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

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

Abstract—**In this paper, the capacity region of the linear deterministic interference channel with noisy channel-output feedback (LD-IC-NF) is fully characterized. The proof of achievability is based on random coding arguments and rate splitting; block-Markov superposition coding; and backward decoding. The proof of converse reuses some of the existing outer bounds and includes new ones obtained using genie-aided models. Following the insight gained from the analysis of the LD-IC-NF, an achievability region and a converse region for the two-user Gaussian interference channel with noisy channel-output feedback (G-IC-NF) are presented. Finally, the achievability region and the converse region are proven to approximate the capacity region of the G-IC-NF to within 4.4 bits.**

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

I. Introduction

Recently, perfect feedback (PF) from the receivers to the corresponding transmitters has been shown to bring an unprecedented gain on the number of generalized degrees of freedom (GDoF) with respect to the case without feedback in the Gaussian interference channel (IC) [3]. Let $\mathcal{C}(\overrightarrow{\text{SNR}}, \text{INR})$ denote a set containing all achievable rates of a symmetric Gaussian IC (G-IC) with parameters $\overrightarrow{\text{SNR}}$ (signal to noise ratio in the forward link) and INR (interference to noise ratio). The number of GDoF [4] is:

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

Victor Quintero is with the Department of Telecommunications, Universidad del Cauca, 19003, Popayán, Cauca, Colombia.

Samir M. Perlaza and Jean-Marie Gorce are with the CITI Laboratory, a joint laboratory between the Institut National de Recherche en Informatique et en Automatique (INRIA), the Université de Lyon and the Institut National de Sciences Apliquées (INSA) de Lyon. 6 Av. des Arts 69621 Villeurbanne, France. ({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, UK. (esnaola@sheffield.ac.uk).

Samir M. Perlaza and Iñaki Esnaola are also with the Department of Electrical Engineering at Princeton University, Princeton, NJ, 08544, USA.

where $\alpha = \frac{\log(\text{INR})}{\log(\overrightarrow{\text{SNR}})}$. In Figure 1, the number of GDoF is plotted as a function of α when $\mathcal{C}(\overrightarrow{\text{SNR}}, \text{INR})$ is calculated without feedback [5]; and with PF from each receiver to their corresponding transmitters [3]. Note that with PF, $\text{GDoF}(\alpha) \to \infty$ when $\alpha \to \infty$, which implies an arbitrarily large increment. Surprisingly, using only one PF link from one of the receivers to the corresponding transmitter provides the same sum-capacity as having four PF links from both receivers to both transmitters [6]–[8] in certain interference regimes. These benefits rely on the fact that feedback provides relevant information about the interference. Hence, such information can be retransmitted to: (a) perform interference cancellation at the intended receiver or (b) provide an alternative communication path between the other transmitter-receiver pair. These promising results are also observed when the system is decentralized, i.e., when each transmitter seeks to unilaterally maximize its own individual information rate [9], [10].

The capacity region of the G-IC with PF has been approximated to within two bits in [3]. The achievability scheme presented therein is based on three well-known techniques: rate splitting [11], [12], block-Markov superposition coding [13], and backward decoding [14], [15]. The converse in [3] is obtained following classical tools among cut-set bounds and genie-aided models. Other achievability schemes have been presented in [16] and [17] using rate-splitting, block-Markov superposition coding, backward decoding, and binning/dirty paper coding in the context of a more general channel, i.e., G-IC with generalized feedback (IC-GF).

From a system analysis perspective, PF might be an exceptionally optimistic model to study the benefits of feedback in the G-IC. Denote by $\overrightarrow{\boldsymbol{y}} = (\overrightarrow{y}_1, \overrightarrow{y}_2, \ldots, \overrightarrow{y}_N)$ a given sequence of N channel outputs at a given receiver. A more realistic model of channel-output feedback is to consider that the feedback signal, denoted by $\overleftarrow{\boldsymbol{Y}}$, satisfies $\overleftarrow{\boldsymbol{Y}} = g(\overrightarrow{\boldsymbol{y}})$ (random transformation in $\mathbb{R}^N$). Hence, a relevant question is: what is a realistic assumption on g? This question has been solved aiming to highlight different impairments that feedback signals might go through. Some of these answers are discussed in the following sections.

A. Rate-Limited Feedback:

Consider that the receiver produces the feedback signal using a deterministic transformation g, such that for a large N, a positive finite $C_F \in \mathbb{R}$ and for all $\overrightarrow{\boldsymbol{y}} \in \mathbb{R}^N$:

$$\overleftarrow{\boldsymbol{y}} = g(\overrightarrow{\boldsymbol{y}}) \in \mathcal{D} \subseteq \mathbb{R}^N, \quad (2)$$

such that for all $\delta > 0$,

$$|\mathcal{D}| < 2^{N(C_F+\delta)}. \tag{3}$$

This model is known in literature as *rate limited feedback* (RLF) [18]–[20], where C_F is the capacity of the feedback link. The choice of the deterministic transformation g subject to (3) is part of the coding scheme, i.e., the transformation g takes the N channel outputs observed during block $t > 0$ and chooses a codeword in the codebook $\mathcal{D}$. Such a codeword is sent back to the transmitter during block $t + 1$. From this standpoint, this model highlights the signal impairments derived from transmitting a signal with continuous support via a channel with finite-capacity. Note that if $C_F = \infty$, then g can be the identity function and thus,

$$\overleftarrow{\boldsymbol{y}} = g(\overrightarrow{\boldsymbol{y}}) = \overrightarrow{\boldsymbol{y}}, \tag{4}$$

which is the case of PF [3]. When $C_F = 0$, then $|\mathcal{D}| = 1$ and thus, no information can be conveyed through the feedback links, which is the case studied in [5], [12], [21]. The main result in [18] is twofold: first, given a fixed C_F, the authors provide a deterministic transformation g using lattice coding [22] and a particular power assignment such that partial or complete decoding of the interference is possible at the transmitter. An achievable region is presented using random coding arguments with rate splitting, block-Markov superposition coding, and backward decoding. Second, the authors provide outer bounds that hold for any g in (2). This result induces a converse region whose sum-rate is shown to be at a constant gap of the achievable sum-rate, at least in the symmetric case. These results are generalized for the K-user G-IC with RLF in the symmetric case in [19], [20], where the analysis focuses on the fundamental limit of the symmetric rate. The main novelty on the extension to $K > 2$ users lies in the joint use of interference alignment and lattice codes for the proof of the achievability. The proof of converse remains an open problem when $K > 2$, even for the symmetric case.

B. Intermittent Feedback

Assume that for all $n \in \{1, 2 \ldots, N\}$, the random transformation g is such that given a channel output $\overrightarrow{y}_n$,

$$\overleftarrow{Y}_n = \begin{cases} \star & \text{with probability} \quad 1-p \\ \overrightarrow{y}_n & \text{with probability} \quad p, \end{cases} \tag{5}$$

where $\star$ represents an erasure and $p \in [0, 1]$. Note that the random transformation g is fully determined by the parameters of the channels, e.g., the probability p. Thus, as opposed to the RLF, the transformation g cannot be optimized as part of the receiver design. This model emphasizes the fact that the usage of the feedback link might be available only during certain channel uses, not necessarily known by the receivers with anticipation. This model is referred to as *intermittent feedback* (IF) [23]. The main result in [23] is an approximation of the capacity region to within a constant gap. The achievability scheme relies upon random coding arguments with forward decoding and a quantize-map-and-foward strategy to retransmit the information obtained through feedback. This is because erasures might constrain either partial or complete decoding of the interference at the transmitter. Nonetheless, even a quantized version of the interference might be useful for interference cancellation or for providing an alternative path.

C. Noisy Feedback

Assume that for all $n \in \{1, 2 \ldots, N\}$, the random transformation g is such that given a channel output $\overrightarrow{y}_n$,

$$\overleftarrow{Y}_n = \overleftarrow{h}\, \overrightarrow{y}_n + Z_n, \tag{6}$$

where $\overrightarrow{h} \in \mathbb{R}_+$ is a parameter of the channel and Z_n is a real Gaussian random variable with zero mean and unit variance. This model is known in literature as *noisy feedback* (NF) or *partial feedback* [24]–[26]. Note that the receiver does not apply any processing to the channel output and sends a re-scaled copy to the transmitter via a noisy channel. From this point of view, as opposed to RLF, this model does not focus on the constraint on the number of codewords that can be used to perform feedback, but rather on the fact that the feedback channel might be noisy. Essentially, the codebook used to perform feedback in NF is $\mathbb{R}^N$. In [25], the capacity of the G-IC with NF has been approximated to within a constant gap for the symmetric case. The achievable scheme in [25] is a particular case of a more general achievability scheme presented in [16], [17]. An outer bound using the Hekstra-Willems dependence-balance arguments [27] has been introduced in [24]. 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. Similar results have been reported in the fully decentralized IC with NF [10], [28]–[30]. Inner and outer bounds on the sum-capacity of the G-IC with PF using the existing connections between channel-output feedback and conferencing transmitters have been presented in [31]. More general channel models, for instance when channel-outputs are fed back to both receivers, have been studied in [7], [32]–[34].

D. A Comparison Between Feedback Models

In both IF and NF, the feedback signal is obtained via a random transformation. In particular, IF models the feedback link as an erasure-channel, whereas NF models the feedback link as an additive white Gaussian noise (AWGN) channel. Alternatively in RLF, the feedback signal is obtained via a deterministic transformation. Let $\overleftarrow{\text{SNR}}$ be the SNR in each of the feedback links from the receiver to the corresponding transmitters in the symmetric G-IC with NF (G-IC-NF) mentioned above. Let also β and β' be

$$\beta = \frac{\log\left(\overleftarrow{\text{SNR}}\right)}{\log\left(\overrightarrow{\text{SNR}}\right)} \tag{7a}$$

and

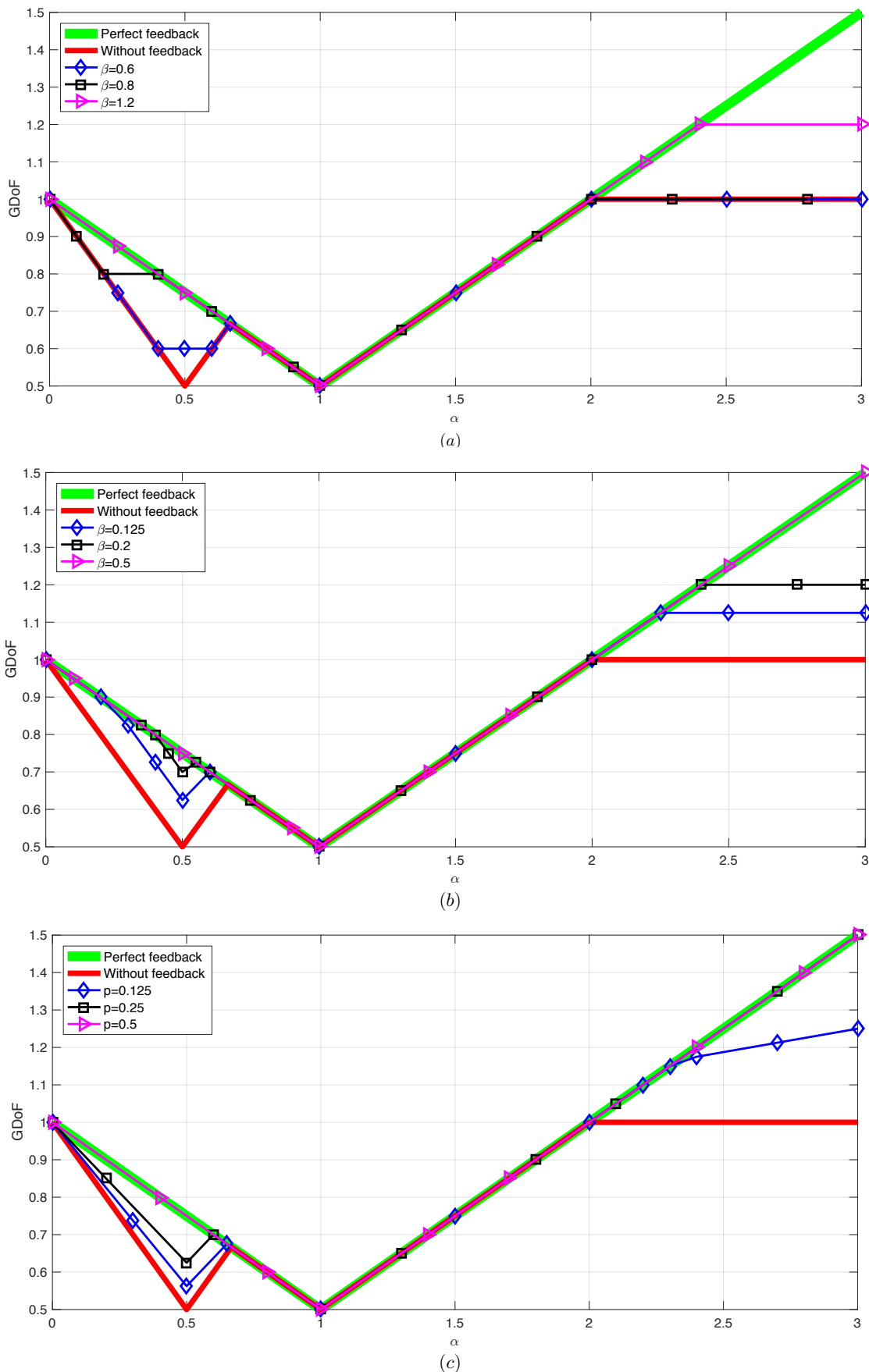$$\beta' = \frac{C_F}{\log\left(\overrightarrow{\text{SNR}}\right)}. \tag{7b}$$

Fig. 1. Number of generalized degrees of freedom (GDoF) of a symmetric two-user G-IC; (a) case with NF with $\beta \in \{0.6, 0.8, 1.2\}$; (b) case with RLF with $\beta \in \{0.125, 0.2, 0.5\}$; and (c) case with IF with $p \in \{0.125, 0.25, 0.5\}$.

These parameters approximate the ratio between the capacity of the feedback link and the capacity of the forward link in the NF and RLF case, respectively. Hence, a fair comparison of RLF and NF must be made with $\beta = \beta'$. The number of GDoF is plotted as a function of α when $\mathcal{C}(\overrightarrow{\mathrm{SNR}}, \mathrm{INR})$ is calculated with NF for several values of β in Figure 1(a); with RLF for different values of β' in Figure 1(b); and with IF for several values of p in Figure 1(c).

The most pessimistic channel-output feedback model between NF and RLF, in terms of the number of GDoF when $\beta = \beta'$, is NF. When $\alpha \in (0, \frac{2}{3})$ or $\alpha \in (2, \infty)$, RLF increases the number of GDoF for all $\beta' > 0$. Note that RLF with $\beta' = \frac{1}{2}$ achieves the same performance as PF, for all $\alpha \in [0, 3]$. In the case of NF, there does not exist any benefit in terms of the number of GDoF for all $0 < \beta < \frac{1}{2}$. A noticeable effect of NF occurs when $\alpha \in (0, \frac{2}{3})$, for all $\beta > \frac{1}{2}$; and when $\alpha \in (2, \infty)$, for all $\beta > 1$. This observation can be explained from the fact that in RLF, receivers extract relevant information about interference and send it via a noiseless channel. Alternatively, NF requires sending to the transmitter an exact copy of the channel output via an AWGN channel. Hence with $\beta = \beta' > 0$, the transmitters are always able to obtain information about the interference in RLF, whereas the same is not always true for NF. Finally, note that in both NF and RLF, the number of GDoF is not monotonically increasing with α in the interval $[2, \infty)$. Instead, it is upper-bounded by $\min\left(\frac{\alpha}{2}, \beta\right)$ in NF and by $\min\left(\frac{\alpha}{2}, 1 + \beta\right)$ in RLF.

The most optimistic model in terms of the number of GDoF, aside from PF, is IF. In particular because for any value of $p > 0$, there always exists an improvement of the number of GDoF for all $\alpha \in (0, \frac{2}{3})$ and $\alpha \in (2, \infty)$. Note that, with $p \geqslant \frac{1}{2}$, IF provides the same number of GDoF as PF. Note also that the number of GDoF remains being monotonically increasing with α in the interval $[2, \infty)$ for any positive value of p in (5), which implies an arbitrarily large increment in the number of GDoF.

E. Contributions

In this paper, the capacity of the G-IC-NF is approximated to within 4.4 bits by a new achievable region and a new converse region. These results generalize the approximate capacity region of the G-IC-NF presented in [25], [26] for the symmetric case. The gap between the new achievable region and the new converse region is slightly improved with respect to the one obtained in [25].

The methodology is the same used in [3], [5], [9], [23], [25], among others, i.e., a linear deterministic (LD) approximation [35] to the G-IC, referred to as LD-IC, is studied to gain insight on the construction of both inner and outer bounds. From this perspective, a byproduct of the main results is the full characterization of the capacity region of the LD-IC with NF (LD-IC-NF).

The achievability scheme presented in this paper as well as the one in [25] use a four-layer block-Markov superposition coding and backward decoding. Note that the achievability scheme used in [25] is obtained as a special case of the one presented in [16], [17]. The achievability scheme presented in this paper is developed independently. The main difference between these achievability schemes lies on the choice of the random variables used to generate the codewords of each of the layers of the codebook. Another difference is the power optimization made to obtain the corresponding achievable regions.

The converse region presented in this paper uses existing bounds from the case of PF in [3] and new bounds that generalize those in [25]. The proof of converse presented in [25] uses standard techniques including cut-set bounds and genie-aided channels, which are the same techniques used in this paper. Nonetheless, such generalization is far from trivial, as suggested in [25, Section IV-D].

F. Organization of the paper

Section II introduces the notation used in this paper. Section III describes the two-user G-IC-NF and LD-IC-NF. Section IV describes the exact capacity region of the LD-IC-NF. Section V introduces the main results, essentially, an achievable region and a converse region for the G-IC-NF. Section VI describes the connections between the IC-NF and an IC with conferencing transmitters (IC-CT). Finally, Section VII concludes this work and highlights some extensions.

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, whereas their realizations are denoted by the corresponding lower case letter, e.g., 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 $\mathbf{X} = (X_1, X_2, ..., X_N)^{\mathsf{T}}$ and a corresponding realization is denoted by $\mathbf{x} = (x_1, x_2, ..., x_N)^{\mathsf{T}} \in \mathcal{X}^N$. Given $\mathbf{X} = (X_1, X_2, ..., X_N)^{\mathsf{T}}$ and $(a, b) \in \mathbb{N}^2$, with $a < b \leqslant N$, the $(b - a + 1)$-dimensional vector of random variables formed by the components a to b of $\boldsymbol{X}$ is denoted by $\mathbf{X}_{(a:b)} = (X_a, X_{a+1}, \ldots, X_b)^{\mathsf{T}}$. 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 P_X of the random variable X. The logarithm function $\log$ is assumed to be base 2.

III. PROBLEM FORMULATION

Consider the two-user G-IC-NF depicted in Figure 2. Transmitter i, with $i \in \{1, 2\}$, communicates with receiver i subject to the interference produced by transmitter j, with $j \in \{1, 2\}\backslash\{i\}$. There are two independent and uniformly distributed messages, $W_i \in \mathcal{W}_i$, with $\mathcal{W}_i = \{1, 2, \ldots, 2^{NR_i}\}$, where $N \in \mathbb{N}$ denotes the block-length in channel uses and $R_i \geqslant 0$ is the transmission rate in bits per channel use. For transmitting a message index W_i, transmitter i sends a given codeword denoted by $\boldsymbol{X}_i = (X_{i,1}, X_{i,2}, \ldots, X_{i,N})^{\mathsf{T}} \in \mathbb{R}^N$. The channel coefficient from transmitter j to receiver i is denoted by h_{ij}; the channel coefficient from transmitter i to
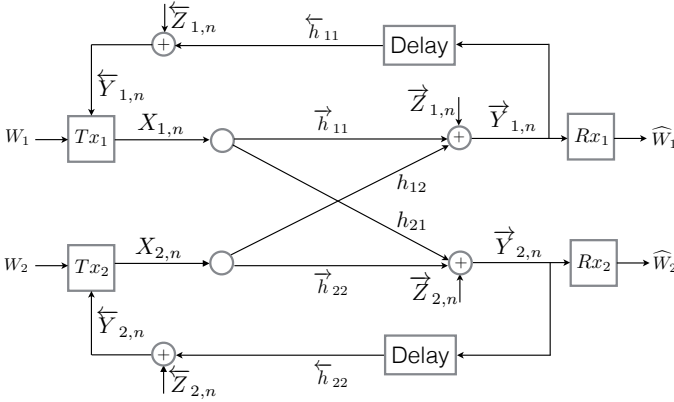
Fig. 2. Gaussian interference channel with noisy channel-output feedback at channel use n.

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

receiver i is denoted by $\overrightarrow{h}_{ii}$; 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 $n \in \{1, 2, \ldots, N\}$, the channel output at receiver i is denoted by $\overrightarrow{Y}_{i,n}$. During channel use n, the input-output relation of the channel model is given by

$$\overrightarrow{Y}_{i,n} = \overrightarrow{h}_{ii} X_{i,n} + h_{ij} X_{j,n} + \overrightarrow{Z}_{i,n}, \qquad (8)$$

where $\overrightarrow{Z}_{i,n}$ 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 $\overrightarrow{Y}_{i,n-d}$. More specifically,

$$\overleftarrow{Y}_{i,n} = \begin{cases} \overleftarrow{Z}_{i,n} & \text{for } n \in \{1,2,\ldots,d\} \\ \overleftarrow{h}_{ii} \overrightarrow{Y}_{i,n-d} + \overleftarrow{Z}_{i,n} & \text{for } n \in \{d+1,d+2,\ldots,N\}, \end{cases} \qquad (9)$$

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 $\overrightarrow{Z}_{i,n}$ and $\overleftarrow{Z}_{i,n}$ are assumed to be independent. In the following, without loss of generality, the feedback delay is assumed to be equal to one channel use, i.e., $d = 1$. The encoder of transmitter i is defined by a set of deterministic functions $\left\{ f_i^{(1)}, f_i^{(2)}, \ldots, f_i^{(N)} \right\}$, with $f_i^{(1)} : \mathcal{W}_i \to \mathbb{R}$ and for all $n \in \{2, 3, \ldots, N\}$, $f_i^{(n)} : \mathcal{W}_i \times \mathbb{R}^{n-1} \to \mathbb{R}$, such that

$$X_{i,1} = f_i^{(1)}(W_i) \text{ and} \qquad (10a)$$
$$X_{i,n} = f_i^{(n)}\left(W_i, \overleftarrow{Y}_{i,1}, \overleftarrow{Y}_{i,2}, \ldots, \overleftarrow{Y}_{i,n-1}\right). \qquad (10b)$$

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

$$\frac{1}{N} \sum_{n=1}^{N} \mathbb{E}_{X_{i,n}}\left[X_{i,n}^2 \right] \le 1. \qquad (11)$$
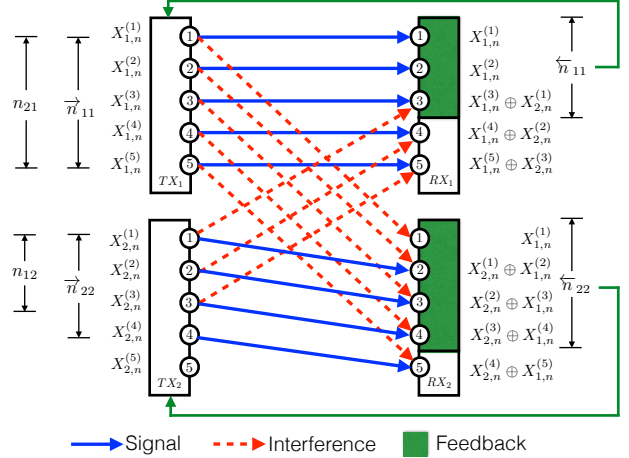
The decoder of receiver i is defined by a deterministic function $\psi_i^{(N)} : \mathbb{R}^N \to \mathcal{W}_i$. At the end of the communication, receiver i uses the vector $\left(\overrightarrow{Y}_{i,1}, \overrightarrow{Y}_{i,2}, \ldots, \overrightarrow{Y}_{i,N} \right)^{\mathsf{T}}$ to obtain an estimate of the message index:

$$\widehat{W}_i = \psi_i^{(N)}\left(\overrightarrow{Y}_{i,1}, \overrightarrow{Y}_{i,2}, \ldots, \overrightarrow{Y}_{i,N} \right), \qquad (12)$$

where $\widehat{W}_i$ is an estimate of the message index W_i. The decoding error probability in the two-user G-IC-NF of a codebook of block-length N, denoted by $P_e(N)$, is given by

$$P_e(N) = \max \left(\Pr\left[\widehat{W}_1 \ne W_1 \right], \Pr\left[\widehat{W}_2 \ne W_2 \right] \right). \qquad (13)$$

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 exist encoding functions $f_i^{(1)}, f_i^{(2)}, \ldots, f_i^{(N)}$ and decoding functions $\psi_i^{(1)}$, $\psi_i^{(2)}, \ldots, \psi_i^{(N)}$ for all $i \in \{1, 2\}$ such that the decoding error probability $P_e(N)$ can be made arbitrarily small by letting the block-length N grow to infinity.

In the next sections, it is shown that the capacity region of the two-user G-IC-NF in Figure 2 can be described by six parameters: $\overrightarrow{\mathrm{SNR}}_i$, $\overleftarrow{\mathrm{SNR}}_i$, and INR_{ij}, with $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$, which are defined as follows:

$$\overrightarrow{\mathrm{SNR}}_i = \overrightarrow{h}_{ii}^2, \qquad (14)$$
$$\mathrm{INR}_{ij} = h_{ij}^2, \text{ and} \qquad (15)$$
$$\overleftarrow{\mathrm{SNR}}_i = \overleftarrow{h}_{ii}^2 \left(\overrightarrow{h}_{ii}^2 + 2\overrightarrow{h}_{ii} h_{ij} + h_{ij}^2 + 1 \right). \qquad (16)$$

IV. LINEAR DETERMINISTIC CHANNELS

This section describes the two-user LD-IC-NF and its exact capacity region. The relevance of this result is that it provides the main insight used to obtain the approximate capacity of the G-IC-NF in Section V.

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

Consider the two-user LD-IC-NF depicted 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 intended receiver is denoted by $\overrightarrow{n}_{ii}$; the number of bit-pipes between transmitter i and its non-intended receiver is denoted by n_{ji}; and the number of bit-pipes between receiver i and its corresponding transmitter is denoted by $\overleftarrow{n}_{ii}$.

At transmitter i, the channel-input $\boldsymbol{X}_{i,n}$ at channel use n, with $n \in \{1, 2, \ldots, N\}$, is a q-dimensional binary vector $\boldsymbol{X}_{i,n} = \left(X_{i,n}^{(1)}, X_{i,n}^{(2)}, \ldots, X_{i,n}^{(q)} \right)^{\mathsf{T}}$, with

$$q = \max\left(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21} \right), \quad (17)$$

and N the block-length. At receiver i, the channel-output $\overrightarrow{\boldsymbol{Y}}_{i,n}$ at channel use n is also a q-dimensional binary vector $\overrightarrow{\boldsymbol{Y}}_{i,n} = \left(\overrightarrow{Y}_{i,n}^{(1)}, \overrightarrow{Y}_{i,n}^{(2)}, \ldots, \overrightarrow{Y}_{i,n}^{(q)} \right)^{\mathsf{T}}$. The input-output relation during channel use n is given by

$$\overrightarrow{\boldsymbol{Y}}_{i,n} = \boldsymbol{S}^{q - \overrightarrow{n}_{ii}} \boldsymbol{X}_{i,n} + \boldsymbol{S}^{q - n_{ij}} \boldsymbol{X}_{j,n}, \quad (18)$$

where, $\boldsymbol{S}$ is a $q \times q$ lower shift matrix.

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

$$\left((0, \ldots, 0), \overleftarrow{\boldsymbol{Y}}_{i,n}^{\mathsf{T}} \right)^{\mathsf{T}} = \boldsymbol{S}^{\left(\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii} \right)^+} \overrightarrow{\boldsymbol{Y}}_{i,n-d}, \quad (19)$$

where d is a finite delay, additions and multiplications are defined over the binary field. The dimension of the vector $(0, \ldots, 0)$ in (19) is $q - \min\left(\overleftarrow{n}_{ii}, \max(\overrightarrow{n}_{ii}, n_{ij}) \right)$ and the vector $\overleftarrow{\boldsymbol{Y}}_{i,n}$ represents the $\min\left(\overleftarrow{n}_{ii}, \max(\overrightarrow{n}_{ii}, n_{ij}) \right)$ least significant bits of $\boldsymbol{S}^{\left(\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii} \right)^+} \overrightarrow{\boldsymbol{Y}}_{i,n-d}$.

The feedback delay is assumed to be equal to one channel use, i.e., $d = 1$. The encoding and decoding operations are performed analogously to those in the G-IC case.

The decoding error probability is calculated following (13). 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 $\mathcal{C}(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ the capacity region of the LD-IC-NF with parameters $\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}$, and $\overleftarrow{n}_{22}$. Theorem 1 (in the top of next page) fully characterizes this capacity region. The proof of Theorem 1 is divided into two parts. The first part describes the proof of the achievability and is presented in Appendix A. The second part describes the proof of converse and is presented in Appendix B.

1) Comments on the Achievability Scheme: Let the channel input of transmitter i during channel use n be $\boldsymbol{X}_{i,n} = \left(X_{i,n}^{(1)}, X_{i,n}^{(2)}, \ldots, X_{i,n}^{(q)} \right)^{\mathsf{T}} \in \{0, 1\}^q$. The achievability scheme is a sequel to the following observation: feedback allows the transmitters to obtain information about the interference produced by its counterpart at the intended receiver. This information could be retransmitted aiming either at performing

interference cancellation at the intended receiver or at providing an alternative communication path to the other transmitter-receiver pair [3], [9], [25]. From this standpoint, there are three types of bit-pipes that start at transmitter i that are particularly relevant:

(i) the set of bit-pipes that are observed above the noise level by both receiver j and transmitter j, i.e.,

$$\boldsymbol{U}_{i,n} = \left(X_{i,n}^{(1)}, X_{i,n}^{(2)}, \ldots, X_{i,n}^{(q_{i,1})} \right)^{\mathsf{T}}, \quad (21)$$

where,

$$q_{i,1} = \left(n_{ji} - \left(\max\left(\overrightarrow{n}_{jj}, n_{ji} \right) - \overleftarrow{n}_{jj} \right)^+ \right)^+; \quad (22)$$

(ii) the set of bit-pipes that are observed above the noise level by receiver j and below the (feedback) noise level by transmitter j, i.e.,

$$\boldsymbol{V}_{i,n} = \left(X_{i,n}^{(q_{i,1}+1)}, X_{i,n}^{(q_{i,1}+2)}, \ldots, X_{i,n}^{(q_{i,1}+q_{i,2})} \right)^{\mathsf{T}}, \quad (23)$$

where,

$$q_{i,2} = \min\left(n_{ji}, \left(\max\left(\overrightarrow{n}_{jj}, n_{ji} \right) - \overleftarrow{n}_{jj} \right)^+ \right); \text{ and } \quad (24)$$

(iii) the set of bit-pipes that are exclusively observed above the noise level by receiver i, i.e.,

$$\boldsymbol{X}_{i,P,n} = \left(X_{i,n}^{(q_{i,1}+q_{i,2}+1)}, X_{i,n}^{(q_{i,1}+q_{i,2}+2)}, \ldots, X_{i,n}^{(q_{i,1}+q_{i,2}+q_{i,3})} \right)^{\mathsf{T}}, \quad (25)$$

where,

$$q_{i,3} = \left(\overrightarrow{n}_{ii} - n_{ji} \right)^+. \quad (26)$$

Note that

$$q_{i,1} + q_{i,2} + q_{i,3} = \max\left(\overrightarrow{n}_{ii}, n_{ji} \right) \leqslant q, \quad (27)$$

and thus,

$$\left(X_{i,n}^{(q_{i,1}+q_{i,2}+q_{i,3}+1)}, X_{i,n}^{(q_{i,1}+q_{i,2}+q_{i,3}+2)}, \ldots, X_{i,n}^{(q)} \right) = (0, \ldots, 0).$$

Hence, for all $n \in \{1, 2, \ldots, N\}$,

$$\boldsymbol{X}_{i,n} = \left(\boldsymbol{U}_{i,n}^{\mathsf{T}}, \boldsymbol{V}_{i,n}^{\mathsf{T}}, \boldsymbol{X}_{i,P,n}^{\mathsf{T}}, (0, \ldots, 0) \right)^{\mathsf{T}}. \quad (28)$$

An example of the concatenation of $\boldsymbol{U}_{i,n}$, $\boldsymbol{V}_{i,n}$ and $\boldsymbol{X}_{i,P,n}$ to form the input symbol $\boldsymbol{X}_{i,n}$ is presented in Figure 4. Note that the vector $(0, \ldots, 0)$ exists only when (27) holds with strict inequality.

Within this context, some key observations are worth highlighting:

(a) The interference produced by the bits $\boldsymbol{U}_{j,n}$ on $\boldsymbol{X}_{i,n}$ at receiver i can be eliminated in two cases: when $\boldsymbol{U}_{j,n}$ consists of bits previously transmitted by transmitter i; or when the bits $\boldsymbol{U}_{j,n}$ are retransmitted by transmitter i at a later channel use such that they can be reliably decoded by receiver i. In both cases, receiver i is able to implement interference cancellation.
(b) The interference produced by the bits $\boldsymbol{V}_{j,n}$ on $\boldsymbol{X}_{i,n}$ at receiver i can be eliminated in a single case: when the bits in $\boldsymbol{V}_{j,n}$ consists of bits previously transmitted by transmitter i.
(c) The top

$$\min\left(\left(\overrightarrow{n}_{ii} - n_{ij} \right)^+, \left(\min\left(\overleftarrow{n}_{jj}, \max(\overrightarrow{n}_{jj}, n_{ji}) \right) - \left(\overrightarrow{n}_{jj} - n_{ji} \right)^+ \right)^+ \right)$$

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

$$R_i \leqslant \min\left(\max\left(\overrightarrow{n}_{ii}, n_{ji}\right), \max\left(\overrightarrow{n}_{ii}, n_{ij}\right)\right), \tag{20a}$$

$$R_i \leqslant \min\left(\max\left(\overrightarrow{n}_{ii}, n_{ji}\right), \max\left(\overrightarrow{n}_{ii}, \overleftarrow{n}_{jj} - \left(\overrightarrow{n}_{jj} - n_{ji}\right)^+\right)\right), \tag{20b}$$

$$R_1 + R_2 \leqslant \min\left(\max\left(\overrightarrow{n}_{22}, n_{12}\right) + \left(\overrightarrow{n}_{11} - n_{12}\right)^+, \max\left(\overrightarrow{n}_{11}, n_{21}\right) + \left(\overrightarrow{n}_{22} - n_{21}\right)^+\right), \tag{20c}$$

$$R_1 + R_2 \leqslant \max\left(\left(\overrightarrow{n}_{11} - n_{12}\right)^+, n_{21}, \overrightarrow{n}_{11} - \left(\max\left(\overrightarrow{n}_{11}, n_{12}\right) - \overleftarrow{n}_{11}\right)^+\right)$$
$$+ \max\left(\left(\overrightarrow{n}_{22} - n_{21}\right)^+, n_{12}, \overrightarrow{n}_{22} - \left(\max\left(\overrightarrow{n}_{22}, n_{21}\right) - \overleftarrow{n}_{22}\right)^+\right), \tag{20d}$$

$$2R_i + R_j \leqslant \max\left(\overrightarrow{n}_{ii}, n_{ji}\right) + \left(\overrightarrow{n}_{ii} - n_{ij}\right)^+ + \max\left(\left(\overrightarrow{n}_{jj} - n_{ji}\right)^+, n_{ij}, \overrightarrow{n}_{jj} - \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+\right). \tag{20e}$$
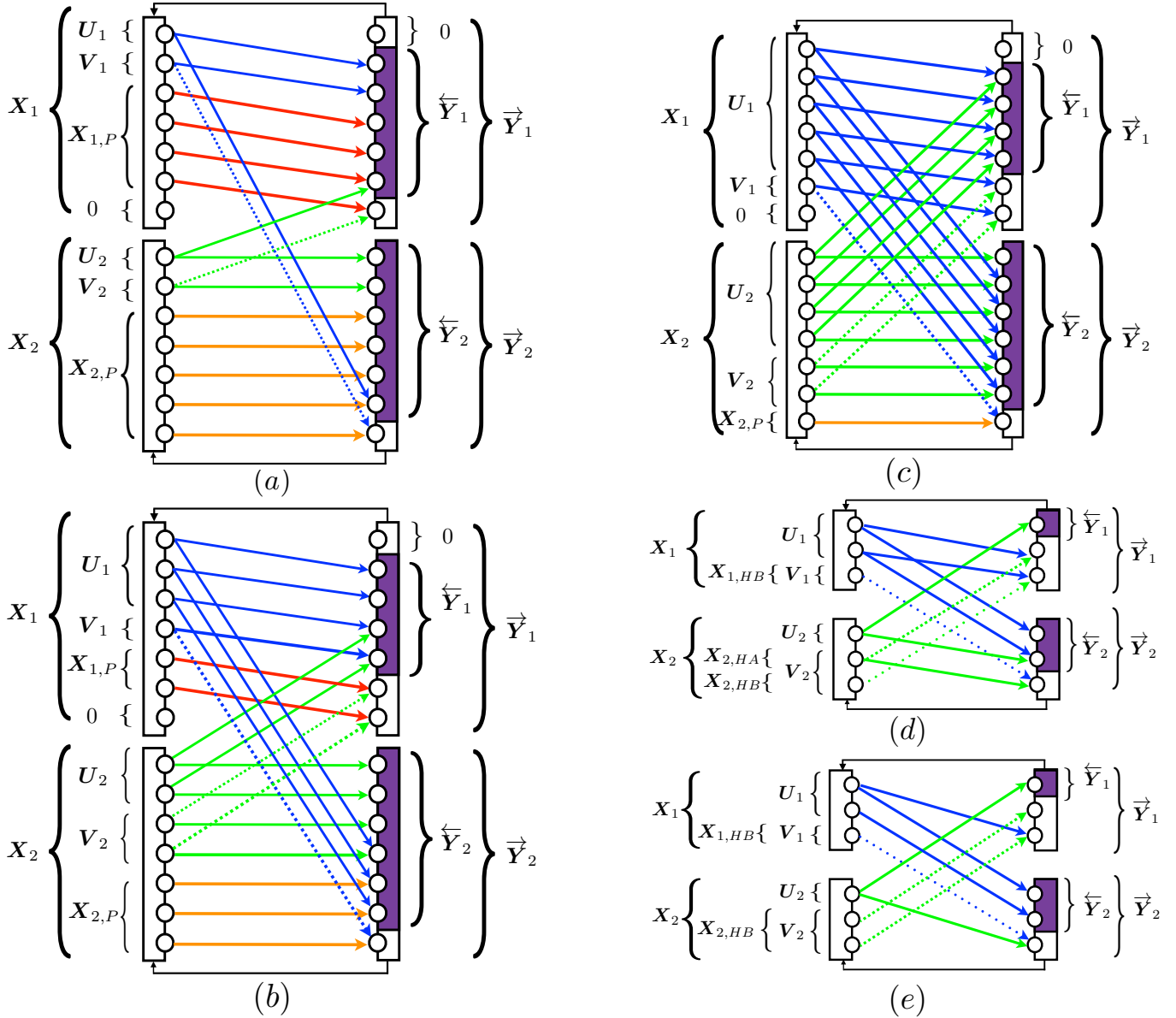
Fig. 4. Concatenation of binary vectors $\boldsymbol{U}_{i,n} \in \{0, 1\}^{q_{i,1}}$, with $q_{i,1}$ defined in (22); $\boldsymbol{V}_{i,n} \in \{0, 1\}^{q_{i,2}}$, with $q_{i,2}$ defined in (24); and $\boldsymbol{X}_{i,P,n} \in \{0, 1\}^{q_{i,3}}$, with $q_{i,3}$ defined in (26) to form the input symbol $\boldsymbol{X}_{i,n}$ in (a) very weak interference regime, (b) weak interference regime, (c) moderate interference regime, (d) strong interference regime, and (e) very strong interference regime.
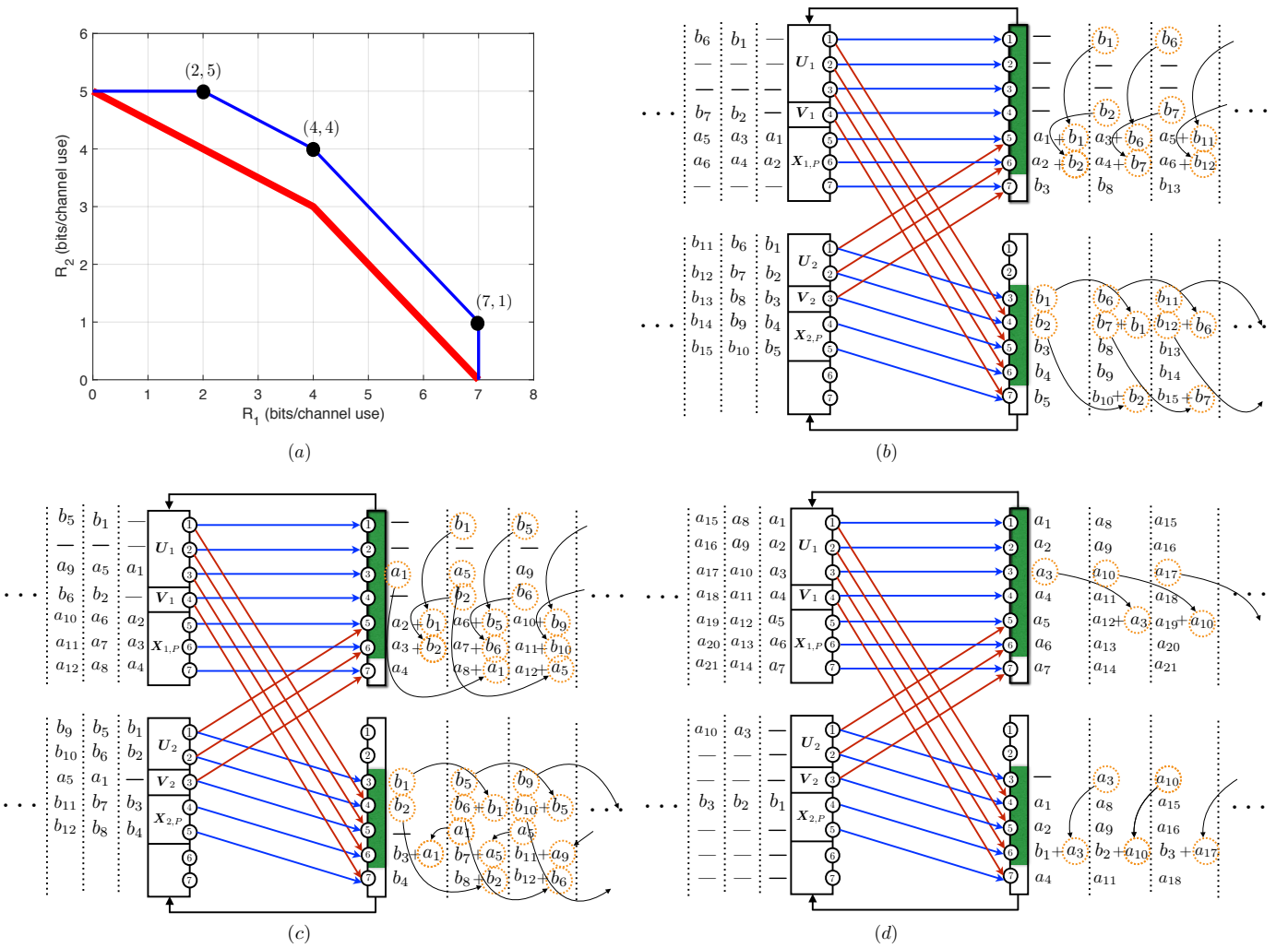
Fig. 5. (a) Capacity regions of $\mathcal{C}(7,5,3,4,0,0)$ (thick red line) and $\mathcal{C}(7,5,3,4,6,4)$ (thin blue line). (b) Achievability of the rate pair $(2,5)$ in $\mathcal{C}(7,5,3,4,6,4)$. (c) Achievability of the rate pair $(4,4)$ in an $\mathcal{C}(7,5,3,4,6,4)$. (d) Achievability of the rate pair $(7,1)$ in an $\mathcal{C}(7,5,3,4,6,4)$.

bits in $\boldsymbol{U}_{i,n}$ are observed interference-free at receiver i and thus, they are reliably decoded at the end of channel use n. Note also that these top bits in $\boldsymbol{U}_{i,n}$ produce interference in receiver j. However, thanks to feedback, transmitter j can retransmit these bits in $\boldsymbol{U}_{j,n+m}$ or $\boldsymbol{V}_{j,n+m}$, for some $m > 0$. Interestingly, the interference produced by these retransmissions can be eliminated at receiver i at the end of channel use $n + m$.

(d) The bottom

$$\min\left((n_{ji} - \overrightarrow{n}_{ii})^+, \left(\overleftarrow{n}_{jj} - \overrightarrow{n}_{ii} - \min\left((\overrightarrow{n}_{jj} - n_{ji})^+, n_{ij} \right) \right. \right.$$
$$\left. \left. - \left((\overrightarrow{n}_{jj} - n_{ij})^+ - n_{ji} \right)^+ \right)^+ \right)$$

bits of $\boldsymbol{U}_{i,n}$ are not observed at receiver i but receiver j. Thus, they can be fed back to transmitter j. If transmitter j retransmits these bits in any of the bits $\boldsymbol{U}_{j,n+m}$ or $\boldsymbol{V}_{j,n+m}$, for some $m > 0$, these bits might be reliably decoded at receiver i forming an alternative communication path to transmitter-receiver pair i.

Taking into account the facts (a) - (d), simple coding schemes can be constructed as shown in the examples in Figure 5. A complete proof of the achievability is presented in Appendix A.

2) Comments on the Converse Region: The outer bounds (20a) and (20c) are cut-set bounds and were first reported in [36] for the case without feedback. These outer bounds are still useful in the case of perfect channel-output feedback [3]. The outer bounds (20b), (20d) and (20e) are new.

The outer bound (20b) is on the individual rate of transmitter-receiver i when at channel use n, receiver i is granted with the knowledge of the message index W_j and all previous feedback signals observed at transmitter j, i.e., $\overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}$. See for instance Figure 6(a) for the case $i = 1$. A complete proof of (20b) is presented in Appendix B.

The outer bound (20d) is on the sum-rate of both transmitter-receiver pairs with enhanced versions of the receivers. More specifically, for all $i \in \{1,2\}$, receiver i is granted at channel use n with the knowledge of all previous feedback signals observed at transmitter i, i.e., $\overleftarrow{\boldsymbol{Y}}_{i,(1:n-1)}$. See for instance Figure 6(b). A complete proof of (20d) is presented in Appendix B.

The outer bounds (20e) are on the sum-rate of three re-

Fig. 6. Genie-aided G-IC-NF 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 $2R_1 + R_2$.
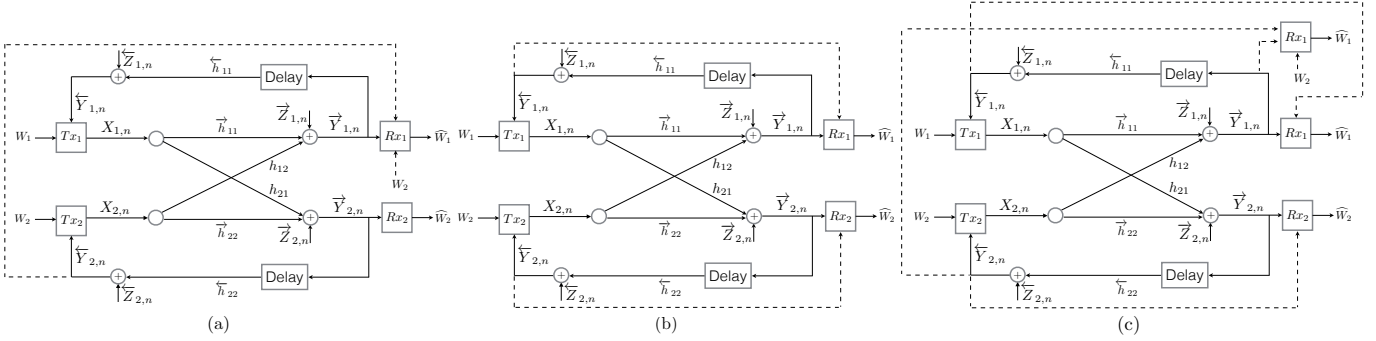
ceivers: for all $i \in \{1, 2\}$, receiver i is granted at channel use n with all previous feedback signals observed at transmitter i, i.e., $\overleftarrow{Y}_{i,(1:n-1)}$. For at most one $i \in \{1, 2\}$, a third receiver has input $\overrightarrow{Y}_{i,n}$ and it is granted with the knowledge of the message index W_j and all previous feedback signals observed at transmitter j, i.e., $\overleftarrow{Y}_{j,(1:n-1)}$. See for instance Figure 6(c) for the case $i = 1$. A complete proof of (20e) is presented in Appendix B.

C. Connections with Previous Results

Theorem 1 generalizes previous results on 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 describes the capacity region of the LD-IC without feedback (Lemma 4 in [36]); when $\overleftarrow{n}_{11} \geqslant \max(\overrightarrow{n}_{11}, n_{12})$ and $\overleftarrow{n}_{22} \geqslant \max(\overrightarrow{n}_{22}, n_{21})$, Theorem 1 describes the capacity region of the LD-IC with perfect channel output feedback (Corollary 1 in [3]); when $\overrightarrow{n}_{11} = \overrightarrow{n}_{22}$, $n_{12} = n_{21}$ 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 [25] and Theorem 4.1, case 1001 in [7]); and when $\overrightarrow{n}_{11} = \overrightarrow{n}_{22}$, $n_{12} = n_{21}$, $\overleftarrow{n}_{ii} \geqslant \max(\overrightarrow{n}_{ii}, n_{ij})$ 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 when only one of the receivers provides perfect channel-output feedback (Theorem 4.1, cases 1000 and 0001 in [7]).

V. GAUSSIAN CHANNELS

This section introduces an achievable region (Theorem 2) and a converse region (Theorem 3), denoted by $\underline{\mathcal{C}}_{\mathrm{G-IC-NF}}$ and $\overline{\mathcal{C}}_{\mathrm{G-IC-NF}}$ respectively, for a two-user G-IC-NF with fixed parameters $\overrightarrow{\mathrm{SNR}}_1$, $\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, INR_{21}, $\overleftarrow{\mathrm{SNR}}_1$, and $\overleftarrow{\mathrm{SNR}}_2$. 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 $\boldsymbol{t} = (t_1, t_2, \ldots, t_m) \in \overline{\mathcal{T}}$, $\left((t_1 - \xi)^+, (t_2 - \xi)^+, \ldots, (t_m - \xi)^+\right) \in \underline{\mathcal{T}}$.*

Denote by $\mathcal{C}_{\mathrm{G-IC-NF}}$ the capacity region of the 2-user G-IC-NF. The achievable region $\underline{\mathcal{C}}_{\mathrm{G-IC-NF}}$ (Theorem 2) and

the converse region $\overline{\mathcal{C}}_{\mathrm{G-IC-NF}}$ (Theorem 3) approximate the capacity region $\mathcal{C}_{\mathrm{G-IC-NF}}$ to within 4.4 bits (Theorem 4).

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

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

$$a_{1,i} = \frac{1}{2} \log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\right) - \frac{1}{2}, \tag{29a}$$

$$a_{2,i}(\rho) = \frac{1}{2} \log\left(b_{1,i}(\rho) + 1\right) - \frac{1}{2}, \tag{29b}$$

$$a_{3,i}(\rho, \mu) = \frac{1}{2} \log\left(\frac{\overleftarrow{\mathrm{SNR}}_i\left(b_{2,i}(\rho) + 2\right) + b_{1,i}(1) + 1}{\overleftarrow{\mathrm{SNR}}_i\left((1-\mu)b_{2,i}(\rho) + 2\right) + b_{1,i}(1) + 1}\right), \tag{29c}$$

$$a_{4,i}(\rho, \mu) = \frac{1}{2} \log\left(\left(1 - \mu\right)b_{2,i}(\rho) + 2\right) - \frac{1}{2}, \tag{29d}$$

$$a_{5,i}(\rho, \mu) = \frac{1}{2} \log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}} + \left(1 - \mu\right)b_{2,i}(\rho)\right) - \frac{1}{2}, \tag{29e}$$

$$a_{6,i}(\rho, \mu) = \frac{1}{2} \log\left(\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\left(\left(1-\mu\right)b_{2,j}(\rho) + 1\right) + 2\right) - \frac{1}{2}, \tag{29f}$$

and

$$a_{7,i}(\rho, \mu_1, \mu_2) = \frac{1}{2} \log\left(\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\left(\left(1 - \mu_i\right)b_{2,j}(\rho) + 1\right)\right.$$
$$\left. + \left(1 - \mu_j\right)b_{2,i}(\rho) + 2\right) - \frac{1}{2}, \tag{29g}$$

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

$$b_{1,i}(\rho) = \overrightarrow{\mathrm{SNR}}_i + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_i\mathrm{INR}_{ij}} + \mathrm{INR}_{ij} \text{ and} \tag{30a}$$

$$b_{2,i}(\rho) = \left(1 - \rho\right)\mathrm{INR}_{ij} - 1, \tag{30b}$$

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

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

$$R_1 \leqslant \min\Big(a_{2,1}(\rho), a_{6,1}(\rho,\mu_1)+a_{3,2}(\rho,\mu_1), a_{1,1}+a_{3,2}(\rho,\mu_1)+a_{4,2}(\rho,\mu_1)\Big), \tag{31a}$$

$$R_2 \leqslant \min\Big(a_{2,2}(\rho), a_{3,1}(\rho,\mu_2)+a_{6,2}(\rho,\mu_2), a_{3,1}(\rho,\mu_2)+a_{4,1}(\rho,\mu_2)+a_{1,2}\Big), \tag{31b}$$

$$R_1+R_2 \leqslant \min\Big(a_{2,1}(\rho)+a_{1,2}, a_{1,1}+a_{2,2}(\rho), a_{3,1}(\rho,\mu_2)+a_{1,1}+a_{3,2}(\rho,\mu_1)+a_{7,2}(\rho,\mu_1,\mu_2),$$

$$a_{3,1}(\rho,\mu_2)+a_{5,1}(\rho,\mu_2)+a_{3,2}(\rho,\mu_1)+a_{5,2}(\rho,\mu_1), a_{3,1}(\rho,\mu_2)+a_{7,1}(\rho,\mu_1,\mu_2)+a_{3,2}(\rho,\mu_1)+a_{1,2}\Big), \tag{31c}$$

$$2R_1+R_2 \leqslant \min\Big(a_{2,1}(\rho)+a_{1,1}+a_{3,2}(\rho,\mu_1)+a_{7,2}(\rho,\mu_1,\mu_2), \tag{31d}$$

$$a_{3,1}(\rho,\mu_2)+a_{1,1}+a_{7,1}(\rho,\mu_1,\mu_2)+2a_{3,2}(\rho,\mu_1)+a_{5,2}(\rho,\mu_1), a_{2,1}(\rho)+a_{1,1}+a_{3,2}(\rho,\mu_1)+a_{5,2}(\rho,\mu_1)\Big),$$

$$R_1+2R_2 \leqslant \min\Big(a_{3,1}(\rho,\mu_2)+a_{5,1}(\rho,\mu_2)+a_{2,2}(\rho)+a_{1,2}, a_{3,1}(\rho,\mu_2)+a_{7,1}(\rho,\mu_1,\mu_2)+a_{2,2}(\rho)+a_{1,2},$$

$$2a_{3,1}(\rho,\mu_2)+a_{5,1}(\rho,\mu_2)+a_{3,2}(\rho,\mu_1)+a_{1,2}+a_{7,2}(\rho,\mu_1,\mu_2)\Big), \tag{31e}$$

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

Finally, using this notation, Theorem 2 is presented at the top of this page.

The proof of the achievability is based on random coding arguments that make use of the insight obtained in the analysis of the LD-IC-NF. The techniques are rate splitting, block-Markov superposition coding, and backward decoding. The complete proof is described in Appendix A. However, a brief description of the ideas leading to the construction of an achievability scheme are discussed hereunder.

One of the central observations in the examples presented for the LD-IC-NF is that transmitters use correlated codewords. This stems from the fact that a fraction of the bits sent by transmitter i are received by receiver j; fed back to transmitter j and finally, retransmitted by transmitter j. See for instance [3], [9], [17], [25], and [33]. This observation is the driving idea in the construction of the achievability schemes presented in previous works and it is central in the proof of Theorem 2.

Let the message index sent by transmitter i during the t-th block be denoted by $W_i^{(t)} \in \{1, 2, \ldots, 2^{NR_i}\}$. Following a rate-splitting argument, assume that $W_i^{(t)}$ is represented by three subindices $(W_{i,C1}^{(t)}, W_{i,C2}^{(t)}, W_{i,P}^{(t)}) \in \{1, 2, \ldots, 2^{NR_{i,C1}}\} \times \{1, 2, \ldots, 2^{NR_{i,C2}}\} \times \{1, 2, \ldots, 2^{NR_{i,P}}\}$, where $R_{i,C1} + R_{i,C2} + R_{i,P} = R_i$. The number of rate splits is based on the fact that each symbol in the LD-IC-NF was divided into three parts.

The codeword generation follows a four-level superposition coding scheme. The number of layers is the number of rate splits plus an additional common layer that accounts for the correlation between codewords. This correlation is induced as follows. 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 $W_{2,C1}^{(t-1)}$. Using these indices both transmitters are able to identify the same codeword in the first code-layer. In the case of the first block $t=1$ and the last

block $t = T$, the indices $W_{1,C1}^{(0)}$, $W_{2,C1}^{(0)}$, $W_{1,C1}^{(T)}$, and $W_{2,C1}^{(T)}$ correspond to indices assumed to be known by all transmitters and receivers.

The first code-layer is a sub-codebook of $2^{N(R_{1,C1}+R_{2,C1})}$ codewords (see Figure 10). Denote by $\boldsymbol{u}\big(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}\big)$ the corresponding codeword in the first code-layer. The second codeword used by transmitter i is selected using $W_{i,C1}^{(t)}$ from the second code-layer, which is a sub-codebook of $2^{NR_{i,C1}}$ codewords specific to $\boldsymbol{u}\big(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}\big)$ as shown in Figure 10. Denote by $\boldsymbol{u}_i\big(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}\big)$ the corresponding codeword in the second code-layer. The third codeword used by transmitter i is selected using $W_{i,C2}^{(t)}$ from the third code-layer, which is a sub-codebook of $2^{NR_{i,C2}}$ codewords specific to $\boldsymbol{u}_i\big(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{t,C1}^{(t)}\big)$ as shown in Figure 10. Denote by $\boldsymbol{v}_i\big(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)}\big)$ the corresponding codeword in the third code-layer. The fourth codeword used by transmitter i is selected using $W_{i,P}^{(t)}$ from the fourth code-layer, which is a sub-codebook of $2^{NR_{i,P}}$ codewords specific to $\boldsymbol{v}_i\big(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)}\big)$ as shown in Figure 10. Denote by $\boldsymbol{x}_{i,P}\big(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)}, W_{i,P}^{(t)}\big)$ the corresponding codeword in the fourth code-layer. Finally, the channel input sequence at transmitter i during block t is denoted by $\boldsymbol{x}_{i,t} = (x_{i,t,1}, x_{i,t,2}, \ldots, x_{i,t,N})$, with $t \in \{1, 2, \ldots, T\}$, and it is a weighted sum of the codewords $\boldsymbol{u}\big(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}\big)$, $\boldsymbol{u}_i\big(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}\big)$, $\boldsymbol{v}_i\big(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)}\big)$ and $\boldsymbol{x}_{i,P}\big(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)}, W_{i,P}^{(t)}\big)$.

At receiver i, the decoding follows a backward decoding approach as in previous works [3], [9], [25], among others.

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

The description of the converse region $\overline{\mathcal{C}}_{\text{G-IC-NF}}$ is determined by two events denoted by $S_{l_1,1}$ and $S_{l_2,2}$, where

$(l_1, l_2) \in \{1, \ldots, 5\}^2$. The events are defined as follows:

$$S_{1,i}: \quad \overrightarrow{\text{SNR}}_j < \min\left(\text{INR}_{ij}, \text{INR}_{ji}\right), \tag{32a}$$

$$S_{2,i}: \quad \text{INR}_{ji} \leqslant \overrightarrow{\text{SNR}}_j < \text{INR}_{ij}, \tag{32b}$$

$$S_{3,i}: \quad \text{INR}_{ij} \leqslant \overrightarrow{\text{SNR}}_j < \text{INR}_{ji}, \tag{32c}$$

$$S_{4,i}: \quad \max\left(\text{INR}_{ij}, \text{INR}_{ji}\right) \leqslant \overrightarrow{\text{SNR}}_j < \text{INR}_{ij}\text{INR}_{ji}, \tag{32d}$$

$$S_{5,i}: \quad \overrightarrow{\text{SNR}}_j \geqslant \text{INR}_{ij}\text{INR}_{ji}. \tag{32e}$$

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 $(l_1, l_2) \in \{1, \ldots, 5\}^2$, that determines 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 (32). Once the exact scenario is identified, the converse region is described using the functions $\kappa_{l,i} : [0, 1] \rightarrow \mathbb{R}_+$, with $l \in \{1, \ldots, 3\}$; $\kappa_l : [0, 1] \rightarrow \mathbb{R}_+$, with $l \in \{4, 5\}$; $\kappa_{6,l} : [0, 1] \rightarrow \mathbb{R}_+$, with $l \in \{1, \ldots, 4\}$; 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 $j \in \{1, 2\} \setminus \{i\}$:

$$\kappa_{1,i}(\rho) = \frac{1}{2} \log\left(b_{1,i}(\rho) + 1\right), \tag{33a}$$

$$\kappa_{2,i}(\rho) = \frac{1}{2} \log\left(1 + b_{5,j}(\rho)\right) + \frac{1}{2} \log\left(1 + \frac{b_{4,i}(\rho)}{1 + b_{5,j}(\rho)}\right), \tag{33b}$$

$$\kappa_{3,i}(\rho) = \frac{1}{2} \log\left(\frac{\left(b_{4,i}(\rho) + b_{5,j}(\rho) + 1\right)\overleftarrow{\text{SNR}}_j}{\left(b_{1,j}(1) + 1\right)\left(b_{4,i}(\rho) + 1\right)} + 1\right)$$
$$+ \frac{1}{2} \log\left(b_{4,i}(\rho) + 1\right), \tag{33c}$$

$$\kappa_4(\rho) = \frac{1}{2} \log\left(1 + \frac{b_{4,1}(\rho)}{1 + b_{5,2}(\rho)}\right) + \frac{1}{2} \log\left(b_{1,2}(\rho) + 1\right), \tag{33d}$$

$$\kappa_5(\rho) = \frac{1}{2} \log\left(1 + \frac{b_{4,2}(\rho)}{1 + b_{5,1}(\rho)}\right) + \frac{1}{2} \log\left(b_{1,1}(\rho) + 1\right), \tag{33e}$$

$$\kappa_6(\rho) = \begin{cases} \kappa_{6,1}(\rho) & \text{if } (S_{1,2} \vee S_{2,2} \vee S_{5,2}) \\ & \wedge (S_{1,1} \vee S_{2,1} \vee S_{5,1}) \\ \kappa_{6,2}(\rho) & \text{if } (S_{1,2} \vee S_{2,2} \vee S_{5,2}) \\ & \wedge (S_{3,1} \vee S_{4,1}) \\ \kappa_{6,3}(\rho) & \text{if } (S_{3,2} \vee S_{4,2}) \\ & \wedge (S_{1,1} \vee S_{2,1} \vee S_{5,1}) \\ \kappa_{6,4}(\rho) & \text{if } (S_{3,2} \vee S_{4,2}) \\ & \wedge (S_{3,1} \vee S_{4,1}), \end{cases} \tag{33f}$$

$$\kappa_{7,i}(\rho) = \begin{cases} \kappa_{7,i,1}(\rho) & \text{if } (S_{1,i} \vee S_{2,i} \vee S_{5,i}) \\ \kappa_{7,i,2}(\rho) & \text{if } (S_{3,i} \vee S_{4,i}), \end{cases} \tag{33g}$$

where,

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

$$+ \frac{1}{2} \log\left(1 + \frac{b_{5,2}(\rho)\overleftarrow{\text{SNR}}_2}{b_{1,2}(1) + 1}\right)$$

$$+ \frac{1}{2} \log\left(b_{1,2}(\rho) + b_{5,1}(\rho)\text{INR}_{21}\right) - \frac{1}{2} \log\left(1 + \text{INR}_{21}\right)$$

$$+ \frac{1}{2} \log\left(1 + \frac{b_{5,1}(\rho)\overleftarrow{\text{SNR}}_1}{b_{1,1}(1) + 1}\right) + \log(2\pi e), \tag{34a}$$

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

$$- \frac{1}{2} \log\left(1 + \text{INR}_{12}\right) + \frac{1}{2} \log\left(1 + \frac{b_{5,1}(\rho)\overleftarrow{\text{SNR}}_1}{b_{1,1}(1) + 1}\right)$$

$$+ \frac{1}{2} \log\left(b_{1,1}(\rho) + b_{5,1}(\rho)\text{INR}_{21}\right) - \frac{1}{2} \log\left(1 + \text{INR}_{21}\right)$$

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

$$- \frac{1}{2} \log\left(1 + \frac{b_{5,1}(\rho)\text{INR}_{21}}{\overrightarrow{\text{SNR}}_2}\right) + \log(2\pi e), \tag{34b}$$

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

$$- \frac{1}{2} \log\left(1 + \text{INR}_{12}\right) + \frac{1}{2} \log\left(1 + \frac{b_{5,2}(\rho)\overleftarrow{\text{SNR}}_2}{b_{1,2}(1) + 1}\right)$$

$$+ \frac{1}{2} \log\left(b_{1,2}(\rho) + b_{5,1}(\rho)\text{INR}_{21}\right) - \frac{1}{2} \log\left(1 + \text{INR}_{21}\right)$$

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

$$- \frac{1}{2} \log\left(1 + \frac{b_{5,1}(\rho)\text{INR}_{21}}{\overrightarrow{\text{SNR}}_1}\right) + \log(2\pi e), \tag{34c}$$

$$\kappa_{6,4}(\rho) = \frac{1}{2} \log\left(b_{6,1}(\rho) + \frac{b_{5,1}(\rho)\text{INR}_{21}}{\overrightarrow{\text{SNR}}_1}\left(\overrightarrow{\text{SNR}}_1 + b_{3,1}\right)\right)$$

$$- \frac{1}{2} \log\left(1 + \text{INR}_{12}\right) - \frac{1}{2} \log\left(1 + \text{INR}_{21}\right)$$

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

$$- \frac{1}{2} \log\left(1 + \frac{b_{5,1}(\rho)\text{INR}_{21}}{\overrightarrow{\text{SNR}}_2}\right)$$

$$- \frac{1}{2} \log\left(1 + \frac{b_{5,1}(\rho)\text{INR}_{21}}{\overrightarrow{\text{SNR}}_1}\right)$$

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

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

$$+ \log(2\pi e), \tag{34d}$$

and

$$\kappa_{7,i,1}(\rho) = \frac{1}{2} \log\left(b_{1,i}(\rho) + 1\right) - \frac{1}{2} \log\left(1 + \text{INR}_{ij}\right)$$

$$+ \frac{1}{2} \log\left(1 + \frac{b_{5,j}(\rho)\overleftarrow{\text{SNR}}_j}{b_{1,j}(1) + 1}\right) + 2\log(2\pi e)$$

$$+ \frac{1}{2} \log\left(b_{1,j}(\rho) + b_{5,i}(\rho)\text{INR}_{ji}\right)$$

$$+\frac{1}{2}\log\left(1+b_{4,i}(\rho)+b_{5,j}(\rho)\right)-\frac{1}{2}\log\left(1+b_{5,j}(\rho)\right)\ (35a)$$

$$\kappa_{7,i,2}(\rho)=\frac{1}{2}\log\left(b_{1,i}(\rho)+1\right)-\frac{1}{2}\log\left(1+\text{INR}_{ij}\right)$$
$$-\frac{1}{2}\log\left(1+b_{5,j}(\rho)\right)+\frac{1}{2}\log\left(1+b_{4,i}(\rho)+b_{5,j}(\rho)\right)$$
$$+\frac{1}{2}\log\left(1+\left(1-\rho^2\right)\frac{\text{INR}_{ji}}{\overrightarrow{\text{SNR}}_j}\left(\text{INR}_{ij}+\frac{b_{3,j}\overleftarrow{\text{SNR}}_j}{b_{1,j}(1)+1}\right)\right)$$
$$-\frac{1}{2}\log\left(1+\frac{b_{5,i}(\rho)\text{INR}_{ji}}{\overrightarrow{\text{SNR}}_j}\right)$$
$$+\frac{1}{2}\log\left(b_{6,j}(\rho)+\frac{b_{5,i}(\rho)\text{INR}_{ji}}{\overrightarrow{\text{SNR}}_j}\left(\overrightarrow{\text{SNR}}_j+b_{3,j}\right)\right)$$
$$+2\log(2\pi e). \tag{35b}$$

The functions $b_{l,i}$, with $(l,i)\in\{1,2\}^2$ are defined in (30); $b_{3,i}$ are constants; and the functions $b_{l,i}:[0,1]\to\mathbb{R}_+$, with $(l,i)\in\{4,5,6\}\times\{1,2\}$ are defined as follows, with $j\in\{1,2\}\setminus\{i\}$:

$$b_{3,i}=\overrightarrow{\text{SNR}}_i-2\sqrt{\overrightarrow{\text{SNR}}_i\text{INR}_{ji}}+\text{INR}_{ji}, \tag{36a}$$

$$b_{4,i}(\rho)=\left(1-\rho^2\right)\overrightarrow{\text{SNR}}_i, \tag{36b}$$

$$b_{5,i}(\rho)=\left(1-\rho^2\right)\text{INR}_{ij}, \tag{36c}$$

$$b_{6,i}(\rho)=\overrightarrow{\text{SNR}}_i+\text{INR}_{ij}+2\rho\sqrt{\text{INR}_{ij}}\left(\sqrt{\overrightarrow{\text{SNR}}_i}-\sqrt{\text{INR}_{ji}}\right)$$
$$+\frac{\text{INR}_{ij}\sqrt{\text{INR}_{ji}}}{\overrightarrow{\text{SNR}}_i}\left(\sqrt{\text{INR}_{ji}}-2\sqrt{\overrightarrow{\text{SNR}}_i}\right). \tag{36d}$$

Finally, using this notation, Theorem 3 is presented below.

Theorem 3: The capacity region $\mathcal{C}_{\text{G}-\text{IC}-\text{NF}}$ is contained within the region $\overline{\mathcal{C}}_{\text{G}-\text{IC}-\text{NF}}$ 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:

$$R_i\leqslant\min\left(\kappa_{1,i}(\rho),\kappa_{2,i}(\rho)\right), \tag{37a}$$

$$R_i\leqslant\kappa_{3,i}(\rho), \tag{37b}$$

$$R_1+R_2\leqslant\min\left(\kappa_4(\rho),\kappa_5(\rho)\right), \tag{37c}$$

$$R_1+R_2\leqslant\kappa_6(\rho), \tag{37d}$$

$$2R_i+R_j\leqslant\kappa_{7,i}(\rho), \tag{37e}$$

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

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

The outer bounds (37a) and (37c) play the same role as the outer bounds (20a) and (20c) in the linear deterministic model and have been previously reported in [3] for the case of perfect channel-output feedback. The bounds (37b), (37d), and (37e) correspond to new outer bounds. The intuition for deriving these outer bounds follows along the same lines of the outer bounds (20b), (20d), and (20e) in the LD-IC-NF, respectively.

C. Connections with Previous Results

Theorem 2 generalizes previous results on the achievable region of the two-user G-IC with channel-output feedback. For instance, when $\overleftarrow{\text{SNR}}_1=0$, $\overleftarrow{\text{SNR}}_2=0$, and $\rho=0$,

Theorem 2 describes the achievable region of the G-IC without feedback [12], [21], [37]; when $\overleftarrow{\text{SNR}}_1\to\infty$ and $\overleftarrow{\text{SNR}}_2\to\infty$, Theorem 2 describes the achievable region of the G-IC-PF (Theorem 2 in [3]); when $\overrightarrow{\text{SNR}}_1=\overrightarrow{\text{SNR}}_2$, $\text{INR}_{12}=\text{INR}_{21}$ and $\overleftarrow{\text{SNR}}_1=\overleftarrow{\text{SNR}}_2$, Theorem 2 describes the achievable region of the symmetric G-IC-NF (Theorem 3 in [25]). Theorem 3 generalizes previous results on the converse region of the two-user G-IC with channel-output feedback. For instance, when $\overleftarrow{\text{SNR}}_1=0$, $\overleftarrow{\text{SNR}}_2=0$, and $\rho=0$, Theorem 3 describes the converse region of the G-IC without feedback [5]; when $\overleftarrow{\text{SNR}}_1\to\infty$ and $\overleftarrow{\text{SNR}}_2\to\infty$, Theorem 3 describes the converse region of the G-IC-PF (Theorem 3 in [3]); when $\overrightarrow{\text{SNR}}_1=\overrightarrow{\text{SNR}}_2$, $\text{INR}_{12}=\text{INR}_{21}$ and $\overleftarrow{\text{SNR}}_1=\overleftarrow{\text{SNR}}_2$, Theorem 3 describes the converse region of the symmetric G-IC-NF (Theorem 2 in [25]).

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

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

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

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

To the extent of the knowledge of the authors, this approximation to the capacity region of the G-IC-NF 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. Figure 7 presents the exact gap existing between the achievable region $\underline{\mathcal{C}}_{\text{G}-\text{IC}-\text{NF}}$ and the converse region $\overline{\mathcal{C}}_{\text{G}-\text{IC}-\text{NF}}$ 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

Despite the fundamental differences discussed in [25], there exist several similarities between the G-IC-NF and the two-user G-IC with conferencing transmitters (G-IC-CT). For instance, there might be cases in which the side information that can be obtained by the transmitters is the same either using transmitter cooperation or channel-output feedback.

Consider for instance the LD-IC with conferencing-transmitters (LD-IC-CT) depicted in Figure 8 and the LD-IC-NF depicted in Figure 9. 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 $\boldsymbol{X}_{i,CF,n}$ and $\boldsymbol{X}_{i,DF,n}$ (see Figure 8 and Figure 9), for all $i\in\{1,2\}$ and $j\in\{1,2\}\setminus\{i\}$. Then, in this particular case, subject to a finite delay, in both channel models (Figure 8 and Figure 9) the corresponding transmitters possess the same side information and the corresponding receivers observe the same channel outputs. From this point of view, any outer
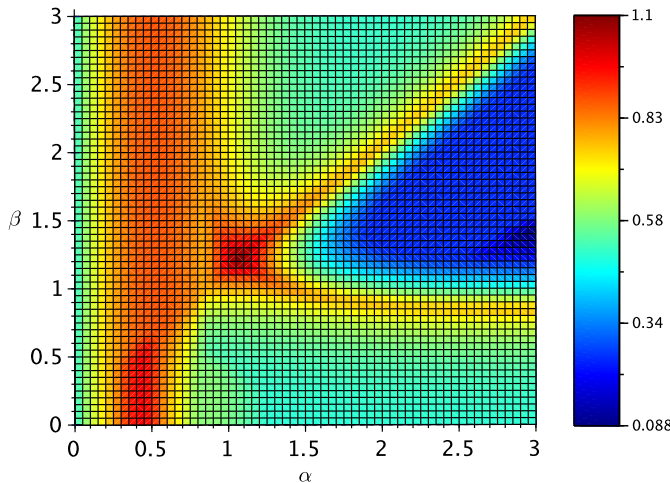
Fig. 7. Gap between the converse region $\overline{\mathcal{C}}_{\mathrm{G-IC-NF}}$ and the achievable region $\underline{\mathcal{C}}_{\mathrm{G-IC-NF}}$ of the two-user G-IC-NF, under symmetric channel conditions, i.e., $\overrightarrow{\mathrm{SNR}}_1 = \overrightarrow{\mathrm{SNR}}_2 = \overrightarrow{\mathrm{SNR}}$, $\mathrm{INR}_{12} = \mathrm{INR}_{21} = \mathrm{INR}$, and $\overleftarrow{\mathrm{SNR}}_1 = \overleftarrow{\mathrm{SNR}}_2 = \overleftarrow{\mathrm{SNR}}$, as a function of $\alpha = \frac{\log \mathrm{INR}}{\log \overrightarrow{\mathrm{SNR}}}$ and $\beta = \frac{\log \overleftarrow{\mathrm{SNR}}}{\log \overrightarrow{\mathrm{SNR}}}$.

bound on the individual or weighted sum-rate for the case of conferencing transmitters is also an outer bound for the case of channel-output feedback and vice-versa. A similar observation can be made for the case of a G-IC-NF and a G-IC with conferencing transmitters (G-IC-CT).

Outer bounds on the sum-rate for the LD-IC-CT and the G-IC-CT have been reported in [31]. The gap between the achievable sum-rate and the corresponding outer bound in the G-IC-CT is 20 bits in [31]. Other outer bounds have been presented in [38], 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-NF has been fully characterized (Theorem 1). Exploiting the insight obtained in the LD-IC-NF, an achievability region (Theorem 2) and a converse region (Theorem 3) have been presented for the two-user G-IC-NF. These two regions approximate the capacity region of the G-IC-NF to within 4.4 bits (Theorem 4).

Despite the contributions made in this paper, several questions remain unsolved in the understanding of the benefits of channel-output feedback in the G-IC-NF. For instance, the case in which the channel-output feedback is observed by both transmitters is still an open problem. Only the case of symmetric channels has been fully studied. Another case in which very little is known about the benefits of channel-output feedback is that of a large number of users (more than two) and large number of antennas (more than one) at each network component.

Appendix A
Proof of the Achievability

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

Codebook Generation: Fix a strictly positive joint probability distribution

$$P_{U\,U_1\,U_2\,V_1\,V_2\,X_{1,P}\,X_{2,P}}(u,u_1,u_2,v_1,v_2,x_{1,P},x_{2,P}) = P_U(u)$$
$$P_{U_1|U}(u_1|u)P_{U_2|U}(u_2|u)P_{V_1|U\,U_1}(v_1|u,u_1)$$
$$P_{V_2|U\,U_2}(v_2|u,u_2)P_{X_{1,P}|U\,U_1\,V_1}(x_{1,P}|u,u_1,v_1)$$
$$P_{X_{2,P}|U\,U_2\,V_2}(x_{2,P}|u,u_2,v_2), \tag{38}$$

for all $(u,u_1,u_2,v_1,v_2,x_{1,P},x_{2,P}) \in (\mathcal{X}_1 \cap \mathcal{X}_2) \times (\mathcal{X}_1 \times \mathcal{X}_2)^3$.

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. codewords of length N $\boldsymbol{u}(s,r) = (u_1(s,r),u_2(s,r),\ldots,u_N(s,r))$ according to

$$P_{\boldsymbol{U}}(\boldsymbol{u}(s,r)) = \prod_{n=1}^{N} P_U(u_n(s,r)), \tag{39}$$

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

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

$$P_{\boldsymbol{U}_1|\boldsymbol{U}}(\boldsymbol{u}_1(s,r,k)|\boldsymbol{u}(s,r)) = \prod_{n=1}^{N} P_{U_1|U}(u_{1,n}(s,r,k)|u_n(s,r)), \tag{40}$$

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

$$P_{\boldsymbol{V}_1|\boldsymbol{U}\,\boldsymbol{U}_1}(\boldsymbol{v}_1(s,r,k,l)|\boldsymbol{u}(s,r),\boldsymbol{u}_1(s,r,k))$$
$$= \prod_{n=1}^{N} P_{V_1|U\,U_1}(v_{1,n}(s,r,k,l)|u_n(s,r),u_{1,n}(s,r,k)), \tag{41}$$

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

$$P_{\boldsymbol{X}_{1,P}|\boldsymbol{U}\,\boldsymbol{U}_1\boldsymbol{V}_1}(\boldsymbol{x}_{1,P}(s,r,k,l,q)|\boldsymbol{u}(s,r),\boldsymbol{u}_1(s,r,k),\boldsymbol{v}_1(s,r,k,l))$$
$$= \prod_{n=1}^{N} P_{X_{1,P}|U\,U_1\,V_1}\Big(x_{1,P,n}(s,r,k,l,q)|u_n(s,r),u_{1,n}(s,r,k),$$
$$v_{1,n}(s,r,k,l)\Big), \tag{42}$$

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

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

$$P_{\boldsymbol{U}_2|\boldsymbol{U}}(\boldsymbol{u}_2(s,r,j)|\boldsymbol{u}(s,r)) = \prod_{n=1}^{N} P_{U_2|U}(u_{2,n}(s,r,j)|u_n(s,r)), \tag{43}$$
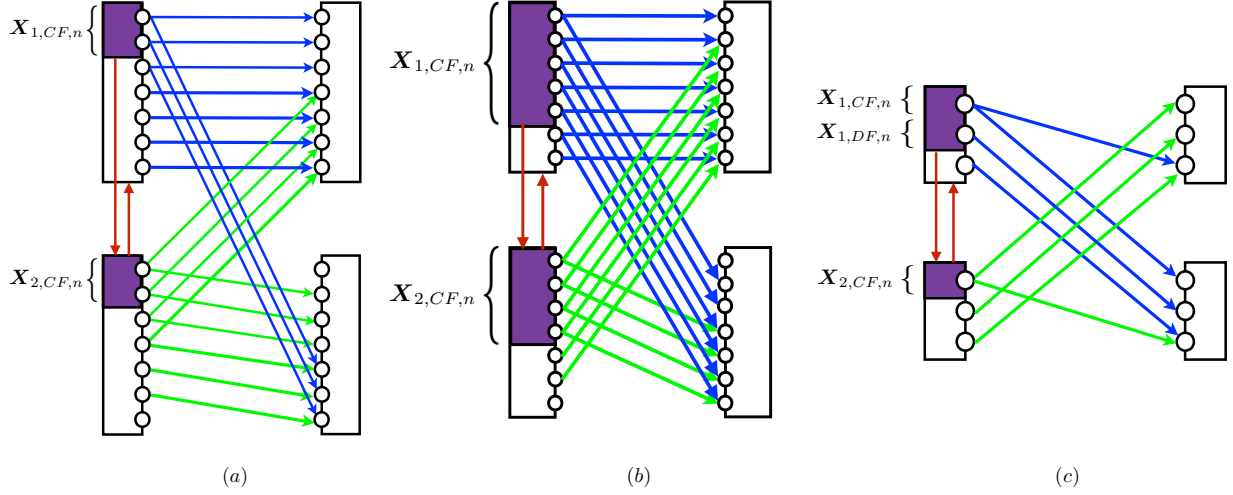
Fig. 8. Linear Deterministic Interference Channels with Conferencing Transmitters.
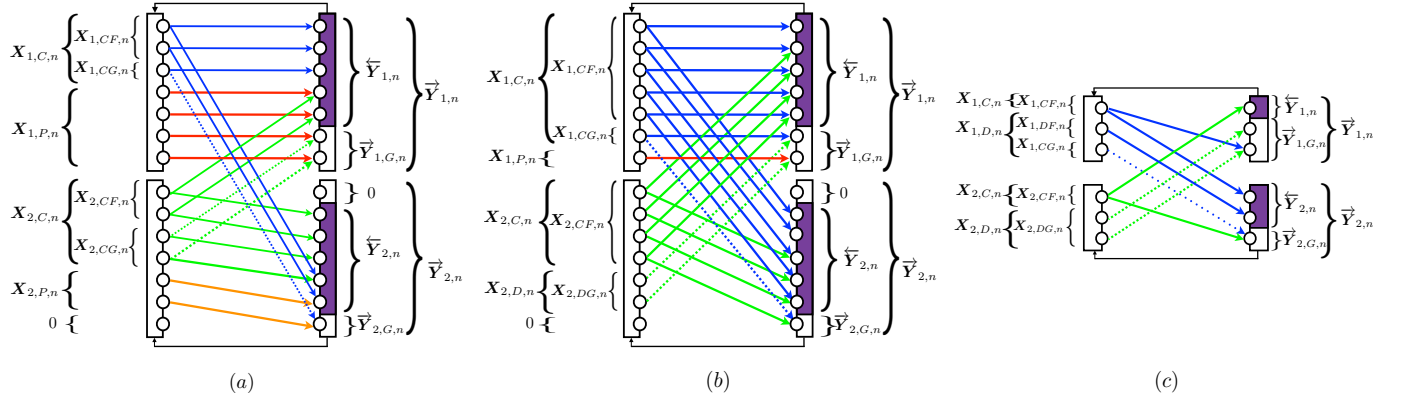
Fig. 9. Example of the notation of the channel inputs and the channel outputs when channel-output feedback is considered.

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

$$P_{\boldsymbol{V}_2|\boldsymbol{U}\boldsymbol{U}_2}\left(\boldsymbol{v}_2(s, r, j, m)|\boldsymbol{u}(s, r), \boldsymbol{u}_2(s, r, j)\right)$$
$$= \prod_{n=1}^{N} P_{V_2|UU_2}(v_{2,n}(s, r, j, m)|u_n(s, r), u_{2,n}(s, r, j)), \quad (44)$$

with $m \in \{1, 2, \ldots, 2^{NR_{2,C2}}\}$. For each tuple of codewords $\left(\boldsymbol{u}(s, r), \boldsymbol{u}_2(s, r, j), \boldsymbol{v}_2(s, r, j, m)\right)$, generate $2^{NR_{2,P}}$ i.i.d. codewords of length N $\boldsymbol{x}_{2,P}(s, r, j, m, b) = \left(x_{2,P,1}(s, r, j, m, b), \quad x_{2,P,2}(s, r, j, m, b), \quad \ldots, x_{2,P,N}(s, r, j, m, b)\right)$ according to

$$P_{\boldsymbol{X}_{2,P}|\boldsymbol{U}\boldsymbol{U}_2\boldsymbol{V}_2}\left(\boldsymbol{x}_{2,P}(s, r, j, m, b)|\boldsymbol{u}(s, r), \boldsymbol{u}_2(s, r, j), \boldsymbol{v}_2(s, r, j, m)\right)$$
$$= \prod_{n=1}^{N} P_{X_{2,P}|UU_2V_2}\left(x_{2,P,n}(s, r, j, m, b)|u_n(s, r), u_{2,n}(s, r, j),\right.$$
$$\left. v_{2,n}(s, r, j, m)\right), \quad (45)$$

with $b \in \{1, 2, \ldots, 2^{NR_{2,P}}\}$. The resulting code structure is shown in Figure 10.

Encoding: Denote by $W_i^{(t)} \in \{1, 2, \ldots, 2^{NR_i}\}$ the message index of transmitter $i \in \{1, 2\}$ during block $t \in \{1, 2, \ldots, T\}$, with T the total number of blocks. Let $W_i^{(t)}$ be composed by the message index $W_{i,C}^{(t)} \in \{1, 2, \ldots, 2^{NR_{i,C}}\}$ and message index $W_{i,P}^{(t)} \in \{1, 2, \ldots, 2^{NR_{i,P}}\}$, i.e., $W_i^{(t)} = \left(W_{i,C}^{(t)}, W_{i,P}^{(t)}\right)$. 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 $W_{i,C1}^{(t)} \in \{1, 2, \ldots, 2^{NR_{i,C1}}\}$ and $W_{i,C2}^{(t)} \in \{1, 2, \ldots, 2^{NR_{i,C2}}\}$, i.e., $W_{i,C}^{(t)} = \left(W_{i,C1}^{(t)}, W_{i,C2}^{(t)}\right)$. The message index $W_{i,C1}^{(t)}$ must be reliably decoded by the other transmitter (via feedback) and by both receivers. The message index $W_{i,C2}^{(t)}$ must be reliably decoded by both receivers, but not transmitter j.

Consider block-Markov encoding over T blocks. At encoding step t, with $t \in \{1, 2, \ldots, T\}$, transmitter 1 sends the

codeword:

$$\boldsymbol{x}_1^{(t)}=\Theta_1\left(\boldsymbol{u}\left(W_{1,C1}^{(t-1)},W_{2,C1}^{(t-1)}\right),\boldsymbol{u}_1\left(W_{1,C1}^{(t-1)},W_{2,C1}^{(t-1)},W_{1,C1}^{(t)}\right),\right.$$
$$\boldsymbol{v}_1\left(W_{1,C1}^{(t-1)},W_{2,C1}^{(t-1)},W_{1,C1}^{(t)},W_{1,C2}^{(t)}\right),$$
$$\left.\boldsymbol{x}_{1,P}\left(W_{1,C1}^{(t-1)},W_{2,C1}^{(t-1)},W_{1,C1}^{(t)},W_{1,C2}^{(t)},W_{1,P}^{(t)}\right)\right),\quad(46)$$

where, $\Theta_1 : (\mathcal{X}_1\cap\mathcal{X}_2)^N \times \mathcal{X}_1^{3N} \rightarrow \mathcal{X}_1^N$ is a function that transforms the codewords $\boldsymbol{u}\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}\right)$, $\boldsymbol{u}_1\left(W_{1,C1}^{(t-1)},W_{2,C1}^{(t-1)},W_{1,C1}^{(t)}\right)$, $\boldsymbol{v}_1\left(W_{1,C1}^{(t-1)},W_{2,C1}^{(t-1)},W_{1,C1}^{(t)},W_{1,C2}^{(t)}\right)$, and $\boldsymbol{x}_{1,P}\left(W_{1,C1}^{(t-1)},\right.$ $W_{2,C1}^{(t-1)}, W_{1,C1}^{(t)}, W_{1,C2}^{(t)}, \left.W_{1,P}^{(t)}\right)$ into the N-dimensional vector $\boldsymbol{x}_1^{(t)}$ 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 $(s^*,r^*) \in \{1,2,\ldots,2^{N\,R_{1,C1}}\} \times \{1,2,\ldots,2^{NR_{2,C1}}\}$ 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 $\overleftarrow{\boldsymbol{y}}_1^{(t-1)}$ 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 $t \in \{1,2,\ldots,T\}$, receiver 1 obtains the message indices $\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, \widehat{W}_{1,C2}^{(T-(t-1))},\right.$ $\left.\widehat{W}_{1,P}^{(T-(t-1))}, \widehat{W}_{2,C2}^{(T-(t-1))}\right) \in \{1, 2,\ldots, 2^{NR_{1,C1}}\} \times \{1, 2,\ldots,2^{NR_{2,C1}}\}\times\{1,2,\ldots,2^{NR_{1,C2}}\}\times\{1,2,\ldots,2^{NR_{1,P}}\}\times \{1, 2,\ldots,2^{NR_{2,C2}}\}$ from the channel output $\overrightarrow{\boldsymbol{y}}_1^{(T-(t-1))}$. The tuple $\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, \widehat{W}_{1,C2}^{(T-(t-1))}, \widehat{W}_{1,P}^{(T-(t-1))},\right.$ $\left.\widehat{W}_{2,C2}^{(T-(t-1))}\right)$ is the unique tuple that satisfies

$$\left(\boldsymbol{u}\left(\widehat{W}_{1,C1}^{(T-t)},\widehat{W}_{2,C1}^{(T-t)}\right),\boldsymbol{u}_1\left(\widehat{W}_{1,C1}^{(T-t)},\widehat{W}_{2,C1}^{(T-t)},W_{1,C1}^{(T-(t-1))}\right),\right.$$
$$\boldsymbol{v}_1\left(\widehat{W}_{1,C1}^{(T-t)},\widehat{W}_{2,C1}^{(T-t)},W_{1,C1}^{(T-(t-1))},\widehat{W}_{1,C2}^{(T-(t-1))}\right),$$
$$\boldsymbol{x}_{1,P}\left(\widehat{W}_{1,C1}^{(T-t)},\widehat{W}_{2,C1}^{(T-t)},W_{1,C1}^{(T-(t-1))},\widehat{W}_{1,C2}^{(T-(t-1))},\widehat{W}_{1,P}^{(T-(t-1))}\right),$$
$$\boldsymbol{u}_2\left(\widehat{W}_{1,C1}^{(T-t)},\widehat{W}_{2,C1}^{(T-t)},W_{2,C1}^{(T-(t-1))}\right),$$
$$\left.\boldsymbol{v}_2\left(\widehat{W}_{1,C1}^{(T-t)},\widehat{W}_{2,C1}^{(T-t)},W_{2,C1}^{(T-(t-1))},\widehat{W}_{2,C2}^{(T-(t-1))}\right),\overrightarrow{\boldsymbol{y}}_1^{(T-(t-1))}\right)$$
$$\in\mathcal{T}_{\left[U\ U_1\ V_1\ X_{1,P}\ U_2\ V_2\ \overrightarrow{Y}_1\right]}^{(N,e)},\qquad(47)$$

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 $\mathcal{T}_{\left[U\ U_1\ V_1\ X_{1,P}\ U_2\ V_2\ \overrightarrow{Y}_1\right]}^{(N,e)}$ represents the set of jointly typical sequences of the random variables $U, U_1, V_1, X_{1,P}, U_2, V_2,$ and $\overrightarrow{Y}_1$, with $e > 0$. Receiver 2 follows a similar decoding scheme.

Error Probability Analysis: An error occurs during encoding step t if the message index $W_{2,C1}^{(t-1)}$ is not correctly decoded at transmitter 1. From the asymptotic equipartion property (AEP) [39], 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:

$$R_{2,C1} \leqslant I\left(\overleftarrow{Y}_1;U_2|U,U_1,V_1,X_1\right)$$
$$= I\left(\overleftarrow{Y}_1;U_2|U,X_1\right).\qquad(48)$$

An error occurs 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 $\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)},\widehat{W}_{1,C2}^{(T-(t-1))}, \widehat{W}_{1,P}^{(T-(t-1))}, \widehat{W}_{2,C2}^{(T-(t-1))}\right)$ that satisfies (47), or (ii) there exist several tuples $\left(\widehat{W}_{1,C1}^{(T-t)},\widehat{W}_{2,C1}^{(T-t)},\widehat{W}_{1,C2}^{(T-(t-1))},\widehat{W}_{1,P}^{(T-(t-1))},\widehat{W}_{2,C2}^{(T-(t-1))}\right)$ that simultaneously satisfy (47). From the AEP, 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 $\left(\boldsymbol{u}(s,r),\ \boldsymbol{u}_1(s,r,W_{1,C1}^{(T-(t-1))}),\ \boldsymbol{v}_1(s,r,W_{1,C1}^{(T-(t-1))},l),\right.$ $\boldsymbol{x}_{1,P}(s,r,W_{1,C1}^{(T-(t-1))},l,q),\ \boldsymbol{u}_2(s,r,W_{2,C1}^{(T-(t-1))}),$ and $\left.\boldsymbol{v}_2(s,r,W_{2,C1}^{(T-(t-1))},m)\right)$ are jointly typical with $\overrightarrow{\boldsymbol{y}}_1^{(T-(t-1))}$ during decoding step t. Assume now that the codeword to be decoded at decoding step t corresponds to the indices $(s,r,l,q,m) = (1,1,1,1,1)$. This is 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 upper-bounded as follows:

$$P_e(N)=\mathrm{Pr}\left[\bigcup_{(s,r,l,q,m)\neq(1,1,1,1,1)}E_{(s,r,l,q,m)}\right]$$
$$\leqslant \sum_{(s,r,l,q,m)\in\mathcal{T}}\mathrm{Pr}\left[E_{(s,r,l,q,m)}\right],\qquad(49)$$

with $\mathcal{T} = \left\{\{1,2,\ldots2^{NR_{1,C1}}\} \times \{1,2,\ldots2^{NR_{2,C1}}\} \times \{1,2,\ldots2^{NR_{1,C2}}\}\times\{1,2,\ldots2^{NR_{1,P}}\}\times\{1,2,\ldots2^{NR_{2,C2}}\}\right\}\backslash \{(1,1,1,1,1)\}$.

From the AEP [39], it follows that

$$P_e(N)\leqslant 2^{N(R_{2,C2}-I(\overrightarrow{Y}_1;V_2|U,U_1,U_2,V_1,X_1)+2\epsilon)}$$
$$+2^{N(R_{1,P}-I(\overrightarrow{Y}_1;X_1|U,U_1,U_2,V_1,V_2)+2\epsilon)}$$
$$+2^{N(R_{2,C2}+R_{1,P}-I(\overrightarrow{Y}_1;V_2,X_1|U,U_1,U_2,V_1)+2\epsilon)}$$
$$+2^{N(R_{1,C2}-I(\overrightarrow{Y}_1;V_1,X_1|U,U_1,U_2,V_2)+2\epsilon)}$$
$$+2^{N(R_{1,C2}+R_{2,C2}-I(\overrightarrow{Y}_1;V_1,V_2,X_1|U,U_1,U_2)+2\epsilon)}$$
$$+2^{N(R_{1,C2}+R_{1,P}-I(\overrightarrow{Y}_1;V_1,X_1|U,U_1,U_2,V_2)+2\epsilon)}$$
$$+2^{N(R_{1,C2}+R_{1,P}+R_{2,C2}-I(\overrightarrow{Y}_1;V_1,V_2,X_1|U,U_1,U_2)+2\epsilon)}$$
$$+2^{N(R_{2,C1}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)}$$
$$+2^{N(R_{2,C}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)}$$
$$+2^{N(R_{2,C1}+R_{1,P}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)}$$
$$+2^{N(R_{2,C}+R_{1,P}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)}$$
$$+2^{N(R_{2,C1}+R_{1,C2}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)}$$
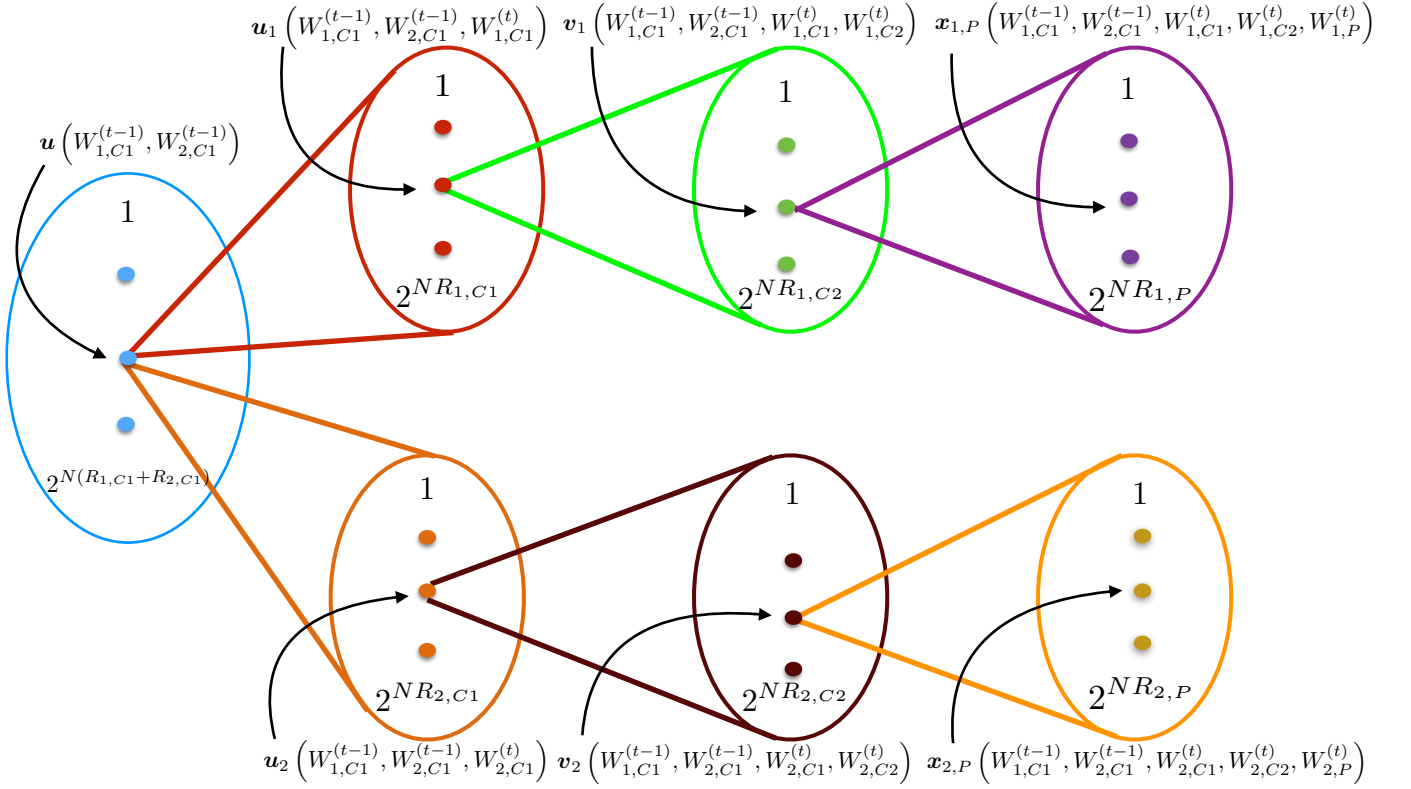
Fig. 10. Structure of the superposition code. The codewords corresponding to the message indices $W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)}, W_{i,P}^{(t)}$ with $i \in \{1,2\}$ as well as the block index t are both highlighted. The (approximate) number of codewords for each code layer is also highlighted.

$$
\begin{aligned}
&+2^{N(R_{2,C}+R_{1,C2}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{2,C1}+R_{1,C2}+R_{1,P}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{2,C}+R_{1,C2}+R_{1,P}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C1}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C1}+R_{2,C2}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C1}+R_{1,P}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C1}+R_{1,P}+R_{2,C2}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C}+R_{2,C2}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_1-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_1+R_{2,C2}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C1}+R_{2,C1}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C1}+R_{2,C}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C1}+R_{2,C2}+R_{1,P}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C1}+R_{2,C}+R_{1,P}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C}+R_{2,C1}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C}+R_{2,C}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_1+R_{2,C1}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_1+R_{2,C}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)}. \quad\quad (50)
\end{aligned}
$$

The same analysis of the probability of error holds for transmitter-receiver pair 2. Hence, in general, from (48) and (50), reliable decoding holds under the following conditions

for transmitter $i \in \{1,2\}$, with $j \in \{1,2\} \setminus \{i\}$:

$$
\begin{aligned}
R_{j,C1} &\leqslant I\left(\overleftarrow{Y}_i; U_j | U, U_i, V_i, X_i\right) \\
&= I\left(\overleftarrow{Y}_i; U_j | U, X_i\right) \\
&\triangleq \theta_{1,i}, \quad\quad (51a) \\
R_i + R_{j,C} &\leqslant I(\overrightarrow{Y}_i; U, U_i, U_j, V_i, V_j, X_i) \\
&= I(\overrightarrow{Y}_i; U, U_j, V_j, X_i) \\
&\triangleq \theta_{2,i}, \quad\quad (51b) \\
R_{j,C2} &\leqslant I(\overrightarrow{Y}_i; V_j | U, U_i, U_j, V_i, X_i) \\
&= I(\overrightarrow{Y}_i; V_j | U, U_j, X_i) \\
&\triangleq \theta_{3,i}, \quad\quad (51c) \\
R_{i,P} &\leqslant I(\overrightarrow{Y}_i; X_i | U, U_i, U_j, V_i, V_j) \\
&\triangleq \theta_{4,i}, \quad\quad (51d) \\
R_{i,P} + R_{j,C2} &\leqslant I(\overrightarrow{Y}_i; V_j, X_i | U, U_i, U_j, V_i) \\
&\triangleq \theta_{5,i}, \quad\quad (51e) \\
R_{i,C2} + R_{i,P} &\leqslant I(\overrightarrow{Y}_i; V_i, X_i | U, U_i, U_j, V_j) \\
&= I(\overrightarrow{Y}_i; X_i | U, U_i, U_j, V_j) \\
&\triangleq \theta_{6,i}, \text{ and} \quad\quad (51f) \\
R_{i,C2} + R_{i,P} + R_{j,C2} &\leqslant I(\overrightarrow{Y}_i; V_i, V_j, X_i | U, U_i, U_j) \\
&= I(\overrightarrow{Y}_i; V_j, X_i | U, U_i, U_j) \\
&\triangleq \theta_{7,i}. \quad\quad (51g)
\end{aligned}
$$

Taking into account that $R_i = R_{i,C1} + R_{i,C2} + R_{i,P}$, a Fourier-

Motzkin elimination process in (51) yields:

$$R_1 \leqslant \min\left(\theta_{2,1}, \theta_{6,1} + \theta_{1,2}, \theta_{4,1} + \theta_{1,2} + \theta_{3,2}\right), \quad (52\text{a})$$

$$R_2 \leqslant \min\left(\theta_{2,2}, \theta_{1,1} + a_{6,2}, \theta_{1,1} + \theta_{3,1} + \theta_{4,2}\right), \quad (52\text{b})$$

$$R_1 + R_2 \leqslant \min(\theta_{2,1} + \theta_{4,2}, \theta_{2,1} + a_{6,2}, \theta_{4,1} + \theta_{2,2},$$
$$\theta_{6,1} + \theta_{2,2}, \theta_{1,1} + \theta_{3,1} + \theta_{4,1} + \theta_{1,2} + \theta_{5,2},$$
$$\theta_{1,1} + \theta_{7,1} + \theta_{1,2} + \theta_{5,2}, \theta_{1,1} + \theta_{4,1} + \theta_{1,2} + \theta_{7,2},$$
$$\theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{3,2} + \theta_{4,2},$$
$$\theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{5,2}, \theta_{1,1} + \theta_{7,1} + \theta_{1,2} + \theta_{4,2}),$$
$$(52\text{c})$$

$$2R_1 + R_2 \leqslant \min(\theta_{2,1} + \theta_{4,1} + \theta_{1,2} + \theta_{7,2},$$
$$\theta_{1,1} + \theta_{4,1} + \theta_{7,1} + 2\theta_{1,2} + \theta_{5,2},$$
$$\theta_{2,1} + \theta_{4,1} + \theta_{1,2} + \theta_{5,2}), \quad (52\text{d})$$

$$R_1 + 2R_2 \leqslant \min(\theta_{1,1} + \theta_{5,1} + \theta_{2,2} + \theta_{4,2},$$
$$\theta_{1,1} + \theta_{7,1} + \theta_{2,2} + \theta_{4,2},$$
$$2\theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{4,2} + \theta_{7,2}), \quad (52\text{e})$$

where $\theta_{l,i}$ are defined in (51) with $(l,i) \in \{1,\ldots,7\} \times \{1,2\}$.

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

Following the discussion in Section IV-B1, consider the random variables U, U_i, V_i, and $X_{i,P}$ independent and uniformly distributed over the sets $\{0\}$, $\{0,1\}^{\left(n_{ji} - \left(\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj}\right)^+\right)^+}$, $\{0,1\}^{\left(\min\left(n_{ji},\left(\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj}\right)^+\right)\right)}$ and $\{0,1\}^{\left(\overrightarrow{n}_{ii} - n_{ji}\right)^+}$, respectively, with $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$.

Let the channel input symbol of transmitter i at each channel use be $X_i = \left(U_i^\mathsf{T}, V_i^\mathsf{T}, X_{i,P}^\mathsf{T}, (0,\ldots,0)\right)^\mathsf{T} \in \{0,1\}^q$. Considering these random variables, the following holds for the terms $\theta_{l,i}$, with $(l,i) \in \{1,\ldots,7\} \times \{1,2\}$, in (51):

$$\theta_{1,i} = I\left(\overleftarrow{Y}_i; U_j | U, X_i\right)$$
$$\overset{(a)}{=} H\left(\overleftarrow{Y}_i | U, X_i\right)$$
$$= H\left(U_j\right)$$
$$= \left(n_{ij} - \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^+\right)^+, \quad (53\text{a})$$

$$\theta_{2,i} = I\left(\overrightarrow{Y}_i; U, U_j, V_j, X_i\right)$$
$$\overset{(b)}{=} H\left(\overrightarrow{Y}_i\right)$$
$$= \max\left(\overrightarrow{n}_{ii}, n_{ij}\right), \quad (53\text{b})$$

$$\theta_{3,i} = I\left(\overrightarrow{Y}_i; V_j | U, U_j, X_i\right)$$
$$\overset{(b)}{=} H\left(\overrightarrow{Y}_i | U, U_j, X_i\right)$$
$$= H\left(V_j\right)$$
$$= \min\left(n_{ij}, \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^+\right), \quad (53\text{c})$$

$$\theta_{4,i} = I\left(\overrightarrow{Y}_i; X_i | U, U_i, U_j, V_i, V_j\right)$$
$$\overset{(b)}{=} H\left(\overrightarrow{Y}_i | U, U_i, U_j, V_i, V_j\right)$$
$$= H\left(X_{i,P}\right)$$
$$= \left(\overrightarrow{n}_{ii} - n_{ji}\right)^+, \text{ and} \quad (53\text{d})$$

$$\theta_{5,i} = I\left(\overrightarrow{Y}_i; V_j, X_i | U, U_i, U_j, V_i\right)$$
$$\overset{(b)}{=} H\left(\overrightarrow{Y}_i | U, U_i, U_j, V_i\right)$$
$$= \max\left(\dim X_{i,P}, \dim V_j\right)$$
$$= \max\left(\left(\overrightarrow{n}_{ii} - n_{ji}\right)^+,\right.$$
$$\left. \min\left(n_{ij}, \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^+\right)\right), \quad (53\text{e})$$

where (a) follows from the fact that $H\left(\overleftarrow{Y}_i | U, U_j, X_i\right) = 0$; and (b) follows from the fact that $H(\overrightarrow{Y}_i | U, U_j, V_j, X_i) = 0$.

For the calculation of the last two mutual information terms in inequalities (51f) and (51g), special notation is used. Let the vector V_i be the concatenation of the vectors $X_{i,HA}$ and $X_{i,HB}$, i.e., $V_i = \left(X_{i,HA}^\mathsf{T}, X_{i,HB}^\mathsf{T}\right)^\mathsf{T}$. The vector $X_{i,HA}$ contains the entries of V_i that are available in both receivers. The vector $X_{i,HB}$ is the part of V_i that is exclusively available in receiver j (see Figure 4). Note also that the vectors $X_{i,HA}$ and $X_{i,HB}$ possess the following dimensions:

$$\dim X_{i,HA} = \min\left(n_{ji}, \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+\right)$$
$$- \min\left(\left(n_{ji} - \overrightarrow{n}_{ii}\right)^+, \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+\right)$$

$$\dim X_{i,HB} = \min\left(\left(n_{ji} - \overrightarrow{n}_{ii}\right)^+, \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+\right).$$

Using this notation, the following holds:

$$\theta_{6,i} = I\left(\overrightarrow{Y}_i; X_i | U, U_i, U_j, V_j\right)$$
$$\overset{(c)}{=} H\left(\overrightarrow{Y}_i | U, U_i, U_j, V_j\right)$$
$$= H\left(X_{i,HA}, X_{i,P}\right)$$
$$= \dim X_{i,HA} + \dim X_{i,P}$$
$$= \min\left(n_{ji}, \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+\right)$$
$$- \min\left(\left(n_{ji} - \overrightarrow{n}_{ii}\right)^+, \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+\right)$$
$$+ \left(\overrightarrow{n}_{ii} - n_{ji}\right)^+, \text{ and} \quad (53\text{f})$$

$$\theta_{7,i} = I\left(\overrightarrow{Y}_i; V_j, X_i | U, U_i, U_j\right)$$
$$= I\left(\overrightarrow{Y}_i; X_i | U, U_i, U_j\right) + I\left(\overrightarrow{Y}_i; V_j | U, U_i, U_j, X_i\right)$$
$$= I\left(\overrightarrow{Y}_i; X_i | U, U_i, U_j\right) + I\left(\overrightarrow{Y}_i; V_j | U, U_j, X_i\right)$$
$$\overset{(c)}{=} H\left(\overrightarrow{Y}_i | U, U_i, U_j\right)$$
$$= \max\left(H\left(V_j\right), H\left(X_{i,HA}\right) + H\left(X_{i,P}\right)\right)$$
$$= \max\left(\dim V_j, \dim X_{i,HA} + \dim X_{i,P}\right)$$
$$= \max\left(\min\left(n_{ij}, \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^+\right),\right.$$
$$\min\left(n_{ji}, \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+\right)$$
$$- \min\left(\left(n_{ji} - \overrightarrow{n}_{ii}\right)^+, \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+\right)$$
$$\left. + \left(\overrightarrow{n}_{ii} - n_{ji}\right)^+\right), \quad (53\text{g})$$

where (c) follows from the fact that $H(\overrightarrow{Y}_i | U, U_j, V_j, X_i) = 0$.

Plugging (53) into (52) (after some algebraic manipulation) yields the system of inequalities in Theorem 1.

The sum-rate bound in (52c) can be simplified as follows:

$$R_1 + R_2 \leqslant \min\Big(\theta_{2,1} + \theta_{4,2}, \theta_{4,1} + \theta_{2,2},$$
$$\theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{5,2}\Big). \qquad (54)$$

Note that this follows from the fact that

$$\max\Big(\theta_{2,1} + \theta_{4,2}, \theta_{4,1} + \theta_{2,2}, \theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{5,2}\Big) \leqslant$$
$$\min\Big(\theta_{2,1} + a_{6,2}, \theta_{6,1} + \theta_{2,2}, \theta_{1,1} + \theta_{3,1} + \theta_{4,1} + \theta_{1,2} + \theta_{5,2},$$
$$\theta_{1,1} + \theta_{7,1} + \theta_{1,2} + \theta_{5,2}, \theta_{1,1} + \theta_{4,1} + \theta_{1,2} + \theta_{7,2},$$
$$\theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{3,2} + \theta_{4,2}, \theta_{1,1} + \theta_{7,1} + \theta_{1,2} + \theta_{4,2}\Big). \qquad (55)$$

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

Consider that transmitter i uses the following channel input:

$$X_i = U + U_i + V_i + X_{i,P}, \qquad (56)$$

where U, U_1, U_2, V_1, V_2, $X_{1,P}$, and $X_{2,P}$ in (56) are mutually independent and distributed as follows:

$$U \sim \mathcal{N}(0, \rho), \qquad (57a)$$
$$U_i \sim \mathcal{N}(0, \mu_i \lambda_{i,C}), \qquad (57b)$$
$$V_i \sim \mathcal{N}(0, (1 - \mu_i)\lambda_{i,C}), \qquad (57c)$$
$$X_{i,P} \sim \mathcal{N}(0, \lambda_{i,P}), \qquad (57d)$$

with

$$\rho + \lambda_{i,C} + \lambda_{i,P} = 1 \text{ and} \qquad (58a)$$

$$\lambda_{i,P} = \min\Big(\frac{1}{\mathrm{INR}_{ji}}, 1\Big), \qquad (58b)$$

where $\mu_i \in [0, 1]$ and $\rho \in \Big[0, \big(1 - \max\big(\frac{1}{\mathrm{INR}_{12}}, \frac{1}{\mathrm{INR}_{21}}\big)\big)^+\Big]$.

The parameters ρ, μ_i, and $\lambda_{i,P}$ define a particular coding scheme for transmitter i. The assignment in (58b) 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 (8), (9), and (56), the right-hand side of the inequalities in (51) can be written in terms of $\overrightarrow{\mathrm{SNR}}_1$, $\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, INR_{21}, $\overleftarrow{\mathrm{SNR}}_1$, $\overleftarrow{\mathrm{SNR}}_2$, ρ, μ_1, and μ_2 as follows:

$$\theta_{1,i} = I\Big(\overleftarrow{Y}_i; U_j | U, X_i\Big)$$
$$= \frac{1}{2}\log\left(\frac{\overleftarrow{\mathrm{SNR}}_i\big(b_{2,i}(\rho) + 2\big) + b_{1,i}(1) + 1}{\overleftarrow{\mathrm{SNR}}_i\big((1 - \mu_j)b_{2,i}(\rho) + 2\big) + b_{1,i}(1) + 1}\right)$$
$$= a_{3,i}(\rho, \mu_j), \qquad (59a)$$
$$\theta_{2,i} = I\Big(\overrightarrow{Y}_i; U, U_j, V_j, X_i\Big)$$
$$= \frac{1}{2}\log\Big(b_{1,i}(\rho) + 1\Big) - \frac{1}{2}$$
$$= a_{2,i}(\rho), \qquad (59b)$$

$$\theta_{3,i} = I\Big(\overrightarrow{Y}_i; V_j | U, U_j, X_i\Big)$$
$$= \frac{1}{2}\log\Big(\big(1 - \mu_j\big)b_{2,i}(\rho) + 2\Big) - \frac{1}{2}$$
$$= a_{4,i}(\rho, \mu_j), \qquad (59c)$$
$$\theta_{4,i} = I\Big(\overrightarrow{Y}_i; X_i | U, U_i, U_j, V_i, V_j\Big)$$
$$= \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}} + 2\right) - \frac{1}{2}$$
$$= a_{1,i}, \qquad (59d)$$
$$\theta_{5,i} = I\Big(\overrightarrow{Y}_i; V_j, X_i | U, U_i, U_j, V_i\Big)$$
$$= \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}} + \big(1 - \mu_j\big)b_{2,i}(\rho)\right) - \frac{1}{2}$$
$$= a_{5,i}(\rho, \mu_j), \qquad (59e)$$
$$\theta_{6,i} = I\Big(\overrightarrow{Y}_i; X_i | U, U_i, U_j, V_j\Big)$$
$$= \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\Big(\big(1 - \mu_i\big)b_{2,j}(\rho) + 1\Big) + 2\right) - \frac{1}{2}$$
$$= a_{6,i}(\rho, \mu_i), \text{ and} \qquad (59f)$$
$$\theta_{7,i} = I\Big(\overrightarrow{Y}_i; V_j, X_i | U, U_i, U_j\Big)$$
$$= \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\Big(\big(1 - \mu_i\big)b_{2,j}(\rho) + 1\Big) + \big(1 - \mu_j\big)b_{2,i}(\rho) + 2\right)$$
$$- \frac{1}{2}$$
$$= a_{7,i}(\rho, \mu_1, \mu_2). \qquad (59g)$$

Finally, plugging (59) into (52) (after some algebraic manipulation) yields the system of inequalities in Theorem 2. The sum-rate bound in (52c) can be simplified to give:

$$R_1 + R_2 \leqslant \min\Big(a_{2,1}(\rho) + a_{1,2}, a_{1,1} + a_{2,2}(\rho),$$
$$a_{3,1}(\rho, \mu_2) + a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{7,2}(\rho, \mu_1, \mu_2),$$
$$a_{3,1}(\rho, \mu_2) + a_{5,1}(\rho, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{5,2}(\rho, \mu_1),$$
$$a_{3,1}(\rho, \mu_2) + a_{7,1}(\rho, \mu_1, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{1,2}\Big). \qquad (60)$$

Note that this follows from the fact that

$$\max\Big(a_{2,1}(\rho) + a_{1,2}, a_{1,1} + a_{2,2}(\rho),$$
$$a_{3,1}(\rho, \mu_2) + a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{7,2}(\rho, \mu_1, \mu_2),$$
$$a_{3,1}(\rho, \mu_2) + a_{5,1}(\rho, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{5,2}(\rho, \mu_1),$$
$$a_{3,1}(\rho, \mu_2) + a_{7,1}(\rho, \mu_1, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{1,2}\Big)$$
$$\leqslant \min\Big(a_{2,1} + a_{6,2}(\rho, \mu_2), a_{6,1}(\rho, \mu_1) + a_{2,2}(\rho),$$
$$a_{3,1}(\rho, \mu_2) + a_{4,1}(\rho, \mu_2) + a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{5,2}(\rho, \mu_1),$$
$$a_{3,1}(\rho, \mu_2) + a_{7,1}(\rho, \mu_1, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{5,2}(\rho, \mu_1),$$
$$a_{3,1}(\rho, \mu_2) + a_{5,1}(\rho, \mu_2) + a_{3,2}(\rho, \mu_1) + \theta_{3,2} + a_{1,2}\Big). \qquad (61)$$

Therefore, the inequalities in (52) simplify into (31) and this completes the proof of Theorem 2.

APPENDIX B
CONVERSE PROOF FOR THEOREM 1

This appendix provides the proof of converse for Theorem 1. The proof of inequalities (20a) and (20c) is presented in [3]. The rest of this appendix provides a proof of the inequalities (20b), (20d) and (20e).

Notation. For all $i \in \{1, 2\}$, the channel input $\boldsymbol{X}_{i,n}$ of the LD-IC-NF in (18) for any channel use $n \in \{1, 2, \ldots, N\}$ is a q-dimensional vector, with q in (27), that can be written as the concatenation of four vectors: $\boldsymbol{X}_{i,C,n}$, $\boldsymbol{X}_{i,P,n}$, $\boldsymbol{X}_{i,D,n}$, and $\boldsymbol{X}_{i,Q,n}$, i.e., $\boldsymbol{X}_{i,n} = \left(\boldsymbol{X}_{i,C,n}^{\mathsf{T}}, \boldsymbol{X}_{i,P,n}^{\mathsf{T}}, \boldsymbol{X}_{i,D,n}^{\mathsf{T}}, \boldsymbol{X}_{i,Q,n}^{\mathsf{T}}\right)^{\mathsf{T}}$, as shown in Figure 9. Note that this notation is independent of the feedback parameters $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$, and it holds for all $n \in \{1, 2, \ldots, N\}$. More specifically,

$\boldsymbol{X}_{i,C,n}$ represents the bits of $\boldsymbol{X}_{i,n}$ that are observed by both receivers. Then,

$$\dim \boldsymbol{X}_{i,C,n} = \min\left(\overrightarrow{n}_{ii}, n_{ji}\right), \tag{62a}$$

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

$$\dim \boldsymbol{X}_{i,P,n} = (\overrightarrow{n}_{ii} - n_{ji})^{+}, \tag{62b}$$

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

$$\dim \boldsymbol{X}_{i,D,n} = (n_{ji} - \overrightarrow{n}_{ii})^{+}, \text{ and} \tag{62c}$$

$\boldsymbol{X}_{i,Q,n} = (0, \ldots, 0)^{\mathsf{T}}$ is included for dimensional matching of the model in (19). Then,

$$\dim \boldsymbol{X}_{i,Q,n} = q - \max\left(\overrightarrow{n}_{ii}, n_{ji}\right). \tag{62d}$$

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

$$\begin{aligned}
H\left(\boldsymbol{X}_{i,n}\right) &= H\left(\boldsymbol{X}_{i,C,n}, \boldsymbol{X}_{i,P,n}, \boldsymbol{X}_{i,D,n}, \boldsymbol{X}_{i,Q,n}\right) \\
&= H\left(\boldsymbol{X}_{i,C,n}, \boldsymbol{X}_{i,P,n}, \boldsymbol{X}_{i,D,n}\right) \\
&\leqslant \dim \boldsymbol{X}_{i,C,n} + \dim \boldsymbol{X}_{i,P,n} + \dim \boldsymbol{X}_{i,D,n}. \tag{62e}
\end{aligned}$$

Note that vectors $\boldsymbol{X}_{i,P,n}$ and $\boldsymbol{X}_{i,D,n}$ do not exist simultaneously. The former exists when $\overrightarrow{n}_{ii} > n_{ji}$, while the latter exists when $\overrightarrow{n}_{ii} < n_{ji}$. Let $\boldsymbol{X}_{i,D,n}$ be written in terms of $\boldsymbol{X}_{i,DF,n}$ and $\boldsymbol{X}_{i,DG,n}$, i.e., $\boldsymbol{X}_{i,D,n} = \left(\boldsymbol{X}_{i,DF,n}^{\mathsf{T}}, \boldsymbol{X}_{i,DG,n}^{\mathsf{T}}\right)^{\mathsf{T}}$. The vector $\boldsymbol{X}_{i,DF,n}$ represents the bits of $\boldsymbol{X}_{i,D,n}$ that are above the noise level in the feedback link from receiver j to transmitter j; and $\boldsymbol{X}_{i,DG,n}$ represents the bits of $\boldsymbol{X}_{i,D,n}$ that are below the noise level in the feedback link from receiver j to transmitter j, as shown in Figure 9. The dimension of vectors $\boldsymbol{X}_{i,DF,n}$ and $\boldsymbol{X}_{i,DG,n}$ are given by

$$\begin{aligned}
\dim \boldsymbol{X}_{i,DF,n} = \min\Bigg(&(n_{ji} - \overrightarrow{n}_{ii})^{+}, \Big(\overleftarrow{n}_{jj} - \overrightarrow{n}_{ii} \\
&- \min\left((\overrightarrow{n}_{jj} - n_{ji})^{+}, n_{ij}\right) \\
&- \left((\overrightarrow{n}_{jj} - n_{ij})^{+} - n_{ji}\right)^{+}\Big)^{+}\Bigg) \text{ and} \tag{63a}
\end{aligned}$$

$$\dim \boldsymbol{X}_{i,DG,n} = \dim \boldsymbol{X}_{i,D,n} - \dim \boldsymbol{X}_{i,DF,n}. \tag{63b}$$

Let $\boldsymbol{X}_{i,C,n}$ be written in terms of $\boldsymbol{X}_{i,CF,n}$ and $\boldsymbol{X}_{i,CG,n}$, i.e., $\boldsymbol{X}_{i,C,n} = \left(\boldsymbol{X}_{i,CF,n}^{\mathsf{T}}, \boldsymbol{X}_{i,CG,n}^{\mathsf{T}}\right)^{\mathsf{T}}$. The vector $\boldsymbol{X}_{i,CF,n}$ represents the bits of $\boldsymbol{X}_{i,C,n}$ that are above the noise level in the feedback link from receiver j to transmitter j; and $\boldsymbol{X}_{i,CG,n}$ represents the bits of $\boldsymbol{X}_{i,C,n}$ that are below the noise level in the feedback link from receiver j to transmitter j, as shown in Figure 9. Define the dimension of the vector $\left(\boldsymbol{X}_{i,CF,n}^{\mathsf{T}}, \boldsymbol{X}_{i,DF,n}^{\mathsf{T}}\right)^{\mathsf{T}}$ as follows:

$$\begin{aligned}
\dim\left(\left(\boldsymbol{X}_{i,CF,n}^{\mathsf{T}}, \boldsymbol{X}_{i,DF,n}^{\mathsf{T}}\right)^{\mathsf{T}}\right) = \Big(&\min\left(\overleftarrow{n}_{jj}, \max\left(\overrightarrow{n}_{jj}, n_{ji}\right)\right) \\
&- (\overrightarrow{n}_{jj} - n_{ji})^{+}\Big)^{+}. \tag{64}
\end{aligned}$$

The dimension of vectors $\boldsymbol{X}_{i,CF,n}$ and $\boldsymbol{X}_{i,CG,n}$ can be obtained as follows:

$$\dim \boldsymbol{X}_{i,CF,n} = \dim\left(\left(\boldsymbol{X}_{i,CF,n}^{\mathsf{T}}, \boldsymbol{X}_{i,DF,n}^{\mathsf{T}}\right)^{\mathsf{T}}\right) - \dim \boldsymbol{X}_{i,DF,n} \tag{65a}$$

and

$$\dim \boldsymbol{X}_{i,CG,n} = \dim \boldsymbol{X}_{i,C,n} - \dim \boldsymbol{X}_{i,CF,n}. \tag{65b}$$

The vector $\boldsymbol{X}_{i,U,n}$ is used to represent the bits of vector $\boldsymbol{X}_{i,n}$ that are observed at receiver j without producing any interference with bits in $\boldsymbol{X}_{j,P,n}$. An example is shown in Figure 11.

Based on its definition, the vector $\boldsymbol{X}_{i,U,n}$ consists of the top

$$\begin{aligned}
\dim \boldsymbol{X}_{i,U,n} = &\min\left(\overrightarrow{n}_{jj}, n_{ij}\right) - \min\left((\overrightarrow{n}_{jj} - n_{ji})^{+}, n_{ij}\right) \\
&+ (n_{ji} - \overrightarrow{n}_{jj})^{+} \tag{66}
\end{aligned}$$

bits of $\boldsymbol{X}_{i,n}$.

Finally, for all $i \in \{1, 2\}$, with $j \in \{1, 2\} \setminus \{i\}$, the channel output $\overrightarrow{\boldsymbol{Y}}_{i,n}$ can be written as the concatenation of three vectors: $\overrightarrow{\boldsymbol{Y}}_{i,Q,n}$, $\overleftarrow{\boldsymbol{Y}}_{i,n}$, and $\overrightarrow{\boldsymbol{Y}}_{i,G,n}$, i.e., $\overrightarrow{\boldsymbol{Y}}_{i,n} = \left(\overrightarrow{\boldsymbol{Y}}_{i,Q,n}^{\mathsf{T}}, \overleftarrow{\boldsymbol{Y}}_{i,n}^{\mathsf{T}}, \overrightarrow{\boldsymbol{Y}}_{i,G,n}^{\mathsf{T}}\right)^{\mathsf{T}}$, as shown in Figure 9. More specifically, the vector $\overleftarrow{\boldsymbol{Y}}_{i,n}$ contains the bits that are above the noise level in the feedback link from receiver i to transmitter i. Then,

$$\dim \overleftarrow{\boldsymbol{Y}}_{i,n} = \min\left(\overleftarrow{n}_{ii}, \max\left(\overrightarrow{n}_{ii}, n_{ij}\right)\right). \tag{67a}$$

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

$$\dim \overrightarrow{\boldsymbol{Y}}_{i,G,n} = \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^{+}. \tag{67b}$$

The vector $\overrightarrow{\boldsymbol{Y}}_{i,Q,n} = (0, \ldots, 0)$ is included for dimensional matching with the model in (19). Then,

$$\begin{aligned}
H\left(\overrightarrow{\boldsymbol{Y}}_{i,n}\right) &= H\left(\overrightarrow{\boldsymbol{Y}}_{i,Q,n}, \overleftarrow{\boldsymbol{Y}}_{i,n}, \overrightarrow{\boldsymbol{Y}}_{i,G,n}\right) \\
&= H\left(\overleftarrow{\boldsymbol{Y}}_{i,n}, \overrightarrow{\boldsymbol{Y}}_{i,G,n}\right) \\
&\leqslant \dim \overleftarrow{\boldsymbol{Y}}_{i,n} + \dim \overrightarrow{\boldsymbol{Y}}_{i,G,n}. \tag{67c}
\end{aligned}$$

Using this notation, the proof continues as follows.

Proof of (20b): First, consider $n_{ji} \leqslant \overrightarrow{n}_{ii}$, i.e., vector $\boldsymbol{X}_{i,P,n}$ exists and vector $\boldsymbol{X}_{i,D,n}$ does not exist. 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 $k \in \{1, 2, \ldots, N\}$:

$$
\begin{aligned}
NR_i &= H(W_i) \\
&\overset{(a)}{=} H(W_i|W_j) \\
&\overset{(b)}{\leqslant} I\left(W_i; \overrightarrow{\boldsymbol{Y}}_i, \overleftarrow{\boldsymbol{Y}}_j | W_j\right) + N\delta(N) \\
&= H\left(\overrightarrow{\boldsymbol{Y}}_i, \overleftarrow{\boldsymbol{Y}}_j | W_j\right) + N\delta(N) \\
&\overset{(c)}{=} \sum_{n=1}^{N} H\left(\overrightarrow{\boldsymbol{Y}}_{i,n}, \overleftarrow{\boldsymbol{Y}}_{j,n} | W_j, \overrightarrow{\boldsymbol{Y}}_{i,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}, \boldsymbol{X}_{j,n}\right) \\
&\quad + N\delta(N) \\
&\leqslant \sum_{n=1}^{N} H\left(\boldsymbol{X}_{i,n}, \overleftarrow{\boldsymbol{Y}}_{j,n} | \boldsymbol{X}_{j,n}\right) + N\delta(N) \\
&\leqslant \sum_{n=1}^{N} H(\boldsymbol{X}_{i,n}) + N\delta(N) \\
&= NH(\boldsymbol{X}_{i,k}) + N\delta(N), \\
&\leqslant N\left(\dim \boldsymbol{X}_{i,C,k} + \dim \boldsymbol{X}_{i,P,k}\right) + N\delta(N). \quad (68)
\end{aligned}
$$

Second, consider the case in which $n_{ji} > \overrightarrow{n}_{ii}$. In this case the vector $\boldsymbol{X}_{i,P,n}$ does not exist and the vector $\boldsymbol{X}_{i,D,n}$ exists. 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, hence the following holds for any $k \in \{1, 2, \ldots, N\}$:

$$
\begin{aligned}
NR_i &= H(W_i) \\
&\overset{(a)}{=} H(W_i|W_j) \\
&\overset{(b)}{\leqslant} I\left(W_i; \overrightarrow{\boldsymbol{Y}}_i, \overleftarrow{\boldsymbol{Y}}_j | W_j\right) + N\delta(N) \\
&= H\left(\overrightarrow{\boldsymbol{Y}}_i, \overleftarrow{\boldsymbol{Y}}_j | W_j\right) + N\delta(N) \\
&\overset{(c)}{=} \sum_{n=1}^{N} H\left(\overrightarrow{\boldsymbol{Y}}_{i,n}, \overleftarrow{\boldsymbol{Y}}_{j,n} | W_j, \overrightarrow{\boldsymbol{Y}}_{i,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}, \boldsymbol{X}_{j,n}\right) \\
&\quad + N\delta(N) \\
&\leqslant \sum_{n=1}^{N} H\left(\boldsymbol{X}_{i,C,n}, \boldsymbol{X}_{i,CF,n}, \boldsymbol{X}_{i,DF,n}\right) + N\delta(N) \\
&= \sum_{n=1}^{N} H\left(\boldsymbol{X}_{i,C,n}, \boldsymbol{X}_{i,DF,n}\right) + N\delta(N) \\
&= NH\left(\boldsymbol{X}_{i,C,k}, \boldsymbol{X}_{i,DF,k}\right) + N\delta(N) \\
&\leqslant N\left(\dim \boldsymbol{X}_{i,C,k} + \dim \boldsymbol{X}_{i,DF,k}\right) + N\delta(N). \quad (69)
\end{aligned}
$$

The inequalities (a), (b), and (c) in (68) and (69) are justified in accordance with: (a) follows from the fact that W_1 and W_2 are independent; (b) follows from considering enhanced receivers (see Figure 6(a) for the case $i = 1$) and Fano's inequality; and (c) follows from the fact that $\boldsymbol{X}_{j,n} = f_j^{(n)}\left(W_j, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}\right)$.

Then, (68) and (69) can be expressed as one inequality in the asymptotic regime, as follows:

$$
R_i \leqslant \dim \boldsymbol{X}_{i,C,k} + \dim \boldsymbol{X}_{i,P,k} + \dim \boldsymbol{X}_{i,DF,k}, \quad (70)
$$

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

Plugging (62a), (62b), and (63a) in (70), and after some algebraic manipulation, the following holds:

$$
R_i \leqslant \min\left(\max(\overrightarrow{n}_{ii}, n_{ji}), \max\left(\overrightarrow{n}_{ii}, \overleftarrow{n}_{jj} - (\overrightarrow{n}_{jj} - n_{ji})^+\right)\right).
$$

This completes the proof of (20b).

Proof of (20d): 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 $k \in \{1, 2, \ldots, N\}$:

$$
\begin{aligned}
N(R_1 + R_2) &= H(W_1) + H(W_2) \\
&\overset{(a)}{\leqslant} I\left(W_1; \overrightarrow{\boldsymbol{Y}}_1, \overleftarrow{\boldsymbol{Y}}_1\right) + I\left(W_2; \overrightarrow{\boldsymbol{Y}}_2, \overleftarrow{\boldsymbol{Y}}_2\right) + N\delta(N) \\
&\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_1\right) - H\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right) - H\left(\boldsymbol{X}_{2,C}|W_1, \overleftarrow{\boldsymbol{Y}}_1, \boldsymbol{X}_1\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_2\right) - H\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right) \\
&\quad - H\left(\boldsymbol{X}_{1,C}|W_2, \overleftarrow{\boldsymbol{Y}}_2, \boldsymbol{X}_2\right) + N\delta(N) \\
&= H\left(\overrightarrow{\boldsymbol{Y}}_1\right) - H\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right) - H\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}|W_1, \overleftarrow{\boldsymbol{Y}}_1, \boldsymbol{X}_1\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_2\right) - H\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right) \\
&\quad - H\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}|W_2, \overleftarrow{\boldsymbol{Y}}_2, \boldsymbol{X}_2\right) + N\delta(N) \\
&= H\left(\overrightarrow{\boldsymbol{Y}}_1\right) + \left[I\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}; W_1, \overleftarrow{\boldsymbol{Y}}_1\right) - H\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}\right)\right] \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_2\right) + \left[I\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}; W_2, \overleftarrow{\boldsymbol{Y}}_2\right) - H\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}\right)\right] \\
&\quad - H\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right) - H\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right) + N\delta(N) \\
&\overset{(b)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_1|\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}\right) - H\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}|\overrightarrow{\boldsymbol{Y}}_1\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_2|\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}\right) - H\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}|\overrightarrow{\boldsymbol{Y}}_2\right) \\
&\quad + I\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}; W_1, \overleftarrow{\boldsymbol{Y}}_1\right) + I\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}; W_2, \overleftarrow{\boldsymbol{Y}}_2\right) \\
&\quad - H\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right) - H\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right) + N\delta(N) \\
&\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_1|\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_2|\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}\right) \\
&\quad + I\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}; W_1, \overleftarrow{\boldsymbol{Y}}_1\right) + I\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}; W_2, \overleftarrow{\boldsymbol{Y}}_2\right) \\
&\quad - H\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right) - H\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right) + N\delta(N) \\
&\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_1|\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_2|\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}\right) \\
&\quad + I\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}, W_2, \overleftarrow{\boldsymbol{Y}}_2; W_1, \overleftarrow{\boldsymbol{Y}}_1\right) \\
&\quad + I\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}, W_1, \overleftarrow{\boldsymbol{Y}}_1; W_2, \overleftarrow{\boldsymbol{Y}}_2\right) - H\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right) \\
&\quad - H\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right) + N\delta(N) \\
&= H\left(\overrightarrow{\boldsymbol{Y}}_1|\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_2|\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}\right) \\
&\quad + I\left(W_2; W_1, \overleftarrow{\boldsymbol{Y}}_1\right) + I\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}, \overleftarrow{\boldsymbol{Y}}_2; W_1, \overleftarrow{\boldsymbol{Y}}_1|W_2\right) \\
&\quad + I\left(W_1; W_2, \overleftarrow{\boldsymbol{Y}}_2\right) + I\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}, \overleftarrow{\boldsymbol{Y}}_1; W_2, \overleftarrow{\boldsymbol{Y}}_2|W_1\right) \\
&\quad - H\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right) - H\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right) + N\delta(N)
\end{aligned}
$$
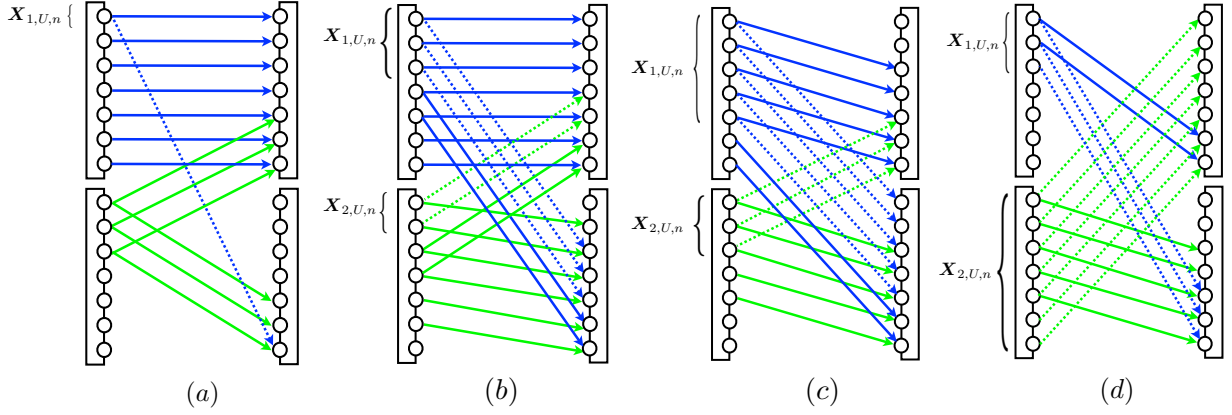
Fig. 11. Vector $\boldsymbol{X}_{i,U,n}$ in different combination of interference regimes.

$$\overset{(c)}{=}H\left(\overrightarrow{\boldsymbol{Y}}_1|\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2|\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U}\right)+H\left(W_1\right)$$
$$+H\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right)-H\left(W_1|W_2\right)-H\left(\overleftarrow{\boldsymbol{Y}}_1|W_2,W_1\right)$$
$$+H\left(\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U},\overleftarrow{\boldsymbol{Y}}_2|W_2\right)+H\left(W_2\right)+H\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right)$$
$$-H\left(W_2|W_1\right)-H\left(\overleftarrow{\boldsymbol{Y}}_2|W_1,W_2\right)$$
$$+H\left(\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U},\overleftarrow{\boldsymbol{Y}}_1|W_1\right)-H\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right)-H\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right)$$
$$+N\delta(N)$$
$$\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_1|\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2|\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U}\right)$$
$$+H\left(\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U},\overleftarrow{\boldsymbol{Y}}_2|W_2\right)$$
$$+H\left(\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U},\overleftarrow{\boldsymbol{Y}}_1|W_1\right)+N\delta(N)$$
$$=\sum_{n=1}^N\left[H\left(\overrightarrow{\boldsymbol{Y}}_{1,n}|\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U},\overrightarrow{\boldsymbol{Y}}_{1,(1:n-1)}\right)+\right.$$
$$H\left(\overrightarrow{\boldsymbol{Y}}_{2,n}|\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U},\overrightarrow{\boldsymbol{Y}}_{2,(1:n-1)}\right)+$$
$$H\left(\boldsymbol{X}_{2,C,n},\boldsymbol{X}_{1,U,n},\overleftarrow{\boldsymbol{Y}}_{2,n}|W_2,\boldsymbol{X}_{2,C,(1:n-1)},\boldsymbol{X}_{1,U,(1:n-1)},\right.$$
$$\overleftarrow{\boldsymbol{Y}}_{2,(1:n-1)}\right)+H\left(\boldsymbol{X}_{1,C,n},\boldsymbol{X}_{2,U,n},\overleftarrow{\boldsymbol{Y}}_{1,n}|W_1,\boldsymbol{X}_{1,C,(1:n-1)},\right.$$
$$\left.\left.\boldsymbol{X}_{2,U,(1:n-1)},\overleftarrow{\boldsymbol{Y}}_{1,(1:n-1)}\right)\right]+N\delta(N)$$
$$\overset{(d)}{=}\sum_{n=1}^N\left[H\left(\overrightarrow{\boldsymbol{Y}}_{1,n}|\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U},\overrightarrow{\boldsymbol{Y}}_{1,(1:n-1)}\right)+\right.$$
$$H\left(\overrightarrow{\boldsymbol{Y}}_{2,n}|\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U},\overrightarrow{\boldsymbol{Y}}_{2,(1:n-1)}\right)+$$
$$H\left(\boldsymbol{X}_{2,C,n},\boldsymbol{X}_{1,U,n},\overleftarrow{\boldsymbol{Y}}_{2,n}|W_2,\boldsymbol{X}_{2,C,(1:n-1)},\boldsymbol{X}_{1,U,(1:n-1)},\right.$$
$$\overleftarrow{\boldsymbol{Y}}_{2,(1:n-1)},\boldsymbol{X}_{2,(1:n)}\right)+H\left(\boldsymbol{X}_{1,C,n},\boldsymbol{X}_{2,U,n},\overleftarrow{\boldsymbol{Y}}_{1,n}|W_1,\right.$$
$$\left.\left.\boldsymbol{X}_{1,C,(1:n-1)},\boldsymbol{X}_{2,U,(1:n-1)},\overleftarrow{\boldsymbol{Y}}_{1,(1:n-1)},\boldsymbol{X}_{1,(1:n)}\right)\right]$$
$$+N\delta(N)$$
$$\overset{(e)}{\leqslant}\sum_{n=1}^N\left[H\left(\overrightarrow{\boldsymbol{Y}}_{1,n}|\boldsymbol{X}_{1,C,n},\boldsymbol{X}_{2,U,n}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_{2,n}|\boldsymbol{X}_{2,C,n},\boldsymbol{X}_{1,U,n}\right)\right.$$
$$\left.+H\left(\boldsymbol{X}_{1,U,n},\overleftarrow{\boldsymbol{Y}}_{2,n}|\boldsymbol{X}_{2,n}\right)+H\left(\boldsymbol{X}_{2,U,n},\overleftarrow{\boldsymbol{Y}}_{1,n}|\boldsymbol{X}_{1,n}\right)\right]$$
$$+N\delta(N)$$
$$\leqslant\sum_{n=1}^N\left[H\left(\boldsymbol{X}_{1,P,n}\right)+H\left(\boldsymbol{X}_{2,P,n}\right)+H\left(\boldsymbol{X}_{1,U,n},\overleftarrow{\boldsymbol{Y}}_{2,n}|\boldsymbol{X}_{2,n}\right)\right.$$

$$+H\left(\boldsymbol{X}_{2,U,n},\overleftarrow{\boldsymbol{Y}}_{1,n}|\boldsymbol{X}_{1,n}\right)\right]+N\delta(N)$$
$$\overset{(e)}{\leqslant}N\left[H\left(\boldsymbol{X}_{1,P,k}\right)+H\left(\boldsymbol{X}_{2,P,k}\right)+H\left(\boldsymbol{X}_{1,U,k}\right)\right.$$
$$+H\left(\overleftarrow{\boldsymbol{Y}}_{2,k}|\boldsymbol{X}_{2,k},\boldsymbol{X}_{1,U,k}\right)+H\left(\boldsymbol{X}_{2,U,k}\right)$$
$$\left.+H\left(\overleftarrow{\boldsymbol{Y}}_{1,k}|\boldsymbol{X}_{1,k},\boldsymbol{X}_{2,U,k}\right)\right]+N\delta(N),$$
$$=N\left[H\left(\boldsymbol{X}_{1,P,k}\right)+H\left(\boldsymbol{X}_{2,P,k}\right)+H\left(\boldsymbol{X}_{1,U,k}\right)\right.$$
$$+H\left(\boldsymbol{X}_{1,CF,k},\boldsymbol{X}_{1,DF,k}|\boldsymbol{X}_{2,k},\boldsymbol{X}_{1,U,k}\right)+H\left(\boldsymbol{X}_{2,U,k}\right)$$
$$\left.+H\left(\boldsymbol{X}_{2,CF,k},\boldsymbol{X}_{2,DF,k}|\boldsymbol{X}_{1,k},\boldsymbol{X}_{2,U,k}\right)\right]+N\delta(N)$$
$$\leqslant N\left[H\left(\boldsymbol{X}_{1,P,k}\right)+H\left(\boldsymbol{X}_{2,P,k}\right)+H\left(\boldsymbol{X}_{1,U,k}\right)\right.$$
$$+H\left(\boldsymbol{X}_{1,CF,k},\boldsymbol{X}_{1,DF,k}|\boldsymbol{X}_{1,U,k}\right)+H\left(\boldsymbol{X}_{2,U,k}\right)$$
$$\left.+H\left(\boldsymbol{X}_{2,CF,k},\boldsymbol{X}_{2,DF,k}|\boldsymbol{X}_{2,U,k}\right)\right]+N\delta(N)$$
$$\leqslant N\left[\dim\boldsymbol{X}_{1,P,k}+\dim\boldsymbol{X}_{2,P,k}+\dim\boldsymbol{X}_{1,U,k}\right.$$
$$+\left(\dim(\boldsymbol{X}_{1,CF,k},\boldsymbol{X}_{1,DF,k})-\dim\boldsymbol{X}_{1,U,k}\right)^++\dim\boldsymbol{X}_{2,U,k}$$
$$\left.+\left(\dim\left(\boldsymbol{X}_{2,CF,k},\boldsymbol{X}_{2,DF,k}\right)-\dim\boldsymbol{X}_{2,U,k}\right)^++\delta(N)\right].$$
$$\tag{71}$$

where, (a) follows from considering enhanced receivers (see Figure 6(b)) and 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 $H\left(\boldsymbol{X}_{i,C},\boldsymbol{X}_{j,U},\overleftarrow{\boldsymbol{Y}}_i|W_i,W_j,\overleftarrow{\boldsymbol{Y}}_j\right)=0$; (d) follows from the fact that $\boldsymbol{X}_{i,n}=f_i^{(n)}\left(W_i,\overleftarrow{\boldsymbol{Y}}_{i,(1:n-1)}\right)$; and (e) follows from the fact that conditioning does not increase entropy. Plugging (62b), (64), and (66) in (71) and after some algebraic manipulation, the following holds in the asymptotic regime:

$$R_1+R_2\leqslant\max\left(\left(\overrightarrow{n}_{11}-n_{12}\right)^+,n_{21},\right.$$
$$\overrightarrow{n}_{11}-\left(\max\left(\overrightarrow{n}_{11},n_{12}\right)-\overleftarrow{n}_{11}\right)^++$$
$$\left.\max\left(\left(\overrightarrow{n}_{22}-n_{21}\right)^+,n_{12},\overrightarrow{n}_{22}-\left(\max\left(\overrightarrow{n}_{22},n_{21}\right)-\overleftarrow{n}_{22}\right)^+\right).\right.$$

This completes the proof of (20d).

Proof of (20e)**:** 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, it follows that for all $i \in \{1,2\}$, with $j \in \{1,2\} \setminus \{i\}$ and for all $k \in \{1,2,\ldots,N\}$:

$$N\big(2R_i + R_j\big) = 2H(W_i) + H(W_j)$$

$$\stackrel{(a)}{\leqslant} I\left(W_i; \overrightarrow{\boldsymbol{Y}}_i, \overleftarrow{\boldsymbol{Y}}_i\right) + I\left(W_i; \overrightarrow{\boldsymbol{Y}}_i, \overleftarrow{\boldsymbol{Y}}_j | W_j\right)$$
$$+ I\left(W_j; \overrightarrow{\boldsymbol{Y}}_j, \overleftarrow{\boldsymbol{Y}}_j\right) + N\delta(N)$$

$$\stackrel{(b)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i | W_i\right) - H\left(\overrightarrow{\boldsymbol{Y}}_i | W_i, \overleftarrow{\boldsymbol{Y}}_i\right)$$
$$+ H\left(\overrightarrow{\boldsymbol{Y}}_i | W_j, \overleftarrow{\boldsymbol{Y}}_j\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j\right) - H\left(\overrightarrow{\boldsymbol{Y}}_j | W_j, \overleftarrow{\boldsymbol{Y}}_j\right)$$
$$+ N\delta(N)$$

$$= H\left(\overrightarrow{\boldsymbol{Y}}_i\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i | W_i\right) - H\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{j,D} | W_i, \overleftarrow{\boldsymbol{Y}}_i\right)$$
$$+ H\left(\overrightarrow{\boldsymbol{Y}}_i | W_j, \overleftarrow{\boldsymbol{Y}}_j\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j\right)$$
$$- H\left(\boldsymbol{X}_{i,C}, \boldsymbol{X}_{i,D} | W_j, \overleftarrow{\boldsymbol{Y}}_j\right) + N\delta(N)$$

$$\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_i\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i | W_i\right) - H\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U} | W_i, \overleftarrow{\boldsymbol{Y}}_i\right)$$
$$+ H\left(\overrightarrow{\boldsymbol{Y}}_i | W_j, \overleftarrow{\boldsymbol{Y}}_j\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j\right)$$
$$- H\left(\boldsymbol{X}_{i,C} | W_j, \overleftarrow{\boldsymbol{Y}}_j\right) + N\delta(N)$$

$$\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_i\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i | W_i\right) + \Big[I\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}; W_i, \overleftarrow{\boldsymbol{Y}}_i\right)$$
$$- H\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) \Big] + H\left(\overrightarrow{\boldsymbol{Y}}_i, \boldsymbol{X}_{i,C} | W_j, \overleftarrow{\boldsymbol{Y}}_j\right)$$
$$+ H\left(\overrightarrow{\boldsymbol{Y}}_j\right) - H\left(\boldsymbol{X}_{i,C} | W_j, \overleftarrow{\boldsymbol{Y}}_j\right) + N\delta(N)$$

$$= H\left(\overrightarrow{\boldsymbol{Y}}_i\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i | W_i\right) + \Big[I\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}; W_i, \overleftarrow{\boldsymbol{Y}}_i\right)$$
$$- H\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) \Big] + H\left(\overrightarrow{\boldsymbol{Y}}_i | W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right)$$
$$+ H\left(\overrightarrow{\boldsymbol{Y}}_j\right) + N\delta(N)$$

$$\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_i\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i | W_i\right) + \Big[I\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}; W_i, \overleftarrow{\boldsymbol{Y}}_i\right)$$
$$- H\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) \Big] + H\left(\overrightarrow{\boldsymbol{Y}}_i | W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right)$$
$$+ H\left(\overrightarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) + N\delta(N)$$

$$\stackrel{(c)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i | W_i\right) + I\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}; W_i, \overleftarrow{\boldsymbol{Y}}_i\right)$$
$$+ H\left(\overrightarrow{\boldsymbol{Y}}_i | W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j | \boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right)$$
$$+ N\delta(N)$$

$$\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_i\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i | W_i\right)$$
$$+ I\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}, W_j, \overleftarrow{\boldsymbol{Y}}_j; W_i, \overleftarrow{\boldsymbol{Y}}_i\right)$$
$$+ H\left(\overrightarrow{\boldsymbol{Y}}_i | W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j | \boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right)$$
$$+ N\delta(N)$$

$$\stackrel{(d)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i | W_j, W_i,\right)$$
$$+ H\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}, \overleftarrow{\boldsymbol{Y}}_j | W_j\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i | W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right)$$
$$+ H\left(\overrightarrow{\boldsymbol{Y}}_j | \boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) + N\delta(N)$$

$$\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_i\right) + H\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}, \overleftarrow{\boldsymbol{Y}}_j | W_j\right)$$
$$+ H\left(\overrightarrow{\boldsymbol{Y}}_i | W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j | \boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right)$$
$$+ N\delta(N)$$

$$\leqslant \sum_{n=1}^{N} \Big[H\left(\overrightarrow{\boldsymbol{Y}}_{i,n}\right) + H\Big(\boldsymbol{X}_{j,C,n}, \boldsymbol{X}_{i,U,n}, \overleftarrow{\boldsymbol{Y}}_{j,n} | W_j,$$
$$\boldsymbol{X}_{j,C,(1:n-1)}, \boldsymbol{X}_{i,U,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}\Big)$$
$$+ H\left(\overrightarrow{\boldsymbol{Y}}_{i,n} | W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}, \overrightarrow{\boldsymbol{Y}}_{i,(1:n-1)}\right)$$
$$+ H\left(\overrightarrow{\boldsymbol{Y}}_{j,n} | \boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}, \overrightarrow{\boldsymbol{Y}}_{j,(1:n-1)}\right) \Big] + N\delta(N)$$

$$= \sum_{n=1}^{N} \Big[H\left(\overrightarrow{\boldsymbol{Y}}_{i,n}\right) + H\Big(\boldsymbol{X}_{j,C,n}, \boldsymbol{X}_{i,U,n}, \overleftarrow{\boldsymbol{Y}}_{j,n} | W_j,$$
$$\boldsymbol{X}_{j,C,(1:n-1)}, \boldsymbol{X}_{i,U,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}, \boldsymbol{X}_{j,(1:n)}\Big)$$
$$+ H\left(\overrightarrow{\boldsymbol{Y}}_{i,n} | W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}, \overrightarrow{\boldsymbol{Y}}_{i,(1:n-1)}, \boldsymbol{X}_{j,(1:n)}\right)$$
$$+ H\left(\overrightarrow{\boldsymbol{Y}}_{j,n} | \boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}, \overrightarrow{\boldsymbol{Y}}_{j,(1:n-1)}\right) \Big] + N\delta(N)$$

$$\leqslant \sum_{n=1}^{N} \Big[H\left(\overrightarrow{\boldsymbol{Y}}_{i,n}\right) + H\left(\boldsymbol{X}_{i,U,n} | \boldsymbol{X}_{j,n}\right)$$
$$+ H\left(\overleftarrow{\boldsymbol{Y}}_{j,n} | \boldsymbol{X}_{j,n}, \boldsymbol{X}_{i,U,n}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_{i,n} | \boldsymbol{X}_{i,C,n}, \boldsymbol{X}_{j,n}\right)$$
$$+ H\left(\overrightarrow{\boldsymbol{Y}}_{j,n} | \boldsymbol{X}_{j,C,n}, \boldsymbol{X}_{i,U,n}\right) \Big] + N\delta(N)$$

$$\leqslant N\Big[H\left(\overrightarrow{\boldsymbol{Y}}_{i,k}\right) + H\left(\boldsymbol{X}_{i,U,k}\right) + H\left(\overleftarrow{\boldsymbol{Y}}_{j,k} | \boldsymbol{X}_{j,k}, \boldsymbol{X}_{i,U,k}\right)$$
$$+ H\left(\boldsymbol{X}_{i,P,k}\right) + H\left(\boldsymbol{X}_{j,P,k}\right) \Big] + N\delta(N)$$

$$= N\Big[H\left(\overrightarrow{\boldsymbol{Y}}_{i,k}\right) + H\left(\boldsymbol{X}_{i,U,k}\right) + H\left(\boldsymbol{X}_{i,CF,k} \boldsymbol{X}_{i,DF,k} | \boldsymbol{X}_{i,U,k}\right)$$
$$+ H\left(\boldsymbol{X}_{i,P,k}\right) + H\left(\boldsymbol{X}_{j,P,k}\right) \Big] + N\delta(N)$$

$$\leqslant N\Big[\dim \overleftarrow{\boldsymbol{Y}}_{i,k} + \dim \overrightarrow{\boldsymbol{Y}}_{i,G,k} + \dim \boldsymbol{X}_{i,U,k}$$
$$+ \left(\dim\left(\boldsymbol{X}_{i,CF,k}, \boldsymbol{X}_{i,DF,k}\right) - \dim \boldsymbol{X}_{i,U,k}\right)^+ + \dim \boldsymbol{X}_{i,P,k}$$
$$+ \dim \boldsymbol{X}_{j,P,k} \Big] + N\delta(N), \tag{72}$$

where, (a) follows from considering enhanced receivers (see Figure 6(c) for the case $i = 1$) and Fano's inequality; (b) follows from the fact that $H\left(\overleftarrow{\boldsymbol{Y}}_i, \overleftarrow{\boldsymbol{Y}}_j | W_i, W_j\right) = 0$; (c) follows from the fact that $H(Y|X) = H(X,Y) - H(X)$; and (d) follows from the fact that $H\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}, \overleftarrow{\boldsymbol{Y}}_j | W_j, W_i, \overleftarrow{\boldsymbol{Y}}_i\right) = 0$.

Plugging (62b), (64), (66), (67a), and (67b) in (72) and after some algebraic manipulation, the following holds in the asymptotic regime:

$$2R_i + R_j \leqslant \max\left(\overrightarrow{n}_{ii}, n_{ji}\right) + \left(\overrightarrow{n}_{ii} - n_{ij}\right)^+$$
$$+ \max\left(\left(\overrightarrow{n}_{jj} - n_{ji}\right)^+, n_{ij}, \overrightarrow{n}_{jj} - \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+\right).$$

This completes the proof of (20e).

APPENDIX C
PROOF OF THEOREM 3

The outer bounds (37a) and (37c) correspond to the outer bounds of the case of perfect channel-output feedback derived in [3]. The bounds (37b), (37d) and (37e) 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:

$$h_{ji,U} = \begin{cases} 0 & \text{if } (S_{1,i} \vee S_{2,i} \vee S_{3,i}) \\ \sqrt{\dfrac{\text{INR}_{ij}\text{INR}_{ji}}{\overleftarrow{\text{SNR}}_j}} & \text{if } (S_{4,i} \vee S_{5,i}), \end{cases} \quad (73)$$

where, the events $S_{1,i}$, $S_{2,i}$, $S_{3,i}$, $S_{4,i}$, and $S_{5,i}$ are defined in (32). Consider also the following signals:

$$X_{i,C,n} = \sqrt{\text{INR}_{ji}} X_{i,n} + \overrightarrow{Z}_{j,n} \text{ and} \quad (74)$$

$$X_{i,U,n} = h_{ji,U} X_{i,n} + \overrightarrow{Z}_{j,n}, \quad (75)$$

where, $X_{i,n}$ and $\overrightarrow{Z}_{j,n}$ are the channel input of transmitter i and the noise observed at receiver j during a given channel use $n \in \{1,2,\ldots,N\}$, as described by (8).

Proof of (37b): 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 $k \in \{1,2,\ldots,N\}$:

$$
\begin{aligned}
NR_i &= H(W_i) \\
&= H(W_i|W_j) \\
&\overset{(a)}{\leqslant} I\left(W_i; \overrightarrow{\boldsymbol{Y}}_i, \overleftarrow{\boldsymbol{Y}}_j | W_j\right) + N\delta(N) \\
&\leqslant \sum_{n=1}^{N} \Big[h\Big(\overrightarrow{Y}_{i,n}, \overleftarrow{Y}_{j,n} | W_j, \overrightarrow{\boldsymbol{Y}}_{i,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}, \boldsymbol{X}_{j,(1:n)}\Big) \\
&\quad - h\left(\overrightarrow{Z}_{i,n}\right) - h\left(\overleftarrow{Z}_{j,n}\right) \Big] + N\delta(N) \\
&\leqslant \sum_{n=1}^{N} \Big[h\left(\overrightarrow{Y}_{i,n}, \overleftarrow{Y}_{j,n} | X_{j,n}\right) - h\left(\overrightarrow{Z}_{i,n}\right) - h\left(\overleftarrow{Z}_{j,n}\right) \Big] \\
&\quad + N\delta(N) \\
&= N\left[h\left(\overrightarrow{Y}_{i,k}, \overleftarrow{Y}_{j,k} | X_{j,k}\right) - \log(2\pi e)\right] + N\delta(N), \quad (76)
\end{aligned}
$$

where (a) follows from considering enhanced receivers (see Figure 6(a) for the case $i=1$) and Fano's inequality.

From (76), the following holds in the asymptotic regime:

$$
\begin{aligned}
R_i &\leqslant h\left(\overrightarrow{Y}_{i,k}, \overleftarrow{Y}_{j,k} | X_{j,k}\right) - \log(2\pi e) \\
&\leqslant \frac{1}{2}\log\left(b_{3,i} + 1\right) \\
&\quad + \frac{1}{2}\log\left(\frac{\left(b_{3,i} + b_{4,j}(\rho) + 1\right)\overleftarrow{\text{SNR}}_j}{\left(b_{1,j}(\rho) + 1\right)\left(b_{3,i} + (1-\rho^2)\right)} + 1 \right). \quad (77)
\end{aligned}
$$

This completes the proof of (37b).

Proof of (37d):

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 $k \in$

$\{1,2,\ldots,N\}$:

$$
\begin{aligned}
N&\left(R_1 + R_2\right) = H(W_1) + H(W_2) \\
&\overset{(a)}{\leqslant} I\left(W_1; \overrightarrow{\boldsymbol{Y}}_1, \overleftarrow{\boldsymbol{Y}}_1\right) + I\left(W_2; \overrightarrow{\boldsymbol{Y}}_2, \overleftarrow{\boldsymbol{Y}}_2\right) + N\delta(N) \\
&= h\left(\overrightarrow{\boldsymbol{Y}}_1\right) + h\left(\overleftarrow{\boldsymbol{Z}}_1 | \overrightarrow{\boldsymbol{Y}}_1\right) - h\left(\overrightarrow{\boldsymbol{Y}}_1 | W_1\right) \\
&\quad - h\left(\overleftarrow{\boldsymbol{Y}}_1 | W_1, \overrightarrow{\boldsymbol{Y}}_1, \boldsymbol{X}_1\right) + h\left(\overrightarrow{\boldsymbol{Y}}_2\right) + h\left(\overleftarrow{\boldsymbol{Z}}_2 | \overrightarrow{\boldsymbol{Y}}_2\right) \\
&\quad - h\left(\overleftarrow{\boldsymbol{Y}}_2 | W_2\right) - h\left(\overrightarrow{\boldsymbol{Y}}_2 | W_2, \overleftarrow{\boldsymbol{Y}}_2, \boldsymbol{X}_2\right) + N\delta(N) \\
&\leqslant h\left(\overrightarrow{\boldsymbol{Y}}_1\right) + h\left(\overleftarrow{\boldsymbol{Z}}_1\right) - h\left(\overleftarrow{\boldsymbol{Y}}_1 | W_1\right) \\
&\quad - h\left(\boldsymbol{X}_{2,C} | W_1, \overleftarrow{\boldsymbol{Y}}_1, \boldsymbol{X}_1\right) + h\left(\overrightarrow{\boldsymbol{Y}}_2\right) + h\left(\overleftarrow{\boldsymbol{Z}}_2\right) \\
&\quad - h\left(\overleftarrow{\boldsymbol{Y}}_2 | W_2\right) - h\left(\boldsymbol{X}_{1,C} | W_2, \overleftarrow{\boldsymbol{Y}}_2, \boldsymbol{X}_2\right) + N\delta(N) \\
&= h\left(\overrightarrow{\boldsymbol{Y}}_1\right) - h\left(\overleftarrow{\boldsymbol{Y}}_1 | W_1\right) - h\left(\boldsymbol{X}_{2,C}, \overrightarrow{\boldsymbol{Z}}_2 | W_1, \overleftarrow{\boldsymbol{Y}}_1, \boldsymbol{X}_1\right) \\
&\quad + h\left(\overrightarrow{\boldsymbol{Z}}_2 | W_1, \overleftarrow{\boldsymbol{Y}}_1, \boldsymbol{X}_1, \boldsymbol{X}_{2,C}\right) + h\left(\overrightarrow{\boldsymbol{Y}}_2\right) - h\left(\overleftarrow{\boldsymbol{Y}}_2 | W_2\right) \\
&\quad - h\left(\boldsymbol{X}_{1,C}, \overrightarrow{\boldsymbol{Z}}_1 | W_2, \overleftarrow{\boldsymbol{Y}}_2, \boldsymbol{X}_2\right) \\
&\quad + h\left(\overrightarrow{\boldsymbol{Z}}_1 | W_2, \overleftarrow{\boldsymbol{Y}}_2, \boldsymbol{X}_2, \boldsymbol{X}_{1,C}\right) + N\log(2\pi e) + N\delta(N) \\
&= h\left(\overrightarrow{\boldsymbol{Y}}_1\right) - h\left(\overleftarrow{\boldsymbol{Y}}_1 | W_1\right) - h\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U} | W_1, \overleftarrow{\boldsymbol{Y}}_1, \boldsymbol{X}_1\right) \\
&\quad + h\left(\overrightarrow{\boldsymbol{Z}}_2 | W_1, \overleftarrow{\boldsymbol{Y}}_1, \boldsymbol{X}_1, \boldsymbol{X}_{2,C}\right) + h\left(\overrightarrow{\boldsymbol{Y}}_2\right) - h\left(\overleftarrow{\boldsymbol{Y}}_2 | W_2\right) \\
&\quad - h\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U} | W_2, \overleftarrow{\boldsymbol{Y}}_2, \boldsymbol{X}_2\right) \\
&\quad + h\left(\overrightarrow{\boldsymbol{Z}}_1 | W_2, \overleftarrow{\boldsymbol{Y}}_2, \boldsymbol{X}_2, \boldsymbol{X}_{1,C}\right) + N\log(2\pi e) + N\delta(N) \\
&= h\left(\overrightarrow{\boldsymbol{Y}}_1\right) - h\left(\overleftarrow{\boldsymbol{Y}}_1 | W_1\right) + \Big[I\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}; W_1, \overleftarrow{\boldsymbol{Y}}_1\right) \\
&\quad - h\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}\right)\Big] + h\left(\overrightarrow{\boldsymbol{Y}}_2\right) - h\left(\overleftarrow{\boldsymbol{Y}}_2 | W_2\right) \\
&\quad + \Big[I\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}; W_2, \overleftarrow{\boldsymbol{Y}}_2\right) - h\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}\right)\Big] \\
&\quad + h\left(\overrightarrow{\boldsymbol{Z}}_1 | W_2, \overleftarrow{\boldsymbol{Y}}_2, \boldsymbol{X}_2, \boldsymbol{X}_{1,C}\right) \\
&\quad + h\left(\overrightarrow{\boldsymbol{Z}}_2 | W_1, \overleftarrow{\boldsymbol{Y}}_1, \boldsymbol{X}_1, \boldsymbol{X}_{2,C}\right) + N\log(2\pi e) + N\delta(N) \\
&\leqslant h\left(\overrightarrow{\boldsymbol{Y}}_1\right) - h\left(\overleftarrow{\boldsymbol{Y}}_1 | W_1\right) + \Big[I\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}; W_1, \overleftarrow{\boldsymbol{Y}}_1\right) \\
&\quad - h\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}\right)\Big] + h\left(\overrightarrow{\boldsymbol{Y}}_2\right) - h\left(\overleftarrow{\boldsymbol{Y}}_2 | W_2\right) \\
&\quad + \Big[I\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}; W_2, \overleftarrow{\boldsymbol{Y}}_2\right) - h\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}\right)\Big] \\
&\quad + \Big[h\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U} | \overrightarrow{\boldsymbol{Y}}_2\right) - h\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U} | \overrightarrow{\boldsymbol{Y}}_2, \boldsymbol{X}_1, \boldsymbol{X}_2\right)\Big] \\
&\quad + \Big[h\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U} | \overrightarrow{\boldsymbol{Y}}_1\right) - h\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U} | \overrightarrow{\boldsymbol{Y}}_1, \boldsymbol{X}_2, \boldsymbol{X}_1\right)\Big] \\
&\quad + h\left(\overrightarrow{\boldsymbol{Z}}_1 | W_2, \overleftarrow{\boldsymbol{Y}}_2, \boldsymbol{X}_2, \boldsymbol{X}_{1,C}\right) \\
&\quad + h\left(\overrightarrow{\boldsymbol{Z}}_2 | W_1, \overleftarrow{\boldsymbol{Y}}_1, \boldsymbol{X}_1, \boldsymbol{X}_{2,C}\right) + N\log(2\pi e) + N\delta(N) \\
&\overset{(b)}{=} h\left(\overrightarrow{\boldsymbol{Y}}_1 | \boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}\right) - h\left(\overleftarrow{\boldsymbol{Y}}_1 | W_1\right) \\
&\quad + I\left(\boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}; W_1, \overleftarrow{\boldsymbol{Y}}_1\right) + h\left(\overrightarrow{\boldsymbol{Y}}_2 | \boldsymbol{X}_{2,C}, \boldsymbol{X}_{1,U}\right) \\
&\quad - h\left(\overleftarrow{\boldsymbol{Y}}_2 | W_2\right) + I\left(\boldsymbol{X}_{1,C}, \boldsymbol{X}_{2,U}; W_2, \overleftarrow{\boldsymbol{Y}}_2\right) \\
&\quad - h\left(\overrightarrow{\boldsymbol{Z}}_1, \overrightarrow{\boldsymbol{Z}}_2 | \overrightarrow{\boldsymbol{Y}}_2, \boldsymbol{X}_1, \boldsymbol{X}_2\right) - h\left(\overrightarrow{\boldsymbol{Z}}_2, \overrightarrow{\boldsymbol{Z}}_1 | \overrightarrow{\boldsymbol{Y}}_1, \boldsymbol{X}_2, \boldsymbol{X}_1\right) \\
&\quad + h\left(\overrightarrow{\boldsymbol{Z}}_1 | W_2, \overleftarrow{\boldsymbol{Y}}_2, \boldsymbol{X}_2, \boldsymbol{X}_{1,C}\right) \\
&\quad + h\left(\overrightarrow{\boldsymbol{Z}}_2 | W_1, \overleftarrow{\boldsymbol{Y}}_1, \boldsymbol{X}_1, \boldsymbol{X}_{2,C}\right) \\
&\quad + N\log(2\pi e) + N\delta(N)
\end{aligned}
$$

$$\overset{(c)}{\leqslant} h\left(\overrightarrow{\boldsymbol{Y}}_1|\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U}\right) - h\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right)$$
$$+ I\left(\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U};W_1,\overleftarrow{\boldsymbol{Y}}_1\right) + h\left(\overrightarrow{\boldsymbol{Y}}_2|\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U}\right)$$
$$- h\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right) + I\left(\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U};W_2,\overleftarrow{\boldsymbol{Y}}_2\right) + N\log\left(2\pi e\right)$$
$$+ N\delta(N)$$
$$\leqslant h\left(\overrightarrow{\boldsymbol{Y}}_1|\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U}\right) - h\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right)$$
$$+ I\left(\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U},W_2,\overleftarrow{\boldsymbol{Y}}_2;W_1,\overleftarrow{\boldsymbol{Y}}_1\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Y}}_2|\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U}\right) - h\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right)$$
$$+ I\left(\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U},W_1,\overleftarrow{\boldsymbol{Y}}_1;W_2,\overleftarrow{\boldsymbol{Y}}_2\right) + N\log\left(2\pi e\right)$$
$$+ N\delta(N)$$
$$\overset{(d)}{\leqslant} \sum_{n=1}^{N}\left[h\left(\overrightarrow{Y}_{1,n}|\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U},\overrightarrow{Y}_{1,(1:n-1)}\right)\right.$$
$$+ h\left(X_{1,U,n}|X_{2,C,n}\right) + h\left(\overleftarrow{Y}_{2,n}|X_{2,n},X_{1,U,n}\right)$$
$$+ h\left(\overrightarrow{Y}_{2,n}|\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U}\overrightarrow{Y}_{2,(1:n-1)}\right) + h\left(X_{2,U,n}|X_{1,C,n}\right)$$
$$\left. + h\left(\overleftarrow{Y}_{1,n}|X_{1,n},X_{2,U,n}\right) - 3\log\left(2\pi e\right)\right] + N\log\left(2\pi e\right)$$
$$+ N\delta(N)$$
$$\leqslant \sum_{n=1}^{N}\left[h\left(\overrightarrow{Y}_{1,n}|X_{1,C,n},X_{2,U,n}\right) + h\left(X_{1,U,n}|X_{2,C,n}\right)\right.$$
$$+ h\left(\overleftarrow{Y}_{2,n}|X_{2,n},X_{1,U,n}\right) + h\left(\overrightarrow{Y}_{2,n}|X_{2,C,n},X_{1,U,n}\right)$$
$$+ h\left(X_{2,U,n}|X_{1,C,n}\right) + h\left(\overleftarrow{Y}_{1,n}|X_{1,n},X_{2,U,n}\right)$$
$$\left. - 3\log\left(2\pi e\right)\right] + N\log\left(2\pi e\right) + N\delta(N)$$
$$= N\left[h\left(\overrightarrow{Y}_{1,k}|X_{1,C,k},X_{2,U,k}\right) + h\left(X_{1,U,k}|X_{2,C,k}\right)\right.$$
$$+ h\left(\overleftarrow{Y}_{2,k}|X_{2,k},X_{1,U,k}\right) + h\left(\overrightarrow{Y}_{2,k}|X_{2,C,k},X_{1,U,k}\right)$$
$$+ h\left(X_{2,U,k}|X_{1,C,k}\right) + h\left(\overleftarrow{Y}_{1,k}|X_{1,k},X_{2,U,k}\right)$$
$$\left. - 3\log\left(2\pi e\right)\right] + N\log\left(2\pi e\right) + N\delta(N), \quad (78)$$

where (a) follows from considering enhanced receivers (see Figure 6(b)) and Fano's inequality; (b) follows from the fact that $h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\boldsymbol{X}_{i,C},\boldsymbol{X}_{j,U}\right) + h\left(\boldsymbol{X}_{i,C},\boldsymbol{X}_{j,U}|\overrightarrow{\boldsymbol{Y}}_i\right) = h\left(\overrightarrow{\boldsymbol{Y}}_i|\boldsymbol{X}_{i,C},\boldsymbol{X}_{j,U}\right)$; (c) follows from the fact that $h\left(\overrightarrow{\boldsymbol{Z}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_j,\boldsymbol{X}_{i,C}\right) - h\left(\overrightarrow{\boldsymbol{Z}}_i,\overrightarrow{\boldsymbol{Z}}_j|\overrightarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_i,\boldsymbol{X}_j\right) \leqslant 0$; and (d) follows from Lemma 1 at the end of this appendix.

From (78), the following holds in the asymptotic regime for any $k \in \{1, 2, \ldots, N\}$:

$$R_1 + R_2 \leqslant h\left(\overrightarrow{Y}_{1,k}|X_{1,C,k},X_{2,U,k}\right) + h\left(X_{1,U,k}|X_{2,C,k}\right)$$
$$+ h\left(\overleftarrow{Y}_{2,k}|X_{2,k},X_{1,U,k}\right) + h\left(\overrightarrow{Y}_{2,k}|X_{2,C,k},X_{1,U,k}\right)$$
$$+ h\left(X_{2,U,k}|X_{1,C,k}\right) + h\left(\overleftarrow{Y}_{1,k}|X_{1,k},X_{2,U,k}\right)$$
$$- 2\log\left(2\pi e\right)$$
$$\leqslant \frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(\overrightarrow{Y}_{1,k},X_{1,C,k},X_{2,U,k}\right)\right)\right)$$
$$- \frac{1}{2}\log\left(\mathrm{INR}_{12}+1\right) - \frac{1}{2}\log\left(\mathrm{INR}_{21}+1\right)$$

$$+ \frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(\overleftarrow{Y}_{2,k},X_{2,k},X_{1,U,k}\right)\right)\right)$$
$$- \frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(X_{2,k},X_{1,U,k}\right)\right)\right)$$
$$+ \frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(\overrightarrow{Y}_{2,k},X_{2,C,k},X_{1,U,k}\right)\right)\right)$$
$$+ \frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(\overleftarrow{Y}_{1,k},X_{1,k},X_{2,U,k}\right)\right)\right)$$
$$- \frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(X_{1,k},X_{2,U,k}\right)\right)\right) + \log\left(2\pi e\right), \quad (79)$$

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

$$\det\left(\mathrm{Var}\left(\overrightarrow{Y}_{j,k},X_{j,C,k},X_{i,U,k}\right)\right) = \overrightarrow{\mathrm{SNR}}_j + \mathrm{INR}_{ji} + h_{ji,U}^2$$
$$- 2h_{ji,U}\sqrt{\mathrm{INR}_{ji}} + \left(1-\rho^2\right)\left(\mathrm{INR}_{ij}\mathrm{INR}_{ji}\right.$$
$$+ h_{ji,U}^2\left(\overrightarrow{\mathrm{SNR}}_j + \mathrm{INR}_{ij}\right) - 2h_{ji,U}\mathrm{INR}_{ij}\sqrt{\mathrm{INR}_{ji}}\Big)$$
$$+ 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_j}\left(\sqrt{\mathrm{INR}_{ji}} - h_{ji,U}\right), \quad (80a)$$
$$\det\left(\mathrm{Var}\left(\overleftarrow{Y}_{j,k},X_{j,k},X_{i,U,k}\right)\right) = 1 + h_{ji,U}^2\left(1-\rho^2\right)$$
$$+ \frac{\overleftarrow{\mathrm{SNR}}_j\left(1-\rho^2\right)\left(h_{ji,U}^2 - 2h_{ji,U}\sqrt{\mathrm{INR}_{ji}} + \mathrm{INR}_{ji}\right)}{\left(\overrightarrow{\mathrm{SNR}}_j + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_j\mathrm{INR}_{ji}} + \mathrm{INR}_{ji} + 1\right)}, \quad (80b)$$

and
$$\det\left(\mathrm{Var}\left(X_{j,k},X_{i,U,k}\right)\right) = 1 + \left(1-\rho^2\right)h_{ji,U}^2. \quad (80c)$$

The expressions in (80) 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 (73). 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 (73), it follows that $h_{12,U} = 0$ and $h_{21,U} = 0$. Therefore, plugging the expression (80) into (79) yields (34a).

Case 2: $(S_{1,2} \vee S_{2,2} \vee S_{5,2}) \wedge (S_{3,1} \vee S_{4,1})$. From (73), it follows that $h_{12,U} = 0$ and $h_{21,U} = \sqrt{\frac{\mathrm{INR}_{12}\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_2}}$. Therefore, plugging the expression (80) into (79) yields (34b).

Case 3: $(S_{3,2} \vee S_{4,2}) \wedge (S_{1,1} \vee S_{2,1} \vee S_{5,1})$. From (73), it follows that $h_{12,U} = \sqrt{\frac{\mathrm{INR}_{12}\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_1}}$ and $h_{21,U} = 0$. Therefore, plugging the expression (80) into (79) yields (34c).

Case 4: $(S_{3,2} \vee S_{4,2}) \wedge (S_{3,1} \vee S_{4,1})$. From (73), it follows that $h_{12,U} = \sqrt{\frac{\mathrm{INR}_{12}\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_1}}$ and $h_{21,U} = \sqrt{\frac{\mathrm{INR}_{12}\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_2}}$. Therefore, plugging the expression (80) into (79) yields (34d). This completes the proof of (37d).

Proof of (37e): 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, for all $i \in \{1, 2\}$, with

$j \in \{1,2\}\backslash\{i\}$, the following holds for any $k \in \{1,2,\ldots,N\}$:

$$N\left(2R_i + R_j\right) = 2H\left(W_i\right) + H\left(W_j\right)$$

$$\overset{(a)}{=} H\left(W_i\right) + H\left(W_i|W_j\right) + H\left(W_j\right)$$

$$\overset{(b)}{\leqslant} I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i,\overleftarrow{\boldsymbol{Y}}_i\right) + I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i,\overleftarrow{\boldsymbol{Y}}_j|W_j\right) + I\left(W_j;\overrightarrow{\boldsymbol{Y}}_j,\overleftarrow{\boldsymbol{Y}}_j\right)$$
$$+ N\delta(N)$$

$$\leqslant h\left(\overrightarrow{\boldsymbol{Y}}_i\right) + h\left(\overleftarrow{\boldsymbol{Z}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) - h\left(\overrightarrow{\boldsymbol{Y}}_i|W_i,\overleftarrow{\boldsymbol{Y}}_i\right)$$
$$+ h\left(\overleftarrow{\boldsymbol{Y}}_j|W_j\right) - h\left(\overleftarrow{\boldsymbol{Y}}_j|W_i,W_j\right) + I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Y}}_j\right) + h\left(\overleftarrow{\boldsymbol{Z}}_j\right) - h\left(\overleftarrow{\boldsymbol{Y}}_j|W_j\right) - h\left(\overrightarrow{\boldsymbol{Y}}_j|W_j,\overleftarrow{\boldsymbol{Y}}_j\right)$$
$$+ N\delta(N)$$

$$= h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) - h\left(\overrightarrow{\boldsymbol{Y}}_i|W_i,\overleftarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_i\right)$$
$$- h\left(\overleftarrow{\boldsymbol{Y}}_j|W_i,W_j\right) + I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j\right) + h\left(\overrightarrow{\boldsymbol{Y}}_j\right)$$
$$- h\left(\overrightarrow{\boldsymbol{Y}}_j|W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_j\right) + N\log\left(2\pi e\right) + N\delta(N)$$

$$\leqslant h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) - h\left(\overrightarrow{\boldsymbol{Y}}_i|W_i,\overleftarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_i\right)$$
$$+ I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j\right) + h\left(\overrightarrow{\boldsymbol{Y}}_j\right) - h\left(\overrightarrow{\boldsymbol{Y}}_j|W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_j\right)$$
$$+ N\log\left(2\pi e\right) + N\delta(N)$$

$$\overset{(c)}{=} h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) - h\left(\boldsymbol{X}_{j,C}|W_i,\overleftarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_i\right)$$
$$+ I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j\right) + h\left(\overrightarrow{\boldsymbol{Y}}_j\right) - h\left(\boldsymbol{X}_{i,C}|W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_j\right)$$
$$+ N\log\left(2\pi e\right) + N\delta(N)$$

$$= h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) - h\left(\boldsymbol{X}_{j,C},\overrightarrow{\boldsymbol{Z}}_j|W_i,\overleftarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_i\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Z}}_j|W_i,\overleftarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_i,\boldsymbol{X}_{j,C}\right) + I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Y}}_j\right) - h\left(\boldsymbol{X}_{i,C}|W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_j\right) + N\log\left(2\pi e\right)$$
$$+ N\delta(N)$$

$$\overset{(d)}{=} h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) - h\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}|W_i,\overleftarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_i\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Z}}_j|W_i,\overleftarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_i,\boldsymbol{X}_{j,C}\right) + I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Y}}_j\right) - h\left(\boldsymbol{X}_{i,C}|W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_j\right) + N\log\left(2\pi e\right)$$
$$+ N\delta(N)$$

$$\leqslant h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) - h\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}|W_i,\overleftarrow{\boldsymbol{Y}}_i\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Z}}_j|W_i,\overleftarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_i,\boldsymbol{X}_{j,C}\right) + I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_{i,C}|W_j,\overleftarrow{\boldsymbol{Y}}_j\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Y}}_j\right) - h\left(\boldsymbol{X}_{i,C}|W_j,\overleftarrow{\boldsymbol{Y}}_j\right) + N\log\left(2\pi e\right) + N\delta(N)$$

$$= h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) - h\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}|W_i,\overleftarrow{\boldsymbol{Y}}_i\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Z}}_j|W_i,\overleftarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_i,\boldsymbol{X}_{j,C}\right) + h\left(\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_{i,C}\right)$$
$$- h\left(\overrightarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_{i,C}|W_i,W_j,\overleftarrow{\boldsymbol{Y}}_j\right) + h\left(\overrightarrow{\boldsymbol{Y}}_j\right) + N\log\left(2\pi e\right)$$
$$+ N\delta(N)$$

$$\overset{(e)}{\leqslant} h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) - h\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}|W_i,\overleftarrow{\boldsymbol{Y}}_i\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Z}}_j|W_i,\overleftarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_i,\boldsymbol{X}_{j,C}\right) + h\left(\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_{i,C}\right)$$
$$- h\left(\overrightarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_{i,C}|W_i,W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_i,\boldsymbol{X}_j\right) + h\left(\overrightarrow{\boldsymbol{Y}}_j\right)$$
$$+ N\log\left(2\pi e\right) + N\delta(N)$$

$$= h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) - h\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}|W_i,\overleftarrow{\boldsymbol{Y}}_i\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Z}}_j|W_i,\overleftarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_i,\boldsymbol{X}_{j,C}\right) + h\left(\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_{i,C}\right)$$

$$- h\left(\overrightarrow{\boldsymbol{Z}}_i,\overrightarrow{\boldsymbol{Z}}_j|W_i,W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_i,\boldsymbol{X}_j\right) + h\left(\overrightarrow{\boldsymbol{Y}}_i\right)$$
$$+ N\log\left(2\pi e\right) + N\delta(N)$$

$$\overset{(f)}{\leqslant} h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) - h\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}|W_i,\overleftarrow{\boldsymbol{Y}}_i\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_{i,C}\right) + h\left(\overrightarrow{\boldsymbol{Y}}_j\right) + N\log\left(2\pi e\right)$$
$$+ N\delta(N)$$

$$\leqslant h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) + I\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U};W_i,\overleftarrow{\boldsymbol{Y}}_i\right)$$
$$- h\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}\right) + h\left(\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_{i,C}\right) + h\left(\overrightarrow{\boldsymbol{Y}}_j\right)$$
$$+ h\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}|\overrightarrow{\boldsymbol{Y}}_j\right) + N\log\left(2\pi e\right) + N\delta(N)$$

$$\overset{(g)}{=} h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) + I\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U};W_i,\overleftarrow{\boldsymbol{Y}}_i\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_{i,C}\right) + h\left(\overrightarrow{\boldsymbol{Y}}_j|\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}\right)$$
$$+ N\log\left(2\pi e\right) + N\delta(N)$$

$$\leqslant h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) + I\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U},W_j,\overleftarrow{\boldsymbol{Y}}_j;W_i,\overleftarrow{\boldsymbol{Y}}_i\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_{i,C}\right) + h\left(\overrightarrow{\boldsymbol{Y}}_j|\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}\right)$$
$$+ N\log\left(2\pi e\right) + N\delta(N)$$

$$\overset{(h)}{\leqslant} h\left(\overrightarrow{\boldsymbol{Y}}_i\right) + \sum_{n=1}^{N}\left[h\left(X_{i,U,n}|X_{j,C,n}\right) + h\left(\overleftarrow{\boldsymbol{Y}}_{j,n}|X_{j,n},X_{i,U,n}\right)\right.$$
$$\left. - \frac{3}{2}\log\left(2\pi e\right)\right] + h\left(\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_{i,C}\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Y}}_j|\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}\right) + N\log\left(2\pi e\right) + N\delta(N)$$

$$\overset{(i)}{\leqslant} h\left(\overrightarrow{\boldsymbol{Y}}_i\right) + \sum_{n=1}^{N}\left[h\left(X_{i,U,n}|X_{j,C,n}\right) + h\left(\overleftarrow{\boldsymbol{Y}}_{j,n}|X_{j,n},X_{i,U,n}\right)\right.$$
$$\left. - \frac{3}{2}\log\left(2\pi e\right)\right] + h\left(\overrightarrow{\boldsymbol{Y}}_i|\boldsymbol{X}_{i,C},\boldsymbol{X}_j\right)$$
$$+ h\left(\overrightarrow{\boldsymbol{Y}}_j|\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}\right) + N\log\left(2\pi e\right) + N\delta(N)$$

$$\leqslant \sum_{n=1}^{N}\left[h\left(\overrightarrow{Y}_{i,n}\right) + h\left(X_{i,U,n}|X_{j,C,n}\right) + h\left(\overleftarrow{Y}_{j,n}|X_{j,n},X_{i,U,n}\right)\right.$$
$$- \frac{3}{2}\log\left(2\pi e\right) + h\left(\overrightarrow{Y}_{i,n}|X_{i,C,n},X_{j,n}\right)$$
$$\left. + h\left(\overrightarrow{Y}_{j,n}|X_{j,C,n},X_{i,U,n}\right)\right] + N\log\left(2\pi e\right) + N\delta(N)$$

$$= N\left[h\left(\overrightarrow{Y}_{i,k}\right) + h\left(X_{i,U,k}|X_{j,C,k}\right) + h\left(\overleftarrow{Y}_{j,k}|X_{j,k},X_{i,U,j}\right)\right.$$
$$- \frac{5}{2}\log\left(2\pi e\right) + h\left(\overrightarrow{Y}_{i,k}|X_{i,C,k},X_{j,k}\right)$$
$$\left. + h\left(\overrightarrow{Y}_{j,k}|X_{j,C,k},X_{i,U,k}\right) + 2\log\left(2\pi e\right) + \delta(N)\right], \quad (81)$$

where, (a) follows from the fact that W_1 and W_2 are mutually independent; (b) follows from considering enhanced receivers (see Figure 6(c) for the case $i = 1$) and Fano's inequality; (c) follows from (8) and (74); (d) follows from (75); (e) follows from (10) and the fact that conditioning does not increase entropy; (f) follows from the fact that $h\left(\overrightarrow{\boldsymbol{Z}}_j|W_j,\overleftarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_i,\boldsymbol{X}_{j,C}\right) - h\left(\overrightarrow{\boldsymbol{Z}}_i,\overrightarrow{\boldsymbol{Z}}_j|W_i,W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_i,\boldsymbol{X}_j\right) \leqslant 0$; (g) follows from the fact that $h\left(\overrightarrow{\boldsymbol{Y}}_j\right) - h\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}\right) + h\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}|\overrightarrow{\boldsymbol{Y}}_j\right) = h\left(\overrightarrow{\boldsymbol{Y}}_j|\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}\right)$; (h) follows from Lemma 1 at the end of this appendix; and (i) follows from the fact that conditioning does not increase entropy.

From (81), the following holds in the asymptotic regime for any $k \in \{1, 2, \ldots N\}$:

$$2R_i + R_j \leqslant h\left(\overrightarrow{Y}_{i,k}\right) + h\left(X_{i,U,k}|X_{j,C,k}\right) + h\left(\overleftarrow{Y}_{j,k}|X_{j,k},X_{i,U,k}\right)$$
$$+ h\left(\overrightarrow{Y}_{i,k}|X_{i,C,k},X_{j,k}\right) + h\left(\overrightarrow{Y}_{j,k}|X_{j,C,k},X_{i,U,k}\right)$$
$$- \frac{1}{2}\log\left(2\pi e\right)$$
$$\leqslant \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_i + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_i\mathrm{INR}_{ij}} + \mathrm{INR}_{ij} + 1\right)$$
$$- \frac{1}{2}\log\left(\mathrm{INR}_{ij}+1\right) + \frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(\overleftarrow{Y}_{j,k},X_{j,k},X_{i,U,k}\right)\right)\right)$$
$$- \frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(X_{j,k},X_{i,U,k}\right)\right)\right)$$
$$+ \frac{1}{2}\log\left(1 + \left(1-\rho^2\right)\left(\overrightarrow{\mathrm{SNR}}_i + \mathrm{INR}_{ji}\right)\right)$$
$$- \frac{1}{2}\log\left(1 + \left(1-\rho^2\right)\mathrm{INR}_{ji}\right)$$
$$+ \frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(\overrightarrow{Y}_{j,k},X_{j,C,k},X_{i,U,k}\right)\right)\right) + 2\log\left(2\pi e\right). \tag{82}$$

The outer bound on (82) 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 (73). 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 (73), it follows that $h_{ji,U} = 0$. Therefore, plugging (80) into (82) yields (35a).

Case 2: $(S_{3,i} \vee S_{4,i})$. From (73), it follows that $h_{ji,U} = \sqrt{\frac{\mathrm{INR}_{ij}\mathrm{INR}_{ji}}{\overrightarrow{\mathrm{SNR}}_j}}$. Therefore, plugging the expressions (80) into (82) yields (35b).

This completes the proof of (37e) and the proof of Theorem 3.

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

$$I\left(\boldsymbol{X}_{i,C}, \boldsymbol{X}_{j,U}, \overleftarrow{\boldsymbol{Y}}_i, W_i; \overleftarrow{\boldsymbol{Y}}_j, W_j\right) \leqslant h\left(\overleftarrow{\boldsymbol{Y}}_j|W_j\right)$$
$$+ \sum_{n=1}^{N}\left[h\left(X_{j,U,n}|X_{i,C,n}\right) + h\left(\overleftarrow{Y}_{i,n}|X_{i,n},X_{j,U,n}\right)\right.$$
$$\left. - \frac{3}{2}\log\left(2\pi e\right)\right]. \tag{83}$$

Proof: Lemma 1 is proved as follows:

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

$$\boldsymbol{X}_{j,U,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{i,(1:n-1)}, \boldsymbol{X}_{i,(1:n)}\Big)$$
$$- h\Big(X_{i,C,n}, X_{j,U,n}, \overleftarrow{Y}_{i,n}, \overleftarrow{Y}_{j,n}|W_i, W_j, \boldsymbol{X}_{i,C,(1:n-1)},$$
$$\left.\boldsymbol{X}_{j,U,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{i,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}, \boldsymbol{X}_{i,(1:n)}, \boldsymbol{X}_{j,(1:n)}\Big)\right]$$
$$\leqslant h\left(\overleftarrow{\boldsymbol{Y}}_j|W_j\right) + \sum_{n=1}^{N}\left[h\left(X_{i,C,n}, X_{j,U,n}, \overleftarrow{Y}_{i,n}|X_{i,n}\right)\right.$$
$$- h\Big(\overrightarrow{Z}_{j,n}, \overrightarrow{Z}_{i,n}, \overleftarrow{Y}_{i,n}, \overleftarrow{Y}_{j,n}|W_i, W_j, \boldsymbol{X}_{i,C,(1:n-1)},$$
$$\left.\boldsymbol{X}_{j,U,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{i,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}, \boldsymbol{X}_{i,(1:n)}, \boldsymbol{X}_{j,(1:n)}\Big)\right]$$
$$= h\left(\overleftarrow{\boldsymbol{Y}}_j|W_j\right) + \sum_{n=1}^{N}\left[h\left(X_{i,C,n}|X_{i,n}\right)\right.$$
$$+ h\Big(X_{j,U,n}|X_{i,n}, X_{i,C,n}\Big) + h\Big(\overleftarrow{Y}_{i,n}|X_{i,n}, X_{i,C,n}, X_{j,U,n}\Big)$$
$$- h\left(\overrightarrow{Z}_{j,n}\right) - h\left(\overrightarrow{Z}_{i,n}\right) - h\Big(\overleftarrow{Y}_{i,n}, \overleftarrow{Y}_{j,n}|W_i, W_j,$$
$$\boldsymbol{X}_{i,C,(1:n-1)}, \boldsymbol{X}_{j,U,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{i,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)},$$
$$\left.\boldsymbol{X}_{i,(1:n)}, \boldsymbol{X}_{j,(1:n)}, \overrightarrow{Z}_{j,n}, \overrightarrow{Z}_{i,n}\Big)\right]$$
$$\leqslant h\left(\overleftarrow{\boldsymbol{Y}}_j|W_j\right) + \sum_{n=1}^{N}\left[h\left(\overrightarrow{Z}_{j,n}|X_{i,n}\right) + h\left(X_{j,U,n}|X_{i,C,n}\right)\right.$$
$$+ h\Big(\overleftarrow{Y}_{i,n}|X_{i,n}, X_{j,U,n}\Big) - h\left(\overrightarrow{Z}_{j,n}\right) - h\left(\overrightarrow{Z}_{i,n}\right)$$
$$- h\Big(\overleftarrow{Z}_{i,n}, \overleftarrow{Z}_{j,n}|W_i, W_j, \boldsymbol{X}_{i,C,(1:n-1)}, \boldsymbol{X}_{j,U,(1:n-1)},$$
$$\left.\overleftarrow{\boldsymbol{Y}}_{i,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}, \boldsymbol{X}_{i,(1:n)}, \boldsymbol{X}_{j,(1:n)}, \overrightarrow{Z}_{j,n}, \overrightarrow{Z}_{i,n}\Big)\right]$$
$$\overset{(a)}{=} h\left(\overleftarrow{\boldsymbol{Y}}_j|W_j\right) + \sum_{n=1}^{N}\left[h\left(X_{j,U,n}|X_{i,C,n}\right)\right.$$
$$+ h\Big(\overleftarrow{Y}_{i,n}|X_{i,n}, X_{j,U,n}\Big) - h\left(\overrightarrow{Z}_{i,n}\right) - h\left(\overleftarrow{Z}_{i,n}\right)$$
$$\left. - h\left(\overleftarrow{Z}_{j,n}\right)\right]$$
$$= h\left(\overleftarrow{\boldsymbol{Y}}_j|W_j\right) + \sum_{n=1}^{N}\left[h\left(X_{j,U,n}|X_{i,C,n}\right)\right.$$
$$\left. + h\left(\overleftarrow{Y}_{i,n}|X_{i,n}, X_{j,U,n}\right) - \frac{3}{2}\log\left(2\pi e\right)\right],$$

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

APPENDIX D
PROOF OF THEOREM 4

The gap between the sets $\overline{\mathcal{C}}_{\mathrm{G-IC-NF}}$ and $\underline{\mathcal{C}}_{\mathrm{G-IC-NF}}$ (Definition 2), denoted by δ, is approximated as follows:

$$\delta = \max\left(\delta_{R_1}^*, \delta_{R_2}^*, \frac{\delta_{2R}^*}{2}, \frac{\delta_{3R_1}^*}{3}, \frac{\delta_{3R_2}^*}{3}\right), \tag{84}$$

where,

$$\delta_{R_1}^* = \sup_{\rho' \in [0,1]} \{\delta_{R_1}(\rho', \rho, \mu_1)\}, \tag{85a}$$

$$\delta_{R_2}^* = \sup_{\rho' \in [0,1]} \{\delta_{R_2}(\rho', \rho, \mu_2)\}, \tag{85b}$$

$$\delta_{2R}^* = \sup_{\rho' \in [0,1]} \{\delta_{2R}(\rho', \rho, \mu_1, \mu_2)\}, \tag{85c}$$

$$\delta_{3R_1}^* = \sup_{\rho' \in [0,1]} \{\delta_{3R_1}(\rho', \rho, \mu_1, \mu_2)\}, \tag{85d}$$

$$\delta_{3R_2}^* = \sup_{\rho' \in [0,1]} \{\delta_{3R_2}(\rho', \rho, \mu_1, \mu_2)\}, \tag{85e}$$

for some ρ, μ_1, μ_2 chosen arbitrarily, with

$$\begin{aligned}
\delta_{R_1}\Big(\rho', \rho, \mu_1\Big) &= \min\Big(\kappa_{1,1}(\rho'), \kappa_{2,1}(\rho'), \kappa_{3,1}(\rho')\Big) \\
&- \min\Big(a_{2,1}(\rho), a_{6,1}(\rho, \mu_1) + a_{3,2}(\rho, \mu_1), \\
&\quad a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{4,2}(\rho, \mu_1)\Big),
\end{aligned} \tag{86a}$$

$$\begin{aligned}
\delta_{R_2}\Big(\rho', \rho, \mu_2\Big) &= \min\Big(\kappa_{1,2}(\rho'), \kappa_{2,2}(\rho'), \kappa_{3,2}(\rho')\Big) \\
&- \min\Big(a_{2,2}(\rho), a_{3,1}(\rho, \mu_2) + a_{6,2}(\rho, \mu_2), \\
&\quad a_{3,1}(\rho, \mu_2) + a_{4,1}(\rho, \mu_2) + a_{1,2}\Big),
\end{aligned} \tag{86b}$$

$$\begin{aligned}
\delta_{2R}\Big(\rho', \rho, \mu_1, \mu_2\Big) &= \min\Big(\kappa_4(\rho'), \kappa_5(\rho'), \kappa_6(\rho')\Big) \\
&- \min\Big(a_{2,1}(\rho) + a_{1,2}, a_{1,1} + a_{2,2}(\rho), \\
&\quad a_{3,1}(\rho, \mu_2) + a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{7,2}(\rho, \mu_1, \mu_2), \\
&\quad a_{3,1}(\rho, \mu_2) + a_{5,1}(\rho, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{5,2}(\rho, \mu_1), \\
&\quad a_{3,1}(\rho, \mu_2) + a_{7,1}(\rho, \mu_1, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{1,2}\Big),
\end{aligned} \tag{86c}$$

$$\begin{aligned}
\delta_{3R_1}\Big(\rho', \rho, \mu_1, \mu_2\Big) &= \kappa_{7,1}(\rho') \\
&- \min\Big(a_{2,1}(\rho) + a_{1,1} + a_{3,2}(\rho, \mu_1) + a_{7,2}(\rho, \mu_1, \mu_2), \\
&\quad a_{3,1}(\rho, \mu_2) + a_{1,1} + a_{7,1}(\rho, \mu_1, \mu_2) + 2a_{3,2}(\rho, \mu_1) \\
&\quad + a_{5,2}(\rho, \mu_1), a_{2,1}(\rho) + a_{1,1} + a_{3,2}(\rho, \mu_1) \\
&\quad + a_{5,2}(\rho, \mu_1)\Big),
\end{aligned} \tag{86d}$$

$$\begin{aligned}
\delta_{3R_2}\Big(\rho', \rho, \mu_1, \mu_2\Big) &= \kappa_{7,2}(\rho') \\
&- \min\Big(a_{3,1}(\rho, \mu_2) + a_{5,1}(\rho, \mu_2) + a_{2,2}(\rho) + a_{1,2}, \\
&\quad a_{3,1}(\rho, \mu_2) + a_{7,1}(\rho, \mu_1, \mu_2) + a_{2,2}(\rho) + a_{1,2}, \\
&\quad 2a_{3,1}(\rho, \mu_2) + a_{5,1}(\rho, \mu_2) + a_{3,2}(\rho, \mu_1) + a_{1,2} \\
&\quad + a_{7,2}(\rho, \mu_1, \mu_2)\Big).
\end{aligned} \tag{86e}$$

Note that $\delta_{R_1}^*$ and $\delta_{R_2}^*$ represent the gap between an achievable individual rate bound (determined by ρ, μ_1, μ_2) and the active converse individual rate bound; δ_{2R}^* represents the gap between an achievable sum-rate bound (determined by ρ, μ_1, μ_2) and the active converse sum-rate bound; $\delta_{3R_1}^*$ and $\delta_{3R_2}^*$ represent the gap between an achievable weighted sum-rate bound (determined by ρ, μ_1, μ_2) and the active converse weighted sum-rate bound.

It is important to highlight that, as suggested in [3], [5], and [25], the gap between $\underline{\mathcal{C}}_{\text{G-IC-NF}}$ and $\overline{\mathcal{C}}_{\text{G-IC-NF}}$ can be

calculated more precisely. However, the approximation in (84) eases the calculations at the expense of less precision. A key point in order to find the gap between the achievable region and the converse region is to choose convenient values of ρ, μ_1, and μ_2, such that the optimization problem in (86) become an optimization only on $\rho' \in [0,1]$. The following describes all the key cases.

Case 1: $\text{INR}_{12} > \overrightarrow{\text{SNR}}_1$ and $\text{INR}_{21} > \overrightarrow{\text{SNR}}_2$. 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.

<u>Case 1.1</u>: $\overleftarrow{\text{SNR}}_2 \leqslant \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{SNR}}_1 \leqslant \overleftarrow{\text{SNR}}_2$. The choice is: $\rho = 0$, $\mu_1 = 0$, and $\mu_2 = 0$.

<u>Case 1.2</u>: $\overleftarrow{\text{SNR}}_2 > \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{SNR}}_1 > \overleftarrow{\text{SNR}}_2$. The choice is: $\rho = 0$, $\mu_1 = 1$, and $\mu_2 = 1$.

<u>Case 1.3</u>: $\overleftarrow{\text{SNR}}_2 \leqslant \overleftarrow{\text{SNR}}_1$ and $\overleftarrow{\text{SNR}}_1 > \overleftarrow{\text{SNR}}_2$. The choice is: $\rho = 0$, $\mu_1 = 0$, and $\mu_2 = 1$.

Case 2: $\text{INR}_{12} \leqslant \overrightarrow{\text{SNR}}_1$ and $\text{INR}_{21} \leqslant \overrightarrow{\text{SNR}}_2$. 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 choice is: $\rho = 0$, $\mu_1 = 0$, and $\mu_2 = 0$.

<u>Case 2.1</u>: $\overleftarrow{\text{SNR}}_1 \leqslant \text{INR}_{21}$, $\overleftarrow{\text{SNR}}_2 \leqslant \text{INR}_{12}$, $\text{INR}_{12}\text{INR}_{21} > \overrightarrow{\text{SNR}}_1$ and $\text{INR}_{12}\text{INR}_{21} > \overrightarrow{\text{SNR}}_2$.

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

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

<u>Case 2.4</u>: $\overleftarrow{\text{SNR}}_1\text{INR}_{12} \leqslant \overrightarrow{\text{SNR}}_1$, $\overleftarrow{\text{SNR}}_2\text{INR}_{21} \leqslant \overrightarrow{\text{SNR}}_2$, $\text{INR}_{12}\text{INR}_{21} < \overrightarrow{\text{SNR}}_1$ and $\text{INR}_{12}\text{INR}_{21} < \overrightarrow{\text{SNR}}_2$.

In the following four subcases, the choice is: $\rho = 0$, $\mu_1 = \frac{\text{INR}_{21}^2\overleftarrow{\text{SNR}}_2}{(\text{INR}_{21}-1)(\text{INR}_{21}\overleftarrow{\text{SNR}}_2+\overrightarrow{\text{SNR}}_2)}$, and $\mu_2 = \frac{\text{INR}_{12}^2\overleftarrow{\text{SNR}}_1}{(\text{INR}_{12}-1)(\text{INR}_{12}\overleftarrow{\text{SNR}}_1+\overrightarrow{\text{SNR}}_1)}$.

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

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

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

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

In the following four subcases, the choice is: $\rho = 0$, $\mu_1 = 0$, and $\mu_2 = \frac{\text{INR}_{12}^2\overleftarrow{\text{SNR}}_1}{(\text{INR}_{12}-1)(\text{INR}_{12}\overleftarrow{\text{SNR}}_1+\overrightarrow{\text{SNR}}_1)}$.

<u>Case 2.9</u>: $\overleftarrow{\text{SNR}}_1 > \text{INR}_{21}$, $\overleftarrow{\text{SNR}}_2 \leqslant \text{INR}_{12}$, $\text{INR}_{12}\text{INR}_{21} > \overrightarrow{\text{SNR}}_1$ and $\text{INR}_{12}\text{INR}_{21} > \overrightarrow{\text{SNR}}_2$.

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

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

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

Case 3: $\text{INR}_{12} > \overrightarrow{\text{SNR}}_1$ and $\text{INR}_{21} \leqslant \overrightarrow{\text{SNR}}_2$. 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 choice is: $\rho = 0$, $\mu_1 = 0$, and $\mu_2 = 0$.

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

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

In the following two subcases, the choice is: $\rho = 0$, $\mu_1 = 1$, 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 $\delta^*_{R_1}$.

From (86a) and considering the choice of values: $\rho = 0$, $\mu_1 = 0$ and $\mu_2 = 0$, it follows that

$$\delta^*_{R_1} \leqslant \sup_{\rho' \in [0,1]} \left\{ \min\left(\kappa_{1,1}(\rho'), \kappa_{2,1}(\rho'), \kappa_{3,1}(\rho') \right) \right\} - \min\left(a_{6,1}(0,0), a_{1,1} + a_{4,2}(0,0) \right). \qquad (87)$$

Note that in this case:

$$\kappa_{1,1}(\rho') = \frac{1}{2} \log\left(b_{1,1}(\rho') + 1 \right)$$
$$\overset{(a)}{\leqslant} \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_1 + 2\sqrt{\overrightarrow{\text{SNR}}_1 \text{INR}_{12}} + \text{INR}_{12} + 1 \right)$$
$$\overset{(b)}{\leqslant} \frac{1}{2} \log\left(2\overrightarrow{\text{SNR}}_1 + 2\text{INR}_{12} + 1 \right)$$
$$\leqslant \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{12} + 1 \right) + \frac{1}{2}, \qquad (88a)$$

$$\kappa_{2,1}(\rho') = \frac{1}{2} \log\left(1 + b_{4,1}(\rho') + b_{5,2}(\rho') \right)$$
$$\leqslant \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1 \right), \qquad (88b)$$

$$\kappa_{3,1}(\rho') = \frac{1}{2} \log\left(b_{4,1}(\rho') + 1 \right)$$
$$+ \frac{1}{2} \log\left(\frac{\overleftarrow{\text{SNR}}_2\left(b_{4,1}(\rho') + b_{5,2}(\rho') + 1\right)}{\left(b_{1,2}(1) + 1\right)\left(b_{4,1}(\rho') + 1\right)} + 1 \right)$$
$$\overset{(c)}{\leqslant} \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_1 + 1 \right)$$
$$+ \frac{1}{2} \log\left(\frac{\overleftarrow{\text{SNR}}_2\left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1\right)}{\left(\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1\right)\left(\overrightarrow{\text{SNR}}_1 + 1\right)} + 1 \right)$$
$$= \frac{1}{2} \log\left(\frac{\overleftarrow{\text{SNR}}_2\left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1\right)}{\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1} + \overrightarrow{\text{SNR}}_1 + 1 \right), \qquad (88c)$$

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

$$\left(\sqrt{\overrightarrow{\text{SNR}}_1} - \sqrt{\text{INR}_{12}} \right)^2 \geqslant 0; \qquad (89)$$

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:

$$a_{1,1} + a_{4,2}(0,0) = \frac{1}{2} \log\left(\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} + 2 \right)$$
$$+ \frac{1}{2} \log\left(\text{INR}_{21} + 1 \right) - 1$$
$$\geqslant \frac{1}{2} \log\left(\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}} + 2 \right) + \frac{1}{2} \log\left(\text{INR}_{21} \right) - 1$$
$$= \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_1 + 2\text{INR}_{21} \right) - 1$$
$$= \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + \text{INR}_{21} \right) - 1$$
$$\geqslant \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1 \right) - 1. \qquad (90)$$

From (87), (88) and (90), assuming that $a_{1,1} + a_{4,2}(0,0) < a_{6,1}(0,0)$, it follows that

$$\delta^*_{R_1} \leqslant \sup_{\rho' \in [0,1]} \left\{ \min\left(\kappa_{1,1}(\rho'), \kappa_{2,1}(\rho'), \kappa_{3,1}(\rho') \right) \right\}$$
$$- \left(a_{1,1} + a_{4,2}(0,0) \right)$$
$$\leqslant \sup_{\rho' \in [0,1]} \kappa_{2,1}(\rho') - \left(a_{1,1} + a_{4,2}(0,0) \right)$$
$$\leqslant 1. \qquad (91)$$

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

$$\delta^*_{R_1} \leqslant \sup_{\rho' \in [0,1]} \left\{ \min\left(\kappa_{1,1}(\rho'), \kappa_{2,1}(\rho'), \kappa_{3,1}(\rho') \right) \right\}$$
$$- a_{6,1}(0,0). \qquad (92)$$

To calculate an upper bound for (92), the following cases are considered:

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

Case 1.1.2: $\overrightarrow{\text{SNR}}_1 < \text{INR}_{21} \wedge \overrightarrow{\text{SNR}}_2 \geqslant \text{INR}_{12}$, 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 (88) and (92), it follows that

$$\delta^*_{R_1} \leqslant \sup_{\rho' \in [0,1]} \kappa_{2,1}(\rho') - a_{6,1}(0,0)$$
$$\leqslant \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1 \right) - \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_1 + 2 \right)$$
$$+ \frac{1}{2}$$
$$\leqslant \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_1 + \overrightarrow{\text{SNR}}_1 + 1 \right) - \frac{1}{2} \log\left(\overrightarrow{\text{SNR}}_1 + 2 \right)$$
$$+ \frac{1}{2}$$
$$\leqslant 1. \qquad (93)$$

In Case 1.1.2, from (88) and (92), it follows that

$$\delta_{R_1}^* \leqslant \sup_{\rho' \in [0,1]} \kappa_{3,1}(\rho') - a_{6,1}(0,0)$$

$$\leqslant \frac{1}{2}\log\left(\frac{\overleftarrow{\text{SNR}}_2\left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1\right)}{\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1} + \overrightarrow{\text{SNR}}_1 + 1\right)$$

$$- \frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1 + 2\right) + \frac{1}{2}$$

$$\leqslant \frac{1}{2}\log\left(\overleftarrow{\text{SNR}}_2 + \overrightarrow{\text{SNR}}_1 + 1\right)$$

$$- \frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1 + 2\right) + \frac{1}{2}$$

$$\leqslant \frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1 + \overrightarrow{\text{SNR}}_1 + 1\right)$$

$$- \frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1 + 2\right) + \frac{1}{2}$$

$$\leqslant 1. \tag{94}$$

In Case 1.1.3 two additional cases are considered:
Case 1.1.3.1: $\overrightarrow{\text{SNR}}_1 \geqslant \overrightarrow{\text{SNR}}_2$ and
Case 1.1.3.2: $\overrightarrow{\text{SNR}}_1 < \overrightarrow{\text{SNR}}_2$.
In Case 1.1.3.1, from (88) and (92), it follows that

$$\delta_{R_1}^* \leqslant \sup_{\rho' \in [0,1]} \kappa_{3,1}(\rho') - a_{6,1}(0,0)$$

$$\leqslant \frac{1}{2}\log\left(\frac{\overleftarrow{\text{SNR}}_2\left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1\right)}{\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1} + \overrightarrow{\text{SNR}}_1 + 1\right)$$

$$- \frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1 + 2\right) + \frac{1}{2}$$

$$= \frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1 + 1\right)$$

$$+ \frac{1}{2}\log\left(\frac{\overleftarrow{\text{SNR}}_2\left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1\right)}{\left(\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1\right)\left(\overrightarrow{\text{SNR}}_1 + 1\right)} + 1\right)$$

$$- \frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1 + 2\right) + \frac{1}{2}$$

$$\leqslant \frac{1}{2}\log\left(\frac{\overrightarrow{\text{SNR}}_1\left(\text{INR}_{21} + \text{INR}_{21} + \text{INR}_{21}\right)}{\text{INR}_{21}\overline{\text{SNR}}_1} + 1\right)$$

$$+ \frac{1}{2}$$

$$= \frac{3}{2}. \tag{95}$$

In Case 1.1.3.2, from (88) and (92), it follows that

$$\delta_{R_1}^* \leqslant \sup_{\rho' \in [0,1]} \kappa_{3,1}(\rho') - a_{6,1}(0,0)$$

$$\leqslant \frac{1}{2}\log\left(\frac{\overleftarrow{\text{SNR}}_2\left(\overrightarrow{\text{SNR}}_1 + \text{INR}_{21} + 1\right)}{\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1} + \overrightarrow{\text{SNR}}_1 + 1\right)$$

$$- \frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1 + 2\right) + \frac{1}{2}$$

$$\leqslant \frac{1}{2}\log\left(\overleftarrow{\text{SNR}}_2 + \overrightarrow{\text{SNR}}_1 + 1\right) - \frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1 + 2\right)$$

$$+ \frac{1}{2}$$

$$\leqslant \frac{1}{2}\log\left(\overleftarrow{\text{SNR}}_1 + \overrightarrow{\text{SNR}}_1 + 1\right) - \frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1 + 2\right)$$

$$+ \frac{1}{2}$$

$$\leqslant 1. \tag{96}$$

Then, from (91), (93), (94), (95), and (96), it follows that in Case 1.1:

$$\delta_{R_1}^* \leqslant \frac{3}{2}. \tag{97}$$

The same procedure holds to calculate $\delta_{R_2}^*$ and it yields:

$$\delta_{R_2}^* \leqslant \frac{3}{2}. \tag{98}$$

2) Calculation of δ_{2R}^*. From (86c) and considering the values $\rho = 0$, $\mu_1 = 0$, and $\mu_2 = 0$, it follows that

$$\delta_{2R}^* \leqslant \sup_{\rho' \in [0,1]} \left\{ \min\left(\kappa_4(\rho'), \kappa_5(\rho'), \kappa_6(\rho')\right)\right\}$$

$$- \min\left(a_{2,1}(0) + a_{1,2}, a_{1,1} + a_{2,2}(0),\right.$$

$$\left. a_{5,1}(0,0) + a_{5,2}(0,0)\right)$$

$$\leqslant \sup_{\rho' \in [0,1]} \left\{ \min\left(\kappa_4(\rho'), \kappa_5(\rho')\right)\right\}$$

$$- \min\left(a_{2,1}(0) + a_{1,2}, a_{1,1} + a_{2,2}(0),\right.$$

$$\left. a_{5,1}(0,0) + a_{5,2}(0,0)\right). \tag{99}$$

Note that

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

$$\leqslant \frac{1}{2}\log\left(1 + \frac{b_{4,1}(\rho')}{b_{5,2}(\rho')}\right) + \frac{1}{2}\log\left(b_{1,2}(\rho') + 1\right)$$

$$= \frac{1}{2}\log\left(1 + \frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}}\right) + \frac{1}{2}\log\left(b_{1,2}(\rho') + 1\right)$$

$$\overset{(h)}{\leqslant} \frac{1}{2}\log\left(1 + \frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}}\right)$$

$$+ \frac{1}{2}\log\left(2\overrightarrow{\text{SNR}}_2 + 2\text{INR}_{21} + 1\right),$$

$$\leqslant \frac{1}{2}\log\left(1 + \frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}}\right)$$

$$+ \frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1\right) + \frac{1}{2}$$

$$\leqslant \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}}\right)$$

$$+ \frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_2 + \text{INR}_{21} + 1\right) + \frac{1}{2}, \tag{100a}$$

and

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

$$\leqslant \frac{1}{2}\log\left(1 + \frac{b_{4,2}(\rho')}{b_{5,1}(\rho')}\right) + \frac{1}{2}\log\left(b_{1,1}(\rho') + 1\right)$$

$$= \frac{1}{2}\log\left(1 + \frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}\right) + \frac{1}{2}\log\left(b_{1,1}(\rho') + 1\right)$$

$$\overset{(i)}{\leqslant} \frac{1}{2}\log\left(1 + \frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}\right)$$

$$+ \frac{1}{2}\log\left(2\overrightarrow{\mathrm{SNR}}_1 + 2\mathrm{INR}_{12} + 1\right)$$

$$\leqslant \frac{1}{2}\log\left(1 + \frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}\right)$$

$$+ \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1\right) + \frac{1}{2}$$

$$\leqslant \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}\right)$$

$$+ \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1\right) + \frac{1}{2}, \qquad \text{(100b)}$$

where (h) follows from the fact that

$$\left(\sqrt{\overrightarrow{\mathrm{SNR}}_2} - \sqrt{\mathrm{INR}_{21}}\right)^2 \geqslant 0; \qquad \text{(101)}$$

and (i) follows from the fact that

$$\left(\sqrt{\overrightarrow{\mathrm{SNR}}_1} - \sqrt{\mathrm{INR}_{12}}\right)^2 \geqslant 0. \qquad \text{(102)}$$

From (99) and (100), 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

$$\delta_{2R}^* \leqslant \sup_{\rho' \in [0,1]}\left\{\min\left(\kappa_4(\rho'), \kappa_5(\rho')\right)\right\} - \left(a_{2,1}(0) + a_{1,2}\right)$$

$$\leqslant \sup_{\rho' \in [0,1]} \kappa_5(\rho') - \left(a_{2,1}(0) + a_{1,2}\right)$$

$$\leqslant \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}\right)$$

$$+ \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1\right) + \frac{1}{2}$$

$$- \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1\right)$$

$$- \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}} + 2\right) + 1$$

$$= \frac{3}{2}. \qquad \text{(103)}$$

From (99) and (100), 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

$$\delta_{2R}^* \leqslant \sup_{\rho' \in [0,1]}\left\{\min\left(\kappa_4(\rho'), \kappa_5(\rho')\right)\right\} - \left(a_{1,1} + a_{2,2}(0)\right)$$

$$\leqslant \sup_{\rho' \in [0,1]} \kappa_4(\rho') - \left(a_{1,1} + a_{2,2}(0)\right)$$

$$\leqslant \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}}\right)$$

$$+ \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_2 + \mathrm{INR}_{21} + 1\right) + \frac{1}{2}$$

$$- \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_2 + \mathrm{INR}_{21} + 1\right)$$

$$- \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}} + 2\right) + 1$$

$$= \frac{3}{2}. \qquad \text{(104)}$$

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:

$$\delta_{2R}^* \leqslant \sup_{\rho' \in [0,1]}\left\{\min\left(\kappa_4(\rho'), \kappa_5(\rho')\right)\right\}$$

$$- \left(a_{5,1}(0,0) + a_{5,2}(0,0)\right). \qquad \text{(105)}$$

To calculate (105), the cases 1.1.1, 1.1.2, and 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:

$$a_{5,1}(0,0) + a_{5,2}(0,0) = \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}} + \mathrm{INR}_{12} + 1\right)$$

$$+ \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}} + \mathrm{INR}_{21} + 1\right) - 1$$

$$\geqslant \frac{1}{2}\log\left(\mathrm{INR}_{12} + 1\right) - 1. \qquad \text{(106)}$$

From (100), (105), and (106), it follows that

$$\delta_{2R}^* \leqslant \sup_{\rho' \in [0,1]}\left\{\min\left(\kappa_4(\rho'), \kappa_5(\rho')\right)\right\}$$

$$- \left(a_{5,1}(0,0) + a_{5,2}(0,0)\right)$$

$$\leqslant \sup_{\rho' \in [0,1]} \kappa_5(\rho') - \left(a_{5,1}(0,0) + a_{5,2}(0,0)\right)$$

$$\leqslant \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}\right) + \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1\right)$$

$$+ \frac{1}{2} - \frac{1}{2}\log\left(\mathrm{INR}_{12} + 1\right) + 1$$

$$\leqslant \frac{1}{2}\log\left(2 + 1\right) + \frac{1}{2}\log\left(\mathrm{INR}_{12} + \mathrm{INR}_{12} + 1\right)$$

$$- \frac{1}{2}\log\left(\mathrm{INR}_{12} + 1\right) + \frac{3}{2}$$

$$\leqslant \frac{1}{2}\log\left(3\right) + 2. \qquad \text{(107)}$$

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

bounded as follows:

$$a_{5,1}(0,0)+a_{5,2}(0,0) = \frac{1}{2}\log\left(\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}}+\text{INR}_{12}+1\right)$$

$$+\frac{1}{2}\log\left(\frac{\overrightarrow{\text{SNR}}_2}{\text{INR}_{12}}+\text{INR}_{21}+1\right)-1$$

$$\geqslant\frac{1}{2}\log\left(\text{INR}_{21}+1\right)-1. \tag{108}$$

From (100), (105), and (108), it follows that

$$\delta_{2R}^* \leqslant \sup_{\rho'\in[0,1]}\left\{\min\left(\kappa_4(\rho'),\kappa_5(\rho')\right)\right\}$$

$$-\left(a_{5,1}(0,0)+a_{5,2}(0,0)\right)$$

$$\leqslant \sup_{\rho'\in[0,1]}\kappa_4(\rho')-\left(a_{5,1}(0,0)+a_{5,2}(0,0)\right)$$

$$\leqslant\frac{1}{2}\log\left(2+\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}}\right)+\frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_2+\text{INR}_{21}+1\right)$$

$$+\frac{1}{2}-\frac{1}{2}\log\left(\text{INR}_{21}+1\right)+1$$

$$\leqslant\frac{1}{2}\log\left(2+1\right)+\frac{1}{2}\log\left(\text{INR}_{21}+\text{INR}_{21}+1\right)$$

$$-\frac{1}{2}\log\left(\text{INR}_{21}+1\right)+\frac{3}{2}$$

$$\leqslant\frac{1}{2}\log\left(3\right)+2. \tag{109}$$

In Case 1.1.3, from (100), (105), and (106), it follows that

$$\delta_{2R}^* \leqslant \sup_{\rho'\in[0,1]}\left\{\min\left(\kappa_4(\rho'),\kappa_5(\rho')\right)\right\}$$

$$-\left(a_{5,1}(0,0)+a_{5,2}(0,0)\right)$$

$$\leqslant \sup_{\rho'\in[0,1]}\kappa_5(\rho')-\left(a_{5,1}(0,0)+a_{5,2}(0,0)\right)$$

$$\leqslant\frac{1}{2}\log\left(2+\frac{\overrightarrow{\text{SNR}}_2}{\text{INR}_{12}}\right)+\frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1+\text{INR}_{12}+1\right)$$

$$+\frac{1}{2}-\frac{1}{2}\log\left(\text{INR}_{12}+1\right)+1$$

$$\leqslant\frac{1}{2}\log\left(2+1\right)+\frac{1}{2}\log\left(\text{INR}_{12}+\text{INR}_{12}+1\right)$$

$$-\frac{1}{2}\log\left(\text{INR}_{12}+1\right)+\frac{3}{2}$$

$$\leqslant\frac{1}{2}\log\left(3\right)+2. \tag{110}$$

Then, from (103), (104), (107), (109), and (110), it follows that in Case 1.1:

$$\delta_{2R}* \leqslant 2+\frac{1}{2}\log\left(3\right). \tag{111}$$

3) Calculation of δ_{3R_1}. From (86d) and considering the choice $\rho = 0$, $\mu_1 = 0$, and $\mu_2 = 0$, it follows that

$$\delta_{3R_1}^* \leqslant \sup_{\rho'\in[0,1]}\kappa_{7,1}(\rho')-\left(a_{1,1}+a_{7,1}(0,0,0)+a_{5,2}(0,0)\right). \tag{112}$$

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

$$a_{1,1}+a_{7,1}(0,0,0)+a_{5,2}(0,0) = \frac{1}{2}\log\left(\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}}+2\right)$$

$$+\frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1+\text{INR}_{12}+1\right)$$

$$+\frac{1}{2}\log\left(\frac{\overrightarrow{\text{SNR}}_2}{\text{INR}_{12}}+\text{INR}_{21}+1\right)-\frac{3}{2}$$

$$\geqslant\frac{1}{2}\log\left(\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}}+2\right)+\frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1+\text{INR}_{12}+1\right)$$

$$+\frac{1}{2}\log\left(\text{INR}_{21}+1\right)-\frac{3}{2}. \tag{113}$$

When the term $\kappa_{7,1}(\rho')$ is active in the converse region, for a given $\rho' \in [0,1]$, it can be upper bounded by the sum $\kappa_{1,1}(\rho') + \kappa_4(\rho')$, i.e., the sum of the single rate and sum-rate outer bounds respectively, and this can be upper bounded as follows:

$$\kappa_{7,1}(\rho')\leqslant\kappa_{1,1}(\rho')+\kappa_4(\rho')$$

$$\leqslant\frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1+\text{INR}_{12}+1\right)+\frac{1}{2}\log\left(2+\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}}\right)$$

$$+\frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_2+\text{INR}_{21}+1\right)+1$$

$$\leqslant\frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1+\text{INR}_{12}+1\right)+\frac{1}{2}\log\left(2+\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}}\right)$$

$$+\frac{1}{2}\log\left(\text{INR}_{21}+\text{INR}_{21}+1\right)+1$$

$$\leqslant\frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1+\text{INR}_{12}+1\right)+\frac{1}{2}\log\left(2+\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}}\right)$$

$$+\frac{1}{2}\log\left(\text{INR}_{21}+1\right)+\frac{3}{2}. \tag{114}$$

From (112), (113) and (114), it follows that in Case 1.1:

$$\delta_{3R_1}^* \leqslant\frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1+\text{INR}_{12}+1\right)+\frac{1}{2}\log\left(2+\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}}\right)$$

$$+\frac{1}{2}\log\left(\text{INR}_{21}+1\right)+\frac{3}{2}-\frac{1}{2}\log\left(\frac{\overrightarrow{\text{SNR}}_1}{\text{INR}_{21}}+2\right)$$

$$-\frac{1}{2}\log\left(\overrightarrow{\text{SNR}}_1+\text{INR}_{12}+1\right)-\frac{1}{2}\log\left(\text{INR}_{21}+1\right)$$

$$+\frac{3}{2}$$

$$=3. \tag{115}$$

The same procedure holds in the calculation of $\delta_{3R_2}^*$ and it yields:

$$\delta_{3R_2}^* \leqslant 3. \tag{116}$$

Therefore, in Case 1.1, from (84), (98), (97), (111), (115) and (116) it follows that

$$\delta=\sup\left(\delta_{R_1}^*,\delta_{R_2}^*,\frac{\delta_{2R}^*}{2},\frac{\delta_{3R_1}^*}{3},\frac{\delta_{3R_2}^*}{3}\right)\leqslant\frac{3}{2}. \tag{117}$$

This completes the calculation of the gap in Case 1.1. Applying the same procedure to all the other cases listed above yields that $\delta \leqslant 4.4$ 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 Inform. Theory Workshop (ITW)*, 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 *Proc. IEEE Inform. Theory Workshop (ITW)*, Cambridge, UK., Sept. 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 Commun. and Inform. Theory*, vol. 7, no. 1, pp. 1–134, 2010.

[5] 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.

[6] A. Sahai, V. Aggarwal, M. Yüksel, and A. Sabharwal, "On channel output feedback in deterministic interference channels," in *Proc. IEEE Inform. Theory Workshop (ITW)*, Taormina, Italy, Oct. 2009.

[7] 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, Nov. 2013.

[8] V. Quintero, S. M. Perlaza, I. Esnaola, and J.-M. Gorce, "When does output feedback enlarge the capacity of the interference channel?" *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 615–628, Feb. 2018.

[9] 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.

[10] V. Quintero, "Noisy channel-output feedback in the interference channel," Ph.D. dissertation, Université de Lyon, Lyon, France, Dec. 2017.

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

[12] 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, Jan. 1981.

[13] 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.

[14] F. M. J. Willems, "Information theoretical results for multiple access channels," Ph.D. dissertation, Katholieke Universiteit, Leuven, Belgium, Oct. 1982.

[15] F. M. J. Willems and E. C. Van Der Meulen, "The discrete memoryless multiple-access channel with cribbing encoders," *IEEE Trans. Inf. Theory*, vol. 31, no. 3, pp. 313–327, May 1985.

[16] D. Tuninetti, "On interference channel with generalized feedback (IFC-GF)," in *Proc. of Int. Symp. on Inform. Theory (ISIT)*, Nice, France, June 2007, pp. 2661–2665.

[17] 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.

[18] A. Vahid, C. Suh, and A. S. Avestimehr, "Interference channels with rate-limited feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2788–2812, May 2012.

[19] M. Ashraphijuo, V. Aggarwal, and X. Wang, "On the symmetric K-user linear deterministic interference channels with limited feedback," in *Proc. 52nd Annu. Allerton Conf. on Communication, Control, and Computing*, Monticello, IL, USA, Oct. 2014.

[20] ——, "On the symmetric K-user interference channels with limited feedback," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 6969–6985, Oct. 2016.

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

[22] Z. Ram, *Lattice Coding for Signals and Networks*. University Printing House, Cambridge CB2 8BS, UK: Cambridge University Press, 2014.

[23] C. Karakus, I.-H. Wang, and S. Diggavi, "Gaussian interference channel with intermittent feedback," *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 4663–4699, Sept. 2015.

[24] M. Gastpar and G. Kramer, "On noisy feedback for interference channels," in *Proc. 40th Annu. Asilomar Conf. on Signals, Syst., and Comput.*, Asilomar, USA., Nov. 2006, pp. 216–220.

[25] 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, July 2015.

[26] ——, "The capacity region of the symmetric linear deterministic interference channel with partial feedback," in *Proc. 50th Annu. Allerton Conf. on Communication, Control, and Computing*, Monticello, IL, USA, Oct. 2012.

[27] 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.

[28] S. M. Perlaza, R. Tandon, and H. V. Poor, "Symmetric decentralized interference channels with noisy feedback," in *Proc. IEEE Intl. Symp. on Inform. Theory (ISIT)*, Honolulu, HI, USA, June 2014.

[29] V. Quintero, S. M. Perlaza, J.-M. Gorce, and H. V. Poor, "Nash region of the linear deterministic interference channel with noisy output feedback," in *Proc. IEEE Int. Symp. on Inform. Theory (ISIT)*, Aachen, Germany, June 2017.

[30] V. Quintero, S. M. Perlaza, and J.-M. Gorce, "On the efficiency of Nash equilibria in the interference channel with noisy feedback," in *Proc. Eur. Wireless Conf.*, Dresden, Germany, May 2017.

[31] V. M. Pabhakaran and P. Viswanath, "Interference channel with source cooperation," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 156–186, Jan. 2011.

[32] G. Kramer, "Feedback strategies for white Gaussian interference networks," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1423–1438, June 2002.

[33] D. Tuninetti, "An outer bound region for interference channels with generalized feedback," in *IEEE Inform. Theory and Appl. Workshop (ITA)*, San Diego, CA, USA., Feb. 2010, pp. 1–5.

[34] ——, "An outer bound for the memoryless two-user interference channel with general cooperation," in *IEEE Inform. Theory Workshop (ITW)*, Lausanne, Switzerland, Sept. 2012, pp. 217–221.

[35] 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.

[36] G. Bresler and D. N. C. Tse, "The two user Gaussian interference channel: A deterministic view," *Eur. Trans. on Telecommun.*, vol. 19, no. 4, pp. 333–354, Apr. 2008.

[37] H.-F. Chong, M. Motani, and H. K. Garg, "A comparison of two achievable rate regions for the interference channel," in *ITA Workshop*, San Diego, CA, USA., Feb. 2006.

[38] A. Host-Madsen, "Capacity bounds for cooperative diversity," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1522–1544, Apr. 2006.

[39] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley-Interscience, 1991.