



HAL
open science

Approximate Capacity Region of the Two-User Gaussian Interference Channel with Noisy Channel-Output Feedback

Victor Quintero, Samir Perlaza, Iñaki Esnaola, Jean-Marie M Gorce

► **To cite this version:**

Victor Quintero, Samir Perlaza, Iñaki Esnaola, Jean-Marie M Gorce. Approximate Capacity Region of the Two-User Gaussian Interference Channel with Noisy Channel-Output Feedback. 2016. hal-01397118v1

HAL Id: hal-01397118

<https://hal.science/hal-01397118v1>

Preprint submitted on 15 Nov 2016 (v1), last revised 19 May 2018 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Approximate Capacity Region of the Two-User Gaussian Interference Channel with Noisy Channel-Output Feedback

Victor Quintero, Samir M. Perlaza, Iñaki Esnaola, Jean-Marie Gorce

Abstract—In this paper, the capacity region of the linear deterministic interference channel with noisy channel-output feedback (LD-IC-NOF) is fully characterized. A capacity-achieving scheme is obtained using a random coding argument and three well-known techniques: rate splitting, superposition coding and backward decoding. The converse region is obtained using some of the existing outer bounds as well as a set of new outer bounds that are obtained by using genie-aided models of the original LD-IC-NOF. Using the insights gained from the analysis of the LD-IC-NOF, an achievability region and a converse region for the two-user Gaussian interference channel with noisy channel-output feedback (G-IC-NOF) are presented. Finally, the achievability region and the converse region approximate the capacity region of the G-IC-NOF to within 4.4 bits.

Index Terms—Linear Deterministic Interference Channel, Gaussian Interference Channel, Feedback and Capacity.

I. INTRODUCTION

The interest in the analysis of channel-output feedback in the interference channel (IC) has been fueled by the significant improvement it brings on the number of generalized degrees of freedom (GDoFs) with respect to the case without feedback [3]. Let $\mathcal{C}(\overline{\text{SNR}}, \overline{\text{INR}})$ denote the set of all achievable rates (capacity region) of a symmetric Gaussian IC with parameters $\overline{\text{SNR}}$ (signal to noise ratio in the forward link) and $\overline{\text{INR}}$ (interference to noise ratio). Note that the number of GDoFs [4], i.e.,

$$\text{GDoF}(\alpha) = \lim_{\overline{\text{SNR}} \rightarrow \infty} \frac{\sup \{R : (R, R) \in \mathcal{C}(\overline{\text{SNR}}, \overline{\text{SNR}}^\alpha)\}}{\log(\overline{\text{SNR}})}, \quad (1)$$

increases for some values of $\alpha = \frac{\log(\overline{\text{INR}})}{\log(\overline{\text{SNR}})}$ with respect to the case without feedback. In particular, when $\alpha \in [0, \frac{2}{3}]$ and

Victor Quintero, Samir M. Perlaza and Jean-Marie Gorce are with the CITI Laboratory of the Institut National de Recherche en Informatique et en Automatique (INRIA), Université de Lyon, and Institut National de Sciences Appliquées (INSA) de Lyon, 6 Av. des Arts 69621 Villeurbanne, France. ({victor.quintero-florez, samir.perlaza, jean-marie.gorce}@inria.fr).

Iñaki Esnaola is with the Department of Automatic Control and Systems Engineering, University of Sheffield, Mappin Street, Sheffield, S1 3JD, United Kingdom. (esnaola@sheffield.ac.uk).

Victor Quintero is also with Universidad del Cauca, Popayán, Colombia.

Samir M. Perlaza and Iñaki Esnaola are also with the Department of Electrical Engineering, University of Sheffield, Princeton, NJ.

This research was supported in part by the European Commission under Marie Skłodowska-Curie Individual Fellowship No. 659316 (CYBERNETS); and the Administrative Department of Science, Technology, and Innovation of Colombia (Colciencias), fellowship No. 617-2013.

Parts of this work were presented at the IEEE International Workshop on Information Theory (ITW), Jeju Island, South Korea, October, 2015 [1], and ITW, Cambridge, United Kingdom, September, 2016 [2].

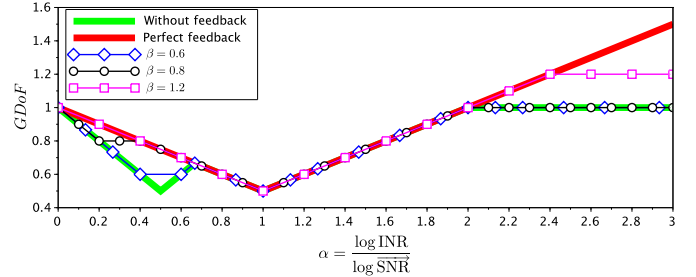


Fig. 1. Generalized Degrees of Freedom (GDoF) of a symmetric two-user Gaussian IC with feedback.

$\alpha \in [2, \infty)$ substantial improvements are observed, as shown in Figure 1. Note that $\text{GDoF}(\alpha) \rightarrow \infty$ when $\alpha \rightarrow \infty$, which implies an arbitrarily large gain in the asymptotic regime of INR. These promising results are also observed when the system is fully decentralized, that is, when each transmitter seeks to unilaterally maximize its own individual data rate [5].

However, recent studies by Le *et al.* [6] show that the number of GDoFs is highly sensitive to additive Gaussian noise on the feedback links. Let $\overline{\text{SNR}}$ be the signal to noise ratio in the feedback links. Figure 1 shows the GDoFs of the G-IC for different values of the ratio $\beta = \frac{\log(\overline{\text{SNR}})}{\log(\overline{\text{SNR}})}$. Note that in the presence of noise, the improvement on the number of GDoFs is no longer monotonically increasing with α in the interval $[2, \infty)$ but is instead upper-bounded by $\min(\frac{\alpha}{2}, \beta)$. This reveals the fact that Gaussian additive noise in the feedback links dramatically limits the benefits of channel-output feedback. Similar results have been reported in the fully decentralized G-IC-NOF [7].

A deeper understanding on the impact of Gaussian additive noise in the feedback links of a given G-IC can be obtained by studying the capacity region of the two-user Gaussian interference channel with noisy channel-output feedback (G-IC-NOF). This capacity region, denoted by $\mathcal{C}_{\text{NOF}}(\overline{\text{SNR}}_1, \overline{\text{SNR}}_2, \overline{\text{SNR}}_1, \overline{\text{SNR}}_2, \overline{\text{INR}}_{12}, \overline{\text{INR}}_{21})$, is fully described by six parameters: the SNR of the forward link between transmitter-receiver pair i , denoted by $\overline{\text{SNR}}_i$; the SNR of the feedback link between transmitter-receiver pair i , denoted by $\overline{\text{SNR}}_i$; and the INR at receiver i , denoted by $\overline{\text{INR}}_{ij}$, with $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$. With the exception of a particular case mentioned later, the region $\mathcal{C}_{\text{NOF}}(\overline{\text{SNR}}_1, \overline{\text{SNR}}_2, \overline{\text{SNR}}_1, \overline{\text{SNR}}_2, \overline{\text{INR}}_{12}, \overline{\text{INR}}_{21})$

is not yet fully known and only approximations to within a constant gap are available for certain cases. Denote by $\overrightarrow{C_{\text{WOF}}}(\overrightarrow{\text{SNR}}_1, \overrightarrow{\text{SNR}}_2, \text{INR}_{12}, \text{INR}_{21})$ and $\overrightarrow{C_{\text{POF}}}(\overrightarrow{\text{SNR}}_1, \overrightarrow{\text{SNR}}_2, \text{INR}_{12}, \text{INR}_{21})$ the capacity region of the G-IC without feedback and with perfect channel-output feedback, respectively. Note also that:

$$\overrightarrow{C_{\text{WOF}}}(\overrightarrow{\text{SNR}}_1, \overrightarrow{\text{SNR}}_2, \text{INR}_{12}, \text{INR}_{21}) = \lim_{\substack{\overrightarrow{\text{SNR}}_1 \rightarrow 0 \\ \overrightarrow{\text{SNR}}_2 \rightarrow 0}} \overrightarrow{C_{\text{NOF}}}(\overrightarrow{\text{SNR}}_1, \overrightarrow{\text{SNR}}_2, \overleftarrow{\text{SNR}}_1, \overleftarrow{\text{SNR}}_2, \text{INR}_{12}, \text{INR}_{21})$$

and

$$\overrightarrow{C_{\text{POF}}}(\overrightarrow{\text{SNR}}_1, \overrightarrow{\text{SNR}}_2, \text{INR}_{12}, \text{INR}_{21}) = \lim_{\substack{\overrightarrow{\text{SNR}}_1 \rightarrow \infty \\ \overrightarrow{\text{SNR}}_2 \rightarrow \infty}} \overrightarrow{C_{\text{NOF}}}(\overrightarrow{\text{SNR}}_1, \overrightarrow{\text{SNR}}_2, \overleftarrow{\text{SNR}}_1, \overleftarrow{\text{SNR}}_2, \text{INR}_{12}, \text{INR}_{21}).$$

The capacity region $\overrightarrow{C_{\text{WOF}}}(\overrightarrow{\text{SNR}}_1, \overrightarrow{\text{SNR}}_2, \text{INR}_{12}, \text{INR}_{21})$ is approximated to within one bit by Etkin and Tse [8] (see Definition 2) using elements of the achievability scheme introduced by Han and Kobayashi in [9]. The capacity region $\overrightarrow{C_{\text{WOF}}}(\overrightarrow{\text{SNR}}_1, \overrightarrow{\text{SNR}}_2, \text{INR}_{12}, \text{INR}_{21})$ is perfectly known in the case in which both links are in high interference regime [10]. In all the other interference regimes, only inner and outer bounds are available [11], [12].

The capacity region $\overrightarrow{C_{\text{POF}}}(\overrightarrow{\text{SNR}}_1, \overrightarrow{\text{SNR}}_2, \text{INR}_{12}, \text{INR}_{21})$ is approximated to within two-bits by Suh and Tse [3] (see Definition 2). The achievability scheme presented in [3] is based on three well-known techniques: rate splitting [9], [13], block Markov superposition coding [14] and backward decoding [15], [16]. Other achievability schemes have been independently presented by Tuninetti [17] and Yang *et al.* [18] using rate-splitting, superposition coding, backward decoding, and binning/dirty paper coding for the interference channel with generalized feedback (IC-GF). It is worth mentioning that these achievability schemes are general and are useful for more involved channel models. Recently Prabhakaran *et al.* [19] have presented both inner and outer bounds for the sum-capacity using the existing connections between channel-output feedback and conferencing transmitters. More advanced models, for instance when channel-outputs are fed back to several receivers, have been studied and inner and outer bounds are available in [20]–[23]. When feedback is impaired by additive Gaussian noise, Kramer and Gastpar [24] have established an outer bound using the Hekstra-Willems dependence-balance arguments for two-way channels [25]. In the G-IC, these results suggest that feedback loses its efficacy on increasing the capacity region roughly when the noise variance on the feedback link is larger than on the forward link. More recently, using similar arguments, Tandon and Ulukus [26] have derived outer bounds that are tighter than the cut-set bound in some interference regimes. In the particular case of the symmetric G-IC-NOF, the capacity region is approximated to within a constant gap in [6]. However, despite this approximate capacity region description, very little can be concluded in the case in which feedback is available in only one of the point-to-point links or simply when the point-to-point links are in different interference regimes.

A. Contribution

In this paper, the exact capacity region of the LD-IC-NOF is fully characterized (Theorem 1). Using key insights obtained from the analysis of the LD-IC-NOF, an achievability region (Theorem 2) and a converse region (Theorem 3) are presented for the two-user G-IC-NOF. These two regions approximate the capacity region $\overrightarrow{C_{\text{NOF}}}(\overrightarrow{\text{SNR}}_1, \overrightarrow{\text{SNR}}_2, \overleftarrow{\text{SNR}}_1, \overleftarrow{\text{SNR}}_2, \text{INR}_{12}, \text{INR}_{21})$ to within 4.4 bits (Theorem 4). To the extent of the knowledge of the authors, this approximation to the capacity region of the G-IC-NOF is the most general with respect to existing literature and the one that guarantees the smallest gap between the achievable and converse regions when feedback links are subject to Gaussian additive noise.

B. Organization of the paper

Section II introduces the notation used in this paper. Section III describes the two-user G-IC-NOF and identifies the parameters that fully describe it. Section IV describes the exact capacity region of the LD-IC-NOF. Using two examples, the main insights for the design of the achievability scheme and the derivation of the new outer bounds are described. Section V introduces the main results, essentially, an achievability and a converse region for the G-IC-NOF. Section VI describes the connections between the IC-NOF and an IC with conferencing transmitters (IC-CT). In particular, it is highlighted that any outer bound for the IC-CT is also an outer bound for the IC-NOF (and *vice-versa*) under certain conditions. Finally, Section VII concludes this work and highlights some interesting extensions of these results.

II. NOTATION

Throughout this paper, sets are denoted with uppercase calligraphic letters, e.g. \mathcal{X} . Random variables are denoted by uppercase letters, e.g., X . The realizations and the set of events from which the random variable X takes values are respectively denoted by x and \mathcal{X} . The probability distribution of X over the set \mathcal{X} is denoted P_X . Whenever a second random variable Y is involved, P_{XY} and $P_{Y|X}$ denote respectively the joint probability distribution of (X, Y) and the conditional probability distribution of Y given X . Let N be a fixed natural number. An N -dimensional vector of random variables is denoted by $\mathbf{X} = (X_1, X_2, \dots, X_N)^T$ and a corresponding realization is denoted by $\mathbf{x} = (x_1, x_2, \dots, x_N)^T \in \mathcal{X}^N$. Given $\mathbf{X} = (X_1, X_2, \dots, X_N)^T$ and $(a, b) \in \mathbb{N}^2$, with $a < b \leq N$, the $(b - a + 1)$ -dimensional vector of random variables formed by the components a to b of \mathbf{X} is denoted by $\mathbf{X}_{(a,b)} = (X_a, X_{a+1}, \dots, X_b)^T$. The notation $(\cdot)^+$ denotes the positive part operator, i.e., $(\cdot)^+ = \max(\cdot, 0)$ and $\mathbb{E}_X[\cdot]$ denotes the expectation with respect to the distribution of the random variable X . The logarithm function log is assumed to be base 2.

III. PROBLEM FORMULATION

This section introduces the two-user Gaussian interference channel with noisy channel-output feedback (G-IC-NOF) and defines an approximation to its corresponding

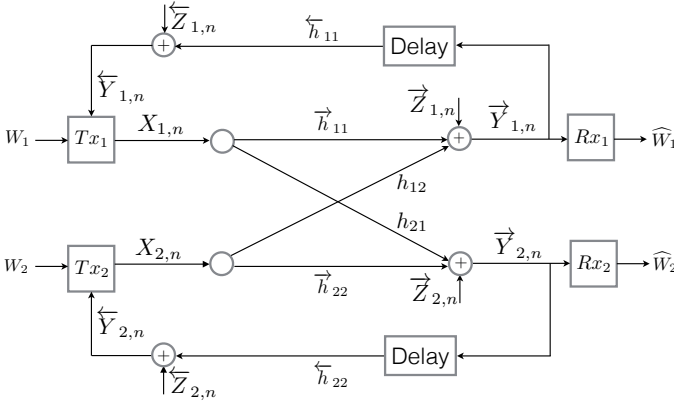


Fig. 2. Gaussian interference channel with noisy channel-output feedback at channel use n .

capacity region. Consider the two-user G-IC-NOF in Figure 2. Transmitter i , with $i \in \{1, 2\}$, communicates with receiver i subject to the interference produced by transmitter j , with $j \in \{1, 2\} \setminus \{i\}$. There are two independent and uniformly distributed messages, $W_i \in \mathcal{W}_i$, with $\mathcal{W}_i = \{1, 2, \dots, 2^{NR_i}\}$, where N denotes the fixed block-length in channel uses and R_i is the transmission rate in bits per channel use. At each block, transmitter i sends the codeword $\mathbf{X}_i = (X_{i,1}, X_{i,2}, \dots, X_{i,N})^T \in \mathcal{X}_i^N$, where \mathcal{X}_i and \mathcal{X}_i^N are respectively the channel-input alphabet and the codebook of transmitter i .

The channel coefficient from transmitter j to receiver i is denoted by h_{ij} ; the channel coefficient from transmitter i to receiver i is denoted by \vec{h}_{ii} ; and the channel coefficient from channel-output i to transmitter i is denoted by \overleftarrow{h}_{ii} . All channel coefficients are assumed to be non-negative real numbers. At a given channel use $n \in \{1, 2, \dots, N\}$, the channel output at receiver i is denoted by $\vec{Y}_{i,n}$. During channel use n , the input-output relation of the channel model is given by

$$\vec{Y}_{i,n} = \vec{h}_{ii} X_{i,n} + h_{ij} X_{j,n} + \vec{Z}_{i,n}, \quad (2)$$

where $\vec{Z}_{i,n}$ is a real Gaussian random variable with zero mean and unit variance that represents the noise at the input of receiver i . Let $d > 0$ be the finite feedback delay measured in channel uses. At the end of channel use n , transmitter i observes $\overleftarrow{Y}_{i,n}$, which consists of a scaled and noisy version of $\vec{Y}_{i,n-d}$. More specifically,

$$\overleftarrow{Y}_{i,n} = \begin{cases} \vec{Z}_{i,n} & \text{for } n \in \{1, 2, \dots, d\} \\ \overleftarrow{h}_{ii} \vec{Y}_{i,n-d} + \overleftarrow{Z}_{i,n} & \text{for } n \in \{d+1, d+2, \dots, N\}, \end{cases} \quad (3)$$

where $\overleftarrow{Z}_{i,n}$ is a real Gaussian random variable with zero mean and unit variance that represents the noise in the feedback link of transmitter-receiver pair i . The random variables $\vec{Z}_{i,n}$ and $\overleftarrow{Z}_{i,n}$ are independent and identically distributed. In the following, without loss of generality, the feedback delay is assumed to be one channel use, i.e., $d = 1$. The encoder of transmitter i is defined by a set of deterministic functions $f_i^{(1)}, f_i^{(2)}, \dots, f_i^{(N)}$, with $f_i^{(t)} : \mathcal{W}_i \rightarrow \mathcal{X}_i$ and for all

$n \in \{2, 3, \dots, N\}$, $f_i^{(n)} : \mathcal{W}_i \times \mathbb{R}^{n-1} \rightarrow \mathcal{X}_i$, such that

$$X_{i,1} = f_i^{(1)}(W_i), \quad \text{and} \quad (4a)$$

$$X_{i,n} = f_i^{(n)}(W_i, \overleftarrow{Y}_{i,1}, \overleftarrow{Y}_{i,2}, \dots, \overleftarrow{Y}_{i,n-1}). \quad (4b)$$

The components of the input vector \mathbf{X}_i are real numbers subject to an average power constraint:

$$\frac{1}{N} \sum_{n=1}^N \mathbb{E}(X_{i,n}^2) \leq 1, \quad (5)$$

where the expectation is taken over the joint distribution of the message indexes W_1, W_2 , and the noise terms, i.e., $\vec{Z}_1, \vec{Z}_2, \overleftarrow{Z}_1$, and \overleftarrow{Z}_2 . The dependence of $X_{i,n}$ on W_1, W_2 , and the previously observed noise realizations is due to the effect of feedback as shown in (3) and (4).

Let $T \in \mathbb{N}$ be fixed. Assume that during a given communication, T blocks, each of N channel uses, are transmitted. Hence, the decoder of receiver i is defined by a deterministic function $\psi_i : \mathbb{R}_i^{NT} \rightarrow \mathcal{W}_i^T$. At the end of the communication, receiver i uses the vector $(\overleftarrow{Y}_{i,1}, \overleftarrow{Y}_{i,2}, \dots, \overleftarrow{Y}_{i,NT})^T$ to obtain an estimate of the message indices:

$$(\widehat{W}_i^{(1)}, \widehat{W}_i^{(2)}, \dots, \widehat{W}_i^{(T)}) = \psi_i(\overleftarrow{Y}_{i,1}, \overleftarrow{Y}_{i,2}, \dots, \overleftarrow{Y}_{i,NT}), \quad (6)$$

where $\widehat{W}_i^{(t)}$ is an estimate of the message index sent during block $t \in \{1, 2, \dots, T\}$. The decoding error probability in the two-user G-IC-NOF during block t of a codebook of block-length N , denoted by $P_e^{(t)}(N)$, is given by

$$P_e^{(t)}(N) = \max \left(\Pr \left[\widehat{W}_1^{(t)} \neq W_1^{(t)} \right], \Pr \left[\widehat{W}_2^{(t)} \neq W_2^{(t)} \right] \right). \quad (7)$$

The definition of an achievable rate pair $(R_1, R_2) \in \mathbb{R}_+^2$ is given below.

Definition 1 (Achievable Rate Pairs): A rate pair $(R_1, R_2) \in \mathbb{R}_+^2$ is achievable if there exists at least one pair of codebooks \mathcal{X}_1^N and \mathcal{X}_2^N with codewords of length N , and the corresponding encoding functions $f_1^{(1)}, f_1^{(2)}, \dots, f_1^{(N)}$ and $f_2^{(1)}, f_2^{(2)}, \dots, f_2^{(N)}$ such that the decoding error probability $P_e^{(t)}(N)$ can be made arbitrarily small by letting the block-length N grow to infinity, for all blocks $t \in \{1, 2, \dots, T\}$.

The two-user G-IC-NOF in Figure 2 can be fully described by six parameters: $\overrightarrow{\text{SNR}}_i, \overleftarrow{\text{SNR}}_i$, and INR_{ij} , with $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$, which are defined as follows:

$$\overrightarrow{\text{SNR}}_i = \vec{h}_{ii}^2, \quad (8)$$

$$\text{INR}_{ij} = h_{ij}^2, \quad \text{and} \quad (9)$$

$$\overleftarrow{\text{SNR}}_i = \overleftarrow{h}_{ii}^2 (\vec{h}_{ii}^2 + 2\vec{h}_{ii}h_{ij} + h_{ij}^2 + 1). \quad (10)$$

The analysis presented in this paper focuses exclusively on the case in which $\text{INR}_{ij} > 1$ for all $(i, j) \in \{1, 2\} \times \{\{1, 2\} \setminus \{i\}\}$. The reason for exclusively considering this case follows from the fact that when $\text{INR}_{ij} \leq 1$, transmitter-receiver pair i is impaired mainly by noise instead of interference. In this case, treating interference as noise is optimal and feedback does not bring a significant rate improvement.

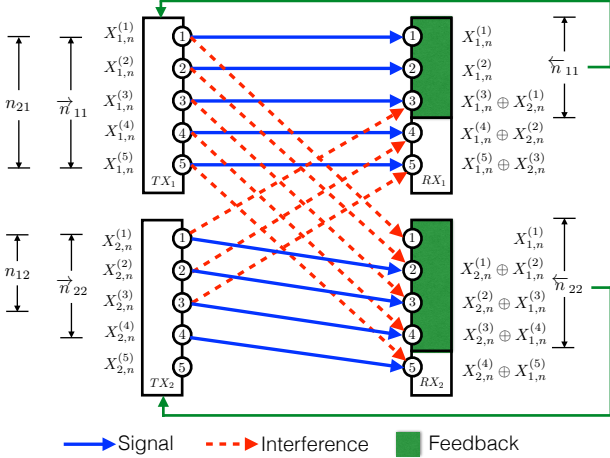


Fig. 3. Two-user linear deterministic interference channel with noisy channel-output feedback at channel use n .

IV. PRELIMINARY RESULTS

This section describes the two-user linear deterministic interference channel with noisy channel-output feedback (LD-IC-NOF) and its exact capacity region. The relevance of this result is that it provides the main insights used to obtain the approximate capacity of the G-IC-NOF in Section V.

A. Two-User Linear Deterministic Interference Channel with Noisy Feedback

Consider the two-user linear deterministic interference channel with noisy channel-output feedback (LD-IC-NOF) described in Figure 3. For all $i \in \{1, 2\}$, with $j \in \{1, 2\} \setminus \{i\}$, the number of bit-pipes between transmitter i and its corresponding intended receiver is denoted by \vec{n}_{ii} ; the number of bit-pipes between transmitter i and its corresponding non-intended receiver is denoted by n_{ji} ; and the number of bit-pipes between receiver i and its corresponding transmitter is denoted by \overleftarrow{n}_{ii} . These six integer non-negative parameters fully describe the LD-IC-NOF in Figure 3.

At transmitter i , the channel-input $\mathbf{X}_{i,n}$ at channel use n , with $n \in \{1, 2, \dots, N\}$, is a q -dimensional binary vector $\mathbf{X}_{i,n} = (X_{i,n}^{(1)}, X_{i,n}^{(2)}, \dots, X_{i,n}^{(q)})^\top$, with

$$q = \max(\vec{n}_{11}, \vec{n}_{22}, n_{12}, n_{21}), \quad (11)$$

and N the block-length. At receiver i , the channel-output $\vec{\mathbf{Y}}_{i,n}$ at channel use n is also a q -dimensional binary vector $\vec{\mathbf{Y}}_{i,n} = (\vec{Y}_{i,n}^{(1)}, \vec{Y}_{i,n}^{(2)}, \dots, \vec{Y}_{i,n}^{(q)})^\top$. The input-output relation during channel use n is given by

$$\vec{\mathbf{Y}}_{i,n} = \mathbf{S}^{q-\vec{n}_{ii}} \mathbf{X}_{i,n} + \mathbf{S}^{q-n_{ij}} \mathbf{X}_{j,n}, \quad (12)$$

and the feedback signal $\overleftarrow{\mathbf{Y}}_{i,n}$ available at transmitter i at the end of channel use n satisfies

$$\left((0, \dots, 0), \overleftarrow{\mathbf{Y}}_{i,n}^\top \right)^\top = \mathbf{S}^{(\max(\vec{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii})^+} \vec{\mathbf{Y}}_{i,n-d}, \quad (13)$$

where d is a finite delay, additions and multiplications are defined over the binary field, and \mathbf{S} is a $q \times q$ lower shift matrix of the form:

$$\mathbf{S} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix}. \quad (14)$$

The dimension of the vector $(0, \dots, 0)$ in (13) is $q - \min(\overleftarrow{n}_{ii}, \max(\vec{n}_{ii}, n_{ij}))$ and the vector $\overleftarrow{\mathbf{Y}}_{i,n}$ represents the $\min(\overleftarrow{n}_{ii}, \max(\vec{n}_{ii}, n_{ij}))$ least significant bits of $\mathbf{S}^{(\max(\vec{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii})^+} \vec{\mathbf{Y}}_{i,n-d}$.

The parameters \vec{n}_{ii} , \overleftarrow{n}_{ii} and n_{ij} correspond to $\lfloor \frac{1}{2} \log_2(\overline{\text{SNR}}_i) \rfloor$, $\lfloor \frac{1}{2} \log_2(\underline{\text{SNR}}_i) \rfloor$ and $\lfloor \frac{1}{2} \log_2(\text{INR}_{ij}) \rfloor$ respectively, where $\overline{\text{SNR}}_i$, $\underline{\text{SNR}}_i$ and INR_{ij} are parameters of the Gaussian interference channel (G-IC) in (8), (9) and (10), respectively. The existing connections between the linear deterministic model and the Gaussian model are thoroughly described in [27].

As in the previous section and without any loss of generality, the feedback delay is assumed to be equal to 1 channel use. Transmitter i sends the message index W_i by sending the codeword $\mathbf{X}_i = (\mathbf{X}_{i,1}, \mathbf{X}_{i,2}, \dots, \mathbf{X}_{i,N}) \in \mathcal{X}_i^N$. The encoder of transmitter i can be modeled as a set of deterministic mappings $f_i^{(1)}, f_i^{(2)}, \dots, f_i^{(N)}$, with $f_i^{(1)}: W_i \rightarrow \{0, 1\}^q$ and for all $n \in \{2, 3, \dots, N\}$, $f_i^{(n)}: \mathcal{W}_i \times \{0, 1\}^{q(n-1)} \rightarrow \{0, 1\}^q$, such that

$$\mathbf{X}_{i,1} = f_i^{(1)}(W_i) \quad \text{and} \quad (16)$$

$$\mathbf{X}_{i,n} = f_i^{(n)}(W_i, \overleftarrow{\mathbf{Y}}_{i,1}, \overleftarrow{\mathbf{Y}}_{i,2}, \dots, \overleftarrow{\mathbf{Y}}_{i,n-1}). \quad (17)$$

Assume that during a given communication, T blocks are transmitted. Hence, the decoder of receiver i is defined by a deterministic function $\psi_i: \{0, 1\}^{q \times N \times T} \rightarrow \mathcal{W}_i^T$. At the end of the communication, receiver i uses the sequence $(\overleftarrow{\mathbf{Y}}_{i,1}, \overleftarrow{\mathbf{Y}}_{i,2}, \dots, \overleftarrow{\mathbf{Y}}_{i,NT})$ to obtain an estimate of the message indices.

The decoding error probability in the two-user LD-IC-NOF, denoted by $P_e^{(i)}(N)$, is calculated following (7). Similarly, a rate pair $(R_1, R_2) \in \mathbb{R}_+^2$ is said to be achievable if it satisfies Definition 1.

B. Capacity Region of the Two-User Linear Deterministic Interference Channel with Noisy Channel-Output Feedback

Denote by $\mathcal{C}(\vec{n}_{11}, \vec{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ the capacity region of the LD-IC-NOF with parameters \vec{n}_{11} , \vec{n}_{22} , n_{12} , n_{21} , \overleftarrow{n}_{11} , and \overleftarrow{n}_{22} . Theorem 1 (in the top of this page) fully characterizes this capacity region. The proof of Theorem 1 is divided into two parts. The first part describes the achievable region and is presented in Appendix A. The second part describes the converse region and is presented in Appendix B.

Theorem 1 generalizes previous results regarding the capacity region of the LD-IC with channel-output feedback. For instance, when $\overleftarrow{n}_{11} = 0$ and $\overleftarrow{n}_{22} = 0$, Theorem 1

Theorem 1: The capacity region $\mathcal{C}(\vec{n}_{11}, \vec{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ of the two-user LD-IC-NOF is the set of non-negative rate pairs (R_1, R_2) that satisfy for all $i \in \{1, 2\}$, with $j \in \{1, 2\} \setminus \{i\}$:

$$R_i \leq \min(\max(\vec{n}_{ii}, n_{ji}), \max(\vec{n}_{ii}, n_{ij})), \quad (15a)$$

$$R_i \leq \min(\max(\vec{n}_{ii}, n_{ji}), \max(\vec{n}_{ii}, \overleftarrow{n}_{jj} - (\vec{n}_{jj} - n_{ji})^+)), \quad (15b)$$

$$R_1 + R_2 \leq \min(\max(\vec{n}_{22}, n_{12}) + (\vec{n}_{11} - n_{12})^+, \max(\vec{n}_{11}, n_{21}) + (\vec{n}_{22} - n_{21})^+), \quad (15c)$$

$$R_1 + R_2 \leq \max\left((\vec{n}_{11} - n_{12})^+, n_{21}, \vec{n}_{11} - (\max(\vec{n}_{11}, n_{12}) - \overleftarrow{n}_{11})^+\right) \\ + \max\left((\vec{n}_{22} - n_{21})^+, n_{12}, \vec{n}_{22} - (\max(\vec{n}_{22}, n_{21}) - \overleftarrow{n}_{22})^+\right), \quad (15d)$$

$$2R_i + R_j \leq \max(\vec{n}_{ii}, n_{ji}) + (\vec{n}_{ii} - n_{ij})^+ + \max\left((\vec{n}_{jj} - n_{ji})^+, n_{ij}, \vec{n}_{jj} - (\max(\vec{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+\right). \quad (15e)$$

describes the capacity region of the LD-IC without feedback (Lemma 4 in [28]); when $\overleftarrow{n}_{11} \geq \max(\vec{n}_{11}, n_{12})$ and $\overleftarrow{n}_{22} \geq \max(\vec{n}_{22}, n_{21})$, Theorem 1 describes the capacity region of the LD-IC with perfect channel output feedback (Corollary 1 in [3]); when $\vec{n}_{11} = \vec{n}_{22}$, $n_{12} = n_{21}$ and $\overleftarrow{n}_{11} = \overleftarrow{n}_{22}$, Theorem 1 describes the capacity region of the symmetric LD-IC with noisy channel output feedback (Theorem 1 in [6] and Theorem 4.1, case 1001 in [23]); and when $\vec{n}_{11} = \vec{n}_{22}$, $n_{12} = n_{21}$, $\overleftarrow{n}_{ii} \geq \max(\vec{n}_{ii}, n_{ij})$ and $\overleftarrow{n}_{jj} = 0$, with $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$, Theorem 1 describes the capacity region of the symmetric LD-IC with only one perfect channel output feedback (Theorem 4.1, cases 1000 and 0001 in [23]).

An interesting observation from Theorem 1 is that feedback is beneficial only when at least one of the feedback parameters \overleftarrow{n}_{11} or \overleftarrow{n}_{22} is beyond a certain threshold [29]. For instance, note that when $\overleftarrow{n}_{ii} \leq (\vec{n}_{ii} - n_{ij})^+$, receiver i is unable to send to its corresponding transmitter via feedback any information about the message sent by transmitter j , and thus, feedback does not play any role for enlarging the capacity region. This is basically because the bit-pipes that are subject to interference at receiver i are not included in the set of bit-pipes that are above the (feedback) noise level. However, the threshold $(\vec{n}_{ii} - n_{ij})^+$ for \overleftarrow{n}_{ii} is necessary but not sufficient for feedback to enlarge the capacity region. Consider for instance the following examples.

Example 1: Consider the LD-IC-NOF with parameters $\vec{n}_{11} = 5$, $\vec{n}_{22} = 1$, $\vec{n}_{12} = 3$, $\vec{n}_{21} = 4$, and $\overleftarrow{n}_{22} = 0$. The capacity regions $\mathcal{C}(5, 1, 3, 4, 0, 0)$ and $\mathcal{C}(5, 1, 3, 4, 4, 0)$ are shown in Figure 4a. In this case, channel-output feedback in transmitter-receiver pair 1 enlarges the capacity region only when $\overleftarrow{n}_{11} > \vec{n}_{22} + (\vec{n}_{11} - n_{12})^+ = 3$. More specifically, for all $\overleftarrow{n}_{11} \in \{0, \dots, 3\}$,

$$\mathcal{C}(5, 1, 3, 4, \overleftarrow{n}_{11}, 0) = \mathcal{C}(5, 1, 3, 4, 0, 0)$$

and for all $\overleftarrow{n}_{11} \in \{4, 5, \dots, \infty\}$,

$$\mathcal{C}(5, 1, 3, 4, 0, 0) \subset \mathcal{C}(5, 1, 3, 4, \overleftarrow{n}_{11}, 0).$$

In Example 1, in the absence of channel-output feedback, the rate R_2 is upper-bounded by 1 bit per channel use, whereas the sum-rate $R_1 + R_2$ is upper-bounded by 5 bits per channel use. Figure 4b shows a simple achievability scheme for the rate pair (3, 1). Note that R_2 cannot be improved letting transmitter 2

use the bit-pipes $X_{2,n}^{(2:5)}$ as they are not observed at receiver 2. When channel-output feedback is available at least at transmitter-receiver pair 1 and the bit-pipe from transmitter 2 ending at $\overrightarrow{Y}_{1,n}^{(4)}$ is included in the feedback signal $\overleftarrow{Y}_{i,n}$, the bit-pipe $X_{2,n}^{(2)}$ can be used by transmitter 2 as feedback provides a path between transmitter 2 and receiver 2: transmitter 2 – receiver 1 – transmitter 1 – receiver 2. For this alternative path to become available at least the $(\vec{n}_{22} + (\vec{n}_{11} - n_{12})^+ + 1)$ -th (feedback) bit-pipe from receiver 1 to transmitter 1 must be above the noise level, i.e., $\overleftarrow{n}_{11} > \vec{n}_{22} + (\vec{n}_{11} - n_{12})^+$.

Example 2: Consider an LD-IC-NOF with parameters $\vec{n}_{11} = 7$, $\vec{n}_{22} = 7$, $n_{12} = 3$, $n_{21} = 5$, and $\overleftarrow{n}_{22} = 0$. The capacity regions $\mathcal{C}(7, 7, 3, 5, 0, 0)$ and $\mathcal{C}(7, 7, 3, 5, 6, 0)$ are shown in Figure 5a. In this case, channel-output feedback in transmitter-receiver pair 1 enlarges the capacity region only when $\overleftarrow{n}_{11} > \max(n_{21}, (\vec{n}_{11} - n_{12})^+) = 5$. More specifically, for all $\overleftarrow{n}_{11} \in \{0, 1, \dots, 5\}$,

$$\mathcal{C}(7, 7, 3, 5, \overleftarrow{n}_{11}, 0) = \mathcal{C}(7, 7, 3, 5, 0, 0),$$

and for all $\overleftarrow{n}_{11} \in \{6, 7, \dots, \infty\}$,

$$\mathcal{C}(7, 7, 3, 5, 0, 0) \subset \mathcal{C}(7, 7, 3, 5, \overleftarrow{n}_{11}, 0).$$

In Example 2, in the absence of feedback, the sum-rate capacity can be achieved by simultaneously using two groups of bit-pipes: (a) All bit-pipes starting at transmitter i and being exclusively observed by receiver i ; and (b) All bit-pipes starting at transmitter i that are observed at receiver j but do not interfere with the first group of bit-pipes, with $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$. Figure 5b shows an achievability scheme that uses this idea and achieves the sum-rate capacity. Note that using other bit-pipes to increase any of the individual rates produces interference that cannot be resolved and thus, impedes reliable decoding. In particular note that $X_{2,n}^{(2)}$ and $X_{2,n}^{(3)}$ must remain unused. When feedback is available at least at transmitter-receiver pair 1 and the bit-pipe from transmitter 2 ending at $\overrightarrow{Y}_{1,n}^{(6)}$ is included in the feedback signal $\overleftarrow{Y}_{1,n}$, the bit-pipe $X_{2,n}^{(2)}$ can be used for transmitting maximum-entropy i.i.d. bits for increasing the individual rate R_2 and the sum-rate (see Figure 5c). This is mainly because the bits $X_{2,n}^{(2)}$ can be decoded by transmitter 1 via feedback and be re-transmitted to resolve interference at receiver 1. Interestingly, during the re-transmission by transmitter 1 these bits produce an interference

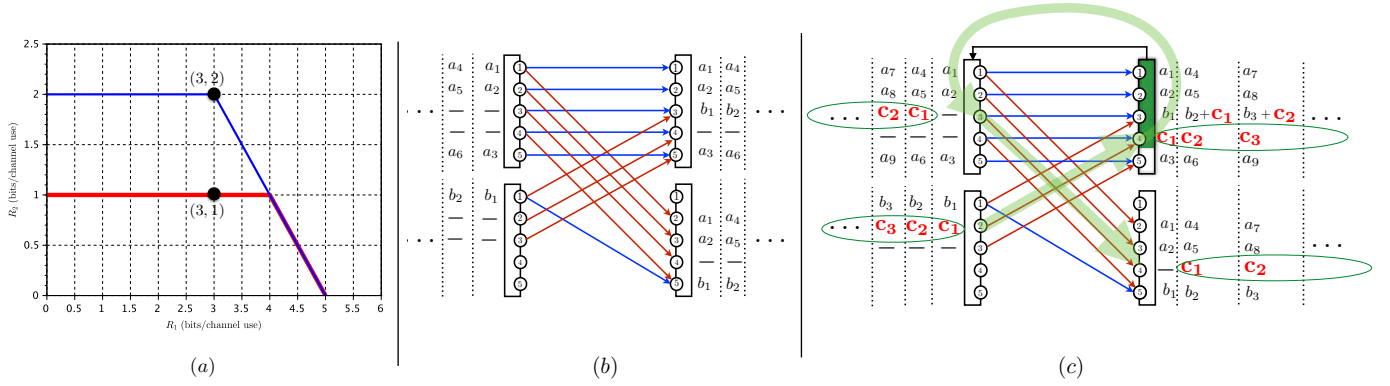


Fig. 4. (a) Capacity regions of $\mathcal{C}(5, 1, 3, 4, 0, 0)$ (thick red line) and $\mathcal{C}(5, 1, 3, 4, 4, 0)$ (thin blue line). (b) Achievability of the rate pair $(3, 1)$ in an LD-IC with parameters $\overline{n}_{11} = 5$, $\overline{n}_{22} = 1$, $n_{12} = 3$, $n_{21} = 4$, $\overline{n}_{11} = 0$ and $\overline{n}_{22} = 0$ (no feedback links). (c) Achievability of the rate pair $(3, 2)$ in an LD-IC with parameters $\overline{n}_{11} = 5$, $\overline{n}_{22} = 1$, $n_{12} = 3$, $n_{21} = 4$, $\overline{n}_{11} = 4$ and $\overline{n}_{22} = 0$.

that can be resolved by receiver 2, as these bits have been received interference-free in the previous channel uses. Note that for this to be possible, at least one of the bit-pipes of transmitter 2 that do not belong to either of the two groups mentioned above, i.e., $X_{2,n}^{(2)}$ and $X_{2,n}^{(3)}$, must be observed above the noise level in the feedback link of transmitter-receiver pair 1, i.e., $\overline{n}_{11} > 5$.

The exact thresholds for the feedback parameters \overline{n}_{11} or \overline{n}_{22} beyond which the capacity region is enlarged is strongly dependent on the parameters, i.e., \overline{n}_{11} , \overline{n}_{22} , n_{12} , and n_{21} . However, the calculation of these thresholds is involved. Thus, this analysis is left out of the scope of this paper. The interested reader might refer to [29] for a full-characterization of these thresholds.

1) *Comments on the Achievability Scheme:* The achievable region is obtained using a coding scheme that combines classical tools such as rate splitting, superposition coding, and backward decoding. This coding scheme is described in Appendix A. In the following, an intuitive description of this coding scheme is presented. Let the message index sent by transmitter i during the t -th block be denoted by $W_i^{(t)} \in \{1, 2, \dots, 2^{NR_i}\}$. Following a rate-splitting argument, assume that $W_i^{(t)}$ is represented by three subindices $(W_{i,C1}^{(t)}, W_{i,C2}^{(t)}, W_{i,P}^{(t)}) \in \{1, 2, \dots, 2^{NR_{i,C1}}\} \times \{1, 2, \dots, 2^{NR_{i,C2}}\} \times \{1, 2, \dots, 2^{NR_{i,P}}\}$, where $R_{i,C1} + R_{i,C2} + R_{i,P} = R_i$. The codeword generation from $(W_{i,C1}^{(t)}, W_{i,C2}^{(t)}, W_{i,P}^{(t)})$ follows a four-level superposition coding scheme. The index $W_{i,C1}^{(t-1)}$ is assumed to be decoded at transmitter j via the feedback link of transmitter-receiver pair j at the end of the transmission of block $t-1$. Therefore, at the beginning of block t , each transmitter possesses the knowledge of the indices $W_{1,C1}^{(t-1)}$ and $W_{2,C1}^{(t-1)}$. In the case of the first block $t=1$, the indices $W_{1,C1}^{(0)}$ and $W_{2,C1}^{(0)}$ correspond to two indices assumed to be known by all transmitters and receivers. Using these indices both transmitters are able to identify the same codeword in the first code-layer. This first code-layer is a sub-codebook of $2^{N(R_{1,C1}+R_{2,C1})}$ codewords (see Figure 8). Denote by $\mathbf{u}(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)})$ the corresponding codeword in

the first code-layer. The second codeword is chosen by transmitter i using $W_{i,C1}^{(t)}$ from the second code-layer, which is a sub-codebook of $2^{NR_{i,C1}}$ codewords corresponding at $\mathbf{u}(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)})$ as shown in Figure 8. Denote by $\mathbf{u}_i(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)})$ the corresponding codeword in the second code-layer. The third codeword is chosen by transmitter i using $W_{i,C2}^{(t)}$ from the third code-layer, which is a sub-codebook of $2^{NR_{i,C2}}$ codewords corresponding at $\mathbf{u}_i(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)})$ as shown in Figure 8. Denote by $\mathbf{v}_i(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)})$ the corresponding codeword in the third code-layer. The fourth codeword is chosen by transmitter i using $W_{i,P}^{(t)}$ from the fourth code-layer, which is a sub-codebook of $2^{NR_{i,P}}$ codewords corresponding at $\mathbf{v}_i(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)})$ as shown in Figure 8. Denote by $\mathbf{x}_{i,P}(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)}, W_{i,P}^{(t)})$ the corresponding codeword in the fourth code-layer. Finally, the generation of the codeword $\mathbf{x}_i = (\mathbf{x}_{i,1}, \mathbf{x}_{i,2}, \dots, \mathbf{x}_{i,N}) \in \mathcal{X}_i^N$ during block $t \in \{1, 2, \dots, T\}$ is a simple concatenation of the codewords $\mathbf{u}_i(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)})$, $\mathbf{v}_i(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)})$ and $\mathbf{x}_{i,P}(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)}, W_{i,P}^{(t)})$, i.e., $\mathbf{x}_i = (\mathbf{u}_i^T, \mathbf{v}_i^T, \mathbf{x}_{i,P}^T)^T$, where the message indices have been dropped for ease of notation.

The intuition to build this code structure follows from the identification of three types of bit-pipes that start at transmitter i : (a) The set of bit-pipes that are observed by receiver j but not necessarily by receiver i and are above the (feedback) noise level; (b) The set of bit-pipes that are observed by receiver j but not necessarily by receiver i and are below the (feedback) noise level; and (c) The set of bit-pipes that are exclusively observed by receiver i . The first set of bit-pipes can be used to convey message index $W_{i,C1}^{(t)}$ from transmitter i to receiver j and to transmitter j during block t . The second set of bit-pipes can be used to convey message index $W_{i,C2}^{(t)}$ from transmitter i to receiver j and not to transmitter j during block t . The third set of bit-pipes can be used to convey message index $W_{i,P}^{(t)}$ from transmitter i to receiver i during block t .

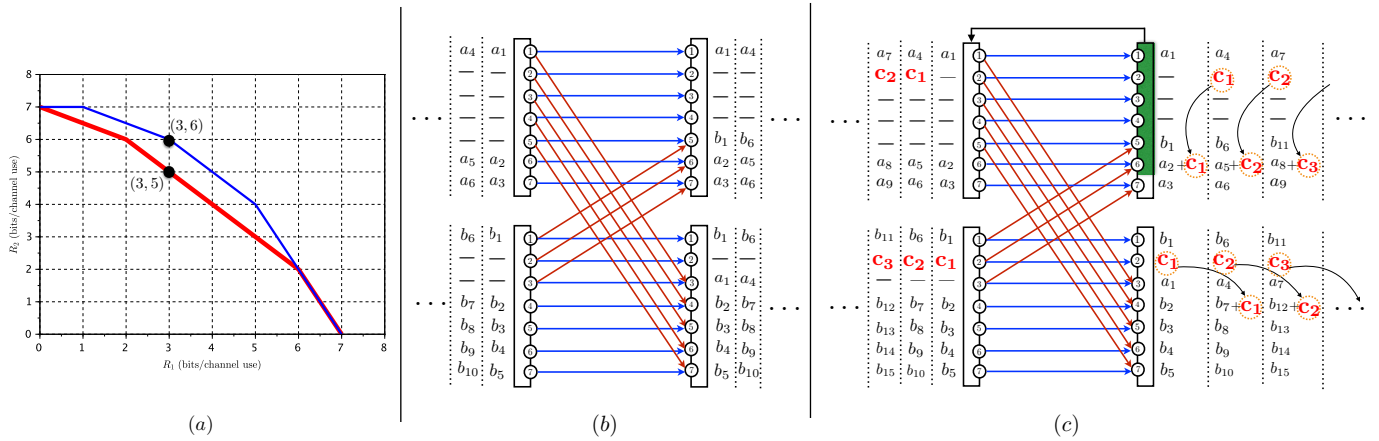


Fig. 5. (a) Capacity regions of $C(7, 7, 3, 5, 0, 0)$ (thick red line) and $C(7, 7, 3, 5, 6, 0)$ (thin blue line). (b) Achievability of the rate pair (3, 5) in an LD-IC with parameters $\vec{n}_{11} = 7$, $\vec{n}_{22} = 7$, $n_{12} = 3$, $n_{21} = 5$, $\vec{n}_{11} = 0$ and $\vec{n}_{22} = 0$ (no feedback links). (c) Achievability of the rate pair (3, 6) in an LD-IC with parameters $\vec{n}_{11} = 7$, $\vec{n}_{22} = 7$, $n_{12} = 3$, $n_{21} = 5$, $\vec{n}_{11} = 6$ and $\vec{n}_{22} = 0$.

These three types of bit-pipes justify the three code-layers super-posed over a common layer, which is justified by the fact that feedback allows both transmitters to decode part of the message sent by each other. The decoder follows a classical backward decoding scheme. This coding/decoding scheme is thoroughly described in Appendix A in the most general case. Later, it is particularized for the case of the LD-IC and G-IC. Other achievable schemes, as reported in [6], can also be obtained as special cases of the more general scheme presented in [17]. However, in this more general case, the resulting code for the IC-NOF counts with a handful of unnecessary superposing code-layers, which demands further optimization. This observation becomes clearer in the analysis of the G-IC-NOF in Section V.

2) *Comments on the Converse region:* The outer bounds (15a) and (15c) are cut-set bounds and were first reported in [28] for the case without feedback. These outer bounds are still useful in the case of perfect channel-output feedback [3]. The outer bounds (15b), (15d) and (15e) are new. Consider the notation used in Appendix B (See Figure 10 and Figure 11). The outer bound (15b) on the individual rate i is a cut-set bound at the input of an enhanced version of receiver i . More specifically, this outer bound is calculated considering that receiver i possesses the message index of transmitter j , i.e., W_j , as side information and observes the channel output \vec{Y}_i and the feedback signal \vec{Y}_j of transmitter-receiver pair j at each channel use. A complete proof of (15b) is presented in Appendix B.

The intuition behind the outer bound (15d) follows from the observation that in the absence of feedback, the sum-rate is upper-bounded by the sum of the bit-pipes from transmitter i that are exclusively observed by receiver i (denoted by $\mathbf{X}_{i,P}$) and the bit-pipes from transmitter i that are observed by receiver j and do not interfere with bit-pipes $\mathbf{X}_{j,P}$ (denoted by $\mathbf{X}_{i,U}$), with $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$. More specifically, in the absence of feedback:

$$R_1 + R_2 \leq \sum_{i=1}^2 \dim \mathbf{X}_{i,P} + \dim \mathbf{X}_{i,U}. \quad (18)$$

When $R_1 + R_2 = \sum_{i=1}^2 \dim \mathbf{X}_{i,P} + \dim \mathbf{X}_{i,U}$ is achievable without feedback, the bit-pipes $\mathbf{X}_{i,P}$ and $\mathbf{X}_{i,U}$ can be used for sending maximum-entropy i.i.d bits from transmitter i to receiver i , which maximizes the sum-rate. Interestingly, any attempt of using any of the other bit-pipes creates interference that cannot be resolved and thus impedes reliable decoding. This observation is formally proved in Appendix B (see proof of (15d)). Note also that this outer bound is not necessarily tight (see Example 1). When feedback is available at least at transmitter-receiver pair i , other bit-pipes different from $\mathbf{X}_{j,P}$ and $\mathbf{X}_{j,U}$ might be used by transmitter j for simultaneously increasing the rate R_j and the sum-rate (see Example 2). This simple observation suggests that there must exist an upper-bound on the sum-rate of the form:

$$R_1 + R_2 \leq \sum_{i=1}^2 \dim \mathbf{X}_{i,P} + \dim \mathbf{X}_{i,U} + F_i, \quad (19)$$

where, $F_i \leq \dim \mathbf{X}_{i,C} + \dim \mathbf{X}_{i,D}$ represents the bit-pipes other than $\mathbf{X}_{i,P}$ and $\mathbf{X}_{i,U}$, whose origin is at transmitter i , that can be used for sending maximum entropy i.i.d bits from transmitter i to receiver i , while generating an interference that can be resolved by the use of feedback. Following this idea, the following outer bound is presented in Appendix B (see proof of (15d)):

$$R_1 + R_2 \leq \sum_{i=1}^2 \dim \mathbf{X}_{i,P} + \dim \mathbf{X}_{i,U} + \dim \mathbf{X}_{i,CF_j} + \dim \mathbf{X}_{i,DF}, \quad (20)$$

where $\dim(\mathbf{X}_{i,CF_j}, \mathbf{X}_{i,DF})$ is the number of the bit-pipes whose origin is at transmitter i and are observed above the noise level in the feedback link of transmitter-receiver pair j . The outer bound (20) is derived considering genie-aided receivers. More specifically, receiver i has inputs \vec{Y}_i and \vec{Y}_i , with $i \in \{1, 2\}$.

A similar reasoning is followed to derive the outer bound (15e) considering three genie-aided receivers. More specifically, receiver i has inputs \vec{Y}_i and \vec{Y}_i , with $i \in \{1, 2\}$, and a

third receiver has inputs $\overrightarrow{\mathbf{Y}}_i$, $\overleftarrow{\mathbf{Y}}_j$, and W_j for at most one $i \in \{1, 2\}$, with $j \in \{1, 2\} \setminus \{i\}$.

V. MAIN RESULTS

This section introduces an achievable region (Theorem 2) and a converse region (Theorem 3), denoted by $\underline{\mathcal{C}}_{\text{G-IC-NOF}}$ and $\overline{\mathcal{C}}_{\text{G-IC-NOF}}$ respectively, for the two-user G-IC-NOF with fixed parameters $\overrightarrow{\text{SNR}}_1$, $\overrightarrow{\text{SNR}}_2$, INR_{12} , INR_{21} , $\overleftarrow{\text{SNR}}_1$, and $\overleftarrow{\text{SNR}}_2$. In general, the capacity region of a given multi-user channel is said to be approximated to within a constant gap according to the following definition.

Definition 2 (Approximation to within ξ units):

A closed and convex set $\mathcal{T} \subset \mathbb{R}_+^m$ is approximated to within ξ units by the sets $\underline{\mathcal{T}}$ and $\overline{\mathcal{T}}$ if $\underline{\mathcal{T}} \subseteq \mathcal{T} \subseteq \overline{\mathcal{T}}$ and for all $\mathbf{t} = (t_1, t_2, \dots, t_m) \in \overline{\mathcal{T}}$, $((t_1 - \xi)^+, (t_2 - \xi)^+, \dots, (t_m - \xi)^+) \in \underline{\mathcal{T}}$.

Denote by $\mathcal{C}_{\text{G-IC-NOF}}$ the capacity region of the 2-user G-IC-NOF. The achievable region $\underline{\mathcal{C}}_{\text{G-IC-NOF}}$ and the converse region $\overline{\mathcal{C}}_{\text{G-IC-NOF}}$ approximate the capacity region $\mathcal{C}_{\text{G-IC-NOF}}$ to within 4.4 bits (Theorem 4).

A. An Achievable Region for the Two-User G-IC-NOF

The description of the achievable region $\underline{\mathcal{C}}_{\text{G-IC-NOF}}$ is presented using the constants $a_{1,i}$; the functions $a_{2,i} : [0, 1] \rightarrow \mathbb{R}_+$, $a_{1,i} : [0, 1]^2 \rightarrow \mathbb{R}_+$, with $l \in \{3, \dots, 6\}$; and $a_{7,i} : [0, 1]^3 \rightarrow \mathbb{R}_+$, which are defined as follows, for all $i \in \{1, 2\}$, with $j \in \{1, 2\} \setminus \{i\}$:

$$a_{1,i} = \frac{1}{2} \log \left(2 + \frac{\overrightarrow{\text{SNR}}_i}{\text{INR}_{ji}} \right) - \frac{1}{2}, \quad (21a)$$

$$a_{2,i}(\rho) = \frac{1}{2} \log \left(b_{1,i}(\rho) + 1 \right) - \frac{1}{2}, \quad (21b)$$

$$a_{3,i}(\rho, \mu) = \frac{1}{2} \log \left(\frac{\overleftarrow{\text{SNR}}_i (b_{2,i}(\rho) + 2) + b_{1,i}(1) + 1}{\overleftarrow{\text{SNR}}_i ((1-\mu)b_{2,i}(\rho) + 2) + b_{1,i}(1) + 1} \right), \quad (21c)$$

$$a_{4,i}(\rho, \mu) = \frac{1}{2} \log \left((1-\mu)b_{2,i}(\rho) + 2 \right) - \frac{1}{2}, \quad (21d)$$

$$a_{5,i}(\rho, \mu) = \frac{1}{2} \log \left(2 + \frac{\overrightarrow{\text{SNR}}_i}{\text{INR}_{ji}} + (1-\mu)b_{2,i}(\rho) \right) - \frac{1}{2}, \quad (21e)$$

$$a_{6,i}(\rho, \mu) = \frac{1}{2} \log \left(\frac{\overrightarrow{\text{SNR}}_i}{\text{INR}_{ji}} \left((1-\mu)b_{2,j}(\rho) + 1 \right) + 2 \right) - \frac{1}{2}, \quad (21f)$$

and

$$a_{7,i}(\rho, \mu_1, \mu_2) = \frac{1}{2} \log \left(\frac{\overrightarrow{\text{SNR}}_i}{\text{INR}_{ji}} \left((1-\mu_i)b_{2,j}(\rho) + 1 \right) + (1-\mu_j)b_{2,i}(\rho) + 2 \right) - \frac{1}{2}, \quad (21g)$$

where the functions $b_{l,i} : [0, 1] \rightarrow \mathbb{R}_+$, with $(l, i) \in \{1, 2\}^2$ are defined as follows:

$$b_{1,i}(\rho) = \overrightarrow{\text{SNR}}_i + 2\rho \sqrt{\overrightarrow{\text{SNR}}_i \text{INR}_{ij}} + \text{INR}_{ij} \quad (22a)$$

$$b_{2,i}(\rho) = (1-\rho)\text{INR}_{ij} - 1, \quad (22b)$$

with $j \in \{1, 2\} \setminus \{i\}$.

Note that the functions in (21) and (22) depend on $\overrightarrow{\text{SNR}}_1$, $\overrightarrow{\text{SNR}}_2$, INR_{12} , INR_{21} , $\overleftarrow{\text{SNR}}_1$, and $\overleftarrow{\text{SNR}}_2$, however as these parameters are fixed in this analysis, this dependence is not emphasized in the definition of these functions. Finally, using this notation, Theorem 2 is presented on the next page.

Proof: The proof of Theorem 2 is presented in Appendix A. ■

The achievability scheme presented in Appendix A is general and thus, it can be used for both the LD-IC-NOF and the G-IC-NOF. The special case of the G-IC-NOF is derived in Appendix A.

B. A Converse Region for the Two-User G-IC-NOF

The description of the converse region $\overline{\mathcal{C}}_{\text{G-IC-NOF}}$ is determined by two events denoted by $S_{l_1,1}$ and $S_{l_2,2}$, where $(l_1, l_2) \in \{1, \dots, 5\}^2$. The events are defined as follows:

$$S_{1,i}: \overrightarrow{\text{SNR}}_j < \min(\text{INR}_{ij}, \text{INR}_{ji}), \quad (24a)$$

$$S_{2,i}: \text{INR}_{ji} \leq \overrightarrow{\text{SNR}}_j < \text{INR}_{ij}, \quad (24b)$$

$$S_{3,i}: \text{INR}_{ij} \leq \overrightarrow{\text{SNR}}_j < \text{INR}_{ji}, \quad (24c)$$

$$S_{4,i}: \max(\text{INR}_{ij}, \text{INR}_{ji}) \leq \overrightarrow{\text{SNR}}_j < \text{INR}_{ij}\text{INR}_{ji}, \quad (24d)$$

$$S_{5,i}: \overrightarrow{\text{SNR}}_j \geq \text{INR}_{ij}\text{INR}_{ji}. \quad (24e)$$

Note that for all $i \in \{1, 2\}$, the events $S_{1,i}$, $S_{2,i}$, $S_{3,i}$, $S_{4,i}$, and $S_{5,i}$ are mutually exclusive. This observation shows that given any 4-tuple $(\overrightarrow{\text{SNR}}_1, \overrightarrow{\text{SNR}}_2, \text{INR}_{12}, \text{INR}_{21})$, there always exists one and only one pair of events $(S_{l_1,1}, S_{l_2,2})$, with $(l_1, l_2) \in \{1, \dots, 5\}^2$, that identifies a unique scenario. Note also that the pairs of events $(S_{2,1}, S_{2,2})$ and $(S_{3,1}, S_{3,2})$ are not feasible. In view of this, twenty-three different scenarios can be identified using the events in (24). Once the exact scenario is identified, the converse region is described using the functions $\kappa_{l,i} : [0, 1] \rightarrow \mathbb{R}_+$, with $l \in \{1, \dots, 3\}$; $\kappa_l : [0, 1] \rightarrow \mathbb{R}_+$, with $l \in \{4, 5\}$; $\kappa_{6,l} : [0, 1] \rightarrow \mathbb{R}_+$, with $l \in \{1, \dots, 4\}$; and $\kappa_{7,i,l} : [0, 1] \rightarrow \mathbb{R}_+$, with $l \in \{1, 2\}$. These functions are defined as follows, for all $i \in \{1, 2\}$, with $j \in \{1, 2\} \setminus \{i\}$:

$$\kappa_{1,i}(\rho) = \frac{1}{2} \log \left(b_{1,i}(\rho) + 1 \right), \quad (25a)$$

$$\kappa_{2,i}(\rho) = \frac{1}{2} \log \left(1 + b_{5,j}(\rho) \right) + \frac{1}{2} \log \left(1 + \frac{b_{4,i}(\rho)}{1 + b_{5,j}(\rho)} \right), \quad (25b)$$

$$\kappa_{3,i}(\rho) = \frac{1}{2} \log \left(\frac{(b_{4,i}(\rho) + b_{5,j}(\rho) + 1)\overleftarrow{\text{SNR}}_j}{(b_{1,j}(1) + 1)(b_{4,i}(\rho) + 1)} + 1 \right) + \frac{1}{2} \log \left(b_{4,i}(\rho) + 1 \right), \quad (25c)$$

$$\kappa_4(\rho) = \frac{1}{2} \log \left(1 + \frac{b_{4,1}(\rho)}{1 + b_{5,2}(\rho)} \right) + \frac{1}{2} \log \left(b_{1,2}(\rho) + 1 \right), \quad (25d)$$

$$\kappa_5(\rho) = \frac{1}{2} \log \left(1 + \frac{b_{4,2}(\rho)}{1 + b_{5,1}(\rho)} \right) + \frac{1}{2} \log \left(b_{1,1}(\rho) + 1 \right), \quad (25e)$$

Theorem 2: The capacity region $\mathcal{C}_{G-IC-NOF}$ contains the region $\underline{\mathcal{C}}_{G-IC-NOF}$ given by the closure of the set of all possible non-negative achievable rate pairs (R_1, R_2) that satisfy

$$R_1 \leq \min \left(a_{2,1}(\rho), a_{6,1}(\rho, \mu_1) + a_{3,2}(\rho, \mu_1), a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{4,2}(\rho, \mu_1) \right), \quad (23a)$$

$$R_2 \leq \min \left(a_{2,2}(\rho), a_{3,1}(\rho, \mu_2) + a_{6,2}(\rho, \mu_2), a_{3,1}(\rho, \mu_2) + a_{4,1}(\rho, \mu_2) + a_{1,2} \right), \quad (23b)$$

$$R_1 + R_2 \leq \min \left(a_{2,1}(\rho) + a_{1,2}, a_{1,1} + a_{2,2}(\rho), a_{3,1}(\rho, \mu_2) + a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{7,2}(\rho, \mu_1, \mu_2), \right. \\ \left. a_{3,1}(\rho, \mu_2) + a_{5,1}(\rho, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{5,2}(\rho, \mu_1), a_{3,1}(\rho, \mu_2) + a_{7,1}(\rho, \mu_1, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{1,2} \right), \quad (23c)$$

$$2R_1 + R_2 \leq \min \left(a_{2,1}(\rho) + a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{7,2}(\rho, \mu_1, \mu_2), \right. \\ \left. a_{3,1}(\rho, \mu_2) + a_{1,1} + a_{7,1}(\rho, \mu_1, \mu_2) + 2a_{3,2}(\rho, \mu_1) + a_{5,2}(\rho, \mu_1), a_{2,1}(\rho) + a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{5,2}(\rho, \mu_1) \right), \quad (23d)$$

$$R_1 + 2R_2 \leq \min \left(a_{3,1}(\rho, \mu_2) + a_{5,1}(\rho, \mu_2) + a_{2,2}(\rho) + a_{1,2}, a_{3,1}(\rho, \mu_2) + a_{7,1}(\rho, \mu_1, \mu_2) + a_{2,2}(\rho) + a_{1,2}, \right. \\ \left. 2a_{3,1}(\rho, \mu_2) + a_{5,1}(\rho, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{1,2} + a_{7,2}(\rho, \mu_1, \mu_2) \right), \quad (23e)$$

with $(\rho, \mu_1, \mu_2) \in \left[0, \left(1 - \max\left(\frac{1}{\text{INR}_{12}}, \frac{1}{\text{INR}_{21}}\right)\right)^+\right] \times [0, 1] \times [0, 1]$.

$$\kappa_6(\rho) = \begin{cases} \kappa_{6,1}(\rho) & \text{if } (S_{1,2} \vee S_{2,2} \vee S_{5,2}) \\ & \wedge (S_{1,1} \vee S_{2,1} \vee S_{5,1}) \\ \kappa_{6,2}(\rho) & \text{if } (S_{1,2} \vee S_{2,2} \vee S_{5,2}) \\ & \wedge (S_{3,1} \vee S_{4,1}) \\ \kappa_{6,3}(\rho) & \text{if } (S_{3,2} \vee S_{4,2}) \\ & \wedge (S_{1,1} \vee S_{2,1} \vee S_{5,1}) \\ \kappa_{6,4}(\rho) & \text{if } (S_{3,2} \vee S_{4,2}) \wedge (S_{3,1} \vee S_{4,1}) \end{cases} \quad (25f)$$

$$\kappa_{7,i}(\rho) = \begin{cases} \kappa_{7,i,1}(\rho) & \text{if } (S_{1,i} \vee S_{2,i} \vee S_{5,i}) \\ \kappa_{7,i,2}(\rho) & \text{if } (S_{3,i} \vee S_{4,i}) \end{cases} \quad (25g)$$

where

$$\kappa_{6,1}(\rho) = \frac{1}{2} \log \left(b_{1,1}(\rho) + b_{5,1}(\rho) \text{INR}_{21} \right) - \frac{1}{2} \log \left(1 + \text{INR}_{12} \right) \\ + \frac{1}{2} \log \left(1 + \frac{b_{5,2}(\rho) \overleftarrow{\text{SNR}}_2}{b_{1,2}(1) + 1} \right) \\ + \frac{1}{2} \log \left(b_{1,2}(\rho) + b_{5,1}(\rho) \text{INR}_{21} \right) - \frac{1}{2} \log \left(1 + \text{INR}_{21} \right) \\ + \frac{1}{2} \log \left(1 + \frac{b_{5,1}(\rho) \overleftarrow{\text{SNR}}_1}{b_{1,1}(1) + 1} \right) + \log(2\pi e), \quad (26a)$$

$$\kappa_{6,2}(\rho) = \frac{1}{2} \log \left(b_{6,2}(\rho) + \frac{b_{5,1}(\rho) \text{INR}_{21}}{\overrightarrow{\text{SNR}}_2} \left(\overrightarrow{\text{SNR}}_2 + b_{3,2} \right) \right) \\ - \frac{1}{2} \log \left(1 + \text{INR}_{12} \right) + \frac{1}{2} \log \left(1 + \frac{b_{5,1}(\rho) \overleftarrow{\text{SNR}}_1}{b_{1,1}(1) + 1} \right) \\ + \frac{1}{2} \log \left(b_{1,1}(\rho) + b_{5,1}(\rho) \text{INR}_{21} \right) - \frac{1}{2} \log \left(1 + \text{INR}_{21} \right) \\ + \frac{1}{2} \log \left(1 + \frac{b_{5,2}(\rho)}{\overrightarrow{\text{SNR}}_2} \left(\text{INR}_{12} + \frac{b_{3,2} \overleftarrow{\text{SNR}}_2}{b_{1,2}(1) + 1} \right) \right) \\ - \frac{1}{2} \log \left(1 + \frac{b_{5,1}(\rho) \text{INR}_{21}}{\overrightarrow{\text{SNR}}_2} \right) + \log(2\pi e), \quad (26b)$$

$$\kappa_{6,3}(\rho) = \frac{1}{2} \log \left(b_{6,1}(\rho) + \frac{b_{5,1}(\rho) \text{INR}_{21}}{\overrightarrow{\text{SNR}}_1} \left(\overrightarrow{\text{SNR}}_1 + b_{3,1} \right) \right) \\ - \frac{1}{2} \log \left(1 + \text{INR}_{12} \right) + \frac{1}{2} \log \left(1 + \frac{b_{5,2}(\rho) \overleftarrow{\text{SNR}}_2}{b_{1,2}(1) + 1} \right)$$

$$+ \frac{1}{2} \log \left(b_{1,2}(\rho) + b_{5,1}(\rho) \text{INR}_{21} \right) - \frac{1}{2} \log \left(1 + \text{INR}_{21} \right) \\ + \frac{1}{2} \log \left(1 + \frac{b_{5,1}(\rho)}{\overrightarrow{\text{SNR}}_1} \left(\text{INR}_{21} + \frac{b_{3,1} \overleftarrow{\text{SNR}}_1}{b_{1,1}(1) + 1} \right) \right) \\ - \frac{1}{2} \log \left(1 + \frac{b_{5,1}(\rho) \text{INR}_{21}}{\overrightarrow{\text{SNR}}_1} \right) + \log(2\pi e), \quad (26c)$$

$$\kappa_{6,4}(\rho) = \frac{1}{2} \log \left(b_{6,1}(\rho) + \frac{b_{5,1}(\rho) \text{INR}_{21}}{\overrightarrow{\text{SNR}}_1} \left(\overrightarrow{\text{SNR}}_1 + b_{3,1} \right) \right) \\ - \frac{1}{2} \log \left(1 + \text{INR}_{12} \right) - \frac{1}{2} \log \left(1 + \text{INR}_{21} \right) \\ + \frac{1}{2} \log \left(1 + \frac{b_{5,2}(\rho)}{\overrightarrow{\text{SNR}}_2} \left(\text{INR}_{12} + \frac{b_{3,2} \overleftarrow{\text{SNR}}_2}{b_{1,2}(1) + 1} \right) \right) \\ - \frac{1}{2} \log \left(1 + \frac{b_{5,1}(\rho) \text{INR}_{21}}{\overrightarrow{\text{SNR}}_2} \right) \\ - \frac{1}{2} \log \left(1 + \frac{b_{5,1}(\rho) \text{INR}_{21}}{\overrightarrow{\text{SNR}}_1} \right) \\ + \frac{1}{2} \log \left(b_{6,2}(\rho) + \frac{b_{5,1}(\rho) \text{INR}_{21}}{\overrightarrow{\text{SNR}}_2} \left(\overrightarrow{\text{SNR}}_2 + b_{3,2} \right) \right) \\ + \frac{1}{2} \log \left(1 + \frac{b_{5,1}(\rho)}{\overrightarrow{\text{SNR}}_1} \left(\text{INR}_{21} + \frac{b_{3,1} \overleftarrow{\text{SNR}}_1}{b_{1,1}(1) + 1} \right) \right) \\ + \log(2\pi e), \quad (26d)$$

and

$$\kappa_{7,i,1}(\rho) = \frac{1}{2} \log \left(b_{1,i}(\rho) + 1 \right) - \frac{1}{2} \log \left(1 + \text{INR}_{ij} \right) \\ + \frac{1}{2} \log \left(1 + \frac{b_{5,j}(\rho) \overleftarrow{\text{SNR}}_j}{b_{1,j}(1) + 1} \right) \\ + \frac{1}{2} \log \left(b_{1,j}(\rho) + b_{5,i}(\rho) \text{INR}_{ji} \right) \\ + \frac{1}{2} \log \left(1 + b_{4,i}(\rho) + b_{5,j}(\rho) \right) - \frac{1}{2} \log \left(1 + b_{5,j}(\rho) \right) \\ + 2 \log(2\pi e), \quad (27a)$$

$$\begin{aligned}
\kappa_{7,i,2}(\rho) &= \frac{1}{2} \log(b_{1,i}(\rho) + 1) - \frac{1}{2} \log(1 + \text{INR}_{ij}) \\
&\quad - \frac{1}{2} \log(1 + b_{5,j}(\rho)) + \frac{1}{2} \log(1 + b_{4,i}(\rho) + b_{5,j}(\rho)) \\
&\quad + \frac{1}{2} \log\left(1 + (1 - \rho^2) \frac{\overrightarrow{\text{INR}}_{ji}}{\overrightarrow{\text{SNR}}_j} \left(\text{INR}_{ij} + \frac{b_{3,j} \overleftarrow{\text{SNR}}_j}{b_{1,j}(1) + 1}\right)\right) - \frac{1}{2} \log\left(1 + \frac{b_{5,i}(\rho) \text{INR}_{ji}}{\overrightarrow{\text{SNR}}_j}\right) \\
&\quad + \frac{1}{2} \log\left(b_{6,j}(\rho) + \frac{b_{5,i}(\rho) \text{INR}_{ji}}{\overrightarrow{\text{SNR}}_j} (\overrightarrow{\text{SNR}}_j + b_{3,j})\right) \\
&\quad + 2 \log(2\pi e), \tag{27b}
\end{aligned}$$

where, the functions $b_{l,i}$, with $(l, i) \in \{1, 2\}^2$ are defined in (22); $b_{3,i}$ are constants; and the functions $b_{l,i} : [0, 1] \rightarrow \mathbb{R}_+$, with $(l, i) \in \{4, 5, 6\} \times \{1, 2\}$ are defined as follows, with $j \in \{1, 2\} \setminus \{i\}$:

$$b_{3,i} = \overrightarrow{\text{SNR}}_i - 2\sqrt{\overrightarrow{\text{SNR}}_i \text{INR}_{ji}} + \text{INR}_{ji}, \tag{28a}$$

$$b_{4,i}(\rho) = (1 - \rho^2) \overrightarrow{\text{SNR}}_i, \tag{28b}$$

$$b_{5,i}(\rho) = (1 - \rho^2) \text{INR}_{ij}, \tag{28c}$$

$$\begin{aligned}
b_{6,i}(\rho) &= \overrightarrow{\text{SNR}}_i + \text{INR}_{ij} + 2\rho\sqrt{\text{INR}_{ij}} \left(\sqrt{\overrightarrow{\text{SNR}}_i} - \sqrt{\text{INR}_{ji}}\right) \\
&\quad + \frac{\text{INR}_{ij} \sqrt{\text{INR}_{ji}}}{\overrightarrow{\text{SNR}}_i} \left(\sqrt{\text{INR}_{ji}} - 2\sqrt{\overrightarrow{\text{SNR}}_i}\right). \tag{28d}
\end{aligned}$$

Note that the functions in (25), (26), (27), and (28) depend on $\overrightarrow{\text{SNR}}_1$, $\overrightarrow{\text{SNR}}_2$, INR_{12} , INR_{21} , $\overleftarrow{\text{SNR}}_1$, and $\overleftarrow{\text{SNR}}_2$. However, these parameters are fixed in this analysis, and therefore, this dependence is not emphasized in the definition of these functions. Finally, using this notation, Theorem 3 is presented below.

Theorem 3: The capacity region $\mathcal{C}_{\text{G-IC-NOF}}$ is contained within the region $\overline{\mathcal{C}}_{\text{G-IC-NOF}}$ given by the closure of the set of non-negative rate pairs (R_1, R_2) that for all $i \in \{1, 2\}$, with $j \in \{1, 2\} \setminus \{i\}$ satisfy:

$$R_i \leq \min(\kappa_{1,i}(\rho), \kappa_{2,i}(\rho)), \tag{29a}$$

$$R_i \leq \kappa_{3,i}(\rho), \tag{29b}$$

$$R_1 + R_2 \leq \min(\kappa_4(\rho), \kappa_5(\rho)), \tag{29c}$$

$$R_1 + R_2 \leq \kappa_6(\rho), \tag{29d}$$

$$2R_i + R_j \leq \kappa_{7,i}(\rho), \tag{29e}$$

with $\rho \in [0, 1]$.

Proof: The proof of Theorem 3 is presented in Appendix C. ■

The outer bounds (29a) and (29c) play the same role as the outer bounds (15a) and (15c) in the linear deterministic model and have been previously reported in [3] for the case of perfect channel-output feedback. The bounds (29b), (29d), and (29e) correspond to new outer bounds. The intuition for deriving these outer bounds follows along the same steps of those used to prove the outer bounds (15b), (15d), and (15e), respectively. Note the duality between the Gaussian signals $X_{i,C}$ and $X_{i,U}$ (in 70 and 69, respectively) and the bit-pipes $(\mathbf{X}_{i,C}, \mathbf{X}_{i,D})$ and $\mathbf{X}_{i,U}$ (in (54a), (54d) and (58), respectively).

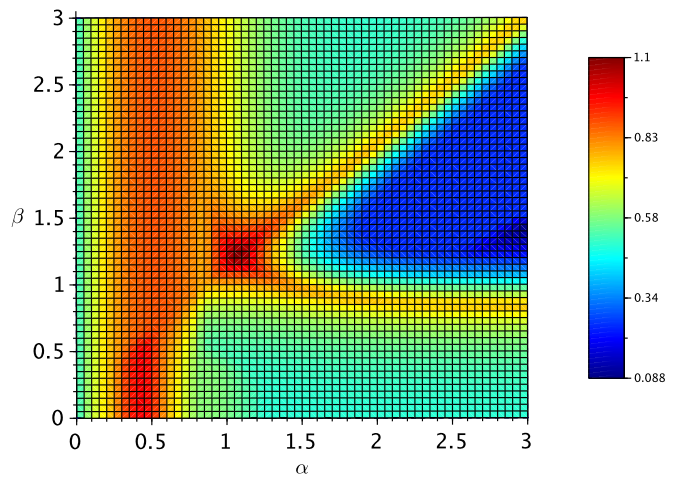


Fig. 6. Gap between the converse region $\overline{\mathcal{C}}_{\text{G-IC-NOF}}$ and the achievable region $\underline{\mathcal{C}}_{\text{G-IC-NOF}}$ of the two-user G-IC-NOF, under symmetric channel conditions, i.e., $\overrightarrow{\text{SNR}}_1 = \overrightarrow{\text{SNR}}_2 = \overrightarrow{\text{SNR}}$, $\text{INR}_{12} = \text{INR}_{21} = \text{INR}$, and $\overleftarrow{\text{SNR}}_1 = \overleftarrow{\text{SNR}}_2 = \overleftarrow{\text{SNR}}$, as a function of $\alpha = \frac{\log \text{INR}}{\log \overrightarrow{\text{SNR}}}$ and $\beta = \frac{\log \overleftarrow{\text{SNR}}}{\log \overrightarrow{\text{SNR}}}$.

C. A Gap Between the Achievable Region and the Converse Region

Theorem 4 describes the gap between the achievable region $\underline{\mathcal{C}}_{\text{G-IC-NOF}}$ and the converse region $\overline{\mathcal{C}}_{\text{G-IC-NOF}}$ (Definition 2).

Theorem 4: The capacity region of the two-user G-IC-NOF is approximated to within 4.4 bits by the achievable region $\underline{\mathcal{C}}_{\text{G-IC-NOF}}$ and the converse region $\overline{\mathcal{C}}_{\text{G-IC-NOF}}$.

Proof: The proof of Theorem 4 is presented in Appendix D. ■

Figure 6 presents the exact gap existing between the achievable region $\underline{\mathcal{C}}_{\text{G-IC-NOF}}$ and the converse region $\overline{\mathcal{C}}_{\text{G-IC-NOF}}$ for the case in which $\overrightarrow{\text{SNR}}_1 = \overrightarrow{\text{SNR}}_2 = \overrightarrow{\text{SNR}}$, $\text{INR}_{12} = \text{INR}_{21} = \text{INR}$, and $\overleftarrow{\text{SNR}}_1 = \overleftarrow{\text{SNR}}_2 = \overleftarrow{\text{SNR}}$ as a function of $\alpha = \frac{\log \text{INR}}{\log \overrightarrow{\text{SNR}}}$ and $\beta = \frac{\log \overleftarrow{\text{SNR}}}{\log \overrightarrow{\text{SNR}}}$. Note that in this case, the maximum gap is 1.1 bits and occurs when $\alpha = 1.05$ and $\beta = 1.2$.

VI. TRANSMITTER COOPERATION VIA FEEDBACK

Channel-Output feedback enables transmitter cooperation. Consider for instance the LD-IC-NOF models in Figure 10 and the notation used in Appendix B. In this model, the feedback signal $\tilde{\mathbf{Y}}_{i,n}$ can be used by transmitter i to obtain the bits sent by the other transmitter through the bit-pipes $\mathbf{X}_{j,CF_i,n-1}$ and $\mathbf{X}_{j,DF,n-1}$, with $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$. Essentially, feedback allows a transmitter to observe at least partially the channel input of the other transmitter, subject to a finite delay. This highlights the strong connections of channel-output feedback and conferencing transmitters.

Consider the LD-IC models with conferencing-transmitters (LD-IC-CT) in Figure 7. Note that the noise level in the link from transmitter i to transmitter j is such that transmitter j observes only the bits sent by transmitter i through the bit-pipes $\mathbf{X}_{i,CF_j,n-1}$ and $\mathbf{X}_{i,DF,n-1}$, for all $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$. Then in both channel models (Figure 7 and

Figure 10) the corresponding transmitters possess the same side information and the corresponding receivers observe the same channel outputs. This observation holds also for the case of a G-IC-NOF and a G-IC with conferencing transmitters (G-IC-CT). For instance, in the case where the noise power in the link from transmitter i to transmitter j in the G-IC-CT is identical to the sum of the noise power at the input of receiver j and the noise power in the feedback link of transmitter-receiver pair j in the G-IC-NOF. Hence, under these conditions, either in the linear deterministic or Gaussian case, any outer bound for the case of conferencing transmitters is also an outer bound for the case of channel-output feedback and vice-versa.

Outer bounds for the LD-IC-CT and the G-IC-CT have been reported in [19] and [21]. However, the gap between the achievable sum-rate and the corresponding converse in the G-IC-CT is 20 bits in [19]. Other outer bounds have been presented in [30] for the G-IC-CT, but a gap between the achievable and converse regions is not reported.

VII. CONCLUDING REMARKS

In this paper, the exact capacity region of the LD-IC-NOF has been fully characterized (Theorem 1). Using key observations in the LD-IC-NOF, an achievability region (Theorem 2) and a converse region (Theorem 3) have been presented for the two-user G-IC-NOF. These two regions approximate the capacity region of the G-IC-NOF to within 4.4 bits (Theorem 4).

Despite the contributions made in this paper, the understanding of the benefits of channel-output feedback are far from being completely understood in the G-IC-NOF. For instance, an interesting extension of these results is to consider the case in which the channel-output feedback is observed by both transmitters subject to (possibly non-independent) additive Gaussian noise processes. This analysis would complement initial attempts (see [23] and references therein) for understanding a larger and more interesting problem suggested by Shannon [31]: the two-way interference channel.

APPENDIX A PROOF OF ACHIEVABILITY

This appendix describes an achievability scheme for the IC-NOF based on a three-part message splitting, superposition coding, and backward decoding. This coding scheme is general and thus, it holds for the two-user LD-IC-NOF and the two-user G-IC-NOF.

Codebook Generation: Fix a strictly positive joint probability distribution

$$\begin{aligned} P_{U_1 U_2 V_1 V_2 X_{1,P} X_{2,P}}(u, u_1, u_2, v_1, v_2, x_{1,P}, x_{2,P}) = \\ P_U(u)P_{U_1|U}(u_1|u)P_{U_2|U}(u_2|u)P_{V_1|U U_1}(v_1|u, u_1) \\ P_{V_2|U U_2}(v_2|u, u_2)P_{X_{1,P}|U U_1 V_1}(x_{1,P}|u, u_1, v_1) \\ P_{X_{2,P}|U U_2 V_2}(x_{2,P}|u, u_2, v_2), \end{aligned} \quad (30)$$

for all $(u, u_1, u_2, v_1, v_2, x_{1,P}, x_{2,P}) \in (\mathcal{X}_1 \cup \mathcal{X}_2) \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_1 \times \mathcal{X}_2$.

Let $R_{1,C1}$, $R_{1,C2}$, $R_{2,C1}$, $R_{2,C2}$, $R_{1,P}$, and $R_{2,P}$ be non-negative real numbers. Let also $R_{1,C} = R_{1,C1} + R_{1,C2}$,

$R_{2,C} = R_{2,C1} + R_{2,C2}$, $R_1 = R_{1,C} + R_{1,P}$, and $R_2 = R_{2,C} + R_{2,P}$.

Generate $2^{N(R_{1,C1} + R_{2,C1})}$ i.i.d. length- N codewords $\mathbf{u}(s, r) = (u_1(s, r), u_2(s, r), \dots, u_N(s, r))$ according to

$$P_U(\mathbf{u}(s, r)) = \prod_{i=1}^N P_{U_i}(u_i(s, r)), \quad (31)$$

with $s \in \{1, 2, \dots, 2^{NR_{1,C1}}\}$ and $r \in \{1, 2, \dots, 2^{NR_{2,C1}}\}$.

For encoder 1, generate for each codeword $\mathbf{u}(s, r)$, $2^{NR_{1,C1}}$ i.i.d. length- N codewords $\mathbf{u}_1(s, r, k) = (u_{1,1}(s, r, k), u_{1,2}(s, r, k), \dots, u_{1,N}(s, r, k))$ according to

$$P_{U_1|U}(\mathbf{u}_1(s, r, k)|\mathbf{u}(s, r)) = \prod_{i=1}^N P_{U_{1,i}|U}(u_{1,i}(s, r, k)|u_i(s, r)), \quad (32)$$

with $k \in \{1, 2, \dots, 2^{NR_{1,C1}}\}$. For each pair of codewords $(\mathbf{u}(s, r), \mathbf{u}_1(s, r, k))$, generate $2^{NR_{1,C2}}$ i.i.d. length- N codewords $\mathbf{v}_1(s, r, k, l) = (v_{1,1}(s, r, k, l), v_{1,2}(s, r, k, l), \dots, v_{1,N}(s, r, k, l))$ according to

$$\begin{aligned} P_{V_1|U U_1}(\mathbf{v}_1(s, r, k, l)|\mathbf{u}(s, r), \mathbf{u}_1(s, r, k)) = \\ \prod_{i=1}^N P_{V_{1,i}|U U_1}(v_{1,i}(s, r, k, l)|u_i(s, r), u_{1,i}(s, r, k)), \end{aligned} \quad (33)$$

with $l \in \{1, 2, \dots, 2^{NR_{1,C2}}\}$. For each tuple of codewords $(\mathbf{u}(s, r), \mathbf{u}_1(s, r, k), \mathbf{v}_1(s, r, k, l))$, generate $2^{NR_{1,P}}$ i.i.d. length- N codewords $\mathbf{x}_{1,P}(s, r, k, l, q) = (x_{1,P,1}(s, r, k, l, q), x_{1,P,2}(s, r, k, l, q), \dots, x_{1,P,N}(s, r, k, l, q))$ according to

$$\begin{aligned} P_{X_{1,P}|U U_1 V_1}(\mathbf{x}_{1,P}(s, r, k, l, q)|\mathbf{u}(s, r), \mathbf{u}_1(s, r, k), \mathbf{v}_1(s, r, k, l)) \\ = \prod_{i=1}^N P_{X_{1,P,i}|U U_1 V_1}(x_{1,P,i}(s, r, k, l, q)|u_i(s, r), u_{1,i}(s, r, k), \\ v_{1,i}(s, r, k, l)), \end{aligned} \quad (34)$$

with $q \in \{1, 2, \dots, 2^{NR_{1,P}}\}$.

For encoder 2, generate for each codeword $\mathbf{u}(s, r)$, $2^{NR_{2,C1}}$ i.i.d. length- N codewords $\mathbf{u}_2(s, r, j) = (u_{2,1}(s, r, j), u_{2,2}(s, r, j), \dots, u_{2,N}(s, r, j))$ according to

$$P_{U_2|U}(\mathbf{u}_2(s, r, j)|\mathbf{u}(s, r)) = \prod_{i=1}^N P_{U_{2,i}|U}(u_{2,i}(s, r, j)|u_i(s, r)), \quad (35)$$

with $j \in \{1, 2, \dots, 2^{NR_{2,C1}}\}$. For each pair of codewords $(\mathbf{u}(s, r), \mathbf{u}_2(s, r, j))$, generate $2^{NR_{2,C2}}$ i.i.d. length- N codewords $\mathbf{v}_2(s, r, j, m) = (v_{2,1}(s, r, j, m), v_{2,2}(s, r, j, m), \dots, v_{2,N}(s, r, j, m))$ according to

$$\begin{aligned} P_{V_2|U U_2}(\mathbf{v}_2(s, r, j, m)|\mathbf{u}(s, r), \mathbf{u}_2(s, r, j)) = \\ \prod_{i=1}^N P_{V_{2,i}|U U_2}(v_{2,i}(s, r, j, m)|u_i(s, r), u_{2,i}(s, r, j)), \end{aligned} \quad (36)$$

with $m \in \{1, 2, \dots, 2^{NR_{2,C2}}\}$. For each tuple of codewords $(\mathbf{u}(s, r), \mathbf{u}_2(s, r, j), \mathbf{v}_2(s, r, j, m))$,

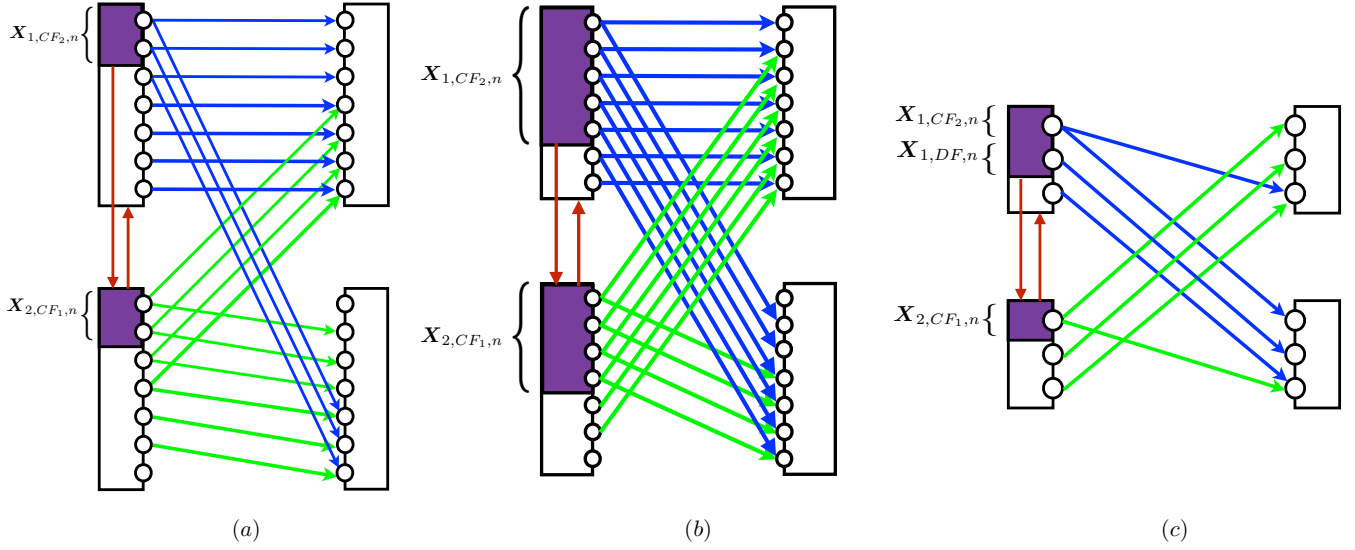


Fig. 7. Linear Deterministic Interference Channels with Conferencing Transmitters.

generate $2^{NR_{2,P}}$ i.i.d. length- N codewords $\mathbf{x}_{2,P}(s, r, j, m, b) = (x_{2,P,1}(s, r, j, m, b), x_{2,P,2}(s, r, j, m, b), \dots, x_{2,P,N}(s, r, j, m, b))$ according to

$$P_{\mathbf{X}_{2,P}} U V_2 V_2^T (\mathbf{x}_{2,P}(s, r, j, m, b) | \mathbf{u}(s, r), \mathbf{u}_2(s, r, j), \mathbf{v}_2(s, r, j, m)) \\ = \prod_{i=1}^N P_{X_{2,P}|U V_2 V_2^T} (x_{2,P,i}(s, r, j, m, b) | u_i(s, r), u_{2,i}(s, r, j), \\ v_{2,i}(s, r, j, m, b)), \quad (37)$$

with $b \in \{1, 2, \dots, 2^{NR_{2,P}}\}$. The resulting code structure is shown in Figure 8.

Encoding: Denote by $W_i^{(t)} \in \{1, 2, \dots, 2^{NR_i}\}$ the message index of transmitter $i \in \{1, 2\}$ during block $t \in \{1, 2, \dots, T\}$, with T the total number of blocks. Let $W_i^{(t)}$ be composed by the message index $W_{i,C}^{(t)} \in \{1, 2, \dots, 2^{NR_{i,C}}\}$ and message index $W_{i,P}^{(t)} \in \{1, 2, \dots, 2^{NR_{i,P}}\}$. That is, $W_i^{(t)} = (W_{i,C}^{(t)}, W_{i,P}^{(t)})$. The message index $W_{i,P}^{(t)}$ must be reliably decoded at receiver i . Let also $W_{i,C}^{(t)}$ be composed by the message indices $W_{i,C1}^{(t)} \in \{1, 2, \dots, 2^{NR_{i,C1}}\}$ and $W_{i,C2}^{(t)} \in \{1, 2, \dots, 2^{NR_{i,C2}}\}$. That is, $W_{i,C}^{(t)} = (W_{i,C1}^{(t)}, W_{i,C2}^{(t)})$. The message index $W_{i,C1}^{(t)}$ must be reliably decoded by the other transmitter (via feedback) and by the non-intended receiver, but not necessarily by the intended receiver. The message index $W_{i,C2}^{(t)}$ must be reliably decoded by the non-intended receiver, but not necessarily by the intended receiver.

Consider Markov encoding over T blocks. At encoding step t , with $t \in \{1, 2, \dots, T\}$, transmitter 1 sends the codeword:

$$\mathbf{x}_1^{(t)} = \Theta_1 \left(\mathbf{u}(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}), \mathbf{u}_1(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{1,C1}^{(t)}), \right. \\ \left. \mathbf{v}_1(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{1,C1}^{(t)}, W_{1,C2}^{(t)}), \right. \\ \left. \mathbf{x}_{1,P}(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{1,C1}^{(t)}, W_{1,C2}^{(t)}, W_{1,P}^{(t)}) \right), \quad (38)$$

where, $\Theta_1 : (\mathcal{X}_1 \cup \mathcal{X}_2)^N \times \mathcal{X}_1^N \times \mathcal{X}_1^N \times \mathcal{X}_1^N \rightarrow \mathcal{X}_1^N$ is a function that transforms the codewords $\mathbf{u}(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)})$, $\mathbf{u}_1(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{1,C1}^{(t)})$, $\mathbf{v}_1(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{1,C1}^{(t)}, W_{1,C2}^{(t)})$, and $\mathbf{x}_{1,P}(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{1,C1}^{(t)}, W_{1,C2}^{(t)}, W_{1,P}^{(t)})$ into the N -dimensional vector $\mathbf{x}_1^{(t)}$ of channel inputs. The indices $W_{1,C1}^{(0)} = W_{1,C1}^{(T)} = s^*$ and $W_{2,C1}^{(0)} = W_{2,C1}^{(T)} = r^*$, and the pair $(s^*, r^*) \in \{1, 2, \dots, 2^{NR_{1,C1}}\} \times \{1, 2, \dots, 2^{NR_{2,C1}}\}$ are pre-defined and known by both receivers and transmitters. It is worth noting that the message index $W_{2,C1}^{(t-1)}$ is obtained by transmitter 1 from the feedback signal $\hat{\mathbf{y}}_1^{(t-1)}$ at the end of the previous encoding step $t-1$.

Transmitter 2 follows a similar encoding scheme.

Decoding: Both receivers decode their message indices at the end of block T in a backward decoding fashion. At each decoding step t , with $t \in \{1, 2, \dots, T\}$, receiver 1 obtains the message indices $(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, \widehat{W}_{1,C2}^{(T-t-1)}, \widehat{W}_{1,P}^{(T-t-1)}, \widehat{W}_{2,C2}^{(T-t-1)}) \in \{1, 2, \dots, 2^{NR_{1,C1}}\} \times \{1, 2, \dots, 2^{NR_{2,C1}}\} \times \{1, 2, \dots, 2^{NR_{1,C2}}\} \times \{1, 2, \dots, 2^{NR_{1,P}}\} \times \{1, 2, \dots, 2^{NR_{2,C2}}\}$ from the channel output $\hat{\mathbf{y}}_1^{(T-t-1)}$. The tuple $(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, \widehat{W}_{1,C2}^{(T-t-1)}, \widehat{W}_{1,P}^{(T-t-1)}, \widehat{W}_{2,C2}^{(T-t-1)})$ is the unique tuple that satisfies

$$\left(\mathbf{u}(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}), \mathbf{u}_1(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, W_{1,C1}^{(T-t-1)}), \right. \\ \left. \mathbf{v}_1(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, W_{1,C1}^{(T-t-1)}, \widehat{W}_{1,C2}^{(T-t-1)}), \right. \\ \left. \mathbf{x}_{1,P}(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, W_{1,C1}^{(T-t-1)}, \widehat{W}_{1,C2}^{(T-t-1)}, \right. \\ \left. \widehat{W}_{1,P}^{(T-t-1)}) \right), \mathbf{u}_2(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, W_{2,C1}^{(T-t-1)}), \\ \left. \mathbf{v}_2(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, W_{2,C1}^{(T-t-1)}, \widehat{W}_{2,C2}^{(T-t-1)}), \right. \\ \left. \hat{\mathbf{y}}_1^{(T-t-1)} \right) \in \mathcal{T}^{(N,e)} \\ [U \ U_1 \ V_1 \ X_{1,P} \ U_2 \ V_2 \ \hat{\mathbf{Y}}_1], \quad (39)$$

where $W_{1,C1}^{(T-(t-1))}$ and $W_{2,C1}^{(T-(t-1))}$ are assumed to be perfectly decoded in the previous decoding step $t-1$. The set $\mathcal{T}^{(N,e)}$ $[U, U_1, V_1, X_{1,P}, U_2, V_2, \vec{Y}_1]$ represents the set of jointly typical sequences of the random variables $U, U_1, V_1, X_{1,P}, U_2, V_2$, and \vec{Y}_1 , with $e > 0$. Receiver 2 follows a similar decoding scheme.

Probability of Error Analysis: An error might occur during encoding step t if the message index $W_{2,C1}^{(t-1)}$ is not correctly decoded at transmitter 1. From the asymptotic equipartition property (AEP) [32], it follows that the message index $W_{2,C1}^{(t-1)}$ can be reliably decoded at transmitter 1 during encoding step t , under the condition:

$$\begin{aligned} R_{2,C1} &\leq I(\vec{Y}_1; U_2 | U, U_1, V_1, X_1) \\ &= I(\vec{Y}_1; U_2 | U, X_1). \end{aligned} \quad (40)$$

An error might occur during the (backward) decoding step t if the message indices $W_{1,C1}^{(T-t)}$, $W_{2,C1}^{(T-t)}$, $W_{1,C2}^{(T-(t-1))}$, $W_{1,P}^{(T-(t-1))}$, and $W_{2,C2}^{(T-(t-1))}$ are not decoded correctly given that the message indices $W_{1,C1}^{(T-(t-1))}$ and $W_{2,C1}^{(T-(t-1))}$ were correctly decoded in the previous decoding step $t-1$. These errors might arise for two reasons: (i) there does not exist a tuple $(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, \widehat{W}_{1,C2}^{(T-(t-1))}, \widehat{W}_{1,P}^{(T-(t-1))}, \widehat{W}_{2,C2}^{(T-(t-1))})$ that satisfies (39), or (ii) there exist several tuples $(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, \widehat{W}_{1,C2}^{(T-(t-1))}, \widehat{W}_{1,P}^{(T-(t-1))}, \widehat{W}_{2,C2}^{(T-(t-1))})$ that simultaneously satisfy (39). From the asymptotic equipartition property (AEP) [32], the probability of an error due to (i) tends to zero when N grows to infinity. Consider the error due to (ii) and define the event $E_{(s,r,l,q,m)}$ that describes the case in which the codewords $(\mathbf{u}(s,r), \mathbf{u}_1(s,r, W_{1,C1}^{(T-(t-1))}), \mathbf{v}_1(s,r, W_{1,C1}^{(T-(t-1))}), l), \mathbf{x}_{1,P}(s,r, W_{1,C1}^{(T-(t-1))}, l, q), \mathbf{u}_2(s,r, W_{2,C1}^{(T-(t-1))}),$ and $\mathbf{v}_2(s,r, W_{2,C1}^{(T-(t-1))}, m))$ are jointly typical with $\vec{y}_1^{(T-(t-1))}$ during decoding step t . Assume now that the codeword to be decoded at decoding step t corresponds to the indices $(s,r,l,q,m) = (1,1,1,1,1)$ without loss of generality due to the symmetry of the code. Then, the probability of error due to (ii) during decoding step t , can be bounded as follows:

$$\begin{aligned} P_e &= \Pr \left[\bigcup_{(s,r,l,q,m) \neq (1,1,1,1,1)} E_{(s,r,l,q,m)} \right] \\ &\leq \sum_{(s,r,l,q,m) \in \mathcal{T}} \Pr [E_{(s,r,l,q,m)}], \end{aligned} \quad (41)$$

with $\mathcal{T} = \{1, 2, \dots, 2^{NR_{1,C1}}\} \times \{1, 2, \dots, 2^{NR_{2,C1}}\} \times \{1, 2, \dots, 2^{NR_{1,C2}}\} \times \{1, 2, \dots, 2^{NR_{1,P}}\} \times \{1, 2, \dots, 2^{NR_{2,C2}}\} \setminus \{(1,1,1,1,1)\}$. From the AEP [32], it follows that

$$\begin{aligned} P_e &\leq 2^{N(R_{2,C2} - I(\vec{Y}_1; V_2 | U, U_1, U_2, V_1, X_1) + 2\epsilon)} \\ &\quad + 2^{N(R_{1,P} - I(\vec{Y}_1; X_1 | U, U_1, U_2, V_1, V_2) + 2\epsilon)} \\ &\quad + 2^{N(R_{2,C2} + R_{1,P} - I(\vec{Y}_1; V_2, X_1 | U, U_1, U_2, V_1) + 2\epsilon)} \\ &\quad + 2^{N(R_{1,C2} - I(\vec{Y}_1; V_1, X_1 | U, U_1, U_2, V_2) + 2\epsilon)} \end{aligned}$$

$$\begin{aligned} &+ 2^{N(R_{1,C2} + R_{2,C2} - I(\vec{Y}_1; V_1, V_2, X_1 | U, U_1, U_2) + 2\epsilon)} \\ &+ 2^{N(R_{1,C2} + R_{1,P} - I(\vec{Y}_1; V_1, X_1 | U, U_1, U_2, V_2) + 2\epsilon)} \\ &+ 2^{N(R_{1,C2} + R_{1,P} + R_{2,C2} - I(\vec{Y}_1; V_1, V_2, X_1 | U, U_1, U_2) + 2\epsilon)} \\ &+ 2^{N(R_{2,C1} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{2,C} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{2,C1} + R_{1,P} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{2,C} + R_{1,P} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{2,C1} + R_{1,C2} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{2,C} + R_{1,C2} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{2,C1} + R_{1,C2} + R_{1,P} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{2,C} + R_{1,C2} + R_{1,P} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{1,C1} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{1,C1} + R_{2,C2} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{1,C1} + R_{1,P} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{1,C1} + R_{1,P} + R_{2,C2} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{1,C} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{1,C} + R_{2,C2} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_1 - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_1 + R_{2,C2} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{1,C1} + R_{2,C1} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{1,C1} + R_{2,C} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{1,C1} + R_{2,C1} + R_{1,P} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{1,C1} + R_{2,C} + R_{1,P} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{1,C} + R_{2,C1} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_{1,C} + R_{2,C} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_1 + R_{2,C1} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)} \\ &+ 2^{N(R_1 + R_{2,C} - I(\vec{Y}_1; U, U_1, U_2, V_1, V_2, X_1) + 2\epsilon)}. \end{aligned} \quad (42)$$

The same analysis of the probability of error holds for transmitter-receiver pair 2. Hence, in general, from (40) and (42), reliable decoding holds under the following conditions for transmitter $i \in \{1, 2\}$, with $j \in \{1, 2\} \setminus \{i\}$:

$$\begin{aligned} R_{j,C1} &\leq I(\vec{Y}_i; U_j | U, U_i, V_i, X_i) \\ &= I(\vec{Y}_i; U_j | U, X_i) \\ &\triangleq \theta_{1,i}, \end{aligned} \quad (43a)$$

$$\begin{aligned} R_i + R_{j,C} &\leq I(\vec{Y}_i; U, U_i, U_j, V_i, V_j, X_i) \\ &= I(\vec{Y}_i; U, U_j, V_j, X_i) \\ &\triangleq \theta_{2,i}, \end{aligned} \quad (43b)$$

$$\begin{aligned} R_{j,C2} &\leq I(\vec{Y}_i; V_j | U, U_i, U_j, V_i, X_i) \\ &= I(\vec{Y}_i; V_j | U, U_j, X_i) \\ &\triangleq \theta_{3,i}, \end{aligned} \quad (43c)$$

$$\begin{aligned} R_{i,P} &\leq I(\vec{Y}_i; X_i | U, U_i, U_j, V_i, V_j) \\ &\triangleq \theta_{4,i}, \end{aligned} \quad (43d)$$

$$\begin{aligned} R_{i,P} + R_{j,C2} &\leq I(\vec{Y}_i; V_j, X_i | U, U_i, U_j, V_i) \\ &\triangleq \theta_{5,i}, \end{aligned} \quad (43e)$$

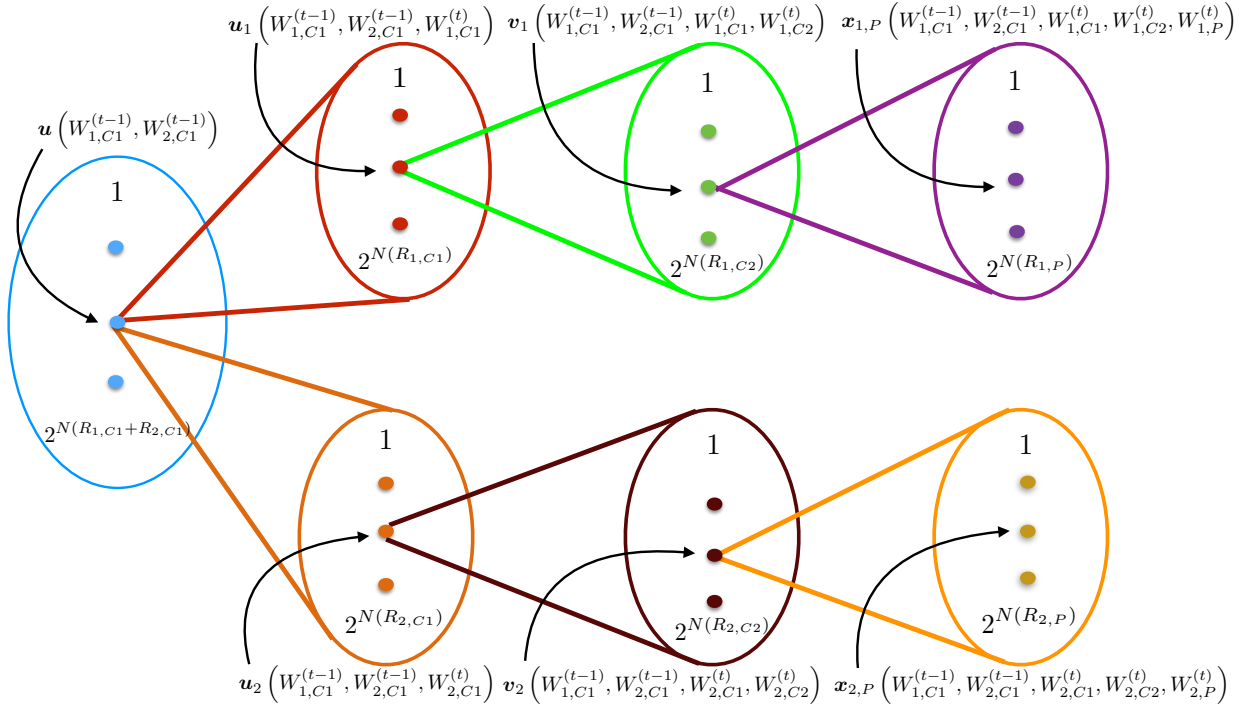


Fig. 8. Structure of the superposition code. The codewords corresponding to the message indices $W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{1,C1}^{(t)}, W_{1,C2}^{(t)}, W_{1,P}^{(t)}$ with $i \in \{1, 2\}$ as well as the block index t are both highlighted. The (approximate) number of codewords for each code layer is also highlighted.

$$\begin{aligned} R_{i,C2} + R_{i,P} &\leq I(\vec{Y}_i; V_i, X_i | U, U_i, U_j, V_j) \\ &= I(\vec{Y}_i; X_i | U, U_i, U_j, V_j) \\ &\triangleq \theta_{6,i}, \text{ and} \end{aligned} \quad (43f)$$

$$\begin{aligned} R_{i,C2} + R_{i,P} + R_{j,C2} &\leq I(\vec{Y}_i; V_i, V_j, X_i | U, U_i, U_j) \\ &= I(\vec{Y}_i; V_j, X_i | U, U_i, U_j) \\ &\triangleq \theta_{7,i}. \end{aligned} \quad (43g)$$

Taking into account that $R_i = R_{i,C1} + R_{i,C2} + R_{i,P}$, a Fourier-Motzkin elimination process in (43) yields:

$$R_1 \leq \min(\theta_{2,1}, \theta_{6,1} + \theta_{1,2}, \theta_{4,1} + \theta_{1,2} + \theta_{3,2}), \quad (44a)$$

$$R_2 \leq \min(\theta_{2,2}, \theta_{1,1} + a_{6,2}, \theta_{1,1} + \theta_{3,1} + \theta_{4,2}), \quad (44b)$$

$$\begin{aligned} R_1 + R_2 &\leq \min(\theta_{2,1} + \theta_{4,2}, \theta_{2,1} + a_{6,2}, \theta_{4,1} + \theta_{2,2}, \\ &\quad \theta_{6,1} + \theta_{2,2}, \theta_{1,1} + \theta_{3,1} + \theta_{4,1} + \theta_{1,2} + \theta_{5,2}, \\ &\quad \theta_{1,1} + \theta_{7,1} + \theta_{1,2} + \theta_{5,2}, \\ &\quad \theta_{1,1} + \theta_{4,1} + \theta_{1,2} + \theta_{7,2}, \\ &\quad \theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{3,2} + \theta_{4,2}, \\ &\quad \theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{5,2}, \\ &\quad \theta_{1,1} + \theta_{7,1} + \theta_{1,2} + \theta_{4,2}), \end{aligned} \quad (44c)$$

$$\begin{aligned} 2R_1 + R_2 &\leq \min(\theta_{2,1} + \theta_{4,1} + \theta_{1,2} + \theta_{7,2}, \\ &\quad \theta_{1,1} + \theta_{4,1} + \theta_{7,1} + 2\theta_{1,2} + \theta_{5,2}, \\ &\quad \theta_{2,1} + \theta_{4,1} + \theta_{1,2} + \theta_{5,2}), \end{aligned} \quad (44d)$$

$$\begin{aligned} R_1 + 2R_2 &\leq \min(\theta_{1,1} + \theta_{5,1} + \theta_{2,2} + \theta_{4,2}, \\ &\quad \theta_{1,1} + \theta_{7,1} + \theta_{2,2} + \theta_{4,2}, \\ &\quad 2\theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{4,2} + \theta_{7,2}), \end{aligned} \quad (44e)$$

where $\theta_{l,i}$ are defined in (43) with $(l, i) \in \{1, \dots, 7\} \times \{1, 2\}$.

A. An Achievable Region for the Two-user Linear Deterministic Interference Channel with Noisy Channel-Output Feedback

In the LD-IC-NOF model, the channel input of transmitter i at each channel use is a q -dimensional vector $\mathbf{X}_i \in \{0, 1\}^q$ with $i \in \{1, 2\}$ and q as defined in (11). Following this observation, the random variables U, U_i, V_i , and $X_{i,P}$ described in (30) in the codebook generation are also vectors, and thus, in this subsection, they are denoted by $\mathbf{U}, \mathbf{U}_i, \mathbf{V}_i$ and $\mathbf{X}_{i,P}$, respectively.

The random variables $\mathbf{U}_i, \mathbf{V}_i$, and $\mathbf{X}_{i,P}$ are assumed to be mutually independent and uniformly distributed over the sets $\{0, 1\}^{(n_{ji} - (\max(\vec{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+)^+}$, $\{0, 1\}^{(\min(n_{ji}, (\max(\vec{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+))}$ and $\{0, 1\}^{(\vec{n}_{ii} - n_{ji})^+}$, respectively. Note that the random variables $\mathbf{U}_i, \mathbf{V}_i$, and $\mathbf{X}_{i,P}$ have the following dimensions:

$$\dim \mathbf{U}_i = (n_{ji} - (\max(\vec{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+)^+, \quad (45a)$$

$$\dim \mathbf{V}_i = \min(n_{ji}, (\max(\vec{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+), \text{ and} \quad (45b)$$

$$\dim \mathbf{X}_{i,P} = (\vec{n}_{ii} - n_{ji})^+. \quad (45c)$$

These dimensions satisfy the following condition:

$$\dim \mathbf{U}_i + \dim \mathbf{V}_i + \dim \mathbf{X}_{i,P} = \max(\vec{n}_{ii}, n_{ji}) \leq q. \quad (46)$$

Note that the random variable \mathbf{U} in (30) is not used, and therefore, is a constant. The input symbol of transmitter i during channel use n is $\mathbf{X}_i = (\mathbf{U}_i^\top, \mathbf{V}_i^\top, \mathbf{X}_{i,P}^\top, (0, \dots, 0))^\top$, where $(0, \dots, 0)$ is put to meet the dimension constraint $\dim \mathbf{X}_i = q$. Hence, during block $t \in \{1, 2, \dots, T\}$, the

codeword $\mathbf{X}_i^{(t)}$ in the LD-IC-NOF is a $q \times N$ matrix, i.e., $\mathbf{X}_i^{(t)} = (\mathbf{X}_{i,1}, \mathbf{X}_{i,2}, \dots, \mathbf{X}_{i,N}) \in \{0, 1\}^{q \times N}$.

The intuition behind this choice is based on the following observations: (a) The vector \mathbf{U}_i represents the bits in \mathbf{X}_i that can be observed by transmitter j via feedback but not necessarily by receiver i ; (b) The vector \mathbf{V}_i represents the bits in \mathbf{X}_i that can be observed by receiver j but not necessarily by receiver i ; and finally, (c) The vector $\mathbf{X}_{i,P}$ is a notational artefact to denote the bits of \mathbf{X}_i that are neither in \mathbf{U}_i nor \mathbf{V}_i . In particular, the bits in $\mathbf{X}_{i,P}$ are only observed by receiver i , as shown in Figure 9. This intuition justifies the dimensions described in (45).

Considering this particular code structure, the following holds for the terms $\theta_{l,i}$, with $(l, i) \in \{1, \dots, 7\} \times \{1, 2\}$, in (43):

$$\begin{aligned} \theta_{1,i} &= I(\overleftarrow{\mathbf{Y}}_i; \mathbf{U}_j | \mathbf{U}, \mathbf{X}_i) \\ &\stackrel{(a)}{=} H(\overleftarrow{\mathbf{Y}}_i | \mathbf{U}, \mathbf{X}_i) \\ &= H(\mathbf{U}_j) \\ &= (n_{ij} - (\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii})^+); \end{aligned} \quad (47a)$$

$$\begin{aligned} \theta_{2,i} &= I(\overrightarrow{\mathbf{Y}}_i; \mathbf{U}, \mathbf{U}_j, \mathbf{V}_j, \mathbf{X}_i) \\ &\stackrel{(b)}{=} H(\overrightarrow{\mathbf{Y}}_i) \\ &= \max(\overrightarrow{n}_{ii}, n_{ij}); \end{aligned} \quad (47b)$$

$$\begin{aligned} \theta_{3,i} &= I(\overrightarrow{\mathbf{Y}}_i; \mathbf{V}_j | \mathbf{U}, \mathbf{U}_j, \mathbf{X}_i) \\ &\stackrel{(b)}{=} H(\overrightarrow{\mathbf{Y}}_i | \mathbf{U}, \mathbf{U}_j, \mathbf{X}_i) \\ &= H(\mathbf{V}_j) \\ &= \min(n_{ij}, (\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii})^+); \end{aligned} \quad (47c)$$

$$\begin{aligned} \theta_{4,i} &= I(\overrightarrow{\mathbf{Y}}_i; \mathbf{X}_i | \mathbf{U}, \mathbf{U}_i, \mathbf{U}_j, \mathbf{V}_i, \mathbf{V}_j) \\ &\stackrel{(b)}{=} H(\overrightarrow{\mathbf{Y}}_i | \mathbf{U}, \mathbf{U}_i, \mathbf{U}_j, \mathbf{V}_i, \mathbf{V}_j) \\ &= H(\mathbf{X}_{i,P}) \\ &= (\overrightarrow{n}_{ii} - n_{ji})^+; \text{ and} \end{aligned} \quad (47d)$$

$$\begin{aligned} \theta_{5,i} &= I(\overrightarrow{\mathbf{Y}}_i; \mathbf{V}_j, \mathbf{X}_i | \mathbf{U}, \mathbf{U}_i, \mathbf{U}_j, \mathbf{V}_i) \\ &\stackrel{(b)}{=} H(\overrightarrow{\mathbf{Y}}_i | \mathbf{U}, \mathbf{U}_i, \mathbf{U}_j, \mathbf{V}_i) \\ &= \max(\dim \mathbf{X}_{i,P}, \dim \mathbf{V}_j) \\ &= \max\left((\overrightarrow{n}_{ii} - n_{ji})^+, \right. \\ &\quad \left. \min(n_{ij}, (\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii})^+)\right), \end{aligned} \quad (47e)$$

where (a) follows from the fact that $H(\overleftarrow{\mathbf{Y}}_i | \mathbf{U}, \mathbf{U}_j, \mathbf{X}_i) = 0$; and (b) follows from the fact that $H(\overrightarrow{\mathbf{Y}}_i | \mathbf{U}, \mathbf{U}_j, \mathbf{V}_j, \mathbf{X}_i) = 0$. For the calculation of the last two mutual information terms in inequalities (43f) and (43g), special notation is used. Let for instance the vector \mathbf{V}_i be the concatenation of the vectors $\mathbf{X}_{i,HA}$ and $\mathbf{X}_{i,HB}$, i.e., $\mathbf{V}_i = (\mathbf{X}_{i,HA}, \mathbf{X}_{i,HB})$. The vector $\mathbf{X}_{i,HA}$ is the part of \mathbf{V}_i that is available in both receivers. The vector $\mathbf{X}_{i,HB}$ is the part of \mathbf{V}_i that is exclusively available in receiver j (see Figure 9). Note that $H(\mathbf{V}_i) = H(\mathbf{X}_{i,HA}) + H(\mathbf{X}_{i,HB})$. Note also that the

vectors $\mathbf{X}_{i,HA}$ and $\mathbf{X}_{i,HB}$ possess the following dimensions:

$$\begin{aligned} \dim \mathbf{X}_{i,HA} &= \min\left(n_{ji}, (\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+\right) \\ &\quad - \min\left((n_{ji} - \overrightarrow{n}_{ii})^+, (\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+\right) \\ \dim \mathbf{X}_{i,HB} &= \min\left((n_{ji} - \overrightarrow{n}_{ii})^+, (\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+\right). \end{aligned}$$

Using this notation, the following holds:

$$\begin{aligned} \theta_{6,i} &= I(\overrightarrow{\mathbf{Y}}_i; \mathbf{X}_i | \mathbf{U}, \mathbf{U}_i, \mathbf{U}_j, \mathbf{V}_j) \\ &\stackrel{(c)}{=} H(\overrightarrow{\mathbf{Y}}_i | \mathbf{U}, \mathbf{U}_i, \mathbf{U}_j, \mathbf{V}_j) \\ &= H(\mathbf{X}_{i,HA}, \mathbf{X}_{i,P}) \\ &= \dim \mathbf{X}_{i,HA} + \dim \mathbf{X}_{i,P} \\ &= \min\left(n_{ji}, (\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+\right) \\ &\quad - \min\left((n_{ji} - \overrightarrow{n}_{ii})^+, (\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+\right) \\ &\quad + (\overrightarrow{n}_{ii} - n_{ji})^+; \text{ and} \end{aligned} \quad (47f)$$

$$\begin{aligned} \theta_{7,i} &= I(\overrightarrow{\mathbf{Y}}_i; \mathbf{V}_j, \mathbf{X}_i | \mathbf{U}, \mathbf{U}_i, \mathbf{U}_j) \\ &= I(\overrightarrow{\mathbf{Y}}_i; \mathbf{X}_i | \mathbf{U}, \mathbf{U}_i, \mathbf{U}_j) + I(\overrightarrow{\mathbf{Y}}_i; \mathbf{V}_j | \mathbf{U}, \mathbf{U}_i, \mathbf{U}_j, \mathbf{X}_i) \\ &= I(\overrightarrow{\mathbf{Y}}_i; \mathbf{X}_i | \mathbf{U}, \mathbf{U}_i, \mathbf{U}_j) + I(\overrightarrow{\mathbf{Y}}_i; \mathbf{V}_j | \mathbf{U}, \mathbf{U}_j, \mathbf{X}_i) \\ &\stackrel{(c)}{=} H(\overrightarrow{\mathbf{Y}}_i | \mathbf{U}, \mathbf{U}_i, \mathbf{U}_j) \\ &= \max(H(\mathbf{V}_j), H(\mathbf{X}_{i,HA}) + H(\mathbf{X}_{i,P})) \\ &= \max(\dim \mathbf{V}_j, \dim \mathbf{X}_{i,HA} + \dim \mathbf{X}_{i,P}) \\ &= \max\left(\min(n_{ij}, (\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii})^+), \right. \\ &\quad \left. \min(n_{ji}, (\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+)\right) \\ &\quad - \min\left((n_{ji} - \overrightarrow{n}_{ii})^+, (\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+\right) \\ &\quad + (\overrightarrow{n}_{ii} - n_{ji})^+; \end{aligned} \quad (47g)$$

where (c) follows from the fact that $H(\overrightarrow{\mathbf{Y}}_i | \mathbf{U}, \mathbf{U}_j, \mathbf{V}_j, \mathbf{X}_i) = 0$.

Plugging (47) into (44) (after some trivial manipulations) yields the system of inequalities in Theorem 1.

The sum-rate bound in (44c) can be simplified as follows:

$$\begin{aligned} R_1 + R_2 &\leq \min(\theta_{2,1} + \theta_{4,2}, \theta_{4,1} + \theta_{2,2}, \\ &\quad \theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{5,2}). \end{aligned} \quad (48)$$

Note that this follows from the realization that $\max(\theta_{2,1} + \theta_{4,2}, \theta_{4,1} + \theta_{2,2}, \theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{5,2}) \leq \min(\theta_{2,1} + \theta_{6,2}, \theta_{6,1} + \theta_{2,2}, \theta_{1,1} + \theta_{3,1} + \theta_{4,1} + \theta_{1,2} + \theta_{5,2}, \theta_{1,1} + \theta_{7,1} + \theta_{1,2} + \theta_{5,2}, \theta_{1,1} + \theta_{4,1} + \theta_{1,2} + \theta_{7,2}, \theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{3,2} + \theta_{4,2}, \theta_{1,1} + \theta_{7,1} + \theta_{1,2} + \theta_{4,2})$.

B. An Achievable Region for the Two-user Gaussian Interference Channel with Noisy Channel-Output Feedback

Consider that transmitter i uses the following Gaussian input distribution:

$$\mathbf{X}_i = \mathbf{U} + \mathbf{U}_i + \mathbf{V}_i + \mathbf{X}_{i,P}, \quad (49)$$

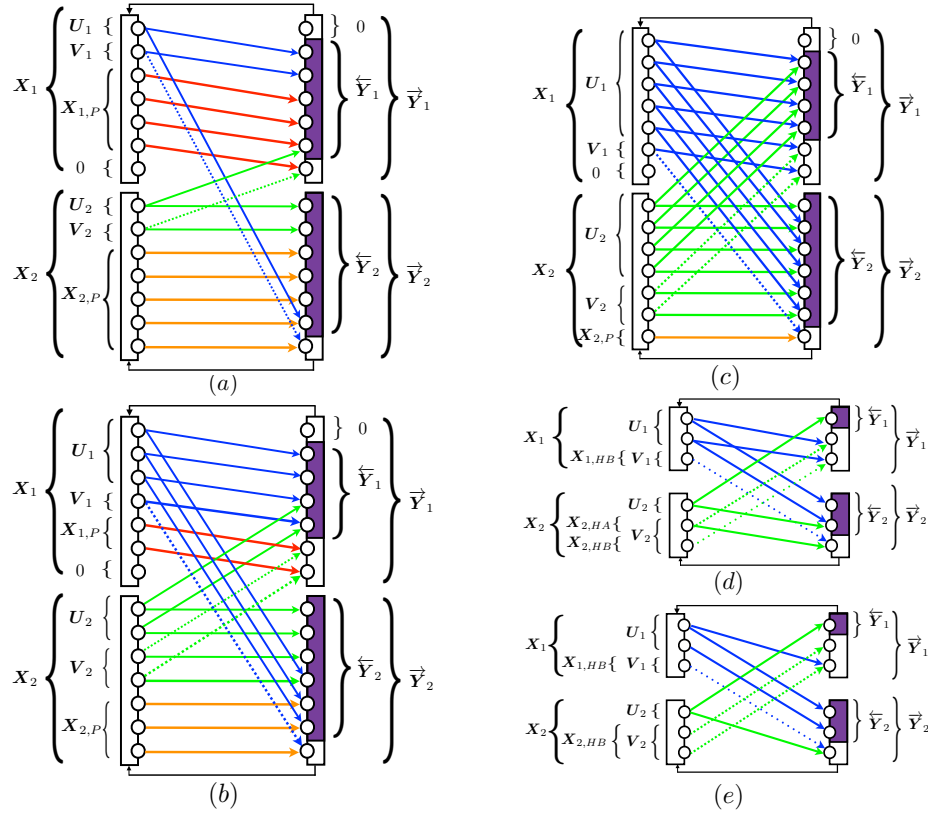


Fig. 9. The auxiliary random variables and their relation with signals when channel-output feedback is considered in (a) very weak interference regime, (b) weak interference regime, (c) moderate interference regime, (d) strong interference regime and (e) very strong interference regime.

where U , U_1 , U_2 , V_1 , V_2 , $X_{1,P}$, and $X_{2,P}$ in (30) are mutually independent and distributed as follows:

$$U \sim \mathcal{N}(0, \rho), \quad (50a)$$

$$U_i \sim \mathcal{N}(0, \mu_i \lambda_{i,C}), \quad (50b)$$

$$V_i \sim \mathcal{N}(0, (1 - \mu_i) \lambda_{i,C}), \quad (50c)$$

$$X_{i,P} \sim \mathcal{N}(0, \lambda_{i,P}), \quad (50d)$$

with

$$\rho + \lambda_{i,C} + \lambda_{i,P} = 1 \text{ and} \quad (51a)$$

$$\lambda_{i,P} = \min\left(\frac{1}{\text{INR}_{ji}}, 1\right), \quad (51b)$$

where $\mu_i \in [0, 1]$ and $\rho \in \left[0, \left(1 - \max\left(\frac{1}{\text{INR}_{12}}, \frac{1}{\text{INR}_{21}}\right)\right)^+\right]$. The random variables U , U_1 , U_2 , V_1 , V_2 , $X_{1,P}$, and $X_{2,P}$ can be interpreted as components of the signals X_1 and X_2 following the insights described in this appendix. The random variable U , which is used in this case, represents the common component of the channel inputs of transmitter 1 and transmitter 2.

The parameters ρ , μ_i , and $\lambda_{i,P}$ define a particular coding scheme for transmitter i . The assignment in (51b) is based on the intuition obtained from the linear deterministic model, in which the power of the signal $X_{i,P}$ from transmitter i to receiver j must be observed at the noise level. From (2), (3), and (49), the right-hand side of the inequalities in (43) can

be written in terms of $\overrightarrow{\text{SNR}}_1$, $\overrightarrow{\text{SNR}}_2$, $\overrightarrow{\text{INR}}_{12}$, $\overrightarrow{\text{INR}}_{21}$, $\overleftarrow{\text{SNR}}_1$, $\overleftarrow{\text{SNR}}_2$, ρ , μ_1 , and μ_2 as follows:

$$\begin{aligned} \theta_{1,i} &= I(\overleftarrow{Y}_i; U_j | U, X_i) \\ &= \frac{1}{2} \log \left(\frac{\overleftarrow{\text{SNR}}_i (b_{2,i}(\rho) + 2) + b_{1,i}(1) + 1}{\overleftarrow{\text{SNR}}_i ((1 - \mu_j) b_{2,i}(\rho) + 2) + b_{1,i}(1) + 1} \right) \\ &= a_{3,i}(\rho, \mu_j); \end{aligned} \quad (52a)$$

$$\begin{aligned} \theta_{2,i} &= I(\overleftarrow{Y}_i; U, U_j, V_j, X_i) \\ &= \frac{1}{2} \log (b_{1,i}(\rho) + 1) - \frac{1}{2} \\ &= a_{2,i}(\rho); \end{aligned} \quad (52b)$$

$$\begin{aligned} \theta_{3,i} &= I(\overleftarrow{Y}_i; V_j | U, U_j, X_i) \\ &= \frac{1}{2} \log \left((1 - \mu_j) b_{2,i}(\rho) + 2 \right) - \frac{1}{2} \\ &= a_{4,i}(\rho, \mu_j); \end{aligned} \quad (52c)$$

$$\begin{aligned} \theta_{4,i} &= I(\overleftarrow{Y}_i; X_i | U, U_i, U_j, V_i, V_j) \\ &= \frac{1}{2} \log \left(\frac{\overrightarrow{\text{SNR}}_i}{\text{INR}_{ji}} + 2 \right) - \frac{1}{2} \\ &= a_{1,i}; \end{aligned} \quad (52d)$$

$$\begin{aligned} \theta_{5,i} &= I(\overleftarrow{Y}_i; V_j, X_i | U, U_i, U_j, V_i) \\ &= \frac{1}{2} \log \left(2 + \frac{\overrightarrow{\text{SNR}}_i}{\text{INR}_{ji}} + (1 - \mu_j) b_{2,i}(\rho) \right) - \frac{1}{2} \\ &= a_{5,i}(\rho, \mu_j); \end{aligned} \quad (52e)$$

$$\begin{aligned}\theta_{6,i} &= I(\vec{Y}_i; X_i | U, U_i, U_j, V_j) \\ &= \frac{1}{2} \log \left(\frac{\overline{\text{SNR}}_i}{\overline{\text{INR}}_{ji}} \left((1 - \mu_i) b_{2,j}(\rho) + 1 \right) + 2 \right) - \frac{1}{2} \\ &= a_{6,i}(\rho, \mu_i); \text{ and}\end{aligned}\quad (52f)$$

$$\begin{aligned}\theta_{7,i} &= I(\vec{Y}_i; V_j, X_i | U, U_i, U_j) \\ &= \frac{1}{2} \log \left(\frac{\overline{\text{SNR}}_i}{\overline{\text{INR}}_{ji}} \left((1 - \mu_i) b_{2,j}(\rho) + 1 \right) + (1 - \mu_j) b_{2,i}(\rho) + 2 \right) \\ &\quad - \frac{1}{2} \\ &= a_{7,i}(\rho, \mu_1, \mu_2).\end{aligned}\quad (52g)$$

Finally, plugging (52) into (44) (after some trivial manipulations) yields the system of inequalities in Theorem 2. The sum-rate bound in (44c) can be simplified as follows:

$$\begin{aligned}R_1 + R_2 &\leq \min \left(a_{2,1}(\rho) + a_{1,2}, a_{1,1} + a_{2,2}(\rho), \right. \\ &\quad a_{3,1}(\rho, \mu_2) + a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{7,2}(\rho, \mu_1, \mu_2), \\ &\quad a_{3,1}(\rho, \mu_2) + a_{5,1}(\rho, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{5,2}(\rho, \mu_1), \\ &\quad \left. a_{3,1}(\rho, \mu_2) + a_{7,1}(\rho, \mu_1, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{1,2} \right).\end{aligned}\quad (53)$$

Note that this follows from the realization that $\max(a_{2,1}(\rho) + a_{1,2}, a_{1,1} + a_{2,2}(\rho), a_{3,1}(\rho, \mu_2) + a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{7,2}(\rho, \mu_1, \mu_2), a_{3,1}(\rho, \mu_2) + a_{5,1}(\rho, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{5,2}(\rho, \mu_1), a_{3,1}(\rho, \mu_2) + a_{7,1}(\rho, \mu_1, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{1,2}) \leq \min(a_{2,1} + a_{6,2}(\rho, \mu_2), a_{6,1}(\rho, \mu_1) + a_{2,2}(\rho), a_{3,1}(\rho, \mu_2) + a_{4,1}(\rho, \mu_2) + a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{5,2}(\rho, \mu_1), a_{3,1}(\rho, \mu_2) + a_{7,1}(\rho, \mu_1, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{5,2}(\rho, \mu_1), a_{3,1}(\rho, \mu_2) + a_{5,1}(\rho, \mu_2) + a_{3,2}(\rho, \mu_1) + \theta_{3,2} + a_{1,2})$. Therefore, the inequalities in (44) simplify into (23) and this completes the proof of Theorem 2.

APPENDIX B CONVERSE PROOF FOR THEOREM 1

This appendix provides a converse proof for Theorem 1. Inequalities (15a) and (15c) correspond to the minimum cut-set bound [33] and the sum-rate bound for the case of the two-user LD-IC-POF. The proofs of these bounds are presented in [3]. The rest of this appendix provides a proof of the inequalities (15b), (15d) and (15e).

Notation. For all $i \in \{1, 2\}$, the channel input $\mathbf{X}_{i,n}$ of the LD-IC-NOF in (12) for any channel use $n \in \{1, 2, \dots, N\}$ is a q -dimensional vector, with q in (11), that can be written as the concatenation of four vectors: $\mathbf{X}_{i,C,n}$, $\mathbf{X}_{i,P,n}$, $\mathbf{X}_{i,D,n}$, and $\mathbf{X}_{i,Q,n}$, i.e., $\mathbf{X}_{i,n} = \left(\mathbf{X}_{i,C,n}^\top, \mathbf{X}_{i,P,n}^\top, \mathbf{X}_{i,D,n}^\top, \mathbf{X}_{i,Q,n}^\top \right)^\top$, as shown in Figure 10. Note that this notation is independent of the feedback parameters \overleftarrow{n}_{11} and \overleftarrow{n}_{22} , and it holds for all $n \in \{1, 2, \dots, N\}$. More specifically, $\mathbf{X}_{i,C,n}$ represents the bits of $\mathbf{X}_{i,n}$ that are observed by both receivers. Then,

$$\dim \mathbf{X}_{i,C,n} = \min(\overleftarrow{n}_{ii}, n_{ji}); \quad (54a)$$

$\mathbf{X}_{i,P,n}$ represents the bits of $\mathbf{X}_{i,n}$ that are observed only at receiver i . Then,

$$\dim \mathbf{X}_{i,P,n} = (\overleftarrow{n}_{ii} - n_{ji})^+; \quad (54b)$$

$\mathbf{X}_{i,D,n}$ represents the bits of $\mathbf{X}_{i,n}$ that are observed only at receiver j . Then,

$$\dim \mathbf{X}_{i,D,n} = (n_{ji} - \overleftarrow{n}_{ii})^+; \text{ and} \quad (54c)$$

$\mathbf{X}_{i,Q,n} = (0, \dots, 0)^\top$ is included for dimensional matching of the model in (13). Then,

$$\dim \mathbf{X}_{i,Q,n} = q - \max(\overleftarrow{n}_{ii}, n_{ji}). \quad (54d)$$

The bits $\mathbf{X}_{i,Q,n}$ are fixed and thus do not carry any information. Hence, the following holds:

$$\begin{aligned}H(\mathbf{X}_{i,n}) &= H(\mathbf{X}_{i,C,n}, \mathbf{X}_{i,P,n}, \mathbf{X}_{i,D,n}, \mathbf{X}_{i,Q,n}) \\ &= H(\mathbf{X}_{i,C,n}, \mathbf{X}_{i,P,n}, \mathbf{X}_{i,D,n}) \\ &\leq \dim \mathbf{X}_{i,C,n} + \dim \mathbf{X}_{i,P,n} + \dim \mathbf{X}_{i,D,n}.\end{aligned}\quad (54e)$$

Note that vectors $\mathbf{X}_{i,P,n}$ and $\mathbf{X}_{i,D,n}$ do not exist simultaneously. The former exists when $\overleftarrow{n}_{ii} > n_{ji}$, while the latter exists when $\overleftarrow{n}_{ii} < n_{ji}$. Moreover, the dimension of $\mathbf{X}_{i,n}$ satisfies

$$\begin{aligned}\dim \mathbf{X}_{i,n} &= \dim \mathbf{X}_{i,C,n} + \dim \mathbf{X}_{i,P,n} + \dim \mathbf{X}_{i,D,n} \\ &\quad + \dim \mathbf{X}_{i,Q,n} \\ &= q.\end{aligned}\quad (54f)$$

For the case in which feedback is taken into account an alternative notation is adopted. Let $\mathbf{X}_{i,D,n}$ be written in terms of $\mathbf{X}_{i,DF,n}$ and $\mathbf{X}_{i,DG,n}$, i.e., $\mathbf{X}_{i,D,n} = \left(\mathbf{X}_{i,DF,n}^\top, \mathbf{X}_{i,DG,n}^\top \right)^\top$. The vector $\mathbf{X}_{i,DF,n}$ represents the bits of $\mathbf{X}_{i,D,n}$ that are above the noise level in the feedback link from receiver j to transmitter j ; and $\mathbf{X}_{i,DG,n}$ represents the bits of $\mathbf{X}_{i,D,n}$ that are below the noise level in the feedback link from receiver j to transmitter j , as shown in Figure 10. The dimension of vectors $\mathbf{X}_{i,DF,n}$ and $\mathbf{X}_{i,DG,n}$ are given by

$$\begin{aligned}\dim \mathbf{X}_{i,DF,n} &= \min \left((n_{ji} - \overleftarrow{n}_{ii})^+, \left(\overleftarrow{n}_{jj} - \overleftarrow{n}_{ii} \right. \right. \\ &\quad \left. \left. - \min \left((\overleftarrow{n}_{jj} - n_{ji})^+, n_{ij} \right) \right. \right. \\ &\quad \left. \left. - \left((\overleftarrow{n}_{jj} - n_{ij})^+ - n_{ji} \right)^+ \right)^+ \right) \text{ and}\end{aligned}\quad (55a)$$

$$\dim \mathbf{X}_{i,DG,n} = \dim \mathbf{X}_{i,D,n} - \dim \mathbf{X}_{i,DF,n}.\quad (55b)$$

Let $\mathbf{X}_{i,C,n}$ be written in terms of $\mathbf{X}_{i,CF_j,n}$ and $\mathbf{X}_{i,CG_j,n}$, i.e., $\mathbf{X}_{i,C,n} = \left(\mathbf{X}_{i,CF_j,n}^\top, \mathbf{X}_{i,CG_j,n}^\top \right)^\top$. The vector $\mathbf{X}_{i,CF_j,n}$ represents the bits of $\mathbf{X}_{i,C,n}$ that are above the noise level in the feedback link from receiver j to transmitter j ; and $\mathbf{X}_{i,CG_j,n}$ represents the bits of $\mathbf{X}_{i,C,n}$ that are below the noise level in the feedback link from receiver j to transmitter j , as shown in Figure 10. Let also, the dimension of vector $\left(\mathbf{X}_{i,CF_j,n}^\top, \mathbf{X}_{i,DF,n}^\top \right)$ be defined as follows:

$$\begin{aligned}\dim \left(\left(\mathbf{X}_{i,CF_j,n}^\top, \mathbf{X}_{i,DF,n}^\top \right) \right) \\ = \left(\min(\overleftarrow{n}_{jj}, \max(\overleftarrow{n}_{jj}, n_{ji})) - (\overleftarrow{n}_{jj} - n_{ji})^+ \right)^+.\end{aligned}\quad (56)$$

The dimension of vectors $\mathbf{X}_{i,CF_j,n}$ and $\mathbf{X}_{i,CG_j,n}$ can be obtained as follows:

$$\begin{aligned}\dim \mathbf{X}_{i,CF_j,n} &= \dim \left(\left(\mathbf{X}_{i,CF_j,n}^\top, \mathbf{X}_{i,DF,n}^\top \right) \right) - \dim \mathbf{X}_{i,DF,n} \\ &\text{and}\end{aligned}\quad (57a)$$

$$\dim \mathbf{X}_{i,CG_j,n} = \dim \mathbf{X}_{i,C,n} - \dim \mathbf{X}_{i,CF_j,n}.\quad (57b)$$

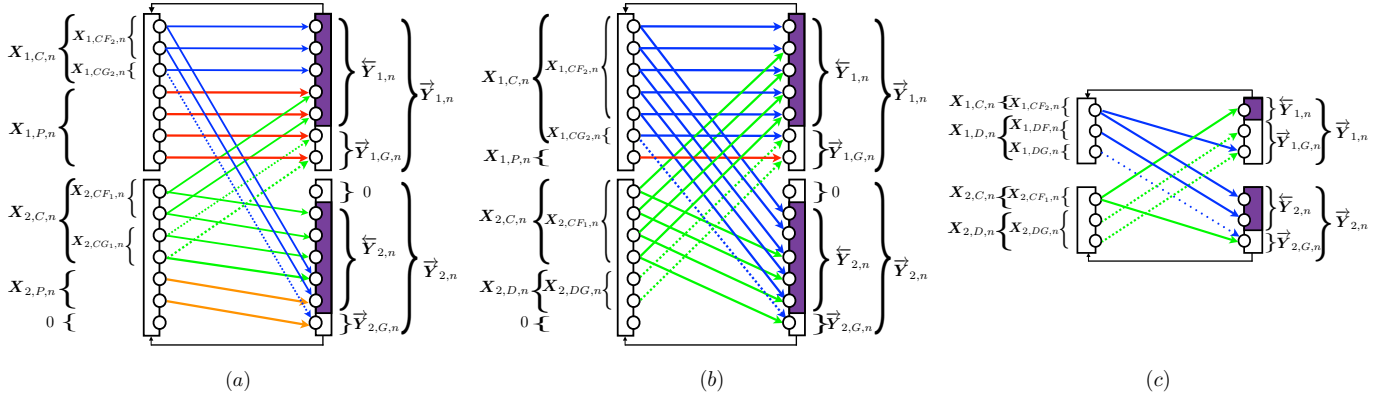


Fig. 10. Example of the notation of the channel inputs and the channel outputs when channel-output feedback is considered.

More generally, when needed, the vector $\mathbf{X}_{iF_k,n}$ is used to represent the bits of $\mathbf{X}_{i,n}$ that are above the noise level in the feedback link from receiver k to transmitter k , with $k \in \{1, 2\}$. The vector $\mathbf{X}_{iG_k,n}$ is used to represent the bits of $\mathbf{X}_{i,n}$ that are below the noise level in the feedback link from receiver k to transmitter k .

The vector $\mathbf{X}_{i,U,n}$ is used to represent the bits of vector $\mathbf{X}_{i,n}$ that interfere with bits of $\mathbf{X}_{j,C,n}$ at receiver j and those bits of $\mathbf{X}_{i,n}$ that are observed by receiver j and do not interfere any bits from transmitter j . An alternative definition of the vector $\mathbf{X}_{i,U,n}$ is the following: the bits of vector $\mathbf{X}_{i,n}$ that are observed by receiver j and do not interfere any bits corresponding to the vector $\mathbf{X}_{j,P,n}$. An example is shown in Figure 11.

Based on its definition, the dimension of vector $\mathbf{X}_{i,U,n}$ is

$$\dim \mathbf{X}_{i,U,n} = \min(\bar{n}_{jj}, n_{ij}) - \min((\bar{n}_{jj} - n_{ji})^+, n_{ij}) + (n_{ji} - \bar{n}_{jj})^+. \quad (58)$$

Finally, for all $i \in \{1, 2\}$, with $j \in \{1, 2\} \setminus \{i\}$, the channel output $\vec{\mathbf{Y}}_{i,n}$ of the LD-IC-NOF in (12) for any channel use $n \in \{1, 2, \dots, N\}$ is a q -dimensional vector, with q in (11), that can be written as the concatenation of three vectors: $\vec{\mathbf{Y}}_{i,Q,n}$, $\vec{\mathbf{Y}}_{i,n}$, and $\vec{\mathbf{Y}}_{i,G,n}$, i.e., $\vec{\mathbf{Y}}_{i,n} = (\vec{\mathbf{Y}}_{i,Q,n}^\top, \vec{\mathbf{Y}}_{i,n}^\top, \vec{\mathbf{Y}}_{i,G,n}^\top)^\top$, as shown in Figure 10.

More specifically, the vector $\vec{\mathbf{Y}}_{i,n}$ contains the bits that are above the noise level in the feedback link from receiver i to transmitter i . Then,

$$\dim \vec{\mathbf{Y}}_{i,n} = \min(\bar{n}_{ii}, \max(\bar{n}_{ii}, n_{ij})). \quad (59a)$$

The vector $\vec{\mathbf{Y}}_{i,G,n}$ contains the bits that are below the noise level in the feedback link from receiver i to transmitter i . Then,

$$\dim \vec{\mathbf{Y}}_{i,G,n} = (\max(\bar{n}_{ii}, n_{ij}) - \bar{n}_{ii})^+. \quad (59b)$$

The vector $\vec{\mathbf{Y}}_{i,Q,n} = (0, \dots, 0)$ is included for dimensional matching with the model in (13). Then,

$$\begin{aligned} H(\vec{\mathbf{Y}}_{i,n}) &= H(\vec{\mathbf{Y}}_{i,Q,n}, \vec{\mathbf{Y}}_{i,n}, \vec{\mathbf{Y}}_{i,G,n}) \\ &= H(\vec{\mathbf{Y}}_{i,n}, \vec{\mathbf{Y}}_{i,G,n}) \\ &\leq \dim \vec{\mathbf{Y}}_{i,n} + \dim \vec{\mathbf{Y}}_{i,G,n}. \end{aligned} \quad (59c)$$

The dimension of $\vec{\mathbf{Y}}_{i,n}$ satisfies $\dim \vec{\mathbf{Y}}_{i,n} = q$.

Using this notation, the proof continues as follows.

Proof of (15b): First, consider $n_{ji} \leq \bar{n}_{ii}$, i.e., vector $\mathbf{X}_{i,P,n}$ exists and vector $\mathbf{X}_{i,D,n}$ does not exist. From the assumption that the message index W_i is i.i.d. following a uniform distribution over the set \mathcal{W}_i , the following holds for any $k \in \{1, 2, \dots, N\}$:

$$\begin{aligned} NR_i &= H(W_i) \\ &\stackrel{(a)}{=} H(W_i | W_j) \\ &\stackrel{(b)}{\leq} I(W_i; \vec{\mathbf{Y}}_i, \vec{\mathbf{Y}}_j | W_j) + N\delta(N) \\ &= H(\vec{\mathbf{Y}}_i, \vec{\mathbf{Y}}_j | W_j) + N\delta(N) \\ &\stackrel{(c)}{=} \sum_{n=1}^N H(\vec{\mathbf{Y}}_{i,n}, \vec{\mathbf{Y}}_{j,n} | W_j, \vec{\mathbf{Y}}_{i,(1:n-1)}, \vec{\mathbf{Y}}_{j,(1:n-1)}, \\ &\quad \mathbf{X}_{j,n}) + N\delta(N) \\ &\leq \sum_{n=1}^N H(\mathbf{X}_{i,n}, \vec{\mathbf{Y}}_{j,n} | \mathbf{X}_{j,n}) + N\delta(N) \\ &\leq \sum_{n=1}^N H(\mathbf{X}_{i,n}) + N\delta(N) \\ &= NH(\mathbf{X}_{i,k}) + N\delta(N), \\ &\leq N(\dim \mathbf{X}_{i,C,k} + \dim \mathbf{X}_{i,P,k}) + N\delta(N), \end{aligned} \quad (60)$$

where, (a) follows from the fact that W_1 and W_2 are independent; (b) follows from Fano's inequality; and (c) follows from the fact that $\mathbf{X}_{j,n} = f_j^{(n)}(W_j, \vec{\mathbf{Y}}_{j,(1:n-1)})$.

Second, consider the case in which $n_{ji} > \bar{n}_{ii}$. In this case the vector $\mathbf{X}_{i,P,n}$ does not exist and the vector $\mathbf{X}_{i,D,n}$ exists. From the assumption that the message index W_i is i.i.d. following a uniform distribution over the set \mathcal{W}_i , hence the following holds for any $k \in \{1, 2, \dots, N\}$:

$$\begin{aligned} NR_i &= H(W_i) \\ &\stackrel{(a)}{=} H(W_i | W_j) \\ &\stackrel{(b)}{\leq} I(W_i; \vec{\mathbf{Y}}_i, \vec{\mathbf{Y}}_j | W_j) + N\delta(N) \end{aligned}$$

$$\begin{aligned}
&= H(\vec{\mathbf{Y}}_i, \overleftarrow{\mathbf{Y}}_j | W_j) + N\delta(N) \\
&\stackrel{(c)}{=} \sum_{n=1}^N H(\vec{\mathbf{Y}}_{i,n}, \overleftarrow{\mathbf{Y}}_{j,n} | W_j, \vec{\mathbf{Y}}_{i,(1:n-1)}, \overleftarrow{\mathbf{Y}}_{j,(1:n-1)}, \\
&\quad \mathbf{X}_{j,n}) + N\delta(N) \\
&\leq \sum_{n=1}^N H(\mathbf{X}_{i,C,n}, \mathbf{X}_{i,CF_j,n}, \mathbf{X}_{i,DF,n}) + N\delta(N) \\
&= \sum_{n=1}^N H(\mathbf{X}_{i,C,n}, \mathbf{X}_{i,DF,n}) + N\delta(N) \\
&= NH(\mathbf{X}_{i,C,k}, \mathbf{X}_{i,DF,k}) + N\delta(N), \\
&\leq N(\dim \mathbf{X}_{i,C,k} + \dim \mathbf{X}_{i,DF,k}) + N\delta(N). \quad (61)
\end{aligned}$$

Then, (60) and (61) can be expressed as one inequality in the asymptotic regime, as follows:

$$R_i \leq \dim \mathbf{X}_{i,C,k} + \dim \mathbf{X}_{i,P,k} + \dim \mathbf{X}_{i,DF,k}, \quad (62)$$

which holds for any $k \in \{1, 2, \dots, N\}$.

Plugging (54a), (54b), and (55a) in (62), and after some trivial manipulations, the following holds:

$$R_i \leq \min \left(\max(\vec{n}_{ii}, n_{ji}), \max(\vec{n}_{ii}, \overleftarrow{n}_{jj} - (\vec{n}_{jj} - n_{ji})^+) \right). \quad (63)$$

This completes the proof of (15b).

Proof of (15d): From the assumption that the message indices W_1 and W_2 are i.i.d. following a uniform distribution over the sets \mathcal{W}_1 and \mathcal{W}_2 respectively, the following holds for any $k \in \{1, 2, \dots, N\}$:

$$\begin{aligned}
&N(R_1 + R_2) = H(W_1) + H(W_2) \\
&\stackrel{(a)}{\leq} I(W_1; \vec{\mathbf{Y}}_1, \overleftarrow{\mathbf{Y}}_1) + I(W_2; \vec{\mathbf{Y}}_2, \overleftarrow{\mathbf{Y}}_2) + N\delta(N) \\
&\leq H(\vec{\mathbf{Y}}_1) - H(\overleftarrow{\mathbf{Y}}_1 | W_1) - H(\mathbf{X}_{2,C} | W_1, \overleftarrow{\mathbf{Y}}_1, \mathbf{X}_1) \\
&\quad + H(\vec{\mathbf{Y}}_2) - H(\overleftarrow{\mathbf{Y}}_2 | W_2) - H(\mathbf{X}_{1,C} | W_2, \overleftarrow{\mathbf{Y}}_2, \mathbf{X}_2) \\
&\quad + N\delta(N) \\
&= H(\vec{\mathbf{Y}}_1) - H(\overleftarrow{\mathbf{Y}}_1 | W_1) - H(\mathbf{X}_{2,C}, \mathbf{X}_{1,U} | W_1, \overleftarrow{\mathbf{Y}}_1, \mathbf{X}_1) \\
&\quad + H(\vec{\mathbf{Y}}_2) - H(\overleftarrow{\mathbf{Y}}_2 | W_2) \\
&\quad - H(\mathbf{X}_{1,C}, \mathbf{X}_{2,U} | W_2, \overleftarrow{\mathbf{Y}}_2, \mathbf{X}_2) + N\delta(N) \\
&= H(\vec{\mathbf{Y}}_1) + \left[I(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}; W_1, \overleftarrow{\mathbf{Y}}_1) \right. \\
&\quad \left. - H(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}) \right] + H(\vec{\mathbf{Y}}_2) \\
&\quad + \left[I(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}; W_2, \overleftarrow{\mathbf{Y}}_2) - H(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}) \right] \\
&\quad - H(\overleftarrow{\mathbf{Y}}_1 | W_1) - H(\overleftarrow{\mathbf{Y}}_2 | W_2) + N\delta(N) \\
&\stackrel{(b)}{=} H(\vec{\mathbf{Y}}_1 | \mathbf{X}_{1,C}, \mathbf{X}_{2,U}) - H(\mathbf{X}_{1,C}, \mathbf{X}_{2,U} | \vec{\mathbf{Y}}_1) \\
&\quad + H(\vec{\mathbf{Y}}_2 | \mathbf{X}_{2,C}, \mathbf{X}_{1,U}) - H(\mathbf{X}_{2,C}, \mathbf{X}_{1,U} | \vec{\mathbf{Y}}_2) \\
&\quad + I(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}; W_1, \overleftarrow{\mathbf{Y}}_1) + I(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}; W_2, \overleftarrow{\mathbf{Y}}_2) \\
&\quad - H(\overleftarrow{\mathbf{Y}}_1 | W_1) - H(\overleftarrow{\mathbf{Y}}_2 | W_2) + N\delta(N)
\end{aligned}$$

$$\begin{aligned}
&\leq H(\vec{\mathbf{Y}}_1 | \mathbf{X}_{1,C}, \mathbf{X}_{2,U}) + H(\vec{\mathbf{Y}}_2 | \mathbf{X}_{2,C}, \mathbf{X}_{1,U}) \\
&\quad + I(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}; W_1, \overleftarrow{\mathbf{Y}}_1) + I(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}; W_2, \overleftarrow{\mathbf{Y}}_2) \\
&\quad - H(\overleftarrow{\mathbf{Y}}_1 | W_1) - H(\overleftarrow{\mathbf{Y}}_2 | W_2) + N\delta(N) \\
&\leq H(\vec{\mathbf{Y}}_1 | \mathbf{X}_{1,C}, \mathbf{X}_{2,U}) + H(\vec{\mathbf{Y}}_2 | \mathbf{X}_{2,C}, \mathbf{X}_{1,U}) \\
&\quad + I(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}, W_2, \overleftarrow{\mathbf{Y}}_2; W_1, \overleftarrow{\mathbf{Y}}_1) \\
&\quad + I(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}, W_1, \overleftarrow{\mathbf{Y}}_1; W_2, \overleftarrow{\mathbf{Y}}_2) - H(\overleftarrow{\mathbf{Y}}_1 | W_1) \\
&\quad - H(\overleftarrow{\mathbf{Y}}_2 | W_2) + N\delta(N) \\
&= H(\vec{\mathbf{Y}}_1 | \mathbf{X}_{1,C}, \mathbf{X}_{2,U}) + H(\vec{\mathbf{Y}}_2 | \mathbf{X}_{2,C}, \mathbf{X}_{1,U}) \\
&\quad + I(W_2; W_1, \overleftarrow{\mathbf{Y}}_1) + I(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}, \overleftarrow{\mathbf{Y}}_2; W_1, \overleftarrow{\mathbf{Y}}_1 | W_2) \\
&\quad + I(W_1; W_2, \overleftarrow{\mathbf{Y}}_2) + I(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}, \overleftarrow{\mathbf{Y}}_1; W_2, \overleftarrow{\mathbf{Y}}_2 | W_1) \\
&\quad - H(\overleftarrow{\mathbf{Y}}_1 | W_1) - H(\overleftarrow{\mathbf{Y}}_2 | W_2) + N\delta(N) \\
&\stackrel{(c)}{=} H(\vec{\mathbf{Y}}_1 | \mathbf{X}_{1,C}, \mathbf{X}_{2,U}) + H(\vec{\mathbf{Y}}_2 | \mathbf{X}_{2,C}, \mathbf{X}_{1,U}) \\
&\quad + H(W_1) + H(\overleftarrow{\mathbf{Y}}_1 | W_1) - H(W_1 | W_2) \\
&\quad - H(\overleftarrow{\mathbf{Y}}_1 | W_2, W_1) + H(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}, \overleftarrow{\mathbf{Y}}_2 | W_2) + H(W_2) \\
&\quad + H(\overleftarrow{\mathbf{Y}}_2 | W_2) - H(W_2 | W_1) - H(\overleftarrow{\mathbf{Y}}_2 | W_1, W_2) \\
&\quad + H(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}, \overleftarrow{\mathbf{Y}}_1 | W_1) - H(\overleftarrow{\mathbf{Y}}_1 | W_1) - H(\overleftarrow{\mathbf{Y}}_2 | W_2) \\
&\quad + N\delta(N) \\
&\leq H(\vec{\mathbf{Y}}_1 | \mathbf{X}_{1,C}, \mathbf{X}_{2,U}) + H(\vec{\mathbf{Y}}_2 | \mathbf{X}_{2,C}, \mathbf{X}_{1,U}) \\
&\quad + H(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}, \overleftarrow{\mathbf{Y}}_2 | W_2) + H(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}, \overleftarrow{\mathbf{Y}}_1 | W_1) \\
&\quad + N\delta(N) \\
&= \sum_{n=1}^N \left[H(\vec{\mathbf{Y}}_{1,n} | \mathbf{X}_{1,C}, \mathbf{X}_{2,U}, \vec{\mathbf{Y}}_{1,(1:n-1)}) \right. \\
&\quad + H(\vec{\mathbf{Y}}_{2,n} | \mathbf{X}_{2,C}, \mathbf{X}_{1,U}, \vec{\mathbf{Y}}_{2,(1:n-1)}) \\
&\quad + H(\mathbf{X}_{2,C,n}, \mathbf{X}_{1,U,n}, \overleftarrow{\mathbf{Y}}_{2,n} | W_2, \mathbf{X}_{2,C,(1:n-1)}, \\
&\quad \mathbf{X}_{1,U,(1:n-1)}, \overleftarrow{\mathbf{Y}}_{2,(1:n-1)}) \\
&\quad \left. + H(\mathbf{X}_{1,C,n}, \mathbf{X}_{2,U,n}, \overleftarrow{\mathbf{Y}}_{1,n} | W_1, \mathbf{X}_{1,C,(1:n-1)}, \right. \\
&\quad \left. \mathbf{X}_{2,U,(1:n-1)}, \overleftarrow{\mathbf{Y}}_{1,(1:n-1)}) \right] + N\delta(N) \\
&\stackrel{(d)}{=} \sum_{n=1}^N \left[H(\vec{\mathbf{Y}}_{1,n} | \mathbf{X}_{1,C}, \mathbf{X}_{2,U}, \vec{\mathbf{Y}}_{1,(1:n-1)}) \right. \\
&\quad + H(\vec{\mathbf{Y}}_{2,n} | \mathbf{X}_{2,C}, \mathbf{X}_{1,U}, \vec{\mathbf{Y}}_{2,(1:n-1)}) \\
&\quad + H(\mathbf{X}_{2,C,n}, \mathbf{X}_{1,U,n}, \overleftarrow{\mathbf{Y}}_{2,n} | W_2, \mathbf{X}_{2,C,(1:n-1)}, \\
&\quad \mathbf{X}_{1,U,(1:n-1)}, \overleftarrow{\mathbf{Y}}_{2,(1:n-1)}, \mathbf{X}_{2,(1:n)}) \\
&\quad \left. + H(\mathbf{X}_{1,C,n}, \mathbf{X}_{2,U,n}, \overleftarrow{\mathbf{Y}}_{1,n} | W_1, \mathbf{X}_{1,C,(1:n-1)}, \right. \\
&\quad \left. \mathbf{X}_{2,U,(1:n-1)}, \overleftarrow{\mathbf{Y}}_{1,(1:n-1)}, \mathbf{X}_{1,(1:n)}) \right] + N\delta(N) \\
&\stackrel{(e)}{\leq} \sum_{n=1}^N \left[H(\vec{\mathbf{Y}}_{1,n} | \mathbf{X}_{1,C,n}, \mathbf{X}_{2,U,n}) \right. \\
&\quad \left. + H(\vec{\mathbf{Y}}_{2,n} | \mathbf{X}_{2,C,n}, \mathbf{X}_{1,U,n}) \right]
\end{aligned}$$

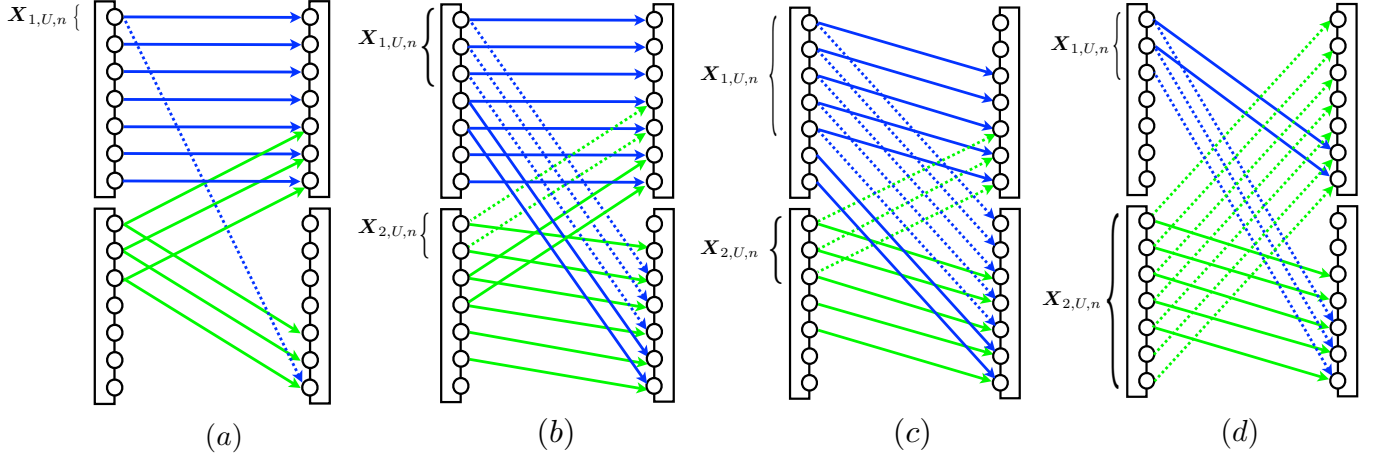


Fig. 11. Vector $\mathbf{X}_{i,U,n}$ in different combination of interference regimes.

$$\begin{aligned}
& +H(\mathbf{X}_{1,U,n}, \hat{\mathbf{Y}}_{2,n} | \mathbf{X}_{2,n}) \\
& +H(\mathbf{X}_{2,U,n}, \hat{\mathbf{Y}}_{1,n} | \mathbf{X}_{1,n}) \Big] + N\delta(N) \\
\leq & \sum_{n=1}^N \left[H(\mathbf{X}_{1,P,n}) + H(\mathbf{X}_{2,P,n}) \right. \\
& +H(\mathbf{X}_{1,U,n}, \hat{\mathbf{Y}}_{2,n} | \mathbf{X}_{2,n}) \\
& \left. +H(\mathbf{X}_{2,U,n}, \hat{\mathbf{Y}}_{1,n} | \mathbf{X}_{1,n}) \right] + N\delta(N) \\
\stackrel{(e)}{\leq} & N \left[H(\mathbf{X}_{1,P,k}) + H(\mathbf{X}_{2,P,k}) + H(\mathbf{X}_{1,U,k}) \right. \\
& +H(\hat{\mathbf{Y}}_{2,k} | \mathbf{X}_{2,k}, \mathbf{X}_{1,U,k}) + H(\mathbf{X}_{2,U,k}) \\
& \left. +H(\hat{\mathbf{Y}}_{1,k} | \mathbf{X}_{1,k}, \mathbf{X}_{2,U,k}) \right] + N\delta(N), \\
= & N \left[H(\mathbf{X}_{1,P,k}) + H(\mathbf{X}_{2,P,k}) + H(\mathbf{X}_{1,U,k}) \right. \\
& +H(\mathbf{X}_{1,CF_2,k}, \mathbf{X}_{1,DF,k} | \mathbf{X}_{2,k}, \mathbf{X}_{1,U,k}) + H(\mathbf{X}_{2,U,k}) \\
& \left. +H(\mathbf{X}_{2,CF_1,k}, \mathbf{X}_{2,DF,k} | \mathbf{X}_{1,k}, \mathbf{X}_{2,U,k}) \right] + N\delta(N) \\
\leq & N \left[H(\mathbf{X}_{1,P,k}) + H(\mathbf{X}_{2,P,k}) + H(\mathbf{X}_{1,U,k}) \right. \\
& +H(\mathbf{X}_{1,CF_2,k}, \mathbf{X}_{1,DF,k} | \mathbf{X}_{1,U,k}) + H(\mathbf{X}_{2,U,k}) \\
& \left. +H(\mathbf{X}_{2,CF_1,k}, \mathbf{X}_{2,DF,k} | \mathbf{X}_{2,U,k}) \right] + N\delta(N), \\
\leq & N \left[\dim \mathbf{X}_{1,P,k} + \dim \mathbf{X}_{2,P,k} + \dim \mathbf{X}_{1,U,k} \right. \\
& + \left(\dim(\mathbf{X}_{1,CF_2,k}, \mathbf{X}_{1,DF,k}) - \dim \mathbf{X}_{1,U,k} \right)^+ \\
& + \dim \mathbf{X}_{2,U,k} + \left(\dim(\mathbf{X}_{2,CF_1,k}, \mathbf{X}_{2,DF,k}) \right. \\
& \left. \left. - \dim \mathbf{X}_{2,U,k} \right)^+ \right] + N\delta(N). \tag{64}
\end{aligned}$$

where, (a) follows from Fano's inequality; (b) follows from the fact that $H(Y) - H(X) = H(Y|X) - H(X|Y)$; (c) follows from the fact that $H(\mathbf{X}_{i,C}, \mathbf{X}_{j,U}, \hat{\mathbf{Y}}_i | W_i, W_j, \hat{\mathbf{Y}}_j) = 0$; (d) follows from the fact that $\mathbf{X}_{i,n} = f_i^{(n)}(W_i, \hat{\mathbf{Y}}_{i,(1:n-1)})$; and (e) follows from the fact that conditioning reduces the entropy. Plugging (54b), (56), and (58) in (64) and after some trivial

manipulations, the following holds in the asymptotic regime:

$$\begin{aligned}
R_1 + R_2 \leq & \max \left((\bar{n}_{11} - n_{12})^+, n_{21}, \right. \\
& \bar{n}_{11} - (\max(\bar{n}_{11}, n_{12}) - \bar{n}_{11})^+ \Big) \\
& + \max \left((\bar{n}_{22} - n_{21})^+, n_{12}, \right. \\
& \left. \bar{n}_{22} - (\max(\bar{n}_{22}, n_{21}) - \bar{n}_{22})^+ \right). \tag{65}
\end{aligned}$$

This completes the proof of (15d).

Proof of (15e): From the assumption that the message indices W_i and W_j are i.i.d. following a uniform distribution over the sets \mathcal{W}_i and \mathcal{W}_j respectively, for all $i \in \{1, 2\}$, with $j \in \{1, 2\} \setminus \{i\}$, the following holds for any $k \in \{1, 2, \dots, N\}$:

$$\begin{aligned}
N(2R_i + R_j) & = 2H(W_i) + H(W_j) \\
& \stackrel{(a)}{\leq} I(W_i; \bar{\mathbf{Y}}_i, \hat{\mathbf{Y}}_i) + I(W_i; \bar{\mathbf{Y}}_i, \hat{\mathbf{Y}}_j | W_j) \\
& \quad + I(W_j; \bar{\mathbf{Y}}_j, \hat{\mathbf{Y}}_j) + N\delta(N) \\
& \stackrel{(b)}{=} H(\bar{\mathbf{Y}}_i) - H(\hat{\mathbf{Y}}_i | W_i) - H(\bar{\mathbf{Y}}_i | W_i, \hat{\mathbf{Y}}_i) \\
& \quad + H(\bar{\mathbf{Y}}_i | W_j, \hat{\mathbf{Y}}_j) + H(\bar{\mathbf{Y}}_j) - H(\bar{\mathbf{Y}}_j | W_j, \hat{\mathbf{Y}}_j) \\
& \quad + N\delta(N) \\
& = H(\bar{\mathbf{Y}}_i) - H(\hat{\mathbf{Y}}_i | W_i) - H(\mathbf{X}_{j,C}, \mathbf{X}_{j,D} | W_i, \hat{\mathbf{Y}}_i) \\
& \quad + H(\bar{\mathbf{Y}}_i | W_j, \hat{\mathbf{Y}}_j) \\
& \quad + H(\bar{\mathbf{Y}}_j) - H(\mathbf{X}_{i,C}, \mathbf{X}_{i,D} | W_j, \hat{\mathbf{Y}}_j) + N\delta(N) \\
& \leq H(\bar{\mathbf{Y}}_i) - H(\hat{\mathbf{Y}}_i | W_i) - H(\mathbf{X}_{j,C}, \mathbf{X}_{i,U} | W_i, \hat{\mathbf{Y}}_i) \\
& \quad + H(\bar{\mathbf{Y}}_i | W_j, \hat{\mathbf{Y}}_j) + H(\bar{\mathbf{Y}}_j) - H(\mathbf{X}_{i,C} | W_j, \hat{\mathbf{Y}}_j) \\
& \quad + N\delta(N) \\
& \leq H(\bar{\mathbf{Y}}_i) - H(\hat{\mathbf{Y}}_i | W_i) + [I(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}; W_i, \hat{\mathbf{Y}}_i) \\
& \quad - H(\mathbf{X}_{j,C}, \mathbf{X}_{i,U})] + H(\bar{\mathbf{Y}}_i, \mathbf{X}_{i,C} | W_j, \hat{\mathbf{Y}}_j) \\
& \quad + H(\bar{\mathbf{Y}}_j) - H(\mathbf{X}_{i,C} | W_j, \hat{\mathbf{Y}}_j) + N\delta(N)
\end{aligned}$$

$$\begin{aligned}
&= H(\vec{\mathbf{Y}}_i) - H(\overleftarrow{\mathbf{Y}}_i|W_i) + \left[I(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}; W_i, \overleftarrow{\mathbf{Y}}_i) \right. \\
&\quad \left. - H(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}) \right] + H(\vec{\mathbf{Y}}_i|W_j, \overleftarrow{\mathbf{Y}}_j, \mathbf{X}_{i,C}) \\
&\quad + H(\vec{\mathbf{Y}}_j) + N\delta(N) \\
&\leq H(\vec{\mathbf{Y}}_i) - H(\overleftarrow{\mathbf{Y}}_i|W_i) + \left[I(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}; W_i, \overleftarrow{\mathbf{Y}}_i) \right. \\
&\quad \left. - H(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}) \right] + H(\vec{\mathbf{Y}}_i|W_j, \overleftarrow{\mathbf{Y}}_j, \mathbf{X}_{i,C}) \\
&\quad + H(\vec{\mathbf{Y}}_j, \mathbf{X}_{j,C}, \mathbf{X}_{i,U}) + N\delta(N) \\
&\stackrel{(c)}{=} H(\vec{\mathbf{Y}}_i) - H(\overleftarrow{\mathbf{Y}}_i|W_i) + I(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}; W_i, \overleftarrow{\mathbf{Y}}_i) \\
&\quad + H(\vec{\mathbf{Y}}_i|W_j, \overleftarrow{\mathbf{Y}}_j, \mathbf{X}_{i,C}) + H(\vec{\mathbf{Y}}_j|\mathbf{X}_{j,C}, \mathbf{X}_{i,U}) \\
&\quad + N\delta(N) \\
&\leq H(\vec{\mathbf{Y}}_i) - H(\overleftarrow{\mathbf{Y}}_i|W_i) \\
&\quad + I(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}, W_j, \overleftarrow{\mathbf{Y}}_j; W_i, \overleftarrow{\mathbf{Y}}_i) \\
&\quad + H(\vec{\mathbf{Y}}_i|W_j, \overleftarrow{\mathbf{Y}}_j, \mathbf{X}_{i,C}) + H(\vec{\mathbf{Y}}_j|\mathbf{X}_{j,C}, \mathbf{X}_{i,U}) \\
&\quad + N\delta(N) \\
&\stackrel{(d)}{=} H(\vec{\mathbf{Y}}_i) - H(\overleftarrow{\mathbf{Y}}_i|W_j, W_i,) \\
&\quad + H(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}, \overleftarrow{\mathbf{Y}}_j|W_j) + H(\vec{\mathbf{Y}}_i|W_j, \overleftarrow{\mathbf{Y}}_j, \mathbf{X}_{i,C}) \\
&\quad + H(\vec{\mathbf{Y}}_j|\mathbf{X}_{j,C}, \mathbf{X}_{i,U}) + N\delta(N) \\
&\leq H(\vec{\mathbf{Y}}_i) + H(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}, \overleftarrow{\mathbf{Y}}_j|W_j) \\
&\quad + H(\vec{\mathbf{Y}}_i|W_j, \overleftarrow{\mathbf{Y}}_j, \mathbf{X}_{i,C}) + H(\vec{\mathbf{Y}}_j|\mathbf{X}_{j,C}, \mathbf{X}_{i,U}) \\
&\quad + N\delta(N) \\
&\leq \sum_{n=1}^N \left[H(\vec{\mathbf{Y}}_{i,n}) + H(\mathbf{X}_{j,C,n}, \mathbf{X}_{i,U,n}, \overleftarrow{\mathbf{Y}}_{j,n}|W_j, \right. \\
&\quad \left. \mathbf{X}_{j,C,(1:n-1)}, \mathbf{X}_{i,U,(1:n-1)}, \overleftarrow{\mathbf{Y}}_{j,(1:n-1)}) \right. \\
&\quad \left. + H(\vec{\mathbf{Y}}_{i,n}|W_j, \overleftarrow{\mathbf{Y}}_j, \mathbf{X}_{i,C}, \vec{\mathbf{Y}}_{i,(1:n-1)}) \right. \\
&\quad \left. + H(\vec{\mathbf{Y}}_{j,n}|\mathbf{X}_{j,C}, \mathbf{X}_{i,U}, \vec{\mathbf{Y}}_{j,(1:n-1)}) \right] \\
&\quad + N\delta(N) \\
&= \sum_{n=1}^N \left[H(\vec{\mathbf{Y}}_{i,n}) + H(\mathbf{X}_{j,C,n}, \mathbf{X}_{i,U,n}, \overleftarrow{\mathbf{Y}}_{j,n}|W_j, \right. \\
&\quad \left. \mathbf{X}_{j,C,(1:n-1)}, \mathbf{X}_{i,U,(1:n-1)}, \overleftarrow{\mathbf{Y}}_{j,(1:n-1)}, \mathbf{X}_{j,(1:n)}) \right. \\
&\quad \left. + H(\vec{\mathbf{Y}}_{i,n}|W_j, \overleftarrow{\mathbf{Y}}_j, \mathbf{X}_{i,C}, \vec{\mathbf{Y}}_{i,(1:n-1)}, \mathbf{X}_{j,(1:n)}) \right. \\
&\quad \left. + H(\vec{\mathbf{Y}}_{j,n}|\mathbf{X}_{j,C}, \mathbf{X}_{i,U}, \vec{\mathbf{Y}}_{j,(1:n-1)}) \right] \\
&\quad + N\delta(N) \\
&\leq \sum_{n=1}^N \left[H(\vec{\mathbf{Y}}_{i,n}) + H(\mathbf{X}_{i,U,n}|\mathbf{X}_{j,n}) \right. \\
&\quad \left. + H(\overleftarrow{\mathbf{Y}}_{j,n}|\mathbf{X}_{j,n}, \mathbf{X}_{i,U,n}) + H(\vec{\mathbf{Y}}_{i,n}|\mathbf{X}_{i,C,n}, \mathbf{X}_{j,n}) \right. \\
&\quad \left. + H(\vec{\mathbf{Y}}_{j,n}|\mathbf{X}_{j,C,n}, \mathbf{X}_{i,U,n}) \right] + N\delta(N) \\
&\leq N \left[H(\vec{\mathbf{Y}}_{i,k}) + H(\mathbf{X}_{i,U,k}) + H(\overleftarrow{\mathbf{Y}}_{j,k}|\mathbf{X}_{j,k}, \mathbf{X}_{i,U,k}) \right. \\
&\quad \left. + H(\mathbf{X}_{i,P,k}) + H(\mathbf{X}_{j,P,k}) \right] + N\delta(N)
\end{aligned}$$

$$\begin{aligned}
&= N \left[H(\vec{\mathbf{Y}}_{i,k}) + H(\mathbf{X}_{i,U,k}) \right. \\
&\quad \left. + H(\mathbf{X}_{i,CF_j,k}, \mathbf{X}_{i,DF,k}|\mathbf{X}_{i,U,k}) + H(\mathbf{X}_{i,P,k}) \right. \\
&\quad \left. + H(\mathbf{X}_{j,P,k}) \right] + N\delta(N), \\
&\leq N \left[\dim \overleftarrow{\mathbf{Y}}_{i,k} + \dim \vec{\mathbf{Y}}_{i,G,k} + \dim \mathbf{X}_{i,U,k} \right. \\
&\quad \left. + (\dim(\mathbf{X}_{i,CF_j,k}, \mathbf{X}_{i,DF,k}) - \dim \mathbf{X}_{i,U,k})^+ \right. \\
&\quad \left. + \dim \mathbf{X}_{i,P,k} + \dim \mathbf{X}_{j,P,k} \right] + N\delta(N), \tag{66}
\end{aligned}$$

where, (a) follows from Fano's inequality; (b) follows from the fact that $H(\vec{\mathbf{Y}}_i, \overleftarrow{\mathbf{Y}}_j|W_i, W_j) = 0$; (c) follows from the fact that $H(Y|X) = H(X, Y) - H(X)$; and (d) follows from the fact that $H(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}, \overleftarrow{\mathbf{Y}}_j|W_j, W_i, \overleftarrow{\mathbf{Y}}_i) = 0$.

Plugging (54b), (56), (58), (59a), and (59b) in (66) and after some trivial manipulations, the following holds in the asymptotic regime:

$$\begin{aligned}
2R_i + R_j &\leq \max(\overleftarrow{n}_{ii}, n_{ji}) + (\overleftarrow{n}_{ii} - n_{ji})^+ \\
&\quad + \max\left((\overleftarrow{n}_{jj} - n_{ji})^+, n_{ij}, \right. \\
&\quad \left. \overleftarrow{n}_{jj} - (\max(\overleftarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+ \right). \tag{67}
\end{aligned}$$

This completes the proof of (15e).

APPENDIX C PROOF OF THEOREM 3

The outer bounds (29a) and (29c) correspond to the outer bounds of the case of perfect channel-output feedback derived in [3]. The bounds (29b), (29d) and (29e) correspond to new outer bounds. Before presenting the proof, consider the parameter $h_{ji,U}$, with $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$, defined as follows:

$$h_{ji,U} = \begin{cases} 0 & \text{if } (S_{1,i} \vee S_{2,i} \vee S_{3,i}) \\ \sqrt{\frac{\text{INR}_{ij} \text{INR}_{ji}}{\text{SNR}_j}} & \text{if } (S_{4,i} \vee S_{5,i}), \end{cases} \tag{68}$$

where, the events $S_{1,i}$, $S_{2,i}$, $S_{3,i}$, $S_{4,i}$, and $S_{5,i}$ are defined in (24). Consider also the following signals:

$$\mathbf{X}_{i,C,n} = \sqrt{\text{INR}_{ji}} \mathbf{X}_{i,n} + \vec{\mathbf{Z}}_{j,n} \text{ and} \tag{69}$$

$$\mathbf{X}_{i,U,n} = h_{ji,U} \mathbf{X}_{i,n} + \vec{\mathbf{Z}}_{j,n}, \tag{70}$$

where, $\mathbf{X}_{i,n}$ and $\vec{\mathbf{Z}}_{j,n}$ are the channel input of transmitter i and the noise observed at receiver j during a given channel use $n \in \{1, 2, \dots, N\}$, as described by (2). The following lemma is instrumental in the present proof of Theorem 3.

Lemma 1: For all $i \in \{1, 2\}$, with $j \in \{1, 2\} \setminus \{i\}$, the following holds:

$$\begin{aligned}
I(\mathbf{X}_{i,C}, \mathbf{X}_{j,U}, \overleftarrow{\mathbf{Y}}_i, W_i; \overleftarrow{\mathbf{Y}}_j, W_j) &\leq h(\overleftarrow{\mathbf{Y}}_j|W_j) \\
&\quad + \sum_{n=1}^N \left[h(\mathbf{X}_{j,U,n}|\mathbf{X}_{i,C,n}) + h(\overleftarrow{\mathbf{Y}}_{i,n}|\mathbf{X}_{i,n}, \mathbf{X}_{j,U,n}) \right. \\
&\quad \left. - \frac{3}{2} \log(2\pi e) \right]. \tag{71}
\end{aligned}$$

Proof: The proof of Lemma 1 is presented in this appendix. ■

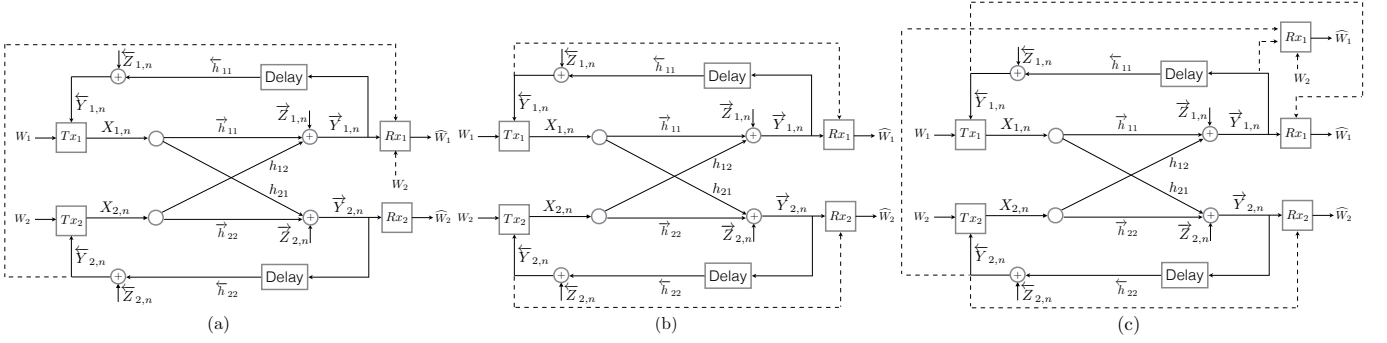


Fig. 12. Genie-Aided G-IC-NOF models for channel use n . (a) Model used to calculate the outer bound on R_1 ; (b) Model used to calculate the outer bound on $R_1 + R_2$; and (c) Model used to calculate the outer bound on $2R_1 + R_2$

Proof of (29b): From the assumption that the message index W_i is i.i.d. following a uniform distribution over the set \mathcal{W}_i , the following holds for any $k \in \{1, 2, \dots, N\}$:

$$\begin{aligned}
NR_i &= H(W_i) \\
&= H(W_i|W_j) \\
&\stackrel{(a)}{\leq} I(W_i; \vec{Y}_i, \check{Y}_j|W_j) + N\delta(N) \\
&\leq \sum_{n=1}^N \left[h(\vec{Y}_{i,n}, \check{Y}_{j,n}|W_j, \vec{Y}_{i,(1:n-1)}, \check{Y}_{j,(1:n-1)}, \mathbf{X}_{j,(1:n)}) - h(\vec{Z}_{i,n}) - h(\check{Z}_{j,n}) \right] + N\delta(N) \\
&\leq \sum_{n=1}^N \left[h(\vec{Y}_{i,n}, \check{Y}_{j,n}|X_{j,n}) - h(\vec{Z}_{i,n}) - h(\check{Z}_{j,n}) \right] + N\delta(N) \\
&= N \left[h(\vec{Y}_{i,k}, \check{Y}_{j,k}|X_{j,k}) - \log(2\pi e) \right] + N\delta(N), \quad (72)
\end{aligned}$$

where (a) follows from Fano's inequality (see Figure 12a). From (72), the following holds in the asymptotic regime:

$$\begin{aligned}
R_i &\leq h(\vec{Y}_{i,k}, \check{Y}_{j,k}|X_{j,k}) - \log(2\pi e) \\
&\leq \frac{1}{2} \log(b_{3,i} + 1) \\
&\quad + \frac{1}{2} \log \left(\frac{(b_{3,i} + b_{4,j}(\rho) + 1) \overline{\text{SNR}}_j}{(b_{1,j}(\rho) + 1)(b_{3,i} + (1 - \rho^2))} + 1 \right). \quad (73)
\end{aligned}$$

This completes the proof of (29b).

Proof of (29d): From the assumption that the message indices W_1 and W_2 are i.i.d. following a uniform distribution over the sets \mathcal{W}_1 and \mathcal{W}_2 respectively, the following holds for any $k \in \{1, 2, \dots, N\}$:

$$\begin{aligned}
N(R_1 + R_2) &= H(W_1) + H(W_2) \\
&\stackrel{(a)}{\leq} I(W_1; \vec{Y}_1, \check{Y}_1) + I(W_2; \vec{Y}_2, \check{Y}_2) + N\delta(N) \\
&= h(\vec{Y}_1) + h(\vec{Z}_1|\vec{Y}_1) - h(\check{Y}_1|W_1) \\
&\quad - h(\vec{Y}_1|W_1, \check{Y}_1, \mathbf{X}_1) + h(\vec{Y}_2) + h(\vec{Z}_2|\vec{Y}_2) \\
&\quad - h(\check{Y}_2|W_2) - h(\vec{Y}_2|W_2, \check{Y}_2, \mathbf{X}_2) + N\delta(N)
\end{aligned}$$

$$\begin{aligned}
&\leq h(\vec{Y}_1) + h(\vec{Z}_1) - h(\check{Y}_1|W_1) \\
&\quad - h(\mathbf{X}_{2,C}|W_1, \check{Y}_1, \mathbf{X}_1) + h(\vec{Y}_2) + h(\vec{Z}_2) \\
&\quad - h(\check{Y}_2|W_2) - h(\mathbf{X}_{1,C}|W_2, \check{Y}_2, \mathbf{X}_2) + N\delta(N) \\
&= h(\vec{Y}_1) - h(\check{Y}_1|W_1) \\
&\quad - h(\mathbf{X}_{2,C}, \vec{Z}_2|W_1, \check{Y}_1, \mathbf{X}_1) \\
&\quad + h(\vec{Z}_2|W_1, \check{Y}_1, \mathbf{X}_1, \mathbf{X}_{2,C}) + h(\vec{Y}_2) \\
&\quad - h(\check{Y}_2|W_2) - h(\mathbf{X}_{1,C}, \vec{Z}_1|W_2, \check{Y}_2, \mathbf{X}_2) \\
&\quad + h(\vec{Z}_1|W_2, \check{Y}_2, \mathbf{X}_2, \mathbf{X}_{1,C}) + N \log(2\pi e) \\
&\quad + N\delta(N) \\
&= h(\vec{Y}_1) - h(\check{Y}_1|W_1) \\
&\quad - h(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}|W_1, \check{Y}_1, \mathbf{X}_1) \\
&\quad + h(\vec{Z}_2|W_1, \check{Y}_1, \mathbf{X}_1, \mathbf{X}_{2,C}) + h(\vec{Y}_2) \\
&\quad - h(\check{Y}_2|W_2) - h(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}|W_2, \check{Y}_2, \mathbf{X}_2) \\
&\quad + h(\vec{Z}_1|W_2, \check{Y}_2, \mathbf{X}_2, \mathbf{X}_{1,C}) + N \log(2\pi e) \\
&\quad + N\delta(N) \\
&= h(\vec{Y}_1) - h(\check{Y}_1|W_1) + \left[I(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}; W_1, \check{Y}_1) \right. \\
&\quad \left. - h(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}) \right] + h(\vec{Y}_2) - h(\check{Y}_2|W_2) \\
&\quad + \left[I(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}; W_2, \check{Y}_2) - h(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}) \right] \\
&\quad + h(\vec{Z}_1|W_2, \check{Y}_2, \mathbf{X}_2, \mathbf{X}_{1,C}) \\
&\quad + h(\vec{Z}_2|W_1, \check{Y}_1, \mathbf{X}_1, \mathbf{X}_{2,C}) + N \log(2\pi e) \\
&\quad + N\delta(N) \\
&\leq h(\vec{Y}_1) - h(\check{Y}_1|W_1) + \left[I(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}; W_1, \check{Y}_1) \right. \\
&\quad \left. - h(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}) \right] + h(\vec{Y}_2) - h(\check{Y}_2|W_2) \\
&\quad + \left[I(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}; W_2, \check{Y}_2) - h(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}) \right] \\
&\quad + \left[h(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}|\vec{Y}_2) - h(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}|\vec{Y}_2, \mathbf{X}_1, \mathbf{X}_2) \right] \\
&\quad + \left[h(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}|\vec{Y}_1) - h(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}|\vec{Y}_1, \mathbf{X}_2, \mathbf{X}_1) \right] \\
&\quad + h(\vec{Z}_1|W_2, \check{Y}_2, \mathbf{X}_2, \mathbf{X}_{1,C}) \\
&\quad + h(\vec{Z}_2|W_1, \check{Y}_1, \mathbf{X}_1, \mathbf{X}_{2,C}) + N \log(2\pi e) + N\delta(N)
\end{aligned}$$

$$\begin{aligned}
&\stackrel{(b)}{=} h(\vec{Y}_1 | \mathbf{X}_{1,C}, \mathbf{X}_{2,U}) - h(\vec{Y}_1 | W_1) \\
&\quad + I(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}; W_1, \vec{Y}_1) + h(\vec{Y}_2 | \mathbf{X}_{2,C}, \mathbf{X}_{1,U}) \\
&\quad - h(\vec{Y}_2 | W_2) + I(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}; W_2, \vec{Y}_2) \\
&\quad - h(\vec{Z}_1, \vec{Z}_2 | \vec{Y}_2, \mathbf{X}_1, \mathbf{X}_2) - h(\vec{Z}_2, \vec{Z}_1 | \vec{Y}_1, \mathbf{X}_2, \mathbf{X}_1) \\
&\quad + h(\vec{Z}_1 | W_2, \vec{Y}_2, \mathbf{X}_2, \mathbf{X}_{1,C}) \\
&\quad + h(\vec{Z}_2 | W_1, \vec{Y}_1, \mathbf{X}_1, \mathbf{X}_{2,C}) + N \log(2\pi e) + N\delta(N) \\
&\stackrel{(c)}{\leq} h(\vec{Y}_1 | \mathbf{X}_{1,C}, \mathbf{X}_{2,U}) - h(\vec{Y}_1 | W_1) \\
&\quad + I(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}; W_1, \vec{Y}_1) + h(\vec{Y}_2 | \mathbf{X}_{2,C}, \mathbf{X}_{1,U}) \\
&\quad - h(\vec{Y}_2 | W_2) + I(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}; W_2, \vec{Y}_2) \\
&\quad + N \log(2\pi e) + N\delta(N) \\
&\leq h(\vec{Y}_1 | \mathbf{X}_{1,C}, \mathbf{X}_{2,U}) - h(\vec{Y}_1 | W_1) \\
&\quad + I(\mathbf{X}_{2,C}, \mathbf{X}_{1,U}, W_2, \vec{Y}_2; W_1, \vec{Y}_1) \\
&\quad + h(\vec{Y}_2 | \mathbf{X}_{2,C}, \mathbf{X}_{1,U}) - h(\vec{Y}_2 | W_2) \\
&\quad + I(\mathbf{X}_{1,C}, \mathbf{X}_{2,U}, W_1, \vec{Y}_1; W_2, \vec{Y}_2) \\
&\quad + N \log(2\pi e) + N\delta(N) \\
&\stackrel{(d)}{\leq} \sum_{n=1}^N \left[h(\vec{Y}_{1,n} | \mathbf{X}_{1,C}, \mathbf{X}_{2,U}, \vec{Y}_{1,(1:n-1)}) \right. \\
&\quad + h(X_{1,U,n} | X_{2,C,n}) + h(\vec{Y}_{2,n} | X_{2,n}, X_{1,U,n}) \\
&\quad + h(\vec{Y}_{2,n} | \mathbf{X}_{2,C}, \mathbf{X}_{1,U}, \vec{Y}_{2,(1:n-1)}) \\
&\quad + h(X_{2,U,n} | X_{1,C,n}) + h(\vec{Y}_{1,n} | X_{1,n}, X_{2,U,n}) \\
&\quad \left. - 3 \log(2\pi e) \right] + N \log(2\pi e) + N\delta(N) \\
&\leq \sum_{n=1}^N \left[h(\vec{Y}_{1,n} | X_{1,C,n}, X_{2,U,n}) + h(X_{1,U,n} | X_{2,C,n}) \right. \\
&\quad + h(\vec{Y}_{2,n} | X_{2,n}, X_{1,U,n}) + h(\vec{Y}_{2,n} | X_{2,C,n}, X_{1,U,n}) \\
&\quad + h(X_{2,U,n} | X_{1,C,n}) + h(\vec{Y}_{1,n} | X_{1,n}, X_{2,U,n}) \\
&\quad \left. - 3 \log(2\pi e) \right] + N \log(2\pi e) + N\delta(N) \\
&= N \left[h(\vec{Y}_{1,k} | X_{1,C,k}, X_{2,U,k}) + h(X_{1,U,k} | X_{2,C,k}) \right. \\
&\quad + h(\vec{Y}_{2,k} | X_{2,k}, X_{1,U,k}) + h(\vec{Y}_{2,k} | X_{2,C,k}, X_{1,U,k}) \\
&\quad + h(X_{2,U,k} | X_{1,C,k}) + h(\vec{Y}_{1,k} | X_{1,k}, X_{2,U,k}) \\
&\quad \left. - 3 \log(2\pi e) \right] + N \log(2\pi e) + N\delta(N), \tag{74}
\end{aligned}$$

where (a) follows from Fano's inequality (see Figure 12b); (b) follows from the fact that $h(\vec{Y}_i) - h(\mathbf{X}_{i,C}, \mathbf{X}_{j,U}) + h(\mathbf{X}_{i,C}, \mathbf{X}_{j,U} | \vec{Y}_i) = h(\vec{Y}_i | \mathbf{X}_{i,C}, \mathbf{X}_{j,U})$; (c) follows from the fact that $h(\vec{Z}_i | W_j, \vec{Y}_j, \mathbf{X}_j, \mathbf{X}_{i,C}) - h(\vec{Z}_i, \vec{Z}_j | \vec{Y}_j, \mathbf{X}_i, \mathbf{X}_j) \leq 0$; and (d) follows from Lemma 1.

From (74), the following holds in the asymptotic regime for any $k \in \{1, 2, \dots, N\}$:

$$R_1 + R_2 \leq h(\vec{Y}_{1,k} | X_{1,C,k}, X_{2,U,k}) + h(X_{1,U,k} | X_{2,C,k})$$

$$\begin{aligned}
&\quad + h(\vec{Y}_{2,k} | X_{2,k}, X_{1,U,k}) + h(\vec{Y}_{2,k} | X_{2,C,k}, X_{1,U,k}) \\
&\quad + h(X_{2,U,k} | X_{1,C,k}) + h(\vec{Y}_{1,k} | X_{1,k}, X_{2,U,k}) \\
&\quad - 2 \log(2\pi e) \\
&\leq \frac{1}{2} \log(\det(\text{Var}(\vec{Y}_{1,k}, X_{1,C,k}, X_{2,U,k}))) \\
&\quad - \frac{1}{2} \log(\text{INR}_{12} + 1) \\
&\quad + \frac{1}{2} \log(\det(\text{Var}(\vec{Y}_{2,k}, X_{2,k}, X_{1,U,k}))) \\
&\quad - \frac{1}{2} \log(\det(\text{Var}(X_{2,k}, X_{1,U,k}))) \\
&\quad + \frac{1}{2} \log(\det(\text{Var}(\vec{Y}_{2,k}, X_{2,C,k}, X_{1,U,k}))) \\
&\quad - \frac{1}{2} \log(\text{INR}_{21} + 1) \\
&\quad + \frac{1}{2} \log(\det(\text{Var}(\vec{Y}_{1,k}, X_{1,k}, X_{2,U,k}))) \\
&\quad - \frac{1}{2} \log(\det(\text{Var}(X_{1,k}, X_{2,U,k}))) + \log(2\pi e), \tag{75}
\end{aligned}$$

where, for all $i \in \{1, 2\}$, with $j \in \{1, 2\} \setminus \{i\}$ the following holds for any $k \in \{1, 2, \dots, N\}$:

$$\begin{aligned}
\det(\text{Var}(\vec{Y}_{j,k}, X_{j,C,k}, X_{i,U,k})) &= \overline{\text{SNR}}_j + \text{INR}_{ji} + h_{ji,U}^2 \\
&\quad - 2h_{ji,U} \sqrt{\text{INR}_{ji}} + (1 - \rho^2) (\text{INR}_{ij} \text{INR}_{ji} \\
&\quad + h_{ji,U}^2 (\overline{\text{SNR}}_j + \text{INR}_{ij}) - 2h_{ji,U} \text{INR}_{ij} \sqrt{\text{INR}_{ji}}) \\
&\quad + 2\rho \sqrt{\overline{\text{SNR}}_j} (\sqrt{\text{INR}_{ji}} - h_{ji,U}), \tag{76a}
\end{aligned}$$

$$\begin{aligned}
\det(\text{Var}(\vec{Y}_{j,k}, X_{j,k}, X_{i,U,k})) &= 1 + h_{ji,U}^2 (1 - \rho^2) \\
&\quad + \frac{\overline{\text{SNR}}_j (1 - \rho^2) (h_{ji,U}^2 - 2h_{ji,U} \sqrt{\text{INR}_{ji}} + \text{INR}_{ji})}{(\overline{\text{SNR}}_j + 2\rho \sqrt{\overline{\text{SNR}}_j} \text{INR}_{ji} + \text{INR}_{ji} + 1)}, \text{ and} \\
&\tag{76b}
\end{aligned}$$

$$\det(\text{Var}(X_{j,k}, X_{i,U,k})) = 1 + (1 - \rho^2) h_{ji,U}^2. \tag{76c}$$

The expressions in (76) depend on $S_{1,i}$, $S_{2,i}$, $S_{3,i}$, $S_{4,i}$, and $S_{5,i}$ via the parameter $h_{ji,U}$ in (68). Hence, the following cases are identified:

Case 1: $(S_{1,2} \vee S_{2,2} \vee S_{5,2}) \wedge (S_{1,1} \vee S_{2,1} \vee S_{5,1})$. From (68), it follows that $h_{12,U} = 0$ and $h_{21,U} = 0$. Therefore, plugging the expression (76) into (75) yields (26a).

Case 2: $(S_{1,2} \vee S_{2,2} \vee S_{5,2}) \wedge (S_{3,1} \vee S_{4,1})$. From (68), it follows that $h_{12,U} = 0$ and $h_{21,U} = \sqrt{\frac{\text{INR}_{12} \text{INR}_{21}}{\text{SNR}_2}}$. Therefore, plugging the expression (76) into (75) yields (26b).

Case 3: $(S_{3,2} \vee S_{4,2}) \wedge (S_{1,1} \vee S_{2,1} \vee S_{5,1})$. From (68), it follows that $h_{12,U} = \sqrt{\frac{\text{INR}_{12} \text{INR}_{21}}{\text{SNR}_1}}$ and $h_{21,U} = 0$. Therefore, plugging the expression (76) into (75) yields (26c).

Case 4: $(S_{3,2} \vee S_{4,2}) \wedge (S_{3,1} \vee S_{4,1})$. From (68), it follows that $h_{12,U} = \sqrt{\frac{\text{INR}_{12} \text{INR}_{21}}{\text{SNR}_1}}$ and $h_{21,U} = \sqrt{\frac{\text{INR}_{12} \text{INR}_{21}}{\text{SNR}_2}}$. Therefore, plugging the expression (76) into (75) yields (26d).

This completes the proof of (29d).

Proof of (29e): From the assumption that the message indices W_i and W_j are i.i.d. following a uniform distribution over the sets \mathcal{W}_i and \mathcal{W}_j respectively, for all $i \in \{1, 2\}$, with $j \in$

$\{1, 2\} \setminus \{i\}$, the following holds for any $k \in \{1, 2, \dots, N\}$:

$$\begin{aligned}
& N \left(2R_i + R_j \right) = 2H(W_i) + H(W_j) \\
& \stackrel{(a)}{=} H(W_i) + H(W_i|W_j) + H(W_j) \\
& \stackrel{(b)}{\leq} I(W_i; \vec{Y}_i, \hat{Y}_i) + I(W_i; \vec{Y}_i, \hat{Y}_j|W_j) \\
& \quad + I(W_j; \vec{Y}_j, \hat{Y}_j) + N\delta(N) \\
& \leq h(\vec{Y}_i) + h(\vec{Z}_i) - h(\hat{Y}_i|W_i) - h(\vec{Y}_i|W_i, \hat{Y}_i) \\
& \quad + h(\hat{Y}_j|W_j) - h(\hat{Y}_j|W_i, W_j) + I(W_i; \vec{Y}_i|W_j, \hat{Y}_j) \\
& \quad + h(\vec{Y}_j) + h(\vec{Z}_j) - h(\hat{Y}_j|W_j) - h(\vec{Y}_j|W_j, \hat{Y}_j) \\
& \quad + N\delta(N) \\
& = h(\vec{Y}_i) - h(\hat{Y}_i|W_i) - h(\vec{Y}_i|W_i, \hat{Y}_i, \mathbf{X}_i) \\
& \quad - h(\hat{Y}_j|W_i, W_j) + I(W_i; \vec{Y}_i|W_j, \hat{Y}_j) + h(\vec{Y}_j) \\
& \quad - h(\vec{Y}_j|W_j, \hat{Y}_j, \mathbf{X}_j) + N \log(2\pi e) + N\delta(N) \\
& \leq h(\vec{Y}_i) - h(\hat{Y}_i|W_i) - h(\vec{Y}_i|W_i, \hat{Y}_i, \mathbf{X}_i) \\
& \quad + I(W_i; \vec{Y}_i|W_j, \hat{Y}_j) + h(\vec{Y}_j) \\
& \quad - h(\vec{Y}_j|W_j, \hat{Y}_j, \mathbf{X}_j) + N \log(2\pi e) + N\delta(N) \\
& \stackrel{(c)}{=} h(\vec{Y}_i) - h(\hat{Y}_i|W_i) - h(\mathbf{X}_{j,C}|W_i, \hat{Y}_i, \mathbf{X}_i) \\
& \quad + I(W_i; \vec{Y}_i|W_j, \hat{Y}_j) + h(\vec{Y}_j) \\
& \quad - h(\mathbf{X}_{i,C}|W_j, \hat{Y}_j, \mathbf{X}_j) + N \log(2\pi e) + N\delta(N) \\
& = h(\vec{Y}_i) - h(\hat{Y}_i|W_i) - h(\mathbf{X}_{j,C}, \vec{Z}_j|W_i, \hat{Y}_i, \mathbf{X}_i) \\
& \quad + h(\vec{Z}_j|W_i, \hat{Y}_i, \mathbf{X}_i, \mathbf{X}_{j,C}) + I(W_i; \vec{Y}_i|W_j, \hat{Y}_j) \\
& \quad + h(\vec{Y}_j) - h(\mathbf{X}_{i,C}|W_j, \hat{Y}_j, \mathbf{X}_j) + N \log(2\pi e) \\
& \quad + N\delta(N) \\
& \stackrel{(d)}{=} h(\vec{Y}_i) - h(\hat{Y}_i|W_i) - h(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}|W_i, \hat{Y}_i, \mathbf{X}_i) \\
& \quad + h(\vec{Z}_j|W_i, \hat{Y}_i, \mathbf{X}_i, \mathbf{X}_{j,C}) + I(W_i; \vec{Y}_i|W_j, \hat{Y}_j) \\
& \quad + h(\vec{Y}_j) \\
& \quad - h(\mathbf{X}_{i,C}|W_j, \hat{Y}_j, \mathbf{X}_j) + N \log(2\pi e) + N\delta(N) \\
& \leq h(\vec{Y}_i) - h(\hat{Y}_i|W_i) - h(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}|W_i, \hat{Y}_i) \\
& \quad + h(\vec{Z}_j|W_i, \hat{Y}_i, \mathbf{X}_i, \mathbf{X}_{j,C}) \\
& \quad + I(W_i; \vec{Y}_i, \mathbf{X}_{i,C}|W_j, \hat{Y}_j) + h(\vec{Y}_j) \\
& \quad - h(\mathbf{X}_{i,C}|W_j, \hat{Y}_j) + N \log(2\pi e) + N\delta(N) \\
& = h(\vec{Y}_i) - h(\hat{Y}_i|W_i) - h(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}|W_i, \hat{Y}_i) \\
& \quad + h(\vec{Z}_j|W_i, \hat{Y}_i, \mathbf{X}_i, \mathbf{X}_{j,C}) + h(\vec{Y}_i|W_j, \hat{Y}_j, \mathbf{X}_{i,C}) \\
& \quad - h(\vec{Y}_i, \mathbf{X}_{i,C}|W_i, W_j, \hat{Y}_j) + h(\vec{Y}_j) \\
& \quad + N \log(2\pi e) + N\delta(N) \\
& \stackrel{(e)}{\leq} h(\vec{Y}_i) - h(\hat{Y}_i|W_i) - h(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}|W_i, \hat{Y}_i) \\
& \quad + h(\vec{Z}_j|W_i, \hat{Y}_i, \mathbf{X}_i, \mathbf{X}_{j,C}) + h(\vec{Y}_i|W_j, \hat{Y}_j, \mathbf{X}_{i,C}) \\
& \quad - h(\vec{Y}_i, \mathbf{X}_{i,C}|W_i, W_j, \hat{Y}_j, \mathbf{X}_i, \mathbf{X}_j) + h(\vec{Y}_j) \\
& \quad + N \log(2\pi e) + N\delta(N)
\end{aligned}$$

$$\begin{aligned}
& = h(\vec{Y}_i) - h(\hat{Y}_i|W_i) - h(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}|W_i, \hat{Y}_i) \\
& \quad + h(\vec{Z}_j|W_i, \hat{Y}_i, \mathbf{X}_i, \mathbf{X}_{j,C}) + h(\vec{Y}_i|W_j, \hat{Y}_j, \mathbf{X}_{i,C}) \\
& \quad - h(\vec{Z}_i, \vec{Z}_j|W_i, W_j, \hat{Y}_j, \mathbf{X}_i, \mathbf{X}_j) + h(\vec{Y}_j) \\
& \quad + N \log(2\pi e) + N\delta(N) \\
& \stackrel{(f)}{\leq} h(\vec{Y}_i) - h(\hat{Y}_i|W_i) - h(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}|W_i, \hat{Y}_i) \\
& \quad + h(\vec{Y}_i|W_j, \hat{Y}_j, \mathbf{X}_{i,C}) + h(\vec{Y}_j) + N \log(2\pi e) \\
& \quad + N\delta(N) \\
& \leq h(\vec{Y}_i) - h(\hat{Y}_i|W_i) + I(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}; W_i, \hat{Y}_i) \\
& \quad - h(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}) + h(\vec{Y}_i|W_j, \hat{Y}_j, \mathbf{X}_{i,C}) + h(\vec{Y}_j) \\
& \quad + h(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}|\vec{Y}_j) + N \log(2\pi e) + N\delta(N) \\
& \stackrel{(g)}{=} h(\vec{Y}_i) - h(\hat{Y}_i|W_i) + I(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}; W_i, \hat{Y}_i) \\
& \quad + h(\vec{Y}_i|W_j, \hat{Y}_j, \mathbf{X}_{i,C}) + h(\vec{Y}_j|\mathbf{X}_{j,C}, \mathbf{X}_{i,U}) \\
& \quad + N \log(2\pi e) + N\delta(N) \\
& \leq h(\vec{Y}_i) - h(\hat{Y}_i|W_i) + I(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}, W_j, \hat{Y}_j; W_i, \hat{Y}_i) \\
& \quad + h(\vec{Y}_i|W_j, \hat{Y}_j, \mathbf{X}_{i,C}) + h(\vec{Y}_j|\mathbf{X}_{j,C}, \mathbf{X}_{i,U}) \\
& \quad + N \log(2\pi e) + N\delta(N) \\
& \stackrel{(h)}{\leq} h(\vec{Y}_i) + \sum_{n=1}^N \left[h(X_{i,U,n}|X_{j,C,n}) + h(\hat{Y}_{j,n}|X_{j,n}, X_{i,U,n}) \right. \\
& \quad \left. - \frac{3}{2} \log(2\pi e) \right] + h(\vec{Y}_i|W_j, \hat{Y}_j, \mathbf{X}_{i,C}) \\
& \quad + h(\vec{Y}_j|\mathbf{X}_{j,C}, \mathbf{X}_{i,U}) + N \log(2\pi e) + N\delta(N) \\
& \stackrel{(i)}{\leq} h(\vec{Y}_i) + \sum_{n=1}^N \left[h(X_{i,U,n}|X_{j,C,n}) + h(\hat{Y}_{j,n}|X_{j,n}, X_{i,U,n}) \right. \\
& \quad \left. - \frac{3}{2} \log(2\pi e) \right] + h(\vec{Y}_i|\mathbf{X}_{i,C}, \mathbf{X}_j) + h(\vec{Y}_j|\mathbf{X}_{j,C}, \mathbf{X}_{i,U}) \\
& \quad + N \log(2\pi e) + N\delta(N) \\
& \leq \sum_{n=1}^N \left[h(\vec{Y}_{i,n}) + h(X_{i,U,n}|X_{j,C,n}) + h(\hat{Y}_{j,n}|X_{j,n}, X_{i,U,n}) \right. \\
& \quad \left. - \frac{3}{2} \log(2\pi e) + h(\vec{Y}_{i,n}|\mathbf{X}_{i,C,n}, X_{j,n}) \right. \\
& \quad \left. + h(\vec{Y}_{j,n}|X_{j,C,n}, X_{i,U,n}) \right] + N \log(2\pi e) + N\delta(N) \\
& = N \left[h(\vec{Y}_{i,k}) + h(X_{i,U,k}|X_{j,C,k}) + h(\hat{Y}_{j,k}|X_{j,k}, X_{i,U,j}) \right. \\
& \quad \left. - \frac{5}{2} \log(2\pi e) + h(\vec{Y}_{i,k}|\mathbf{X}_{i,C,k}, X_{j,k}) \right. \\
& \quad \left. + h(\vec{Y}_{j,k}|X_{j,C,k}, X_{i,U,k}) + 2 \log(2\pi e) + \delta(N) \right], \quad (77)
\end{aligned}$$

where, (a) follows from the fact that W_1 and W_2 are mutually independent; (b) follows from Fano's inequality (see Figure 12c); (c) follows from (2) and (69); (d) follows from (70); (e) follows from (4) and the fact that conditioning reduces the entropy; (f) follows from the fact that $h(\vec{Z}_j|W_j, \hat{Y}_i, \mathbf{X}_i, \mathbf{X}_{j,C}) - h(\vec{Z}_i, \vec{Z}_j|W_i, W_j, \hat{Y}_j, \mathbf{X}_i, \mathbf{X}_j) \leq 0$; (g) follows from the fact that $h(\vec{Y}_j) - h(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}) + h(\mathbf{X}_{j,C}, \mathbf{X}_{i,U}|\vec{Y}_j) = h(\vec{Y}_j|\mathbf{X}_{j,C}, \mathbf{X}_{i,U})$; (h) follows from Lemma 1; and (i)

follows from the fact that conditioning reduces the entropy. From (77), the following holds in the asymptotic regime for any $k \in \{1, 2, \dots, N\}$:

$$\begin{aligned}
2R_i + R_j &\leq h(\vec{Y}_{i,k}) + h(X_{i,U,k}|X_{j,C,k}) \\
&\quad + h(\overleftarrow{Y}_{j,k}|X_{j,k}, X_{i,U,k}) + h(\vec{Y}_{i,k}|X_{i,C,k}, X_{j,k}) \\
&\quad + h(\vec{Y}_{j,k}|X_{j,C,k}, X_{i,U,k}) - \frac{1}{2} \log(2\pi e) \\
&\leq \frac{1}{2} \log(\overrightarrow{\text{SNR}}_i + 2\rho\sqrt{\overrightarrow{\text{SNR}}_i \text{INR}_{ij}} + \text{INR}_{ij} + 1) \\
&\quad - \frac{1}{2} \log(\text{INR}_{ij} + 1) \\
&\quad + \frac{1}{2} \log(\det(\text{Var}(\overleftarrow{Y}_{j,k}, X_{j,k}, X_{i,U,k}))) \\
&\quad - \frac{1}{2} \log(\det(\text{Var}(X_{j,k}, X_{i,U,k}))) \\
&\quad + \frac{1}{2} \log(1 + (1 - \rho^2)(\overrightarrow{\text{SNR}}_i + \text{INR}_{ji})) \\
&\quad - \frac{1}{2} \log(1 + (1 - \rho^2)\text{INR}_{ji}) \\
&\quad + \frac{1}{2} \log(\det(\text{Var}(\vec{Y}_{j,k}, X_{j,C,k}, X_{i,U,k}))) \\
&\quad + 2 \log(2\pi e). \tag{78}
\end{aligned}$$

The outer bound on (78) depends on $S_{1,i}$, $S_{2,i}$, $S_{3,i}$, $S_{4,i}$, and $S_{5,i}$ via the parameter $h_{ji,U}$ in (68). Hence, as in the previous part, the following cases are identified:

Case 1: ($S_{1,i} \vee S_{2,i} \vee S_{5,i}$). From (68), it follows that $h_{ji,U} = 0$. Therefore, plugging the expressions (76) into (78) yields (27a).

Case 2: ($S_{3,i} \vee S_{4,i}$). From (68), it follows that $h_{ji,U} = \sqrt{\frac{\text{INR}_{ij}\text{INR}_{ji}}{\text{SNR}_j}}$. Therefore, plugging the expressions (76) into (78) yields (27b).

This completes the proof of (29e) and the proof of Theorem 3.

A. Proof of Lemma 1

Lemma 1 is proved as follows:

$$\begin{aligned}
&I(\mathbf{X}_{i,C}, \mathbf{X}_{j,U}, \overleftarrow{Y}_i, W_i; \overleftarrow{Y}_j, W_j) \\
&= I(W_i; \overleftarrow{Y}_j, W_j) + I(\mathbf{X}_{i,C}, \mathbf{X}_{j,U}, \overleftarrow{Y}_i; \overleftarrow{Y}_j, W_j|W_i) \\
&= h(\overleftarrow{Y}_j, W_j) - h(\overleftarrow{Y}_j, W_j|W_i) + h(\mathbf{X}_{i,C}, \mathbf{X}_{j,U}, \overleftarrow{Y}_i|W_i) \\
&\quad - h(\mathbf{X}_{i,C}, \mathbf{X}_{j,U}, \overleftarrow{Y}_i|W_i, W_j, \overleftarrow{Y}_j) \\
&= h(\overleftarrow{Y}_j|W_j) - h(\overleftarrow{Y}_j|W_i, W_j) + h(\mathbf{X}_{i,C}, \mathbf{X}_{j,U}, \overleftarrow{Y}_i|W_i) \\
&\quad - h(\mathbf{X}_{i,C}, \mathbf{X}_{j,U}, \overleftarrow{Y}_i|W_i, W_j, \overleftarrow{Y}_j) \\
&= h(\overleftarrow{Y}_j|W_j) + h(\mathbf{X}_{i,C}, \mathbf{X}_{j,U}, \overleftarrow{Y}_i|W_i) \\
&\quad - h(\mathbf{X}_{i,C}, \mathbf{X}_{j,U}, \overleftarrow{Y}_i, \overleftarrow{Y}_j|W_i, W_j) \\
&= h(\overleftarrow{Y}_j|W_j) + \sum_{n=1}^N \left[h(X_{i,C,n}, X_{j,U,n}, \overleftarrow{Y}_{i,n}|W_i, \right. \\
&\quad \left. \mathbf{X}_{i,C,(1:n-1)}, \mathbf{X}_{j,U,(1:n-1)}, \overleftarrow{Y}_{i,(1:n-1)}, \mathbf{X}_{i,(1:n)} \right)
\end{aligned}$$

$$\begin{aligned}
&- h(X_{i,C,n}, X_{j,U,n}, \overleftarrow{Y}_{i,n}, \overleftarrow{Y}_{j,n}|W_i, W_j, \mathbf{X}_{i,C,(1:n-1)}, \\
&\quad \mathbf{X}_{j,U,(1:n-1)}, \overleftarrow{Y}_{i,(1:n-1)}, \overleftarrow{Y}_{j,(1:n-1)}, \mathbf{X}_{i,(1:n)}, \mathbf{X}_{j,(1:n)}) \Big] \\
&\leq h(\overleftarrow{Y}_j|W_j) + \sum_{n=1}^N \left[h(X_{i,C,n}, X_{j,U,n}, \overleftarrow{Y}_{i,n}|X_{i,n}) \right. \\
&\quad \left. - h(\vec{Z}_{j,n}, \vec{Z}_{i,n}, \overleftarrow{Y}_{i,n}, \overleftarrow{Y}_{j,n}|W_i, W_j, \mathbf{X}_{i,C,(1:n-1)}, \right. \\
&\quad \left. \mathbf{X}_{j,U,(1:n-1)}, \overleftarrow{Y}_{i,(1:n-1)}, \overleftarrow{Y}_{j,(1:n-1)}, \mathbf{X}_{i,(1:n)}, \mathbf{X}_{j,(1:n)}) \right] \\
&= h(\overleftarrow{Y}_j|W_j) + \sum_{n=1}^N \left[h(X_{i,C,n}|X_{i,n}) \right. \\
&\quad \left. + h(X_{j,U,n}|X_{i,n}, X_{i,C,n}) + h(\overleftarrow{Y}_{i,n}|X_{i,n}, X_{i,C,n}, X_{j,U,n}) \right. \\
&\quad \left. - h(\vec{Z}_{j,n}) - h(\vec{Z}_{i,n}) - h(\overleftarrow{Y}_{i,n}, \overleftarrow{Y}_{j,n}|W_i, W_j, \right. \\
&\quad \left. \mathbf{X}_{i,C,(1:n-1)}, \mathbf{X}_{j,U,(1:n-1)}, \overleftarrow{Y}_{i,(1:n-1)}, \overleftarrow{Y}_{j,(1:n-1)}, \right. \\
&\quad \left. \mathbf{X}_{i,(1:n)}, \mathbf{X}_{j,(1:n)}, \vec{Z}_{j,n}, \vec{Z}_{i,n}) \right] \\
&\leq h(\overleftarrow{Y}_j|W_j) + \sum_{n=1}^N \left[h(\vec{Z}_{j,n}|X_{i,n}) + h(X_{j,U,n}|X_{i,C,n}) \right. \\
&\quad \left. + h(\overleftarrow{Y}_{i,n}|X_{i,n}, X_{j,U,n}) - h(\vec{Z}_{j,n}) - h(\vec{Z}_{i,n}) \right. \\
&\quad \left. - h(\overleftarrow{Z}_{i,n}, \overleftarrow{Z}_{j,n}|W_i, W_j, \mathbf{X}_{i,C,(1:n-1)}, \mathbf{X}_{j,U,(1:n-1)}, \right. \\
&\quad \left. \overleftarrow{Y}_{i,(1:n-1)}, \overleftarrow{Y}_{j,(1:n-1)}, \mathbf{X}_{i,(1:n)}, \mathbf{X}_{j,(1:n)}, \vec{Z}_{j,n}, \vec{Z}_{i,n}) \right] \\
&\stackrel{(a)}{=} h(\overleftarrow{Y}_j|W_j) + \sum_{n=1}^N \left[h(X_{j,U,n}|X_{i,C,n}) \right. \\
&\quad \left. + h(\overleftarrow{Y}_{i,n}|X_{i,n}, X_{j,U,n}) - h(\vec{Z}_{i,n}) \right. \\
&\quad \left. - h(\overleftarrow{Z}_{i,n}) - h(\vec{Z}_{j,n}) \right] \\
&= h(\overleftarrow{Y}_j|W_j) + \sum_{n=1}^N \left[h(X_{j,U,n}|X_{i,C,n}) \right. \\
&\quad \left. + h(\overleftarrow{Y}_{i,n}|X_{i,n}, X_{j,U,n}) - \frac{3}{2} \log(2\pi e) \right],
\end{aligned}$$

where (a) follows from the fact that $\overleftarrow{Z}_{i,n}$ and $\vec{Z}_{j,n}$ are independent of W_i , W_j , $\mathbf{X}_{i,C,(1:n-1)}$, $\mathbf{X}_{j,U,(1:n-1)}$, $\overleftarrow{Y}_{i,(1:n-1)}$, $\overleftarrow{Y}_{j,(1:n-1)}$, $\mathbf{X}_{i,(1:n)}$, $\mathbf{X}_{j,(1:n)}$, $\vec{Z}_{j,n}$, and $\vec{Z}_{i,n}$. This completes the proof of Lemma 1.

APPENDIX D PROOF OF THEOREM 4

This appendix presents a proof of the Theorem 4. The gap, denoted by δ , between the sets $\overline{\mathcal{C}}_{\text{G-IC-NOF}}$ and $\underline{\mathcal{C}}_{\text{G-IC-NOF}}$ (Definition 2) is approximated as follows:

$$\delta = \max\left(\delta_{R_1}, \delta_{R_2}, \frac{\delta_{2R}}{2}, \frac{\delta_{3R_1}}{3}, \frac{\delta_{3R_2}}{3}\right), \tag{79}$$

where,

$$\delta_{R_1} = \min \left(\kappa_{1,1}(\rho'), \kappa_{2,1}(\rho'), \kappa_{3,1}(\rho') \right) - \min \left(a_{2,1}(\rho), a_{6,1}(\rho, \mu_1) + a_{3,2}(\rho, \mu_1), a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{4,2}(\rho, \mu_1) \right), \quad (80a)$$

$$\delta_{R_2} = \min \left(\kappa_{1,2}(\rho'), \kappa_{2,2}(\rho'), \kappa_{3,2}(\rho') \right) - \min \left(a_{2,2}(\rho), a_{3,1}(\rho, \mu_2) + a_{6,2}(\rho, \mu_2), a_{3,1}(\rho, \mu_2) + a_{4,1}(\rho, \mu_2) + a_{1,2} \right), \quad (80b)$$

$$\delta_{2R} = \min \left(\kappa_4(\rho'), \kappa_5(\rho'), \kappa_6(\rho') \right) - \min \left(a_{2,1}(\rho) + a_{1,2}, a_{1,1} + a_{2,2}(\rho), a_{3,1}(\rho, \mu_2) + a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{7,2}(\rho, \mu_1, \mu_2), a_{3,1}(\rho, \mu_2) + a_{5,1}(\rho, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{5,2}(\rho, \mu_1), a_{3,1}(\rho, \mu_2) + a_{7,1}(\rho, \mu_1, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{1,2} \right), \quad (80c)$$

$$\delta_{3R_1} = \kappa_{7,1}(\rho') - \min \left(a_{2,1}(\rho) + a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{7,2}(\rho, \mu_1, \mu_2), a_{3,1}(\rho, \mu_2) + a_{1,1} + a_{7,1}(\rho, \mu_1, \mu_2) + 2a_{3,2}(\rho, \mu_1) + a_{5,2}(\rho, \mu_1), a_{2,1}(\rho) + a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{5,2}(\rho, \mu_1) \right), \quad (80d)$$

$$\delta_{3R_2} = \kappa_{7,2}(\rho') - \min \left(a_{3,1}(\rho, \mu_2) + a_{5,1}(\rho, \mu_2) + a_{2,2}(\rho) + a_{1,2}, a_{1,2}, 2a_{3,1}(\rho, \mu_2) + a_{5,1}(\rho, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{1,2} + a_{7,2}(\rho, \mu_1, \mu_2) \right), \quad (80e)$$

where, $\rho' \in [0, 1]$ and $(\rho, \mu_1, \mu_2) \in [0, (1 - \max(\frac{1}{\text{INR}_{21}}, \frac{1}{\text{INR}_{21}}))^+] \times [0, 1] \times [0, 1]$.

Note that δ_{R_1} and δ_{R_2} represent the gap between the active achievable single-rate bound and the active converse single-rate bound; δ_{2R} represents the gap between the active achievable sum-rate bound and the active converse sum-rate bound; and, δ_{3R_1} and δ_{3R_2} represent the gap between the active achievable weighted sum-rate bound and the active converse weighted sum-rate bound.

It is important to highlight that, as suggested in [3], [6], and [8], the gap between $\underline{C}_{G-IC-NOF}$ and $\bar{C}_{G-IC-NOF}$ can be calculated more precisely. However, the choice in (79) eases the calculations at the expense of less precision. Note also that whether the bounds are active (achievable or converse) in either of the equalities in (80) depend on the exact values of $\overleftarrow{\text{SNR}}_1$, $\overleftarrow{\text{SNR}}_2$, $\overleftarrow{\text{INR}}_{12}$, $\overleftarrow{\text{INR}}_{21}$, $\overleftarrow{\text{SNR}}_1$, and $\overleftarrow{\text{SNR}}_2$. Hence a key point in order to find the gap between the achievable region and the converse region is to choose a convenient coding scheme for the achievable region, i.e., the values of ρ , μ_1 , and μ_2 , according to the definitions in (80) for all $i \in \{1, 2\}$. This particular coding scheme is chosen such that the expressions in (80) become simpler to upper bound at the expense of a looser outer bound. This particular coding scheme is different for each interference regime. The following describes all the key cases and the corresponding coding schemes.

Case 1: $\overleftarrow{\text{INR}}_{12} > \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_2$. This case corresponds to the scenario in which both transmitter-receiver

pairs are in high interference regime (HIR). Three subcases follow considering the SNR in the feedback links.

Case 1.1: $\overleftarrow{\text{SNR}}_2 \leq \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{SNR}}_1 \leq \overleftarrow{\text{SNR}}_2$. In this case the coding scheme is: $\rho = 0$, $\mu_1 = 0$ and $\mu_2 = 0$.

Case 1.2: $\overleftarrow{\text{SNR}}_2 > \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{SNR}}_1 > \overleftarrow{\text{SNR}}_2$. In this case the coding scheme is: $\rho = 0$, $\mu_1 = 1$, and $\mu_2 = 1$.

Case 1.3: $\overleftarrow{\text{SNR}}_2 \leq \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{SNR}}_1 > \overleftarrow{\text{SNR}}_2$. In this case the coding scheme is: $\rho = 0$, $\mu_1 = 0$, and $\mu_2 = 1$.

Case 2: $\overleftarrow{\text{INR}}_{12} \leq \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{INR}}_{21} \leq \overleftarrow{\text{SNR}}_2$. This case corresponds to the scenario in which both transmitter-receiver pairs are in low interference regime (LIR). There are twelve subcases that must be studied separately.

In the following four subcases, the achievability scheme presented above is used considering the following coding scheme: $\rho = 0$, $\mu_1 = 0$, and $\mu_2 = 0$.

Case 2.1: $\overleftarrow{\text{SNR}}_1 \leq \overleftarrow{\text{INR}}_{21}$, $\overleftarrow{\text{SNR}}_2 \leq \overleftarrow{\text{INR}}_{12}$, $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_2$.

Case 2.2: $\overleftarrow{\text{SNR}}_1 \leq \overleftarrow{\text{INR}}_{21}$, $\overleftarrow{\text{SNR}}_2\overleftarrow{\text{INR}}_{21} \leq \overleftarrow{\text{SNR}}_2$, $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} < \overleftarrow{\text{SNR}}_2$.

Case 2.3: $\overleftarrow{\text{SNR}}_1\overleftarrow{\text{INR}}_{12} \leq \overleftarrow{\text{SNR}}_1$, $\overleftarrow{\text{SNR}}_2 \leq \overleftarrow{\text{INR}}_{12}$, $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} < \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_2$.

Case 2.4: $\overleftarrow{\text{SNR}}_1\overleftarrow{\text{INR}}_{12} \leq \overleftarrow{\text{SNR}}_1$, $\overleftarrow{\text{SNR}}_2\overleftarrow{\text{INR}}_{21} \leq \overleftarrow{\text{SNR}}_2$, $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} < \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} < \overleftarrow{\text{SNR}}_2$.

In the following four subcases, the achievability scheme presented above is used considering the following coding scheme: $\rho = 0$, $\mu_1 = \frac{\overleftarrow{\text{INR}}_{21}^2 \overleftarrow{\text{SNR}}_2}{(\overleftarrow{\text{INR}}_{21} - 1)(\overleftarrow{\text{INR}}_{21} \overleftarrow{\text{SNR}}_2 + \overleftarrow{\text{SNR}}_2)}$, and

$$\mu_2 = \frac{\overleftarrow{\text{INR}}_{12}^2 \overleftarrow{\text{SNR}}_1}{(\overleftarrow{\text{INR}}_{12} - 1)(\overleftarrow{\text{INR}}_{12} \overleftarrow{\text{SNR}}_1 + \overleftarrow{\text{SNR}}_1)}.$$

Case 2.5: $\overleftarrow{\text{SNR}}_1 > \overleftarrow{\text{INR}}_{21}$, $\overleftarrow{\text{SNR}}_2 > \overleftarrow{\text{INR}}_{12}$, $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_2$.

Case 2.6: $\overleftarrow{\text{SNR}}_1 > \overleftarrow{\text{INR}}_{21}$, $\overleftarrow{\text{SNR}}_2\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_2$, $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} < \overleftarrow{\text{SNR}}_2$.

Case 2.7: $\overleftarrow{\text{SNR}}_1\overleftarrow{\text{INR}}_{12} > \overleftarrow{\text{SNR}}_1$, $\overleftarrow{\text{SNR}}_2 > \overleftarrow{\text{INR}}_{12}$, $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} < \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_2$.

Case 2.8: $\overleftarrow{\text{SNR}}_1\overleftarrow{\text{INR}}_{12} > \overleftarrow{\text{SNR}}_1$, $\overleftarrow{\text{SNR}}_2\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_2$, $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} < \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} < \overleftarrow{\text{SNR}}_2$.

In the following four subcases, the achievability scheme presented above is used considering the following coding scheme: $\rho = 0$, $\mu_1 = 0$, and $\mu_2 = \frac{\overleftarrow{\text{INR}}_{12}^2 \overleftarrow{\text{SNR}}_1}{(\overleftarrow{\text{INR}}_{12} - 1)(\overleftarrow{\text{INR}}_{12} \overleftarrow{\text{SNR}}_1 + \overleftarrow{\text{SNR}}_1)}$.

Case 2.9: $\overleftarrow{\text{SNR}}_1 > \overleftarrow{\text{INR}}_{21}$, $\overleftarrow{\text{SNR}}_2 \leq \overleftarrow{\text{INR}}_{12}$, $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_2$.

Case 2.10: $\overleftarrow{\text{SNR}}_1 > \overleftarrow{\text{INR}}_{21}$, $\overleftarrow{\text{SNR}}_2\overleftarrow{\text{INR}}_{21} \leq \overleftarrow{\text{SNR}}_2$, $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} < \overleftarrow{\text{SNR}}_2$.

Case 2.11: $\overleftarrow{\text{SNR}}_1\overleftarrow{\text{INR}}_{12} > \overleftarrow{\text{SNR}}_1$, $\overleftarrow{\text{SNR}}_2 \leq \overleftarrow{\text{INR}}_{12}$, $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} < \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_2$.

Case 2.12: $\overleftarrow{\text{SNR}}_1\overleftarrow{\text{INR}}_{12} > \overleftarrow{\text{SNR}}_1$, $\overleftarrow{\text{SNR}}_2\overleftarrow{\text{INR}}_{21} \leq \overleftarrow{\text{SNR}}_2$, $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} < \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} < \overleftarrow{\text{SNR}}_2$.

Case 3: $\overleftarrow{\text{INR}}_{12} > \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{INR}}_{21} \leq \overleftarrow{\text{SNR}}_2$. This case corresponds to the scenario in which transmitter-receiver pair 1 is in HIR and transmitter-receiver pair 2 is in LIR. There are four subcases that must be studied separately.

In the following two subcases, the achievability scheme presented above is used considering the following coding scheme: $\rho = 0$, $\mu_1 = 0$, and $\mu_2 = 0$.

Case 3.1: $\overleftarrow{\text{SNR}}_2 \leq \overleftarrow{\text{INR}}_{12}$ and $\overleftarrow{\text{INR}}_{12}\overleftarrow{\text{INR}}_{21} > \overleftarrow{\text{SNR}}_2$.

Case 3.2: $\overleftarrow{\text{SNR}}_2 \text{INR}_{21} \leq \overrightarrow{\text{SNR}}_2$ and $\text{INR}_{12} \text{INR}_{21} < \overrightarrow{\text{SNR}}_2$.

In the following two subcases, the achievability scheme presented above is used considering the following coding scheme: $\rho = 0$, $\mu_1 = 1$, and $\mu_2 = 0$.

Case 3.3: $\overleftarrow{\text{SNR}}_2 > \text{INR}_{12}$ and $\text{INR}_{12} \text{INR}_{21} > \overrightarrow{\text{SNR}}_2$.

Case 3.4: $\overleftarrow{\text{SNR}}_2 \text{INR}_{21} > \overrightarrow{\text{SNR}}_2$ and $\text{INR}_{12} \text{INR}_{21} < \overrightarrow{\text{SNR}}_2$.

The following is the calculation of the gap δ in Case 1.1.

- 1) Calculation of δ_{R_1} . From (80a) and considering the corresponding coding scheme for the achievable region ($\rho = 0$, $\mu_1 = 0$ and $\mu_2 = 0$), it follows that

$$\delta_{R_1} \leq \min \left(\kappa_{1,1}(\rho'), \kappa_{2,1}(\rho'), \kappa_{3,1}(\rho') \right) - \min \left(a_{6,1}(0,0), a_{1,1} + a_{4,2}(0,0) \right), \quad (81)$$

where the exact value of ρ' is chosen to provide at least an outer bound for (81).

Note that in this case:

$$\begin{aligned} \kappa_{1,1}(\rho') &= \frac{1}{2} \log \left(b_{1,1}(\rho') + 1 \right) \\ &\stackrel{(a)}{\leq} \frac{1}{2} \log \left(\overrightarrow{\text{SNR}}_1 + 2\sqrt{\overrightarrow{\text{SNR}}_1 \text{INR}_{12}} + \text{INR}_{12} + 1 \right) \\ &\stackrel{(b)}{\leq} \frac{1}{2} \log \left(2\overrightarrow{\text{SNR}}_1 + 2\text{INR}_{12} + 1 \right) \\ &\leq \frac{1}{2} \log \left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{12} + 1 \right) + \frac{1}{2}, \end{aligned} \quad (82a)$$

$$\begin{aligned} \kappa_{2,1}(\rho') &= \frac{1}{2} \log \left(1 + b_{4,1}(\rho') + b_{5,2}(\rho') \right) \\ &\leq \frac{1}{2} \log \left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1 \right), \end{aligned} \quad (82b)$$

$$\begin{aligned} \kappa_{3,1}(\rho') &= \frac{1}{2} \log \left(b_{4,1}(\rho') + 1 \right) \\ &\quad + \frac{1}{2} \log \left(\frac{\overleftarrow{\text{SNR}}_2 (b_{4,1}(\rho') + b_{5,2}(\rho') + 1)}{(b_{1,2}(1) + 1)(b_{4,1}(\rho') + 1)} + 1 \right) \\ &\stackrel{(c)}{\leq} \frac{1}{2} \log \left(\overrightarrow{\text{SNR}}_1 + 1 \right) \\ &\quad + \frac{1}{2} \log \left(\frac{\overleftarrow{\text{SNR}}_2 (\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1)}{(\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1)(\overrightarrow{\text{SNR}}_1 + 1)} + 1 \right) \\ &= \frac{1}{2} \log \left(\frac{\overleftarrow{\text{SNR}}_2 (\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1)}{\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1} + \overrightarrow{\text{SNR}}_1 + 1 \right), \end{aligned} \quad (82c)$$

where (a) follows from the fact that $0 \leq \rho' \leq 1$; (b) follows from the fact that

$$\left(\sqrt{\overrightarrow{\text{SNR}}_1} - \sqrt{\text{INR}_{12}} \right)^2 \geq 0; \quad (83)$$

and (c) follows from the fact that $\kappa_{3,1}(\rho')$ is a monotonically decreasing function of ρ' .

Note also that the achievable bound $a_{1,1} + a_{4,2}(0,0)$ can

be lower bounded as follows:

$$\begin{aligned} a_{1,1} + a_{4,2}(0,0) &= \frac{1}{2} \log \left(\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} + 2 \right) + \frac{1}{2} \log \left(\text{INR}_{21} + 1 \right) \\ &\quad - 1 \\ &\geq \frac{1}{2} \log \left(\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} + 2 \right) + \frac{1}{2} \log \left(\text{INR}_{21} \right) - 1 \\ &= \frac{1}{2} \log \left(\overrightarrow{\text{SNR}}_1 + 2\text{INR}_{21} \right) - 1 \\ &= \frac{1}{2} \log \left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + \text{INR}_{21} \right) - 1 \\ &\geq \frac{1}{2} \log \left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1 \right) - 1. \end{aligned} \quad (84)$$

From (81), (82) and (84), assuming that $a_{1,1} + a_{4,2}(0,0) < a_{6,1}(0,0)$, it follows that

$$\begin{aligned} \delta_{R_1} &\leq \min \left(\kappa_{1,1}(\rho'), \kappa_{2,1}(\rho'), \kappa_{3,1}(\rho) \right) - \left(a_{1,1} + a_{4,2}(0,0) \right) \\ &\leq \kappa_{2,1}(\rho') - \left(a_{1,1} + a_{4,2}(0,0) \right) \\ &\leq 1. \end{aligned} \quad (85)$$

Now, assuming that $a_{6,1}(0,0) < a_{1,1} + a_{4,2}(0,0)$, the following holds:

$$\delta_{R_1} \leq \min \left(\kappa_{1,1}(\rho'), \kappa_{2,1}(\rho'), \kappa_{3,1}(\rho) \right) - a_{6,1}(0,0). \quad (86)$$

To calculate an upper bound for (86), the following cases are considered:

Case 1.1.1: $\overrightarrow{\text{SNR}}_1 \geq \text{INR}_{21} \wedge \overrightarrow{\text{SNR}}_2 < \text{INR}_{12}$;

Case 1.1.2: $\overrightarrow{\text{SNR}}_1 < \text{INR}_{21} \wedge \overrightarrow{\text{SNR}}_2 \geq \text{INR}_{12}$; and

Case 1.1.3: $\overrightarrow{\text{SNR}}_1 < \text{INR}_{21} \wedge \overrightarrow{\text{SNR}}_2 < \text{INR}_{12}$.

In Case 1.1.1, from (82) and (86), it follows that

$$\begin{aligned} \delta_{R_1} &\leq \kappa_{2,1}(\rho') - a_{6,1}(0,0) \\ &\leq \frac{1}{2} \log \left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1 \right) - \frac{1}{2} \log \left(\overrightarrow{\text{SNR}}_1 + 2 \right) \\ &\quad + \frac{1}{2} \\ &\leq \frac{1}{2} \log \left(\overrightarrow{\text{SNR}}_1 + \overrightarrow{\text{SNR}}_1 + 1 \right) - \frac{1}{2} \log \left(\overrightarrow{\text{SNR}}_1 + 2 \right) \\ &\quad + \frac{1}{2} \\ &\leq 1. \end{aligned} \quad (87)$$

In Case 1.1.2, from (82) and (86), it follows that

$$\begin{aligned} \delta_{R_1} &\leq \kappa_{3,1}(\rho') - a_{6,1}(0,0) \\ &\leq \frac{1}{2} \log \left(\frac{\overleftarrow{\text{SNR}}_2 (\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1)}{\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1} + \overrightarrow{\text{SNR}}_1 + 1 \right) \\ &\quad - \frac{1}{2} \log \left(\overrightarrow{\text{SNR}}_1 + 2 \right) + \frac{1}{2} \\ &\leq \frac{1}{2} \log \left(\overleftarrow{\text{SNR}}_2 + \overrightarrow{\text{SNR}}_1 + 1 \right) - \frac{1}{2} \log \left(\overrightarrow{\text{SNR}}_1 + 2 \right) \\ &\quad + \frac{1}{2} \\ &\leq \frac{1}{2} \log \left(\overrightarrow{\text{SNR}}_1 + \overrightarrow{\text{SNR}}_1 + 1 \right) - \frac{1}{2} \log \left(\overrightarrow{\text{SNR}}_1 + 2 \right) \\ &\quad + \frac{1}{2} \\ &\leq 1. \end{aligned} \quad (88)$$

In Case 1.1.3 two additional cases are considered:

Case 1.1.3.1: $\overrightarrow{\text{SNR}}_1 \geq \overrightarrow{\text{SNR}}_2$; and

Case 1.1.3.2: $\overrightarrow{\text{SNR}}_1 < \overrightarrow{\text{SNR}}_2$.

In Case 1.1.3.1, from (82) and (86), it follows that

$$\begin{aligned}
\delta_{R_1} &\leq \kappa_{3,1}(\rho') - a_{6,1}(0,0) \\
&\leq \frac{1}{2} \log \left(\frac{\overrightarrow{\text{SNR}}_2 (\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1)}{\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1} + \overrightarrow{\text{SNR}}_1 + 1 \right) \\
&\quad - \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + 2) + \frac{1}{2} \\
&= \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + 1) \\
&\quad + \frac{1}{2} \log \left(\frac{\overrightarrow{\text{SNR}}_2 (\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1)}{(\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1) (\overrightarrow{\text{SNR}}_1 + 1)} + 1 \right) \\
&\quad - \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + 2) + \frac{1}{2} \\
&\leq \frac{1}{2} \log \left(\frac{\overrightarrow{\text{SNR}}_1 (\text{INR}_{21} + \text{INR}_{21} + \text{INR}_{21})}{\text{INR}_{21} \overrightarrow{\text{SNR}}_1} + 1 \right) \\
&\quad + \frac{1}{2} \\
&= \frac{3}{2}. \tag{89}
\end{aligned}$$

In Case 1.1.3.2, from (82) and (86), it follows that

$$\begin{aligned}
\delta_{R_1} &\leq \kappa_{3,1}(\rho') - a_{6,1}(0,0) \\
&\leq \frac{1}{2} \log \left(\frac{\overrightarrow{\text{SNR}}_2 (\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1)}{\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1} + \overrightarrow{\text{SNR}}_1 + 1 \right) \\
&\quad - \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + 2) + \frac{1}{2} \\
&\leq \frac{1}{2} \log (\overrightarrow{\text{SNR}}_2 + \overrightarrow{\text{SNR}}_1 + 1) - \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + 2) \\
&\quad + \frac{1}{2} \\
&\leq \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + \overrightarrow{\text{SNR}}_1 + 1) - \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + 2) \\
&\quad + \frac{1}{2} \\
&\leq 1. \tag{90}
\end{aligned}$$

Then, from (85), (87), (88), (89), and (90), it follows that in Case 1.1:

$$\delta_{R_1} \leq \frac{3}{2}. \tag{91}$$

The same procedure holds to calculate δ_{R_2} and it yields:

$$\delta_{R_2} \leq \frac{3}{2}. \tag{92}$$

2) Calculation of δ_{2R} . From (80c) and considering the corresponding coding scheme for the achievable region

($\rho = 0$, $\mu_1 = 0$ and $\mu_2 = 0$), it follows that

$$\begin{aligned}
\delta_{2R} &\leq \min \left(\kappa_4(\rho'), \kappa_5(\rho'), \kappa_6(\rho') \right) - \min \left(a_{2,1}(0) + a_{1,2}, \right. \\
&\quad \left. a_{1,1} + a_{2,2}(0), a_{5,1}(0,0) + a_{5,2}(0,0) \right) \\
&\leq \min \left(\kappa_4(\rho'), \kappa_5(\rho') \right) - \min \left(a_{2,1}(0) + a_{1,2}, \right. \\
&\quad \left. a_{1,1} + a_{2,2}(0), a_{5,1}(0,0) + a_{5,2}(0,0) \right). \tag{93}
\end{aligned}$$

Note that

$$\begin{aligned}
\kappa_4(\rho') &= \frac{1}{2} \log \left(1 + \frac{b_{4,1}(\rho')}{1 + b_{5,2}(\rho')} \right) + \frac{1}{2} \log (b_{1,2}(\rho') + 1) \\
&\leq \frac{1}{2} \log \left(1 + \frac{b_{4,1}(\rho')}{b_{5,2}(\rho')} \right) + \frac{1}{2} \log (b_{1,2}(\rho') + 1) \\
&= \frac{1}{2} \log \left(1 + \frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} \right) + \frac{1}{2} \log (b_{1,2}(\rho') + 1) \\
&\stackrel{(h)}{\leq} \frac{1}{2} \log \left(1 + \frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} \right) \\
&\quad + \frac{1}{2} \log (2\overrightarrow{\text{SNR}}_2 + 2\text{INR}_{21} + 1), \\
&\leq \frac{1}{2} \log \left(1 + \frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} \right) \\
&\quad + \frac{1}{2} \log (\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1) + \frac{1}{2} \\
&\leq \frac{1}{2} \log \left(2 + \frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} \right) \\
&\quad + \frac{1}{2} \log (\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1) + \frac{1}{2}, \tag{94a}
\end{aligned}$$

and

$$\begin{aligned}
\kappa_5(\rho') &= \frac{1}{2} \log \left(1 + \frac{b_{4,2}(\rho')}{1 + b_{5,1}(\rho')} \right) + \frac{1}{2} \log (b_{1,1}(\rho') + 1) \\
&\leq \frac{1}{2} \log \left(1 + \frac{b_{4,2}(\rho')}{b_{5,1}(\rho')} \right) + \frac{1}{2} \log (b_{1,1}(\rho') + 1) \\
&= \frac{1}{2} \log \left(1 + \frac{\overrightarrow{\text{SNR}}_2}{\text{INR}_{12}} \right) + \frac{1}{2} \log (b_{1,1}(\rho') + 1) \\
&\stackrel{(i)}{\leq} \frac{1}{2} \log \left(1 + \frac{\overrightarrow{\text{SNR}}_2}{\text{INR}_{12}} \right) \\
&\quad + \frac{1}{2} \log (2\overrightarrow{\text{SNR}}_1 + 2\text{INR}_{12} + 1) \\
&\leq \frac{1}{2} \log \left(1 + \frac{\overrightarrow{\text{SNR}}_2}{\text{INR}_{12}} \right) \\
&\quad + \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + \text{INR}_{12} + 1) + \frac{1}{2} \\
&\leq \frac{1}{2} \log \left(2 + \frac{\overrightarrow{\text{SNR}}_2}{\text{INR}_{12}} \right) \\
&\quad + \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + \text{INR}_{12} + 1) + \frac{1}{2}, \tag{94b}
\end{aligned}$$

where (h) follows from the fact that

$$\left(\sqrt{\overrightarrow{\text{SNR}}_2} - \sqrt{\text{INR}_{21}} \right)^2 \geq 0; \tag{95}$$

and (i) follows from the fact that

$$\left(\sqrt{\overrightarrow{\text{SNR}}_1} - \sqrt{\text{INR}_{12}}\right)^2 \geq 0. \quad (96)$$

From (93) and (94), assuming that $a_{2,1}(0) + a_{1,2} < \min\left(a_{1,1} + a_{2,2}(0), a_{5,1}(0,0) + a_{5,2}(0,0)\right)$, it follows that

$$\begin{aligned} \delta_{2R} &\leq \min\left(\kappa_4(\rho'), \kappa_5(\rho')\right) - \left(a_{2,1}(0) + a_{1,2}\right) \\ &\leq \kappa_5(\rho') - \left(a_{2,1}(0) + a_{1,2}\right) \\ &\leq \frac{1}{2} \log\left(2 + \frac{\overrightarrow{\text{SNR}}_2}{\text{INR}_{12}}\right) \\ &\quad + \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{12} + 1\right) + \frac{1}{2} \\ &\quad - \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{12} + 1\right) \\ &\quad - \frac{1}{2} \log\left(\frac{\overrightarrow{\text{SNR}}_2}{\text{INR}_{12}} + 2\right) + 1 \\ &= \frac{3}{2}. \end{aligned} \quad (97)$$

From (93) and (94), assuming that $a_{1,1} + a_{2,2}(0) < \min\left(a_{2,1}(0) + a_{1,2}, a_{5,1}(0,0) + a_{5,2}(0,0)\right)$, it follows that

$$\begin{aligned} \delta_{2R} &\leq \min\left(\kappa_4(\rho'), \kappa_5(\rho')\right) - \left(a_{1,1} + a_{2,2}(0)\right) \\ &\leq \kappa_4(\rho') - \left(a_{1,1} + a_{2,2}(0)\right) \\ &\leq \frac{1}{2} \log\left(2 + \frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}}\right) \\ &\quad + \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1\right) + \frac{1}{2} \\ &\quad - \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1\right) \\ &\quad - \frac{1}{2} \log\left(\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} + 2\right) + 1 \\ &= \frac{3}{2}. \end{aligned} \quad (98)$$

Now, assume that $a_{5,1}(0,0) + a_{5,2}(0,0) < \min(a_{2,1}(0) + a_{1,2}, a_{1,1} + a_{2,2}(0))$. In this case, the following holds:

$$\delta_{2R} \leq \min\left(\kappa_4(\rho'), \kappa_5(\rho')\right) - \left(a_{5,1}(0,0) + a_{5,2}(0,0)\right). \quad (99)$$

To calculate an upper bound for (99), the cases 1.1.1 - 1.1.3 defined above are analyzed hereunder.

In Case 1.1.1, $a_{5,1}(0,0) + a_{5,2}(0,0)$ can be lower bounded as follows:

$$\begin{aligned} a_{5,1}(0,0) + a_{5,2}(0,0) &= \frac{1}{2} \log\left(\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} + \text{INR}_{12} + 1\right) \\ &\quad + \frac{1}{2} \log\left(\frac{\overrightarrow{\text{SNR}}_2}{\text{INR}_{12}} + \text{INR}_{21} + 1\right) - 1 \\ &\geq \frac{1}{2} \log(\text{INR}_{12} + 1) - 1. \end{aligned} \quad (100)$$

From (94), (99), and (100), it follows that

$$\begin{aligned} \delta_{2R} &\leq \min\left(\kappa_4(\rho'), \kappa_5(\rho')\right) - \left(a_{5,1}(0,0) + a_{5,2}(0,0)\right) \\ &\leq \kappa_5(\rho') - \left(a_{5,1}(0,0) + a_{5,2}(0,0)\right) \\ &\leq \frac{1}{2} \log\left(2 + \frac{\overrightarrow{\text{SNR}}_2}{\text{INR}_{12}}\right) + \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{12} + 1\right) \\ &\quad + \frac{1}{2} - \frac{1}{2} \log(\text{INR}_{12} + 1) + 1 \\ &\leq \frac{1}{2} \log(2 + 1) + \frac{1}{2} \log(\text{INR}_{12} + \text{INR}_{12} + 1) \\ &\quad - \frac{1}{2} \log(\text{INR}_{12} + 1) + \frac{3}{2} \\ &\leq \frac{1}{2} \log(3) + 2. \end{aligned} \quad (101)$$

In Case 1.1.2, $a_{5,1}(0,0) + a_{5,2}(0,0)$ can be lower bounded as follows:

$$\begin{aligned} a_{5,1}(0,0) + a_{5,2}(0,0) &= \frac{1}{2} \log\left(\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} + \text{INR}_{12} + 1\right) \\ &\quad + \frac{1}{2} \log\left(\frac{\overrightarrow{\text{SNR}}_2}{\text{INR}_{12}} + \text{INR}_{21} + 1\right) - 1 \\ &\geq \frac{1}{2} \log(\text{INR}_{21} + 1) - 1. \end{aligned} \quad (102)$$

From (94), (99), and (102), it follows that

$$\begin{aligned} \delta_{2R} &\leq \min\left(\kappa_4(\rho'), \kappa_5(\rho')\right) - \left(a_{5,1}(0,0) + a_{5,2}(0,0)\right) \\ &\leq \kappa_4(\rho') - \left(a_{5,1}(0,0) + a_{5,2}(0,0)\right) \\ &\leq \frac{1}{2} \log\left(2 + \frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}}\right) + \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1\right) \\ &\quad + \frac{1}{2} - \frac{1}{2} \log(\text{INR}_{21} + 1) + 1 \\ &\leq \frac{1}{2} \log(2 + 1) + \frac{1}{2} \log(\text{INR}_{21} + \text{INR}_{21} + 1) \\ &\quad - \frac{1}{2} \log(\text{INR}_{21} + 1) + \frac{3}{2} \\ &\leq \frac{1}{2} \log(3) + 2. \end{aligned} \quad (103)$$

In Case 1.1.3, from (94), (99), and (100), it follows that

$$\begin{aligned} \delta_{2R} &\leq \min\left(\kappa_4(\rho'), \kappa_5(\rho')\right) - \left(a_{5,1}(0,0) + a_{5,2}(0,0)\right) \\ &\leq \kappa_5(\rho') - \left(a_{5,1}(0,0) + a_{5,2}(0,0)\right) \\ &\leq \frac{1}{2} \log\left(2 + \frac{\overrightarrow{\text{SNR}}_2}{\text{INR}_{12}}\right) + \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{12} + 1\right) \\ &\quad + \frac{1}{2} - \frac{1}{2} \log(\text{INR}_{12} + 1) + 1 \\ &\leq \frac{1}{2} \log(2 + 1) + \frac{1}{2} \log(\text{INR}_{12} + \text{INR}_{12} + 1) \\ &\quad - \frac{1}{2} \log(\text{INR}_{12} + 1) + \frac{3}{2} \\ &\leq \frac{1}{2} \log(3) + 2. \end{aligned} \quad (104)$$

Then, from (97), (98), (101), (103), and (104), it follows that in Case 1.1:

$$\delta_{2R} \leq 2 + \frac{1}{2} \log(3). \quad (105)$$

3) Calculation of δ_{3R_1} . From (80d) and considering the corresponding coding scheme for the achievable region ($\rho = 0$, $\mu_1 = 0$ and $\mu_2 = 0$), it follows that

$$\delta_{3R_1} \leq \kappa_{7,1}(\rho') - (a_{1,1} + a_{7,1}(0,0,0) + a_{5,2}(0,0)). \quad (106)$$

The sum $a_{1,1} + a_{7,1}(0,0,0) + a_{5,2}(0,0)$ can be lower bounded as follows:

$$\begin{aligned} a_{1,1} + a_{7,1}(0,0,0) + a_{5,2}(0,0) &= \frac{1}{2} \log \left(\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} + 2 \right) \\ &+ \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + \text{INR}_{12} + 1) \\ &+ \frac{1}{2} \log \left(\frac{\overrightarrow{\text{SNR}}_2}{\text{INR}_{12}} + \text{INR}_{21} + 1 \right) - \frac{3}{2} \\ &\geq \frac{1}{2} \log \left(\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} + 2 \right) + \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + \text{INR}_{12} + 1) \\ &+ \frac{1}{2} \log (\text{INR}_{21} + 1) - \frac{3}{2}. \end{aligned} \quad (107)$$

If the term $\kappa_{7,1}(\rho')$ is active in the converse region, this can be upper bounded by the sum $\kappa_{1,1}(\rho') + \kappa_4(\rho')$, which corresponds to the sum of the single rate and sum-rate outer bounds respectively, and this can be upper bounded as follows:

$$\begin{aligned} \kappa_{7,1}(\rho') &\leq \kappa_{1,1}(\rho') + \kappa_4(\rho') \\ &\leq \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + \text{INR}_{12} + 1) + \frac{1}{2} \log \left(2 + \frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} \right) \\ &+ \frac{1}{2} \log (\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1) + 1 \\ &\leq \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + \text{INR}_{12} + 1) + \frac{1}{2} \log \left(2 + \frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} \right) \\ &+ \frac{1}{2} \log (\text{INR}_{21} + \text{INR}_{21} + 1) + 1 \\ &\leq \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + \text{INR}_{12} + 1) + \frac{1}{2} \log \left(2 + \frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} \right) \\ &+ \frac{1}{2} \log (\text{INR}_{21} + 1) + \frac{3}{2}. \end{aligned} \quad (108)$$

From (106), (107) and (108), it follows that in Case 1.1:

$$\begin{aligned} \delta_{3R_1} &\leq \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + \text{INR}_{12} + 1) + \frac{1}{2} \log \left(2 + \frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} \right) \\ &+ \frac{1}{2} \log (\text{INR}_{21} + 1) + \frac{3}{2} - \frac{1}{2} \log \left(\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} + 2 \right) \\ &- \frac{1}{2} \log (\overrightarrow{\text{SNR}}_1 + \text{INR}_{12} + 1) - \frac{1}{2} \log (\text{INR}_{21} + 1) \\ &+ \frac{3}{2} \\ &= 3. \end{aligned} \quad (109)$$

The same procedure holds in the calculation of δ_{3R_2} and it yields:

$$\delta_{3R_2} \leq 3. \quad (110)$$

Therefore, in Case 1.1, from (79), (92), (91), (105), (109) and (110) it follows that

$$\delta = \max \left(\delta_{R_1}, \delta_{R_2}, \frac{\delta_{2R}}{2}, \frac{\delta_{3R_1}}{3}, \frac{\delta_{3R_2}}{3} \right) \leq \frac{3}{2}. \quad (111)$$

This completes the calculation of the gap in Case 1.1. Applying the same procedure to all the other cases listed above yields that $\delta \leq 4.4$ bits.

REFERENCES

- [1] V. Quintero, S. M. Perlaza, and J.-M. Gorce, "Noisy channel-output feedback capacity of the linear deterministic interference channel," in *IEEE Information Theory Workshop*, Jeju Island, Korea, Oct. 2015.
- [2] V. Quintero, S. M. Perlaza, I. Esnaola, and J.-M. Gorce, "Approximate capacity of the Gaussian interference channel with noisy channel-output feedback," in *IEEE Information Theory Workshop (ITW)*, Sep. 2016.
- [3] C. Suh and D. N. C. Tse, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2667–2685, May. 2011.
- [4] S. A. Jafar, "Interference alignment: A new look at signal dimensions in a communication network," *Foundations and Trends in Communications and Information Theory*, vol. 7, no. 1, pp. 1–134, 2010.
- [5] S. M. Perlaza, R. Tandon, H. V. Poor, and Z. Han, "Perfect output feedback in the two-user decentralized interference channel," *IEEE Trans. Inf. Theory*, vol. 61, no. 10, pp. 5441–5462, Oct. 2015.
- [6] S.-Q. Le, R. Tandon, M. Motani, and H. V. Poor, "Approximate capacity region for the symmetric Gaussian interference channel with noisy feedback," *IEEE Trans. Inf. Theory*, vol. 61, no. 7, pp. 3737–3762, Jul. 2015.
- [7] S. M. Perlaza, R. Tandon, and H. V. Poor, "Symmetric decentralized interference channels with noisy feedback," in *Proc. IEEE Intl. Symposium on Information Theory (ISIT)*, Honolulu, HI, USA, Jun. 2014.
- [8] R. H. Etkin, D. N. C. Tse, and W. Hua, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.
- [9] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, 1981.
- [10] A. Carleial, "A case where interference does not reduce capacity," *IEEE Trans. Inf. Theory*, vol. 21, no. 5, pp. 569–570, Sep. 1975.
- [11] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 3, pp. 581–586, Mar. 2004.
- [12] I. Sason, "On achievable rate regions for the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1345–1356, Jun. 2004.
- [13] A. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 1, pp. 60–70, Sep. 1978.
- [14] T. M. Cover and C. S. K. Leung, "An achievable rate region for the multiple-access channel with feedback," *IEEE Trans. Inf. Theory*, vol. 27, no. 3, pp. 292–298, May. 1981.
- [15] F. M. J. Willems, "Information theoretical results for multiple access channels," Ph.D. dissertation, Katholieke Universiteit, Leuven, Belgium, Oct. 1982.
- [16] F. M. J. Willems and E. C. Van Der Meulen, "The discrete memoryless multiple-access channel with cribbing encoders," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 3, pp. 313–327, May. 1985.
- [17] D. Tuninetti, "On interference channel with generalized feedback (IFC-GF)," in *Proc. of International Symposium on Information Theory (ISIT)*, Nice, France, Jun. 2007, pp. 2661–2665.
- [18] S. Yang and D. Tuninetti, "Interference channel with generalized feedback (a.k.a. with source cooperation): Part I: Achievable region," *IEEE Trans. Inf. Theory*, vol. 5, no. 57, pp. 2686–2710, May. 2011.
- [19] V. M. Pabbakaran and P. Viswanath, "Interference channel with source cooperation," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 156–186, Jan. 2011.

- [20] G. Kramer, "Feedback strategies for white Gaussian interference networks," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1423–1438, Jun. 2002.
- [21] D. Tuninetti, "An outer bound region for interference channels with generalized feedback," in *IEEE Information Theory and Applications Workshop (ITA)*, Feb. 2010, pp. 1–5.
- [22] —, "An outer bound for the memoryless two-user interference channel with general cooperation," in *IEEE Information Theory Workshop (ITW)*, Sep. 2012, pp. 217–221.
- [23] 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 Trans. Inf. Theory*, vol. 59, no. 11, pp. 6957–6979, 2013.
- [24] M. Gastpar and G. Kramer, "On noisy feedback for interference channels," in *Asilomar Conference on Signals, Systems and Computers*, Oct–Nov. 2006.
- [25] A. P. Hekstra and F. M. J. Willems, "Dependence balance bounds for single-output two-way channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 44–53, Jan. 1989.
- [26] R. Tandon and S. Ulukus, "Dependence balance based outer bounds for Gaussian networks with cooperation and feedback," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4063–4086, Jul. 2011.
- [27] S. Avestimehr, S. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.
- [28] G. Bresler and D. N. C. Tse, "The two user Gaussian interference channel: A deterministic view," *European Transactions on Telecommunications*, vol. 19, no. 4, pp. 333–354, Apr. 2008.
- [29] V. Quintero, S. M. Perlaza, I. Esnaola, and J.-M. Gorce, "When does channel-output feedback enlarge the capacity region of the two-user linear deterministic interference channel?" INRIA Grenoble - Rhône-Alpes, Lyon, France, Tech. Rep. 8862, Mar. 2016.
- [30] A. Host-Madsen, "Capacity bounds for cooperative diversity," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1522–1544, Apr. 2006.
- [31] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, CA, USA, 1961, pp. 611–644.
- [32] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley-Interscience, 1991.
- [33] C. E. Shannon, "The zero-error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 8–19, Sep. 1956.