The Two-user Causal Cognitive Interference Channel: Novel Outer Bounds and Constant Gap Result for the Symmetric Gaussian Noise Channel in Weak Interference

Martina Cardone, Daniela Tuninetti and Raymond Knopp

Abstract

This paper studies the two-user Causal Cognitive Interference Channel (CCIC), where two transmitters aim to communicate independent messages to two different receivers via a common channel. One source, referred to as the *cognitive*, is *capable* of overhearing the other source, referred to as the *primary*, through a noisy in-band link and thus can assist in sending the primary's data. Two novel outer bounds of the type $2R_\mathsf{p} + R_\mathsf{c}$ and $R_\mathsf{p} + 2R_\mathsf{c}$ are derived for the class of injective semi-deterministic CCICs where the noises at the different source-destination pairs are independent. An achievable rate region is derived based on Gelfand-Pinsker binning, superposition coding and simultaneous decoding at the receivers.

The lower and outer bounds are then specialized to the practically relevant Gaussian noise case. The authors of this paper recently characterized to within a constant gap the capacity of the symmetric Gaussian CCIC in (a) the strong interference regime, and (b) for a subset of the weak interference regime when the cooperation link is larger than a given threshold. This work characterizes to within

a constant gap the capacity for the symmetric Gaussian CCIC in the regime that was still open. In particular, it is shown that the novel outer bounds are necessary to characterize the capacity to within a constant gap when the cooperation link is weaker than the direct links, that is, in this regime unilateral cooperation leaves some system resources underutilized.

Index Terms

Achievable rate region, causal cooperation, channel capacity, cognitive radio, constant gap, interference channel, outer bound region, unilateral source cooperation.

I. INTRODUCTION

This work considers the two-user Causal Cognitive Interference Channel (CCIC), a wireless network where one *primary* source PTx and one *cognitive / capable* source CTx aim to reliably communicate with two different receivers, namely the PRx and the CRx, via a common channel. Differently from the classical non-cooperative IC, in the CCIC the CTx operates in full-duplex mode and is able to overhear the PTx through a noisy in-band link; the CTx can thus exploit this side information to boost the performance of the two (primary and cognitive) systems.

The major feature of the CCIC is the concept of *causal cognition / source cooperation*, which represents both an interference management tool and a practical model for the cognitive radio technology. Actually, unilateral source cooperation offers a way to smartly cope with interference. In today's wireless systems, the general approach to deal with interference is either to avoid it, by trying to orthogonalize (in time / frequency / space) users' transmission, or to treat it as noise. However, these approaches may severely limit the system capacity since perfect user orthogonalization is not possible in practice. In contrast, in the CCIC the CTx, which can causally learn the primary's data through a noisy link, may protect both its own (by precoding against some known interference) and the primary's (by allocating some of its transmission resources to assist the PTx to convey data to the PRx) information from interference. Thus, the transmission techniques designed for the CCIC aim to leverage the structure of the interference, instead of just simply disregarding it and treating it as noise. The CCIC also represents a more practically relevant model for the cognitive *overlay* paradigm, compared to the case where the CTx is assumed to a priori (before the transmission begins) know the message of the PTx [3], which may be granted only in limited scenarios. In contrast, in the CCIC the CTx causally learns the PTx's data through a noisy link. Thus, the transmission techniques designed for the CCIC

account for the time the CTx needs for decoding and for the (possible) further rate losses that may incur in decoding the PTx's message through a noisy link of limited capacity.

A. Related past work

The CCIC studied in this work models an IC with unilateral source cooperation and represents a practical scenario for cognitive radios, where one source has superior capabilities with respect to the other source. Moreover, closely related to the IC with unilateral source cooperation is the IC with perfect output feedback model, where the received signal is fed back through a perfect channel from one receiver to the corresponding transmitter. Lately, these scenarios have received significant attention, as summarized next.

1) IC with source cooperation: The CCIC is a particular case of the IC with bilateral source cooperation, where one of the cooperation link (in our case the one from the CTx to the PTx) is absent. For the IC with bilateral source cooperation several outer bounds on the capacity region were derived in [4], [5], [6], [7] and a number of transmission strategies were designed in [8].

In [4], the author firstly derived inner and outer bounds on the capacity for the IC with bilateral source cooperation and for the IC with bilateral destination cooperation. The outer bound region was obtained from the max-flow-min-cut theorem, by further strengthening the sum-rate bound $R_{\mathsf{p}} + R_{\mathsf{c}}$ (where R_{p}, respectively R_{c}, is the transmission rate for the PRx, respectively the CRx) as proposed in [9], while the lower bound region was derived by designing a scheme based on Gelfand-Pinsker's binning [10] (i.e., dirty-paper-coding in Gaussian noise [11]) and superposition encoding, decode-and-forward relaying and simultaneous decoding.

In [5], the author derived a novel general outer bound for the IC with bilateral source cooperation that applies to any channel, i.e., the noises can be arbitrarily correlated. The outer bound rate region in [5, Theorem II.1] has constraints on the single rates R_{p} and R_{c} and on the sum-rate $R_{\mathsf{p}} + R_{\mathsf{c}}$. In particular, the constraints on the sum-rate were obtained by extending the idea of Kramer in [12, Theorem 1], beyond the Gaussian noise case. Moreover, the outer bounds in [5, Theorem II.1] were recently shown to hold for the case of bilateral destination cooperation as well [13]. When evaluated for the symmetric Gaussian noise case, the derived region is achievable to within 2 bits in the strong cooperation regime [14].

In [6], the authors derived a novel sum-rate outer bound for a class of Injective Semi-Deterministic (ISD) ICs with bilateral source cooperation and with independent noises at all

terminals; when evaluated for the Gaussian noise case with symmetric cooperation links, the derived region is achievable to within 19 bits.

In [7], the authors derived a novel outer bound on the capacity region of the Gaussian IC with bilateral source cooperation that are based on the idea of 'dependence balance' proposed in [15]. In [7] it was proved that the novel bound region may be tighter than the cut-set bound.

In [8] two transmission strategies for the IC with bilateral source cooperation were designed. The two schemes employ partial-decode-and-forward relaying, rate splitting and simultaneous decoding. While the first strategy (see [8, Section IV]) uses only superposition coding, the second scheme (see [8, Section V]) also employs Gelfand-Pinsker's binning in the encoding phase.

The CCIC was specifically studied in [16], where novel outer and inner bounds were derived. Although, to the best of our knowledge, the work in [16] gives the best known bounds for this channel, their evaluation is not straightforward. For example, the outer bound involves several auxiliary random variables that are jointly distributed with the inputs and for which no cardinality bounds on the corresponding alphabets are known; the inner bound is characterized by many constraints and auxiliary random variables whose optimization is not immediate.

In [17], the capacity of the Gaussian CCIC (GCCIC) was characterized to within a constant gap for a set of channel parameters that, roughly speaking, excludes the case of weak interference at both receivers. For the symmetric case (i.e., the two direct links and the two interfering links are of the same strength) a constant gap result of 1 bit was proved in strong interference and in weak interference when the cooperation link is 'sufficiently strong'. Moreover, in [17], the capacity of the GCCIC was characterized to within 2 bits for two special cases: the Z-channel, where the CRx does not experience interference from the PTx, and the S-channel, where the PRx does not experience interference from the CTx. These constant gap results were obtained by using known outer bounds on the single rates R_p and R_c and on the sum-rate $R_p + R_c$.

In [17] we pointed out that in weak interference the capacity region of the GCCIC should also have outer bounds of the type $2R_p + R_c$ and $R_p + 2R_c$, similarly to the classical non-cooperative IC [18]. To the best of our knowledge, these bounds are not available in the existing literature of cooperative ICs and their derivation represents the main contribution of this work. These novel bounds are proved to be active for the symmetric GCCIC in weak interference and when the cooperation link is weaker than the direct link, thus closing a problem left open in [17].

2) IC with output feedback: In [19], the authors studied the Gaussian IC where each source has a perfect output feedback from the intended destination and characterized the capacity to

within 2 bits. In [19, Theorems 2-3], the capacity region has constraints on the single rates R_{p} and R_{c} and on the sum-rate $R_{\mathsf{p}} + R_{\mathsf{c}}$, but not bounds of the type $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$, which appear in the capacity region of the classical Gaussian IC in weak interference [18]. The authors interpreted the bounds on $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$ in the capacity region of the classical IC as a measure of the amount of 'resource holes', or system under utilizations, due to the distributed nature of the non-cooperative IC. The authors thus concluded that output feedback eliminates these 'resource holes', i.e., the system resources are fully utilized.

In [20], the symmetric Gaussian IC with all 9 possible output feedback configurations was analyzed. The authors proved that the bounds derived in [19] suffice to approximately characterize the capacity of all the 9 configurations except for the case where only one source receives feedback from the corresponding destination, i.e., the 'single direct-link feedback model', or the feedback-model-(1000). For the feedback-model-(1000), it was shown that an outer bound of the type $2R_{\mathsf{p}} + R_{\mathsf{c}}$ is needed to capture the fact that the second source (whose transmission rate is R_{c}) does not receive feedback. In the language of [19] we thus have that a 'single direct-link feedback' does not suffice to cover all the 'resource holes' whose presence is captured by the bound on $2R_{\mathsf{p}} + R_{\mathsf{c}}$. Such a bound was shown to be active for the feedback-model-(1000) in the Gaussian noise case.

In [21], the authors characterized the capacity of the two-user symmetric linear deterministic approximation of the Gaussian IC with bilateral noisy feedback, i.e., where only some of the received bits are fed back to the corresponding transmitter. In [22], the same authors evaluated the bounds for the symmetric Gaussian noise channel and proved that they are at most 11.7 bits far from one another, universally over all channel parameters. The capacity characterization was accomplished by proving novel outer bounds on $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$ that rely on carefully chosen side information random variables tailored to the symmetric Gaussian setting and whose generalization to non-symmetric or non-Gaussian scenarios does not appear straightforward.

In this work we first provide a general framework to derive outer bounds of the type $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$ for the ISD CCIC, which recover as special cases those derived in [20], [21]. We then show that these novel outer bounds are active for the symmetric GCCIC in weak interference when the cooperation link is weaker than the direct link, i.e., in this regime unilateral cooperation does not enable sufficient coordination among the sources for full utilization of the channel resources.

3) Non-causal CIC: In the original cognitive radio *overlay* paradigm, first introduced in [3], the superior capabilities of the CTx were modeled by assuming that the CTx a priori knows the PTx's message. The largest known achievable rate region for the general memoryless non-causal CIC is [23, Theorem 7], which in [24] was evaluated for the Gaussian noise case and shown to be at most 1 bit apart (see [24, Theorem VI.1]) from an outer bound region, which is characterized by constraints on the single rates R_p and R_c and on the sum-rate $R_\mathsf{p} + R_\mathsf{c}$. In other words, the capacity region of the non-causal CIC does not have bounds on $2R_\mathsf{p} + R_\mathsf{c}$ and $R_\mathsf{p} + 2R_\mathsf{c}$, i.e., the assumption of full a priori knowledge of the PTx's message at the CTx allows to fully exploit the available system resources.

In [17], we removed the ideal non-causal message knowledge assumption by considering the CCIC and we identified the set of parameters where unilateral cooperation attains the ultimate performance limits (i.e., generalized Degrees of Freedom (gDoF)) predicted by the non-causal model. In this work we show that the capacity of the CCIC differs in general from that of the ideal non-causal CIC since its capacity region has bounds of the type $2R_\mathsf{p} + R_\mathsf{c}$ and $R_\mathsf{p} + 2R_\mathsf{c}$, i.e., in general removing the ideal full a priori message knowledge at the CTx leaves some 'resource holes' in the system.

B. Contributions

Our main contributions can be summarized as follows:

1) We develop a general framework to derive outer bounds of the type $2R_\mathsf{p} + R_\mathsf{c}$ and $R_\mathsf{p} + 2R_\mathsf{c}$ on the capacity of the general ISD CCIC when the noises at the different source-destination pairs are independent; this framework includes for example feedback from the intended destination. As a special case, we recover and strengthen the bounds derived in [6], [20], [21]. The key technical ingredient is the proof of two novel Markov chains.

2) We design a transmission strategy for the general memoryless CCIC and we derive its achievable rate region. The proposed scheme uses superposition and binning encoding, partial-decode-and-forward relaying and simultaneous decoding at the receivers. Since the CCIC shares common features with the classical non-cooperative IC [25], both *common* (decoded also at the non-intended receiver) and *private* (treated as noise at the non-intended receiver) messages are used. Moreover, we use both *cooperative* (first partially decoded, then re-encoded and finally forwarded by the CTx) and *non-cooperative* (sent without the help of the CTx) messages for the PTx, while the messages of the CTx are only *non-cooperative*.

3) We evaluate the outer bound and the achievable rate regions for the practically relevant Gaussian noise channel. We prove that for the symmetric case, i.e., when the two direct links and the two cross / interfering links are of the same strength, the achievable region is a constant (uniformly over all channel gains) number of bits apart from the outer bound region in the regimes that were left open in [17], i.e., weak interference and weak cooperation. This result, jointly with [17, Theorem 1], fully characterizes the capacity of the symmetric GCCIC to within a constant gap. Moreover, this result sheds light on the regimes where unilateral cooperation is too weak and leaves some system resources underutilized.

C. Paper organization

The rest of the paper is organized as follows. Section II describes the general memoryless CCIC, the ISD CCIC and the Gaussian CCIC, for which the concepts of capacity to within a constant gap and gDoF are defined. Section III is dedicated to the outer bounds. In Section III-A we first summarize known outer bounds and then generalize and tighten the sum-rate bound in [6, bound u_1] by proving two novel Markov chains for the involved random variables. Then, Section III-B derives outer bounds of the type $2R_\mathsf{p} + R_\mathsf{c}$ and $R_\mathsf{p} + 2R_\mathsf{c}$ for the ISD CCIC with independent noises at the different source-destination pairs. Section IV focuses on the Gaussian noise CCIC and shows that the two novel outer bounds suffice to prove a constant gap result for the symmetric case in weak interference when the cooperation link is not 'strong enough', i.e., the regime which was left open in [17]. In particular, in Section IV-A the outer bounds are evaluated for the GCCIC, in Section IV-B a novel transmission strategy is designed and its achievable rate region is derived and finally in Section IV-C the achievable rate region is shown to be a constant number of bits apart from the outer bound region. Section V concludes the paper. Some proofs may be found in Appendix.

Although this paper considers only the symmetric Gaussian noise CCIC, we believe that the results can be extended to the general non-symmetric case albeit with more tedious and involved computations, especially for the achievable region, than those reported here.

D. Notation

Throughout the paper we adopt the following notation convention. The subscript c (in sans serif font) is used for quantities related to the cognitive pair, while the subscript p (in sans serif font) for those related to the primary pair. The subscript f or F (in sans serif font) is used to refer
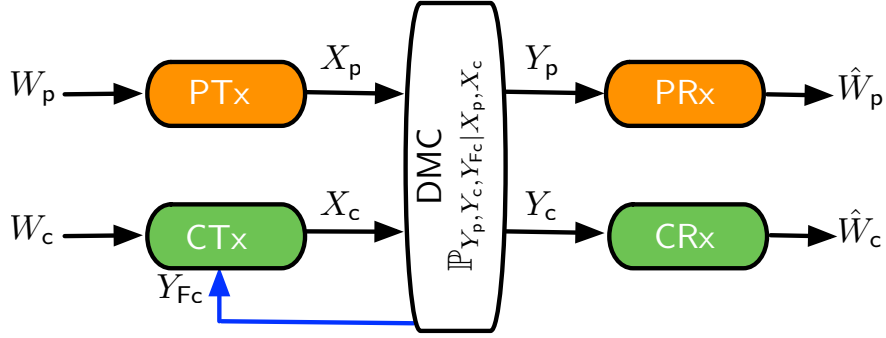
Fig. 1. The general memoryless CCIC.

to generalized feedback information received at the CTx. The subscript c (in roman font) is used to denote both common and cooperative messages, the subscript p (in roman font) to denote private messages and the subscript n (in roman font) to denote non-cooperative messages. With $[n_1 : n_2]$ we denote the set of integers from n_1 to $n_2 \geq n_1$ and $[x]^+ := \max\{0, x\}$ for $x \in \mathbb{R}$; Y^j is a vector of length j with components $(Y_1, \ldots, Y_j)$ and $\mathbb{E}[\cdot]$ indicates the expected value; a^* denotes the complex conjugate of a and $|a|$ is the absolute value of a; $\emptyset$ is the empty set. The notation eq(n) is used to indicate the rightmost side of the equation number n.

II. SYSTEM MODEL

A. The general memoryless channel

A *general memoryless CCIC*, shown in Fig. 1, consists of two input alphabets $(\mathcal{X}_\mathsf{p}, \mathcal{X}_\mathsf{c})$, three output alphabets $(\mathcal{Y}_\mathsf{Fc}, \mathcal{Y}_\mathsf{p}, \mathcal{Y}_\mathsf{c})$ and a memoryless transition probability $P_{Y_\mathsf{Fc}, Y_\mathsf{p}, Y_\mathsf{c} | X_\mathsf{p}, X_\mathsf{c}}$. The PTx has a message $W_\mathsf{p} \in [1 : 2^{NR_\mathsf{p}}]$ for the PRx and the CTx has a message $W_\mathsf{c} \in [1 : 2^{NR_\mathsf{c}}]$ for the CRx, where $N \in \mathbb{N}$ denotes the codeword length and $R_\mathsf{p} \in \mathbb{R}_+$ and $R_\mathsf{c} \in \mathbb{R}_+$ the transmission rates for the PTx and the CTx, respectively, in bits per channel use (logarithms are in base 2). The messages W_p and W_c are independent and uniformly distributed on their respective domains. At time i, $i \in [1 : N]$, the PTx maps its message W_p into a channel input symbol $X_{\mathsf{p}i}(W_\mathsf{p})$ and the CTx maps its message W_c and its past channel observations into a channel input symbol $X_{\mathsf{c}i}(W_\mathsf{c}, Y_\mathsf{Fc}^{i-1})$. At time N, the PRx outputs an estimate of its intended message based on all its channel observations as $\widehat{W}_\mathsf{p}(Y_\mathsf{p}^N)$, and similarly the CRx outputs $\widehat{W}_\mathsf{c}(Y_\mathsf{c}^N)$. The capacity region is
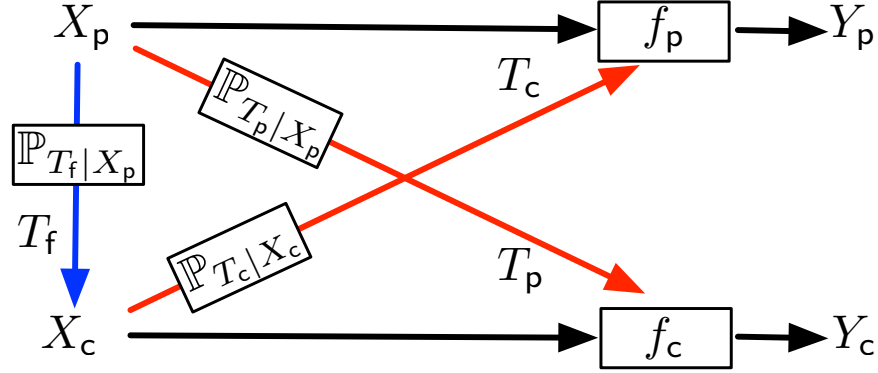
Fig. 2. The ISD CCIC.

the convex closure of all non-negative rate pairs $(R_\mathsf{p}, R_\mathsf{c})$ such that $\max_{u \in \{\mathsf{c}, \mathsf{p}\}} \mathbb{P}[\widehat{W}_u \neq W_u] \to 0$ as $N \to +\infty$.

B. The ISD channel

The *ISD* model, shown in Fig. 2 and first introduced in [26] for the classical IC, assumes that the input X_p, respectively X_c, before reaching the destinations, is passed through a memoryless channel to obtain T_p, respectively T_c. The channel outputs are therefore given by

$$Y_\mathsf{p} = f_\mathsf{p}\left(X_\mathsf{p}, T_\mathsf{c}\right), \tag{1a}$$

$$Y_\mathsf{c} = f_\mathsf{c}\left(X_\mathsf{c}, T_\mathsf{p}\right), \tag{1b}$$

where f_u, $u \in \{\mathsf{p}, \mathsf{c}\}$, is a deterministic function that is invertible given X_u, or in other words, T_p, respectively T_c, is a deterministic function of $(Y_\mathsf{c}, X_\mathsf{c})$, respectively $(Y_\mathsf{p}, X_\mathsf{p})$.

In the CCIC, the generalized feedback signal at the CTx satisfies

$$Y_\mathsf{Fc} = f_\mathsf{f}\left(X_\mathsf{c}, T_\mathsf{f}\right), \tag{1c}$$

for some deterministic function f_f that is invertible given X_c, i.e., T_f is a deterministic function of $(Y_\mathsf{Fc}, X_\mathsf{c})$, where T_f is obtained by passing X_p through a noisy channel [6].

In this work we assume that the noises seen by the different source-destination pairs are independent, that is,

$$\mathbb{P}_{Y_\mathsf{Fc}, Y_\mathsf{p}, Y_\mathsf{c} | X_\mathsf{p}, X_\mathsf{c}} = \mathbb{P}_{Y_\mathsf{p} | X_\mathsf{p}, X_\mathsf{c}} \mathbb{P}_{Y_\mathsf{Fc}, Y_\mathsf{c} | X_\mathsf{p}, X_\mathsf{c}}. \tag{2}$$
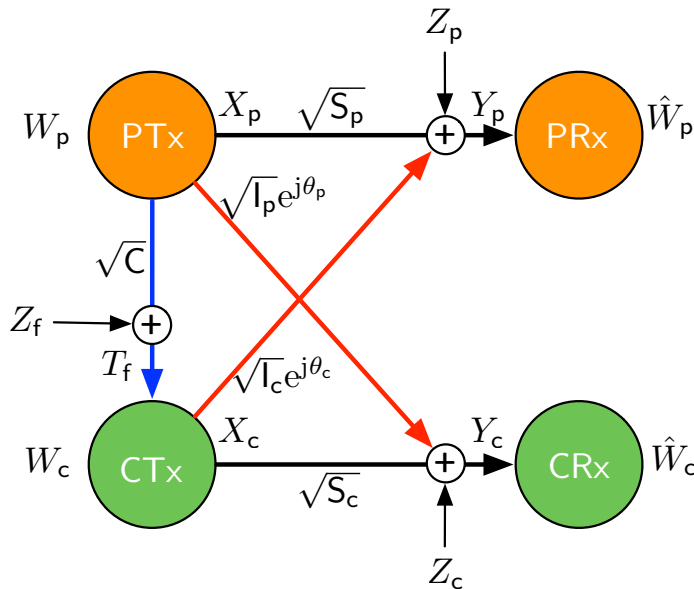
Fig. 3. The GCCIC.

In other words, we assume that the noises at the PRx and at the CRx are independent, but we do not impose any constraint on the noises at the CTx and CRx. For example, the case of output feedback in [20, model-(1000)] is obtained by setting $T_{\mathsf{f}} = T_{\mathsf{p}}$, which would not have been possible within the 'all noises are independent' setting studied in [6].

C. The Gaussian noise channel

A single-antenna GCCIC, shown in Fig. 3, is a special case of the ISD model described in (1) and it is defined by the input / output relationship

$$T_{\mathsf{p}} := \sqrt{\mathsf{I_p}}\mathrm{e}^{\mathrm{j}\theta_{\mathsf{p}}} X_{\mathsf{p}} + Z_{\mathsf{c}}, \tag{3a}$$

$$T_{\mathsf{c}} := \sqrt{\mathsf{I_c}}\mathrm{e}^{\mathrm{j}\theta_{\mathsf{c}}} X_{\mathsf{c}} + Z_{\mathsf{p}}, \tag{3b}$$

$$T_{\mathsf{f}} = \sqrt{\mathsf{C}} X_{\mathsf{p}} + Z_{\mathsf{f}}, \tag{3c}$$

$$Y_{\mathsf{p}} = \sqrt{\mathsf{S_p}} X_{\mathsf{p}} + T_{\mathsf{c}}, \tag{3d}$$

$$Y_{\mathsf{c}} = T_{\mathsf{p}} + \sqrt{\mathsf{S_c}} X_{\mathsf{c}}, \tag{3e}$$

$$Y_{\mathsf{Fc}} = T_{\mathsf{f}}, \tag{3f}$$

where $T_{\mathsf{f}} = Y_{\mathsf{Fc}}$ in (3f) is without loss of generality since the CTx can remove the contribution of its transmit signal X_{c} from its received signal Y_{Fc}. The channel gains are assumed to be

constant for the whole transmission duration, and hence known to all nodes. Without loss of generality, certain channel gains can be taken to be real-valued and non-negative since a node can compensate for the phase of one of its channel gains. The channel inputs are subject to a unitary average power constraint, i.e., $\mathbb{E}\left[|X_i|^2\right] \leq 1, i \in \{\mathsf{p}, \mathsf{c}\}$. This assumption is without loss of generality, since non-unitary power constraints can be incorporated into the channel gains. The noises are circularly symmetric Gaussian random variables with, without loss of generality, zero mean and unitary variance. We assume that the noise Z_p is independent of $(Z_\mathsf{c}, Z_\mathsf{f})$, while $(Z_\mathsf{c}, Z_\mathsf{f})$ can be arbitrarily correlated.

For the Gaussian noise case, it is customary to approximate the channel capacity as follows.

Definition 1 (Capacity region to within a constant gap) *The capacity region of the GCCIC is said to be known to within* GAP *bits if one can show an inner bound region* $\mathcal{I}$ *and an outer bound region* $\mathcal{O}$ *such that*

$$(R_\mathsf{p}, R_\mathsf{c}) \in \mathcal{O} \implies ([R_\mathsf{p} - \mathsf{GAP}]^+, [R_\mathsf{c} - \mathsf{GAP}]^+) \in \mathcal{I}.$$

For the two particular cases of $\mathsf{C} = 0$ (i.e., non-cooperative IC) and of $\mathsf{C} \to +\infty$ (non-causal CIC), the capacity is known to within 1 bit [18], [24].

The approximate (i.e., to within a constant gap) characterization of the capacity region implies the exact knowledge of its gDoF region. The gDoF metric, first introduced in [18] for the non-cooperative IC, captures the high-SNR behavior of the capacity as a function of the relative strengths of the direct, cooperation and interfering links. The gDoF represents a more refined characterization of the capacity in the high-SNR regime compared to the classical DoF since it captures the fact that, in wireless networks, the channel gains can differ by several orders of magnitude. In this work, we consider the symmetric case parameterized as

$$\mathsf{S}_\mathsf{p} = \mathsf{S}_\mathsf{c} := \mathsf{S}^1, \ \mathsf{S} \geq 0, \ \text{direct links}, \tag{4a}$$

$$\mathsf{I}_\mathsf{p} = \mathsf{I}_\mathsf{c} := \mathsf{S}^\alpha, \ \alpha \geq 0, \ \text{interfering links}, \tag{4b}$$

$$\mathsf{C} := \mathsf{S}^\beta, \ \beta \geq 0, \ \text{cooperation link}, \tag{4c}$$

where α measures the strength of the interference links compared to the direct link, while β the strength of the cooperation link compared to the direct link. Thus, the symmetric GCCIC is

parameterized by the triplet $(\mathsf{S}, \alpha, \beta)$, where S is referred to as the (direct link) SNR, α as the interference exponent and β as the cooperation exponent.[1]

Definition 2 (gDoF) *Given the parameterization in* (4)*, the gDoF is defined as*

$$\mathsf{d}(\alpha, \beta) := \lim_{\mathsf{S} \to +\infty} \frac{\max\{R_\mathsf{p} + R_\mathsf{c}\}}{2 \log(1 + \mathsf{S})}, \tag{5}$$

where the maximization is intended over all possible achievable rate pairs $(R_\mathsf{p}, R_\mathsf{c})$*.*

The gDoF of the classical IC ($\mathsf{C} = 0$) is the "W-curve" first characterized in [18] and given by $\mathsf{d}(\alpha, 0)$. The gDoF of the non-causal CIC ($\mathsf{C} \to \infty$) is the "V-curve", which can be evaluated from the capacity characterization to within 1 bit of [24], and is given by $\mathsf{d}(\alpha, \infty)$. An interesting question this work answers is whether there are values of $\beta > 0$ such that $\mathsf{d}(\alpha, \beta) = \mathsf{d}(\alpha, 0)$ — in which case unilateral causal cooperation is not helpful in terms of gDoF — or values of $\beta < \infty$ such that $\mathsf{d}(\alpha, \beta) = \mathsf{d}(\alpha, \infty)$ — in which case unilateral causal cooperation is equivalent to non-causal message knowledge in terms of gDoF.

Following the naming convention of the non-cooperative IC [18], we say that the symmetric GCCIC has strong interference if $\mathsf{S} \leq \mathsf{I}$, that is $1 \leq \alpha$, and weak interference otherwise. Similarly, we say that the symmetric GCCIC has strong cooperation if $\mathsf{S} \leq \mathsf{C}$, that is $1 \leq \beta$, and weak cooperation otherwise.

III. OUTER BOUNDS ON THE CAPACITY REGION FOR THE CCIC

This section is dedicated to the study of outer bounds on the capacity region for the CCIC. First, in Section III-A, some known outer bounds are summarized. Moreover, the outer bound originally derived in [6] for the ISD CCIC with independent noises at all terminals, is generalized to the case where only the noises at the different source-destination pairs are independent as in (2). Then, in Section III-B, two novel outer bounds of the type $R_\mathsf{p} + 2R_\mathsf{c}$ and $2R_\mathsf{p} + R_\mathsf{c}$ are derived for the ISD CCIC with independent noises at the different source-destination pairs as in (2). As we shall see in Section IV, these novel outer bounds allow to characterize the capacity to within a constant gap for the symmetric GCCIC in the regimes that were left open in [17].

---

[1]In principle the system performance also depends on the phases of the interfering links $(\theta_\mathsf{c}, \theta_\mathsf{p})$. However, as far as gDoF and capacity to within a constant gap are concerned, the phases $(\theta_\mathsf{c}, \theta_\mathsf{p})$ only matter if the IC channel matrix $\begin{bmatrix} \sqrt{\mathsf{S}_\mathsf{p}} & \sqrt{\mathsf{I}_\mathsf{c}}e^{j\theta_\mathsf{c}} \\ \sqrt{\mathsf{I}_\mathsf{p}}e^{j\theta_\mathsf{p}} & \sqrt{\mathsf{S}_\mathsf{c}} \end{bmatrix}$ is rank deficient [19], in which case one received signal is a noisier version of the other. In this work, we assume that the phases are such that the IC channel matrix is full rank as in [19].

A. Known outer bounds and some generalizations

In the literature, several outer bounds are known for the IC with bilateral source coopera-tion [5], [6], which we specialize here to the CCIC. In particular, for a joint input distribution $\mathbb{P}_{X_\mathsf{p},X_\mathsf{c}}$, we have:

1) For the *general memoryless CCIC*, described in Section II-A, the cut-set upper bound [27] gives

$$R_\mathsf{p} \leq I\left(X_\mathsf{p}; Y_\mathsf{p}, Y_\mathsf{Fc} | X_\mathsf{c}\right), \tag{6a}$$

$$R_\mathsf{p} \leq I\left(X_\mathsf{p}, X_\mathsf{c}; Y_\mathsf{p}\right), \tag{6b}$$

$$R_\mathsf{c} \leq I\left(X_\mathsf{c}; Y_\mathsf{c} | X_\mathsf{p}\right), \tag{6c}$$

and from [5] we have

$$R_\mathsf{p} + R_\mathsf{c} \leq I\left(X_\mathsf{p}; Y_\mathsf{p}, Y_\mathsf{Fc} | Y_\mathsf{c}, X_\mathsf{c}\right) + I\left(X_\mathsf{p}, X_\mathsf{c}; Y_\mathsf{c}\right), \tag{6d}$$

$$R_\mathsf{p} + R_\mathsf{c} \leq I\left(X_\mathsf{c}; Y_\mathsf{c} | Y_\mathsf{p}, X_\mathsf{p}\right) + I\left(X_\mathsf{p}, X_\mathsf{c}; Y_\mathsf{p}\right). \tag{6e}$$

Notice that in the bounds in (6a)-(6e), Y_Fc always appears conditioned on X_c. This implies that, for the ISD channel described in Section II-B, Y_Fc can be replaced with T_f without loss of generality.

2) For the *memoryless ISD CCIC*, described in Section II-B, with independent noises at the different source-destination pairs as in (2), we have

$$R_\mathsf{p} + R_\mathsf{c} \leq I\left(Y_\mathsf{p}; X_\mathsf{p}, X_\mathsf{c} | T_\mathsf{p}, T_\mathsf{f}\right) + I\left(Y_\mathsf{c}, T_\mathsf{f}; X_\mathsf{p}, X_\mathsf{c} | T_\mathsf{c}\right). \tag{6f}$$

The details of the proof of the bound in (6f) can be found in Appendix B.

We note that a bound as the one in (6f) was originally derived in [6, Appendix IV pages 177-179] for the ISD IC with bilateral source cooperation when all noises are independent; for the case of unilateral source cooperation, this follows from the following Markov chain

$$\left(W_\mathsf{p}, X_\mathsf{p}^{\ i}\right) - \left(T_\mathsf{f}^{\ i-1}\right) - \left(W_\mathsf{c}, X_\mathsf{c}^{\ i}, T_\mathsf{c}^{\ i}\right), \quad \forall i \in [1:N], \tag{7a}$$

$$\left(W_\mathsf{c}, X_\mathsf{c}^{\ i}\right) - \left(T_\mathsf{f}^{\ i-1}\right) - \left(W_\mathsf{p}, X_\mathsf{p}^{\ i}, T_\mathsf{p}^{\ i}\right), \quad \forall i \in [1:N]. \tag{7b}$$

A careful analysis of the bounding steps in [6, Appendix IV pages 177-179] shows that the derivation of the bound in (6f) is valid even when $\mathbb{P}_{Y_\mathsf{Fc},Y_\mathsf{c},Y_\mathsf{p}|X_\mathsf{p},X_\mathsf{c}}$ factors as in (2), i.e., the independent noises assumption at all terminals captured by the product distribution $\mathbb{P}_{Y_\mathsf{Fc},Y_\mathsf{c},Y_\mathsf{p}|X_\mathsf{p},X_\mathsf{c}} = \mathbb{P}_{Y_\mathsf{Fc}|X_\mathsf{p},X_\mathsf{c}}\mathbb{P}_{Y_\mathsf{c}|X_\mathsf{p},X_\mathsf{c}}\mathbb{P}_{Y_\mathsf{p}|X_\mathsf{p},X_\mathsf{c}}$ is not necessary for the bound to hold by

suitably modifying the Markov chains in (7) — see Lemma 1. Among the advantages of the bound in (6f) is that the case of output feedback from the intended destination is a special case of the more general framework and can be obtained by $Y_{\mathsf{Fc}} = Y_{\mathsf{c}}$.

We note that our bound in (6f) is not only more general but also tighter than the one in [6, Appendix IV pages 177-179] since $I\left(T_{\mathsf{f}}; X_{\mathsf{p}}, X_{\mathsf{c}}|T_{\mathsf{c}}\right) \leq I\left(T_{\mathsf{f}}; X_{\mathsf{p}}\right)$ and $H\left(Y_{\mathsf{p}}|T_{\mathsf{p}}, T_{\mathsf{f}}\right) \leq H\left(Y_{\mathsf{p}}|T_{\mathsf{p}}\right)$.

The key step of the proof for the bound in (6f) is the following Lemma:

Lemma 1 *For the ISD CCIC with the noise structure in* (2)*, the following Markov chains hold for all* $i \in [1 : N]$*:*

$$(W_{\mathsf{p}}, T_{\mathsf{p}}^{\;i-1}, X_{\mathsf{p}}^{\;i}) - (T_{\mathsf{c}}^{\;i-1}, T_{\mathsf{f}}^{\;i-1}) - (T_{\mathsf{c}i}), \tag{8a}$$

$$(W_{\mathsf{c}}, T_{\mathsf{c}}^{\;i-1}, X_{\mathsf{c}}^{\;i}) - (T_{\mathsf{p}}^{\;i-1}, T_{\mathsf{f}}^{\;i-1}) - (T_{\mathsf{p}_i}, T_{\mathsf{f}i}). \tag{8b}$$

Proof: The proof is based on the Functional Dependence Graph (FDG) [28] and can be found in Appendix A. ■

3) For the *memoryless ISD IC with output feedback* $Y_{\mathsf{Fc}} = Y_{\mathsf{c}}$ in (2), from [20, model-(1000)] we have

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq I\left(Y_{\mathsf{c}}; X_{\mathsf{p}}, X_{\mathsf{c}}\right) + I\left(Y_{\mathsf{c}}; X_{\mathsf{c}}|Y_{\mathsf{p}}, X_{\mathsf{p}}\right) + I\left(Y_{\mathsf{p}}; X_{\mathsf{p}}, X_{\mathsf{c}}|T_{\mathsf{p}}\right). \tag{9}$$

To the best of our knowledge, the bound in (9) is the only upper bound of the type $R_{\mathsf{p}} + 2R_{\mathsf{c}}$ available in the literature for the cooperative IC (which includes feedback models as a special case), but it is only valid for the case of output feedback. Our goal in the next section is to derive bounds of the type of (9) for the class of ISD CCICs described in Section II-B with independent noises at the different source-destination pairs as in (2).

B. Novel outer bounds

In this section we derive two novel outer bounds of the type $R_{\mathsf{p}} + 2R_{\mathsf{c}}$ and $2R_{\mathsf{p}} + R_{\mathsf{c}}$ on the capacity region for the ISD CCIC described in Section II-B with independent noises at the different source-destination pairs as in (2). These two outer bounds generalize to the CCIC those of the same type in [26, Theorem 1], derived for the classical non-cooperative IC.

Our main result in this section is as follows.

Theorem 1 *For the ISD CCIC satisfying the condition in* (2)*, the capacity region is outer bounded by*

$$2R_{\mathsf{p}} + R_{\mathsf{c}} \leq I\left(Y_{\mathsf{p}}; X_{\mathsf{p}}, X_{\mathsf{c}}\right) + I\left(Y_{\mathsf{p}}; X_{\mathsf{p}}|T_{\mathsf{f}}, Y_{\mathsf{c}}, X_{\mathsf{c}}\right) + I\left(Y_{\mathsf{c}}, T_{\mathsf{f}}; X_{\mathsf{p}}, X_{\mathsf{c}}|T_{\mathsf{c}}\right), \qquad (10)$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq I\left(Y_{\mathsf{c}}; X_{\mathsf{p}}, X_{\mathsf{c}}\right) + I\left(Y_{\mathsf{c}}; X_{\mathsf{c}}|T_{\mathsf{f}}, Y_{\mathsf{p}}, X_{\mathsf{p}}\right) + I\left(Y_{\mathsf{p}}, T_{\mathsf{f}}; X_{\mathsf{p}}, X_{\mathsf{c}}|T_{\mathsf{p}}\right). \qquad (11)$$

for some joint input distribution $\mathbb{P}_{X_{\mathsf{p}}, X_{\mathsf{c}}}$.

Note that, when evaluated for the case of output feedback with independent noises, i.e., $T_{\mathsf{f}} = T_{\mathsf{p}}$, the outer bound in (11) reduces to the one in (9).

Proof: By Fano's inequality, by considering that the messages W_{p} and W_{c} are independent and by giving side information similarly to [6], we have

$$N(2R_{\mathsf{p}} + R_{\mathsf{c}} - 3\epsilon_N)$$

$$\leq 2I\left(W_{\mathsf{p}}; Y_{\mathsf{p}}^N\right) + I\left(W_{\mathsf{c}}; Y_{\mathsf{c}}^N\right)$$

$$\leq I\left(W_{\mathsf{p}}; Y_{\mathsf{p}}^N\right) + I\left(W_{\mathsf{p}}; Y_{\mathsf{p}}^N, T_{\mathsf{p}}^N, T_{\mathsf{f}}^N|W_{\mathsf{c}}\right) + I\left(W_{\mathsf{c}}; Y_{\mathsf{c}}^N, T_{\mathsf{c}}^N, T_{\mathsf{f}}^N\right)$$

$$= H\left(Y_{\mathsf{p}}^N\right) - H\left(Y_{\mathsf{p}}^N, T_{\mathsf{p}}^N, T_{\mathsf{f}}^N|W_{\mathsf{p}}, W_{\mathsf{c}}\right) \qquad (12\text{a})$$

$$+ H\left(Y_{\mathsf{p}}^N, T_{\mathsf{p}}^N, T_{\mathsf{f}}^N|W_{\mathsf{c}}\right) - H\left(Y_{\mathsf{c}}^N, T_{\mathsf{c}}^N, T_{\mathsf{f}}^N|W_{\mathsf{c}}\right) \qquad (12\text{b})$$

$$+ H\left(Y_{\mathsf{c}}^N, T_{\mathsf{c}}^N, T_{\mathsf{f}}^N\right) - H\left(Y_{\mathsf{p}}^N|W_{\mathsf{p}}\right). \qquad (12\text{c})$$

We now analyze and bound each pair of terms.

Pair in (12a)*:* We have

$$H\left(Y_{\mathsf{p}}^N\right) - H\left(Y_{\mathsf{p}}^N, T_{\mathsf{p}}^N, T_{\mathsf{f}}^N|W_{\mathsf{p}}, W_{\mathsf{c}}\right)$$

$$\overset{(a)}{=} \sum_{i \in [1:N]} H\left(Y_{\mathsf{p}_i}|Y_{\mathsf{p}}^{i-1}\right) - H\left(Y_{\mathsf{p}_i}, T_{\mathsf{p}_i}, T_{\mathsf{f}_i}|W_{\mathsf{p}}, W_{\mathsf{c}}, Y_{\mathsf{p}}^{i-1}, T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}, X_{\mathsf{p}}^i, X_{\mathsf{c}}^i\right)$$

$$\overset{(b)}{\leq} \sum_{i \in [1:N]} H\left(Y_{\mathsf{p}_i}\right) - H\left(Y_{\mathsf{p}_i}, T_{\mathsf{p}_i}, T_{\mathsf{f}i}|W_{\mathsf{p}}, W_{\mathsf{c}}, Y_{\mathsf{p}}^{i-1}, T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}, X_{\mathsf{p}}^i, X_{\mathsf{c}}^i\right)$$

$$\overset{(c)}{=} \sum_{i \in [1:N]} H\left(Y_{\mathsf{p}_i}\right) - H\left(Y_{\mathsf{p}_i}, T_{\mathsf{p}_i}, T_{\mathsf{f}i}|X_{\mathsf{p}_i}, X_{\mathsf{c}i}\right),$$

where: the equality in (a) follows by applying the chain rule of the entropy and since, for the ISD CCIC, the encoding function $X_{\mathsf{c}i}(W_{\mathsf{c}}, Y_{\mathsf{Fc}}^{i-1})$ is equivalent to $X_{\mathsf{c}i}(W_{\mathsf{c}}, T_{\mathsf{f}}^{i-1})$ and since, given W_{p}, X_{p} is uniquely determined; the inequality in (b) is due to the conditioning reduces entropy principle; the equality in (c) follows because of the ISD property of the channel and since the channel is memoryless.

Pair in (12b)*:* We have

$$H\left(Y_{\mathsf{p}}^{N}, T_{\mathsf{p}}^{N}, T_{\mathsf{f}}^{N}|W_{\mathsf{c}}\right) - H\left(Y_{\mathsf{c}}^{N}, T_{\mathsf{c}}^{N}, T_{\mathsf{f}}^{N}|W_{\mathsf{c}}\right)$$

$$\stackrel{(d)}{=} \sum_{i\in[1:N]} H\left(Y_{\mathsf{p}i}, T_{\mathsf{p}i}, T_{\mathsf{f}i}|Y_{\mathsf{p}}^{i-1}, T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}, W_{\mathsf{c}}, X_{\mathsf{c}}^{i}\right) - H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i}|Y_{\mathsf{c}}^{i-1}, T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i-1}, W_{\mathsf{c}}, X_{\mathsf{c}}^{i}\right)$$

$$\stackrel{(e)}{=} \sum_{i\in[1:N]} H\left(Y_{\mathsf{p}i}, T_{\mathsf{p}i}, T_{\mathsf{f}i}|Y_{\mathsf{p}}^{i-1}, T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}, W_{\mathsf{c}}, X_{\mathsf{c}}^{i}\right) - H\left(T_{\mathsf{p}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i}|T_{\mathsf{p}}^{i-1}, T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i-1}, W_{\mathsf{c}}, X_{\mathsf{c}}^{i}\right)$$

$$\stackrel{(f)}{\leq} \sum_{i\in[1:N]} \underbrace{H\left(T_{\mathsf{p}i}, T_{\mathsf{f}i}|T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}, W_{\mathsf{c}}, X_{\mathsf{c}}^{i}\right) - H\left(T_{\mathsf{p}i}, T_{\mathsf{f}i}|T_{\mathsf{p}}^{i-1}, T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i-1}, W_{\mathsf{c}}, X_{\mathsf{c}}^{i}\right)}_{=\,0 \text{ because of (8b)}}$$

$$+ \sum_{i\in[1:N]} H\left(Y_{\mathsf{p}i}|T_{\mathsf{p}i}, T_{\mathsf{f}i}, X_{\mathsf{c}i}\right) - H\left(T_{\mathsf{c}i}|T_{\mathsf{p}}^{i}, T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i}, W_{\mathsf{c}}, X_{\mathsf{c}}^{i}, X_{\mathsf{p}}^{i}\right)$$

$$\stackrel{(g)}{=} \sum_{i\in[1:N]} H\left(Y_{\mathsf{p}i}|T_{\mathsf{p}i}, T_{\mathsf{f}i}, X_{\mathsf{c}i}\right) - H\left(Y_{\mathsf{p}i}|T_{\mathsf{p}i}, T_{\mathsf{f}i}, X_{\mathsf{c}i}, X_{\mathsf{p}i}\right),$$

where: the equality in (d) follows by applying the chain rule of the entropy and since, for the ISD CCIC, the encoding function $X_{\mathsf{c}i}(W_{\mathsf{c}}, Y_{\mathsf{Fc}}^{i-1})$ is equivalent to $X_{\mathsf{c}i}(W_{\mathsf{c}}, T_{\mathsf{f}}^{i-1})$; the equality in (e) is due to the fact that Y_{c} is a deterministic function of $(X_{\mathsf{c}}, T_{\mathsf{p}})$, which is invertible given X_{c}; the inequality in (f) is due to the conditioning reduces entropy principle; the equality in (g) follows because of the ISD property of the channel and since the channel is memoryless.

Pair in (12c)*:* We have

$$H\left(Y_{\mathsf{p}}^{N}|W_{\mathsf{p}}\right)$$

$$\stackrel{(h)}{=} \sum_{i\in[1:N]} H\left(Y_{\mathsf{p}i}|Y_{\mathsf{p}}^{i-1}, W_{\mathsf{p}}, X_{\mathsf{p}}^{i}\right)$$

$$\stackrel{(i)}{=} \sum_{i\in[1:N]} H\left(T_{\mathsf{c}i}|T_{\mathsf{c}}^{i-1}, W_{\mathsf{p}}, X_{\mathsf{p}}^{i}\right)$$

$$\stackrel{(j)}{\geq} \sum_{i\in[1:N]} H\left(T_{\mathsf{c}i}|T_{\mathsf{c}}^{i-1}, W_{\mathsf{p}}, X_{\mathsf{p}}^{i}, T_{\mathsf{f}}^{i-1}\right)$$

$$\stackrel{(k)}{=} \sum_{i\in[1:N]} H\left(T_{\mathsf{c}i}|T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i-1}\right) - \underbrace{I\left(T_{\mathsf{c}i}; W_{\mathsf{p}}, X_{\mathsf{p}}^{i}|T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i-1}\right)}_{=\,0 \text{ because of (8a)}},$$

where: the equality in (h) follows by applying the chain rule of the entropy and since, given W_{p}, X_{p} is uniquely determined; the equality in (i) is due to the fact that Y_{p} is a deterministic function of $(X_{\mathsf{p}}, T_{\mathsf{c}})$, which is invertible given X_{p}; the inequality in (j) follows since conditioning reduces the entropy; the equality in (k) follows from the definition of mutual information. Therefore,

$$H\left(Y_{\mathsf{c}}^{N}, T_{\mathsf{c}}^{N}, T_{\mathsf{f}}^{N}\right) - H\left(Y_{\mathsf{p}}^{N}|W_{\mathsf{p}}\right)$$

$$\overset{(l)}{\leq} \sum_{i \in [1:N]} H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i}|Y_{\mathsf{c}}^{i-1}, T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i-1}\right) - H\left(T_{\mathsf{c}i}|T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i-1}\right)$$

$$\overset{(m)}{\leq} \sum_{i \in [1:N]} H\left(T_{\mathsf{c}i}|T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i-1}\right) - H\left(T_{\mathsf{c}i}|T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i-1}\right) + H\left(Y_{\mathsf{c}i}, T_{\mathsf{f}i}|Y_{\mathsf{c}}^{i-1}, T_{\mathsf{c}}^{i}, T_{\mathsf{f}}^{i-1}\right)$$

$$\overset{(n)}{\leq} \sum_{i \in [1:N]} 0 + H\left(Y_{\mathsf{c}i}, T_{\mathsf{f}i}, |T_{\mathsf{c}i}\right),$$

where the inequality in (l) is a consequence of the inequality in (k) above and the inequalities in (m) and (n) are due to the conditioning reduces entropy principle.

Final step: By combining everything together, by introducing the time sharing random variable uniformly distributed over $[1:N]$ and independent of everything else, by dividing both sides by N and taking the limit for $N \to \infty$ we get the bound in (10). We finally notice that by dropping the time sharing we do not decrease the bound. Note also that, since for the ISD model defined in (1) T_{p}, respectively T_{c}, is a deterministic function of $(Y_{\mathsf{c}}, X_{\mathsf{c}})$, respectively $(Y_{\mathsf{p}}, X_{\mathsf{p}})$, we have $H\left(T_{\mathsf{p}}, T_{\mathsf{f}}|Y_{\mathsf{p}}, X_{\mathsf{p}}, X_{\mathsf{c}}\right) = H\left(Y_{\mathsf{c}}, T_{\mathsf{f}}|T_{\mathsf{c}}, X_{\mathsf{p}}, X_{\mathsf{c}}\right)$.

By following similar steps as in the derivation of (10) and by using the Markov chains in (8a) and (8b), one can derive the upper bound in (11). For completeness, we report the proof of (11) in Appendix C. ∎

In the next section we will evaluate the outer bounds in (6) and those in Theorem 1 for the Gaussian noise channel described in Section II-C and show that they allow to characterize the capacity to within a constant gap in the regimes which were left open in [17].

IV. The capacity region to within a constant gap for the symmetric GCCIC

In this section we analyze the practically relevant Gaussian noise channel described in Section II-C. In particular, we focus on the symmetric case defined in (4). For such a scenario, in [17, Theorem 1], we proved a constant gap of 1 bit in the strong interference regime, i.e., $\mathsf{I} \geq \mathsf{S}$ (equivalent to $\alpha \geq 1$), and in the weak interference regime, i.e., $\mathsf{I} < \mathsf{S}$ (equivalent to $\alpha < 1$) when the cooperation link is 'sufficiently strong' as quantified by $\mathsf{C} \geq \Delta_{\mathrm{th}}$ with

$$\Delta_{\mathrm{th}} := \left(\mathsf{S} + \mathsf{I} + 2\sqrt{\mathsf{IS}\frac{\mathsf{I}}{1+\mathsf{I}}}\right)(1+\mathsf{I}), \quad \text{(equivalent to } \beta \geq \alpha + 1). \tag{13}$$

The regimes for which we proved the constant gap result of 1 bit in [17, Theorem 1] are depicted in blue in Fig. 4, where the whole set of parameters has been partitioned into multiple sub-regions depending upon different levels of cooperation (β) and interference (α) strengths.
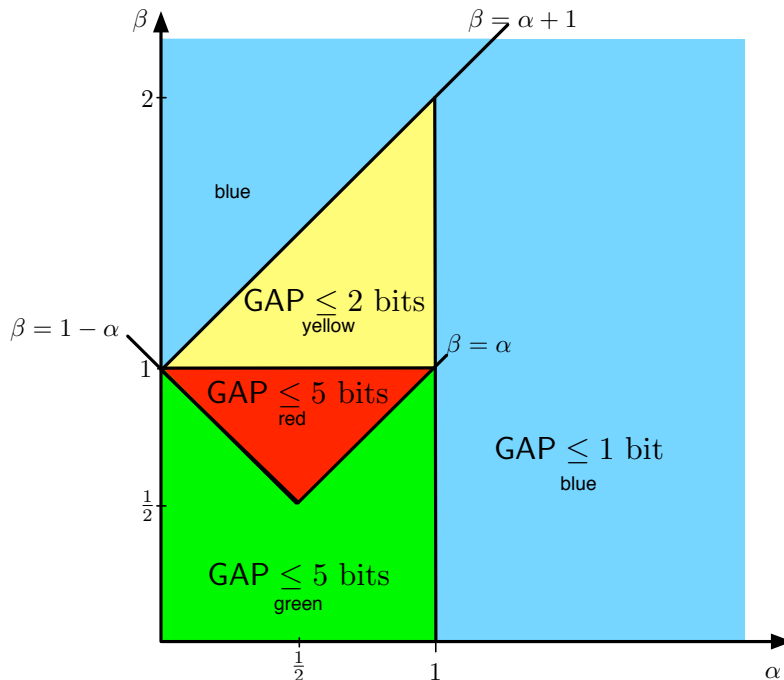
Fig. 4. Gap for different parameter regimes for the symmetric GCCIC.

Therefore, the case of weak interference ($\alpha < 1$) and weak cooperation ($\beta < \alpha + 1$) (i.e., yellow, red and green regions in Fig. 4) was left open in [17], where we speculated that for this regime, novel outer bounds of the type $2R_\mathsf{p} + R_\mathsf{c}$ and $R_\mathsf{p} + 2R_\mathsf{c}$ would be needed since, when the cooperation link is not 'sufficiently strong', the performance of the GCCIC should reduce to the classical non-cooperative IC, whose capacity has bounds of the type $2R_\mathsf{p} + R_\mathsf{c}$ and $R_\mathsf{p} + 2R_\mathsf{c}$ [18]. With the two novel outer bounds in Theorem 1 we can prove the main result of this section:

Theorem 2 *The capacity region outer bound for the symmetric GCCIC in weak interference, i.e., $\mathsf{I} < \mathsf{S}$, is achievable to within a constant gap. In particular:*

1) *$\mathsf{C} \leq \mathsf{S}$ (i.e., $\beta \leq 1$), which corresponds to the green and red regions in Fig. 4:* $\mathsf{GAP} \leq 5$ *bits.*

2) *$\mathsf{S} < \mathsf{C} < \Delta_{th}$ (i.e., $1 < \beta < 1 + \alpha$) for Δ_{th} in (13), which corresponds to the yellow region in Fig. 4:* $\mathsf{GAP} \leq 2$ *bits.*

3) *For the remaining parameter regimes, which correspond to the blue region in Fig. 4:* $\mathsf{GAP} \leq 1$ *bits [17].*

Theorem 2 and [17, Theorem 1] characterize the whole capacity region for the symmetric

GCCIC to within 5 bits. The rest of this section is dedicated to the proof of Theorem 2. In particular, in Section IV-A we evaluate the outer bounds in (6) and those in Theorem 1 for the symmetric GCCIC, in Section IV-B we derive a novel achievable rate region and finally in Section IV-C we show that the achievable rate region is a constant number of bits apart from the outer bound region.

A. Outer bound region

We evaluate the bounds in (6), (10) and (11) for the Gaussian noise channel in (3). We define $\mathbb{E}\left[X_{\mathsf{p}} X_{\mathsf{c}}^{*}\right] := \rho : |\rho| \in [0, 1]$. We also assume that all the noises are independent, which represents a particular case for which our outer bounds hold. By the 'Gaussian maximizes entropy' principle, jointly Gaussian inputs exhaust the outer bounds in (6), (10) and (11). Thus, we start by evaluating each mutual information term in (6), (10) and (11) by using jointly Gaussian inputs. Then, we further upper bound each mutual information term over the input correlation coefficient $\rho : |\rho| \in [0, 1]$. By doing so we obtain:

Lemma 2 *The capacity region of the symmetric GCCIC is contained into*

$$R_{\mathsf{p}} \leq \log\left(1 + \mathsf{C} + \mathsf{S}\right), \tag{14a}$$

$$R_{\mathsf{p}} \leq \log\left(1 + \left(\sqrt{\mathsf{S}} + \sqrt{\mathsf{I}}\right)^2\right), \tag{14b}$$

$$R_{\mathsf{c}} \leq \log\left(1 + \mathsf{S}\right), \tag{14c}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1 + \frac{\mathsf{S}}{1 + \mathsf{I}}\right) + \log\left(1 + \left(\sqrt{\mathsf{S}} + \sqrt{\mathsf{I}}\right)^2\right), \tag{14d}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \underbrace{\log\left(1 + \frac{\mathsf{S} + \mathsf{C}}{1 + \mathsf{I}}\right) + \log\left(1 + \left(\sqrt{\mathsf{S}} + \sqrt{\mathsf{I}}\right)^2\right)}_{\geq \mathrm{eq}(14\mathrm{d})}, \tag{14e}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \underbrace{\log\left(1 + \mathsf{C} + \mathsf{I} + \frac{\mathsf{S}}{1 + \mathsf{I}}\right) + \log\left(1 + \mathsf{I} + \mathsf{S}\frac{1 + \mathsf{C}}{1 + \mathsf{C} + \mathsf{I}}\right)}_{\stackrel{\text{for } \mathsf{C} \geq \mathsf{I}}{\geq} \log\left(1 + 2\mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{2}\right) \geq \mathrm{eq}(14\mathrm{d}) - 2\log(2)} + 2\log(2), \tag{14f}$$

$$2R_{\mathsf{p}} + R_{\mathsf{c}} \leq \underbrace{\log\left(1 + \left(\sqrt{\mathsf{S}} + \sqrt{\mathsf{I}}\right)^2\right) + \log\left(1 + \frac{\mathsf{S}}{1 + \mathsf{I}}\right)}_{\text{sum-rate bound in (14d)}} + \Delta_{(14\mathrm{g})} + \log(2), \tag{14g}$$

$$\Delta_{(14\mathrm{g})} := \log\left(1 + \frac{\mathsf{C}}{1 + \mathsf{I} + \mathsf{S}}\right) + \log\left(1 + \mathsf{I} + \mathsf{S}\frac{1 + \mathsf{C}}{1 + \mathsf{I} + \mathsf{C}}\right) \stackrel{\text{for } \mathsf{C} \geq \mathsf{I}}{\geq} \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{2}\right), \tag{14h}$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \le \underbrace{\log\left(1 + \left(\sqrt{\mathsf{S}} + \sqrt{\mathsf{I}}\right)^2\right) + \log\left(1 + \frac{\mathsf{S}}{1 + \mathsf{I}}\right)}_{\text{sum-rate bound in (14d)}} + \Delta_{(14i)} + \log(2), \tag{14i}$$

$$\Delta_{(14i)} := \log\left(1 + \mathsf{C} + \mathsf{I} + \frac{\mathsf{S}}{1 + \mathsf{I}}\right) \overset{\text{for } \mathsf{C} \ge \mathsf{S}}{\ge} \log(1 + \mathsf{S}). \tag{14j}$$

Proof: The outer bound region for the general GCCIC is given in (25) in Appendix D that, specialized to the symmetric case, gives the region in (14).

We note that the outer bound expressions in (14) are slightly different from those reported in [1, eq.(1)] and [2, eq.(6)] because of different bounding steps over the input correlation factor $\rho, |\rho| \in [0, 1]$; the form presented in (14), although not being the tightest, is in our opinion the most amenable for easy closed-form gap computations. ∎

The following discussion applies 'up to a constant gap', that is, by excluding terms that are not a function of the channel gains. In weak interference the bound in (14a) can be dropped since the condition $\mathsf{I} \le \mathsf{S}$ in (14b) implies $R_{\mathsf{p}} \le \log(1 + \mathsf{S}) + 2\log(2)$. The bound in (14e) can also be dropped because it is looser than the one in (14d). These observations imply that in the weak interference regime only the 'ISD bounds' in (14f)-(14i) depend on the strength of the cooperation link C. The channel conditions for which the whole outer bound in (14) does not depend on C can be characterized as follows:

1) $\mathsf{C} \ge \max\{\mathsf{S}, \mathsf{I}\}$: when the outer bound in (14) is an outer bound for the non-causal Gaussian CIC. The bounds in (14g) and (14i) are redundant when $\mathsf{C} \ge \max\{\mathsf{S}, \mathsf{I}\}$. Moreover, for $\mathsf{C} \ge \mathsf{I}$ the sum-rate in (14f) is looser than the one in (14d). We therefore conclude that for $\mathsf{C} \ge \max\{\mathsf{S}, \mathsf{I}\}$ the outer bound in (14) does not depend on C and reduces to the region

$$R_{\mathsf{p}} \le \log\left(1 + \left(\sqrt{\mathsf{S}} + \sqrt{\mathsf{I}}\right)^2\right), \tag{15a}$$

$$R_{\mathsf{c}} \le \log(1 + \mathsf{S}), \tag{15b}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \le \log\left(1 + \frac{\mathsf{S}}{1 + \mathsf{I}}\right) + \log\left(1 + \left(\sqrt{\mathsf{S}} + \sqrt{\mathsf{I}}\right)^2\right), \tag{15c}$$

which is an outer bound on the capacity region of the non-causal Gaussian CIC (see [24, Theorem III.1]). Thus, our outer bound predicts that, when $\mathsf{C} \ge \max\{\mathsf{S}, \mathsf{I}\}$ (yellow and part of the blue regions in Fig. 4), the symmetric GCCIC should behave, up to constant gap, as the symmetric non-causal Gaussian CIC; in Section IV-C3, we will formally prove this for

the yellow regime in Fig. 4, i.e., for $\mathsf{I} \leq \mathsf{S} \leq \mathsf{C} < \Delta_{\text{th}}$ with Δ_{th} defined in (13) (the blue regimes in Fig. 4, i.e., $\mathsf{S} \leq \mathsf{I}$ and $\mathsf{S} > \mathsf{I}$ with $\mathsf{C} \geq \Delta_{\text{th}}$, were proved in [17]).

2) $\mathsf{C} \leq \min\left\{\mathsf{S}, \frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}}\right\}$: when the outer bound in (14) is an outer bound for the classical Gaussian IC. In [17], we proved this behavior in strong interference (i.e., $\mathsf{S} \leq \mathsf{I}$), in which case we have $\min\left\{\mathsf{S}, \frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}}\right\} = \mathsf{S}$. On the other hand, in weak interference we have $\min\left\{\mathsf{S}, \frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}}\right\} = \frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}}$, in which case the condition $\mathsf{C} \leq \frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}}$ implies $\mathsf{C} \leq \mathsf{I}$ and $\mathsf{S}\frac{1+\mathsf{C}}{1+\mathsf{C}+\mathsf{I}} \leq \frac{\mathsf{S}}{1+\mathsf{I}} + \mathsf{I}$; thus, the outer bound in (14) reduces to the region

$$R_{\mathsf{p}} \leq \log\left(1 + \mathsf{S}\right) + 2\log(2), \tag{16a}$$

$$R_{\mathsf{c}} \leq \log\left(1 + \mathsf{S}\right), \tag{16b}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1 + \mathsf{S}\right) + \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + 2\log(2), \tag{16c}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + 4\log(2), \tag{16d}$$

$$2R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1 + \mathsf{S} + \mathsf{I}\right) + \log\left(\frac{1+\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + 5\log(2), \tag{16e}$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq \log\left(1 + \mathsf{S} + \mathsf{I}\right) + \log\left(\frac{1+\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + 4\log(2), \tag{16f}$$

which is an outer bound on the capacity region of the classical Gaussian IC (see [18, Theorem 3]). Thus, our outer bound predicts that, when $\mathsf{C} \leq \min\left\{\mathsf{S}, \frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}}\right\}$, the symmetric GCCIC should behave, up to constant gap, as the non-cooperative Gaussian IC; we will formally prove this for the weak interference regime $\mathsf{C} \leq \frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}}, \mathsf{I} \leq \mathsf{S}$ in Section IV-C1 (the strong interference regime $\mathsf{C} \leq \mathsf{S} \leq \mathsf{I}$ was proved in [17]).

B. Transmission strategy and achievable rate region

We next design a transmission strategy for the CCIC. In particular, since we aim to derive a scheme that is approximately optimal in weak interference, we consider both *private* and *common* messages for the PTx and the CTx, in the spirit of [18] for the classical IC. Moreover, depending on the strength of the cooperation link, the PTx might take advantage of the help of the CTx in transmitting its message, i.e., the messages of the PTx are both *cooperative* and *non-cooperative*.

In order to enable cooperation, a block Markov coding scheme is used as follows. Transmission is over a frame of $B \gg 1$ slots. In slot $t \in [1 : B]$, the PTx sends its old (cooperative common and

private) messages $(W_{\mathsf{p}cc,t-1}, W_{\mathsf{p}pc,t-1})$ and superposes to them the new (cooperative common and private) messages $(W_{\mathsf{p}cc,t}, W_{\mathsf{p}pc,t})$ and the new non-cooperative common and private messages $(W_{\mathsf{p}cn,t}, W_{\mathsf{p}pn,t})$. At the end of slot t, the CTx jointly decodes the new cooperative messages $(W_{\mathsf{p}cc,t}, W_{\mathsf{p}pc,t})$ after subtracting the contribution of the old messages $(W_{\mathsf{p}cc,t-1}, W_{\mathsf{p}pc,t-1})$. At the beginning of slot $t \in [1 : B]$, the CTx knows the PTx's old private cooperative message $W_{\mathsf{p}pc,t-1}$ and, depending on the strength of the cooperation link compared to the interference link, the CTx might precode both its private and common non-cooperative messages $(W_{\mathsf{c}cn,t}, W_{\mathsf{c}pn,t})$ against the known interference, in such a way that the CRx does not experience interference from the codewords conveying these messages. The destinations wait until the whole frame has been received and then proceed to simultaneous backward decode all messages.

The detailed derivation of the achievable rate region can be found in Appendix E. The achievable rate region in compact form can be obtained by applying the Fourier-Motzkin Elimination (FME) procedure on the rate constraints in (27). Since the rate region in (27) is specified by 8 auxiliary random variables and by 20 rate constraints, the FME turns out to be quite involved. However, depending on the strength of the cooperation link compared to the direct and interfering links, the PTx might not use some of the messages, i.e., the corresponding auxiliary random variables are set to a deterministic constant. In particular:

1) When $\beta \leq \max\{\alpha, 1 - \alpha\}$ (green region in Fig. 4), the cooperation link is quite weak; we therefore expect the CCIC to 'behave' as the classical non-cooperative IC [18] for which both *private* and *common non-cooperative* messages are approximately optimal. Differently from the classical IC, the PTx also conveys part of its message through the CTx. This *cooperative* message is *common*, and thus also decoded at the CRx. Actually, since the cooperation link is weak, the amount of information that can be decoded, and hence delivered, by the CTx is limited. Thus, there is no need to employ binning, i.e., the scheme is based on superposition only. In other words, for this regime, the PTx does not make use of the private cooperative message, i.e., with reference to the transmission strategy in Appendix E, we set $S_1 = Z_1 = \emptyset$.

2) When $\max\{\alpha, 1 - \alpha\} < \beta \leq 1$ (red region in Fig. 4), the cooperation link is stronger than the interfering link, but weaker than the direct link. Thus, on the one hand the PTx takes advantage of these channel conditions by using *cooperative* messages; on the other hand, the cooperation link is not strong enough (i.e., weaker than the direct link) to allow the CTx to fully decode the PTx's message, and hence the PTx also uses a *non-cooperative*

message. In particular, the *non-cooperative* message of the PTx is *private*; this is because the interference is too weak and forcing the CRx to fully decode the PTx's message would constrain the rate too much. At the same time, the CTx can also benefit from the strength of the cooperation link and boost its rate performance by precoding its message against the private cooperative message of the PTx, i.e., the scheme is based both on superposition and binning. In other words, for this regime, the PTx does not make use of the common non-cooperative message, i.e., with reference to the transmission strategy in Appendix E, we set $U_1 = \emptyset$.

3) When $1 < \beta < \alpha+1$ (yellow region in Fig. 4), the cooperation link is stronger than both the interfering and the direct links. Thus, the PTx takes advantage of the strong cooperation link and sends its message to the PRx with the help of the CTx, i.e., the messages of the PTx are only *cooperative*. Moreover, since the interference is weak, the messages of the CTx and of the PTx are both *common* and *private*. Also for this regime we use binning at the CTx. In other words, for this regime, the PTx does not make use of the non-cooperative messages, i.e., with reference to the transmission strategy in Appendix E, we set $U_1 = T_1 = \emptyset$.

As per our discussion above, depending on the strength of the cooperation link compared to the direct and interfering links, some types of messages are not needed to achieve the outer bound to within a constant gap. Thus, instead of performing the FME directly on the rate constraints in (27), we apply the FME for two special cases: when $S_1 = Z_1 = \emptyset$ (to obtain a scheme for the green region in Fig. 4) and when $U_1 = \emptyset$ (to obtain a scheme for the red and yellow regions in Fig. 4; for the yellow region we then further set $T_1 = \emptyset$). The details can be found in Appendix E-A and Appendix E-B, respectively. In the next section, we will show that the two achievable rate regions achieve the outer bound in (14) to within a constant gap.

Remark 1 *Although in this work we focus on the symmetric GCCIC, i.e., the two direct and the two interfering links are of the same strength, the derived outer bound and the designed transmission strategy are valid for a general GCCIC, which is described by 5 different channel gains. Extensions to the general case are part of future work.*

Remark 2 *The novelty of the transmission strategy designed in Section IV-B compared to those proposed in [17] lies in the fact that the PTx uses cooperative and non-cooperative messages together. In particular: (i) the scheme based on superposition coding in [17, Appendix B] used $U_1 = \emptyset$, i.e., the PTx does not use a non-cooperative common message, while the superposition*

based scheme here proposed has $U_1 \neq \emptyset$; (ii) the scheme based on binning and superposition coding in [17, Appendix C] used $U_1 = T_1 = \emptyset$, i.e., the PTx does not use non-cooperative common and private messages, while the scheme here proposed has $U_1 \neq \emptyset$ and $T_1 \neq \emptyset$.

C. Constant gap characterization

In the following we analyze the green, red and yellow regions in Fig. 4 separately.

1) Regime $\mathsf{C} \leq \max\left\{\mathsf{I}, \frac{\mathsf{S}}{1+\mathsf{I}}\right\}, \mathsf{I} \leq \mathsf{S}$ (green region in Fig. 4): As remarked in item 1 in Section IV-B, for this region we set $S_1 = Z_1 = \emptyset$ in the transmission strategy in Appendix E, i.e., the PTx does not use private cooperative messages. With this choice and after performing the FME, we obtain the rate region in (28) in Appendix E-A, which evaluated for the Gaussian noise case gives the region in (29). In (29), we set $\mathsf{I}_\mathsf{p} = \mathsf{I}_\mathsf{c} = \mathsf{I}$, $\mathsf{S}_\mathsf{p} = \mathsf{S}_\mathsf{c} = \mathsf{S}$ (i.e., we consider the symmetric case) and $|b_2|^2 = 1 - |a_2|^2 = \frac{1}{1+\mathsf{I}}$ so that the private message of CTx (conveyed by T_2) is received below the noise level at the PRx in the spirit of [18]. Regarding the choice of the power splits for the PTx, we further split the green region into two subregions: subregion (i) for which $\mathsf{C} \leq \frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}}$ (i.e., $\beta \leq [2\alpha - 1]^+$) and subregion (ii) for which $\mathsf{C} > \frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}}$ (i.e., $\beta > [2\alpha - 1]^+$). We now analyze these two subregions separately.

Subregion (i): when $\beta \leq [2\alpha - 1]^+$, the cooperation link is very weak and thus we expect the GCCIC to behave as the non-cooperative Gaussian IC [18]. Therefore, we set the power of the cooperative common message (carried by V_1) to $|b_1|^2 = 0$ in (29) and $|c_1|^2 = 1 - |a_1|^2 = \frac{1}{1+\mathsf{I}}$ so that the private message of PTx (conveyed by T_1) is received below the noise level at the CRx in the spirit of [18]. With these choices and by removing the redundant constraints in (29) (i.e., eq(29a), eq(29e), eq(29f), eq(29g), eq(29j) and eq(29k)), we get that the achievable rate region in (29) is contained into

$$\mathcal{I}^{\text{green-(i)}}: \quad R_\mathsf{p} \leq \log\left(1 + \mathsf{S}\right) - \log(2), \tag{17a}$$

$$R_\mathsf{c} \leq \log\left(1 + \mathsf{S}\right) - \log(2), \tag{17b}$$

$$R_\mathsf{p} + R_\mathsf{c} \leq \log\left(1 + \mathsf{S} + \mathsf{I}\right) + \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right) - 2\log(2), \tag{17c}$$

$$R_\mathsf{p} + R_\mathsf{c} \leq \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}}\right) - 2\log(2), \tag{17d}$$

$$2R_\mathsf{p} + R_\mathsf{c} \leq \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \mathsf{S} + \mathsf{I}\right) + \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}}\right) - 3\log(2), \tag{17e}$$

$$R_\mathsf{p} + 2R_\mathsf{c} \leq \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \mathsf{S} + \mathsf{I}\right) - 3\log(2). \tag{17f}$$

Notice that the rate region in (17) is the achievable region for the classical symmetric non-cooperative IC in weak interference, which is optimal up to a gap of 1 bit/user [18].

For this regime, the outer bound in (14) can be further upper bounded (by removing the constraints in (14b) and (14e)) as

$$\mathcal{O}^{\text{green-(i)}}: \quad R_{\mathsf{p}} \leq \log\left(1 + \mathsf{S}\right) + \log(2), \tag{18a}$$

$$R_{\mathsf{c}} \leq \log\left(1 + \mathsf{S}\right), \tag{18b}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \mathsf{S} + \mathsf{I}\right) + \log(2), \tag{18c}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \overset{0 \leq \mathsf{C} \leq \mathsf{I}}{\leq} \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \mathsf{I} + \mathsf{S}\frac{1+\mathsf{C}}{1+\mathsf{I}}\right) + 3\log(2)$$

$$\overset{\mathsf{C} \leq \frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}}}{\leq} \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + 4\log(2), \tag{18d}$$

$$2R_{\mathsf{p}} + R_{\mathsf{c}} \overset{0 \leq \mathsf{C} \leq \mathsf{I}}{\leq} \log\left(1 + \mathsf{S} + \mathsf{I}\right) + \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \mathsf{I} + \mathsf{S}\frac{1+\mathsf{C}}{1+\mathsf{I}}\right) + 3\log(2)$$

$$\overset{\mathsf{C} \leq \frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}}}{\leq} \log\left(1 + \mathsf{S} + \mathsf{I}\right) + \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + 4\log(2), \tag{18e}$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \overset{0 \leq \mathsf{C} \leq \mathsf{I}}{\leq} \log\left(1 + \mathsf{S} + \mathsf{I}\right) + \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + 3\log(2). \tag{18f}$$

It is easy to see that the outer bound region in (18) and the achievable rate region in (17) are to within 3 bits/user of one another. Notice that in order to prove a constant gap we compared: eq(17a) with eq(18a), eq(17b) with eq(18b), eq(17c) with eq(18c), eq(17d) with eq(18d), eq(17e) with eq(18e), and eq(17f) with eq(18f).

Subregion (ii): when $\beta > [2\alpha - 1]^+$, the GCCIC starts to benefit from cooperation and indeed the outer bound region depends on C. Therefore the cooperative common message carried by V_1 can boost the rate performance of the system. In (29), we set the power of the common non-cooperative message (carried by U_1) to $|a_1|^2 = \frac{1}{2(1+\min\{\mathsf{C},\mathsf{I}\})}$. This choice is motivated by the fact that, in order to approximately match the outer bound, the single rate constraint on R_{p} in (29b) must approximately behave as an interference-free point-to-point channel. Therefore, the fact that the CTx can now decode part of the message of the PTx (carried by V_1) must not limit (up to a constant gap) the performance of the PTx. In other words, since C is 'quite large'

but not 'huge', the rate of V_1 cannot be too large. Moreover, we set $|c_1|^2 = \frac{1}{2(1+\mathsf{I})}$ so that the private message of PTx (conveyed by T_1) is received below the noise level at the CRx in the spirit of [18]. Thus, if $\mathsf{I} \leq \mathsf{C}$ we have $|b_1|^2 = \frac{1}{1+\mathsf{I}}$, while if $\mathsf{I} > \mathsf{C}$ we have $|b_1|^2 = \frac{\mathsf{C}+\mathsf{I}+2\mathsf{C}\mathsf{I}}{2(1+\mathsf{C})(1+\mathsf{I})}$. With these choices and by removing the redundant constraints in (29) (i.e., eq(29a), eq(29d), eq(29f), eq(29g), and eq(29l)), we get that the achievable rate region in (29) is contained into

$$\mathcal{I}^{\text{green-(ii)}}: \quad R_{\mathsf{p}} \leq \log\left(1+\mathsf{S}\right) - 4\log(2), \tag{19a}$$

$$R_{\mathsf{c}} \leq \log\left(1+\mathsf{S}\right) - \log(2), \tag{19b}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1+\mathsf{S}+\mathsf{I}\right) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) - 3\log(2), \tag{19c}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)$$
$$+ \log\left(1+\min\{\mathsf{I},\mathsf{C}\}\right) - 5\log(2), \tag{19d}$$

$$2R_{\mathsf{p}} + R_{\mathsf{c}} \leq 2\log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\mathsf{S}+\mathsf{I}\right)$$
$$+ \log\left(1+\min\{\mathsf{I},\mathsf{C}\}\right) - 6\log(2), \tag{19e}$$

$$2R_{\mathsf{p}} + R_{\mathsf{c}} \leq 2\log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\min\{\mathsf{I},\mathsf{C}\}}\right)$$
$$+ 2\log\left(1+\min\{\mathsf{I},\mathsf{C}\}\right) - 9\log(2), \tag{19f}$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq \log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)$$
$$+ \log\left(1+\mathsf{S}+\mathsf{I}\right) - 4\log(2). \tag{19g}$$

For this regime, the outer bound in (14) can be further upper bounded (by removing the constraints in (14b) and (14e)) as

$$\mathcal{O}^{\text{green-(ii)}}: \quad R_{\mathsf{p}} \leq \log\left(1+\mathsf{S}\right) + \log(2), \tag{20a}$$

$$R_{\mathsf{c}} \leq \log\left(1+\mathsf{S}\right), \tag{20b}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1+\mathsf{S}+\mathsf{I}\right) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log(2), \tag{20c}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \overset{0\leq\mathsf{C}\leq\max\left\{\mathsf{I},\frac{\mathsf{S}}{1+\mathsf{I}}\right\}}{\leq} \log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)$$
$$+ \log\left(1+\mathsf{I}+\frac{\mathsf{S}(1+\mathsf{C})}{1+\mathsf{I}+\mathsf{C}}\right) + 3\log(2), \tag{20d}$$

$$2R_{\mathsf{p}} + R_{\mathsf{c}} \overset{0\leq\mathsf{C}\leq\mathsf{S}}{\leq} \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\mathsf{S}+\mathsf{I}\right)$$

$$+ \log\left(1 + \mathsf{I} + \frac{\mathsf{S}(1+\mathsf{C})}{1+\mathsf{C}+\mathsf{I}}\right) + 3\log(2), \tag{20e}$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \overset{0 \leq \mathsf{C} \leq \max\left\{\mathsf{I}, \frac{\mathsf{S}}{1+\mathsf{I}}\right\}}{\leq} \log\left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right)$$

$$+ \log\left(1 + \mathsf{S} + \mathsf{I}\right) + 3\log(2). \tag{20f}$$

It is easy to see that the outer bound region in (20) and the achievable rate region in (19) are to within 5 bits/user of one another. Notice that in order to prove a constant gap we compared: eq(19a) with eq(20a), eq(19b) with eq(20b), eq(19c) with eq(20c), eq(19d) with eq(20d), eq(19e) with eq(20e), eq(19f) with eq(20e), and eq(19g) with eq(20f).

2) Regime $\max\left\{\mathsf{I}, \frac{\mathsf{S}}{1+\mathsf{I}}\right\} < \mathsf{C} \leq \mathsf{S}$ *(red region in Fig. 4):* As remarked in item 2 in Section IV-B, for this region we set $U_1 = \emptyset$ in the transmission strategy in Appendix E, i.e., the PTx does not use a common non-cooperative message. With this choice and after performing the FME, we obtain the rate region in (30) in Appendix E-B, which evaluated for the Gaussian noise case gives the region in (31). For $\mathsf{I}_{\mathsf{p}} = \mathsf{I}_{\mathsf{c}} = \mathsf{I}$ and $\mathsf{S}_{\mathsf{p}} = \mathsf{S}_{\mathsf{c}} = \mathsf{S}$, we set $|a_2| = 0$, and $|c_2|^2 = 1 - |b_2|^2 = \frac{1}{1+\mathsf{I}}$ in the region in (31). With this choice the private message of the CTx (conveyed by T_2) is received below the noise level at the PRx in the spirit of [18]. Note also that the CTx does not cooperate with the PTx in conveying information to the PRx, but it just exploits the information it learns through the cooperation link to smartly pre-encode its messages. For the PTx we let $|a_1|^2 = |b_1|^2 = \frac{\mathsf{I}+\mathsf{C}+2\mathsf{IC}}{4(1+\mathsf{I})(1+\mathsf{C})}$, $|c_1|^2 = \frac{1}{2(1+\mathsf{I})}$ and $|d_1|^2 = \frac{1}{2(1+\mathsf{C})}$; with this choice of the power splits and since we are in the regime $\mathsf{C} > \mathsf{I}$, the two private messages of the PTx (i.e., the cooperative one carried by Z_1 and the non-cooperative one carried by T_1) are received at most at the level of the noise at the CRx. Moreover, the non-cooperative private message (carried by T_1) is received at the level of the noise at the CTx.

With these choices we get that the achievable rate region in (31) is contained into (by considering $\min\{k_1, k_2\} \geq 0$ in (31))

$$\mathcal{I}^{\text{red}}: \quad R_{\mathsf{p}} \leq \log\left(1 + \mathsf{C} + \mathsf{S}\right) - 5\log(2), \tag{21a}$$

$$R_{\mathsf{p}} \leq \log\left(1 + \mathsf{S} + \mathsf{I}\right) - \log(2), \tag{21b}$$

$$R_{\mathsf{c}} \leq \log\left(1 + \mathsf{S}\right) - \log(3), \tag{21c}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1 + \mathsf{C}\right) + \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{C}} + \mathsf{I}\right)$$

$$+ \log \left(1 + \frac{\mathsf{S}}{1+\mathsf{I}} \right) - 5 \log(2) - \log(3), \tag{21d}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log \left(1 + \mathsf{S} + \mathsf{I} \right) + \log \left(1 + \frac{\mathsf{S}}{1+\mathsf{I}} \right) - \log(2) - \log(3), \tag{21e}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log \left(1 + \frac{\mathsf{C}}{1+\mathsf{I}} \right) + \log \left(1 + \frac{\mathsf{S}}{1+\mathsf{C}} \right) + \log \left(1 + \mathsf{S} \right) - 4 \log(2) - \log(3), \tag{21f}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log \left(1 + \mathsf{S} \right) + \log \left(1 + \frac{\mathsf{S}}{1+\mathsf{I}} + \frac{\mathsf{S}}{1+\mathsf{C}} \right) - 2 \log(2) - \log(3), \tag{21g}$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq \log \left(1 + \mathsf{S} + \mathsf{I} \right) + \log \left(1 + \frac{\mathsf{S}}{1+\mathsf{I}} \right) + \log \left(1 + \mathsf{S} \right) - 3 \log(2) - 2 \log(3), \tag{21h}$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq \log \left(1 + \mathsf{S} \right) + \log \left(1 + \frac{\mathsf{S}}{1+\mathsf{I}} \right) + \log \left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{C}} \right)$$
$$+ \log \left(1 + \frac{\mathsf{C}}{1+\mathsf{I}} \right) - 4 \log(2) - 2 \log(3), \tag{21i}$$

$$R_{\mathsf{p}} + 3R_{\mathsf{c}} \leq \log \left(1 + \mathsf{S} + \mathsf{I} \right) + 2 \log \left(1 + \frac{\mathsf{S}}{1+\mathsf{I}} \right) + \log \left(1 + \mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{C}} \right)$$
$$\log \left(1 + \mathsf{S} \right) - 5 \log(2) - 3 \log(3). \tag{21j}$$

For this regime, the outer bound in (14) can be further upper bounded (by considering the constraints in (14a), (14b), (14c), (14d) and (14i))

$$\mathcal{O}^{\text{red}}: \quad R_{\mathsf{p}} \leq \log \left(1 + \mathsf{C} + \mathsf{S} \right), \tag{22a}$$

$$R_{\mathsf{p}} \leq \log \left(1 + \mathsf{S} + \mathsf{I} \right) + \log(2), \tag{22b}$$

$$R_{\mathsf{c}} \leq \log \left(1 + \mathsf{S} \right), \tag{22c}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log \left(1 + \frac{\mathsf{S}}{1+\mathsf{I}} \right) + \log \left(1 + \mathsf{S} + \mathsf{I} \right) + \log(2), \tag{22d}$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq \log \left(1 + \mathsf{S} + \mathsf{I} \right) + \log \left(1 + \frac{\mathsf{S}}{1+\mathsf{I}} \right) + \log \left(1 + \mathsf{C} \right) + 2 \log(2) + \log(3), \tag{22e}$$

where the inequality in (22e) follows since $\mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}} \leq 2 \max \left\{ \mathsf{I}, \frac{\mathsf{S}}{1+\mathsf{I}} \right\} \leq 2\mathsf{C}$.

It is easy to see that the outer bound region in (22) and the achievable rate region in (21) are to within 5 bits/user of one another. Notice that in order to prove a constant gap we compared: eq(21a) with eq(22a), eq(21b) with eq(22b), eq(21c) with eq(22c), eq(21d) with eq(22d), eq(21e) with eq(22d), eq(21f) with eq(22d), eq(21g) with eq(22d), eq(21h) with eq(22c)+eq(22d), eq(21i) with eq(22e), and eq(21j) with eq(22c)+eq(22e).

3) Regime $\mathsf{I} \leq \mathsf{S} < \mathsf{C} < \Delta_{th}$ *with* Δ_{th} *defined in* (13) *(yellow region in Fig. 4):* As remarked in item 3 in Section IV-B, for this region we set $U_1 = T_1 = \emptyset$ in the transmission strategy in

Appendix E, i.e., the PTx uses only cooperative messages. With this choice and after performing the FME, we obtain the rate region in (30) in Appendix E-B, which evaluated for the Gaussian noise case gives the region in (31). In (31), for $I_p = I_c = I$ and $S_p = S_c = S$, we set $|a_1|^2 = |b_1|^2 = \frac{I}{2(1+I)}$, $|b_2|^2 = \frac{I}{1+I}$ and $|c_1|^2 = |c_2|^2 = \frac{1}{1+I}$. Notice that with this choice of the power splits, we have $|d_1| = 0$, i.e., the power allocated for T_1 is zero. With these choices, we get that the achievable rate region in (31) can be further lower bounded (by considering $\min\{k_1, k_2\} \geq 0$ and that the constraint in (31e) is redundant in (31)) as

$$\mathcal{I}^{\text{yellow}}: \quad R_p \leq \log(1 + C) - \log(2), \tag{23a}$$

$$R_p \leq \log(1 + S + I) - \log(2), \tag{23b}$$

$$R_c \leq \log(1 + S) - \log(2), \tag{23c}$$

$$R_p + R_c \leq \log(1 + C) + \log(1 + S + I) - 3\log(2), \tag{23d}$$

$$R_p + R_c \leq \log\left(1 + \frac{C}{1 + I}\right) + \log(1 + S) - \log(2), \tag{23e}$$

$$R_p + R_c \leq \log\left(1 + \frac{S}{1 + I}\right) + \log(1 + S) - 2\log(2), \tag{23f}$$

$$R_p + 2R_c \leq \log(1 + S) + \log\left(1 + \frac{S}{1 + I}\right) + \log(1 + S + I) - 4\log(2), \tag{23g}$$

$$R_p + 2R_c \leq \log\left(1 + \frac{C}{1 + I}\right) + \log(1 + S + I) + \log(1 + S) - 3\log(2), \tag{23h}$$

$$R_p + 3R_c \leq 2\log(1 + S + I) + \log(1 + S) + \log\left(1 + \frac{S}{1 + I}\right) - 6\log(2). \tag{23i}$$

For this regime, the outer bound in (14) can be further upper bounded (by considering the constraints in (14a), (14b), (14c), and (14d))

$$\mathcal{O}^{\text{yellow}}: \quad R_p \leq \log(1 + C) + \log(2), \tag{24a}$$

$$R_p \leq \log(1 + S + I) + \log(2), \tag{24b}$$

$$R_c \leq \log(1 + S), \tag{24c}$$

$$R_p + R_c \leq \log\left(1 + \frac{S}{1 + I}\right) + \log(1 + S + I) + \log(2). \tag{24d}$$

It is easy to see that the outer bound region in (24) and the achievable rate region in (23) are to within 2 bits/user of one another. Notice that in order to prove a constant gap we compared: eq(23a) with eq(24a), eq(23b) with eq(24b), eq(23c) with eq(24c), eq(23d) with

eq(24a)+eq(24c), eq(23e) with eq(24d), eq(23f) with eq(24d), eq(23g) with eq(24c)+eq(24d), eq(23h) with eq(24c)+eq(24d), and eq(23i) with 2eq(24c)+eq(24d).

D. Discussion

The two novel outer bounds $2R_{\sf p} + R_{\sf c}$ in (14g) and $R_{\sf p} + 2R_{\sf c}$ in (14i) are active when $\sf S \geq \max\{C, I\}$ (weak interference and weak cooperation, which corresponds to the red and green regions in Fig. 4). In [19], the authors interpreted the need of this type of bounds as a measure of the amount of the 'resource holes', or inefficiency, due to the distributed nature of the non-cooperative classical IC [18]. Thus, in line with the work in [19], we conclude that when $\sf S \geq \max\{C, I\}$ unilateral cooperation is too weak to allow for a full utilization of the channel resources, i.e., it leaves some system resources underutilized. In particular:

- **Strong interference** (i.e., $\sf I > S$): in this regime neither the capacity region of the non-cooperative Gaussian IC [18] nor the capacity region of the non-causal Gaussian CIC [24], have bounds of the type $2R_{\sf p} + R_{\sf c}$ and $R_{\sf p} + 2R_{\sf c}$. It turns out that these types of bounds are not necessary for the GCCIC either.

- **Weak interference and strong cooperation** (i.e., $\sf I \leq S < C$): for this regime the outer bound in (14) equals (to within a constant gap) the outer bound on the capacity region for the non-causal Gaussian CIC [24, Theorem III.1], which does not have bounds of the type $2R_{\sf p} + R_{\sf c}$ and $R_{\sf p} + 2R_{\sf c}$ (see discussion in item 1 in Section IV-A). In other words, in this regime the ideal non-causal cognition assumption at the CTx just provides a bounded rate increase compared to the more practical case of causal learning for the CTx through a noisy link. It hence follows that for this regime unilateral cooperation allows to fully utilize the channel resources [19], i.e., the bounds of the type $2R_{\sf p} + R_{\sf c}$ and $R_{\sf p} + 2R_{\sf c}$ are not active.

- **Weak interference and weak cooperation** (i.e., $\sf S \geq \max\{C, I\}$): for this regime the capacity region of the non-cooperative Gaussian IC has bounds of the type $2R_{\sf p} + R_{\sf c}$ and $R_{\sf p} + 2R_{\sf c}$ [18], while the one of the non-causal Gaussian CIC does not [24]. From our constant gap result in this region, it follows that $2R_{\sf p} + R_{\sf c}$ in (14g) and $R_{\sf p} + 2R_{\sf c}$ in (14i) are both active. In other words, in this regime unilateral cooperation does not allow enough coordination among the sources which results in some 'resource holes' as in the non-cooperative Gaussian IC.

V. CONCLUSIONS

In this work we studied the two-user CCIC, an interference channel where one capable full-duplex source, i.e., the *cognitive* CTx, cooperates with / assists the other source, i.e., the *primary* PTx, to convey information. In contrast to the original overlay cognitive paradigm, where the CTx a priori knows the message of the PTx, in the CCIC the CTx causally learns the primary's data through a noisy in-band link. We first derived two novel outer bounds of the type $2R_\mathsf{p} + R_\mathsf{c}$ and $R_\mathsf{p} + 2R_\mathsf{c}$ on the capacity region of the injective semi-deterministic channel with independent noises at the two pairs. We then designed a transmission strategy based on binning and superposition encoding, partial-decode-and-forward relaying and simultaneous decoding and we derived its achievable rate region. We finally evaluated the outer and lower bounds on the capacity for the practically relevant Gaussian noise case and we proved that our bounds are a constant number of bits apart from one another for the symmetric case (i.e., the two direct links and the two interfering links are of the same strength) in weak interference when the cooperation link is weaker than a given threshold. We showed that the two novel outer bounds of the type $2R_\mathsf{p} + R_\mathsf{c}$ and $R_\mathsf{p} + 2R_\mathsf{c}$ are active in weak interference when the cooperation link is weaker than the direct link, i.e., in this regime unilateral cooperation is too weak to allow for a full utilization of the channel resources.

APPENDIX A

PROOF OF THE MARKOV CHAINS IN (8a) AND (8b)

We start by proving the two Markov chains in (8a)-(8b) by using the FDG [28]. Fig. 5 proves the Markov chain in (8a), while Fig. 6 the one in (8b). The two proofs, without loss of generality, consider the time instant $i = 3$. According to [28], we proceed through the following steps.

1) Draw the directed graph $\mathcal{G}_1$, which takes into consideration the dependence between the different random variables involved in the ISD CCIC considered. In particular, we define

$$Z_{\mathsf{f}i}^\star = \begin{bmatrix} Z_{\mathsf{f}i} \\ Z_{\mathsf{c}i} \end{bmatrix}$$

to consider the fact that the noises at the CTx and at the CRx can be arbitrarily correlated and we have

$$X_{\mathsf{p}i} = f(W_\mathsf{p}), \ Y_{\mathsf{p}i} = f(X_{\mathsf{p}i}, T_{\mathsf{c}i}), \ T_{\mathsf{c}i} = f(X_{\mathsf{c}i}, Z_{\mathsf{p}i}), \ T_{\mathsf{f}i} = f(X_{\mathsf{p}i}, Z_{\mathsf{f}i}^\star),$$

$$X_{\mathsf{c}i} = f(W_\mathsf{c}, T_\mathsf{f}^{i-1}), \ Y_{\mathsf{c}i} = f(X_{\mathsf{c}i}, T_{\mathsf{p}i}), \ T_{\mathsf{p}i} = f(X_{\mathsf{p}i}, Z_{\mathsf{f}i}^\star),$$
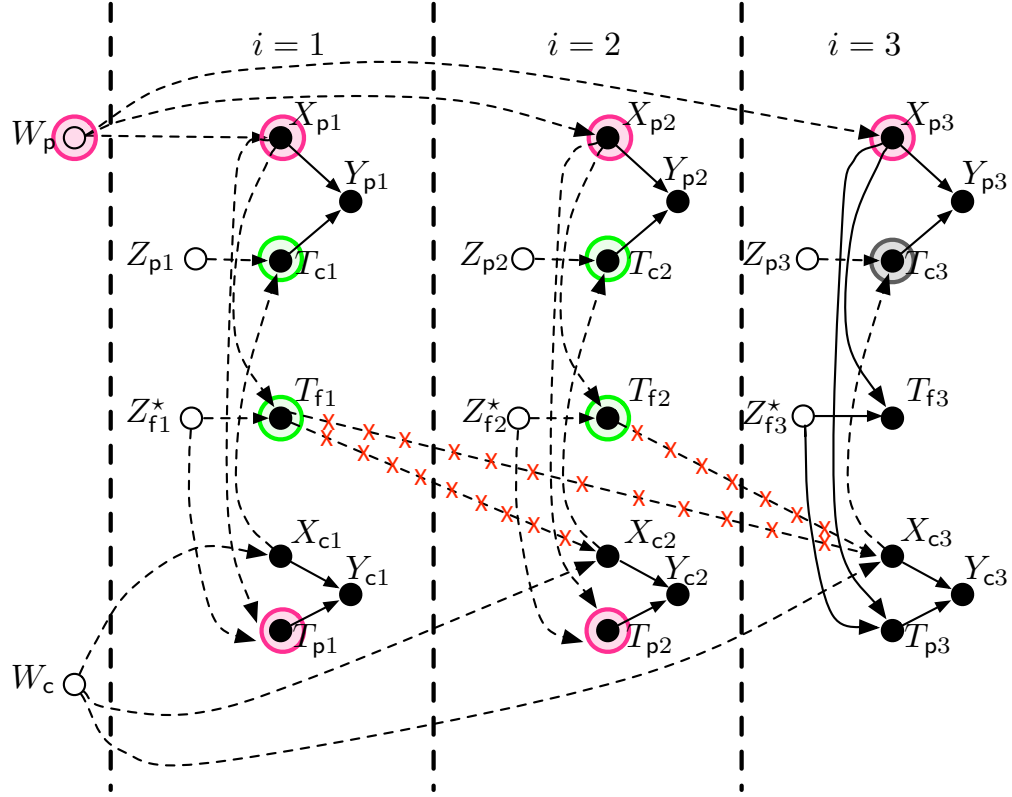
Fig. 5. Proof of the Markov chain in (8a) using the FDG.

where with f we indicate that the left-hand side of the equality is a function of the random variables into the bracket.

2) In $\mathcal{G}_1$, highlight all the different nodes / random variables involved in the two Markov chains in (8a)-(8b) we aim to prove. In particular, the random variables circled in magenta, given those circled in green, should be proved to be independent of those circled in grey.

3) From the graph $\mathcal{G}_1$, consider the subgraph $\mathcal{G}_2$ which contains those edges and vertices encountered when moving backwards one or more edges starting from the colored (magenta, green and grey) random variables. The edges of the subgraph $\mathcal{G}_2$ are depicted with dashed black lines in Fig. 5 and Fig. 6 and the vertices in $\mathcal{G}_2$ are all those touched by a dashed black line.

4) From the graph $\mathcal{G}_2$, remove all the edges coming out from the random variables in green (those which are supposed to d-separate the random variables colored in magenta and grey). In Fig. 5 and Fig. 6, this step is highlighted with red crosses on the edges which are removed. We let $\mathcal{G}_3$ be the subgraph obtained from $\mathcal{G}_2$ by removing all the edges with red crosses.
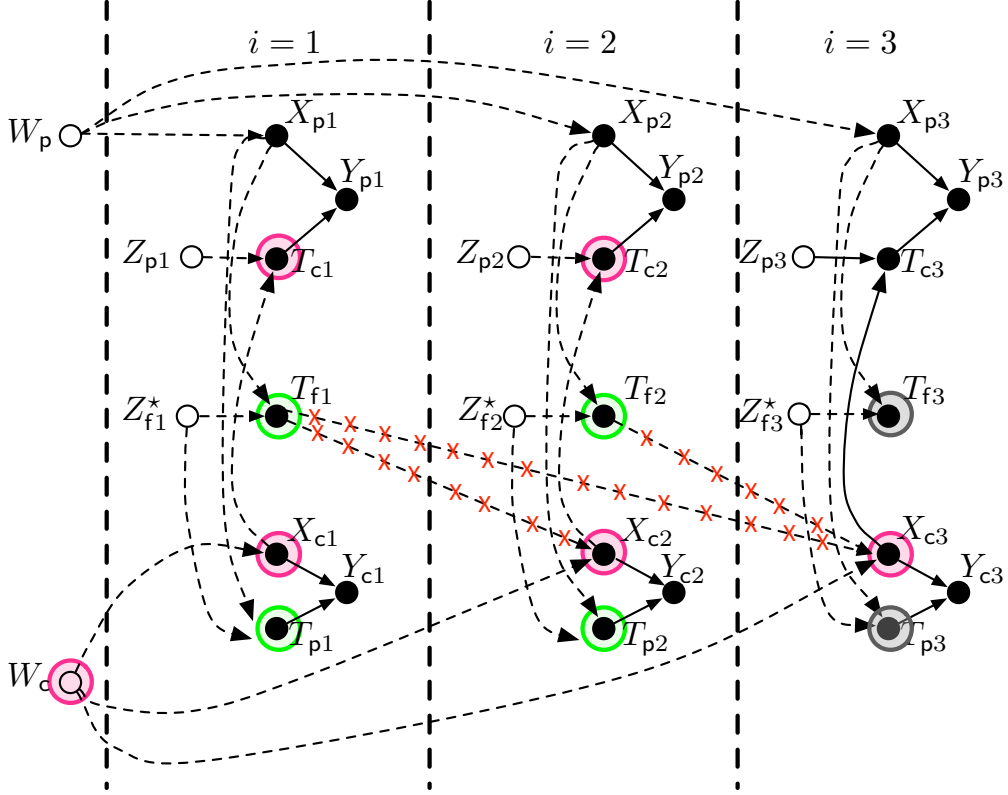
Fig. 6. Proof of the Markov chain in (8b) using the FDG.

5) From $\mathcal{G}_3$, remove all the arrows on the edges, and obtain the undirected subgraph $\mathcal{G}_4$. In $\mathcal{G}_4$ it is easy to see that, by starting from any grey node, it is not possible to reach any magenta node. This concludes the proof of the two Markov chains in (8a)-(8b).

APPENDIX B

PROOF OF THE SUM-RATE OUTER BOUND IN (6f)

By using the two Markov chains in (8a)-(8b) we can now derive the sum-rate outer bound in (6f). This bound was originally derived in [6] for the case of independent noises; here we extend it to the case when only the noises at the different source-destination pairs are independent, i.e., $\mathbb{P}_{Y_{\mathsf{Fc}}, Y_{\mathsf{c}}|X_{\mathsf{p}}, X_{\mathsf{c}}}$ in (2) is not a product distribution. By using Fano's inequality and by providing the same genie side information as in [6], we have

$$N(R_{\mathsf{p}} + R_{\mathsf{c}} - 2\epsilon_N)$$
$$\leq I\left(W_{\mathsf{p}}; Y_{\mathsf{p}}^N\right) + I\left(W_{\mathsf{c}}; Y_{\mathsf{c}}^N\right)$$

$$\leq I\left(W_{\mathsf{p}};Y_{\mathsf{p}}{}^{N},T_{\mathsf{p}}{}^{N},T_{\mathsf{f}}{}^{N}\right)+I\left(W_{\mathsf{c}};Y_{\mathsf{c}}{}^{N},T_{\mathsf{c}}{}^{N},T_{\mathsf{f}}{}^{N}\right)$$

$$=H\left(Y_{\mathsf{p}}{}^{N},T_{\mathsf{p}}{}^{N},T_{\mathsf{f}}{}^{N}\right)-H\left(Y_{\mathsf{c}}{}^{N},T_{\mathsf{c}}{}^{N},T_{\mathsf{f}}{}^{N}|W_{\mathsf{c}}\right)$$

$$+H\left(Y_{\mathsf{c}}{}^{N},T_{\mathsf{c}}{}^{N},T_{\mathsf{f}}{}^{N}\right)-H\left(Y_{\mathsf{p}}{}^{N},T_{\mathsf{p}}{}^{N},T_{\mathsf{f}}{}^{N}|W_{\mathsf{p}}\right).$$

We now analyze and bound the two pairs of terms. First pair:

$$H\left(Y_{\mathsf{p}}{}^{N},T_{\mathsf{p}}{}^{N},T_{\mathsf{f}}{}^{N}\right)-H\left(Y_{\mathsf{c}}{}^{N},T_{\mathsf{c}}{}^{N},T_{\mathsf{f}}{}^{N}|W_{\mathsf{c}}\right)$$

$$\overset{(a)}{=}\sum_{i\in[1:N]}H\left(Y_{\mathsf{p}i},T_{\mathsf{p}i},T_{\mathsf{f}i}|Y_{\mathsf{p}}{}^{i-1},T_{\mathsf{p}}{}^{i-1},T_{\mathsf{f}}{}^{i-1}\right)-H\left(Y_{\mathsf{c}i},T_{\mathsf{c}i},T_{\mathsf{f}i}|Y_{\mathsf{c}}{}^{i-1},T_{\mathsf{c}}{}^{i-1},T_{\mathsf{f}}{}^{i-1},W_{\mathsf{c}},X_{\mathsf{c}}{}^{i}\right)$$

$$\overset{(b)}{=}\sum_{i\in[1:N]}H\left(Y_{\mathsf{p}i},T_{\mathsf{p}i},T_{\mathsf{f}i}|Y_{\mathsf{p}}{}^{i-1},T_{\mathsf{p}}{}^{i-1},T_{\mathsf{f}}{}^{i-1}\right)-H\left(T_{\mathsf{p}i},T_{\mathsf{c}i},T_{\mathsf{f}i}|T_{\mathsf{p}}{}^{i-1},T_{\mathsf{c}}{}^{i-1},T_{\mathsf{f}}{}^{i-1},W_{\mathsf{c}},X_{\mathsf{c}}{}^{i}\right)$$

$$\overset{(c)}{\leq}\sum_{i\in[1:N]}\underbrace{H\left(T_{\mathsf{p}i},T_{\mathsf{f}i}|T_{\mathsf{p}}{}^{i-1},T_{\mathsf{f}}{}^{i-1}\right)-H\left(T_{\mathsf{p}i},T_{\mathsf{f}i}|T_{\mathsf{p}}{}^{i-1},T_{\mathsf{f}}{}^{i-1},W_{\mathsf{c}},T_{\mathsf{c}}{}^{i-1},X_{\mathsf{c}}{}^{i}\right)}_{=\,0\text{ because of (8b)}}$$

$$+\sum_{i\in[1:N]}H\left(Y_{\mathsf{p}i}|T_{\mathsf{p}i},T_{\mathsf{f}i}\right)-H\left(T_{\mathsf{c}i}|T_{\mathsf{p}}{}^{i},T_{\mathsf{f}}{}^{i},W_{\mathsf{c}},T_{\mathsf{c}}{}^{i-1},X_{\mathsf{c}}{}^{i},X_{\mathsf{p}}{}^{i}\right)$$

$$\overset{(d)}{=}\sum_{i\in[1:N]}H\left(Y_{\mathsf{p}i}|T_{\mathsf{p}i},T_{\mathsf{f}i}\right)-H\left(Y_{\mathsf{p}i}|T_{\mathsf{p}i},T_{\mathsf{f}i},X_{\mathsf{p}i},X_{\mathsf{c}i}\right),$$

where: the equality in (a) follows by applying the chain rule of the entropy and since, for the ISD CCIC, the encoding function $X_{\mathsf{c}i}(W_{\mathsf{c}},Y_{\mathsf{Fc}}^{i-1})$ is equivalent to $X_{\mathsf{c}i}(W_{\mathsf{c}},T_{\mathsf{f}}{}^{i-1})$; the equality in (b) is due to the fact that Y_{c} is a deterministic function of $(X_{\mathsf{c}},T_{\mathsf{p}})$, which is invertible given X_{c}; the inequality in (c) is due to the conditioning reduces entropy principle; the equality in (d) follows because of the ISD property of the channel and since the channel is memoryless. Second pair:

$$H\left(Y_{\mathsf{c}}{}^{N},T_{\mathsf{c}}{}^{N},T_{\mathsf{f}}{}^{N}\right)-H\left(Y_{\mathsf{p}}{}^{N},T_{\mathsf{p}}{}^{N},T_{\mathsf{f}}{}^{N}|W_{\mathsf{p}}\right)$$

$$\overset{(e)}{=}\sum_{i\in[1:N]}H\left(Y_{\mathsf{c}i},T_{\mathsf{c}i},T_{\mathsf{f}i}|Y_{\mathsf{c}}{}^{i-1},T_{\mathsf{c}}{}^{i-1},T_{\mathsf{f}}{}^{i-1}\right)-H\left(Y_{\mathsf{p}i},T_{\mathsf{p}i},T_{\mathsf{f}i}|Y_{\mathsf{p}}{}^{i-1},T_{\mathsf{p}}{}^{i-1},T_{\mathsf{f}}{}^{i-1},W_{\mathsf{p}},X_{\mathsf{p}}{}^{i}\right)$$

$$\overset{(f)}{=}\sum_{i\in[1:N]}H\left(Y_{\mathsf{c}i},T_{\mathsf{c}i},T_{\mathsf{f}i}|Y_{\mathsf{c}}{}^{i-1},T_{\mathsf{c}}{}^{i-1},T_{\mathsf{f}}{}^{i-1}\right)-H\left(T_{\mathsf{c}i},T_{\mathsf{p}i},T_{\mathsf{f}i}|T_{\mathsf{c}}{}^{i-1},T_{\mathsf{p}}{}^{i-1},T_{\mathsf{f}}{}^{i-1},W_{\mathsf{p}},X_{\mathsf{p}}{}^{i}\right)$$

$$\overset{(g)}{\leq}\sum_{i\in[1:N]}\underbrace{H\left(T_{\mathsf{c}i}|T_{\mathsf{c}}{}^{i-1},T_{\mathsf{f}}{}^{i-1}\right)-H\left(T_{\mathsf{c}i}|T_{\mathsf{c}}{}^{i-1},T_{\mathsf{f}}{}^{i-1},W_{\mathsf{p}},T_{\mathsf{p}}{}^{i-1},X_{\mathsf{p}}{}^{i}\right)}_{=\,0\text{ because of (8a)}}$$

$$+\sum_{i\in[1:N]}H\left(T_{\mathsf{f}i}|T_{\mathsf{c}i}\right)-H\left(T_{\mathsf{f}i}|T_{\mathsf{c}}{}^{i},T_{\mathsf{f}}{}^{i-1},W_{\mathsf{p}},T_{\mathsf{p}}{}^{i-1},X_{\mathsf{p}}{}^{i},X_{\mathsf{c}}{}^{i}\right)$$

$$+\sum_{i\in[1:N]}H\left(Y_{\mathsf{c}i}|T_{\mathsf{c}i},T_{\mathsf{f}i}\right)-H\left(T_{\mathsf{p}i}|T_{\mathsf{c}}{}^{i},T_{\mathsf{f}}{}^{i},W_{\mathsf{p}},T_{\mathsf{p}}{}^{i-1},X_{\mathsf{p}}{}^{i},X_{\mathsf{c}}{}^{i}\right)$$

$$\overset{\text{(h)}}{=} \sum_{i \in [1:N]} H\left(T_{\mathsf{f}i}|T_{\mathsf{c}i}\right) - H\left(T_{\mathsf{f}i}|T_{\mathsf{c}i}, X_{\mathsf{p}_i}, X_{\mathsf{c}i}\right)$$

$$+ \sum_{i \in [1:N]} H\left(Y_{\mathsf{c}i}|T_{\mathsf{c}i}, T_{\mathsf{f}i}\right) - H\left(Y_{\mathsf{c}i}|T_{\mathsf{c}i}, T_{\mathsf{f}i}, X_{\mathsf{p}_i}, X_{\mathsf{c}i}\right),$$

where: the equality in (e) follows by applying the chain rule of the entropy and since, given W_{p}, X_{p} is uniquely determined; the equality in (f) is due to the fact that Y_{p} is a deterministic function of $(X_{\mathsf{p}}, T_{\mathsf{c}})$, which is invertible given X_{p}; the inequality in (g) is due to the conditioning reduces entropy principle; the equality in (h) follows because of the ISD property of the channel and since the channel is memoryless.

By combining all the terms together, by introducing the time sharing random variable uniformly distributed over $[1:N]$ and independent of everything else, by dividing both sides by N and taking the limit for $N \to \infty$ we get the bound in (6f). We finally notice that by dropping the time sharing we do not decrease the bound.

APPENDIX C

PROOF OF THE OUTER BOUND IN (11)

By Fano's inequality, by considering that the messages W_{p} and W_{c} are independent and by giving side information similarly to [6], we have

$$N(R_{\mathsf{p}} + 2R_{\mathsf{c}} - 3\epsilon_N)$$

$$\leq I\left(W_{\mathsf{p}}; Y_{\mathsf{p}}^N\right) + 2I\left(W_{\mathsf{c}}; Y_{\mathsf{c}}^N\right)$$

$$\leq I\left(W_{\mathsf{p}}; Y_{\mathsf{p}}^N, T_{\mathsf{p}}^N, T_{\mathsf{f}}^N\right) + I\left(W_{\mathsf{c}}; Y_{\mathsf{c}}^N\right) + I\left(W_{\mathsf{c}}; Y_{\mathsf{c}}^N, T_{\mathsf{c}}^N, T_{\mathsf{f}}^N | W_{\mathsf{p}}\right)$$

$$\leq H\left(Y_{\mathsf{c}}^N\right) - H\left(Y_{\mathsf{c}}^N, T_{\mathsf{c}}^N, T_{\mathsf{f}}^N | W_{\mathsf{p}}, W_{\mathsf{c}}\right)$$

$$+ H\left(Y_{\mathsf{c}}^N, T_{\mathsf{c}}^N, T_{\mathsf{f}}^N | W_{\mathsf{p}}\right) - H\left(Y_{\mathsf{p}}^N, T_{\mathsf{p}}^N, T_{\mathsf{f}}^N | W_{\mathsf{p}}\right)$$

$$+ H\left(Y_{\mathsf{p}}^N, T_{\mathsf{p}}^N, T_{\mathsf{f}}^N\right) - H\left(Y_{\mathsf{c}}^N | W_{\mathsf{c}}\right).$$

We now analyze each pair of terms. In particular, we proceed similarly as we did to prove the outer bound $2R_{\mathsf{p}} + R_{\mathsf{c}}$ in (10).

First pair:

$$H\left(Y_{\mathsf{c}}^N\right) - H\left(Y_{\mathsf{c}}^N, T_{\mathsf{c}}^N, T_{\mathsf{f}}^N | W_{\mathsf{p}}, W_{\mathsf{c}}\right)$$

$$= \sum_{i \in [1:N]} H\left(Y_{\mathsf{c}i}|Y_{\mathsf{c}}^{i-1}\right) - H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i} | W_{\mathsf{p}}, W_{\mathsf{c}}, Y_{\mathsf{c}}^{i-1}, T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i-1}, X_{\mathsf{p}}^{i}, X_{\mathsf{c}}^{i}\right)$$

$$\leq \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i}\right) - H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i} | W_{\mathsf{p}}, W_{\mathsf{c}}, Y_{\mathsf{c}}^{i-1}, T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i-1}, X_{\mathsf{p}}^{\ i}, X_{\mathsf{c}}^{\ i}\right)$$

$$= \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i}\right) - H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i} | X_{\mathsf{p}_i}, X_{\mathsf{c}i}\right).$$

Second pair:

$$H\left(Y_{\mathsf{c}}^{\ N}, T_{\mathsf{c}}^{\ N}, T_{\mathsf{f}}^{\ N} | W_{\mathsf{p}}\right) - H\left(Y_{\mathsf{p}}^{\ N}, T_{\mathsf{p}}^{\ N}, T_{\mathsf{f}}^{\ N} | W_{\mathsf{p}}\right)$$

$$= \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i} | Y_{\mathsf{c}}^{i-1}, T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i-1}, W_{\mathsf{p}}, X_{\mathsf{p}}^{\ i}\right) - H\left(Y_{\mathsf{p}_i}, T_{\mathsf{p}_i}, T_{\mathsf{f}i} | Y_{\mathsf{p}}^{i-1}, T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}, W_{\mathsf{p}}, X_{\mathsf{p}}^{\ i}\right)$$

$$= \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i} | Y_{\mathsf{c}}^{i-1}, T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i-1}, W_{\mathsf{p}}, X_{\mathsf{p}}^{\ i}\right) - H\left(T_{\mathsf{c}i}, T_{\mathsf{p}_i}, T_{\mathsf{f}i} | T_{\mathsf{c}}^{i-1}, T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}, W_{\mathsf{p}}, X_{\mathsf{p}}^{\ i}\right)$$

$$\leq \sum_{i\in[1:N]} \underbrace{H\left(T_{\mathsf{c}i} | T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i-1}\right) - H\left(T_{\mathsf{c}i} | T_{\mathsf{c}}^{i-1}, T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}, W_{\mathsf{p}}, X_{\mathsf{p}}^{\ i}\right)}_{= 0 \text{ because of (8a)}}$$

$$+ \sum_{i\in[1:N]} H\left(T_{\mathsf{f}i} | T_{\mathsf{c}i}, X_{\mathsf{p}_i}\right) - H\left(T_{\mathsf{f}i} | T_{\mathsf{c}}^{\ i}, T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}, W_{\mathsf{p}}, X_{\mathsf{p}}^{\ i}\right)$$

$$+ \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i} | T_{\mathsf{c}i}, T_{\mathsf{f}i}, X_{\mathsf{p}_i}\right) - H\left(T_{\mathsf{p}_i} | T_{\mathsf{c}}^{\ i}, T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{\ i}, W_{\mathsf{p}}, X_{\mathsf{p}}^{\ i}, X_{\mathsf{c}}^{\ i}\right)$$

$$= \sum_{i\in[1:N]} H\left(T_{\mathsf{f}i} | T_{\mathsf{c}i}, X_{\mathsf{p}_i}\right) - H\left(T_{\mathsf{f}i} | T_{\mathsf{c}i}, X_{\mathsf{p}_i}\right)$$

$$+ \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i} | T_{\mathsf{c}i}, T_{\mathsf{f}i}, X_{\mathsf{p}_i}\right) - H\left(Y_{\mathsf{c}i} | T_{\mathsf{c}i}, T_{\mathsf{f}i}, X_{\mathsf{c}i}, X_{\mathsf{p}_i}\right)$$

$$= \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i} | T_{\mathsf{c}i}, T_{\mathsf{f}i}, X_{\mathsf{p}_i}\right) - H\left(Y_{\mathsf{c}i} | T_{\mathsf{c}i}, T_{\mathsf{f}i}, X_{\mathsf{c}i}, X_{\mathsf{p}_i}\right).$$

Third pair: since

$$H\left(Y_{\mathsf{c}}^{\ N} | W_{\mathsf{c}}\right)$$

$$= \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i} | Y_{\mathsf{c}}^{i-1}, W_{\mathsf{c}}\right)$$

$$\geq \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i} | Y_{\mathsf{c}}^{i-1}, W_{\mathsf{c}}, T_{\mathsf{f}}^{i-1}, X_{\mathsf{c}}^{\ i}\right)$$

$$= \sum_{i\in[1:N]} H\left(T_{\mathsf{p}_i} | T_{\mathsf{p}}^{i-1}, W_{\mathsf{c}}, T_{\mathsf{f}}^{i-1}, X_{\mathsf{c}}^{\ i}\right)$$

$$= \sum_{i\in[1:N]} H\left(T_{\mathsf{p}_i} | T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}\right),$$

where the last equality follows because of (8b), then

$$H\left(Y_{\mathsf{p}}^{\ N}, T_{\mathsf{p}}^{\ N}, T_{\mathsf{f}}^{\ N}\right) - H\left(Y_{\mathsf{c}}^{\ N} | W_{\mathsf{c}}\right)$$

$$\leq \sum_{i \in [1:N]} H\left(Y_{\mathsf{p}i}, T_{\mathsf{p}i}, T_{\mathsf{f}i} | Y_{\mathsf{p}}^{i-1}, T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}\right) - H\left(T_{\mathsf{p}i} | T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}\right)$$

$$\leq \sum_{i \in [1:N]} H\left(Y_{\mathsf{p}i}, T_{\mathsf{f}i} | T_{\mathsf{p}i}\right).$$

By combining everything together, by introducing the time sharing random variable uniformly distributed over $[1 : N]$ and independent of everything else, by dividing both sides by N and taking the limit for $N \to \infty$ we get the bound in (11). We finally notice that by dropping the time sharing we do not decrease the bound.

APPENDIX D
EVALUATION OF THE OUTER BOUNDS IN (6), (10) AND (11) FOR THE GCCIC

By defining $\mathbb{E}\left[X_{\mathsf{p}} X_{\mathsf{c}}^*\right] := \rho : |\rho| \in [0, 1]$ we obtain: from the cut-set bounds in (6a)-(6c)

$$R_{\mathsf{p}} \leq \log\left(1 + (\mathsf{C} + \mathsf{S}_{\mathsf{p}})\left(1 - |\rho|^2\right)\right)$$
$$\overset{|\rho|=0}{\leq} \log\left(1 + \mathsf{C} + \mathsf{S}_{\mathsf{p}}\right), \tag{25a}$$

$$R_{\mathsf{p}} \leq \log\left(1 + \mathsf{S}_{\mathsf{p}} + \mathsf{I}_{\mathsf{c}} + 2\sqrt{\mathsf{S}_{\mathsf{p}}\mathsf{I}_{\mathsf{c}}}\,\Re\left\{\rho e^{-j\theta_{\mathsf{c}}}\right\}\right)$$
$$\overset{\rho=e^{j\theta_{\mathsf{c}}}}{\leq} \log\left(1 + \left(\sqrt{\mathsf{S}_{\mathsf{p}}} + \sqrt{\mathsf{I}_{\mathsf{c}}}\right)^2\right), \tag{25b}$$

$$R_{\mathsf{c}} \leq \log\left(1 + \left(1 - |\rho|^2\right)\mathsf{S}_{\mathsf{c}}\right)$$
$$\overset{|\rho|=0}{\leq} \log\left(1 + \mathsf{S}_{\mathsf{c}}\right). \tag{25c}$$

From the bounds in (6d)-(6e) we get

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1 + \frac{(\mathsf{S}_{\mathsf{p}} + \mathsf{C})\left(1 - |\rho|^2\right)}{1 + \mathsf{I}_{\mathsf{p}}\left(1 - |\rho|^2\right)}\right) + \log\left(1 + \mathsf{S}_{\mathsf{c}} + \mathsf{I}_{\mathsf{p}} + 2\sqrt{\mathsf{S}_{\mathsf{c}}\mathsf{I}_{\mathsf{p}}}\,\Re\left\{\rho e^{j\theta_{\mathsf{p}}}\right\}\right)$$
$$\overset{(a)}{\leq} \log\left(1 + \frac{\mathsf{S}_{\mathsf{p}} + \mathsf{C}}{1 + \mathsf{I}_{\mathsf{p}}}\right) + \log\left(1 + \left(\sqrt{\mathsf{S}_{\mathsf{c}}} + \sqrt{\mathsf{I}_{\mathsf{p}}}\right)^2\right) \tag{25d}$$
$$= \underbrace{\log\left(1 + \frac{\mathsf{S}_{\mathsf{p}}}{1 + \mathsf{I}_{\mathsf{p}}}\right) + \log\left(1 + \left(\sqrt{\mathsf{S}_{\mathsf{c}}} + \sqrt{\mathsf{I}_{\mathsf{p}}}\right)^2\right)}_{\text{as no cooperation / } \mathsf{C} = 0}$$
$$+ \underbrace{\log\left(1 + \frac{\mathsf{C}}{1 + \mathsf{I}_{\mathsf{p}} + \mathsf{S}_{\mathsf{p}}}\right)}_{\text{increasing in } \mathsf{C}},$$

and

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1 + \frac{\mathsf{S}_{\mathsf{c}}\left(1 - |\rho|^2\right)}{1 + \mathsf{I}_{\mathsf{c}}\left(1 - |\rho|^2\right)}\right) + \log\left(1 + \mathsf{S}_{\mathsf{p}} + \mathsf{I}_{\mathsf{c}} + 2\sqrt{\mathsf{S}_{\mathsf{p}}\mathsf{I}_{\mathsf{c}}}\,\Re\left\{\rho e^{-j\theta_{\mathsf{c}}}\right\}\right)$$

$$\overset{(b)}{\leq} \underbrace{\log\left(1 + \frac{\mathsf{S_c}}{1 + \mathsf{I_c}}\right) + \log\left(1 + \left(\sqrt{\mathsf{S_p}} + \sqrt{\mathsf{I_c}}\right)^2\right)}_{\text{as no cooperation / } \mathsf{C} = 0}, \tag{25e}$$

where the inequality in (a) follows by evaluating the first logarithm in $|\rho| = 0$ and the second logarithm in $\rho = \mathrm{e}^{-\mathrm{j}\theta_\mathsf{p}}$ and the inequality in (b) follows by evaluating the first logarithm in $|\rho| = 0$ and the second logarithm in $\rho = \mathrm{e}^{\mathrm{j}\theta_\mathsf{c}}$. Finally, from the bound in (6f) we obtain

$$R_\mathsf{p} + R_\mathsf{c} \leq \log\left(1 + \frac{\mathsf{S_p} + \mathsf{I_c} + 2\sqrt{\mathsf{S_p I_c}}\mathfrak{R}\left\{\rho\mathrm{e}^{-\mathrm{j}\theta_\mathsf{c}}\right\} + (1 - |\rho|^2)\left(\mathsf{I_p I_c} + \mathsf{C I_c}\right)}{1 + \mathsf{C} + \mathsf{I_p}}\right)$$

$$+ \log\left(1 + \frac{\mathsf{S_c} + \mathsf{I_p} + 2\sqrt{\mathsf{S_c I_p}}\mathfrak{R}\left\{\rho\mathrm{e}^{\mathrm{j}\theta_\mathsf{p}}\right\} + (1 - |\rho|^2)\left(\mathsf{I_p I_c} + \mathsf{C S_c}\right)}{1 + \mathsf{C} + \mathsf{I_c} + \mathsf{I_c C}(1 - |\rho|^2)}\right)$$

$$+ \log\left(1 + \frac{\mathsf{C} + \mathsf{C I_c}(1 - |\rho|^2)}{1 + \mathsf{I_c}}\right)$$

$$= \log\left(1 + \frac{\mathsf{S_p} + \mathsf{I_c} + 2\sqrt{\mathsf{S_p I_c}}\mathfrak{R}\left\{\rho\mathrm{e}^{-\mathrm{j}\theta_\mathsf{c}}\right\} + (1 - |\rho|^2)\left(\mathsf{I_p I_c} + \mathsf{C I_c}\right)}{1 + \mathsf{C} + \mathsf{I_p}}\right)$$

$$+ \log\left(1 + \frac{\mathsf{C} + \mathsf{S_c} + \mathsf{I_p} + 2\sqrt{\mathsf{S_c I_p}}\mathfrak{R}\left\{\rho\mathrm{e}^{\mathrm{j}\theta_\mathsf{p}}\right\} + (1 - |\rho|^2)\left(\mathsf{I_p I_c} + \mathsf{C S_c} + \mathsf{I_c C}\right)}{1 + \mathsf{I_c}}\right)$$

$$\overset{(c)}{\leq} \log\left(1 + \mathsf{I_c} + \frac{\mathsf{S_p}}{1 + \mathsf{C} + \mathsf{I_p}}\right) + \log\left(1 + \mathsf{C} + \mathsf{I_p} + \frac{\mathsf{S_c}(1 + \mathsf{C})}{1 + \mathsf{I_c}}\right) + 2\log(2), \tag{25f}$$

$$= \underbrace{\log\left(1 + \mathsf{C} + \mathsf{I_p} + \frac{\mathsf{S_p}}{1 + \mathsf{I_c}}\right) + \log\left(1 + \mathsf{I_c} + \mathsf{S_c}\frac{1 + \mathsf{C}}{1 + \mathsf{C} + \mathsf{I_p}}\right)}_{\text{increasing in } \mathsf{C}} + 2\log(2),$$

where the inequality in (c) follows by: (i) evaluating the first term of the first logarithm in $\rho = \mathrm{e}^{\mathrm{j}\theta_\mathsf{c}}$ and the second term of the first logarithm in $|\rho| = 0$; (ii) evaluating the first term of the second logarithm in $\rho = \mathrm{e}^{-\mathrm{j}\theta_\mathsf{p}}$ and the second term of the second logarithm in $|\rho| = 0$; (iii) since $\log\left(1 + \left(\sqrt{|a|^2} + \sqrt{|b|^2}\right)^2\right) \leq \log\left(1 + |a|^2 + |b|^2\right) + \log(2)$.

We now evaluate the new outer bounds in Theorem 1 and we get

$$2R_\mathsf{p} + R_\mathsf{c} \leq \log\left(1 + \mathsf{S_p} + \mathsf{I_c} + 2\sqrt{\mathsf{S_p I_c}}\mathfrak{R}\left\{\rho\mathrm{e}^{-\mathrm{j}\theta_\mathsf{c}}\right\}\right) + \log\left(1 + \frac{\mathsf{S_p}(1 - |\rho|^2)}{1 + (\mathsf{C} + \mathsf{I_p})(1 - |\rho|^2)}\right)$$

$$+ \log\left(1 + \frac{\mathsf{C} + \mathsf{S_c} + \mathsf{I_p} + 2\sqrt{\mathsf{S_c I_p}}\mathfrak{R}\left\{\rho\mathrm{e}^{\mathrm{j}\theta_\mathsf{p}}\right\} + (1 - |\rho|^2)\left(\mathsf{I_p I_c} + \mathsf{C S_c} + \mathsf{I_c C}\right)}{1 + \mathsf{I_c}}\right)$$

$$\overset{(d)}{\leq} \log\left(1 + \left(\sqrt{\mathsf{S_p}} + \sqrt{\mathsf{I_c}}\right)^2\right) + \log\left(1 + \frac{\mathsf{S_p}}{1 + \mathsf{I_p} + \mathsf{C}}\right)$$

$$+ \log\left(1 + \mathsf{C} + \mathsf{I_p} + \frac{\mathsf{S_c}(1 + \mathsf{C})}{1 + \mathsf{I_c}}\right) + \log(2), \tag{25g}$$

$$= \underbrace{\log\left(1 + \frac{S_c}{1 + I_c}\right) + \log\left(1 + \left(\sqrt{S_p} + \sqrt{I_c}\right)^2\right)}_{\text{as no cooperation / } C = 0}$$

$$+ \underbrace{\log\left(\frac{1 + I_p + S_p}{1 + I_c + S_c}\right) + \log\left(1 + \frac{C}{1 + I_p + S_p}\right) + \log\left(1 + I_c + S_c\frac{1 + C}{1 + I_p + C}\right)}_{\text{increasing in } C} + \log(2),$$

where the inequality in (d) follows by (i) evaluating the first logarithm in $\rho = e^{j\theta_c}$, (ii) evaluating the second logarithm in $|\rho| = 0$, (iii) evaluating the first term of the third logarithm in $\rho = e^{-j\theta_p}$ and the second term of the third logarithm in $|\rho| = 0$ and (iv) since $\log\left(1 + \left(\sqrt{|a|^2} + \sqrt{|b|^2}\right)^2\right) \leq \log\left(1 + |a|^2 + |b|^2\right) + \log(2)$. Similarly,

$$R_p + 2R_c \leq \log\left(1 + S_c + I_p + 2\sqrt{S_c I_p}\Re\left\{\rho e^{j\theta_p}\right\}\right) + \log\left(1 + \frac{S_c\left(1 - |\rho|^2\right)}{1 + I_c\left(1 - |\rho|^2\right)}\right)$$

$$+ \log\left(1 + \frac{S_p + I_c + 2\sqrt{S_p I_c}\Re\left\{\rho e^{-j\theta_c}\right\} + \left(1 - |\rho|^2\right)\left(I_p I_c + C I_c\right)}{1 + C + I_p}\right)$$

$$+ \log\left(1 + \frac{C}{1 + I_p}\right)$$

$$\overset{(e)}{\leq} \log\left(1 + \left(\sqrt{S_c} + \sqrt{I_p}\right)^2\right) + \log\left(1 + \frac{S_c}{1 + I_c}\right)$$

$$+ \log\left(1 + I_c + \frac{S_p}{1 + C + I_p}\right) + \log\left(1 + \frac{C}{1 + I_p}\right) + \log(2), \tag{25h}$$

$$= \underbrace{\log\left(1 + \frac{S_p}{1 + I_p}\right) + \log\left(1 + \left(\sqrt{S_c} + \sqrt{I_p}\right)^2\right)}_{\text{as no cooperation / } C = 0}$$

$$+ \underbrace{\log\left(\frac{1 + I_c + S_c}{1 + I_p + S_p}\right) + \log\left(1 + C + I_p + \frac{S_p}{1 + I_c}\right)}_{\text{increasing in } C} + \log(2),$$

where the inequality in (e) follows by: (i) evaluating the first logarithm in $\rho = e^{-j\theta_p}$, (ii) evaluating the second logarithm in $|\rho| = 0$ and (iii) evaluating the first term of the third logarithm in $\rho = e^{j\theta_c}$ and the second term of the third logarithm in $|\rho| = 0$ and again (iv) since $\log\left(1 + \left(\sqrt{|a|^2} + \sqrt{|b|^2}\right)^2\right) \leq \log\left(1 + |a|^2 + |b|^2\right) + \log(2)$.

APPENDIX E

ACHIEVABLE SCHEME BASED ON SUPERPOSITION CODING AND BINNING

We specialize the 'binning+superposition' achievable scheme in [8, Section V]. In [8, Thereom V.1] the network comprises four nodes numbered from 1 to 4; nodes 1 and 2 are sources and
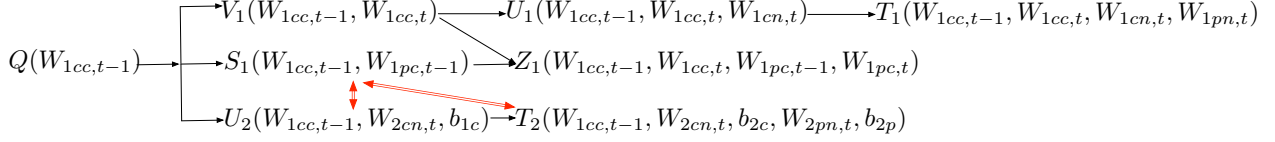
$$V_1(W_{1cc,t-1}, W_{1cc,t}) \longrightarrow U_1(W_{1cc,t-1}, W_{1cc,t}, W_{1cn,t}) \longrightarrow T_1(W_{1cc,t-1}, W_{1cc,t}, W_{1cn,t}, W_{1pn,t})$$

$$Q(W_{1cc,t-1}) \longrightarrow S_1(W_{1cc,t-1}, W_{1pc,t-1}) \longrightarrow Z_1(W_{1cc,t-1}, W_{1cc,t}, W_{1pc,t-1}, W_{1pc,t})$$

$$U_2(W_{1cc,t-1}, W_{2cn,t}, b_{1c}) \longrightarrow T_2(W_{1cc,t-1}, W_{2cn,t}, b_{2c}, W_{2pn,t}, b_{2p})$$

Fig. 7. Achievable scheme based on binning and superposition coding.

nodes 3 and 4 are destinations; source node $j \in [1:2]$, with input to the channel X_j and output from the channel Y_j, has a message W_j for node $j+2$; destination node $j \in [3:4]$ has channel output Y_j from which it decodes the message W_{j-2}. Both users use rate splitting, where the messages of user 1 / primary are both non-cooperative and cooperative, while the messages of user 2 / cognitive are non-cooperative. In [8, Section V], we set $Y_1 = S_2 = V_2 = Z_2 = \emptyset$, i.e., then $R_1 = R_{11c} + R_{10c} + R_{10n} + R_{11n}$, $R_2 = R_{22n} + R_{20n}$, to obtain a scheme that comprises: a cooperative common message (carried by the pair (Q, V_1) at rate R_{10c}) for user 1, a cooperative private message (carried by the pair (S_1, Z_1) at rate R_{11c}) for user 1, a non-cooperative common message (carried by U_1 at rate R_{10n}) for user 1, a non-cooperative private message (carried by T_1 at rate R_{11n}) for user 1, a non-cooperative common message (carried by U_2 at rate R_{20n}) for user 2 and a non-cooperative private message (carried by T_2 at rate R_{22n}) for user 2. Here the pair (Q, S_1) carries the 'past cooperative messages', and the pair (V_1, Z_1) the 'new cooperative messages' in a block Markov encoding scheme. The channel inputs are functions of the auxiliary random variables, where X_1 is a function of $(Q, S_1, Z_1, V_1, U_1, T_1)$ and X_2 is a function of (Q, S_1, U_2, T_2).

a) Input distributions: The set of possible input distributions is

$$P_{Q,S_1,V_1,Z_1,U_1,T_1,X_1,U_2,T_2,X_2}$$

$$= P_Q P_{V_1|Q} P_{U_1,T_1|Q,V_1} P_{S_1|Q} P_{Z_1|Q,S_1,V_1} P_{U_2,T_2|S_1,Q} P_{X_1|Q,S_1,Z_1,V_1,U_1,T_1} P_{X_2|Q,S_1,U_2,T_2}. \tag{26}$$

A schematic representation of the achievable scheme is given in Fig. 7, where a black arrow indicates superposition coding and a red arrow indicates binning. Note that the difference between Fig. 7 and [17, Fig. 10] is that in the achievable scheme in [17, Fig. 10] we set $X_1 = T_1$, $X_2 = T_2$ and $U_1 = S_1 = Z_1 = \emptyset$, i.e., user 1 / primary only used a common cooperative message and a private non-cooperative message, while here also private cooperative as well as common non-cooperative messages are used. Similarly, the difference between Fig. 7 and [17, Fig. 11] is that in the achievable scheme in [17, Fig. 11] we set $U_1 = T_1 = \emptyset$, i.e., the messages of user 1 /

primary were only cooperative, while here we consider also non-cooperative (both private and common) messages for user 1 / primary.

b) Encoding: The codebooks are generated as follows: first the codebook Q is generated; then the codebook V_1 is superposed to Q, after the codebook U_1 is superposed to (Q, V_1) and finally the codebook T_1 is superposed to (Q, V_1, U_1); independently of (V_1, U_1, T_1), the codebook S_1 is superposed to Q and then the codebook Z_1 is superposed to (Q, S_1, V_1); independently of $(S_1, Z_1, V_1, U_1, T_1)$, the codebook U_2 is superposed to Q and then the codebook T_2 is superposed to (Q, U_2). With this random coding codebook generation, the pair (U_2, T_2) is independent of S_1 conditioned on Q. [8, Theorem V.1] involves several binning steps to allow for a large set of input distributions. Here the only binning steps are for (U_2, T_2) against S_1.

We use a block Markov coding scheme to convey the message of user 1 to user 2. In particular, at the end of any given time slot in a block Markov coding scheme, encoder 2 knows (Q, S_1, U_2, T_2) and decodes (V_1, Z_1) from its channel output; the decoded pair (V_1, Z_1) becomes the pair (Q, S_1) of the next time slot; then, at the beginning of each time slot, encoder 2, by binning, finds the new pair (U_2, T_2) that is jointly typical with (Q, S_1); for this to be possible, we must generate several (U_2, T_2) sequences for each message of user 2 so as to be able to find one pair to send with the correct joint distribution with (Q, S_1); this entails the rate penalties in [8, eq(20)] for user 1 and then again [8, eq(20)] for user 2 by swapping the role of the subscripts 1 and 2, with $S_2 = Z_2 = V_2 = \emptyset$, i.e.,

$$R'_{20n} + R'_{22n} \geq I(U_2, T_2; S_1|Q), \tag{27a}$$

$$R'_{20n} \geq I(U_2; S_1|Q). \tag{27b}$$

c) Decoding: The cooperative source uses the partial-decode-and-forward strategy and the destinations backward decoding. There are three decoding nodes in the network and therefore three groups of rate constraints. These are:

- Node2 / CTx jointly decodes (V_1, Z_1) from its channel output with knowledge of the indices in (Q, S_1, U_2, T_2, X_2). Successful decoding is possible if (use [8, eq(21)] by swapping the role of the subscripts 1 and 2, with $S_2 = Z_2 = V_2 = \emptyset$ and with V_1 independent of S_1)

$$R_{10c} + R_{11c} \leq I(Y_2; Z_1, V_1|U_2, T_2, X_2, S_1, Q), \tag{27c}$$

$$R_{11c} \leq I(Y_2; Z_1|U_2, T_2, X_2, S_1, Q, V_1). \tag{27d}$$

- Node3 / PRx jointly decodes (Q, S_1, U_2, U_1, T_1) from its channel output, with knowledge of some message indices in (V_1, Z_1), by treating T_2 as noise. Successful decoding is possible if (see [8, eq(22)] with $S_2 = Z_2 = V_2 = \emptyset$)

$$R_{10c} + R_{10n} + R_{11n} + R_{20n} + R_{11c} \leq I(Y_3; Q, V_1, U_1, T_1, S_1, Z_1, U_2)$$
$$- (R'_{20n} - I(U_2; S_1|Q)), \tag{27e}$$

$$R_{10n} + R_{11n} + R_{20n} + R_{11c} \leq I(Y_3; U_1, T_1, S_1, Z_1, U_2|Q, V_1)$$
$$- (R'_{20n} - I(U_2; S_1|Q)), \tag{27f}$$

$$R_{10n} + R_{11n} + R_{11c} \leq I(Y_3; U_1, T_1, S_1, Z_1|Q, V_1, U_2), \tag{27g}$$

$$R_{11n} + R_{20n} + R_{11c} \leq I(Y_3; T_1, S_1, Z_1, U_2|Q, V_1, U_1) - (R'_{20n} - I(U_2; S_1|Q)), \tag{27h}$$

$$R_{10n} + R_{11n} + R_{20n} \leq I(Y_3; U_1, T_1, U_2|Q, S_1, Z_1, V_1) - (R'_{20n} - I(U_2; S_1|Q)), \tag{27i}$$

$$R_{11n} + R_{11c} \leq I(Y_3; T_1, S_1, Z_1|Q, V_1, U_1, U_2), \tag{27j}$$

$$R_{20n} + R_{11c} \leq I(Y_3; S_1, Z_1, U_2|Q, V_1, U_1, T_1) - (R'_{20n} - I(U_2; S_1|Q)), \tag{27k}$$

$$R_{10n} + R_{11n} \leq I(Y_3; U_1, T_1|Q, S_1, Z_1, V_1, U_2), \tag{27l}$$

$$R_{11n} + R_{20n} \leq I(Y_3; T_1, U_2|Q, S_1, Z_1, V_1, U_1) - (R'_{20n} - I(U_2; S_1|Q)), \tag{27m}$$

$$R_{11c} \leq I(Y_3; S_1, Z_1|Q, V_1, U_1, T_1, U_2), \tag{27n}$$

$$R_{11n} \leq I(Y_3; T_1|Q, S_1, Z_1, V_1, U_1, U_2). \tag{27o}$$

- Node4 / CRx jointly decodes (Q, U_1, U_2, T_2) from its channel output, with knowledge of some message index in V_1, by treating Z_1 and T_1 as noise (recall that the pair (U_2, T_2) has been precoded/binned against S_1). Successful decoding is possible if (see [8, eq(22)], with the role of the users swapped, where only the bounds in [8, eq(22a)], [8, eq(22h)], [8, eq(22i)], [8, eq(22j)], and [8, eq(22k)] remain after setting several auxiliary random variables to zero and removing the redundant constraints)

$$R_{10c} + R_{20n} + R_{22n} + R_{10n} \leq I(Y_4; Q, U_2, T_2, V_1, U_1) - (R'_{20n} + R'_{22n}), \tag{27p}$$

$$R_{20n} + R_{22n} + R_{10n} \leq I(Y_4; U_2, T_2, U_1|Q, V_1) - (R'_{20n} + R'_{22n}), \tag{27q}$$

$$R_{20n} + R_{22n} \leq I(Y_4; U_2, T_2|Q, V_1, U_1) - (R'_{20n} + R'_{22n}), \tag{27r}$$

$$R_{22n} + R_{10n} \leq I(Y_4; T_2, U_1|Q, U_2, V_1) - R'_{22n}, \tag{27s}$$

$$R_{22n} \leq I(Y_4; T_2|Q, U_2, V_1, U_1) - R'_{22n}. \tag{27t}$$

d) Compact region: Instead of applying the Fourier-Motzkin Elimination (FME) directly on the general achievable rate region, in the following we apply the FME on two particular cases, namely the case when $S_1 = Z_1 = \emptyset$, i.e., no private cooperative messages for the PTx and the case $U_1 = \emptyset$, i.e., no common non-cooperative message for the PTx. The first case is analyzed in Section E-A, while the second case is analyzed in Section E-B. For both cases we take the constraints in (27a) and (27b) to hold with equality (i.e., $R'_{20n} = I(U_2; S_1|Q)$, $R'_{22n} = I(S_1; T_2|Q, U_2)$).

A. FME on the achievable rate region when $S_1 = Z_1 = \emptyset$

We set $S_1 = Z_1 = \emptyset$ in the achievable rate region. After FME of the achievable region in (27) with $S_1 = Z_1 = \emptyset$ (see also [8, eq(8)]), we get

$$R_1 \leq \text{eq(27e)}, \tag{28a}$$

$$R_1 \leq \text{eq(27c)} + \text{eq(27g)}, \tag{28b}$$

$$R_2 \leq \text{eq(27r)}, \tag{28c}$$

$$R_1 + R_2 \leq \text{eq(27e)} + \text{eq(27t)}, \tag{28d}$$

$$R_1 + R_2 \leq \text{eq(27j)} + \text{eq(27p)}, \tag{28e}$$

$$R_1 + R_2 \leq \text{eq(27c)} + \text{eq(27i)} + \text{eq(27t)}, \tag{28f}$$

$$R_1 + R_2 \leq \text{eq(27c)} + \text{eq(27j)} + \text{eq(27q)}, \tag{28g}$$

$$R_1 + R_2 \leq \text{eq(27c)} + \text{eq(27m)} + \text{eq(27s)}, \tag{28h}$$

$$2R_1 + R_2 \leq \text{eq(27c)} + \text{eq(27j)} + \text{eq(27e)} + \text{eq(27s)}, \tag{28i}$$

$$2R_1 + R_2 \leq 2 \cdot \text{eq(27c)} + \text{eq(27j)} + \text{eq(27i)} + \text{eq(27s)}, \tag{28j}$$

$$R_1 + 2R_2 \leq \text{eq(27m)} + \text{eq(27s)} + \text{eq(27p)}, \tag{28k}$$

$$R_1 + 2R_2 \leq \text{eq(27c)} + \text{eq(27m)} + \text{eq(27t)} + \text{eq(27q)}, \tag{28l}$$

for all distributions that factor as (26) and by setting $S_1 = Z_1 = \emptyset$ in all the mutual information terms.

We identify Node1 with the PTx (i.e., $X_{\mathsf{p}} = X_1$), Node2 with the CTx (i.e., $X_{\mathsf{c}} = X_2, T_{\mathsf{f}} = Y_2$), Node3 with the PRx (i.e., $Y_{\mathsf{p}} = Y_3$) and Node4 with the CRx (i.e., $Y_{\mathsf{c}} = Y_4$). For the Gaussian

noise channel, in the achievable region in (28), we choose $Q = \emptyset$, we let V_1, U_1, T_1, U_2, T_2 be i.i.d. $\mathcal{N}(0, 1)$, and

$$X_\mathsf{p} = a_1 U_1 + b_1 V_1 + c_1 T_1 : \qquad\qquad |a_1|^2 + |b_1|^2 + |c_1|^2 = 1,$$

$$X_\mathsf{c} = a_2 U_2 + b_2 T_2 : \qquad\qquad |a_2|^2 + |b_2|^2 = 1.$$

With these choices, the channel outputs are

$$T_\mathsf{f} = \sqrt{\mathsf{C}}\,(a_1 U_1 + b_1 V_1 + c_1 T_1) + Z_\mathsf{f},$$

$$Y_\mathsf{p} = \sqrt{\mathsf{S_p}}\,(a_1 U_1 + b_1 V_1 + c_1 T_1) + \sqrt{\mathsf{I_c}}\mathsf{e}^{\mathsf{j}\theta_\mathsf{c}}\,(a_2 U_2 + b_2 T_2) + Z_\mathsf{p},$$

$$Y_\mathsf{c} = \sqrt{\mathsf{I_p}}\mathsf{e}^{\mathsf{j}\theta_\mathsf{p}}\,(a_1 U_1 + b_1 V_1 + c_1 T_1) + \sqrt{\mathsf{S_c}}\,(a_2 U_2 + b_2 T_2) + Z_\mathsf{c},$$

and the achievable region in (28) becomes

$$R_\mathsf{p} \leq \log\left(\frac{1 + \mathsf{S_p} + \mathsf{I_c}}{1 + \mathsf{I_c}|b_2|^2}\right), \tag{29a}$$

$$R_\mathsf{p} \leq \log\left(\frac{1 + \mathsf{C}}{1 + \mathsf{C}\,(|a_1|^2 + |c_1|^2)}\right) + \log\left(1 + \frac{\mathsf{S_p}\,(|a_1|^2 + |c_1|^2)}{1 + \mathsf{I_c}|b_2|^2}\right), \tag{29b}$$

$$R_\mathsf{c} \leq \log\left(1 + \frac{\mathsf{S_c}}{1 + \mathsf{I_p}|c_1|^2}\right), \tag{29c}$$

$$R_\mathsf{p} + R_\mathsf{c} \leq \log\left(\frac{1 + \mathsf{S_p} + \mathsf{I_c}}{1 + \mathsf{I_c}|b_2|^2}\right) + \log\left(1 + \frac{\mathsf{S_c}|b_2|^2}{1 + \mathsf{I_p}|c_1|^2}\right), \tag{29d}$$

$$R_\mathsf{p} + R_\mathsf{c} \leq \log\left(1 + \frac{\mathsf{S_p}|c_1|^2}{1 + \mathsf{I_c}|b_2|^2}\right) + \log\left(\frac{1 + \mathsf{S_c} + \mathsf{I_p}}{1 + \mathsf{I_p}|c_1|^2}\right), \tag{29e}$$

$$R_\mathsf{p} + R_\mathsf{c} \leq \log\left(\frac{1 + \mathsf{C}}{1 + \mathsf{C}\,(|a_1|^2 + |c_1|^2)}\right) + \log\left(\frac{1 + \mathsf{S_p}\,(|a_1|^2 + |c_1|^2) + \mathsf{I_c}}{1 + \mathsf{I_c}|b_2|^2}\right)$$
$$+ \log\left(1 + \frac{\mathsf{S_c}|b_2|^2}{1 + \mathsf{I_p}|c_1|^2}\right), \tag{29f}$$

$$R_\mathsf{p} + R_\mathsf{c} \leq \log\left(\frac{1 + \mathsf{C}}{1 + \mathsf{C}\,(|a_1|^2 + |c_1|^2)}\right) + \log\left(1 + \frac{\mathsf{S_p}|c_1|^2}{1 + \mathsf{I_c}|b_2|^2}\right) + \log\left(1 + \frac{\mathsf{S_c} + \mathsf{I_p}|a_1|^2}{1 + \mathsf{I_p}|c_1|^2}\right), \tag{29g}$$

$$R_\mathsf{p} + R_\mathsf{c} \leq \log\left(\frac{1 + \mathsf{C}}{1 + \mathsf{C}\,(|a_1|^2 + |c_1|^2)}\right) + \log\left(\frac{1 + \mathsf{S_p}|c_1|^2 + \mathsf{I_c}}{1 + \mathsf{I_c}|b_2|^2}\right)$$
$$+ \log\left(1 + \frac{\mathsf{I_p}|a_1|^2 + \mathsf{S_c}|b_2|^2}{1 + \mathsf{I_p}|c_1|^2}\right), \tag{29h}$$

$$2R_\mathsf{p} + R_\mathsf{c} \leq \log\left(\frac{1 + \mathsf{C}}{1 + \mathsf{C}\,(|a_1|^2 + |c_1|^2)}\right) + \log\left(1 + \frac{\mathsf{S_p}|c_1|^2}{1 + \mathsf{I_c}|b_2|^2}\right) + \log\left(\frac{1 + \mathsf{S_p} + \mathsf{I_c}}{1 + \mathsf{I_c}|b_2|^2}\right)$$
$$+ \log\left(1 + \frac{\mathsf{I_p}|a_1|^2 + \mathsf{S_c}|b_2|^2}{1 + \mathsf{I_p}|c_1|^2}\right), \tag{29i}$$

$$2R_{\mathsf{p}} + R_{\mathsf{c}} \leq 2 \cdot \log\left(\frac{1+\mathsf{C}}{1+\mathsf{C}\left(|a_1|^2+|c_1|^2\right)}\right) + \log\left(1+\frac{\mathsf{S_p}|c_1|^2}{1+\mathsf{I_c}|b_2|^2}\right)$$

$$+ \log\left(\frac{1+\mathsf{S_p}\left(|a_1|^2+|c_1|^2\right)+\mathsf{I_c}}{1+\mathsf{I_c}|b_2|^2}\right) + \log\left(1+\frac{\mathsf{I_p}|a_1|^2+\mathsf{S_c}|b_2|^2}{1+\mathsf{I_p}|c_1|^2}\right), \tag{29j}$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq \log\left(\frac{1+\mathsf{S_p}|c_1|^2+\mathsf{I_c}}{1+\mathsf{I_c}|b_2|^2}\right) + \log\left(1+\frac{\mathsf{I_p}|a_1|^2+\mathsf{S_c}|b_2|^2}{1+\mathsf{I_p}|c_1|^2}\right) + \log\left(\frac{1+\mathsf{S_c}+\mathsf{I_p}}{1+\mathsf{I_p}|c_1|^2}\right),$$
$$\tag{29k}$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq \log\left(\frac{1+\mathsf{C}}{1+\mathsf{C}\left(|a_1|^2+|c_1|^2\right)}\right) + \log\left(\frac{1+\mathsf{S_p}|c_1|^2+\mathsf{I_c}}{1+\mathsf{I_c}|b_2|^2}\right) + \log\left(1+\frac{\mathsf{S_c}|b_2|^2}{1+\mathsf{I_p}|c_1|^2}\right)$$

$$+ \log\left(1+\frac{\mathsf{S_c}+\mathsf{I_p}|a_1|^2}{1+\mathsf{I_p}|c_1|^2}\right). \tag{29l}$$

B. FME on the achievable rate region when $U_1 = \emptyset$

We set $U_1 = \emptyset$ in the achievable rate region. After FME of the achievable region in (27) with $U_1 = \emptyset$, we get

$$R_1 \leq \text{eq(27c)} + \text{eq(27o)}, \tag{30a}$$

$$R_1 \leq \text{eq(27e)}, \tag{30b}$$

$$R_2 \leq \text{eq(27r)}, \tag{30c}$$

$$R_1 + R_2 \leq \text{eq(27c)} + \text{eq(27i)} + \text{eq(27t)}, \tag{30d}$$

$$R_1 + R_2 \leq \text{eq(27e)} + \text{eq(27t)}, \tag{30e}$$

$$R_1 + R_2 \leq \text{eq(27d)} + \text{eq(27o)} + \text{eq(27p)}, \tag{30f}$$

$$R_1 + R_2 \leq \text{eq(27g)} + \text{eq(27p)}, \tag{30g}$$

$$R_1 + 2R_2 \leq \text{eq(27f)} + \text{eq(27p)} + \text{eq(27t)}, \tag{30h}$$

$$R_1 + 2R_2 \leq \text{eq(27d)} + \text{eq(27i)} + \text{eq(27p)} + \text{eq(27t)}, \tag{30i}$$

$$R_1 + 3R_2 \leq \text{eq(27k)} + \text{eq(27i)} + \text{eq(27p)} + 2 \cdot \text{eq(27t)}, \tag{30j}$$

for all distributions that factor as (26) and by setting $U_1 = \emptyset$ in all the mutual information terms.

Remark 3 *After FME, the following rate constraints (with $U_1 = \emptyset$) also appear*

$$R_2 \leq \text{eq(27m)} + \text{eq(27t)},$$

$$R_2 \leq \text{eq(27k)} + \text{eq(27t)}.$$

By following similar steps as in [29, Lemma 2] and [8, Appendix A], it is possible to show that these constraints are redundant. In other words, if these constraints are active a larger rate region is attained by not sending any common message, i.e., by setting $U_2 = \emptyset$.

We identify Node1 with the PTx (i.e., $X_{\mathsf{p}} = X_1$), Node2 with the CTx (i.e., $X_{\mathsf{c}} = X_2, T_{\mathsf{f}} = Y_2$), Node3 with the PRx (i.e., $Y_{\mathsf{p}} = Y_3$) and Node4 with the CRx (i.e., $Y_{\mathsf{c}} = Y_4$). For the Gaussian noise channel, in the achievable region in (30), we choose $Q = \emptyset$, we let $S_1, V_1, T_1, Z_1, U_2', T_2'$ be i.i.d. $\mathcal{N}(0,1)$, and

$$X_{\mathsf{p}} = |a_1|e^{j\theta_c}S_1 + b_1V_1 + c_1Z_1 + d_1T_1 \ : |a_1|^2 + |b_1|^2 + |c_1|^2 + |d_1|^2 = 1,$$

$$X_{\mathsf{c}} = |a_2|S_1 + b_2U_2' + c_2T_2' \ : |a_2|^2 + |b_2|^2 + |c_2|^2 = 1,$$

$$U_2 = U_2' + \lambda_U S_1 \ : \lambda_U = \frac{\mathsf{S_c}|b_2|^2}{\mathsf{S_c}|b_2|^2 + \mathsf{S_c}|c_2|^2 + 1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)} \frac{\sqrt{\mathsf{I_p}}e^{j\theta_p}e^{j\theta_c}|a_1| + \sqrt{\mathsf{S_c}}|a_2|}{\sqrt{\mathsf{S_c}}b_2},$$

$$T_2 = T_2' + \lambda_T S_1 \ : \lambda_T = \frac{\mathsf{S_c}|c_2|^2}{\mathsf{S_c}|c_2|^2 + 1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)} \frac{\sqrt{\mathsf{I_p}}e^{j\theta_p}e^{j\theta_c}|a_1| + \sqrt{\mathsf{S_c}}|a_2| - \sqrt{\mathsf{S_c}}b_2\lambda_U}{\sqrt{\mathsf{S_c}}c_2},$$

where the choice of λ_U is so as to "pre-cancel" S_1 from Y_{c} in decoding U_2, i.e., so as to have $I(Y_{\mathsf{c}}; U_2|V_1, Q) - I(S_1; U_2|Q) = I(Y_{\mathsf{c}}; U_2|V_1, Q, S_1)$ and the choice of λ_T is so as to "pre-cancel" S_1 from Y_{c} in decoding T_2, i.e., so as to have $I(Y_{\mathsf{c}}; T_2|V_1, Q, U_2) - I(S_1; T_2|Q, U_2) = I(Y_{\mathsf{c}}; T_2|V_1, Q, U_2, S_1)$. With these choices, the channel outputs are

$$T_{\mathsf{f}} = \sqrt{\mathsf{C}} \left(|a_1|e^{j\theta_c}S_1 + b_1V_1 + c_1Z_1 + d_1T_1 \right) + Z_{\mathsf{f}},$$

$$Y_{\mathsf{p}} = (\sqrt{\mathsf{S_p}}|a_1| + \sqrt{\mathsf{I_c}}|a_2|)e^{j\theta_c}S_1 + \sqrt{\mathsf{S_p}}(b_1V_1 + c_1Z_1 + d_1T_1) + \sqrt{\mathsf{I_c}}e^{j\theta_c}(b_2U_2' + c_2T_2') + Z_{\mathsf{p}},$$

$$Y_{\mathsf{c}} = (\sqrt{\mathsf{I_p}}e^{j\theta_p}e^{j\theta_c}|a_1| + \sqrt{\mathsf{S_c}}|a_2|)S_1 + \sqrt{\mathsf{I_p}}e^{j\theta_p}(b_1V_1 + c_1Z_1 + d_1T_1) + \sqrt{\mathsf{S_c}}(b_2U_2' + c_2T_2') + Z_{\mathsf{c}},$$

and the achievable region in (30) becomes

$$R_{\mathsf{p}} \leq \log\left(1 + \frac{\mathsf{C}(|b_1|^2 + |c_1|^2)}{1 + \mathsf{C}|d_1|^2}\right) + \log\left(1 + \frac{\mathsf{S_p}|d_1|^2}{1 + \mathsf{I_c}|c_2|^2}\right), \tag{31a}$$

$$R_{\mathsf{p}} \leq \log\left(\frac{1 + \mathsf{S_p} + \mathsf{I_c} + 2\sqrt{\mathsf{S_p}\mathsf{I_c}|a_1|^2|a_2|^2}}{1 + \mathsf{I_c}|c_2|^2}\right), \tag{31b}$$

$$R_{\mathsf{c}} \leq \log\left(1 + \frac{\mathsf{S_c}(|b_2|^2 + |c_2|^2)}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right), \tag{31c}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1 + \frac{\mathsf{C}(|b_1|^2 + |c_1|^2)}{1 + \mathsf{C}|d_1|^2}\right) + \log\left(1 + \frac{\mathsf{S_p}|d_1|^2 + \mathsf{I_c}|b_2|^2}{1 + \mathsf{I_c}|c_2|^2}\right)$$

$$+ \log\left(1 + \frac{\mathsf{S_c}|c_2|^2}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right), \tag{31d}$$

$$R_\mathsf{p} + R_\mathsf{c} \leq \log\left(\frac{1 + \mathsf{S_p} + \mathsf{I_c} + 2\sqrt{\mathsf{S_p}\mathsf{I_c}|a_1|^2|a_2|^2}}{1 + \mathsf{I_c}|c_2|^2}\right) + \log\left(1 + \frac{\mathsf{S_c}|c_2|^2}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right), \tag{31e}$$

$$R_\mathsf{p} + R_\mathsf{c} \leq \log\left(1 + \frac{\mathsf{C}|c_1|^2}{1 + \mathsf{C}|d_1|^2}\right) + \log\left(1 + \frac{\mathsf{S_p}|d_1|^2}{1 + \mathsf{I_c}|c_2|^2}\right)$$
$$+ \log\left(1 + \frac{\mathsf{S_c}(|b_2|^2 + |c_2|^2)}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right) + k_1, \tag{31f}$$

$$R_\mathsf{p} + R_\mathsf{c} \leq \log\left(1 + \frac{\mathsf{S_p}(|c_1|^2 + |d_1|^2)}{1 + \mathsf{I_c}|c_2|^2}\right) + \log\left(1 + \frac{\mathsf{S_c}(|b_2|^2 + |c_2|^2)}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right) + k_1 + k_2, \tag{31g}$$

$$R_\mathsf{p} + 2R_\mathsf{c} \leq \log\left(1 + \frac{\mathsf{S_p}(|c_1|^2 + |d_1|^2) + \left(\sqrt{\mathsf{S_p}}|a_1| + \sqrt{\mathsf{I_c}}|a_2|\right)^2 + \mathsf{I_c}|b_2|^2}{1 + \mathsf{I_c}|c_2|^2}\right)$$
$$+ \log\left(1 + \frac{\mathsf{S_c}(|b_2|^2 + |c_2|^2)}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right) + \log\left(1 + \frac{\mathsf{S_c}|c_2|^2}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right) + k_1, \tag{31h}$$

$$R_\mathsf{p} + 2R_\mathsf{c} \leq \log\left(1 + \frac{\mathsf{C}|c_1|^2}{1 + \mathsf{C}|d_1|^2}\right) + \log\left(1 + \frac{\mathsf{S_p}|d_1|^2 + \mathsf{I_c}|b_2|^2}{1 + \mathsf{I_c}|c_2|^2}\right)$$
$$+ \log\left(1 + \frac{\mathsf{S_c}(|b_2|^2 + |c_2|^2)}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right) + \log\left(1 + \frac{\mathsf{S_c}|c_2|^2}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right) + k_1, \tag{31i}$$

$$R_\mathsf{p} + 3R_\mathsf{c} \leq \log\left(1 + \frac{\mathsf{S_p}|c_1|^2 + \left(\sqrt{\mathsf{S_p}}|a_1| + \sqrt{\mathsf{I_c}}|a_2|\right)^2 + \mathsf{I_c}|b_2|^2}{1 + \mathsf{I_c}|c_2|^2}\right) + \log\left(1 + \frac{\mathsf{S_p}|d_1|^2 + \mathsf{I_c}|b_2|^2}{1 + \mathsf{I_c}|c_2|^2}\right)$$
$$+ \log\left(1 + \frac{\mathsf{S_c}(|b_2|^2 + |c_2|^2)}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right) + 2\log\left(1 + \frac{\mathsf{S_c}|c_2|^2}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right) + k_1, \tag{31j}$$

where we defined $k_1 := I(Y_\mathsf{c}; V_1)$ and $k_2 := I(Y_\mathsf{p}; S_1 | V_1, U_2)$ without evaluating them for the Gaussian noise case. Actually, further lower bounding k_1 and k_2 by zero suffices to prove a constant gap for the yellow and red regions in Fig. 4.

References

[1] M. Cardone, D. Tuninetti, R. Knopp, and U. Salim, "On the Gaussian interference channel with unilateral generalized feedback," in *6th International Symposium on Communications, Control and Signal Processing (ISCCSP), 2014*, May 2014.

[2] ——, "New outer bounds for the interference channel with unilateral source cooperation," in *IEEE International Symposium on Information Theory (ISIT), 2014*, July 2014.

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

[4] A. Host-Madsen, "Capacity bounds for cooperative diversity," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1522 –1544, April 2006.

[5] D. Tuninetti, "An outer bound region for interference channels with generalized feedback," in *Information Theory and Applications Workshop (ITA), 2010*, February 2010.

[6] V. Prabhakaran and P. Viswanath, "Interference channels with source cooperation," *IEEE Transactions on Information Theory*, vol. 57, no. 1, pp. 156 –186, January 2011.

[7] R. Tandon and S. Ulukus, "Dependence balance based outer bounds for Gaussian networks with cooperation and feedback," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4063 –4086, July 2011.

[8] S. Yang and D. Tuninetti, "Interference channel with generalized feedback (a.k.a. with source cooperation): Part i: Achievable region," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2686 –2710, May 2011.

[9] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.

[10] S. Gelfand and M. Pinsker, "Coding for channel with random parameters," *Problems of control and information theory*, vol. 9, no. 1, pp. 19 – 31, 1980.

[11] M. Costa, "Writing on dirty paper (corresp.)," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439 – 441, May 1983.

[12] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 20, no. 3, pp. 581–586, March 2004.

[13] D. Tuninetti, "An outer bound for the memoryless two-user interference channel with general cooperation," in *IEEE Information Theory Workshop (ITW), 2012*, September 2012, pp. 217–221.

[14] S. Yang and D. Tuninetti, "Interference channels with source cooperation in the strong cooperation regime: Symmetric capacity to within 2 bits/s/hz with dirty paper coding," in *Asilomar Conference on Signals, Systems, and Computers*, November 2011, pp. 2140 –2144.

[15] A. Hekstra and F. Willems, "Dependence balance bounds for single-output two-way channels," *IEEE Transactions on Information Theory*, vol. 35, no. 1, pp. 44 –53, January 1989.

[16] M. Mirmohseni, B. Akhbari, and M. Aref, "On the capacity of interference channel with causal and noncausal generalized feedback at the cognitive transmitter," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2813–2837, May 2012.

[17] M. Cardone, D. Tuninetti, R. Knopp, and U. Salim, "On the capacity of the two-user Gaussian causal cognitive interference channel," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2512–2541, May 2014.

[18] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Transactions on Information Theory*, vol. 54, no. 12, pp. 5534 –5562, December 2008.

[19] C. Suh and D. Tse, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2667–2685, May 2011.

[20] A. Sahai, V. Aggarwal, M. Yuksel, and A. Sabharwal, "Capacity of all nine models of channel output feedback for the two-user interference channel," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 6957–6979, November 2013.

[21] S.-Q. Le, R. Tandon, M. Motani, and H. Poor, "The capacity region of the symmetric linear deterministic interference channel with partial feedback," in *50th Annual Allerton Conference on Communication, Control, and Computing, 2012*, October 2012, pp. 1864–1871.

[22] ——, "Approximate capacity region for the symmetric Gaussian interference channel with noisy feedback," *submitted to IEEE Transactions on Information Theory, available at http://people.cs.vt.edu/tandonr/journals.html*, December 2012.

[23] S. Rini, D. Tuninetti, and N. Devroye, "New inner and outer bounds for the memoryless cognitive interference channel and some new capacity results," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4087–4109, July 2011.

[24] ——, "Inner and outer bounds for the Gaussian cognitive interference channel and new capacity results," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 820 –848, February 2012.

[25] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 49 – 60, January 1981.

[26] E. Telatar and D. Tse, "Bounds on the capacity region of a class of interference channels," in *IEEE International Symposium on Information Theory (ISIT), 2007*, June 2007, pp. 2871–2874.

[27] A. E. Gamal and Y.-H. Kim, *Network Information Theory*.   Cambridge U.K.: Cambridge Univ. Press,, 2011.

[28] G. Kramer, "Directed information for channel with feedback," Ph.D. dissertation, Swiss Federal Institute of Technology, Zurich, 1998.

[29] H.-F. Chong, M. Motani, H. K. Garg, and H. El Gamal, "On the Han-Kobayashi region for the interference channel," *IEEE Transactions on Information Theory*, vol. 54, no. 7, pp. 3188–3195, July 2008.