Decentralized Interference Channels with Noisy Output Feedback

Victor Quintero, Samir Perlaza, Jean-Marie Gorce, H Vincent Poor

Decentralized Interference Channels with Noisy Output Feedback

Victor Quintero, Samir M. Perlaza, Jean-Marie Gorce, H. Vincent Poor

RESEARCH REPORT
N° 9011
January 2017
Project-Team Socrate

ISSN 0249-6399 ISRN INRIA/RR--9011--FR+ENG

Decentralized Interference Channels with Noisy Output Feedback

Victor Quintero, Samir M. Perlaza, Jean-Marie Gorce, H. Vincent Poor

Project-Team Socrate

Research Report n° 9011 — January 2017 — 37 pages

Abstract: In this research report, the η-Nash equilibrium (η-NE) region of the two-user linear deterministic interference channel with noisy channel-output feedback is characterized for all $\eta > 0$ arbitrarily small. It also characterizes the η-Nash achievable region of the two-user Gaussian interference with noisy channel output feedback for all $\eta > 1$. The η-NE region, a subset of the capacity region, contains the set of all achievable information rate pairs that are stable in the sense of an η-NE. More specifically, given an η-NE coding scheme, there does not exist an alternative coding scheme for either transmitter-receiver pair that increases the individual rate by more than η bits per channel use. Existing results such as the η-NE region of the linear deterministic interference channel and the Gaussian interference channel without feedback and with perfect output feedback are obtained as particular cases of the result presented in this research report.

Key-words: Decentralized interference channel, noisy channel-output feedback, Nash equilibrium.

Victor Quintero, Samir M. Perlaza and Jean-Marie Gorce are with the CITI Laboratory, a joint lab between the Institut National de Recherche en Informatique et en Automatique (INRIA), Université de Lyon and Institut National de Sciences Apliquées (INSA) de Lyon. 6 Av. des Arts 69621 Villeurbanne, France. ({victor.quintero-florez, samir.perlaza}@inria.fr). H. Vincent Poor is with the Department of Electrical Engineering at Princeton University, Princeton, NJ. (poor@princeton.edu). Victor Quintero is also with Universidad del Cauca, Popayán, Colombia. This research was supported in part by the European Commission under Marie Skłodowska-Curie Individual Fellowship No. 659316; in part by the Administrative Department of Science, Technology, and Innovation of Colombia (Colciencias), fellowship No. 617-2013; and in part by the U. S. National Science Foundation under Grants CCF-1420575 and ECCS-1343210.

RESEARCH CENTRE GRENOBLE – RHÔNE-ALPES

Inovallée
655 avenue de l'Europe Montbonnot
38334 Saint Ismier Cedex

Résumé : Ce rapport de recherche présente la région d' équilibre η-Nash (η-Nash) du canal linéaire déterministe à interférences avec rétroalimentation dégradée par bruit additif pour tout $\eta > 0$ arbitrairement petits. Il caractérise également la région d' équilibre η-Nash atteignable du canal Gaussien à interférences avec rétroalimentation dégradée par bruit additif pour tout $\eta > 1$. La région d' équilibre η-Nash, un sous-ensemble de la région de capacité, contient l'ensemble de toutes les paires de taux d'information réalisables qui sont stables au sens d'un η-NE. Plus précisément, étant donné un schéma de codage η-NE, il n'existe pas de schéma de codage alternatif poor l'une ou l'autre paire émetteur-récepteur qui augmente le taux individuel de plus de η bits par utilisation du canal. Les résultats existants, tels que la région d' équilibre η-NE du canal linéaire déterministe à interférences et du canal Gaussien à interférences sans rétroalimentation et avec rétroalimentation parfaite, sont obtenus comme cas particuliers du résultat présenté dans ce rapport de recherche.

Mots-clés : Canal à interférences décentralisé, rétroalimentation dégradée, équilibre de Nash.

Contents

1 Notation

Throughout this research report, sets are denoted with uppercase calligraphic letters, e.g. $\mathcal{X}$. Random variables are denoted by uppercase letters, e.g., X. The realizations and the set of events from which the random variable X takes values are respectively denoted by x and $\mathcal{X}$. The probability distribution of X over the set $\mathcal{X}$ is denoted P_X. Whenever a second random variable Y is involved, $P_{X\,Y}$ 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 of the random variable X. The logarithm function log is assumed to be base 2.

2 Problem Formulation

2.1 Channel Model

Consider the two-user Gaussian interference channel (GIC) with noisy channel-output feedback in Figure 1. 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 $\mathcal{W}_i = \{1, 2, \ldots, 2^{N_i R_i}\}$, where N_i denotes the fixed block-length in channel uses and R_i the transmission rate in bits per channel use. At each block, transmitter i sends the codeword $\boldsymbol{X}_i = (X_{i,1}, X_{i,2}, \ldots, X_{i,N_i})^{\mathsf{T}} \in \mathcal{X}_i^{N_i}$, where $\mathcal{X}_i$ and $\mathcal{X}_i^{N_i}$ are respectively the channel-input alphabet and the codebook of transmitter i.

The channel coefficient from transmitter j to receiver i is denoted by h_{ij}; the channel coefficient from transmitter i to receiver i is denoted by $\overrightarrow{h}_{ii}$; 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 $n \in \{1, 2, \ldots, N\}$, with

$$N = \max(N_1, N_2), \tag{1}$$

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

$$\overrightarrow{Y}_{i,n} = \overrightarrow{h}_{ii} X_{i,n} + h_{ij} X_{j,n} + \overrightarrow{Z}_{i,n}, \tag{2}$$

where $X_{i,n} = 0$ for all $n > N_i$ and $\overrightarrow{Z}_{i,n}$ 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,

$$\overleftarrow{Y}_{i,n} = \begin{cases} \overleftarrow{Z}_{i,n} & \text{for } n \in \{1, 2, \ldots, d\} \\ \overleftarrow{h}_{ii} \overrightarrow{Y}_{i,n-d} + \overleftarrow{Z}_{i,n}, & \text{for } n \in \{d+1, d+2, \ldots, N\}, \end{cases} \tag{3}$$

where $\overleftarrow{Z}_{i,n}$ is a real Gaussian random variable with zero mean and unit variance that represents the noise in the feedback link of transmitter-receiver pair i. The random variables $\overrightarrow{Z}_{i,n}$ and

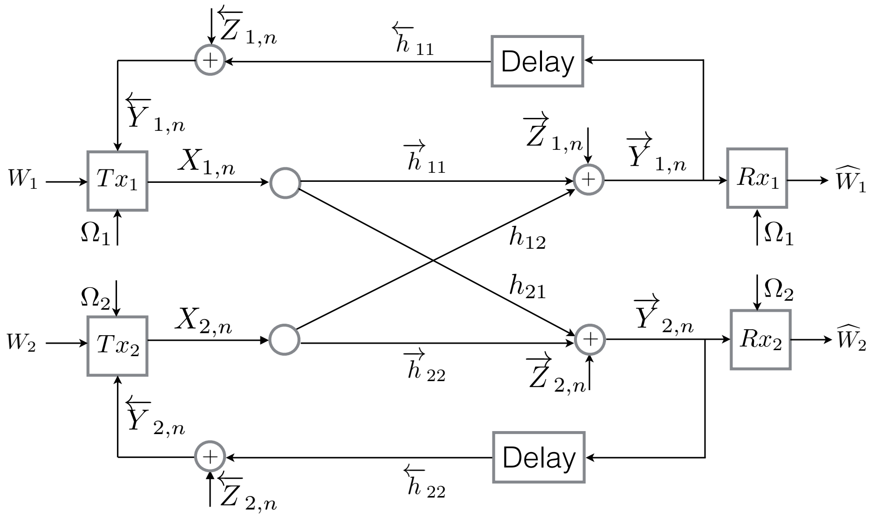

Figure 1: Two-User Gaussian interference channel with noisy channel-output feedback at channel use n.

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

$$X_{i,1} = f_{i,1}^{(N)}\left(W_i, \Omega_i\right), \text{ and} \tag{4a}$$

$$X_{i,n} = f_{i,n}^{(N)}\left(W_i, \Omega_i, \overleftarrow{Y}_{i,1}, \overleftarrow{Y}_{i,2}, \ldots, \overleftarrow{Y}_{i,n-1}\right), \tag{4b}$$

where Ω_i is an additional index randomly generated. The index Ω_i is assumed to be known by both transmitter i and receiver i, while totally unknown by transmitter j and receiver j.

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

$$\frac{1}{N_i} \sum_{n=1}^{N_i} \mathbb{E}\left(X_{i,n}^2\right) \leq 1, \tag{5}$$

where the expectation is taken over the joint distribution of the message indexes W_1, W_2, the random indices Ω_1 and Ω_2, and the noise terms, i.e., $\overrightarrow{Z}_1$, $\overrightarrow{Z}_2$, $\overleftarrow{Z}_1$, and $\overleftarrow{Z}_2$. The dependence of $X_{i,n}$ on W_1, W_2, Ω_1, Ω_2 and the previously observed noise realizations is due to the effect of feedback as shown in (3) and (4).

Assume that during a given communication, transmitter i sends $T_i \in \mathbb{N}$ blocks, each of N_i channel uses. Let $T = \max(T_1, T_2)$. Hence, the decoder of receiver i is defined by a deterministic function $\psi_i^{(N,T)} : \mathbb{R}_i^{NT} \to \mathcal{W}_i^T$. At the end of the communication, receiver i uses the vector $\left(\overrightarrow{Y}_{i,1}, \overrightarrow{Y}_{i,2}, \ldots, \overrightarrow{Y}_{i,NT}\right)^{\mathsf{T}}$ to obtain an estimate of the message indices

$$\left(\widehat{W}_i^{(1)}, \widehat{W}_i^{(2)}, \ldots, \widehat{W}_i^{(T)}\right) = \psi_i^{(N,T)}\left(\overrightarrow{Y}_{i,1}, \overrightarrow{Y}_{i,2}, \ldots, \overrightarrow{Y}_{i,NT}\right), \tag{6}$$

where $\widehat{W}_i^{(t)}$ is an estimate of the message index sent during block $t \in \{1, 2, \ldots, T\}$. The decoding error probability in the two-user Gaussian interference channel with noisy channel-output feedback (GIC-NOF) during block t, denoted by $P_e^{(t)}$, is given by

$$P_e^{(t)}{=}\max\left(\Pr\left(\widehat{W_1}^{(t)}{\neq} W_1^{(t)}\right), \Pr\left(\widehat{W_2}^{(t)}{\neq} W_2^{(t)}\right)\right). \tag{7}$$

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

Definition 1 (Achievable Rate Pairs) A rate pair $(R_1, R_2) \in \mathbb{R}_+^2$ is achievable if there exists at least one pair of codebooks $\mathcal{X}_1^{N_1}$ and $\mathcal{X}_2^{N_2}$ with codewords of length N_1 and N_2, respectively, and the corresponding encoding functions $f_{1,1}^{(N)}, f_{1,2}^{(N)}, \ldots, f_{1,N_1}^{(N)}$ and $f_{2,1}^{(N)}, f_{2,2}^{(N)}, \ldots, f_{2,N_2}^{(N)}$ such that the decoding error probability $P_e^{(t)}$ can be made arbitrarily small by letting the block-lengths N_1 and N_2 grow to infinity, for all blocks $t \in \{1, 2, \ldots, T\}$, respectively.

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

$$\overrightarrow{\mathrm{SNR}}_i{=}\overrightarrow{h}_{ii}^2, \tag{8}$$

$$\mathrm{INR}_{ij}{=}h_{ij}^2 \text{ and} \tag{9}$$

$$\overleftarrow{\mathrm{SNR}}_i{=}\overleftarrow{h}_{ii}^2\left(\overrightarrow{h}_{ii}^2 + 2\overrightarrow{h}_{ii}h_{ij} + h_{ij}^2 + 1\right). \tag{10}$$

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

The aim of transmitter i is to autonomously choose its transmit-receive configuration s_i in order to maximize its achievable rate R_i. More specifically, the transmit-receive configuration s_i can be described in terms of the block-length N_i, the codebook $\mathcal{X}_i^{N_i}$, the encoding functions $f_{i,1}^{(N)}, f_{i,2}^{(N)}, \ldots, f_{i,N_i}^{(N)}$, the decoding function $\psi_i^{(N,T)}$, the number of blocks T_i, etc. Note that the rate achieved by receiver i depends on both configurations s_1 and s_2 due to mutual interference. This reveals the competitive interaction between both links in the decentralized interference channel. The following section models this interaction using tools from game theory.

2.2 Game Formulation

The competitive interaction of the two transmitter-receiver pairs in the interference channel can be modeled by the following game in normal-form:

$$\mathcal{G} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}\right). \tag{11}$$

The set $\mathcal{K} = \{1, 2\}$ is the set of players, that is, the set of transmitter-receiver pairs. The sets $\mathcal{A}_1$ and $\mathcal{A}_2$ are the sets of actions of player 1 and 2, respectively. An action of a player $i \in \mathcal{K}$, which is denoted by $s_i \in \mathcal{A}_i$, is basically its transmit-receive configuration as described above. The utility function of player i is $u_i : \mathcal{A}_1 \times \mathcal{A}_2 \to \mathbb{R}_+$ and it is defined as the achieved rate of transmitter i,

$$u_i(s_1, s_2) = \begin{cases} R_i(s_1, s_2), & \text{if } \quad P_e^{(t)} < \epsilon, \forall t \in \{1, 2, \ldots, T_i\} \\ 0, & \text{otherwise,} \end{cases} \tag{12}$$

where $\epsilon > 0$ is an arbitrarily small number and $R_i(s_1, s_2)$ denotes a transmission rate achievable with the configurations s_1 and s_2. Often, the rate $R_i(s_1, s_2)$ is written as R_i for the sake of simplicity. However, every non-negative achievable rate is associated with the particular transmit-receive configuration pair (s_1, s_2) that achieves it. It is worth noting that there might exist several transmit-receive configurations that achieve the same rate pair (R_1, R_2) and distinction between the different transmit-receive configuration is made only when needed.

A class of transmit-receive configurations $\boldsymbol{s}^* = (s_1^*, s_2^*) \in \mathcal{A}_1 \times \mathcal{A}_2$ that are particularly important in the analysis of this game is referred to as the set of η-Nash equilibria (η-NE). This type of configurations satisfy the following definition.

Definition 2 (η-Nash equilibrium [1]) In the game $\mathcal{G} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}\right)$, an action profile (s_1^*, s_2^*) is an η-Nash equilibrium if for all $i \in \mathcal{K}$ and for all $s_i \in \mathcal{A}_i$, there exits an $\eta > 0$ such that

$$u_i(s_i, s_j^*) \leqslant u_i(s_i^*, s_j^*) + \eta. \tag{13}$$

From Definition 2, it becomes clear that if (s_1^*, s_2^*) is an η-Nash equilibrium, then none of the transmitters can increase its own transmission rate more than η bits per channel use by changing its own transmit-receive configuration and keeping the average bit error probability arbitrarily close to zero. Thus, at a given η-NE, every transmitter achieves a utility (transmission rate) that is η-close to its maximum achievable rate given the transmit-receive configuration of the other transmitter. Note that if $\eta = 0$, then the classical definition of Nash equilibrium is obtained [2]. The relevance of the notion of equilibrium is that at any NE, every transmitter-receiver pair's configuration is optimal with respect to the configuration of the other transmitter-receiver pairs.

The set of rate pairs that can be achieved at an NE is known as the Nash region.

Definition 3 (η-Nash Region) Let $\eta > 0$. An achievable rate pair (R_1, R_2) is said to be in the η-Nash region of the game $\mathcal{G} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}\right)$ if there exists a pair $(s_1^*, s_2^*) \in \mathcal{A}_1 \times \mathcal{A}_2$ that is an η-NE and the following holds:

$$u_1(s_1^*, s_2^*) = R_1 \quad \text{and} \quad u_2(s_1^*, s_2^*) = R_2. \tag{14}$$

Following along the same lines in [3], if there exists a strategy pair (s_1, s_2) that achieves a rate pair (R_1, R_2) using codes of block lengths N_1 and N_2 respectively, then there exists a strategy pair (s_1', s_2') that achieves the same rate pair using the same block length for both users, e.g., $N = \max(N_1, N_2)$. The resulting probability of error with (s_1', s_2') is smaller than or equal to the probability of error obtained by the strategy pair (s_1, s_2). For this reason, without loss of generality, the same block length is considered for both users in the remaining of this report.

3 Linear Deterministic Interference Channel with Noisy Channel Output Feedback

3.1 Channel Model

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

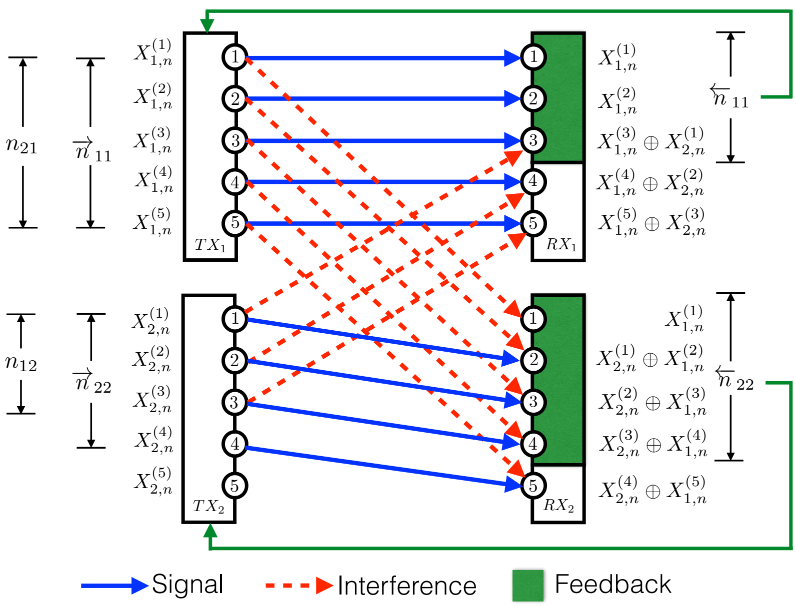

Figure 2: Two-user linear deterministic interference channel with noisy channel-output feedback at channel use n.

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

$$q = \max\left(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}\right), \tag{15}$$

and $N_i \in \mathbb{N}$ the block-length. At receiver i, the channel-output $\overrightarrow{\boldsymbol{Y}}_{i,n}$ at channel use n, with $n \in \{1, 2, \ldots, \max\left(N_1, N_2\right)\}$, is also a q-dimensional binary vector $\overrightarrow{\boldsymbol{Y}}_{i,n} = \left(\overrightarrow{Y}^{(1)}_{i,n}, \overrightarrow{Y}^{(2)}_{i,n}, \ldots, \overrightarrow{Y}^{(q)}_{i,n}\right)^{\mathsf{T}}$. The input-output relation during channel use n is given by

$$\overrightarrow{\boldsymbol{Y}}_{i,n} = \boldsymbol{S}^{q-\overrightarrow{n}_{ii}} \boldsymbol{X}_{i,n} + \boldsymbol{S}^{q-n_{ij}} \boldsymbol{X}_{j,n}, \tag{16}$$

and the feedback signal $\overleftarrow{\boldsymbol{Y}}_{i,n}$ available at transmitter i at the end of channel use n is a vector containing the $\min\left(\overleftarrow{n}_{ii}, \max(\overrightarrow{n}_{ii}, n_{ij})\right)$ least significant bits of $\boldsymbol{S}^{\left(\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii}\right)^+} \overrightarrow{\boldsymbol{Y}}_{i,n-d}$. That is,

$$\left(\left(0, \ldots, 0\right), \overleftarrow{\boldsymbol{Y}}^{\mathsf{T}}_{i,n}\right)^{\mathsf{T}} = \boldsymbol{S}^{\left(\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii}\right)^+} \overrightarrow{\boldsymbol{Y}}_{i,n-d}, \tag{17}$$

where d is a finite delay, additions and multiplications are defined over the binary field, $\boldsymbol{S}$ is a

$q \times q$ lower shift matrix of the form:

$$\boldsymbol{S} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}, \tag{18}$$

and the vector $(0,\ldots,0)$ in (17) is a $q - \min\left(\overleftarrow{n}_{ii}, \max(\overrightarrow{n}_{ii}, n_{ij})\right)$ dimensional vector.

The parameters $\overrightarrow{n}_{ii}$, $\overleftarrow{n}_{ii}$ and n_{ij} correspond to $\left\lfloor \frac{1}{2}\log_2\left(\overrightarrow{\mathrm{SNR}}_i\right)\right\rfloor$, $\left\lfloor \frac{1}{2}\log_2\left(\overleftarrow{\mathrm{SNR}}_i\right)\right\rfloor$ and $\left\lfloor \frac{1}{2}\log_2\left(\mathrm{INR}_{ij}\right)\right\rfloor$ respectively, where $\overrightarrow{\mathrm{SNR}}_i$, $\overleftarrow{\mathrm{SNR}}_i$ and INR_{ij} are parameters of the two-user GIC-NOF in (8), (9) and (10), respectively. The existing connections between the linear deterministic model and the Gaussian model are thoroughly described in [4].

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

$$\boldsymbol{X}_{i,1} = f_{i,1}^{(N)}\left(W_i, \Omega_i\right) \text{ and} \tag{19a}$$

$$\boldsymbol{X}_{i,n} = f_{i,n}^{(N)}\left(W_i, \Omega_i, \overleftarrow{\boldsymbol{Y}}_{i,1}, \overleftarrow{\boldsymbol{Y}}_{i,2}, \ldots, \overleftarrow{\boldsymbol{Y}}_{i,n-1}\right), \tag{19b}$$

where Ω_i is an additional index randomly generated. The index Ω_i is assumed to be known by both transmitter i and receiver i, while totally unknown by transmitter j and receiver j.

Assume that during a given communication, transmitter i sends $T_i \in \mathbb{N}$ blocks, each of N_i channel uses. Let $T = \max(T_1, T_2)$. Hence, the decoder of receiver i is defined by a deterministic function $\psi_i^{(N,T)} : \{0,1\}^{q\times N\times T} \rightarrow \mathcal{W}_i^T$. At the end of the communication, receiver i uses the sequence $\left(\overrightarrow{\boldsymbol{Y}}_{i,1}, \overrightarrow{\boldsymbol{Y}}_{i,2}, \ldots, \overrightarrow{\boldsymbol{Y}}_{i,NT}\right)$ to obtain an estimate of the message index.

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

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-NOF with parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$, and $\overleftarrow{n}_{22}$.

3.2 Nash Region of the Two-User Linear Deterministic Interference Channel with Noisy Channel-Output Feedback

This section characterizes the η-NE region (Definition 3) of the two-user linear deterministic interference channel with noisy channel-output feedback (LD-IC-NOF).

The η-NE region of the two-user LD-IC-NOF, given the fixed parameters $\left(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}\right) \in \mathbb{N}^6$, is denoted by $\mathcal{N}_\eta\left(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}\right)$. The region $\mathcal{N}_\eta\left(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}\right)$ is characterized in terms of two regions: the capacity region, denoted by $\mathcal{C}(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ and a convex closed region, denoted by $\mathcal{B}_\eta(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$. In the following, the analysis of these regions is made for fix parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$, and $\overleftarrow{n}_{22}$. Then, the tuple $\left(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}\right)$ is used only when needed.

The capacity region $\mathcal{C}$ of the two-user LD-IC-NOF is described in Theorem 1 in [5], which is a generalization of the cases with and without perfect channel-output feedback, studied respectively in [6] and [7]. For all $\eta > 0$, the convex region $\mathcal{B}_\eta$ is defined as follows:

$$\mathcal{B}_\eta = \Big\{ (R_1, R_2) : L_i \leqslant R_i \leqslant U_i, \text{ for all } i \in \{1,2\} \Big\}, \tag{20}$$

where,

$$L_i = \left((\overrightarrow{n}_{ii} - n_{ij})^+ - \eta \right)^+ \text{ and} \tag{21a}$$

$$U_i = \max(\overrightarrow{n}_{ii}, n_{ij}) - \Bigg(\min\left((\overrightarrow{n}_{jj} - n_{ji})^+, n_{ij} \right) - \left(\min\left((\overrightarrow{n}_{jj} - n_{ij})^+, n_{ji} \right) - (\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+ \right)^+ \Bigg)^+ + \eta, \tag{21b}$$

with $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$. Theorem 1 uses the region $\mathcal{B}_\eta$ in (20) and the capacity region $\mathcal{C}$ to describe the η-NE region.

Theorem 1 The η-Nash region of the two-user LD-IC-NOF ($\mathcal{N}_\eta$) with parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$, $\overleftarrow{n}_{22}$, is

$$\mathcal{N}_\eta = \mathcal{C} \cap \mathcal{B}_\eta. \tag{22}$$

The proof of Theorem 1 is presented in Appendix C.

The following describes some interesting observations from Theorem 1 using a particular example of LD-IC-NOF as shown in Figure 3.

Let transmitter-receiver pair 1 be in very weak interference regime and transmitter-receiver pair 2 be in moderate interference regime ($\overrightarrow{n}_{11} = 7$, $\overrightarrow{n}_{22} = 5$, $n_{12} = 2$, and $n_{21} = 4$). Figures 3a, 3b, 3c, and 3d respectively show the capacity region without channel-output feedback, i.e., $\mathcal{C}(7,5,2,4,0,0)$; the capacity region with noisy channel-output feedback, i.e., $\mathcal{C}(7,5,2,4,\overleftarrow{n}_{11},\overleftarrow{n}_{22})$; the η-Nash region without channel output feedback, i.e., $\mathcal{N}_\eta(7,5,2,4,0,0)$; and the η-Nash region with noisy channel-output feedback, i.e. $\mathcal{N}_\eta(7,5,2,4,\overleftarrow{n}_{11},\overleftarrow{n}_{22})$ for different values in the feedback parameters. It is worth noting here that the capacity and the η-Nash regions can not be enlarged for certain values in the feedback parameters (see Figure 3a) and the capacity and the η-Nash region can be enlarged for certain values in the feedback parameters (see Figure 3b, Figure 3c, and Figure 3d). Note that when $\overleftarrow{n}_{11} = 6$, for all $\overleftarrow{n}_{22} \in \{0,1,2\}$ (Figure 3b), it follows that $\mathcal{N}_\eta(7,5,2,4,6,\overleftarrow{n}_{22}) = \mathcal{N}_\eta(7,5,2,4,6,0)$. That is, there exists a threshold $\overleftarrow{n}^*_{ii}$ for which feedback from receiver i to transmitter i with $\overleftarrow{n}_{ii} \leqslant \overleftarrow{n}^*_{ii}$ does not have any impact in the η-NE. Similarly, when $\overleftarrow{n}_{11} = 7$, for all $\overleftarrow{n}_{22} \in \{3,4,\ldots\}$, it follows that $\mathcal{N}_\eta(7,5,2,4,6,\overleftarrow{n}_{22}) = \mathcal{N}_\eta(7,5,2,4,6,3)$. This implies the existence of a threshold $\overleftarrow{n}^\dagger_{ii}$ for which feedback from receiver i to transmitter i with $\overleftarrow{n}_{ii} > \overleftarrow{n}^\dagger_{ii}$ does not enlarge the η-region. Despite the fact that the exact values of the thresholds $\overleftarrow{n}^*_{ii}$ and $\overleftarrow{n}^\dagger_{ii}$ are implicit in (22), the calculation of the exact values is beyond of the scope of this research report. Nonetheless, the following observations are highlighted. Note that the bound $R_i \leqslant U_i$ is not always active. For instance, when $\overrightarrow{n}_{jj} \leqslant \min(n_{ji}, n_{ij})$, then $U_i = \max(\overrightarrow{n}_{ii}, n_{ij})$, which is redundant with the bounds given by the capacity region $\mathcal{C}$ (see Theorem 1 in [5]). When $\overrightarrow{n}_{jj} > \max(n_{ji}, n_{ij})$, the bound $R_i \leqslant U_i$ might be active. In the case it is active, the following is a necessary condition to observe a larger η-NE with respect to the case in which feedback is not available:

$$\left(\min\left((\overrightarrow{n}_{jj} - n_{ij})^+, n_{ji} \right) - (\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+ \right)^+ > 0,$$

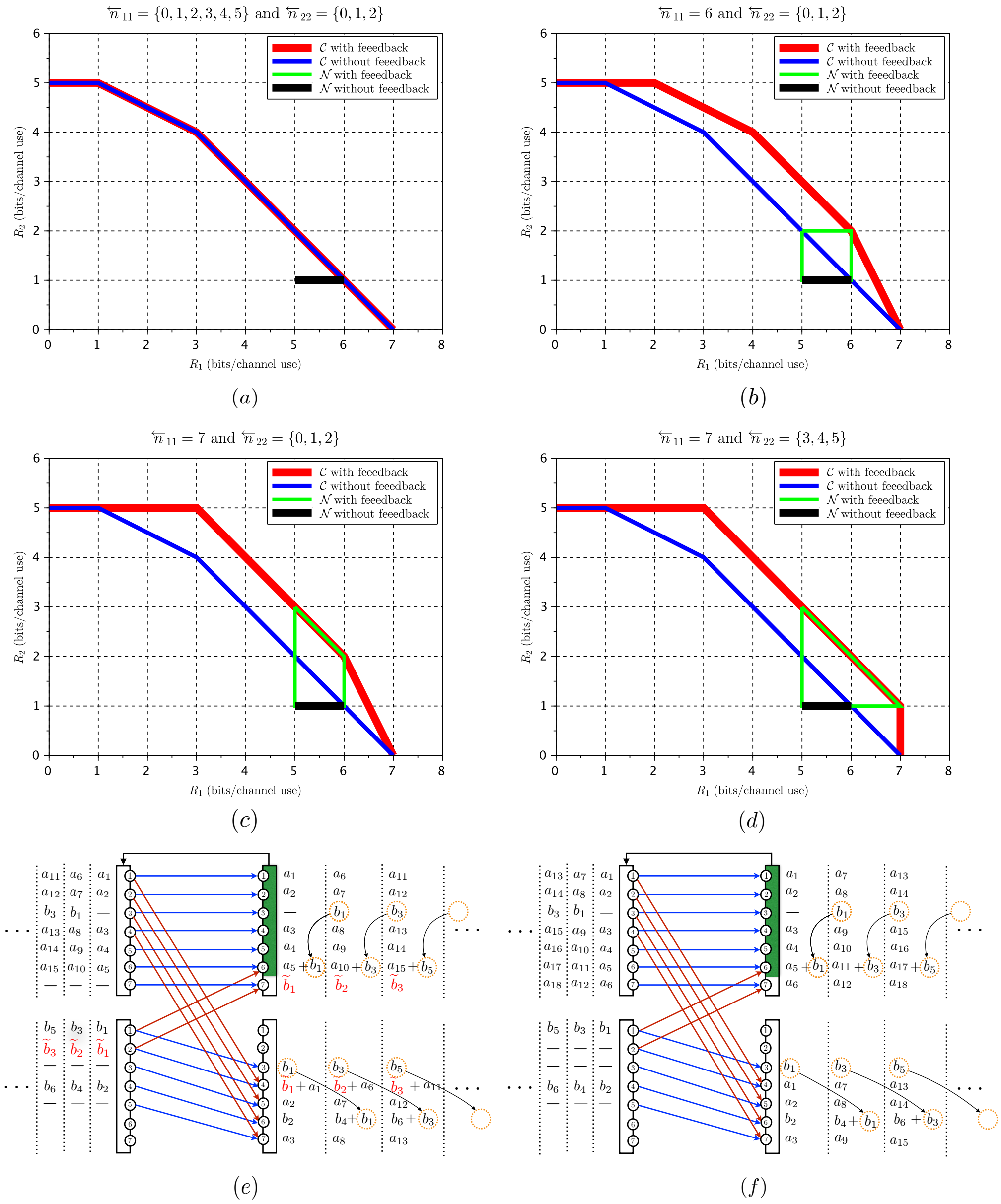

Figure 3: Transmitter-receiver 1 in the very weak interference regime ($\alpha_1 \leqslant \frac{1}{2}$) and transmitter-receiver 2 in the moderate interference regime ($\frac{2}{3} \leqslant \alpha_2 \leqslant 1$). In (a), (b), (c), and (d) Illustration of $\mathcal{C}(7, 5, 2, 4, 0, 0)$ (blue line), $\mathcal{C}(7, 5, 2, 4, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ (red line), $\mathcal{N}_\eta(7, 5, 2, 4, 0, 0)$ (black line), and $\mathcal{N}_\eta(7, 5, 2, 4, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ (green line) for different values in the feedback parameters. In (e) illustration of the achievability scheme for the equilibrium rate pair $(5, 2)$ of the η-NE region $\mathcal{N}_\eta(7, 5, 2, 4, 6, 0)$. Noisy channel-output feedback in transmitter-receiver 1. Note that $\widetilde{b}_1$, $\widetilde{b}_2$, and $\widetilde{b}_3$ are known at receiver 2, then $\widetilde{b}_1$, $\widetilde{b}_2$, and $\widetilde{b}_3$ do not produce any interference at receiver 2. In (f) illustration of the achievability scheme for the equilibrium rate pair $(6, 2)$ of the η-NE region $\mathcal{N}_\eta(7, 5, 2, 4, 6, 0)$. Noisy channel-output feedback in transmitter-receiver 1.

which implies that $\overleftarrow{n}_{jj}$ must satisfy

$$\overleftarrow{n}_{jj} > \max\left(n_{ij}, \overrightarrow{n}_{jj} - n_{ji}\right). \tag{23}$$

Note that condition (23) is identical to the condition needed to observe an enlargement of the capacity region [8]. This observations lead to the following remarks.
Remark 1: The existence of a feedback link in at least one of the two transmitter-receiver pairs is not sufficient to observe an enlargement of the η-NE region. The quality of the feedback link, measured in terms of the number of available bit-pipes between the receiver and the transmitter, must be beyond a threshold that depends on the parameteres $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12} and n_{21}.
Remark 2: For all $i \in \{1, 2\}$, the upper-bound on the rate R_i at an η-NE, i.e, U_i, is independent of the feedback parameter $\overleftarrow{n}_{ii}$ and dependent on the feedback parameter $\overleftarrow{n}_{jj}$. That is, the maximum achievable rate R_i at an η-NE does not depend on whether or not feedback is implemented in transmitter-receiver pair i, but rather transmitter-receiver pair j.

The η-NE region $\mathcal{N}_\eta$ without feedback, i.e., when $\overleftarrow{n}_{11} = 0$ and $\overleftarrow{n}_{22} = 0$, is described by Theorem 1 in [3]. This result is obtained as a corollary of Theorem 1.

Corollary 1 (Theorem 1 **in [3])** The η-Nash region of the LD-IC without feedback, with parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, and n_{21}, is $\mathcal{N}_\eta(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, 0, 0)$.

The η-NE region with perfect feedback i.e., $\overleftarrow{n}_{11} = \max(\overrightarrow{n}_{11}, n_{12})$ and $\overleftarrow{n}_{22} = \max(\overrightarrow{n}_{22}, n_{21})$, is described by Theorem 1 in [9]. This result can also be obtained as a corollary of Theorem 1.

Corollary 2 (Theorem 1 **in [9])** The η-Nash region of the LD-IC with perfect channel-output feedback, with parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, and n_{21}, is $\mathcal{N}_\eta(\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\max(\overrightarrow{n}_{11}, n_{12})$, $\max(\overrightarrow{n}_{22}, n_{21}))$.

The η-NE region with noisy feedback under symmetric conditions i.e., $\overrightarrow{n}_{11} = \overrightarrow{n}_{22} = \overrightarrow{n}$, $n_{12} = n_{21} = m$, and $\overleftarrow{n}_{11} = \overleftarrow{n}_{22} = \overleftarrow{n}$, is described by Theorem 1 in [10]. This result can also be obtained as a corollary of Theorem 1.

Corollary 3 (Symmetric Noisy Channel-Output Feedback) The η-Nash region of the symmetric LD-IC with noisy channel-output feedback ($\overrightarrow{n}_{11} = \overrightarrow{n}_{22} = \overrightarrow{n}$, $n_{12} = n_{21} = m$, and $\overleftarrow{n}_{11} = \overleftarrow{n}_{22} = \overleftarrow{n}$), with parameters $\overrightarrow{n}$, m, and $\overleftarrow{n}$, is $\mathcal{N}_\eta(\overrightarrow{n}, \overrightarrow{n}, m, m, \overleftarrow{n}, \overleftarrow{n})$.

It is interesting to highlight the following set of inclusions.

Corollary 4 *Given an LD-IC with fixed parameters* $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21})$*, the following holds for all* $(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \in \mathbb{N}^2$*:*

$$\mathcal{N}_\eta\Big(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, 0, 0\Big) \subseteq \mathcal{N}_\eta\Big(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}\Big) \subseteq$$
$$\mathcal{N}_\eta\Big(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \max\left(\overrightarrow{n}_{11}, n_{12}\right), \max\left(\overrightarrow{n}_{22}, n_{21}\right)\Big).$$

Considering again the example in which $\overrightarrow{n}_{11} = 7$, $\overrightarrow{n}_{22} = 5$, $n_{12} = 2$, and $n_{21} = 4$, Figure 3e illustrates the achievability scheme for the equilibrium rate pair $(6, 2)$ of the η-NE region $\mathcal{N}_\eta(7, 5, 2, 4, 6, 0)$ and Figure 3f illustrates the achievability scheme for the equilibrium rate pair $(6, 2)$ of the η-NE region $\mathcal{N}_\eta(7, 5, 2, 4, 6, 0)$.

Consider the case in which $\overleftarrow{n}_{11} = 6$ and $\overleftarrow{n}_{22} = 0$ (See Figure 3b). In this case, the η-Nash region $\mathcal{N}_\eta$ is the convex hull of the rate pairs $(5, 1)$, $(5, 2)$, $(6, 2)$, and $(6, 1)$.

The rate pair $(5, 2)$ is achieved at an η-NE thanks to the use of feedback in transmitter-receiver pair 1 when transmitter 1 uses its third bit-pipe of its own codeword $\boldsymbol{X}_{1,n}$ to re-transmit during channel use n, 1 bit that has been previously transmitted by transmitter 2 and have produced interference at receiver 1 during channel use $n-1$. Note that there are 2 bit-pipes at receiver 1 impaired by the interference of transmitter 2, however, only 1 bits can be fed back due to the effect of noise in the feedback channel (See Figure 3e). At channel use n, transmitter 1 re-transmits the interfering bit that is simultaneously received by receiver 1 and receiver 2. At receiver 2, this bit is seen at bit-pipe 6. However, this bit does not represent any interference for receiver 2 since it was received interference free at channel use $n-1$, and thus, it can be cancelled at channel use n. At receiver 1, this bit is seen during channel use n at bit-pipe 3 inside of its top $(\overrightarrow{n}_{11} - n_{12})^+ = 5$ bit-pipes and thus, interference free. Hence, at channel use n, receiver 1 can cancel the interference it produced during channel use $n-1$. In this case, transmitter 1 and transmitter 2 are able to send 5 and 2 bits per channel use, respectively. Note that transmitter 2 also sends randomly generated bits, denoted by $\widetilde{b}_1, \widetilde{b}_2, \ldots$ in Figure 3e. These bits are assumed to be known at both transmitter 2 and receiver 2 and thus, they do not increase the transmission rate of transmitter-receiver 2, however, they produce interference at receiver 1. In this case, the sole objective of transmitting randomly generated bits by transmitter 2 is to prevent the transmitter 1 from sending new information bits and thus, increasing its transmission rate. Then, any attempt of transmitter i to transmit additional information bits would bound its probability of error away from zero. Thus, the rate pair $(5, 2)$ is achieved at an η-NE.

The achievability of the rate pair $(6, 2)$ follows the same explanation of the achievability of the η-NE rate pair $(5, 2)$ with the only difference that for this rate pair it is not necessary that transmitter 2 sends randomly generated bits (See Figure 3f). In this case, transmitter 1 and transmitter 2 are able to send 6 and 2 bits per channel use, respectively. Any attempt of transmitter i to transmit additional information bits would bound its probability of error away from zero. Thus, the rate pair $(6, 2)$ is achieved at an η-NE.

4 Two-User Gaussian Interference Channel with Noisy Feedback

Denote by $\mathcal{C}_{\mathrm{GIC-NOF}}$ the capacity region of the two-user GIC-NOF with fixed parameters $\overrightarrow{\mathrm{SNR}}_1$, $\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, INR_{21}, $\overleftarrow{\mathrm{SNR}}_1$, and $\overleftarrow{\mathrm{SNR}}_2$. The achievable region $\underline{\mathcal{C}}_{\mathrm{GIC-NOF}}$ and the converse region $\overline{\mathcal{C}}_{\mathrm{GIC-NOF}}$ approximate the capacity region $\mathcal{C}_{\mathrm{GIC-NOF}}$ to within 4.4 bits [5]. The achievable region $\underline{\mathcal{C}}_{\mathrm{GIC-NOF}}$ and the converse region $\overline{\mathcal{C}}_{\mathrm{GIC-NOF}}$ are defined by Theorem 1 and Theorem 2 in [5], respectively.

4.1 Nash Achievable Region of the two-user Gaussian Interference Channel with Noisy Channel-Output Feedback

Let the η-NE region (Def. 3) of the GIC-NOF be denoted by $\mathcal{N}_{\mathrm{GIC-NOF}}$. This section introduces a region $\underline{\mathcal{N}}_\eta \subseteq \mathcal{N}_{\mathrm{GIC-NOF}}$ that is achievable using the randomized Han-Kobayashi scheme with noisy channel-output feedback (RHK-NOF). This coding scheme is presented in Appendices A and B. The RHK-NOF is proved to be an η-NE action profile with $\eta > 1$. That is, any unilateral deviation from the RHK-NOF by any of the transmitter-receiver pairs might lead to an individual rate improvement that is upper bounded by one bit per channel use. The description of the achievable η-Nash region $\underline{\mathcal{N}}_\eta$ is presented using the constants $a_{1,i}$; the functions $a_{2,i} : [0,1] \rightarrow \mathbb{R}_+$, $a_{l,i} : [0,1]^2 \rightarrow \mathbb{R}_+$, with $l \in \{3, \ldots, 6\}$; and $a_{7,i} : [0,1]^3 \rightarrow \mathbb{R}_+$, which are defined as follows, for all $i \in \{1, 2\}$, with $j \in \{1, 2\} \setminus \{i\}$:

$$a_{1,i}=\frac{1}{2}\log\left(2+\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\right)-\frac{1}{2}, \tag{24a}$$

$$a_{2,i}(\rho)=\frac{1}{2}\log\Big(b_{1,i}(\rho)+1\Big)-\frac{1}{2}, \tag{24b}$$

$$a_{3,i}(\rho,\mu)=\frac{1}{2}\log\left(\frac{\overleftarrow{\mathrm{SNR}}_i\Big(b_{2,i}(\rho)+2\Big)+b_{1,i}(1)+1}{\overleftarrow{\mathrm{SNR}}_i\Big((1-\mu)\,b_{2,i}(\rho)+2\Big)+b_{1,i}(1)+1}\right), \tag{24c}$$

$$a_{4,i}(\rho,\mu)=\frac{1}{2}\log\Big(\Big(1-\mu\Big)b_{2,i}(\rho)+2\Big)-\frac{1}{2}, \tag{24d}$$

$$a_{5,i}(\rho,\mu)=\frac{1}{2}\log\left(2+\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}+\Big(1-\mu\Big)b_{2,i}(\rho)\right)-\frac{1}{2}, \tag{24e}$$

$$a_{6,i}(\rho,\mu)=\frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\Big(\Big(1-\mu\Big)b_{2,j}(\rho)+1\Big)+2\right)-\frac{1}{2},\text{ and} \tag{24f}$$

$$a_{7,i}(\rho,\mu_1,\mu_2)=\frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\Big(\Big(1-\mu_i\Big)b_{2,j}(\rho)+1\Big)+\Big(1-\mu_j\Big)b_{2,i}(\rho)+2\right)-\frac{1}{2}, \tag{24g}$$

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

$$b_{1,i}(\rho)=\overrightarrow{\mathrm{SNR}}_i+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_i\mathrm{INR}_{ij}}+\mathrm{INR}_{ij}\text{ and} \tag{25a}$$

$$b_{2,i}(\rho)=\Big(1-\rho\Big)\mathrm{INR}_{ij}-1. \tag{25b}$$

Note that the functions in (24) and (25) depend on $\overrightarrow{\mathrm{SNR}}_1$, $\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, INR_{21}, $\overleftarrow{\mathrm{SNR}}_1$, and $\overleftarrow{\mathrm{SNR}}_2$, however as these parameters are fixed in this analysis, this dependence is not emphasized in the definition of these functions. Finally, using this notation, the main result is presented by Theorem 2. The inequalities in (81) are additional conditions to those defining the region $\underline{\mathcal{C}}_{\mathrm{GIC-NOF}}$ [5, Theorem 2]. More specifically, the η-NE region is described by the intersection of the achievable region $\underline{\mathcal{C}}_{\mathrm{GIC-NOF}}$ and the set of rate pairs (R_1,R_2) satisfying (81).

Theorem 2 Let $\eta\geqslant 1$. The η-Nash achievable region $\underline{\mathcal{N}}_\eta$ is given by the closure of the set of all possible non-negative achievable rate pairs $(R_1,R_2)\in\underline{\mathcal{C}}_{\mathrm{GIC-NOF}}$ that satisfy, for all $i\in\{1,2\}$ and $j\in\{1,2\}\setminus\{i\}$, the following conditions:

$$R_i\geqslant a_{2,i}(\rho)-a_{3,i}(\rho,\mu_j)-a_{4,i}(\rho,\mu_j), \tag{26a}$$

$$\begin{aligned} R_i\leqslant \min\Big(& a_{2,i}(\rho)+a_{3,j}(\rho,\mu_i)+a_{5,j}(\rho,\mu_i)-a_{2,j}(\rho),\\ & a_{3,i}(\rho,\mu_j)+a_{7,i}(\rho,\mu_1,\mu_2)+2a_{3,j}(\rho,\mu_i)+a_{5,j}(\rho,\mu_i)-a_{2,j}(\rho),\\ & a_{2,i}(\rho)+a_{3,i}(\rho,\mu_j)+2a_{3,j}(\rho,\mu_i)+a_{5,j}(\rho,\mu_i)+a_{7,j}(\rho,\mu_1,\mu_2)-2a_{2,j}(\rho)\Big), \end{aligned} \tag{26b}$$

$$R_1+R_2\leqslant a_{1,i}+a_{3,i}(\rho,\mu_j)+a_{7,i}(\rho,\mu_1,\mu_2)+a_{2,j}(\rho)+a_{3,j}(\rho,\mu_1)-a_{2,i}(\rho), \tag{26c}$$

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

The proof of Theorem 2 is presented in Appendix D

The following describes some interesting observations from Theorem 2. Figure 4 shows an inner-bound on the capacity region [5, Theorem 2] and the achievable η-NE region in Theorem

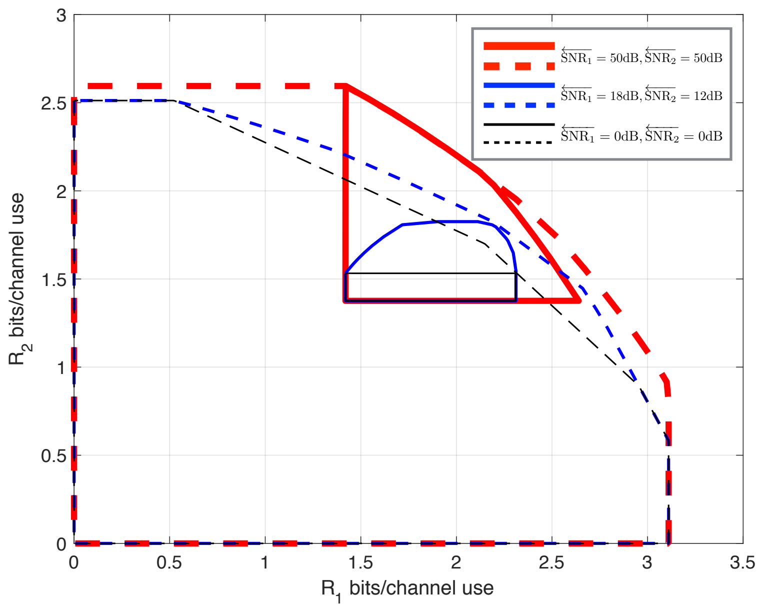

Figure 4: Achievable regions (dashed-lines) and η-NE achievable regions (solid lines) of a GIC with parameters $\overrightarrow{\mathrm{SNR}}_1 = 24$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 18$ dB, $\mathrm{INR}_{12} = 16$ dB, $\mathrm{INR}_{21} = 10$ dB, $\overleftarrow{\mathrm{SNR}}_1 \in \{0, 18, 50\}$ dB and $\overleftarrow{\mathrm{SNR}}_2 \in \{0, 12, 50\}$ dB.

2 for a GIC-NOF channel with parameters $\overrightarrow{\mathrm{SNR}}_1 = 24$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 18$ dB, $\mathrm{INR}_{12} = 16$ dB, $\mathrm{INR}_{21} = 10$ dB, $\overleftarrow{\mathrm{SNR}}_1 \in \{0, 18, 50\}$ dB and $\overleftarrow{\mathrm{SNR}}_2 \in \{0, 12, 50\}$ dB. At low values of $\overleftarrow{\mathrm{SNR}}_1$ and $\overleftarrow{\mathrm{SNR}}_2$, the η-NE region approaches the rectangular region reported in [3] for the case of GIC without feedback. Alternatively, for high values of $\overleftarrow{\mathrm{SNR}}_1$ and $\overleftarrow{\mathrm{SNR}}_2$, the η-NE region approaches the region reported in [9] for the case of GIC with perfect output feedback. These observations are formalized by the following corollaries.

Denote by $\underline{\mathcal{N}}_{\eta\mathrm{PF}}$ the η-NE of GIC with perfect output feedback presented in [9]. The region $\underline{\mathcal{N}}_{\eta\mathrm{PF}}$ can be obtained as a special case of Theorem 2 as shown by the following corollary.

Corollary 5 (η-NE Region with Perfect Output-Feedback) *Let* $\underline{\mathcal{N}}_{\eta\mathrm{PF}}(\overrightarrow{\mathrm{SNR}}_1, \overrightarrow{\mathrm{SNR}}_2, \mathrm{INR}_{12}, \mathrm{INR}_{21})$ *denote the η-NE of the GIC with perfect channel-output feedback with fixed parameters* $\overrightarrow{\mathrm{SNR}}_i$ *and* INR_{ij}, $i \in \{1, 2\}$ *and* $j \in \{1, 2\} \backslash \{i\}$. *Then, the following holds:*

$$\underline{\mathcal{N}}_{\eta\mathrm{PF}}(\overrightarrow{\mathrm{SNR}}_1, \overrightarrow{\mathrm{SNR}}_2, \mathrm{INR}_{12}, \mathrm{INR}_{21}) = \lim_{\substack{\overleftarrow{\mathrm{SNR}}_1 \to \infty \\ \overleftarrow{\mathrm{SNR}}_2 \to \infty}} \underline{\mathcal{N}}_{\eta}(\overrightarrow{\mathrm{SNR}}_1, \overrightarrow{\mathrm{SNR}}_2, \overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2, \mathrm{INR}_{12}, \mathrm{INR}_{21}).$$

Denote by $\underline{\mathcal{N}}_{\eta\mathrm{WF}}$ the η-NE of the GIC without output feedback presented in [3]. The region $\underline{\mathcal{N}}_{\eta\mathrm{WF}}$ can be obtained as a special case of Theorem 2 as shown by the following corollary.

Corollary 6 (η-NE Region without Output-Feedback) *Let* $\underline{\mathcal{N}}_{\eta\mathrm{WF}}(\overrightarrow{\mathrm{SNR}}_1, \overrightarrow{\mathrm{SNR}}_2, \mathrm{INR}_{12}, \mathrm{INR}_{21})$ *denote the η-NE of the GIC without feedback and fixed*

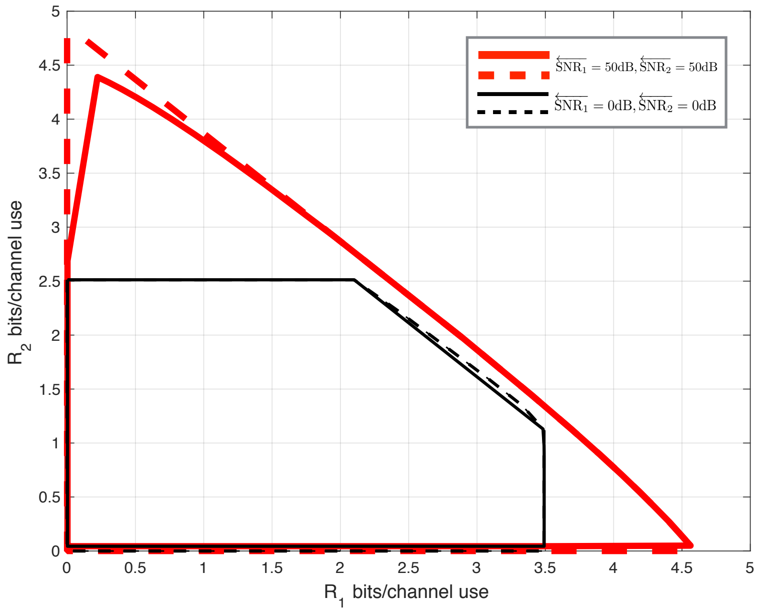

Figure 5: Achievable regions (dashed-lines) and η-NE achievable regions (solid lines) of a GIC with parameters $\overrightarrow{\mathrm{SNR}}_1 = 24$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 18$ dB, $\mathrm{INR}_{12} = 48$ dB, $\mathrm{INR}_{21} = 30$ dB, $\overleftarrow{\mathrm{SNR}}_1 \in \{0, 50\}$ dB and $\overleftarrow{\mathrm{SNR}}_2 \in \{0, 50\}$ dB.

parameters $\overrightarrow{\mathrm{SNR}}_i$ *and* INR_{ij}*, with* $i \in \{1, 2\}$ *and* $j \in \{1, 2\} \backslash \{i\}$*. Then, the following holds:*

$$\underline{\mathcal{N}}_{\eta\mathrm{WF}}(\overrightarrow{\mathrm{SNR}}_1, \overrightarrow{\mathrm{SNR}}_2, \mathrm{INR}_{12}, \mathrm{INR}_{21}) = \lim_{\substack{\overleftarrow{\mathrm{SNR}}_1 \rightarrow 0 \\ \overleftarrow{\mathrm{SNR}}_2 \rightarrow 0 \\ \rho \rightarrow 0.}} \underline{\mathcal{N}}_{\eta}(\overrightarrow{\mathrm{SNR}}_1, \overrightarrow{\mathrm{SNR}}_2, \overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2, \mathrm{INR}_{12}, \mathrm{INR}_{21}).$$

Figure 5 shows an inner-region on the capacity region [5, Theorem 2] and the achievable η-NE region in Theorem 2 for a GIC channel with parameters $\overrightarrow{\mathrm{SNR}}_1 = 24$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 18$ dB, $\mathrm{INR}_{12} = 48$ dB, $\mathrm{INR}_{21} = 30$ dB, $\overleftarrow{\mathrm{SNR}}_1 \in \{0, 50\}$ dB and $\overleftarrow{\mathrm{SNR}}_2 \in \{0, 50\}$ dB. In this case, the η-NE achievable region and the inner-region on the capacity region [5, Theorem 2] are almost identical, which implies that in the cases in which $\overrightarrow{\mathrm{SNR}}_i < \mathrm{INR}_{ij}$, for both $i \in \{1, 2\}$, with $j \in \{1, 2\} \setminus \{i\}$, the η-NE region is almost the same as the region achievable in the centralized case studied in [5].

4.2 Nash Converse Region of the two-user Gaussian Interference Channel with Noisy Channel-Output Feedback

This section will appear in the next version of this research report.

Appendices

A Proofs of Lemma 3

This appendix provides a description of the RHK-NOF and a proof of Lemma 3. This scheme is based on a three-part message splitting, superposition coding, and backward decoding. This coding scheme is general and thus, it holds for the two-user LD-IC-NOF and the two-user GIC-NOF.

Codebook Generation: fix a strictly positive joint probability distribution

$$P_{U\,U_1\,U_2\,V_1\,V_2\,X_{1,P}\,X_{2,P}}(u, u_1, u_2, v_1, v_2, x_{1,P}, x_{2,P}) = P_U(u)P_{U_1|U}(u_1|u)P_{U_2|U}(u_2|u)$$
$$P_{V_1|U\,U_1}(v_1|u, u_1)P_{V_2|U\,U_2}(v_2|u, u_2)P_{X_{1,P}|U\,U_1\,V_1}(x_{1,P}|u, u_1, v_1)P_{X_{2,P}|U\,U_2\,V_2}(x_{2,P}|u, u_2, v_2), \tag{27}$$

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

Let $R_{1,C1}$, $R_{1,R1}$, $R_{1,C2}$, $R_{1,R2}$, $R_{2,C1}$, $R_{2,R1}$, $R_{2,C2}$, $R_{1,R2}$, $R_{1,P}$, and $R_{2,P}$ be non-negative real numbers. Let $R_{1,C} = R_{1,C1} + R_{1,C2}$, $R_{2,C} = R_{2,C1} + R_{2,C2}$, $R_{1,R} = R_{1,R1} + R_{1,R2}$, $R_{2,R} = R_{2,R1} + R_{2,R2}$. Let also $R_1 = R_{1,C} + R_{1,P}$ and $R_2 = R_{2,C} + R_{2,P}$. Note that the rate R_i is not considering the rate $R_{i,R}$, this is due to the fact that it corresponds to a message that is assumed to be known by transmitter i and receiver i. Consider without any loss of generality that $N = N_1 = N_2$.

Generate $2^{N(R_{1,C1}+R_{1,R1}+R_{2,C1}+R_{2,R1})}$ i.i.d. N-length codewords $\boldsymbol{u}(s, r) = \big(u_1(s, r), u_2(s, r), \ldots, u_N(s, r)\big)$ according to

$$P_{\boldsymbol{U}}\big(\boldsymbol{u}(s, r)\big) = \prod_{i=1}^{N} P_U(u_i(s, r)),$$

with $s \in \{1, 2, \ldots, 2^{N(R_{1,C1}+R_{1,R1})}\}$ and $r \in \{1, 2, \ldots, 2^{N(R_{2,C1}+R_{2,R1})}\}$.

For encoder 1, generate for each codeword $\boldsymbol{u}(s, r)$, $2^{N(R_{1,C1}+R_{1,R1})}$ i.i.d. N-length codewords $\boldsymbol{u}_1(s, r, k) = \big(u_{1,1}(s, r, k),\, u_{1,2}(s, r, k),\, \ldots,\, u_{1,N}(s, r, k)\big)$ according to

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

with $k \in \{1, 2, \ldots, 2^{N(R_{1,C1}+R_{1,R1})}\}$. For each pair of codewords $\big(\boldsymbol{u}(s, r), \boldsymbol{u}_1(s, r, k)\big)$, generate $2^{N(R_{1,C2}+R_{1,R2})}$ i.i.d. N-length codewords $\boldsymbol{v}_1(s, r, k, l, d) = \big(v_{1,1}(s, r, k, l),\, v_{1,2}(s, r, k, l),\, \ldots,\, v_{1,N}(s, r, k, l)\big)$ according to

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

with $l \in \{1, 2, \ldots, 2^{N(R_{1,C2}+R_{1,R2})}\}$. For each tuple of codewords $\big(\boldsymbol{u}(s, r),\ \boldsymbol{u}_1(s, r, k),\ \boldsymbol{v}_1(s, r, k, l)\big)$, generate $2^{NR_{1,P}}$ i.i.d. N-length codewords $\boldsymbol{x}_{1,P}(s, r, k, l, q) = \big(x_{1,P,1}(s, r, k, l, q),\, x_{1,P,2}(s, r, k, l, q),\, \ldots,\, x_{1,P,N}(s, r, k, l, q)\big)$ according to

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

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

For encoder 2, generate for each codeword $\boldsymbol{u}(s,r)$, $2^{N(R_{2,C1}+R_{2,R1})}$ i.i.d. N-length codewords $\boldsymbol{u}_2(s,r,j) = \big(u_{2,1}(s,r,j),\, u_{2,2}(s,r,j),\, \ldots,\, u_{2,N}(s,r,j)\big)$ according to

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

with $j \in \{1, 2, \ldots, 2^{N(R_{2,C1}+R_{2,R1})}\}$. For each pair of codewords $\big(\boldsymbol{u}(s,r), \boldsymbol{u}_2(s,r,j)\big)$, generate $2^{N(R_{2,C2}+R_{2,R2})}$ i.i.d. length-N codewords $\boldsymbol{v}_2(s,r,j,m) = \big(v_{2,1}(s,r,j,m),\, v_{2,2}(s,r,j,m),\, \ldots,\, v_{2,N}(s,r,j,m)\big)$ according to

$$P_{\boldsymbol{V}_2|\boldsymbol{U}\,\boldsymbol{U}_2}\big(\boldsymbol{v}_2(s,r,j,m)|\boldsymbol{u}(s,r), \boldsymbol{u}_2(s,r,j)\big) = \prod_{i=1}^{N} P_{V_2|U\,U_2}(v_{2,i}(s,r,j,m)|u_i(s,r), u_{2,i}(s,r,j)),$$

with $m \in \{1, 2, \ldots, 2^{N(R_{2,C2}+R_{2,R2})}\}$. For each tuple of codewords $\big(\boldsymbol{u}(s,r), \boldsymbol{u}_2(s,r,j), \boldsymbol{v}_2(s,r,j,m)\big)$, generate $2^{NR_{2,P}}$ i.i.d. N-length codewords $\boldsymbol{x}_{2,P}(s,r,j,m,b) = \big(x_{2,P,1}(s,r,j,m,b), x_{2,P,2}(s,r,j,m,b),\, \ldots,\, x_{2,P,N}(s,r,j,m)\big)$ according to

$$P_{\boldsymbol{X}_{2,P}|\boldsymbol{U}\,\boldsymbol{U}_2\,\boldsymbol{V}_2}\big(\boldsymbol{x}_{2,P}(s,r,j,m,b)|\boldsymbol{u}(s,r), \boldsymbol{u}_2(s,r,j), \boldsymbol{v}_2(s,r,j,m)\big) = \\ \prod_{i=1}^{N} P_{X_{2,P}|U\,U_2\,V_2}\big(x_{2,P,i}(s,r,j,m,b)|u_i(s,r), u_{2,i}(s,r,j), v_{2,i}(s,r,j,m,b)\big),$$

with $b \in \{1, 2, \ldots, 2^{NR_{2,P}}\}$. The resulting code structure is shown in Figure 6.

Encoding: denote by $\left(W_i^{(t)}, \Omega_i^{(t)}\right) \in \{1, 2, \ldots, 2^{N(R_{i,C}+R_{i,P})}\} \times \{1, 2, \ldots, 2^{NR_{i,R}}\}$ the index that comprises the message index and the random message index of transmitter i 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 the message index $W_{i,P}^{(t)} \in \{1,\ 2,\ \ldots, 2^{NR_{i,P}}\}$. That is, $W_i^{(t)} = \left(W_{i,C}^{(t)}, W_{i,P}^{(t)}\right)$. The message index $W_{i,P}^{(t)}$ must be reliably decoded at receiver i. Let $W_{i,C}^{(t)}$ be composed by the message indices $W_{i,C1}^{(t)} \in \{1, 2, \ldots, 2^{NR_{i,C1}}\}$ and $W_{i,C2}^{(t)} \in \{1, 2, \ldots, 2^{NR_{i,C2}}\}$. That is, $W_{i,C}^{(t)} = \left(W_{i,C1}^{(t)}, W_{i,C2}^{(t)}\right)$. Let $\Omega_i^{(t)}$ be composed by the message indices $\Omega_{i,R1}^{(t)} \in \{1, 2, \ldots, 2^{NR_{i,R1}}\}$ and $\Omega_{i,R2}^{(t)} \in \{1, 2, \ldots, 2^{NR_{i,R2}}\}$. That is, $\Omega_i^{(t)} = \left(\Omega_{i,R1}^{(t)}, \Omega_{i,R2}^{(t)}\right)$. The index $\left(W_{i,C1}^{(t)}, \Omega_{i,R1}^{(t)}\right)$ must be reliably decoded by transmitter j (via feedback) but no necessarily by receiver i. The index $\left(W_{i,C2}^{(t)}, \Omega_{i,R2}^{(t)}\right)$ must be reliably decoded by receiver j but no necessarily by receiver i.

Consider Markov encoding over T blocks. At encoding step t, with $t \in \{1, 2, \ldots, T\}$, transmitter 1 sends the codeword $\boldsymbol{x}_1^{(t)} = \Theta_1\Bigg(\boldsymbol{u}\Big(\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}\right),\ \left(W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right)\Big)$, $\boldsymbol{u}_1\Big(\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}\right),\ \left(W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right),\ \left(W_{1,C1}^{(t)}, \Omega_{1,R1}^{(t)}\right)\Big)$, $\boldsymbol{v}_1\Big(\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}\right)$, $\left(W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right),\ \left(W_{1,C1}^{(t)}, \Omega_{1,R1}^{(t)}\right),\ \left(W_{1,C2}^{(t)}, \Omega_{1,R2}^{(t)}\right)\Big)$, $\boldsymbol{x}_{1,P}\Big(\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}\right),\ \left(W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right)$, $\left(W_{1,C1}^{(t)}, \Omega_{1,R1}^{(t)}\right),\ \left(W_{1,C2}^{(t)}, \Omega_{1,R2}^{(t)}\right),\ W_{1,P}^{(t)}\Big)\Bigg)$, where $\Theta_1 : (\mathcal{X}_1 \cup \mathcal{X}_2)^N \times \mathcal{X}_1^N \times \mathcal{X}_1^N \times \mathcal{X}_1^N \rightarrow \mathcal{X}_1^N$ is a function that transforms the codewords $\boldsymbol{u}\Big(\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}\right),\ \left(W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right)\Big)$,

$\boldsymbol{u}_1\Big(\big(W_{1,C1}^{(t-1)},\Omega_{1,R1}^{(t-1)}\big), \big(W_{2,C1}^{(t-1)},\Omega_{2,R1}^{(t-1)}\big), \big(W_{1,C1}^{(t)},\Omega_{1,R1}^{(t)}\big)\Big)$, $\boldsymbol{v}_1\Big(\big(W_{1,C1}^{(t-1)},\Omega_{1,R1}^{(t-1)}\big), \big(W_{2,C1}^{(t-1)},\Omega_{2,R1}^{(t-1)}\big), \big(W_{1,C1}^{(t)},\Omega_{1,R1}^{(t)}\big), \big(W_{1,C2}^{(t)},\Omega_{1,R2}^{(t)}\big)\Big)$, and $\boldsymbol{x}_{1,P}\Big(\big(W_{1,C1}^{(t-1)},\Omega_{1,R1}^{(t-1)}\big), \big(W_{2,C1}^{(t-1)},\Omega_{2,R1}^{(t-1)}\big),\big(W_{1,C1}^{(t)},\Omega_{1,R1}^{(t)}\big), \big(W_{1,C2}^{(t)},\Omega_{1,R2}^{(t)}\big), W_{1,P}^{(t)}\Big)$ into the N-dimensional vector $\boldsymbol{x}_1^{(t)}$ of channel inputs. The indices $\big(W_{1,C1}^{(0)},\Omega_{1,R1}^{(0)}\big) = \big(W_{1,C1}^{(T)},\Omega_{1,R1}^{(T)}\big) = s^*$ and $\big(W_{2,C1}^{(0)}=\Omega_{2,R1}^{(0)}\big) = \big(W_{2,C1}^{(T)},\Omega_{2,R1}^{(T)}\big) = r^*$, and the pair $(s^*,r^*) \in \{1,2,\ldots,2^{N(R_{1,C1}+R_{1,R1})}\} \times \{1,2,\ldots,2^{N(R_{2,C1}+R_{2,R1})}\}$ are pre-defined and known by both receivers and transmitters. It is worth noting that the index $\big(W_{2,C1}^{(t-1)},\Omega_{2,R1}^{(t-1)}\big)$ are 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 indices $\Big(\big(\widehat{W}_{1,C1}^{(T-t)},\widehat{\Omega}_{1,R1}^{(T-t)}\big), \big(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\big), \big(\widehat{W}_{1,C2}^{(T-(t-1))},\widehat{\Omega}_{1,R2}^{(T-(t-1))}\big), \widehat{W}_{1,P}^{(T-(t-1))}, \big(\widehat{W}_{2,C2}^{(T-(t-1))},\widehat{\Omega}_{2,R2}^{(T-(t-1))}\big)\Big) \in \{1,\ 2,\ldots,\ 2^{NR_{1,C1}}\} \times \{1,\ 2,\ldots,\ 2^{NR_{1,R1}}\} \times \{1,\ 2,\ldots,2^{NR_{2,C1}} \times \{1,\ 2,\ldots,2^{NR_{2,R1}}\} \times \{1,\ 2,\ldots,2^{NR_{1,C2}}\} \times \{1,\ 2,\ldots,2^{NR_{1,R2}}\} \times \{1,\ 2,\ldots,2^{NR_{1,P}}\} \times \{1,\ 2,\ldots,2^{NR_{2,C2}}\} \times \{1,\ 2,\ldots,2^{NR_{2,R2}}\}$ from the channel output $\overrightarrow{\boldsymbol{y}}_1^{(T-(t-1))}$. The tuple $\Big(\big(\widehat{W}_{1,C1}^{(T-t)},\widehat{\Omega}_{1,R1}^{(T-t)}\big), \big(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\big), \big(\widehat{W}_{1,C2}^{(T-(t-1))},\widehat{\Omega}_{1,R2}^{(T-(t-1))}\big), \widehat{W}_{1,P}^{(T-(t-1))}, \big(\widehat{W}_{2,C2}^{(T-(t-1))},\widehat{\Omega}_{2,R2}^{(T-(t-1))}\big)\Big)$ is the unique tuple that satisfies:

$$\Bigg(\boldsymbol{u}\Big(\big(\widehat{W}_{1,C1}^{(T-t)},\widehat{\Omega}_{1,R1}^{(T-t)}\big),\big(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\big)\Big),\boldsymbol{u}_1\Big(\big(\widehat{W}_{1,C1}^{(T-t)},\widehat{\Omega}_{1,R1}^{(T-t)}\big),\big(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\big), \big(W_{1,C1}^{(T-(t-1))},\Omega_{1,R1}^{(T-(t-1))}\big)\Big),\boldsymbol{v}_1\Big(\big(\widehat{W}_{1,C1}^{(T-t)},\widehat{\Omega}_{1,R1}^{(T-t)}\big),\big(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\big), \big(W_{1,C1}^{(T-(t-1))},\Omega_{1,R1}^{(T-(t-1))}\big),\big(\widehat{W}_{1,C2}^{(T-(t-1))},\widehat{\Omega}_{1,R2}^{(T-(t-1))}\big)\Big),\boldsymbol{x}_{1,P}\Big(\big(\widehat{W}_{1,C1}^{(T-t)},\widehat{\Omega}_{1,R1}^{(T-t)}\big), \big(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\big),\big(W_{1,C1}^{(T-(t-1))},\Omega_{1,R1}^{(T-(t-1))}\big),\big(\widehat{W}_{1,C2}^{(T-(t-1))},\widehat{\Omega}_{1,R2}^{(T-(t-1))}\big),\widehat{W}_{1,P}^{(T-(t-1))}\Big), \boldsymbol{u}_2\Big(\big(\widehat{W}_{1,C1}^{(T-t)},\widehat{\Omega}_{1,R1}^{(T-t)}\big),\big(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\big),\big(W_{2,C1}^{(T-(t-1))},\Omega_{2,R1}^{(T-(t-1))}\big)\Big), \boldsymbol{v}_2\Big(\big(\widehat{W}_{1,C1}^{(T-t)},\widehat{\Omega}_{1,R1}^{(T-t)}\big),\big(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\big),\big(W_{2,C1}^{(T-(t-1))},\Omega_{2,R1}^{(T-(t-1))}\big), \big(\widehat{W}_{2,C2}^{(T-(t-1))},\widehat{\Omega}_{2,R2}^{(T-(t-1))}\big)\Big),\overrightarrow{\boldsymbol{y}}_1^{(T-(t-1))}\Bigg) \in \mathcal{T}^{(N,e)}_{\left[U\ U_1\ V_1\ X_{1,P}\ U_2\ V_2\ \overrightarrow{Y}_1\right]}, \tag{28}$$

where $\big(W_{1,C1}^{(T-(t-1))},\Omega_{1,R1}^{(T-(t-1))}\big)$ and $\big(W_{2,C1}^{(T-(t-1))},\Omega_{2,R1}^{(T-(t-1))}\big)$ are assumed to be perfectly decoded in the previous decoding step $t-1$. The set $\mathcal{T}^{(N,e)}_{\left[U\ U_1\ V_1\ X_{1,P}\ U_2\ V_2\ \overrightarrow{Y}_1\right]}$ 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 $e > 0$. Finally, receiver 2 follows a similar decoding scheme.

Probability of Error Analysis: an error might occur during encoding step t if the index $\big(W_{2,C1}^{(t-1)},\Omega_{2,R1}^{(t-1)}\big)$ is not correctly decoded at transmitter 1. From the asymptotic equipartion property (AEP) [12], it follows that the index $\big(W_{2,C1}^{(t-1)},\Omega_{2,R1}^{(t-1)}\big)$ can be reliably decoded at trans-

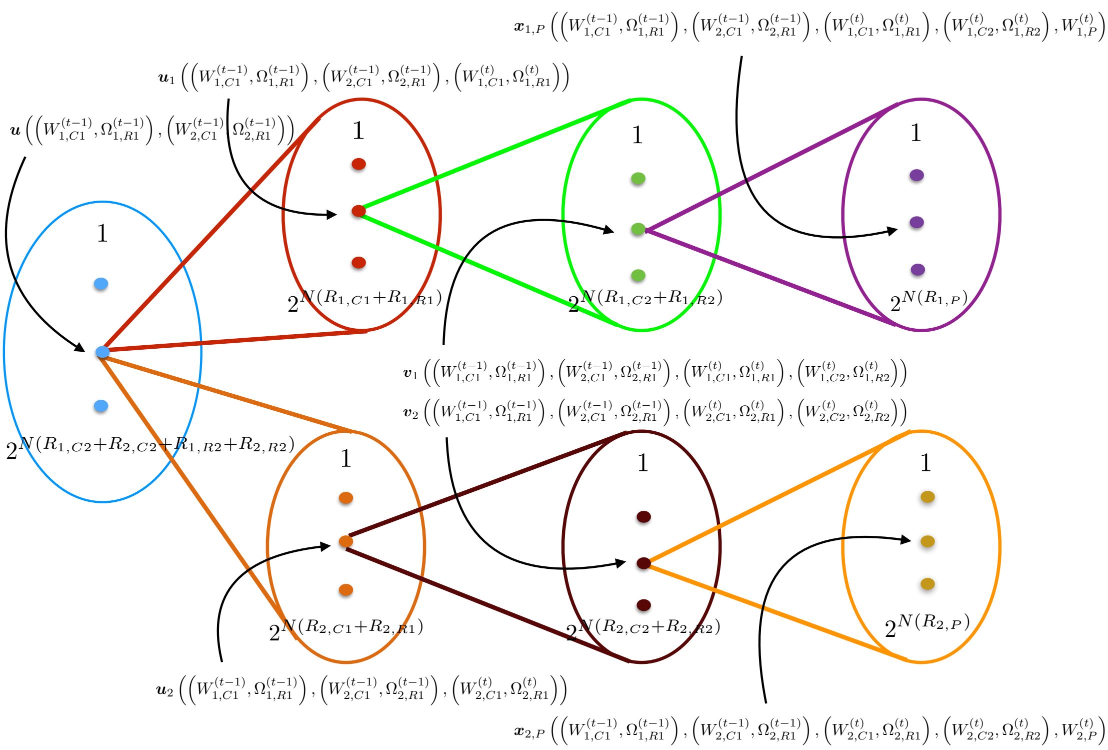

Figure 6: Structure of the superposition code. The codewords corresponding to the message indices $W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)}, W_{i,P}^{(t)}$ with $i \in \{1,2\}$ as well as the block index t are both highlighted. The (approximate) number of codewords for each code layer is also highlighted.

mitter 1 during encoding step t, under the condition:

$$
\begin{aligned}
R_{2,C1}+R_{2,R1} &\leqslant I\left(\overleftarrow{Y}_1;U_2|U,U_1,V_1,X_1\right)\\
&= I\left(\overleftarrow{Y}_1;U_2|U,X_1\right). \qquad (29)
\end{aligned}
$$

An error might occur during the (backward) decoding step t if the indices $\left(W_{1,C1}^{(T-t)},\Omega_{1,R1}^{(T-t)}\right)$, $\left(W_{2,C1}^{(T-t)},\Omega_{2,R1}^{(T-t)}\right)$, $\left(W_{1,C2}^{(T-(t-1))},\Omega_{1,R2}^{(T-(t-1))}\right)$, $W_{1,P}^{(T-(t-1))}$, and $\left(W_{2,C2}^{(T-(t-1))},\Omega_{2,R2}^{(T-(t-1))}\right)$ are not decoded correctly given that the indices $\left(W_{1,C1}^{(T-(t-1))},\Omega_{1,R1}^{(T-(t-1))}\right)$ and $\left(W_{2,C1}^{(T-(t-1))},\Omega_{2,R1}^{(T-(t-1))}\right)$ were correctly decoded in the previous decoding step $t-1$. These errors might arise for two reasons: (i) there does not exist a tuple $\left(\left(\widehat{W}_{1,C1}^{(T-t)},\widehat{\Omega}_{1,R1}^{(T-t)}\right),\right.$ $\left(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\right),\left(\widehat{W}_{1,C2}^{(T-(t-1))},\widehat{\Omega}_{1,R2}^{(T-(t-1))}\right),\widehat{W}_{1,P}^{(T-(t-1))},$ $\left.\left(\widehat{W}_{2,C2}^{(T-(t-1))},\widehat{\Omega}_{2,R2}^{(T-(t-1))}\right)\right)$ that satisfies (28), or (ii) there exist several tuples $\left(\left(\widehat{W}_{1,C1}^{(T-t)},\ \widehat{\Omega}_{1,R1}^{(T-t)}\right),\ \left(\widehat{W}_{2,C1}^{(T-t)},\ \widehat{\Omega}_{2,R1}^{(T-t)}\right),\right.$ $\left(\widehat{W}_{1,C2}^{(T-(t-1))},\ \widehat{\Omega}_{1,R2}^{(T-(t-1))}\right),\ \widehat{W}_{1,P}^{(T-(t-1))},\ \left.\left(\widehat{W}_{2,C2}^{(T-(t-1))},\ \widehat{\Omega}_{2,R2}^{(T-(t-1))}\right)\right)$ that simultaneously satisfy (28). From the asymptotic equipartion property (AEP) [12], 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 $\Big(\boldsymbol{u}\,(s,r),\ \ \boldsymbol{u}_1\Big(s,r,\big(W_{1,C1}^{(T-(t-1))},\Omega_{1,R1}^{(T-(t-1))}\big)\Big),\ \ \boldsymbol{v}_1\Big(s,r,\Big(W_{1,C1}^{(T-(t-1))},\Omega_{1,R1}^{(T-(t-1))}\Big),l\Big),$ $\boldsymbol{x}_{1,P}\Big(s,r,\Big(W_{1,C1}^{(T-(t-1))},\Omega_{1,R1}^{(T-(t-1))}\Big),l,q\Big),\ \ \boldsymbol{u}_2\Big(s,r,\Big(W_{2,C1}^{(T-(t-1))},\Omega_{2,R1}^{(T-(t-1))}\Big)\Big),$ and $\boldsymbol{v}_2\Big(s,r,\Big(W_{2,C1}^{(T-(t-1))},\Omega_{2,R1}^{(T-(t-1))}\Big),m\Big)\Big)$ 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 bounded as follows:

$$
\begin{aligned}
P_e=&\Pr\left(\bigcup_{(s,r,l,q,m)\neq(1,1,1,1,1)} E_{(s,r,l,q,m)}\right)\\
\leqslant&\sum_{(s,r,l,q,m)\in\mathcal{T}}\Pr\left(E_{(s,r,l,q,m)}\right),
\end{aligned}
\tag{30}
$$

with $\mathcal{T}=\Big\{\{1,2,\dots 2^{N(R_{1,C1}+R_{1,R1})}\}\times\{1,2,\dots 2^{N(R_{2,C1}+R_{2,R1})}\}\times\{1,2,\dots 2^{N(R_{1,C2}+R_{1,R2})}\}\times\{1,2,\dots 2^{NR_{1,P}}\}\times\{1,2,\dots 2^{N(R_{2,C2}+R_{2,R2})}\}\Big\}\setminus\{(1,1,1,1,1)\}$. Therefore,

$$
\begin{aligned}
P_e\leqslant&\sum_{s=1,r=1,l=1,q=1,m\neq1}\Pr\left(E_{(s,r,l,q,m)}\right)+\sum_{s=1,r=1,l=1,q\neq1,m=1}\Pr\left(E_{(s,r,l,q,m)}\right)\\
&+\sum_{s=1,r=1,l=1,q\neq1,m\neq1}\Pr\left(E_{(s,r,l,q,m)}\right)+\sum_{s=1,r=1,l\neq1,q=1,m=1}\Pr\left(E_{(s,r,l,q,m)}\right)\\
&+\sum_{s=1,r=1,l\neq1,q=1,m\neq1}\Pr\left(E_{(s,r,l,q,m)}\right)+\sum_{s=1,r=1,l\neq1,q\neq1,m=1}\Pr\left(E_{(s,r,l,q,m)}\right)\\
&+\sum_{s=1,r=1,l\neq1,q\neq1,m\neq1}\Pr\left(E_{(s,r,l,q,m)}\right)+\sum_{s=1,r\neq1,l=1,q=1,m=1}\Pr\left(E_{(s,r,l,q,m)}\right)\\
&+\sum_{s=1,r\neq1,l=1,q=1,m\neq1}\Pr\left(E_{(s,r,l,q,m)}\right)+\sum_{s=1,r\neq1,l=1,q\neq1,m=1}\Pr\left(E_{(s,r,l,q,m)}\right)\\
&+\sum_{s=1,r\neq1,l=1,q\neq1,m\neq1}\Pr\left(E_{(s,r,l,q,m)}\right)+\sum_{s=1,r\neq1,l\neq1,q=1,m=1}\Pr\left(E_{(s,r,l,q,m)}\right)\\
&+\sum_{s=1,r\neq1,l\neq1,q=1,m\neq1}\Pr\left(E_{(s,r,l,q,m)}\right)+\sum_{s=1,r\neq1,l\neq1,q\neq1,m=1}\Pr\left(E_{(s,r,l,q,m)}\right)\\
&+\sum_{s=1,r\neq1,l\neq1,q\neq1,m\neq1}\Pr\left(E_{(s,r,l,q,m)}\right)+\sum_{s\neq1,r=1,l=1,q=1,m=1}\Pr\left(E_{(s,r,l,q,m)}\right)\\
&+\sum_{s\neq1,r=1,l=1,q=1,m\neq1}\Pr\left(E_{(s,r,l,q,m)}\right)+\sum_{s\neq1,r=1,l=1,q\neq1,m=1}\Pr\left(E_{(s,r,l,q,m)}\right)\\
&+\sum_{s\neq1,r=1,l=1,q\neq1,m\neq1}\Pr\left(E_{(s,r,l,q,m)}\right)+\sum_{s\neq1,r=1,l\neq1,q=1,m=1}\Pr\left(E_{(s,r,l,q,m)}\right)\\
&+\sum_{s\neq1,r=1,l\neq1,q=1,m\neq1}\Pr\left(E_{(s,r,l,q,m)}\right)+\sum_{s\neq1,r=1,l\neq1,q\neq1,m=1}\Pr\left(E_{(s,r,l,q,m)}\right)\\
&+\sum_{s\neq1,r=1,l\neq1,q\neq1,m\neq1}\Pr\left(E_{(s,r,l,q,m)}\right)+\sum_{s\neq1,r\neq1,l=1,q=1,m=1}\Pr\left(E_{(s,r,l,q,m)}\right)
\end{aligned}
$$

$$
\begin{aligned}
&+ \sum_{s\neq 1, r\neq 1, l=1, q=1, m\neq 1} \Pr\left(E_{(s,r,l,q,m)}\right) + \sum_{s\neq 1, r\neq 1, l=1, q\neq 1, m=1} \Pr\left(E_{(s,r,l,q,m)}\right)\\
&+ \sum_{s\neq 1, r\neq 1, l=1, q\neq 1, m\neq 1} \Pr\left(E_{(s,r,l,q,m)}\right) + \sum_{s\neq 1, r\neq 1, l\neq 1, q=1, m=1} \Pr\left(E_{(s,r,l,q,m)}\right)\\
&+ \sum_{s\neq 1, r\neq 1, l\neq 1, q=1, m\neq 1} \Pr\left(E_{(s,r,l,q,m)}\right) + \sum_{s\neq 1, r\neq 1, l\neq 1, q\neq 1, m=1} \Pr\left(E_{(s,r,l,q,m)}\right)\\
&+ \sum_{s\neq 1, r\neq 1, l\neq 1, q\neq 1, m\neq 1} \Pr\left(E_{(s,r,l,q,m)}\right). \qquad (31)
\end{aligned}
$$

From the asymptotic equipartion property (AEP) [12], it follows that

$$
\begin{aligned}
P_e \leqslant\; & 2^{N(R_{2,C2}+R_{2,R2}-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_{2,R2}+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}+R_{2,R2}-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}+R_{2,R2}-I(\overrightarrow{Y}_1;V_1,V_2,X_1|U,U_1,U_2)+2\epsilon)}\\
&+2^{N(R_{2,C1}+R_{2,R1}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)}\\
&+2^{N(R_{2,C}+R_{2,R}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)}\\
&+2^{N(R_{2,C1}+R_{2,R1}+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_{2,R}+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_{2,R1}+R_{1,C2}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)}\\
&+2^{N(R_{2,C}+R_{2,R}+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_{2,R1}+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_{2,R}+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}+R_{2,R2}-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}+R_{2,R2}-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}+R_{2,R2}-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}+R_{2,R2}-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}+R_{2,R1}-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_{2,R}-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}+R_{2,R1}+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_{2,R}+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}+R_{2,R1}-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}+R_{2,R}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)}
\end{aligned}
$$

$$+2^{N(R_1+R_{2,C1}+R_{2,R1}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)}$$
$$+2^{N(R_1+R_{2,C}+R_{2,R}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)}. \tag{32}$$

Note that the random message indices $\Omega_{i,R1}^{(t)}$ and $\Omega_{i,R2}^{(t)}$ are assumed to be known at both transmitter i and receiver i.

The same analysis of the probability of error holds for transmitter-receiver pair 2. Hence, in general, from (29) and (32), reliable decoding holds under the following conditions for transmitter $i \in \{1,2\}$, with $j \in \{1,2\} \setminus \{i\}$:

$$\begin{aligned}
R_{j,C1}+R_{j,R1} &\leqslant I\left(\overleftarrow{Y}_i;U_j|U,U_i,V_i,X_i\right)\\
&= I\left(\overleftarrow{Y}_i;U_j|U,X_i\right)\\
&\triangleq \theta_{1,i}, && \text{(33a)}\\
R_i+R_{j,C}+R_{j,R} &\leqslant I(\overrightarrow{Y}_i;U,U_i,U_j,V_i,V_j,X_i)\\
&= I(\overrightarrow{Y}_i;U,U_j,V_j,X_i)\\
&\triangleq \theta_{2,i}, && \text{(33b)}\\
R_{j,C2}+R_{j,R2} &\leqslant I(\overrightarrow{Y}_i;V_j|U,U_i,U_j,V_i,X_i)\\
&= I(\overrightarrow{Y}_i;V_j|U,U_j,X_i)\\
&\triangleq \theta_{3,i}, && \text{(33c)}\\
R_{i,P} &\leqslant I(\overrightarrow{Y}_i;X_i|U,U_i,U_j,V_i,V_j)\\
&\triangleq \theta_{4,i}, && \text{(33d)}\\
R_{i,P}+R_{j,C2}+R_{j,R2} &\leqslant I(\overrightarrow{Y}_i;V_j,X_i|U,U_i,U_j,V_i)\\
&\triangleq \theta_{5,i}, && \text{(33e)}\\
R_{i,C2}+R_{i,P} &\leqslant I(\overrightarrow{Y}_i;V_i,X_i|U,U_i,U_j,V_j)\\
&= I(\overrightarrow{Y}_i;X_i|U,U_i,U_j,V_j),\\
&\triangleq \theta_{6,i}, \text{ and,} && \text{(33f)}\\
R_{i,C2}+R_{i,P}+R_{j,C2}+R_{j,R2} &\leqslant I(\overrightarrow{Y}_i;V_i,V_j,X_i|U,U_i,U_j)\\
&= I(\overrightarrow{Y}_i;V_j,X_i|U,U_i,U_j)\\
&\triangleq \theta_{7,i}. && \text{(33g)}
\end{aligned}$$

From the probability of error analysis, it follows that the rate-pairs achievable with the proposed randomized coding scheme with noisy channel-output feedback are those simultaneously satisfying conditions (33) with $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$. Indeed, when $R_{1,R} = R_{2,R} = 0$, the coding scheme described above reduces to the coding scheme presented in [5]. For the two-user linear deterministic interference channel model, $\theta_{1,i}, \theta_{2,i}, \ldots, \theta_{7,i}$ in (33) are defined in (58), which completes the proof of Lemma 3.

B Proof of Lemma 7

This appendix provides a proof of Lemma 7. For the two-user Gaussian interference channel model, consider that transmitter i uses the following Gaussian input distribution [5]:

$$X_i = U + U_i + V_i + X_{i,P}, \tag{34}$$

where U, U_1, U_2, V_1, V_2, $X_{1,P}$, and $X_{2,P}$ in (27) are mutually independent and distributed as follows:

$$U \sim \mathcal{N}(0, \rho), \tag{35a}$$
$$U_i \sim \mathcal{N}(0, \mu_i \lambda_{i,C}), \tag{35b}$$
$$V_i \sim \mathcal{N}(0, (1-\mu_i)\lambda_{i,C}), \tag{35c}$$
$$X_{i,P} \sim \mathcal{N}(0, \lambda_{i,P}), \tag{35d}$$

with

$$\rho + \lambda_{i,C} + \lambda_{i,P} = 1 \text{ and} \tag{36a}$$

$$\lambda_{i,P} = \min\left(\frac{1}{\mathrm{INR}_{ji}}, 1\right), \tag{36b}$$

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

Then, the following holds in (33) for the two-user Gaussian interference channel with noisy channel-output feedback:

$$\theta_{1,i} \triangleq a_{3,i}(\rho, \mu_j), \tag{37a}$$
$$\theta_{2,i} \triangleq a_{2,i}(\rho), \tag{37b}$$
$$\theta_{3,i} \triangleq a_{4,i}(\rho, \mu_j), \tag{37c}$$
$$\theta_{4,i} \triangleq a_{1,i}, \tag{37d}$$
$$\theta_{5,i} \triangleq a_{5,i}(\rho, \mu_j), \tag{37e}$$
$$\theta_{6,i} \triangleq a_{6,i}(\rho, \mu_i), \text{ and} \tag{37f}$$
$$\theta_{7,i} \triangleq a_{7,i}(\rho, \mu_1, \mu_2), \tag{37g}$$

where the functions $a_{1,i}$, $a_{2,i}(\rho)$, $a_{3,i}(\rho, \mu_j)$, $a_{4,i}(\rho, \mu_j)$, $a_{5,i}(\rho, \mu_j)$, $a_{6,i}(\rho, \mu_i)$, and $a_{7,i}(\rho, \mu_1, \mu_2)$ are defined in (24). This completes the proof of Lemma 7.

C Proof of Theorem 1

To prove Theorem 1, the first step is to show that a rate pair (R_1, R_2), with $R_i < L_i$ or $R_i > U_i$ for at least one $i \in \{1, 2\}$, is not achievable at an η-equilibrium for an arbitrarily small $\eta > 0$. That is,

$$\mathcal{N}_\eta \subseteq \mathcal{C} \cap \mathcal{B}_\eta. \tag{38}$$

The second step is to show that any point in $\mathcal{C} \cap \mathcal{B}_\eta$ can be achievable at an η-equilibrium for all $\eta > 0$. That is,

$$\mathcal{N}_\eta \supseteq \mathcal{C} \cap \mathcal{B}_\eta, \tag{39}$$

which proves Theorem 1.

Proof of (38) The proof of (38) is completed by the following lemmas.

Lemma 1 A rate pair $(R_1, R_2) \in \mathcal{C}$, with either $R_1 < L_1$ or $R_2 < L_2$ is not achievable at an η-equilibrium, with $\eta > 0$ arbitrarily small.

Proof: Let (s_1^*, s_2^*) be an η-NE transmit-receive configuration pair such that users 1 and 2 achieves the rates $R_1(s_1^*, s_2^*)$ and $R_2(s_1^*, s_2^*)$, respectively. Assume, without loss of generality, that $R_1(s_1^*, s_2^*) < L_1$. Let $s_1' \in \mathcal{A}_1$ be a transmit receive configuration in which transmitter 1 uses its $(\overrightarrow{n}_{11} - n_{12})^+$ most significant bit-pipes, which are interference free, to transmit new bits at each channel use n. Hence, it achieves a rate $R_1(s_1', s_2^*) \geqslant (\overrightarrow{n}_{11} - n_{12})^+$ and thus, a utility improvement of at least η bits per channel use is always possible, i.e., $R_1(s_1', s_2^*) - R_1(s_1^*, s_2^*) > \eta$, independently of the current transmit-receive configuration s_2^* of user 2. This implies that the transmit-receive configuration pair (s_1^*, s_2^*) is not an η-equilibrium, which contradicts the initial assumption. This proves that if (s_1^*, s_2^*) is an η-NE, then $R_1(s_1^*, s_2^*) \geqslant L_1$ an $R_2(s_1^*, s_2^*) \geqslant L_2$. This completes the proof. ■

Lemma 2 A rate pair $(R_1, R_2) \in \mathcal{C}$, with either $R_1 > U_1$ or $R_2 > U_2$ is not achievable at an η-equilibrium, with $\eta > 0$ arbitrarily small.

Proof: Let $(s_1^*, s_2^*) \in \mathcal{A}_1 \times \mathcal{A}_2$ be an η-NE transmit-receive configuration pair that achieves the rate pair $(R_1(s_1^*, s_2^*), R_2(s_1^*, s_2^*)) \in \mathcal{N}_\eta$. Hence, the following holds for transmitter-receiver i:

$$\begin{aligned} N\, R_i(s_1^*, s_2^*) &= H(W_i) \\ &\overset{(a)}{=} H(W_i | \Omega_i) \\ &\overset{(b)}{\leqslant} I\left(W_i; \overrightarrow{\boldsymbol{Y}}_i | \Omega_i\right) + N\delta_i(N) \end{aligned} \tag{40}$$

where, (a) follows from the independence between the indices W_i and Ω_i; and (b) follows from Fano's inequality, as the rate $R_i(s_1^*, s_2^*)$ is achievable from the assumptions of the lemma. In particular, for transmitter-receiver pair 1 in (40), the following holds:

$$N\, R_1(s_1^*, s_2^*) \overset{(c)}{\leqslant} N \max(\overrightarrow{n}_{11}, n_{12}) - \sum_{n=1}^{N} H\left(\overrightarrow{\boldsymbol{Y}}_{1,n} | \Omega_1, W_1, \overrightarrow{\boldsymbol{Y}}_{1,(1:n-1)}\right) + N\delta_1(N), \tag{41}$$

where, (c) follows from $H\left(\overrightarrow{\boldsymbol{Y}}_{1,n} | \Omega_1, \overrightarrow{\boldsymbol{Y}}_{1,(1:n-1)}\right) \leqslant H\left(\overrightarrow{\boldsymbol{Y}}_{1,n}\right) \leqslant \max(\overrightarrow{n}_{11}, n_{12})$, for all $n \in \{1, 2, \ldots, N\}$. Note that $\boldsymbol{X}_{1,n} = f_{1,n}^{(N)}\left(W_1, \Omega_1, \overleftarrow{\boldsymbol{Y}}_{1,(1:n-1)}\right)$ from the definition of the coding function in (19). Moreover, for all $n \in \{1, 2, \ldots, N\}$, the channel input $\boldsymbol{X}_{i,n}$ can be written as

$$\boldsymbol{X}_{i,n} = \left(\boldsymbol{X}_{i,C,n}, \boldsymbol{X}_{i,D,n}, \boldsymbol{X}_{i,P,n}, \boldsymbol{X}_{i,Q,n}\right), \tag{42}$$

where for all $i \in \{1, 2\}$, the vector $\boldsymbol{X}_{i,C,n}$ represents the bits of $\boldsymbol{X}_{i,n}$ that are observed by both receivers, i.e.,

$$\dim \boldsymbol{X}_{i,C,n} = \min(\overrightarrow{n}_{ii}, n_{ji}); \tag{43}$$

the vector $\boldsymbol{X}_{i,P,n}$ represents the bits of $\boldsymbol{X}_{i,n}$ that are exclusively observed by receiver i, i.e.,

$$\dim \boldsymbol{X}_{i,P,n} = (\overrightarrow{n}_{ii} - n_{ji})^+; \tag{44}$$

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

$$\dim \boldsymbol{X}_{i,D,n} = (n_{ji} - \overrightarrow{n}_{ii})^+; \tag{45}$$

finally, $\boldsymbol{X}_{i,Q,n} = (0, \ldots, 0)^{\mathsf{T}}$ is included for dimensional matching of the model in (17), i.e.,

$$\dim \boldsymbol{X}_{i,Q,n} = q - \max(\overrightarrow{n}_{ii}, n_{ji}). \tag{46}$$

Using this notation, the following holds from (41):

$$\begin{aligned} R_1(s_1^*, s_2^*) \leqslant & \max\left(\overrightarrow{n}_{11}, n_{12}\right) - \frac{1}{N}\sum_{n=1}^{N} H\left(\boldsymbol{X}_{2,C,n}, \boldsymbol{X}_{2,D,n} | \Omega_1, W_1, \overrightarrow{\boldsymbol{Y}}_{1,(1:n-1)}\right) + \delta_1(N), \\ = & \max\left(\overrightarrow{n}_{11}, n_{12}\right) - H\left(\widetilde{\boldsymbol{X}}_{2,C,n}, \widetilde{\boldsymbol{X}}_{2,D,n}\right) + \delta_1(N), \text{for any } n \in \{1, 2, \ldots, N\} \end{aligned} \quad (47)$$

where $\widetilde{\boldsymbol{X}}_{2,C} = \left(\widetilde{\boldsymbol{X}}_{2,C,1}, \widetilde{\boldsymbol{X}}_{2,C,2}, \ldots, \widetilde{\boldsymbol{X}}_{2,C,N}\right)$ and $\widetilde{\boldsymbol{X}}_{2,D} = \left(\widetilde{\boldsymbol{X}}_{2,D,1}, \widetilde{\boldsymbol{X}}_{2,D,2}, \ldots, \widetilde{\boldsymbol{X}}_{2,D,N}\right)$; and for all $n \in \{1, 2, \ldots, N\}$, $\widetilde{\boldsymbol{X}}_{2,C,n}$ and $\widetilde{\boldsymbol{X}}_{2,D,n}$ are respectively the bits in $\boldsymbol{X}_{2,C,n}$ and $\boldsymbol{X}_{2,D,n}$ that are independent of W_1, Ω_1, and $\overrightarrow{\boldsymbol{Y}}_{1,(1:n-1)}$. That is, the bits other than those depending on bits previously transmitted by transmitter 1. The last inequality in (47) follows from the signal construction in (16).

The following step is to obtain a lower bound for $H\left(\widetilde{\boldsymbol{X}}_{2,C,n}, \widetilde{\boldsymbol{X}}_{2,D,n}\right)$ at an η-NE. For this purpose, two cases are considered:

Case 1($\overrightarrow{n}_{22} > n_{12}$): Under the condition $\overrightarrow{n}_{22} > n_{12}$, it follows that $\dim \boldsymbol{X}_{2,P,n} > 0$, whereas $\dim \boldsymbol{X}_{2,D,n} = 0$. Hence, the following inequality holds for transmitter-receiver pair 2:

$$\begin{aligned} I\left(W_2; \overrightarrow{\boldsymbol{Y}}_2 | \Omega_2\right) \leqslant & I\left(\boldsymbol{X}_2; \overrightarrow{\boldsymbol{Y}}_2 | \Omega_2\right) \\ = & I\left(\boldsymbol{X}_{2,C1}, \boldsymbol{X}_{2,C2}, \boldsymbol{X}_{2P}; \overrightarrow{\boldsymbol{Y}}_2 | \Omega_2\right) \\ = & H\left(\boldsymbol{X}_{2,C1} | \Omega_2, \boldsymbol{X}_{2,C2}, \boldsymbol{X}_{2,P}\right) + I\left(\boldsymbol{X}_{2,C2}, \boldsymbol{X}_{2,P}; \overrightarrow{\boldsymbol{Y}}_2 | \Omega_2\right), \end{aligned} \quad (48)$$

where, for all $i \in \{1, 2\}$, $\boldsymbol{X}_{i,C,n} = \left(\boldsymbol{X}_{i,C1,n}^{\mathsf{T}}, \boldsymbol{X}_{i,C2,n}^{\mathsf{T}}\right)^{\mathsf{T}}$ and $\boldsymbol{X}_{i,C1,n}$ satisfies:

$$\dim \boldsymbol{X}_{i,C1,n} = \left(\min\left((\overrightarrow{n}_{ii} - n_{ij})^+, n_{ji}\right) - \left(\min\left((\overrightarrow{n}_{ii} - n_{ji})^+, n_{ij}\right) - \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^+\right)^+\right)^+. \quad (49)$$

The dimension of $\boldsymbol{X}_{i,C1,n}$ is chosen as the non-negative difference of two values: (a) All the bits in $\boldsymbol{X}_{i,C,n}$ that are observed at both receivers and the observation at receiver i is interference-free, i.e., $\min\left((\overrightarrow{n}_{ii} - n_{ij})^+, n_{ji}\right)$; and (b) the number of bits in $\boldsymbol{X}_{i,n}$ that are only observed at receiver i, interfered by transmitter j, and can be sent via feedback from receiver i to transmitter i, i.e., $\left(\min\left((\overrightarrow{n}_{ii} - n_{ji})^+, n_{ij}\right) - \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^+\right)^+$. The vector $\boldsymbol{X}_{i,C2,n}$ contains the bits in $\boldsymbol{X}_{i,C,n}$ that are not in $\boldsymbol{X}_{i,C1,n}$. That is,

$$\dim \boldsymbol{X}_{i,C2,n} = \min\left(\overrightarrow{n}_{ii}, n_{ji}\right) - \dim \boldsymbol{X}_{i,C1,n}. \quad (50)$$

Using this notation, it holds from (40) and (48) that there always exists a positive monotonically decreasing function $\delta : \mathbb{N} \to \mathbb{R}_+$, such that

$$R_2(s_1^*, s_2^*) = \frac{1}{N} H\left(\boldsymbol{X}_{2,C1} | \Omega_2, \boldsymbol{X}_{2,C2}, \boldsymbol{X}_{2,P}\right) + \frac{1}{N} I\left(\boldsymbol{X}_{2,C2}, \boldsymbol{X}_{2,P}; \overrightarrow{\boldsymbol{Y}}_2 | \Omega_2\right) + \delta(N). \quad (51)$$

Assume now that there exists another transmit-receive configuration for receiver-transmitter pair 2 and denote it by s_2'. Assume also that using s_2', for all $n \in \{1, 2, \ldots, N\}$, transmitter-receiver pair 2 continues to generate the symbols $\boldsymbol{X}_{2,C2,n}$ and $\boldsymbol{X}_{2,P,n}$ as with the equilibrium transmit-receive configuration s_2^*. Alternatively, for all $n \in \{1,2,\ldots,N\}$, the bits $\boldsymbol{X}_{2,C1,n}$ are generated at maximum entropy and independently of any other previously symbol transmitted by any transmitter. More specifically, the bits $\boldsymbol{X}_{2,C1,n}$ are used to send new information bits at each channel use n, i.e.,

$$R_2(s_1^*, s_2') = \dim \boldsymbol{X}_{2,C1,n} + \frac{1}{N} I\left(\boldsymbol{X}_{2,C2}, \boldsymbol{X}_{2,P}; \overrightarrow{\boldsymbol{Y}}_2 | \Omega_2\right) + \delta'(N), \tag{52}$$

with $\delta' : \mathbb{N} \rightarrow \mathbb{R}_+$ a positive monotonically decreasing function. From Definition 2, it follows that $R_2(s_1^*, s_2') - R_2(s_1^*, s_2^*) < \eta$. Hence, from (51) and (52), it follows that

$$H\big(\boldsymbol{X}_{2,C1} | \Omega_2, \boldsymbol{X}_{2,C2}, \boldsymbol{X}_{2,P}\big) \geqslant \dim \boldsymbol{X}_{2,C1,n} - N\eta - N\delta(N) + N\delta'(N). \tag{53}$$

It is worth noting here that $H\big(\boldsymbol{X}_{2,C1} | \Omega_2, \boldsymbol{X}_{2,C2}, \boldsymbol{X}_{2,P}\big) = NH\big(\boldsymbol{X}_{2,C1,n} | \Omega_2, \boldsymbol{X}_{2,C2,n}, \boldsymbol{X}_{2,P,n}\big)$ for any $n \in \{1,2,\ldots,N\}$, then,

$$\begin{aligned} H\big(\boldsymbol{X}_{2,C1,n} | \Omega_2, \boldsymbol{X}_{2,C2,n}, \boldsymbol{X}_{2,P,n}\big) \geqslant & \dim \boldsymbol{X}_{2,C1,n} + \frac{1}{N} I\left(\boldsymbol{X}_{2,C2}, \boldsymbol{X}_{2,P}; \overrightarrow{\boldsymbol{Y}}'_2 | \Omega_2\right) \\ & - \frac{1}{N} I\left(\boldsymbol{X}_{2,C2}, \boldsymbol{X}_{2,P}; \overrightarrow{\boldsymbol{Y}}_2 | \Omega_2\right) - \eta - \delta(N) + \delta'(N). \end{aligned} \tag{54}$$

Note also that

$$\begin{aligned} H\left(\widetilde{\boldsymbol{X}}_{2,C,n}\right) \geqslant & H\left(\boldsymbol{X}_{2,C1,n} | \Omega_2, \boldsymbol{X}_{2,C2,n}, \boldsymbol{X}_{2,P,n}\right) \\ = & \dim \boldsymbol{X}_{2,C1,n}. \end{aligned} \tag{55}$$

Replacing (55) in (47), it follows that an η-NE,

$$\begin{aligned} R_1(s_1^*, s_2^*) \leqslant & \max\left(\overrightarrow{n}_{11}, n_{12}\right) - \Bigg(\min\left(\left(\overrightarrow{n}_{22} - n_{21}\right)^+, n_{12}\right) \\ & - \left(\min\left(\left(\overrightarrow{n}_{22} - n_{12}\right)^+, n_{21}\right) - \left(\max\left(\overrightarrow{n}_{22}, n_{21}\right) - \overleftarrow{n}_{22}\right)^+ \right)^+ \Bigg)^+ + \eta + \delta(N) - \delta'(N), \end{aligned} \tag{56}$$

which proves that $U_1^{(a)} = \max\left(\overrightarrow{n}_{11}, n_{12}\right) - \Bigg(\min\left(\left(\overrightarrow{n}_{22} - n_{21}\right)^+, n_{12}\right) - \left(\min\left(\left(\overrightarrow{n}_{22} - n_{12}\right)^+, n_{21}\right) - \left(\max\left(\overrightarrow{n}_{22}, n_{21}\right) - \overleftarrow{n}_{22}\right)^+ \right)^+ \Bigg)^+ + \eta$. The same analysis holds for the case of $U_2^{(a)}$ under the condition $\overrightarrow{n}_{11} > n_{21}$.

Case 2 ($\overrightarrow{n}_{22} \leqslant n_{12}$): Under the condition $\overrightarrow{n}_{22} \leqslant n_{12}$, it follows that $\dim \boldsymbol{X}_{2,P,n} = 0$, whereas $\dim \boldsymbol{X}_{2,D,n} \geqslant 0$. Moreover, $R_2(s_1^*, s_2^*) = \eta$, with η arbitrarily small. Hence, $H\left(\boldsymbol{X}_{2,C,n}, \boldsymbol{X}_{2,D,n} | \Omega_1, W_1, \overrightarrow{\boldsymbol{Y}}_{1,(1:n-1)}\right) = 0$, since all bits transmitted by transmitter 2 can be re-transmissions of bits previously transmitted by transmitter 1, which yields, from (47), $R_1(s_1^*, s_2^*) \leqslant \max\left(\overrightarrow{n}_{11}, n_{12}\right)$. This proves that $U_1^{(b)} = \max\left(\overrightarrow{n}_{11}, n_{12}\right) + \eta$. The same analysis holds to obtain $U_2^{(b)}$ under the condition $\overrightarrow{n}_{11} \leqslant n_{21}$. It can be seen that $U_1^{(b)} = U_1^{(a)}$ and $U_2^{(b)} = U_2^{(a)}$ under conditions $\overrightarrow{n}_{22} \leqslant n_{12}$ and $\overrightarrow{n}_{11} \leqslant n_{21}$, respectively. Then, $U_i = \max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \Bigg(\min\left(\left(\overrightarrow{n}_{jj} - n_{ji}\right)^+, n_{ij}\right) - \left(\min\left(\left(\overrightarrow{n}_{jj} - n_{ij}\right)^+, n_{ji}\right) - \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+ \right)^+ \Bigg)^+ + \eta$, and this completes the proof of Lemma 2. ■

Proof of (39) To continue with the second part of the proof of Theorem 1, consider a modification of the coding scheme with noisy feedback presented in [5]. The novelty with respect to [5] consists of allowing users to introduce common randomness as suggested in [3] and [9].

Consider without any loss of generality that $N = N_1 = N_2$. Let the message index and the random message index sent by transmitter i during the t-th block, with $t \in \{1, 2, \ldots, T\}$, be denoted by $W_i^{(t)} \in \{1, 2, \ldots, 2^{NR_i}\}$ and $\Omega_i^{(t)} \in \{1, 2, \ldots, 2^{NR_{i,R}}\}$. Following a rate-splitting argument, assume that $\left(W_i^{(t)}, \Omega_i^{(t)}\right)$ is represented by the indices $\left(W_{i,C1}^{(t)}, \Omega_{i,R1}^{(t)}, W_{i,C2}^{(t)}, \Omega_{i,R2}^{(t)}, W_{i,P}^{(t)}\right) \in \{1, 2, \ldots, 2^{NR_{i,C1}}\} \times \{1, 2, \ldots, 2^{NR_{i,R1}}\} \times \{1, 2, \ldots, 2^{NR_{i,C2}}\} \times \{1, 2, \ldots, 2^{NR_{i,R2}}\} \times \{1, 2, \ldots, 2^{NR_{i,P}}\}$, where $R_i = R_{i,C1} + R_{i,C2} + R_{i,P}$ and $R_{i,R} = R_{i,R1} + R_{i,R2}$. The rate $R_{i,R}$ is the number of transmitted bits that are known by both transmitter i and receiver i per channel use and thus, it does not have an impact on the effective information rate R_i.

The codeword generation follows a four-level superposition coding scheme. The indices $W_{i,C1}^{(t-1)}$ and $\Omega_{i,R1}^{(t-1)}$ are 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 $W_{1,C1}^{(t-1)}$, $\Omega_{1,R1}^{(t-1)}$, $W_{2,C1}^{(t-1)}$ and $\Omega_{2,R1}^{(t-1)}$. In the case of the first block $t = 1$, the indices $W_{1,C1}^{(0)}$, $\Omega_{1,R1}^{(0)}$, $W_{2,C1}^{(0)}$ and $\Omega_{1,R2}^{(0)}$ are assumed to be known by all transmitters and receivers. Using these indices both transmitters are able to identify the same codeword in the first code-layer. This first code-layer, which is common for both transmitter-receiver pairs, is a sub-codebook of $2^{N(R_{1,C1}+R_{2,C1}+R_{1,R1}+R_{2,R1})}$ codewords. Denote by $\boldsymbol{u}\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}, W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right)$ the corresponding codeword in the first code-layer. The second codeword is chosen by transmitter i using $\left(W_{i,C1}^{(t)}, \Omega_{i,R1}^{(t)}\right)$ from the second code-layer, which is a sub-codebook of $2^{N(R_{i,C1}+R_{i,R1})}$ codewords corresponding to the codeword $\boldsymbol{u}\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}, W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right)$. Denote by $\boldsymbol{u}_i\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}, W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}, W_{i,C1}^{(t)}, \Omega_{i,R1}^{(t)}\right)$ the corresponding codeword in the second code-layer. The third codeword is chosen by transmitter i using $\left(W_{i,C2}^{(t)}, \Omega_{i,R2}^{(t)}\right)$ from the third code-layer, which is a sub-codebook of $2^{N(R_{i,C2}+R_{i,R2})}$ codewords corresponding to the codeword $\boldsymbol{u}_i\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}, W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}, W_{i,C1}^{(t)}, \Omega_{i,R1}^{(t)}\right)$. Denote by $\boldsymbol{v}_i\Big(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}, W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}, W_{i,C1}^{(t)}, \Omega_{i,R1}^{(t)}, W_{i,C2}^{(t)}, \Omega_{i,R2}^{(t)}\Big)$ the corresponding codeword in the third code-layer. The fourth codeword is chosen by transmitter i using $W_{i,P}^{(t)}$ from the fourth code-layer, which is a sub-codebook of $2^{N\,R_{i,P}}$ codewords corresponding to the codeword $\boldsymbol{v}_i\Big(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}, W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}, W_{i,C1}^{(t)}, \Omega_{i,R1}^{(t)}, W_{i,C2}^{(t)}, \Omega_{i,R2}^{(t)}\Big)$. Denote by $\boldsymbol{x}_{i,P}\Big(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}, W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}, W_{i,C1}^{(t)}, \Omega_{i,R1}^{(t)}, W_{i,C2}^{(t)}, \Omega_{i,R2}^{(t)}, W_{i,P}^{(t)}\Big)$ the corresponding codeword in the fourth code-layer. Finally, the codeword $\boldsymbol{x}_i\Big(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}, W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}, W_{i,C1}^{(t)}, \Omega_{i,R1}^{(t)}, W_{i,C2}^{(t)}, \Omega_{i,R2}^{(t)}, W_{i,P}^{(t)}\Big)$ to be sent during block $t \in \{1, 2, \ldots, T\}$ is a simple concatenation of the previous codewords, i.e., $\boldsymbol{x}_i = \left(\boldsymbol{u}_i^{\mathsf{T}}, \boldsymbol{v}_i^{\mathsf{T}}, \boldsymbol{x}_{i,P}^{\mathsf{T}}\right)^{\mathsf{T}} \in \{0,1\}^{q \times N}$, where the message indices have been dropped for the ease of notation.

The decoder follows a backward decoding scheme. In the following, this coding scheme is referred to as a randomized Han-Kobayashi coding scheme with noisy feedback (RHK-NOF). This coding/decoding scheme is thoroughly described in Appendix A.

The proof of (39) uses the following results:

- Lemma 3 proves that the RHK-NOF achieves all the rate pairs $(R_1, R_2) \in \mathcal{C}$;
- Lemma 4 provides the maximum rate improvement that a transmitter-receiver pair can

obtain when it deviates from the RHK-NOF coding scheme;

- Lemma 5 proves that when the rates of the random components $R_{1,R1}$, $R_{1,R2}$, $R_{2,R1}$, and $R_{2,R2}$ are properly chosen, the RHK-NOF is an η-NE, with $\eta > 0$ arbitrarily small; and
- Lemma 6 shows that for all rate pairs in $\mathcal{C} \cap \mathcal{B}_\eta$ there always exists a RHK-NOF that is an η-NE and achieves such a rate pair.

This verifies that $\mathcal{N}_\eta \supseteq \mathcal{C} \cap \mathcal{B}_\eta$ and completes the proof of (39).

Lemma 3 The achievable region of the randomized Han-Kobayashi coding scheme for the LD-IC-NOF is the set of non-negative rates $\Big(R_{1,C1}, R_{1,R1}, R_{1,C2}, R_{1,R2}, R_{1,P}, R_{2,C1}, R_{2,R1}, R_{2,C2}, R_{2,R2}, R_{2,P}\Big)$ that satisfy the following conditions:

$$R_{j,C1} + R_{j,R1} \leqslant \theta_{1,i}, \tag{57a}$$
$$R_i + R_{j,C} + R_{j,R} \leqslant \theta_{2,i}, \tag{57b}$$
$$R_{j,C2} + R_{j,R2} \leqslant \theta_{3,i}, \tag{57c}$$
$$R_{i,P} \leqslant \theta_{4,i}, \tag{57d}$$
$$R_{i,P} + R_{j,C2} + R_{j,R2} \leqslant \theta_{5,i}, \tag{57e}$$
$$R_{i,C2} + R_{i,P} \leqslant \theta_{6,i}, \text{ and} \tag{57f}$$
$$R_{i,C2} + R_{i,P} + R_{j,C2} + R_{j,R2} \leqslant \theta_{7,i}, \tag{57g}$$

where,

$$\theta_{1,i} = \Big(n_{ij} - \big(\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii}\big)^+\Big)^+, \tag{58a}$$
$$\theta_{2,i} = \max(\overrightarrow{n}_{ii}, n_{ij}), \tag{58b}$$
$$\theta_{3,i} = \min\Big(n_{ij}, \big(\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii}\big)^+\Big), \tag{58c}$$
$$\theta_{4,i} = (\overrightarrow{n}_{ii} - n_{ji})^+, \tag{58d}$$
$$\theta_{5,i} = \max\Big((\overrightarrow{n}_{ii} - n_{ji})^+, \min\Big(n_{ij}, \big(\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii}\big)^+\Big)\Big),$$
$$\theta_{6,i} = \min\Big(n_{ji}, \big(\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj}\big)^+\Big) - \min\Big((n_{ji} - \overrightarrow{n}_{ii})^+, \big(\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj}\big)^+ \Big) + (\overrightarrow{n}_{ii} - n_{ji})^+, \text{ and} \tag{58e}$$
$$\theta_{7,i} = \max\Big(\min\Big(n_{ij}, \big(\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii}\big)^+ \Big), \min\Big(n_{ji}, \big(\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj}\big)^+ \Big) - \min\Big((n_{ji} - \overrightarrow{n}_{ii})^+, \big(\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj}\big)^+ \Big) + (\overrightarrow{n}_{ii} - n_{ji})^+ \Big).$$

Proof: The proof of Lemma 3 is presented in Appendix A. ■

The set of inequalities in (57) can be written in terms of the transmission rates $R_1 = R_{1,C1} + R_{1,C2} + R_{1,P}$, $R_2 = R_{2,C1} + R_{2,C2} + R_{2,P}$, $R_{1,R} = R_{1,R1} + R_{1,R2}$ and $R_{2,R} = R_{2,R1} + R_{2,R2}$. When $R_{1,R} = R_{2,R} = 0$, the region characterized by (57) in terms of R_1 and R_2, corresponds to the region $\mathcal{C}$ (Theorem 1 in [5]). Therefore, the relevance of Lemma 3 relies on the implication that any rate pair $(R_1, R_2) \in \mathcal{C}$ is achievable by the RHK-NOF, under the assumption that the random common rates $R_{1,R1}$, $R_{1,R2}$, $R_{2,R1}$, and $R_{2,R2}$ are chosen accordingly to the conditions in (57).

The following lemma shows than when both transmitter-receiver links use the RHK-NOF and one of them unilaterally changes its coding scheme, it obtains a rate improvement that can be upper bounded.

Lemma 4 Let $\eta > 0$ be an arbitrarily small number and let the rate tuple $\boldsymbol{R} = (R_{1,C} - \frac{\eta}{6}, R_{1,R} - \frac{\eta}{6}, R_{1,P} - \frac{\eta}{6}, R_{2,C} - \frac{\eta}{6}, R_{2,R} - \frac{\eta}{6}, R_{2,P} - \frac{\eta}{6})$ be achievable with the RHK-NOF such that $R_1 = R_{1,P} + R_{1,C} - \frac{1}{3}\eta$ and $R_2 = R_{2,P} + R_{2,C} - \frac{1}{3}\eta$. Then, any unilateral deviation of transmitter-receiver pair i by using any other coding scheme leads to a transmission rate R_i' that satisfies:

$$R_i' \leqslant \max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \left(R_{j,C} + R_{j,R}\right) + \frac{2}{3}\eta. \tag{59}$$

Proof: Without loss of generality, let $i = 1$ be the deviating user in the following analysis. After the deviation, the new coding scheme used by transmitter 1 can be of any type. Indeed, with such a new coding scheme, the deviating transmitter might or might not use feedback to generate its codewords. It can also use or not use random symbols and it might possibly have a different block-length $N_1' \neq N_1$. Let $\overrightarrow{\boldsymbol{Y}}_1' = \left(\overrightarrow{\boldsymbol{Y}}_{1,1}'^{\mathsf{T}}, \overrightarrow{\boldsymbol{Y}}_{1,2}'^{\mathsf{T}}, \ldots, \overrightarrow{\boldsymbol{Y}}_{1,N}'^{\mathsf{T}}\right)^{\mathsf{T}}$ be the super vector of channel outputs at receiver 1 during $N = \max(N_1', N_2)$ consecutive channel uses in the model in (16). Hence, an upper bound for R_1' is obtained from the following inequalities:

$$\begin{aligned}
NR_1' &= H\left(W_1\right) \\
&= H\left(W_1|\Omega_1\right) \\
&= I\left(W_1; \overrightarrow{\boldsymbol{Y}}_1'|\Omega_1\right) + H\left(W_1|\overrightarrow{\boldsymbol{Y}}_1', \Omega_1\right) \\
&\overset{(a)}{\leqslant} I\left(W_1; \overrightarrow{\boldsymbol{Y}}_1'|\Omega_1\right) + N\delta_1(N) \\
&= H\left(\overrightarrow{\boldsymbol{Y}}_1'|\Omega_1\right) - H\left(\overrightarrow{\boldsymbol{Y}}_1'|W_1, \Omega_1\right) + N\delta_1(N) \\
&= \sum_{n=1}^{N} H\left(\overrightarrow{\boldsymbol{Y}}_{1,n}'|\overrightarrow{\boldsymbol{Y}}_{1,(1:n-1)}', \Omega_1\right) - H\left(\overrightarrow{\boldsymbol{Y}}_1'|W_1, \Omega_1\right) + N\delta_1(N) \\
&\overset{(b)}{\leqslant} N\max\left(\overrightarrow{n}_{11}, n_{12}\right) - H\left(\overrightarrow{\boldsymbol{Y}}_1'|W_1, \Omega_1\right) + N\delta_1(N),
\end{aligned} \tag{60}$$

where, (a) follows from Fano's inequality, as the rate R_1' is achievable from the assumptions of the lemma; and (b) follows from $H(\overrightarrow{\boldsymbol{Y}}_{1,n}'|\Omega_1) \leqslant \dim \overrightarrow{\boldsymbol{Y}}_{1,n}' = \max\left(\overrightarrow{n}_{11}, n_{12}\right)$, for all $n \in \{1, 2, \ldots, N\}$. To refine this upper bound, the term $H\left(\overrightarrow{\boldsymbol{Y}}_1'|W_1, \Omega_1\right)$ in (60) can be lower bounded as follows:

$$\begin{aligned}
N\left(R_{2,C} + R_{2,R}\right) &= H(W_{2,C}, \Omega_2) \\
&\overset{(c)}{=} H(W_{2,C}, \Omega_2|W_1, \Omega_1) \\
&= I(W_{2,C}, \Omega_2; \overrightarrow{\boldsymbol{Y}}_1'|W_1, \Omega_1) + H(W_{2,C}, \Omega_2|W_1, \Omega_1, \overrightarrow{\boldsymbol{Y}}_1') \\
&\overset{(d)}{\leqslant} I(W_{2,C}, \Omega_2; \overrightarrow{\boldsymbol{Y}}_1'|W_1, \Omega_1) + N\delta_2(N) \\
&= H(\overrightarrow{\boldsymbol{Y}}_1'|W_1, \Omega_1) - H(\overrightarrow{\boldsymbol{Y}}_1'|W_1, \Omega_1, W_{2,C}, \Omega_2) + N\delta_2(N) \\
&\leqslant H(\overrightarrow{\boldsymbol{Y}}_1'|W_1, \Omega_1) + N\delta_2(N),
\end{aligned} \tag{61}$$

where (c) follows from the mutual independence between $W_{2,C}$, Ω_2, W_1 and Ω_1; and (d) follows from Fano's inequality as $(W_{2,C}, \Omega_2)$ can be decoded from $\overrightarrow{\boldsymbol{Y}}_1'$. Hence, it follows from (61) that

$$H(\overrightarrow{\boldsymbol{Y}}_1'|W_1, \Omega_1) \geqslant N\left(R_{2,C} + R_{2,R}\right) - N\delta_2(N). \tag{62}$$

Finally, plugging (62) into (60) yields the following upper bound:

$$R_1' \leqslant \max\left(\overrightarrow{n}_{11}, n_{12}\right) - \left(R_{2,C} + R_{2,R}\right) + \delta(N), \tag{63}$$

where, there always exist a block-length $N = \max\left(N_1', N_2\right)$ such that $\delta(N) = \delta_1(N) + \delta_2(N)$ can be made arbitrarily small and thus, $\delta(N) < \frac{2}{3}\eta$. The same can be proved for the other transmitter-receiver pair. This completes the proof. ■

Lemma 4 reveals the relevance of the random symbols Ω_1 and Ω_2 used by the RHK-NOF. Even though the random symbols used by transmitter j do not increase the effective transmission rate of the transmitter-receiver pair j, they strongly limit the rate improvement transmitter-receiver pair i can obtain by deviating from the RHK-NOF coding scheme. This observation can be used to show that the RHK-NOF can be an η-NE, when both $R_{1,R}$ and $R_{2,R}$ are properly chosen. For instance, for any achievable rate pair $(R_1, R_2) \in \mathcal{C} \cap \mathcal{B}_\eta$, there exists a RHK-NOF that achieves the rate tuple $\boldsymbol{R} = (R_{1,C} - \frac{\eta}{6}, R_{1,R} - \frac{\eta}{6}, R_{1,P} - \frac{\eta}{6}, R_{2,C} - \frac{\eta}{6}, R_{2,R} - \frac{\eta}{6}, R_{2,P} - \frac{\eta}{6})$, with $R_i = R_{i,P} + R_{i,C} - \frac{1}{3}\eta$ and η arbitrarily small. Denote by $R'_{i,\max} = \max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - (R_{j,C} + R_{j,R}) + \frac{2}{3}\eta$ the maximum rate transmitter-receiver pair i can obtain by unilaterally deviating from its RHK-NOF. Then, when the rates $R_{1,R}$ and $R_{2,R}$ are chosen such that $R'_{i,\max} - R_i \leqslant \eta$, any improvement obtained by either transmitter deviating from its RHK-NOF is bounded by η. The following lemma formalizes this observation.

Lemma 5 Let $\eta > 0$ be an arbitrarily small number and let the rate tuple $\boldsymbol{R} = (R_{1,C} - \frac{\eta}{6}, R_{1,R} - \frac{\eta}{6}, R_{1,P} - \frac{\eta}{6}, R_{2,C} - \frac{\eta}{6}, R_{2,R} - \frac{\eta}{6}, R_{2,P} - \frac{\eta}{6})$ be achievable with the RHK-NOF and satisfy for all $i \in \{1, 2\}$,

$$R_{i,C} + R_{i,P} + R_{j,C} + R_{j,R} = \max(\overrightarrow{n}_{ii}, n_{ij}) + \frac{2}{3}\eta. \tag{64}$$

Then, the rate pair (R_1, R_2), with $R_i = R_{i,C} + R_{i,P} - \frac{1}{3}\eta$ is achievable at an η-Nash equilibrium.

Proof: Let $(s_1^*, s_2^*) \in \mathcal{A}_1 \times \mathcal{A}_2$ be an transmit-receive configuration pair, in which the individual strategy s_i^* is a RHK-NOF satisfying condition (64). From the assumptions of the lemma, it follows that (s_1^*, s_2^*) is an η-NE at which $u_1(s_1^*, s_2^*) = R_{1,C} + R_{1,P} - \frac{1}{3}\eta$ and $u_2(s_1^*, s_2^*) = R_{2,C} + R_{2,P} - \frac{1}{3}\eta$. Consider that such a transmit-receive configuration pair (s_1^*, s_2^*) is not an η-Nash equilibrium. Then, from Definition 2, there exists at least one $i \in \{1, 2\}$ and at least one strategy $s_i \in \mathcal{A}_i$ such that the utility u_i is improved by at least η bits per channel use when transmitter-receiver pair i deviates from s_i^* to s_i. Without loss of generality, let $i = 1$ be the deviating user and denote by R_1' the rate achieved after the deviation. Then,

$$u_1(s_1, s_2^*) = R_1' \geqslant u_1(s_1^*, s_2^*) + \eta = R_{1,C} + R_{1,P} + \frac{2}{3}\eta. \tag{65}$$

However, from Lemma 4, it follows that

$$R_1' \leqslant \max\left(\overrightarrow{n}_{11}, n_{12}\right) - \left(R_{2,C} + R_{2,R}\right) + \frac{2}{3}\eta, \tag{66}$$

and from the assumption in (64), with $i = 1$, i.e.,

$$R_{2,C} + R_{2,R} = \max(\overrightarrow{n}_{11}, n_{12}) - (R_{1,C} + R_{1,P}) + \frac{2}{3}\eta, \tag{67}$$

it follows that

$$R_1' \leqslant R_{1,C} + R_{1,P}. \tag{68}$$

The result in (68) contradicts condition (65) for any $\eta > 0$ and shows that there exists no other coding scheme that brings an individual utility improvement higher than η. The same can be proved for the other transmitter-receiver pair. This completes the proof. ■

The following lemma shows that all the rate pairs $(R_1, R_2) \in \mathcal{C} \cap \mathcal{B}_\eta$ can be achieved by at least one η-NE.

Lemma 6 Let $\eta > 0$ be an arbitrarily small number. Then, for all rate pairs $(R_1, R_2) \in \mathcal{C} \cap \mathcal{B}_\eta$, there always exists at least one η-NE transmit-receive configuration pair $(s_1^*, s_2^*) \in \mathcal{A}_1 \times \mathcal{A}_2$, such that $u_1(s_1^*, s_2^*) = R_1$ and $u_2(s_1^*, s_2^*) = R_2$.

Proof: From Lemma 5, it is known that a transmit-receive configuration pair (s_1^*, s_2^*) in which each player's transmit-receive configuration is the RHK-NOF satisfying condition (64) is an η-NE and achieves any rate tuple ($R_{1,C1}$, $R_{1,R1}$, $R_{1,C2}$, $R_{1,R2}$, $R_{1,P}$, $R_{2,C1}$, $R_{2,R1}$, $R_{2,C2}$, $R_{2,R2}$, $R_{2,P}$). Thus, from the conditions in (57) and (64), the following holds:

$$\begin{aligned} R_{j,C1} + R_{j,R1} &\leqslant \theta_{1,i}, \\ R_i + R_{j,C} + R_{j,R} &= \theta_{2,i}, \\ R_{j,C2} + R_{j,R2} &\leqslant \theta_{3,i}, \\ R_{i,P} &\leqslant \theta_{4,i}, \\ R_{i,P} + R_{j,C2} + R_{j,R2} &\leqslant \theta_{5,i}, \\ R_{i,C2} + R_{i,P} &\leqslant \theta_{6,i}, \text{ and} \\ R_{i,C2} + R_{i,P} + R_{j,C2} + R_{j,R2} &\leqslant \theta_{7,i}. \end{aligned} \tag{69}$$

The region characterized by (69) can be written in terms of $R_1 = R_{1,C1} + R_{1,C2} + R_{1,P}$ and $R_2 = R_{2,C1} + R_{2,C2} + R_{2,P}$ following a Fourier-Motzkin elimination process:

$$\begin{aligned} R_1 &\geqslant \theta_{2,1} - \theta_{1,1} - \theta_{3,1}, \\ R_1 &\leqslant \min\left(\theta_{6,1} + \theta_{1,2}, \theta_{2,1} + \theta_{1,2} + \theta_{5,2} - \theta_{2,2}, \theta_{2,1}\right), \\ R_2 &\geqslant \theta_{2,2} - \theta_{1,2} - \theta_{3,2}, \\ R_2 &\leqslant \min\left(\theta_{1,1} + \theta_{6,2}, \theta_{2,2}, \theta_{1,1} + \theta_{5,1} + \theta_{2,2} - \theta_{2,1}\right), \\ R_1 + R_2 &\leqslant \min\left(\theta_{4,1} + \theta_{2,2}, \theta_{2,1} + \theta_{4,2}, \theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{5,2}\right), \\ R_1 + 2R_2 &\leqslant \min\left(\theta_{1,1} + \theta_{5,1} + \theta_{2,2} + \theta_{4,2}, \theta_{1,1} + \theta_{2,1} + \theta_{4,2} + \theta_{6,2}\right), \\ 2R_1 + R_2 &\leqslant \min\left(\theta_{4,1} + \theta_{6,1} + \theta_{1,2} + \theta_{2,2}, \theta_{2,1} + \theta_{4,1} + \theta_{1,2} + \theta_{5,2}\right). \end{aligned} \tag{70}$$

The region described by (70) is identical to $\mathcal{C} \cap \mathcal{B}_\eta$. This completes the proof. ■

D Proof of Theorem 2

The proof of Theorem 2 consists of constructing a coding scheme that satisfies Definition 2. The coding scheme is a generalization to continuous channel inputs of the coding scheme introduced in Appendix C for the linear deterministic interference channel. The difference is that the generation of the codeword $\boldsymbol{x}_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,N}) \in \mathbb{R}^N$ during block $t \in \{1, 2, \ldots, T\}$ is obtained by adding the described codewords , i.e., $\boldsymbol{x}_i = \boldsymbol{u} + \boldsymbol{u}_i + \boldsymbol{v}_i + \boldsymbol{x}_{i,p}$, whose message indices and random indices are dropped by the ease of notation. The rest of the proof consists of showing that this code construction is an η-NE for certain values of η. This is immediate from the following lemmas. Lemma 7 describes all the rate pairs (R_1, R_2) that can be achieved with the RHK-NOF scheme.

Lemma 7 The RHK-NOF scheme achieves the set of non-negative tuples $\Big(R_{1,C1},\ R_{1,R1},\ R_{1,C2},\ R_{1,R2},\ R_{1,P},\ R_{2,C1},\ R_{2,R1},\ R_{2,C2},\ R_{2,R2},\ R_{2,P}\Big)$ that satisfy the following conditions:

$$R_{i,P} \leqslant a_{1,i}, \tag{71a}$$
$$R_i + R_{j,C} + R_{j,R} \leqslant a_{2,i}(\rho), \tag{71b}$$
$$R_{j,C1} + R_{j,R1} \leqslant a_{3,i}(\rho, \mu_j), \tag{71c}$$
$$R_{j,C2} + R_{j,R2} \leqslant a_{4,i}(\rho, \mu_j), \tag{71d}$$
$$R_{i,P} + R_{j,C2} + R_{j,R2} \leqslant a_{5,i}(\rho, \mu_j), \tag{71e}$$
$$R_{i,C2} + R_{i,P} \leqslant a_{6,i}(\rho, \mu_i), \text{ and} \tag{71f}$$
$$R_{i,C2} + R_{i,P} + R_{j,C2} + R_{j,R2} \leqslant a_{7,i}(\rho, \mu_1, \mu_2), \tag{71g}$$

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

Proof: The proof of Lemma 7 is presented in in Appendices A and B. ■

The set of inequalities in (71) can be written in terms of the transmission rates $R_1 = R_{1,C1} + R_{1,C2} + R_{1,P}$, $R_2 = R_{2,C1} + R_{2,C2} + R_{2,P}$, $R_{1,R} = R_{1,R1} + R_{1,R2}$, and $R_{2,R} = R_{2,R1} + R_{2,R2}$ following a Fourier-Motzkin elimination process. The resulting region, when $R_{1,R1} = R_{1,R2} = R_{2,R1} = R_{2,R2} = 0$ corresponds to the region $\underline{\mathcal{C}}_{\mathrm{GIC-NOF}}$ (Theorem 1 in [5]). Therefore, the relevance of Lemma 7 relies on the implication that any rate pair $(R_1, R_2) \in \underline{\mathcal{C}}_{\mathrm{GIC-NOF}}$ is achievable by the RHK-NOF coding scheme, under the assumption that the random rates $R_{1,R1}$, $R_{1,R2}$, $R_{2,R1}$, and $R_{2,R2}$ are properly chosen.

Lemma 8 provides the maximum rate improvement that a given transmitter-receiver pair achieves by unilateral deviation from the R-KH-NOF coding scheme.

Lemma 8 Assume that the rate tuple $\boldsymbol{R} = (R_{1,C1},\ R_{1,R1},\ R_{1,C2},\ R_{1,R2},\ R_{1,P},\ R_{2,C1},\ R_{2,R1},\ R_{2,C2},\ R_{2,R2},\ R_{2,P})$ is achievable with the RHK-NOF. Then, any unilateral deviation of transmitter-receiver pair i by using any other coding scheme leads to a transmission rate R_i' that satisfies:

$$R_i' \leqslant \frac{1}{2}\log\left(1 + \overrightarrow{\mathrm{SNR}}_i + \mathrm{INR}_{ij} + 2\sqrt{\overrightarrow{\mathrm{SNR}}_i \mathrm{INR}_{ij}}\right) - (R_{j,C} + R_{j,R}).$$

Proof: From Lemma 7, it is known that for all rate tuples $(R_1, R_2) \in \underline{\mathcal{C}}_{\mathrm{GIC-NOF}}$ there always exists a rate tuple $\boldsymbol{R} = (R_{1,C1},\ R_{1,R1},\ R_{1,C2},\ R_{1,R2},\ R_{1,P},\ R_{2,C1},\ R_{2,R1},\ R_{2,C2},\ R_{2,R2},\ R_{2,P})$ that satisfies (71). Assume that both transmitters achieve the rates $\boldsymbol{R}$ by using the RHK-NOF coding scheme following the code construction in Appendix B.

Without loss of generality, let transmitter 1 change its transmit-receive configuration while the transmitter-receiver pair 2 remains unchanged. Note that the new transmit-receive configuration of transmitter-receiver pair 1 can be arbitrary, i.e., it may or may not use feedback, and it may or may not use any random symbols. It can also use a new block length $N_1' \neq N_1$. Denote respectively by W_1 and Ω_1 the message index and the random index of transmitter-receiver pair 1 after its deviation. Let also $\boldsymbol{X}_1' = \left(X_{1,1}', X_{1,2}', \ldots, X_{1,N}'\right)$ and $\overrightarrow{\boldsymbol{Y}}'_1 = \left(\overrightarrow{Y}'_{1,1}, \overrightarrow{Y}'_{1,2}, \ldots, \overrightarrow{Y}'_{1,N}\right)$ respectively be the vector of outputs of transmitter 1 and inputs to receiver 1, with $N = \max(N_1', N_2)$.

Hence, an upper bound for R_1' is obtained from the following inequalities:

$$\begin{aligned} R_1' &= H\left(W_1|\Omega_1\right), \\ &\overset{(a)}{\leqslant} I\left(W_1; \overrightarrow{\boldsymbol{Y}}'_1|\Omega_1\right) + N\delta_1(N), \\ &= h\left(\overrightarrow{\boldsymbol{Y}}'_1|\Omega_1\right) - h\left(\overrightarrow{\boldsymbol{Y}}'_1|W_1,\Omega_1\right) + N\delta_1(N), \\ &\overset{(b)}{\leqslant} \frac{N}{2}\log\left(2\pi e\left(\overrightarrow{\mathrm{SNR}}_1 + 2\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right)\right) - h\left(\overrightarrow{\boldsymbol{Y}}'_1|W_1,\Omega_1\right) + N\delta_1(N), \end{aligned} \quad (72)$$

where, (a) follows from Fano's inequality, as the rate R_1' is achievable by assumption and (b) follows from the fact that for all for all $n \in \{1, 2, \ldots, N\}$, $h(\overrightarrow{Y}'_{1,n}|\overrightarrow{Y}'_{1,1}, \overrightarrow{Y}'_{1,2}\ldots, \overrightarrow{Y}'_{1,n-1}, \Omega_1) \leqslant h(\overrightarrow{Y}'_{1,n}) \leqslant \frac{1}{2}\log\left(2\pi e\left(\overrightarrow{\mathrm{SNR}}_1 + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right)\right)$. To refine this upper bound, the term $h\left(\overrightarrow{\boldsymbol{Y}}'_1|W_1,\Omega_1\right)$ in (72) can be lower bounded. Denote by $W_{2,C}$ and Ω_2 the common message index and the random index of transmitter-receiver pair 2 after the deviation of transmitter-receiver pair 1. Hence, the following holds:

$$\begin{aligned} N_2(R_{2,C} + R_{2,R}) &= H(W_{2,C},\Omega_2), \\ &\overset{(d)}{=} H(W_{2,C},\Omega_2|W_1,\Omega_1), \\ &= I(W_{2,C},\Omega_2;\overrightarrow{\boldsymbol{Y}}'_1|W_1,\Omega_1) + H\left(W_{2,C},\Omega_2|\overrightarrow{\boldsymbol{Y}}'_1,W_1,\Omega_1\right), \\ &\overset{(e)}{\leqslant} I(W_{2,C},\Omega_2;\overrightarrow{\boldsymbol{Y}}'_1|W_1,\Omega_1) + N\delta_2(N), \\ &= h(\overrightarrow{\boldsymbol{Y}}'_1|W_1,\Omega_1) - h(\overrightarrow{\boldsymbol{Y}}'_1|W_1,\Omega_1,W_{2,C},\Omega_2) + N\delta_2(N) \\ &\overset{(f)}{\leqslant} h(\overrightarrow{\boldsymbol{Y}}'_1|W_1,\Omega_1) + N\left(\delta_2(N) - \frac{1}{2}\log(2\pi e)\right), \end{aligned} \quad (73)$$

where, (d) follows from the independence of the indices W_1, Ω_1, W_2, and Ω_2; (e) follows from Fano's inequality as the indices $W_{2,C}$ and Ω_2 can be reliably decoded by receiver 1 using the signals $\overrightarrow{\boldsymbol{Y}}'_1$, W_1, and Ω_1 as transmitter-receiver pair 2 did not change its transmit-receive configuration and by assumption of the lemma that the rate tuple $\boldsymbol{R}$ is achievable; and finally, (f) follows from the fact that $h\left(\overrightarrow{\boldsymbol{Y}}'_1|W_1,\Omega_1,W_{2,C},\Omega_2\right) > \frac{N}{2}\log(2\pi e)$. Substituting (73) into (72), it follows that

$$\begin{aligned} R_1' &\leqslant \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) - (R_{2,C} + R_{2,R}) - \frac{1}{2}\log(2\pi e) + \delta(N), \\ &\leqslant \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) - (R_{2,C} + R_{2,R}) + \delta(N). \end{aligned} \quad (74)$$

Note that $\delta(N) = \delta_1(N) + \delta_2(N)$ is a monotonically decreasing functions of N. Hence, in the asymptotic regime, it follows that

$$R_1' \leqslant \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) - (R_{2,C} + R_{2,R}).$$

The same can be proved for the other transmitter-receiver pair 2 and this completes the proof. ■

Note that if there exists an $\eta > 0$ and a rate tuple $\boldsymbol{R} = (R_{1,C1}, R_{1,R1}, R_{1,C2}, R_{1,R2}, R_{1,P}, R_{2,C1}, R_{2,R1}, R_{2,C2}, R_{2,R2}, R_{2,P})$ achievable with the RHK-NOF coding scheme, such that

$R_i' - (R_{i,C} + R_{i,P}) < \eta$, then the rate pair (R_1, R_2), with $R_{1,C} = R_{1,C1} + R_{1,C2}$, $R_{2,C} = R_{2,C1} + R_{2,C2}$, $R_1 = R_{1,P} + R_{1,C}$ and $R_2 = R_{2,P} + R_{2,C}$, is achievable at an η-NE. The following lemma formalizes this observation.

Lemma 9 Let $\eta > 1$ and let the rate tuple $\boldsymbol{R} = (R_{1,C1}, R_{1,R1}, R_{1,C2}, R_{1,R2}, R_{1,P}, R_{2,C1}, R_{2,R1}, R_{2,C2}, R_{2,R2}, R_{2,P})$ be achievable with the RHK-NOF scheme. Let also $\rho \in [0,1]$ and for all $i \in \{1,2\}$,

$$R_{i,C} + R_{i,P} + R_{j,C} + R_{j,R} = \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_i + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_i \mathrm{INR}_{ij}} + \mathrm{INR}_{ij} + 1\right) - \frac{1}{2}. \tag{75}$$

Then, the rate pair (R_1, R_2), with $R_{i,C} = R_{i,C1} + R_{i,C2}$ and $R_i = R_{i,P} + R_{i,C}$ is achievable at an η-NE.

The proof of Lemma 9 follows the same steps as in the proof of Lemma 5.

Proof: Let $s_i^* \in \mathcal{A}_i$ be a transmit-receive configuration in which communication takes place using the RHK-NOF coding scheme and $R_{1,R1}$, $R_{1,R2}$, $R_{2,R1}$, and $R_{2,R2}$ are chosen according to condition (75), with $i = 1$ and $i = 2$, respectively. From the assumptions of the lemma such a transmit-receive configuration pair (s_1^*, s_2^*) is an η-NE and

$$\begin{aligned} u_i(s_1^*, s_2^*) =& R_i \\ =& R_{i,C} + R_{i,P} \\ =& \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_i + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_i \mathrm{INR}_{ij}} + \mathrm{INR}_{ij} + 1\right) - (R_{j,C} + R_{j,R}) - \frac{1}{2}, \end{aligned} \tag{76}$$

where the last equality holds from (75). Then, from Definition 2, it holds that for all $i \in \{1,2\}$ and for all transmit-receive configurations $s_i \neq s_i^* \in \mathcal{A}_i$, the utility improvement is upper bounded by η, that is,

$$u_i(s_i, s_j^*) - u_i(s_i^*, s_j^*) \leqslant \eta. \tag{77}$$

Without loss of generality, let $i = 1$ be the deviating transmitter-receiver pair and assume it achieves the highest improvement (Lemma 8), that is,

$$u_1(s_1, s_2^*) = \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) - (R_{2,C} + R_{2,R}). \tag{78}$$

Hence, replacing (76) and (78) in (77) yields

$$\begin{aligned} u_1(s_1, s_2^*) - u_1(s_1^*, s_2^*) =& \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) \\ & - \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) + \frac{1}{2} \\ \overset{(a)}{\leqslant}& 1 \\ \leqslant& \eta, \end{aligned} \tag{79}$$

where (a) follows from the fact that $\Delta = \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) - \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) + \frac{1}{2}$ satisfies the following inequality:

$$\begin{aligned} \Delta =& \frac{1}{2}\log\left(1 + \frac{2(1-\rho)\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}}}{\overrightarrow{\mathrm{SNR}}_1 + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1}\right) + \frac{1}{2} \\ \leqslant& \frac{1}{2}\log\left(1 + \frac{2\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}}}{\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1}\right) + \frac{1}{2} \end{aligned}$$

$$
\begin{aligned}
&\leqslant \frac{1}{2}\log\left(1+\frac{\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{12}}{\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{12}+1}\right)+\frac{1}{2}\\
&\leqslant \frac{1}{2}\log(2)+\frac{1}{2}\\
&=1\\
&\leqslant \eta.
\end{aligned} \tag{80}
$$

This verifies that any rate improvement by unilateral deviation of the transmit-receive configuration (s_1^*, s_2^*) is upper bounded by any η arbitrarily close to 1, i.e., $\eta \geqslant 1$. The same can be proved for the other transmitter-receiver pair and this completes the proof. ■

Finally, Lemma 10 shows the η-Nash achievable region $\underline{\mathcal{N}}_\eta$ and this completes the proof of Theorem 2.

Lemma 10 For all rate pairs $(R_1, R_2) \in \underline{\mathcal{N}}_\eta$, there always exists at least one η-NE transmit-receive configuration pair $(s_1^*, s_2^*) \in \mathcal{A}_1 \times \mathcal{A}_2$, with $\eta \geqslant 1$, such that $u_1(s_1^*, s_2^*) = R_1$ and $u_2(s_1^*, s_2^*) = R_2$.

Proof: A rate tuple ($R_{1,C1}$, $R_{1,R1}$, $R_{1,C2}$, $R_{1,R2}$, $R_{1,P}$, $R_{2,C1}$, $R_{2,R1}$, $R_{2,C2}$, $R_{2,R2}$, $R_{2,P}$) that is achievable with the RHK-NOF coding scheme satisfies the inequalities in (71). Additionally, any rate tuple ($R_{1,C1}$, $R_{1,R1}$, $R_{1,C2}$, $R_{1,R2}$, $R_{1,P}$, $R_{2,C1}$, $R_{2,R1}$, $R_{2,C2}$, $R_{2,R2}$, $R_{2,P}$) that satisfies (71) and (75) is an η-NE (Lemma 9). A Fourier-Motzkin elimination of inequalities (71) and (75) leads to a region in terms of the rates R_1 and R_2, as follows:

$$
\begin{aligned}
R_1 \geqslant\ & a_{2,1}(\rho) - a_{3,1}(\rho,\mu_2) - a_{4,1}(\rho,\mu_2),\\
R_1 \leqslant\ & \min\Big(a_{2,1}(\rho), a_{6,1}(\rho,\mu_1) + a_{3,2}(\rho,\mu_1), a_{2,1}(\rho) + a_{3,2}(\rho,\mu_1) + a_{5,2}(\rho,\mu_1) - a_{2,2}(\rho),\\
& a_{3,1}(\rho,\mu_2) + a_{7,1}(\rho,\mu_1,\mu_2) + 2a_{3,2}(\rho,\mu_1) + a_{5,2}(\rho,\mu_1) - a_{2,2}(\rho),\\
& a_{2,1}(\rho) + a_{3,1}(\rho,\mu_2) + 2a_{3,2}(\rho,\mu_1) + a_{5,2}(\rho,\mu_1) + a_{7,2}(\rho,\mu_1,\mu_2) - 2a_{2,2}(\rho)\Big),\\
R_2 \geqslant\ & a_{2,2}(\rho) - a_{3,2}(\rho,\mu_1) - a_{4,2}(\rho,\mu_1),\\
R_2 \leqslant\ & \min\Big(a_{3,1}(\rho,\mu_2) + a_{6,2}(\rho,\mu_2), a_{2,2}(\rho), a_{3,1}(\rho,\mu_2) + a_{5,1}(\rho,\mu_2) + a_{2,2}(\rho) - a_{2,1}(\rho),\\
& 2a_{3,1}(\rho,\mu_2) + a_{5,1}(\rho,\mu_2) + a_{7,1}(\rho,\mu_1,\mu_2) + a_{2,2}(\rho) + a_{3,2}(\rho,\mu_1) - 2a_{2,1}(\rho),\\
& 2a_{3,1}(\rho,\mu_2) + a_{5,1}(\rho,\mu_2) + a_{3,2}(\rho,\mu_1) + a_{7,2}(\rho,\mu_1,\mu_2) - a_{2,1}(\rho)\Big),\\
R_1 + R_2 \leqslant\ & \min\Big(a_{1,1} + a_{2,2}(\rho), a_{1,2} + a_{2,1}(\rho),\\
& a_{1,1} + a_{3,1}(\rho,\mu_2) + a_{7,1}(\rho,\mu_1,\mu_2) + a_{2,2}(\rho) + a_{3,2}(\rho,\mu_1) - a_{2,1}(\rho),\\
& a_{3,1}(\rho,\mu_2) + a_{5,1}(\rho,\mu_2) + a_{3,2}(\rho,\mu_1) + a_{5,2}(\rho,\mu_1),\\
& a_{1,1} + a_{3,1}(\rho,\mu_2) + a_{3,2}(\rho,\mu_1) + a_{7,2}(\rho,\mu_1,\mu_2),\\
& a_{3,1}(\rho,\mu_2) + a_{7,1}(\rho,\mu_1,\mu_2) + a_{1,2} + a_{3,2}(\rho,\mu_1),\\
& a_{2,1}(\rho) + a_{3,1}(\rho,\mu_2) + a_{1,2} + a_{3,2}(\rho,\mu_1) + a_{7,2}(\rho,\mu_1,\mu_2) - a_{2,2}(\rho)\Big),\\
R_1 + 2R_2 \leqslant\ & \min\Big(a_{3,1}(\rho,\mu_2) + a_{5,1}(\rho,\mu_2) + a_{1,2} + a_{2,2}(\rho),\\
& a_{3,1}(\rho,\mu_2) + a_{7,1}(\rho,\mu_1,\mu_2) + a_{1,2} + a_{2,2}(\rho),\\
& 2a_{3,1}(\rho,\mu_2) + a_{5,1}(\rho,\mu_2) + a_{1,2} + a_{3,2}(\rho,\mu_1) + a_{7,2}(\rho,\mu_1,\mu_2)\Big),
\end{aligned}
$$

$$2R_1 + R_2 \leqslant \min\Big(a_{1,1} + a_{2,1}(\rho) + a_{3,2}(\rho,\mu_1) + a_{5,2}(\rho,\mu_1),\\ a_{1,1} + a_{3,1}(\rho,\mu_2) + a_{7,1}(\rho,\mu_1,\mu_2) + 2a_{3,2}(\rho,\mu_1) + a_{5,2}(\rho,\mu_1),\\ a_{1,1} + a_{2,1}(\rho) + a_{3,2}(\rho,\mu_1) + a_{7,2}(\rho,\mu_1,\mu_2)\Big). \quad (81)$$

The region (81) corresponds to the Nash achievable region for the two-user GIC-NOF, i.e. $\underline{\mathcal{N}}_\eta$. Finally, the Nash achievable region in (81) can be presented as the intersection of the achievable region *agicnof* (Theorem 1 in [5]) and the region defined in (81). This completes the proof. ■

References

[1] N. Nisan, T. Roughgarden, Éva Tardos, and V. V. Vazirani, *Algorithmic Game Theory.* Cambridge, UK: Cambridge University Press, 2007.

[2] J. F. Nash, "Equilibrium points in n-person games," *Proc. National Academy of Sciences of the United States of America*, vol. 36, no. 1, pp. 48–49, Jan. 1950.

[3] R. A. Berry and D. N. C. Tse, "Shannon meets Nash on the interference channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2821–2836, May. 2011.

[4] 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.

[5] V. Quintero, S. M. Perlaza, I. Esnaola, and J.-M. Gorce, "Approximate capacity region of the two-user Gaussian interference channel with noisy channel-output feedback," (Submitted to) *IEEE Trans. Inf. Theory*, Nov 2016. [Online]. Available: https://arxiv.org/pdf/1611.05322.pdf

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

[7] 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.

[8] V. Quintero, S. M. Perlaza, I. Esnaola, and J.-M. Gorce, "When does channel-output feedback enlarge the capacity region of the two-user linear deterministic interference channel?" INRIA Grenoble - Rhône-Alpes, Lyon, France, Tech. Rep. 8862, Mar. 2016.

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

[11] E. Koutsoupias and C. Papadimitriou, "Worst-case equilibria," in *Proc. 16th Annual Symposium on Theoretical Aspects of Computer Science*, Trier, Germany, Mar. 1999.

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

RESEARCH CENTRE
GRENOBLE – RHÔNE-ALPES

Inovallée
655 avenue de l'Europe Montbonnot
38334 Saint Ismier Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399