a deterministic model that captures the behavior of a Gaussian network for large transmit powers [25].

Z-channel. The special case where the cognitive transmitter does not create interference to the primary receiver is called the *Z cognitive interference channel*; inner and outer bounds when the cognitive-primary link is noiseless are obtained in [26], [19]. The Gaussian causal case is considered in [19], and is related to the general causal cognitive [27], [9]. For the case where the primary transmitter does not create interference to the cognitive receiver (also known as the S-channel) capacity is known in “weak interference” [11], [6]. For the “strong interference” regime, independently and concurrently to the submission of the authors’ conference papers [2], [28] on the capacity of the Gaussian Z-cognitive interference channel, similar results appeared in the online work [29], [30].

B. Contributions

In this work, we focus on the *Gaussian* cognitive interference channel in a comprehensive and comparative manner. In particular, our main contributions are:

- 1) **We evaluate the outer bound of [13, Th. 4.1] for the Gaussian channel.** We show that it unifies the previously proposed outer bounds for the “weak interference” and the “strong interference” regimes of [6] and [14], respectively.
- 2) **We derive a new outer bound based on the broadcast channel with degraded message sets.** The capacity region of the Gaussian MIMO (multiple-input multiple-output) broadcast channel with degraded message sets is an outer bound for the cognitive interference channel in “strong interference”. We show that the new bound may be strictly tighter than the “strong interference” outer bound of [14].
- 3) **Derive a new outer bound by transformation or inclusion into channels with known capacity.** We determine the conditions under which the capacity region of a Gaussian channel is contained in that of a channel with known capacity. The capacity of the latter channel thus provides an outer bound for the former.
- 4) **We specialize the largest known inner bound of [13, Th. 5.1] to the Gaussian channel.** We first present a compact Fourier–Motzkin eliminated version of our original achievable rate region (provably the largest known to date). We utilize this as a unified framework to derive and compare various achievable schemes for specific jointly Gaussian input. These schemes are used to show achievability in our capacity results and numerical comparisons with outer bounds.
- 5) **We prove a new capacity result for the “primary decodes cognitive” regime.** This regime is a subset of the “strong interference” regime that is not included in the “very strong interference” regime for which capacity was known [12]. In this regime capacity is achieved by having the primary receiver decode the message of the cognitive user in addition to its own message, as the name suggests.
- 6) **We prove a new capacity result for the S-channel in strong interference.** In the S-channel the primary transmission does not interfere at the cognitive receiver. For this channel we show the achievability of our outer bound based on the capacity of the broadcast channel with degraded message sets.
- 7) **We show capacity to within one bit/s/Hz and to within a factor two.** These two results characterize the capacity region of the Gaussian channel at high and low SNR, respectively. To this end, we use a transmission scheme inspired by the capacity achieving scheme for the semi-deterministic cognitive interference channel of [13, Th. 8.1]. The multiplicative gap is shown by using a simple time sharing argument between achievable rate pairs.
- 8) **We provide insights on the capacity region of the Gaussian channel for the regimes in which capacity is still unknown.** We do so by showing that very simple transmission strategies can achieve capacity to within a constant gap for large sets of parameters. We conclude by showing that a constant gap result may alternatively be

proved by trading off interference pre-coding at the cognitive encoder and interference decoding at the primary receiver.

C. Paper Organization

The rest of the paper is organized as follows. Section II formally defines the cognitive interference channel model and summarizes known results for the Gaussian channel. Section III presents new outer bounds for the Gaussian channel. Section IV gives a compact Fourier–Motzkin eliminated version of the largest known achievable region and specialize it to the Gaussian noise case, where several simplified schemes with jointly Gaussian input are evaluated in closed form. Section V proves the two new capacity results. Section VI characterizes the capacity of the Gaussian channel to within one bit/s/Hz and to within a factor two. Section VII shows some relevant numerical results. Section VIII concludes the paper. Most of the proofs may be found in the Appendixes.

II. GAUSSIAN CHANNEL MODEL AND KNOWN RESULTS

A. Notation

We use the following convention:

- The symbol $X \sim \mathcal{N}_{\mathbb{C}}(\mu, \sigma^2)$ indicates that the random variable (RV) X is a complex-valued proper Gaussian RV with mean μ and covariance σ^2 .
- We define $\mathcal{C}(x) := \log(1 + x)$ for $x \in \mathbb{R}^+$.
- We define $[x]^+ := \max\{0, x\}$ for $x \in \mathbb{R}$.
- We define $\bar{x} := 1 - x$ for $x \in [0, 1]$.
- We use $[1 : n]$ to denote the set of natural numbers from 1 to n .
- The notation $A \stackrel{(n)}{=} B$ to indicate that the expression B is obtained from A with the assignment of variables given in equation number n .
- The notation $(n) \leq (m)$ indicates that the expression (usually on the right hand side of) equation number n is smaller or equal than the expression (usually on the right hand side of) equation number m .
- For an integer N , the symbol X^N indicates the length- N vector (X_1, \dots, X_N) .
- For the plots, the logarithms are in base 2, i.e., rates are expressed in bits/s/Hz.

B. General Memoryless Cognitive Interference Channel

A two-user Interference Channel (IFC) is a multiterminal network with two input alphabets \mathcal{X}_1 and \mathcal{X}_2 , two output alphabets \mathcal{Y}_1 and \mathcal{Y}_2 , and a channel transition probability $P_{Y_1 Y_2 | X_1 X_2}(y_1, y_2 | x_1, x_2) : \mathcal{Y}_1 \times \mathcal{Y}_2 \rightarrow [0, 1]$ for all $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$. Each transmitter $i, i \in \{1, 2\}$, wishes to communicate a message W_i , uniformly distributed on $[1 : 2^{NR_i}]$, to receiver i in N channel uses at rate R_i . The two messages are independent. In the classical IFC, the two transmitters operate independently having no knowledge of each others' messages. Here we consider a variation of this setup assuming that transmitter 1, in addition to its own message, also knows the message of transmitter 2 prior to transmission. We

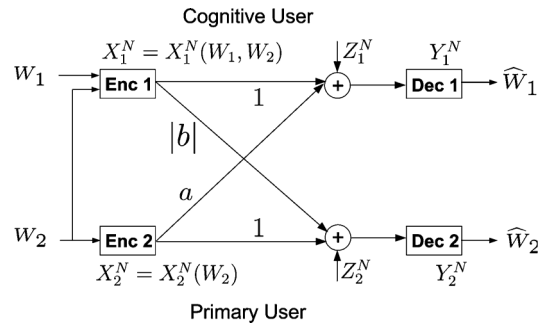


Fig. 1. Gaussian cognitive interference channel (G-CIFC).

refer to transmitter/receiver 1 as the *cognitive* pair and to transmitter/receiver 2 as the *primary* pair. This model is commonly known as the Cognitive Interference Channel (CIFC).

The CIFC is an idealized model for the *unilateral source cooperation* of transmitter 1 with transmitter 2. The receivers however do not cooperate. This implies that the capacity region of the CIFC, similar to the broadcast channel (BC) [17], only depends on the output conditional marginal distributions $P_{Y_1 | X_1 X_2}$ and $P_{Y_2 | X_1 X_2}$, and not on the output joint marginal distribution $P_{Y_1 Y_2 | X_1 X_2}$.

A non-negative rate pair (R_1, R_2) is achievable if there exist a sequence of encoding functions

$$\begin{aligned} X_1^N &= X_1^N(W_1, W_2), \\ X_2^N &= X_2^N(W_2) \end{aligned}$$

and a sequence of decoding functions

$$\widehat{W}_i = \widehat{W}_i(Y_i^N), \quad i \in \{1, 2\}$$

such that the probability of error satisfies

$$\max_{i \in \{1, 2\}} \mathbb{P}[\widehat{W}_i \neq W_i] \rightarrow 0, \quad N \rightarrow \infty.$$

The capacity region is defined as the convex closure of the region of achievable (R_1, R_2) -pairs [16].

C. Gaussian CIFC

A Gaussian CIFC (G-CIFC) in *standard form* (see Appendix A) is described by the input/output relationship

$$Y_1 = X_1 + aX_2 + Z_1, \quad (1a)$$

$$Y_2 = |b|X_1 + X_2 + Z_2 \quad (1b)$$

where the channel gains a and b are complex-valued, constant, and known to all terminals, the channel inputs are subject to the average power constraint

$$\mathbb{E}[|X_i|^2] \leq P_i, \quad P_i \in \mathbb{R}^+, \quad i \in \{1, 2\} \quad (1c)$$

and the channel noise $Z_i \sim \mathcal{N}_{\mathbb{C}}(0, 1), i \in \{1, 2\}$. Since the capacity only depends on the output conditional marginal distributions, the correlation coefficient among Z_1 and Z_2 is irrelevant. The capacity of the channel in (1) is indicated in the rest of the paper by $\mathcal{C}(a, |b|, P_1, P_2)$. A graphical representation of a G-CIFC is found in Fig. 1.

A G-CIFC is said to be a

- **Z-channel** if $|b| = 0$ and referred to as the Z-G-CIFC. In this case the primary decoder does not experience interference from the cognitive transmitter. Capacity is given by

$$C(a, 0, P_1, P_2) = \{R_1 \leq C(P_1), R_2 \leq C(P_2)\}.$$

The Z-G-CIFC is a channel with “weak interference” ($|b| \leq 1$) for which capacity is known [11], [6].

- **S-channel** if $a = 0$ and referred to as the S-G-CIFC. In this channel the cognitive decoder does not experience interference from the primary transmitter. For this channel capacity is only known in “weak interference” ($|b| \leq 1$) [11], [6].
- **Degraded channel** if $a|b| = 1$. In this case one channel output is a degraded version of the other. In particular, in “strong interference” ($|b| > 1$) Y_1 is a degraded version of Y_2 since

$$Y_1 = X_1 + \frac{1}{|b|}X_2 + Z_1 \sim \frac{1}{|b|}Y_2 + Z_0$$

for $Z_0 \sim \mathcal{N}_{\mathbb{C}}(0, |b|^2 - 1)$ independent of everything else. Similarly, in “weak interference” ($|b| \leq 1$) Y_2 is a degraded version of Y_1 . Capacity is known in “weak interference” ($|b| \leq 1$) [11], [6].

D. Known Results for the G-CIFC

The capacity of the G-CIFC is not known in general. However some capacity results exist, as summarized next.

Theorem II.1: “Weak interference” capacity of [6, Lemma 3.6] and [11, Th. 4.1]. If

$$|b| \leq 1, \text{ (the “weak interference” regime/condition)} \quad (2)$$

the capacity $C(a, |b|, P_1, P_2)$ is

$$R_1 \leq C(\alpha P_1), \quad (3a)$$

$$R_2 \leq C\left(|b|^2 P_1 + P_2 + 2\sqrt{\alpha}|b|^2 P_1 P_2\right) - C(|b|^2 \alpha P_1) \quad (3b)$$

taken over the union of all $\alpha \in [0, 1]$.

Theorem II.2: “Strong interference” outer bound of [14, Th. 4]. When

$$|b| > 1, \text{ (the “strong interference” regime/condition)} \quad (4)$$

the capacity $C(a, |b|, P_1, P_2)$ is included into the region $\mathcal{O}^{(SI)}$ defined as

$$R_1 \leq C(\alpha P_1), \quad (5a)$$

$$R_1 + R_2 \leq C(|b|^2 P_1 + P_2 + 2\sqrt{\alpha}|b|^2 P_1 P_2) \quad (5b)$$

taken over the union of all $\alpha \in [0, 1]$.

Theorem II.3: “Very strong interference” capacity of [12, Th. 6] extended to complex-valued channels (see Appendix B). When

$$(|a|^2 - 1)P_2 - (|b|^2 - 1)P_1 - 2|a - |b||\sqrt{P_1 P_2} \geq 0,$$

and $|b| > 1$ (“very strong interference” regime/condition) (6)

the capacity $C(a, |b|, P_1, P_2)$ coincides with the outer bound $\mathcal{O}^{(SI)}$ in (5) of Theorem II.2.

A plot of the capacity results of Theorem II.1 and Theorem II.3 for $a \in \mathbb{R}$ and $P_1 = P_2$ is depicted in Fig. 2. The channel gains a and $|b|$ for which capacity is known are shaded, while those for which capacity is unknown are white; in these cases capacity is known to within a constant gap [7] as further elaborated on in Section VI.

Remark on naming convention. We note that our naming convention is not entirely consistent with past uses of the term “strong/weak interference”. Here, as in our previous work on the CIFC [13], we use “strong/weak interference” to denote regimes inspired by similar results for the IFC under which we may obtain either a tighter or simpler outer bound for the channel of interest, and use the terms “very strong/very weak” to denote regimes in which additional conditions (therefore forming subsets of the “strong/weak” regimes) are imposed on top of the “strong/weak” conditions that allow these outer bounds to be achieved. We note that this distinction between “weak” and “very weak” is not needed at this point for the G-CIFC.

III. OUTER BOUNDS

In this section we prove several outer bounds.

- 1) First we evaluate the outer bound of [13, Th. 4.1] for the Gaussian channel and show that it coincides with the outer bounds of Theorem II.1 and Theorem II.2 in “weak” ($|b| \leq 1$) and “strong interference” ($|b| > 1$), respectively.
- 2) Then we tighten it by using the observation of [4] that the capacity region of a G-CIFC is included into the capacity region of the Gaussian MIMO BC obtained by allowing full cooperation among the transmitters. We further tighten the outer bound in “strong interference” ($|b| > 1$), where we show that the capacity region of a Gaussian broadcast channel with degraded message sets forms an outer bound to the capacity of the G-CIFC.
- 3) Finally, we propose an outer bound based on enhancing the original channel so as to transform it into a channel for which capacity is known.

A. A Unifying Framework for Theorem II.1 and Theorem II.2

Our objective is to obtain an outer bound for $C(a, |b|, P_1, P_2)$ in the “strong interference” ($|b| > 1$) regime that improves on the “strong interference” outer bound $\mathcal{O}^{(SI)}$ in (5) of Theorem II.2. Although the following theorem does not result in such a bound, it is of interest because it provides a simple unifying framework for Theorem II.1 and Theorem II.2, whose proof techniques are quite different. On the one hand, the bound in Theorem II.1 is valid for a general channel under the “weak interference” condition in [6, Th. 3.7] and is inspired by the converse for “more capable BC” [31]. On the other hand, the bound in Theorem II.2 is valid for Gaussian channels with “strong interference” only and is inspired by the converse of IFC in “strong interference” [32]–[34]. We will show next that both results may be derived within the framework proposed in [13].

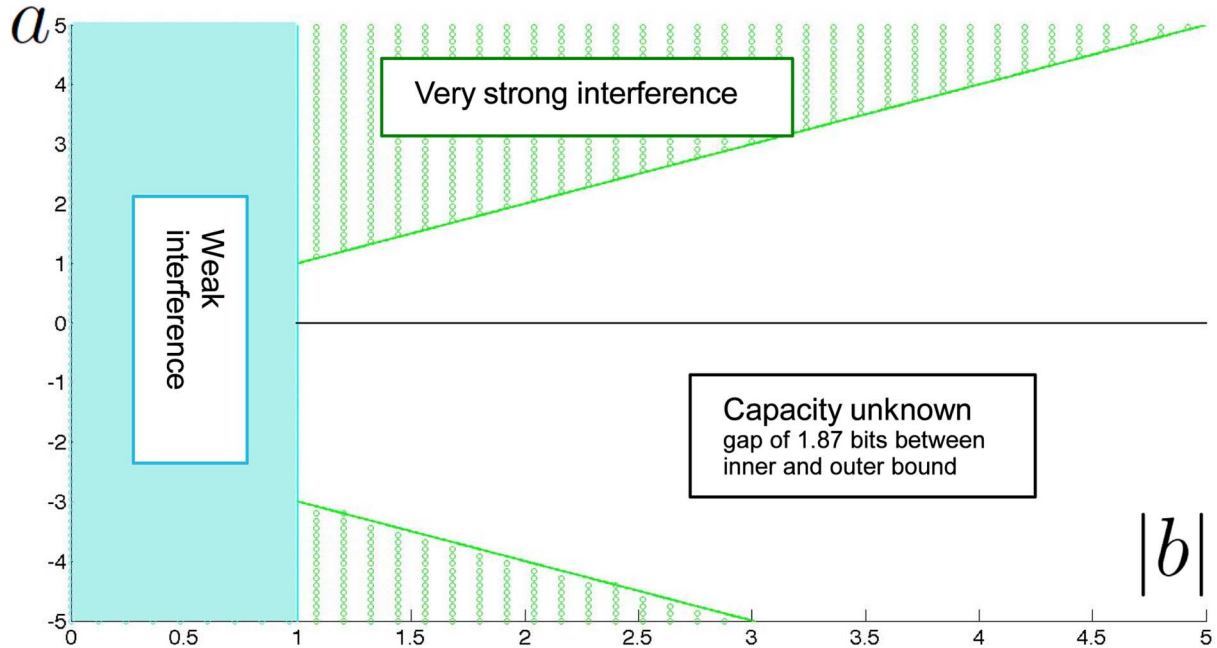


Fig. 2. Representation of known results on the capacity of the G-CIFC prior to this work for $P_1 = P_2 = 10$ and $(a, |b|) \in [-5, 5] \times [0, 5]$. The set of channel parameters for which capacity is known are shaded (corresponding to Theorem II.1 and Theorem II.3), while those for which capacity is still unknown are white (in these cases capacity is known to within 1.87 bits [7]).

The proof of [13, Th. 4.1] uses the argument originally devised by Sato for the BC [17] that, for channels without receiver cooperation, the capacity only depends on the output conditional marginal distributions. The bound in [13, Th. 4.1] is valid for a general CIFC.

Theorem III.1: Unifying outer bound. The capacity region $\mathcal{C}(a, |b|, P_1, P_2)$ is contained into the region

$$R_1 \leq \mathcal{C}(\alpha P_1), \quad (7a)$$

$$R_2 \leq \mathcal{C}(|b|^2 P_1 + P_2 + 2\sqrt{\alpha}|b|^2 P_1 P_2), \quad (7b)$$

$$R_1 + R_2 \leq \mathcal{C}(|b|^2 P_1 + P_2 + 2\sqrt{\alpha}|b|^2 P_1 P_2) + [\mathcal{C}(\alpha P_1) - \mathcal{C}(|b|^2 \alpha P_1)]^+ \quad (7c)$$

taken over the union of all $\alpha \in [0, 1]$. In “strong interference” ($|b| > 1$) the region in (7) reduces to Theorem II.2, and in “weak interference” ($|b| \leq 1$) to Theorem II.1.

Proof: In [13, Th. 4.1], we showed that the capacity of a general CIFC is contained in the region \mathcal{O}^{RTD} defined as

$$R_1 \leq I(Y_1; X_1 | X_2), \quad (8a)$$

$$R_2 \leq I(X_1, X_2; Y_2), \quad (8b)$$

$$R_1 + R_2 \leq I(X_1, X_2; Y_2) + I(Y_1; X_1 | Y_2, X_2) \quad (8c)$$

taken over the union of all joint distributions P_{X_1, X_2} and where the sum-rate bound in (8c) can be tightened by choosing the worst conditional joint distribution $P_{Y_1, Y_2 | X_1, X_2}$ with the correct conditional marginal distributions $P_{Y_1 | X_1, X_2}$ and $P_{Y_2 | X_1, X_2}$. For the G-CIFC this latter “same marginals” condition amounts to optimizing the sum-rate with respect to the correlation coefficient between the Gaussian additive noises, that is, with respect to $\gamma := \mathbb{E}[Z_1 Z_2^*] \in \mathbb{C}$ such that $|\gamma| \leq 1$.

First we show that a proper-complex jointly Gaussian input exhausts the region in (8). It is clear that the optimal input must be zero-mean and meet the power constraint with equality. Let ρ

parameterize the correlation coefficient between the inputs, that is, $\mathbb{E}[X_1 X_2^*] := \rho \sqrt{P_1 P_2}$ such that $|\rho| \leq 1$. Let (X_{1G}, X_{2G}) be a zero-mean jointly Gaussian input with covariance matrix

$$\mathbf{S} := \begin{bmatrix} P_1 & \rho \sqrt{P_1 P_2} \\ \rho^* \sqrt{P_1 P_2} & P_2 \end{bmatrix}, \quad |\rho| \leq 1. \quad (9)$$

By using the “Gaussian maximizes entropy” principle (see [16] and also [35, Eq. (3.29)]), we conclude that, for a given input covariance \mathbf{S} in (9), the region \mathcal{O}^{RTD} in (8) is outer bounded by

$$(8a) \leq I(Y_1; X_{1G} | X_{2G})$$

$$= (7a)|_{\alpha=1-|\rho|^2}$$

$$(8b) \leq I(Y_2; X_{1G}, X_{2G})$$

$$= \log(1 + P_2 + |b|^2 P_1 + 2|b|\Re\{\rho\} \sqrt{P_1 P_2})$$

$$\leq (7b)|_{\alpha=1-|\rho|^2},$$

$$(8c) \leq I(Y_2; X_{1G}, X_{2G}) + \min_{|\gamma| \leq 1} I(Y_1; X_{1G} | Y_2, X_{2G})$$

$$\leq (7b) + \min_{|\gamma| \leq 1} \log \left(\frac{1 + (1 - |\rho|^2) P_1 \frac{|b|^2 + 1 - 2|b|\Re\{\gamma\}}{1 - |\gamma|^2}}{1 + |b|^2 (1 - |\rho|^2) P_1} \right)$$

$$= (7c)|_{\alpha=1-|\rho|^2}. \quad (10)$$

This shows that a jointly Gaussian input is optimal in (8). Note that the minimizing γ in (10) is

$$\arg \min_{|\gamma| \leq 1} \frac{|b|^2 + 1 - 2|b|\Re\{\gamma\}}{1 - |\gamma|^2} = \min \left\{ |b|, \frac{1}{|b|} \right\},$$

that is, the worst conditional marginal distribution is such that one of Y_1 or Y_2 is the degraded version of the other when conditioned on X_2 .

Finally, in “strong interference” ($|b| > 1$) the region in (7) reduces to Theorem II.2 because the bound in (7b) is redundant due to (7c), while in “weak interference” ($|b| \leq 1$) it reduces to

Theorem II.1 because the closure of the region is determined by the rates pairs for which (7a) and (7c) are met with equality as argued in [36, Ex. 4.3]. ■

B. BC-Based Outer Bounds

In this subsection we propose an outer bound that is tighter than the “strong interference” outer bound $\mathcal{O}^{(SI)}$ in (5) in the “strong interference” regime ($|b| > 1$). The following observation is key: if we provide the primary transmitter with the cognitive message, the G-CIFC becomes a Gaussian MIMO BC with private rates (BC-PR), with two antennas at the transmitter and one antenna at each receiver and where the input is subject to a per-antenna power constraint, as originally used in [4, p. 1819].

Theorem III.2: BC-PR-based outer bound for the general CIFC. The capacity of a general CIFC (not necessarily Gaussian) is contained into the region

$$\mathcal{O}^{(BC-PR)} \cap \mathcal{O}^{(RTD)} \quad (11)$$

where $\mathcal{O}^{(BC-PR)}$ is the capacity region (or an outer bound) for the BC with private rates only obtained by allowing the transmitters of the CIFC to fully cooperate and where $\mathcal{O}^{(RTD)}$ is the outer bound in [13, Th. 4.1] given in (8).

Proof: The theorem follows from the fact that allowing transmitter cooperation enlarges the capacity region of the CIFC and results in a BC. ■

For the G-CIFC, the closed form expression of $\mathcal{O}^{(BC-PR)}$ is given in [37] and is reported in Appendix C for completeness, while $\mathcal{O}^{(RTD)}$ is given in (7). The BC-based outer bound of Theorem III.2 may not only be completely inside $\mathcal{O}^{(SI)}$ in (5) but it can actually be capacity. As an example of this, consider the G-CIFC with “strong interference” ($|b| > 1$) and with $P_2 = 0$; this channel is equivalent to a (degraded) BC with input X_1 whose capacity is $\mathcal{O}^{(BC-PR)}$ [38]; thus we have

$$\begin{aligned} & \mathcal{C}(a, |b|, P_1, 0) \\ &= \bigcup_{\alpha \in [0,1]} \left\{ R_1 \leq \mathcal{C} \left(\frac{\alpha P_1}{\bar{\alpha} P_1 + 1} \right), R_2 \leq \mathcal{C}(\alpha |b|^2 P_1) \right\} \\ &= \mathcal{O}^{(BC-PR)} \Big|_{P_2=0} \\ &\subset \bigcup_{\alpha \in [0,1]} \left\{ R_1 \leq \mathcal{C}(P_1), R_2 \leq \mathcal{C}(|b|^2 P_1) \right\} = \mathcal{O}^{(SI)} \Big|_{P_2=0}. \end{aligned}$$

Theorem III.2 is valid for a general CIFC. It may be further tightened for the Gaussian channel in “strong interference” ($|b| > 1$). As previously noted in [14, Sec. 6.1], in the “strong interference” regime ($|b| > 1$) there is no loss of optimality in having the primary receiver decode the cognitive message in addition to its own message. Indeed, after decoding W_2 , receiver 2 can reconstruct $X_2^N(W_2)$ and compute the following estimate of the receiver 1 output

$$\tilde{Y}_1^N := \frac{Y_2^N - X_2^N}{|b|} + aX_2^N + \sqrt{1 - \frac{1}{|b|^2}} Z_0^N \sim Y_1^N \quad (12)$$

where $Z_0^N \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{I})$ and independent of everything else. Hence, if receiver 1 can decode W_1 from Y_1^N , so can receiver

2 from \tilde{Y}_1^N . For this reason the capacity region of the G-CIFC for $|b| > 1$ is unchanged if receiver 2 is required to decode both messages. If we further allow the two transmitters to fully cooperate, the resulting channel is a *Gaussian MIMO BC with degraded message sets* (BC-DMS), with per-antenna power constraint, where message W_2 is to be decoded at receiver 2 only and message W_1 at both receivers.² This implies that the bound in Theorem III.2 may be tightened for the G-CIFC in “strong interference” ($|b| > 1$) by using the capacity of the Gaussian MIMO BC-DMS instead of the capacity of the Gaussian MIMO BC-PR:

Theorem III.3: BC-DMS-based outer bound for the Gaussian CIFC in strong interference. The capacity of a G-CIFC in “strong interference” ($|b| > 1$) satisfies

$$\mathcal{C}(a, |b|, P_1, P_2) \subseteq \mathcal{O}^{(BC-DMS)} \cap \mathcal{O}^{(SI)}, \quad (13)$$

where $\mathcal{O}^{(BC-DMS)}$ is the capacity of the *Gaussian MIMO BC with degraded message sets* determined in [39] and $\mathcal{O}^{(SI)}$ is the “strong interference” outer bound given in (5).

The analytical evaluation of the outer bound region in (11) of Theorem III.2 and in (13) of Theorem III.3 is quite involved in general. Our contribution here is to determine an expression for the capacity region of the Gaussian MIMO BC-DMS that is simpler than the one in [39]. In particular, in Appendix D we prove the optimality of Gaussian inputs by directly using the region of [40]. We further simplify the BC-DMS-based outer bound so as to have only one free parameter and obtain closed form expressions for the degraded G-CIFC and for the S-G-CIFC.

Corollary III.4: BC-DMS-based outer bound for the degraded G-CIFC in strong interference. For a degraded G-CIFC with $1/a = |b| > 1$, the outer bounds of Theorem III.2 and of Theorem III.3 coincide and reduce to

$$R_1 \leq \mathcal{C}(\alpha P_1), \quad (14a)$$

$$R_2 \leq \mathcal{C} \left(\frac{P_2 + \bar{\alpha} |b|^2 P_1 + 2\sqrt{|b|^2 P_1 P_2}}{1 + \alpha P_1} \right), \quad (14b)$$

$$R_1 + R_2 \leq \mathcal{C}(P_2 + |b|^2 P_1 + 2\sqrt{\bar{\alpha} |b|^2 P_1 P_2}) \quad (14c)$$

for $\alpha \in [0, 1]$. Moreover, the R_2 -bound from the MIMO BC-DMS capacity region (in (14b)) is more stringent than the R_2 -bound from the “strong interference” outer bound (from the difference of (14c) and (14a)) if

$$|b| \geq \sqrt{\frac{P_2}{P_1}} + \sqrt{1 + \frac{P_2}{P_1}}. \quad (15)$$

Proof: See Appendix D. ■

Corollary III.5: BC-DMS-based outer bound for the S-G-CIFC in strong interference. For a S-G-CIFC with $a = 0$ and $|b| \geq 1$ the outer bound of Theorem III.3 is contained into the region

²That the capacity of the general BC-DMS is an outer bound for the capacity of the general CIFC in “strong interference” was also pointed out in the independent work [29]. An alternative way to derive the outer bound of [29] for a general CIFC is by loosening the outer bound in [14, Th. 4] by dropping [14, eq. (33)] and letting $U = [V, U_1]$ in [14, Th. 4].

$$R_1 \leq \mathcal{C}(\alpha P_1), \quad (16a)$$

$$R_2 \leq \mathcal{C}\left(\left(\sqrt{\frac{\bar{\alpha}}{1 + \alpha P_1}}|b|^2 P_1 + \sqrt{P_2}\right)^2\right), \quad (16b)$$

$$R_1 + R_2 \leq \mathcal{C}(P_2 + |b|^2 P_1 + 2\sqrt{\bar{\alpha}}|b|^2 P_1 P_2) \quad (16c)$$

for $\alpha \in [0, 1]$. Moreover, the R_2 -bound from the MIMO BC-DMS capacity region (from (16b)) is more stringent than the R_2 -bound from the “strong interference” outer bound (from the difference of (16c) and (16a)) if

$$|b| \geq \sqrt{1 + P_2}. \quad (17)$$

Proof: See Appendix D. ■

C. Outer Bounds by Transformation

Further outer bounds for the G-CIFC may be obtained by transforming the original G-CIFC into a different channel for which capacity is known. In the transformed channel the transmitters can reproduce the channel outputs of the original channel: this ensures that the transformation enlarges the capacity region thus providing an outer bound for the original channel.

Theorem III.6: Outer bound by channel transformation or inclusion. The capacity region $\mathcal{C}(a, |b|, P_1, P_2)$ is contained into the region

$$\bigcap_{(A,B,C) \in \mathbb{C}^3: |A| \geq 1, |AC/(A-Bb)| \geq 1} \mathcal{C}\left(\frac{aA-B}{C}, \left|\frac{Cb}{A-Bb}\right|, (\sqrt{|A|^2 P_1} + \sqrt{|B|^2 P_2})^2, |C|^2 P_2\right).$$

Proof: See Appendix E. ■

The outer bound in Theorem III.6 may be used to transform the given G-CIFC into channels with known capacity regions. In particular, through careful choice of the parameters (A, B, C) one may show that $\mathcal{C}(a, |b|, P_1, P_2)$ is included into those of

- 1) a S-G-CIFC: by imposing $\frac{aA-B}{C} = 0$,
- 2) a G-IFC in “weak interference”: by imposing $\left|\frac{Cb}{A-Bb}\right| \leq 1$, and
- 3) a G-CIFC in “very strong interference”: by imposing $\frac{aA-B}{C} = \left|\frac{Cb}{A-Bb}\right|$ and $(\sqrt{|A|^2 P_1} + \sqrt{|B|^2 P_2})^2 = |C|^2 P_2$, since the “very strong interference” condition in (6) is trivially verified by $a = |b|$ and $P_1 = P_2$.

This observation will be used in the numerical evaluations later on, where we see that this outer bound is tighter than some of the other individual bounds such as the “strong interference” and BC-based outer bounds for certain rate pairs.

IV. INNER BOUNDS

In [13, Th. 4.1], we introduced the largest known achievable rate region to date for the general CIFC, which employed new

transmission features that were crucial in proving capacity for the semi-deterministic DM-CIFC [13, Sec. VIII]. The rate region of [13, Th. 4.1] is expressed before Fourier–Motzkin elimination. The contribution here is to show that a specific choice of a rate-split in [13, Th. 4.1] is without loss of generality (see Appendix F); with this choice certain rate-bounds become redundant which results in a manageable number of rate bounds after Fourier–Motzkin elimination. We present our simplified region in Theorem IV.1. We then use this as a unified framework from which we derive few simple schemes that will be used in the following sections to prove capacity results, constant gap results as well as numerical evaluations.

As the Gaussian CIFC encompasses classical interference, multiple-access and broadcast channels as special cases, the achievable rate region of [13] incorporates a combination of the transmission techniques devised for these channels. In particular:

- **Rate splitting.** Both the primary and the cognitive message are split into private and common parts, as in the Han and Kobayashi scheme [41] for the IFC. Although rate-splitting was shown to be unnecessary in the “weak interference” [6] and “very strong interference” [5] regimes of (2) and (6), respectively, it allows significant rate improvement in the “strong interference” regime.
- **Superposition coding.** The cognitive common message is superposed to the primary common message and parts of the cognitive message are superposed to parts of the primary message. A simple superposition of the primary and cognitive messages (all common) is capacity achieving in the “very strong interference” regime [5].
- **Binning.** Gel’fand–Pinsker coding [42], often referred to as pre-coding or Dirty Paper Coding (DPC) for Gaussian channels [44], allows a transmitter to “pre-cancel” portions of the interference known to be experienced at the receiver. In [13], binning is performed at the cognitive encoder for both the common and the private message and it allows for the cancellation of interference from the primary transmitter.
- **Broadcasting.** In [13] we introduced the idea of having the cognitive encoder transmit part of the primary private message not transmitted by the primary encoder. The additional primary message is superposed to the cognitive common message and also pre-coded against the cognitive private message. The incorporation of the broadcast feature at the cognitive transmitter was initially motivated by the fact that in certain regimes, this strategy was shown to be capacity achieving for the high-SNR linear deterministic approximation of the CIFC [24].

The achievable scheme may be described as follows:

- **Rate splitting.** The independent messages W_1 and W_2 are rate split into the messages $W_\ell, \ell \in \{1c, 2c, 1pb, 2pb, 2pa\}$, all independent and uniformly distributed on $[1 : 2^{NR_\ell}]$, each encoded into the RV U_ℓ (notice the same subscript), such that $R_1 = R_{1c} + R_{1pb}$ $R_2 = R_{2c} + R_{2pa} + R_{2pb}$.

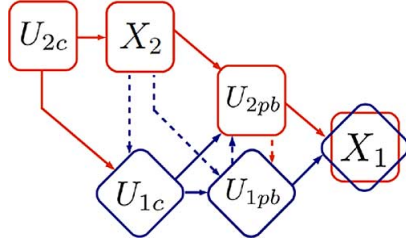


Fig. 3. Graphical representation of the coding scheme for the inner bound region in Theorem IV.1. The RVs for message 1 are in blue diamond boxes while the RVs for message 2 are in red square boxes. A solid line among RVs indicates that the RVs are superposed while a dashed line that the RVs are binned against each other.

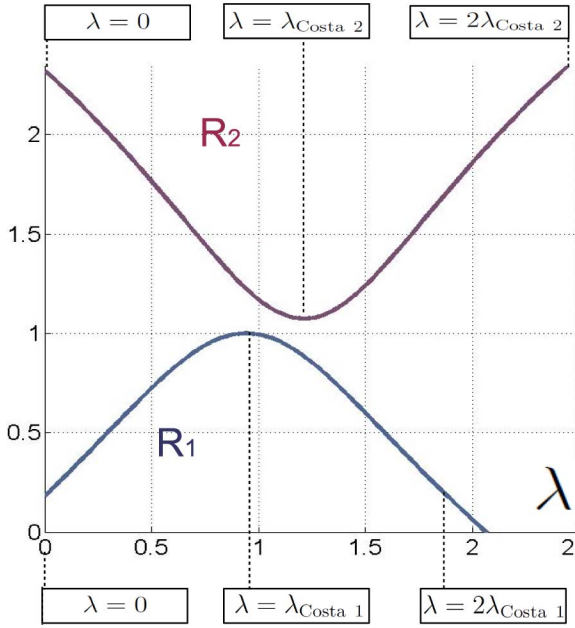


Fig. 4. Bound for R_1 in (29a) (bottom) and the bound for R_2 in (29b) (top) as a function of $\lambda \in \mathbb{R}$, for $P_1 = P_2 = 6$, $|b| = \sqrt{2}$, $a = \sqrt{0.3}$, $\alpha = 0.5$.

- **Primary encoder.** Transmitter 2 sends X_2 that carries the private message W_{2pa} (“p” for private, “a” for alone) superposed to the common message W_{2c} carried by U_{2c} (“c” for common).
- **Cognitive encoder.** The common message of transmitter 1, encoded by U_{1c} , is binned against X_2 conditioned on U_{2c} . The private message of transmitter 2, W_{2pb} , encoded by U_{2pb} (“b” for broadcast) and a portion of the private message of transmitter 1, W_{1pb} , encoded as U_{1pb} , are

binned against each other as in Marton’s region [20] conditioned on U_{1c}, U_{2c}, X_2 . Transmitter 1 sends X_1 , which is a function of all the RVs.

- **Primary decoder.** Receiver 2 jointly decodes U_{2c} (carrying W_{2c}), U_{1c} (carrying W_{1c}), U_{2pb} (carrying W_{2pb}), and X_2 (carrying W_{2pa}).
- **Cognitive decoder.** Receiver 1 jointly decodes U_{1c} (carrying W_{1c}), U_{2c} (carrying W_{2c}), and U_{1pb} (carrying W_{1pb}).
- **Analysis.** The codebook generation, encoding, decoding and the error analysis are provided in [13].

We now present a compact, Fourier–Motzkin eliminated version of the achievable rate region in [13, Th. 4.1], whose graphical representation is given in Fig. 3. In Fig. 3 the RVs for message 1 are in blue diamond boxes while the RVs for message 2 are in red square boxes; a solid line among RVs indicates that the RVs are superposed, while a dashed line that the RVs are binned against each other.

Theorem IV.1: The achievable region [13, Th. 4.1] in compact form. A non-negative rate pair (R_1, R_2) that satisfies the inequalities in (19), shown at the bottom of the page, for some input distribution $P_{U_{1c}, U_{2c}, U_{1pb}, U_{2pb}, X_1, X_2}$ is achievable for a general CIFIC.

Proof: The proof may be found in Appendix F. ■

An achievable rate region for the G-CIFIC could be obtained by considering a jointly Gaussian input $(X_1, X_2, U_{1c}, U_{2c}, U_{1pb}, U_{2pb})$; however, this approach would require the specification of a 6×6 covariance matrix, that is, of 15 complex-valued correlation coefficients; such a region is not tractable analytically because of the large number of free parameters. For the sake of analytical tractability, we now present few different subschemes obtained from the achievable rate region in (19) by reducing the number of auxiliary RVs to at most three. The resulting transmission schemes are used in the rest of the paper for achievability proofs (for capacity and constant gap results) and numerical evaluations. Table I helps illustrate the different subschemes. It indicates, for each subscheme, which auxiliary RVs are retained (by a • in the corresponding column), and where this subscheme is used in the remainder of the paper.

A. Scheme (A). Achievable Scheme With U_{2pb} and U_{1pb} : Capacity Achieving for the Degraded Broadcast Channel

Motivation: Achieve the capacity to within a finite gap in some parameter regime by having transmitter 2 silent.

$$R_1 \leq I(Y_1; U_{1pb}, U_{1c} | U_{2c}) - I(U_{1pb}, U_{1c}; X_2 | U_{2c}), \quad (19a)$$

$$R_1 \leq I(Y_1; U_{1pb} | U_{1c}, U_{2c}) + I(Y_2; U_{2pb}, U_{1c}, X_2 | U_{2c}) - I(U_{1pb}; X_2 | U_{1c}, U_{2c}), \quad (19b)$$

$$R_2 \leq I(Y_2; U_{2pb}, X_2, U_{1c}, U_{2c}) + I(U_{1c}; X_2 | U_{2c}), \quad (19c)$$

$$R_1 + R_2 \leq I(Y_1; U_{1pb}, U_{1c}, U_{2c}) + I(Y_2; U_{2pb}, X_2 | U_{1c}, U_{2c}) - I(U_{1pb}; U_{2pb}, X_2 | U_{1c}, U_{2c}), \quad (19d)$$

$$R_1 + R_2 \leq I(Y_1; U_{1pb} | U_{1c}, U_{2c}) + I(Y_2; U_{2pb}, X_2, U_{1c}, U_{2c}) - I(U_{1pb}; U_{2pb}, X_2 | U_{1c}, U_{2c}), \quad (19e)$$

$$2R_1 + R_2 \leq I(Y_1; U_{1pb}, U_{1c}, U_{2c}) + I(Y_1; U_{1pb} | U_{1c}, U_{2c}) + I(Y_2; U_{2pb}, U_{1c}, X_2 | U_{2c}) - I(U_{1pb}; U_{2pb}, X_2 | U_{1c}, U_{2c}) - I(U_{1pb}, U_{1c}; X_2 | U_{2c}) \quad (19f)$$

TABLE I
ROLE OF THE DIFFERENT ACHIEVABLE SCHEMES IN THE FOLLOWING SECTIONS

Scheme	U_{2c}	U_{1c}	X_2	U_{1pb}	U_{2pb}	Role	Where
(A)				•	•	constant gap in a subspace of the parameter region	Th. VI.4
(B)			•	•		capacity in “weak interference”	Th. II.1, Th. VI.1, Th. VI.3
(C)			•	•	•	constant gap in the whole parameter region	Th. VI.1
(D)	•	•				capacity in “very strong interference”	Th. II.3
						constant gap in a subspace of the parameter region	Th. VI.4
(E)		•	•			capacity in the “primary decodes cognitive” regime	Th. V.1, Th. V.2
						constant gap in a subspace of the parameter region	Th. VI.3, Th. VI.4
(F)	•	•	•			Numerical results	Sec.VII

Consider the case where transmitter 2 is silent and transmitter 1 broadcasts to both receivers. In this case, the G-CIFC with $P_2 = 0$ reduces to a degraded BC with input X_1 whose capacity was determined in [43]. In particular, in “strong interference” ($|b| > 1$), Y_1 is a degraded version of Y_2 and the maximum achievable rate region when transmitter 2 is silent is

$$R_1 \leq I(Y_1; U_{1pb}) \stackrel{[43]}{=} C\left(\frac{\bar{\alpha}P_1}{1 + \alpha P_1}\right), \quad (20a)$$

$$R_2 \leq I(Y_2; U_{2pb}) - I(U_{1pb}; U_{2pb}) \stackrel{[43]}{=} C(\alpha|b|^2 P_1) \quad (20b)$$

taken over the union of all $\alpha \in [0, 1]$, obtained from Theorem IV.1 with $U_{1c} = U_{2c} = X_2 = \emptyset$ (by recalling that for a SISO BC, both superposition coding and Costa’s DPC are optimal).

B. Scheme (B). Achievable Scheme With X_2 and U_{1pb} : Capacity Achieving in the “Weak Interference” Regime

Motivation: Completeness.

In this scheme both messages are private and receiver 2 treats the interference from transmitter 1 as noise while transmitter 1 performs Costa’s DPC [44] against the interference from transmitter 2. This scheme achieves capacity in “weak interference” ($|b| \leq 1$), see Theorem II.1.

C. Scheme (C). Achievable Scheme With X_2 , U_{1pb} and U_{2pb} : Capacity Achieving in the Semi-Deterministic DM-CIFC

Motivation: Achieve the “strong interference” outer bound to within a constant gap in the whole “strong interference” regime.

This achievable strategy is obtained by combining the previous two transmission schemes, scheme (A) and (B), and it corresponds to the capacity achieving scheme for the semi-deterministic G-CIFC [13]. The achievable rate region is

$$R_1 \leq I(Y_1; U_{1pb}) - I(U_{1pb}; X_2) \stackrel{(22)}{=} \log(\sigma_{1pb}^2 + \alpha P_1) - \log\left(\sigma_{1pb}^2 + \frac{\text{Var}[X_1 + aX_2]}{1 + \text{Var}[X_1 + aX_2]}\right), \quad (21a)$$

$$R_2 \leq I(Y_2; U_{2pb}, X_2) \stackrel{(22)}{=} \log(1 + \text{Var}[|b|X_1 + X_2]) - \log\left(1 + \frac{\sigma_{2pb}^2 \text{Var}[|b|X_1|X_2]}{\sigma_{2pb}^2 + \text{Var}[|b|X_1|X_2]}\right), \quad (21b)$$

$$R_1 + R_2 \leq I(Y_2; U_{2pb}, X_2) + I(Y_1; U_{1pb}) - I(U_{1pb}; U_{2pb}, X_2) \stackrel{(22)}{=} (21a) + (21b)$$

$$+ \log\left(1 - \frac{\left|[|b|P_1\alpha - \sqrt{\sigma_{1pb}^2\sigma_{2pb}^2}\right]|^2}{(|b|^2 P_1\alpha + \sigma_{2pb}^2)(P_1\alpha + 1)}\right) \quad (21c)$$

where $\text{Var}X$ indicates the variance of the RV X and where the region in (21) is obtained from Theorem IV.1 with $U_{1c} = U_{2c} = \emptyset$ and

$$X_{1pb} \sim \mathcal{N}_{\mathbb{C}}(0, \alpha P_1), \quad \alpha \in [0, 1], \quad (22a)$$

$$X_2 \sim \mathcal{N}_{\mathbb{C}}(0, P_2), \quad \text{independent of } X_{1pb}, \quad (22b)$$

$$X_1 = X_{1pb} + \sqrt{\frac{\bar{\alpha}P_1}{P_2}} X_2 \quad (22c)$$

$$U_{1pb} = X_1 + aX_2 + Z_{1pb} \quad (22d)$$

$$U_{2pb} = |b|X_1 + X_2 + Z_{2pb}, \quad (22e)$$

$$\begin{bmatrix} Z_{1pb} \\ Z_{2pb} \end{bmatrix} \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{0}, \begin{bmatrix} \sigma_{1pb}^2 & \rho_{pb} \sqrt{\sigma_{1pb}^2 \sigma_{2pb}^2} \\ \rho_{pb}^* \sqrt{\sigma_{1pb}^2 \sigma_{2pb}^2} & \sigma_{2pb}^2 \end{bmatrix}\right) \quad (22f)$$

for $|\rho_{pb}| \leq 1$. The assignment in (22) is inspired by the capacity achieving scheme for the semi-deterministic CIFC of [13] where U_{1pb} and U_{2pb} are set to be equal to Y_1 and Y_2 , respectively. The inequality in (21) is obtained by optimizing over ρ_{pb} as detailed in Theorem VI.1.

D. Scheme (D). Achievable Scheme With U_{1c} and U_{2c} : Capacity Achieving in the “Very Strong Interference” Regime

Motivation: Completeness.

This scheme achieves the “strong interference” outer bound of Theorem II.2 under the “very strong interference” condition of Theorem II.3 [12]. The achievable rate region is

$$R_1 \leq I(Y_1; X_1 | X_2) \quad (23a)$$

$$\stackrel{(24)}{=} C((1 - |\rho|^2)P_1), \quad (23b)$$

$$R_1 + R_2 \leq I(Y_1; X_1, X_2) \quad (23c)$$

$$\stackrel{(24)}{=} C(P_1 + |a|^2 P_2 + 2\Re\{a^* \rho\} \sqrt{P_1 P_2}), \quad (23d)$$

$$R_1 + R_2 \leq I(Y_2; X_1, X_2) \quad (23e)$$

$$\stackrel{(24)}{=} C(|b|^2 P_1 + P_2 + 2|b|\Re\{\rho\} \sqrt{P_1 P_2}) \quad (23f)$$

obtained from Theorem IV.1 with $U_{1pb} = U_{2pb} = \emptyset$ and

$$U_{1c} \sim \mathcal{N}_{\mathbb{C}}(0, (1 - |\rho|^2)P_1), \quad (24a)$$

$$U_{2c} \sim \mathcal{N}_{\mathbb{C}}(0, P_2), \quad \text{independent of } U_{1c}, \quad (24b)$$

$$X_2 = U_{2c}, \quad (24c)$$

$$X_1 = U_{1c} + \rho \sqrt{\frac{P_1}{P_2}} U_{2c} \quad (24d)$$

for some $|\rho| \leq 1$, as originally proposed in [12] for real-valued G-CIFC.

E. Scheme (E). Achievable Scheme With X_2 and U_{1c} : Capacity Achieving in the “Primary Decodes Cognitive” Regime

Motivation: Achieve capacity in the “primary decodes cognitive” regime.

In this scheme the primary message is private while the cognitive message is public and binned against the interference created by the primary user at the cognitive decoder. The achievable rate region is

$$R_1 \leq I(Y_1; U_{1c}) - I(U_{1c}; X_2) \stackrel{(28)}{=} f\left(a + \sqrt{\frac{\alpha P_1}{P_2}}, 1; \lambda\right), \quad (25a)$$

$$\begin{aligned} R_1 + R_2 &\leq I(Y_1; U_{1c}) + I(Y_2; X_2 | U_{1c}) \\ &= I(Y_2; X_2, U_{1c}) \\ &\quad + [I(Y_1; U_{1c}) - I(U_{1c}; X_2)] \\ &\quad - [I(Y_2; U_{1c}) - I(U_{1c}; X_2)] \\ &\stackrel{(28)}{=} \mathcal{C}(P_2 + |b|^2 P_1 + 2\sqrt{\alpha}|b|^2 P_1 P_2) \\ &\quad + f\left(a + \sqrt{\frac{\alpha P_1}{P_2}}, 1; \lambda\right) \\ &\quad - f\left(\frac{1}{|b|} + \sqrt{\frac{\alpha P_1}{P_2}}, \frac{1}{|b|^2}; \lambda\right), \end{aligned} \quad (25b)$$

$$R_1 + R_2 \leq I(Y_2; U_{1c}, X_2) \stackrel{(28)}{=} \mathcal{C}(P_2 + |b|^2 P_1 + 2\sqrt{\alpha}|b|^2 P_1 P_2) \quad (25c)$$

for

$$f(h, \sigma^2; \lambda) := I(X_{1c} + hX_2 + \sigma Z_1; U_{1c}) - I(U_{1c}; X_2) \stackrel{(28)}{=} \log\left(\frac{\sigma^2 + \alpha P_1}{\sigma^2 + \frac{\alpha P_1 |h|^2 P_2}{\alpha P_1 + |h|^2 P_2 + \sigma^2} \left|\frac{\lambda}{\lambda_{\text{Costa}}(h, \sigma^2)} - 1\right|^2}\right) \quad (26)$$

with

$$\lambda_{\text{Costa}}(h, \sigma^2) := \frac{\alpha P_1}{\alpha P_1 + \sigma^2} h \quad (27)$$

and where the region in (29) is obtained from Theorem IV.1 with $U_{2c} = U_{1pb} = U_{2pb} = \emptyset$ and

$$X_{1c} \sim \mathcal{N}_{\mathbb{C}}(0, \alpha P_1), \quad \alpha \in [0, 1], \quad (28a)$$

$$X_2 \sim \mathcal{N}_{\mathbb{C}}(0, P_2) \text{ independent of } \quad (28b)$$

$$X_1 = X_{1c} + \sqrt{\frac{\alpha P_1}{P_2}} X_2, \quad (28c)$$

$$U_{1c} = X_{1c} + \lambda X_2, \quad \lambda \in \mathbb{C}. \quad (28d)$$

Note that $f(h, \sigma^2; \lambda)$ in (26) is non-negative if $\left|\frac{\lambda}{\lambda_{\text{Costa}}(h, \sigma^2)} - 1\right|^2 \leq 1 + \frac{\alpha P_1 + \sigma^2}{|h|^2 P_2}$.

It is easily shown that the region in (25) is equivalent to

$$R_1 \leq (25a), \quad (29a)$$

$$R_2 \leq (25b) - (25a), \quad (29b)$$

$$R_1 + R_2 \leq (25c) \quad (29c)$$

although the region in (29) appears to be smaller than (25). In the following we refer to the formulation in (29) because the bound in (29b) provides important insights on the capacity result of Section V and the role of binning at the cognitive receiver in the “strong interference” regime ($|b| > 1$).

F. Scheme (F). Achievable Scheme With U_{2c} , X_2 , and U_{1c}

Motivation: Achieve capacity in the largest subset of the “strong interference” regime.

This scheme is obtained by combining the previous two schemes, scheme (D) and (E), and will be used for numerical evaluations later on.

V. NEW CAPACITY RESULTS

We now present two new capacity results for the G-CIFC. The first capacity result uses scheme (E) of Section IV-E to achieve the “strong interference” outer bound $\mathcal{O}^{(\text{SI})}$ in (5) in what we term the “primary decodes cognitive” regime, a subset of the “strong interference” regime that is not included in the “very strong interference” regime of Theorem II.3, for which capacity is already known. The second capacity result focuses on the S-G-CIFC where we show that the BC-DMS-based outer bound of Corollary III.5 is achieved by scheme (E) for a large set of parameters where capacity was previously unknown.

A. Intermezzo

Although the two capacity results involve the same achievable scheme (E), in the first result the cognitive transmitter performs Costa’s DPC of the interference from the primary receiver while, in the second result, no DPC is necessary. In scheme (E) the pre-coding operation has an interesting effect on the rate region that we investigate in detail in Section V-B. Before presenting the new results, we describe scheme (E) in more detail.

The achievable rate region in (25) is expressed as a function of two parameters: α and λ .

The parameter α denotes the fraction of power that the cognitive encoder employs to transmit its own message versus the power to broadcast the primary message. For $\alpha = 0$, transmitter 1 uses all its power to broadcast X_2 as in a virtual Multiple Input Single Output (MISO) channel. When $\alpha = 1$, transmitter 1 utilizes all its power to transmit its common message X_{1c} .

The parameter λ controls the amount of interference (created by X_2 at receiver 1) “pre-cancellation” achievable using DPC at transmitter 1. With $\lambda = 0$, no DPC is performed at transmitter 1 and the interference due to X_2 is treated as noise. On the other hand, with $\lambda = \lambda_{\text{Costa } 1}$ for

$$\lambda_{\text{Costa } 1} := \lambda_{\text{Costa}}\left(a + \sqrt{\frac{\alpha P_1}{P_2}}, 1\right)$$

with $\lambda_{\text{Costa}}(\cdot, \cdot)$ defined in (27), the interference due to X_2 at receiver 1 is completely “pre-canceled”, thus achieving the maximum possible rate R_1 . Different values of λ are not usually investigated because, as long as the interference is a nuisance (i.e., no node in the network has information to extract from the interference), the best is to completely “pre-cancel” it by using

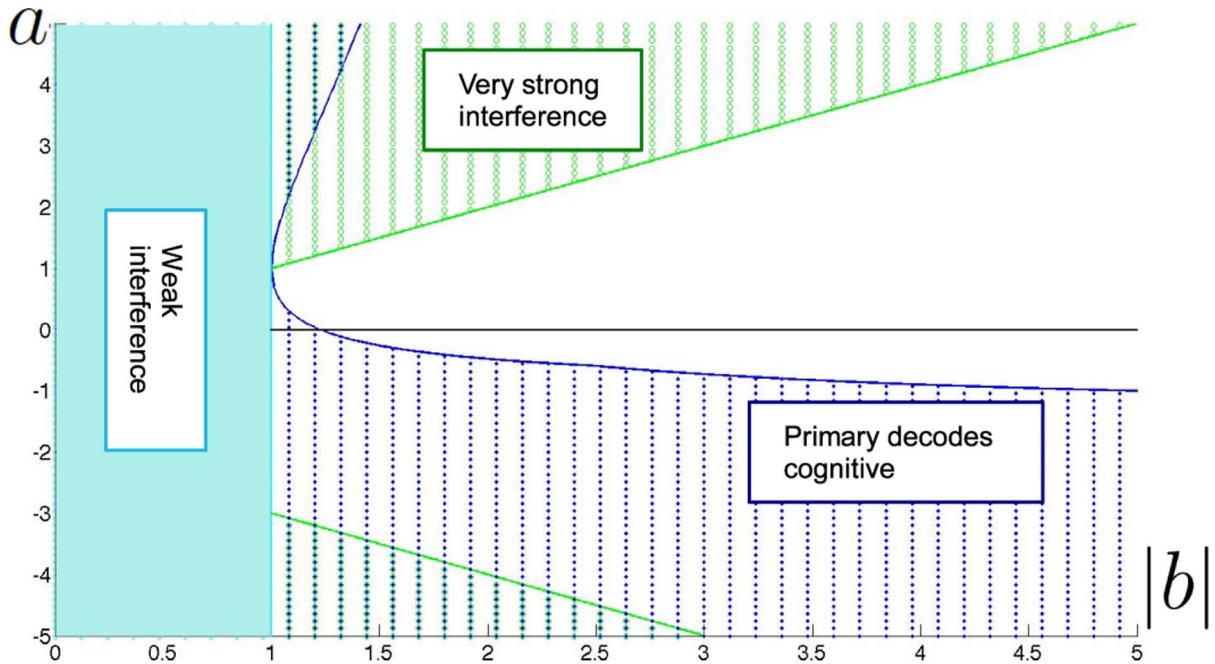


Fig. 5. Representation of the capacity result of Theorem V.1 for a G-CIFC with $a \in \mathbb{R}$, $P_1 = P_2 = 10$ and $(a, |b|) \in [-5, 5] \times [0, 5]$.

$\lambda = \lambda_{\text{Costa } 1}$. However, λ influences not only the rate R_1 in (29a), but also the rate R_2 in (29b). An interesting question is whether $\lambda \neq \lambda_{\text{Costa } 1}$, although it does not achieve the largest possible R_1 , would improve the achievable rate region by sufficiently boosting the rate R_2 . We comment on this question later on in Section VII-D. At this point we make the following observation: R_1 is a concave function in λ , symmetric around $\lambda = \lambda_{\text{Costa } 1}$ and with a global maximum at $\lambda = \lambda_{\text{Costa } 1}$, while R_2 is a convex function in λ , symmetric around $\lambda = \lambda_{\text{Costa } 2}$ and with a global minimum at $\lambda = \lambda_{\text{Costa } 2}$, where

$$\lambda_{\text{Costa } 2} := \lambda_{\text{Costa}} \left(\frac{1}{|b|} + \sqrt{\frac{\bar{\alpha} P_1}{P_2}}, \frac{1}{|b|^2} \right).$$

Fig. 4 shows R_1 in (29a) and R_2 in (29b) as a function of $\lambda \in \mathbb{R}$, for $P_1 = P_2 = 6$, $|b| = \sqrt{2}$, $a = \sqrt{0.3}$, and $\alpha = 0.5$. For the chosen parameters, we observe a trade-off among the rates: $\lambda = \lambda_{\text{Costa } 1}$ achieves the maximum for R_1 , but it achieves close to the minimum for R_2 . This observation will help in understanding why scheme (E) does not perform well in certain parameter regimes as will be pointed out in Section V-B.

B. New Capacity Results for the G-CIFC

Theorem V.1: Capacity in the “primary decodes cognitive” regime. When $|b| > 1$ and

$$P_2|1 - a|b|^2 \geq (|b|^2 - 1)(1 + P_1 + |a|^2 P_2) - P_1 P_2 |1 - a|b|^2, \quad (30a)$$

$$P_2|1 - a|b|^2 \geq (|b|^2 - 1)(1 + P_1 + |a|^2 P_2 + 2\Re\{a\} \sqrt{P_1 P_2}) \quad (30b)$$

the “strong interference” outer bound $O^{(\text{SI})}$ in (5) of Theorem II.2 is achieved by scheme (E) in Section IV-E.

The “primary decodes cognitive” regime, illustrated in Fig. 5 in the $(a, |b|)$ -plane for $a \in \mathbb{R}$ and $P_1 = P_2 = 10$, covers parts

of the “strong interference” regime ($|b| > 1$) where capacity was not known. It also shows that the scheme (E) is capacity achieving for part of the “very strong interference” region in (6), thus providing an alternative capacity achieving scheme to superposition coding [12] [i.e., scheme (D)].

Proof: We compare the achievable rate with scheme (E) in Section IV-E with the “strong interference” outer bound $O^{(\text{SI})}$ in (5) of Theorem II.2. Scheme (E) for $|b| > 1$, $\lambda = \lambda_{\text{Costa } 1}$ and the assignment in (28) achieves (7a) = (29a) and (7c) = (29c) (and (7b) is redundant). Therefore the “strong interference” outer bound $O^{(\text{SI})}$ in (5) is achievable when (29a) + (29b) \geq (29c), that is, when

$$\begin{aligned} \mathcal{C}(\alpha P_1) = f \left(a + \sqrt{\frac{\bar{\alpha} P_1}{P_2}}, 1; \lambda_{\text{Costa } 1} \right) &\geq \\ f \left(\frac{1}{|b|} + \sqrt{\frac{\bar{\alpha} P_1}{P_2}}, \frac{1}{|b|^2}; \lambda_{\text{Costa } 1} \right), \forall \alpha \in [0, 1], & \\ \Leftrightarrow \alpha P_1 + |\lambda_{\text{Costa } 1} a|^2 P_2 - \frac{\alpha P_1}{\alpha P_1 + 1} &\geq \end{aligned} \quad (31)$$

$$\frac{|b| \alpha P_1 + \lambda_{\text{Costa } 1} \left(P_2 + \sqrt{\bar{\alpha} |b|^2 P_1 P_2} \right)^2}{|b|^2 P_1 + P_2 + 2\sqrt{\bar{\alpha} |b|^2 P_1 P_2} + 1}, \forall \alpha \in [0, 1], \quad (32)$$

$$\Leftrightarrow \frac{Q(\alpha)}{1 + |b|^2 P_1 + P_2 + 2\sqrt{\bar{\alpha} |b|^2 P_1 P_2}} \geq 0, \forall \alpha \in [0, 1] \quad (33)$$

where the function $Q(\alpha)$ in (33) is defined as

$$\begin{aligned} Q(\alpha) := &P_2|1 - a|b|^2(\alpha P_1 + 1) \\ &- (|b|^2 - 1) \left(P_1 + |a|^2 P_2 + 2\Re\{a\} \sqrt{\bar{\alpha} P_1 P_2} + 1 \right). \end{aligned}$$

Clearly the condition in (33) is verified if for all $\alpha \in [0, 1]$ we have $Q(\alpha) \geq 0$. $Q(\alpha)$ is a quadratic function in $x = \sqrt{1 - \alpha}$

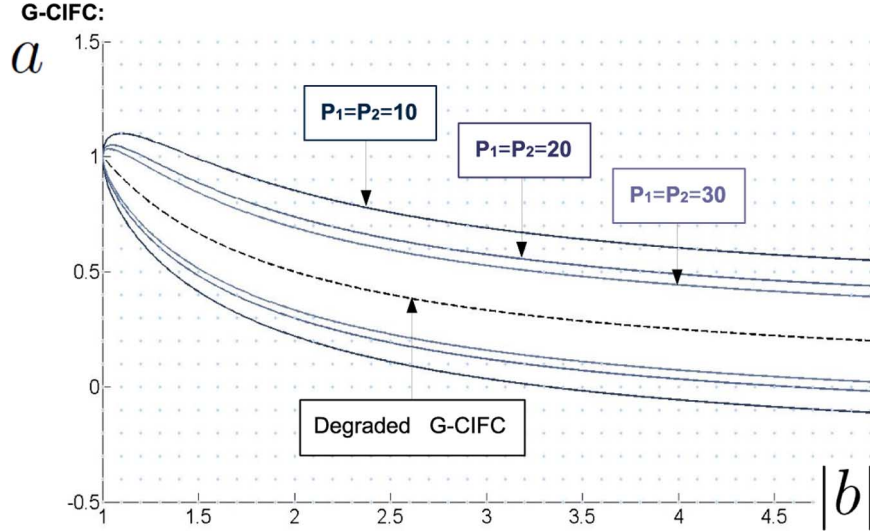


Fig. 6. Condition (34) for different values of $P_1 = P_2 = P \in [10, 20, 30]$ for a G-CIFC with $a \in \mathbb{R}$ and $(a, |b|) \in [-0.5, 1, 5] \times [1, 5]$.

of the form $c_1 x^2 + c_2 x + c_3$ with $c_1 = -P_1 P_2 |1 - a|b||^2 \leq 0$, which implies that $Q(\alpha)$ is concave in α . Hence, the inequality in (33) is verified for every $\alpha \in [0, 1]$ if it is verified for $\alpha = 1$ and $\alpha = 0$. The condition $Q(0) \geq 0$ corresponds to (30b) while the condition $Q(1) \geq 0$ corresponds to (30a). ■

We conclude the section with the following remark. Previous capacity results for the G-CIFC imposed conditions on the channel parameters that lent themselves well to “natural” interpretations. For example, the “weak interference” condition $I(Y_1; X_1|X_2) \geq I(Y_2; X_1|X_2)$ of [6] in (2) suggests that decoding X_1 at receiver 2, even after having decoded the intended message in X_2 , would constrain the rate R_1 too much, thus preventing it from achieving the interference-free rate in (7a). The “very strong interference” condition $I(Y_1; X_1, X_2) \geq I(Y_2; X_1, X_2)$ of [12] in (4) suggests that requiring receiver 1 to decode both messages should not prevent achieving the maximum sum-rate at receiver 2 given by (7c). A similar intuition about the new “primary decodes cognitive” capacity condition in (30) unfortunately does not emerge from the proof of Theorem V.1.

To provide some insight on the achievability conditions of Theorem V.1, we focus on the condition in (30a). When (30a) is verified, scheme (E) of Section IV-E achieves the “strong interference” outer bound at the point corresponding to $\alpha = 0$ in (5); to achieve more points on the “strong interference” outer bound stricter conditions than (30a) alone are necessary; to achieve all the points on the outer bound, both conditions (30a) and (30b) must be verified. A representation of the region where the condition in (30a) holds is depicted in Fig. 6 for the case $a \in \mathbb{R}$ and $P_1 = P_2 = P$ with increasing P , in which case (30a) becomes

$$P(P+1)|1 - a|b||^2 \geq (|b|^2 - 1)(P+1 + |a|^2 P). \quad (34)$$

We observe that, as P increases, the region where the condition in (34) is not verified shrinks. Indeed, as $P \rightarrow \infty$, the condition in (34) is always verified unless the channel is degraded (i.e., $a|b| = 1$). For a degraded channel with “strong interference”, the primary receiver is able to reconstruct Y_1 from Y_2 once W_2 has been decoded, as seen in (12). This means that U_{1c} may be

decoded at the primary receiver with no rate penalty for the cognitive user. Under this condition, the scheme with a common cognitive message and a private primary one seems a natural choice, reminiscent of the capacity achieving scheme in the degraded BC. Despite this intuition, in a degraded channel with large power P , $\lambda_{\text{Costa } 1}$ approaches $\lambda_{\text{Costa } 2}$ (similarly to the case depicted in Fig. 4) and thus the maximum of the rate R_1 in (29a) approaches the minimum of the rate R_2 in (29b). This rate penalty for the R_2 -bound prevents us from achieving the “strong interference” outer bound point for $\alpha = 0$ in (5) when $a|b| = 1$.

Another consideration provides further insight on the condition in (30a): take a channel where

$$\frac{1}{|b|} = a \frac{P_1}{P_1 + 1}. \quad (35)$$

Then, as $P_1 \rightarrow \infty$ in (35) and for $\alpha = 0$, this condition approaches the degraded condition $a|b| = 1$. For this choice of a , Y_2 may be rewritten as $Y_2 = |b|U_{1c} + Z_2$, so that the R_2 -bound of (29b) for $\alpha = 0$ becomes

$$R_2 \leq I(Y_2, U_{1c}; X_2) = I(U_{1c}; X_2) = \mathcal{C} \left(\frac{P_2}{|b|^2 P_1} \right).$$

This observation reveals an interesting aspect of the RV U_{1c} . U_{1c} is DPC-ed against X_2 with the objective to remove (some of) the interference created by X_2 at Y_1 . However, decoder 2 is not interested in removing X_2 from Y_2 (it must decode X_2 !). Hence, for decoder 2, U_{1c} acts as “side information” when decoding X_2 . Now, both U_{1c} and Y_2 contain X_2 , but for this specific choice of parameters Y_2 is a noisy version of U_{1c} . This shows why the scheme performs poorly close to the degraded line: there is no gain for receiver 2 from having two observations (i.e., Y_2 and U_{1c}) of the intended message X_2 as they are noisy versions of each other.

C. New Capacity Results for the S-G-CIFC

Theorem V.2: Capacity for S-G-CIFC. For an S-G-CIFC (i.e., $a = 0$) with

$$|b| \leq \sqrt{1 + \frac{P_2}{P_1 + 1}} \quad (36)$$

or with

$$|b| \geq \sqrt{1 + P_2 + P_1 P_2} + \sqrt{P_1 P_2} \quad (37)$$

the BC-DMS-based outer bound of Corollary III.5 is achieved by scheme (E) of Section IV-E.

Proof: When $|b| \leq 1$, capacity is known so we focus only on the case $|b| > 1$. By setting $a = 0$ in Theorem V.1 we obtain that scheme (E) with $\lambda = \lambda_{\text{Costa } 1}$ achieves the “strong interference” outer bound for

$$(|b|^2 - 1)(1 + P_1) \leq \min\{P_2, P_2(1 + P_1)\} = P_2$$

which is equivalent to (36).

Scheme (E) with $\lambda = 0$ achieves

$$R_1 \leq C\left(\frac{\alpha P_1}{1 + \bar{\alpha} P_1}\right), \quad (38a)$$

$$R_2 \leq C((\sqrt{P_2} + \sqrt{\bar{\alpha}|b|^2 P_1})^2), \quad (38b)$$

$$R_1 + R_2 \leq C(\alpha|b|^2 P_1 + (\sqrt{P_2} + \sqrt{\bar{\alpha}|b|^2 P_1})^2). \quad (38c)$$

In this case the the BC-DMS-based outer bound of Corollary III.5 may be achieved when the sum-rate bound in (16c) and (38c) are both redundant. As pointed out in Corollary III.5, the bound in (16c) is redundant when the condition in (15) holds. Similarly, the sum-rate bound in (38c) can be dropped when (see the equations at the bottom of the page), which corresponds to (37). ■

We would like to conclude the section we a couple of remarks.

On extensions of Theorem V.2. The range of the parameter $|b|$ for which the BC-DMS-based outer bound of Corollary III.5 is achieved may be numerically shown to be strictly larger than the range in Theorem V.2, as follows.

The outer bound of Corollary III.5 in (16) was derived by outer bounding the actual MIMO BC-DMS capacity region in (67) so as to have one free parameter only instead of two. For notation convenience let $O^{\text{eq.}(16)}$ denote the outer bound in (16) and $O^{\text{eq.}(67)}$ denote the outer bound in (67), where $O^{\text{eq.}(67)} \subseteq O^{\text{eq.}(16)}$.

One sees immediately by comparing $O^{\text{eq.}(67)}$ with the achievable rate region in (38), that the points on the closure of $O^{\text{eq.}(67)}$ for which $\alpha_2 = 0$ is optimal are achievable, thus they are on the closure of the capacity region. Corollary III.5 gives conditions under which $\alpha_2 = 0$ is optimal for all points on the closure of

$O^{\text{eq.}(16)}$. By imposing that $\alpha_2 = 0$ is optimal for all points on the closure of $O^{\text{eq.}(67)}$ one can improve on Theorem V.2. One possible way to do so is: the outer bound $O^{\text{eq.}(67)}$ is optimal for the S-G-CIFC in “strong interference” ($|b| > 1$) if

$$(67a) + (67b) \leq (67c) \quad \forall (\alpha_1, \alpha_2) \in [0, 1]^2. \quad (39)$$

In fact, if the condition in (39) holds, the sum-rate in (67c) can be dropped and the R_2 -bound in (67b) is maximum when $\alpha_2 = 0$, and the resulting region is therefore achievable by scheme (E) in (38). Unfortunately, we have not been able to solve in closed form the problem in (39).

The condition in (39) is an implicit characterization (no explicit, closed form result) for a class of S-G-CIFC which, as numerically may be verified, holds for a larger range of the parameter $|b|$ than the one stated (in closed form) in (37).

On S-channels excluded by Theorem V.2. Theorem V.2 implies that the capacity of the S-G-CIFC remains (analytically) unknown for

$$\sqrt{1 + \frac{P_2}{P_1 + 1}} < |b| < \sqrt{1 + P_2 + P_1 P_2} + \sqrt{P_1 P_2}.$$

As mentioned above, we were able to tighten this somewhat, but not to the whole set of parameters $|b|$ given in (17) for which the BC-DMS-based outer bound outperforms the “strong interference” outer bound. The work [30] claims that a BC-based approach similar to the one of this paper suffices to prove capacity in the whole set of parameters $|b|$ given in (17). We remind the reader that our outer bound in Corollary III.5 was derived by outer bounding the MIMO BC-DMS capacity region so as to have one free parameter only. We have not attempted here to work with the MIMO BC-DMS capacity region itself, besides for (39), because it is defined as function of two free parameters and it is therefore quite complicated to manipulate analytically.

VI. CAPACITY TO WITHIN A CONSTANT GAP

In the last couple of years a novel approach to the difficult task of determining the capacity region of a multiuser Gaussian network has been suggested. Rather than proving an equality between inner and outer bounds, the authors of [45] (and references therein) advocate a powerful new method for obtaining achievable rate regions that lie within a bounded distance from capacity region outer bounds, thereby determining the capacity region to *within a constant gap* for any channel configuration. Two measures are used to determine the distance between inner and outer bounds: the additive gap and the multiplicative gap. An additive gap corresponds to a finite difference between inner

$$\begin{aligned} 1 + P_2 + |b|^2 P_1 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2} &\geq \\ \frac{1 + P_1}{1 + \bar{\alpha} P_1} (1 + P_2 + |b|^2 P_1 - \alpha|b|^2 P_1 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2}), &\quad \forall \alpha \in [0, 1] \\ \iff |b|^2 \geq 1 + P_2 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2}, &\quad \forall \alpha \in [0, 1] \\ \iff |b|^2 \geq 1 + P_2 + 2\sqrt{|b|^2 P_1 P_2} \end{aligned}$$

and outer bound, while a multiplicative gap corresponds to a finite ratio. The additive gap is useful at high SNR, where the difference between inner and outer bound is small in comparison to the magnitude of the capacity region, while the multiplicative gap is useful at low SNR, where the ratio between inner and outer bounds is a more indicative measure of their distance.

In this section we show the capacity to within an additive gap of one bit/s/Hz and to within a multiplicative gap of a factor two. We also determine additional additive constant gap results that suggest which strategies approach the “strong interference” outer bound in different parameter regimes. Since the expressions of the BC-PR-based outer bound of Theorem III.2 and of the BC-DMS-based outer bound of Theorem III.3 involve many parameters over which to optimize, it is not analytically straightforward to determine conditions for achievability. For this reason we restrict our attention to the “strong interference” outer bound $\mathcal{O}^{(SI)}$ in (5) of Theorem II.2. These results are derived for the complex-valued channel and rather than for the real-valued channel as done in [7].

A. Additive Gap

Theorem VI.1: Additive gap. Capacity is known to within one bit/s/Hz.

Proof: The capacity for weak interference ($|b| \leq 1$) was determined in [6], so we only need to concentrate on the strong interference regime ($|b| > 1$). We show the achievability of the “strong interference” outer bound $\mathcal{O}^{(SI)}$ in (5) to within a constant additive gap using the scheme (C) of Section IV-C with the assignment in (22). The assignment proposed in (22) is inspired by the capacity achieving scheme for the deterministic C-IFC in [13], where we showed that setting $U_{ipb} = Y_i, i \in \{1, 2\}$, is optimal. In a noisy channel, it is not possible to choose $U_{ipb} = Y_i$. We mimic this by setting $U_{ipb} \sim Y_i, i \in \{1, 2\}$. To the best of the authors’ knowledge, this is the first gap result that uses a binning-based achievable scheme.

Consider the achievable rate region in (21) and note that

$$\begin{aligned} \text{Var}[X_1 + aX_2] &= P_1 + |a|^2 P_2 + 2\Re\{a\}\sqrt{\alpha P_1 P_2}, \\ \text{Var}[|b|X_1 + X_2] &= |b|^2 P_1 + P_2 + 2\sqrt{\alpha}|b|\sqrt{P_1 P_2}. \end{aligned}$$

The inequality in (21) follows by choosing

$$\begin{aligned} \rho_{pb} &= \arg \min_{\rho} I(U_{1pb}; U_{2pb} | X_2) \\ &= \arg \min_{\rho} |\mathbb{E}[U_{1pb} U_{2pb}^* | X_2]|^2 \\ &= \arg \min_{\rho} \left| |b| P_1 \alpha + \rho \sqrt{\sigma^2_{1pb} \sigma^2_{2pb}} \right|^2 \\ &= - \min \left\{ 1, \frac{|b| P_1 \alpha}{\sqrt{\sigma^2_{1pb} \sigma^2_{2pb}}} \right\}. \end{aligned}$$

With $\sigma^2_{2pb} = 0$ and $\sigma^2_{1pb} = 1$ in (21) we have

$$R_1 \leq \log(1 + \alpha P_1) - \text{GAP}(\alpha), \quad (40a)$$

$$R_1 + R_2 \leq \log(1 + \text{Var}[|b|X_1 + X_2]) - \text{GAP}(\alpha) \quad (40b)$$

with $\text{GAP}(\alpha)$ bounded as

$$\text{GAP}(\alpha) = \log \left(1 + \frac{\text{Var}[X_1 + aX_2]}{1 + \text{Var}[X_1 + aX_2]} \right) \leq \log(2) = 1$$

as claimed. Notice that with $\sigma^2_{2pb} = 0$, the R_2 -bound in (21b) is equivalent to the sum-rate outer bound in (5b) and it is thus redundant. ■

B. Multiplicative Gap

To prove the multiplicative gap result, we utilize a looser version of Theorem III.1 that we present in the next lemma.

Lemma VI.2: “Piecewise linear strong interference” outer bound. The outer bound of Theorem III.1 for $|b| > 1$ is contained in the region $\mathcal{O}^{(PL-SI)}$ defined as

$$R_1 \leq \mathcal{C}(P_1), \quad (41a)$$

$$R_1 + R_2 \leq \mathcal{C}((\sqrt{|b|^2 P_1} + \sqrt{P_2})^2). \quad (41b)$$

Proof: The bound in (41a) (respectively (41b)) is obtained by considering the maximum value of (7a) (respectively (7c)) over $\alpha \in [0, 1]$. ■

The region $\mathcal{O}^{(PL-SI)}$ in (41) has two Pareto optimal points:

$$A = (R_1^{(A)}, R_2^{(A)}) = (0, \mathcal{C}((\sqrt{|b|^2 P_1} + \sqrt{P_2})^2)), \quad (42)$$

$$B = (R_1^{(B)}, R_2^{(B)}) = (\mathcal{C}(P_1), \mathcal{C}((\sqrt{|b|^2 P_1} + \sqrt{P_2})^2) - \mathcal{C}(P_1)). \quad (43)$$

The point A in (42) is on the boundary of the “strong interference” outer bound region $\mathcal{O}^{(SI)}$ of Theorem II.2 while point B in (43) has the same R_1 -coordinate as the point for $\alpha = 1$ in $\mathcal{O}^{(SI)}$, given by

$$C = (R_1^{(C)}, R_2^{(C)}) = (\mathcal{C}(P_1), \mathcal{C}(|b|^2 P_1 + P_2) - \mathcal{C}(P_1)) \quad (44)$$

but lies outside $\mathcal{O}^{(SI)}$. However, it is easy to see that the two points are no more than one bit away, i.e., $R_2^{(B)} \leq \log(2) + R_2^{(C)}$.

Theorem VI.3: Multiplicative gap. For a Gaussian C-IFC, the capacity is known to within a factor two.

Proof: The capacity for weak interference ($|b| \leq 1$) was determined in [6], thus we only need to concentrate on the strong interference regime ($|b| > 1$).

For outer bound, we rewrite $\mathcal{O}^{(PL-SI)}$ in (41) as

$$\begin{aligned} R_2 &\leq \mathcal{C} \left((\sqrt{|b|^2 P_1} + \sqrt{P_2})^2 \right) - R_1 \\ &=: R_2^{(PL-SI)}(R_1) \end{aligned} \quad (45)$$

for $R_1 \in [0, \log(1 + P_1)]$. For achievability, we consider the following TDMA strategy. The rate-point

$$(\mathcal{C}(P_1), 0) \quad (46)$$

is achievable by silencing the primary transmitter, while the rate-point A in (42) is achievable by beam-forming. Hence, the following region is achievable by time sharing

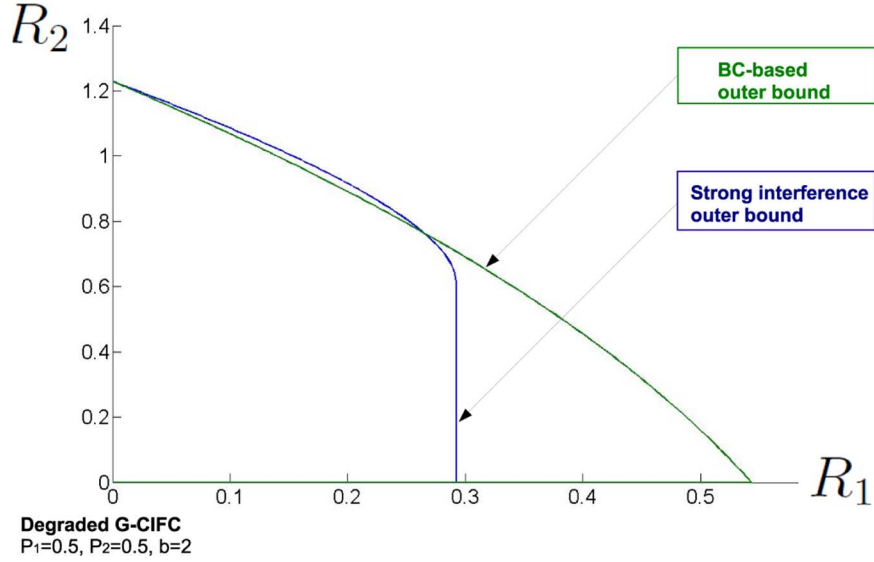


Fig. 8. The “strong interference” outer bound and the BC-based outer bound for the degraded G-CIFC.

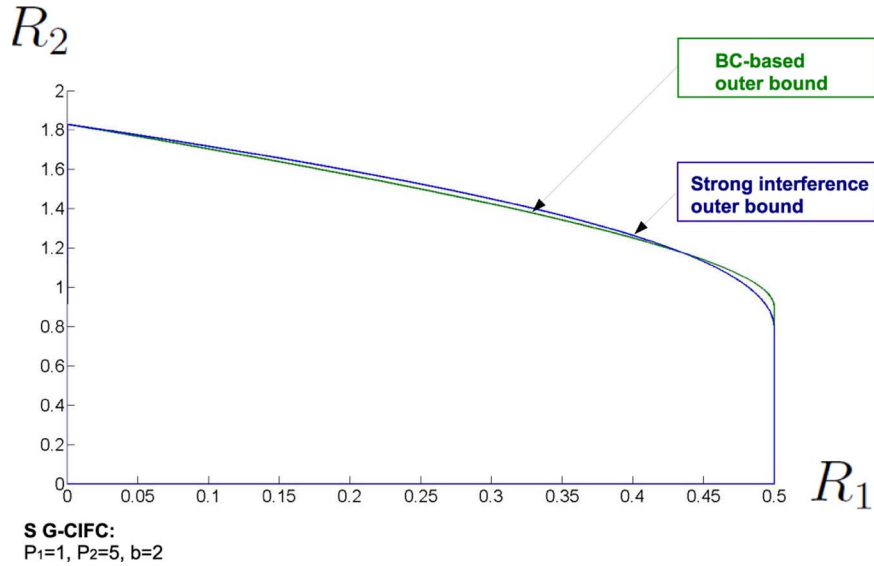


Fig. 9. The “strong interference” outer bound and the BC-based outer bound for the S-G-CIFC.

A. Section III: Outer Bounds

In Section III, we introduced the tightest known outer bound for a G-CIFC in “strong interference”, obtained as the intersection of the “strong interference” outer bound of Theorem II.2 and the BC-based outer bound of Theorem III.2. This outer bound has a simple closed form expression for the degraded G-CIFC and the S-G-CIFC: Figs. 8 and 9 present the result of Corollaries III.4 and III.5, respectively, where the intersection of the “strong interference” outer bound and the BC-based outer bound for the degraded G-CIFC and the S-G-CIFC is derived. Note that we chose two channels where the two bounds intersect for some $R_1 \in (0, \mathcal{C}(P_1)]$ and neither bound strictly includes the other. The maximum rate R_1 in the “strong interference” outer bound and the BC-based outer bound for the S-G-CIFC are the same: in this channel transmitter 2 does not influence the output at receiver 1 and hence full receiver cooperation does not increase the maximum attainable rate R_1 .

For a general G-CIFC the intersection between the “strong interference” and the BC-based outer bound has no simple closed

form expression. Consequently, it is difficult to determine where one dominates and find their intersection analytically. In Fig. 10 we show that the two bounds can intersect up to two times.

The outer bounds of Theorem III.6 are presented in Fig. 11 which shows that these outer bounds may be tighter than either the “strong interference” or the BC-based outer bounds. Unfortunately, in the examples we considered, we did not find an instance where the outer bounds of Theorem III.6 were tighter than the intersection of the “strong interference” and the BC-based outer bound. Despite this, we believe that our approach in transforming the channel provides a general, useful tool to derive outer bounds for channels with cognition.

B. Section IV: Inner Bounds

In Section IV, we introduced the $R^{(\text{RTD})}$ achievable rate region and derived five subschemes from this general inner bound: in the following we plot these subschemes for the degraded channel, the S-channel and a general G-CIFC. The “strong interference” and the BC-based outer bounds are provided for

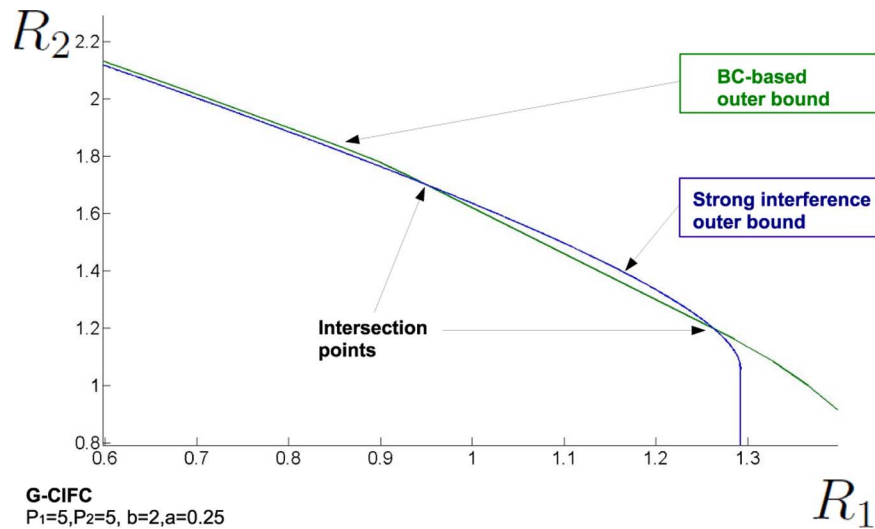


Fig. 10. The “strong interference” outer bound and the BC-based outer bound for a general G-CIFC.

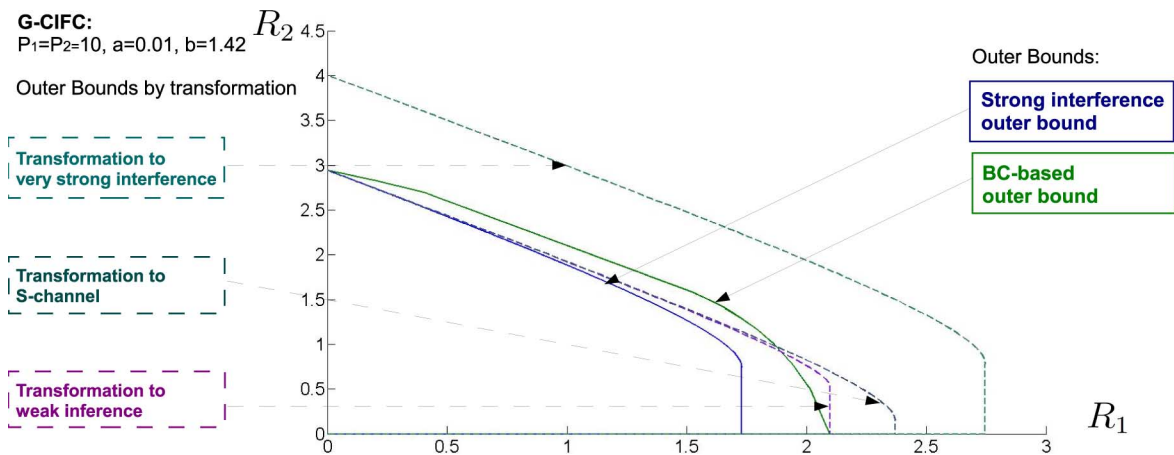


Fig. 11. Outer bounds of Theorem III.6 for special choices of transformations alongside the “strong interference” outer bound and the BC-based outer bound.

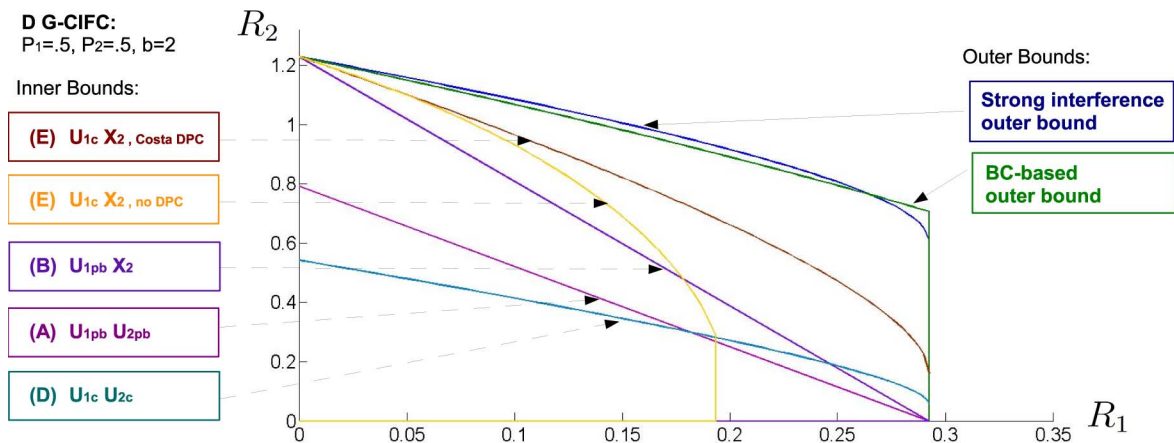


Fig. 12. Achievable schemes of Section IV for the degraded G-CIFC.

reference. Note that both the achievable rate regions and the outer bounds are expressed as a function of one parameter only, $\alpha \in [0, 1]$, that controls the amount of cooperation between the cognitive and the primary transmitters.

We begin by considering the degraded G-CIFC in Fig. 12. The scheme that yields the largest achievable rate region is scheme (E) with the choice $\lambda = \lambda_{\text{Costa } 1}$. Despite its superior performance (to other presented schemes) we may analytically show

that this scheme cannot achieve either the “strong interference” or the BC-based outer bound. Both schemes (A) and (B) treat the interference as noise at receiver 1 and thus the maximum R_1 may be achieved only by silencing transmitter 2. For this reason $R_2 \rightarrow 0$ as $R_1 \rightarrow \mathcal{C}(P_1)$ for these two schemes.

We next consider the S-G-CIFC in Fig. 13. The channel parameters are chosen to show that scheme (E), with the choice $\lambda = \lambda_{\text{Costa } 1}$, achieves the “strong interference” outer bound

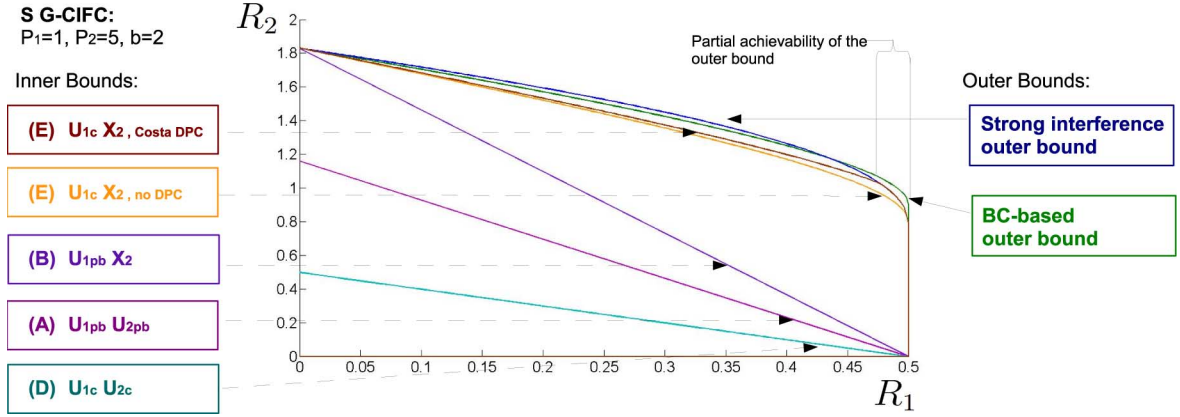


Fig. 13. Achievable schemes of Section IV for the S-G-CIFC.

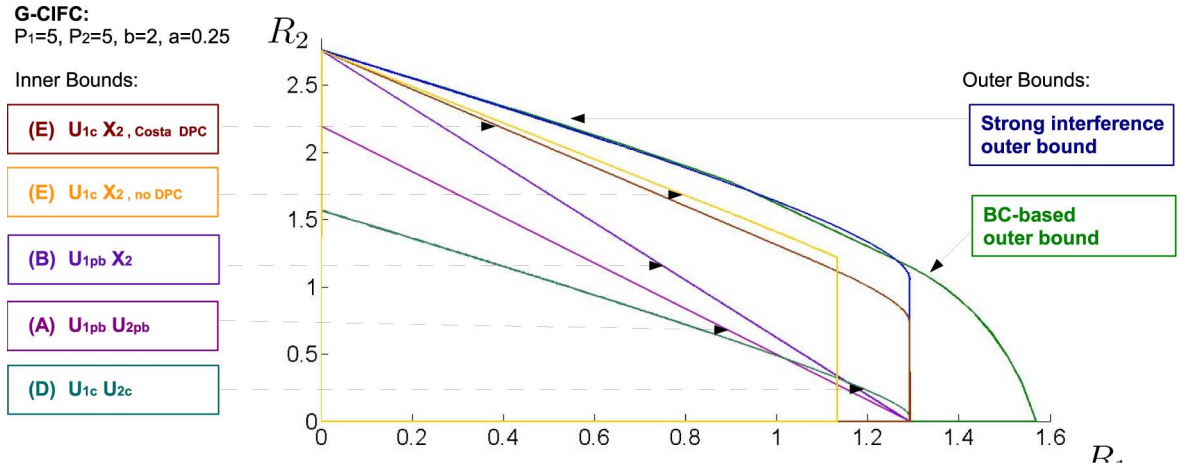


Fig. 14. Achievable schemes of Section IV for a general G-CIFC.

for a subset of $R_1 \in (0, \mathcal{C}(P_1)]$ where the inequality in (33) holds. The figure also shows how, in the S-channel, it is possible to achieve the outer bound for $R_1 = \mathcal{C}(P_1)$ with scheme (E) without DPC. This is possible only in this channel, since X_2 does not influence Y_1 and no rate loss occurs at the cognitive receiver by treating the interference as noise. Note that scheme (D) performs the worst among all the achievable schemes: in this scheme the cognitive receiver is *required* to decode both messages—a very stringent constraint since Y_1 does not contain X_2 . In particular, $R_2 \rightarrow 0$ when $R_1 \rightarrow \mathcal{C}(P_1)$ as in schemes (A) and (B): this is so because $R_1 = \mathcal{C}(P_1)$ may be achieved with scheme (D) only for Y_1 independent of X_2 .

A general G-CIFC is considered in Fig. 14. In this example, scheme (E) with $\lambda = 0$ performs better than the scheme with $\lambda = \lambda_{\text{Costa } 1}$ for small R_1 while the opposite is true for large R_1 . This is the first instance in which we see that a single choice of λ does not yield the largest inner bound: for small INR, it is better for the cognitive user to treat the interference as noise, while for large INR it is more advantageous to perform Costa's DPC. From Section III-B we know that, for $|b| > 1$, the primary receiver can decode the cognitive message with no additional rate penalties; this may be observed by comparing scheme (E) with Costa's DPC and scheme (B). The primary message is private in both schemes while the cognitive message is common in scheme (E) and private in scheme (B). Since the primary receiver can decode the cognitive message at no

cost, scheme (E) with Costa's DPC achieves larger rates than scheme (B). When no DPC is used ($\lambda = 0$) in scheme (E), the cognitive receiver observes an equivalent additive Gaussian noise of variance $1 + |a|^2 P_2$: for this region rate R_1 is always bounded by $R_1 \leq \mathcal{C}(P_1 / (1 + |a|^2 P_2))$ and thus scheme (B) outperforms scheme (E) with no DPC in the interval $R_1 \in [\mathcal{C}(P_1 / (1 + |a|^2 P_2)), \mathcal{C}(P_1)]$.

Finally, in Fig. 15 we consider scheme (F) (not previously defined but listed in Table I) which takes U_{2c}, X_2 and U_{1c} to be non-zero. As for scheme (C), this scheme is obtained by combining the schemes (D) and (E). The achievable rate region we consider is obtained from Theorem IV.1 with $U_{1pb} = U_{2pb} = \emptyset$ and

$$X_{2c}, X_{2p}, X_{1c} \sim \text{iid } \mathcal{N}_{\mathbb{C}}(0, 1), \quad (49a)$$

$$X_2 = \sqrt{\beta P_2} X_{2c} + \sqrt{\beta P_2} X_{2p}, \quad \beta \in [0, 1], \quad (49b)$$

$$X_1 = \sqrt{\alpha P_1} X_{1c} + \sqrt{\alpha P_1} (\sqrt{\gamma} X_{2c} + \sqrt{\gamma} X_{2p}), \quad (49c)$$

$$U_{2c} = X_{2c}, \quad (49d)$$

$$U_{1c} = X_{1c} + \lambda X_{2p} \quad (49e)$$

where $(\alpha, \gamma) \in [0, 1]^2$ and $\lambda \in \mathbb{C}$. This scheme unifies the two schemes that achieve capacity in two different parameter

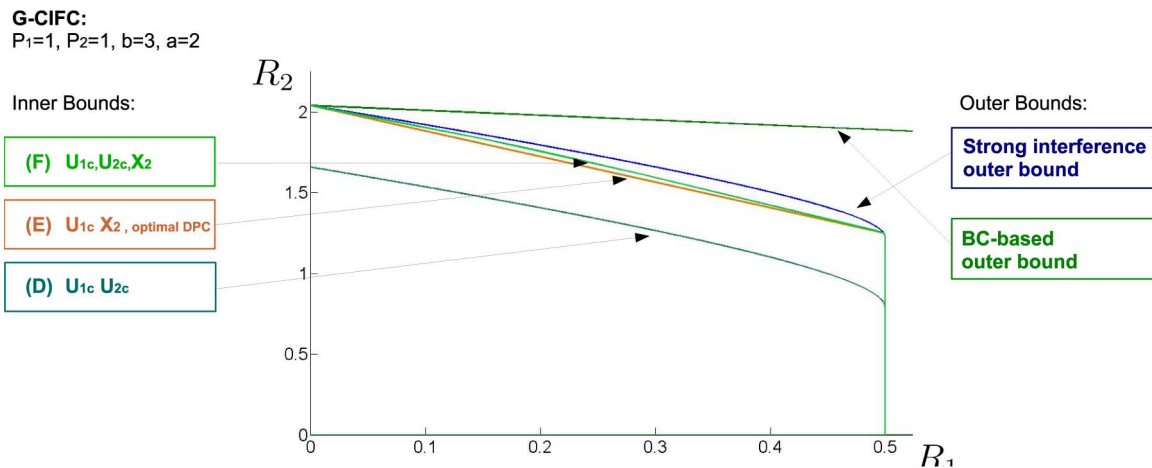


Fig. 15. Achievable rate region of schemes (D), (E) and (F) for a general G-CIFC.

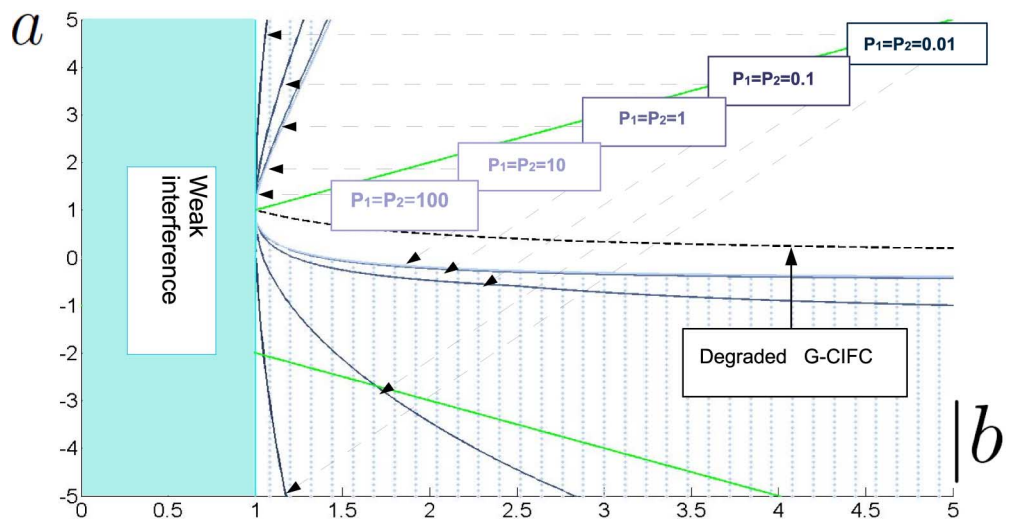


Fig. 16. The “primary decodes cognitive” region for different powers $P_1 = P_2 = P$ for a G-CIFC with $a \in \mathbb{R}$ and $(a, |b|) \in [-5, 5] \times [0, 5]$.

regimes of $|b| > 1$ and hence is the scheme that achieves capacity in the largest subset of the “strong interference” regime ($|b| > 1$). The rate expressions with the assignment in (49) are quite complex and are omitted here for sake of space. This scheme unifies capacity achieving schemes in the “very strong interference” and the “primary decodes cognitive” regimes. It is possible that by unifying the two schemes, we may show capacity in a larger region than the union of the two regimes. Unfortunately determining the achievability conditions in closed form is not straightforward as it requires the optimization of the four parameters in (49). In Fig. 15, we show through numerical evaluation that scheme (F) indeed achieves a larger region than the union of the schemes (E) and (D). Whether this scheme achieves capacity for a larger parameter region remains an open question.

C. Section V: New Capacity Results

In Section V we determined new capacity results for the “primary decodes cognitive” regime both for a general G-CIFC and the S-G-CIFC. In Fig. 16 we plot the “primary decodes cognitive” regime in (30) for different transmitter powers $P_1 = P_2 = P$. Note that the “weak interference” and the “very strong interference” regimes do not depend on P so their plot does not vary.

As the power P increases, the “primary decodes cognitive” region expands from the line $|b| = 1$ to cover a larger region around the degraded line. Interestingly the “primary decodes cognitive” regime intersects with the “very strong interference” regime, thus showing that the “strong interference” outer bound may be achieved with two different transmission schemes for some channels.

In a similar fashion, Fig. 17 shows the capacity results of Theorem V.2 for the case $P_1 = P_2 = P$ on the plane $P \times |b|$. For equal transmitter powers, the conditions in (36) and in (37) become

$$|b|^2 \sim \frac{2P+1}{P+1} \approx 2 \quad (50a)$$

$$|b|^2 \sim P + \sqrt{P^2 + P + 1} \approx 2P \quad (50b)$$

and these two asymptotic behaviors are clearly visible in Fig. 17.

D. Section VI: Capacity to Within a Constant Gap

In Theorem VI.1 we established the capacity of a general G-CIFC to within one bit/s/Hz with a specific assignment in the region of (21). This specific assignment was chosen to mimic the capacity achieving scheme in the deterministic CIFC of [13] and

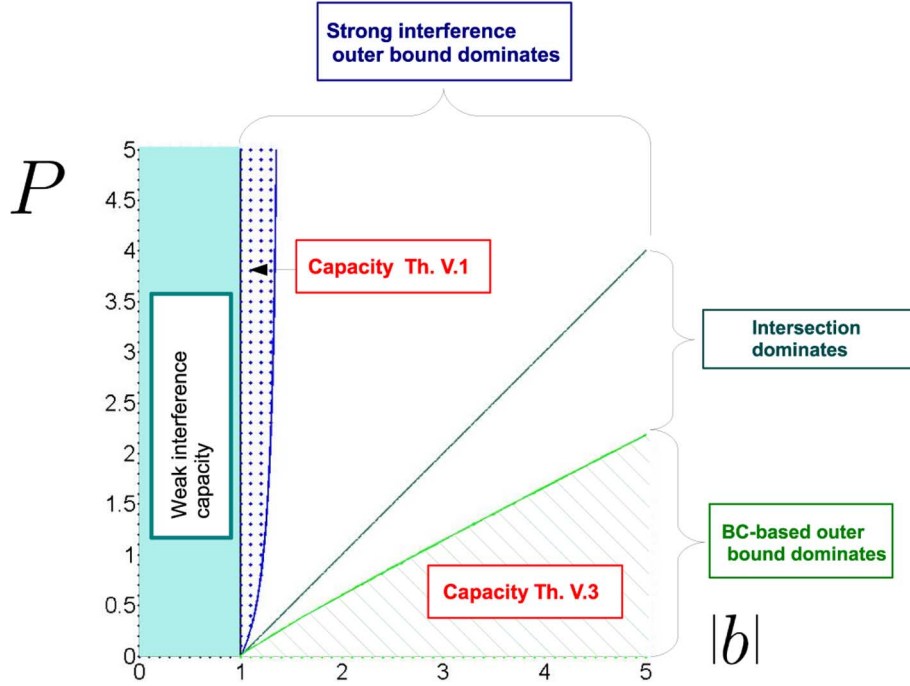


Fig. 17. Capacity results for the S-G-CIFC for the case $P_1 = P_2 = P$ for $(P, |b|) \in [0, 5] \times [0, 5]$.

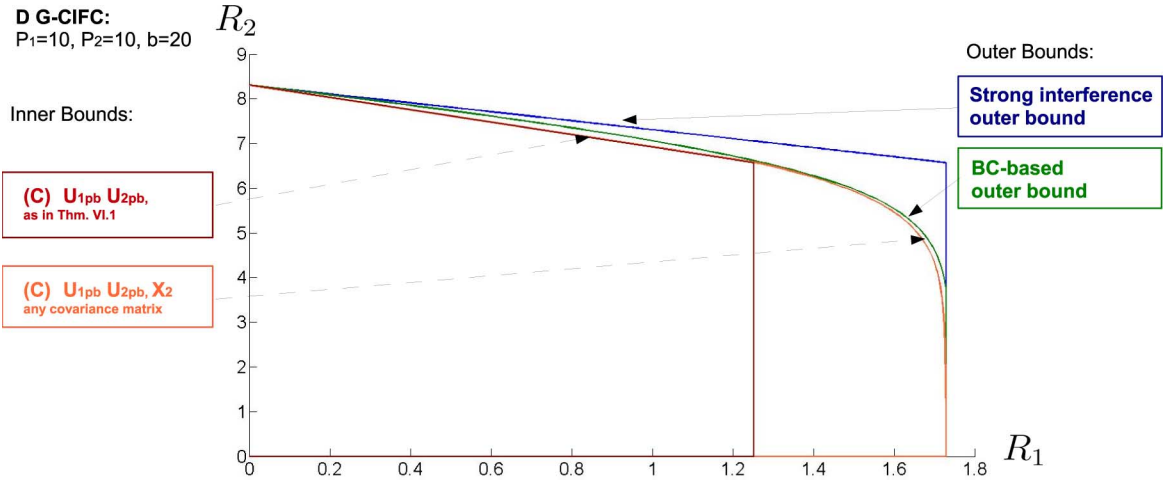


Fig. 18. Achievable rate region of scheme (C) with the assignment of RVs in (22) and in (51).

partially optimized to yield the smallest gap. A larger achievable rate region could be obtained by considering the scheme (C) with the assignment of random variables

$$U_{1pb} = X_1 + c_1 X_2 + Z_{1pb} \quad (51a)$$

$$U_{2pb} = c_2 X_1 + X_2 + Z_{2pb} \quad (51b)$$

in (22). The region in (21) considers only the case $c_1 = a$, $c_2 = |b|$ while the assignment in (51) parameterizes any covariance matrix $\text{Cov}([X_1 \ X_2 \ U_{1pb} \ U_{2pb}])$. Unfortunately, this scheme is parametrized by five coefficients and the algebraic optimization of the additional parameters is quite involved. Instead, in Fig. 18, we may use numerical evaluations to investigate the rate improvements that may be obtained with the more general achievable scheme of (51). We consider a degraded G-CIFC and show that this choice of RVs greatly improves on the result in Theorem VI.1. With the assignment in (21) it is not possible to approach the outer bound of Theorem III.4 for large R_1 . On the

other hand, with the more general formulation of the auxiliary RVs in (51), it is possible to greatly reduce the distance between inner and outer bounds.

Although the scheme (E) in Section IV-E does not achieve capacity outside the “primary decodes cognitive” regime, we next show by numerical evaluation that scheme (E) is close to optimal for a general channel in “strong interference”, especially when considering the union over all $\lambda \in \mathbb{C}$ instead of the choice $\lambda = \lambda_{\text{Costa } 1}$. Fig. 19 shows the position of point

$$D(\lambda) = ((29a), \min\{(29b), (29c) - (29a)\})$$

in the range $\lambda \in [0, 2\lambda_{\text{Costa } 1}]$, for a fixed $\alpha^{(\text{in})}$, together with the outer bound point C for $\alpha^{(\text{out})} = \alpha^{(\text{in})}$. Under the “primary decodes cognitive” condition, $D(\lambda_{\text{Costa } 1}) = C$ for every $\alpha \in [0, 1]$. However, here we show a channel where the condition in (30a) is not satisfied. In this case the choice $\lambda = \lambda_{\text{Costa } 1}$ minimizes the distance of the R_1 -coordinate between D and C,

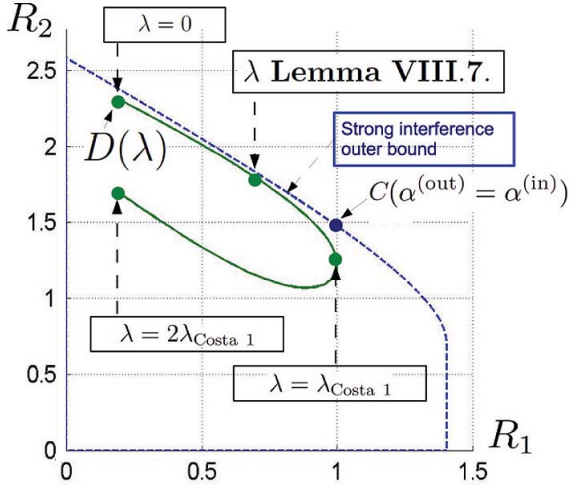


Fig. 19. Plot of points $C(\lambda)$ and $D(\lambda)$ for $\alpha = 0.5$, together with the “strong interference” bound for the G-CIFC with parameters $P_1 = P_2 = 6$, $|b| = \sqrt{2}$ and $a = \sqrt{0.3}$.

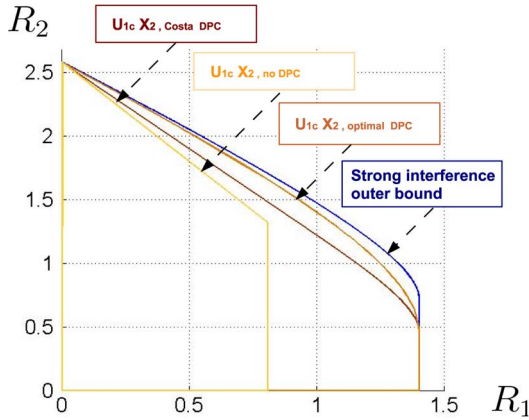


Fig. 20. Achievable rate region in (29) for $\lambda = 0$, $\lambda = \lambda_{\text{Costa } 1}$ and any $\lambda \in [0, 2\lambda_{\text{Costa } 1}]$ for the G-CIFC with parameters $P_1 = P_2 = 6$, $|b| = \sqrt{2}$ and $a = \sqrt{0.3}$.

but it does not minimize the Euclidean distance between the two points.

The rate improvements that may be obtained with any $\lambda \in \mathbb{C}$ are exemplified in Fig. 20. In this figure we plot the achievable rate regions of (29) obtained for $\lambda = 0$, $\lambda = \lambda_{\text{Costa } 1}$ and any $\lambda \in [0, 2\lambda_{\text{Costa } 1}]$. Unlike Fig. 14, the scheme for $\lambda = \lambda_{\text{Costa } 1}$ strictly outperforms the scheme for $\lambda = 0$; the choice $\lambda \in [0, 2\lambda_{\text{Costa } 1}]$ not only includes the previous regions but improves on the case $\lambda = \lambda_{\text{Costa } 1}$ as well. The inner bound point for $R_1 = 0$ corresponds to point A in (42) and is always achievable; the inner bound point for $R_1 = C(P_1)$ may be achieved only for $\lambda = \lambda_{\text{Costa } 1}$.

VIII. CONCLUSION AND FUTURE WORK

In this paper, we presented outer bounds, inner bounds, and new capacity results for the Gaussian cognitive interference channel. We derived the tightest known outer bound for the cognitive interference channel in “strong interference,” which is based on the capacity of the MIMO BC with degraded message sets. We showed the achievability of this outer bound in the subset of the channel parameter space which we term the “primary decodes cognitive” regime and for some “S-channels

in strong interference.” We also proved capacity to within both an additive and a multiplicative gap, thus providing an approximate characterization of the capacity region in both high and low SNR.

Despite the new results presented, the capacity of the Gaussian cognitive interference channel remains unknown in general. The achievable rate region of [13], which is Fourier-Motzkin eliminated here, resulting in an easier to understand region, provides a comprehensive inner bound that may yield new capacity results: only some specific choices of parameters for this region have been considered so far and we expect that new results may be derived from this region. We have shown that the tightest outer bound for the Gaussian cognitive interference channel in “strong interference” is obtained as the intersection of different bounds. The expression of this outer bound does not have a simple closed form expression except in some special cases such as the S and the degraded channels. Even in these two subcases, capacity is not known in general. Another interesting open question is how much rate improvement is attainable with binning at the cognitive encoder: we have shown how dirty paper coding may be used to boost the rate of both the primary and the cognitive user; whether non perfect interference cancellation achieves capacity is still unknown.

APPENDIX A

THE G-CIFC IN STANDARD FORM

A general G-CIFC has outputs

$$\begin{aligned}\tilde{Y}_1 &= h_{11}\tilde{X}_1 + h_{12}\tilde{X}_2 + \tilde{Z}_1, \\ \tilde{Y}_2 &= h_{21}\tilde{X}_1 + h_{22}\tilde{X}_2 + \tilde{Z}_2\end{aligned}$$

where

$$\tilde{Z}_i \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_i^2), \quad \sigma_i^2 > 0, \quad i \in \{1, 2\}$$

and the inputs are subject to the power constraint

$$\mathbb{E}[|\tilde{X}_i|^2] \leq \tilde{P}_i, \quad \tilde{P}_i \geq 0, \quad i \in \{1, 2\}.$$

When $h_{11} \neq 0$ and $h_{22} \neq 0$, we may scale the channel outputs and redefine the inputs as follows:

$$Y_1 := \frac{\tilde{Y}_1}{\sigma_1}, \quad (52a)$$

$$Y_2 := \frac{\tilde{Y}_2}{\sigma_2} e^{j(\angle h_{11} - \angle h_{12})}, \quad (52b)$$

$$X_1 := \frac{h_{11}}{\sigma_1} \tilde{X}_1$$

such that $\mathbb{E}[|X_1|^2] \leq P_1 := \frac{|h_{11}|^2}{\sigma_1^2} \tilde{P}_1$, (52c)

$$X_2 := \frac{h_{22}}{\sigma_2} e^{j(\angle h_{11} - \angle h_{12})} \tilde{X}_2$$

such that $\mathbb{E}[|X_2|^2] \leq P_2 := \frac{|h_{22}|^2}{\sigma_2^2} \tilde{P}_2$, (52d)

$$a := \frac{h_{12}}{\sigma_1} \frac{\sigma_2}{h_{22}} e^{j(-\angle h_{11} + \angle h_{12})} \in \mathbb{C}, \quad (52e)$$

$$b := \frac{|h_{21}|}{\sigma_2} \frac{\sigma_1}{|h_{11}|} \in \mathbb{R}^+ \quad (52f)$$

to obtain the equivalent channel outputs that have additive noise of unit variance and unit gain on the direct link, as in (1). To remind the reader that b as defined in (52f) is always real-valued and non-negative we use the notation $|b|$.

When $h_{22} = 0$, transmitter 2 can only create interference at receiver 1 and thus the channel reduces to a BC where the cognitive transmitter is sending a message to each receiver. The case $h_{22} = 0$ in the general channel is equivalent to $P_2 = 0$ in the channel in standard form.

The same is not true when $h_{11} = 0$. When $h_{11} = 0$, receiver 1 only receives interference and noise, hence the rate for user 1 is zero; this is possible in the channel in standard form with $P_1 = 0$. However, receiver 2 can get information from both transmitters as in a MISO point-to-point channel and therefore its rate depends on both P_1 and P_2 in the channel in standard form. This means that one cannot set $P_1 = 0$ in the channel in standard form and recover the case $h_{11} = 0$ in the general channel. In [11, Sec. II-B]], this fact is overlooked and the transformation in (52) is claimed to be without loss of generality.

APPENDIX B

PROOF OF THEOREM II.3

For a complex-valued G-CIFC with $|b| > 1$, the outer bound of Theorem II.2 is achievable by the superposition coding only (scheme (D) of Section IV-D) if $I(Y_1; X_1, X_2) \geq I(Y_2; X_1, X_2)$ for all input distributions [12], that is, if

$$\begin{aligned} \mathbb{E}[|Y_1|^2] - \mathbb{E}[|Y_2|^2] &= (|a|^2 - 1)P_2 - (|b|^2 - 1)P_1 + \\ &+ 2\sqrt{P_1 P_2}(\Re\{a^* \rho\} - |b|\Re\{\rho\}) \geq 0, \quad \forall |\rho| \leq 1. \end{aligned} \quad (53)$$

Let $\rho = |\rho|e^{j\phi_\rho}$ and $a = |a|e^{j\phi_a}$. We have

$$\begin{aligned} \Re\{a^* \rho\} - |b|\Re\{\rho\} &= |\rho||a| \cos(\phi_\rho - \phi_a) - |\rho||b| \cos(\phi_\rho) \\ &= |\rho| [|a| \cos(\phi_a) - |b|] \cos(\phi_\rho) + |\rho| [|a| \sin(\phi_a)] \sin(\phi_\rho) \\ &= |\rho| \sqrt{(|a| \cos(\phi_a) - |b|)^2 + [|a| \sin(\phi_a)]^2} \cos(\phi) \\ &= |a - |b|| \cdot |\rho| \cos(\phi) \end{aligned}$$

for some angle ϕ . The condition in (53) is thus verified for all $|\rho| \cos(\phi) \in [-1, +1]$ if it is verified for $|\rho| \cos(\phi) = -1$ as claimed in Theorem II.3. ■

APPENDIX C

CLOSED FORM EXPRESSION FOR $\mathcal{O}^{(\text{BC-PR})}$ FOR THE GAUSSIAN MIMO BC

The closed form expression of $\mathcal{O}^{(\text{BC-PR})}$ was obtained in [37] and is presented here for completeness.

Consider an input covariance matrix defined as follows

$$\mathbf{S} := \begin{bmatrix} P_1 & \rho\sqrt{P_1 P_2} \\ \rho^*\sqrt{P_1 P_2} & P_2 \end{bmatrix} : |\rho| \leq 1. \quad (54)$$

The capacity region of a Gaussian *MIMO BC with private rates only* with a per-antenna power constraint is given by [37]

$$\mathcal{O}^{(\text{BC-PR})} = \mathcal{CH} \bigcup_{\mathbf{S}} \mathcal{O}^{(\text{BC-PR})}(\mathbf{S})$$

where \mathcal{CH} denotes the convex-hull operation, $\bigcup_{\mathbf{S}}$ denotes the union over all input covariance matrices \mathbf{S} in (54),

$$\mathcal{O}^{(\text{BC-PR})}(\mathbf{S}) = \bigcup_{u \in \{1,2\}} \mathcal{R}^{(\text{DPC } u)}(\mathbf{S})$$

where $\mathcal{R}^{(\text{DPC } u)}(\mathbf{S})$ is the DPC region for the encoding order where user u is pre-coded against the interference created by the other user at its intended receiver, which is given by

$$\begin{aligned} \mathcal{R}^{(\text{DPC } u)}(\mathbf{S}) &= \bigcup_{0 \leq \mathbf{B}_1, 0 \leq \mathbf{B}_2, \mathbf{B}_1 + \mathbf{B}_2 = \mathbf{S}} \mathcal{R}^{(\text{DPC } u)}(\mathbf{B}_1, \mathbf{B}_2), \quad u \in \{1, 2\} \end{aligned}$$

and where, for

$$\mathbf{B}_1 := \begin{bmatrix} \alpha_1 P_1 & \rho_1 \sqrt{\alpha_1 P_1 \alpha_2 P_2} \\ \rho_1^* \sqrt{\alpha_1 P_1 \alpha_2 P_2} & \alpha_2 P_2 \end{bmatrix}, \quad (55a)$$

$$\mathbf{B}_2 := \begin{bmatrix} \bar{\alpha}_1 P_1 & \rho_2 \sqrt{\bar{\alpha}_1 P_1 \bar{\alpha}_2 P_2} \\ \rho_2^* \sqrt{\bar{\alpha}_1 P_1 \bar{\alpha}_2 P_2} & \bar{\alpha}_2 P_2 \end{bmatrix} \quad (55b)$$

with

$$(\alpha_1, \alpha_2, |\rho_1|, |\rho_2|) \in [0, 1]^4 : \rho_1 \sqrt{\alpha_1 \alpha_2} + \rho_2 \sqrt{\bar{\alpha}_1 \bar{\alpha}_2} = \rho \quad (56)$$

the region $\mathcal{R}^{(\text{DPC } 1)}(\mathbf{B}_1, \mathbf{B}_2)$ is given by

$$R_1 \leq \mathcal{C}(\alpha_1 P_1 + |a|^2 \alpha_2 P_2 + 2\Re\{a^* \rho_1\} \sqrt{\alpha_1 \alpha_2 P_1 P_2}), \quad (57a)$$

$$R_2 \leq \mathcal{C} \left(\frac{\bar{\alpha}_1 |b|^2 P_1 + \bar{\alpha}_2 P_2 + 2\Re\{\rho_2\} \sqrt{\bar{\alpha}_1 \bar{\alpha}_2 |b|^2 P_1 P_2}}{1 + |b|^2 \alpha_1 P_1 + \alpha_2 P_2 + 2\Re\{\rho_1\} \sqrt{\alpha_1 \alpha_2 |b|^2 P_1 P_2}} \right) \quad (57b)$$

and $\mathcal{R}^{(\text{DPC } 2)}(\mathbf{B}_1, \mathbf{B}_2)$ is given by

$$R_1 \leq \mathcal{C} \left(\frac{\alpha_1 P_1 + |a|^2 \alpha_2 P_2 + 2\Re\{a^* \rho_1\} \sqrt{\alpha_1 \alpha_2 P_1 P_2}}{1 + \bar{\alpha}_1 P_1 + |a|^2 \bar{\alpha}_2 P_2 + 2\Re\{a^* \rho_2\} \sqrt{\bar{\alpha}_1 \bar{\alpha}_2 P_1 P_2}} \right), \quad (58a)$$

$$R_2 \leq \mathcal{C}(\bar{\alpha}_1 |b|^2 P_1 + \bar{\alpha}_2 P_2 + 2\Re\{\rho_2\} \sqrt{\bar{\alpha}_1 \bar{\alpha}_2 |b|^2 P_1 P_2}). \quad (58b)$$

The quantity $\alpha_u, u \in \{1, 2\}$, represents the fraction of power P_u used to send the cognitive message W_1 on antenna u . The requirement $(\alpha_1, \alpha_2) \in [0, 1]^2$ guarantees that the per-antenna power constraints are verified.

APPENDIX D

PROOF OF COROLLARIES III.4 AND III.5

A. Proof of Corollary III.4

When allowing full transmitter cooperation for a channel with $a|b| = 1$ and $|b| > 1$, we obtain an equivalent degraded BC with input $X_{\text{eq}} = |b|X_1 + X_2$ and outputs

$$\begin{aligned} Y_2 &= (|b|X_1 + X_2) + Z_2 = X_{\text{eq}} + Z_2, \\ |b|Y_1 &= (|b|X_1 + X_2) + |b|Z_1 \sim Y_2 + \sqrt{|b|^2 - 1} Z_0 \end{aligned}$$

with $Z_0 \sim \mathcal{N}_{\mathbb{C}}(0, 1)$ and independent of everything else. The input of the equivalent BC is subject to the power constraint

$$\mathbb{E}[|X_{\text{eq}}|^2] \leq (\sqrt{|b|^2 P_1} + \sqrt{P_2})^2 := P_{\text{eq}}.$$

For this order of degradedness among the users, the capacity region of the degraded BC with private rates equals the capacity with degraded message sets. In general $\mathcal{O}^{(\text{BC-DMS})} \subseteq \mathcal{O}^{(\text{BC-PR})}$, but since here Y_1 is a degraded version of Y_2 , decoder 2 can decode the message of decoder 1 without imposing any rate penalty to user 1, thus $\mathcal{O}^{(\text{BC-PR})}$ is achievable. This implies $\mathcal{O}^{(\text{BC-DMS})} = \mathcal{O}^{(\text{BC-PR})}$.

The capacity region of the equivalent BC is [38]

$$R_1 \leq \mathcal{C} \left(\frac{\alpha' P_{\text{eq}}}{(1 - \alpha') P_{\text{eq}} + |b|^2} \right) =: R_1^{(\text{BC-deg})}(\alpha'), \quad (59a)$$

$$R_2 \leq \mathcal{C}((1 - \alpha') P_{\text{eq}}) =: R_2^{(\text{BC-deg})}(\alpha') \quad (59b)$$

taken over the union of all $\alpha' \in [0, 1]$, i.e., that is $\rho_1 = \rho_2 = 1$, $\alpha_1 = \alpha_2 = \alpha'$ and $\mathcal{O}^{(\text{BC-PR})} = \mathcal{R}^{(\text{DPC } 2)}$ in (58).

To intersect the region in (59) with the ‘‘strong interference’’ outer bound $\mathcal{O}^{(\text{SI})}$ in (5) of Theorem II.2 we equate the R_1 -bounds in (59a) and (5a) to obtain

$$\alpha' = \frac{\alpha P_1}{1 + \alpha P_1} \left(1 + \frac{|b|^2}{(\sqrt{|b|^2 P_1} + \sqrt{P_2})^2} \right). \quad (60)$$

Notice that α' in (60) satisfies $\alpha' \leq 1$ (the maximum value of 1 is obtained for $P_2 = 0$ and $\alpha = 1$). By substituting α' from (60) in (59b), we obtain the bound in (14b).

Let $R_{\text{sum}}^{(\text{SI})}(\alpha)$ denote the right-hand-side of (5b) as a function of α . The BC-DMS-based outer bound is more stringent than the ‘‘strong interference’’ outer bound if

$$\begin{aligned} R_1^{(\text{BC-deg})}(\alpha) + R_2^{(\text{BC-deg})}(\alpha) &\leq R_{\text{sum}}^{(\text{SI})}(\alpha) \quad \forall \alpha \in [0, 1] \\ \iff \alpha P_1 + P_2 + (1 - \alpha)|b|^2 P_1 + 2\sqrt{|b|^2 P_1 P_2} &\leq \\ &P_2 + |b|^2 P_1 + 2\sqrt{\bar{\alpha}}|b|^2 P_1 P_2 \quad \forall \alpha \in [0, 1] \\ \iff 2\sqrt{|b|^2 P_1 P_2}(1 - \sqrt{\bar{\alpha}}) &\leq \alpha P_1(|b|^2 - 1) \quad \forall \alpha \in [0, 1] \\ \iff 2\frac{\sqrt{|b|^2 P_1 P_2}}{P_1(|b|^2 - 1)} &\leq 1 + \sqrt{\bar{\alpha}} \quad \forall \alpha \in [0, 1] \\ &(\text{since } \alpha = (1 - \sqrt{\bar{\alpha}})(1 + \sqrt{\bar{\alpha}})) \\ \iff 2\frac{\sqrt{|b|^2 P_1 P_2}}{P_1(|b|^2 - 1)} &\leq \min_{\bar{\alpha} \in [0, 1]} \{1 + \sqrt{\bar{\alpha}}\} = 1 \\ \iff 1 + \frac{P_2}{P_1} &\leq \left(|b| - \sqrt{\frac{P_2}{P_1}} \right)^2 \\ \iff |b| &\geq \sqrt{1 + \frac{P_2}{P_1}} + \sqrt{\frac{P_2}{P_1}} \end{aligned}$$

as claimed in (15).

The capacity of the equivalent degraded BC may be achieved both by using superposition coding and binning. An achievable scheme inspired by the degraded BC and employing superposition coding is scheme (E) with $\lambda = 0$. An achievable scheme inspired by the degraded BC and employing binning coding is scheme (B). Both schemes achieve the outer bound only in point A in (42). The capacity region of the degraded CIFC is therefore unknown in general it remains an interesting open problem.

B. Proof of Corollary III.5

To establish the result in Corollary III.5 we proceed as follows. Firstly, we prove that the capacity region of the Gaussian MIMO BC-DMS may be obtained from the region in [40] by considering jointly Gaussian inputs and auxiliary RV. Secondly, we perform a partial optimization of the region in [40] in the Gaussian case and obtain a looser outer bound that may be expressed as a function of a single free parameter. Finally, we intersect this outer bound with the ‘‘strong interference’’ outer bound $\mathcal{O}^{(\text{SI})}$ in (5) to obtain the expression in (16).

The capacity region of the general BC-DMS is found in [40] and is expressed as the union over all possible distributions of the input and one auxiliary RV. A closed form expression of the capacity region of the Gaussian MIMO BC-DMS is derived in [39] and is expressed as the intersection of the capacity region of a general BC-PR and an additional sum-rate constraint. Here, we derive a simpler expression of the capacity region of the Gaussian MIMO BC-DMS than [39]; we do so by showing that we may restrict the union in [40] over all jointly Gaussian inputs and auxiliary RV.

Consider the Gaussian MIMO BC-DMS defined as

$$Y_i = \mathbf{H}_i X + Z_i \quad \forall i \in \{1, 2\} \quad (61)$$

where

- X is a real-valued input vector of size $n \times 1$ subject to the second moment constraint $\text{Cov}[X] = \mathbf{K}_X \preceq \mathbf{S}$ for some $\mathbf{S} \succeq 0$,
- Y_i is a real-valued output vector of size $m_i \times 1$ received by user $i \in \{1, 2\}$,
- \mathbf{H}_i is a fixed real-valued gain matrix imposed on user $i \in \{1, 2\}$. This is a matrix of size $m_i \times n$,
- Z_i is a length- m_i real-valued Gaussian random vector with zero mean and covariance matrix $\text{Cov}[Z_i] = \mathbf{K}_Z \succ 0$.

As for the Gaussian MIMO BC-PR of [39], we consider real-valued channels first. The extension to complex-valued channels is easily obtained by doubling the number of real dimensions. Also, as for the Gaussian MIMO BC-PR of [39], we first derive the capacity of a Gaussian MIMO BC-DMS for the case where \mathbf{H}_i is square and invertible. We then argue that the case for a general \mathbf{H}_i may be obtained by the series of channel transformations in [37].

Theorem D.1: The capacity region of the Gaussian MIMO BC-DMS in (61) with input covariance constraint $\text{Cov}[(X_1 \ X_2)] \preceq \mathbf{S}$ is

$$R_1 \leq I(U; Y_1), \quad (62a)$$

$$R_2 \leq I(X; Y_2|U), \quad (62b)$$

$$R_1 + R_2 \leq I(X; Y_2) \quad (62c)$$

taken over the union of all jointly Gaussian U and X vectors of size n such that the input covariance constraint is satisfied.

Proof: The region in (62) was originally obtained in [40] for a general BC-DMS and is expressed as the union over all distribution P_{UX} . To prove the theorem we need to show that only jointly Gaussian U and X need to be considered. First, we notice that (62c) is always maximized by having X Gaussian with covariance \mathbf{S} by the ‘‘Gaussian maximizes entropy’’ principle [16]. Since (62c) is maximized by Gaussian input, we have to show that the region obtained by considering (62a) and (62b) only is optimized by a jointly Gaussian U and X . To this end, the points on the convex closure of the region with (62a) and (62b) satisfy for some $\mu \in [0, 1]$ and $\bar{\mu} = 1 - \mu$

$$\begin{aligned} & \max\{\mu R_1 + \bar{\mu} R_2\} \\ &= \max_{\substack{P_{X|U}: \\ \text{Cov}[X|U] \preceq \mathbf{S}}} \mu I(U; Y_1) + \bar{\mu} I(X; Y_2|U) \\ &\leq \mu h(\mathbf{H}_1 X_{G,\mathbf{S}} + Z_1) - \bar{\mu} h(Z_2) \\ &+ \bar{\mu} \max_{\substack{P_{X|U}: \\ \text{Cov}[X|U] \preceq \mathbf{S}}} \left(h(\mathbf{H}_2 X + Z_2|U) - \frac{\mu}{\bar{\mu}} h(\mathbf{H}_1 X + Z_1|U) \right) \end{aligned} \quad (63)$$

where $X_{G,\mathbf{S}}$ indicates a zero-mean jointly Gaussian vector with covariance \mathbf{S} . In (63), the notation $h(X)$ indicates the differential entropy of the RV X .

Next we show that it suffices to consider $\mu \in (1/2, 1]$ in (63) rather than $\mu \in [0, 1]$. The region in (62) is contained in the triangular region

$$\{(R_1, R_2) \in \mathbb{R}_+^2 : R_1 + R_2 \leq I(\mathbf{H}_2 X_{G,\mathbf{S}} + Z_2; X_{G,\mathbf{S}})\} \quad (64)$$

whose boundary point

$$A' = (R_1^{(A')}, R_2^{(A')}) = (0, I(\mathbf{H}_2 X_{G,\mathbf{S}} + Z_2; X_{G,\mathbf{S}})).$$

is also in (62). Since (62) is convex and intersects (64) in A' , the region in (62) cannot contain any rate point with tangent greater than -1 . Hence, there is no loss of generality in restricting μ in (63) to the interval $(1/2, 1]$.

We now show that solution of the optimization problem in (63), for $\mu \in (1/2, 1]$, must be a jointly Gaussian (X, U) by using the extremal inequality of [46]. We first focus on channels where \mathbf{H}_i , $i \in \{1, 2\}$ is square and invertible, then show how this result may be extended to a general channel using the perturbation techniques of [37]. With \mathbf{H}_i , $i \in \{1, 2\}$, square and invertible, let $|\mathbf{H}_i|$ denote the (non-zero) determinant of \mathbf{H}_i and write

$$\begin{aligned} & \max_{\substack{P_{X|U}: \\ \text{Cov}[X|U] \preceq \mathbf{S}}} h(\mathbf{H}_2 X + Z_2|U) - \frac{\mu}{\bar{\mu}} h(\mathbf{H}_1 X + Z_1|U) \\ &= \frac{\mu}{\bar{\mu}} (\log |\mathbf{H}_1|)^{-1} - (\log |\mathbf{H}_2|)^{-1} \end{aligned}$$

$$+ \max_{\substack{P_{X|U}: \\ \text{Cov}[X|U] \preceq \mathbf{S}}} h(X + \mathbf{H}_2^{-1} Z_2|U) - \frac{\mu}{\bar{\mu}} h(X + \mathbf{H}_1^{-1} Z_1|U). \quad (65)$$

The result in [46, Th. 8] grants that the optimal solution of the optimization problem in (65) is a Gaussian X conditioned on U with $\text{Cov}[X|U] \preceq \mathbf{S}$, since $\mu/\bar{\mu} > 1$ for $\mu \in (1/2, 1]$; this is possible if X and U are jointly Gaussian. We therefore conclude that the region in (62) is exhausted by jointly Gaussian X and U . Note that U has the same dimension of the input X .

Finally, the perturbation technique in [37, Section V-B] allows us to extend this result to a general channel where \mathbf{H}_i in not square or invertible. The derivation in [37, Sec. V-B] was originally devised for the general BC-PR but it extends in a straightforward manner to the BC-DMS since it solely relies on the channel matrix and the covariance of the noise and not on the message set. ■

Theorem D.1 shows that for the G-CIFC in (1) the Gaussian MIMO BC-DMS $\mathcal{O}^{(\text{BC-DMS})}$, with the parameterization given in (55), is

$$R_1 \leq (58a), \quad (66a)$$

$$R_2 \leq (58b), \quad (66b)$$

$$\begin{aligned} R_1 + R_2 \leq & \mathcal{C}(|b|^2 P_1 + P_2 + \\ & 2\Re\{\rho_1 \sqrt{\alpha_1 \bar{\alpha}_2} + \rho_2 \sqrt{\bar{\alpha}_1 \alpha_2}\} \sqrt{|b|^2 P_1 P_2}) \end{aligned} \quad (66c)$$

taken over the union of all $(\alpha_1, \alpha_2, \rho_1, \rho_2)$ that satisfy (56).

For the S-G-CIFC ($a = 0$), it is straightforward to see that $\rho_1 = \rho_2 = 1$ is optimal in (66), and the the outer bound simplifies to

$$R_1 \leq \mathcal{C}\left(\frac{\alpha_1 P_1}{1 + \bar{\alpha}_1 P_1}\right), \quad (67a)$$

$$R_2 \leq \mathcal{C}\left(\bar{\alpha}_1 |b|^2 P_1 + \bar{\alpha}_2 P_2 + 2\sqrt{\bar{\alpha}_1 \bar{\alpha}_2 |b|^2 P_1 P_2}\right), \quad (67b)$$

$$\begin{aligned} R_1 + R_2 \leq & \mathcal{C}(|b|^2 P_1 + P_2 + \\ & 2(\sqrt{\alpha_1 \bar{\alpha}_2} + \sqrt{\bar{\alpha}_1 \alpha_2}) \sqrt{|b|^2 P_1 P_2}) \end{aligned} \quad (67c)$$

taken over the union of all $(\alpha_1, \alpha_2) \in [0, 1]^2$. Since the region in (67) is still expressed as a function of two free parameters, we further upper bound it by dropping the sum-rate bound in (67c) and by choosing the α_2 that maximizes the R_2 -bound in (67b), that is, $\alpha_2 = 0$; we thus obtain that the region $\mathcal{O}^{(\text{BC-DMS})}$ for the S-G-CIFC is included into

$$R_1 \leq \mathcal{C}\left(\frac{\alpha_1 P_1}{1 + \bar{\alpha}_1 P_1}\right), \quad (68a)$$

$$R_2 \leq \mathcal{C}\left((\sqrt{P_2} + \sqrt{\bar{\alpha}_1 |b|^2 P_1})^2\right) \quad (68b)$$

taken over the union of all $\alpha_1 \in [0, 1]$. By setting

$$\alpha = \frac{\alpha_1}{1 + \bar{\alpha}_1 P_1}$$

we have that the region in (68) may be rewritten as

$$R_1 \leq C(\alpha P_1) =: R_1^{(\text{BC-S})}(\alpha), \quad (69a)$$

$$R_2 \leq C \left(\left(\sqrt{|b|^2 \frac{\bar{\alpha} P_1}{1 + \alpha P_1}} + \sqrt{P_2} \right)^2 \right) =: R_2^{(\text{BC-S})}(\alpha) \quad (69b)$$

taken over the union of all $\alpha \in [0, 1]$.

The BC-DMS-based outer bound of (69) is more stringent than the ‘‘strong interference’’ outer bound in (5) if

$$\begin{aligned} R_1^{(\text{BC-S})}(\alpha) + R_2^{(\text{BC-S})}(\alpha) &\leq R_{\text{sum}}^{(\text{SI})}(\alpha) \quad \forall \alpha \in [0, 1] \\ \iff \alpha P_1 + \left(\sqrt{|b|^2(1-\alpha)P_1} + \sqrt{P_2(1+\alpha P_1)} \right)^2 &\leq \\ &P_2 + |b|^2 P_1 + 2\sqrt{(1-\alpha)|b|^2 P_1 P_2} \quad \forall \alpha \in [0, 1] \\ \iff 2\sqrt{(1-\alpha)|b|^2 P_1 P_2}(\sqrt{1+\alpha P_1} - 1) &+ \\ &+ \alpha P_1(1 + P_2 - |b|^2) \leq 0 \quad \forall \alpha \in [0, 1] \\ \iff 1 + P_2 - |b|^2 \leq 0 \end{aligned}$$

as claimed in (17).

APPENDIX E PROOF OF LEMMA III.6

Let $X_1^N(W_1, W_2), X_2^N(W_2)$ be a good code for the channel with parameters (a, b, P_1, P_2) , where we note that we have left b complex but note that only its magnitude $|b|$ affects the capacity region. Consider now the inputs

$$\begin{aligned} X'_1 &= AX_1 + BX_2, \\ X'_2 &= CX_2 \end{aligned}$$

on a channel with parameters (a', b', P'_1, P'_2) resulting in the outputs

$$\begin{aligned} Y'_1 &= X'_1 + a'X'_2 + Z_1 \propto X_1 + \frac{B + a'C}{A}X_2 + \frac{Z_1}{A}, \quad (70a) \\ Y'_2 &= b'X'_1 + X'_2 + Z_2 \propto \frac{b'A}{b'B + C}X_1 + X_2 + \frac{Z_2}{b'B + C} \end{aligned}$$

for some $(A, B, C) \in \mathbb{C}^3$. If we impose

$$a = \frac{B + a'C}{A}, \quad (71a)$$

$$b = \frac{b'A}{b'B + C}, \quad (71b)$$

$$|A|^2 \geq 1, \quad (71c)$$

$$|b'B + C|^2 \geq 1, \quad (71d)$$

$$P'_1 \geq (\sqrt{|A|^2 P_1} + \sqrt{|B|^2 P_2})^2, \quad (71e)$$

$$P'_2 \geq |C|^2 P_2 \quad (71f)$$

the output of the channel (a, b, P_1, P_2) may be reconstructed in the channel (a', b', P'_1, P'_2) since the latter is less noisy than the former. This implies the relationship claimed in (18).

APPENDIX F PROOF OF THEOREM IV.1

In [13, Th. 4.1] we showed that a non-negative rate pair (R_1, R_2) such that

$$R_1 = R_{1c} + R_{1pb}, \quad (72a)$$

$$R_2 = R_{2c} + R_{2pa} + R_{2pb} \quad (72b)$$

is achievable for a general CIFC if

$$(R'_{1c}, R'_{1pb}, R'_{2pb}, R_{1c}, R_{1pb}, R_{2c}, R_{2pa}, R_{2pb}) \in \mathbb{R}_+^8$$

satisfies the inequalities in (73a–k), shown at the bottom of the page, for some input distribution $P_{U_{1c}, U_{2c}, U_{1pb}, U_{2pb}, X_1, X_2}$. Note that we have used notation $(a_1, d_1, g_1, a_2, d_2, g_2)$ reminiscent of that of [47]. We first show that without loss of generality one can take $R_{2pb} = 0$ in (73). For notational convenience, let $\mathbb{R}^{(\text{RTD})}$ indicate the region in (73), $\mathbb{R}^{(\text{RTDup})}$ indicate the region in (73) after dropping the rate constraints in (73g) and in (73h), and $\mathbb{R}^{(\text{RTDlow})}$ indicate the region in (73) with $R_{2pb} = 0$. Note that $\mathbb{R}^{(\text{RTDlow})}$ and $\mathbb{R}^{(\text{RTD})}$ are achievable while $\mathbb{R}^{(\text{RTDup})}$ might not be. In general $\mathbb{R}^{(\text{RTDlow})} \subseteq \mathbb{R}^{(\text{RTD})} \subseteq \mathbb{R}^{(\text{RTDup})}$.

$$R'_{1c} \geq I(U_{1c}; X_2 | U_{2c}), \quad (73a)$$

$$R'_{1c} + R'_{1pb} \geq I(U_{1pb}; X_2 | U_{1c}, U_{2c}) + I(U_{1c}; X_2 | U_{2c}), \quad (73b)$$

$$R'_{1c} + R'_{1pb} + R'_{2pb} \geq I(U_{1pb}; U_{2pb}, X_2 | U_{1c}, U_{2c}) + I(U_{1c}; X_2 | U_{2c}), \quad (73c)$$

$$R_{2c} + R_{2pa} + (R_{1c} + R'_{1c}) + (R_{2pb} + R'_{2pb}) \leq I(Y_2; U_{2pb}, U_{1c}, X_2, U_{2c}) + I(U_{1c}; X_2 | U_{2c}) := g_2, \quad (73d)$$

$$R_{2pa} + (R_{1c} + R'_{1c}) + (R_{2pb} + R'_{2pb}) \leq I(Y_2; U_{2pb}, U_{1c}, X_2 | U_{2c}) + I(U_{1c}; X_2 | U_{2c}) := d_2, \quad (73e)$$

$$R_{2pa} + (R_{2pb} + R'_{2pb}) \leq I(Y_2; U_{2pb}, X_2 | U_{1c}, U_{2c}) + I(U_{1c}; X_2 | U_{2c}) := a_2, \quad (73f)$$

$$(R_{1c} + R'_{1c}) + (R_{2pb} + R'_{2pb}) \leq I(Y_2; U_{2pb}, U_{1c} | X_2, U_{2c}) + I(U_{1c}; X_2 | U_{2c}) \quad (73g)$$

$$(R_{2pb} + R'_{2pb}) \leq I(Y_2; U_{2pb} | U_{1c}, X_2, U_{2c}), \quad (73h)$$

$$R_{2c} + (R_{1c} + R'_{1c}) + (R_{1pb} + R'_{1pb}) \leq I(Y_1; U_{1pb}, U_{1c}, U_{2c}) := g_1, \quad (73i)$$

$$(R_{1c} + R'_{1c}) + (R_{1pb} + R'_{1pb}) \leq I(Y_1; U_{1pb}, U_{1c} | U_{2c}) := d_1, \quad (73j)$$

$$(R_{1pb} + R'_{1pb}) \leq I(Y_1; U_{1pb} | U_{1c}, U_{2c}) := a_1. \quad (73k)$$

The key observation here is that $\mathcal{R}^{(\text{RTD}_{\text{up}})}$ depends on R_{2pa} and R_{2pb} only through their sum $R_{2pa} + R_{2pb}$. By Fourier-Motzkin elimination of the region $\mathcal{R}^{(\text{RTD}_{\text{up}})}$ on the subspace

$$\begin{aligned}\widehat{R}_1 &:= R_{1c} + R_{1pb} + R'_{1c} + R'_{1pb} = R_1 + R'_{1c} + R'_{1pb}, \\ \widehat{R}_2 &:= R_{2c} + R_{2pa} + R_{2pb} + R'_{2pb} = R_2 + R'_{2pb}\end{aligned}$$

for fixed binning rates $(R'_{1c}, R'_{1pb}, R'_{2pb})$ we obtain

$$\widehat{R}_1 \leq \min\{d_1, a_1 + d_2\}, \quad (74a)$$

$$\widehat{R}_2 \leq g_2, \quad (74b)$$

$$\widehat{R}_1 + \widehat{R}_2 \leq \min\{g_1 + a_2, g_2 + a_1\}, \quad (74c)$$

$$2\widehat{R}_1 + \widehat{R}_2 \leq g_1 + a_1 + d_2 \quad (74d)$$

Next, we note that in $\mathcal{R}^{(\text{RTD}_{\text{low}})}$ the rate constraints in (73g) and in (73h) with $R_{2pb} = 0$ can be dropped without loss of generality because they correspond to an error at receiver 2 on the common message of user 1 and on a bin index, respectively. However, in either case all the intended message indices at receiver 2 are correct. This means that in $\mathcal{R}^{(\text{RTD}_{\text{low}})}$ the rate constraints in (73g) and in (73h) with $R_{2pb} = 0$ are not needed to drive the probability of error to zero. By Fourier-Motzkin elimination of the region $\mathcal{R}^{(\text{RTD}_{\text{low}})}$ on the subspace

$$\begin{aligned}\widehat{R}_1 &:= R_{1c} + R_{1pb} + R'_{1c} + R'_{1pb}, \\ \widehat{R}_2 &:= R_{2c} + R_{2pa} + R'_{2pb}\end{aligned}$$

for fixed binning rates $(R'_{1c}, R'_{1pb}, R'_{2pb})$ we obtain the same region as in (74). This shows that $\mathcal{R}^{(\text{RTD}_{\text{low}})} = \mathcal{R}^{(\text{RTD}_{\text{up}})}$, that is, that $R_{2pb} = 0$ is without loss of generality in $\mathcal{R}^{(\text{RTD})}$ and that the Fourier-Motzkin elimination of the region $\mathcal{R}^{(\text{RTD})}$ for fixed binning rates $(R'_{1c}, R'_{1pb}, R'_{2pb})$ is given by (74).

As a final step, we express the region in (74) as function of R_1 and R_2 only by Fourier-Motzkin elimination of the binning rates $(R'_{1c}, R'_{1pb}, R'_{2pb})$. We next parameterize the bounds in (73a), (73b), and (73c) as

$$\begin{aligned}R'_{1c} &= I_0, & I_0 &:= I(U_{1c}; X_2 | U_{2c}), \\ R'_{1pb} &= \beta I_2 + I_1, & I_1 &:= I(U_{1pb}; X_2 | U_{1c}, U_{2c}), \\ R'_{2pb} &= \beta I_2, & I_2 &:= I(U_{1pb}; U_{2pb} | X_2, U_{1c}, U_{2c}), \\ & & I_{01} &:= I_0 + I_1, \\ & & I_{012} &:= I_0 + I_1 + I_2\end{aligned}$$

so that in (74) we can substitute

$$\begin{aligned}\widehat{R}_1 &= R_1 + I_{012} - \beta I_2, \\ \widehat{R}_2 &= R_2 + \beta I_2.\end{aligned}$$

By eliminating β from (74) we obtain

$$R_1 \leq \min\{d_1, a_1 + d_2\} - I_{01}, \quad (75a)$$

$$R_2 \leq g_2, \quad (75b)$$

$$R_1 + R_2 \leq \min\{g_1 + a_2, g_2 + a_1\} - I_{012}, \quad (75c)$$

$$2R_1 + R_2 \leq g_1 + a_1 + d_2 - I_{012} - I_{01} \quad (75d)$$

together with two redundant sum-rate constraints

$$\begin{aligned}R_1 + R_2 &\leq \min\{d_1, a_1 + d_2\} + g_2 - I_{012}, \\ R_1 + R_2 &\leq \frac{g_1 + a_1 + d_2 + g_2}{2} - I_{012}.\end{aligned}$$

The region in (75) is the same as the region in (19).

APPENDIX G

PROOF OF COROLLARY VI.4

In the following we use the fact that point B in (43) is to within one bits/s/Hz and a factor two from point C in (44). This is the case as, for the additive gap, $R_1^{(B)} = R_1^{(C)}$ and

$$\begin{aligned}R_2^{(B)} - R_2^{(C)} &= \mathcal{C} \left(\frac{2\sqrt{|b|^2 P_1 P_2}}{1 + |b|^2 P_1 + P_2} \right) \\ &\leq \mathcal{C} \left(\sqrt{\frac{P_2}{P_2 + 1}} \right) \leq \log(2) = 1\end{aligned} \quad (76)$$

where we use the fact that $R_2^{(B)} - R_2^{(C)}$ has a maximum in $|b|^2 P_1 = P_2 + 1$.

A representation of the ‘‘strong interference’’ outer bound and the ‘‘piecewise linear strong interference’’ outer bound is shown in Fig. 21. The ‘‘strong interference’’ outer bound coincides with the ‘‘piecewise linear strong interference’’ outer bound at point A and the largest distance between the two outer bounds is attained between points B and C . This figure also introduces a new corner point of the inner bound: point D , the inner bound point with the largest R_2 rate when $R_1 = \mathcal{C}(P_1) - \Delta_1$ where Δ_1 will be defined later on (see next page).

1) *Perfect Interference Cancellation:* In the proof of Theorem V.1 we have seen that under condition (30a) it is possible to achieve point C in (44) with scheme (E) with Costa’s DPC. This result may be used to show achievability of the ‘‘strong interference’’ outer bound to within half a bit/s/Hz per real dimension.

Theorem G.1: If condition in (30a) holds, the ‘‘strong interference’’ outer bound of Theorem II.2 is achievable to within half a bit/s/Hz per real dimension.

Proof: Under the condition in (30a), point C is achievable. This point lies to within half a bit/s/Hz per real dimension from the outer bound. ■

2) *Non-Perfect Interference Cancellation:* Although it is not possible to achieve point C using scheme (E) and perfect interference cancellation, it is possible to achieve this point to within a bounded distance using non perfect interference cancellation in the strong interference ($|b| > 1$) and strong signal ($P_2 \geq |b|^2 P_1$) regimes.

Theorem G.2: When $|b| > 1$ and $P_2 \geq |b|^2 P_1$, the outer bound of Theorem II.2 may be achieved to within 1.87 bits/s/Hz per real dimension.

Proof: To prove this theorem we show the achievability of point D in Fig. 21 which lies at a bounded distance from point C using scheme (E) in (29) for $\alpha = 1$. Fig. 21 shows the different additive gaps between inner and outer bound points in the following proof. If (29a) is tight there are two possible scenarios: the corner point D is determined by 1) the intersection between (29c) and (29a) or by 2) the intersection of (29b) and (29a). We choose λ so that both (29a) and (29b) lie within a finite distance from $R_1^{(B)}$ and $R_2^{(B)}$ respectively. The sum rate bound (29c) does not depend on the choice of λ . We divide the

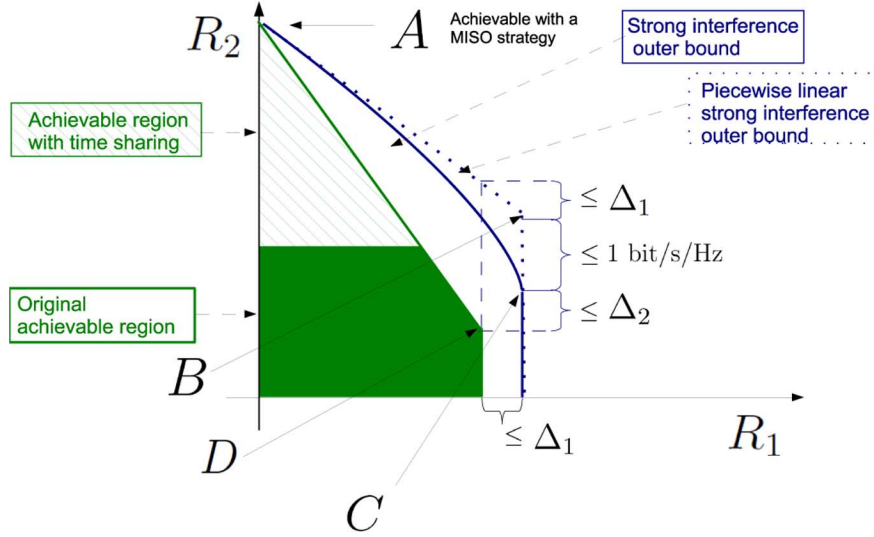


Fig. 21. Graphical representation of the relationship between inner and outer bound for Theorem G.2, G.5 and G.5.

proof in two subcases $\Re\{a\} \geq |b|^{-1}$.

Sub-case $\Re\{a\} \leq |b|^{-1}$: When $P_1 \leq 1$ a gap of 1 bit per dimension is achievable by having both transmitters transmit to receiver 2 at rate $R_2^{(C)}$, which we recall is given in (44). In this case the distance along the rate R_2 is zero and on the rate R_1 is $R_1^{(C)} - 0 \leq \log(1+1) < 2$. For $P_1 \geq 1$ let $\lambda = \frac{P_1 - \sqrt{P_1}}{P_1 + 1}a$, in (29). The distance between inner and outer bound for R_1 is

$$\Delta_1 := R_1^{(C)} - R_1^{(D)} = \log \left(\frac{1 + P_1 + 2|a|^2 P_2}{1 + P_1 + |a|^2 P_2} \right) \leq 1$$

where we have used the inequality $P_2 \geq |b|^2 P_1$. Similarly letting (29b) hold with equality, we obtain

$$\begin{aligned} \Delta_2 &:= R_2^{(C)} - R_2^{(D)} \\ &\leq \max_{a: \Re\{a\} \leq |b|^{-1}} \log \left(\frac{\frac{1+2P_2}{1+P_1}}{1 + P_2 \left| 1 - \frac{(P_1 - \sqrt{P_1})a|b|}{1+P_1} \right|^2} \right) \\ &\leq \log \left(\frac{(1+P_1)(1+2P_2)}{(1+P_1)(1+P_2+P_1) + 2P_2\sqrt{P_1}} \right) \\ &\leq \log \left(\frac{1+2P_2}{1+P_2+P_1} \right) \leq 1 \end{aligned}$$

where we have used that the expression has a global maximum in $a^* > \frac{1}{|b|}$. The largest gap between the inner bound and point B is thus bounded by $\max\{1 + \Delta_1, \Delta_1 + \Delta_2\} = 2$, and so the overall gap between the specified achievable scheme of (29) and the outer bound is within $1 + 2 = 3$ bits/s/Hz for a complex valued channel.

Subcase $\Re\{a\} > |b|^{-1}$: When $P_1 \leq 3$ a gap of 1 bit per dimension is achievable by having transmitter 1 remain silent (rate $R_1 = 0$) since in this case $R_1^{(B)} - 0 \leq \log(3+1)$. When $P_1 > 3$ let $\lambda = \frac{P_1 + 2\sqrt{P_1}}{P_1 + 1}$ in (29). The gap for R_1 may be bounded as

$$\Delta_1 := R_1^{(C)} - R_1^{(D)} = \log \left(\frac{1 + P_1 + 5|a|^2 P_2}{1 + P_1 + |a|^2 P_2} \right) \leq \log(5)$$

while that for the rate R_2 of transmitter 2 may be bounded as

$$\Delta_2 := R_2^{(C)} - R_2^{(D)} \quad (77a)$$

$$\leq \max_{a: \Re\{a\} \leq |b|^{-1}} \log \left(\frac{1 + 2P_2}{(1 + P_1) \left(P_2 \left| 1 - \frac{P_1 + 2a|b|\sqrt{P_1}}{1+P_1} \right|^2 + 1 \right)} \right) \quad (77b)$$

$$\sim \log \left(\frac{(1 + P_1)(1 + 2P_2)}{P_2 - 4P_2\sqrt{P_1} + 4P_2P_1 + (1 + P_1)^2} \right) \quad (77c)$$

$$\leq \log \left(\frac{(1 + P_1)(1 + 2P_2)}{2P_1P_2 + (1 + P_1)^2} \right) \quad (77d)$$

$$\leq \log \left(\frac{P_1 + 1}{P_1} \right) \quad (77e)$$

$$\leq \log \left(\frac{4}{3} \right) \quad (77f)$$

where (77c) follows since the expression has a global maximum for $a^{(opt)} < \frac{1}{|b|}$ and (77d) follows since $4P_1 - 4\sqrt{P_1} + 1 > 2P_1$ for $P_1 > 3$. Finally (77e) and (77f) follow since the expression is monotonically increasing in P_2 and decreasing in P_1 . As in the subcase $\Re\{a\} \leq |b|^{-1}$, the maximum distance between points C and D is bounded by $\max\{1 + \Delta_1, \Delta_1 + \Delta_2\} \leq \log\left(\frac{20}{3}\right)$ so that the overall gap is bounded by $\log\left(\frac{40}{3}\right) \approx 3.74$ bits/s/Hz for a complex valued channel. ■

3) *Cognitive Broadcasting:* The outer bound Thm II.1 is achievable in “weak interference”: the capacity achieving scheme in this regime is scheme (B) in Section IV-B and it employs Costa’s DPC at the cognitive transmitter to “pre-cancel” the known interference generated by the primary user. While capacity is known in this regime, we show that the very simple broadcast strategy of scheme (A) in Section IV-A achieves capacity to within a constant gap from the outer bound when the INR is larger than the SNR at the primary receiver (i.e., $|b|^2 P_1 > P_2$). When the INR is larger than the SNR at the primary receiver, scheme (A) achieves a constant gap from the outer bound in “strong interference” as well. Although the

resulting gap does not improve on the result of Theorem VI.1, this result suggests that, in a general scheme, rate improvement may be obtained by having the cognitive transmitter send part of the primary message.

Theorem G.3: When $|b| < 1$ and $|b|^2 P_1 \geq P_2$, the outer bound of Theorem II.1 may be achieved within 1 bit/s/Hz per real dimension.

Proof: Consider the scheme (A) in Section IV-A for $|b| \leq 1$. With $P_2 = 0$, the channel reduces to a degraded BC with input X_1 [43] and Y_2 is a degraded version of Y_1 , therefore

$$R_1 \leq I(Y_1; U_{1pb}) - I(U_{1pb}; U_{2pb}) \stackrel{[43]}{=} \mathcal{C}(\alpha P_1), \quad (78a)$$

$$R_2 \leq I(Y_2; U_{2pb}) \stackrel{[43]}{=} \mathcal{C}\left(\frac{\bar{\alpha}|b|^2 P_1}{1 + \alpha|b|^2 P_1}\right) \quad (78b)$$

taken over the union over of all $\alpha \in [0, 1]$, is achievable. Then since (3a) and (78a) are the same for every α there is zero gap for the rate R_1 . By considering the difference between (3b) and (78b), the gap for the rate R_2 is bounded as

$$\begin{aligned} (3b)-(78b) &\leq \mathcal{C}\left(|b|^2 P_1 + P_2 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2}\right) - \mathcal{C}\left(|b|^2 P_1\right) \\ &\leq \mathcal{C}\left(\frac{P_2 + 2\sqrt{|b|^2 P_1 P_2}}{1 + |b|^2 P_1}\right) \\ &\leq \mathcal{C}\left(\frac{3|b|^2 P_1}{1 + |b|^2 P_1}\right) \\ &\leq \log(4) = 2. \end{aligned}$$

Theorem G.4: When $|b| > 1$ and $|b|^2 P_1 \geq P_2$, the outer bound of Theorem II.2 may be achieved within 1.5 bits/s/Hz per real dimension.

Proof: Consider scheme (A) in Section IV-A for $|b| > 1$ and $\alpha = \min\{1, 1/P_1\}$ in (20). Then the gap for user 1 is

$$\Delta_1 := R_1^{(B)} - R_1^{(C)} = \mathcal{C}(\min\{1, P_1\}) \leq \log(2) = 1$$

while the gap for user 2 (using $P_2 \leq |b|^2 P_1$ and $|b|^2 \geq 1$) is

$$\begin{aligned} \Delta_2 := R_2^{(B)} - R_2^{(C)} &\leq \mathcal{C}\left(\frac{1 + 2|b|^2 P_1}{(1 + P_1)(1 + |b|^2 \min\{1, P_1\})}\right) \\ &\leq \max\left\{\log\left(\frac{2|b|^2}{1 + |b|^2}\right), \log\left(\frac{2}{1 + P_1}\right)\right\} \\ &\leq \log(2) = 1. \end{aligned}$$

As shown in Fig. 21, the achievable point C in (44) is at most at $1 + \Delta_1 + \Delta_2 \leq 3$ bits from the outer bound. By time sharing between points A and C, we have an achievable rate region that is at most at $\max\{1, 3\} = 3$ bits/Hz/s from the outer bound for complex valued channels. ■

4) *Interference Stripping:* With interference stripping we have:

Theorem G.5: When $|a| \geq 1$, $|b| \geq 1$ and $|b|^2 P_1 \leq P_2$, the outer bound of Theorem II.2 may be achieved within 1.5 bits/s/Hz per real dimension.

Proof: We consider scheme (D)'s performance in the "strong interference" regime when $|b|^2 > 1$, $|a|^2 \geq 1$. When we set $\alpha = 1$, it achieves the rate

$$\begin{aligned} R_1 &\leq \mathcal{C}(P_1) \\ R_1 + R_2 &\leq \mathcal{C}(\min\{|a|^2 P_2 + P_1, P_2 + |b|^2 P_1\}). \end{aligned}$$

Referring again to Fig. 21, the gap between points B and C may be bounded as

$$\begin{aligned} \Delta_1 := R_1^{(B)} - R_1^{(C)} &\leq \log(2) = 1, \\ \Delta_2 := R_2^{(B)} - R_2^{(C)} &\leq \mathcal{C}\left(\frac{1 + |b|^2 P_1 + P_2}{1 + \min\{|a|^2 P_2 + P_1, P_2 + |b|^2 P_1\}}\right) \\ &\leq \mathcal{C}\left(\max\left\{1, \frac{1 + |b|^2 P_1 + P_2}{1 + |a|^2 P_2 + P_1}\right\}\right) \\ &\leq \mathcal{C}\left(\max\left\{1, \frac{1 + 2P_2}{1 + |a|^2 P_2 + P_1}\right\}\right) \\ &\leq \mathcal{C}\left(\max\left\{1, \frac{1 + 2P_2}{1 + |a|^2 P_2 + P_1}\right\}\right) \\ &\leq \log(2) = 1. \end{aligned}$$

We thus achieve a rate pair that lies within $1 + \Delta_1 + \Delta_2 = 3$ bits/s/Hz of the outer bound for complex valued channel. ■

ACKNOWLEDGMENT

The authors would like to thank Prof. Shlomo Shamai and Prof. Gerhard Kramer for insightful discussions on the problem during IZS 2010 and ITA 2010, respectively.

REFERENCES

- [1] S. Rini, D. Tuninetti, and N. Devroye, "State of the cognitive interference channel: A new unified inner bound," in *Proc. Int. Zurich Seminar Commun.*, Mar. 2010.
- [2] S. Rini, D. Tuninetti, and N. Devroye, "New results on the capacity of the Gaussian cognitive interference channel," in *Proc. Allerton Conf. Commun., Control and Comp.*, Sep. 2010.
- [3] S. Rini, D. Tuninetti, and N. Devroye, "The capacity of the semi-deterministic cognitive interference channel and its application to constant gap results for the Gaussian channel," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2011.
- [4] 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.
- [5] I. Maric, R. Yates, and G. Kramer, "The strong interference channel with unidirectional cooperation," in *Proc. UCSD Workshop on Information Theory and Applications*, Feb. 2006.
- [6] 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.
- [7] S. Rini, D. Tuninetti, and N. Devroye, "The capacity region of Gaussian cognitive radio channels to within 1.87 bits," in *Proc. IEEE Inf. Theory Workshop*, Jan. 2010.
- [8] H. Sato, "Two-user communication channels," *IEEE Trans. Inf. Theory*, vol. 23, no. 3, pp. 295–304, 1977.
- [9] E. Yang and D. Tuninetti, "Interference channel with generalized feedback (a.k.a. with source cooperation)—Part I: Achievable region," *IEEE Trans. Inf. Theory, Special Issue on Interference Networks*, vol. 57, no. 5, pp. 2686–2710, May 2011.
- [10] T. Cover, "Comments on broadcast channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2524–2530, Oct. 1998.

- [11] A. Jovicic and P. Viswanath, "Cognitive radio: An information-theoretic perspective," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3945–3958, Aug. 2009.
- [12] I. Maric, R. 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.
- [13] 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, July 2011.
- [14] I. Maric, A. Goldsmith, G. Kramer, and S. Shamai, "On the capacity of interference channels with one cooperating transmitter," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 405–420, Apr. 2008.
- [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. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley-Interscience, 1991.
- [17] H. Sato, "An outer bound to the capacity region of broadcast channels (corresp.)," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 374–377, 1978.
- [18] J. Jiang and Y. Xin, "A new achievable rate region for the cognitive radio channel," in *Proc. IEEE Int. Conf. Commun.*, 2008, pp. 1055–1059.
- [19] Y. Cao and B. Chen, "Interference Channels with one cognitive transmitter," in *Proc. Asilomar Conference on Signals, Systems and Computers*, Nov. 2008, pp. 1593–1597.
- [20] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.
- [21] J. Jiang, I. Maric, A. Goldsmith, and S. Cui, "Achievable rate regions for broadcast channels with cognitive relays," in *Proc. IEEE Inf. Theory Workshop*, 2009.
- [22] G. Hodtani and M. Aref, "On the Devroye-Mitran-Tarokh rate region for the cognitive radio channel," *IEEE Trans. Wireless Comm.*, vol. 8, no. 7, pp. 3458–3461, Jul. 2009.
- [23] N. Devroye, "Information Theoretic Limits of Cognition and Cooperation in Wireless Networks," Ph.D. dissertation, Harvard University, Cambridge, MA, 2007.
- [24] S. Rini, D. Tuninetti, and N. Devroye, "The capacity region of Gaussian cognitive radio channels at high SNR," in *Proc. IEEE Inf. Theory Workshop*, Nov. 2009.
- [25] A. Avestimehr, S. Diggavi, and D. Tse, "A deterministic approach to wireless relay networks," in *Proc. Allerton Conf. Commun., Control and Comp.*, Monticello, IL, Sep. 2007.
- [26] N. Liu, I. Maric, A. Goldsmith, and S. Shamai, "The capacity region of the cognitive z-interference channel with one noiseless component," Arxiv Preprint arXiv:0812.0617 2008.
- [27] S. Seyedmehdi, Y. Xin, and Y. Lian, "An achievable rate region for the causal cognitive radio," in *Proc. Allerton Conf. Communications Control and Computer*, 2007.
- [28] S. Rini, D. Tuninetti, and N. Devroye, "A new capacity result for the Z-Gaussian cognitive interference channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Aug. 2011.
- [29] J. Jiang, I. Maric, A. Goldsmith, S. Shamai, and S. Cui, "On the capacity of a class of cognitive Z-Interference channels," in *Proc. IEEE Int. Conf. Communication*, Jul. 2011.
- [30] M. Vaezi and M. Vu, "Superposition coding-based bounds and capacity for the cognitive z-interference channels," Arxiv Preprint arXiv:1101.1920 2011.
- [31] A. El Gamal, "The capacity of a class of broadcast channels," *IEEE Trans. Inf. Theory*, vol. 25, no. 2, pp. 166–169, 1979.
- [32] H. Sato, "The capacity of Gaussian interference channel under strong interference," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 6, pp. 786–788, Nov. 1981.
- [33] A. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 1, pp. 60–70, Jan. 1978.
- [34] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 3, pp. 581–586, 2004.
- [35] G. Kramer, *Topics in Multi-User Information Theory ser. Foundations and Trends in Communications and Information Theory*, vol. 4, no. 45, pp. 265–444, 2008.
- [36] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Channels*. Budapest, Hungary: Akademiai Kiado, 1981.
- [37] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, 2006.
- [38] P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inf. Theory*, vol. 19, no. 2, pp. 197–207, 1973.
- [39] H. Weingarten, Y. Steinberg, and S. Shamai, "On the capacity region of the multi-antenna broadcast channel with common messages," in *Proc. IEEE Int. Symp. Inf. Theory*, 2006, pp. 2195–2199.
- [40] J. Korner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 23, no. 1, pp. 60–64, Jan. 1977.
- [41] 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.
- [42] S. Gel'fand and M. Pinsker, "Coding for channel with random parameters," *Probl. Contr. Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.
- [43] R. Gallager, "Capacity and coding for degraded broadcast channels," *Probl. Transm. Inf.*, vol. 10, no. 3, pp. 3–14, 1974.
- [44] M. Costa, "Writing on dirty paper (corresp.)," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, 1983.
- [45] A. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow," in *Proc. Allerton Conf. Commun., Control and Comp.*, Sep. 2007.
- [46] T. Liu and P. Viswanath, "An extremal inequality motivated by multi-terminal information-theoretic problems," *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1839–1851, 2007.
- [47] K. Kobayashi and T. Han, "A further consideration on the HK and the CMG regions for the interference channel," in *Proc. ITA Workshop*, 2007.

Stefano Rini received the B.A. degree in computer science from the Politecnico di Milano, Italy, in 2005, the M.A. degree in statistics from the University of Illinois at Chicago (UIC) in 2008 and the Ph.D. degree in electrical and computer engineering from UIC in 2010.

He was a postdoctoral fellow with the Department of Electrical Engineering at Stanford University, Stanford, during the fall and summer of 2011 and he is currently a postdoctoral fellow at the Institute for Communication Engineering at the Technische Universität München, Munich, Germany.

Daniela Tuninetti received the M.S. degree in telecommunication engineering from the Politecnico di Torino, Turin, Italy, in 1998 and the Ph.D. in electrical engineering from ENST/Telecom ParisTech, Paris, France (work done at the Eurecom Institute in Sophia Antipolis, France) in 2002.

Since January 2005, she is with the Department of Electrical and Computer Engineering at the University of Illinois at Chicago, where she is currently an Associate Professor. Her research interests are in the ultimate performance limits of wireless interference networks, with special emphasis on cognition and user cooperation.

Dr. Tuninetti was the editor-in-chief of the IEEE Information Theory Society Newsletter from 2006 to 2008 and an associate editor for the IEEE COMMUNICATION LETTERS from 2006 to 2009. She currently is an editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. She regularly serves on the Technical Program Committee 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.

Natasha Devroye received the a Honors B.Eng. degree in electrical engineering from McGill University, Montreal, QC, Canada, in 2001, the M.Sc. degree from Harvard University, Cambridge, MA, in 2003, and the Ph.D. degree in engineering sciences from the School of Engineering and Applied Sciences at Harvard in 2007.

She has been an Assistant Professor in the Department of Electrical and Computer Engineering, University of Illinois at Chicago, since January 2009. From July 2007 until July 2008, she was a Lecturer at Harvard. She has worked as a research intern at Intel Corporation, Santa Clara, CA, as well as with Mitsubishi Electric Research Labs, in Cambridge. Her research focuses on multi-user information theory and applications to cognitive and software-defined radio, radar, two-way and wireless communications in general.

Dr. Devroye was a recipient of an NSF CAREER award in 2011.