Noisy channel-output feedback in the interference channel

Victor Quintero Florez

N° d'ordre NNT : 2017LYSEI128

THÈSE de DOCTORAT DE L'UNIVERSITÉ DE LYON
opérée au sein de
INRIA – Équipe SOCRATE

École Doctorale N° 160
Électronique, Électrotechnique et Automatique
(ED EEA)

Spécialité / discipline de doctorat :
Génie Électrique

Soutenue publiquement le 12/12/2017, par :
Victor Manuel Quintero Florez

Noisy Channel-Output Feedback in the Interference Channel

Devant le jury composé de :

Rapporteurs :		
Michèle Wigger	Professeure	Télécom ParisTech, France.
Abdellatif Zaidi	Maitre de Conférences	Université Paris-Est Marne la Vallée, France.
Examinateurs :		
Inbar Fijalkow	Professeure	Université de Cergy-Pontoise, France.
David Gesbert	Professeur	Eurécom, France.
Gerhard Kramer	Professeur	Technische Universität München, Germany.
H. Vincent Poor	Professeur	Princeton University, USA.
Directeurs de Thèse :		
Jean-Marie Gorce	Professeur	Université de Lyon, France.
Samir M. Perlaza	Chargé de Recherche	INRIA, France.
Membre invité :		
Iñaki Esnaola	Professeur Associé	University of Sheffield, UK.

UNIVERSITY OF LYON
Doctoral School of Electronics, Electrotechnics and Automation
(ED EEA)

THESIS

presented in partial fulfilment of the requirements
for the degree of Doctor of Philosophy from the University of Lyon,
the 12/12/2017.

Specialization: Electrical Engineering

Victor Manuel Quintero Florez

Noisy Channel-Output Feedback in the Interference Channel

Members of the Jury:

Reviewers:		
Michèle Wigger	Professor	Télécom ParisTech, France.
Abdellatif Zaidi	Assistant Professor	Université Paris-Est Marne la Vallée, France.
Examiners:		
Inbar Fijalkow	Professor	Université de Cergy-Pontoise, France.
David Gesbert	Professor	Eurécom, France.
Gerhard Kramer	Professor	Technische Universität München, Germany.
H. Vincent Poor	Professor	Princeton University, USA.
Thesis supervisors:		
Jean-Marie Gorce	Professor	Université de Lyon, France.
Samir M. Perlaza	Research Scientist	INRIA, France.
Guest:		
Iñaki Esnaola	Assistant Professor	University of Sheffield, UK.

UNIVERSITÉ DE LYON
Électronique, Électrotechnique et Automatique
(ED EEA)

THÈSE

présentée publiquement pour l'obtention
du diplôme de Docteur de l'Université de Lyon,
le 12/12/2017.

Spécialité : Génie Électrique

Victor Manuel Quintero Florez

Noisy Channel-Output Feedback in the Interference Channel

Devant le jury composé de :

Rapporteurs :		
Michèle Wigger	Professeure	Télécom ParisTech, France.
Abdellatif Zaidi	Maitre de Conférences	Université Paris-Est Marne la Vallée, France.
Examinateurs :		
Inbar Fijalkow	Professeure	Université de Cergy-Pontoise, France.
David Gesbert	Professeur	Eurécom, France.
Gerhard Kramer	Professeur	Technische Universität München, Germany.
H. Vincent Poor	Professeur	Princeton University, USA.
Directeurs de Thèse :		
Jean-Marie Gorce	Professeur	Université de Lyon, France.
Samir M. Perlaza	Chargé de Recherche	INRIA, France.
Membre invité :		
Iñaki Esnaola	Professeur Associé	University of Sheffield, UK.

Acknowledgements

This thesis is dedicated to the memory of my parents and to all my family, especially Yolanda and Matilde for their encouragement, patience, and unconditional love every day in this amazing journey; and my brother Luis who has always supported me and believed in me.

There have been many people who have walked by my side during these incredible four years, whom I would like to thank. I would like to particularly thank my thesis director, Prof. Jean-Marie Gorce, who has always been encouraging and supporting me with all his knowledge and enthusiasm; my thesis advisor, Dr. Samir M. Perlaza, for his entire dedication during all the days and nights of this PhD, and his enormous contribution with all his knowledge and energy to assure the quality of this thesis; and Prof. Iñaki Esnaola, for all his time and the invaluable contributions and advices to improve this thesis. This thesis would not have been completed without their commitment and participation.

My deepest gratitude goes to all the people in the CITI Laboratory. Especially, I would like to thank Dr. Selma Belhadj Amor, Dr. Leonardo S. Cardoso, Dr. Malcolm Egan, David Kibloff, and Nizar Kalfhet for all their technical contributions and their friendship. I am also grateful to Gaelle Tworkowski, Joelle Charnay, Margarita Raimbaud, Aude Montillet, Alexia Cottin, Caroline Tremeaud, Sophie Karsai, Delphine Diner, Marina Da Graca, Giulietta Randisi, and Sandrine Dahan who have made great the experience of being part of the CITI Laboratory.

Abstract

In this thesis, the two-user Gaussian interference channel with noisy channel-output feedback (GIC-NOF) is studied from two perspectives: centralized and decentralized networks.

From the perspective of centralized networks, the fundamental limits of the two-user GIC-NOF are characterized by the capacity region. One of the main contributions of this thesis is an approximation to within a constant number of bits of the capacity region of the two-user GIC-NOF. This result is obtained through the analysis of a simpler channel model, *i.e.*, a two-user linear deterministic interference channel with noisy channel-output feedback (LDIC-NOF). The analysis to obtain the capacity region of the two-user LDIC-NOF provides the main insights required to analyze the two-user GIC-NOF. From the perspective of decentralized networks, the fundamental limits of the two-user decentralized GIC-NOF (D-GIC-NOF) are characterized by the η-Nash equilibrium (η-NE) region. Another contribution of this thesis is an approximation of the η-NE region of the two-user GIC-NOF, with $\eta > 1$. As in the centralized case, the two-user decentralized LDIC-NOF (D-LDIC-NOF) is studied first and the lessons learnt are applied in the two-user D-GIC-NOF.

The final contribution of this thesis consists in a closed-form answer to the question: "When does channel-output feedback enlarge the capacity or η-NE regions of the two-user GIC-NOF or two-user D-GIC-NOF?". This answer is of the form: *Implementing channel-output feedback in transmitter-receiver i enlarges the capacity or η-NE regions if the feedback SNR is beyond* SNR_i^*, with $i \in \{1, 2\}$. The approximate value of SNR_i^* is shown to be a function of all the other parameters of the two-user GIC-NOF or two-user D-GIC-NOF.

Résumé

Dans cette thèse, le canal Gaussien à interférence à deux utilisateurs avec voie de retour dégradée par un bruit additif (GIC-NOF) est étudié sous deux perspectives : les réseaux centralisés et décentralisés.

Du point de vue des réseaux centralisés, les limites fondamentales du GIC-NOF sont caractérisées par la région de capacité. L'une des principales contributions de cette thèse est une approximation à un nombre constant de bits près de la région de capacité du GIC-NOF. Ce résultat est obtenu grâce à l'analyse d'un modèle de canal plus simple, le canal linéaire déterministe à interférence à deux utilisateurs avec voie de retour dégradée par un bruit additif (LDIC-NOF). L'analyse pour obtenir la région de capacité du LDIC-NOF fournit les idées principales pour l'analyse du GIC-NOF.

Du point de vue des réseaux décentralisés, les limites fondamentales du GIC-NOF sont caractérisées par la région d'η-équilibre de Nash (η-EN). Une autre contribution de cette thèse est une approximation de la région η-EN du GIC-NOF, avec $\eta > 1$. Comme dans le cas centralisé, le cas décentralisé LDIC-NOF (D-LDIC-NOF) est étudié en premier et les observations sont appliquées dans le cas décentralisé GIC-NOF (D-GIC-NOF).

La contribution finale de cette thèse répond à la question suivante : "À quelles conditions la voie de retour permet d'agrandir la région de capacité, la région η-EN du GIC-NOF ou du D-GIC-NOF ? ". La réponse obtenue est de la forme : L'implémentation de la voie de retour de la sortie du canal dans l'émetteur-récepteur i agrandit la région de capacité ou la région η-EN si le rapport signal sur bruit de la voie de retour est supérieure à SNR_i^*, avec $i \in \{1, 2\}$. La valeur approximative de SNR_i^* est une fonction de tous les autres paramètres du GIC-NOF ou du D-GIC-NOF.

Publications

Journals

[1] "*When Does Output Feedback Enlarge the Capacity of the Interference Channel?*". Victor Quintero, Samir M. Perlaza, Iñaki Esnaola, and Jean-Marie Gorce. IEEE Transactions on Communications, vol. 66 no. 2, pp. 615-628, Feb., 2018.

[2] "*Approximate Capacity Region of the Two-User Gaussian Interference Channel with Noisy Channel-Output Feedback*". Victor Quintero, Samir M. Perlaza, Iñaki Esnaola, and Jean-Marie Gorce. This work will appear in the IEEE Transactions on Information Theory. Submitted on Nov. 10, 2016; revised on Sep. 23, 2017; and accepted on Mar. 4, 2018.

Conferences

[1] "*Noisy Channel-Output Feedback Capacity of the Linear Deterministic Interference Channel*". Victor Quintero, Samir M. Perlaza, and Jean-Marie Gorce. IEEE Information Theory Workshop (ITW), Jeju, Korea, Oct., 2015.

[2] "*Approximate Capacity of the Gaussian Interference Channel with Noisy Channel-Output Feedback*". Victor Quintero, Samir M. Perlaza, Iñaki Esnaola, and Jean-Marie Gorce. IEEE Information Theory Workshop (ITW), Cambridge, UK, Sep., 2016.

[3] "*Nash Region of the Linear Deterministic Interference Channel with Noisy Output Feedback*". Victor Quintero, Samir M. Perlaza, Jean-Marie Gorce, and H. Vincent Poor. IEEE Intl. Symposium on Information Theory (ISIT), Aachen, Germany, Jun., 2017.

INRIA Technical Reports

[1] "*Noisy Channel-Output Feedback Capacity of the Linear Deterministic Interference Channel*". Victor Quintero, Samir M. Perlaza, and Jean-Marie Gorce. Technical Report, INRIA, No. 456, Lyon, France, Jan., 2015.

[2] "*Approximate Capacity of the Two-User Gaussian Interference Channel with Noisy Channel-Output Feedback*". Victor Quintero, Samir M. Perlaza, Iñaki Esnaola, and Jean-Marie Gorce. Technical Report, INRIA, No. 8861, Lyon, France, Mar., 2016.

[3] "*When Does Channel-Output Feedback Enlarge the Capacity Region of the Interference Channel?*". Victor Quintero, Samir M. Perlaza, Iñaki Esnaola, and Jean-Marie Gorce. Technical Report, INRIA, No. 8862, Lyon, France, Mar., 2016.

[4] "*Decentralized Interference Channels with Noisy Output Feedback*". Victor Quintero, Samir M. Perlaza, Jean-Marie Gorce, and H. Vincent Poor. Technical Report, INRIA, No. RR-9011, Lyon, France, Jan., 2017.

Other Publications

[1] "*On the Benefits of Channel-Output Feedback in 5G*". Victor Quintero, Samir M. Perlaza, and Jean-Marie Gorce. GDR-ISIS meeting: 5G & Beyond: Promises and Challenges, Paris, Oct., 2014. (Invited talk).

[2] "*Noisy Channel-Output Feedback Capacity of the Linear Deterministic Interference Channel*". Victor Quintero, Samir M. Perlaza, and Jean-Marie Gorce. European School of Information Theory (ESIT), Zandvoort, The Netherlands, Apr., 2015. (Poster).

[3] "*Capacité du Canal Linéaire Déterministe à Interférences avec Voies de Retour Bruitées*". Victor Quintero, Samir M. Perlaza, and Jean-Marie Gorce. Colloque GRETSI, Lyon, France, Sep., 2015.

[4] "*Does the Channel-Output Feedback improve the performance of the Two-User Linear Deterministic Interference Channel?*". Victor Quintero, Samir M. Perlaza, Iñaki Esnaola, and Jean-Marie Gorce. INRIA-GDR-ISIS meeting: Recent Advances in Network Information Theory and Coding Theory, Villeurbanne, France, Nov., 2015. (Poster).

[5] "*When Does Channel-Output Feedback Increase the Capacity Region of the Two-User Linear Deterministic Interference Channel?*". Victor Quintero, Samir M. Perlaza, Iñaki Esnaola, and Jean-Marie Gorce. 11th International Conference on Cognitive Radio Oriented Wireless Networks (CROWNCOM), Grenoble, France, May, 2016.

[6] "*On the Efficiency of Nash Equilibria in the Interference Channel with Noisy Feedback*". Victor Quintero, Samir M. Perlaza, and Jean-Marie Gorce. European Wireless Conference. Workshop COCOA - COmpetitive and COoperative Approaches for 5G networks, Dresden, Germany, May, 2017. (Invited Paper).

[7] "*Région d'η-Équilibre de Nash du Canal Linéaire Déterministe à Interférences avec Rétroalimentation Dégradée*". Victor Quintero, Samir M. Perlaza, and Jean-Marie Gorce. Colloque GRETSI, Juan-les-Pins, France, Sep., 2017.

Contents

List of Figures

Acronyms

η-NE	η-Nash Equilibrium
AEP	Asymptotic Equipartition Property
AWGN	Additive White Gaussian Noise
BC	Broadcast Channel
D-GIC	Decentralized Gaussian Interference Channel
D-GIC-NOF	Decentralized Gaussian Interference Channel with Noisy Channel-Output Feedback
D-LDIC	Decentralized Linear Deterministic Interference Channel
D-LDIC-NOF	Decentralized Linear Deterministic Interference Channel with Noisy Channel-Output Feedback
FDM	Frequency Division Multiplexing
GDoF	Generalized Degrees of Freedom
GIC	Gaussian Interference Channel
GIC-NOF	Gaussian Interference Channel with Noisy Channel-Output Feedback
GIC-POF	Gaussian Interference Channel with Perfect Channel-Output Feedback
GIC-RLF	Gaussian Interference Channel with Rate-Limited Feedback
HIR	High-Interference Regime
i.i.d.	Independent and Identically Distributed
IC	Interference Channel
IC-CT	Interference Channel with Conferencing Transmitters
IC-GF	Interference Channel with Generalized Feedback
IC-NOF	Interference Channel with Noisy Channel-Output Feedback
IF	Intermittent Feedback
INR	Interference-to-Noise Ratio

LDIC	Linear Deterministic Interference Channel
LDIC-NOF	Linear Deterministic Interference Channel with Noisy Channel-Output Feedback
LDIC-POF	Linear Deterministic Interference Channel with Perfect Channel-Output Feedback
LIR	Low-Interference Regime
MAC	Multiple Access Channel
NE	Nash Equilibrium
NOF	Noisy Channel-Output Feedback
pdf	Probability Density Function
pmf	Probability Mass Function
POF	Perfect Channel-Output Feedback
RC	Relay Channel
RHK-NOF	Randomized Han-Kobayashi scheme with Noisy Channel-Output Feedback
RLF	Rate-Limited Feedback
SIC	Successive Interference Cancellation
SNR	Signal-to-Noise Ratio
TDM	Time Division Multiplexing
TIN	Treating Interference as Noise
WC	Wiretap Channel

Notation

THROUGHOUT this thesis, sets are denoted with uppercase calligraphic letters, e.g. $\mathcal{X}$. Random variables are denoted by uppercase letters, *e.g.*, X. The realization and the set of events from which the random variable X takes values are respectively denoted by x and $\mathcal{X}$.

For discrete random variables, the probability mass function (pmf) of X over the set $\mathcal{X}$ is denoted by $P_X : \mathcal{X} \rightarrow [0,1]$. The support of P_X is $\text{supp}\,(P_X) = \{x \in \mathcal{X} : P_X\,(x) > 0\}$. Whenever a second discrete random variable Y is considered, P_{XY} and $P_{Y|X}$ denote respectively the joint pmf of (X,Y), *i.e.*, $P_{XY} : \mathcal{X} \times \mathcal{Y} \rightarrow [0,1]$, and the conditional pmf of Y given X, *i.e.*, $P_{Y|X} : \mathcal{X} \times \mathcal{Y} \rightarrow [0,1]$. $\mathbb{E}_X[\cdot]$ denotes the expectation with respect to the distribution of the random variable X. For real-valued random variables, the probability density function (pdf) of X is denoted by $f_X : \mathbb{R} \rightarrow [0,\infty)$. The support of f_X is $\text{supp}\,(f_X) = \{x \in \mathbb{R} : f_X\,(x) > 0\}$. Whenever a second real-valued random variable Y is considered, f_{XY} and $f_{Y|X}$ denote respectively the joint pdf of (X,Y), *i.e.*, $f_{XY} : \mathbb{R}^2 \rightarrow [0,\infty)$, and the conditional pdf of Y given X, *i.e.*, $f_{Y|X} : \mathbb{R}^2 \rightarrow [0,\infty)$.

Let N be a fixed natural number. An N-dimensional vector of random variables is denoted by $\boldsymbol{X} = (X_1, X_2, ..., X_N)^{\mathsf{T}}$ and a corresponding realization is denoted by $\boldsymbol{x} = (x_1, x_2, ..., x_N)^{\mathsf{T}} \in \mathcal{X}^N$. Given $\boldsymbol{X} = (X_1, X_2, ..., X_N)^{\mathsf{T}}$ 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 $\boldsymbol{X}_{(a:b)} = (X_a, X_{a+1}, \ldots, X_b)^{\mathsf{T}}$. If the component a of the N-dimensional vector of random variables $\boldsymbol{X}$ is also a q-dimensional vector, it is denoted by $\boldsymbol{X}_a$. Given $(c,d) \in \mathbb{N}^2$, with $c < d \leqslant q$, the $(d-c+1)$-dimensional vector formed by the components c to d of $\boldsymbol{X}_a$ is denoted by $\boldsymbol{X}_a^{(c:d)} = \left(X_a^{(c)}, X_a^{(c+1)}, \ldots, X_a^{(d)}\right)^{\mathsf{T}}$. The notation $(\cdot)^+$ denotes the positive part operator, *i.e.*, $(\cdot)^+ = \max(\cdot, 0)$ The logarithm function log is assumed to be in base 2.

— 0 —

Synthèse des Contributions

Cette thèse a pour objet d'étude le GIC-NOF asymétrique à deux utilisateurs. L'analyse est réalisée en tenant compte de deux scénarios : (1) centralisé, dans lequel la totalité du réseau est contrôlée par une entité centrale qui configure les deux paires émetteur-récepteur ; et (2) décentralisé, dans lequel chaque paire émetteur-récepteur configure de façon autonome ses paramètres de transmission-réception. L'analyse dans ces deux scénarios permet d'approximer la région de capacité ainsi que la région d'η-équilibre de Nash du GIC-NOF à deux utilisateurs. Ces résultats permettent également d'identifier les conditions dans lesquelles une voie de retour peut agrandir la région de capacité et la région d'équilibre.

Voici les principales contributions de cette thèse :

- Une description complète de la région de capacité du LDIC-NOF à deux utilisateurs [70, 72]. Cette contribution généralise les résultats pour les cas du LDIC sans voie de retour [20], avec voie de retour parfaite (LDIC-POF) [88], avec voie de retour dégradée par un bruit additif (LDIC-NOF) dans des conditions symétriques [53], et les cas incluant une voie de retour des récepteurs vers leurs émetteurs appairés [80].

- Une région atteignable et une région d'impossibilité du LDIC-NOF à deux utilisateurs [69, 70]. Ces deux régions encadrent la région de capacité du GIC-NOF à deux utilisateurs avec un écart maximal de 4.4 bits. La région atteignable est obtenue à l'aide d'un argument de codage aléatoire combinant le fractionnement de message, un codage de superposition de bloc Markov et un décodage à rebours, comme suggéré précédemment en [88, 94, 102]. La région d'impossibilité est obtenue en utilisant des bornes supérieures existantes dans le cas du CGI à deux utilisateurs avec POF (GIC-POF) [88] ainsi qu'un ensemble de nouvelles bornes supérieures obtenues à l'aide de modèles assistés par génies.

Cette contribution généralise les résultats obtenus à partir des cas sans voie de retour (CGI) [31], avec POF (GIC-POF) [88], et avec NOF (GIC-NOF) dans des conditions symétriques [53].

- Une description complète de la région d'η-équilibre de Nash du LDIC-NOF à deux utilisateurs [74]. Cette contribution généralise les résultats obtenus à partir des cas du canal linéaire déterministe à interférence (LDIC) sans voie de retour [15], avec POF (GIC-POF) [66], et avec NOF (GIC-NOF) dans des conditions symétriques [68].
- Une région d'η-équilibre de Nash atteignable et une région de deséquilibre avec $\eta \geqslant 1$ pour le GIC-NOF à deux utilisateurs [73]. La région d'η-équilibre de Nash atteignable est obtenue en effectuant une modification du schéma de codage utilisé dans la partie centralisée. Cette modification implique l'introduction d'un caractère aléatoire commun dans le schéma de codage comme suggéré par [16] et [66], ce qui permet aux deux paires d'émetteur-récepteur de limiter l'amélioration de débit de chacune lorsque l'une d'elles s'écarte de l'équilibre. La région de deséquilibre est obtenue avec $\eta \geqslant 1$ en utilisant les connaissances obtenues à partir de l'analyse du modèle linéaire déterministe.
- Identification des scénarios dans lesquels l'utilisation d'une voie de retour agrandit la région de capacité et la région d'η-équilibre de Nash [71].

1. Canaux à Interférence Centralisés

Considérons le IC-NOF continu à deux utilisateurs representé par la Figure 1. L'émetteur i, $i \in \{1, 2\}$, souhaite communiquer au récepteur i un message W_i. Ce message est représenté par une variable aléatoire indépendante et uniformément distribuée prenant des valeurs d'indices dans $\mathcal{W}_i = \{1, 2, \ldots, 2^{NR_i}\}$, durant $N \in \mathbb{N}$ utilisations du canal, $R_i \in \mathbb{R}_+$ indiquant le taux de transmission de l'émetteur-récepteur i en bits par utilisation du canal. À cet égard, l'émetteur i envoie le mot-code $\boldsymbol{X}_i = (X_{i,1}, X_{i,2}, \ldots, X_{i,N})^{\mathsf{T}} \in \mathcal{C}_i \subseteq \mathbb{R}^N$, $\mathcal{C}_i$ étant le dictionnaire de l'émetteur.

Pour une utilisation du canal donné $n \in \{1, 2, \ldots, N\}$, les émetteurs 1 et 2 envoient les symboles $X_{1,n} \in \mathbb{R}$ et $X_{2,n} \in \mathbb{R}$, respectivement, génèrent les sorties de canal $\overrightarrow{Y}_{1,n} \in \mathbb{R}$, $\overrightarrow{Y}_{2,n} \in \mathbb{R}$, $\overleftarrow{Y}_{1,n} \in \mathbb{R}$, et $\overleftarrow{Y}_{2,n} \in \mathbb{R}$ conformément à la fonction de densité de probabilité conditionnelle $f_{\overrightarrow{Y}_1, \overrightarrow{Y}_2, \overleftarrow{Y}_1, \overleftarrow{Y}_2 | X_1, X_2}\left(\overrightarrow{y}_1, \overrightarrow{y}_2, \overleftarrow{y}_1, \overleftarrow{y}_2 | x_1, x_2\right)$, pour tout $(\overrightarrow{y}_1, \overrightarrow{y}_2, \overleftarrow{y}_1, \overleftarrow{y}_2, x_1, x_2) \in \mathbb{R}^6$.

L'émetteur i génère le symbole $X_{i,n} \in \mathbb{R}$ en prenant en compte l'indice de message W_i et tous les résultats précédents de la voie de retour i, à savoir, $\left(\overleftarrow{Y}_{i,1}, \overleftarrow{Y}_{i,2}, \ldots, \overleftarrow{Y}_{i,n-1}\right)$. L'émetteur i observe $\overleftarrow{Y}_{i,n}$ à la fin de la n-ième utilisation du canal. L'émetteur i est défini par l'ensemble des fonctions déterministes $\{f_{i,1}, f_{i,2}, \ldots, f_{i,N}\}$, avec $f_{i,1} : \mathcal{W}_i \rightarrow \mathbb{R}$ et pour $n \in \{2, 3, \ldots, N\}$, $f_{i,n} : \mathcal{W}_i \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}$, de sorte que

$$X_{i,1} = f_{i,1}(W_i) \text{ et} \tag{1a}$$

$$X_{i,n} = f_{i,n}\left(W_i, \overleftarrow{Y}_{i,1}, \overleftarrow{Y}_{i,2}, \ldots, \overleftarrow{Y}_{i,n-1}\right) \text{ pour tout } n > 1. \tag{1b}$$

À la fin de la transmission, le récepteur i utilise toutes les sorties de canal $\left(\overrightarrow{Y}_{i,1}, \overrightarrow{Y}_{i,2}, \ldots, \overrightarrow{Y}_{i,N}\right)^{\mathsf{T}}$ pour obtenir une estimation de l'indice de message W_i, notée $\widehat{W}_i$.

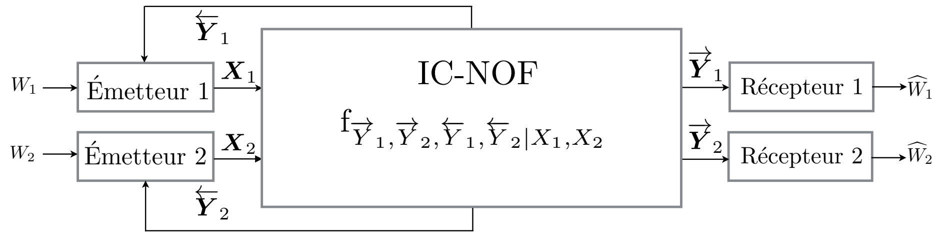


FIGURE 1. : Canal à interférence continu à deux utilisateurs avec voie de retour dégradée par un bruit additif.

Ainsi, la chaîne de Markov suivante existe :

$$\left(W_i, \overleftarrow{\boldsymbol{Y}}_{i,(1;n-1)}\right) \rightarrow X_{i,n} \rightarrow \overrightarrow{Y}_{i,n}. \tag{2}$$

Soit $T \in \mathbb{N}$ fixé. Supposons que, durant une communication, soient transmis T blocs, chaque bloc consistant en N utilisations du canal. Le récepteur i est défini par la fonction déterministe $\psi_i : \mathbb{R}^{NT} \rightarrow \mathcal{W}_i^T$. À la fin de la communication, le récepteur i utilise le vecteur $\left(\overrightarrow{Y}_{i,1}, \overrightarrow{Y}_{i,2}, \ldots, \overrightarrow{Y}_{i,NT}\right)^{\mathsf{T}}$ pour obtenir

$$\left(\widehat{W}_i^{(1)}, \widehat{W}_i^{(2)}, \ldots, \widehat{W}_i^{(T)}\right) = \psi_i\left(\overrightarrow{Y}_{i,1}, \overrightarrow{Y}_{i,2}, \ldots, \overrightarrow{Y}_{i,NT}\right), \tag{3}$$

où $\widehat{W}_i^{(t)}$ est une estimation de l'indice de message $W_i^{(t)}$ envoyé au cours du bloc $t \in \{1, 2, \ldots, T\}$.

La probabilité d'erreur de décodage dans le IC continu à deux utilisateurs durant le bloc t, notée $P_{\mathrm{e}}^{(t)}(N)$, est donnée par

$$P_{\mathrm{e}}^{(t)}(N) = \max\left(\Pr\left[\widehat{W_1}^{(t)} \neq W_1^{(t)}\right], \Pr\left[\widehat{W_2}^{(t)} \neq W_2^{(t)}\right]\right). \tag{4}$$

La définition d'une paire de débit atteignable $(R_1, R_2) \in \mathbb{R}_+^2$ figure ci-dessous.

Définition i (Paires de débit atteignables)**.** *Une paire de débit $(R_1, R_2) \in \mathbb{R}_+^2$ est atteignable s'il existe des ensembles des fonctions d'encodage $\left\{f_1^{(1)}, f_1^{(2)}, \ldots, f_1^{(N)}\right\}$ et $\left\{f_2^{(1)}, f_2^{(2)}, \ldots, f_2^{(N)}\right\}$, et des fonctions de décodage ψ_1 et ψ_2, telles que la probabilité d'erreur $P_{\mathrm{e}}^{(t)}(N)$ peut être arbitrairement réduite lorsque le nombre d'utilisations du canal N tend vers l'infini pour tous les blocs $t \in \{1, 2, \ldots, T\}$.*

Dans un système centralisé, un contrôleur central détermine les configurations de toutes les paires d'émetteur-récepteur. Le contrôleur central a une vision globale du réseau et permet de sélectionner des configurations optimales étant donné une métrique, par exemple, la somme des débits, l'efficacité énergétique, etc. Les limites fondamentales dans un système centralisé sont caractérisées par la région de capacité.

Définition ii (Région de capacité d'un IC à deux utilisateurs)**.** *La région de capacité d'un IC à deux utilisateurs est la fermeture de l'ensemble de toutes les paires possibles de débit atteignables $(R_1, R_2) \in \mathbb{R}_+^2$.*

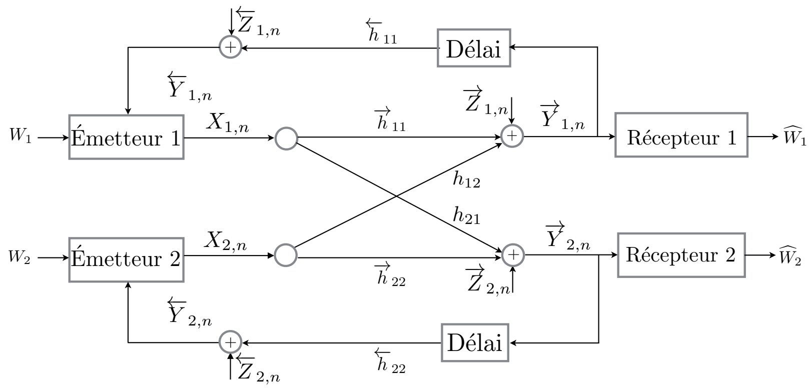


FIGURE 2. : Canal Gaussien à interférence avec voie de retour dégradée par un bruit additif à la n-ième utilisation du canal.

1.1. Canal Gaussien à Interférence

Le canal Gaussien à interférence (GIC) représente un cas particulier du IC-NOF décrit ci-dessus dans la perspective des réseaux centralisés. Considérons le GIC-NOF à deux utilisateurs décrit par la Figure 2. Le coefficient de canal de l'émetteur j au récepteur i est noté h_{ij} ; le coefficient de canal à partir de l'émetteur i au récepteur i est noté $\overrightarrow{h}_{ii}$; et le coefficient de canal de la sortie de canal i l'émetteur i est noté $\overleftarrow{h}_{ii}$. Tous les coefficients de canal sont supposés être des nombres réels non négatifs. A la n-ième utilisation du canal,

$$\overrightarrow{Y}_{i,n}=\overrightarrow{h}_{ii}X_{i,n}+h_{ij}X_{j,n}+\overrightarrow{Z}_{i,n}, \tag{5}$$

et

$$\overleftarrow{Y}_{i,n}=\begin{cases}\overleftarrow{Z}_{i,n} & \text{pour } n\in\{1,2,\ldots,d\}\\ \overleftarrow{h}_{ii}\overrightarrow{Y}_{i,n-d}+\overleftarrow{Z}_{i,n} & \text{pour } n\in\{d+1,d+2,\ldots,N\}\end{cases}, \tag{6}$$

où $\overrightarrow{Z}_{i,n}$ et $\overleftarrow{Z}_{i,n}$ sont des variables aléatoires Gaussiennes standard et où $d>0$ est le délai de la voie de retour, fini, mesuré en utilisations du canal. Sans perte de généralité, le délai de la voie de retour est supposé fixé et égal à une utilisation du canal, c'est à dire, $d=1$. Le vecteur d'entrée $\boldsymbol{X}_i$, dont les composantes sont réelles, est soumis à une contrainte de puissance moyenne :

$$\frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left[X_{i,n}^2\right]\leq 1, \tag{7}$$

l'espérance étant calculée d'après la distribution conjointe des messages W_1, W_2, et les termes de bruits, à savoir $\overrightarrow{Z}_1$, $\overrightarrow{Z}_2$, $\overleftarrow{Z}_1$, et $\overleftarrow{Z}_2$. La dépendance de $X_{i,n}$ avec W_1, W_2, et les réalisations de bruit observées précédemment sont dues à l'effet de la voie de retour comme l'indiquent (1) et (6).

Le GIC-NOF à deux utilisateurs représenté par la Figure 2 peut être décrit par six paramètres : $\overrightarrow{\mathrm{SNR}}_i$, $\overleftarrow{\mathrm{SNR}}_i$, et INR_{ij}, avec $i\in\{1,2\}$ et $j\in\{1,2\}\setminus\{i\}$, définis comme suit :

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

$$\mathrm{INR}_{ij} = h_{ij}^2, \text{ et} \tag{8b}$$

$$\overleftarrow{\mathrm{SNR}}_i = \overleftarrow{h}_{ii}^{\,2} \left(\overrightarrow{h}_{ii}^{\,2} + 2\overrightarrow{h}_{ii} h_{ij} + h_{ij}^2 + 1 \right). \tag{8c}$$

Lorsque $\mathrm{INR}_{ij} \leqslant 1$, la paire d'émetteur-récepteur i est altérée principalement par le bruit au lieu de l'interférence. Dans ce cas, le traitement de l'interférence comme un bruit est optimal et la voie de retour n'apporte pas d'amélioration significative du débit. C'est la raison pour laquelle l'analyse développée dans cette thèse traite exclusivement le cas dans lequel $\mathrm{INR}_{ij} > 1$ pour tout $i \in \{1, 2\}$ et $j \in \{1, 2\} \setminus \{i\}$.

1.2. Canal Linéaire Déterministe à Interférence

On considère le canal linéaire déterministe à interférence avec rétroalimentation dégradée par un bruit additif (LDIC-NOF) représenté sur la Figure 3. Pour tout $i \in \{1, 2\}$ et $j \in \{1, 2\} \setminus \{i\}$, le nombre de tubes pour les bits entre l'émetteur i et son récepteur associé (récepteur i) est noté par $\overrightarrow{n}_{ii}$; le nombre de voies pour les bits entre l'émetteur i et l'autre récepteur (récepteur j) est noté par n_{ji} ; enfin le nombre de tubes pour les bits entre le récepteur i et son émetteur correspondant est noté par $\overleftarrow{n}_{ii}$. Ces six paramètres sont des entiers positifs et décrivent le LDIC-NOF représenté sur la Figure 3.

A l'émetteur i, durant la n-ième utilisation du canal, où $n \in \{1, 2, \ldots, N\}$, le symbole en entrée, noté $\boldsymbol{X}_{i,n}$ est un vecteur binaire de longueur q, c'est-à-dire $\boldsymbol{X}_{i,n} = \left(X_{i,n}^{(1)}, X_{i,n}^{(2)}, \ldots, X_{i,n}^{(q)} \right)^{\mathsf{T}} \in \mathcal{X}_i$, avec $\mathcal{X}_i = \{0, 1\}^q$, $q = \max\left(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21} \right)$, et $N \in \mathbb{N}$ est la longueur de bloc. Au récepteur i, durant la n-ième utilisation du canal, où $n \in \{1, 2, \ldots, \max(N_1, N_2)\}$, le symbole à la sortie du canal i, noté $\overrightarrow{\boldsymbol{Y}}_{i,n}$, est également un vecteur binaire de longueur q, c'est-à-dire $\overrightarrow{\boldsymbol{Y}}_{i,n} = \left(\overrightarrow{Y}_{i,n}^{(1)}, \overrightarrow{Y}_{i,n}^{(2)}, \ldots, \overrightarrow{Y}_{i,n}^{(q)} \right)^{\mathsf{T}}$. Soit $\boldsymbol{S}$ une matrice $q \times q$ à décalage.

Les entrées-sorties de ce canal durant la n-ième utilisation du canal sont liées par les relations suivantes :

$$\overrightarrow{\boldsymbol{Y}}_{i,n} = \boldsymbol{S}^{q - \overrightarrow{n}_{ii}} \boldsymbol{X}_{i,n} + \boldsymbol{S}^{q - n_{ij}} \boldsymbol{X}_{j,n}, \tag{9}$$

pour tout $i \in \{1, 2\}$ et $j \in \{1, 2\} \setminus \{i\}$.

Le signal dégradé par un bruit additif dans la voie de retour est disponible à l'émetteur i à la fin de la n-ième utilisation du canal est :

$$\left((0, \ldots, 0), \overleftarrow{\boldsymbol{Y}}_{i,n}^{\mathsf{T}} \right)^{\mathsf{T}} = \boldsymbol{S}^{\left(\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii} \right)^+} \overrightarrow{\boldsymbol{Y}}_{i,n-d}, \tag{10}$$

où $d \in \mathbb{N}$ est un délai limité. On notera que les additions et les multiplications sont en binaire. La dimension du vecteur $(0, \ldots, 0)$ dans (10) est $q - \min\left(\overleftarrow{n}_{ii}, \max(\overrightarrow{n}_{ii}, n_{ij}) \right)$ et le vecteur $\overleftarrow{\boldsymbol{Y}}_{i,n}$ représente les $\min\left(\overleftarrow{n}_{ii}, \max(\overrightarrow{n}_{ii}, n_{ij}) \right)$ bits moins significatifs de $\boldsymbol{S}^{\left(\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii} \right)^+} \overrightarrow{\boldsymbol{Y}}_{i,n-d}$.

Sans perte de généralité, le délai de la voie de retour est supposé correspondre à une seule utilisation du canal.

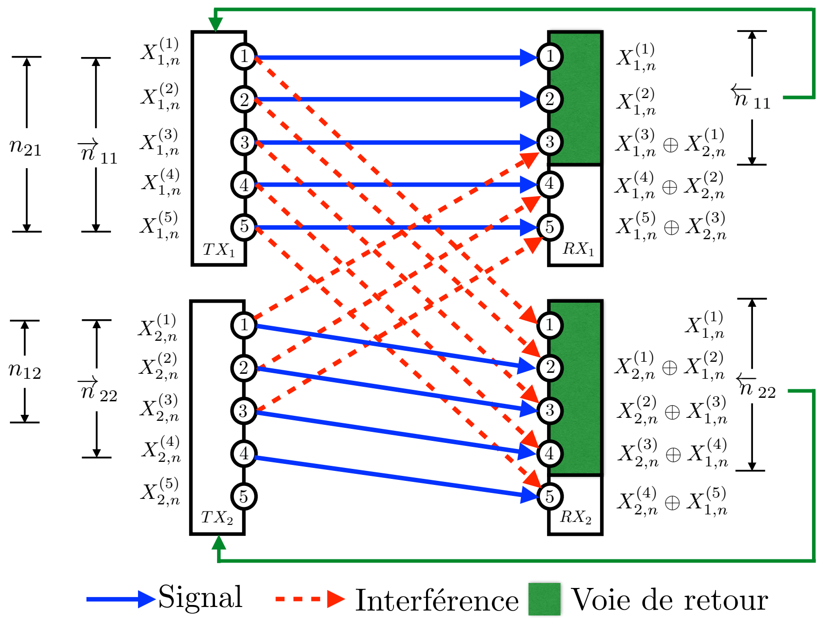


FIGURE 3. : Canal linéaire déterministe à interférence à deux utilisateurs avec voie de retour dégradée par un bruit additif à la n-ième utilisation du canal.

2. Canaux à Interférence Décentralisés

Dans un système décentralisé, il n'existe pas de contrôleur central, et chaque paire d'émetteur-récepteur est responsable de la sélection de sa propre configuration d'émission-réception pour maximiser son débit de transmission de données. La configuration de transmission-réception, notée s_i, avec $i \in \{1, 2\}$, peut être caractérisé par la longueur de bloc N_i, le nombre de bits par bloc M_i, l'alphabet d'entrée $\mathcal{X}_i$, le dictionnaire, les fonctions d'encodage $f_i^{(1)}$, $f_i^{(2)}, \ldots, f_i^{(N_i)}$, les fonctions de decodage $\psi_i^{(N)}$, etc. Le but de l'émetteur i est de choisir en toute autonomie sa configuration de transmission-réception, pour maximiser son taux de transmission atteignable R_i. Il convient de noter que le taux de transmission atteignable par l'émetteur-récepteur i dépend des configurations s_1 et s_2 en raison de l'interférence mutuelle. Ceci est révélateur de l'interaction compétitive entre les deux liens de communication du canal à interférence décentralisé.

Les modèles de systèmes pour le IC-NOF continu décentralisé à deux utilisateurs, le D-GIC-NOF à deux utilisateurs et le D-LDIC-NOF à deux utilisateurs sont en général les mêmes que pour le cas centralisé. Les principales différences sont les suivantes :

- Chaque émetteur-récepteur définit le nombre d'utilisations du canal par bloc, à savoir les utilisations du canal N_1 et N_2.
- La transmission d'un bloc est constituée de N utilisations du canal, où $N = \max(N_1, N_2)$. Alors, $X_{i,n} = 0$ pour tout $n > N_i$.
- L'encodeur i génère le symbole $x_{i,n}$ en tenant compte non seulement de l'indice de message $W_i \in \mathcal{W}_i = \{1, 2, \ldots, 2^{N_i R_i}\}$ et de toutes les sorties précédentes de la voie de

retour i, à savoir, $\left(\overleftarrow{y}_{i,1}, \overleftarrow{y}_{i,2}, \ldots, \overleftarrow{y}_{i,n-1}\right)$, mais aussi de l'indice de message aléatoire $\Omega_i \in \mathbb{N}$. L'indice Ω_i est un indice supplémentaire généré de façon aléatoire supposé être connu à l'émetteur i et au récepteur i, mais inconnu à l'émetteur j et au récepteur j (caractère aléatoire commun).

- À la fin de la transmission, le décodeur i utilise toutes les sorties de canal, à savoir, $\left(\overrightarrow{y}_{i,1}, \overrightarrow{y}_{i,2}, \ldots, \overrightarrow{y}_{i,N}\right)$ et l'indice de message aléatoire Ω_i pour estimer l'indice de message W_i, indiquée par $\widehat{W}_i$.

- La chaîne de Markov suivante existe :

$$\left(W_i, \Omega_i, \overleftarrow{\boldsymbol{Y}}_{i,(1;n-1)}\right) \to X_{i,n} \to \overrightarrow{Y}_{i,n}. \tag{11}$$

- Le calcul de la probabilité d'erreur est effectué pour chacune des paires émetteur-récepteur. Soit $W_i^{(t)}$ indiqué comme $c_{i,1}^{(t)}\, c_{i,2}^{(t)} \ldots c_{i,M_i}^{(t)}$ sous forme binaire. Soit aussi $\widehat{W}_i^{(t)}$ indiqué comme $\widehat{c}_{i,1}^{(t)}\, \widehat{c}_{i,2}^{(t)} \ldots \widehat{c}_{i,M_i}^{(t)}$ sous forme binaire. La probabilité d'erreur moyenne de bit au niveau du décodeur i compte tenu des configurations s_1 et s_2, notée $p_i(s_1, s_2)$, est alors donnée par

$$p_i(s_1, s_2)=\frac{1}{M_i}\sum_{\ell=1}^{M_i} \mathbb{1}_{\left\{\widehat{c}_{i,\ell}^{(t)} \neq c_{i,\ell}^{(t)}\right\}}. \tag{12}$$

Les limites fondamentales dans un système IC-NOF décentralisé à deux utilisateurs sont définies par la région d'η-équilibre de Nash $\mathcal{N}_\eta$.

Définition iii (Région d'η-Équilibre de Nash d'un IC à Deux Utilisateurs)**.** *La région d'η-équilibre de Nash $\mathcal{N}_\eta$ d'un IC à deux utilisateurs est la fermeture de l'ensemble de toutes les paires possibles de débit atteignables $(R_1, R_2) \in \mathbb{R}_+^2$ qui sont stables dans le sens d'un η-équilibre de Nash. Plus précisément, étant donné un schéma de codage atteignant l'équilibre de Nash, il n'existe pas de schéma de codage alternatif pour une ou l'autre des paires émetteur-récepteur augmentant leurs débits individuels de plus de η bits par utilisation du canal.*

2.1. Formulation du Jeu

L'intéraction compétitive des deux paires émetteur-récepteur dans le canal à interférence décentralisé peut être modélisée par un jeu sous forme normale suivant :

$$\mathcal{G} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}\right). \tag{13}$$

L'ensemble $\mathcal{K} = \{1, 2\}$ est l'ensemble des joueurs, c'est-à-dire, l'ensemble des paires émetteur-récepteur. Les ensembles $\mathcal{A}_1$ et $\mathcal{A}_2$ sont les ensembles d'actions des joueurs 1 et 2, respectivement. Une action du joueur $i \in \mathcal{K}$, notée par $s_i \in \mathcal{A}_i$, correspond à sa configuration de transmission-réception comme indiqué précédemment. La fonction d'utilité du joueur i est $u_i : \mathcal{A}_1 \times \mathcal{A}_2 \to \mathbb{R}_+$, et est définie comme le taux de transmission de l'émetteur i,

$$u_i(s_1, s_2) = \begin{cases} R_i = \frac{M_i}{N_i} & \text{if} \quad p_i(s_1, s_2) < \epsilon \\ 0 & \text{otherwise} \end{cases}, \tag{14}$$

où $\epsilon > 0$ est un nombre arbitrairement petit. Ce jeu sous forme normale fut proposé en premier lieu par [103] et [15].

Une classe de configurations de transmission-réception $\boldsymbol{s}^* = (s_1^*, s_2^*) \in \mathcal{A}_1 \times \mathcal{A}_2$, particulièrement importantes dans l'analyse de ce jeu est désignée comme l'ensemble d'η-équilibre de Nash. Cette classe de configurations satisfait la définition suivante :

Définition iv (η-équilibre de Nash)**.** *Dans le jeu* $\mathcal{G} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}\right)$*, un profil d'actions* (s_1^*, s_2^*) *est un* η*-équilibre de Nash si pour tout* $i \in \mathcal{K}$ *et pour tout* $s_i \in \mathcal{A}_i$*, il existe un* $\eta > 0$ *tels que*

$$u_i(s_i, s_j^*) \leqslant u_i(s_i^*, s_j^*) + \eta. \tag{15}$$

Soit (s_1^*, s_2^*) un profil d'actions d'un η-équilibre de Nash. Il s'en suit qu'aucun des émetteurs ne peut augmenter son taux de transmission de plus de η bits par utilisation du canal en changeant sa configuration de transmission-réception tout en conservant la probabilité d'erreur binaire moyenne arbitrairement proche de zéro. Il convient de noter que si η est suffisamment élevé, il suit de la Définition iv, que toute paire de configurations peut être un η-équilibre de Nash. D'autre part, pour $\eta = 0$, la définition classique de l'équilibre de Nash est obtenue [59]. Dans ce cas, si une paire de configurations est un équilibre de Nash ($\eta = 0$), alors chaque configuration individuelle est optimale étant donnée à la configuration de l'autre paire émetteur-récepteur. Par conséquent, l'objet est de décrire l'ensemble de toutes les paires de taux de transmission (R_1, R_2) η-équilibre de Nash du jeu définit par (13) avec la plus petite valeur possible de η pour lequel il existe au moins une configuration d'équilibre.

L'ensemble des paires de taux de transmission atteignables dans un η-équilibre de Nash est connu comme la région d'η-équilibre de Nash.

Définition v (La Région d'η-Équilibre de Nash)**.** *Soit* $\eta > 0$ *fixé. Une paire de taux de transmission atteignables* (R_1, R_2) *appartient à la région d'*η*-équilibre de Nash du jeu* $\mathcal{G} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}\right)$ *s'il existe une paire* $(s_1^*, s_2^*) \in \mathcal{A}_1 \times \mathcal{A}_2$ *qui est telle que*

$$u_1(s_1^*, s_2^*) = R_1 \quad \textit{et} \quad u_2(s_1^*, s_2^*) = R_2. \tag{16}$$

3. Connexions entre Canaux Linéaires Déterministes à Interférence et Canaux Gaussiens à Interférence

La région de capacité du GIC-NOF à deux utilisateurs avec les paramètres $\overrightarrow{\mathrm{SNR}}_1$, $\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, INR_{21}, $\overleftarrow{\mathrm{SNR}}_1$ et $\overleftarrow{\mathrm{SNR}}_2$ peut-être décrite approximativement par la région de capacité d'un LDIC-NOF avec des paramètres $\overrightarrow{n}_{ii} = \left\lfloor \frac{1}{2}\log(\overrightarrow{\mathrm{SNR}}_i) \right\rfloor$; $n_{ij} = \left\lfloor \frac{1}{2}\log(\mathrm{INR}_{ij}) \right\rfloor$; $\overleftarrow{n}_{ii} = \left\lfloor \frac{1}{2}\log(\overleftarrow{\mathrm{SNR}}_i) \right\rfloor$, avec $i \in \{1, 2\}$ et $j \in \{1, 2\} \setminus \{i\}$.

4. Principaux Résultats du Canal Linéaire Déterministe à Interférence Centralisé

Cette section présente les principaux résultats sur le LDIC-NOF centralisé à deux utilisateurs décrit en Section 1.2.

4.1. Region de Capacité

Notons $\mathcal{C}(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ la région de capacité du LDIC-NOF à deux utilisateurs avec les paramètres $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$, et $\overleftarrow{n}_{22}$, caractérisée par le Théorème i.

Théorème i. Region de capacité

La région de capacité $\mathcal{C}(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ du LDIC-NOF à deux utilisateurs est l'ensemble des paires de débit $(R_1, R_2) \in \mathbb{R}_+^2$ qui pour tout $i \in \{1,2\}$, avec $j \in \{1,2\} \setminus \{i\}$:

$$R_i \leqslant \min\left(\max\left(\overrightarrow{n}_{ii}, n_{ji}\right), \max\left(\overrightarrow{n}_{ii}, n_{ij}\right)\right), \tag{17a}$$

$$R_i \leqslant \min\left(\max\left(\overrightarrow{n}_{ii}, n_{ji}\right), \max\left(\overrightarrow{n}_{ii}, \overleftarrow{n}_{jj} - \left(\overrightarrow{n}_{jj} - n_{ji}\right)^+\right)\right), \tag{17b}$$

$$\begin{aligned} R_1 + R_2 \leqslant & \min\left(\max\left(\overrightarrow{n}_{22}, n_{12}\right) + \left(\overrightarrow{n}_{11} - n_{12}\right)^+, \max\left(\overrightarrow{n}_{11}, n_{21}\right) + \left(\overrightarrow{n}_{22} - n_{21}\right)^+\right), \\ R_1 + R_2 \leqslant & \max\left(\left(\overrightarrow{n}_{11} - n_{12}\right)^+, n_{21}, \overrightarrow{n}_{11} - \left(\max\left(\overrightarrow{n}_{11}, n_{12}\right) - \overleftarrow{n}_{11}\right)^+\right) \\ & + \max\left(\left(\overrightarrow{n}_{22} - n_{21}\right)^+, n_{12}, \overrightarrow{n}_{22} - \left(\max\left(\overrightarrow{n}_{22}, n_{21}\right) - \overleftarrow{n}_{22}\right)^+\right), \end{aligned} \tag{17c}$$

$$\begin{aligned} 2R_i + R_j \leqslant & \max\left(\overrightarrow{n}_{ii}, n_{ji}\right) + \left(\overrightarrow{n}_{ii} - n_{ij}\right)^+ \\ & + \max\left(\left(\overrightarrow{n}_{jj} - n_{ji}\right)^+, n_{ij}, \overrightarrow{n}_{jj} - \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+\right). \end{aligned} \tag{17d}$$

4.2. Cas où la Voie de Retour Agrandit la Région de Capacité

Soit $\alpha_i \in \mathbb{Q}$, avec $i \in \{1,2\}$ et $j \in \{1,2\} \setminus \{i\}$ défini comme

$$\alpha_i = \frac{n_{ij}}{\overrightarrow{n}_{ii}}. \tag{18}$$

Pour chaque paire d'émetteur-récepteur i, il existe cinq régimes d'interférence (IR) possibles, comme suggéré en [31] : le IR très faible (VWIR), à savoir, $\alpha_i \leqslant \frac{1}{2}$, le IR faible (WIR), à savoir, $\frac{1}{2} < \alpha_i \leqslant \frac{2}{3}$, le IR modéré (MIR), à savoir, $\frac{2}{3} < \alpha_i < 1$, le IR fort (SIR), à savoir, $1 \leqslant \alpha_i \leqslant 2$ et le IR très fort (VSIR), à savoir, $\alpha_i > 2$. Les scénarios dans lesquels le signal souhaité est plus fort que l'interférence ($\alpha_i < 1$), notamment les VWIR, WIR, et MIR, sont cités comme les régimes à faible interférence (LIR). A l'inverse, les scénarios dans lesquels le signal souhaité est plus faible ou égal à l'interférence ($\alpha_i \geqslant 1$), à savoir le SIR et le VSIR, sont considérés comme des régimes à hautes interférence (HIR).

Les résultats de cette section sont présentés au moyen d'un ensemble d'événements (variables booléennes) déterminés par les paramètres $\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}$, et n_{21}. Étant donné un quadruplet déterminé ($\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}), les événements sont définis ci-dessous :

$$E_1 : \quad \alpha_1 < 1 \wedge \alpha_2 < 1, \tag{19}$$

$$E_{2,i} : \quad \alpha_i \leqslant \frac{1}{2} \wedge 1 \leqslant \alpha_j \leqslant 2, \tag{20}$$

$$E_{3,i} : \quad \alpha_i \leqslant \frac{1}{2} \wedge \alpha_j > 2, \tag{21}$$

$$E_{4,i} : \quad \frac{1}{2} < \alpha_i \leqslant \frac{2}{3} \wedge \alpha_j \geqslant 1, \tag{22}$$

$$
\begin{aligned}
E_{5,i}: &\quad \frac{2}{3} < \alpha_i < 1 \wedge \alpha_j \geqslant 1, & (23)\\
E_{6,i}: &\quad \frac{1}{2} < \alpha_i \leqslant 1 \wedge \alpha_j > 1, & (24)\\
E_{7,i}: &\quad \alpha_i \geqslant 1 \wedge \alpha_j \leqslant 1, & (25)\\
E_{8,i}: &\quad \overrightarrow{n}_{ii} > n_{ji}, & (26)\\
E_9: &\quad \overrightarrow{n}_{11} + \overrightarrow{n}_{22} > n_{12} + n_{21}, & (27)\\
E_{10,i}: &\quad \overrightarrow{n}_{ii} + \overrightarrow{n}_{jj} > n_{ij} + 2n_{ji}, & (28)\\
E_{11,i}: &\quad \overrightarrow{n}_{ii} + \overrightarrow{n}_{jj} < n_{ij}. & (29)
\end{aligned}
$$

Dans le cas de $E_{8,i} : \overrightarrow{n}_{ii} > n_{ji}$, la notation $\widetilde{E}_{8,i}$ indique $\overrightarrow{n}_{ii} < n_{ji}$; la notation $\overline{E}_{8,i}$ indique $\overrightarrow{n}_{ii} \leqslant n_{ji}$ (complément logique) ; et la notation $\check{E}_{8,i}$ indique $\overrightarrow{n}_{ii} \geqslant n_{ji}$. Dans le cas $E_1 : \alpha_1 < 1 \wedge \alpha_2 < 1$, la notation $\widetilde{E}_1$ indique $\alpha_1 > 1 \wedge \alpha_2 > 1$; et la notation $\overline{E}_1$ indique $\alpha_1 \geqslant 1 \wedge \alpha_2 \geqslant 1$. Dans le cas $E_9 : \overrightarrow{n}_{11} + \overrightarrow{n}_{22} > n_{12} + n_{21}$, la notation $\overline{E}_9$ indique $\overrightarrow{n}_{11} + \overrightarrow{n}_{22} \leqslant n_{12} + n_{21}$.

En combinant les événements (19)-(29), cinq scénarios principaux sont identifiés :

$$
\begin{aligned}
S_{1,i}: &(E_1 \wedge E_{8,i}) \vee (E_{2,i} \wedge E_{8,i}) \vee (E_{3,i} \wedge E_{8,i} \wedge E_9) \vee (E_{4,i} \wedge E_{8,i} \wedge E_9) \vee (E_{5,i} \wedge E_{8,i} \wedge E_9), & (30)\\
S_{2,i}: &\left(E_{3,i} \wedge \widetilde{E}_{8,j} \wedge \overline{E}_9\right) \vee \left(E_{6,i} \wedge \widetilde{E}_{8,j} \wedge \overline{E}_9\right) \vee \left(\widetilde{E}_1 \wedge \widetilde{E}_{8,j}\right), & (31)\\
S_{3,i}: &\left(E_1 \wedge \overline{E}_{8,i}\right) \vee \left(E_{2,i} \wedge \overline{E}_{8,i}\right) \vee \left(E_{3,i} \wedge \check{E}_{8,j} \wedge \overline{E}_{8,i}\right) \vee \left(E_{4,i} \wedge \check{E}_{8,j} \wedge \overline{E}_{8,i}\right)\\
&\vee \left(E_{5,i} \wedge \check{E}_{8,j} \wedge \overline{E}_{8,i}\right) \vee \left(\overline{E}_1 \wedge \check{E}_{8,j}\right) \vee (E_{7,i}), & (32)\\
S_4 \ : &E_1 \wedge E_{8,1} \wedge E_{8,2} \wedge E_{10,1} \wedge E_{10,2}, & (33)\\
S_5 \ : &\overline{E}_1 \wedge E_{11,1} \wedge E_{11,2}. & (34)
\end{aligned}
$$

Pour tout $i \in \{1, 2\}$, les événements $S_{1,i}$, $S_{2,i}$, $S_{3,i}$, S_4 et S_5 montrent les propriétés énumérées par les corollaires suivants.

Corollaire i. Pour tout $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$, étant donné un $i \in \{1, 2\}$ déterminé, seul un des événements $S_{1,i}$, $S_{2,i}$ et $S_{3,i}$ est vrai.

Corollaire ii. Pour tout $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$, lorsque l'un des événements S_4 ou S_5 est vrai, l'autre est nécessairement faux.

Notez que le Corollaire ii n'exclut pas le cas dans lequel les deux S_4 et S_5 sont simultanément faux.

Corollaire iii. Pour tout $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$, si S_4 est vrai, alors $S_{1,1}$ et $S_{1,2}$ sont tous deux vrais ; et si S_5 est vrai, alors $S_{2,1}$ et $S_{2,2}$ sont tous deux vrais.

Métrique de l'Amélioration de Débit

Étant donné un quadruplet $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21})$, soit $\mathcal{C}(\overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ la région de capacité d'un LDIC-NOF avec les paramètres $\overleftarrow{n}_{11}$ et $\overleftarrow{n}_{22}$. L'amélioration maximum des débits individuels R_1 et R_2, notée respectivement par $\Delta_1(\overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ et $\Delta_2(\overleftarrow{n}_{11}, \overleftarrow{n}_{22})$, due à l'effet de la voie de retour est :

$$\Delta_1(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) = \max_{0<R_2<R_2^*} \left\{ \sup \left\{ R_1 : (R_1, R_2) \in \mathcal{C}(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \right\} - \sup \left\{ R_1^\dagger : (R_1^\dagger, R_2) \in \mathcal{C}(0,0) \right\} \right\}$$

et (35)

$$\Delta_2(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) = \max_{0<R_1<R_1^*} \left\{ \sup \left\{ R_2 : (R_1, R_2) \in \mathcal{C}(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \right\} - \sup \left\{ R_2^\dagger : (R_1, R_2^\dagger) \in \mathcal{C}(0,0) \right\} \right\}, \quad (36)$$

avec

$$R_1^* = \sup \left\{ r_1 : (r_1, r_2) \in \mathcal{C}(0,0) \right\} \text{ et} \quad (37)$$

$$R_2^* = \sup \left\{ r_2 : (r_1, r_2) \in \mathcal{C}(0,0) \right\}. \quad (38)$$

Notons que pour tout $i \in \{1, 2\}$, $\Delta_i(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) > 0$ si et seulement s'il est possible d'atteindre une paire de débits $(R_1, R_2) \in \mathbb{R}_+^2$ avec une voie de retour telle que R_i est supérieur au débit maximum atteignable par l'émetteur-récepteur i sans voie de retour lorsque le débit de la paire émetteur-récepteur j est fixé à R_j. Étant donné les paramètres déterminés $\overleftarrow{n}_{11}$ et $\overleftarrow{n}_{22}$, l'assertion *"le débit R_i est amélioré à l'aide de la voie de retour"* sert à indiquer que $\Delta_i(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) > 0$.

L'amélioration maximum du débit somme $\Sigma(\overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ concernant le cas sans voie de retour est :

$$\Sigma(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) = \sup \left\{ R_1 + R_2 : (R_1, R_2) \in \mathcal{C}(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \right\} - \sup \left\{ R_1^\dagger + R_2^\dagger : (R_1^\dagger, R_2^\dagger) \in \mathcal{C}(0,0) \right\}. \quad (39)$$

Notons que $\Sigma(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) > 0$ si et seulement s'il existe une paire de débits avec voie de retour dont la somme est supérieure au débit somme maximum atteignable sans voie de retour. Étant donné les paramètres déterminés $\overleftarrow{n}_{11}$ et $\overleftarrow{n}_{22}$, l'assertion *"le débit somme est amélioré à l'aide de la voie de retour"* sert à impliquer que $\Sigma(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) > 0$.

Lorsque la voie de retour est exclusivement utilisée par la paire d'émetteur-récepteur i, à savoir, $\overleftarrow{n}_{ii} > 0$ et $\overleftarrow{n}_{jj} = 0$, l'amélioration maximum du débit individuel de l'émetteur-récepteur k, avec $k \in \{1, 2\}$, et l'amélioration maximum du débit somme sont indiquées par $\Delta_k(\overleftarrow{n}_{ii})$ et $\Sigma(\overleftarrow{n}_{ii})$, respectivement. Par conséquent, cette notation $\Delta_k(\overleftarrow{n}_{ii})$ remplace soit $\Delta_k(\overleftarrow{n}_{11}, 0)$ ou $\Delta_k(0, \overleftarrow{n}_{22})$, lorsque $i = 1$ ou $i = 2$, respectivement. Il en va de même pour la notation $\Sigma(\overleftarrow{n}_{ii})$ qui remplace $\Sigma(\overleftarrow{n}_{11}, 0)$ ou $\Sigma(0, \overleftarrow{n}_{22})$, lorsque $i = 1$ ou $i = 2$, respectivement.

Agrandissement de la Région de Capacité

Compte tenu des paramètres fixés $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21})$, $i \in \{1, 2\}$, et $j \in \{1, 2\} \setminus \{i\}$, la région de capacité d'un LDIC-NOF à deux utilisateurs, lorsque la voie de retour est disponible uniquement au niveau de la paire émetteur-récepteur i, à savoir, $\overleftarrow{n}_{ii} > 0$ et $\overleftarrow{n}_{jj} = 0$, est noté $\mathcal{C}(\overleftarrow{n}_{ii})$ au lieu de $\mathcal{C}(\overleftarrow{n}_{11}, 0)$ ou $\mathcal{C}(0, \overleftarrow{n}_{22})$, quand $i = 1$ ou $i = 2$, respectivement. À la suite de cette notation, le Théorème ii identifie les valeurs exactes de $\overleftarrow{n}_{ii}$ pour lesquelles l'inclusion stricte de $\mathcal{C}(0,0) \subset \mathcal{C}(\overleftarrow{n}_{ii})$ vaut pour tout $i \in \{1, 2\}$.

Théorème ii. Agrandissement de la Région de Capacité

Soit $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$ un quadruplet fixe. Soit aussi $i \in \{1,2\}$, $j \in \{1,2\} \setminus \{i\}$ et $\overleftarrow{n}_{ii}^{*} \in \mathbb{N}$ des nombres entiers fixes, avec

$$\overleftarrow{n}_{ii}^{*} = \begin{cases} \max\left(n_{ji}, (\overrightarrow{n}_{ii} - n_{ij})^{+}\right) & \text{si } S_{1,i} \text{ est vrai} \\ \overrightarrow{n}_{jj} + (\overrightarrow{n}_{ii} - n_{ij})^{+} & \text{si } S_{2,i} \text{ est vrai} \end{cases}. \quad (40)$$

Supposons que $S_{3,i}$ est vrai. Alors, pour tout $\overleftarrow{n}_{ii} \in \mathbb{N}$, $\mathcal{C}\left(0,0\right) = \mathcal{C}\left(\overleftarrow{n}_{ii}\right)$. Supposons aussi que $S_{1,i}$ est vrai ou $S_{2,i}$ est vrai. Alors, pour tout $\overleftarrow{n}_{ii} \leqslant \overleftarrow{n}_{ii}^{*}$, $\mathcal{C}\left(0,0\right) = \mathcal{C}\left(\overleftarrow{n}_{ii}\right)$ et pour tout $\overleftarrow{n}_{ii} > \overleftarrow{n}_{ii}^{*}$, $\mathcal{C}\left(0,0\right) \subset \mathcal{C}\left(\overleftarrow{n}_{ii}\right)$.

Le Théorème ii indique qu'étant donné l'événement $S_{3,i}$ en (32), l'implémentation de la voie de retour à la paire émetteur-récepteur i, avec tout $\overleftarrow{n}_{ii} > 0$ et $\overleftarrow{n}_{jj} = 0$, n'agrandit pas la région de capacité. Étant donné les événements $S_{1,i}$ en (30) et $S_{2,i}$ en (31), la région de capacité peut être agrandie lorsque $\overleftarrow{n}_{ii} > \overleftarrow{n}_{ii}^{*}$. Il importe de souligner que dans les cas dans lesquels la voie de retour agrandit la région de capacité du LDIC-NOF à deux utilisateurs, étant donné les événements $S_{1,1}$, $S_{2,1}$, $S_{1,2}$ ou $S_{2,2}$, pour tout $i \in \{1,2\}$ et $j \in \{1,2\} \setminus \{i\}$, ce qui suit est toujours vrai :

$$\overleftarrow{n}_{ii}^{*} > (\overrightarrow{n}_{ii} - n_{ij})^{+}. \quad (41)$$

En effet, l'inégalité (41) dévoile une condition nécessaire mais non suffisante pour agrandir la région de capacité à l'aide de la voie de retour. Cette condition indique que pour au moins un $i \in \{1,2\}$, avec $j \in \{1,2\} \setminus \{i\}$, l'émetteur i puisse décoder un sous-ensemble de bits d'information envoyés par l'émetteur j à chaque utilisation du canal.

Il est également intéressant d'observer que, le seuil $\overleftarrow{n}_{ii}^{*}$ au-delà duquel la voie de retour est utile est différent étant donné l'événement $S_{1,i}$ en (30) et l'événement $S_{2,i}$ en (31). En général, lorsque $S_{1,i}$ est vrai, l'agrandissement de la région de capacité est due au fait que la voie de retour permet d'utiliser *l'interférence comme information secondaire* [86]. Si $S_{2,i}$ en (31) est vrai, l'agrandissement de la région de capacité survient est comme une conséquence du fait que certains bits ne pouvant être transmis directement de l'émetteur j au récepteur j, peuvent parvenir au récepteur j par un chemin alternatif : émetteur j - récepteur i - émetteur i - récepteur j.

Amélioration du Débit Individuel R_i à l'Aide de la Voie de Retour dans le Lien i

Étant donné les paramètres fixés $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21})$, et $i \in \{1,2\}$, l'implémentation de la voie de retour à la paire émetteur-récepteur i augmente le débit individuel R_i, à savoir, $\Delta_i(\overleftarrow{n}_{ii}) > 0$ pour certaines valeurs de $\overleftarrow{n}_{ii}$. Le Théorème iii identifie les valeurs exactes de $\overleftarrow{n}_{ii}$ pour lesquelles $\Delta_i(\overleftarrow{n}_{ii}) > 0$.

Théorème iii. Amélioration de R_i à l'Aide de la Voie de Retour dans le Lien i

Soit $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$ un quadruplet fixe. Soit aussi $i \in \{1, 2\}$, $j \in \{1, 2\} \setminus \{i\}$ et $\overleftarrow{n}_{ii}^{\dagger} \in \mathbb{N}$ des nombres entiers fixes, avec

$$\overleftarrow{n}_{ii}^{\dagger} = \max\left(n_{ji}, (\overrightarrow{n}_{ii} - n_{ij})^+\right). \quad (42)$$

Supposons que $S_{2,i}$ est vrai ou $S_{3,i}$ est vrai. Alors, pour tout $\overleftarrow{n}_{ii} \in \mathbb{N}$, $\Delta_i(\overleftarrow{n}_{ii}) = 0$. Supposons que $S_{1,i}$ est vrai. Alors, si $\overleftarrow{n}_{ii} \leqslant \overleftarrow{n}_{ii}^{\dagger}$, cela signifie que $\Delta_i(\overleftarrow{n}_{ii}) = 0$; et sinon $\overleftarrow{n}_{ii} > \overleftarrow{n}_{ii}^{\dagger}$, cela signifie que $\Delta_i(\overleftarrow{n}_{ii}) > 0$.

Le Théorème iii souligne qu'étant donné les événements $S_{2,i}$ en (31) et $S_{3,i}$ en (32), le débit individuel R_i ne peut être amélioré à l'aide de la voie de retour à la paire émetteur-récepteur i, à savoir, $\Delta_i(\overleftarrow{n}_{ii}) = 0$. Étant donné l'événement $S_{1,i}$ en (30), le débit individuel R_i peut être amélioré, à savoir, $\Delta_i\left(\overleftarrow{n}_{ii}\right) > 0$, si $\overleftarrow{n}_{ii} > \max\left(n_{ji}, (\overrightarrow{n}_{ii} - n_{ij})^+\right)$.

Amélioration du Débit Individuel R_j à l'Aide de la Voie de Retour dans le Lien i

Étant donné les paramètres fixés $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21})$, $i \in \{1, 2\}$, et $j \in \{1, 2\} \setminus \{i\}$, l'implémentation de la voie de retour à la paire émetteur-récepteur i augmente le débit individuel R_j, à savoir, $\Delta_j(\overleftarrow{n}_{ii}) > 0$ pour certaines valeurs de $\overleftarrow{n}_{ii}$. Le Théorème iv identifie les valeurs exactes de $\overleftarrow{n}_{ii}$ pour lesquelles $\Delta_j(\overleftarrow{n}_{ii}) > 0$.

Théorème iv. Amélioration de R_j à l'Aide de la Voie de Retour dans le Lien i

Soit $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$ un quadruplet fixe. Soit aussi $i \in \{1, 2\}$, $j \in \{1, 2\} \setminus \{i\}$ et $\overleftarrow{n}_{ii}^{*} \in \mathbb{N}$ en (40), des nombres entiers fixes. Supposons que $S_{3,i}$ est vrai. Alors, pour tout $\overleftarrow{n}_{ii} \in \mathbb{N}$, $\Delta_j(\overleftarrow{n}_{ii}) = 0$. Supposons que $S_{1,i}$ est vrai ou $S_{2,i}$ est vrai. Alors, si $\overleftarrow{n}_{ii} \leqslant \overleftarrow{n}_{ii}^{*}$, cela signifie que $\Delta_j(\overleftarrow{n}_{ii}) = 0$; et sinon $\overleftarrow{n}_{ii} > \overleftarrow{n}_{ii}^{*}$, cela signifie que $\Delta_j(\overleftarrow{n}_{ii}) > 0$.

Le Théorème iv indique qu'étant donné l'événement $S_{3,i}$ en (32), l'implémentation de la voie de retour à la paire émetteur-récepteur i n'apporte aucune amélioration au débit R_j. Étant donné les événements $S_{1,i}$ en (30) et $S_{2,i}$ en (31), le débit individuel R_j peut être amélioré, à savoir, $\Delta_j(\overleftarrow{n}_{ii}) > 0$ pour tout $\overleftarrow{n}_{ii} > \overleftarrow{n}_{ii}^{*}$.

Amélioration du Débit Somme

Étant donné les paramètres fixés $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21})$, et $i \in \{1, 2\}$, l'implémentation de la voie de retour à la paire émetteur-récepteur i augmente le débit somme, à savoir, $\Sigma(\overleftarrow{n}_{ii}) > 0$ pour certaines valeurs de $\overleftarrow{n}_{ii}$. Le Théorème v identifie les valeurs exactes de $\overleftarrow{n}_{ii}$ pour lesquelles $\Sigma(\overleftarrow{n}_{ii}) > 0$.

Théorème v. Amélioration de la Somme des Capacités

Soit $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$ un quadruplet fixe. Soit aussi $i \in \{1, 2\}$, $j \in \{1, 2\} \setminus \{i\}$ et $\overleftarrow{n}_{ii}^{+} \in \mathbb{N}$ des nombres entiers fixes, avec

$$\overleftarrow{n}_{ii}^{+} = \begin{cases} \max\left(n_{ji}, (\overrightarrow{n}_{ii} - n_{ij})^{+}\right) & \text{si } S_4 \text{ est vrai} \\ \overrightarrow{n}_{jj} + (\overrightarrow{n}_{ii} - n_{ij})^{+} & \text{si } S_5 \text{ est vrai} \end{cases}. \tag{43}$$

Supposons que S_4 et S_5 soient faux. Alors, $\Sigma(\overleftarrow{n}_{ii}) = 0$ pour tout $\overleftarrow{n}_{ii} \in \mathbb{N}$. Supposons que S_4 est vrai ou S_5 est vrai. Alors, si $\overleftarrow{n}_{ii} \leqslant \overleftarrow{n}_{ii}^{+}$, cela signifie que $\Sigma(\overleftarrow{n}_{ii}) = 0$; et sinon $\overleftarrow{n}_{ii} > \overleftarrow{n}_{ii}^{+}$, cela signifie que $\Sigma(\overleftarrow{n}_{ii}) > 0$.

Le Théorème v introduit une condition nécessaire mais non suffisante pour l'amélioration du débit somme par l'implémentation de la voie de retour à la paire émetteur-récepteur i.

Remarque 1 : *Une condition nécessaire mais non suffisante pour observer* $\Sigma(\overleftarrow{n}_{ii}) > 0$ *est de satisfaire à l'une des conditions suivantes :* (a) *les deux paires émetteur-transmetteur sont en LIR (Événement* E_1*) ; ou* (b) *les deux paires émetteur-transmetteur sont en HIR (Événement* $\overline{E}_1$*).*

Enfin, il résulte du Corollaire iii que si S_4 ou S_5 est vrai, avec $i \in \{1, 2\}$ et $\overleftarrow{n}_{ii} > \overleftarrow{n}_{ii}^{+}$, outre $\Sigma(\overleftarrow{n}_{ii}) > 0$, cela signifie aussi que $\Delta_1(\overleftarrow{n}_{ii}) > 0$ et $\Delta_2(\overleftarrow{n}_{ii}) > 0$.

4.3. Degrés de Liberté Généralisés

Cette section se concentre sur l'analyse du nombre de degrés de liberté généralisés (GDoF) du LDIC-NOF à deux utilisateurs pour étudier le cas dans lequel la voie de retour est implémentée simultanément dans les deux paires émetteur-récepteur. L'analyse n'est de plus réalisée que pour le cas symétrique, à savoir, $\overrightarrow{n} = \overrightarrow{n}_{11} = \overrightarrow{n}_{22}$, $m = n_{12} = n_{21}$, et $\overleftarrow{n} = \overleftarrow{n}_{11} = \overleftarrow{n}_{22}$, avec $(\overrightarrow{n}, m, \overleftarrow{n}) \in \mathbb{N}^3$. Les résultats du théorème i permettent une analyse plus générale du nombre de GDoF, à savoir, le cas non symétrique. Le cas symétrique donne toutefois certains des aperçus les plus importants concernant l'agrandissement de la région de capacité lorsque la voie de retour est utilisée dans les deux paires d'émetteurs-récepteurs.

Étant donné les paramètres $\overrightarrow{n}$, m et $\overleftarrow{n}$, avec $\alpha = \frac{m}{\overrightarrow{n}}$ et $\beta = \frac{\overleftarrow{n}}{\overrightarrow{n}}$, le nombre de GDoF, notée $D(\alpha, \beta)$, représente le rapport entre la capacité symétrique, à savoir, $C_{\text{sym}}(\overrightarrow{n}, m, \overleftarrow{n}) = \sup\{R : (R, R) \in \mathcal{C}(\overrightarrow{n}, \overrightarrow{n}, m, m, \overleftarrow{n}, \overleftarrow{n})\}$, et la capacité individuelle sans interférence, à savoir, $\overrightarrow{n}$, lorsque $(\overrightarrow{n}, m, \overleftarrow{n}) \to (\infty, \infty, \infty)$. Le nombre de GDoF est plus précisément :

$$D(\alpha, \beta) = \lim_{(\overrightarrow{n}, m, \overleftarrow{n}) \to (\infty, \infty, \infty)} \frac{C_{\text{sym}}(\overrightarrow{n}, m, \overleftarrow{n})}{\overrightarrow{n}}. \tag{44}$$

Le Théorème vi détermine le nombre de GDoF pour le LDIC-NOF symétrique à deux utilisateurs.

Théorème vi. Le Nombre de GDoF

Le nombre de GDoF pour le LDIC-NOF symétrique à deux utilisateurs avec paramètres

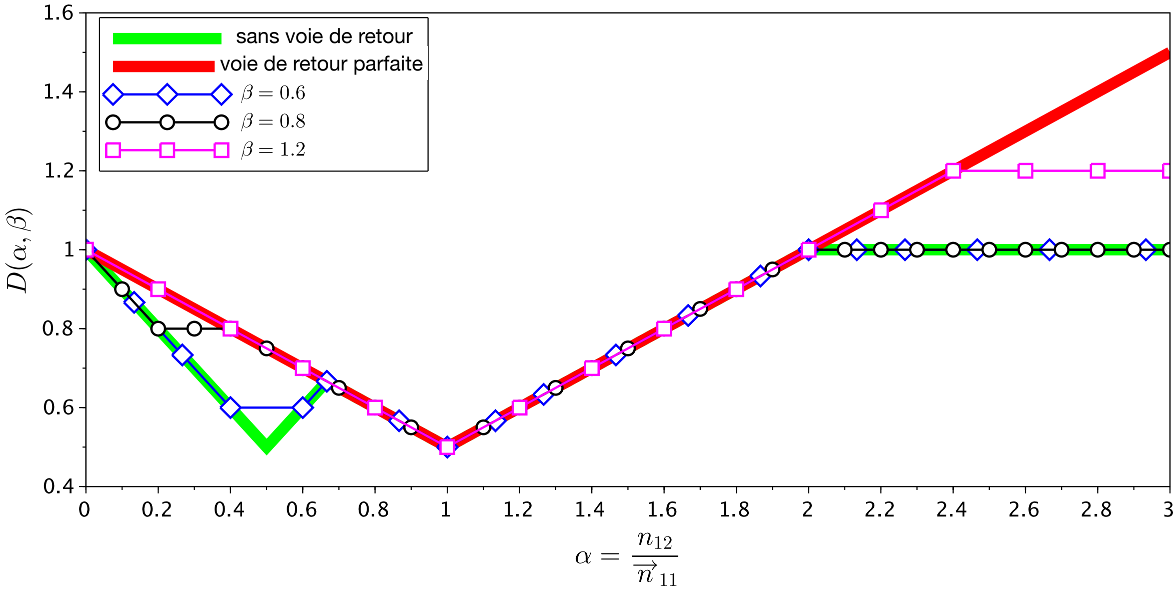


FIGURE 4. : Degrés de liberté généralisés comme fonction des paramètres α et β, avec $0 \leqslant \alpha \leqslant 3$ et $\beta \in \{\frac{3}{5}, \frac{4}{5}, \frac{6}{5}\}$, du LDIC-NOF symétrique à deux utilisateurs. Le point sans voie de retour est obtenu à partir de [31] et le point avec voie de retour parfaite est obtenu à partir de [88].

α et β est donné par

$$D(\alpha,\beta) = \min\Bigg(\max(1,\alpha), \max\left(1, \beta - (1-\alpha)^+\right), \frac{1}{2}\left(\max(1,\alpha) + (1-\alpha)^+\right), \max\left((1-\alpha)^+, \alpha, 1 - (\max(1,\alpha) - \beta)^+\right)\Bigg). \tag{45}$$

Le résultat du Théorème vi peut également être obtenu à partir du Théorème 1 en [53]. Les propriétés suivantes sont une conséquence directe du Théorème vi.

Corollaire iv. *Le nombre de GDoF pour le LDIC-NOF symétrique à deux utilisateurs avec paramètres α et β répond aux propriétés suivantes :*

$$\forall \alpha \in \left[0, \frac{2}{3}\right] \text{ et } \beta \leqslant 1, \qquad \max\left(\frac{1}{2}, \beta\right) \leqslant D(\alpha,\beta) \leqslant 1, \tag{46a}$$

$$\forall \alpha \in \left[0, \frac{2}{3}\right] \text{ et } \beta > 1, \qquad D(\alpha,\beta) = 1 - \frac{\alpha}{2}, \tag{46b}$$

$$\forall \alpha \in \left(\frac{2}{3}, 2\right] \text{ et } \beta \in [0,\infty), \qquad D(\alpha,0) = D(\alpha,\beta) = D(\alpha, \max(1,\alpha)), \tag{46c}$$

$$\forall \alpha \in (2, \infty) \ \ et\ \beta \geqslant 1, \quad 1 \leqslant D(\alpha, \beta) \leqslant \min\left(\frac{\alpha}{2}, \beta\right), \tag{46d}$$

$$\forall \alpha \in (2, \infty) \ \ et\ \beta < 1, \quad D(\alpha, \beta) = 1. \tag{46e}$$

Les propriétés (46a) et (46b) mettent en évidence le fait que l'existence de voies de retour dans le LDIC-NOF symétrique à deux utilisateurs dans le VWIR et WIR n'ont aucun impact sur le nombre de GDoF si $\beta \leqslant \frac{1}{2}$, et le nombre de GDoF est égal au cas avec une voie de retour parfaite si $\beta > 1$. La propriété (46c) souligne que dans le LDIC-NOF symétrique à deux utilisateurs en MIR et SIR, le nombre de GDoF est identique dans les deux cas extrêmes : sans voie de retour ($\beta = 0$) et avec voie de retour parfaite $\big(\beta = \max(1, \alpha)\big)$. Il résulte enfin qu'à partir de (46d) et (46e), pour observer une amélioration du nombre de GDoF du LDIC-NOF symétrique à deux utilisateurs en VSIR, la condition suivante doit être satisfaite : $\beta > 1$. Autrement dit, la capacité dans les voies de retour doit être supérieur a la capacité dans les liens directs.

La Figure 4 montre le nombre de GDoF pour le LDIC-NOF symétrique à deux utilisateurs pour le cas dans lequel $0 \leqslant \alpha \leqslant 3$ et $\beta \in \{\frac{3}{5}, \frac{4}{5}, \frac{6}{5}\}$.

5. Principaux Résultats du Canal Gaussien à Interférence Centralisé

Cette section présente les principaux résultats sur le GIC-NOF centralisé décrit en Section 1.1. Ceux-ci incluent une région atteignable (Théorème vii) et une région d'impossibilité (Théorème viii), notée $\underline{\mathcal{C}}$ et $\overline{\mathcal{C}}$ respectivement, pour le GIC-NOF à deux utilisateurs et paramètres fixés $\overrightarrow{\mathrm{SNR}}_1$, $\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, INR_{21}, $\overleftarrow{\mathrm{SNR}}_1$, et $\overleftarrow{\mathrm{SNR}}_2$. En général, la région de capacité d'un canal multi-utilisateurs donné est approximé avec un écart constant conformément à la Définition vi.

Définition vi (Approximation à ξ unités près)**.**
Un ensemble fermé et convexe $\mathcal{T} \subset \mathbb{R}_+^m$ est obtenu par approximation à ξ unités près par les ensembles $\underline{\mathcal{T}}$ et $\overline{\mathcal{T}}$ si $\underline{\mathcal{T}} \subseteq \mathcal{T} \subseteq \overline{\mathcal{T}}$ et pour tout $\boldsymbol{t} = (t_1, t_2, \ldots, t_m) \in \overline{\mathcal{T}}$, $\Big((t_1 - \xi)^+, (t_2 - \xi)^+, \ldots, (t_m - \xi)^+ \Big) \in \underline{\mathcal{T}}$.

Notons $\mathcal{C}$ la région de capacité du GIC-NOF à deux utilisateurs. La région atteignable $\underline{\mathcal{C}}$ et la région d'impossibilité $\overline{\mathcal{C}}$ correspondent approximativement à la région de capacité $\mathcal{C}$ à 4.4 bits près. (Théorème ix).

5.1. Une Région Atteignable

La description de la région atteignable $\underline{\mathcal{C}}$ est présentée à l'aide des constantes $a_{1,i}$; des fonctions $a_{2,i} : [0,1] \to \mathbb{R}_+$, $a_{l,i} : [0,1]^2 \to \mathbb{R}_+$, avec $l \in \{3, \ldots, 6\}$; et $a_{7,i} : [0,1]^3 \to \mathbb{R}_+$, qui sont définis comme suit, pour tout $i \in \{1, 2\}$, avec $j \in \{1, 2\} \setminus \{i\}$:

$$a_{1,i} = \frac{1}{2} \log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\right) - \frac{1}{2}, \tag{47a}$$

$$a_{2,i}(\rho) = \frac{1}{2} \log\Big(b_{1,i}(\rho) + 1\Big) - \frac{1}{2}, \tag{47b}$$

$$a_{3,i}(\rho,\mu)=\frac{1}{2}\log\left(\frac{\overleftarrow{\mathrm{SNR}}_i\Big(b_{2,i}(\rho)+2\Big)+b_{1,i}(1)+1}{\overleftarrow{\mathrm{SNR}}_i\Big((1-\mu)\,b_{2,i}(\rho)+2\Big)+b_{1,i}(1)+1}\right), \tag{47c}$$

$$a_{4,i}(\rho,\mu)=\frac{1}{2}\log\left(\Big(1-\mu\Big)b_{2,i}(\rho)+2\right)-\frac{1}{2}, \tag{47d}$$

$$a_{5,i}(\rho,\mu)=\frac{1}{2}\log\left(2+\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}+\Big(1-\mu\Big)b_{2,i}(\rho)\right)-\frac{1}{2}, \tag{47e}$$

$$a_{6,i}(\rho,\mu)=\frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\left(\Big(1-\mu\Big)b_{2,j}(\rho)+1\right)+2\right)-\frac{1}{2},\ \text{et} \tag{47f}$$

$$a_{7,i}(\rho,\mu_1,\mu_2)=\frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\left(\Big(1-\mu_i\Big)b_{2,j}(\rho)+1\right)+\Big(1-\mu_j\Big)b_{2,i}(\rho)+2\right)-\frac{1}{2}, \tag{47g}$$

où les fonctions $b_{l,i}:[0,1]\rightarrow\mathbb{R}_+$, avec $(l,i)\in\{1,2\}^2$ sont définies comme suit :

$$b_{1,i}(\rho)=\overrightarrow{\mathrm{SNR}}_i+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_i\mathrm{INR}_{ij}}+\mathrm{INR}_{ij}\ \text{et} \tag{48a}$$

$$b_{2,i}(\rho)=\Big(1-\rho\Big)\mathrm{INR}_{ij}-1, \tag{48b}$$

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

Notons que les fonctions dans (47) et (48) dépendent de $\overrightarrow{\mathrm{SNR}}_1$, $\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, INR_{21}, $\overleftarrow{\mathrm{SNR}}_1$, et $\overleftarrow{\mathrm{SNR}}_2$. Toutefois, comme ces paramètres sont fixés dans cette analyse, cette dépendance n'est pas mise en évidence dans la définition de ces fonctions. Enfin, en utilisant cette notation, le Théorème vii est présenté comme suit :

Théorème vii. Région Atteignable

La région de capacité $\mathcal{C}$ contient la région $\underline{\mathcal{C}}$ donnée par la fermeture de l'ensemble de toutes les paires possibles de débit atteignables $(R_1,R_2)\in\mathbb{R}_+^2$ satisfaisant :

$$R_1\leqslant\min\Big(a_{2,1}(\rho),a_{6,1}(\rho,\mu_1)+a_{3,2}(\rho,\mu_1),a_{1,1}+a_{3,2}(\rho,\mu_1)+a_{4,2}(\rho,\mu_1)\Big), \tag{49a}$$

$$R_2\leqslant\min\Big(a_{2,2}(\rho),a_{3,1}(\rho,\mu_2)+a_{6,2}(\rho,\mu_2),a_{3,1}(\rho,\mu_2)+a_{4,1}(\rho,\mu_2)+a_{1,2}\Big), \tag{49b}$$

$$\begin{aligned}R_1+R_2\leqslant\min\Big(&a_{2,1}(\rho)+a_{1,2},a_{1,1}+a_{2,2}(\rho),\\&a_{3,1}(\rho,\mu_2)+a_{1,1}+a_{3,2}(\rho,\mu_1)+a_{7,2}(\rho,\mu_1,\mu_2),\\&a_{3,1}(\rho,\mu_2)+a_{5,1}(\rho,\mu_2)+a_{3,2}(\rho,\mu_1)+a_{5,2}(\rho,\mu_1),\\&a_{3,1}(\rho,\mu_2)+a_{7,1}(\rho,\mu_1,\mu_2)+a_{3,2}(\rho,\mu_1)+a_{1,2}\Big),\end{aligned} \tag{49c}$$

$$\begin{aligned}2R_1+R_2\leqslant\min\Big(&a_{2,1}(\rho)+a_{1,1}+a_{3,2}(\rho,\mu_1)+a_{7,2}(\rho,\mu_1,\mu_2),\\&a_{3,1}(\rho,\mu_2)+a_{1,1}+a_{7,1}(\rho,\mu_1,\mu_2)+2a_{3,2}(\rho,\mu_1)+a_{5,2}(\rho,\mu_1),\\&a_{2,1}(\rho)+a_{1,1}+a_{3,2}(\rho,\mu_1)+a_{5,2}(\rho,\mu_1)\Big),\end{aligned} \tag{49d}$$

$$R_1 + 2R_2 \leqslant \min\Big(a_{3,1}(\rho,\mu_2) + a_{5,1}(\rho,\mu_2) + a_{2,2}(\rho) + a_{1,2},$$
$$a_{3,1}(\rho,\mu_2) + a_{7,1}(\rho,\mu_1,\mu_2) + a_{2,2}(\rho) + a_{1,2},$$
$$2a_{3,1}(\rho,\mu_2) + a_{5,1}(\rho,\mu_2) + a_{3,2}(\rho,\mu_1) + a_{1,2} + a_{7,2}(\rho,\mu_1,\mu_2)\Big), \tag{49e}$$

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

5.2. Une Région d'Impossibilité

La description de la région d'impossibilité $\overline{\mathcal{C}}$ est déterminée par deux événements indiqués par $S_{l_1,1}$ et $S_{l_2,2}$, où $(l_1,l_2) \in \{1,\ldots,5\}^2$. Les événements sont définis comme suit pour tout $i \in \{1,2\}$:

$$S_{1,i}: \quad \overrightarrow{\mathrm{SNR}}_j < \min\left(\mathrm{INR}_{ij}, \mathrm{INR}_{ji}\right), \tag{50a}$$
$$S_{2,i}: \quad \mathrm{INR}_{ji} \leqslant \overrightarrow{\mathrm{SNR}}_j < \mathrm{INR}_{ij}, \tag{50b}$$
$$S_{3,i}: \quad \mathrm{INR}_{ij} \leqslant \overrightarrow{\mathrm{SNR}}_j < \mathrm{INR}_{ji}, \tag{50c}$$
$$S_{4,i}: \quad \max\left(\mathrm{INR}_{ij}, \mathrm{INR}_{ji}\right) \leqslant \overrightarrow{\mathrm{SNR}}_j < \mathrm{INR}_{ij}\mathrm{INR}_{ji}, \tag{50d}$$
$$S_{5,i}: \quad \overrightarrow{\mathrm{SNR}}_j \geqslant \mathrm{INR}_{ij}\mathrm{INR}_{ji}. \tag{50e}$$

Notons que pour tout $i \in \{1,2\}$, les événements $S_{1,i}$, $S_{2,i}$, $S_{3,i}$, $S_{4,i}$, et $S_{5,i}$ sont mutuellement exclusifs. Cette observation indique que pour tout quadruplet $(\overrightarrow{\mathrm{SNR}}_1, \overrightarrow{\mathrm{SNR}}_2, \mathrm{INR}_{12}, \mathrm{INR}_{21})$, il existe toujours une unique paire d'événements $(S_{l_1,1}, S_{l_2,2})$, avec $(l_1,l_2) \in \{1,\ldots,5\}^2$, qui identifie un scénario unique. Notons également que les paires d'événements $(S_{2,1}, S_{2,2})$ et $(S_{3,1}, S_{3,2})$ ne sont pas réalisables. Compte tenu de ceci, vingt-trois scénarios différents peuvent être identifiés à l'aide des événements en (50). Une fois le scénario exact identifié, la région d'impossibilité est décrite à l'aide des fonctions $\kappa_{l,i} : [0,1] \to \mathbb{R}_+$, avec $l \in \{1,\ldots,3\}$; $\kappa_l : [0,1] \to \mathbb{R}_+$, avec $l \in \{4,5\}$; $\kappa_{6,l} : [0,1] \to \mathbb{R}_+$, avec $l \in \{1,\ldots,4\}$; et $\kappa_{7,i,l} : [0,1] \to \mathbb{R}_+$, avec $l \in \{1,2\}$. Ces fonctions sont définies comme suit, pour tout $i \in \{1,2\}$, avec $j \in \{1,2\}\backslash\{i\}$:

$$\kappa_{1,i}(\rho) = \frac{1}{2}\log\Big(b_{1,i}(\rho) + 1\Big), \tag{51a}$$

$$\kappa_{2,i}(\rho) = \frac{1}{2}\log\Big(1 + b_{5,j}(\rho)\Big) + \frac{1}{2}\log\left(1 + \frac{b_{4,i}(\rho)}{1 + b_{5,j}(\rho)}\right), \tag{51b}$$

$$\kappa_{3,i}(\rho) = \frac{1}{2}\log\left(\frac{\Big(b_{4,i}(\rho) + b_{5,j}(\rho) + 1\Big)\overleftarrow{\mathrm{SNR}}_j}{\Big(b_{1,j}(1)+1\Big)\Big(b_{4,i}(\rho)+1\Big)} + 1\right) + \frac{1}{2}\log\Big(b_{4,i}(\rho) + 1\Big), \tag{51c}$$

$$\kappa_4(\rho) = \frac{1}{2}\log\left(1 + \frac{b_{4,1}(\rho)}{1 + b_{5,2}(\rho)}\right) + \frac{1}{2}\log\Big(b_{1,2}(\rho) + 1\Big), \tag{51d}$$

$$\kappa_5(\rho) = \frac{1}{2}\log\left(1+\frac{b_{4,2}(\rho)}{1+b_{5,1}(\rho)}\right)+\frac{1}{2}\log\left(b_{1,1}(\rho)+1\right), \tag{51e}$$

$$\kappa_6(\rho) = \begin{cases} \kappa_{6,1}(\rho) & \text{si } (S_{1,2}\vee S_{2,2}\vee S_{5,2})\wedge(S_{1,1}\vee S_{2,1}\vee S_{5,1}) \\ \kappa_{6,2}(\rho) & \text{si } (S_{1,2}\vee S_{2,2}\vee S_{5,2})\wedge(S_{3,1}\vee S_{4,1}) \\ \kappa_{6,3}(\rho) & \text{si } (S_{3,2}\vee S_{4,2})\wedge(S_{1,1}\vee S_{2,1}\vee S_{5,1}) \\ \kappa_{6,4}(\rho) & \text{si } (S_{3,2}\vee S_{4,2})\wedge(S_{3,1}\vee S_{4,1}) \end{cases}, \tag{51f}$$

$$\kappa_{7,i}(\rho) = \begin{cases} \kappa_{7,i,1}(\rho) & \text{si } (S_{1,i}\vee S_{2,i}\vee S_{5,i}) \\ \kappa_{7,i,2}(\rho) & \text{si } (S_{3,i}\vee S_{4,i}) \end{cases}, \tag{51g}$$

où

$$\begin{aligned} \kappa_{6,1}(\rho) = {} & \frac{1}{2}\log\left(b_{1,1}(\rho)+b_{5,1}(\rho)\mathrm{INR}_{21}\right)-\frac{1}{2}\log\left(1+\mathrm{INR}_{12}\right)+\frac{1}{2}\log\left(1+\frac{b_{5,2}(\rho)\overleftarrow{\mathrm{SNR}}_2}{b_{1,2}(1)+1}\right) \\ & +\frac{1}{2}\log\left(b_{1,2}(\rho)+b_{5,1}(\rho)\mathrm{INR}_{21}\right)-\frac{1}{2}\log\left(1+\mathrm{INR}_{21}\right)+\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)\overleftarrow{\mathrm{SNR}}_1}{b_{1,1}(1)+1}\right)+\log(2\pi e), \end{aligned} \tag{52a}$$

$$\begin{aligned} \kappa_{6,2}(\rho) = {} & \frac{1}{2}\log\left(b_{6,2}(\rho)+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_2}\left(\overrightarrow{\mathrm{SNR}}_2+b_{3,2}\right)\right)-\frac{1}{2}\log\left(1+\mathrm{INR}_{12}\right) \\ & +\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)\overleftarrow{\mathrm{SNR}}_1}{b_{1,1}(1)+1}\right)+\frac{1}{2}\log\left(b_{1,1}(\rho)+b_{5,1}(\rho)\mathrm{INR}_{21}\right)-\frac{1}{2}\log\left(1+\mathrm{INR}_{21}\right) \\ & +\frac{1}{2}\log\left(1+\frac{b_{5,2}(\rho)}{\overrightarrow{\mathrm{SNR}}_2}\left(\mathrm{INR}_{12}+\frac{b_{3,2}\overleftarrow{\mathrm{SNR}}_2}{b_{1,2}(1)+1}\right)\right)-\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_2}\right)+\log(2\pi e), \end{aligned} \tag{52b}$$

$$\begin{aligned} \kappa_{6,3}(\rho) = {} & \frac{1}{2}\log\left(b_{6,1}(\rho)+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_1}\left(\overrightarrow{\mathrm{SNR}}_1+b_{3,1}\right)\right)-\frac{1}{2}\log\left(1+\mathrm{INR}_{12}\right) \\ & +\frac{1}{2}\log\left(1+\frac{b_{5,2}(\rho)\overleftarrow{\mathrm{SNR}}_2}{b_{1,2}(1)+1}\right)+\frac{1}{2}\log\left(b_{1,2}(\rho)+b_{5,1}(\rho)\mathrm{INR}_{21}\right)-\frac{1}{2}\log\left(1+\mathrm{INR}_{21}\right) \\ & +\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)}{\overrightarrow{\mathrm{SNR}}_1}\left(\mathrm{INR}_{21}+\frac{b_{3,1}\overleftarrow{\mathrm{SNR}}_1}{b_{1,1}(1)+1}\right)\right)-\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_1}\right)+\log(2\pi e), \end{aligned} \tag{52c}$$

$$\begin{aligned} \kappa_{6,4}(\rho) = {} & \frac{1}{2}\log\left(b_{6,1}(\rho)+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_1}\left(\overrightarrow{\mathrm{SNR}}_1+b_{3,1}\right)\right)-\frac{1}{2}\log\left(1+\mathrm{INR}_{12}\right) \\ & +\frac{1}{2}\log\left(1+\frac{b_{5,2}(\rho)}{\overrightarrow{\mathrm{SNR}}_2}\left(\mathrm{INR}_{12}+\frac{b_{3,2}\overleftarrow{\mathrm{SNR}}_2}{b_{1,2}(1)+1}\right)\right)-\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_2}\right) \\ & -\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_1}\right)+\frac{1}{2}\log\left(b_{6,2}(\rho)+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_2}\left(\overrightarrow{\mathrm{SNR}}_2+b_{3,2}\right)\right) \\ & -\frac{1}{2}\log\left(1+\mathrm{INR}_{21}\right)+\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)}{\overrightarrow{\mathrm{SNR}}_1}\left(\mathrm{INR}_{21}+\frac{b_{3,1}\overleftarrow{\mathrm{SNR}}_1}{b_{1,1}(1)+1}\right)\right)+\log(2\pi e), \end{aligned} \tag{52d}$$

et

$$\kappa_{7,i,1}(\rho) = \frac{1}{2}\log\left(b_{1,i}(\rho)+1\right) - \frac{1}{2}\log\left(1+\mathrm{INR}_{ij}\right) + \frac{1}{2}\log\left(1+\frac{b_{5,j}(\rho)\overleftarrow{\mathrm{SNR}}_j}{b_{1,j}(1)+1}\right)$$
$$+\frac{1}{2}\log\left(b_{1,j}(\rho)+b_{5,i}(\rho)\mathrm{INR}_{ji}\right)+\frac{1}{2}\log\left(1+b_{4,i}(\rho)+b_{5,j}(\rho)\right)-\frac{1}{2}\log\left(1+b_{5,j}(\rho)\right)$$
$$+2\log(2\pi e), \tag{53a}$$

$$\kappa_{7,i,2}(\rho) = \frac{1}{2}\log\left(b_{1,i}(\rho)+1\right) - \frac{1}{2}\log\left(1+\mathrm{INR}_{ij}\right) - \frac{1}{2}\log\left(1+b_{5,j}(\rho)\right)$$
$$+\frac{1}{2}\log\left(1+b_{4,i}(\rho)+b_{5,j}(\rho)\right)+\frac{1}{2}\log\left(1+\left(1-\rho^2\right)\frac{\mathrm{INR}_{ji}}{\overrightarrow{\mathrm{SNR}}_j}\left(\mathrm{INR}_{ij}+\frac{b_{3,j}\overleftarrow{\mathrm{SNR}}_j}{b_{1,j}(1)+1}\right)\right)$$
$$-\frac{1}{2}\log\left(1+\frac{b_{5,i}(\rho)\mathrm{INR}_{ji}}{\overrightarrow{\mathrm{SNR}}_j}\right)+\frac{1}{2}\log\left(b_{6,j}(\rho)+\frac{b_{5,i}(\rho)\mathrm{INR}_{ji}}{\overrightarrow{\mathrm{SNR}}_j}\left(\overrightarrow{\mathrm{SNR}}_j+b_{3,j}\right)\right)+2\log(2\pi e), \tag{53b}$$

où, les fonctions $b_{l,i}$, avec $(l,i)\in\{1,2\}^2$ sont définies en (48) ; les paramètres $b_{3,i}$ sont des constantes ; et les fonctions $b_{l,i}:[0,1]\to\mathbb{R}_+$, avec $(l,i)\in\{4,5,6\}\times\{1,2\}$ sont définies comme suit, avec $j\in\{1,2\}\setminus\{i\}$:

$$b_{3,i}=\overrightarrow{\mathrm{SNR}}_i-2\sqrt{\overrightarrow{\mathrm{SNR}}_i\mathrm{INR}_{ji}}+\mathrm{INR}_{ji}, \tag{54a}$$
$$b_{4,i}(\rho)=\left(1-\rho^2\right)\overrightarrow{\mathrm{SNR}}_i, \tag{54b}$$
$$b_{5,i}(\rho)=\left(1-\rho^2\right)\mathrm{INR}_{ij}, \tag{54c}$$
$$b_{6,i}(\rho)=\overrightarrow{\mathrm{SNR}}_i+\mathrm{INR}_{ij}+2\rho\sqrt{\mathrm{INR}_{ij}}\left(\sqrt{\overrightarrow{\mathrm{SNR}}_i}-\sqrt{\mathrm{INR}_{ji}}\right)$$
$$+\frac{\mathrm{INR}_{ij}\sqrt{\mathrm{INR}_{ji}}}{\overrightarrow{\mathrm{SNR}}_i}\left(\sqrt{\mathrm{INR}_{ji}}-2\sqrt{\overrightarrow{\mathrm{SNR}}_i}\right). \tag{54d}$$

Notons que les fonctions en (51), (52), (53), et (54) dépendent de $\overrightarrow{\mathrm{SNR}}_1$, $\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, INR_{21}, $\overleftarrow{\mathrm{SNR}}_1$, et $\overleftarrow{\mathrm{SNR}}_2$. Toutefois, comme ces paramètres sont fixés dans cette analyse, cette dépendance n'est pas mise en évidence dans la définition de ces fonctions. Enfin, à l'aide de cette notation, le Théorème viii est présenté comme suit.

Théorème viii. Région d'Impossibilité

La région de capacité $\mathcal{C}$ est contenue dans la région $\overline{\mathcal{C}}$ donnée par la fermeture de l'ensemble de paires de $(R_1,R_2)\in\mathbb{R}_+^2$ qui pour tout $i\in\{1,2\}$, avec $j\in\{1,2\}\setminus\{i\}$ définies :

$$R_i\leqslant\min\left(\kappa_{1,i}(\rho),\kappa_{2,i}(\rho)\right), \tag{55a}$$
$$R_i\leqslant\kappa_{3,i}(\rho), \tag{55b}$$
$$R_1+R_2\leqslant\min\left(\kappa_4(\rho),\kappa_5(\rho)\right), \tag{55c}$$
$$R_1+R_2\leqslant\kappa_6(\rho), \tag{55d}$$
$$2R_i+R_j\leqslant\kappa_{7,i}(\rho), \tag{55e}$$

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

5.3. Écart entre la Région Atteignable et la Région d'Impossibilité

Le Théorème ix décrit l'écart entre la région atteignable $\underline{\mathcal{C}}$ et la région d'impossibilité $\overline{\mathcal{C}}$ (Définition vi).

Théorème ix. Écart

La région de capacité du GIC-NOF à deux utilisateurs est approximée à 4.4 bits près par la région atteignable $\underline{\mathcal{C}}$ et la région d'impossibilité $\overline{\mathcal{C}}$.

5.4. Cas où la Voie de Retour Agrandit la Région de Capacité

Cette section examine l'application des résultats obtenus à la Section 4.2 dans le GIC-NOF à deux utilisateurs. En Section 3, ont été présentées les connexions entre le LDIC-NOF à deux utilisateurs et le GIC-NOF à deux utilisateurs. À l'aide de ces connexions, un CGI avec des paramètres fixés $\left(\overrightarrow{\mathrm{SNR}}_1, \overrightarrow{\mathrm{SNR}}_2, \mathrm{INR}_{12}, \mathrm{INR}_{21}\right)$ est approximé par un LDIC avec des paramètres $\overrightarrow{n}_{11} = \left\lfloor \frac{1}{2}\log(\overrightarrow{\mathrm{SNR}}_1)\right\rfloor$, $\overrightarrow{n}_{22} = \left\lfloor \frac{1}{2}\log(\overrightarrow{\mathrm{SNR}}_2)\right\rfloor$, $n_{12} = \left\lfloor \frac{1}{2}\log(\mathrm{INR}_{12})\right\rfloor$ et $n_{21} = \left\lfloor \frac{1}{2}\log(\mathrm{INR}_{21})\right\rfloor$. À partir de cette observation, les résultats des Théorème ii - Théorème v peuvent servir à déterminer les seuils de SNR de la voie de retour au-delà duquel un des débits individuels ou le débit somme est amélioré dans le GIC-NOF original. La procédure consiste à utiliser les égalités $\overleftarrow{n}_{ii} = \left\lfloor \frac{1}{2}\log\left(\overleftarrow{\mathrm{SNR}}_i\right)\right\rfloor$, avec $i \in \{1,2\}$. Par conséquent, les seuils correspondants dans le CGI à deux utilisateurs peuvent être calculés approximativement par :

$$\overleftarrow{\mathrm{SNR}}_i^{*} = 2^{2\overleftarrow{n}_{ii}^{*}}, \tag{56a}$$

$$\overleftarrow{\mathrm{SNR}}_i^{\dagger} = 2^{2\overleftarrow{n}_{ii}^{\dagger}}, \text{ et} \tag{56b}$$

$$\overleftarrow{\mathrm{SNR}}_i^{+} = 2^{2\overleftarrow{n}_{ii}^{+}}. \tag{56c}$$

Lorsque le LDIC-NOF correspondant est tel que sa région de capacité peut être agrandie, à savoir, si $\overleftarrow{n}_{ii} > \overleftarrow{n}_{ii}^{*}$ (Théorème ii) pour un $i \in \{1,2\}$, on s'attend à ce que, soit la faisabilité, soit les régions d'impossibilité du GIC-NOF original s'agrandissent lorsque $\overleftarrow{\mathrm{SNR}}_i > \overleftarrow{\mathrm{SNR}}_i^{*}$. De même, si le LDIC-NOF correspondant est tel que $\Delta_i(\overleftarrow{n}_{ii}) > 0$ ou $\Delta_i(\overleftarrow{n}_{jj}) > 0$, on s'attend à observer une amélioration du débit individuel R_i soit en utilisant la voie de retour à la paire émetteur-récepteur i, avec $\overleftarrow{\mathrm{SNR}}_i > \overleftarrow{\mathrm{SNR}}_i^{\dagger}$ soit en utilisant la voie de retour à la paire émetteur-récepteur j, avec $\overleftarrow{\mathrm{SNR}}_j > \overleftarrow{\mathrm{SNR}}_j^{*}$. Lorsque le LDIC-NOF correspondant est tel que $\Sigma(\overleftarrow{n}_{ii}) > 0$ en utilisant la voie de retour à la paire émetteur-récepteur i, avec $\overleftarrow{n}_{ii} > \overleftarrow{n}_{ii}^{+}$ (Théorème v), on s'attend à observer une amélioration au niveau du débit somme en utilisant la voie de retour à la paire émetteur-récepteur i, avec $\overleftarrow{\mathrm{SNR}}_i > \overleftarrow{\mathrm{SNR}}_i^{+}$. Enfin, lorsqu'on n'observe aucune amélioration dans une métrique donnée dans le LDIC-NOF à deux utilisateurs, à savoir, $\Delta_1(\overleftarrow{n}_{11}) = 0$, $\Delta_1(\overleftarrow{n}_{22}) = 0$, $\Delta_2(\overleftarrow{n}_{11}) = 0$, $\Delta_2(\overleftarrow{n}_{22}) = 0$, $\Sigma(\overleftarrow{n}_{11}) = 0$, ou $\Sigma(\overleftarrow{n}_{22}) = 0$, on n'observe qu'une amélioration négligeable (le cas échéant) dans la métrique correspondante du GIC-NOF à deux utilisateurs.

6. Principaux Résultats du Canal Linéaire Déterministe à Interférence Décentralisé

Cette section présente les principaux résultats sur le D-LDIC-NOF à deux utilisateurs. Ce modèle est décrit dans la Section 1.2 et peut être analysé par un jeu comme suggéré dans la Section 2.1. Notons $\mathcal{C}(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ la région de capacité du LDIC-NOF à deux utilisateurs avec les paramètres $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$, et $\overleftarrow{n}_{22}$, décrits dans le Théorème i.

6.1. Région d'η-Équilibre de Nash

Cette section présente la région d'η-équilibre de Nash $\mathcal{N}_\eta$ (Définition v) du D-LDIC-NOF à deux utilisateurs.

La région d'η-équilibre de Nash $\mathcal{N}_\eta$ du D-LDIC-NOF, à deux utilisateurs, étant donné les paramètres fixés $\left(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}\right) \in \mathbb{N}^6$, est notée, $\mathcal{N}_\eta\left(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}\right)$. La région d'$\eta$-équilibre de Nash $\mathcal{N}_\eta$ est caractérisée par deux régions : la région de capacité, notée par $\mathcal{C}\left(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}\right)$ et une région convexe, notée par $\mathcal{B}_\eta\left(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}\right)$. Dans les sections suivantes, le sextuplet $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ sera explicité seulement s'il est nécessaire.

La région de capacité $\mathcal{C}$ du LDIC-NOF est décrite par le Théorème i. Pour tout $\eta > 0$, la région convexe $\mathcal{B}_\eta$ est définie de la manière suivante :

$$\mathcal{B}_\eta=\left\{(R_1, R_2) \in \mathbb{R}_+^2 : L_i \leqslant R_i \leqslant U_i, \text{ pour tout } i \in \mathcal{K} = \{1,2\}\right\}, \tag{57}$$

où,

$$L_i=\left((\overrightarrow{n}_{ii} - n_{ij})^+ - \eta\right)^+ \text{ et} \tag{58a}$$

$$U_i=\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \left(\min\left((\overrightarrow{n}_{jj}-n_{ji})^+, n_{ij}\right) - \left(\min\left((\overrightarrow{n}_{jj}-n_{ij})^+, n_{ji}\right) - (\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj})^+\right)^+\right)^+ + \eta, \tag{58b}$$

avec $i \in \{1,2\}$ et $j \in \{1,2\} \setminus \{i\}$. Le Théorème x utilise la région $\mathcal{B}_\eta$ dans (57) et la région de capacité $\mathcal{C}$ pour décrire la région d'η-équilibre de Nash $\mathcal{N}_\eta$.

Théorème x. Région d'η-Équilibre de Nash

Soit $\eta > 0$ fixé. La région d'η-équilibre de Nash $\mathcal{N}_\eta$ du canal linéaire déterministe à interférence avec rétro-alimentation dégradée par un bruit additif décrit par les paramètres $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$ et $\overleftarrow{n}_{22}$, est

$$\mathcal{N}_\eta = \mathcal{C} \cap \mathcal{B}_\eta. \tag{59}$$

6.2. Agrandissement de la Région d'η-Équilibre de Nash avec Voie de Retour

La métrique, les conditions et les valeurs pour les paramètres de la voie de retour au-delà desquelles la région d'η-équilibre de Nash $\mathcal{N}_\eta$ du LDIC-NOF à deux utilisateurs peut être agrandie sont les mêmes que dans le cas centralisé, en tenant compte du fait que ceux-ci se réfèrent à la région d'η-équilibre de Nash $\mathcal{N}_\eta$ au lieu de la région de capacité.

6.3. Efficacité d'η-Équilibre de Nash

Cette section caractérise l'efficacité de l'ensemble des équilibres dans le D-LDIC-NOF à deux utilisateurs à l'aide de deux indicateurs : le prix de l'anarchie (PoA) et le prix de la stabilité (PoS).

Définitions

Les résultats de cette section sont présentés au moyen d'une liste d'événements (variables booléennes) déterminés par les paramètres $\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}$, et n_{21}. Soient $i \in \{1, 2\}$ et $j \in \{1, 2\} \setminus \{i\}$, et définissons les évènements suivants :

$$\begin{aligned}
A_{1,i} &: \quad \overrightarrow{n}_{ii} - n_{ij} \geqslant n_{ji}, && \text{(60a)}\\
A_{2,i} &: \quad \overrightarrow{n}_{ii} \geqslant n_{ji}, && \text{(60b)}\\
B_1 &: \quad A_{1,1} \wedge A_{1,2}, && \text{(60c)}\\
B_{2,i} &: \quad A_{1,i} \wedge \overline{A}_{1,j} \wedge A_{2,j}, && \text{(60d)}\\
B_{3,i} &: \quad A_{1,i} \wedge \overline{A}_{1,j} \wedge \overline{A}_{2,j}, && \text{(60e)}\\
B_4 &: \quad \overline{A}_{1,1} \wedge \overline{A}_{1,2} \wedge A_{2,1} \wedge A_{2,2}, && \text{(60f)}\\
B_{5,i} &: \quad \overline{A}_{1,1} \wedge \overline{A}_{1,2} \wedge \overline{A}_{2,i} \wedge A_{2,j}, && \text{(60g)}\\
B_6 &: \quad \overline{A}_{1,1} \wedge \overline{A}_{1,2} \wedge \overline{A}_{2,1} \wedge \overline{A}_{2,2}, && \text{(60h)}\\
B_7 &: \quad A_{1,1}, && \text{(60i)}\\
B_8 &: \quad \overline{A}_{1,1} \wedge A_{2,1} \wedge A_{2,2}, && \text{(60j)}\\
B_9 &: \quad \overline{A}_{1,1} \wedge \overline{A}_{2,1} \wedge A_{2,2}, && \text{(60k)}\\
B_{10} &: \quad \overline{A}_{1,1} \wedge \overline{A}_{2,2}. && \text{(60l)}
\end{aligned}$$

Lorsque les deux paires émetteur-récepteur sont en LIR, à savoir, $\overrightarrow{n}_{11} > n_{12}$ et $\overrightarrow{n}_{22} > n_{21}$, les événements B_1, $B_{2,1}$, $B_{2,2}$, $B_{3,1}$, $B_{3,2}$, B_4, $B_{5,1}$, $B_{5,2}$, et B_6 satisfont la propriété déclarée par le lemme suivant.

Lemme i. *Pour un quadruplet fixe* $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$ *avec* $\overrightarrow{n}_{11} > n_{12}$ *et* $\overrightarrow{n}_{22} > n_{21}$*, seul un des événements* B_1*,* $B_{2,1}$*,* $B_{2,2}$*,* $B_{3,1}$*,* $B_{3,2}$*,* B_4*,* $B_{5,1}$*,* $B_{5,2}$*, et* B_6 *est vrai.*

Lorsque la paire émetteur-transmetteur 1 est en LIR et la paire émetteur-transmetteur 2 est en HIR, à savoir, $\overrightarrow{n}_{11} > n_{12}$ et $\overrightarrow{n}_{22} \leqslant n_{21}$, les événements B_7, B_8, B_9, et B_{10} satisfont la propriété déclarée par le lemme suivant.

Lemme ii. *Pour un quadruplet fixe* $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$ *avec* $\overrightarrow{n}_{11} > n_{12}$ *et* $\overrightarrow{n}_{22} \leqslant n_{21}$*, seul un des événements* B_7*,* B_8*,* B_9*, et* B_{10} *est vrai.*

Prix de l'Anarchie

Soit $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2$ l'ensemble de toutes les paires de configuration possibles et $\mathcal{A}_{\eta-\text{EN}} \subset \mathcal{A}$ l'ensemble de paires de configuration d'η-équilibre de Nash du jeu dans (13) (Définition iv).

Définition vii (Prix de l'Anarchie [45])**.** *Soit $\eta > 0$. Le PoA du jeu $\mathcal{G}$, noté* $\mathrm{PoA}\,(\eta, \mathcal{G})$*, est donné par :*

$$\mathrm{PoA}\,(\eta, \mathcal{G}) = \frac{\displaystyle\max_{(s_1,s_2)\in\mathcal{A}} \sum_{i=1}^{2} R_i(s_1, s_2)}{\displaystyle\min_{(s_1^*,s_2^*)\in\mathcal{A}_{\eta-\mathrm{EN}}} \sum_{i=1}^{2} R_i(s_1^*, s_2^*)}. \tag{61}$$

Soit $\overline{\Sigma}_C$ la solution du problème d'optimisation dans le numérateur de (61), qui correspond au débit somme maximum dans le cas centralisé. Soit aussi $\underline{\Sigma}_N$ la solution du problème d'optimisation dans le dénominateur de (61).

Les théorèmes suivants explicitent l'expression du $\mathrm{PoA}\,(\eta, \mathcal{G})$ dans les régimes d'interférence particuliers du D-LDIC-NOF à deux utilisateurs. Dans tous les cas, on suppose que $\overleftarrow{n}_{ii} \leqslant \max\left(\overrightarrow{n}_{ii}, n_{ij}\right)$ pour tout $i \in \{1,2\}$ et $j \in \{1,2\} \setminus \{i\}$. Si $\overleftarrow{n}_{11} > \max\left(\overrightarrow{n}_{11}, n_{12}\right)$ ou $\overleftarrow{n}_{22} > \max\left(\overrightarrow{n}_{22}, n_{21}\right)$, les résultats sont les mêmes que ceux dans le cas de POF, à savoir, $\overleftarrow{n}_{11} = \max\left(\overrightarrow{n}_{11}, n_{12}\right)$ ou $\overleftarrow{n}_{22} = \max\left(\overrightarrow{n}_{22}, n_{21}\right)$.

Théorème xi. Les Deux Paires Émetteur-Récepteur en LIR

Pour tout $i \in \{1,2\}$, $j \in \{1,2\} \setminus \{i\}$ et pour tout $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \in \mathbb{N}^6$ avec $\overrightarrow{n}_{11} > n_{12}$ et $\overrightarrow{n}_{22} > n_{21}$, le $\mathrm{PoA}\,(\eta, \mathcal{G})$ satisfait :

$$\mathrm{PoA}\,(\eta, \mathcal{G}) = \begin{cases} \dfrac{\overline{\Sigma}_{C1}}{\overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21} - 2\eta} & \text{si} \quad B_1 \text{ est vrai} \\ \dfrac{\overline{\Sigma}_{C2,i}}{\overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21} - 2\eta} & \text{si} \quad B_{2,i} \text{ est vrai} \\ \dfrac{\overrightarrow{n}_{ii}}{\overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21} - 2\eta} & \text{si} \quad B_{3,i} \vee B_{5,i} \text{ est vrai} \\ \dfrac{\overline{\Sigma}_{C3}}{\overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21} - 2\eta} & \text{si} \quad B_4 \text{ est vrai} \\ \dfrac{\min(\overrightarrow{n}_{11}, \overrightarrow{n}_{22})}{\overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21} - 2\eta} & \text{si} \quad B_6 \text{ est vrai} \end{cases}, \tag{62}$$

où,

$$\begin{aligned} \overline{\Sigma}_{C1} = \min\Big(&\overrightarrow{n}_{22} + \overrightarrow{n}_{11} - n_{12}, \overrightarrow{n}_{11} + \overrightarrow{n}_{22} - n_{21}, \\ &\max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right) + \max\left(\overrightarrow{n}_{22} - n_{21}, \overleftarrow{n}_{22}\right), \\ &2\overrightarrow{n}_{11} - n_{12} + \max\left(\overrightarrow{n}_{22} - n_{21}, \overleftarrow{n}_{22}\right), 2\overrightarrow{n}_{22} - n_{21} + \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right)\Big); \end{aligned} \tag{63a}$$

$$\overline{\Sigma}_{C2,i}=\min\Big(\overrightarrow{n}_{22}+\overrightarrow{n}_{11}-n_{12},\overrightarrow{n}_{11}+\overrightarrow{n}_{22}-n_{21},$$
$$\max\left(\overrightarrow{n}_{11}-n_{12},\overleftarrow{n}_{11}\right)+\max\left(\overrightarrow{n}_{22}-n_{21},\overleftarrow{n}_{22}\right),$$
$$2\overrightarrow{n}_{ii}-n_{ij}+\max\left(n_{ij},\overleftarrow{n}_{jj}\right),2\overrightarrow{n}_{jj}-n_{ji}+\max\left(\overrightarrow{n}_{ii}-n_{ij},\overleftarrow{n}_{ii}\right)\Big);\text{ et} \quad (63b)$$

$$\overline{\Sigma}_{C3}=\min\Big(\overrightarrow{n}_{22}+\overrightarrow{n}_{11}-n_{12},\overrightarrow{n}_{11}+\overrightarrow{n}_{22}-n_{21},\max\left(n_{21},\overleftarrow{n}_{11}\right)+\max\left(n_{12},\overleftarrow{n}_{22}\right),$$
$$2\overrightarrow{n}_{11}-n_{12}+\max\left(n_{12},\overleftarrow{n}_{22}\right),2\overrightarrow{n}_{22}-n_{21}+\max\left(n_{21},\overleftarrow{n}_{11}\right)\Big). \quad (63c)$$

Théorème xii. Émetteur-Récepteur 1 en LIR et Émetteur-Récepteur 2 en HIR

Pour tout $(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})\in\mathbb{N}^6$ avec $\overrightarrow{n}_{11}>n_{12}$ et $\overrightarrow{n}_{22}\leqslant n_{21}$, le $\text{PoA}(\eta,\mathcal{G})$ satisfait :

$$\text{PoA}(\eta,\mathcal{G})=\begin{cases}\dfrac{\overrightarrow{n}_{11}}{\overrightarrow{n}_{11}-n_{12}-\eta} & \text{si} \quad B_7\vee B_8\vee B_{10}\text{ est vrai}\\[2ex] \dfrac{\min(\overrightarrow{n}_{22}+\overrightarrow{n}_{11}-n_{12},n_{21})}{\overrightarrow{n}_{11}-n_{12}-\eta} & \text{si} \quad B_9\text{ est vrai}\end{cases}. \quad (64)$$

Notons que dans les cas dans lesquels la paire émetteur-récepteur 1 est en LIR et la paire émetteur-récepteur 2 est en HIR, le $\text{PoA}(\eta,\mathcal{G})$ ne dépend pas des paramètres de la voie de retour. Ceci en résulte dans la mesure où l'utilisation de la voie de retour dans ce scénario peut agrandir la région de capacité mais n'augmente pas la capacité du débit somme (Théorème v).

Dans le cas dans lequel la paire émetteur-récepteur 1 est en HIR et la paire émetteur-récepteur 2 est en LIR, à savoir, $\overrightarrow{n}_{11}\leqslant n_{12}$ et $\overrightarrow{n}_{22}>n_{21}$, le $\text{PoA}(\eta,\mathcal{G})$ pour le D-LDIC-NOF à deux utilisateurs est qualifié comme dans le Théorème xii interchangeant les indices des paramètres.

Théorème xiii. Les Deux Paires Émetteur-Récepteur en HIR

Pour tout $(\overrightarrow{n}_{11},\ \overrightarrow{n}_{22},\ n_{12},\ n_{21},\ \overleftarrow{n}_{11},\ \overleftarrow{n}_{22})\in\mathbb{N}^6$ avec $\overrightarrow{n}_{11}\leqslant n_{12}$ et $\overrightarrow{n}_{22}\leqslant n_{21}$, le $\text{PoA}(\eta,\mathcal{G})$ satisfait :

$$\text{PoA}(\eta,\mathcal{G})=\infty. \quad (65)$$

Le résultat du Théorème xiii est dû au fait que $\left(\left(\overrightarrow{n}_{11}-n_{12}\right)^+-\eta\right)^++\left(\left(\overrightarrow{n}_{22}-n_{21}\right)^+-\eta\right)^+=0$. Autrement dit, lorsque $\overrightarrow{n}_{11}\leqslant n_{12}$ et $\overrightarrow{n}_{22}\leqslant n_{21}$, aucune des paires d'émetteur-récepteur n'est capable de transmettre à un débit strictement positif au pire η-équilibre de Nash (le plus petit débit somme dans la région d'η-équilibre de Nash $\mathcal{N}_\eta$).

D'une façon générale, dans tout régime d'interférence dans lequel le $\text{PoA}(\eta,\mathcal{G})$ dépend des paramètres de la voie de retour $\overleftarrow{n}_{11}$ ou $\overleftarrow{n}_{22}$, il existe une valeur de paramètre de la voie de retour $\overleftarrow{n}_{11}$ ou de paramètre de la voie de retour $\overleftarrow{n}_{22}$ au-delà duquel le PoA $(\eta, \mathcal{G})$ augmente. Ces valeurs correspondent à celles au-delà desquelles la capacité de somme peut être augmentée (Théorème v).

Prix de la Stabilité

Dans cette section, l'efficacité d'η-équilibre de Nash du jeu $\mathcal{G}$ en (13) est analysée à l'aide du PoS.

Définition viii (Prix de la Stabilité [4])**.** *Soit $\eta > 0$. Le PoS du jeu $\mathcal{G}$, dénoté par* $\mathrm{PoS}\,(\eta, \mathcal{G})$, *est donné par :*

$$\mathrm{PoS}\,(\eta, \mathcal{G}) = \frac{\max_{(s_1,s_2)\in\mathcal{A}} \sum_{i=1}^{2} R_i(s_1, s_2)}{\max_{(s_1^*,s_2^*)\in\mathcal{A}_{\eta-\mathrm{EN}}} \sum_{i=1}^{2} R_i(s_1^*, s_2^*)}. \quad (66)$$

Soit $\overline{\Sigma}_N$ la solution du problème d'optimisation dans le dénominateur de (66).

La proposition suivante caractérise le PoS du jeu $\mathcal{G}$ dans (13) pour le D-LDIC-NOF à deux utilisateurs.

Proposition 1 (PoS)**.** Pour tout $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \in \mathbb{N}^6$ et pour tout $\eta > 0$ arbitrairement petit, le PoS dans le jeu $\mathcal{G}$ du D-LDIC-NOF à deux utilisateurs est :

$$\mathrm{PoS}\,(\eta, \mathcal{G}) = 1. \quad (67)$$

Notons que le fait que le prix de la stabilité soit égal à un, indépendamment des paramètres $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$ et $\overleftarrow{n}_{22}$, implique que malgré le comportement anarchique des deux paires émetteur-récepteur, le débit somme d'η-équilibre de Nash le plus élevé est égal à la capacité du débit somme, à savoir, $\overline{\Sigma}_C = \overline{\Sigma}_N$. Cela implique que dans tous les régimes d'interférence, il existe toujours un η-équilibre de Nash optimal pour le débit somme (optimum de Pareto d'η-équilibre de Nash). Les seuils sur les paramètres de la voie de retour au-delà desquels la somme des capacités et le débit somme maximum dans la région d'η-équilibre de Nash $\mathcal{N}_\eta$ peuvent être améliorés sont obtenus d'après le Théorème v.

7. Principaux Résultats du Canal Gaussien à Interférence Décentralisé

Cette section présente les principaux résultats sur le D-GIC-NOF à deux utilisateurs. Ce modèle est décrit dans la Section 1.1 et peut être modélisé par un jeu comme suggéré dans la Section 2.1. Notons $\mathcal{C}$ la région de capacité du GIC-NOF à deux utilisateurs avec les paramètres fixés $\overrightarrow{\mathrm{SNR}}_1$, $\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, INR_{21}, $\overleftarrow{\mathrm{SNR}}_1$, et $\overleftarrow{\mathrm{SNR}}_2$. La région de capacité atteignable $\underline{\mathcal{C}}$ et la région d'impossibilité $\overline{\mathcal{C}}$ correspondent approximent la région de capacité $\mathcal{C}$ à 4.4 bits près (Théorème ix). La région de capacité atteignable $\underline{\mathcal{C}}$ et la région d'impossibilité $\overline{\mathcal{C}}$ sont définies par le Théorème vii et Théorème viii, respectivement.

7.1. Région d'η-Équilibre de Nash Atteignable

Soit la région d'η-équilibre de Nash (Définition v) du D-GIC-NOF à deux utilisateurs indiquée par $\mathcal{N}_\eta$. Cette section présente une région $\underline{\mathcal{N}}_\eta \subseteq \mathcal{N}_\eta$ qui est atteignable en utilisant le système Han-Kobayashi randomisé avec la voie de retour dégradée par un bruit additif (RHK-NOF). Le RHK-NOF s'avère être une paire de configuration d'η-équilibre de Nash avec $\eta \geqslant 1$. Autrement dit, toute déviation unilatérale du RHK-NOF par toute paire émetteur-récepteur peut entraîner une amélioration du débit individuel d'au maximum un bit par utilisation du canal. La description de la région d'η-équilibre de Nash atteignable $\underline{\mathcal{N}}_\eta$ est présentée à l'aide des constantes $a_{1,i}$; des fonctions $a_{2,i} : [0,1] \rightarrow \mathbb{R}_+$, $a_{l,i} : [0,1]^2 \rightarrow \mathbb{R}_+$, avec $l \in \{3,\ldots,6\}$; et $a_{7,i} : [0,1]^3 \rightarrow \mathbb{R}_+$, défini dans (47), pour tout $i \in \{1,2\}$, avec $j \in \{1,2\} \setminus \{i\}$, et les fonctions $b_{l,i} : [0,1] \rightarrow \mathbb{R}_+$, avec $(l,i) \in \{1,2\}^2$, défini en (48). À l'aide de cette notation, le résultat principal est présenté par le Théorème xiv.

Théorème xiv. Région d'η-Équilibre de Nash Atteignable

Soit $\eta \geqslant 1$. La région d'η-équilibre de Nash atteignable $\underline{\mathcal{N}}_\eta$ est donnée par la fermeture de toutes les paires possibles de débit atteignables $(R_1, R_2) \in \underline{\mathcal{C}}$ satisfaisant, pour tout $i \in \{1,2\}$ et $j \in \{1,2\} \setminus \{i\}$, les conditions suivantes :

$$R_i \geqslant \Big(a_{2,i}(\rho) - a_{3,i}(\rho,\mu_j) - a_{4,i}(\rho,\mu_j) - \eta\Big)^+, \tag{68a}$$

$$\begin{aligned} R_i \leqslant \min \Big(& a_{2,i}(\rho) + a_{3,j}(\rho,\mu_i) + a_{5,j}(\rho,\mu_i) - a_{2,j}(\rho) + \eta, \\ & a_{3,i}(\rho,\mu_j) + a_{7,i}(\rho,\mu_1,\mu_2) + 2a_{3,j}(\rho,\mu_i) + a_{5,j}(\rho,\mu_i) - a_{2,j}(\rho) + \eta, \\ & a_{2,i}(\rho) + a_{3,i}(\rho,\mu_j) + 2a_{3,j}(\rho,\mu_i) + a_{5,j}(\rho,\mu_i) + a_{7,j}(\rho,\mu_1,\mu_2) - 2a_{2,j}(\rho) + 2\eta\Big), \end{aligned} \tag{68b}$$

$$R_1 + R_2 \leqslant a_{1,i} + a_{3,i}(\rho,\mu_j) + a_{7,i}(\rho,\mu_1,\mu_2) + a_{2,j}(\rho) + a_{3,j}(\rho,\mu_1) - a_{2,i}(\rho) + \eta, \tag{68c}$$

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

7.2. Région de Deséquilibre

Soit la région d'η-équilibre de Nash (Définition v) du D-GIC-NOF à deux utilisateurs indiquée par $\mathcal{N}_\eta$. Cette section présente une région $\overline{\mathcal{N}}_\eta \supseteq \mathcal{N}_\eta$ donnée en termes de région convexe $\overline{\mathcal{B}}_\eta$. Ici, pour le cas du D-GIC-NOF à deux utilisateurs, la région convexe $\overline{\mathcal{B}}_\eta$ est donnée par la fermeture des paires de débit non négatif (R_1, R_2) qui pour tout $i \in \{1,2\}$, avec $j \in \{1,2\} \setminus \{i\}$ vérifient :

$$\overline{\mathcal{B}}_\eta = \Big\{(R_1,R_2) \in \mathbb{R}_+^2 : L_i \leqslant R_i \leqslant \overline{U}_i, \text{ pour tout } i \in \mathcal{K} = \{1,2\}\Big\}, \tag{69}$$

où

$$L_i = \left(\frac{1}{2}\log\left(1 + \frac{\overrightarrow{\mathrm{SNR}}_i}{1 + \mathrm{INR}_{ij}}\right) - \eta\right)^+ \quad \text{et} \tag{70}$$

$$
\begin{aligned}
\overline{U}_i = & \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_i + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_i \mathrm{INR}_{ij}} + \mathrm{INR}_{ij} + 1\right)\\
& -\frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_j + 2\left(\rho - \rho_{X_iV_j}\sqrt{\gamma_j}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_j \mathrm{INR}_{ji}} + \mathrm{INR}_{ji} + 1}{(1-\gamma_j)\,\overrightarrow{\mathrm{SNR}}_j + 2\left(\rho - \rho_{X_iV_j}\sqrt{\gamma_j}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_j \mathrm{INR}_{ji}} + \mathrm{INR}_{ji} + 1}\right)\\
& -\frac{1}{2}\log\left(\mathrm{INR}_{ji}\left(1-\rho^2\right)+1\right) + \frac{1}{2}\log\left(\mathrm{INR}_{ji}\left(1-\left(\rho - \rho_{X_iV_j}\sqrt{\gamma_j}\right)^2\right)+1\right)\\
& +\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_j + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_j \mathrm{INR}_{ji}} + \mathrm{INR}_{ji} + 1\right)\\
& -\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_j\left(\gamma_j - \gamma_j^2\right) + 2\gamma_j\left(\rho - \rho_{X_iV_j}\sqrt{\gamma_j}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_j \mathrm{INR}_{ji}} + \gamma_j \mathrm{INR}_{ji}\left(1-\rho^2_{X_iV_j}\right) + \gamma_j\right)\\
& -\frac{1}{2}\log\left(1-\rho^2_{X_iV_j}\right) - \frac{1}{2}\log\left(\mathrm{INR}_{ij}\left(1-\rho^2\right)+1\right)\\
& +\frac{1}{2}\log\left(\gamma_j\left(\mathrm{INR}_{ij}\left(1-\rho^2\right)+1\right) - \rho^2_{X_iV_j}\gamma_j\left(\mathrm{INR}_{ij}+1\right) + \gamma_2 \mathrm{INR}_{ij}\left(2\rho_{X_iV_j}\rho\sqrt{\gamma_j} - \gamma_j\right)\right)\\
& +\eta. \qquad (71)
\end{aligned}
$$

avec

$$
\gamma_j = \begin{cases}
\min\left(\dfrac{\overrightarrow{\mathrm{SNR}}_j}{\mathrm{INR}_{ij}\mathrm{INR}_{ji}}, \dfrac{1}{\mathrm{INR}_{ji}}\right) & \text{si} \quad C_{1,j} \vee \left(C_{2,j} \wedge C_{3,j}\right) \text{ est vrai}\\
\min\left(\dfrac{1}{\mathrm{INR}_{ji}}, \dfrac{\mathrm{INR}_{ij}}{\mathrm{INR}_{ji}\overrightarrow{\mathrm{SNR}}_j}\right) & \text{si} \quad C_{4,j} \text{ est vrai}\\
\min\left(1, \dfrac{\mathrm{INR}_{ij}}{\overrightarrow{\mathrm{SNR}}_j}\right) & \text{si} \quad C_{5,j} \wedge C_{6,j} \text{ est vrai}\\
\min\left(\dfrac{\overrightarrow{\mathrm{SNR}}_j}{\overleftarrow{\mathrm{SNR}}_j \mathrm{INR}_{ji}}, \dfrac{\mathrm{INR}_{ij}}{\overleftarrow{\mathrm{SNR}}_j \mathrm{INR}_{ji}}, 1\right) & \text{si} \quad C_{7,j} \vee C_{8,j} \text{ est vrai}\\
0 & \quad \text{otherwise}
\end{cases}, \qquad (72)
$$

$$C_{1,j} : \mathrm{INR}_{ji} < \overrightarrow{\mathrm{SNR}}_j \leqslant \mathrm{INR}_{ij}, \qquad (73a)$$

$$C_{2,j} : \max\left(\mathrm{INR}_{ij}, \mathrm{INR}_{ji}, \overleftarrow{\mathrm{SNR}}_j\right) < \overrightarrow{\mathrm{SNR}}_j < \mathrm{INR}_{ij}\mathrm{INR}_{ji}, \qquad (73b)$$

$$C_{3,j} : \overleftarrow{\mathrm{SNR}}_j \leqslant \mathrm{INR}_{ij}, \qquad (73c)$$

$$C_{4,j} : \mathrm{INR}_{ji} < \mathrm{INR}_{ij} < \overrightarrow{\mathrm{SNR}}_j \leqslant \overleftarrow{\mathrm{SNR}}_j, \qquad (73d)$$

$$C_{5,j} : \overrightarrow{\mathrm{SNR}}_j > \max\left(\mathrm{INR}_{ij}, \mathrm{INR}_{ji}, \overleftarrow{\mathrm{SNR}}_j\right), \qquad (73e)$$

$$C_{6,j} : \overrightarrow{\mathrm{SNR}}_j \geqslant \max\left(\mathrm{INR}_{ij}\mathrm{INR}_{ji}, \overleftarrow{\mathrm{SNR}}_j \mathrm{INR}_{ji}\right), \qquad (73f)$$

$$C_{7,j} : \max\left(\mathrm{INR}_{ij}, \mathrm{INR}_{ji}, \overleftarrow{\mathrm{SNR}}_j, \frac{\overleftarrow{\mathrm{SNR}}_j \mathrm{INR}_{ji}}{\mathrm{INR}_{ij}}\right) < \overrightarrow{\mathrm{SNR}}_j < \overleftarrow{\mathrm{SNR}}_j \mathrm{INR}_{ji} \leqslant \frac{\overleftarrow{\mathrm{SNR}}_j \overrightarrow{\mathrm{SNR}}_j}{\mathrm{INR}_{ij}}, \qquad (73g)$$

$$C_{8,j} : \max\left(\mathrm{INR}_{ij}, \mathrm{INR}_{ji}, \overleftarrow{\mathrm{SNR}}_j, \frac{\overleftarrow{\mathrm{SNR}}_j \mathrm{INR}_{ji}}{\mathrm{INR}_{ij}}\right) < \overrightarrow{\mathrm{SNR}}_j < \mathrm{INR}_{ij}\mathrm{INR}_{ji} < \overleftarrow{\mathrm{SNR}}_j \mathrm{INR}_{ji}, \qquad (73h)$$

,$\rho \in [0,1]$, et $\rho_{X_i V_j} \in [0,1)$.

Notons que L_i est le débit obtenu par la paire d'émetteur-récepteur i quand il sature la contrainte de puissance en (7) et traite l'interférence comme un bruit. Conformément à cette notation, la région de deséquilibre du GIC-NOF à deux utilisateurs, à savoir, $\overline{\mathcal{N}}_\eta$, est caracterisé par le thèoréme suivant.

Théorème xv. Région de Deséquilibre

Soit $\eta \geqslant 1$. La région de deséquilibre $\overline{\mathcal{N}}_\eta$ du D-GIC-NOF à deux utilisateurs est donnée par la fermeture de toutes les paires de débits non négatifs possibles $(R_1, R_2) \in \overline{\mathcal{C}} \cap \overline{\mathcal{B}}_\eta$ pour tout $\rho \in [0,1]$.

— 1 —

Introduction

The interference channel (IC) is one of the simplest yet insightful multi-user channels in network information theory. An important class of ICs is the two-user Gaussian interference channel (GIC) in which there exist two point-to-point links subject to mutual interference. In this model, each output signal is a noisy version of the sum of the two transmitted signals affected by the corresponding channel gains. The two-user GIC is a model that forms a basis to analyze not only the effect of the noise but also the effect of the interference in a multi-user communication system.

Some of the techniques often used to deal with interference have been to avoid it, suppress it, or treat it as noise. However, these techniques are not necessarily optimal in all cases. These approaches follow the long-established convention of communication networks in which nodes act as stand alone systems without considering the messages transmitted by other nodes [49]. From this perspective, the determination of the capacity region of a two-user GIC remains as a long standing open problem. The capacity region of the two-user GIC is known in the very strong interference regime [22] and in the strong interference regime [37, 81]. In both of these cases, each receiver must decode the messages coming from both transmitters. The best known achievable region for the two-user GIC is given in [37], which is simplified in [24]. The strategy in [37] uses rate-splitting [23], whereas the strategy in [24] uses both rate-splitting [23, 37] and block-Markov superposition coding [27]. These strategies split each user's message into two parts: (1) a common part that can be decoded at both receivers; and (2) a private part that is only decoded at the intended receiver. That is, only part of the other transmitter message is decoded. Partial decoding provides a means of controlling, at least partially, the interference. The capacity region of the two-user GIC is at most one bit away from the achievable region described in [24]. That is, the capacity region is approximated to within one bit [31]. However, the aforementioned strategies do not allow users to work together to deal with offending interference. To obtain further performance gains, intelligent cooperation among users to control interference is required. How to carry out this cooperation is therefore an important question and forms the basic question addressed in this thesis.

One way to achieve cooperation is through channel-output feedback. Channel-output feedback is an interference management technique that aims to improve the reliability and the performance of a communication network. From a general perspective, channel-output feedback enables a transmitter in a wireless network to observe the channel-output at its intended receiver. This allows the transmitters to exploit a coding strategy to control the interference, namely use interference as side information, and at the same time to benefit from the broadcast nature of the wireless channel making use of all possible links, establishing new paths for the communication.

Perfect observation of the channel-output at the intended receiver by each one of the corresponding transmitters is studied in [88]. The achievability scheme presented in [88] is based upon: rate-splitting [23, 37], block Markov superposition coding [14, 27], and backward decoding [98, 99]. The capacity region of the two-user GIC with perfect channel-output feedback (GIC-POF) is at most two bits from the achievable region. One of the most important observations made in [88] is that there exists a multiplicative gain in the capacity in certain interference regimes, particularly when both transmitter-receiver pairs are in the very strong interference regime. The next step towards a more general model was to consider the effect of the noise in the feedback links of a two-user symmetric GIC [53]. The results on the interference channel with generalized feedback (IC-GF) in [94, 102] are applied to obtain an achievable region in this channel model. The capacity region of the two-user symmetric GIC with noisy channel-ouput feedback is at most 4.7 bits away from the achievable region. The results provide a means of identifying certain values of the signal-to-noise ratios (SNRs) in the feedback links beyond which the capacity region can be enlarged with respect to the case without feedback. An important observation from these results is that the benefits of feedback are limited by noise in the feedback links.

The benefits of channel-output feedback in communication systems have been also observed in other network topologies. More specifically, the effect of feedback in the multiple access channel (MAC) has been studied in [13, 27, 33, 50, 61, 91, 97] and references therein; in the broadcast channel (BC) in [13, 17, 18, 29, 36, 62, 87, 96, 100, 101] and references therein; in the relay channel (RC) in [21, 26, 34] and references therein; and in the wiretap channel (WC) in [5]. Channel-output feedback has been also shown to be beneficial in the simultaneous transmission of both information and energy in the MAC [12] as well as in the IC [42, 43].

From the perspective of decentralized networks, very little is known about the benefits of feedback. Some works highlighting these benefits in the MAC are described in [11] and in the IC in [65, 66, 67, 68]. The case of decentralized communications systems without feedback is a bit better understood [38, 51]. For instance, the NEs of games arising in the MAC are described in [9, 10, 56, 64] and in the IC are described in [16, 76, 78].

This thesis considers the two-user asymmetric GIC-NOF. The analysis is performed considering two general scenarios: (1) centralized, in which the entire network is controlled by a central entity that configures both transmitter-receiver pairs; and (2) decentralized, in which each transmitter-receiver pair autonomously configures their transmission-reception parameters. The analysis in these two scenarios provides the characterization of the approximate capacity region and the approximate η-Nash equilibrium (η-NE) region of the two-user GIC-NOF. These results also provide the identification of the scenarios and the conditions in which one feedback link can enlarge the capacity region and the equilibrium region, respectively.

1.1. Motivation

This thesis focuses on the case of the GIC with NOF (GIC-NOF). The analysis of channel-output feedback in the IC has been fueled by the significant improvement it gives to the number of generalized degrees of freedom (GDoF) [40] with respect to the case without feedback. In particular, one of the main benefits of feedback is that the number of GDoF with perfect feedback increases monotonically with the interference-to-noise ratio (INR) in the very strong interference regime [88]. However, in the presence of additive Gaussian noise in the feedback links, the number of GDoF is bounded [53]. A significant improvement of the Nash equilibrium (NE) region of the Gaussian IC is also observed in the decentralized IC [66], *i.e.*, the case in which the transmitter-receiver pairs autonomously choose their own transmit-receive configurations to achieve their best data transmission rate. More specifically, the NE region is enlarged with respect to the case in which feedback is not available.

The GDoF gain due to feedback in the IC depends on the topology of the network and the number of transmitter-receiver pairs in the network. In the symmetric K-user cyclic Z-interference channel, the GDoF gain does not increase with K [90]. In particular, in the very strong interference regime, the GDoF gain is shown to be monotonically decreasing with K. In the fully connected symmetric K-user IC with perfect feedback, the number of GDoF per user is shown to be identical to the one in the two-user case, with an exception in a particular singularity, and totally independent of the exact number of transmitter-receiver pairs [57]. It is important to highlight that the network topology, the number of transmitter-receiver pairs, and the interference regimes are not the only parameters determining the effect of feedback. Indeed, the presence of noise in the feedback links turns out to be another relevant factor.

The main motivation to study the two-user GIC-NOF is to analyze the effect of the noise in the feedback links on the capacity region and the NE region of the two-user GIC-NOF under asymmetric conditions. This implies the identification of the scenarios in which the capacity region and the NE region can be enlarged by the use of one noisy feedback link and how the feedback parameters are related to the parameters of the GIC.

1.2. Contributions

The following are the main contributions of this thesis:

- A full characterization of the capacity region of the two-user LDIC-NOF [70, 72]. This contribution generalizes the results for the cases of the LDIC without feedback [20], with perfect channel-output feedback (LDIC-POF) [88], with noisy channel-output feedback (LDIC-NOF) under symmetric conditions [53], and the cases involving channel-output feedback from the intended receivers to the corresponding transmitters [80].

- An achievable region and a converse region for the two-user GIC-NOF [69, 70]. These two regions approximate the capacity region of the two-user GIC-NOF within 4.4 bits. The achievable region is obtained using a random coding argument combining rate-splitting [23, 37], block Markov superposition coding [14, 27], and backward decoding [98, 99], as first suggested in [88, 94, 102]. The converse region is obtained using some existing outer bounds from the case of the two-user GIC with POF (GIC-POF) [88] as well as a set of new outer bounds that are obtained by using genie-aided models. This contribution generalizes the results obtained for the cases without feedback (GIC) [31], with POF (GIC-POF) [88], and with NOF (GIC-NOF) under symmetric conditions [53].

- A full characterization of the η-NE region of the two-user LDIC-NOF [74]. This contribution generalizes the results for the cases of the linear deterministic interference channel (LDIC) without feedback [15], with POF (LDIC-POF) [66], and with NOF (LDIC-NOF) under symmetric conditions [68].

- An achievable η-NE region and a non-equilibrium region with $\eta \geqslant 1$ for the two-user GIC-NOF [73]. The achievable η-NE region is obtained introducing a modification of the achievability coding scheme considered in the centralized part. This modification implies the introduction of common randomness in the coding scheme as suggested in [16] and [66], which allows both transmitter-receiver pairs to limit the rate improvement of each other when either of them deviates from equilibrium. The non-equilibrium region with $\eta \geqslant 1$ is obtained using the insights from the analysis of the linear deterministic model.

- An identification of the scenarios in which the use of one feedback link enlarges the capacity region and the η-NE region [71].

1.3. Outlines

This thesis contains 5 parts as follows:

- **Part I**. This part describes the system model of the two-user continuous IC as well as the particular cases studied in this thesis: the two-user GIC-NOF and the two-user LDIC-NOF. It also establishes the differences between centralized and decentralized systems.
  - **Chapter 2**. This chapter presents the IC-NOF and more particularly, it describes the two-user GIC-NOF and the two-user LDIC-NOF as centralized systems. This chapter also presents the fundamental limits in both models in the cases without feedback, with perfect channel-output feedback (POF), and noisy channel-output feedback (NOF) under symmetric conditions.
  - **Chapter 3**. This chapter establishes the differences between the centralized and decentralized systems. It establishes a formulation of the game for the decentralized system. Finally, this chapter also presents the fundamental limits in the two-user GIC-NOF and the two-user LDIC-NOF as decentralized systems in the cases without feedback, with POF, and NOF under symmetric conditions.
  - **Chapter 3.4**. This chapter establishes the connections between the two-user GIC-NOF and the two-user LDIC-NOF.

- **Part II**. This part presents the main results derived from the analysis of the two-user LDIC-NOF and the two-user GIC-NOF considering a centralized control of the communication network.
  - **Chapter 4**. This chapter presents the main results for the two-user LDIC-NOF, *i.e.*, the capacity region, and analyzes the cases in which the capacity region can be enlarged by the use of feedback;

  - **Chapter 5**. This chapter presents the main results for the two-user GIC-NOF, *i.e.*, an achievable region, a converse region, the gap between both regions, and analyzes the cases in which the approximate capacity region might be enlarged.

- **Part III**. This part presents the main results derived from the analysis of the two-user LDIC-NOF and the two-user GIC-NOF considering a decentralized control of the communication network.
  - **Chapter 6**. This chapter presents the main results for the two-user D-LDIC-NOF, *i.e.*, the η-NE region, and analyzes the efficiency of the equilibrium region.
  - **Chapter 7**. This chapter presents the main results for the two-user D-GIC-NOF, *i.e.*, an achievable η-NE region and a non-equilibrium region with $\eta \geqslant 1$.

- **Part IV**. This part contains the conclusions of this thesis.

- **Part V**. This part contains fundamental concepts on information theory and network information theory that are used along this thesis and the proofs of the main results in parts II and III.
  - **Appendix A**. This appendix contains the description of the achievability scheme for the two-user LDIC-NOF and two-user GIC-NOF.
  - **Appendix B**. This appendix contains an outer bound for the two-user LDIC-NOF.
  - **Appendix C**. This appendix contains the calculation of the thresholds in the feedback parameters, beyond which the capacity region of the two-user LDIC-NOF can be enlarged with respect to the case without feedback. This calculation is made for the case in which both transmitter-receiver pairs are in very weak interference regime.
  - **Appendix D**. This appendix contains the calculation of the threshold in the feedback parameter i with $i \in \{1, 2\}$, beyond which the individual rate R_i can be improved in the two-user LDIC-NOF with respect to the case without feedback.
  - **Appendix E**. This appendix contains the calculation of the threshold in one feedback parameter, beyond which the sum-rate capacity can be improved in the two-user LDIC-NOF with respect to the case without feedback.
  - **Appendix F**. This appendix contains a proof of the number of GDoF for the two-user LDIC-NOF.
  - **Appendix G**. This appendix contains an outer bound for the two-user GIC-NOF.
  - **Appendix H**. This appendix contains the proof of the gap between the inner bound and the outer bound of the two-user GIC-NOF. The proof is for the case in which both transmitter-receiver pairs are in high interference regime (HIR). This appendix gives the values of the parameters of the coding scheme that must be considered in the other cases.
  - **Appendix I**. This appendix contains a proof of the η-Nash Equilibrium (NE) region for the two-user LDIC-NOF.
  - **Appendix J**. This appendix contains an inner bound on the η-NE region for the two-user GIC-NOF.

– **Appendix K**. This appendix contains a proof of the non-equilibrium region for the two-user GIC-NOF.

– **Appendix L**. This appendix contains a proof of a Lemma 21 in Appendix G.

– **Appendix M**. This appendix contains a proof of a Lemma I for the two-user LDIC-NOF in Appendix I.

– **Appendix N**. This appendix contains a proof of an inner bound of the η-Nash equilibrium (NE) region for the two-user GIC-NOF.

– **Appendix O**. This appendix presents the sum-rate capacity and the maximum and minimum sum-rate in the decentralized case for the two-user LDIC-NOF.

Part I.

INTERFERENCE CHANNELS

— 2 —

Centralized Interference Channels

Consider the two-user continuous IC-NOF in Figure 2.1. Transmitter i, $i \in \{1,2\}$, wishes to reliably communicate an independent and uniformly distributed message index $W_i \in \mathcal{W}_i = \{1, 2, \ldots, 2^{NR_i}\}$ to receiver i, during $N \in \mathbb{N}$ channel uses, where $R_i \in \mathbb{R}_+$ denotes the transmission rate of transmitter-receiver i in bits per channel use. In this respect, the transmitter i sends the codeword $\boldsymbol{X}_i = (X_{i,1}, X_{i,2}, \ldots, X_{i,N})^{\mathsf{T}} \in \mathcal{C}_i \subseteq \mathbb{R}^N$, where $\mathcal{C}_i$ is the codebook of transmitter i.

For a given channel use $n \in \{1, 2, \ldots, N\}$, the transmitters 1 and 2 send the channel inputs $X_{1,n} \in \mathbb{R}$ and $X_{2,n} \in \mathbb{R}$, respectively, which generate the channel-outputs $\overrightarrow{Y}_{1,n} \in \mathbb{R}$, $\overrightarrow{Y}_{2,n} \in \mathbb{R}$, $\overleftarrow{Y}_{1,n} \in \mathbb{R}$, and $\overleftarrow{Y}_{2,n} \in \mathbb{R}$ according to the conditional pdf $f_{\overrightarrow{Y}_1, \overrightarrow{Y}_2, \overleftarrow{Y}_1, \overleftarrow{Y}_2 | X_1, X_2}\left(\overrightarrow{y}_1, \overrightarrow{y}_2, \overleftarrow{y}_1, \overleftarrow{y}_2 | x_1, x_2\right)$, for all $(\overrightarrow{y}_1, \overrightarrow{y}_2, \overleftarrow{y}_1, \overleftarrow{y}_2, x_1, x_2) \in \mathbb{R}^6$.

The transmitter i generates the symbol $X_{i,n} \in \mathbb{R}$ considering the message index W_i and all previous outputs from the feedback link i, *i.e.*, $\left(\overleftarrow{Y}_{i,1}, \overleftarrow{Y}_{i,2}, \ldots, \overleftarrow{Y}_{i,n-1}\right)$. The transmitter i observes $\overleftarrow{Y}_{i,n}$ at the end of the channel use n. The transmitter i is defined by the set of deterministic functions $\{f_{i,1}, f_{i,2}, \ldots, f_{i,N}\}$, with $f_{i,1} : \mathcal{W}_i \to \mathbb{R}$ and for $n \in \{2, 3, \ldots, N\}$, $f_{i,n} : \mathcal{W}_i \times \mathbb{R}^{n-1} \to \mathbb{R}$, such that

$$X_{i,1} = f_{i,1}(W_i) \text{ and} \tag{2.1a}$$
$$X_{i,n} = f_{i,n}\left(W_i, \overleftarrow{Y}_{i,1}, \overleftarrow{Y}_{i,2}, \ldots, \overleftarrow{Y}_{i,n-1}\right) \text{ for all } n > 1. \tag{2.1b}$$

At the end of the transmission, the receiver i uses all the channel-outputs $\left(\overrightarrow{Y}_{i,1}, \overrightarrow{Y}_{i,2}, \ldots, \overrightarrow{Y}_{i,N}\right)$ to obtain an estimate of the message index W_i, denoted by $\widehat{W}_i$.

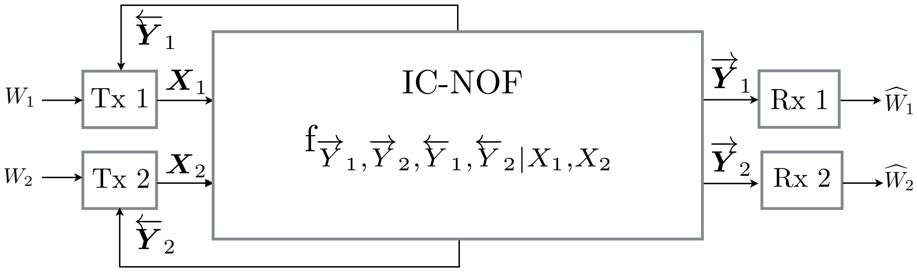


Figure 2.1.: Two-user continuous interference channel with noisy channel-output feedback.

Thus, the following Markov chain holds:

$$\left(W_i, \overleftarrow{\boldsymbol{Y}}_{i,(1;n-1)}\right) \rightarrow X_{i,n} \rightarrow \overrightarrow{Y}_{i,n}. \tag{2.2}$$

Let $T \in \mathbb{N}$ be fixed. Assume that during a communication, T blocks, each of N channel uses, are transmitted. The receiver i is defined by the deterministic function $\psi_i : \mathbb{R}^{NT} \rightarrow \mathcal{W}_i^T$. 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

$$\left(\widehat{W}_i^{(1)}, \widehat{W}_i^{(2)}, \ldots, \widehat{W}_i^{(T)}\right) = \psi_i \left(\overrightarrow{Y}_{i,1}, \overrightarrow{Y}_{i,2}, \ldots, \overrightarrow{Y}_{i,NT}\right), \tag{2.3}$$

where $\widehat{W}_i^{(t)}$ is an estimate of the message index $W_i^{(t)}$ sent during block $t \in \{1, 2, \ldots, T\}$.

The decoding error probability in the two-user continuous IC during the block t, denoted by $P_{\mathrm{e}}^{(t)}(N)$, is given by

$$P_{\mathrm{e}}^{(t)}(N) = \max\left(\Pr\left[\widehat{W_1}^{(t)} \neq W_1^{(t)}\right], \Pr\left[\widehat{W_2}^{(t)} \neq W_2^{(t)}\right]\right). \tag{2.4}$$

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

Definition 1 (Achievable Rate Pairs)**.** *A rate pair $(R_1, R_2) \in \mathbb{R}_+^2$ is achievable if there exist sets of encoding functions $\left\{f_1^{(1)}, f_1^{(2)}, \ldots, f_1^{(N)}\right\}$ and $\left\{f_2^{(1)}, f_2^{(2)}, \ldots, f_2^{(N)}\right\}$, and decoding functions ψ_1 and ψ_2, such that the error probability $P_{\mathrm{e}}^{(t)}(N)$ can be made arbitrarily small by letting the block-length N grow to infinity, for all blocks $t \in \{1, 2, \ldots, T\}$.*

In a centralized system, a central controller determines the configurations of all transmitter-receiver pairs. The central controller has a global view of the network and can select optimal configurations with respect to a given metric, *e.g.*, sum-rate, energy-efficiency, etc. The fundamental limits in a centralized system are characterized by the capacity region.

Definition 2 (Capacity region of a two-user IC)**.** *The capacity region of a two-user IC is the closure of the set of all possible achievable rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$.*

2.1. Gaussian Interference Channel

A special case of the IC-NOF described above from the perspective of centralized networks is the Gaussian IC-NOF. Consider the two-user GIC-NOF depicted in Figure 2.2. The channel

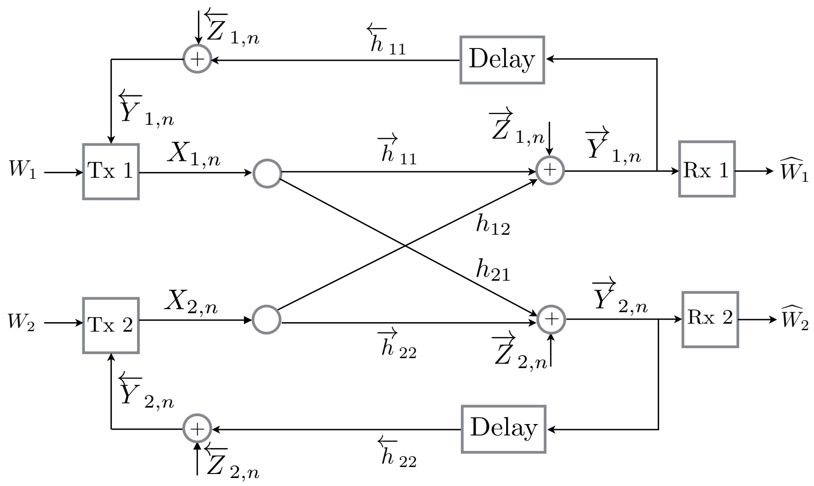


Figure 2.2.: Gaussian interference channel with noisy channel-output feedback at channel use n.

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. During channel use n, the input-output relations of the channel model are given for all $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$ by

$$\overrightarrow{Y}_{i,n}=\overrightarrow{h}_{ii}X_{i,n} + h_{ij}X_{j,n} + \overrightarrow{Z}_{i,n}, \tag{2.5}$$

and

$$\overleftarrow{Y}_{i,n}=\begin{cases}\overleftarrow{Z}_{i,n} & \text{for } n\in \{1,2,\ldots,d\}\\ \overleftarrow{h}_{ii}\overrightarrow{Y}_{i,n-d}+\overleftarrow{Z}_{i,n} & \text{for } n\in \{d+1,d+2,\ldots,N\}\end{cases}, \tag{2.6}$$

where $\overrightarrow{Z}_{i,n}$ and $\overleftarrow{Z}_{i,n}$ are independent real Gaussian random variables with zero mean and unit variance, and $d > 0$ is the finite feedback delay measured in channel uses.

In the remainder of this thesis, without loss of generality, the feedback delay is assumed to be one channel use, *i.e.*, $d = 1$. The components of the input vector $\boldsymbol{X}_i$ are real numbers subject to an average power constraint:

$$\frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left[X_{i,n}^2\right] \leq 1, \tag{2.7}$$

where the expectation is taken over the joint distribution of the message indices W_1 and W_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, and the previously observed noise realizations is due to the effect of feedback as shown in (2.1) and (2.6).

The two-user GIC-NOF in Figure 2.2 can be described by six parameters: $\overrightarrow{\mathrm{SNR}}_i$, $\overleftarrow{\mathrm{SNR}}_i$, and

INR_{ij}, with $i \in \{1,2\}$ and $j \in \{1,2\}\setminus\{i\}$, which are defined as follows:

$$\overrightarrow{\mathrm{SNR}}_i = \overrightarrow{h}_{ii}^{2}, \tag{2.8a}$$

$$\mathrm{INR}_{ij} = h_{ij}^{2}, \text{ and} \tag{2.8b}$$

$$\overleftarrow{\mathrm{SNR}}_i = \overleftarrow{h}_{ii}^{2}\left(\overrightarrow{h}_{ii}^{2} + 2\overrightarrow{h}_{ii}h_{ij} + h_{ij}^{2} + 1\right). \tag{2.8c}$$

When $\mathrm{INR}_{ij} \leqslant 1$, transmitter-receiver pair i is impaired mainly by noise instead of interference. In this case, treating interference as noise (TIN) is optimal and feedback does not bring a significant rate improvement. Therefore, the analysis developed in this thesis focuses exclusively on the case in which $\mathrm{INR}_{ij} > 1$ for all $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$.

In this special case, the pdf of the IC-NOF can be factorized as follows:

$$\mathrm{f}_{\overrightarrow{Y}_1,\overrightarrow{Y}_2,\overleftarrow{Y}_1,\overleftarrow{Y}_2|X_1,X_2} = \mathrm{f}_{\overrightarrow{Y}_1|X_1,X_2}\mathrm{f}_{\overrightarrow{Y}_2|X_1,X_2}\mathrm{f}_{\overleftarrow{Y}_1|\overrightarrow{Y}_1}\mathrm{f}_{\overleftarrow{Y}_2|\overrightarrow{Y}_2}, \tag{2.9}$$

given that for all $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$, $\overrightarrow{Y}_i$ is independent of $\overrightarrow{Y}_j$ conditioning on X_i and X_j; and $\overleftarrow{Y}_i$ is independent of X_i, X_j, and $\overrightarrow{Y}_j$ conditioning on $\overrightarrow{Y}_i$. Based on the input-output relation in (2.5), for all $i \in \{1,2\}$ and given the channel-inputs x_1 and x_2 during a specific channel use, the pdf $\mathrm{f}_{\overrightarrow{Y}_i|X_1,X_2}$ in (2.9) can be expressed as follows:

$$\mathrm{f}_{\overrightarrow{Y}_i|X_1,X_2}(\overrightarrow{y}_i|x_1,x_2) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\overrightarrow{y}_i - \overrightarrow{h}_{ii}x_i - h_{ij}x_j\right)^2\right). \tag{2.10}$$

Similarly, based on the input-output relation in (2.6), for all $i \in \{1,2\}$ and given the channel-outputs $\overrightarrow{y}_1$ and $\overrightarrow{y}_2$ during a specific channel use, the pdf $\mathrm{f}_{\overleftarrow{Y}_i|\overrightarrow{Y}_i}$ in (2.9) can be expressed as follows:

$$\mathrm{f}_{\overleftarrow{Y}_i|\overrightarrow{Y}_i}(\overleftarrow{y}_i|\overrightarrow{y}_i) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\overleftarrow{y}_i - \overleftarrow{h}_{ii}\overrightarrow{y}_{ii}\right)^2\right). \tag{2.11}$$

2.1.1. Case without Feedback

Assessing the capacity region of the two-user GIC is also a long-standing problem in network information theory. The capacity region is perfectly known in the very strong interference regime [22], which is the same capacity region of two non-interfering point to point links. In this case, the interference in both receivers is stronger than the intended signals and therefore the interference can be decoded and substracted from the received signals to decode the intended signals in each receiver (successive interference cancellation, SIC). The capacity region of the GIC is also known in the case of strong interference regime and it was independently obtained by [37] and [81]. The capacity region of the GIC for the case of strong interference regime in [81] is obtained considering that each receiver must decode both messages. Thus, each transmitter with both receivers can be seen as a multiple access channel (MAC) and the capacity region of the GIC under strong interference can be obtained as the intersection of the capacity regions of the two MACs [1]. This capacity region was initially introduced in [2]. This approach considered the joint decoding instead of sequential decoding as in [22].

In the other interference regimes, different strategies have been investigated, including considering partial decoding of the interference and TIN.

Fundamental results on the GIC are described in [23]. Particularly, two general coding schemes are presented. The first one is based on time division multiplexing and frequency division multiplexing (TDM/FDM), in which transmitter 1 and transmitter 2 use a fraction α and $1-\alpha$ of the bandwidth with powers P_1/α and $P_2/(1-\alpha)$, respectively. The second coding scheme is rate-splitting, in which transmitter $i \in \{1,2\}$ splits the message index $W_i \in \mathcal{W}_i = \{1,2,\ldots,2^{NR_i}\}$ into two message indices $W_{i,1} \in \mathcal{W}_{i,1} = \{1,2,\ldots,2^{NR_{i,1}}\}$ and $W_{i,2} \in \mathcal{W}_{i,2} = \{1,2,\ldots,2^{NR_{i,2}}\}$, with $R_i = R_{i,1} + R_{i,2}$. Transmitter i generates two codebooks with independent codewords to represent all message indices in $\mathcal{W}_{i,1}$ and $\mathcal{W}_{i,2}$. Transmitter i encodes the message index W_i summing the two independent codewords corresponding to the indices $W_{i,1}$ and $W_{i,2}$, *i.e.*, $\boldsymbol{x}_i = \boldsymbol{u}_i(W_{i,1}) + \boldsymbol{v}_i(W_{i,2})$, where $\boldsymbol{u}_i(W_{i,1})$ and $\boldsymbol{v}_i(W_{i,2})$ represent the corresponding codewords for the message indices $W_{i,1}$ and $W_{i,2}$ in transmitter i, respectively. The general idea is to decode the interfering signals in order to facilitate the decoding of the intended signals (this can be seen as a kind of cooperation), which can allow both transmitter-receiver pairs to achieve higher rates.

The best known achievable region for the two-GIC is given in [37]. This achievable region is simplified in [24]. The strategy in [37] uses rate-splitting [23], which implies dividing the transmitted information of both users into two parts: common information that can be decoded at both receivers and private information to be decoded only at the intended receiver. This strategy also implies to arbitrarily split the user signal power into the common and private parts of the message. In reception, this strategy uses joint typical decoding.

The following lemma presents the achievable region for the two-user GIC obtained in [37].

Lemma 1 (Han-Kobayashi Achievable Region for the two-user GIC)**.** *Let $\mathcal{C} \subset \mathbb{R}_+^2$ denote the capacity region of the two-user GIC. Then, $\mathcal{C}$ contains all the rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy the following inequalities:*

$$R_1 \le \sigma_1(\lambda_{1,P}, \lambda_{2,P}) + \frac{1}{2}\log\left(1 + \frac{\lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1}{1+\lambda_{2,P}\mathrm{INR}_{12}}\right), \tag{2.12a}$$

$$R_2 \le \sigma_2(\lambda_{1,P}, \lambda_{2,P}) + \frac{1}{2}\log\left(1 + \frac{\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2}{1+\lambda_{1,P}\mathrm{INR}_{21}}\right), \tag{2.12b}$$

$$R_1 + R_2 \le \sigma_0(\lambda_{1,P}, \lambda_{2,P}) + \frac{1}{2}\log\left(1 + \frac{\lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1}{1+\lambda_{2,P}\mathrm{INR}_{12}}\right) + \frac{1}{2}\log\left(1 + \frac{\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2}{1+\lambda_{1,P}\mathrm{INR}_{21}}\right), \tag{2.12c}$$

$$\begin{aligned}
2R_1 + R_2 \le\; & 2\sigma_1(\lambda_{1,P}, \lambda_{2,P}) + \log\left(1 + \frac{\lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1}{1+\lambda_{2,P}\mathrm{INR}_{12}}\right) + \frac{1}{2}\log\left(1 + \frac{\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2}{1+\lambda_{1,P}\mathrm{INR}_{21}}\right) \\
& - \left(\sigma_1(\lambda_{1,P}, \lambda_{2,P}) - \frac{1}{2}\log\left(1 + \frac{(1-\lambda_{1,P})\,\mathrm{INR}_{21}}{1+\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2 + \lambda_{1,P}\mathrm{INR}_{21}}\right)\right)^+ \\
& + \min\left(\frac{1}{2}\log\left(1 + \frac{(1-\lambda_{2,P})\,\overrightarrow{\mathrm{SNR}}_2}{1+\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2 + \lambda_{1,P}\mathrm{INR}_{21}}\right), \frac{1}{2}\log\left(1 + \frac{(1-\lambda_{2,P})\,\overrightarrow{\mathrm{SNR}}_2}{1+\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2 + \mathrm{INR}_{21}}\right)\right. \\
& + \left(\frac{1}{2}\log\left(1 + \frac{(1-\lambda_{1,P})\,\mathrm{INR}_{21}}{1+\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2 + \lambda_{1,P}\mathrm{INR}_{21}}\right) - \sigma_1(\lambda_{1,P}, \lambda_{2,P})\right)^+, \\
& \frac{1}{2}\log\left(1 + \frac{(1-\lambda_{2,P})\,\mathrm{INR}_{12}}{1+\lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1 + \lambda_{2,P}\mathrm{INR}_{12}}\right), \\
& \left.\frac{1}{2}\log\left(1 + \frac{(1-\lambda_{1,P})\,\overrightarrow{\mathrm{SNR}}_1 + (1-\lambda_{2,P})\,\mathrm{INR}_{12}}{1+\lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1 + \lambda_{2,P}\mathrm{INR}_{12}}\right) - \sigma_1(\lambda_{1,P}, \lambda_{2,P})\right),
\end{aligned} \tag{2.12d}$$

$$R_1+2R_2 \leq 2\sigma_2\left(\lambda_{1,P},\lambda_{2,P}\right)+\frac{1}{2}\log\left(1+\frac{\lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1}{1+\lambda_{2,P}\mathrm{INR}_{12}}\right)+\frac{1}{2}\log\left(1+\frac{\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2}{1+\lambda_{1,P}\mathrm{INR}_{21}}\right)$$
$$-\left(\sigma_2\left(\lambda_{1,P},\lambda_{2,P}\right)-\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{2,P}\right)\overrightarrow{\mathrm{SNR}}_2}{1+\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2+\lambda_{1,P}\mathrm{INR}_{21}}\right)\right)^{+}$$
$$+\min\left(\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{1,P}\right)\overrightarrow{\mathrm{SNR}}_1}{1+\lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1+\lambda_{2,P}\mathrm{INR}_{12}}\right),\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{1,P}\right)\overrightarrow{\mathrm{SNR}}_1}{1+\lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{12}}\right)\right.$$
$$+\left(\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{2,P}\right)\mathrm{INR}_{12}}{1+\lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1+\lambda_{2,P}\mathrm{INR}_{12}}\right)-\sigma_2\left(\lambda_{1,P},\lambda_{2,P}\right)\right)^{+},$$
$$\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{1,P}\right)\mathrm{INR}_{21}}{1+\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2+\lambda_{1,P}\mathrm{INR}_{21}}\right),\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{2,P}\right)\overrightarrow{\mathrm{SNR}}_2+\lambda_{1,C}\mathrm{INR}_{21}}{1+\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2+\lambda_{1,P}\mathrm{INR}_{21}}\right)$$
$$\left.-\sigma_2\left(\lambda_{1,P},\lambda_{2,P}\right)\right), \tag{2.12e}$$

with

$$\sigma_1\left(\lambda_{1,P},\lambda_{2,P}\right)=\min\left(\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{1,P}\right)\overrightarrow{\mathrm{SNR}}_1}{1+\lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1+\lambda_{2,P}\mathrm{INR}_{12}}\right),\right.$$
$$\left.\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{1,P}\right)\mathrm{INR}_{21}}{1+\lambda_{1,P}\mathrm{INR}_{21}}\right)\right), \tag{2.13a}$$

$$\sigma_2\left(\lambda_{1,P},\lambda_{2,P}\right)=\min\left(\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{2,P}\right)\overrightarrow{\mathrm{SNR}}_2}{1+\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2+\lambda_{1,P}\mathrm{INR}_{21}}\right),\right.$$
$$\left.\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{2,P}\right)\mathrm{INR}_{12}}{1+\lambda_{2,P}\mathrm{INR}_{12}}\right)\right), \tag{2.13b}$$

$$\sigma_0\left(\lambda_{1,P},\lambda_{2,P}\right)=\min\left(\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{1,P}\right)\overrightarrow{\mathrm{SNR}}_1+\left(1-\lambda_{2,P}\right)\mathrm{INR}_{12}}{1+\lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1+\lambda_{2,P}\mathrm{INR}_{12}}\right),\right.$$
$$\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{2,P}\right)\overrightarrow{\mathrm{SNR}}_2+\left(1-\lambda_{1,P}\right)\mathrm{INR}_{21}}{1+\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2+\lambda_{1,P}\mathrm{INR}_{21}}\right),$$
$$\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{2,P}\right)\mathrm{INR}_{12}}{1+\lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1+\lambda_{2,P}\mathrm{INR}_{12}}\right)$$
$$+\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{2,P}\right)\overrightarrow{\mathrm{SNR}}_2}{1+\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2+\lambda_{1,P}\mathrm{INR}_{21}}\right),$$
$$\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{1,P}\right)\mathrm{INR}_{21}}{1+\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2+\lambda_{1,P}\mathrm{INR}_{21}}\right)$$
$$\left.+\frac{1}{2}\log\left(1+\frac{\left(1-\lambda_{2,P}\right)\overrightarrow{\mathrm{SNR}}_2}{1+\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2+\lambda_{1,P}\mathrm{INR}_{21}}\right)\right), \tag{2.13c}$$

with $\lambda_{i,P} \in [0,1]$ *for all* $i \in \{1,2\}$.

The strategy in [24] uses rate-splitting [23, 37] and superposition coding [27]. The superposition coding is a technique that was introduced in the study of the broadcast channel (BC) in [25]. Consider the message index sent by transmitter i denoted by $W_i \in \{1, 2, \ldots, 2^{NR_i}\}$. Following a rate-splitting argument, assume that W_i is represented by two sub-indices $(W_{i,C}, W_{i,P}) \in \{1, 2, \ldots, 2^{NR_{i,C}}\} \times \{1, 2, \ldots, 2^{NR_{i,P}}\}$, where $R_{i,C} + R_{i,P} = R_i$. The message index $W_{i,C}$ is assumed to be decoded at both receivers (common part of the message) and the message index $W_{i,P}$ is assumed to be decoded at the intended receiver (private part of the message) at the end of the transmission. Using the index $W_{i,C}$, transmitter i identifies a codeword in the first code-layer. The first code-layer is a sub-codebook of $2^{NR_{i,C}}$ codewords (cloud centers). Denote by $\boldsymbol{u}_i\left(W_{i,C}\right)$ the corresponding codeword in the first code-layer. The second codeword used by transmitter i is selected using $W_{i,P}$ from the second code-layer, which is a sub-codebook of $2^{NR_{i,P}}$ codewords corresponding to $\boldsymbol{u_i}\left(W_{i,C}\right)$. Denote by $\boldsymbol{x}_i\left(W_{i,C}, W_{i,P}\right)$ the corresponding codeword in the second code-layer. Finally, transmitter i sends the codeword $\boldsymbol{x}_i\left(W_{i,C}, W_{i,P}\right)$. The simplification of the Han-Kobayashi achievable region for the two-user IC in [24] is due to an observation of the authors in which each receiver is not interested in decoding the common message index coming from the non-corresponding transmitter. The consideration of decoding in each receiver the common message index coming from the non-corresponding transmitter in [37] generated a pair of inequalities in the evaluation of the error probability that are not necessary, which is proved in [44].

The following lemma presents the achievable region for the two-user GIC in [43] obtained from the results in [24].

Lemma 2 (Chong-Motani-Garg-El Gamal Achievable Region for the two-user GIC)**.** *Let $\mathcal{C} \subset \mathbb{R}_+^2$ denote the capacity region of the two-user GIC. Then, $\mathcal{C}$ contains all the rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy the following inequalities:*

$$R_1 \leq \frac{1}{2}\log\left(1 + \frac{\overrightarrow{\mathrm{SNR}}_1}{1 + \lambda_{2,P}\mathrm{INR}_{12}}\right), \tag{2.14a}$$

$$R_2 \leq \frac{1}{2}\log\left(1 + \frac{\overrightarrow{\mathrm{SNR}}_2}{1 + \lambda_{1,P}\mathrm{INR}_{21}}\right), \tag{2.14b}$$

$$R_1 + R_2 \leq \frac{1}{2}\log\left(\frac{1 + \overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12}}{1 + \lambda_{2,P}\mathrm{INR}_{12}}\right) + \frac{1}{2}\log\left(1 + \frac{\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2}{1 + \lambda_{1,P}\mathrm{INR}_{21}}\right), \tag{2.14c}$$

$$R_1 + R_2 \leq \frac{1}{2}\log\left(\frac{1 + \overrightarrow{\mathrm{SNR}}_2 + \mathrm{INR}_{21}}{1 + \lambda_{1,P}\mathrm{INR}_{21}}\right) + \frac{1}{2}\log\left(1 + \frac{\lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1}{1 + \lambda_{2,P}\mathrm{INR}_{12}}\right), \tag{2.14d}$$

$$R_1 + R_2 \leq \frac{1}{2}\log\left(\frac{1 + \lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12}}{1 + \lambda_{2,P}\mathrm{INR}_{12}}\right) + \frac{1}{2}\log\left(\frac{1 + \lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2 + \mathrm{INR}_{21}}{1 + \lambda_{1,P}\mathrm{INR}_{21}}\right), \tag{2.14e}$$

$$\begin{aligned} 2R_1 + R_2 \leq & \frac{1}{2}\log\left(\frac{1 + \overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12}}{1 + \lambda_{2,P}\mathrm{INR}_{12}}\right) + \frac{1}{2}\log\left(\frac{1 + \lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2 + \mathrm{INR}_{21}}{1 + \lambda_{1,P}\mathrm{INR}_{21}}\right) \\ & + \frac{1}{2}\log\left(1 + \frac{\lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1}{1 + \lambda_{2,P}\mathrm{INR}_{12}}\right), \end{aligned} \tag{2.14f}$$

$$R_1 + 2R_2 \leq \frac{1}{2}\log\left(\frac{1+\overrightarrow{\mathrm{SNR}}_2+\mathrm{INR}_{21}}{1+\lambda_{1,P}\mathrm{INR}_{21}}\right)+\frac{1}{2}\log\left(\frac{1+\lambda_{1,P}\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{12}}{1+\lambda_{2,P}\mathrm{INR}_{12}}\right)$$
$$+\frac{1}{2}\log\left(1+\frac{\lambda_{2,P}\overrightarrow{\mathrm{SNR}}_2}{1+\lambda_{1,P}\mathrm{INR}_{21}}\right), \tag{2.14g}$$

with $\lambda_{i,P} \in [0,1]$ *for all* $i \in \{1,2\}$.

There are several outer bounds on the capacity region of the GIC [3, 31, 46, 58, 82, 83]. Some outer bounds correspond to the capacity region of other network models that are seen as simplified models of the GIC under certain conditions. Some other outer bounds are obtained based on genie-aided models. Some of these outer bounds provide the sum-rate capacity or at least some corner points of the capacity region for specific conditions in the GIC. In the cases in which both transmitter-receiver pairs are in low-interference regime and the interference parameters are below certain thresholds, TIN achieves the sum-capacity of the GIC (this is also denominated the noisy interference regime) [3, 58, 83].

The authors in [31] obtained an outer bound based on genie-aided models which was used to prove that the achievable region in [24] (Lemma 2) is at most one bit per channel use away from the capacity region of the two-user GIC. Note that the authors in [31] assumed $\lambda_{i,P} = \frac{1}{\mathrm{INR}_{ji}}$ for all $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$ in the achievable region introduced in [24], considering that the private part of a message has not to be decoded in the non-intended receiver because it can be under the noise level. The authors in [31] considered three different interference regimes: weak interference channel ($\mathrm{INR}_{12} < \overrightarrow{\mathrm{SNR}}_2$ and $\mathrm{INR}_{21} < \overrightarrow{\mathrm{SNR}}_1$); mixed interference channel ($\mathrm{INR}_{12} \geqslant \overrightarrow{\mathrm{SNR}}_2$ and $\mathrm{INR}_{21} < \overrightarrow{\mathrm{SNR}}_1$, or $\mathrm{INR}_{12} < \overrightarrow{\mathrm{SNR}}_2$ and $\mathrm{INR}_{21} \geqslant \overrightarrow{\mathrm{SNR}}_1$); and strong interference channel ($\mathrm{INR}_{12} \geqslant \overrightarrow{\mathrm{SNR}}_2$ and $\mathrm{INR}_{21} \geqslant \overrightarrow{\mathrm{SNR}}_1$), where the outer bound for the last interference regime is not shown given that the capacity region is already known [22, 37, 81]. The following two lemmas present the outer bounds on the capacity region of the two-user GIC for the weak interference channel and for the mixed interference channel.

Lemma 3 (Outer bound for weak GIC [31, Theorem 3])**.** *Let* $\mathcal{C} \subset \mathbb{R}_+^2$ *denote the capacity region of the GIC. Then,* $\mathcal{C}$ *is contained within the set of rate pairs* $(R_1, R_2) \in \mathbb{R}_+^2$ *that satisfy the following inequalities:*

$$R_1 \leqslant \frac{1}{2}\log\left(1+\overrightarrow{\mathrm{SNR}}_1\right), \tag{2.15a}$$

$$R_2 \leqslant \frac{1}{2}\log\left(1+\overrightarrow{\mathrm{SNR}}_2\right), \tag{2.15b}$$

$$R_1 + R_2 \leqslant \frac{1}{2}\log\left(1+\overrightarrow{\mathrm{SNR}}_1\right)+\frac{1}{2}\log\left(1+\frac{\overrightarrow{\mathrm{SNR}}_2}{1+\mathrm{INR}_{21}}\right), \tag{2.15c}$$

$$R_1 + R_2 \leqslant \frac{1}{2}\log\left(1+\overrightarrow{\mathrm{SNR}}_2\right)+\frac{1}{2}\log\left(1+\frac{\overrightarrow{\mathrm{SNR}}_1}{1+\mathrm{INR}_{12}}\right), \tag{2.15d}$$

$$R_1 + R_2 \leqslant \frac{1}{2}\log\left(1+\mathrm{INR}_{12}+\frac{\overrightarrow{\mathrm{SNR}}_1}{1+\mathrm{INR}_{21}}\right)+\frac{1}{2}\log\left(1+\mathrm{INR}_{21}+\frac{\overrightarrow{\mathrm{SNR}}_2}{1+\mathrm{INR}_{12}}\right), \tag{2.15e}$$

$$2R_1 + R_2 \leqslant \frac{1}{2}\log\left(1 + \overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12}\right) + \frac{1}{2}\log\left(1 + \mathrm{INR}_{21} + \frac{\overrightarrow{\mathrm{SNR}}_2}{1 + \mathrm{INR}_{12}}\right)$$
$$+\frac{1}{2}\log\left(\frac{1 + \overrightarrow{\mathrm{SNR}}_1}{1 + \mathrm{INR}_{21}}\right), \quad (2.15f)$$
$$R_1 + 2R_2 \leqslant \frac{1}{2}\log\left(1 + \overrightarrow{\mathrm{SNR}}_2 + \mathrm{INR}_{21}\right) + \frac{1}{2}\log\left(1 + \mathrm{INR}_{12} + \frac{\overrightarrow{\mathrm{SNR}}_1}{1 + \mathrm{INR}_{21}}\right)$$
$$+\frac{1}{2}\log\left(\frac{1 + \overrightarrow{\mathrm{SNR}}_2}{1 + \mathrm{INR}_{12}}\right). \quad (2.15g)$$

Lemma 4 (Outer bound for mixed GIC [31, Theorem 4])**.** *Let $\mathcal{C} \subset \mathbb{R}_+^2$ denote the capacity region of the GIC. Then, $\mathcal{C}$ is contained within the set of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy the following inequalities:*

$$R_1 \leqslant \frac{1}{2}\log\left(1 + \overrightarrow{\mathrm{SNR}}_1\right), \quad (2.16a)$$
$$R_2 \leqslant \frac{1}{2}\log\left(1 + \overrightarrow{\mathrm{SNR}}_2\right), \quad (2.16b)$$
$$R_1 + R_2 \leqslant \frac{1}{2}\log\left(1 + \overrightarrow{\mathrm{SNR}}_1\right) + \frac{1}{2}\log\left(1 + \frac{\overrightarrow{\mathrm{SNR}}_2}{1 + \mathrm{INR}_{21}}\right), \quad (2.16c)$$
$$R_1 + R_2 \leqslant \frac{1}{2}\log\left(1 + \overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12}\right), \quad (2.16d)$$
$$R_1 + 2R_2 \leqslant \frac{1}{2}\log\left(1 + \overrightarrow{\mathrm{SNR}}_2 + \mathrm{INR}_{21}\right) + \frac{1}{2}\log\left(1 + \mathrm{INR}_{12} + \frac{\overrightarrow{\mathrm{SNR}}_1}{1 + \mathrm{INR}_{21}}\right) \quad (2.16e)$$
$$+\frac{1}{2}\log\left(1 + \frac{\overrightarrow{\mathrm{SNR}}_2}{1 + \mathrm{INR}_{12}}\right). \quad (2.16f)$$

2.1.2. Case with Perfect Channel-Output Feedback

The two-user GIC-POF is analyzed in [88], and its capacity region is characterized to within two bits per channel use. The achievability scheme presented in [88] is based upon: rate-splitting [23, 37], block Markov superposition coding [14, 27], and backward decoding [98, 99]. The outer bound is obtained considering genie-aided models. One of the most important conclusions in [88] is that feedback can provide an arbitrary multiplicative gain in the high SNR regime for certain channel conditions in the two-user GIC, *i.e.*, the very strong interference regime.

The following two lemmas present an inner bound and an outer bound on the capacity region of the two-user GIC-POF.

Lemma 5 (Inner bound two-user GIC-POF [88, Theorem 2])**.** *Let $\mathcal{C} \subset \mathbb{R}_+^2$ denote the capacity region of the two-user GIC-POF. Then, $\mathcal{C}$ contains all the rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy the following inequalities:*

$$R_1 \leqslant \frac{1}{2}\log\left(1 + \overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}}\right) - \frac{1}{2}, \quad (2.17a)$$
$$R_1 \leqslant \frac{1}{2}\log\left(1 + (1-\rho)\,\mathrm{INR}_{21}\right) + \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}}\right) - 1, \quad (2.17b)$$

$$R_2 \leqslant \frac{1}{2}\log\left(1+\overrightarrow{\mathrm{SNR}}_2+\mathrm{INR}_{21}+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}\right)-\frac{1}{2}, \tag{2.17c}$$

$$R_2 \leqslant \frac{1}{2}\log\left(1+(1-\rho)\,\mathrm{INR}_{12}\right)+\frac{1}{2}\log\left(2+\frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}\right)-1, \tag{2.17d}$$

$$R_1+R_2 \leqslant \frac{1}{2}\log\left(2+\frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}}\right)+\frac{1}{2}\log\left(1+\overrightarrow{\mathrm{SNR}}_2+\mathrm{INR}_{21}+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}\right)-1, \tag{2.17e}$$

$$R_1+R_2 \leqslant \frac{1}{2}\log\left(2+\frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}\right)+\frac{1}{2}\log\left(1+\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{12}+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}}\right)-1, \tag{2.17f}$$

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

Lemma 6 (Outer bound two-user GIC-POF, [88, Theorem 3])**.** *Let* $\mathcal{C} \subset \mathbb{R}_+^2$ *denote the capacity region of the GIC-POF. Then,* $\mathcal{C}$ *is contained within the set of rate pairs* $(R_1, R_2) \in \mathbb{R}_+^2$ *that satisfy the following inequalities:*

$$R_1 \leqslant \frac{1}{2}\log\left(1+\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{12}+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}}\right), \tag{2.18a}$$

$$R_1 \leqslant \frac{1}{2}\log\left(1+(1-\rho)\,\mathrm{INR}_{21}\right)+\frac{1}{2}\log\left(1+\frac{(1-\rho^2)\,\overrightarrow{\mathrm{SNR}}_1}{1+(1-\rho^2)\,\mathrm{INR}_{21}}\right), \tag{2.18b}$$

$$R_2 \leqslant \frac{1}{2}\log\left(1+\overrightarrow{\mathrm{SNR}}_2+\mathrm{INR}_{21}+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}\right), \tag{2.18c}$$

$$R_2 \leqslant \frac{1}{2}\log\left(1+(1-\rho)\,\mathrm{INR}_{12}\right)+\frac{1}{2}\log\left(1+\frac{(1-\rho^2)\,\overrightarrow{\mathrm{SNR}}_2}{1+(1-\rho^2)\,\mathrm{INR}_{12}}\right), \tag{2.18d}$$

$$\begin{aligned} R_1+R_2 \leqslant & \frac{1}{2}\log\left(1+\frac{(1-\rho^2)\,\overrightarrow{\mathrm{SNR}}_1}{1+(1-\rho^2)\,\mathrm{INR}_{21}}\right) \\ & +\frac{1}{2}\log\left(1+\overrightarrow{\mathrm{SNR}}_2+\mathrm{INR}_{21}+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}\right), \end{aligned} \tag{2.18e}$$

$$\begin{aligned} R_1+R_2 \leqslant & \frac{1}{2}\log\left(1+\frac{(1-\rho^2)\,\overrightarrow{\mathrm{SNR}}_2}{1+(1-\rho^2)\,\mathrm{INR}_{12}}\right) \\ & +\frac{1}{2}\log\left(1+\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{12}+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}}\right), \end{aligned} \tag{2.18f}$$

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

In [57], it is shown that the number of GDoF of a symmetric fully connected K-user GIC with POF is the same as in the case of the two-user IC-POF, except for the case in which the power of the signal in each receiver is equal to the power of the interfering signal. Then, feedback can improve the performance of the networks except under the aforementioned condition. The coding scheme takes advantage of the network symmetry and is based on interference alignment and interference decoding. Thus, given the alignment of the interference (it is necessary to decode the interference to remove it, which is suppressed in standard approaches), the interference received from all other users can be seen as a single message using a lattice code approach. In [90], an approximate capacity region of the cyclic K-user GIC is presented. The network involves K-users where each intended signal is only interfered by one of the neighboring transmitters in a cyclic fashion. It is shown that the number of GDoF of a cyclic symmetric K-user symmetric GIC with POF is a function of K, *i.e.*, the capacity gain for each user is inversely proportional to the number of users K. Thus, the improvement in the capacity per user of a cyclic and symmetric K-user GIC vanishes as K grows, and when K tends to infinity the number of GDoF with feedback is equal to the number of GDoF without feedback. It is worth noting that the GDoF of the symmetric and cyclic K-user without feedback are the same as for the two-user GIC [105]. Other feedback coding schemes for K-user Gaussian interference networks have been analyzed in [48, 47].

In [80] the impact of nine different POF architectures are studied for the symmetric LDIC and the symmetric GIC. The exact capacity region is obtained for the linear deterministic model and an approximate capacity region is obtained for the Gaussian case in which the capacity region is approximated to within 4.59 bits from the inner-bound. The authors proposed two achievable strategies: one based on rate-splitting [23, 37] and the other one based on block-Markov coding (at one transmitter) and dirty paper coding at the other transmitter. The authors also proposed two new outer bounds that are tighter than the cut-set bound in some interference regimes.

The authors in [63] presented an inner bound and an outer bound on the sum-capacity of a symmetric IC with source cooperation (IC-CT). The inner bound is obtained using block-Markov superposition coding [14, 27], backward decoding [98, 99], and a decode-and-forward strategy. The coding scheme splits the message index into four message indices, considering that common and private messages can be split into cooperative and non-cooperative. The outer bound is shown to be at most 20 bits away from the sum-rate capacity. Even though the IC-NOF is a model that differs from the symmetric IC-CT, there exists a connection between these two models. In this sense, the authors in [63] show that using their results on the symmetric IC-CT, the sum-capacity of the two-user symmetric GIC-POF is approximated to within a constant gap of 19 bits.

2.1.3. Symmetric Case with Noisy Channel-Output Feedback

The two-user symmetric GIC-NOF is analyzed in [35, 53, 52], and its capacity region is characterized to within 4.7 bits per channel use in [53]. The achievability scheme in [53] is a particular case of a more general achievability scheme presented in [94, 102]. An outer bound using the Hekstra-Willems dependence-balance arguments [39] has been introduced in [35]. In the GIC, these results suggest that feedback loses its efficacy on increasing the capacity region roughly when the noise variance on the feedback link is larger than on the forward link. Similar results have been reported in the fully decentralized IC with NOF [68, 74, 106]. More general channel models, for instance when channel-outputs are fed back to both receivers, have been studied and inner and outer bounds are presented in [48, 93, 92, 80]. Despite the fact that the capacity region was approximated, very little can be concluded in the case in which feedback is available in only one of the point-to-point links or simply when the point-to-point links are in different interference regimes. The results on the interference channel with generalized feedback (IC-GF) in [94, 102] are applied to obtain an inner bound in this channel model. The outer bound is derived using genie-aided models thanks to insights from the analysis of the corresponding linear deterministic model.

The following two lemmas present an inner bound and an outer bound on the capacity region of the two-user symmetric GIC-NOF.

Lemma 7 (Inner bound two-user symmetric GIC-NOF [53, Theorem 3])**.** *Let $\mathcal{C} \subset \mathbb{R}_+^2$ denote the capacity region of the two-user symmetric GIC-NOF. Then, $\mathcal{C}$ contains all the rate pairs*

$(R_1, R_2) \in \mathbb{R}_+^2$ *that satisfy the following inequalities:*

$$R_1 \leqslant \min\Big(\tau_6(\rho, \lambda_P), \tau_4(\lambda_{NC}, \lambda_P) + \tau_1(\lambda_{CC}, \lambda_{NC}, \lambda_P), \tau_1(\lambda_{CC}, \lambda_{NC}, \lambda_P) + \tau_2(\lambda_P) + \tau_3(\lambda_{NC}, \lambda_P)\Big), \quad (2.19a)$$

$$R_2 \leqslant \min\Big((\tau_6(\rho, \lambda_P), \tau_4(\lambda_{NC}, \lambda_P) + \tau_1(\lambda_{CC}, \lambda_{NC}, \lambda_P), \tau_1(\lambda_{CC}, \lambda_{NC}, \lambda_P) + \tau_2(\lambda_P) + \tau_3(\lambda_{NC}, \lambda_P)\Big), \quad (2.19b)$$

$$R_1 + R_2 \leqslant \min\Big(\tau_2(\lambda_P) + \tau_6(\rho, \lambda_P), 2\tau_1(\lambda_{CC}, \lambda_{NC}, \lambda_P) + \tau_5(\lambda_{NC}, \lambda_P) + \tau_2(\lambda_P), 2\tau_1(\lambda_{CC}, \lambda_{NC}, \lambda_P) + 2\tau_3(\lambda_{NC}, \lambda_P)\Big), \quad (2.19c)$$

$$2R_1 + R_2 \leqslant \min\Big(\tau_6(\rho, \lambda_P) + \tau_1(\lambda_{CC}, \lambda_{NC}, \lambda_P) + \tau_2(\lambda_P) + \tau_3(\lambda_{NC}, \lambda_P), 3\tau_1(\lambda_{CC}, \lambda_{NC}, \lambda_P) + \tau_2(\lambda_P) + \tau_3(\lambda_{NC}, \lambda_P) + \tau_5(\lambda_{NC}, \lambda_P)\Big), \quad (2.19d)$$

$$R_1 + 2R_2 \leqslant \min\Big(\tau_6(\rho, \lambda_P) + \tau_1(\lambda_{CC}, \lambda_{NC}, \lambda_P) + \tau_2(\lambda_P) + \tau_3(\lambda_{NC}, \lambda_P), 3\tau_1(\lambda_{CC}, \lambda_{NC}, \lambda_P) + \tau_2(\lambda_P) + \tau_3(\lambda_{NC}, \lambda_P) + \tau_5(\lambda_{NC}, \lambda_P)\Big), \quad (2.19e)$$

with

$$\tau_6(\rho, \lambda_P) \triangleq \frac{1}{2}\log\left(\frac{1 + \overrightarrow{\mathrm{SNR}} + \mathrm{INR} + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}\mathrm{INR}}}{\lambda_P \mathrm{INR} + 1}\right), \quad (2.20a)$$

$$\tau_5(\lambda_{NC}, \lambda_P) \triangleq \frac{1}{2}\log\left(\frac{(\lambda_{NC} + \lambda_P)\mathrm{SNR} + (\lambda_{NC} + \lambda_P)\mathrm{INR} + 1}{\lambda_P \mathrm{INR} + 1}\right), \quad (2.20b)$$

$$\tau_4(\lambda_{NC}, \lambda_P) \triangleq \frac{1}{2}\log\left(\frac{(\lambda_{NC} + \lambda_P)\mathrm{SNR} + \lambda_P \mathrm{INR} + 1}{\lambda_P \mathrm{INR} + 1}\right), \quad (2.20c)$$

$$\tau_3(\lambda_{NC}, \lambda_P) \triangleq \frac{1}{2}\log\left(\frac{\lambda_P \mathrm{SNR} + (\lambda_{NC} + \lambda_P)\mathrm{INR} + 1}{\lambda_P \mathrm{INR} + 1}\right), \quad (2.20d)$$

$$\tau_2(\lambda_P) \triangleq \frac{1}{2}\log\left(\frac{\lambda_P \mathrm{SNR} + \lambda_P \mathrm{INR} + 1}{\lambda_P \mathrm{INR} + 1}\right), \quad (2.20e)$$

$$\tau_1(\lambda_{CC}, \lambda_{NC}, \lambda_P) \triangleq \frac{1}{2}\log\left(\frac{\tau_{1n}(\lambda_{CC}, \lambda_{NC}, \lambda_P)}{\tau_{1d}(\lambda_{NC}, \lambda_P)}\right), \quad (2.20f)$$

$$\tau_{1n}(\lambda_{CC}, \lambda_{NC}, \lambda_P) \triangleq \frac{1}{2}\log\left(\frac{\overleftarrow{\mathrm{SNR}}((\lambda_{CC} + \lambda_{NC} + \lambda_P)\mathrm{INR} + 1)}{1 + \overrightarrow{\mathrm{SNR}} + \mathrm{INR} + 2\sqrt{\overrightarrow{\mathrm{SNR}}\mathrm{INR}}}\right) + 1, \quad (2.20g)$$

$$\tau_{1d}(\lambda_{NC}, \lambda_P) \triangleq \frac{1}{2}\log\left(\frac{\overleftarrow{\mathrm{SNR}}((\lambda_{NC} + \lambda_P)\mathrm{INR} + 1)}{1 + \overrightarrow{\mathrm{SNR}} + \mathrm{INR} + 2\sqrt{\overrightarrow{\mathrm{SNR}}\mathrm{INR}}}\right) + 1, \quad (2.20h)$$

$\rho \in [0, 1]$ *and for all coding schemes that satisfy* $\lambda_{CC} + \lambda_{NC} + \lambda_P = 1 - \rho$.

Lemma 8 (Outer bound two-user symmetric GIC-NOF [53, Theorem 2])**.** *Let* $\mathcal{C} \subset \mathbb{R}_+^2$ *denote the capacity region of the two-user symmetric GIC-NOF. Then,* $\mathcal{C}$ *is contained within the set* *of rate pairs* $(R_1, R_2) \in \mathbb{R}_+^2$ *that satisfy the following inequalities:*

$$R_1 \leqslant \min\left(\Upsilon_1(\rho), \Upsilon_2\right), \tag{2.21a}$$

$$R_2 \leqslant \min\left(\Upsilon_1(\rho), \Upsilon_2\right), \tag{2.21b}$$

$$R_1 + R_2 \leqslant \min\left(\Upsilon_3(\rho), \Upsilon_4\right), \tag{2.21c}$$

$$2R_1 + R_2 \leqslant \Upsilon_5(\rho), \tag{2.21d}$$

$$R_1 + 2R_2 \leqslant \Upsilon_5(\rho), \tag{2.21e}$$

with

$$\Upsilon_1(\rho) \triangleq \frac{1}{2}\log\left(1 + \overrightarrow{\mathrm{SNR}} + \mathrm{INR} + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}\mathrm{INR}}\right), \tag{2.22a}$$

$$\Upsilon_2 \triangleq \frac{1}{2}\log\left(1 + \overrightarrow{\mathrm{SNR}}\right) + \frac{1}{2}\log\left(1 + \frac{\overleftarrow{\mathrm{SNR}}}{1 + \overrightarrow{\mathrm{SNR}}}\right), \tag{2.22b}$$

$$\Upsilon_3(\rho) \triangleq \frac{1}{2}\log\left(1 + \overrightarrow{\mathrm{SNR}} + \mathrm{INR} + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}\mathrm{INR}}\right) + \frac{1}{2}\log\left(1 + \frac{\overrightarrow{\mathrm{SNR}}}{1 + \mathrm{INR}}\right), \tag{2.22c}$$

$$\Upsilon_4 \triangleq \begin{cases} \log\left(1 + \frac{\mathrm{INR}^2}{\overrightarrow{\mathrm{SNR}}}\right) + \log\left(1 + \frac{\overleftarrow{\mathrm{SNR}}}{\mathrm{INR}}\right) + \log\left(\frac{\overrightarrow{\mathrm{SNR}}}{\mathrm{INR}}\right) + \log 3 & \text{if } \frac{1}{2} \leqslant \alpha_G < 1 \\ \log\left(\frac{\mathrm{INR}^2 + \overrightarrow{\mathrm{SNR}} + 2\mathrm{INR} + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}\mathrm{INR}} + 1}{\mathrm{INR} + 1}\right) & \\ + \log\left(1 + \frac{\overleftarrow{\mathrm{SNR}}\left(\mathrm{INR}+1\right)}{\overrightarrow{\mathrm{SNR}} + \mathrm{INR} + 1}\right) & \text{otherwise} \end{cases}, \tag{2.22d}$$

$$\Upsilon_5(\rho) \triangleq \begin{cases} \frac{1}{2}\left(1 + \frac{\mathrm{INR}^2}{\overrightarrow{\mathrm{SNR}}}\right) + \frac{1}{2}\log\left(1 + \frac{\overleftarrow{\mathrm{SNR}}}{\mathrm{INR}}\right) + \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}}{\mathrm{INR}}\right) + \frac{1}{2}\log 3 & \\ + \frac{1}{2}\log\left(1 + \frac{\overrightarrow{\mathrm{SNR}}}{1 + \mathrm{INR}}\right) + \frac{1}{2}\log\left(1 + \overrightarrow{\mathrm{SNR}} + \mathrm{INR} + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}\mathrm{INR}}\right) & \text{if } \frac{1}{2} \leqslant \alpha_G < 1 \\ \frac{1}{2}\log\left(\frac{\mathrm{INR}^2 + \overrightarrow{\mathrm{SNR}} + 2\mathrm{INR} + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}\mathrm{INR}} + 1}{\mathrm{INR} + 1}\right) & \\ + \frac{1}{2}\log\left(1 + \frac{\overleftarrow{\mathrm{SNR}}\left(\mathrm{INR}+1\right)}{\overrightarrow{\mathrm{SNR}} + \mathrm{INR} + 1}\right) + \frac{1}{2}\log\left(1 + \frac{\overrightarrow{\mathrm{SNR}}}{\mathrm{INR} + 1}\right) & \\ + \frac{1}{2}\log\left(1 + \overrightarrow{\mathrm{SNR}} + \mathrm{INR} + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}\mathrm{INR}}\right) & \text{otherwise} \end{cases} \tag{2.22e}$$

$\alpha_G \triangleq \frac{\log \mathrm{INR}}{\log \overrightarrow{\mathrm{SNR}}}$, *and* $\rho \in [0, 1]$.

An other outer bound for the two-user IC-NOF is introduced in [39], which considers the Hekstra-Willems dependence-balance arguments used in the analysis of two-way channels. In the GIC, these results suggest that feedback loses its ability to increase the capacity region when the noise variance on the feedback link is larger than that on the forward link. Using similar arguments, new outer bounds that are tighter than the cut-set bound in some interference regimes are presented in [89].

2.1.4. Rate-Limited Feedback

The two-user GIC with rate-limited feedback (GIC-RLF), in which the feedback links have finite capacity instead of the case in which the feedback is perfect or noiseless, is analyzed in [6, 7, 95]. This corresponds to a more realistic feedback model in which the receivers can use all the information they have received to feed information back through an orthogonal channel of finite capacity. The rate-limited feedback (RLF) increases the complexity of the receivers, given that these must encode the information they transmit over the capacity-limited feedback channels. Under symmetric conditions in this channel model, the symmetric capacity is approximated to within 7.4 bits from the inner bound. The problem is analyzed using three different IC models: the El Gamal-Costa deterministic model [30], the linear deterministic model [8, 20], and the Gaussian model. In the analysis of the deterministic models the coding strategies are based upon: rate-splitting [23, 37], quantize-and-binning, and decode-and-forward. In the analysis of the Gaussian model, the coding strategy is based upon block-Markov superposition coding [14, 27], backward decoding [98, 99], and lattice coding, which enable receivers to decode superposition of codewords. Outer bounds are developed based upon the insights from the analysis of the deterministic models.

In [53, 57, 63, 88, 90, 95], the key insights for the analysis of the Gaussian cases are obtained from previous analysis of the respective linear deterministic models.

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

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

$$\overleftarrow{\boldsymbol{y}} = g(\overrightarrow{\boldsymbol{y}}) \in \mathcal{D} \subset \mathbb{R}^N, \tag{2.23}$$

such that for all $\delta > 0$,

$$|\mathcal{D}| < 2^{N(C_F+\delta)}. \tag{2.24}$$

The choice of the deterministic transformation g subject to (2.24) is part of the coding scheme, that is, the transformation g processes the N channel outputs observed during block $t > 0$ and chooses a codeword in the codebook $\mathcal{D}$. Such a codeword is sent back to the transmitter during block $t+1$. From this standpoint, this model highlights the signal impairments derived from transmitting a signal with continuous support via a channel with finite-capacity. Note that if $C_F = \infty$, then g can be the identity function and thus, $\overleftarrow{\boldsymbol{y}} = g(\overrightarrow{\boldsymbol{y}}) = \overrightarrow{\boldsymbol{y}}$, which is the case of POF [88]. When $C_F = 0$, then $|\mathcal{D}| = 1$ and thus, no information can be conveyed through the feedback links, which is the case studied in [24, 31, 37]. The main result in [95] is twofold: first, given a fixed C_F, the authors provide a deterministic transformation g using lattice coding [75] and a particular power assignment such that partial or complete decoding of the interference is possible at the transmitter. An achievable region is presented using random coding arguments with rate splitting, block-Markov superposition coding, and backward decoding. Second, the authors provide outer bounds that hold for any g in (2.23). This result induces a converse region whose sum-rate is shown to be at a constant gap of the achievable sum-rate, at least in the symmetric case. These results are generalized for the K-user GIC with RLF in the symmetric case in [6, 7], where the analysis focuses on the fundamental limit of the symmetric rate. The main novelty on the extension to $K > 2$ users lies in the joint use of interference alignment and lattice codes for the proof of the achievability. The proof of converse remains an open problem when $K > 2$, even for the symmetric case.

2.1.5. Intermittent Feedback

This model emphasizes the fact that the usage of the feedback link might be available only during certain channel uses, not necessarily known by the receivers with anticipation. This model is referred to as *intermittent feedback* (IF) [41]. The main result in [41] is an approximation of the capacity region to within a constant gap. The achievability scheme relies upon random coding arguments with forward decoding and a quantize-map-and-foward strategy to retransmit the information obtained through feedback. This is because erasures might constrain either partial or complete decoding of the interference at the transmitter. Nonetheless, even a quantized version of the interference might be useful at the intended receiver for interference cancellation or at the non-intended receiver for providing an alternative path.

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

$$\overleftarrow{Y}_n = \begin{cases} \star & \text{with probability} \quad 1-p \\ \overrightarrow{y}_n & \text{with probability} \quad p \end{cases}, \tag{2.25}$$

where $\star$ represents an erasure and $(1-p) \in [0, 1]$ is its probability of ocurrence. Note that the random transformation g is fully determined by the parameters of the channels, e.g., the probability p. Thus, as opposed to the RLF, the transformation g can not be optimized as part of the receiver design.
In the case of the two-user symmetric GIC-NOF, assume that for all $n \in \{1, 2 \ldots, N\}$, the random transformation g is such that given a channel output $\overrightarrow{y}_n$,

$$\overleftarrow{Y}_n = \overleftarrow{h}\,\overrightarrow{y}_n + Z_n, \tag{2.26}$$

where $\overrightarrow{h} \in \mathbb{R}_+$ is a parameter of the channel and Z_n is a real Gaussian random variable with zero mean and unit variance. Note that the receiver does not apply any processing to the channel output and sends a re-scaled copy to the transmitter via a noisy channel. From this point of view, as opposed to RLF, the GIC-NOF model does not focus on the constraint on the number of codewords that can be used to perform feedback, but rather on the fact that the feedback channel might be noisy. Essentially, the codebook used to perform feedback in NOF is $\mathbb{R}^N$.

2.1.6. A Comparison Between Feedback Models

Let $\mathcal{C}(\overrightarrow{\mathrm{SNR}}, \mathrm{INR})$ denote a set containing all achievable rates of a symmetric GIC with parameters $\overrightarrow{\mathrm{SNR}}$ (signal to noise ratio in the forward link) and INR (interference to noise ratio). The number of GDoF [40] is:

$$\text{GDoF}(\alpha) = \lim_{\overrightarrow{\text{SNR}} \to \infty} \frac{\sup\left\{R : (R,R) \in \mathcal{C}(\overrightarrow{\text{SNR}}, \overrightarrow{\text{SNR}}^{\alpha})\right\}}{\log\left(\overrightarrow{\text{SNR}}\right)}, \tag{2.27}$$

where $\alpha = \frac{\log(\text{INR})}{\log(\overrightarrow{\text{SNR}})}$. In Figure 2.3, the number of GDoF is plotted as a function of α when $\mathcal{C}(\overrightarrow{\text{SNR}}, \text{INR})$ is calculated without feedback (dashed line)[31]; and with PF from each receiver to their corresponding transmitters (solid line)[88]. Note that with PF, the number of GDoF goes to infinity when α goes to infinity, which implies an arbitrarily large gain. Surprisingly, using only one PF link from one of the receivers to the corresponding transmitter provides the same sum-capacity as having four PF links from both receivers to both transmitters [79, 80, 71] in certain interference regimes. These benefits rely on the fact that feedback from the intended receiver to the corresponding transmitter provides relevant information about the interference. Hence, such information can be retransmitted to: (a) perform interference cancellation at the intended receiver or (b) provide an alternative communication path between the other transmitter-receiver pair. These promising results are also observed when the system is decentralized, that is, when each transmitter seeks to unilaterally maximize its own individual information rate [66].

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

$$\beta = \frac{\log\left(\overleftarrow{\text{SNR}}\right)}{\log\left(\overrightarrow{\text{SNR}}\right)} \text{ and} \tag{2.28a}$$

$$\beta' = \frac{C_F}{\log\left(\overrightarrow{\text{SNR}}\right)} \tag{2.28b}$$

be fixed parameters. These parameters approximate the ratio between the capacity of the feedback link and the capacity of the forward link in the NOF and RLF case, respectively. Hence, a fair comparison of RLF and NOF must be made with $\beta = \beta'$. The number of GDoF is plotted as a function of α when $\mathcal{C}(\overrightarrow{\text{SNR}}, \text{INR})$ is calculated with NOF for several values of β in Figure 2.3(a); with RLF for different values of β' in Figure 2.3(b); and with IF for several values of p in Figure 2.3(c).

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

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

2.2. Linear Deterministic Interference Channel

A deterministic channel model is introduced by [8] as an approximation to the Gaussian channel models in the very high SNR regime. This model captures the key properties of the wireless communication systems: the signal strength; the broadcast nature of the wireless channel in which the signal sent by one transmitter can be overheard by many receivers at different signal strengths; and multiple signals can arrive to one receiver coming from different transmitters. Networks are affected not only by noise but also by interference. This linear deterministic approximation considers that the network is operating in an interference-limited regime, where the noise power is small compared to the signal powers. Thus, this model focuses on the interactions of the signals rather than the noise. Therefore, the noise as well as the parts of the signal affected by the noise are neglected. the LDIC is a special class of the El Gamal-Costa deterministic IC [30] and a special class of the IC-NOF.

Consider the two-user LDIC-NOF depicted in Figure 2.4. 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 non-negative integer parameters describe the two-user LDIC-NOF in Figure 2.4.

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

$$q = \max\left(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}\right). \tag{2.29}$$

At receiver i, the channel-output $\overrightarrow{\boldsymbol{Y}}_{i,n}$ is also a q-dimensional binary vector $\overrightarrow{\boldsymbol{Y}}_{i,n} = \left(\overrightarrow{Y}_{i,n}^{(1)}, \overrightarrow{Y}_{i,n}^{(2)}, \ldots, \overrightarrow{Y}_{i,n}^{(q)}\right)^{\mathsf{T}}$. Let $\boldsymbol{S}$ be a $q \times q$ lower shift matrix of the form:

$$\boldsymbol{S} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}. \tag{2.30}$$

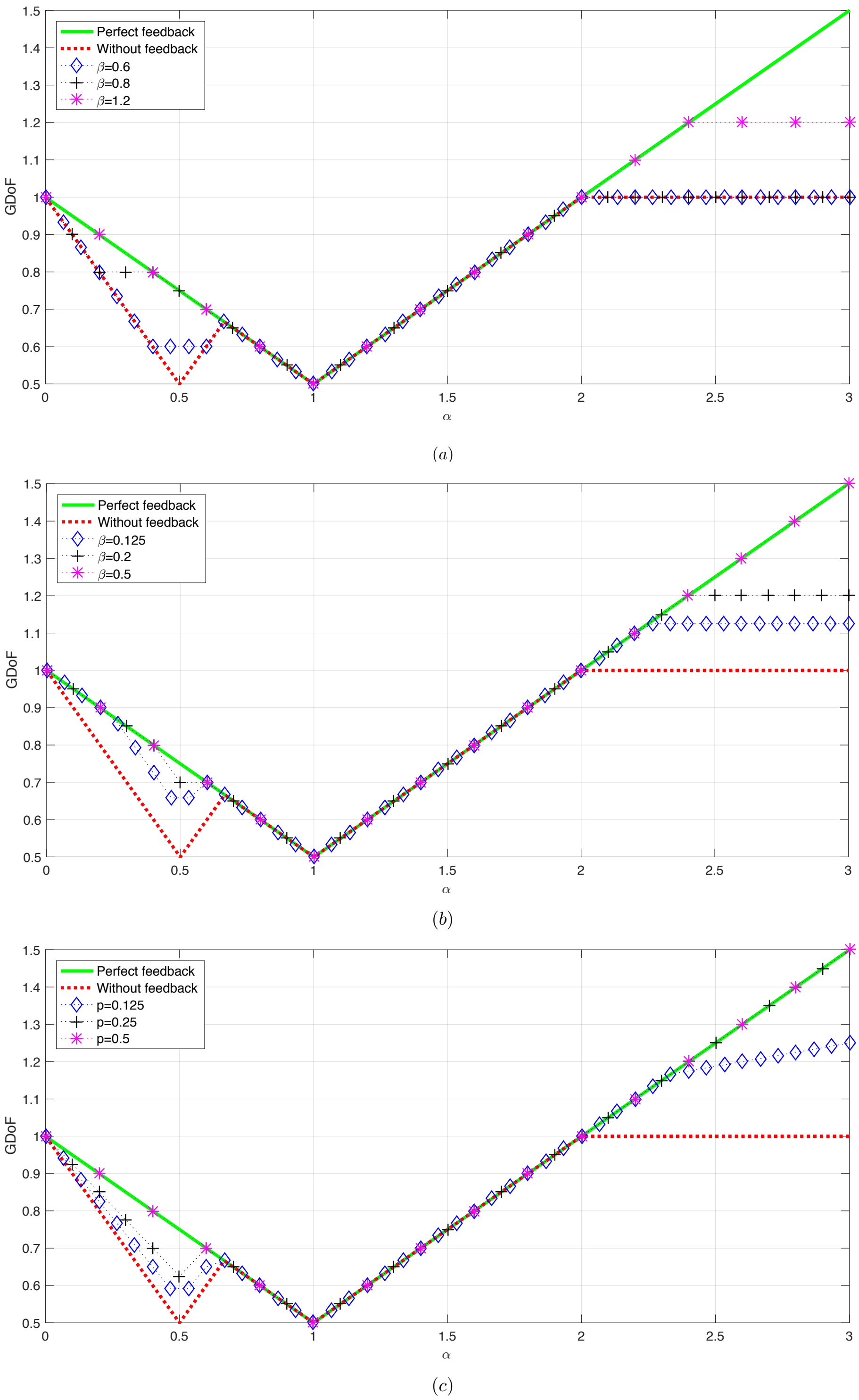


Figure 2.3.: Number of generalized degrees of freedom (GDoF) of a symmetric two-user GIC; (a) case with NOF with $\beta \in \{0.6, 0.8, 1.2\}$; (b) case with RLF with $\beta \in \{0.125, 0.2, 0.5\}$; and (c) case with IF with $p \in \{0.125, 0.25, 0.5\}$

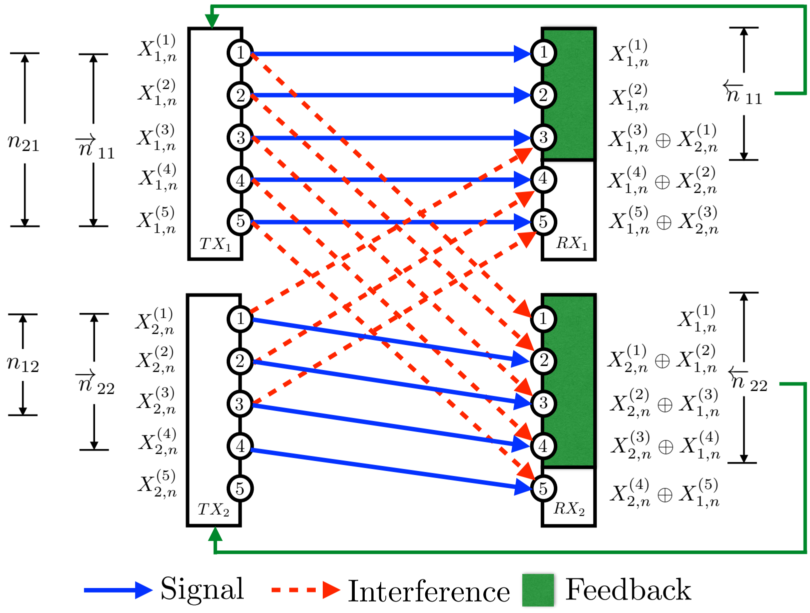


Figure 2.4.: Two-user linear deterministic interference channel with noisy channel-output feedback at channel use n.

The input-output relation during channel use n is given for all $i \in \{1,2\}$, with $j \in \{1,2\} \setminus \{i\}$ by

$$\overrightarrow{\boldsymbol{Y}}_{i,n} = \boldsymbol{S}^{q-\overrightarrow{n}_{ii}} \boldsymbol{X}_{i,n} + \boldsymbol{S}^{q-n_{ij}} \boldsymbol{X}_{j,n}, \tag{2.31}$$

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

$$\left((0,\ldots,0), \overleftarrow{\boldsymbol{Y}}_{i,n}^{\mathsf{T}}\right)^{\mathsf{T}} = \boldsymbol{S}^{\left(\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii}\right)^+} \overrightarrow{\boldsymbol{Y}}_{i,n-d}, \tag{2.32}$$

where d is a finite delay and additions and multiplications between matrices and vectors are defined over the Galois Field of cardinality two, GF(2).

The dimension of the vector $(0,\ldots,0)$ in (2.32) is $q - \min\left(\overleftarrow{n}_{ii}, \max(\overrightarrow{n}_{ii}, n_{ij})\right)$ and the vector $\overleftarrow{\boldsymbol{Y}}_{i,n}$ represents the $\min\left(\overleftarrow{n}_{ii}, \max(\overrightarrow{n}_{ii}, n_{ij})\right)$ least significant bits of $\boldsymbol{S}^{\left(\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii}\right)^+} \overrightarrow{\boldsymbol{Y}}_{i,n-d}$.

Without any loss of generality, the feedback delay is assumed to be equal to 1 channel use.

In this special case, the pdf of the IC-NOF can be factorized as in (2.9). Based on the input-output relation in (2.31), for all $i \in \{1,2\}$ and given the channel-inputs $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ during a specific channel use, the pdf $f_{\overrightarrow{\boldsymbol{Y}}_i|\boldsymbol{X}_1,\boldsymbol{X}_2}$ in (2.9) can be expressed as follows:

$$f_{\overrightarrow{\boldsymbol{Y}}_i|\boldsymbol{X}_1,\boldsymbol{X}_2}(\overrightarrow{\boldsymbol{y}}_i|\boldsymbol{x}_1,\boldsymbol{x}_2) = \mathbb{1}_{\left\{\overrightarrow{\boldsymbol{y}}_i = \boldsymbol{S}^{q-\overrightarrow{n}_{ii}}\boldsymbol{x}_i + \boldsymbol{S}^{q-n_{ij}}\boldsymbol{x}_j\right\}}, \tag{2.33}$$

for all $\overrightarrow{\boldsymbol{y}}_i, \boldsymbol{x}_1, \boldsymbol{x}_2$.

Similarly, based on the input-output relation in (2.32), for all $i \in \{1,2\}$ and given the channel-outputs $\overrightarrow{\boldsymbol{y}}_1$ and $\overrightarrow{\boldsymbol{y}}_2$ during a specific channel use, the pdf $f_{\overleftarrow{\boldsymbol{Y}}_i|\overrightarrow{\boldsymbol{Y}}_i}$ in (2.9) can be expressed as follows:

$$\mathrm{f}_{\overleftarrow{\boldsymbol{Y}}_i|\overrightarrow{\boldsymbol{Y}}_i}\left(\overleftarrow{\boldsymbol{y}}_i|\overrightarrow{\boldsymbol{y}}_i\right) = \mathbb{1}_{\left\{\overleftarrow{\boldsymbol{y}}_i^{\mathsf{T}} = \boldsymbol{S}^{\left(\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii}\right)^+} \overrightarrow{\boldsymbol{y}}_i\right\}}, \tag{2.34}$$

for all $\overleftarrow{\boldsymbol{y}}_i, \overrightarrow{\boldsymbol{y}}_i$.

2.2.1. Case without Feedback

The following lemma presents the capacity region for the two-user LDIC without channel-output feedback.

Lemma 9 (Capacity region two-user LDIC [20, Lemma 4])**.** *Let $\mathcal{C} \subset \mathbb{R}_+^2$ denote the capacity region of the two-user LDIC without channel-output feedback. Then, $\mathcal{C}$ contains all the rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy the following inequalities:*

$$R_1 \leqslant \overrightarrow{n}_{11}, \tag{2.35a}$$

$$R_2 \leqslant \overrightarrow{n}_{22}, \tag{2.35b}$$

$$R_1 + R_2 \leqslant \left(\overrightarrow{n}_{11} - n_{12}\right)^+ + \max\left(\overrightarrow{n}_{22}, n_{12}\right), \tag{2.35c}$$

$$R_1 + R_2 \leqslant \left(\overrightarrow{n}_{22} - n_{21}\right)^+ + \max\left(\overrightarrow{n}_{11}, n_{21}\right), \tag{2.35d}$$

$$R_1 + R_2 \leqslant \max\left(n_{21}, \left(\overrightarrow{n}_{11} - n_{12}\right)^+\right) + \max\left(n_{12}, \left(\overrightarrow{n}_{22} - n_{21}\right)^+\right), \tag{2.35e}$$

$$2R_1 + R_2 \leqslant \max\left(\overrightarrow{n}_{11}, n_{21}\right) + \left(\overrightarrow{n}_{11} - n_{12}\right)^+ + \max\left(n_{12}, \left(\overrightarrow{n}_{22} - n_{21}\right)^+\right), \tag{2.35f}$$

$$R_1 + 2R_2 \leqslant \max\left(\overrightarrow{n}_{22}, n_{12}\right) + \left(\overrightarrow{n}_{22} - n_{21}\right)^+ + \max\left(n_{21}, \left(\overrightarrow{n}_{11} - n_{12}\right)^+\right). \tag{2.35g}$$

2.2.2. Case with Perfect Channel-Output Feedback

The following lemma presents the capacity region for the two-user LDIC-POF.

Lemma 10 (Capacity region two-user LDIC-POF [88, Corollary 1])**.** *Let $\mathcal{C} \subset \mathbb{R}_+^2$ denote the capacity region of the two-user LDIC-POF. Then, $\mathcal{C}$ contains all the rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy the following inequalities:*

$$R_1 \leqslant \min\left(\max\left(\overrightarrow{n}_{11}, n_{21}\right), \max\left(\overrightarrow{n}_{11}, n_{12}\right)\right), \tag{2.36a}$$

$$R_2 \leqslant \min\left(\max\left(\overrightarrow{n}_{22}, n_{12}\right), \max\left(\overrightarrow{n}_{22}, n_{21}\right)\right), \tag{2.36b}$$

$$R_1 + R_2 \leqslant \min\left(\max\left(\overrightarrow{n}_{22}, n_{21}\right) + \left(\overrightarrow{n}_{11} - n_{21}\right)^+, \max\left(\overrightarrow{n}_{11}, n_{12}\right) + \left(\overrightarrow{n}_{22} - n_{12}\right)^+\right). \tag{2.36c}$$

2.2.3. Symmetric Case with Noisy Channel-Output Feedback

The following lemma presents the capacity region for the two-user symmetric LDIC-NOF, in which $\overrightarrow{n}_{11} = \overrightarrow{n}_{22} = \overrightarrow{n}$, $n_{12} = n_{21} = m$, and $\overleftarrow{n}_{11} = \overleftarrow{n}_{22} = \overleftarrow{n}$.

Lemma 11 (Capacity region two-user symmetric LDIC-NOF [53, Theorem 1])**.** *Let $\mathcal{C} \subset \mathbb{R}_+^2$ denote the capacity region of the two-user symmetric LDIC-NOF. Then, $\mathcal{C}$ contains all the* *rate pairs* $(R_1, R_2) \in \mathbb{R}_+^2$ *that satisfy the following inequalities:*

$$
\begin{aligned}
R_1 &\leqslant \max\left(\overrightarrow{n}, m\right), && \text{(2.37a)}\\
R_2 &\leqslant \max\left(\overrightarrow{n}, m\right), && \text{(2.37b)}\\
R_1 &\leqslant \overrightarrow{n} + \left(\overleftarrow{n} - \overrightarrow{n}\right)^+, && \text{(2.37c)}\\
R_2 &\leqslant \overrightarrow{n} + \left(\overleftarrow{n} - \overrightarrow{n}\right)^+, && \text{(2.37d)}\\
R_1 + R_2 &\leqslant \max\left(\overrightarrow{n}, m\right) + \left(\overrightarrow{n} - m\right)^+, && \text{(2.37e)}\\
R_1 + R_2 &\leqslant 2\max\Big(\left(\overrightarrow{n} - m\right)^+, m\Big) + 2\min\Big(\left(\overrightarrow{n} - m\right)^+, \Big(\overleftarrow{n} - \max\Big(m, \left(\overrightarrow{n} - m\right)^+\Big)\Big)^+\Big), && \text{(2.37f)}\\
2R_1 + R_2 &\leqslant \max\left(\overrightarrow{n}, m\right) + \left(\overrightarrow{n} - m\right)^+ + \max\Big(\left(\overrightarrow{n} - m\right)^+, m\Big)\\
&\quad + \min\Big(\left(\overrightarrow{n} - m\right)^+, \Big(\overleftarrow{n} - \max\Big(m, \left(\overrightarrow{n} - m\right)^+\Big)\Big)^+\Big), && \text{(2.37g)}\\
R_1 + 2R_2 &\leqslant \max\left(\overrightarrow{n}, m\right) + \left(\overrightarrow{n} - m\right)^+ + \max\Big(\left(\overrightarrow{n} - m\right)^+, m\Big)\\
&\quad + \min\Big(\left(\overrightarrow{n} - m\right)^+, \Big(\overleftarrow{n} - \max\Big(m, \left(\overrightarrow{n} - m\right)^+\Big)\Big)^+\Big). && \text{(2.37h)}
\end{aligned}
$$

2.2.4. Symmetric Case with only one Perfect Channel-Output Feedback

The following lemma presents the capacity region for the two-user symmetric LDIC with only one POF, in which $\overrightarrow{n}_{11} = \overrightarrow{n}_{22} = \overrightarrow{n}$, $n_{12} = n_{21} = m$ and $\overleftarrow{n}_{11} = \max\left(\overrightarrow{n}, m\right)$, and $\overleftarrow{n}_{22} = 0$.

Lemma 12 (Capacity region two-user LDIC with only one POF, [**\cite**]Sahai-TIT-2013)**.** *Let* $\mathcal{C} \subset \mathbb{R}_+^2$ *denote the capacity region of the two-user symmetric LDIC with only one POF between receiver* 1 *and transmitter* 1*. Then,* $\mathcal{C}$ *contains all the rate pairs* $(R_1, R_2) \in \mathbb{R}_+^2$ *that satisfy the following inequalities:*

$$
\begin{aligned}
R_1 &\leqslant \overrightarrow{n}, && \text{(2.38)}\\
R_2 &\leqslant \max\left(\overrightarrow{n}, m\right), && \text{(2.39)}\\
R_1 + R_2 &\leqslant \max\left(\overrightarrow{n}, m\right) + \left(\overrightarrow{n} - m\right)^+, && \text{(2.40)}\\
2R_1 + R_2 &\leqslant \max\left(\overrightarrow{n}, m\right) + \left(\overrightarrow{n} - m\right)^+ + \max\left(\overrightarrow{n} - m, m\right). && \text{(2.41)}
\end{aligned}
$$

Note that the model 0001 in [80] corresponds to the two-user symmetric LDIC with only one POF between receiver 2 and transmitter 2. Note also that the model 1001 in [80] corresponds to the two-user symmetric LDIC with POF and it is equivalent to the Lemma 10 for symmetric parameters in the LDIC.

2.2.5. Sum-Capacity with Source Cooperation

In the two-user IC-NOF, a transmitter sees a noisy version of the sum of its own transmitted signal and the interfering signal from the other transmitter. Hence, subject to a finite delay, one transmitter knows, at least partially, the information transmitted by the other transmitter in the network. This observation highlights the connections between the IC with feedback and the IC with source cooperation studied in [63]. These two channel models are related but they are not the same. There are two main differences between the two channel models. First, the channel-output signal observed by the transmitter in the case of IC-NOF is impaired by the noise in the feedback link and the noise in the forward channel. In the case of source cooperation, the cooperation signal is only affected by the noise in the cooperative link. Second, the cooperation between transmitters is direct and symmetric in the case of source cooperation.

Conversely, in the case of IC-NOF, the signal that is observed by the transmitter is affected by the delay in the feedback link, and the part of the signal that was transmitted by the other transmitter is obtained from the substraction between the signal observed by the transmitter and its own signal that was transmitted previously. Then, the cooperation is not direct [53].

The two-user IC with source cooperation has two transmitters, *i.e.*, 1 and 2, two receivers, *i.e.*, 3 and 4, and it also has noisy connections between the two transmitters [63]. The following lemma presents the sum-capacity of the LDIC with source cooperation.

Lemma 13 (Sum-capacity two-user LDIC with source cooperation [63, Theorem 1])**.** *The sum-capacity region of the two-user LDIC with source cooperation is the minimum of the following inequalities:*

$$R_1 + R_2 = \max\left(n_{1,3} - n_{1,4} + n_c, n_{2,3}, n_c\right) + \max\left(n_{2,4} - n_{2,3} + n_c, n_{1,4}, n_c\right), \tag{2.42a}$$

$$R_1 + R_2 = \max\left(n_{1,3}, n_{2,3}\right) + \left(\max\left(n_{2,4}, n_{2,3}, n_c\right) - n_{2,3}\right), \tag{2.42b}$$

$$R_1 + R_2 = \max\left(n_{2,4}, n_{1,4}\right) + \left(\max\left(n_{1,3}, n_{1,4}, n_c\right) - n_{1,4}\right), \tag{2.42c}$$

$$R_1 + R_2 = \max\left(n_{1,3}, n_c\right) + \max\left(n_{2,4}, n_c\right), \tag{2.42d}$$

$$R_1 + R_2 = \begin{cases} \max\left(n_{1,3} + n_{2,4}, n_{1,4} + n_{2,3}\right) & \textit{if } n_{1,3} - n_{2,3} \neq n_{1,4} - n_{2,4}, \\ \max\left(n_{1,3}, n_{2,4}, n_{1,4}, n_{2,3}\right) & \textit{otherwise} \end{cases}. \tag{2.42e}$$

In order to establish a connection between (2.42) and the sum-rate capacity of the two-user LDIC-NOF the following identities must be introduced: $n_{1,3} = \overrightarrow{n}_{11}$, $n_{2,4} = \overrightarrow{n}_{22}$, $n_{2,3} = n_{12}$, $n_{1,4} = n_{21}$, and $n_c = \overleftarrow{n}_{11} - (\overrightarrow{n}_{11} - n_{12})^+ = \overleftarrow{n}_{22} - (\overrightarrow{n}_{22} - n_{21})^+$. The last equality implies that the feedback must include the signal levels that contain information about the non-intended source in order to establish cooperation between the sources.

— 3 —

Decentralized Interference Channels

In a decentralized system, a central controller does not exist and each transmitter-receiver pair is responsible for the selection of its own transmit-receive configuration to maximize its data transmission rate. A transmit-receive configuration for transmitter-receiver pair i, with $i \in \{1, 2\}$, denoted by s_i, can be described in terms of the block-length N_i, the number of bits per block $M_i = \lceil \log_2 |\mathcal{W}_i| \rceil$, the channel-input alphabet $\mathcal{X}_i$, the codebook $\mathcal{C}_i$, the encoding function f_i, the decoding function ψ_i, etc. The aim of transmitter i is to autonomously choose its transmit-receive configuration s_i, in order to maximize its achievable rate R_i. Note that the rate achieved by transmitter-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 system models for the two-user decentralized continuous IC-NOF; the two-user D-GIC-NOF; and the two-user D-LDIC-NOF are in general the same as in the centralized case. The main differences are the following:

- Each transmitter-receiver defines the number of channel-uses per block, *i.e.*, N_1 and N_2 channel uses.
- The transmission of a block consists of N channel uses, where $N = \max(N_1, N_2)$. Then, $X_{i,n} = 0$ for all $n > N_i$.
- Encoder i generates the symbol $x_{i,n}$ considering not only the message index $W_i \in \mathcal{W}_i = \{1, 2, \ldots, 2^{N_i R_i}\}$ and all previous outputs from the feedback link i, *i.e.*, $\left(\overleftarrow{y}_{i,1}, \overleftarrow{y}_{i,2}, \ldots, \overleftarrow{y}_{i,n-1}\right)$, but also the random message index $\Omega_i \in \mathbb{N}$. The index Ω_i is an additional index randomly generated which is assumed to be known by both transmitter i and receiver i, while unknown by transmitter j and receiver j (common randomness).

- At the end of the transmission, decoder i uses all the channel-outputs, *i.e.*, $\left(\overrightarrow{y}_{i,1}, \overrightarrow{y}_{i,2}, \ldots, \overrightarrow{y}_{i,N}\right)$ and the random message index Ω_i to obtain an estimate of the message index W_i, denoted by $\widehat{W}_i$.

- The following Markov chain holds:

$$\left(W_i, \Omega_i, \overleftarrow{\boldsymbol{Y}}_{i,(1;n-1)}\right) \rightarrow X_{i,n} \rightarrow \overrightarrow{Y}_{i,n}. \tag{3.1}$$

- The calculation of the probability of error is made for each of the transmitter-receiver pairs. Let $W_i^{(t)}$ be written as $c_{i,1}^{(t)}\, c_{i,2}^{(t)} \ldots c_{i,M_i}^{(t)}$ in binary form. Let also $\widehat{W}_i^{(t)}$ be written as $\widehat{c}_{i,1}^{(t)}\, \widehat{c}_{i,2}^{(t)} \ldots \widehat{c}_{i,M_i}^{(t)}$ in binary form. Then, the average bit error probability at decoder i given the configurations s_1 and s_2, denoted by $p_i(s_1, s_2)$, is given by

$$p_i(s_1, s_2) = \frac{1}{M_i} \sum_{\ell=1}^{M_i} \mathbb{1}_{\left\{\widehat{c}_{i,\ell}^{(t)} \neq c_{i,\ell}^{(t)}\right\}}. \tag{3.2}$$

The fundamental limits in a decentralized two-user IC-NOF system are defined by the η-NE region.

Definition 3 (η-NE region of a two-user IC)**.** *The η-NE region of a two-user IC is the closure of the set of all possible achievable rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that are stable in the sense of a Nash equilibrium. More specifically, given an η-NE coding scheme, there does not exist an alternative coding scheme for either transmitter-receiver pair that increases their individual rates by more than η bits per channel-use.*

3.1. Game Formulation

The competitive interaction between the two transmitter-receiver pairs in the IC can be modeled by the following game in normal-form:

$$\mathcal{G} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}\right). \tag{3.3}$$

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. The choice of one transmit-receive configuration by player $i \in \mathcal{K}$ is an action, which is denoted by $s_i \in \mathcal{A}_i$. The utility function of player i is $u_i : \mathcal{A}_1 \times \mathcal{A}_2 \rightarrow \mathbb{R}_+$ and it is defined as the achieved rate of transmitter i,

$$u_i(s_1, s_2) = \begin{cases} R_i = \dfrac{M_i}{N_i} & \text{if} \quad p_i(s_1, s_2) < \epsilon \\ 0 & \quad\quad \text{otherwise} \end{cases}, \tag{3.4}$$

where $\epsilon > 0$ is an arbitrarily small number and R_i denotes a transmission rate achievable with the configurations s_1 and s_2. This game formulation was first proposed in [15] and [103].

A class of transmit-receive configuration pairs $\boldsymbol{s}^* = (s_1^*, s_2^*) \in \mathcal{A}_1 \times \mathcal{A}_2$ that are particularly important in the analysis of this game is referred to as the set of η-Nash equilibria (η-NE), with $\eta > 0$. These pairs of configurations satisfy the following definition:

Definition 4 (η-NE [60])**.** *In the game $\mathcal{G} = (\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}})$, a configuration pair $(s_1^*, s_2^*) \in \mathcal{A}_1 \times \mathcal{A}_2$ is an η-NE if for all $i \in \mathcal{K}$ and for all $s_i \in \mathcal{A}_i$, there exits an $\eta > 0$ such that*

$$u_i(s_i, s_j^*) \leqslant u_i(s_i^*, s_j^*) + \eta. \tag{3.5}$$

Let (s_1^*, s_2^*) be an η-NE configuration pair of the game in (3.3). Then, none of the transmitters can increase its own information 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. Note that for η sufficiently large, from Definition 4, any pair of configurations can be an η-NE. Alternatively, for $\eta = 0$, the classical definition of an NE is obtained [59]. In this case, if a pair of configurations is an NE ($\eta = 0$), then each individual configuration is optimal with respect to each other. Hence, the interest is to describe the set of all possible η-NE rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ of the game in (3.3) with the smallest η for which there exists at least one equilibrium configuration pair.

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

Definition 5 (η-NE Region)**.** *Let $\eta > 0$ be fixed. An achievable rate pair $(R_1, R_2) \in \mathbb{R}_+^2$ is said to be in the η-NE 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:*

$$u_1(s_1^*, s_2^*) = R_1 \quad \text{and} \quad u_2(s_1^*, s_2^*) = R_2. \tag{3.6}$$

Following along the same lines as in [16], if there exists a configuration pair (s_1, s_2) that achieves a rate pair $(R_1, R_2) \in \mathbb{R}_+^2$ using codes of block lengths N_1 and N_2 respectively, then it can be shown that there exists a configuration 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 configuration 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 thesis.

3.2. Gaussian Interference Channel

3.2.1. Case without Feedback

The following lemma presents an approximate NE region for the two-user GIC without channel-output feedback.

Lemma 14 (Approximate NE region two-user GIC, Theorem 2 in [16])**.** *Let $\mathcal{N} \subset \mathbb{R}_+^2$ denote the NE region of the two-user GIC without channel-output feedback. Then,*

$$\underline{\mathcal{C}} \cap \underline{\mathcal{B}} \subseteq \mathcal{N} \subseteq \mathcal{C} \cap \mathcal{B}, \tag{3.7}$$

with $\mathcal{C}$ the capacity region of the two-user GIC, $\underline{\mathcal{C}}$ the achievable region of the two-user GIC (Lemma 1), and, $\mathcal{B}$ and $\underline{\mathcal{B}}$ given by

$$\mathcal{B} = \left\{(R_1, R_2) \in \mathbb{R}_+^2 : L_i \leqslant R_i \leqslant U_i, \text{ for all } i \in \{1, 2\}\right\}, \tag{3.8a}$$

$$\underline{\mathcal{B}} = \left\{(R_1, R_2) \in \mathbb{R}_+^2 : L_i \leqslant R_i \leqslant \max\left(U_i - 1, L_i\right), \text{ for all } i \in \{1, 2\}\right\}, \tag{3.8b}$$

where for all $i \in \{1,2\}$

$$L_i = \frac{1}{2}\log\left(1+\frac{\overrightarrow{\mathrm{SNR}}_i}{1+\mathrm{INR}_{ij}}\right) \text{ and } \tag{3.9a}$$

$$U_i = \min\left(\frac{1}{2}\log\left(1+\overrightarrow{\mathrm{SNR}}_i+\mathrm{INR}_{ij}\right)-\frac{1}{2}\log\left(1+\frac{\left(\overrightarrow{\mathrm{SNR}}_j-\max\left(\mathrm{INR}_{ji},\frac{\overrightarrow{\mathrm{SNR}}_j}{\mathrm{INR}_{ij}}\right)\right)^+}{1+\mathrm{INR}_{ji}+\max\left(\mathrm{INR}_{ji},\frac{\overrightarrow{\mathrm{SNR}}_j}{\mathrm{INR}_{ij}}\right)}\right),\frac{1}{2}\log\left(1+\overrightarrow{\mathrm{SNR}}_i\right)\right). \tag{3.9b}$$

The region defined by $\underline{\mathcal{B}}$ differs from $\mathcal{B}$ by at most one bit, given that the achievable region in Lemma 1 is at most one bit away from the capacity region [31].

3.2.2. Case with Perfect Channel-Output Feedback

The following lemma presents an approximate NE region for the two-user GIC-POF.

Lemma 15 (Approximate NE region two-user GIC-POF, Theorem 2 in [66]). *Let* $\eta \geqslant 1$ *and let* $\mathcal{N} \subset \mathbb{R}_+^2$ *denote the NE region of the two-user GIC-POF. Then,*

$$\underline{\mathcal{C}} \cap \mathcal{B}_\eta \subseteq \mathcal{N} \subseteq \overline{\mathcal{C}} \cap \mathcal{B}_\eta, \tag{3.10}$$

with $\underline{\mathcal{C}}$ *the achievable region of the two-user GIC-POF (Lemma 5),* $\overline{\mathcal{C}}$ *the converse region of the two-user GIC-POF (Lemma 6), and* $\mathcal{B}_\eta$ *given by*

$$\mathcal{B}_\eta = \left\{(R_1,R_2)\in\mathbb{R}_+^2 : (L_i-\eta)^+ \leqslant R_i, \text{ for all } i\in\{1,2\}\right\}, \tag{3.11}$$

where for all $i \in \{1,2\}$, L_i *is given by* (3.9a).

3.3. Linear Deterministic Interference Channel

3.3.1. Case without Feedback

The following lemma presents the NE region for the two-user LDIC without channel-output feedback.

Lemma 16 (NE region two-user LDIC, Theorem 1 in [16]). *Let* $\mathcal{N} \subset \mathbb{R}_+^2$ *denote the NE region of the two-user LDIC without channel-output feedback. Then,* $\mathcal{N}$ *contains all the rate pairs* $(R_1,R_2)\in\mathbb{R}_+^2$ *that satisfy:*

$$\mathcal{N} = \mathcal{C} \cap \mathcal{B}, \tag{3.12}$$

with $\mathcal{C}$ *the capacity region of the two-user LDIC (Lemma 9) and* $\mathcal{B}$ *given by*

$$\mathcal{B} = \left\{(R_1,R_2)\in\mathbb{R}_+^2 : L_i \leqslant R_i \leqslant U_i, \text{ for all } i\in\{1,2\}\right\}, \tag{3.13}$$

where for all $i \in \{1,2\}$,

$$L_i = (\overrightarrow{n}_{ii} - n_{ij})^+ \ \ and \tag{3.14}$$

$$U_i = \begin{cases} \overrightarrow{n}_{ii} - \min(L_j, n_{ij}) & \text{if } n_{ij} \leqslant \overrightarrow{n}_{ii} \\ \min\left((n_{ij} - L_j)^+, \overrightarrow{n}_{ii}\right) & \text{if } n_{ij} > \overrightarrow{n}_{ii} \end{cases}. \tag{3.15}$$

3.3.2. Case with Perfect Channel-Output Feedback

The following lemma presents the NE region for the two-user LDIC-POF.

Lemma 17 (NE region two-user LDIC-POF, Theorem 1 in [66]). *Let* $\eta \geqslant 0$ *and let* $\mathcal{N} \subset \mathbb{R}_+^2$ *denote the NE region of the two-user LDIC-POF. Then,* $\mathcal{N}$ *contains all the rate pairs* $(R_1, R_2) \in \mathbb{R}_+^2$ *that satisfy:*

$$\mathcal{N} = \mathcal{C} \cap \mathcal{B}_\eta, \tag{3.16}$$

with $\mathcal{C}$ *the capacity region of the two-user LDIC-POF (Lemma 10) and* $\mathcal{B}$ *given by*

$$\mathcal{B}_\eta = \left\{(R_1, R_2) \in \mathbb{R}_+^2 : (L_i - \eta)^+ \leqslant R_i, \ \text{for all } i \in \{1,2\}\right\}, \tag{3.17}$$

where for all $i \in \{1,2\}$

$$L_i = (\overrightarrow{n}_{ii} - n_{ij})^+. \tag{3.18}$$

3.3.3. Symmetric Case with Noisy Channel-Output Feedback

The following lemma presents the NE region for the two-user symmetric LDIC-NOF, in which $\overrightarrow{n}_{11} = \overrightarrow{n}_{22} = \overrightarrow{n}$, $n_{12} = n_{21} = m$, and $\overleftarrow{n}_{11} = \overleftarrow{n}_{22} = \overleftarrow{n}$.

Lemma 18 (NE region two-user symmetric LDIC-NOF, Theorem 1 in [68]). *Let* $\mathcal{N} \subset \mathbb{R}_+^2$ *denote the NE region of the two-user symmetric LDIC-NOF. Then,* $\mathcal{N}$ *contains all the rate pairs* $(R_1, R_2) \in \mathbb{R}_+^2$ *that satisfy:*

$$\mathcal{N} = \mathcal{C} \cap \mathcal{B}, \tag{3.19}$$

with $\mathcal{C}$ *the capacity region of the two-user symmetric LDIC-NOF (Lemma 11) and* $\mathcal{B}$ *given by*

$$\mathcal{B} = \left\{(R_1, R_2) \in \mathbb{R}_+^2 : L \leqslant R_i \leqslant U, \ \text{for all } i \in \{1,2\}\right\}, \tag{3.20}$$

where for all $i \in \{1,2\}$

$$L_i = (\overrightarrow{n} - m)^+ \ \ and \tag{3.21}$$

$$U_i = \begin{cases} \min\left(\max\left(\overrightarrow{n}, \overleftarrow{n}\right), m\right) & \text{if } m \geqslant \overrightarrow{n} \\ \max\left(\overrightarrow{n}, m\right) - \min\left((\overrightarrow{n} - m)^+, m\right) & \\ + \left(\min\left((\overrightarrow{n} - m)^+, m\right) - \left(\max\left(\overrightarrow{n}, m\right) - \overleftarrow{n}\right)\right)^+ & \text{if } m < \overrightarrow{n} \end{cases}. \tag{3.22}$$

3.4. Connecting Linear Deterministic and Gaussian Interference Channels

THE capacity region of the two-user GIC-NOF with parameters $\overrightarrow{\mathrm{SNR}}_1$, $\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, INR_{21}, $\overleftarrow{\mathrm{SNR}}_1$ and $\overleftarrow{\mathrm{SNR}}_2$ can be approximated by the capacity region of an LDIC-NOF with parameters $\overrightarrow{n}_{ii} = \left\lfloor \frac{1}{2}\log(\overrightarrow{\mathrm{SNR}}_i)\right\rfloor$; $n_{ij} = \left\lfloor \frac{1}{2}\log(\mathrm{INR}_{ij})\right\rfloor$; $\overleftarrow{n}_{ii} = \left\lfloor \frac{1}{2}\log(\overleftarrow{\mathrm{SNR}}_i)\right\rfloor$, with $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$. For instance, in the case without feedback, the capacity region of any GIC with parameters $\overrightarrow{\mathrm{SNR}}_1 > 1$, $\overrightarrow{\mathrm{SNR}}_2 > 1$, $\mathrm{INR}_{12} > 1$ and $\mathrm{INR}_{21} > 1$ is within 18.6 bits per channel use per user of the capacity of an LDIC with parameters $\overrightarrow{n}_{11} = \lfloor \frac{1}{2}\log(\overrightarrow{\mathrm{SNR}}_1)\rfloor$, $\overrightarrow{n}_{22} = \lfloor \frac{1}{2}\log(\overrightarrow{\mathrm{SNR}}_2)\rfloor$, $n_{12} = \lfloor \frac{1}{2}\log(\mathrm{INR}_{12})\rfloor$, and $n_{21} = \lfloor \frac{1}{2}\log(\mathrm{INR}_{21})\rfloor$ (Theorem 2 in [20]). More specifically, if the capacity region of the two-user GIC and the two-user LDIC without feedback are denoted by $\mathcal{C}_{\mathrm{G}}$ and $\mathcal{C}_{\mathrm{LD}}$, respectively, the following holds:

$$\mathcal{C}_{\mathrm{LD}} \subseteq \mathcal{C}_{\mathrm{G}} + (5,5) \text{ and} \tag{3.23a}$$

$$\mathcal{C}_{\mathrm{G}} \subseteq \mathcal{C}_{\mathrm{LD}} + (13.6, 13.6). \tag{3.23b}$$

In a more general setting, for instance in the case with NOF, the two-user LDIC is known to be a close approximation of the two-user GIC. In Section 5.4, this approximation is used to simplify the identification of the cases in which channel-output feedback, even subject to additive noise, enlarges the capacity region of the two-user GIC.

Part II.

CONTRIBUTIONS TO CENTRALIZED INTERFERENCE CHANNELS

— 4 —

Linear Deterministic Interference Channel

THis chapter presents the main results on the two-user centralized LDIC-NOF described in Section 2.2.

4.1. Capacity Region

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 two-user LDIC-NOF with parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$, and $\overleftarrow{n}_{22}$, characterized in Theorem 1.

Theorem 1. Capacity Region

The capacity region $\mathcal{C}(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ of the two-user LDIC-NOF is the set of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that for all $i \in \{1,2\}$, with $j \in \{1,2\} \setminus \{i\}$ satisfy:

$$R_i \leqslant \min\left(\max\left(\overrightarrow{n}_{ii}, n_{ji}\right), \max\left(\overrightarrow{n}_{ii}, n_{ij}\right)\right), \tag{4.1a}$$

$$R_i \leqslant \min\left(\max\left(\overrightarrow{n}_{ii}, n_{ji}\right), \max\left(\overrightarrow{n}_{ii}, \overleftarrow{n}_{jj} - \left(\overrightarrow{n}_{jj} - n_{ji}\right)^+\right)\right), \tag{4.1b}$$

$$\begin{aligned} R_1 + R_2 \leqslant & \min\left(\max\left(\overrightarrow{n}_{22}, n_{12}\right) + \left(\overrightarrow{n}_{11} - n_{12}\right)^+, \max\left(\overrightarrow{n}_{11}, n_{21}\right) + \left(\overrightarrow{n}_{22} - n_{21}\right)^+\right), \\ R_1 + R_2 \leqslant & \max\left(\left(\overrightarrow{n}_{11} - n_{12}\right)^+, n_{21}, \overrightarrow{n}_{11} - \left(\max\left(\overrightarrow{n}_{11}, n_{12}\right) - \overleftarrow{n}_{11}\right)^+\right) \\ & + \max\left(\left(\overrightarrow{n}_{22} - n_{21}\right)^+, n_{12}, \overrightarrow{n}_{22} - \left(\max\left(\overrightarrow{n}_{22}, n_{21}\right) - \overleftarrow{n}_{22}\right)^+\right), \end{aligned} \tag{4.1c}$$

$$\begin{aligned} 2R_i + R_j \leqslant & \max\left(\overrightarrow{n}_{ii}, n_{ji}\right) + \left(\overrightarrow{n}_{ii} - n_{ij}\right)^+ \\ & + \max\left(\left(\overrightarrow{n}_{jj} - n_{ji}\right)^+, n_{ij}, \overrightarrow{n}_{jj} - \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+\right). \end{aligned} \tag{4.1d}$$

The proof of Theorem 1 is divided into two parts. The first part describes the achievable region and is presented in Appendix A. The second part describes the converse region and is presented in Appendix B.

Theorem 1 generalizes previous results regarding the capacity region of the two-user LDIC with channel-output feedback. For instance, when $\overleftarrow{n}_{11} = 0$ and $\overleftarrow{n}_{22} = 0$, Theorem 1 describes the capacity region of the two-user LDIC without feedback (Lemma 4 in [20]); when $\overleftarrow{n}_{11} \geqslant \max(\overrightarrow{n}_{11}, n_{12})$ and $\overleftarrow{n}_{22} \geqslant \max(\overrightarrow{n}_{22}, n_{21})$, Theorem 1 describes the capacity region of the two-user LDIC with perfect channel output feedback (LDIC-POF) (Corollary 1 in [88]); when $\overrightarrow{n}_{11} = \overrightarrow{n}_{22}$, $n_{12} = n_{21}$ and $\overleftarrow{n}_{11} = \overleftarrow{n}_{22}$, Theorem 1 describes the capacity region of the two-user symmetric LDIC-NOF (Theorem 1 in [53] and Theorem 4.1, Case 1001 in [80]); and when $\overrightarrow{n}_{11} = \overrightarrow{n}_{22}$, $n_{12} = n_{21}$, $\overleftarrow{n}_{ii} \geqslant \max(\overrightarrow{n}_{ii}, n_{ij})$ and $\overleftarrow{n}_{jj} = 0$, with $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$, Theorem 1 describes the capacity region of the two-user symmetric LDIC with only one perfect channel output feedback (Theorem 4.1, Cases 1000 and 0001 in [80]).

An interesting observation from Theorem 1 is that feedback is beneficial only when at least one of the feedback parameters $\overleftarrow{n}_{11}$ or $\overleftarrow{n}_{22}$ is beyond a certain threshold (see Section 4.2). For instance, note that when $\overleftarrow{n}_{ii} \leqslant (\overrightarrow{n}_{ii} - n_{ij})^+$, receiver i is unable to send to its corresponding transmitter via feedback any information about the message sent by transmitter j, and thus, feedback does not play any role to enlarge the capacity region. This is basically because the bit-pipes that are subject to interference at receiver i are not included in the set of bit-pipes that are above the (feedback) noise level. However, the threshold $(\overrightarrow{n}_{ii} - n_{ij})^+$ for $\overleftarrow{n}_{ii}$ is necessary but not sufficient for feedback to enlarge the capacity region. Consider for instance the following examples:

Example 1. Consider the two-user LDIC-NOF with parameters $\overrightarrow{n}_{11} = 5$, $\overrightarrow{n}_{22} = 1$, $\overrightarrow{n}_{12} = 3$, $\overrightarrow{n}_{21} = 4$, and $\overleftarrow{n}_{22} = 0$. The capacity regions $\mathcal{C}(5, 1, 3, 4, 0, 0)$ and $\mathcal{C}(5, 1, 3, 4, 4, 0)$ are shown in Figure 4.1a. In this case, channel-output feedback in the transmitter-receiver pair 1 enlarges the capacity region only when $\overleftarrow{n}_{11} > \overrightarrow{n}_{22} + (\overrightarrow{n}_{11} - n_{12})^+ = 3$. More specifically, for all $\overleftarrow{n}_{11} \in \{0, \ldots, 3\}$,

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

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

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

In Example 1, in the absence of channel-output feedback, the rate R_2 cannot exceed 1 bit per channel use, whereas the sum-rate $R_1 + R_2$ is not greater than by 5 bits per channel use. Figure 4.1b shows a simple achievability scheme for the rate pair $(3, 1)$. Note that R_2 cannot be improved by letting transmitter 2 use the bit-pipes $\boldsymbol{X}_{2,n}^{(2:5)}$ as they are not observed at receiver 2. When channel-output feedback is available at least at transmitter-receiver pair 1 and the bit-pipe from transmitter 2 ending at $\overrightarrow{Y}_{1,n}^{(4)}$ is included in the feedback signal $\overleftarrow{\boldsymbol{Y}}_{i,n}$, the bit-pipe $X_{2,n}^{(2)}$ can be used by transmitter 2 as feedback provides a path between transmitter 2 and receiver 2: transmitter 2 – receiver 1 – transmitter 1 – receiver 2. For this alternative path to become available at least the $\left(\overrightarrow{n}_{22} + (\overrightarrow{n}_{11} - n_{12})^+ + 1\right)$-th (feedback) bit-pipe from receiver 1 to transmitter 1 must be above the noise level, *i.e.*, $\overleftarrow{n}_{11} > \overrightarrow{n}_{22} + (\overrightarrow{n}_{11} - n_{12})^+$.

Example 2. Consider an LDIC-NOF with parameters $\overrightarrow{n}_{11} = 7$, $\overrightarrow{n}_{22} = 7$, $n_{12} = 3$, $n_{21} = 5$, and $\overleftarrow{n}_{22} = 0$. The capacity regions $\mathcal{C}(7, 7, 3, 5, 0, 0)$ and $\mathcal{C}(7, 7, 3, 5, 6, 0)$ are shown in Figure 4.2a. In this case, channel-output feedback in the transmitter-receiver pair 1 enlarges

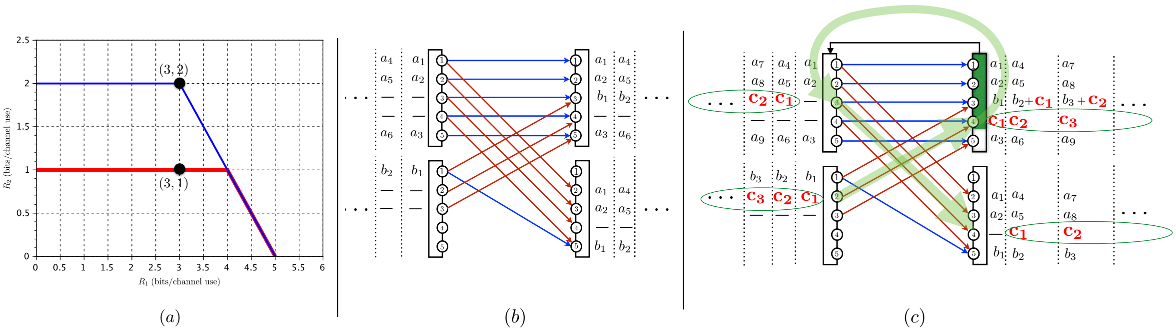


Figure 4.1.: (a) Capacity regions of $\mathcal{C}(5,1,3,4,0,0)$ (thick red line) and $\mathcal{C}(5,1,3,4,4,0)$ (thin blue line). (b) Achievability of the rate pair $(3,1)$ in an LDIC-NOF with parameters $\overrightarrow{n}_{11}=5$, $\overrightarrow{n}_{22}=1$, $n_{12}=3$, $n_{21}=4$, $\overleftarrow{n}_{11}=0$ and $\overleftarrow{n}_{22}=0$ (no feedback links). (c) Achievability of the rate pair $(3,2)$ in an LDIC-NOF with parameters $\overrightarrow{n}_{11}=5$, $\overrightarrow{n}_{22}=1$, $n_{12}=3$, $n_{21}=4$, $\overleftarrow{n}_{11}=4$ and $\overleftarrow{n}_{22}=0$.

the capacity region only when $\overleftarrow{n}_{11} > \max\left(n_{21}, (\overrightarrow{n}_{11}-n_{12})^{+}\right) = 5$. More specifically, for all $\overleftarrow{n}_{11} \in \{0,1,\ldots,5\}$,

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

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

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

In Example 2, in the absence of feedback, the sum-rate capacity can be achieved by simultaneously using two groups of bit-pipes: (a) all bit-pipes starting at transmitter i and being exclusively observed by receiver i; and (b) all bit-pipes starting at transmitter i that are observed at receiver j but do not interfere with the first group of bit-pipes, with $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$. Figure 4.2b shows an achievability scheme that uses this idea and achieves the sum-rate capacity. Note that using other bit-pipes to increase any of the individual rates produces interference that cannot be resolved and thus, impedes reliable decoding. In particular note that $X_{2,n}^{(2)}$ and $X_{2,n}^{(3)}$ must remain unused. When feedback is available at least at transmitter-receiver pair 1 and the bit-pipe from transmitter 2 ending at $\overrightarrow{Y}_{1,n}^{(6)}$ is included in the feedback signal $\overleftarrow{\boldsymbol{Y}}_{1,n}$, the bit-pipe $X_{2,n}^{(2)}$ can be used for transmitting maximum-entropy i.i.d. bits to increase the individual rate R_2 and the sum-rate (see Figure 4.2c). This is mainly because the bits $X_{2,n}^{(2)}$ can be decoded by transmitter 1 via feedback and be re-transmitted to resolve interference at receiver 1. Interestingly, during the re-transmission by transmitter 1 these bits produce an interference that can be resolved by receiver 2, as these bits have been received interference-free in the previous channel uses. Note that for this to be possible, at least one of the bit-pipes of transmitter 2 that does not belong to either of the two groups mentioned above, *i.e.*, $X_{2,n}^{(2)}$ and $X_{2,n}^{(3)}$, must be observed above the noise level in the feedback link of the transmitter-receiver pair 1, *i.e.*, $\overleftarrow{n}_{11} > 5$.

The exact thresholds for the feedback parameters $\overleftarrow{n}_{11}$ or $\overleftarrow{n}_{22}$ beyond which the capacity region is enlarged are strongly dependent on the parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, and n_{21}. A

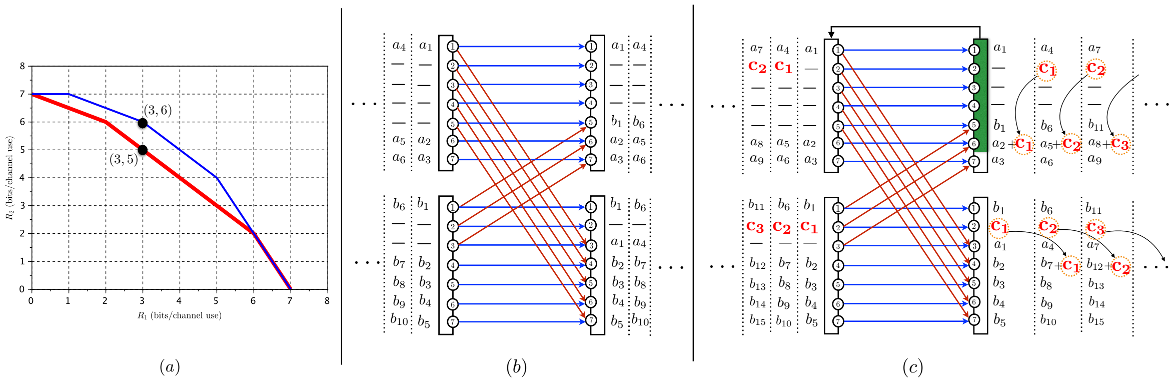


Figure 4.2.: (a) Capacity regions of $\mathcal{C}(7,7,3,5,0,0)$ (thick red line) and $\mathcal{C}(7,7,3,5,6,0)$ (thin blue line). (b) Achievability of the rate pair $(3,5)$ in an LDIC-NOF with parameters $\overrightarrow{n}_{11}=7$, $\overrightarrow{n}_{22}=7$, $n_{12}=3$, $n_{21}=5$, $\overleftarrow{n}_{11}=0$ and $\overleftarrow{n}_{22}=0$ (no feedback links). (c) Achievability of the rate pair $(3,6)$ in an LDIC-NOF with parameters $\overrightarrow{n}_{11}=7$, $\overrightarrow{n}_{22}=7$, $n_{12}=3$, $n_{21}=5$, $\overleftarrow{n}_{11}=6$ and $\overleftarrow{n}_{22}=0$.

characterization of these thresholds is presented in Section 4.2.

4.1.1. Comments on the Achievability Scheme

The achievable region is obtained using a coding scheme that combines classical tools such as rate-splitting [23, 37], block Markov superposition coding [14, 27], and backward decoding [98, 99]. This coding scheme is described in Appendix A. In the following, a brief description of this coding scheme is presented. Let the message index sent by transmitter i during the t-th block be denoted by $W_i^{(t)} \in \{1,2,\ldots,2^{NR_i}\}$. Following a rate-splitting argument, assume that $W_i^{(t)}$ is represented by three sub-indices $(W_{i,C1}^{(t)}, W_{i,C2}^{(t)}, W_{i,P}^{(t)}) \in \{1,2,\ldots,2^{NR_{i,C1}}\} \times \{1,2,\ldots,2^{NR_{i,C2}}\} \times \{1,2,\ldots,2^{NR_{i,P}}\}$, where $R_{i,C1}+R_{i,C2}+R_{i,P}=R_i$. The codeword generation follows a four-level superposition coding scheme. The index $W_{i,C1}^{(t-1)}$ is assumed to be decoded at transmitter j via the feedback link of the transmitter-receiver pair j at the end of the transmission of block $t-1$. Therefore, at the beginning of block t, each transmitter possesses the knowledge of the indices $W_{1,C1}^{(t-1)}$ and $W_{2,C1}^{(t-1)}$. In the case of the first block $t=1$, the indices $W_{1,C1}^{(0)}$ and $W_{2,C1}^{(0)}$ correspond to two indices assumed to be known by all transmitters and receivers. Using these indices both transmitters are able to identify the same codeword in the first code-layer. This first code-layer is a sub-codebook of $2^{N(R_{1,C1}+R_{2,C1})}$ codewords (see Figure A.1). Denote by $\boldsymbol{u}\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}\right)$ the corresponding codeword in the first code-layer. The second codeword used by transmitter i is selected using $W_{i,C1}^{(t)}$ from the second code-layer, which is a sub-codebook of $2^{N\,R_{i,C1}}$ codewords associated to $\boldsymbol{u}\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}\right)$ as shown in Figure A.1. Denote by $\boldsymbol{u}_i\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}\right)$ the corresponding codeword in the second code-layer. The third codeword used by transmitter i is selected using $W_{i,C2}^{(t)}$ from the third code-layer, which is a sub-codebook of $2^{N\,R_{i,C2}}$ codewords associated to $\boldsymbol{u}_i\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{t,C1}^{(t)}\right)$ as shown in Figure A.1. Denote by $\boldsymbol{v}_i\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)}\right)$ the corresponding codeword in the third code-layer. The fourth codeword used by transmitter i is selected using $W_{i,P}^{(t)}$ from the fourth code-layer, which is a sub-codebook of $2^{N R_{i,P}}$ codewords associated to $\boldsymbol{v}_i\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)}\right)$ as shown in Figure A.1. Denote by $\boldsymbol{x}_{i,P}\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)}, W_{i,P}^{(t)}\right)$ the corresponding codeword in the fourth code-layer. Finally, the generation of the codeword $\boldsymbol{x}_i = (\boldsymbol{x}_{i,1}, \boldsymbol{x}_{i,2}, \ldots, \boldsymbol{x}_{i,N}) \in \mathcal{C}_i \subseteq \mathcal{X}_i^N$ during block $t \in \{1, 2, \ldots, T\}$ is a simple concatenation of the codewords $\boldsymbol{u}_i\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}\right)$, $\boldsymbol{v}_i\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)}\right)$ and $\boldsymbol{x}_{i,P}\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)}, W_{i,P}^{(t)}\right)$, *i.e.*, $\boldsymbol{x}_i = \left(\boldsymbol{u}_i^{\mathsf{T}}, \boldsymbol{v}_i^{\mathsf{T}}, \boldsymbol{x}_{i,P}^{\mathsf{T}}\right)^{\mathsf{T}}$, where the message indices have been dropped for ease of notation.

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

These three types of bit-pipes justify the three code-layers superposed over a common layer, which is justified by the fact that feedback allows both transmitters to decode part of the message sent by each other. The decoder follows a classical backward decoding scheme. This coding/decoding scheme is thoroughly described in Appendix A in the most general case. Later, it is particularized for the case of the two-user LDIC-NOF and two-user GIC-NOF.

Other achievable schemes, as reported in [53], can also be obtained as special cases of the more general scheme presented in [94]. However, in this more general case, the resulting code for the IC-NOF counts with a handful of unnecessary superposing code-layers, which demands further optimization. This observation becomes clearer in the analysis of the two-user GIC-NOF in Chapter 5.

4.1.2. Comments on the Converse Region

The outer bounds (4.1a) and (4.1c) are cut-set bounds and were first reported in [20] for the case without feedback. These outer bounds are still useful in the case of POF [88]. The outer bounds (4.1b), (4.1c) and (4.1d) are new.

Consider the notation used in Appendix B (See Figure B.1 and Figure B.2). The outer bound (4.1b) on the individual rate i is a cut-set bound at the input of an enhanced version of receiver i. More specifically, this outer bound is calculated considering that receiver i possesses the message index of transmitter j, *i.e.*, W_j, as side information and observes the channel output $\overrightarrow{\boldsymbol{Y}}_i$ and the feedback signal $\overleftarrow{\boldsymbol{Y}}_j$ of the transmitter-receiver pair j at each channel use. A complete proof of (4.1b) is presented in Appendix B.

The intuition behind the outer bound (4.1c) follows from the observation that in the absence of feedback, the sum-rate is upper-bounded by the sum of the bit-pipes from transmitter i that are exclusively observed by receiver i (denoted by $\boldsymbol{X}_{i,P}$) and the bit-pipes from transmitter i that are observed by receiver j and do not interfere with bit-pipes $\boldsymbol{X}_{j,P}$ (denoted by $\boldsymbol{X}_{i,U}$), with $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$. More specifically, in the absence of feedback:

$$R_1 + R_2 \leqslant \sum_{i=1}^{2} \dim \boldsymbol{X}_{i,P} + \dim \boldsymbol{X}_{i,U}. \tag{4.2}$$

When $R_1 + R_2 = \sum_{i=1}^{2} \dim \boldsymbol{X}_{i,P} + \dim \boldsymbol{X}_{i,U}$ is achievable without feedback, the bit-pipes $\boldsymbol{X}_{i,P}$ and $\boldsymbol{X}_{i,U}$ can be used for sending maximum-entropy i.i.d. bits from transmitter i to receiver i, which maximizes the sum-rate. Interestingly, any attempt of using any of the other bit-pipes creates interference that cannot be resolved and thus impedes reliable decoding. This observation is formally proved in Appendix B (see proof of (4.1c)). Note also that this outer bound is not necessarily tight (see Example 1). When feedback is available at least at transmitter-receiver pair i, other bit-pipes different from $\boldsymbol{X}_{j,P}$ and $\boldsymbol{X}_{j,U}$ might be used by transmitter j for simultaneously increasing the rate R_j and the sum-rate (see Example 2). This simple observation suggests that there must exist an upper-bound on the sum-rate of the form:

$$R_1 + R_2 \leqslant \sum_{i=1}^{2} \dim \boldsymbol{X}_{i,P} + \dim \boldsymbol{X}_{i,U} + F_i, \tag{4.3}$$

where, $F_i \leqslant \dim \boldsymbol{X}_{i,C} + \dim \boldsymbol{X}_{i,D}$ represents the bit-pipes other than $\boldsymbol{X}_{i,P}$ and $\boldsymbol{X}_{i,U}$, whose origin is at transmitter i, that can be used for sending maximum-entropy i.i.d. bits from transmitter i to receiver i, while generating an interference that can be resolved by the use of feedback. Following this idea, the following outer bound is presented in Appendix B (see proof of (4.1c)):

$$R_1 + R_2 \leqslant \sum_{i=1}^{2} \dim \boldsymbol{X}_{i,P} + \dim \boldsymbol{X}_{i,U} + \dim \boldsymbol{X}_{i,CF_j} + \dim \boldsymbol{X}_{i,DF}, \tag{4.4}$$

where $\dim \left(\boldsymbol{X}_{i,CF_j}, \boldsymbol{X}_{i,DF} \right)$ is the number of the bit-pipes whose origin is at transmitter i and are observed above the noise level in the feedback link of transmitter-receiver pair j. The outer bound (4.4) is derived considering genie-aided receivers. More specifically, receiver i has inputs $\overrightarrow{\boldsymbol{Y}}_i$ and $\overleftarrow{\boldsymbol{Y}}_i$, with $i \in \{1,2\}$.

A similar reasoning is followed to derive the outer bound (4.1d) considering three genie-aided receivers. More specifically, receiver i has inputs $\overrightarrow{\boldsymbol{Y}}_i$ and $\overleftarrow{\boldsymbol{Y}}_i$, with $i \in \{1,2\}$, and a third receiver has inputs $\overrightarrow{\boldsymbol{Y}}_i$, $\overleftarrow{\boldsymbol{Y}}_j$, and W_j for at most one $i \in \{1,2\}$, with $j \in \{1,2\} \setminus \{i\}$.

4.2. Cases in which Feedback Enlarges the Capacity Region

Let $\alpha_i \in \mathbb{Q}$, with $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$ be defined as

$$\alpha_i = \frac{n_{ij}}{\overrightarrow{n}_{ii}}. \tag{4.5}$$

For each transmitter-receiver pair i, there exist five possible interference regimes (IRs), as suggested in [31]: the very weak IR (VWIR), *i.e.*, $\alpha_i \leqslant \frac{1}{2}$, the weak IR (WIR), *i.e.*, $\frac{1}{2} < \alpha_i \leqslant \frac{2}{3}$, the moderate IR (MIR), *i.e.*, $\frac{2}{3} < \alpha_i < 1$, the strong IR (SIR), *i.e.*, $1 \leqslant \alpha_i \leqslant 2$ and the very strong IR (VSIR), *i.e.*, $\alpha_i > 2$. The scenarios in which the desired signal is stronger than the interference ($\alpha_i < 1$), namely the VWIR, the WIR, and the MIR, are referred to as the low-interference regimes (LIRs). Conversely, the scenarios in which the desired signal is weaker than or equal to the interference ($\alpha_i \geqslant 1$), namely the SIR and the VSIR, are referred to as the high-interference regimes (HIRs).

The main results of this section are presented using a set of events (Boolean variables) that are determined by the parameters $\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}$, and n_{21}. Given a fixed 4-tuple ($\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}), the events are defined below:

$$
\begin{aligned}
E_1 :& \quad \alpha_1 < 1 \wedge \alpha_2 < 1, && (4.6)\\
E_{2,i} :& \quad \alpha_i \leqslant \frac{1}{2} \wedge 1 \leqslant \alpha_j \leqslant 2, && (4.7)\\
E_{3,i} :& \quad \alpha_i \leqslant \frac{1}{2} \wedge \alpha_j > 2, && (4.8)\\
E_{4,i} :& \quad \frac{1}{2} < \alpha_i \leqslant \frac{2}{3} \wedge \alpha_j \geqslant 1, && (4.9)\\
E_{5,i} :& \quad \frac{2}{3} < \alpha_i < 1 \wedge \alpha_j \geqslant 1, && (4.10)\\
E_{6,i} :& \quad \frac{1}{2} < \alpha_i \leqslant 1 \wedge \alpha_j > 1, && (4.11)\\
E_{7,i} :& \quad \alpha_i \geqslant 1 \wedge \alpha_j \leqslant 1, && (4.12)\\
E_{8,i} :& \quad \overrightarrow{n}_{ii} > n_{ji}, && (4.13)\\
E_9 :& \quad \overrightarrow{n}_{11} + \overrightarrow{n}_{22} > n_{12} + n_{21}, && (4.14)\\
E_{10,i} :& \quad \overrightarrow{n}_{ii} + \overrightarrow{n}_{jj} > n_{ij} + 2n_{ji}, && (4.15)\\
E_{11,i} :& \quad \overrightarrow{n}_{ii} + \overrightarrow{n}_{jj} < n_{ij}. && (4.16)
\end{aligned}
$$

In the following, in the case of $E_{8,i} : \overrightarrow{n}_{ii} > n_{ji}$, the notation $\widetilde{E}_{8,i}$ indicates $\overrightarrow{n}_{ii} < n_{ji}$; the notation $\overline{E}_{8,i}$ indicates $\overrightarrow{n}_{ii} \leqslant n_{ji}$ (logical complement); and the notation $\check{E}_{8,i}$ indicates $\overrightarrow{n}_{ii} \geqslant n_{ji}$. In the case $E_1 : \alpha_1 < 1 \wedge \alpha_2 < 1$, the notation $\widetilde{E}_1$ indicates $\alpha_1 > 1 \wedge \alpha_2 > 1$; and the notation $\overline{E}_1$ indicates $\alpha_1 \geqslant 1 \wedge \alpha_2 \geqslant 1$. In the case $E_9 : \overrightarrow{n}_{11} + \overrightarrow{n}_{22} > n_{12} + n_{21}$, the notation $\overline{E}_9$ indicates $\overrightarrow{n}_{11} + \overrightarrow{n}_{22} \leqslant n_{12} + n_{21}$.

Combining the events (4.6)-(4.16), five main scenarios are identified:

$$
\begin{aligned}
S_{1,i}:& (E_1 \wedge E_{8,i}) \vee (E_{2,i} \wedge E_{8,i}) \vee (E_{3,i} \wedge E_{8,i} \wedge E_9) \vee (E_{4,i} \wedge E_{8,i} \wedge E_9) \vee (E_{5,i} \wedge E_{8,i} \wedge E_9), && (4.17)\\
S_{2,i}:& \left(E_{3,i} \wedge \widetilde{E}_{8,j} \wedge \overline{E}_9\right) \vee \left(E_{6,i} \wedge \widetilde{E}_{8,j} \wedge \overline{E}_9\right) \vee \left(\widetilde{E}_1 \wedge \widetilde{E}_{8,j}\right), && (4.18)\\
S_{3,i}:& \left(E_1 \wedge \overline{E}_{8,i}\right) \vee \left(E_{2,i} \wedge \overline{E}_{8,i}\right) \vee \left(E_{3,i} \wedge \check{E}_{8,j} \wedge \overline{E}_{8,i}\right) \vee \left(E_{4,i} \wedge \check{E}_{8,j} \wedge \overline{E}_{8,i}\right)\\
& \vee \left(E_{5,i} \wedge \check{E}_{8,j} \wedge \overline{E}_{8,i}\right) \vee \left(\overline{E}_1 \wedge \check{E}_{8,j}\right) \vee (E_{7,i}), && (4.19)\\
S_4 \ :& E_1 \wedge E_{8,1} \wedge E_{8,2} \wedge E_{10,1} \wedge E_{10,2}, && (4.20)\\
S_5 \ :& \overline{E}_1 \wedge E_{11,1} \wedge E_{11,2}. && (4.21)
\end{aligned}
$$

For all $i \in \{1, 2\}$, the events $S_{1,i}$, $S_{2,i}$, $S_{3,i}$, S_4 and S_5 exhibit the properties stated by the following corollaries:

Corollary 1. For all $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$, given a fixed $i \in \{1, 2\}$, only one of the events $S_{1,i}$, $S_{2,i}$ and $S_{3,i}$ holds true.

Corollary 2. For all $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$, when one of the events S_4 or S_5 holds true, then the other necessarily holds false.

Note that Corollary 2 does not exclude the case in which both S_4 and S_5 simultaneously hold false.

Corollary 3. For all $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$, when S_4 holds true, then both $S_{1,1}$ and $S_{1,2}$ hold true; and when S_5 holds true, then both $S_{2,1}$ and $S_{2,2}$ hold true.

4.2.1. Rate Improvement Metrics

Given a fixed 4-tuple $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21})$, let $\mathcal{C}(\overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ be the capacity region of an LDIC-NOF with parameters $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$. The maximum improvement of the individual rates R_1 and R_2, denoted by $\Delta_1(\overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ and $\Delta_2(\overleftarrow{n}_{11}, \overleftarrow{n}_{22})$, due to the effect of channel-output feedback with respect to the case without feedback is:

$$\Delta_1(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) = \max_{0<R_2<R_2^*} \left\{ \sup\left\{R_1 : (R_1, R_2) \in \mathcal{C}(\overleftarrow{n}_{11}, \overleftarrow{n}_{22})\right\} - \sup\left\{R_1^\dagger : (R_1^\dagger, R_2) \in \mathcal{C}(0,0)\right\} \right\}$$

and (4.22)

$$\Delta_2(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) = \max_{0<R_1<R_1^*} \left\{ \sup\left\{R_2 : (R_1, R_2) \in \mathcal{C}(\overleftarrow{n}_{11}, \overleftarrow{n}_{22})\right\} - \sup\left\{R_2^\dagger : (R_1, R_2^\dagger) \in \mathcal{C}(0,0)\right\} \right\}, \quad (4.23)$$

with

$$R_1^* = \sup\left\{r_1 : (r_1, r_2) \in \mathcal{C}(0,0)\right\} \text{ and} \quad (4.24)$$

$$R_2^* = \sup\left\{r_2 : (r_1, r_2) \in \mathcal{C}(0,0)\right\}. \quad (4.25)$$

Note that for a fixed $i \in \{1, 2\}$, $\Delta_i(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) > 0$ if and only if it is possible to achieve a rate pair $(R_1, R_2) \in \mathbb{R}_+^2$ with channel-output feedback such that R_i is greater than the maximum rate achievable by transmitter-receiver i without feedback when the rate of transmitter-receiver pair j is fixed at R_j. In the following, given fixed parameters $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$, the statement *"the rate R_i is improved by using feedback"* is used to indicate that $\Delta_i(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) > 0$.

Alternatively, the maximum improvement of the sum-rate $\Sigma(\overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ with respect to the case without feedback is:

$$\Sigma(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) = \sup\left\{R_1 + R_2 : (R_1, R_2) \in \mathcal{C}(\overleftarrow{n}_{11}, \overleftarrow{n}_{22})\right\} - \sup\left\{R_1^\dagger + R_2^\dagger : (R_1^\dagger, R_2^\dagger) \in \mathcal{C}(0,0)\right\}. \quad (4.26)$$

Note that $\Sigma(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) > 0$ if and only if there exists a rate pair with feedback whose sum is greater than the maximum sum-rate achievable without feedback. In the following, given fixed parameters $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$, the statement *"the sum-rate is improved by using feedback"* is used to imply that $\Sigma(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) > 0$. When feedback is exclusively used by transmitter-receiver pair i, *i.e.*, $\overleftarrow{n}_{ii} > 0$ and $\overleftarrow{n}_{jj} = 0$, then the maximum improvement of the individual rate of transmitter-receiver k, with $k \in \{1, 2\}$, and the maximum improvement of the sum-rate are denoted by $\Delta_k(\overleftarrow{n}_{ii})$ and $\Sigma(\overleftarrow{n}_{ii})$, respectively. Hence, this notation $\Delta_k(\overleftarrow{n}_{ii})$ replaces either $\Delta_k(\overleftarrow{n}_{11}, 0)$ or $\Delta_k(0, \overleftarrow{n}_{22})$, when $i = 1$ or $i = 2$, respectively. The same holds for the notation $\Sigma(\overleftarrow{n}_{ii})$ that replaces $\Sigma(\overleftarrow{n}_{11}, 0)$ or $\Sigma(0, \overleftarrow{n}_{22})$, when $i = 1$ or $i = 2$, respectively.

4.2.2. Enlargement of the Capacity Region

Given fixed parameters $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21})$, $i \in \{1,2\}$, and $j \in \{1,2\} \setminus \{i\}$, the capacity region of a two-user LDIC-NOF, when feedback is available only at transmitter-receiver pair i, *i.e.*, $\overleftarrow{n}_{ii} > 0$ and $\overleftarrow{n}_{jj} = 0$, is denoted by $\mathcal{C}(\overleftarrow{n}_{ii})$ instead of $\mathcal{C}(\overleftarrow{n}_{11}, 0)$ or $\mathcal{C}(0, \overleftarrow{n}_{22})$, when $i = 1$ or $i = 2$, respectively. Following this notation, Theorem 2 identifies the exact values of $\overleftarrow{n}_{ii}$ for which the strict inclusion $\mathcal{C}(0,0) \subset \mathcal{C}(\overleftarrow{n}_{ii})$ holds for $i \in \{1,2\}$.

Theorem 2. Enlargement of the Capacity Region

Let $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$ be a fixed 4-tuple. Let also $i \in \{1,2\}$, $j \in \{1,2\} \setminus \{i\}$ and $\overleftarrow{n}^*_{ii} \in \mathbb{N}$ be fixed integers, with

$$\overleftarrow{n}^*_{ii} = \begin{cases} \max\left(n_{ji}, (\overrightarrow{n}_{ii} - n_{ij})^+\right) & \text{if } S_{1,i} \text{ holds true} \\ \overrightarrow{n}_{jj} + (\overrightarrow{n}_{ii} - n_{ij})^+ & \text{if } S_{2,i} \text{ holds true} \end{cases}. \tag{4.27}$$

Assume that $S_{3,i}$ holds true. Then, for all $\overleftarrow{n}_{ii} \in \mathbb{N}$, $\mathcal{C}\Big(0,0\Big) = \mathcal{C}\Big(\overleftarrow{n}_{ii}\Big)$. Assume that either $S_{1,i}$ holds true or $S_{2,i}$ holds true. Then, for all $\overleftarrow{n}_{ii} \leqslant \overleftarrow{n}^*_{ii}$, $\mathcal{C}\Big(0,0\Big) = \mathcal{C}\Big(\overleftarrow{n}_{ii}\Big)$ and for all $\overleftarrow{n}_{ii} > \overleftarrow{n}^*_{ii}$, $\mathcal{C}\Big(0,0\Big) \subset \mathcal{C}\Big(\overleftarrow{n}_{ii}\Big)$.

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

Theorem 2 shows that under event $S_{3,i}$ in (4.19), implementing feedback in transmitter-receiver pair i, with any $\overleftarrow{n}_{ii} > 0$ and $\overleftarrow{n}_{jj} = 0$, does not enlarge the capacity region. Note that when both $E_{8,i}$ and $\widetilde{E}_{8,j}$ hold false, then both $S_{1,i}$ and $S_{2,i}$ hold false, which implies that $S_{3,i}$ holds true (Corollary 1). The following remark is a consequence of this observation.

Remark 1: *A necessary but not sufficient condition for enlarging the capacity region by using feedback in transmitter-receiver pair i is: there exists at least one transmitter able to send more information bits to receiver i than to receiver j,* i.e., $\overrightarrow{n}_{ii} > n_{ji}$ *(Event* $E_{8,i}$*) or* $n_{ij} > \overrightarrow{n}_{jj}$ *(Event* $\widetilde{E}_{8,j}$*).*

Alternatively, under events $S_{1,i}$ in (4.17) and $S_{2,i}$ in (4.18), the capacity region can be enlarged when $\overleftarrow{n}_{ii} > \overleftarrow{n}^*_{ii}$. It is important to highlight that in the cases in which feedback enlarges the capacity region of the two-user LDIC-NOF, that is, in events $S_{1,1}$, $S_{2,1}$, $S_{1,2}$ or $S_{2,2}$, for all $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$, the following always holds true:

$$\overleftarrow{n}^*_{ii} > (\overrightarrow{n}_{ii} - n_{ij})^+. \tag{4.28}$$

Essentially, the inequality in (4.28) unveils a necessary but not sufficient condition to enlarge the capacity region using channel-output feedback. This condition is that for at least one $i \in \{1,2\}$, with $j \in \{1,2\} \setminus \{i\}$, transmitter i decodes a subset of the information bits sent by transmitter j at each channel use.

Another interesting observation is that the threshold $\overleftarrow{n}^*_{ii}$ beyond which feedback is useful is different under event $S_{1,i}$ in (4.17) and event $S_{2,i}$ in (4.18). In general when $S_{1,i}$ holds true, the enlargement of the capacity region is due to the fact that feedback allows transmitter-receiver pair i using *interference as side information* [86]. Alternatively, when $S_{2,i}$ in (4.18) holds true, the enlargement of the capacity region occurs as a consequence of the fact that some of the bits that cannot be transmitted directly from transmitter j to receiver j, can arrive to receiver j via an alternative path: transmitter j - receiver i - transmitter i - receiver j.

4.2.3. Improvement of the Individual Rate R_i by Using Feedback in Link i

Given fixed parameters $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21})$, and $i \in \{1,2\}$, implementing channel-output feedback in transmitter-receiver pair i increases the individual rate R_i, *i.e.*, $\Delta_i(\overleftarrow{n}_{ii}) > 0$ for some values of $\overleftarrow{n}_{ii}$. Theorem 3 identifies the exact values of $\overleftarrow{n}_{ii}$ for which $\Delta_i(\overleftarrow{n}_{ii}) > 0$.

Theorem 3. Improvement of R_i by Using Feedback in Link i

Let $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$ be a fixed 4-tuple. Let also $i \in \{1,2\}$, $j \in \{1,2\} \setminus \{i\}$ and $\overleftarrow{n}^{\dagger}_{ii} \in \mathbb{N}$ be fixed integers, with

$$\overleftarrow{n}^{\dagger}_{ii} = \max\left(n_{ji}, (\overrightarrow{n}_{ii} - n_{ij})^{+}\right). \tag{4.29}$$

Assume that either $S_{2,i}$ holds true or $S_{3,i}$ holds true. Then, for all $\overleftarrow{n}_{ii} \in \mathbb{N}$, $\Delta_i(\overleftarrow{n}_{ii}) = 0$. Assume that $S_{1,i}$ holds true. Then, when $\overleftarrow{n}_{ii} \leqslant \overleftarrow{n}^{\dagger}_{ii}$, it holds that $\Delta_i(\overleftarrow{n}_{ii}) = 0$; and when $\overleftarrow{n}_{ii} > \overleftarrow{n}^{\dagger}_{ii}$, it holds that $\Delta_i(\overleftarrow{n}_{ii}) > 0$.

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

Theorem 3 highlights that under events $S_{2,i}$ in (4.18) and $S_{3,i}$ in (4.19), the individual rate R_i cannot be improved by using feedback in transmitter-receiver pair i, *i.e.*, $\Delta_i(\overleftarrow{n}_{ii}) = 0$. Alternatively, under event $S_{1,i}$ in (4.17), the individual rate R_i can be improved, *i.e.*, $\Delta_i\left(\overleftarrow{n}_{ii}\right) > 0$, whenever $\overleftarrow{n}_{ii} > \max\left(n_{ji}, (\overrightarrow{n}_{ii} - n_{ij})^{+}\right)$. Hence, given the definition of $S_{1,i}$, the following remark is relevant.

Remark 2: *A necessary but not sufficient condition for $\Delta_i\left(\overleftarrow{n}_{ii}\right) > 0$ is: the number of bit-pipes from transmitter i to receiver i is greater than the number of bit-pipes from transmitter i to receiver j,* i.e., *$\overrightarrow{n}_{ii} > n_{ji}$ (Event $E_{8,i}$)*

4.2.4. Improvement of the Individual Rate R_j by Using Feedback in Link i

Given fixed parameters $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21})$, $i \in \{1,2\}$, and $j \in \{1,2\} \setminus \{i\}$, implementing channel-output feedback in transmitter-receiver pair i increases the individual rate R_j, *i.e.*, $\Delta_j(\overleftarrow{n}_{ii}) > 0$ for some values of $\overleftarrow{n}_{ii}$. Theorem 4 identifies the exact values of $\overleftarrow{n}_{ii}$ for which $\Delta_j(\overleftarrow{n}_{ii}) > 0$.

Theorem 4. Improvement of R_j by Using Feedback in Link i

Let $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$ be a fixed 4-tuple. Let also $i \in \{1,2\}$, $j \in \{1,2\} \setminus \{i\}$ and $\overleftarrow{n}^{*}_{ii} \in \mathbb{N}$ given in (4.27), be fixed integers. Assume that $S_{3,i}$ holds true. Then, for all $\overleftarrow{n}_{ii} \in \mathbb{N}$, $\Delta_j(\overleftarrow{n}_{ii}) = 0$. Assume that either $S_{1,i}$ holds true or $S_{2,i}$ holds true. Then, when $\overleftarrow{n}_{ii} \leqslant \overleftarrow{n}^{*}_{ii}$, it holds that $\Delta_j(\overleftarrow{n}_{ii}) = 0$; and when $\overleftarrow{n}_{ii} > \overleftarrow{n}^{*}_{ii}$, it holds that $\Delta_j(\overleftarrow{n}_{ii}) > 0$.

Proof: The proof of Theorem 4 follows along the same lines as the proof of Theorem 3 in Appendix D. ■

Theorem 4 shows that under event $S_{3,i}$ in (4.19), implementing feedback in transmitter-receiver pair i does not bring any improvement on the rate R_j. This is in line with the results of Theorem 2. In contrast, under events $S_{1,i}$ in (4.17) and $S_{2,i}$ in (4.18), the individual rate

R_j can be improved, *i.e.*, $\Delta_j(\overleftarrow{n}_{ii}) > 0$ for all $\overleftarrow{n}_{ii} > \overleftarrow{n}^*_{ii}$. From the definition of events $S_{1,i}$ and $S_{2,i}$, the following remark holds:

Remark 3: *A necessary but not sufficient condition for* $\Delta_j\left(\overleftarrow{n}_{ii}\right) > 0$ *is: there exists at least one transmitter able to send more information bits to receiver i than to receiver j,* i.e., $\overrightarrow{n}_{ii} > n_{ji}$ *(Event $E_{8,i}$) or* $n_{ij} > \overrightarrow{n}_{jj}$ *(Event $\widetilde{E}_{8,j}$).*

It is important to highlight that under event $S_{1,i}$, the threshold on $\overleftarrow{n}_{ii}$ for increasing the individual rate R_i, *i.e.*, $\overleftarrow{n}^{\dagger}_{ii}$, and R_j, *i.e.*, $\overleftarrow{n}^*_{ii}$, are identical, see Theorem 3 and Theorem 4. This implies that in this case, the use of feedback in transmitter-receiver pair i, with $\overleftarrow{n}_{ii} > \overleftarrow{n}^{\dagger}_{ii} = \overleftarrow{n}^*_{ii}$, benefits both transmitter-receiver pairs, *i.e.*, $\Delta_i(\overleftarrow{n}_{ii}) > 0$ and $\Delta_j(\overleftarrow{n}_{ii}) > 0$. Under event $S_{2,i}$, using feedback in transmitter-receiver pair i, with $\overleftarrow{n}_{ii} > \overleftarrow{n}^*_{ii}$, exclusively benefits transmitter-receiver pair j, *i.e.*, $\Delta_i(\overleftarrow{n}_{ii}) = 0$ and $\Delta_j(\overleftarrow{n}_{ii}) > 0$.

4.2.5. Improvement of the Sum-Rate

Given fixed parameters $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21})$, and $i \in \{1,2\}$, implementing channel-output feedback in transmitter-receiver pair i increases the sum-rate, *i.e.*, $\Sigma(\overleftarrow{n}_{ii}) > 0$ for some values of $\overleftarrow{n}_{ii}$. Theorem 5 identifies the exact values of $\overleftarrow{n}_{ii}$ for which $\Sigma(\overleftarrow{n}_{ii}) > 0$.

Theorem 5. Improvement of the Sum-Capacity

Let $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$ be a fixed 4-tuple. Let also $i \in \{1,2\}$, $j \in \{1,2\} \setminus \{i\}$ and $\overleftarrow{n}^+_{ii} \in \mathbb{N}$ be fixed integers, with

$$\overleftarrow{n}^+_{ii} = \begin{cases} \max\left(n_{ji}, (\overrightarrow{n}_{ii} - n_{ij})^+\right) & \text{if } S_4 \text{ holds true} \\ \overrightarrow{n}_{jj} + (\overrightarrow{n}_{ii} - n_{ij})^+ & \text{if } S_5 \text{ holds true} \end{cases}. \tag{4.30}$$

Assume that S_4 holds false and S_5 holds false. Then, $\Sigma(\overleftarrow{n}_{ii}) = 0$ for all $\overleftarrow{n}_{ii} \in \mathbb{N}$. Assume that S_4 holds true or S_5 holds true. Then, when $\overleftarrow{n}_{ii} \leqslant \overleftarrow{n}^+_{ii}$, it holds that $\Sigma(\overleftarrow{n}_{ii}) = 0$; and when $\overleftarrow{n}_{ii} > \overleftarrow{n}^+_{ii}$, it holds that $\Sigma(\overleftarrow{n}_{ii}) > 0$.

Proof: The proof of Theorem 5 is presented in Appendix E. ■

Theorem 5 introduces a necessary but not sufficient condition for improving the sum-rate by implementing feedback in transmitter-receiver pair i.

Remark 4: *To observe* $\Sigma(\overleftarrow{n}_{ii}) > 0$*, it is necessary but not sufficient to satisfy one of the following conditions:* (a) *both transmitter-receiver pairs are in LIR (Event E_1); or* (b) *both transmitter-receiver pairs are in HIR (Event $\overline{E}_1$).*

Finally, it follows from Corollary 3 that when S_4 or S_5 holds true, with $i \in \{1,2\}$ and $\overleftarrow{n}_{ii} > \overleftarrow{n}^+_{ii}$, in addition to $\Sigma(\overleftarrow{n}_{ii}) > 0$, it also holds that $\Delta_1(\overleftarrow{n}_{ii}) > 0$ and $\Delta_2(\overleftarrow{n}_{ii}) > 0$.

4.2.6. Examples

Example 3. Consider an LDIC-NOF with parameters $\overrightarrow{n}_{11} = 7$, $\overrightarrow{n}_{22} = 7$, $n_{12} = 3$, and $n_{21} = 5$.

In Example 3, both $S_{1,1}$ and $S_{1,2}$ hold true. Hence, from Theorem 2, when $\overleftarrow{n}_{11} > 5$ or $\overleftarrow{n}_{22} > 3$, there always exists an enlargement of the capacity region. More specifically, it follows from Theorem 3 and Theorem 4 that using feedback in transmitter-receiver pair 1, with $\overleftarrow{n}_{11} > 5$ or using feedback in transmitter-receiver pair 2, with $\overleftarrow{n}_{22} > 3$, both individual

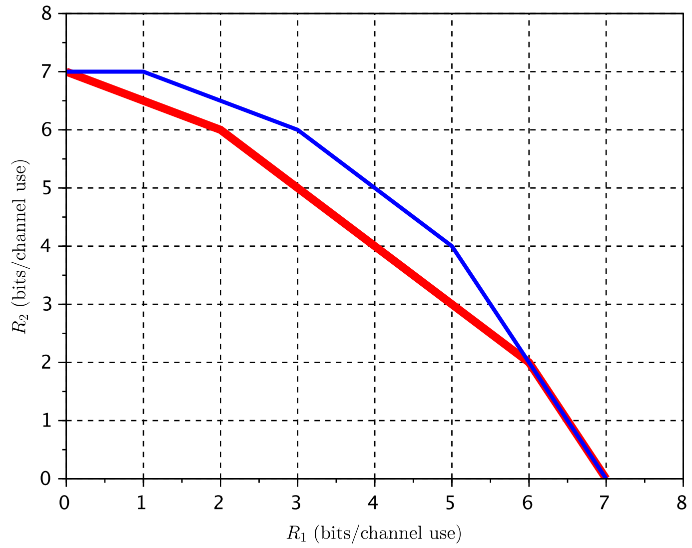


Figure 4.3.: Capacity regions $\mathcal{C}(0,0)$ (thick red line) and $\mathcal{C}(6,0)$ (thin blue line), with $\overrightarrow{n}_{11} = 7$, $\overrightarrow{n}_{22} = 7$, $n_{12} = 3$, $n_{21} = 5$.

rates can be simultaneously improved, *i.e.*, $\Delta_1(\overleftarrow{n}_{ii}) > 0$ and $\Delta_2(\overleftarrow{n}_{ii}) > 0$ with $i = 1$ or $i = 2$ respectively. Alternatively, note that S_4 holds true. Hence, it follows from Theorem 5 that using feedback in transmitter-receiver pair 1, with $\overleftarrow{n}_{11} > 5$ or using feedback in transmitter-receiver pair 2, with $\overleftarrow{n}_{22} > 3$, improves the sum-rate, *i.e.*, $\Sigma(\overleftarrow{n}_{ii}) > 0$ with $i = 1$ or $i = 2$, respectively. These conclusions are observed in Figure 4.3, for the case $\overleftarrow{n}_{11} = 6$ and $\overleftarrow{n}_{22} = 0$, where the capacity regions $\mathcal{C}(0,0)$ (thick red line) and $\mathcal{C}(6,0)$ (thin blue line) are plotted. Note that, when $\overleftarrow{n}_{11} = 6$, there always exists a rate pair $(R_1', R_2') \in \mathcal{C}\,(0,0)$ and a rate pair $(R_1, R_2) \in \mathcal{C}(6,0) \setminus \mathcal{C}(0,0)$ such that $R_1' < R_1$ and $R_2' = R_2$ (Theorem 3). Simultaneously, there always exists a rate pair $(R_1', R_2') \in \mathcal{C}\,(0,0)$ and a rate pair $(R_1, R_2) \in \mathcal{C}(6,0) \setminus \mathcal{C}(0,0)$ such that $R_2' < R_2$ and $R_1' = R_1$ (Theorem 4). Finally, note that for all rate pairs $(R_1', R_2') \in \mathcal{C}\,(0,0)$ there always exists a rate pair $(R_1, R_2) \in \mathcal{C}(6,0)$, for which $R_1 + R_2 > R_1' + R_2'$ (Theorem 5).

Example 4. Consider an LDIC-NOF with parameters $\overrightarrow{n}_{11} = 7$, $\overrightarrow{n}_{22} = 8$, $n_{12} = 6$, and $n_{21} = 5$.

In Example 4, the events $S_{1,1}$ and $S_{1,2}$ hold true, and the events S_4 and S_5 hold false. Hence, it follows from Theorem 5 that using feedback in either transmitter-receiver pair does not improve the sum-rate, *i.e.*, for all $i \in \{1, 2\}$ and for all $\overleftarrow{n}_{ii} > 0$, $\Sigma(\overleftarrow{n}_{ii}) = 0$. These conclusions are observed in Figure 4.4, for the case $\overleftarrow{n}_{11} = 0$ and $\overleftarrow{n}_{22} = 7$, where the capacity regions $\mathcal{C}(0,0)$ (thick red line) and $\mathcal{C}(0,7)$ (thin blue line) are plotted. From Example 4, it becomes evident that when $S_{1,1}$ and $S_{1,2}$ hold true, S_4 and S_5 do not necessarily hold true. That is, the improvements on the individual rates, despite of the fact that they can be observed simultaneously, are not enough to improve the sum-rate beyond what is already achievable without feedback.

Example 5. Consider an LDIC-NOF with parameters $\overrightarrow{n}_{11} = 5$, $\overrightarrow{n}_{22} = 1$, $n_{12} = 3$, and $n_{21} = 4$.

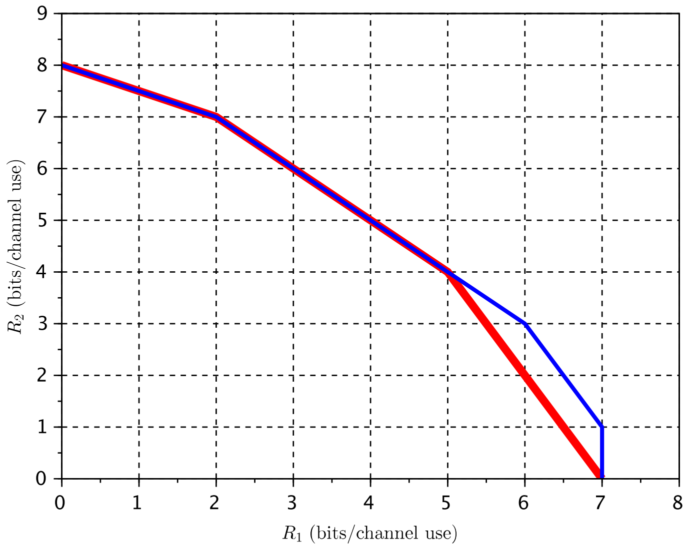


Figure 4.4.: Capacity regions $\mathcal{C}(0,0)$ (thick red line) and $\mathcal{C}(0,7)$ (thin blue line), with $\overrightarrow{n}_{11} = 7$, $\overrightarrow{n}_{22} = 8$, $n_{12} = 6$, $n_{21} = 5$.

In Example 5, both $S_{2,1}$ in (4.18) and $S_{3,2}$ in (4.19) hold true. Hence, it follows from Theorem 2 that the capacity region can be enlarged by using feedback in transmitter-receiver pair 1 when $\overleftarrow{n}_{11} > 3$, whereas using feedback in transmitter-receiver pair 2 does not enlarge the capacity region. More specifically, it follows from Theorem 3 and Theorem 4 that using feedback in transmitter-receiver pair 1 does not improve the individual rate R_1 but R_2, *i.e.*, $\Delta_1(\overleftarrow{n}_{11}) = 0$ and $\Delta_2(\overleftarrow{n}_{11}) > 0$. Note also that S_4 and S_5 hold false. Hence, it follows from Theorem 5 that using feedback in either transmitter-receiver pair does not improve the sum-rate, *i.e.*, $\Sigma(\overleftarrow{n}_{11}) = 0$ and $\Sigma(\overleftarrow{n}_{22}) = 0$. These conclusions are observed in Figure 4.5, for the case $\overleftarrow{n}_{11} = 4$ and $\overleftarrow{n}_{22} = 0$, where the capacity regions $\mathcal{C}(0,0)$ (thick red line) and $\mathcal{C}(4,0)$ (thin blue line) are plotted.

4.3. Generalized Degrees of Freedom

This section focuses on the analysis of the number of GDoF of the two-user LDIC-NOF for studying the case in which feedback is simultaneously implemented in both transmitter-receiver pairs. Moreover, the analysis is only performed for the symmetric case, *i.e.*, $\overrightarrow{n} = \overrightarrow{n}_{11} = \overrightarrow{n}_{22}$, $m = n_{12} = n_{21}$, and $\overleftarrow{n} = \overleftarrow{n}_{11} = \overleftarrow{n}_{22}$, with $(\overrightarrow{n}, m, \overleftarrow{n}) \in \mathbb{N}^3$. The results in Theorem 1 provide a more general analysis of the number of GDoF, *e.g.*, non-symmetric case. However, the symmetric case captures some of the most important insights regarding how the capacity region is enlarged when feedback is used in both transmitter-receiver pairs.

Essentially, given the parameters $\overrightarrow{n}$, m and $\overleftarrow{n}$, with $\alpha = \frac{m}{\overrightarrow{n}}$ and $\beta = \frac{\overleftarrow{n}}{\overrightarrow{n}}$, the number of GDoF, denoted by $D(\alpha, \beta)$, is the ratio between the symmetric capacity, *i.e.*, $C_{\text{sym}}(\overrightarrow{n}, m, \overleftarrow{n}) = \sup\{R : (R,R) \in \mathcal{C}(\overrightarrow{n}, \overrightarrow{n}, m, m, \overleftarrow{n}, \overleftarrow{n})\}$, and the individual interference-free point-to-point capacity, *i.e.*, $\overrightarrow{n}$, when $(\overrightarrow{n}, m, \overleftarrow{n}) \rightarrow (\infty, \infty, \infty)$ at constant ratios $\alpha = \frac{m}{\overrightarrow{n}}$ and $\beta = \frac{\overleftarrow{n}}{\overrightarrow{n}}$. More

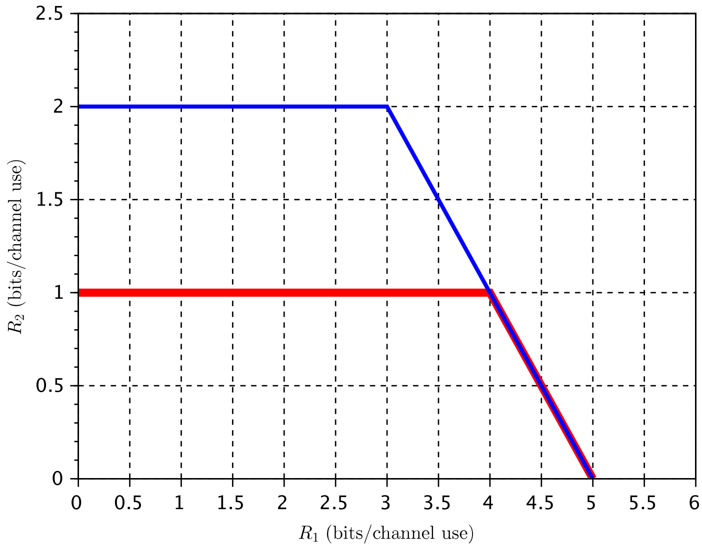


Figure 4.5.: Capacity regions $\mathcal{C}(0,0)$ (thick red line) and $\mathcal{C}(4,0)$ (thin blue line), with $\overrightarrow{n}_{11} = 5$, $\overrightarrow{n}_{22} = 1$, $n_{12} = 3$, $n_{21} = 4$.

specifically, the number of GDoF is:

$$D(\alpha, \beta) = \lim_{(\overrightarrow{n}, m, \overleftarrow{n}) \to (\infty,\infty,\infty)} \frac{C_{\text{sym}}(\overrightarrow{n}, m, \overleftarrow{n})}{\overrightarrow{n}}. \tag{4.31}$$

Theorem 6 determines the number of GDoF for the two-user symmetric LDIC-NOF.

Theorem 6. The number of GDoF

The number of GDoF for the two-user symmetric LDIC-NOF with parameters α and β is given by

$$D(\alpha, \beta) = \min\Bigg(\max(1, \alpha), \max\left(1, \beta - (1-\alpha)^+\right), \frac{1}{2}\left(\max(1,\alpha) + (1-\alpha)^+\right), \max\left((1-\alpha)^+, \alpha, 1 - (\max(1,\alpha) - \beta)^+\right) \Bigg). \tag{4.32}$$

Proof: The proof of Theorem 6 is presented in Appendix F. ■

The result in Theorem 6 can also be obtained from Theorem 1 in [53]. The following corollary is a direct consequence of Theorem 6.

Corollary 4. *The number of GDoF for the two-user symmetric LDIC-NOF with parameters*

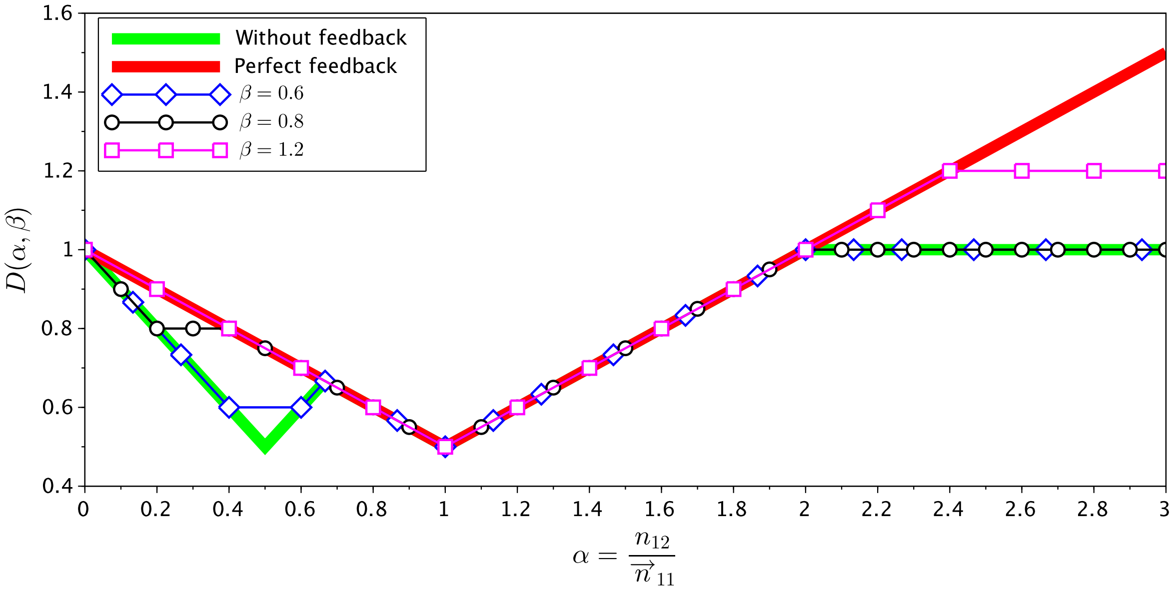


Figure 4.6.: Generalized Degrees of Freedom as a function of the parameters α and β, with $0 \leqslant \alpha \leqslant 3$ and $\beta \in \{\frac{3}{5}, \frac{4}{5}, \frac{6}{5}\}$, of the two-user symmetric LDIC-NOF. The plot without feedback is obtained from [31] and the plot with perfect-output feedback is obtained from [88].

α and β satisfies the following properties:

$$\forall \alpha \in \left[0, \frac{2}{3}\right] \text{ and } \beta \leqslant 1, \qquad \max\left(\frac{1}{2}, \beta\right) \leqslant D(\alpha, \beta) \leqslant 1, \tag{4.33a}$$

$$\forall \alpha \in \left[0, \frac{2}{3}\right] \text{ and } \beta > 1, \qquad D(\alpha, \beta) = 1 - \frac{\alpha}{2}, \tag{4.33b}$$

$$\forall \alpha \in \left(\frac{2}{3}, 2\right] \text{ and } \beta \in [0, \infty), \qquad D(\alpha, 0) = D(\alpha, \beta) = D(\alpha, \max(1, \alpha)), \tag{4.33c}$$

$$\forall \alpha \in (2, \infty) \text{ and } \beta \geqslant 1, \qquad 1 \leqslant D(\alpha, \beta) \leqslant \min\left(\frac{\alpha}{2}, \beta\right), \tag{4.33d}$$

$$\forall \alpha \in (2, \infty) \text{ and } \beta < 1, \qquad D(\alpha, \beta) = 1. \tag{4.33e}$$

Properties (4.33a) and (4.33b) highlight the fact that the existence of feedback links in the two-user symmetric LDIC-NOF in the VWIR and WIR does not have any impact in the number of GDoF when $\beta \leqslant \frac{1}{2}$, and the number of GDoF is equal to the case with perfect-output feedback when $\beta > 1$. Property (4.33c) underlines that in the two-user symmetric LDIC-NOF in MIR and SIR, the number of GDoF is identical in both extreme cases: without feedback ($\beta = 0$) and with perfect-output feedback $\big(\beta = \max(1, \alpha)\big)$. Finally, from (4.33d) and (4.33e), it follows that for observing an improvement in the number of GDoF of the two-user symmetric LDIC-NOF in VSIR, the following condition must be met: $\beta > 1$. That is, the number of bit-pipes in the feedback links must be greater than the number of bit-pipes in the direct links.

Figure 4.6 shows the number of GDoF for the two-user symmetric LDIC-NOF for the case in which $0 \leqslant \alpha \leqslant 3$ and $\beta \in \{\frac{3}{5}, \frac{4}{5}, \frac{6}{5}\}$.

— 5 —

Gaussian Interference Channel

This chapter presents the main results on the centralized GIC-NOF described in Section 2.1. These include an achievable region (Theorem 7) and a converse region (Theorem 8), denoted by $\underline{\mathcal{C}}$ and $\overline{\mathcal{C}}$ respectively, for 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 $\overleftarrow{\mathrm{SNR}}_2$. In general, the capacity region of a given multi-user channel is said to be approximated to within a constant gap according to the following definition:

Definition 6 (Approximation to within ξ units)**.**
A closed and convex set $\mathcal{T} \subset \mathbb{R}_+^m$ is approximated to within ξ units by the sets $\underline{\mathcal{T}}$ and $\overline{\mathcal{T}}$ if $\underline{\mathcal{T}} \subseteq \mathcal{T} \subseteq \overline{\mathcal{T}}$ and for all $\boldsymbol{t} = (t_1, t_2, \ldots, t_m) \in \overline{\mathcal{T}}$, $\left((t_1-\xi)^+, (t_2-\xi)^+, \ldots, (t_m-\xi)^+\right) \in \underline{\mathcal{T}}$.

Denote by $\mathcal{C}$ the capacity region of the 2-user GIC-NOF. The achievable region $\underline{\mathcal{C}}$ and the converse region $\overline{\mathcal{C}}$ approximate the capacity region $\mathcal{C}$ to within 4.4 bits (Theorem 9).

5.1. An Achievable Region

The description of the achievable region $\underline{\mathcal{C}}$ 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 $j \in \{1,2\} \setminus \{i\}$:

$$a_{1,i} = \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\right) - \frac{1}{2}, \tag{5.1a}$$

$$a_{2,i}(\rho) = \frac{1}{2}\log\left(b_{1,i}(\rho) + 1\right) - \frac{1}{2}, \tag{5.1b}$$

$$a_{3,i}(\rho,\mu)=\frac{1}{2}\log\left(\frac{\overleftarrow{\mathrm{SNR}}_i\Big(b_{2,i}(\rho)+2\Big)+b_{1,i}(1)+1}{\overleftarrow{\mathrm{SNR}}_i\Big((1-\mu)\,b_{2,i}(\rho)+2\Big)+b_{1,i}(1)+1}\right), \tag{5.1c}$$

$$a_{4,i}(\rho,\mu)=\frac{1}{2}\log\left(\Big(1-\mu\Big)b_{2,i}(\rho)+2\right)-\frac{1}{2}, \tag{5.1d}$$

$$a_{5,i}(\rho,\mu)=\frac{1}{2}\log\left(2+\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}+\Big(1-\mu\Big)b_{2,i}(\rho)\right)-\frac{1}{2}, \tag{5.1e}$$

$$a_{6,i}(\rho,\mu)=\frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\left(\Big(1-\mu\Big)b_{2,j}(\rho)+1\right)+2\right)-\frac{1}{2},\text{ and} \tag{5.1f}$$

$$a_{7,i}(\rho,\mu_1,\mu_2)=\frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\left(\Big(1-\mu_i\Big)b_{2,j}(\rho)+1\right)+\Big(1-\mu_j\Big)b_{2,i}(\rho)+2\right)-\frac{1}{2}, \tag{5.1g}$$

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

$$b_{1,i}(\rho)=\overrightarrow{\mathrm{SNR}}_i+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_i\mathrm{INR}_{ij}}+\mathrm{INR}_{ij}\text{ and} \tag{5.2a}$$

$$b_{2,i}(\rho)=\Big(1-\rho\Big)\mathrm{INR}_{ij}-1, \tag{5.2b}$$

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

Note that the functions in (5.1) and (5.2) 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, Theorem 7 is presented as follows:

Theorem 7. Achievable Region

The capacity region $\mathcal{C}$ contains the region $\underline{\mathcal{C}}$ given by the closure of the set of all possible achievable rate pairs $(R_1,R_2)\in\mathbb{R}_+^2$ that satisfy:

$$R_1\leqslant\min\Big(a_{2,1}(\rho),a_{6,1}(\rho,\mu_1)+a_{3,2}(\rho,\mu_1),a_{1,1}+a_{3,2}(\rho,\mu_1)+a_{4,2}(\rho,\mu_1)\Big), \tag{5.3a}$$

$$R_2\leqslant\min\Big(a_{2,2}(\rho),a_{3,1}(\rho,\mu_2)+a_{6,2}(\rho,\mu_2),a_{3,1}(\rho,\mu_2)+a_{4,1}(\rho,\mu_2)+a_{1,2}\Big), \tag{5.3b}$$

$$\begin{aligned}R_1+R_2\leqslant\min\Big(&a_{2,1}(\rho)+a_{1,2},a_{1,1}+a_{2,2}(\rho),\\&a_{3,1}(\rho,\mu_2)+a_{1,1}+a_{3,2}(\rho,\mu_1)+a_{7,2}(\rho,\mu_1,\mu_2),\\&a_{3,1}(\rho,\mu_2)+a_{5,1}(\rho,\mu_2)+a_{3,2}(\rho,\mu_1)+a_{5,2}(\rho,\mu_1),\\&a_{3,1}(\rho,\mu_2)+a_{7,1}(\rho,\mu_1,\mu_2)+a_{3,2}(\rho,\mu_1)+a_{1,2}\Big),\end{aligned} \tag{5.3c}$$

$$\begin{aligned}2R_1+R_2\leqslant\min\Big(&a_{2,1}(\rho)+a_{1,1}+a_{3,2}(\rho,\mu_1)+a_{7,2}(\rho,\mu_1,\mu_2),\\&a_{3,1}(\rho,\mu_2)+a_{1,1}+a_{7,1}(\rho,\mu_1,\mu_2)+2a_{3,2}(\rho,\mu_1)+a_{5,2}(\rho,\mu_1),\\&a_{2,1}(\rho)+a_{1,1}+a_{3,2}(\rho,\mu_1)+a_{5,2}(\rho,\mu_1)\Big),\end{aligned} \tag{5.3d}$$

$$R_1 + 2R_2 \leqslant \min\Big(a_{3,1}(\rho,\mu_2) + a_{5,1}(\rho,\mu_2) + a_{2,2}(\rho) + a_{1,2},$$
$$a_{3,1}(\rho,\mu_2) + a_{7,1}(\rho,\mu_1,\mu_2) + a_{2,2}(\rho) + a_{1,2},$$
$$2a_{3,1}(\rho,\mu_2) + a_{5,1}(\rho,\mu_2) + a_{3,2}(\rho,\mu_1) + a_{1,2} + a_{7,2}(\rho,\mu_1,\mu_2)\Big), \quad (5.3e)$$

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

Proof: The proof of Theorem 7 is presented in Appendix A. ■

The achievability scheme presented in Appendix A is general and thus, it can be used for both the two-user LDIC-NOF and the two-user GIC-NOF. The special case of the two-user GIC-NOF is derived in Appendix A.

5.2. A Converse Region

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

$$S_{1,i}: \quad \overrightarrow{\mathrm{SNR}}_j < \min\left(\mathrm{INR}_{ij}, \mathrm{INR}_{ji}\right), \quad (5.4a)$$
$$S_{2,i}: \quad \mathrm{INR}_{ji} \leqslant \overrightarrow{\mathrm{SNR}}_j < \mathrm{INR}_{ij}, \quad (5.4b)$$
$$S_{3,i}: \quad \mathrm{INR}_{ij} \leqslant \overrightarrow{\mathrm{SNR}}_j < \mathrm{INR}_{ji}, \quad (5.4c)$$
$$S_{4,i}: \quad \max\left(\mathrm{INR}_{ij}, \mathrm{INR}_{ji}\right) \leqslant \overrightarrow{\mathrm{SNR}}_j < \mathrm{INR}_{ij}\mathrm{INR}_{ji}, \quad (5.4d)$$
$$S_{5,i}: \quad \overrightarrow{\mathrm{SNR}}_j \geqslant \mathrm{INR}_{ij}\mathrm{INR}_{ji}. \quad (5.4e)$$

Note that for all $i \in \{1,2\}$, the events $S_{1,i}$, $S_{2,i}$, $S_{3,i}$, $S_{4,i}$, and $S_{5,i}$ are mutually exclusive. This observation shows that given any 4-tuple $(\overrightarrow{\mathrm{SNR}}_1, \overrightarrow{\mathrm{SNR}}_2, \mathrm{INR}_{12}, \mathrm{INR}_{21})$, there always exists one and only one pair of events $(S_{l_1,1}, S_{l_2,2})$, with $(l_1,l_2) \in \{1,\ldots,5\}^2$, that identifies a unique scenario. Note also that the pairs of events $(S_{2,1}, S_{2,2})$ and $(S_{3,1}, S_{3,2})$ are not feasible. In view of this, twenty-three different scenarios can be identified using the events in (5.4). Once the exact scenario is identified, the converse region is described using the functions $\kappa_{l,i} : [0,1] \to \mathbb{R}_+$, with $l \in \{1,\ldots,3\}$; $\kappa_l : [0,1] \to \mathbb{R}_+$, with $l \in \{4,5\}$; $\kappa_{6,l} : [0,1] \to \mathbb{R}_+$, with $l \in \{1,\ldots,4\}$; and $\kappa_{7,i,l} : [0,1] \to \mathbb{R}_+$, with $l \in \{1,2\}$. These functions are defined as follows, for all $i \in \{1,2\}$, with $j \in \{1,2\} \setminus \{i\}$:

$$\kappa_{1,i}(\rho) = \frac{1}{2}\log\Big(b_{1,i}(\rho)+1\Big), \quad (5.5a)$$

$$\kappa_{2,i}(\rho) = \frac{1}{2}\log\Big(1+b_{5,j}(\rho)\Big) + \frac{1}{2}\log\left(1+\frac{b_{4,i}(\rho)}{1+b_{5,j}(\rho)}\right), \quad (5.5b)$$

$$\kappa_{3,i}(\rho) = \frac{1}{2}\log\left(\frac{\Big(b_{4,i}(\rho)+b_{5,j}(\rho)+1\Big)\overleftarrow{\mathrm{SNR}}_j}{\Big(b_{1,j}(1)+1\Big)\Big(b_{4,i}(\rho)+1\Big)}+1\right) + \frac{1}{2}\log\Big(b_{4,i}(\rho)+1\Big), \quad (5.5c)$$

$$\kappa_4(\rho) = \frac{1}{2}\log\left(1+\frac{b_{4,1}(\rho)}{1+b_{5,2}(\rho)}\right) + \frac{1}{2}\log\Big(b_{1,2}(\rho)+1\Big), \quad (5.5d)$$

$$\kappa_5(\rho)=\frac{1}{2}\log\left(1+\frac{b_{4,2}(\rho)}{1+b_{5,1}(\rho)}\right)+\frac{1}{2}\log\left(b_{1,1}(\rho)+1\right), \tag{5.5e}$$

$$\kappa_6(\rho)=\begin{cases}\kappa_{6,1}(\rho) & \text{if } (S_{1,2}\vee S_{2,2}\vee S_{5,2})\wedge(S_{1,1}\vee S_{2,1}\vee S_{5,1})\\ \kappa_{6,2}(\rho) & \text{if } (S_{1,2}\vee S_{2,2}\vee S_{5,2})\wedge(S_{3,1}\vee S_{4,1})\\ \kappa_{6,3}(\rho) & \text{if } (S_{3,2}\vee S_{4,2})\wedge(S_{1,1}\vee S_{2,1}\vee S_{5,1})\\ \kappa_{6,4}(\rho) & \text{if } (S_{3,2}\vee S_{4,2})\wedge(S_{3,1}\vee S_{4,1})\end{cases}, \tag{5.5f}$$

$$\kappa_{7,i}(\rho)=\begin{cases}\kappa_{7,i,1}(\rho) & \text{if } (S_{1,i}\vee S_{2,i}\vee S_{5,i})\\ \kappa_{7,i,2}(\rho) & \text{if } (S_{3,i}\vee S_{4,i})\end{cases}, \tag{5.5g}$$

where

$$\kappa_{6,1}(\rho)=\frac{1}{2}\log\Big(b_{1,1}(\rho)+b_{5,1}(\rho)\mathrm{INR}_{21}\Big)-\frac{1}{2}\log\Big(1+\mathrm{INR}_{12}\Big)+\frac{1}{2}\log\left(1+\frac{b_{5,2}(\rho)\overleftarrow{\mathrm{SNR}}_2}{b_{1,2}(1)+1}\right) \tag{5.6a}$$
$$+\frac{1}{2}\log\Big(b_{1,2}(\rho)+b_{5,1}(\rho)\mathrm{INR}_{21}\Big)-\frac{1}{2}\log\Big(1+\mathrm{INR}_{21}\Big)+\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)\overleftarrow{\mathrm{SNR}}_1}{b_{1,1}(1)+1}\right)+\log(2\pi e),$$

$$\kappa_{6,2}(\rho)=\frac{1}{2}\log\left(b_{6,2}(\rho)+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_2}\Big(\overrightarrow{\mathrm{SNR}}_2+b_{3,2}\Big)\right)-\frac{1}{2}\log\Big(1+\mathrm{INR}_{12}\Big) \tag{5.6b}$$
$$+\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)\overleftarrow{\mathrm{SNR}}_1}{b_{1,1}(1)+1}\right)+\frac{1}{2}\log\Big(b_{1,1}(\rho)+b_{5,1}(\rho)\mathrm{INR}_{21}\Big)-\frac{1}{2}\log\Big(1+\mathrm{INR}_{21}\Big)$$
$$+\frac{1}{2}\log\left(1+\frac{b_{5,2}(\rho)}{\overrightarrow{\mathrm{SNR}}_2}\left(\mathrm{INR}_{12}+\frac{b_{3,2}\overleftarrow{\mathrm{SNR}}_2}{b_{1,2}(1)+1}\right)\right)-\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_2}\right)+\log(2\pi e),$$

$$\kappa_{6,3}(\rho)=\frac{1}{2}\log\left(b_{6,1}(\rho)+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_1}\Big(\overrightarrow{\mathrm{SNR}}_1+b_{3,1}\Big)\right)-\frac{1}{2}\log\Big(1+\mathrm{INR}_{12}\Big) \tag{5.6c}$$
$$+\frac{1}{2}\log\left(1+\frac{b_{5,2}(\rho)\overleftarrow{\mathrm{SNR}}_2}{b_{1,2}(1)+1}\right)+\frac{1}{2}\log\Big(b_{1,2}(\rho)+b_{5,1}(\rho)\mathrm{INR}_{21}\Big)-\frac{1}{2}\log\Big(1+\mathrm{INR}_{21}\Big)$$
$$+\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)}{\overrightarrow{\mathrm{SNR}}_1}\left(\mathrm{INR}_{21}+\frac{b_{3,1}\overleftarrow{\mathrm{SNR}}_1}{b_{1,1}(1)+1}\right)\right)-\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_1}\right)+\log(2\pi e),$$

$$\kappa_{6,4}(\rho)=\frac{1}{2}\log\left(b_{6,1}(\rho)+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_1}\Big(\overrightarrow{\mathrm{SNR}}_1+b_{3,1}\Big)\right)-\frac{1}{2}\log\Big(1+\mathrm{INR}_{12}\Big) \tag{5.6d}$$
$$+\frac{1}{2}\log\left(1+\frac{b_{5,2}(\rho)}{\overrightarrow{\mathrm{SNR}}_2}\left(\mathrm{INR}_{12}+\frac{b_{3,2}\overleftarrow{\mathrm{SNR}}_2}{b_{1,2}(1)+1}\right)\right)-\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_2}\right)$$
$$-\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_1}\right)+\frac{1}{2}\log\left(b_{6,2}(\rho)+\frac{b_{5,1}(\rho)\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_2}\Big(\overrightarrow{\mathrm{SNR}}_2+b_{3,2}\Big)\right)$$
$$-\frac{1}{2}\log\Big(1+\mathrm{INR}_{21}\Big)+\frac{1}{2}\log\left(1+\frac{b_{5,1}(\rho)}{\overrightarrow{\mathrm{SNR}}_1}\left(\mathrm{INR}_{21}+\frac{b_{3,1}\overleftarrow{\mathrm{SNR}}_1}{b_{1,1}(1)+1}\right)\right)+\log(2\pi e),$$

and

$$\kappa_{7,i,1}(\rho) = \frac{1}{2}\log\left(b_{1,i}(\rho)+1\right) - \frac{1}{2}\log\left(1+\mathrm{INR}_{ij}\right) + \frac{1}{2}\log\left(1+\frac{b_{5,j}(\rho)\overleftarrow{\mathrm{SNR}}_j}{b_{1,j}(1)+1}\right) \tag{5.7a}$$
$$+\frac{1}{2}\log\left(b_{1,j}(\rho)+b_{5,i}(\rho)\mathrm{INR}_{ji}\right)+\frac{1}{2}\log\left(1+b_{4,i}(\rho)+b_{5,j}(\rho)\right)-\frac{1}{2}\log\left(1+b_{5,j}(\rho)\right)$$
$$+2\log(2\pi e),$$
$$\kappa_{7,i,2}(\rho) = \frac{1}{2}\log\left(b_{1,i}(\rho)+1\right) - \frac{1}{2}\log\left(1+\mathrm{INR}_{ij}\right) - \frac{1}{2}\log\left(1+b_{5,j}(\rho)\right) \tag{5.7b}$$
$$+\frac{1}{2}\log\left(1+b_{4,i}(\rho)+b_{5,j}(\rho)\right)+\frac{1}{2}\log\left(1+\left(1-\rho^2\right)\frac{\mathrm{INR}_{ji}}{\overrightarrow{\mathrm{SNR}}_j}\left(\mathrm{INR}_{ij}+\frac{b_{3,j}\overleftarrow{\mathrm{SNR}}_j}{b_{1,j}(1)+1}\right)\right)$$
$$-\frac{1}{2}\log\left(1+\frac{b_{5,i}(\rho)\mathrm{INR}_{ji}}{\overrightarrow{\mathrm{SNR}}_j}\right)+\frac{1}{2}\log\left(b_{6,j}(\rho)+\frac{b_{5,i}(\rho)\mathrm{INR}_{ji}}{\overrightarrow{\mathrm{SNR}}_j}\left(\overleftarrow{\mathrm{SNR}}_j+b_{3,j}\right)\right)+2\log(2\pi e),$$

where, the functions $b_{l,i}$, with $(l,i)\in\{1,2\}^2$ are defined in (5.2); $b_{3,i}$ are constants; and the functions $b_{l,i}:[0,1]\to\mathbb{R}_+$, with $(l,i)\in\{4,5,6\}\times\{1,2\}$ are defined as follows, with $j\in\{1,2\}\setminus\{i\}$:

$$b_{3,i}=\overrightarrow{\mathrm{SNR}}_i-2\sqrt{\overrightarrow{\mathrm{SNR}}_i\mathrm{INR}_{ji}}+\mathrm{INR}_{ji}, \tag{5.8a}$$
$$b_{4,i}(\rho)=\left(1-\rho^2\right)\overrightarrow{\mathrm{SNR}}_i, \tag{5.8b}$$
$$b_{5,i}(\rho)=\left(1-\rho^2\right)\mathrm{INR}_{ij}, \tag{5.8c}$$
$$b_{6,i}(\rho)=\overrightarrow{\mathrm{SNR}}_i+\mathrm{INR}_{ij}+2\rho\sqrt{\mathrm{INR}_{ij}}\left(\sqrt{\overrightarrow{\mathrm{SNR}}_i}-\sqrt{\mathrm{INR}_{ji}}\right) \tag{5.8d}$$
$$+\frac{\mathrm{INR}_{ij}\sqrt{\mathrm{INR}_{ji}}}{\overrightarrow{\mathrm{SNR}}_i}\left(\sqrt{\mathrm{INR}_{ji}}-2\sqrt{\overrightarrow{\mathrm{SNR}}_i}\right).$$

Note that the functions in (5.5), (5.6), (5.7), and (5.8) depend on $\overrightarrow{\mathrm{SNR}}_1$, $\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, INR_{21}, $\overleftarrow{\mathrm{SNR}}_1$, and $\overleftarrow{\mathrm{SNR}}_2$. However, these parameters are fixed in this analysis, and therefore, this dependence is not emphasized in the definition of these functions. Finally, using this notation, Theorem 8 is presented below.

Theorem 8. Converse Region

The capacity region $\mathcal{C}$ is contained within the region $\overline{\mathcal{C}}$ given by the closure of the set of rate pairs $(R_1,R_2)\in\mathbb{R}_+^2$ that for all $i\in\{1,2\}$, with $j\in\{1,2\}\setminus\{i\}$ satisfy:

$$R_i\leqslant\min\left(\kappa_{1,i}(\rho),\kappa_{2,i}(\rho)\right), \tag{5.9a}$$
$$R_i\leqslant\kappa_{3,i}(\rho), \tag{5.9b}$$
$$R_1+R_2\leqslant\min\left(\kappa_4(\rho),\kappa_5(\rho)\right), \tag{5.9c}$$
$$R_1+R_2\leqslant\kappa_6(\rho), \tag{5.9d}$$
$$2R_i+R_j\leqslant\kappa_{7,i}(\rho), \tag{5.9e}$$

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

Proof: The proof of Theorem 8 is presented in Appendix G. ■

The outer bounds (5.9a) and (5.9c) play the same role as the outer bounds (4.1a) and (4.1c) in the linear deterministic model and have been previously reported in [88] for the case of POF. The bounds (5.9b), (5.9d), and (5.9e) correspond to new outer bounds. The intuition for deriving these outer bounds follows along the same steps as those used to prove the outer bounds (4.1b), (4.1c), and (4.1d), respectively. Note the duality between the Gaussian signals $X_{i,C}$ and $X_{i,U}$ (in (G.2) and (G.3), respectively) and the bit-pipes $(\boldsymbol{X}_{i,C}, \boldsymbol{X}_{i,D})$ and $\boldsymbol{X}_{i,U}$ (in (B.1a), (B.1d) and (B.5), respectively).

5.3. Gap between the Achievable Region and the Converse Region

Theorem 9 describes the gap between the achievable region $\underline{\mathcal{C}}$ and the converse region $\overline{\mathcal{C}}$ (Definition 6).

Theorem 9. Gap

The capacity region of the two-user GIC-NOF is approximated to within 4.4 bits by the achievable region $\underline{\mathcal{C}}$ and the converse region $\overline{\mathcal{C}}$.

Proof: The proof of Theorem 9 is presented in Appendix H. ■

Figure 5.1 presents the exact gap existing between the achievable region $\underline{\mathcal{C}}$ and the converse region $\overline{\mathcal{C}}$ for the case in which $\overrightarrow{\mathrm{SNR}}_1 = \overrightarrow{\mathrm{SNR}}_2 = \overrightarrow{\mathrm{SNR}}$, $\mathrm{INR}_{12} = \mathrm{INR}_{21} = \mathrm{INR}$, and $\overleftarrow{\mathrm{SNR}}_1 = \overleftarrow{\mathrm{SNR}}_2 = \overleftarrow{\mathrm{SNR}}$ as a function of $\alpha = \frac{\log \mathrm{INR}}{\log \overrightarrow{\mathrm{SNR}}}$ and $\beta = \frac{\log \overleftarrow{\mathrm{SNR}}}{\log \overrightarrow{\mathrm{SNR}}}$. Note that in this case, the maximum gap is 1.1 bits and occurs when $\alpha = 1.05$ and $\beta = 1.2$.

5.4. Cases in which Feedback Enlarges the Capacity Region

This section considers the application of the obtained results in Section 4.2.2 into the two-user GIC-NOF. Therefore, this section defines for a given two-user GIC the approximate thresholds for the feedback parameters beyond which its capacity region can be enlarged.

5.4.1. Rate Improvement Metrics

In order to quantify the benefits of channel-output feedback in enlarging the achievable region $\underline{\mathcal{C}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2)$ or the converse region $\overline{\mathcal{C}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2)$, consider the following improvement metrics, which are similar to those defined in Section 4.2.1 for the two-user LDIC-NOF. The improvement metrics on the individual rates are defined as:

$$\Delta_1^{\mathrm{A}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2) = \max_{0 < R_2 < R_2^*} \Big\{ \sup \left\{ R_1 : (R_1, R_2) \in \underline{\mathcal{C}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2) \right\} - \sup \left\{ R_1^\dagger : (R_1^\dagger, R_2) \in \underline{\mathcal{C}}(0,0) \right\} \Big\}, \quad (5.10)$$

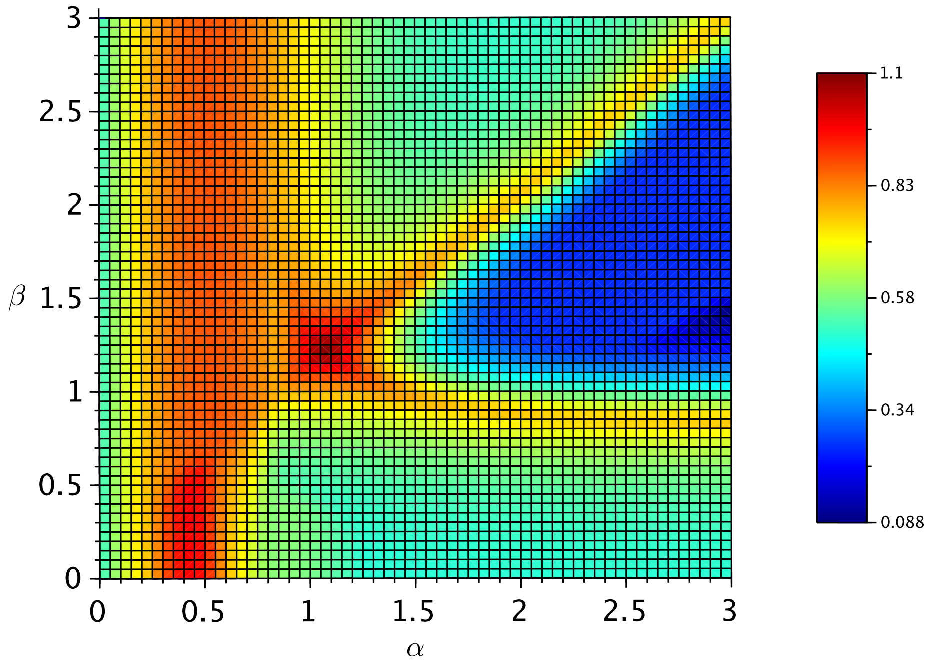


Figure 5.1.: Gap between the converse region $\overline{\mathcal{C}}$ and the achievable region $\underline{\mathcal{C}}$ of the two-user GIC-NOF under symmetric channel conditions, *i.e.*, $\overrightarrow{\mathrm{SNR}}_1 = \overrightarrow{\mathrm{SNR}}_2 = \overrightarrow{\mathrm{SNR}}$, $\mathrm{INR}_{12} = \mathrm{INR}_{21} = \mathrm{INR}$, and $\overleftarrow{\mathrm{SNR}}_1 = \overleftarrow{\mathrm{SNR}}_2 = \overleftarrow{\mathrm{SNR}}$, as a function of $\alpha = \frac{\log \mathrm{INR}}{\log \overrightarrow{\mathrm{SNR}}}$ and $\beta = \frac{\log \overleftarrow{\mathrm{SNR}}}{\log \overrightarrow{\mathrm{SNR}}}$.

$$\Delta_2^{\mathrm{A}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2) = \max_{0<R_1<R_1^*} \Big\{ \sup \left\{ R_2 : (R_1, R_2) \in \underline{\mathcal{C}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2) \right\} - \sup \left\{ R_2^\dagger : (R_1, R_2^\dagger) \in \underline{\mathcal{C}}(0,0) \right\} \Big\}, \tag{5.11}$$

$$\Delta_1^{\mathrm{C}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2) = \max_{0<R_2<R_2^\dagger} \Big\{ \sup \left\{ R_1 : (R_1, R_2) \in \overline{\mathcal{C}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2) \right\} - \sup \left\{ R_1^\dagger : (R_1^\dagger, R_2) \in \overline{\mathcal{C}}(0,0) \right\} \Big\}, \text{ and} \tag{5.12}$$

$$\Delta_2^{\mathrm{C}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2) = \max_{0<R_1<R_1^\dagger} \Big\{ \sup \left\{ R_2 : (R_1, R_2) \in \overline{\mathcal{C}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2) \right\} - \sup \left\{ R_2^\dagger : (R_1, R_2^\dagger) \in \overline{\mathcal{C}}(0,0) \right\} \Big\}, \tag{5.13}$$

with

$$R_1^* = \sup\{r_1 : (r_1, r_2) \in \underline{\mathcal{C}}(0,0)\}, \tag{5.14}$$
$$R_2^* = \sup\{r_2 : (r_1, r_2) \in \underline{\mathcal{C}}(0,0)\}, \tag{5.15}$$

$$R_1^{\dagger}{=}\sup\{r_1 : (r_1, r_2) \in \overline{\mathcal{C}}(0,0)\}, \text{ and} \tag{5.16}$$
$$R_2^{\dagger}{=}\sup\{r_2 : (r_1, r_2) \in \overline{\mathcal{C}}(0,0)\}. \tag{5.17}$$

Alternatively, the maximum improvements of the sum-rate $\Sigma^{\mathrm{A}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2)$ and $\Sigma^{\mathrm{C}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2)$ with respect to the case without feedback are:

$$\begin{aligned}\Sigma^{\mathrm{A}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2){=}&\sup\left\{R_1 + R_2 : (R_1, R_2) \in \underline{\mathcal{C}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2)\right\}\\ &- \sup\left\{R_1^{\dagger} + R_2^{\dagger} : (R_1^{\dagger}, R_2^{\dagger}) \in \underline{\mathcal{C}}(0,0)\right\}, \text{ and}\end{aligned} \tag{5.18}$$

$$\begin{aligned}\Sigma^{\mathrm{C}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2){=}&\sup\left\{R_1 + R_2 : (R_1, R_2) \in \overline{\mathcal{C}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2)\right\}\\ &- \sup\left\{R_1^{\dagger} + R_2^{\dagger} : (R_1^{\dagger}, R_2^{\dagger}) \in \overline{\mathcal{C}}(0,0)\right\}.\end{aligned} \tag{5.19}$$

5.4.2. Improvements

In Chapter 3.4, the connections between the two-user LDIC-NOF and the two-user GIC-NOF were discussed. Using these connections, a GIC with fixed parameters $\left(\overrightarrow{\mathrm{SNR}}_1, \overrightarrow{\mathrm{SNR}}_2, \mathrm{INR}_{12}, \mathrm{INR}_{21}\right)$ is approximated by an LDIC with parameters $\overrightarrow{n}_{11} = \left\lfloor \frac{1}{2}\log(\overrightarrow{\mathrm{SNR}}_1)\right\rfloor$, $\overrightarrow{n}_{22} = \left\lfloor \frac{1}{2}\log(\overrightarrow{\mathrm{SNR}}_2)\right\rfloor$, $n_{12} = \left\lfloor \frac{1}{2}\log(\mathrm{INR}_{12})\right\rfloor$ and $n_{21} = \left\lfloor \frac{1}{2}\log(\mathrm{INR}_{21})\right\rfloor$. From this observation, the results from Theorem 2 - Theorem 5 can be used to determine the feedback SNR thresholds beyond which either an individual rate or the sum-rate is improved in the original GIC-NOF. The procedure consists in using the equalities $\overleftarrow{n}_{ii} = \left\lfloor \frac{1}{2}\log\left(\overleftarrow{\mathrm{SNR}}_i\right)\right\rfloor$, with $i \in \{1, 2\}$. Hence, the corresponding thresholds in the two-user GIC can be approximated by:

$$\overleftarrow{\mathrm{SNR}}_i^{*} = 2^{2\overleftarrow{n}_{ii}^{*}}, \tag{5.20a}$$
$$\overleftarrow{\mathrm{SNR}}_i^{\dagger} = 2^{2\overleftarrow{n}_{ii}^{\dagger}}, \text{ and} \tag{5.20b}$$
$$\overleftarrow{\mathrm{SNR}}_i^{+} = 2^{2\overleftarrow{n}_{ii}^{+}}. \tag{5.20c}$$

When the corresponding LDIC-NOF is such that its capacity region can be improved, *i.e.*, when $\overleftarrow{n}_{ii} > \overleftarrow{n}_{ii}^{*}$ (Theorem 2) for a given $i \in \{1, 2\}$, it is expected that either the achievability or converse regions of the original GIC-NOF become larger when $\overleftarrow{\mathrm{SNR}}_i > \overleftarrow{\mathrm{SNR}}_i^{*}$. Similarly, when the corresponding LDIC-NOF is such that $\Delta_i(\overleftarrow{n}_{ii}) > 0$ or $\Delta_i(\overleftarrow{n}_{jj}) > 0$, it is expected to observe an improvement on the individual rate R_i by either using feedback in transmitter-receiver pair i, with $\overleftarrow{\mathrm{SNR}}_i > \overleftarrow{\mathrm{SNR}}_i^{\dagger}$ or by using feedback in transmitter-receiver pair j, with $\overleftarrow{\mathrm{SNR}}_j > \overleftarrow{\mathrm{SNR}}_j^{*}$. When the corresponding LDIC-NOF is such that $\Sigma(\overleftarrow{n}_{ii}) > 0$ using feedback in transmitter-receiver pair i, with $\overleftarrow{n}_{ii} > \overleftarrow{n}_{ii}^{+}$ (Theorem 5), it is expected to observe an improvement on the sum-rate by using feedback in transmitter-receiver pair i, with $\overleftarrow{\mathrm{SNR}}_i > \overleftarrow{\mathrm{SNR}}_i^{+}$. Finally, when no improvement in a given metric is observed in the two-user LDIC-NOF, *i.e.*, $\Delta_1(\overleftarrow{n}_{11}) = 0$, $\Delta_1(\overleftarrow{n}_{22}) = 0$, $\Delta_2(\overleftarrow{n}_{11}) = 0$, $\Delta_2(\overleftarrow{n}_{22}) = 0$, $\Sigma(\overleftarrow{n}_{11}) = 0$, or $\Sigma(\overleftarrow{n}_{22}) = 0$, only a negligible improvement (if any) is observed in the corresponding metric of the two-user GIC-NOF. For instance, when $\Delta_1(\overleftarrow{n}_{11}) = 0$, it is expected that

$\Delta_1^{\mathrm{A}}(\overleftarrow{\mathrm{SNR}}_1) < \epsilon$ and $\Delta_1^{\mathrm{C}}(\overleftarrow{\mathrm{SNR}}_1) < \epsilon$, with $\epsilon > 0$. Similarly, when $\Delta_2(\overleftarrow{n}_{11}) = 0$, it is expected that $\Delta_2^{\mathrm{A}}(\overleftarrow{\mathrm{SNR}}_1) < \epsilon$ and $\Delta_2^{\mathrm{C}}(\overleftarrow{\mathrm{SNR}}_1) < \epsilon$. Finally, when $\Sigma(\overleftarrow{n}_{11}) = 0$, it is expected that $\Sigma^{\mathrm{A}}(\overleftarrow{\mathrm{SNR}}_1) < \epsilon$ and $\Sigma^{\mathrm{C}}(\overleftarrow{\mathrm{SNR}}_1) < \epsilon$.

5.4.3. Examples

The following examples highlight the relevance of the approximations in (5.20).

Example 6. Consider a GIC with parameters $\overrightarrow{\mathrm{SNR}}_1 = 44$dB, $\overrightarrow{\mathrm{SNR}}_2 = 44$dB, $\mathrm{INR}_{12} = 20$dB, and $\mathrm{INR}_{21} = 33$dB.

The two-user GIC in Example 6 can be approximated by the LIDC presented in Example 2. Hence, $\overleftarrow{n}_{11}^{*} = \overleftarrow{n}_{11}^{\dagger} = \overleftarrow{n}_{11}^{+} = 5$ and $\overleftarrow{n}_{22}^{*} = \overleftarrow{n}_{22}^{\dagger} = \overleftarrow{n}_{22}^{+} = 3$. This implies that $\overleftarrow{\mathrm{SNR}}_1^{*} = \overleftarrow{\mathrm{SNR}}_1^{\dagger} = \overleftarrow{\mathrm{SNR}}_1^{+} = 30$dB and $\overleftarrow{\mathrm{SNR}}_2^{*} = \overleftarrow{\mathrm{SNR}}_2^{\dagger} = \overleftarrow{\mathrm{SNR}}_2^{+} = 18$dB.

Figure 5.2 shows that significant improvements on the metrics $\Delta_i^{\mathrm{A}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2)$, $\Delta_i^{\mathrm{C}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2)$, $\Sigma^{\mathrm{A}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2)$ and $\Sigma^{\mathrm{C}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2)$ are obtained when the feedback SNRs are beyond the corresponding thresholds. More importantly, negligible effects are observed when $\overleftarrow{\mathrm{SNR}}_1 < \overleftarrow{\mathrm{SNR}}_1^{*}$ and $\overleftarrow{\mathrm{SNR}}_2 < \overleftarrow{\mathrm{SNR}}_2^{*}$.

Example 7. Consider a GIC with parameters $\overrightarrow{\mathrm{SNR}}_1 = 45$dB, $\overrightarrow{\mathrm{SNR}}_2 = 50$dB, $\mathrm{INR}_{12} = 40$dB, and $\mathrm{INR}_{21} = 33$dB.

The two-user GIC in Example 7 can be approximated by the LIDC presented in Example 1. Hence, $\overleftarrow{n}_{11}^{*} = \overleftarrow{n}_{11}^{\dagger} = 5$ and $\overleftarrow{n}_{22}^{*} = \overleftarrow{n}_{22}^{\dagger} = 6$. This implies that $\overleftarrow{\mathrm{SNR}}_1^{*} = \overleftarrow{\mathrm{SNR}}_1^{\dagger} = 30$dB and $\overleftarrow{\mathrm{SNR}}_2^{*} = \overleftarrow{\mathrm{SNR}}_2^{\dagger} = 36$dB.

Figure 5.3 shows that significant improvements on the metrics $\Delta_i^{\mathrm{A}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2)$, and $\Delta_i^{\mathrm{C}}(\overleftarrow{\mathrm{SNR}}_1, \overleftarrow{\mathrm{SNR}}_2)$ are obtained when the feedback SNRs are beyond the corresponding thresholds. More importantly, negligible effects are observed when $\overleftarrow{\mathrm{SNR}}_1 < \overleftarrow{\mathrm{SNR}}_1^{*}$ and $\overleftarrow{\mathrm{SNR}}_2 < \overleftarrow{\mathrm{SNR}}_2^{*}$. Note also that using feedback in either transmitter-receiver pair does not improve the sum-rate in the two-user LDIC-NOF, *i.e.*, $\Sigma(\overleftarrow{n}_{11}) = \Sigma(\overleftarrow{n}_{22}) = 0$. This is also verified in the two-user GIC-NOF, and is illustrated by Figure 5.3e and Figure 5.3d, where $\Sigma^{\mathrm{A}}\left(\overleftarrow{\mathrm{SNR}}_1, -100\right) < 0.45$, $\Sigma^{\mathrm{C}}\left(\overleftarrow{\mathrm{SNR}}_1, -100\mathrm{dB}\right) < 0.05$, $\Sigma^{\mathrm{A}}\left(-100\mathrm{dB}, \overleftarrow{\mathrm{SNR}}_2\right) < 0.45$, and $\Sigma^{\mathrm{C}}\left(-100\mathrm{dB}, \overleftarrow{\mathrm{SNR}}_2\right) < 0.05$.

Example 8. Consider a GIC with parameters $\overrightarrow{\mathrm{SNR}}_1 = 33$dB, $\overrightarrow{\mathrm{SNR}}_2 = 9$dB, $\mathrm{INR}_{12} = 20$dB, and $\mathrm{INR}_{21} = 27$dB.

The two-user GIC in Example 8 can be approximated by the LIDC presented in Example 2. Hence, $\overleftarrow{n}_{11}^{*} = 3$, which implies that $\overleftarrow{\mathrm{SNR}}_1^{*} = 18$dB. It follows from the two-user LDIC-NOF that using feedback in transmitter-receiver pair 1 exclusively increases the individual rate R_2. This is observed in Figure 5.4c. Note that the improvement in the individual rate R_2 for all $\overleftarrow{\mathrm{SNR}}_1 < \overleftarrow{\mathrm{SNR}}_1^{*}$ is negligible. Significant improvement is observed only beyond the threshold $\overleftarrow{\mathrm{SNR}}_1^{*}$.

Note also that using feedback in either transmitter-receiver pair does not improve the rate R_1 in the two-user LDIC-NOF, *i.e.*, $\Delta_1(\overleftarrow{n}_{11}) = \Delta_1(\overleftarrow{n}_{22}) = 0$. This is also verified in the GIC-NOF, and is illustrated by Figure 5.4a, Figure 5.4b, and Figure 5.4d, where $\Delta_1^{\mathrm{A}}\left(-100\mathrm{dB}, \overleftarrow{\mathrm{SNR}}_2\right) < 0.15$ and $\Delta_1^{\mathrm{C}}\left(-100\mathrm{dB}, \overleftarrow{\mathrm{SNR}}_2\right) < 0.1$.

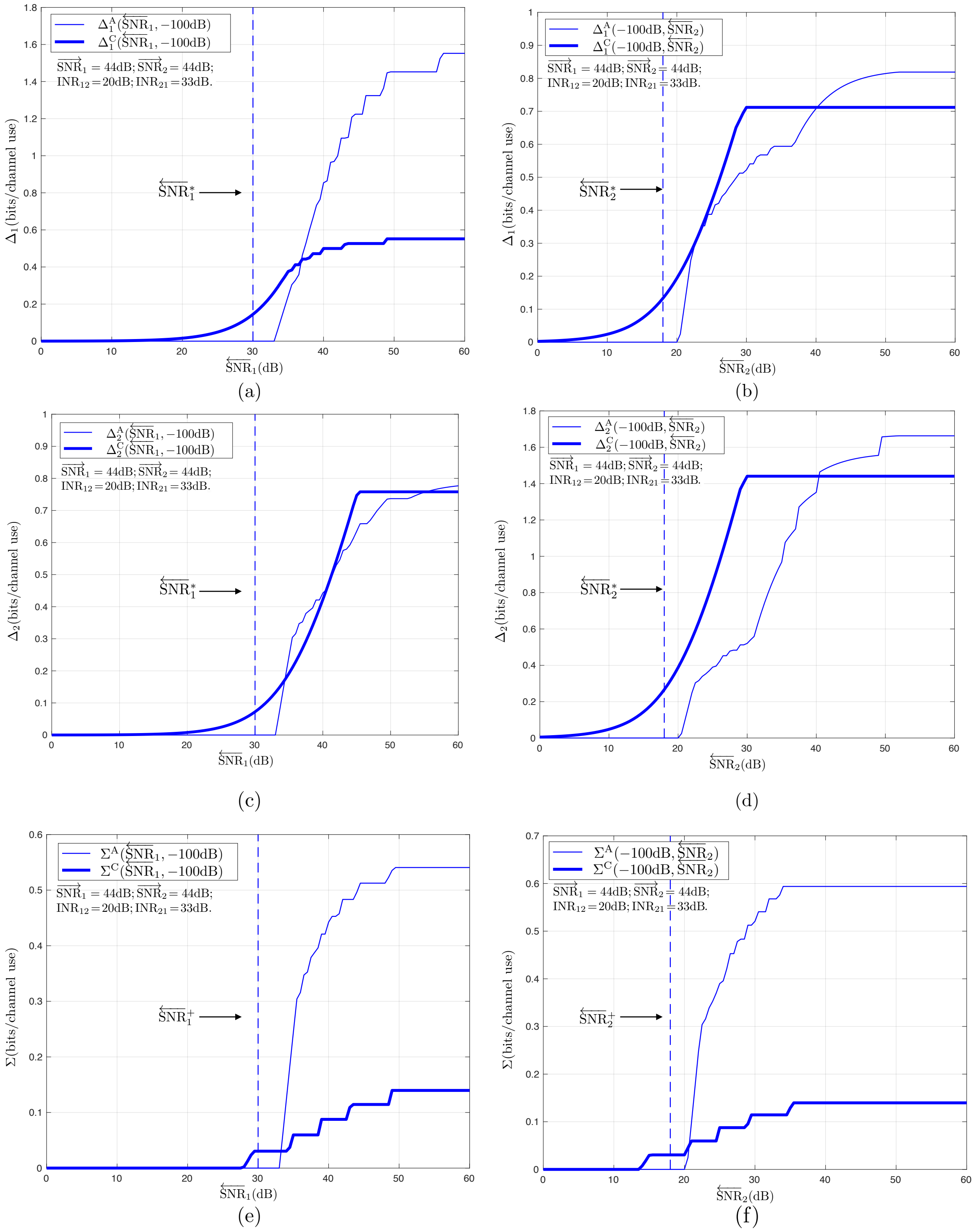


Figure 5.2.: Improvement metrics Δ_i^{A}, Δ_i^{C}, Σ^{A}, and Σ^{C}, with $i \in \{1, 2\}$, as functions of $\overleftarrow{\mathrm{SNR}}_1$ and $\overleftarrow{\mathrm{SNR}}_2$ for Example 6.

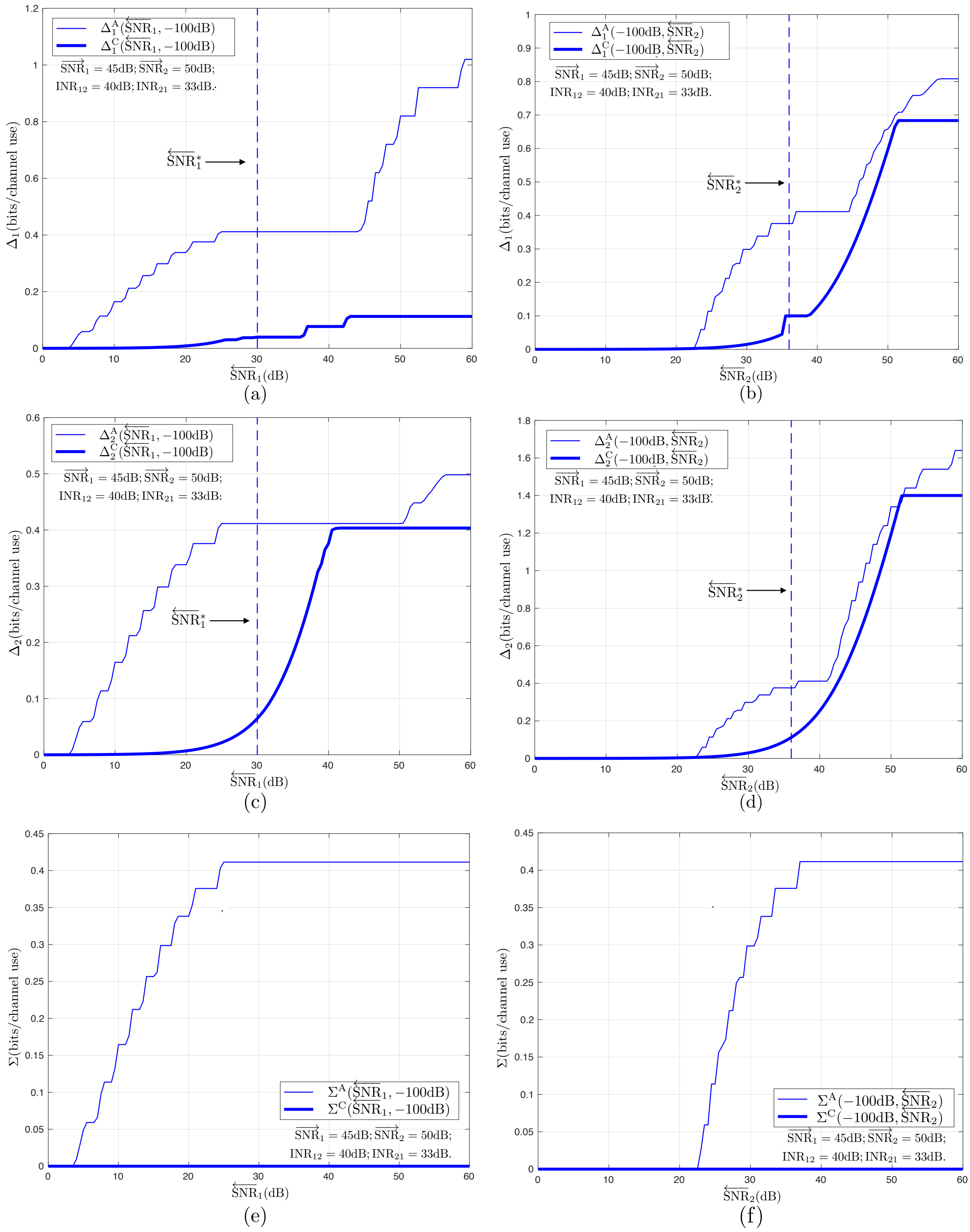


Figure 5.3.: Improvement metrics Δ_i^{A}, Δ_i^{C}, Σ^{A}, and Σ^{C}, with $i \in \{1, 2\}$, as functions of $\overleftarrow{\mathrm{SNR}}_1$ and $\overleftarrow{\mathrm{SNR}}_2$ for Example 7.

Finally, note that using feedback in either transmitter-receiver pair does not increase the sum-

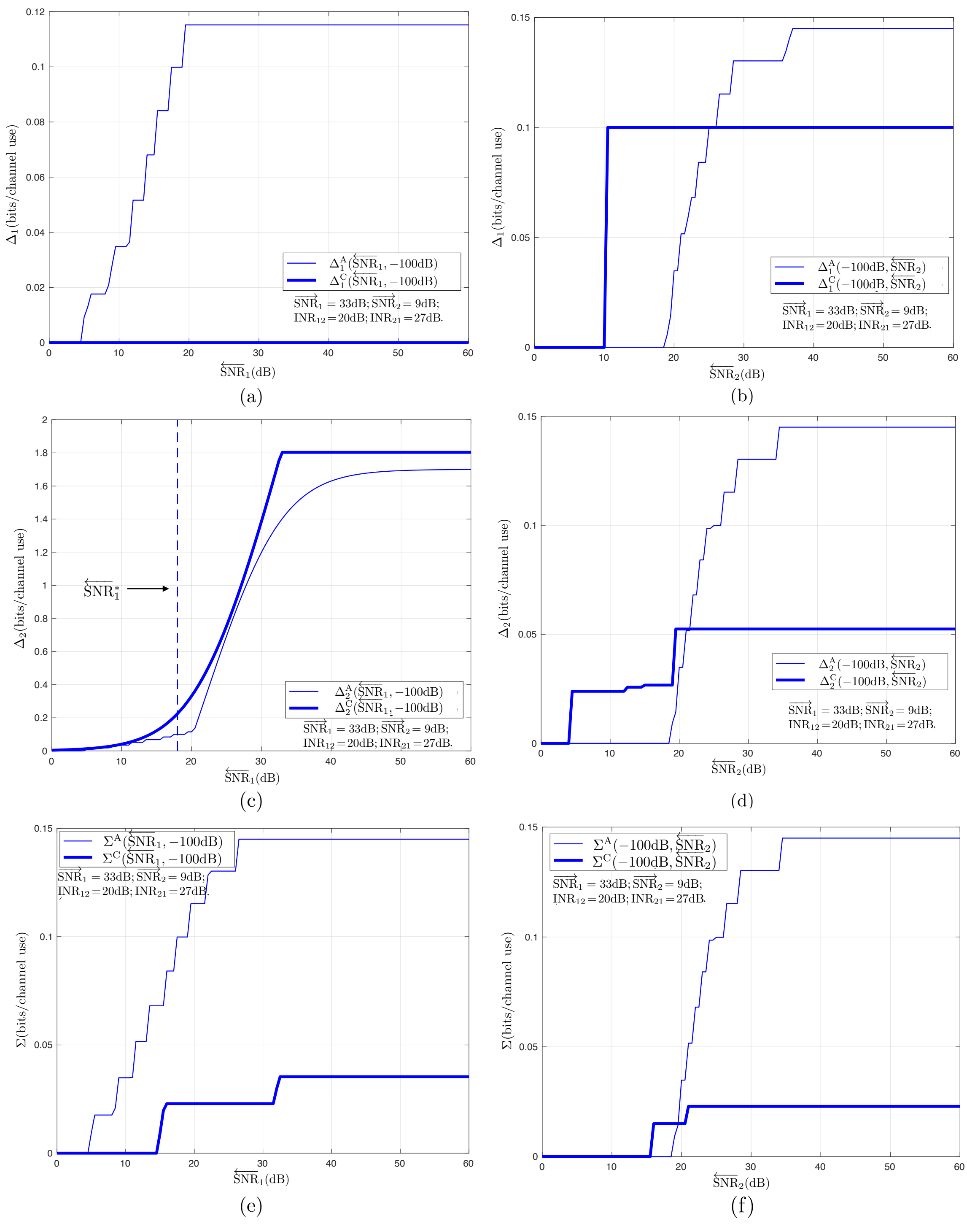


Figure 5.4.: Improvement metrics Δ_i^{A}, Δ_i^{C}, Σ^{A}, and Σ^{C}, with $i \in \{1,2\}$, as functions of $\overleftarrow{\mathrm{SNR}}_1$ and $\overleftarrow{\mathrm{SNR}}_2$ for Example 8.

rate in the two-user LDIC-NOF, *i.e.*, $\Sigma(\overleftarrow{n}_{11}) = \Sigma(\overleftarrow{n}_{22}) = 0$. This is also verified in the two-user

GIC-NOF, and is illustrated by Figure 5.4e and Figure 5.4f, where $\Sigma^{\mathrm{A}}\left(\overleftarrow{\mathrm{SNR}}_1, -100\mathrm{dB}\right) < 0.15$, $\Sigma^{\mathrm{C}}\left(\overleftarrow{\mathrm{SNR}}_1, -100\mathrm{dB}\right) < 0.05$, $\Sigma^{\mathrm{A}}\left(-100\mathrm{dB}, \overleftarrow{\mathrm{SNR}}_2\right) < 0.15$, and $\Sigma^{\mathrm{C}}\left(-100\mathrm{dB}, \overleftarrow{\mathrm{SNR}}_2\right) < 0.05$.

Part III.

CONTRIBUTIONS TO DECENTRALIZED INTERFERENCE CHANNELS

— 6 —

Linear Deterministic Interference Channel

This chapter presents the main results on the two-user D-LDIC-NOF. This model was described in Section 2.2 and can be analyzed by a game as suggested in Section 3.1. 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 two-user LDIC-NOF with parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$, and $\overleftarrow{n}_{22}$, characterized in Theorem 1.

6.1. η-Nash Equilibrium Region

This section characterizes the η-NE region (Definition 5) of the two-user D-LDIC-NOF.

The η-NE region of the two-user D-LDIC-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 $\mathcal{N}_\eta\left(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}\right)$. It 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 region, denoted by $\mathcal{B}_\eta(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$. This region was first characterized in [16] for the case without feedback, in [66] for the case of POF, and in [68] for the case of NOF under symmetric conditions.

In the following, the analysis of these regions is presented for fixed parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$, and $\overleftarrow{n}_{22}$, and thus, the 6-tuple $\left(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}\right)$ is explicited only when needed. The capacity region $\mathcal{C}$ of the two-user LDIC-NOF is described in Theorem 1, which is a generalization of the cases with and without POF, studied respectively in [20] and [88]. For all $\eta > 0$, the convex region $\mathcal{B}_\eta$ is defined as follows:

$$\mathcal{B}_\eta = \left\{(R_1, R_2) \in \mathbb{R}_+^2 : L_i \leqslant R_i \leqslant U_i, \text{ for all } i \in \mathcal{K} = \{1, 2\}\right\}, \tag{6.1}$$

where,

$$L_i = \left((\overrightarrow{n}_{ii} - n_{ij})^+ - \eta \right)^+ \text{ and} \tag{6.2a}$$

$$U_i = \max(\overrightarrow{n}_{ii}, n_{ij}) \tag{6.2b}$$
$$- \left(\min\left((\overrightarrow{n}_{jj} - n_{ji})^+, n_{ij} \right) - \left(\min\left((\overrightarrow{n}_{jj} - n_{ij})^+, n_{ji} \right) - (\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+ \right)^+ \right)^+ + \eta,$$

with $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$. Theorem 10 uses the capacity region $\mathcal{C}$ (Theorem 1) and the region $\mathcal{B}_\eta$ in (6.1) to describe the η-NE region.

Theorem 10. η-NE region

Let $\eta > 0$ be fixed. The η-NE region $\mathcal{N}_\eta$ of the two-user D-LDIC-NOF with parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$, $\overleftarrow{n}_{22}$, is:

$$\mathcal{N}_\eta = \mathcal{C} \cap \mathcal{B}_\eta. \tag{6.3}$$

Proof: The proof of Theorem 10 is presented in Appendix I. ∎

The following describes some interesting observations from Theorem 10. Figure 6.1 shows the capacity region $\mathcal{C}$ and the η-NE region $\mathcal{N}_\eta$ of a channel with parameters $\overrightarrow{n}_{11} = 7$, $\overrightarrow{n}_{22} = 6$, $n_{12} = 4$, $n_{21} = 4$ and different values for $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$, with η arbitrarily small. Note that when $\overleftarrow{n}_{11} \in \{0,1,2,3,4\}$ and $\overleftarrow{n}_{22} \in \{0,1,2,3,4\}$ (Figure 6.1a), it follows that $\mathcal{N}_\eta(7,6,4,4,\overleftarrow{n}_{11},\overleftarrow{n}_{22}) = \mathcal{N}_\eta(7,6,4,4,0,0)$. Thus, in this case the use of feedback in any of the transmitter-receiver pairs does not enlarge the η-NE region. Alternatively, when $\overleftarrow{n}_{11} > 4$ and $\overleftarrow{n}_{22} \in \{0,1,2,3,4\}$ (Figures 6.1b, 6.1c and 6.1d), the resulting η-NE region is larger than in the previous case. A similar effect is observed in Figures 6.1e and 6.1f. This observation implies the existence of a threshold on each feedback parameter $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$ beyond which the η-NE region is enlarged. The exact values of $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$, given a fixed 4-tuple ($\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}), beyond which the η-NE region can be enlarged is presented in Section 6.2.

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). When $\overrightarrow{n}_{jj} > \max(n_{ji}, n_{ij})$ and the condition

$$\Big(\left(\overrightarrow{n}_{jj} > n_{ij} + n_{ji} \wedge \max\left(n_{ji}, \overrightarrow{n}_{ii} - (\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii})^+\right) > (\overrightarrow{n}_{ii} - n_{ij})^+ \right) \vee$$
$$\left(\overrightarrow{n}_{jj} \leqslant n_{ij} + n_{ji} \wedge \overrightarrow{n}_{ii} < n_{ij} + n_{ji} \wedge \overrightarrow{n}_{ii} > n_{ij} \wedge n_{ji} > \overrightarrow{n}_{ii} - (\overrightarrow{n}_{ii} - \overleftarrow{n}_{ii})^+ \right) \vee$$
$$\left(\overrightarrow{n}_{jj} \leqslant n_{ij} + n_{ji} \wedge \overrightarrow{n}_{ii} < n_{ij} + n_{ji} \wedge \overrightarrow{n}_{ii} > n_{ij} \wedge n_{ij} > (\overrightarrow{n}_{ii} - \overleftarrow{n}_{ii})^+ \right) \vee$$
$$\left(\overrightarrow{n}_{jj} \leqslant n_{ij} + n_{ji} \wedge \overrightarrow{n}_{ii} < n_{ij} + n_{ji} \wedge \overrightarrow{n}_{ii} \leqslant n_{ij} \wedge \overrightarrow{n}_{ii} > n_{ij} - \overrightarrow{n}_{jj} + n_{ji} \right) \Big) \tag{6.4}$$

holds, the bound $R_i \leqslant U_i$ is active. In this case,

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

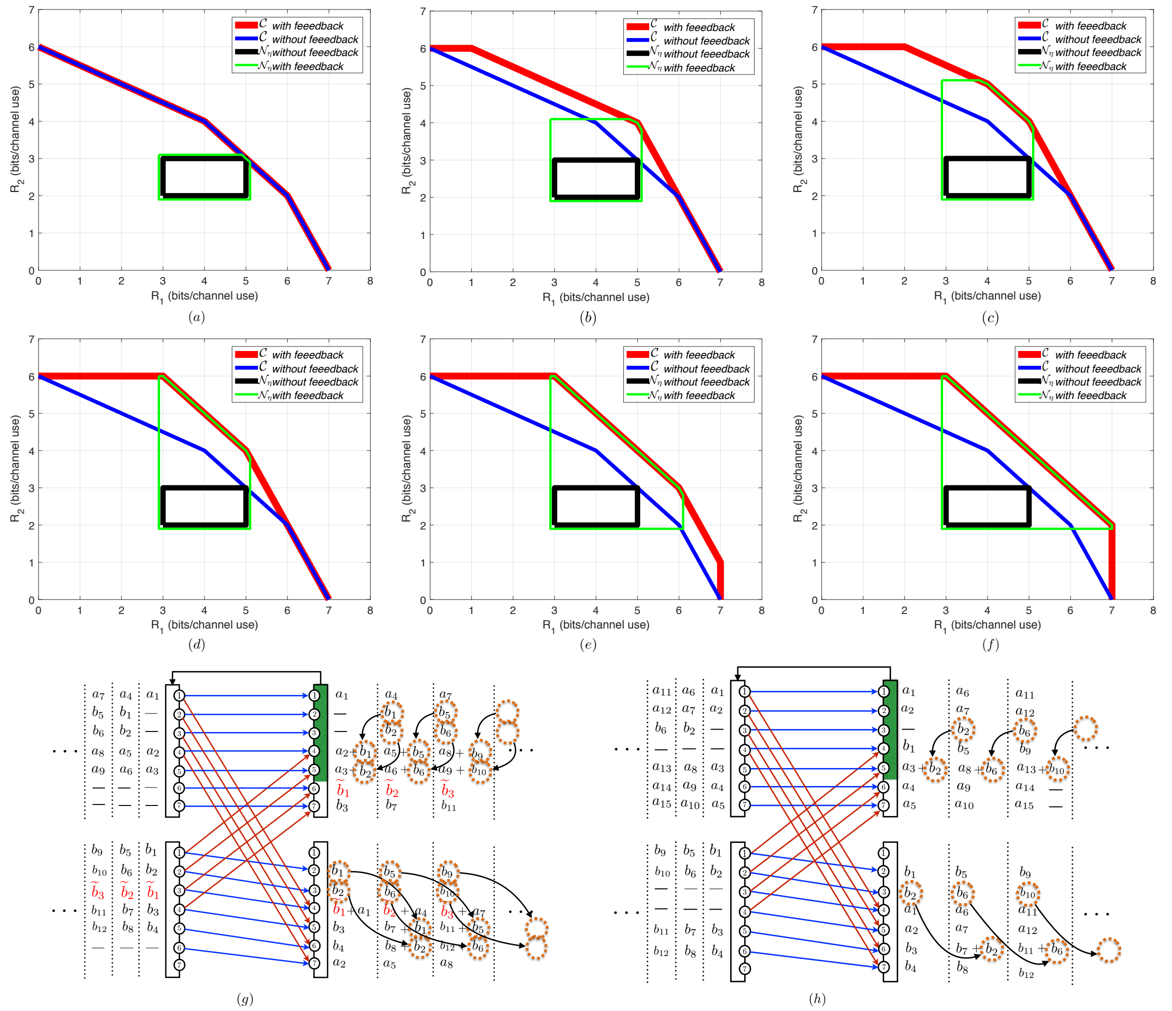


Figure 6.1.: Capacity region $\mathcal{C}(7,6,4,4,0,0)$ (thin blue line) and η-NE region $\mathcal{N}_\eta(7,6,4,4,0,0)$ (thick black line) with η arbitrarily small. Fig. 6.1a shows the capacity region $\mathcal{C}(7,6,4,4,\overleftarrow{n}_{11},\overleftarrow{n}_{22})$ (thick red line) and the η-NE region $\mathcal{N}_\eta(7,6,4,4,\overleftarrow{n}_{11},\overleftarrow{n}_{22})$ (thin green line), with $\overleftarrow{n}_{11} \in \{0,1,2,3,4\}$ and $\overleftarrow{n}_{22} \in \{0,1,2,3,4\}$. Fig. 6.1b shows the capacity region $\mathcal{C}(7,6,4,4,5,\overleftarrow{n}_{22})$ (thick red line) and the η-NE region $\mathcal{N}_\eta(7,6,4,4,5,\overleftarrow{n}_{22})$ (thin green line), with $\overleftarrow{n}_{22} \in \{0,1,2,3,4\}$. Fig. 6.1c shows the capacity region $\mathcal{C}(7,6,4,4,6,\overleftarrow{n}_{22})$ (thick red line) and the η-NE region $\mathcal{N}_\eta(7,6,4,4,6,\overleftarrow{n}_{22})$ (thin green line), with $\overleftarrow{n}_{22} \in \{0,1,2,3,4\}$. Fig. 6.1d shows the capacity region $\mathcal{C}(7,6,4,4,7,\overleftarrow{n}_{22})$ (thick red line) and the η-NE region $\mathcal{N}_\eta(7,6,4,4,7,\overleftarrow{n}_{22})$ (thin green line), with $\overleftarrow{n}_{22} \in \{0,1,2,3,4\}$. Fig. 6.1e shows the capacity region $\mathcal{C}(7,6,4,4,7,5)$ (thick red line) and the η-NE region $\mathcal{N}_\eta(7,6,4,4,7,5)$ (thin green line). Fig. 6.1f shows the capacity region $\mathcal{C}(7,6,4,4,7,6)$ (thick red line) and the η-NE region $\mathcal{N}_\eta(7,6,4,4,7,6)$ (thin green line). Fig. 6.1g and Fig. 6.1h illustrate the achievability scheme for the equilibrium rate pair $(3,4)$ and $(5,4)$ in $\mathcal{N}_\eta(7,6,4,4,5,0)$.

and the following is a necessary condition to observe a larger η-NE region with respect to the case in which feedback in transmitter-receiver j, *i.e.*, $\overleftarrow{n}_{jj}$, is not available:

$$\overleftarrow{n}_{jj} > \max\left(n_{ij}, \overrightarrow{n}_{jj} - n_{ji}\right). \tag{6.5}$$

Note that condition (6.5) is identical to the condition needed to observe an enlargement of the capacity region in this case (see Section 4.2).

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 [16]. This result is obtained as a corollary of Theorem 10.

Corollary 5 (Theorem 1 in [16])**.** The η-NE region of the two-user decentralized linear deterministic interference channel (D-LDIC) without channel-output feedback, with parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, and n_{21}, is $\mathcal{N}_\eta(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, 0, 0)$.

The η-NE region with POF, *i.e.*, $\overleftarrow{n}_{11} \geqslant \max(\overrightarrow{n}_{11}, n_{12})$ and $\overleftarrow{n}_{22} \geqslant \max(\overrightarrow{n}_{22}, n_{21})$, is described by Theorem 1 in [66]. This result can also be obtained as a corollary of Theorem 10.

Corollary 6 (Theorem 1 in [66])**.** The η-NE region of the two-user D-LDIC 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 [68]. This result can also be obtained as a corollary of Theorem 10.

Corollary 7 (Theorem 1 in [68])**.** The η-NE region of the two-user symmetric D-LDIC-NOF, *e.g.*, $\overrightarrow{n}_{11} = \overrightarrow{n}_{22} = \overrightarrow{n}$, $n_{12} = n_{21} = m$, and $\overleftarrow{n}_{11} = \overleftarrow{n}_{22} = \overleftarrow{n}$, is $\mathcal{N}_\eta(\overrightarrow{n}, \overrightarrow{n}, m, m, \overleftarrow{n}, \overleftarrow{n})$.

From the comments above, it is interesting to highlight the following inclusions:

$$\begin{aligned}&\mathcal{N}_\eta\Big(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, 0, 0\Big) \subseteq \mathcal{N}_\eta\Big(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}\Big) \subseteq \\ &\mathcal{N}_\eta\Big(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \max\left(\overrightarrow{n}_{11}, n_{12}\right), \max\left(\overrightarrow{n}_{22}, n_{21}\right)\Big),\end{aligned} \tag{6.6}$$

for all $\eta > 0$. The inclusions above might appear trivial, however, enlarging the set of actions often leads to paradoxes (Braess' Paradox [19]) in which the new game possesses equilibria at which players obtain smaller individual benefits and/or smaller total benefit. Nonetheless, letting both transmitter-receiver pairs to use feedback does not induce this type of paradoxes with respect to the case without feedback.

Consider again the example in which $\overrightarrow{n}_{11} = 7$, $\overrightarrow{n}_{22} = 6$, $n_{12} = 4$, $n_{21} = 4$, $\overleftarrow{n}_{11} = 5$ and $\overleftarrow{n}_{22} = 0$ (see Figure 6.1b). In this case, the η-NE region $\mathcal{N}_\eta$ is the convex hull of the rate pairs $(3, 2)$, $(3, 4)$, $(5, 4)$, and $(5, 2)$. The rate pair $(3, 4)$ is achieved at an η-NE thanks to the use of feedback in transmitter-receiver pair 1. Transmitter 1 uses the bit-pipes 2 and 3 of the channel input $\boldsymbol{X}_{1,n}$ to re-transmit during channel use n two bits that have been previously transmitted by transmitter 2 and have produced interference at receiver 1 during channel use $n-1$ (see Figure 6.1g). Note that there are four bit-pipes at receiver 1 impaired by interference from transmitter 2, however, only two bits can be fed back due to the effect of noise in the feedback channel. At channel use n, transmitter 1 re-transmits the interfering bits through bit-pipes 2 and 3 that are simultaneously received by receiver 1 and receiver 2. At receiver 2, these bits are seen at bit-pipes 5 and 6. However, these bits do not represent any interference for receiver 2 since they were received interference-free at channel use $n-1$, and thus, they can be cancelled at channel use n. At receiver 1, these bits are seen during channel use n at bit-pipes 2 and 3 and thus, interference-free. Hence, at channel use n, receiver 1 can cancel the interference produced during channel use $n-1$. In this case, transmitter 1 and transmitter 2 are able to send three and four 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 6.1g. 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, from 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 $(3, 4)$ is achieved at an η-NE. The use of common randomness is also observed in [16, 66, 68]. Common randomness reflects a competitive behavior between both transmitter-receiver pairs.

The achievability of the rate pair $(5, 4)$ follows the same explanation of the achievability of the η-NE rate pair $(3, 4)$ with the difference that for this rate pair, it is not necessary that transmitter 2 sends randomly generated bits (see Figure 6.1h), and thus, transmitter-receiver pair 1 achieves a greater rate at an η-NE with respect to the previous example. This suggests a more altruistic behavior. In this case, transmitter 1 and transmitter 2 are able to send five and four 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 $(5, 4)$ is achieved at an η-NE.

6.2. Enlargement of the η-Nash Equilibrium Region with Feedback

The metrics, the conditions, and the values on the feedback parameters beyond which the η-NE region of the two-user LDIC-NOF can be enlarged are the same as in the centralized case, taking into account that these are referred to the η-NE region instead of the capacity region.

6.3. Efficiency of the η-NE

This section characterizes the efficiency of the set of equilibria in the two-user D-LDIC-NOF using two metrics: price of anarchy (PoA) and price of stability (PoS). The PoA measures the loss of performance due to decentralization by comparing the maximum sum-rate achieved by a centralized two-user LDIC-NOF with the minimum sum-rate achieved by a decentralized two-user LDIC-NOF at an η-NE. That is, the ratio between the sum-rate capacity and the smallest sum-rate at an η-NE region. Alternatively, the PoS measures the loss of performance due to decentralization by comparing the maximum sum-rate achieved by a centralized two-user LDIC-NOF with the maximum sum-rate achieved by a decentralized two-user LDIC-NOF at an η-NE region. That is, the ratio between the sum-rate capacity and the biggest sum-rate at an η-NE region [67].

6.3.1. Definitions

The results of this section are presented using a list of events (Boolean variables) that are determined by the parameters $\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}$, and n_{21}. Let $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$, and define the following events:

$$\begin{aligned}
A_{1,i}:&\quad \overrightarrow{n}_{ii} - n_{ij} \geqslant n_{ji}, && (6.7a)\\
A_{2,i}:&\quad \overrightarrow{n}_{ii} \geqslant n_{ji}, && (6.7b)\\
B_1:&\quad A_{1,1} \wedge A_{1,2}, && (6.7c)\\
B_{2,i}:&\quad A_{1,i} \wedge \overline{A}_{1,j} \wedge A_{2,j}, && (6.7d)\\
B_{3,i}:&\quad A_{1,i} \wedge \overline{A}_{1,j} \wedge \overline{A}_{2,j}, && (6.7e)\\
B_4:&\quad \overline{A}_{1,1} \wedge \overline{A}_{1,2} \wedge A_{2,1} \wedge A_{2,2}, && (6.7f)\\
B_{5,i}:&\quad \overline{A}_{1,1} \wedge \overline{A}_{1,2} \wedge \overline{A}_{2,i} \wedge A_{2,j}, && (6.7g)\\
B_6:&\quad \overline{A}_{1,1} \wedge \overline{A}_{1,2} \wedge \overline{A}_{2,1} \wedge \overline{A}_{2,2}, && (6.7h)\\
B_7:&\quad A_{1,1}, && (6.7i)\\
B_8:&\quad \overline{A}_{1,1} \wedge A_{2,1} \wedge A_{2,2}, && (6.7j)\\
B_9:&\quad \overline{A}_{1,1} \wedge \overline{A}_{2,1} \wedge A_{2,2}, && (6.7k)\\
B_{10}:&\quad \overline{A}_{1,1} \wedge \overline{A}_{2,2}. && (6.7l)
\end{aligned}$$

When both transmitter-receiver pairs are in LIR, *i.e.*, $\overrightarrow{n}_{11} > n_{12}$ and $\overrightarrow{n}_{22} > n_{21}$, the events B_1, $B_{2,1}$, $B_{2,2}$, $B_{3,1}$, $B_{3,2}$, B_4, $B_{5,1}$, $B_{5,2}$, and B_6 exhibit the property stated by:

Lemma 19. *For a fixed 4-tuple* $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$ *with* $\overrightarrow{n}_{11} > n_{12}$ *and* $\overrightarrow{n}_{22} > n_{21}$, *only one of the events* B_1, $B_{2,1}$, $B_{2,2}$, $B_{3,1}$, $B_{3,2}$, B_4, $B_{5,1}$, $B_{5,2}$, *and* B_6 *holds true.*

Proof: The proof follows from verifying that when both transmitter-receiver pairs are in LIR, *i.e.*, $\overrightarrow{n}_{11} > n_{12}$ and $\overrightarrow{n}_{22} > n_{21}$, the events (6.7c)-(6.7h) are mutually exclusive. This completes the proof. ■

When transmitter-receiver pair 1 is in LIR and transmitter-receiver pair 2 is in HIR, *i.e.*, $\overrightarrow{n}_{11} > n_{12}$ and $\overrightarrow{n}_{22} \leqslant n_{21}$, the events B_7, B_8, B_9, and B_{10} exhibit the property stated by:

Lemma 20. *For a fixed 4-tuple* $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}) \in \mathbb{N}^4$ *with* $\overrightarrow{n}_{11} > n_{12}$ *and* $\overrightarrow{n}_{22} \leqslant n_{21}$, *only one of the events* B_7, B_8, B_9, *and* B_{10} *holds true.*

Proof: The proof of Lemma 20 follows along the same lines as the proof of Lemma 19. ■

6.3.2. Price of Anarchy

Let $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2$ be the set of all possible configuration pairs and $\mathcal{A}_{\eta-\mathrm{NE}} \subset \mathcal{A}$ be the set of η-NE configuration pairs of the game in (3.3) (Definition 4).

Definition 7 (Price of Anarchy [45])**.** *Let* $\eta > 0$. *The PoA of the game* $\mathcal{G}$ *in* (3.3), *denoted by* $\mathrm{PoA}(\eta, \mathcal{G})$, *is given by:*

$$\mathrm{PoA}(\eta, \mathcal{G}) = \frac{\displaystyle\max_{(s_1,s_2)\in\mathcal{A}} \sum_{i=1}^{2} R_i(s_1, s_2)}{\displaystyle\min_{(s_1^*,s_2^*)\in\mathcal{A}_{\eta-\mathrm{NE}}} \sum_{i=1}^{2} R_i(s_1^*, s_2^*)}. \tag{6.8}$$

Let $\overline{\Sigma}_C$ denote the solution to the optimization problem in the numerator of (6.8), which corresponds to the maximum sum-rate in the centralized case. Let also $\underline{\Sigma}_N$ denote the solution to the optimization problem in the denominator of (6.8). Closed-form expressions of the maximum sum-rate in the centralized case, *i.e.*, $\overline{\Sigma}_C$ and the minimum sum-rate in the decentralized case, *i.e.*, $\underline{\Sigma}_N$, are presented in Appendix O.

The following theorems describe the PoA $(\eta, \mathcal{G})$ in particular interference regimes of the two-user D-LDIC-NOF. In all the cases, it is assumed that $\overleftarrow{n}_{ii} \leqslant \max(\overrightarrow{n}_{ii}, n_{ij})$ for all $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$. If $\overleftarrow{n}_{11} > \max(\overrightarrow{n}_{11}, n_{12})$ or $\overleftarrow{n}_{22} > \max(\overrightarrow{n}_{22}, n_{21})$, the results are the same as those in the case of POF, *i.e.*, $\overleftarrow{n}_{11} = \max(\overrightarrow{n}_{11}, n_{12})$ or $\overleftarrow{n}_{22} = \max(\overrightarrow{n}_{22}, n_{21})$.

Theorem 11. Both transmitter-receiver pairs in LIR

Let $\eta > 0$ be fixed. For all $i \in \{1,2\}$, $j \in \{1,2\} \setminus \{i\}$ and for all $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \in \mathbb{N}^6$ with $\overrightarrow{n}_{11} > n_{12}$ and $\overrightarrow{n}_{22} > n_{21}$, the PoA $(\eta, \mathcal{G})$ satisfies:

$$\text{PoA}(\eta, \mathcal{G}) = \begin{cases} \dfrac{\overline{\Sigma}_{C1}}{\overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21} - 2\eta} & \text{if} \quad B_1 \text{ holds true} \\[2ex] \dfrac{\overline{\Sigma}_{C2,i}}{\overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21} - 2\eta} & \text{if} \quad B_{2,i} \text{ holds true} \\[2ex] \dfrac{\overrightarrow{n}_{ii}}{\overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21} - 2\eta} & \text{if} \quad B_{3,i} \vee B_{5,i} \text{ holds true} \\[2ex] \dfrac{\overline{\Sigma}_{C3}}{\overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21} - 2\eta} & \text{if} \quad B_4 \text{ holds true} \\[2ex] \dfrac{\min(\overrightarrow{n}_{11}, \overrightarrow{n}_{22})}{\overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21} - 2\eta} & \text{if} \quad B_6 \text{ holds true} \end{cases}, \tag{6.9}$$

where,

$$\begin{aligned} \overline{\Sigma}_{C1} = \min\Big(& \overrightarrow{n}_{22} + \overrightarrow{n}_{11} - n_{12}, \overrightarrow{n}_{11} + \overrightarrow{n}_{22} - n_{21}, \\ & \max\big(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\big) + \max\big(\overrightarrow{n}_{22} - n_{21}, \overleftarrow{n}_{22}\big), \\ & 2\overrightarrow{n}_{11} - n_{12} + \max\left(\overrightarrow{n}_{22} - n_{21}, \overleftarrow{n}_{22}\right), 2\overrightarrow{n}_{22} - n_{21} + \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right)\Big), \end{aligned} \tag{6.10a}$$

$$\overline{\Sigma}_{C2,i}=\min\Big(\overrightarrow{n}_{22}+\overrightarrow{n}_{11}-n_{12}, \overrightarrow{n}_{11}+\overrightarrow{n}_{22}-n_{21}, \tag{6.10b}$$
$$\max\Big(\overrightarrow{n}_{11}-n_{12}, \overleftarrow{n}_{11}\Big)+\max\Big(\overrightarrow{n}_{22}-n_{21}, \overleftarrow{n}_{22}\Big),$$
$$2\overrightarrow{n}_{ii}-n_{ij}+\max\left(n_{ij}, \overleftarrow{n}_{jj}\right), 2\overrightarrow{n}_{jj}-n_{ji}+\max\left(\overrightarrow{n}_{ii}-n_{ij}, \overleftarrow{n}_{ii}\right)\Big), \text{ and}$$
$$\overline{\Sigma}_{C3}=\min\left(\overrightarrow{n}_{22}+\overrightarrow{n}_{11}-n_{12}, \overrightarrow{n}_{11}+\overrightarrow{n}_{22}-n_{21}, \max\left(n_{21}, \overleftarrow{n}_{11}\right)+\max\left(n_{12}, \overleftarrow{n}_{22}\right),\right.$$
$$\left.2\overrightarrow{n}_{11}-n_{12}+\max\left(n_{12}, \overleftarrow{n}_{22}\right), 2\overrightarrow{n}_{22}-n_{21}+\max\left(n_{21}, \overleftarrow{n}_{11}\right)\right). \tag{6.10c}$$

Proof: The proof is presented in Appendix O. ∎

From Theorem 11, the following conclusions can be drawn. When both transmitter-receiver pairs are in LIR, and at least one of the conditions $B_{3,i}$, $B_{5,i}$, or B_6 holds true, with $i \in \{1,2\}$, then the PoA$(\eta, \mathcal{G})$ does not depend on the feedback parameters $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$. However, under other conditions, *i.e.*, B_1, $B_{2,i}$, or B_4, this is not always the case as shown in the following corollaries:

Corollary 8. *For any* $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \in \mathbb{N}^6$ *with* $\overrightarrow{n}_{11} > n_{12}$ *and* $\overrightarrow{n}_{22} > n_{21}$, *such that* B_1 *holds true, it follows that:*

$$1 < \frac{\overrightarrow{n}_{11}+\overrightarrow{n}_{22}-n_{12}-n_{21}}{\overrightarrow{n}_{11}-n_{12}+\overrightarrow{n}_{22}-n_{21}-2\eta} \leqslant \text{PoA}\left(\eta, \mathcal{G}\right) \leqslant \frac{\overrightarrow{n}_{11}+\overrightarrow{n}_{22}-\max\left(n_{12}, n_{21}\right)}{\overrightarrow{n}_{11}-n_{12}+\overrightarrow{n}_{22}-n_{21}-2\eta}. \tag{6.11}$$

The lower bound in (6.11) is obtained assuming that $\overleftarrow{n}_{11}=0$ and $\overleftarrow{n}_{22}=0$ in (6.9). That is, when feedback is not available. The upper bound in (6.11) is obtained assuming that $\overleftarrow{n}_{11}=\max\left(\overrightarrow{n}_{11}, n_{12}\right)=\overrightarrow{n}_{11}$ and $\overleftarrow{n}_{22}=\max\left(\overrightarrow{n}_{22}, n_{21}\right)=\overrightarrow{n}_{22}$ in (6.9). That is, when POF is available at both transmitter-receiver pairs.

Note also that for any η, when both transmitter-receiver pairs are in LIR, condition B_1 holds true, $\overleftarrow{n}_{11} \leqslant \overrightarrow{n}_{11}-n_{12}$, and $\overleftarrow{n}_{22} \leqslant \overrightarrow{n}_{22}-n_{21}$, the sum-rate capacity approaches to the minimum sum-rate at an η-NE region (PoA$(\eta, \mathcal{G}) \approx 1$). Alternatively, when both transmitter-receiver pairs are in LIR, condition B_1 holds true, and at least one the following conditions: $\overleftarrow{n}_{11} > \overrightarrow{n}_{11}-n_{12}$ or $\overleftarrow{n}_{22} > \overrightarrow{n}_{22}-n_{21}$ holds true, the use of feedback in transmitter-receiver pair 1 or transmitter-receiver pair 2, respectively, enlarges both the capacity region and the η-NE region. Nonetheless, the PoA increases as the smallest sum-rate at an η-NE region remains unchanged with respect to the case without feedback.

Corollary 9. *For any* $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \in \mathbb{N}^6$ *with* $\overrightarrow{n}_{11} > n_{12}$ *and* $\overrightarrow{n}_{22} > n_{21}$, *such that* $B_{2,i}$ *holds true for all* $i \in \{1,2\}$ *and* $j \in \{1,2\} \setminus \{i\}$, *it follows that:*

$$1 < \frac{\overrightarrow{n}_{ii}}{\overrightarrow{n}_{11}-n_{12}+\overrightarrow{n}_{22}-n_{21}-2\eta} \leqslant \text{PoA}\left(\eta, \mathcal{G}\right) \leqslant \frac{\overrightarrow{n}_{11}+\overrightarrow{n}_{22}-\max\left(n_{12}, n_{21}\right)}{\overrightarrow{n}_{11}-n_{12}+\overrightarrow{n}_{22}-n_{21}-2\eta}. \tag{6.12}$$

Note that when both transmitter-receiver pairs are in LIR and for a given $i \in \{1,2\}$ condition $B_{2,i}$ holds true, $\overleftarrow{n}_{ii} \leqslant \overrightarrow{n}_{ii}-n_{ij}$; and $\overleftarrow{n}_{jj} \leqslant n_{ij}$, the use of feedback in either transmitter-receiver pair does not enlarge the capacity region or the η-NE region. Then, the PoA$(\eta, \mathcal{G})$ is equal to the lower bound in (6.12), *i.e.*, $\text{PoA}(\eta,\mathcal{G}) = \frac{\overrightarrow{n}_{ii}}{\overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21} - 2\eta}$. Conversely, when both transmitter-receiver pairs are in LIR and for a given $i \in \{1,2\}$ condition $B_{2,i}$ holds true, and at least one of the following conditions: $\overleftarrow{n}_{ii} > \overrightarrow{n}_{ii} - n_{ij}$ or $\overleftarrow{n}_{jj} > n_{ij}$ holds true, the use of feedback enlarges both the capacity region and the η-NE region.

The lower and upper bounds in (6.12) are obtained as in the case of (6.11).

Corollary 10. *For any* $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \in \mathbb{N}^6$ *with* $\overrightarrow{n}_{11} > n_{12}$ *and* $\overrightarrow{n}_{22} > n_{21}$, *such that* B_4 *holds true, it follows that:*

$$1 < \frac{\min\left(\overrightarrow{n}_{11} + \overrightarrow{n}_{22} - \max(n_{12}, n_{21}), n_{12} + n_{21}\right)}{\overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21} - 2\eta} \leqslant \text{PoA}(\eta,\mathcal{G}) \leqslant \frac{\overrightarrow{n}_{11} + \overrightarrow{n}_{22} - \max(n_{12}, n_{21})}{\overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21} - 2\eta}. \tag{6.13}$$

Note that when both transmitter-receiver pairs are in LIR, condition B_4 holds true, and $\overline{\Sigma}_{C5} \leqslant n_{12} + n_{21}$, then the $\text{PoA}(\eta,\mathcal{G})$ does not depend on the feedback parameters $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$. When both transmitter-receiver pairs are in LIR, condition B_4 holds true, $\overrightarrow{n}_{11} + \overrightarrow{n}_{22} - \max(n_{12}, n_{21}) > n_{12} + n_{21}$; $\overleftarrow{n}_{11} \leqslant n_{21}$, and $\overleftarrow{n}_{22} \leqslant n_{12}$, then the $\text{PoA}(\eta,\mathcal{G})$ is equal to the lower bound in (6.13), *i.e.*, $\text{PoA}(\eta,\mathcal{G}) = \frac{n_{12} + n_{21}}{\overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21} - 2\eta}$. Conversely, when both transmitter-receiver pairs are in LIR, condition B_4 holds true, $\overrightarrow{n}_{11} + \overrightarrow{n}_{22} - \max(n_{12}, n_{21}) > n_{12} + n_{21}$, and at least one of the following conditions: $\overleftarrow{n}_{11} > n_{21}$ or $\overleftarrow{n}_{22} > n_{12}$ holds true, the use of feedback in transmitter-receiver pair 1 or transmitter-receiver pair 2, respectively, enlarges the capacity region and the η-NE region.

Theorem 12. Transmitter-receiver pair 1 in LIR and transmitter-receiver pair 2 in HIR

Let $\eta > 0$ be fixed. For all $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \in \mathbb{N}^6$ with $\overrightarrow{n}_{11} > n_{12}$ and $\overrightarrow{n}_{22} \leqslant n_{21}$, the $\text{PoA}(\eta,\mathcal{G})$ satisfies:

$$\text{PoA}(\eta,\mathcal{G}) = \begin{cases} \frac{\overrightarrow{n}_{11}}{\overrightarrow{n}_{11} - n_{12} - \eta} & \text{if} \quad B_7 \vee B_8 \vee B_{10} \text{ holds true} \\ \frac{\min(\overrightarrow{n}_{22} + \overrightarrow{n}_{11} - n_{12}, n_{21})}{\overrightarrow{n}_{11} - n_{12} - \eta} & \text{if} \quad B_9 \text{ holds true} \end{cases}. \tag{6.14}$$

Proof: The proof is presented in Appendix O. ∎

Note that in the cases in which transmitter-receiver pair 1 is in LIR and transmitter-receiver pair 2 is in HIR, the $\text{PoA}(\eta,\mathcal{G})$ does not depend on the feedback parameters. This follows since the use of feedback in this scenario can enlarge the capacity region but it does not increase the sum-rate capacity (Theorem 5).

In the case in which transmitter-receiver pair 1 is in HIR and transmitter-receiver pair 2 is in LIR, *i.e.*, $\overrightarrow{n}_{11} \leqslant n_{12}$ and $\overrightarrow{n}_{22} > n_{21}$, the $\text{PoA}(\eta,\mathcal{G})$ for the two-user D-LDIC-NOF is characterized as in Theorem 12 interchanging the indices of the parameters.

Theorem 13. Both transmitter-receiver pairs in HIR

Let $\eta > 0$ be fixed. For all $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \in \mathbb{N}^6$ with $\overrightarrow{n}_{11} \leqslant n_{12}$ and $\overrightarrow{n}_{22} \leqslant n_{21}$, the PoA $(\eta, \mathcal{G})$ satisfies:

$$\mathrm{PoA}\,(\eta, \mathcal{G}) = \infty. \tag{6.15}$$

The result in Theorem 13 is due to the fact that $\left(\left(\overrightarrow{n}_{11}-n_{12}\right)^+-\eta\right)^+ + \left(\left(\overrightarrow{n}_{22}-n_{21}\right)^+-\eta\right)^+ = 0$. That is, when $\overrightarrow{n}_{11} \leqslant n_{12}$ and $\overrightarrow{n}_{22} \leqslant n_{21}$, none of the transmitter-receiver pairs is able to transmit at a strictly positive rate at the worst η-NE (the smallest sum-rate at the η-NE region).

In general, in any interference regime in which the PoA $(\eta, \mathcal{G})$ depends on the feedback parameters $\overleftarrow{n}_{11}$ or $\overleftarrow{n}_{22}$, there exists a value in the feedback parameter $\overleftarrow{n}_{11}$ or the feedback parameter $\overleftarrow{n}_{22}$ beyond which the PoA $(\eta, \mathcal{G})$ increases. These values correspond to those values beyond which the sum-capacity can be enlarged (Theorem 5).

6.3.3. Price of Stability

In this section, the efficiency of the η-NE of the game $\mathcal{G}$ in (3.3) is analyzed by using the PoS.

Definition 8 (Price of stability [4])**.** *Let $\eta > 0$. The PoS of the game $\mathcal{G}$ in* (3.3)*, denoted by* $\mathrm{PoS}\,(\eta, \mathcal{G})$*, is given by:*

$$\mathrm{PoS}\,(\eta, \mathcal{G}) = \frac{\displaystyle\max_{(s_1,s_2)\in\mathcal{A}} \sum_{i=1}^{2} R_i(s_1, s_2)}{\displaystyle\max_{(s_1^*,s_2^*)\in\mathcal{A}_{\eta-\mathrm{NE}}} \sum_{i=1}^{2} R_i(s_1^*, s_2^*)}. \tag{6.16}$$

Let $\overline{\Sigma}_N$ denote the solution to the optimization problem in the denominator of (6.16). A closed-form expression of the maximum sum-rate in the decentralized case, *i.e.*, $\overline{\Sigma}_N$ is presented in Appendix O and it can be obtained from Theorem 10.

The following proposition characterizes the PoS of the game $\mathcal{G}$ in (3.3) for the two-user D-LDIC-NOF.

Proposition 2 (PoS)**.** For all $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \in \mathbb{N}^6$ and for all $\eta > 0$ arbitrary small, the PoS in the game $\mathcal{G}$ of the two-user D-LDIC-NOF is:

$$\mathrm{PoS}\,(\eta, \mathcal{G}) = 1. \tag{6.17}$$

Proof: The proof of Proposition 2 is obtained from Lemma 34 and Lemma 35 in Appendix O. ■

Note that the fact that the price of stability is equal to one, independently of the parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$, implies that despite the anarchical behavior of both transmitter-receiver pairs, the largest η-NE sum-rate is equal to the sum-rate capacity, *i.e.*, $\overline{\Sigma}_C = \overline{\Sigma}_N$. This implies that in all interference regimes, there always exists an η-NE that is sum-rate optimal (Pareto optimal η-NE). The thresholds on the feedback parameters beyond which the sum-capacity and the maximum sum-rate in the η-NE region can be improved can be obtained from Theorem 5.

In conclusion, with $\eta > 0$ and when both transmitter-receiver pairs are in LIR, the PoA can be made arbitrarily close to one as η approaches zero, subject to a particular condition. This immediately implies that in this regime even the worst η-NE (in terms of sum-rate) is arbitrarily close to the Pareto boundary of the capacity region. The use of feedback increases the PoA in some interference regimes. This is basically because in these regimes, the use of feedback increases the sum-capacity, whereas the smallest sum-rate at an η-NE region is not changed. In some cases the PoA can be infinite due to the fact that when both transmitter-receiver pairs are in HIR, the smallest sum-rate at an η-NE region is zero bit per channel use. In other regimes, the use of feedback does not have any impact on the PoA as it does not increase the sum-capacity. Finally, the PoS is shown to be equal to one in all interference regimes. This implies that there always exists an η-NE in the Pareto boundary of the capacity region. These results highlight the relevance of designing equilibrium selection methods such that decentralized networks can operate at efficient η-NE points. The need of these methods becomes more relevant when channel-output feedback is available as it might increase the PoA.

— 7 —

Gaussian Interference Channel

This chapter presents the main results on the two-user D-GIC-NOF. This model was described in Section 2.1 and can be modeled by a game as suggested in Section 3.1. Denote by $\mathcal{C}$ 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 capacity region $\underline{\mathcal{C}}$ and the converse region $\overline{\mathcal{C}}$ approximate the capacity region $\mathcal{C}$ to within 4.4 bits (Theorem 9). The achievable capacity region $\underline{\mathcal{C}}$ and the converse region $\overline{\mathcal{C}}$ are defined by Theorem 7 and Theorem 8, respectively.

7.1. Achievable η-Nash Equilibrium Region

Let the η-NE region (Definition 5) of the two-user D-GIC-NOF be denoted by $\mathcal{N}_\eta$. This section introduces a region $\underline{\mathcal{N}}_\eta \subseteq \mathcal{N}_\eta$ that is achievable using the randomized Han-Kobayashi scheme with noisy channel-output feedback (RHK-NOF). This coding scheme is presented in Appendix M and Appendix N. The RHK-NOF is proved to be an η-NE configuration pair with $\eta \geqslant 1$. 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 at most one bit per channel use. The description of the achievable η-NE 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 $a_{7,i} : [0,1]^3 \rightarrow \mathbb{R}_+$, defined in (5.1), for all $i \in \{1,2\}$, with $j \in \{1,2\} \setminus \{i\}$, and the functions $b_{l,i} : [0,1] \rightarrow \mathbb{R}_+$, with $(l,i) \in \{1,2\}^2$, defined in (5.2).

Note that the functions in (5.1) and (5.2) 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 in Theorem 14.

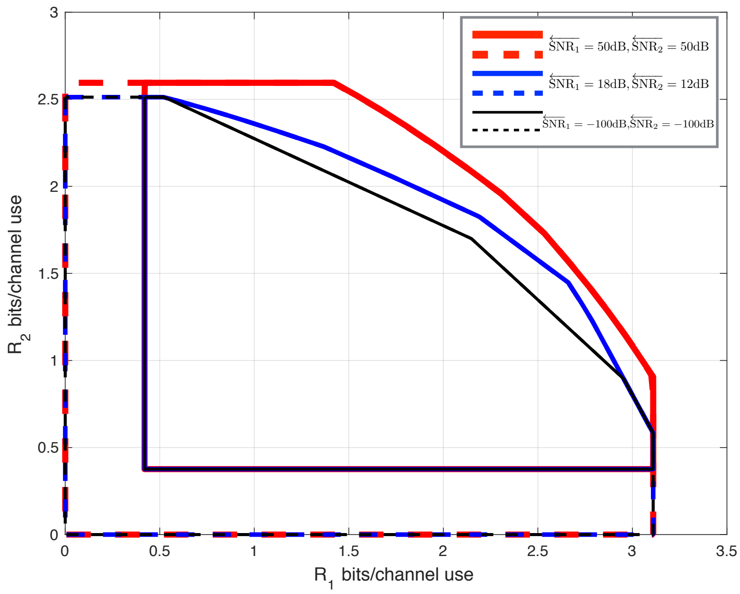


Figure 7.1.: Achievable capacity regions (dashed-lines) and achievable η-NE regions (solid lines) of the two-user GIC-NOF and two-user D-GIC-NOF with parameters $\overrightarrow{\mathrm{SNR}}_1 = 24$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 18$ dB, $\mathrm{INR}_{12} = 16$ dB, $\mathrm{INR}_{21} = 10$ dB, $\overleftarrow{\mathrm{SNR}}_1 \in \{-100, 18, 50\}$ dB, $\overleftarrow{\mathrm{SNR}}_2 \in \{-100, 12, 50\}$ dB and $\eta = 1$.

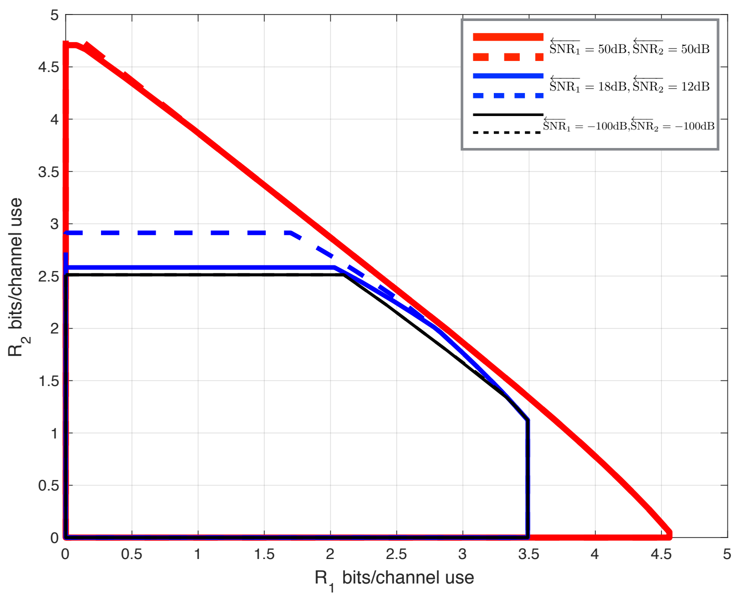


Figure 7.2.: Achievable capacity regions (dashed-lines) and achievable η-NE regions (solid lines) of the two-user GIC-NOF and two-user D-GIC-NOF with parameters $\overrightarrow{\mathrm{SNR}}_1 = 24$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 18$ dB, $\mathrm{INR}_{12} = 48$ dB, $\mathrm{INR}_{21} = 30$ dB, $\overleftarrow{\mathrm{SNR}}_1 \in \{-100, 18, 50\}$ dB, $\overleftarrow{\mathrm{SNR}}_2 \in \{-100, 12, 50\}$ dB and $\eta = 1$.

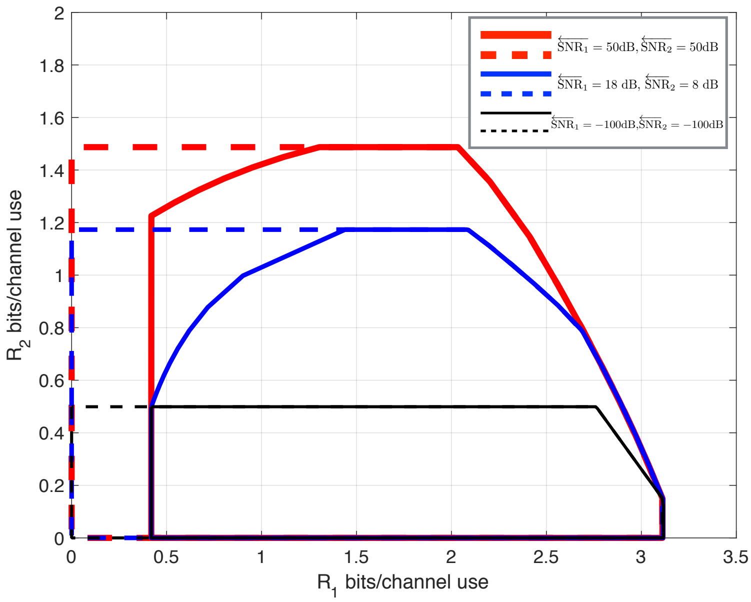


Figure 7.3.: Achievable capacity regions (dashed-lines) and achievable η-NE regions (solid lines) of the two-user GIC-NOF and two-user D-GIC-NOF with parameters $\overrightarrow{\mathrm{SNR}}_1 = 24$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 3$ dB, $\mathrm{INR}_{12} = 16$ dB, $\mathrm{INR}_{21} = 9$ dB, $\overleftarrow{\mathrm{SNR}}_1 \in \{-100, 18, 50\}$ dB, $\overleftarrow{\mathrm{SNR}}_2 \in \{-100, 8, 50\}$ dB and $\eta = 1$.

Theorem 14. Achievable η-NE region

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

$$R_i \geqslant \Big(a_{2,i}(\rho) - a_{3,i}(\rho, \mu_j) - a_{4,i}(\rho, \mu_j) - \eta\Big)^+, \tag{7.1a}$$

$$\begin{aligned} R_i \leqslant \min\Big(& a_{2,i}(\rho) + a_{3,j}(\rho, \mu_i) + a_{5,j}(\rho, \mu_i) - a_{2,j}(\rho) + \eta, \\ & a_{3,i}(\rho, \mu_j) + a_{7,i}(\rho, \mu_1, \mu_2) + 2a_{3,j}(\rho, \mu_i) + a_{5,j}(\rho, \mu_i) - a_{2,j}(\rho) + \eta, \\ & a_{2,i}(\rho) + a_{3,i}(\rho, \mu_j) + 2a_{3,j}(\rho, \mu_i) + a_{5,j}(\rho, \mu_i) + a_{7,j}(\rho, \mu_1, \mu_2) - 2a_{2,j}(\rho) + 2\eta\Big), \end{aligned} \tag{7.1b}$$

$$R_1 + R_2 \leqslant a_{1,i} + a_{3,i}(\rho, \mu_j) + a_{7,i}(\rho, \mu_1, \mu_2) + a_{2,j}(\rho) + a_{3,j}(\rho, \mu_1) - a_{2,i}(\rho) + \eta, \tag{7.1c}$$

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 Theorem 14 is presented in Appendix J. ■

The following describes some interesting observations from Theorem 14. Figure 7.1 shows an inner bound on the capacity region (Theorem 8) and the achievable η-NE region in Theorem 14 for a two-user D-GIC-NOF with parameters $\overrightarrow{\mathrm{SNR}}_1 = 24$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 18$ dB, $\mathrm{INR}_{12} = 16$ dB, $\mathrm{INR}_{21} = 10$ dB, $\overleftarrow{\mathrm{SNR}}_1 \in \{-100, 18, 50\}$ dB and $\overleftarrow{\mathrm{SNR}}_2 \in \{-100, 12, 50\}$ dB. At low values of $\overleftarrow{\mathrm{SNR}}_1$ and $\overleftarrow{\mathrm{SNR}}_2$, the achievable η-NE region approaches the region reported in [16] for the case of the two-user decentralized GIC (D-GIC) without channel-output feedback. Alternatively,

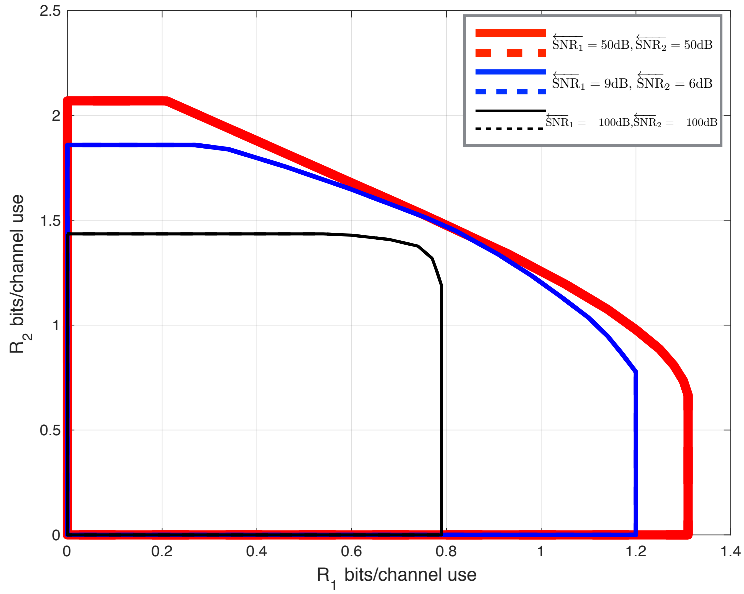


Figure 7.4.: Converse regions (dashed-lines) and non-equilibrium regions with $\eta \geqslant 1$ (solid lines) of the two-user GIC-NOF and two-user D-GIC-NOF with parameters $\overrightarrow{\mathrm{SNR}}_1 = 3$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 8$ dB, $\mathrm{INR}_{12} = 16$ dB, $\mathrm{INR}_{21} = 5$ dB, $\overleftarrow{\mathrm{SNR}}_1 \in \{-100, 9, 50\}$ dB and $\overleftarrow{\mathrm{SNR}}_2 \in \{-100, 6, 50\}$ dB.

for high values of $\overleftarrow{\mathrm{SNR}}_1$ and $\overleftarrow{\mathrm{SNR}}_2$, the achievable η-NE region approaches the region reported in [66] for the case of the two-user D-GIC with POF.

Denote by $\underline{\mathcal{N}}_{\eta\mathrm{PF}}$ the achievable η-NE region of the two-user GIC-POF presented in [66]. Figure 7.2 shows an inner bound on the capacity region (Theorem 7) and the achievable η-NE region in Theorem 14 for a two-user D-GIC-NOF 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, $\overleftarrow{\mathrm{SNR}}_1 \in \{-100, 18, 50\}$ dB and $\overleftarrow{\mathrm{SNR}}_2 \in \{-100, 12, 50\}$ dB. In this case, the achievable η-NE region and the inner bound on the capacity region (Theorem 7) 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 achievable region in the centralized case (Theorem 7).

Figure 7.3 shows an inner bound on the capacity region (Theorem 7) and the achievable η-NE region in Theorem 14 for a two-user D-GIC-NOF channel with parameters $\overrightarrow{\mathrm{SNR}}_1 = 24$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 3$ dB, $\mathrm{INR}_{12} = 16$ dB, $\mathrm{INR}_{21} = 9$ dB, $\overleftarrow{\mathrm{SNR}}_1 \in \{-100, 18, 50\}$ dB and $\overleftarrow{\mathrm{SNR}}_2 \in \{-100, 8, 50\}$ dB. Note that in this case, the feedback parameter $\overleftarrow{\mathrm{SNR}}_2$ does not have an effect on the achievable η-NE region and the inner bound on the capacity region (Theorem 7). This is due to the fact that when one transmitter-receiver pair is in LIR and the other transmitter-receiver pair is in HIR, feedback is useless on the transmitter-receiver pair in HIR.

7.2. Non-Equilibrium Region

Let the η-NE region (Def. 5) of the two-user D-GIC-NOF be denoted by $\mathcal{N}_\eta$. This section introduces a region $\overline{\mathcal{N}}_\eta \supseteq \mathcal{N}_\eta$ which is given in terms of the convex region $\overline{\mathcal{B}}_\eta$. Here, for the case of the two-user D-GIC-NOF, the convex region $\overline{\mathcal{B}}_\eta$ is given by the closure of non-negative rate pairs (R_1, R_2) that satisfy for all $i \in \{1,2\}$, with $j \in \{1,2\} \setminus \{i\}$:

$$\overline{\mathcal{B}}_\eta = \Big\{(R_1, R_2) \in \mathbb{R}_+^2 : L_i \leqslant R_i \leqslant \overline{U}_i, \text{ for all } i \in \mathcal{K} = \{1,2\}\Big\}, \tag{7.2}$$

where

$$L_i = \left(\frac{1}{2}\log\left(1 + \frac{\overrightarrow{\mathrm{SNR}}_i}{1+\mathrm{INR}_{ij}}\right) - \eta\right)^+ \text{ and} \tag{7.3}$$

$$\begin{aligned}
\overline{U}_i =& \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_i + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_i\mathrm{INR}_{ij}} + \mathrm{INR}_{ij} + 1\right)\\
&-\frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_j + 2\left(\rho - \rho_{X_iV_j}\sqrt{\gamma_j}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_j\mathrm{INR}_{ji}} + \mathrm{INR}_{ji} + 1}{(1-\gamma_j)\,\overrightarrow{\mathrm{SNR}}_j + 2\left(\rho - \rho_{X_iV_j}\sqrt{\gamma_j}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_j\mathrm{INR}_{ji}} + \mathrm{INR}_{ji} + 1}\right)\\
&-\frac{1}{2}\log\left(\mathrm{INR}_{ji}\left(1-\rho^2\right)+1\right) + \frac{1}{2}\log\left(\mathrm{INR}_{ji}\left(1 - \left(\rho - \rho_{X_iV_j}\sqrt{\gamma_j}\right)^2\right) + 1\right)\\
&+\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_j + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_j\mathrm{INR}_{ji}} + \mathrm{INR}_{ji} + 1\right)\\
&-\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_j\left(\gamma_j - \gamma_j^2\right) + 2\gamma_j\left(\rho - \rho_{X_iV_j}\sqrt{\gamma_j}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_j\mathrm{INR}_{ji}} + \gamma_j\mathrm{INR}_{ji}\left(1 - \rho_{X_iV_j}^2\right) + \gamma_j\right)\\
&-\frac{1}{2}\log\left(1-\rho_{X_iV_j}^2\right) - \frac{1}{2}\log\left(\mathrm{INR}_{ij}\left(1-\rho^2\right)+1\right)\\
&+\frac{1}{2}\log\left(\gamma_j\left(\mathrm{INR}_{ij}\left(1-\rho^2\right)+1\right) - \rho_{X_iV_j}^2\gamma_j\left(\mathrm{INR}_{ij}+1\right) + \gamma_2\mathrm{INR}_{ij}\left(2\rho_{X_iV_j}\rho\sqrt{\gamma_j} - \gamma_j\right)\right)\\
&+\eta,
\end{aligned} \tag{7.4}$$

with

$$\gamma_j = \begin{cases}
\min\left(\frac{\overrightarrow{\mathrm{SNR}}_j}{\mathrm{INR}_{ij}\mathrm{INR}_{ji}}, \frac{1}{\mathrm{INR}_{ji}}\right) & \text{if } C_{1,j} \vee \left(C_{2,j} \wedge C_{3,j}\right) \text{ holds true}\\
\min\left(\frac{1}{\mathrm{INR}_{ji}}, \frac{\mathrm{INR}_{ij}}{\mathrm{INR}_{ji}\overrightarrow{\mathrm{SNR}}_j}\right) & \text{if } C_{4,j} \text{ holds true}\\
\min\left(1, \frac{\mathrm{INR}_{ij}}{\overrightarrow{\mathrm{SNR}}_j}\right) & \text{if } C_{5,j} \wedge C_{6,j} \text{ holds true}\\
\min\left(\frac{\overrightarrow{\mathrm{SNR}}_j}{\overleftarrow{\mathrm{SNR}}_j\mathrm{INR}_{ji}}, \frac{\mathrm{INR}_{ij}}{\overleftarrow{\mathrm{SNR}}_j\mathrm{INR}_{ji}}, 1\right) & \text{if } C_{7,j} \vee C_{8,j} \text{ holds true}\\
0 & \text{otherwise}
\end{cases}, \tag{7.5}$$

$$C_{1,j} : \mathrm{INR}_{ji} < \overrightarrow{\mathrm{SNR}}_j \leqslant \mathrm{INR}_{ij}, \tag{7.6a}$$

$$C_{2,j} : \max\left(\mathrm{INR}_{ij}, \mathrm{INR}_{ji}, \overleftarrow{\mathrm{SNR}}_j\right) < \overrightarrow{\mathrm{SNR}}_j < \mathrm{INR}_{ij}\mathrm{INR}_{ji}, \tag{7.6b}$$

$$C_{3,j} : \overleftarrow{\mathrm{SNR}}_j \leqslant \mathrm{INR}_{ij}, \tag{7.6c}$$

$$C_{4,j} : \mathrm{INR}_{ji} < \mathrm{INR}_{ij} < \overrightarrow{\mathrm{SNR}}_j \leqslant \overleftarrow{\mathrm{SNR}}_j, \tag{7.6d}$$

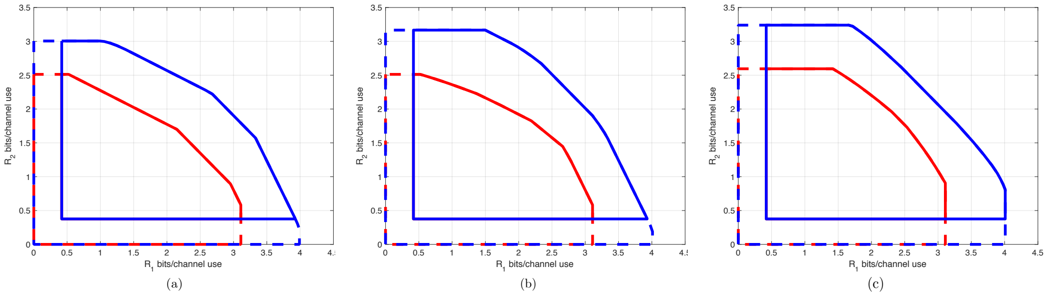


Figure 7.5.: Converse region (blue dashed-line), non-equilibrium region with $\eta \geqslant 1$ (blue solid lines), achievable capacity regions (red dashed-line), and achievable η-NE regions (red solid lines) of the two-user GIC-NOF and two-user D-GIC-NOF with parameters $\overrightarrow{\mathrm{SNR}}_1 = 24$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 18$ dB, $\mathrm{INR}_{12} = 16$ dB, $\mathrm{INR}_{21} = 10$ dB, (a) $\overleftarrow{\mathrm{SNR}}_1 = -100$ dB and $\overleftarrow{\mathrm{SNR}}_2 = -100$ dB, (b) $\overleftarrow{\mathrm{SNR}}_1 = 18$ dB and $\overleftarrow{\mathrm{SNR}}_2 = 12$dB, and (c) $\overleftarrow{\mathrm{SNR}}_1 = 50$ dB and $\overleftarrow{\mathrm{SNR}}_2 = 50$dB.

$$C_{5,j} : \overrightarrow{\mathrm{SNR}}_j > \max\left(\mathrm{INR}_{ij}, \mathrm{INR}_{ji}, \overleftarrow{\mathrm{SNR}}_j\right), \tag{7.6e}$$

$$C_{6,j} : \overrightarrow{\mathrm{SNR}}_j \geqslant \max\left(\mathrm{INR}_{ij}\mathrm{INR}_{ji}, \overleftarrow{\mathrm{SNR}}_j\mathrm{INR}_{ji}\right), \tag{7.6f}$$

$$C_{7,j} : \max\left(\mathrm{INR}_{ij}, \mathrm{INR}_{ji}, \overleftarrow{\mathrm{SNR}}_j, \frac{\overleftarrow{\mathrm{SNR}}_j\mathrm{INR}_{ji}}{\mathrm{INR}_{ij}}\right) < \overrightarrow{\mathrm{SNR}}_j < \overleftarrow{\mathrm{SNR}}_j\mathrm{INR}_{ji} \leqslant \frac{\overleftarrow{\mathrm{SNR}}_j\overrightarrow{\mathrm{SNR}}_j}{\mathrm{INR}_{ij}}, \tag{7.6g}$$

$$C_{8,j} : \max\left(\mathrm{INR}_{ij}, \mathrm{INR}_{ji}, \overleftarrow{\mathrm{SNR}}_j, \frac{\overleftarrow{\mathrm{SNR}}_j\mathrm{INR}_{ji}}{\mathrm{INR}_{ij}}\right) < \overrightarrow{\mathrm{SNR}}_j < \mathrm{INR}_{ij}\mathrm{INR}_{ji} < \overleftarrow{\mathrm{SNR}}_j\mathrm{INR}_{ji}, \tag{7.6h}$$

,$\rho \in [0,1]$, and $\rho_{X_i V_j} \in [0,1)$.

Note that L_i is the rate achieved by the transmitter-receiver pair i when it saturates the power constraint in (2.7) and treats interference as noise. Following this notation, the non-equilibrium region of the two-user GIC-NOF, i.e., $\overline{\mathcal{N}}_\eta$, can be described as follows.

Theorem 15. The non-equilibrium region

Given $\eta \geqslant 1$, the non-equilibrium region $\overline{\mathcal{N}}_\eta$ of the two-user D-GIC-NOF is the closure of all possible non-negative rate pairs $(R_1, R_2) \in \overline{\mathcal{C}} \cap \overline{\mathcal{B}}_\eta$ for all $\rho \in [0,1]$.

Proof: The proof of Theorem 15 is presented in Appendix K. ■

It is worth noting that Theorem 15 has a strong connection with Theorem 10. The relevance of Theorem 15 relies on two important implications: (a) if the pair of configurations (s_1, s_2) is an η-NE, then transmitter-receiver pair 1 and transmitter-receiver pair 2 always achieve a rate equal to or larger than L_1 and L_2, with L_1 and L_2 as in (7.3), respectively; and (b) there always exists an η-NE transmit-receive configuration pair (s_1, s_2) that achieves a rate pair $(R_1(s_1, s_2), R_2(s_1, s_2))$ that is at most η bits per channel use per user away from the outer bound on the converse region.

Figure 7.4 shows an outer bound on the capacity region (Theorem 8) and the non-equilibrium region $\overline{\mathcal{N}}_\eta$ with $\eta \geqslant 1$ in Theorem 15 for a two-user D-GIC-NOF channel with parameters $\overrightarrow{\mathrm{SNR}}_1 = 3$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 8$ dB, $\mathrm{INR}_{12} = 16$ dB, $\mathrm{INR}_{21} = 5$ dB, $\overleftarrow{\mathrm{SNR}}_1 \in \{-100, 9, 50\}$ dB and

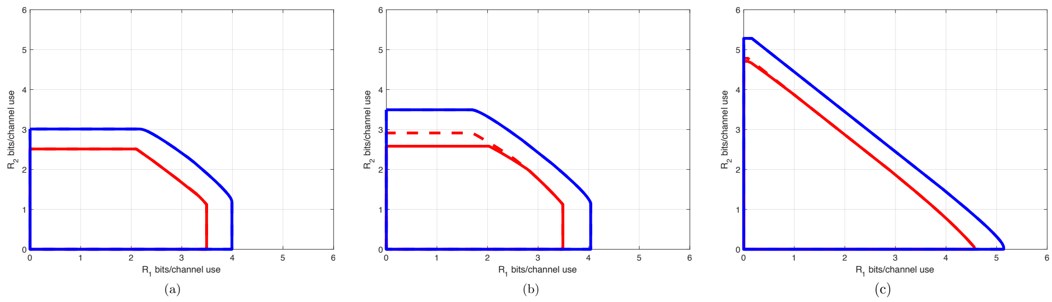


Figure 7.6.: Converse region (blue dashed-line), non-equilibrium region with $\eta \geqslant 1$ (blue solid lines), achievable capacity regions (red dashed-line), and achievable η-NE regions (red solid lines) of the two-user GIC-NOF and two-user D-GIC-NOF with parameters $\overrightarrow{\mathrm{SNR}}_1 = 24$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 18$ dB, $\mathrm{INR}_{12} = 48$ dB, $\mathrm{INR}_{21} = 30$ dB, (a) $\overleftarrow{\mathrm{SNR}}_1 = -100$ dB and $\overleftarrow{\mathrm{SNR}}_2 = -100$ dB, (b) $\overleftarrow{\mathrm{SNR}}_1 = 18$ dB and $\overleftarrow{\mathrm{SNR}}_2 = 12$dB, and (c) $\overleftarrow{\mathrm{SNR}}_1 = 50$ dB and $\overleftarrow{\mathrm{SNR}}_2 = 50$dB.

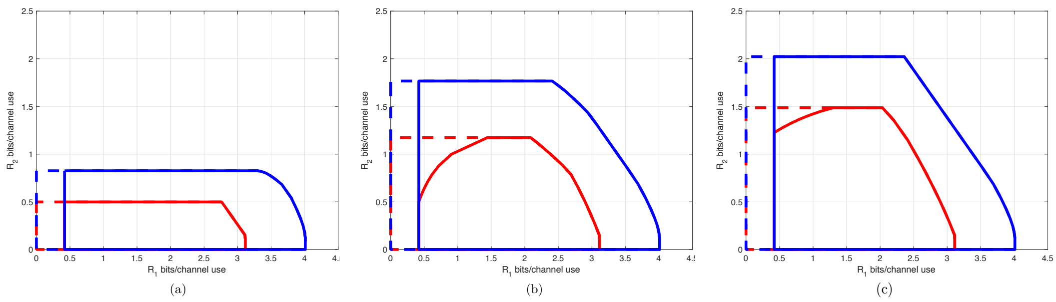


Figure 7.7.: Converse region (blue dashed-line), non-equilibrium region with $\eta \geqslant 1$ (blue solid lines), achievable capacity regions (red dashed-line), and achievable η-NE regions (red solid lines) of the two-user GIC-NOF and two-user D-GIC-NOF with parameters $\overrightarrow{\mathrm{SNR}}_1 = 24$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 3$ dB, $\mathrm{INR}_{12} = 16$ dB, $\mathrm{INR}_{21} = 9$ dB, (a) $\overleftarrow{\mathrm{SNR}}_1 = -100$ dB and $\overleftarrow{\mathrm{SNR}}_2 = -100$ dB, (b) $\overleftarrow{\mathrm{SNR}}_1 = 18$ dB and $\overleftarrow{\mathrm{SNR}}_2 = 8$dB, and (c) $\overleftarrow{\mathrm{SNR}}_1 = 50$ dB and $\overleftarrow{\mathrm{SNR}}_2 = 50$dB.

$\overleftarrow{\mathrm{SNR}}_2 \in \{-100, 6, 50\}$ dB.

Figure 7.5 shows an inner bound (Theorem 7), an outer bound on the capacity region (Theorem 8), the achievable η-NE region $\underline{\mathcal{N}}_\eta$ (Theorem 14), the non-equilibrium region $\overline{\mathcal{N}}_\eta$ with $\eta \geqslant 1$ (Theorem 15) for a two-user D-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, $\overleftarrow{\mathrm{SNR}}_1 \in \{-100, 18, 50\}$ dB and $\overleftarrow{\mathrm{SNR}}_2 \in \{-100, 12, 50\}$ dB.

Figure 7.6 shows an inner bound (Theorem 7), an outer bound on the capacity region (Theorem 8), the achievable η-NE region $\underline{\mathcal{N}}_\eta$ (Theorem 14), the non-equilibrium region $\overline{\mathcal{N}}_\eta$ with $\eta \geqslant 1$ (Theorem 15) for a two-user D-GIC-NOF 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, $\overleftarrow{\mathrm{SNR}}_1 \in \{-100, 18, 50\}$ dB and $\overleftarrow{\mathrm{SNR}}_2 \in \{-100, 12, 50\}$ dB.

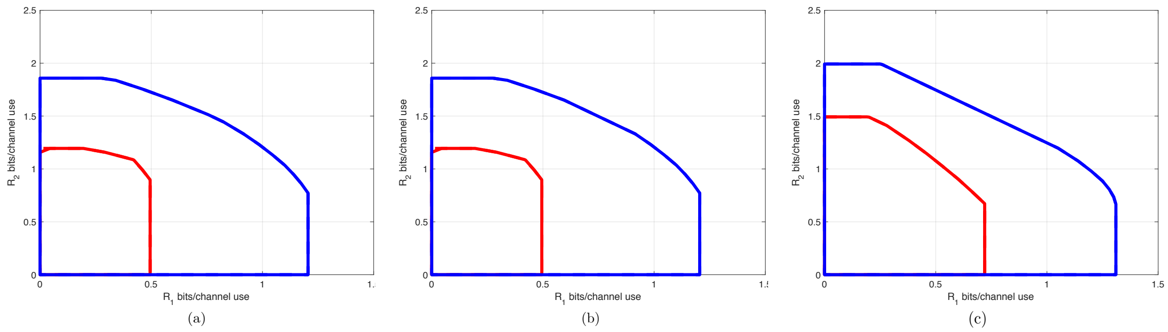


Figure 7.8.: Converse region (blue dashed-line), non-equilibrium region with $\eta \geqslant 1$ (blue solid lines), achievable capacity regions (red dashed-line), and achievable η-NE regions (red solid lines) of the two-user GIC-NOF and two-user D-GIC-NOF with parameters $\overrightarrow{\mathrm{SNR}}_1 = 3$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 8$ dB, $\mathrm{INR}_{12} = 16$ dB, $\mathrm{INR}_{21} = 5$ dB, (a) $\overleftarrow{\mathrm{SNR}}_1 = -100$ dB and $\overleftarrow{\mathrm{SNR}}_2 = -100$ dB, (b) $\overleftarrow{\mathrm{SNR}}_1 = 9$ dB and $\overleftarrow{\mathrm{SNR}}_2 = 6$dB, and (c) $\overleftarrow{\mathrm{SNR}}_1 = 50$ dB and $\overleftarrow{\mathrm{SNR}}_2 = 50$dB.

Figure 7.7 shows an inner bound (Theorem 7), an outer bound on the capacity region (Theorem 8), the achievable η-NE region $\underline{\mathcal{N}}_\eta$ (Theorem 14), the non-equilibrium region $\overline{\mathcal{N}}_\eta$ with $\eta \geqslant 1$ (Theorem 15) for a two-user D-GIC-NOF channel with parameters $\overrightarrow{\mathrm{SNR}}_1 = 24$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 3$ dB, $\mathrm{INR}_{12} = 16$ dB, $\mathrm{INR}_{21} = 9$ dB, $\overleftarrow{\mathrm{SNR}}_1 \in \{-100, 18, 50\}$ dB and $\overleftarrow{\mathrm{SNR}}_2 \in \{-100, 8, 50\}$ dB.

Figure 7.8 shows an inner bound (Theorem 7), an outer bound on the capacity region (Theorem 8), the achievable η-NE region $\underline{\mathcal{N}}_\eta$ (Theorem 14), the non-equilibrium region $\overline{\mathcal{N}}_\eta$ with $\eta \geqslant 1$ (Theorem 15) for a two-user D-GIC-NOF channel with parameters $\overrightarrow{\mathrm{SNR}}_1 = 3$ dB, $\overrightarrow{\mathrm{SNR}}_2 = 8$ dB, $\mathrm{INR}_{12} = 16$ dB, $\mathrm{INR}_{21} = 5$ dB, $\overleftarrow{\mathrm{SNR}}_1 \in \{-100, 9, 50\}$ dB and $\overleftarrow{\mathrm{SNR}}_2 \in \{-100, 6, 50\}$ dB.

Part IV.

CONCLUSIONS

— 8 —

Conclusions

This thesis presented the fundamental limits in the asymptotic regime of the two-user IC with channel-output feedback using tools of information theory and network information theory. More specifically, the focus of this thesis was on the effect of the noise in the feedback links on these fundamental limits under asymmetric conditions on the IC-NOF. The results obtained in this thesis can be seen as a generalization of the results on the two-user IC for the cases without channel-output feedback, with POF, and with NOF under symmetric conditions. To the best of the author's knowledge, this approximation is the most general with respect to existing literature and the one that guarantees the smallest gap between the achievable and converse regions on the GIC when feedback links are subject to Gaussian additive noise. Additionally, the results of this work brought information about the identification of the scenarios in which the use of only one channel-output feedback can bring benefits in terms of rate improvements in a two-user IC despite of the effect of the noise in the feedback link. Particularly, in the case in which one transmitter-receiver pair is in HIR and the other is in LIR, the use of feedback in the transmitter-receiver in HIR does not enlarge the capacity region, even in the case of POF. Additionally, a necessary but no sufficient condition on the GIC for improving the sum-rate is that both transmitter-receiver pairs must be in LIR or in HIR. These improvements were observed and analyzed from the perspective of centralized and decentralized networks.

An achievable region for the two-user LDIC-NOF and an achievable region for the two-user GIC-NOF were obtained using well-known techniques on information theory: rate-splitting, block-Markov superposition coding, and backward decoding. The converse region was the result of using genie-aided models. The genie-aided models and the insights that were used in the Gaussian case were obtained from the analysis of the linear deterministic model. The linear deterministic model is an approximation to the Gaussian case in a very high SNR regime. Therefore, it allowed the analysis of the IC as an interference-limited network focusing more the attention on the interactions of the signals. The achievable and converse regions coincided for the two-user LDIC-NOF. Thus, the capacity region of the two-user LDIC-NOF was characterized. The achievable and converse regions for the two-user GIC-NOF were also characterized and they did not coincide. Nonetheless, It was shown that the capacity region is at most 4.4 bits away from the achievable region, which is a very good approximation given that the capacity region of the two-user GIC is only known in certain specific cases. The capacity regions of the two-user LDIC-NOF and the two-user LDIC were compared to identify the values in the feedback parameters of the two-user LDIC-NOF beyond which the capacity region can be enlarged. This provided the identification of the scenarios in which the capacity region can be enlarged, and more specifically the scenarios in which both individual transmission rates can be improved and not the sum-rate capacity, the scenarios in which the sum-capacity can be improved, and the scenarios in which the use of feedback in one transmitter-receiver pair allows any of the transmitter-receiver pairs to improve the individual transmission rate. The scenarios in which feedback does not enlarge the capacity region were also identified. Given the established connection between the Gaussian and the linear deterministic models, an approximate value in the feedback parameter beyond which the approximate capacity region can be enlarged is also identified. This is a very important result from an engineering point of view, because it establishes the scenarios or the conditions in which to implement channel-output feedback at least in one transmitter-receiver pair of the two-user GIC is useful. These results confirmed the fact that the interference regimes are not the only factor determining the effect of feedback. Indeed, the quality of the feedback links turns out to be another relevant factor.

An achievable η-NE region for the two-user LDIC-NOF and an achievable η-NE region for the two-user GIC-NOF were obtained including common randomness in the coding schemes introduced in the centralized part. This common randomness allowed both transmitter-receiver pairs to limit the rate improvement of each other when either of them deviates from an equilibrium rate pair. A non-equilibrium region was obtained for the two-user LDIC-NOF and a non-equilibrium region was also obtained for the two-user GIC-NOF based on the insights from the analysis in the linear deterministic model. This provided a definition of an η-NE region for the two-user LDIC-NOF with $\eta > 0$ and an approximate η-NE region for the two-user GIC-NOF with $\eta \geqslant 1$. The efficiency of the η-NE region of the two-user LDIC-NOF was characterized using well-known metrics in game theory: price of anarchy (PoA) and price of stability (PoS). These metrics compare the sum-capacity of the two-user LDIC-NOF with the smallest and the best sum-rate at an η-NE region. The PoS is equal to one in all the interference regimes which implies that there always exists an η-NE in the Pareto boundary of the capacity region. It is worth noting here that feedback plays a key role in increasing the PoA in the interference regimes in which feedback can enlarge the sum-rate capacity.

The scenarios, conditions, and values in the feedback parameters beyond which the capacity region of the two-user LDIC-NOF can be enlarged are the same scenarios, conditions, and values to enlarge its η-NE region. In the decentralized case, despite of the anarchical behavior of each transmitter-receiver pair, feedback can be seen as an altruistic technique. The latter is because implementing feedback in one transmitter-receiver pair can enlarge the η-NE region improving the individual rate of the other transmitter-receiver pair.

Future works in this area must consider the cost of feedback. This implies the definition of metrics to analyze if the improvements on the individual rates justify feedback and the additional functionalities that must be implemented overall in the transmitters. This also implies the analysis in case the use of feedback allows the transmitter-receiver pairs to use less transmission power to achieve the same rate pairs as in the case without feedback. The analysis of the effect of feedback in the capacity region when fading is considered into the system model is also an important future work. These studies can be complemented with a real implementation of a basic network using channel-output feedback. Another step to go to a more general model will be the analysis of the two-user GIC in which feedback can also be implemented between each receiver and the non-corresponding transmitter. In the decentralized part, a future work will be the analysis of equilibrium selection methods to reduce the effect of anarchical behavior in a network with channel-output feedback.

Interference is increasing due to the massive use of wireless devices operating in licensed and unlicensed bands. Thus, channel-output feedback might be an effective technique to manage the interference by taking advantage of its structure.

Part V.

APPENDICES

— A —

Achievability Proof of Theorem 1 and Proof of Theorem 7

This appendix describes an achievability scheme for the two-user LDIC-NOF and two-user GIC-NOF based on a three-part message splitting, superposition coding, and backward decoding.

Codebook Generation: Fix a strictly positive joint probability distribution

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

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

Let $R_{1,C1}$, $R_{1,C2}$, $R_{2,C1}$, $R_{2,C2}$, $R_{1,P}$, and $R_{2,P}$ be non-negative real numbers. Define also $R_{1,C} = R_{1,C1} + R_{1,C2}$, $R_{2,C} = R_{2,C1} + R_{2,C2}$, $R_1 = R_{1,C} + R_{1,P}$, and $R_2 = R_{2,C} + R_{2,P}$.

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

$$P_{\boldsymbol{U}}\big(\boldsymbol{u}(s,r)\big) = \prod_{n=1}^{N} P_U(u_n(s,r)), \tag{A.2}$$

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

For encoder 1, generate for each codeword $\boldsymbol{u}(s,r)$, $2^{NR_{1,C1}}$ i.i.d. 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 the conditional distribution

$$P_{\boldsymbol{U}_1|\boldsymbol{U}}\big(\boldsymbol{u}_1(s,r,k)|\boldsymbol{u}(s,r)\big) = \prod_{n=1}^{N} P_{U_1|U}\big(u_{1,n}(s,r,k)|u_n(s,r)\big), \tag{A.3}$$

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

$$P_{\boldsymbol{V}_1|\boldsymbol{U}\boldsymbol{U}_1}\big(\boldsymbol{v}_1(s,r,k,l)|\boldsymbol{u}(s,r), \boldsymbol{u}_1(s,r,k)\big) = \prod_{n=1}^{N} P_{V_1|U\,U_1}\big(v_{1,n}(s,r,k,l)|u_n(s,r), u_{1,n}(s,r,k)\big), \tag{A.4}$$

with $l \in \{1, 2, \ldots, 2^{NR_{1,C2}}\}$. For each triplet 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 the conditional distribution

$$\begin{aligned} & P_{\boldsymbol{X}_{1,P}|\boldsymbol{U}\boldsymbol{U}_1\boldsymbol{V}_1}\big(\boldsymbol{x}_{1,P}(s,r,k,l,q)|\boldsymbol{u}(s,r), \boldsymbol{u}_1(s,r,k), \boldsymbol{v}_1(s,r,k,l)\big) \\ & = \prod_{n=1}^{N} P_{X_{1,P}|U\,U_1\,V_1}\big(x_{1,P,n}(s,r,k,l,q)|u_n(s,r), u_{1,n}(s,r,k), v_{1,n}(s,r,k,l)\big), \end{aligned} \tag{A.5}$$

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

For encoder 2, generate for each codeword $\boldsymbol{u}(s,r)$, $2^{NR_{2,C1}}$ i.i.d. 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 the conditional distribution

$$P_{\boldsymbol{U}_2|\boldsymbol{U}}\big(\boldsymbol{u}_2(s,r,j)|\boldsymbol{u}(s,r)\big) = \prod_{n=1}^{N} P_{U_2|U}\big(u_{2,n}(s,r,j)|u_n(s,r)\big), \tag{A.6}$$

with $j \in \{1, 2, \ldots, 2^{NR_{2,C1}}\}$. For each pair of codewords $\big(\boldsymbol{u}(s,r), \boldsymbol{u}_2(s,r,j)\big)$, generate $2^{NR_{2,C2}}$ i.i.d. N-length 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 the conditional distribution

$$P_{\boldsymbol{V}_2|\boldsymbol{U}\boldsymbol{U}_2}\big(\boldsymbol{v}_2(s,r,j,m)|\boldsymbol{u}(s,r), \boldsymbol{u}_2(s,r,j)\big) = \prod_{n=1}^{N} P_{V_2|U\,U_2}(v_{2,n}(s,r,j,m)|u_n(s,r), u_{2,n}(s,r,j)), \tag{A.7}$$

with $m \in \{1, 2, \ldots, 2^{NR_{2,C2}}\}$. For each triplet 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 $\boldsymbol{x}_{2,P}(s,r,j,m,b) = \big(x_{2,P,1}(s,r,j,m,b),\ x_{2,P,2}(s,r,j,m,b),\ \ldots,\ x_{2,P,N}(s,r,j,m,b)\big)$ according to

$$\begin{aligned} & P_{\boldsymbol{X}_{2,P}|\boldsymbol{U}\boldsymbol{U}_2\boldsymbol{V}_2}\big(\boldsymbol{x}_{2,P}(s,r,j,m,b)|\boldsymbol{u}(s,r), \boldsymbol{u}_2(s,r,j), \boldsymbol{v}_2(s,r,j,m)\big) \\ & = \prod_{n=1}^{N} P_{X_{2,P}|U\,U_2\,V_2}\big(x_{2,P,n}(s,r,j,m,b)|u_n(s,r), u_{2,n}(s,r,j), v_{2,n}(s,r,j,m)\big), \end{aligned} \tag{A.8}$$

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

Encoding: Denote by $W_i^{(t)} \in \mathcal{W}_i = \{1, 2, \ldots, 2^{NR_i}\}$ the message index of transmitter

$i \in \{1,2\}$ during block $t \in \{1,2,\ldots,T\}$, with $T \in \mathbb{N}$ the total number of blocks. Let $W_i^{(t)}$ be composed of the message index $W_{i,C}^{(t)} \in \mathcal{W}_{i,C} = \{1,2,\ldots,2^{NR_{i,C}}\}$ and the message index $W_{i,P}^{(t)} \in \mathcal{W}_{i,P} = \{1,\, 2,\, \ldots, 2^{NR_{i,P}}\}$. 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 also $W_{i,C}^{(t)}$ be composed of the message indices $W_{i,C1}^{(t)} \in \mathcal{W}_{i,C1} = \{1,2,\ldots,2^{NR_{i,C1}}\}$ and $W_{i,C2}^{(t)} \in \mathcal{W}_{i,C2} = \{1,2,\ldots,2^{NR_{i,C2}}\}$. That is, $W_{i,C}^{(t)} = \left(W_{i,C1}^{(t)}, W_{i,C2}^{(t)}\right)$. The message index $W_{i,C1}^{(t)}$ must be reliably decoded by transmitter j, with $j \in \{1,2\} \setminus \{i\}$ (via feedback), and by both receivers. The message index $W_{i,C2}^{(t)}$ must be reliably decoded by both receivers but not by transmitter j.

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

$$\begin{aligned}\boldsymbol{x}_1^{(t)} = \Theta_1 \Bigg(& \boldsymbol{u}\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}\right), \boldsymbol{u}_1\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{1,C1}^{(t)}\right), \boldsymbol{v}_1\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{1,C1}^{(t)}, W_{1,C2}^{(t)}\right), \\ & \boldsymbol{x}_{1,P}\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{1,C1}^{(t)}, W_{1,C2}^{(t)}, W_{1,P}^{(t)}\right)\Bigg), \end{aligned} \tag{A.9}$$

where, $\Theta_1 : (\mathcal{X}_1 \cap \mathcal{X}_2)^N \times \mathcal{X}_1^{3N} \to \mathcal{X}_1^N$ is a function that transforms the codewords $\boldsymbol{u}\Big(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}\Big)$, $\boldsymbol{u}_1\Big(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{1,C1}^{(t)}\Big)$, $\boldsymbol{v}_1\Big(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{1,C1}^{(t)}, W_{1,C2}^{(t)}\Big)$, and $\boldsymbol{x}_{1,P}\Big(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{1,C1}^{(t)}, W_{1,C2}^{(t)}, W_{1,P}^{(t)}\Big)$ into the N-dimensional vector $\boldsymbol{x}_1^{(t)}$ of channel inputs. The indices $W_{1,C1}^{(0)} = W_{1,C1}^{(T)} = s^*$ and $W_{2,C1}^{(0)} = W_{2,C1}^{(T)} = r^*$, and the pair $(s^*, r^*) \in \{1,2,\ldots,2^{N\,R_{1,C1}}\} \times \{1,2,\ldots,2^{NR_{2,C1}}\}$ are pre-defined and known by both receivers and transmitters. It is worth noting that the message index $W_{2,C1}^{(t-1)}$ is obtained by transmitter 1 from the feedback signal $\overleftarrow{\boldsymbol{y}}_1^{(t-1)}$ at the end of the previous encoding step $t-1$.

Transmitter 2 follows a similar encoding scheme.

Decoding: Both receivers decode their message indices at the end of block T in a backward decoding fashion. At each decoding step t, with $t \in \{1,2,\ldots,T\}$, receiver 1 obtains the message indices $\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, \widehat{W}_{1,C2}^{(T-(t-1))}, \widehat{W}_{1,P}^{(T-(t-1))}, \widehat{W}_{2,C2}^{(T-(t-1))}\right) \in \mathcal{W}_{1,C1} \times \mathcal{W}_{2,C1} \times \mathcal{W}_{1,C2} \times \mathcal{W}_{1,P} \times \mathcal{W}_{2,C2}$ from the channel output $\overrightarrow{\boldsymbol{y}}_1^{(T-(t-1))}$. The 5-tuple $\Big(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, \widehat{W}_{1,C2}^{(T-(t-1))}, \widehat{W}_{1,P}^{(T-(t-1))}, \widehat{W}_{2,C2}^{(T-(t-1))}\Big)$ is the unique 5-tuple that satisfies

$$\begin{aligned} \Big(& \boldsymbol{u}\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}\right), \boldsymbol{u}_1\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, W_{1,C1}^{(T-(t-1))}\right), \\ & \boldsymbol{v}_1\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, W_{1,C1}^{(T-(t-1))}, \widehat{W}_{1,C2}^{(T-(t-1))}\right), \\ & \boldsymbol{x}_{1,P}\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, W_{1,C1}^{(T-(t-1))}, \widehat{W}_{1,C2}^{(T-(t-1))}, \widehat{W}_{1,P}^{(T-(t-1))}\right), \\ & \boldsymbol{u}_2\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, W_{2,C1}^{(T-(t-1))}\right), \boldsymbol{v}_2\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, W_{2,C1}^{(T-(t-1))}, \widehat{W}_{2,C2}^{(T-(t-1))}\right), \\ & \overrightarrow{\boldsymbol{y}}_1^{(T-(t-1))}\Big) \in \mathcal{T}^{(N,\epsilon)}_{\left[U\ U_1\ V_1\ X_{1,P}\ U_2\ V_2\ \overrightarrow{Y}_1\right]}, \end{aligned} \tag{A.10}$$

where $W_{1,C1}^{(T-(t-1))}$ and $W_{2,C1}^{(T-(t-1))}$ are assumed to be perfectly decoded in the previous decoding

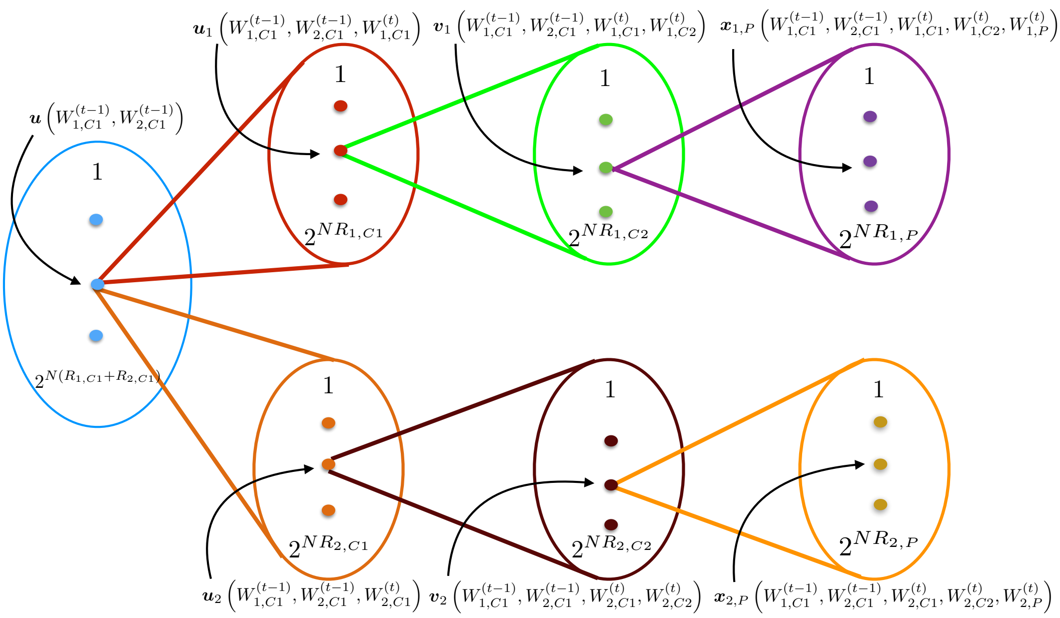


Figure A.1.: Structure of the superposition code. The codewords corresponding to the message indices $W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{i,C1}^{(t)}, W_{i,C2}^{(t)}, W_{i,P}^{(t)}$ with $i \in \{1,2\}$ as well as the block index t are both highlighted. The (approximate) number of codewords for each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

step $t-1$. The set $\mathcal{T}^{(N,\epsilon)}_{\left[U\ U_1\ V_1\ X_{1,P}\ U_2\ V_2\ \overrightarrow{Y}_1\right]}$ 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 $\epsilon > 0$. Receiver 2 follows a similar decoding scheme.

Error Probability Analysis: An error might occur during encoding step t if the message index $W_{2,C1}^{(t-1)}$ is not correctly decoded at transmitter 1 at the end of the step $t-1$. From the AEP [28], it follows that the message index $W_{2,C1}^{(t-1)}$ can be reliably decoded at transmitter 1 during encoding step t, under the condition:

$$\begin{aligned} R_{2,C1} &\leqslant I\left(\overleftarrow{Y}_1; U_2 | U, U_1, V_1, X_1\right) \\ &= I\left(\overleftarrow{Y}_1; U_2 | U, X_1\right). \end{aligned} \tag{A.11}$$

An error might occur during the (backward) decoding step t if the message indices $W_{1,C1}^{(T-t)}$, $W_{2,C1}^{(T-t)}$, $W_{1,C2}^{(T-(t-1))}, W_{1,P}^{(T-(t-1))}$, and $W_{2,C2}^{(T-(t-1))}$ are not decoded correctly given that the message indices $W_{1,C1}^{(T-(t-1))}$ and $W_{2,C1}^{(T-(t-1))}$ were correctly decoded in the previous decoding step $t-1$. These errors might arise for two reasons: (i) there does not exist a 5-tuple $\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, \widehat{W}_{1,C2}^{(T-(t-1))}, \widehat{W}_{1,P}^{(T-(t-1))}, \widehat{W}_{2,C2}^{(T-(t-1))}\right)$ that satisfies (A.10), or (ii) there exist several 5-tuples $\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, \widehat{W}_{1,C2}^{(T-(t-1))}, \widehat{W}_{1,P}^{(T-(t-1))}, \widehat{W}_{2,C2}^{(T-(t-1))}\right)$ that simultaneously satisfy (A.10). From the AEP [28], 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 $E_{(s,r,l,q,m)}^{(t)}$ that describes the case in which the codewords $\boldsymbol{u}(s,r)$, $\boldsymbol{u}_1(s,r,W_{1,C1}^{(T-(t-1))})$, $\boldsymbol{v}_1(s,r,W_{1,C1}^{(T-(t-1))},l)$, $\boldsymbol{x}_{1,P}(s,r,W_{1,C1}^{(T-(t-1))},l,q)$, $\boldsymbol{u}_2(s,r,W_{2,C1}^{(T-(t-1))})$, and $\boldsymbol{v}_2(s,r,W_{2,C1}^{(T-(t-1))},m)$ are jointly typical with $\overrightarrow{\boldsymbol{y}}_1^{(T-(t-1))}$ during decoding step t. Without loss of generality assume that the codeword to be decoded at decoding step t corresponds to the indices $(s,r,l,q,m)=(1,1,1,1,1)$ due to the symmetry of the code. Then, during decoding step t, Boole's inequality yields the following upper-bound on the probability of error due to (ii):

$$\Pr\left[\bigcup_{(s,r,l,q,m)\neq(1,1,1,1,1)} E_{(s,r,l,q,m)}^{(t)}\right] \leqslant \sum_{(s,r,l,q,m)\in\mathcal{T}} \Pr\left[E_{(s,r,l,q,m)}^{(t)}\right], \tag{A.12}$$

with $\mathcal{T}=\left\{\mathcal{W}_{1,C1}\times\mathcal{W}_{2,C1}\times\mathcal{W}_{1,C2}\times\mathcal{W}_{1,P}\times\mathcal{W}_{2,C2}\right\}\setminus\{(1,1,1,1,1)\}$.

From the AEP [28], it follows that

$$\begin{aligned}
P_{e1}^{(t)}(N) \leqslant\ & 2^{N(R_{2,C2}-I(\overrightarrow{Y}_1;V_2|U,U_1,U_2,V_1,X_1)+2\epsilon)} + 2^{N(R_{1,P}-I(\overrightarrow{Y}_1;X_1|U,U_1,U_2,V_1,V_2)+2\epsilon)} \\
&+2^{N(R_{2,C2}+R_{1,P}-I(\overrightarrow{Y}_1;V_2,X_1|U,U_1,U_2,V_1)+2\epsilon)} + 2^{N(R_{1,C2}-I(\overrightarrow{Y}_1;V_1,X_1|U,U_1,U_2,V_2)+2\epsilon)} \\
&+2^{N(R_{1,C2}+R_{2,C2}-I(\overrightarrow{Y}_1;V_1,V_2,X_1|U,U_1,U_2)+2\epsilon)} + 2^{N(R_{1,C2}+R_{1,P}-I(\overrightarrow{Y}_1;V_1,X_1|U,U_1,U_2,V_2)+2\epsilon)} \\
&+2^{N(R_{1,C2}+R_{1,P}+R_{2,C2}-I(\overrightarrow{Y}_1;V_1,V_2,X_1|U,U_1,U_2)+2\epsilon)} + 2^{N(R_{2,C1}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{2,C}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} + 2^{N(R_{2,C1}+R_{1,P}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{2,C}+R_{1,P}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} + 2^{N(R_{2,C1}+R_{1,C2}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{2,C}+R_{1,C2}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} + 2^{N(R_{2,C1}+R_{1,C2}+R_{1,P}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{2,C}+R_{1,C2}+R_{1,P}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} + 2^{N(R_{1,C1}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C1}+R_{2,C2}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} + 2^{N(R_{1,C1}+R_{1,P}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C1}+R_{1,P}+R_{2,C2}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} + 2^{N(R_{1,C}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C}+R_{2,C2}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} + 2^{N(R_1-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_1+R_{2,C2}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} + 2^{N(R_{1,C1}+R_{2,C1}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C1}+R_{2,C}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} + 2^{N(R_{1,C1}+R_{2,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_{2,C}+R_{1,P}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} + 2^{N(R_{1,C}+R_{2,C1}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_{1,C}+R_{2,C}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} + 2^{N(R_1+R_{2,C1}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
&+2^{N(R_1+R_{2,C}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)}.
\end{aligned} \tag{A.13}$$

The same analysis of the probability of error holds for transmitter-receiver pair 2. Hence, in general, from (A.11) and (A.13), reliable decoding holds under the following conditions for transmitter $i\in\{1,2\}$, with $j\in\{1,2\}\setminus\{i\}$:

$$\begin{aligned}
R_{j,C1} \leqslant & \theta_{1,i} \\
& \triangleq I\left(\overleftarrow{Y}_i;U_j|U,U_i,V_i,X_i\right) \\
& = I\left(\overleftarrow{Y}_i;U_j|U,X_i\right),
\end{aligned} \tag{A.14a}$$

$$
\begin{aligned}
R_i + R_{j,C} \leqslant & \theta_{2,i} \\
& \triangleq I(\overrightarrow{Y}_i; U, U_i, U_j, V_i, V_j, X_i) \\
& = I(\overrightarrow{Y}_i; U, U_j, V_j, X_i), && \text{(A.14b)} \\
R_{j,C2} \leqslant & \theta_{3,i} \\
& \triangleq I(\overrightarrow{Y}_i; V_j | U, U_i, U_j, V_i, X_i) \\
& = I(\overrightarrow{Y}_i; V_j | U, U_j, X_i), && \text{(A.14c)} \\
R_{i,P} \leqslant & \theta_{4,i} \\
& \triangleq I(\overrightarrow{Y}_i; X_i | U, U_i, U_j, V_i, V_j), && \text{(A.14d)} \\
R_{i,P} + R_{j,C2} \leqslant & \theta_{5,i} \\
& \triangleq I(\overrightarrow{Y}_i; V_j, X_i | U, U_i, U_j, V_i), && \text{(A.14e)} \\
R_{i,C2} + R_{i,P} \leqslant & \theta_{6,i} \\
& \triangleq I(\overrightarrow{Y}_i; V_i, X_i | U, U_i, U_j, V_j) \\
& = I(\overrightarrow{Y}_i; X_i | U, U_i, U_j, V_j), \text{ and} && \text{(A.14f)} \\
R_{i,C2} + R_{i,P} + R_{j,C2} \leqslant & \theta_{7,i} \\
& \triangleq I(\overrightarrow{Y}_i; V_i, V_j, X_i | U, U_i, U_j) \\
& = I(\overrightarrow{Y}_i; V_j, X_i | U, U_i, U_j). && \text{(A.14g)}
\end{aligned}
$$

Taking into account that $R_i = R_{i,C1} + R_{i,C2} + R_{i,P}$, a Fourier-Motzkin elimination process in (A.14) yields:

$$
\begin{aligned}
R_1 \leqslant & \min\left(\theta_{2,1}, \theta_{6,1} + \theta_{1,2}, \theta_{4,1} + \theta_{1,2} + \theta_{3,2}\right), && \text{(A.15a)} \\
R_2 \leqslant & \min\left(\theta_{2,2}, \theta_{1,1} + a_{6,2}, \theta_{1,1} + \theta_{3,1} + \theta_{4,2}\right), && \text{(A.15b)} \\
R_1 + R_2 \leqslant & \min(\theta_{2,1} + \theta_{4,2}, \theta_{2,1} + a_{6,2}, \theta_{4,1} + \theta_{2,2}, \theta_{6,1} + \theta_{2,2}, \theta_{1,1} + \theta_{3,1} + \theta_{4,1} + \theta_{1,2} + \theta_{5,2}, \\
& \theta_{1,1} + \theta_{7,1} + \theta_{1,2} + \theta_{5,2}, \theta_{1,1} + \theta_{4,1} + \theta_{1,2} + \theta_{7,2}, \theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{3,2} + \theta_{4,2}, \\
& \theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{5,2}, \theta_{1,1} + \theta_{7,1} + \theta_{1,2} + \theta_{4,2}), && \text{(A.15c)} \\
2R_1 + R_2 \leqslant & \min(\theta_{2,1} + \theta_{4,1} + \theta_{1,2} + \theta_{7,2}, \theta_{1,1} + \theta_{4,1} + \theta_{7,1} + 2\theta_{1,2} + \theta_{5,2}, \theta_{2,1} + \theta_{4,1} + \theta_{1,2} + \theta_{5,2}), \\
& && \text{(A.15d)} \\
R_1 + 2R_2 \leqslant & \min(\theta_{1,1} + \theta_{5,1} + \theta_{2,2} + \theta_{4,2}, \theta_{1,1} + \theta_{7,1} + \theta_{2,2} + \theta_{4,2}, 2\theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{4,2} + \theta_{7,2}), \\
& && \text{(A.15e)}
\end{aligned}
$$

where $\theta_{l,i}$ are defined in (A.14) with $(l, i) \in \{1, \ldots, 7\} \times \{1, 2\}$.

A.1. An Achievable Region for the Two-User Linear Deterministic Interference Channel with Noisy Channel-Output Feedback

In the two-user LDIC-NOF, the channel input of transmitter i at each channel use is a q-dimensional binary vector $\boldsymbol{X}_i \in \{0, 1\}^q$ with $i \in \{1, 2\}$ and q as defined in (2.29). Following this observation, the random variables U, U_i, V_i, and $X_{i,P}$ described in (A.1) in the codebook generation are also vectors, and thus, in this subsection, they are denoted by $\boldsymbol{U}$, $\boldsymbol{U}_i$, $\boldsymbol{V}_i$ and $\boldsymbol{X}_{i,P}$, respectively. The random variables $\boldsymbol{U}_i$, $\boldsymbol{V}_i$, and $\boldsymbol{X}_{i,P}$ are assumed to be mutually independent and uniformly distributed over the sets $\{0,1\}^{\left(n_{ji}-\left(\max\left(\overrightarrow{n}_{jj},n_{ji}\right)-\overleftarrow{n}_{jj}\right)^{+}\right)^{+}}$, $\{0,1\}^{\left(\min\left(n_{ji},\left(\max\left(\overrightarrow{n}_{jj},n_{ji}\right)-\overleftarrow{n}_{jj}\right)^{+}\right)\right)}$ and $\{0,1\}^{\left(\overrightarrow{n}_{ii}-n_{ji}\right)^{+}}$, respectively. Note that the random variables $\boldsymbol{U}_i$, $\boldsymbol{V}_i$, and $\boldsymbol{X}_{i,P}$ have the following dimensions:

$$\dim \boldsymbol{U}_i = \left(n_{ji} - \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^{+}\right)^{+}, \tag{A.16a}$$

$$\dim \boldsymbol{V}_i = \min\left(n_{ji}, \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^{+}\right), \text{ and} \tag{A.16b}$$

$$\dim \boldsymbol{X}_{i,P} = \left(\overrightarrow{n}_{ii} - n_{ji}\right)^{+}. \tag{A.16c}$$

These dimensions satisfy the following condition:

$$\dim \boldsymbol{U}_i + \dim \boldsymbol{V}_i + \dim \boldsymbol{X}_{i,P} = \max\left(\overrightarrow{n}_{ii}, n_{ji}\right) \leqslant q. \tag{A.17}$$

Note that the random variable $\boldsymbol{U}$ in (A.1) is not used, and therefore, is a constant. The input symbol of transmitter i during channel use n is $\boldsymbol{X}_i = \left(\boldsymbol{U}_i^{\mathsf{T}}, \boldsymbol{V}_i^{\mathsf{T}}, \boldsymbol{X}_{i,P}^{\mathsf{T}}, (0,\ldots,0)\right)^{\mathsf{T}}$, where $(0,\ldots,0)$ is appended to meet the dimension constraint $\dim \boldsymbol{X}_i = q$. Hence, during block $t \in \{1,2,\ldots,T\}$, the codeword $\boldsymbol{X}_i^{(t)}$ in the two-user LDIC-NOF is a $q \times N$ matrix, *i.e.*, $\boldsymbol{X}_i^{(t)} = \left(\boldsymbol{X}_{i,1}, \boldsymbol{X}_{i,2}, \ldots, \boldsymbol{X}_{i,N}\right) \in \{0,1\}^{q \times N}$.

The intuition behind this choice is based on the following observations: (a) the vector $\boldsymbol{U}_i$ represents the bits in $\boldsymbol{X}_i$ that can be observed by transmitter j via feedback but no necessarily by receiver i; (b) the vector $\boldsymbol{V}_i$ represents the bits in $\boldsymbol{X}_i$ that can be observed by receiver j but no necessarily by receiver i; and finally, (c) the vector $\boldsymbol{X}_{i,P}$ is a notational artefact to denote the bits of $\boldsymbol{X}_i$ that are neither in $\boldsymbol{U}_i$ nor $\boldsymbol{V}_i$. In particular, the bits in $\boldsymbol{X}_{i,P}$ are only observed by receiver i, as shown in Figure A.2. This intuition justifies the dimensions described in (A.16).

Considering this particular code structure, the following holds for the terms $\theta_{l,i}$, with $(l,i) \in \{1,\ldots,7\} \times \{1,2\}$, in (A.14):

$$\begin{aligned}
\theta_{1,i} &= I\left(\overleftarrow{\boldsymbol{Y}}_i; \boldsymbol{U}_j | \boldsymbol{U}, \boldsymbol{X}_i\right) \\
&\overset{(a)}{=} H\left(\overleftarrow{\boldsymbol{Y}}_i | \boldsymbol{U}, \boldsymbol{X}_i\right) \\
&= H\left(\boldsymbol{U}_j\right) \\
&= \left(n_{ij} - \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^{+}\right)^{+},
\end{aligned} \tag{A.18a}$$

$$\begin{aligned}
\theta_{2,i} &= I\left(\overrightarrow{\boldsymbol{Y}}_i; \boldsymbol{U}, \boldsymbol{U}_j, \boldsymbol{V}_j, \boldsymbol{X}_i\right) \\
&\overset{(b)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i\right) \\
&= \max\left(\overrightarrow{n}_{ii}, n_{ij}\right),
\end{aligned} \tag{A.18b}$$

$$\begin{aligned}
\theta_{3,i} &= I\left(\overrightarrow{\boldsymbol{Y}}_i; \boldsymbol{V}_j | \boldsymbol{U}, \boldsymbol{U}_j, \boldsymbol{X}_i\right) \\
&\overset{(b)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i | \boldsymbol{U}, \boldsymbol{U}_j, \boldsymbol{X}_i\right) \\
&= H\left(\boldsymbol{V}_j\right) \\
&= \min\left(n_{ij}, \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^{+}\right),
\end{aligned} \tag{A.18c}$$

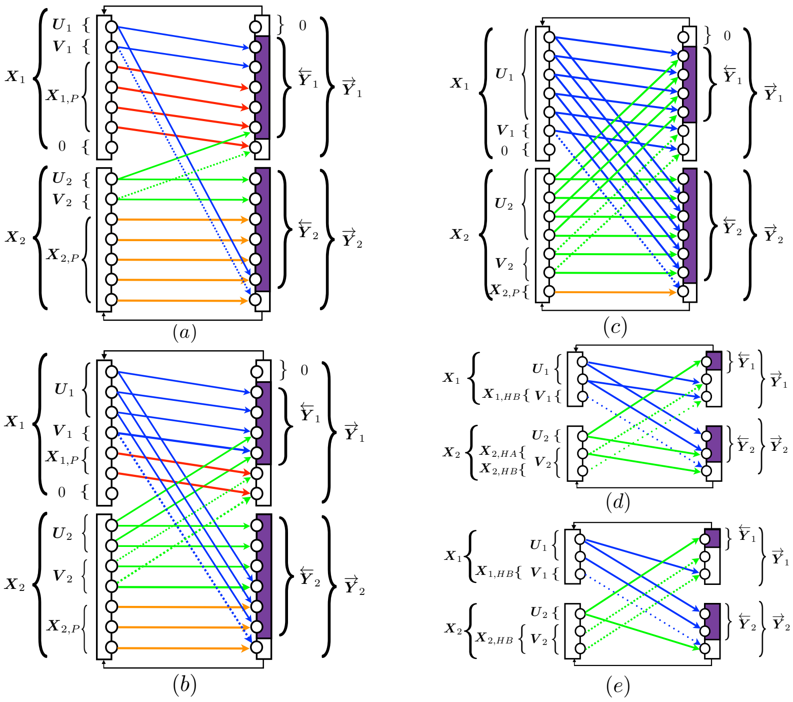


Figure A.2.: The auxiliary random variables and their relation with signals when channel-output feedback is considered in (a) very weak interference regime, (b) weak interference regime, (c) moderate interference regime, (d) strong interference regime and (e) very strong interference regime.

$$
\begin{aligned}
\theta_{4,i} &= I\left(\overrightarrow{\boldsymbol{Y}}_i; \boldsymbol{X}_i | \boldsymbol{U}, \boldsymbol{U}_i, \boldsymbol{U}_j, \boldsymbol{V}_i, \boldsymbol{V}_j\right) \\
&\stackrel{(b)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i | \boldsymbol{U}, \boldsymbol{U}_i, \boldsymbol{U}_j, \boldsymbol{V}_i, \boldsymbol{V}_j\right) \\
&= H\left(\boldsymbol{X}_{i,P}\right) \\
&= \left(\overrightarrow{n}_{ii} - n_{ji}\right)^+, \text{ and}
\end{aligned} \tag{A.18d}
$$

$$
\begin{aligned}
\theta_{5,i} &= I\left(\overrightarrow{\boldsymbol{Y}}_i; \boldsymbol{V}_j, \boldsymbol{X}_i | \boldsymbol{U}, \boldsymbol{U}_i, \boldsymbol{U}_j, \boldsymbol{V}_i\right) \\
&\stackrel{(b)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i | \boldsymbol{U}, \boldsymbol{U}_i, \boldsymbol{U}_j, \boldsymbol{V}_i\right) \\
&= \max\left(\dim \boldsymbol{X}_{i,P}, \dim \boldsymbol{V}_j\right) \\
&= \max\left(\left(\overrightarrow{n}_{ii} - n_{ji}\right)^+, \min\left(n_{ij}, \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^+\right)\right),
\end{aligned} \tag{A.18e}
$$

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

For the calculation of the last two mutual information terms in inequalities (A.14f) and (A.14g), special notation is used. Let the vector $\boldsymbol{V}_i$ be the concatenation of the vectors $\boldsymbol{X}_{i,HA}$ and $\boldsymbol{X}_{i,HB}$, *i.e.*, $\boldsymbol{V}_i = \left(\boldsymbol{X}_{i,HA}, \boldsymbol{X}_{i,HB}\right)$. The vector $\boldsymbol{X}_{i,HA}$ is the part of $\boldsymbol{V}_i$ that is available in both receivers. The vector $\boldsymbol{X}_{i,HB}$ is the part of $\boldsymbol{V}_i$ that is exclusively available in receiver j (see Figure A.2). Note that $H\left(\boldsymbol{V}_i\right) = H\left(\boldsymbol{X}_{i,HA}\right) + H\left(\boldsymbol{X}_{i,HB}\right)$. Note also that the vectors

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

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

Using this notation, the following holds:

$$\begin{aligned}
\theta_{6,i}=&I\left(\overrightarrow{\boldsymbol{Y}}_i;\boldsymbol{X}_i|\boldsymbol{U},\boldsymbol{U}_i,\boldsymbol{U}_j,\boldsymbol{V}_j\right)\\
\overset{(c)}{=}&H\left(\overrightarrow{\boldsymbol{Y}}_i|\boldsymbol{U},\boldsymbol{U}_i,\boldsymbol{U}_j,\boldsymbol{V}_j\right)\\
=&H\left(\boldsymbol{X}_{i,HA},\boldsymbol{X}_{i,P}\right)\\
=&\dim \boldsymbol{X}_{i,HA}+\dim \boldsymbol{X}_{i,P}\\
=&\min\left(n_{ji},\left(\max\left(\overrightarrow{n}_{jj},n_{ji}\right)-\overleftarrow{n}_{jj}\right)^{+}\right)-\min\left(\left(n_{ji}-\overrightarrow{n}_{ii}\right)^{+},\left(\max\left(\overrightarrow{n}_{jj},n_{ji}\right)-\overleftarrow{n}_{jj}\right)^{+}\right)\\
&+\left(\overrightarrow{n}_{ii}-n_{ji}\right)^{+}\ \text{and} \qquad\qquad (A.18f)\\
\theta_{7,i}=&I\left(\overrightarrow{\boldsymbol{Y}}_i;\boldsymbol{V}_j,\boldsymbol{X}_i|\boldsymbol{U},\boldsymbol{U}_i,\boldsymbol{U}_j\right)\\
=&I\left(\overrightarrow{\boldsymbol{Y}}_i;\boldsymbol{X}_i|\boldsymbol{U},\boldsymbol{U}_i,\boldsymbol{U}_j\right)+I\left(\overrightarrow{\boldsymbol{Y}}_i;\boldsymbol{V}_j|\boldsymbol{U},\boldsymbol{U}_i,\boldsymbol{U}_j,\boldsymbol{X}_i\right)\\
=&I\left(\overrightarrow{\boldsymbol{Y}}_i;\boldsymbol{X}_i|\boldsymbol{U},\boldsymbol{U}_i,\boldsymbol{U}_j\right)+I\left(\overrightarrow{\boldsymbol{Y}}_i;\boldsymbol{V}_j|\boldsymbol{U},\boldsymbol{U}_j,\boldsymbol{X}_i\right)\\
\overset{(c)}{=}&H\left(\overrightarrow{\boldsymbol{Y}}_i|\boldsymbol{U},\boldsymbol{U}_i,\boldsymbol{U}_j\right)\\
=&\max\left(H\left(\boldsymbol{V}_j\right),H\left(\boldsymbol{X}_{i,HA}\right)+H\left(\boldsymbol{X}_{i,P}\right)\right)\\
=&\max\left(\dim \boldsymbol{V}_j,\dim \boldsymbol{X}_{i,HA}+\dim \boldsymbol{X}_{i,P}\right)\\
=&\max\Big(\min\left(n_{ij},\left(\max\left(\overrightarrow{n}_{ii},n_{ij}\right)-\overleftarrow{n}_{ii}\right)^{+}\right),\min\left(n_{ji},\left(\max\left(\overrightarrow{n}_{jj},n_{ji}\right)-\overleftarrow{n}_{jj}\right)^{+}\right)\\
&-\min\left(\left(n_{ji}-\overrightarrow{n}_{ii}\right)^{+},\left(\max\left(\overrightarrow{n}_{jj},n_{ji}\right)-\overleftarrow{n}_{jj}\right)^{+}\right)+\left(\overrightarrow{n}_{ii}-n_{ji}\right)^{+}\Big), \qquad (A.18g)
\end{aligned}$$

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

Plugging (A.18) into (A.15) (after some algebraic manipulations) yields the system of inequalities in Theorem 1.

The sum-rate bound in (A.15c) can be simplified as follows:

$$R_1+R_2\leqslant\min(\theta_{2,1}+\theta_{4,2},\theta_{4,1}+\theta_{2,2},\theta_{1,1}+\theta_{5,1}+\theta_{1,2}+\theta_{5,2}). \qquad (A.19)$$

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

This completes the proof of the achievability in Theorem 1.

A.2. An Achievable Region for the Two-User Gaussian Interference Channel with Noisy Channel-Output Feedback

Consider that transmitter i uses the following random variable:

$$X_i=U+U_i+V_i+X_{i,P}, \qquad (A.20)$$

where U, U_1, U_2, V_1, V_2, $X_{1,P}$, and $X_{2,P}$ in (A.20) are mutually independent and distributed as follows:

$$U \sim \mathcal{N}(0, \rho), \tag{A.21a}$$
$$U_i \sim \mathcal{N}(0, \mu_i \lambda_{i,C}), \tag{A.21b}$$
$$V_i \sim \mathcal{N}(0, (1-\mu_i)\lambda_{i,C}), \tag{A.21c}$$
$$X_{i,P} \sim \mathcal{N}(0, \lambda_{i,P}), \tag{A.21d}$$

with

$$\rho + \lambda_{i,C} + \lambda_{i,P} = 1 \text{ and} \tag{A.22a}$$

$$\lambda_{i,P} = \min\left(\frac{1}{\mathrm{INR}_{ji}}, 1\right), \tag{A.22b}$$

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

The parameters ρ, μ_i, and $\lambda_{i,P}$ define a particular coding scheme for transmitter i. The assignment in (A.22b) is based on the intuition obtained from the linear deterministic model, in which the power of the signal $X_{i,P}$ from transmitter i to receiver j must be observed at the noise level. From (2.5), (2.6), and (A.20), the right-hand side of the inequalities in (A.14) can be written in terms of $\overrightarrow{\mathrm{SNR}}_1$, $\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, INR_{21}, $\overleftarrow{\mathrm{SNR}}_1$, $\overleftarrow{\mathrm{SNR}}_2$, ρ, μ_1, and μ_2 as follows:

$$\begin{aligned}\theta_{1,i} &= I\left(\overleftarrow{Y}_i; U_j | U, X_i\right)\\ &= \frac{1}{2}\log\left(\frac{\overleftarrow{\mathrm{SNR}}_i\left(b_{2,i}(\rho)+2\right)+b_{1,i}(1)+1}{\overleftarrow{\mathrm{SNR}}_i\left((1-\mu_j)\,b_{2,i}(\rho)+2\right)+b_{1,i}(1)+1}\right)\\ &= a_{3,i}(\rho,\mu_j),\end{aligned} \tag{A.23a}$$

$$\begin{aligned}\theta_{2,i} &= I\left(\overrightarrow{Y}_i; U, U_j, V_j, X_i\right)\\ &= \frac{1}{2}\log\left(b_{1,i}(\rho)+1\right) - \frac{1}{2}\\ &= a_{2,i}(\rho),\end{aligned} \tag{A.23b}$$

$$\begin{aligned}\theta_{3,i} &= I\left(\overrightarrow{Y}_i; V_j | U, U_j, X_i\right)\\ &= \frac{1}{2}\log\left(\left(1-\mu_j\right)b_{2,i}(\rho)+2\right) - \frac{1}{2}\\ &= a_{4,i}(\rho,\mu_j),\end{aligned} \tag{A.23c}$$

$$\begin{aligned}\theta_{4,i} &= I\left(\overrightarrow{Y}_i; X_i | U, U_i, U_j, V_i, V_j\right)\\ &= \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}} + 2\right) - \frac{1}{2}\\ &= a_{1,i},\end{aligned} \tag{A.23d}$$

$$\begin{aligned}
\theta_{5,i}=&I\left(\overrightarrow{Y}_i;V_j,X_i|U,U_i,U_j,V_i\right)\\
=&\frac{1}{2}\log\left(2+\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}+\Big(1-\mu_j\Big)b_{2,i}(\rho)\right)-\frac{1}{2}\\
=&a_{5,i}(\rho,\mu_j), && \text{(A.23e)}\\
\theta_{6,i}=&I\left(\overrightarrow{Y}_i;X_i|U,U_i,U_j,V_j\right)\\
=&\frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\left(\Big(1-\mu_i\Big)b_{2,j}(\rho)+1\right)+2\right)-\frac{1}{2}\\
=&a_{6,i}(\rho,\mu_i),\text{ and} && \text{(A.23f)}\\
\theta_{7,i}=&I\left(\overrightarrow{Y}_i;V_j,X_i|U,U_i,U_j\right)\\
=&\frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_i}{\mathrm{INR}_{ji}}\left(\Big(1-\mu_i\Big)b_{2,j}(\rho)+1\right)+\Big(1-\mu_j\Big)b_{2,i}(\rho)+2\right)-\frac{1}{2}\\
=&a_{7,i}(\rho,\mu_1,\mu_2). && \text{(A.23g)}
\end{aligned}$$

Finally, plugging (A.23) into (A.15) (after some algebraic manipulations) yields the system of inequalities in Theorem 7. The sum-rate bound in (A.15c) can be simplified as follows:

$$\begin{aligned}
R_1+R_2\leqslant\min\Big(&a_{2,1}(\rho)+a_{1,2},a_{1,1}+a_{2,2}(\rho),a_{3,1}(\rho,\mu_2)+a_{1,1}+a_{3,2}(\rho,\mu_1)+a_{7,2}(\rho,\mu_1,\mu_2),\\
&a_{3,1}(\rho,\mu_2)+a_{5,1}(\rho,\mu_2)+a_{3,2}(\rho,\mu_1)+a_{5,2}(\rho,\mu_1),\\
&a_{3,1}(\rho,\mu_2)+a_{7,1}(\rho,\mu_1,\mu_2)+a_{3,2}(\rho,\mu_1)+a_{1,2}\Big). && \text{(A.24)}
\end{aligned}$$

Note that this follows from the fact that $\max(a_{2,1}(\rho)+a_{1,2},a_{1,1}+a_{2,2}(\rho),a_{3,1}(\rho,\mu_2)+a_{1,1}+a_{3,2}(\rho,\mu_1)+a_{7,2}(\rho,\mu_1,\mu_2),a_{3,1}(\rho,\mu_2)+a_{5,1}(\rho,\mu_2)+a_{3,2}(\rho,\mu_1)+a_{5,2}(\rho,\mu_1),a_{3,1}(\rho,\mu_2)+a_{7,1}(\rho,\mu_1,\mu_2)+a_{3,2}(\rho,\mu_1)+a_{1,2})\leqslant\min(a_{2,1}+a_{6,2}(\rho,\mu_2),a_{6,1}(\rho,\mu_1)+a_{2,2}(\rho),a_{3,1}(\rho,\mu_2)+a_{4,1}(\rho,\mu_2)+a_{1,1}+a_{3,2}(\rho,\mu_1)+a_{5,2}(\rho,\mu_1),a_{3,1}(\rho,\mu_2)+a_{7,1}(\rho,\mu_1,\mu_2)+a_{3,2}(\rho,\mu_1)+a_{5,2}(\rho,\mu_1),a_{3,1}(\rho,\mu_2)+a_{5,1}(\rho,\mu_2)+a_{3,2}(\rho,\mu_1)+\theta_{3,2}+a_{1,2})$. Therefore, the inequalities in (A.15) simplify into (5.3) and this completes the proof of Theorem 7.

— B —

Converse Proof of Theorem 1

This appendix provides a converse proof of Theorem 1. Inequalities (4.1a) and (4.1c) correspond to the minimum cut-set bound [84] and the sum-rate bound for the case of the two-user LDIC with POF. The proofs of these bounds are presented in [88]. The rest of this appendix provides a proof of the inequalities (4.1b), (4.1c) and (4.1d).

Notation. For all $i \in \{1, 2\}$, the channel input $\boldsymbol{X}_{i,n}$ of the two-user LDIC-NOF in (2.31) for any channel use $n \in \{1, 2, \ldots, N\}$ is a q-dimensional binary vector, with q in (2.29), that can be written as the concatenation of four vectors: $\boldsymbol{X}_{i,C,n}$, $\boldsymbol{X}_{i,P,n}$, $\boldsymbol{X}_{i,D,n}$, and $\boldsymbol{X}_{i,Q,n}$, *i.e.*, $\boldsymbol{X}_{i,n} = \left(\boldsymbol{X}_{i,C,n}^{\mathsf{T}}, \boldsymbol{X}_{i,P,n}^{\mathsf{T}}, \boldsymbol{X}_{i,D,n}^{\mathsf{T}}, \boldsymbol{X}_{i,Q,n}^{\mathsf{T}}\right)^{\mathsf{T}}$, as shown in Figure B.1. Note that this notation is independent of the feedback parameters $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$, and it holds for all $n \in \{1, 2, \ldots, N\}$. More specifically,

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

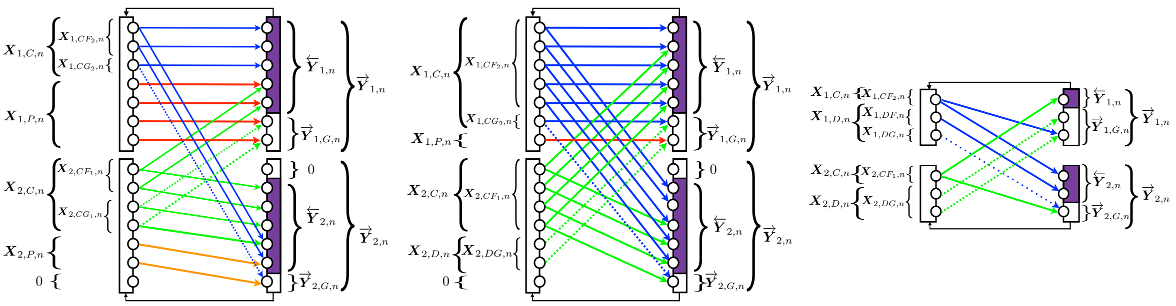
$$\dim \boldsymbol{X}_{i,C,n} = \min\left(\overrightarrow{n}_{ii}, n_{ji}\right); \tag{B.1a}$$

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

$$\dim \boldsymbol{X}_{i,P,n} = (\overrightarrow{n}_{ii} - n_{ji})^{+}; \tag{B.1b}$$

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

$$\dim \boldsymbol{X}_{i,D,n} = (n_{ji} - \overrightarrow{n}_{ii})^{+}; \text{ and} \tag{B.1c}$$



(a) (b) (c)

Figure B.1.: Example of the notation of the channel inputs and the channel outputs when channel-output feedback is considered.

$\boldsymbol{X}_{i,Q,n} = (0,\ldots,0)^{\mathsf{T}}$ is included for dimensional matching of the model in (2.32). Then,

$$\dim \boldsymbol{X}_{i,Q,n} = q - \max\left(\overrightarrow{n}_{ii}, n_{ji}\right). \tag{B.1d}$$

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

$$\begin{aligned} H\left(\boldsymbol{X}_{i,n}\right) &= H\left(\boldsymbol{X}_{i,C,n}, \boldsymbol{X}_{i,P,n}, \boldsymbol{X}_{i,D,n}, \boldsymbol{X}_{i,Q,n}\right) \\ &= H\left(\boldsymbol{X}_{i,C,n}, \boldsymbol{X}_{i,P,n}, \boldsymbol{X}_{i,D,n}\right) \\ &\leqslant \dim \boldsymbol{X}_{i,C,n} + \dim \boldsymbol{X}_{i,P,n} + \dim \boldsymbol{X}_{i,D,n}. \end{aligned} \tag{B.1e}$$

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

$$\begin{aligned} \dim \boldsymbol{X}_{i,n} &= \dim \boldsymbol{X}_{i,C,n} + \dim \boldsymbol{X}_{i,P,n} + \dim \boldsymbol{X}_{i,D,n} + \dim \boldsymbol{X}_{i,Q,n} \\ &= q. \end{aligned} \tag{B.1f}$$

For the case in which feedback is taken into account an alternative notation is adopted. Let $\boldsymbol{X}_{i,D,n}$ be written in terms of $\boldsymbol{X}_{i,DF,n}$ and $\boldsymbol{X}_{i,DG,n}$, *i.e.*, $\boldsymbol{X}_{i,D,n} = \left(\boldsymbol{X}_{i,DF,n}^{\mathsf{T}}, \boldsymbol{X}_{i,DG,n}^{\mathsf{T}}\right)^{\mathsf{T}}$. The vector $\boldsymbol{X}_{i,DF,n}$ represents the bits of $\boldsymbol{X}_{i,D,n}$ that are above the noise level in the feedback link from receiver j to transmitter j; and $\boldsymbol{X}_{i,DG,n}$ represents the bits of $\boldsymbol{X}_{i,D,n}$ that are below the noise level in the feedback link from receiver j to transmitter j, as shown in Figure B.1. The dimension of the vectors $\boldsymbol{X}_{i,DF,n}$ and $\boldsymbol{X}_{i,DG,n}$ are given by

$$\dim \boldsymbol{X}_{i,DF,n} = \min\left(\left(n_{ji} - \overrightarrow{n}_{ii}\right)^+, \left(\overleftarrow{n}_{jj} - \overrightarrow{n}_{ii} - \min\left(\left(\overrightarrow{n}_{jj} - n_{ji}\right)^+, n_{ij}\right) - \left(\left(\overrightarrow{n}_{jj} - n_{ij}\right)^+ - n_{ji}\right)^+\right)^+\right)$$

and (B.2a)

$$\dim \boldsymbol{X}_{i,DG,n} = \dim \boldsymbol{X}_{i,D,n} - \dim \boldsymbol{X}_{i,DF,n}. \tag{B.2b}$$

Let $\boldsymbol{X}_{i,C,n}$ be written in terms of $\boldsymbol{X}_{i,CF_j,n}$ and $\boldsymbol{X}_{i,CG_j,n}$, *i.e.*, $\boldsymbol{X}_{i,C,n} = \left(\boldsymbol{X}_{i,CF_j,n}^{\mathsf{T}}, \boldsymbol{X}_{i,CG_j,n}^{\mathsf{T}}\right)^{\mathsf{T}}$. The vector $\boldsymbol{X}_{i,CF_j,n}$ represents the bits of $\boldsymbol{X}_{i,C,n}$ that are above the noise level in the feedback link from receiver j to transmitter j; and $\boldsymbol{X}_{i,CG_j,n}$ represents the bits of $\boldsymbol{X}_{i,C,n}$ that are below the noise level in the feedback link from receiver j to transmitter j, as shown in Figure B.1.

Let also, the dimension of the vector $\left(\boldsymbol{X}_{i,CF_j,n}^{\mathsf{T}}, \boldsymbol{X}_{i,DF,n}^{\mathsf{T}}\right)$ be defined as follows:

$$\dim\left(\left(\boldsymbol{X}_{i,CF_j,n}^{\mathsf{T}}, \boldsymbol{X}_{i,DF,n}^{\mathsf{T}}\right)\right)=\left(\min\left(\overleftarrow{n}_{jj}, \max\left(\overrightarrow{n}_{jj}, n_{ji}\right)\right) - \left(\overrightarrow{n}_{jj} - n_{ji}\right)^{+}\right)^{+}. \quad \text{(B.3)}$$

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

$$\dim \boldsymbol{X}_{i,CF_j,n}=\dim\left(\left(\boldsymbol{X}_{i,CF_j,n}^{\mathsf{T}}, \boldsymbol{X}_{i,DF,n}^{\mathsf{T}}\right)\right) - \dim \boldsymbol{X}_{i,DF,n} \quad \text{and} \quad \text{(B.4a)}$$
$$\dim \boldsymbol{X}_{i,CG_j,n}=\dim \boldsymbol{X}_{i,C,n} - \dim \boldsymbol{X}_{i,CF_j,n}. \quad \text{(B.4b)}$$

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

The vector $\boldsymbol{X}_{i,U,n}$ is used to represent the bits of the vector $\boldsymbol{X}_{i,n}$ that interfere with bits of $\boldsymbol{X}_{j,C,n}$ at receiver j and the bits of $\boldsymbol{X}_{i,n}$ that are observed by receiver j and do not interfere any bits from transmitter j. An alternative definition of the vector $\boldsymbol{X}_{i,U,n}$ is the following: the bits of the vector $\boldsymbol{X}_{i,n}$ that are observed by receiver j and do not interfere with any bit corresponding to the vector $\boldsymbol{X}_{j,P,n}$. An example is shown in Figure B.2.

Therefore, the dimension of the vector $\boldsymbol{X}_{i,U,n}$ is

$$\dim \boldsymbol{X}_{i,U,n}=\min\left(\overrightarrow{n}_{jj}, n_{ij}\right) - \min\left(\left(\overrightarrow{n}_{jj} - n_{ji}\right)^{+}, n_{ij}\right) + \left(n_{ji} - \overrightarrow{n}_{jj}\right)^{+}. \quad \text{(B.5)}$$

Finally, for all $i \in \{1,2\}$, with $j \in \{1,2\} \setminus \{i\}$, the channel output $\overrightarrow{\boldsymbol{Y}}_{i,n}$ of the two-user LDIC-NOF in (2.31) for any channel use $n \in \{1,2,\ldots,N\}$ is a q-dimensional binary vector, with q in (2.29), that can be written as the concatenation of three vectors: $\overrightarrow{\boldsymbol{Y}}_{i,Q,n}$, $\overleftarrow{\boldsymbol{Y}}_{i,n}$, and $\overrightarrow{\boldsymbol{Y}}_{i,G,n}$, *i.e.*, $\overrightarrow{\boldsymbol{Y}}_{i,n} = \left(\overrightarrow{\boldsymbol{Y}}_{i,Q,n}^{\mathsf{T}}, \overleftarrow{\boldsymbol{Y}}_{i,n}^{\mathsf{T}}, \overrightarrow{\boldsymbol{Y}}_{i,G,n}^{\mathsf{T}}\right)^{\mathsf{T}}$, as shown in Figure B.1. More specifically, the vector $\overleftarrow{\boldsymbol{Y}}_{i,n}$ contains the bits that are above the noise level in the feedback link from receiver i to transmitter i. Then,

$$\dim \overleftarrow{\boldsymbol{Y}}_{i,n}=\min\left(\overleftarrow{n}_{ii}, \max\left(\overrightarrow{n}_{ii}, n_{ij}\right)\right). \quad \text{(B.6a)}$$

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

$$\dim \overrightarrow{\boldsymbol{Y}}_{i,G,n}=\left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^{+}. \quad \text{(B.6b)}$$

The vector $\overrightarrow{\boldsymbol{Y}}_{i,Q,n} = (0,\ldots,0)$ is included for dimensional matching with the model in (2.32). Then,

$$\begin{aligned} H\left(\overrightarrow{\boldsymbol{Y}}_{i,n}\right)&=H\left(\overrightarrow{\boldsymbol{Y}}_{i,Q,n}, \overleftarrow{\boldsymbol{Y}}_{i,n}, \overrightarrow{\boldsymbol{Y}}_{i,G,n}\right)\\ &=H\left(\overleftarrow{\boldsymbol{Y}}_{i,n}, \overrightarrow{\boldsymbol{Y}}_{i,G,n}\right)\\ &\leqslant \dim \overleftarrow{\boldsymbol{Y}}_{i,n} + \dim \overrightarrow{\boldsymbol{Y}}_{i,G,n}. \end{aligned} \quad \text{(B.6c)}$$

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

Using this notation, the proof continues as follows.

Proof of (4.1b)**:** First, consider $n_{ji} \leqslant \overrightarrow{n}_{ii}$, *i.e.*, the vector $\boldsymbol{X}_{i,P,n}$ exists and the vector $\boldsymbol{X}_{i,D,n}$ does not exist. From the assumption that the message index W_i is i.i.d. following a uniform distribution over the set $\mathcal{W}_i$, the following holds for any $k \in \{1, 2, \ldots, N\}$:

$$
\begin{aligned}
NR_i &= H\left(W_i\right) \\
&\stackrel{(a)}{=} H\left(W_i | W_j\right) \\
&\stackrel{(b)}{\leqslant} I\left(W_i; \overrightarrow{\boldsymbol{Y}}_i, \overleftarrow{\boldsymbol{Y}}_j | W_j\right) + N\delta(N) \\
&= H\left(\overrightarrow{\boldsymbol{Y}}_i, \overleftarrow{\boldsymbol{Y}}_j | W_j\right) + N\delta(N) \\
&\stackrel{(c)}{=} \sum_{n=1}^{N} H\left(\overrightarrow{\boldsymbol{Y}}_{i,n}, \overleftarrow{\boldsymbol{Y}}_{j,n} | W_j, \overrightarrow{\boldsymbol{Y}}_{i,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}, \boldsymbol{X}_{j,n}\right) + N\delta(N) \\
&\leqslant \sum_{n=1}^{N} H\left(\boldsymbol{X}_{i,n}, \overleftarrow{\boldsymbol{Y}}_{j,n} | \boldsymbol{X}_{j,n}\right) + N\delta(N) \\
&\leqslant \sum_{n=1}^{N} H\left(\boldsymbol{X}_{i,n}\right) + N\delta(N) \\
&= NH\left(\boldsymbol{X}_{i,k}\right) + N\delta(N) \\
&\leqslant N\left(\dim \boldsymbol{X}_{i,C,k} + \dim \boldsymbol{X}_{i,P,k}\right) + N\delta(N),
\end{aligned} \tag{B.7}
$$

where, (a) follows from the fact that W_1 and W_2 are mutually independent; (b) follows from Fano's inequality with $\delta : \mathbb{N} \rightarrow \mathbb{R}_+$ a positive monotonically decreasing function (Lemma 58); and (c) follows from the fact that $\boldsymbol{X}_{j,n} = f_j^{(n)}\left(W_j, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}\right)$ with $f_j^{(n)}$ a deterministic injective function.

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

$$
\begin{aligned}
NR_i &= H\left(W_i\right) \\
&\stackrel{(a)}{=} H\left(W_i | W_j\right) \\
&\stackrel{(b)}{\leqslant} I\left(W_i; \overrightarrow{\boldsymbol{Y}}_i, \overleftarrow{\boldsymbol{Y}}_j | W_j\right) + N\delta(N) \\
&= H\left(\overrightarrow{\boldsymbol{Y}}_i, \overleftarrow{\boldsymbol{Y}}_j | W_j\right) + N\delta(N) \\
&\stackrel{(c)}{=} \sum_{n=1}^{N} H\left(\overrightarrow{\boldsymbol{Y}}_{i,n}, \overleftarrow{\boldsymbol{Y}}_{j,n} | W_j, \overrightarrow{\boldsymbol{Y}}_{i,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}, \boldsymbol{X}_{j,n}\right) + N\delta(N) \\
&\leqslant \sum_{n=1}^{N} H\left(\boldsymbol{X}_{i,C,n}, \boldsymbol{X}_{i,CF_j,n}, \boldsymbol{X}_{i,DF,n}\right) + N\delta(N) \\
&= \sum_{n=1}^{N} H\left(\boldsymbol{X}_{i,C,n}, \boldsymbol{X}_{i,DF,n}\right) + N\delta(N) \\
&= NH\left(\boldsymbol{X}_{i,C,k}, \boldsymbol{X}_{i,DF,k}\right) + N\delta(N) \\
&\leqslant N\left(\dim \boldsymbol{X}_{i,C,k} + \dim \boldsymbol{X}_{i,DF,k}\right) + N\delta(N).
\end{aligned} \tag{B.8}
$$

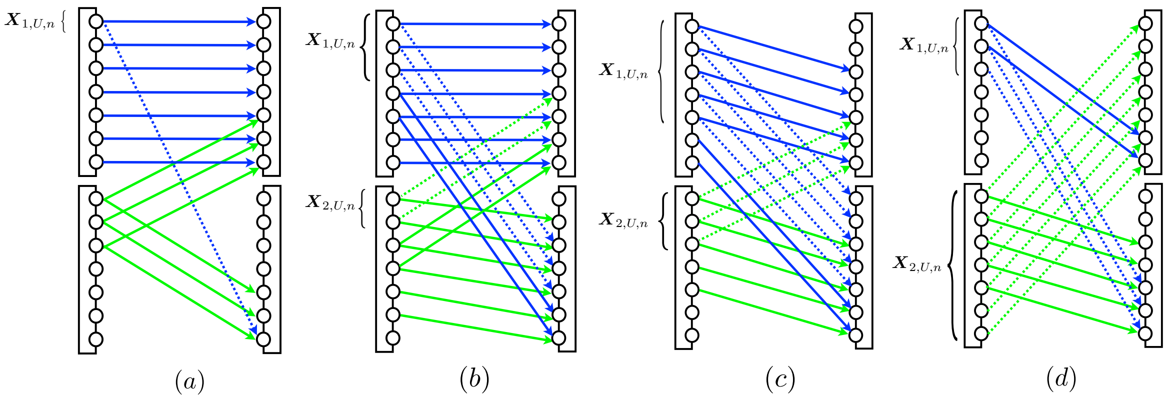


Figure B.2.: $\boldsymbol{X}_{i,U,n}$ in different combinations of interference regimes.

Then, (B.7) and (B.8) can be expressed as one inequality in the asymptotic block-length regime as follows:

$$R_i \leqslant \dim \boldsymbol{X}_{i,C,k} + \dim \boldsymbol{X}_{i,P,k} + \dim \boldsymbol{X}_{i,DF,k}, \tag{B.9}$$

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

Plugging (B.1a), (B.1b), and (B.2a) into (B.9), and after some algebraic manipulations the following holds:

$$R_i \leqslant \min\Big(\max(\overrightarrow{n}_{ii}, n_{ji}), \max\left(\overrightarrow{n}_{ii}, \overleftarrow{n}_{jj} - (\overrightarrow{n}_{jj} - n_{ji})^+\right)\Big). \tag{B.10}$$

This completes the proof of (4.1b).

Proof of (4.1c)**:** From the assumption that the message indices W_1 and W_2 are i.i.d. following a uniform distribution over the sets $\mathcal{W}_1$ and $\mathcal{W}_2$ respectively, the following holds for any $k \in \{1, 2, \ldots, N\}$:

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

$$
\begin{aligned}
&\overset{(b)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_1|\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U}\right) - H\left(\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U}|\overrightarrow{\boldsymbol{Y}}_1\right) + H\left(\overrightarrow{\boldsymbol{Y}}_2|\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U}\right)\\
&\quad -H\left(\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U}|\overrightarrow{\boldsymbol{Y}}_2\right) + I\left(\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U};W_1,\overleftarrow{\boldsymbol{Y}}_1\right) + I\left(\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U};W_2,\overleftarrow{\boldsymbol{Y}}_2\right)\\
&\quad -H\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right) - H\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right) + N\delta(N)\\
&\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_1|\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_2|\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U}\right) + I\left(\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U};W_1,\overleftarrow{\boldsymbol{Y}}_1\right)\\
&\quad +I\left(\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U};W_2,\overleftarrow{\boldsymbol{Y}}_2\right) - H\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right) - H\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right) + N\delta(N)\\
&\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_1|\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_2|\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U}\right) + I\left(\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U},W_2,\overleftarrow{\boldsymbol{Y}}_2;W_1,\overleftarrow{\boldsymbol{Y}}_1\right)\\
&\quad +I\left(\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U},W_1,\overleftarrow{\boldsymbol{Y}}_1;W_2,\overleftarrow{\boldsymbol{Y}}_2\right) - H\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right) - H\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right) + N\delta(N)\\
&= H\left(\overrightarrow{\boldsymbol{Y}}_1|\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_2|\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U}\right) + I\left(W_2;W_1,\overleftarrow{\boldsymbol{Y}}_1\right)\\
&\quad +I\left(\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U},\overleftarrow{\boldsymbol{Y}}_2;W_1,\overleftarrow{\boldsymbol{Y}}_1|W_2\right) + I\left(W_1;W_2,\overleftarrow{\boldsymbol{Y}}_2\right) + I\left(\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U},\overleftarrow{\boldsymbol{Y}}_1;W_2,\overleftarrow{\boldsymbol{Y}}_2|W_1\right)\\
&\quad -H\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right) - H\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right) + N\delta(N)\\
&\overset{(c)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_1|\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_2|\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U}\right) + H\left(W_1\right) + H\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right) - H\left(W_1|W_2\right)\\
&\quad -H\left(\overleftarrow{\boldsymbol{Y}}_1|W_2,W_1\right) + H\left(\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U},\overleftarrow{\boldsymbol{Y}}_2|W_2\right) + H\left(W_2\right) + H\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right) - H\left(W_2|W_1\right)\\
&\quad -H\left(\overleftarrow{\boldsymbol{Y}}_2|W_1,W_2\right) + H\left(\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U},\overleftarrow{\boldsymbol{Y}}_1|W_1\right) - H\left(\overleftarrow{\boldsymbol{Y}}_1|W_1\right) - H\left(\overleftarrow{\boldsymbol{Y}}_2|W_2\right) + N\delta(N)\\
&\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_1|\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_2|\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U}\right) + H\left(\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U},\overleftarrow{\boldsymbol{Y}}_2|W_2\right)\\
&\quad +H\left(\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U},\overleftarrow{\boldsymbol{Y}}_1|W_1\right) + N\delta(N)\\
&= \sum_{n=1}^{N}\Big[H\left(\overrightarrow{\boldsymbol{Y}}_{1,n}|\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U},\overrightarrow{\boldsymbol{Y}}_{1,(1:n-1)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_{2,n}|\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U},\overrightarrow{\boldsymbol{Y}}_{2,(1:n-1)}\right)\\
&\quad +H\Big(\boldsymbol{X}_{2,C,n},\boldsymbol{X}_{1,U,n},\overleftarrow{\boldsymbol{Y}}_{2,n}|W_2,\boldsymbol{X}_{2,C,(1:n-1)},\boldsymbol{X}_{1,U,(1:n-1)},\overleftarrow{\boldsymbol{Y}}_{2,(1:n-1)}\Big)\\
&\quad +H\Big(\boldsymbol{X}_{1,C,n},\boldsymbol{X}_{2,U,n},\overleftarrow{\boldsymbol{Y}}_{1,n}|W_1,\boldsymbol{X}_{1,C,(1:n-1)},\boldsymbol{X}_{2,U,(1:n-1)},\overleftarrow{\boldsymbol{Y}}_{1,(1:n-1)}\Big)\Big] + N\delta(N)\\
&\overset{(d)}{=} \sum_{n=1}^{N}\Big[H\left(\overrightarrow{\boldsymbol{Y}}_{1,n}|\boldsymbol{X}_{1,C},\boldsymbol{X}_{2,U},\overrightarrow{\boldsymbol{Y}}_{1,(1:n-1)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_{2,n}|\boldsymbol{X}_{2,C},\boldsymbol{X}_{1,U},\overrightarrow{\boldsymbol{Y}}_{2,(1:n-1)}\right)\\
&\quad +H\Big(\boldsymbol{X}_{2,C,n},\boldsymbol{X}_{1,U,n},\overleftarrow{\boldsymbol{Y}}_{2,n}|W_2,\boldsymbol{X}_{2,C,(1:n-1)},\boldsymbol{X}_{1,U,(1:n-1)},\overleftarrow{\boldsymbol{Y}}_{2,(1:n-1)},\boldsymbol{X}_{2,n}\Big)\\
&\quad +H\Big(\boldsymbol{X}_{1,C,n},\boldsymbol{X}_{2,U,n},\overleftarrow{\boldsymbol{Y}}_{1,n}|W_1,\boldsymbol{X}_{1,C,(1:n-1)},\boldsymbol{X}_{2,U,(1:n-1)},\overleftarrow{\boldsymbol{Y}}_{1,(1:n-1)},\boldsymbol{X}_{1,n}\Big)\Big] + N\delta(N)\\
&\overset{(e)}{\leqslant} \sum_{n=1}^{N}\Big[H\left(\overrightarrow{\boldsymbol{Y}}_{1,n}|\boldsymbol{X}_{1,C,n},\boldsymbol{X}_{2,U,n}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_{2,n}|\boldsymbol{X}_{2,C,n},\boldsymbol{X}_{1,U,n}\right) + H\left(\boldsymbol{X}_{1,U,n},\overleftarrow{\boldsymbol{Y}}_{2,n}|\boldsymbol{X}_{2,n}\right)\\
&\quad +H\left(\boldsymbol{X}_{2,U,n},\overleftarrow{\boldsymbol{Y}}_{1,n}|\boldsymbol{X}_{1,n}\right)\Big] + N\delta(N)\\
&\leqslant \sum_{n=1}^{N}\Big[H(\boldsymbol{X}_{1,P,n}) + H(\boldsymbol{X}_{2,P,n}) + H\Big(\boldsymbol{X}_{1,U,n},\overleftarrow{\boldsymbol{Y}}_{2,n}|\boldsymbol{X}_{2,n}\Big) + H\left(\boldsymbol{X}_{2,U,n},\overleftarrow{\boldsymbol{Y}}_{1,n}|\boldsymbol{X}_{1,n}\right)\Big] + N\delta(N)\\
&\overset{(e)}{\leqslant} N\Big[H\left(\boldsymbol{X}_{1,P,k}\right) + H\left(\boldsymbol{X}_{2,P,k}\right) + H\left(\boldsymbol{X}_{1,U,k}\right) + H\left(\overleftarrow{\boldsymbol{Y}}_{2,k}|\boldsymbol{X}_{2,k},\boldsymbol{X}_{1,U,k}\right) + H\left(\boldsymbol{X}_{2,U,k}\right)\\
&\quad +H\left(\overleftarrow{\boldsymbol{Y}}_{1,k}|\boldsymbol{X}_{1,k},\boldsymbol{X}_{2,U,k}\right)\Big] + N\delta(N)\\
&= N\Big[H\left(\boldsymbol{X}_{1,P,k}\right) + H\left(\boldsymbol{X}_{2,P,k}\right) + H\left(\boldsymbol{X}_{1,U,k}\right) + H\left(\boldsymbol{X}_{1,CF_2,k},\boldsymbol{X}_{1,DF,k}|\boldsymbol{X}_{2,k},\boldsymbol{X}_{1,U,k}\right)\\
&\quad +H\left(\boldsymbol{X}_{2,U,k}\right) + H\left(\boldsymbol{X}_{2,CF_1,k},\boldsymbol{X}_{2,DF,k}|\boldsymbol{X}_{1,k},\boldsymbol{X}_{2,U,k}\right)\Big] + N\delta(N)
\end{aligned}
$$

$$\begin{aligned}
&\leqslant N\Big[H\left(\boldsymbol{X}_{1,P,k}\right)+H\left(\boldsymbol{X}_{2,P,k}\right)+H\left(\boldsymbol{X}_{1,U,k}\right)+H\left(\boldsymbol{X}_{1,CF_2,k},\boldsymbol{X}_{1,DF,k}|\boldsymbol{X}_{1,U,k}\right)+H\left(\boldsymbol{X}_{2,U,k}\right)\\
&\quad+H\left(\boldsymbol{X}_{2,CF_1,k},\boldsymbol{X}_{2,DF,k}|\boldsymbol{X}_{2,U,k}\right)\Big]+N\delta(N)\\
&\leqslant N\Big[\dim\boldsymbol{X}_{1,P,k}+\dim\boldsymbol{X}_{2,P,k}+\dim\boldsymbol{X}_{1,U,k}+\Big(\dim\left(\boldsymbol{X}_{1,CF_2,k},\boldsymbol{X}_{1,DF,k}\right)-\dim\boldsymbol{X}_{1,U,k}\Big)^+\\
&\quad+\dim\boldsymbol{X}_{2,U,k}+\Big(\dim\left(\boldsymbol{X}_{2,CF_1,k},\boldsymbol{X}_{2,DF,k}\right)-\dim\boldsymbol{X}_{2,U,k}\Big)^+\Big]+N\delta(N). \qquad \text{(B.11)}
\end{aligned}$$

where, (a) follows from Fano's inequality with $\delta:\mathbb{N}\to\mathbb{R}_+$ a positive monotonically decreasing function (Lemma 58); (b) follows from the fact that $H(Y)-H(X)=H(Y|X)-H(X|Y)$; (c) follows from the fact that $H\Big(\boldsymbol{X}_{i,C},\,\boldsymbol{X}_{j,U},\,\overleftarrow{\boldsymbol{Y}}_i|W_i,\,W_j,\,\overleftarrow{\boldsymbol{Y}}_j\Big)=0$; (d) follows from the fact that $\boldsymbol{X}_{i,n}=f_i^{(n)}\left(W_i,\overleftarrow{\boldsymbol{Y}}_{i,(1:n-1)}\right)$ from the definition of the encoding function in (2.1) and $\left(W_i,\overleftarrow{\boldsymbol{Y}}_{i,(1:n-1)}\right)\to X_{i,n}\to\overrightarrow{Y}_{i,n}$; and (e) follows from the fact that conditioning does not increase entropy (Lemma 40).

Plugging (B.1b), (B.3), and (B.5) into (B.11) and after some algebraic manipulations, the following holds in the asymptotic block-length regime:

$$\begin{aligned}
R_1+R_2\leqslant&\max\Big(\left(\overrightarrow{n}_{11}-n_{12}\right)^+,n_{21},\overrightarrow{n}_{11}-\left(\max\left(\overrightarrow{n}_{11},n_{12}\right)-\overleftarrow{n}_{11}\right)^+\Big)\\
&+\max\Big(\left(\overrightarrow{n}_{22}-n_{21}\right)^+,n_{12},\overrightarrow{n}_{22}-\left(\max\left(\overrightarrow{n}_{22},n_{21}\right)-\overleftarrow{n}_{22}\right)^+\Big). \qquad \text{(B.12)}
\end{aligned}$$

This completes the proof of (4.1c).

Proof of (4.1d)**:** From the assumption that the message indices W_i and W_j are i.i.d. following a uniform distribution over the sets $\mathcal{W}_i$ and $\mathcal{W}_j$ respectively, for all $i\in\{1,2\}$ and $j\in\{1,2\}\setminus\{i\}$, the following holds for any $k\in\{1,2,\ldots,N\}$:

$$\begin{aligned}
N\Big(2R_i+R_j\Big)&=2H\left(W_i\right)+H\left(W_j\right)\\
&\overset{(a)}{\leqslant}I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i,\overleftarrow{\boldsymbol{Y}}_i\right)+I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i,\overleftarrow{\boldsymbol{Y}}_j|W_j\right)+I\left(W_j;\overrightarrow{\boldsymbol{Y}}_j,\overleftarrow{\boldsymbol{Y}}_j\right)+N\delta(N)\\
&\overset{(b)}{=}H\left(\overrightarrow{\boldsymbol{Y}}_i\right)-H\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right)-H\left(\overrightarrow{\boldsymbol{Y}}_i|W_i,\overleftarrow{\boldsymbol{Y}}_i\right)+H\left(\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j\right)+H\left(\overrightarrow{\boldsymbol{Y}}_j\right)\\
&\quad-H\left(\overrightarrow{\boldsymbol{Y}}_j|W_j,\overleftarrow{\boldsymbol{Y}}_j\right)+N\delta(N)\\
&=H\left(\overrightarrow{\boldsymbol{Y}}_i\right)-H\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right)-H\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{j,D}|W_i,\overleftarrow{\boldsymbol{Y}}_i\right)+H\left(\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j\right)+H\left(\overrightarrow{\boldsymbol{Y}}_j\right)\\
&\quad-H\left(\boldsymbol{X}_{i,C},\boldsymbol{X}_{i,D}|W_j,\overleftarrow{\boldsymbol{Y}}_j\right)+N\delta(N)\\
&\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_i\right)-H\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right)-H\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}|W_i,\overleftarrow{\boldsymbol{Y}}_i\right)\\
&\quad+H\left(\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j\right)+H\left(\overrightarrow{\boldsymbol{Y}}_j\right)-H\left(\boldsymbol{X}_{i,C}|W_j,\overleftarrow{\boldsymbol{Y}}_j\right)+N\delta(N)\\
&\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_i\right)-H\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right)+\left[I\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U};W_i,\overleftarrow{\boldsymbol{Y}}_i\right)-H\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}\right)\right]\\
&\quad+H\left(\overrightarrow{\boldsymbol{Y}}_i,\boldsymbol{X}_{i,C}|W_j,\overleftarrow{\boldsymbol{Y}}_j\right)+H\left(\overrightarrow{\boldsymbol{Y}}_j\right)-H\left(\boldsymbol{X}_{i,C}|W_j,\overleftarrow{\boldsymbol{Y}}_j\right)+N\delta(N)\\
&=H\left(\overrightarrow{\boldsymbol{Y}}_i\right)-H\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right)+\left[I\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U};W_i,\overleftarrow{\boldsymbol{Y}}_i\right)-H\left(\boldsymbol{X}_{j,C},\boldsymbol{X}_{i,U}\right)\right]\\
&\quad+H\left(\overrightarrow{\boldsymbol{Y}}_i|W_j,\overleftarrow{\boldsymbol{Y}}_j,\boldsymbol{X}_{i,C}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_j\right)+N\delta(N)
\end{aligned}$$

$$
\begin{aligned}
&\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_i\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) + \left[I\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}; W_i, \overleftarrow{\boldsymbol{Y}}_i\right) - H\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right)\right] \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_i|W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) + N\delta(N) \\
&\overset{(c)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) + I\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}; W_i, \overleftarrow{\boldsymbol{Y}}_i\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i|W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_j|\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) + N\delta(N) \\
&\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_i\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) + I\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}, W_j, \overleftarrow{\boldsymbol{Y}}_j; W_i, \overleftarrow{\boldsymbol{Y}}_i\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i|W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_j|\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) + N\delta(N) \\
&\overset{(d)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i|W_j, W_i,\right) + H\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}, \overleftarrow{\boldsymbol{Y}}_j|W_j\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i|W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_j|\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) + N\delta(N) \\
&\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_i\right) + H\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}, \overleftarrow{\boldsymbol{Y}}_j|W_j\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i|W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j|\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) + N\delta(N) \\
&\leqslant \sum_{n=1}^{N} \Big[H\left(\overrightarrow{\boldsymbol{Y}}_{i,n}\right) + H\left(\boldsymbol{X}_{j,C,n}, \boldsymbol{X}_{i,U,n}, \overleftarrow{\boldsymbol{Y}}_{j,n}|W_j, \boldsymbol{X}_{j,C,(1:n-1)}, \boldsymbol{X}_{i,U,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_{i,n}|W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}, \overrightarrow{\boldsymbol{Y}}_{i,(1:n-1)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_{j,n}|\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}, \overrightarrow{\boldsymbol{Y}}_{j,(1:n-1)}\right)\Big] + N\delta(N) \\
&= \sum_{n=1}^{N} \Big[H\left(\overrightarrow{\boldsymbol{Y}}_{i,n}\right) + H\left(\boldsymbol{X}_{j,C,n}, \boldsymbol{X}_{i,U,n}, \overleftarrow{\boldsymbol{Y}}_{j,n}|W_j, \boldsymbol{X}_{j,C,(1:n-1)}, \boldsymbol{X}_{i,U,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}, \boldsymbol{X}_{j,n}\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_{i,n}|W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}, \overrightarrow{\boldsymbol{Y}}_{i,(1:n-1)}, \boldsymbol{X}_{j,n}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_{j,n}|\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}, \overrightarrow{\boldsymbol{Y}}_{j,(1:n-1)}\right)\Big] + N\delta(N) \\
&\leqslant \sum_{n=1}^{N} \Big[H\left(\overrightarrow{\boldsymbol{Y}}_{i,n}\right) + H\left(\boldsymbol{X}_{i,U,n}|\boldsymbol{X}_{j,n}\right) + H\left(\overleftarrow{\boldsymbol{Y}}_{j,n}|\boldsymbol{X}_{j,n}, \boldsymbol{X}_{i,U,n}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_{i,n}|\boldsymbol{X}_{i,C,n}, \boldsymbol{X}_{j,n}\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_{j,n}|\boldsymbol{X}_{j,C,n}, \boldsymbol{X}_{i,U,n}\right)\Big] + N\delta(N) \\
&\leqslant N\Big[H\left(\overrightarrow{\boldsymbol{Y}}_{i,k}\right) + H\left(\boldsymbol{X}_{i,U,k}\right) + H\left(\overleftarrow{\boldsymbol{Y}}_{j,k}|\boldsymbol{X}_{j,k}, \boldsymbol{X}_{i,U,k}\right) + H\left(\boldsymbol{X}_{i,P,k}\right) + H\left(\boldsymbol{X}_{j,P,k}\right)\Big] + N\delta(N) \\
&= N\Big[H\left(\overrightarrow{\boldsymbol{Y}}_{i,k}\right) + H(\boldsymbol{X}_{i,U,k}) + H\left(\boldsymbol{X}_{i,CF_j,k}, \boldsymbol{X}_{i,DF,k}|\boldsymbol{X}_{i,U,k}\right) + H(\boldsymbol{X}_{i,P,k}) + H\left(\boldsymbol{X}_{j,P,k}\right)\Big] + N\delta(N) \\
&\leqslant N\Big[\dim \overleftarrow{\boldsymbol{Y}}_{i,k} + \dim \overrightarrow{\boldsymbol{Y}}_{i,G,k} + \dim \boldsymbol{X}_{i,U,k} + \left(\dim\left(\boldsymbol{X}_{i,CF_j,k}, \boldsymbol{X}_{i,DF,k}\right) - \dim \boldsymbol{X}_{i,U,k}\right)^+ \\
&\quad + \dim \boldsymbol{X}_{i,P,k} + \dim \boldsymbol{X}_{j,P,k}\Big] + N\delta(N), \qquad \text{(B.13)}
\end{aligned}
$$

where, (a) follows from Fano's inequality with $\delta : \mathbb{N} \to \mathbb{R}_+$ a positive monotonically decreasing function (Lemma 58); (b) follows from the fact that $H\left(\overrightarrow{\boldsymbol{Y}}_i, \overleftarrow{\boldsymbol{Y}}_j|W_i, W_j\right) = 0$; (c) follows from the fact that $H(Y|X) = H(X,Y) - H(X)$; and (d) follows from the fact that $H\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}, \overleftarrow{\boldsymbol{Y}}_j|W_j, W_i, \overleftarrow{\boldsymbol{Y}}_i\right) = 0$.

Plugging (B.1b), (B.3), (B.5), (B.6a), and (B.6b) into (B.13) and after some algebraic manipulations, the following holds in the asymptotic block-length regime:

$$
\begin{aligned}
2R_i + R_j \leqslant & \max\left(\overrightarrow{n}_{ii}, n_{ji}\right) + \left(\overrightarrow{n}_{ii} - n_{ij}\right)^+ \\
& + \max\left(\left(\overrightarrow{n}_{jj} - n_{ji}\right)^+, n_{ij}, \overrightarrow{n}_{jj} - \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+\right). \qquad \text{(B.14)}
\end{aligned}
$$

This completes the proof of (4.1d).

— C —

Proof of Theorem 2

THE proof of Theorem 2 is obtained by comparing $\mathcal{C}(\overleftarrow{n}_{11}, 0)$ $\big($resp. $\mathcal{C}(0, \overleftarrow{n}_{22})\big)$ and $\mathcal{C}(0, 0)$, with fixed parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, and n_{21}. More specifically, for each 4-tuple $\left(\overrightarrow{n}_{11},\ \overrightarrow{n}_{22},\ n_{12},\ n_{21}\right)$, the exact value $\overleftarrow{n}^{*}_{11}$ (resp $\overleftarrow{n}^{*}_{22}$) for which any $\overleftarrow{n}_{11} > \overleftarrow{n}^{*}_{11}$ (resp $\overleftarrow{n}_{22} > \overleftarrow{n}^{*}_{22}$) ensures $\mathcal{C}(0, 0) \subset \mathcal{C}(\overleftarrow{n}_{11}, 0)$ (resp. $\mathcal{C}(0, 0) \subset \mathcal{C}(0, \overleftarrow{n}_{22})$) is calculated. This procedure is so long and repetitive. Then, in this appendix only one combination of interference regimes is studied, namely, VWIR - VWIR.

Proof:

Consider that both transmitter-receiver pairs are in VWIR, that is,

$$\alpha_1 = \frac{n_{12}}{\overrightarrow{n}_{11}} \leqslant \frac{1}{2} \text{ and } \alpha_2 = \frac{n_{21}}{\overrightarrow{n}_{22}} \leqslant \frac{1}{2}. \tag{C.1}$$

When the conditions in (C.1) are fulfilled, it follows from Theorem 1 that $\mathcal{C}(0, 0)$ is the set of rate pairs $(R_1, R_2) \in \mathbb{R}^2_+$ that satisfy:

$$R_1 \leqslant \theta_1 \triangleq \overrightarrow{n}_{11}, \tag{C.2a}$$
$$R_2 \leqslant \theta_2 \triangleq \overrightarrow{n}_{22}, \tag{C.2b}$$
$$R_1 + R_2 \leqslant \theta_3 \triangleq \min\left(\max\left(\overrightarrow{n}_{22}, n_{12}\right) + \overrightarrow{n}_{11} - n_{12}, \max\left(\overrightarrow{n}_{11}, n_{21}\right) + \overrightarrow{n}_{22} - n_{21}\right), \tag{C.2c}$$
$$R_1 + R_2 \leqslant \theta_4 \triangleq \max\left(\overrightarrow{n}_{11} - n_{12}, n_{21}\right) + \max\left(\overrightarrow{n}_{22} - n_{21}, n_{12}\right), \tag{C.2d}$$
$$2R_1 + R_2 \leqslant \theta_5 \triangleq \max\left(\overrightarrow{n}_{11}, n_{21}\right) + \overrightarrow{n}_{11} - n_{12} + \max\left(\overrightarrow{n}_{22} - n_{21}, n_{12}\right), \tag{C.2e}$$
$$R_1 + 2R_2 \leqslant \theta_6 \triangleq \max\left(\overrightarrow{n}_{22}, n_{12}\right) + \overrightarrow{n}_{22} - n_{21} + \max\left(n_{21}, \overrightarrow{n}_{11} - n_{12}\right). \tag{C.2f}$$

Note that for all $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{22}) \in \mathbb{N}^5$ and $\overleftarrow{n}_{11} > \max\left(\overrightarrow{n}_{11}, n_{12}\right)$, it follows that $\mathcal{C}(\overleftarrow{n}_{11}, \overleftarrow{n}_{22}) = \mathcal{C}(\max(\overrightarrow{n}_{11}, n_{12}), \overleftarrow{n}_{22})$. Hence, in the following, the analysis is restricted to the following condition:

$$\overleftarrow{n}_{11} \leqslant \max\left(\overrightarrow{n}_{11}, n_{12}\right). \tag{C.3}$$

Under conditions (C.1) and (C.3), it follows from Theorem 1 that $\mathcal{C}(\overleftarrow{n}_{11}, 0)$ is the set of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy:

$$R_1 \leqslant \overrightarrow{n}_{11}, \tag{C.4a}$$
$$R_2 \leqslant \overrightarrow{n}_{22}, \tag{C.4b}$$
$$R_1 + R_2 \leqslant \min\left(\max\left(\overrightarrow{n}_{22}, n_{12}\right) + \overrightarrow{n}_{11} - n_{12}, \max\left(\overrightarrow{n}_{11}, n_{21}\right) + \overrightarrow{n}_{22} - n_{21}\right), \tag{C.4c}$$
$$R_1 + R_2 \leqslant \theta_7 \triangleq \max\left(\overrightarrow{n}_{11} - n_{12}, n_{21}, \overleftarrow{n}_{11}\right) + \max\left(\overrightarrow{n}_{22} - n_{21}, n_{12}\right), \tag{C.4d}$$
$$2R_1 + R_2 \leqslant \max\left(\overrightarrow{n}_{11}, n_{21}\right) + \overrightarrow{n}_{11} - n_{12} + \max\left(\overrightarrow{n}_{22} - n_{21}, n_{12}\right), \tag{C.4e}$$
$$R_1 + 2R_2 \leqslant \theta_8 \triangleq \max\left(\overrightarrow{n}_{22}, n_{12}\right) + \overrightarrow{n}_{22} - n_{21} + \max\left(\overrightarrow{n}_{11} - n_{12}, n_{21}, \overleftarrow{n}_{11}\right). \tag{C.4f}$$

When comparing $\mathcal{C}(0,0)$ and $\mathcal{C}(\overleftarrow{n}_{11}, 0)$, note that (C.2a), (C.2b), (C.2c), and (C.2e) are equivalent to (C.4a), (C.4b), (C.4c), and (C.4e), respectively. That being the case, the region $\mathcal{C}(\overleftarrow{n}_{11}, 0)$ is larger than the region $\mathcal{C}(0,0)$ if at least one of the following conditions holds true:

$$\min(\theta_3, \theta_4, \theta_1 + \theta_2, \theta_5, \theta_6) < \theta_7 < \min(\theta_3, \theta_1 + \theta_2, \theta_5, \theta_8), \tag{C.5a}$$
$$\min(\theta_6, \theta_1 + 2\theta_2, \theta_2 + \theta_3, \theta_4 + \theta_2) < \theta_8 < \min\left(\theta_1 + 2\theta_2, \theta_2 + \theta_3, \theta_2 + \theta_7\right). \tag{C.5b}$$

Condition (C.5a) implies that the active sum-rate bound in $\mathcal{C}(\overleftarrow{n}_{11}, 0)$ is greater than the active sum-rate bound in $\mathcal{C}(0,0)$. Condition (C.5b) implies that the active weighted sum-rate bound on $R_1 + 2R_2$ in $\mathcal{C}(\overleftarrow{n}_{11}, 0)$ is greater than the active weighted sum-rate bound on $R_1 + 2R_2$ in $\mathcal{C}(0,0)$.

To simplify the inequalities containing the operator $\max(\cdot,\cdot)$ in (C.4) and (C.2), the following 4 cases are identified:

$$\text{Case 1}: \overrightarrow{n}_{11} - n_{12} < n_{21} \text{ and } \overrightarrow{n}_{22} - n_{21} < n_{12}; \tag{C.6}$$
$$\text{Case 2}: \overrightarrow{n}_{11} - n_{12} < n_{21} \text{ and } \overrightarrow{n}_{22} - n_{21} \geqslant n_{12}; \tag{C.7}$$
$$\text{Case 3}: \overrightarrow{n}_{11} - n_{12} \geqslant n_{21} \text{ and } \overrightarrow{n}_{22} - n_{21} < n_{12}; \text{ and} \tag{C.8}$$
$$\text{Case 4}: \overrightarrow{n}_{11} - n_{12} \geqslant n_{21} \text{ and } \overrightarrow{n}_{22} - n_{21} \geqslant n_{12}. \tag{C.9}$$

<u>Case 1:</u> Under condition (C.1), the Case 1, *i.e.*, (C.6), is not possible.

<u>Case 2</u>: Under condition (C.1), the Case 2, *i.e.*, (C.7), is possible.

Plugging (C.7) into (C.4) yields:

$$R_1 + R_2 \leqslant \min\left(\overrightarrow{n}_{22} + \overrightarrow{n}_{11} - n_{12}, \max\left(\overrightarrow{n}_{11}, n_{21}\right) + \overrightarrow{n}_{22} - n_{21}\right), \tag{C.10a}$$
$$R_1 + R_2 \leqslant \max\left(n_{21}, \overleftarrow{n}_{11}\right) + \overrightarrow{n}_{22} - n_{21}, \tag{C.10b}$$
$$R_1 + 2R_2 \leqslant 2\overrightarrow{n}_{22} - n_{21} + \max\left(n_{21}, \overleftarrow{n}_{11}\right). \tag{C.10c}$$

Plugging (C.7) into (C.2) yields:

$$R_1 + R_2 \leqslant \overrightarrow{n}_{22}, \tag{C.11a}$$
$$R_1 + 2R_2 \leqslant 2\overrightarrow{n}_{22}. \tag{C.11b}$$

To simplify the inequalities containing the operator $\max(\cdot,\cdot)$ in (C.10), the following 2 cases are identified:

$$\text{Case 2a}: \overrightarrow{n}_{11} > n_{21} \text{ and} \tag{C.12}$$
$$\text{Case 2b}: \overrightarrow{n}_{11} \leqslant n_{21}. \tag{C.13}$$

<u>Case 2*a*</u>: Plugging (C.12) into (C.10) yields:

$$R_1 + R_2 \leqslant \overrightarrow{n}_{11} + \overrightarrow{n}_{22} - n_{21}, \tag{C.14a}$$
$$R_1 + R_2 \leqslant \max\left(n_{21}, \overleftarrow{n}_{11}\right) + \overrightarrow{n}_{22} - n_{21}, \tag{C.14b}$$
$$R_1 + 2R_2 \leqslant 2\overrightarrow{n}_{22} - n_{21} + \max\left(n_{21}, \overleftarrow{n}_{11}\right). \tag{C.14c}$$

Comparing inequalities (C.14a) and (C.14b) with inequality (C.11a), it can be verified that $\min\left(\overrightarrow{n}_{11} + \overrightarrow{n}_{22} - n_{21},\ \max\left(n_{21},\ \overleftarrow{n}_{11}\right) + \overrightarrow{n}_{22} - n_{21}\right) > \overrightarrow{n}_{22}$, *i.e.*, condition (C.5a) holds, when $\overleftarrow{n}_{11} > n_{21}$. Comparing inequalities (C.14c) and (C.11b), it can be verified that $2\overrightarrow{n}_{22} - n_{21} + \max\left(n_{21}, \overleftarrow{n}_{11}\right) > 2\overrightarrow{n}_{22}$, *i.e.*, condition (C.5b) holds, when $\overleftarrow{n}_{11} > n_{21}$. Therefore, $\overleftarrow{n}^{*}_{11} = n_{21}$ under conditions (C.1), (C.3), (C.7), and (C.12).
<u>Case 2*b*:</u> Plugging (C.13) into (C.10) yields:

$$R_1 + R_2 \leqslant \overrightarrow{n}_{22}, \tag{C.15a}$$
$$R_1 + R_2 \leqslant \max\left(n_{21}, \overleftarrow{n}_{11}\right) + \overrightarrow{n}_{22} - n_{21}, \tag{C.15b}$$
$$R_1 + 2R_2 \leqslant 2\overrightarrow{n}_{22} - n_{21} + \max\left(n_{21}, \overleftarrow{n}_{11}\right). \tag{C.15c}$$

Comparing inequalities (C.15a) and (C.15b) with inequality (C.11a), it can be verified that $\min\left(\overrightarrow{n}_{22},\ \max\left(n_{21},\ \overleftarrow{n}_{11}\right) + \overrightarrow{n}_{22} - n_{21}\right) = \overrightarrow{n}_{22}$, *i.e.*, condition (C.5a) does not hold, for all $\overleftarrow{n}_{11} \in \mathbb{N}$. Comparing inequalities (C.15c) and (C.11b) it can be verified that $2\overrightarrow{n}_{22} - n_{21} + \max\left(n_{21}, \overleftarrow{n}_{11}\right) > 2\overrightarrow{n}_{22}$, when $\overleftarrow{n}_{11} > n_{21}$, which implies that $\overleftarrow{n}_{11} > \max\left(\overrightarrow{n}_{11}, n_{12}\right)$. However, under the conditions (C.1), (C.3), (C.7), and (C.13), the bounds (C.11b) and (C.15c) are not active. Hence, condition (C.5b) does not hold. Therefore, for all $\overleftarrow{n}_{11} \in \mathbb{N}$, the capacity region cannot be enlarged under conditions (C.1), (C.3), (C.7), and (C.13).
<u>Case 3:</u> Under condition (C.1), the Case 3, *i.e.*, (C.8), is possible.
Plugging (C.8) into (C.4) yields:

$$R_1 + R_2 \leqslant \min\left(\max\left(\overrightarrow{n}_{22}, n_{12}\right) + \overrightarrow{n}_{11} - n_{12}, \overrightarrow{n}_{11} + \overrightarrow{n}_{22} - n_{21}\right), \tag{C.16a}$$
$$R_1 + R_2 \leqslant \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right) + n_{12}, \tag{C.16b}$$
$$R_1 + 2R_2 \leqslant \max\left(\overrightarrow{n}_{22}, n_{12}\right) + \overrightarrow{n}_{22} - n_{21} + \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right). \tag{C.16c}$$

Plugging (C.8) into (C.2) yields:

$$R_1 + R_2 \leqslant \overrightarrow{n}_{11}, \tag{C.17a}$$
$$R_1 + 2R_2 \leqslant \max\left(\overrightarrow{n}_{22}, n_{12}\right) + \overrightarrow{n}_{22} - n_{21} + \overrightarrow{n}_{11} - n_{12}. \tag{C.17b}$$

To simplify the inequalities containing the operator $\max(\cdot, \cdot)$ in (C.16) and (C.17), the following 2 cases are identified:

$$\text{Case 3a}: \overrightarrow{n}_{22} > n_{12} \text{ and} \tag{C.18}$$
$$\text{Case 3b}: \overrightarrow{n}_{22} \leqslant n_{12}. \tag{C.19}$$

Case 3a: Plugging (C.18) into (C.16) yields:

$$R_1 + R_2 \leqslant \overrightarrow{n}_{22} + \overrightarrow{n}_{11} - n_{12}, \tag{C.20a}$$
$$R_1 + R_2 \leqslant \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right) + n_{12}, \tag{C.20b}$$
$$R_1 + 2R_2 \leqslant 2\overrightarrow{n}_{22} - n_{21} + \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right). \tag{C.20c}$$

Plugging (C.18) into (C.17) yields:

$$R_1 + R_2 \leqslant \overrightarrow{n}_{11}, \tag{C.21a}$$
$$R_1 + 2R_2 \leqslant 2\overrightarrow{n}_{22} - n_{21} + \overrightarrow{n}_{11} - n_{12}. \tag{C.21b}$$

Comparing inequalities (C.20a) and (C.20b) with inequality (C.21a), it can be verified that $\min\left(\overrightarrow{n}_{22} + \overrightarrow{n}_{11} - n_{12}, \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right) + n_{12}\right) > \overrightarrow{n}_{11}$, *i.e.*, condition (C.5a) holds when $\overleftarrow{n}_{11} > \overrightarrow{n}_{11} - n_{12}$. Comparing inequalities (C.20c) and (C.21b), it can be verified that $2\overrightarrow{n}_{22} - n_{21} + \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right) > 2\overrightarrow{n}_{22} - n_{21} + \overrightarrow{n}_{11} - n_{12}$, *i.e.*, condition (C.5b) holds when $\overleftarrow{n}_{11} > \overrightarrow{n}_{11} - n_{12}$. Therefore, $\overleftarrow{n}^*_{11} = \overrightarrow{n}_{11} - n_{12}$ under conditions (C.1), (C.3), (C.8), and (C.18).

Case 3b: Plugging (C.19) into (C.16) yields:

$$R_1 + R_2 \leqslant \overrightarrow{n}_{11}, \tag{C.22a}$$
$$R_1 + R_2 \leqslant \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right) + n_{12}, \tag{C.22b}$$
$$R_1 + 2R_2 \leqslant n_{12} + \overrightarrow{n}_{22} - n_{21} + \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right). \tag{C.22c}$$

Plugging (C.18) into (C.17) yields:

$$R_1 + R_2 \leqslant \overrightarrow{n}_{11}, \tag{C.23a}$$
$$R_1 + 2R_2 \leqslant \overrightarrow{n}_{22} - n_{21} + \overrightarrow{n}_{11}. \tag{C.23b}$$

Comparing inequalities (C.22a) and (C.22b) with inequality (C.23a), it can be verified that $\min\left(\overrightarrow{n}_{11}, \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right) + n_{12}\right) = \overrightarrow{n}_{11}$, *i.e.*, condition (C.5a) does not hold, for all $\overleftarrow{n}_{11} \in \mathbb{N}$. Comparing inequalities (C.22c) and (C.23b), it can be verified that $n_{12} + \overrightarrow{n}_{22} - n_{21} + \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right) > \overrightarrow{n}_{22} - n_{21} + \overrightarrow{n}_{11}$, *i.e.*, condition (C.5b) holds when $\overleftarrow{n}_{11} > \overrightarrow{n}_{11} - n_{12}$. Therefore, $\overleftarrow{n}^*_{11} = \overrightarrow{n}_{11} - n_{12}$ under conditions (C.1), (C.3), (C.8), and (C.19).

Case 4: Under condition (C.1), the Case 4, *i.e.*, (C.9), is possible.

Plugging (C.9) into (C.4) yields:

$$R_1 + R_2 \leqslant \min\left(\overrightarrow{n}_{22} + \overrightarrow{n}_{11} - n_{12}, \overrightarrow{n}_{11} + \overrightarrow{n}_{22} - n_{21}\right), \tag{C.24a}$$
$$R_1 + R_2 \leqslant \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right) + \overrightarrow{n}_{22} - n_{21}, \tag{C.24b}$$
$$R_1 + 2R_2 \leqslant 2\overrightarrow{n}_{22} - n_{21} + \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right). \tag{C.24c}$$

Plugging (C.9) into (C.2) yields:

$$R_1 + R_2 \leqslant \overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21}, \tag{C.25a}$$
$$R_1 + 2R_2 \leqslant 2\overrightarrow{n}_{22} - n_{21} + \overrightarrow{n}_{11} - n_{12}. \tag{C.25b}$$

Comparing inequalities (C.24a) and (C.24b) with inequality (C.25a), it can be verified that

$\min\Big(\min\left(\overrightarrow{n}_{22}+\overrightarrow{n}_{11}-n_{12},\ \overrightarrow{n}_{11}+\overrightarrow{n}_{22}-n_{21}\right),\ \max\left(\overrightarrow{n}_{11}-n_{12},\ \overleftarrow{n}_{11}\right)+\overrightarrow{n}_{22}-n_{21}\Big) > \overrightarrow{n}_{11}-n_{12}+\overrightarrow{n}_{22}-n_{21}$, *i.e.*, condition (C.5a) holds when $\overleftarrow{n}_{11} > \overrightarrow{n}_{11}-n_{12}$. Comparing inequalities (C.24c) and (C.25b), it can be verified that: $2\overrightarrow{n}_{22}-n_{21}+\max\left(\overrightarrow{n}_{11}-n_{12}, \overleftarrow{n}_{11}\right) > 2\overrightarrow{n}_{22}-n_{21}+\overrightarrow{n}_{11}-n_{12}$, *i.e.*, condition (C.5b) holds when $\overleftarrow{n}_{11} > \overrightarrow{n}_{11}-n_{12}$.

Therefore, $\overleftarrow{n}^*_{11} = \overrightarrow{n}_{11}-n_{12}$ under conditions (C.1), (C.3), and (C.9).

From all the observations above, when both transmitter-receiver pairs are in VWIR (event E_1 in (4.6) holds true), it follows that when $\overleftarrow{n}_{11} > \overleftarrow{n}^*_{11}$ and $\overrightarrow{n}_{11} > n_{21}$ (event $E_{8,1}$ in (4.13) with $i=1$ holds true) with $\overleftarrow{n}^*_{11} = \max\left(\overrightarrow{n}_{11}-n_{12}, n_{21}\right)$, then $\mathcal{C}(0,0) \subset \mathcal{C}(\overleftarrow{n}_{11},0)$. Otherwise, $\mathcal{C}(0,0) = \mathcal{C}(\overleftarrow{n}_{11},0)$. Note that when events E_1 and $E_{8,1}$ hold simultaneously true, then the event $S_{1,1}$ in (4.17) with $i=1$ holds true, which verifies the statement of Theorem 2. The same procedure can be applied for all the other combinations of interference regimes. This completes the proof. ■

— D —

Proof of Theorem 3

The proof of Theorem 3 is obtained by comparing $\mathcal{C}(\overleftarrow{n}_{11}, 0)$ $\big($resp. $\mathcal{C}(0, \overleftarrow{n}_{22})\big)$ and $\mathcal{C}(0, 0)$, for all possible parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, and $\overleftarrow{n}_{11}$ (resp. $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, and $\overleftarrow{n}_{22}$). More specifically, for each 4-tuple $\big(\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, $n_{21}\big)$, the exact value $\overleftarrow{n}_{11}^{\dagger}$ (resp $\overleftarrow{n}_{22}^{\dagger}$) for which any $\overleftarrow{n}_{11} > \overleftarrow{n}_{11}^{\dagger}$ (resp $\overleftarrow{n}_{22} > \overleftarrow{n}_{22}^{\dagger}$) ensures an improvement on R_1 (resp. R_2), *i.e.*, $\Delta_1(\overleftarrow{n}_{11}, 0) > 0$ (resp. $\Delta_2(0, \overleftarrow{n}_{22}) > 0$), is calculated. This procedure is so long and repetitive. Then, in this appendix only one combination of interference regimes is studied, namely, VWIR - VWIR.

Proof:

Consider that both transmitter-receiver pairs are in VWIR, *i.e.*, conditions (C.1) hold. Under these conditions, the capacity regions $\mathcal{C}(0, 0)$ and $\mathcal{C}(\overleftarrow{n}_{11}, 0)$ are given by (C.2) and (C.4), respectively. When comparing $\mathcal{C}(0, 0)$ and $\mathcal{C}(\overleftarrow{n}_{11}, 0)$, note that (C.2a), (C.2b), (C.2c), and (C.2e) are equivalent to (C.4a), (C.4b), (C.4c), and (C.4e), respectively. In this case any improvement on R_1 is produced by an improvement on $R_1 + R_2$ (condition (C.5a)) or $2R_1 + R_2$ (condition (C.5a)), and thus, the proof of Theorem 3 in these particular interference regimes follows exactly the same steps as in Theorem 2. This completes the proof. ∎

E

Proof of Theorem 5

The proof of Theorem 5 is obtained by comparing $\mathcal{C}(\overleftarrow{n}_{11}, 0)$ $\big($resp. $\mathcal{C}(0, \overleftarrow{n}_{22})\big)$ and $\mathcal{C}(0, 0)$, for all possible parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, and $\overleftarrow{n}_{11}$ (resp. $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, and $\overleftarrow{n}_{22}$). More specifically, for each 4-tuple $\big(\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, $n_{21}\big)$, the exact value $\overleftarrow{n}_{11}^{+}$ (resp $\overleftarrow{n}_{22}^{+}$) for which any $\overleftarrow{n}_{11} > \overleftarrow{n}_{11}^{+}$ (resp $\overleftarrow{n}_{22} > \overleftarrow{n}_{22}^{+}$) ensures an improvement on $R_1 + R_2$, *i.e.*, $\Sigma(\overleftarrow{n}_{11}, 0) > 0$ (resp. $\Sigma(0, \overleftarrow{n}_{22}) > 0$), is calculated. This procedure is so long and repetitive. Then, in this appendix only one combination of interference regimes is studied, namely, VWIR - VWIR.

Proof:

Consider that both transmitter-receiver pairs are in VWIR, *i.e.*, conditions (C.1) hold. Under these conditions, the capacity regions $\mathcal{C}(0, 0)$ and $\mathcal{C}(\overleftarrow{n}_{11}, 0)$ are given by (C.2) and (C.4), respectively. When comparing $\mathcal{C}(0, 0)$ and $\mathcal{C}(\overleftarrow{n}_{11}, 0)$, note that (C.2a), (C.2b), (C.2c), and (C.2e) are equivalent to (C.4a), (C.4b), (C.4c), and (C.4e), respectively.

In this case, the proof is focused on any improvement on $R_1 + R_2$ (condition (C.5a)), and thus, the proof of Theorem 5 in these particular interference regimes follows exactly the same steps as in Theorem 2.

From the analysis presented in Appendix C, it follows that:

Case 2*a*: condition (C.5a) holds true, when $\overleftarrow{n}_{11} > n_{21}$ under conditions (C.1), (C.3), (C.7), and (C.12).

Case 2*b*: condition (C.5a) does not hold true, under conditions (C.1), (C.7), and (C.13).

Case 3*a*: condition (C.5a) holds true, when $\overleftarrow{n}_{11} > \overrightarrow{n}_{11} - n_{12}$ under conditions (C.1), (C.3), (C.8), and (C.18).

Case 3*b*: condition (C.5a) does not hold true, when $\overleftarrow{n}_{11} > \overrightarrow{n}_{11} - n_{12}$ under conditions (C.1), (C.3), (C.8), and (C.19).

Case 4: condition (C.5a) holds true, when $\overleftarrow{n}_{11} > \overrightarrow{n}_{11} - n_{12}$ under conditions (C.1), (C.3), and (C.9).

From all the observations above, when both transmitter-receiver pairs are in VWIR (event E_1 in (4.6) holds true), it follows that when $\overleftarrow{n}_{11} > \overleftarrow{n}_{11}^{+}$, $\overrightarrow{n}_{11} > n_{21}$ (event $E_{8,1}$ in (4.13) with $i = 1$

holds true), $\overrightarrow{n}_{22} > n_{12}$ (event $E_{8,2}$ in (4.13) with $i = 2$ holds true), $\overrightarrow{n}_{11} + \overrightarrow{n}_{22} > n_{12} + 2n_{21}$ (event $E_{10,1}$ in (4.15) with $i = 1$ holds true), and $\overrightarrow{n}_{11} + \overrightarrow{n}_{22} > n_{21} + 2n_{12}$ (event $E_{10,2}$ in (4.15) with $i = 2$ holds true) with $\overleftarrow{n}_{11}^{+} = \max\left(\overrightarrow{n}_{11} - n_{12}, n_{21}\right)$, then $\Sigma(\overleftarrow{n}_{11}, 0) > 0$. Otherwise, $\Sigma(\overleftarrow{n}_{11}, 0) = 0$. Note that when events E_1, $E_{8,1}$, $E_{8,2}$, $E_{10,1}$, and $E_{10,2}$ hold simultaneously true, then the event S_4 in (4.20) holds true, which verifies the statement of Theorem 5. The same procedure can be applied for all the other combinations of interference regimes. This completes the proof. ■

— F —

Proof of Theorem 6

THIS appendix provides a proof to Theorem 6 for the two-user LDIC-NOF.

Proof:

Under symmetric conditions, *i.e.*, $\overrightarrow{n} = \overrightarrow{n}_{11} = \overrightarrow{n}_{22}$, $m = n_{12} = n_{21}$ and $\overleftarrow{n} = \overleftarrow{n}_{11} = \overleftarrow{n}_{22}$, from (4.1a) and (4.1b) with $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$, it follows that:

$$R_1 \leqslant a_1 \triangleq \min\left(\max\left(\overrightarrow{n}, m\right), \max\left(\overrightarrow{n}, \overleftarrow{n} - \left(\overrightarrow{n} - m\right)^+\right)\right), \tag{F.1}$$
$$R_2 \leqslant a_1 \triangleq \min\left(\max\left(\overrightarrow{n}, m\right), \max\left(\overrightarrow{n}, \overleftarrow{n} - \left(\overrightarrow{n} - m\right)^+\right)\right); \tag{F.2}$$

from (4.1c) and (4.1c), it follows that:

$$\begin{aligned} R_1+R_2 \leqslant{} & a_2 \\ \triangleq{} & \min\left(\max\left(\overrightarrow{n}, m\right) + \left(\overrightarrow{n} - m\right)^+, 2\max\left(\left(\overrightarrow{n} - m\right)^+, m, \overrightarrow{n} - \left(\max\left(\overrightarrow{n}, m\right) - \overleftarrow{n}\right)^+\right)\right); \end{aligned} \tag{F.3}$$

and from (4.1d), it follows that:

$$\begin{aligned} 2R_1+R_2 \leqslant{} & a_3 \\ \triangleq{} & \max\left(\overrightarrow{n}, m\right) + \left(\overrightarrow{n} - m\right)^+ + \max\left(\left(\overrightarrow{n} - m\right)^+, m, \overrightarrow{n} - \left(\max\left(\overrightarrow{n}, m\right) - \overleftarrow{n}\right)^+\right), \end{aligned} \tag{F.4}$$
$$\begin{aligned} R_1+2R_2 \leqslant{} & a_3 \\ \triangleq{} & \max\left(\overrightarrow{n}, m\right) + \left(\overrightarrow{n} - m\right)^+ + \max\left(\left(\overrightarrow{n} - m\right)^+, m, \overrightarrow{n} - \left(\max\left(\overrightarrow{n}, m\right) - \overleftarrow{n}\right)^+\right). \end{aligned} \tag{F.5}$$

The sum-capacity can be obtained considering the sum of (F.1) and (F.2); (F.3); and the sum-rate bound that can be obtained from (F.4) and (F.5) with $R1 \geqslant 0$ and $R2 \geqslant 0$, respectively. Then,

$$\begin{aligned} R_1 + R_2 \leqslant & \min\left(2a_1, a_2, a_3\right) \\ = & \min\left(2a_1, a_2\right), \end{aligned} \tag{F.6}$$

given that $a_3 \geqslant a_2$.

The symmetric capacity, $C_{\text{sym}}(\overrightarrow{n}, m, \overleftarrow{n}) = \sup\{R \in \mathbb{R}_+ : (R, R) \in \mathcal{C}(\overrightarrow{n},\ \overrightarrow{n},\ m,\ m,\ \overleftarrow{n},\ \overleftarrow{n})\}$, can be obtained from (F.6), (F.1), and (F.3) as follows:

$$\begin{aligned} C_{\text{sym}} = & \min\left(a_1, \frac{a_2}{2}\right) \\ = & \min\Bigg(\max\left(\overrightarrow{n}, m\right), \max\left(\overrightarrow{n}, \overleftarrow{n} - \left(\overrightarrow{n} - m\right)^+\right), \frac{1}{2}\left(\max\left(\overrightarrow{n}, m\right) + \left(\overrightarrow{n} - m\right)^+\right), \\ & \max\left(\left(\overrightarrow{n} - m\right)^+, m, \overrightarrow{n} - \left(\max\left(\overrightarrow{n}, m\right) - \overleftarrow{n}\right)^+\right)\Bigg). \end{aligned} \tag{F.7}$$

Plugging (F.7) into (4.31) yields:

$$\begin{aligned} D_{\text{sym}}(\alpha, \beta) = \min\Bigg(& \max\left(1, \alpha\right), \max\left(1, \beta - \left(1 - \alpha\right)^+\right), \frac{1}{2}\left(\max\left(1, \alpha\right) + \left(1 - \alpha\right)^+\right), \\ & \max\left(\left(1 - \alpha\right)^+, \alpha, 1 - \left(\max\left(1, \alpha\right) - \beta\right)^+\right)\Bigg), \end{aligned} \tag{F.8}$$

where $\alpha = \frac{m}{\overrightarrow{n}}$ and $\beta = \frac{\overleftarrow{n}}{\overrightarrow{n}}$. This completes the proof. ■

— G —

Proof of Theorem 8

The outer bounds (5.9a) and (5.9c) correspond to the outer bounds of the case of POF derived in [88]. The bounds (5.9b), (5.9d) and (5.9e) correspond to new outer bounds. Before presenting the proof, consider the parameter $h_{ji,U}$, with $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$, defined as follows:

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

where, the events $S_{1,i}$, $S_{2,i}$, $S_{3,i}$, $S_{4,i}$, and $S_{5,i}$ are defined in (5.4). Consider also the following signals:

$$X_{i,C,n} = \sqrt{\mathrm{INR}_{ji}} X_{i,n} + \overrightarrow{Z}_{j,n} \qquad \text{and} \tag{G.2}$$

$$X_{i,U,n} = h_{ji,U} X_{i,n} + \overrightarrow{Z}_{j,n}, \tag{G.3}$$

where, $X_{i,n}$ and $\overrightarrow{Z}_{j,n}$ are the channel input of transmitter i and the noise observed at receiver j during a given channel use $n \in \{1, 2, \ldots, N\}$, as described by (2.5). The following lemma is instrumental in the present proof of Theorem 8.

Lemma 21. *For all $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$, the following holds:*

$$I\Big(\boldsymbol{X}_{i,C}, \boldsymbol{X}_{j,U}, \overleftarrow{\boldsymbol{Y}}_i, W_i; \overleftarrow{\boldsymbol{Y}}_j, W_j\Big) \leqslant h\left(\overleftarrow{\boldsymbol{Y}}_j | W_j\right) + \sum_{n=1}^{N} \Big[h\left(X_{j,U,n} | X_{i,C,n}\right) + h\left(\overleftarrow{Y}_{i,n} | X_{i,n}, X_{j,U,n}\right) - \frac{3}{2}\log(2\pi e) \Big]. \tag{G.4}$$

Proof: The proof of Lemma 21 is presented in appendix L. ∎

Proof of (5.9b): From the assumption that the message index W_i is i.i.d. following a

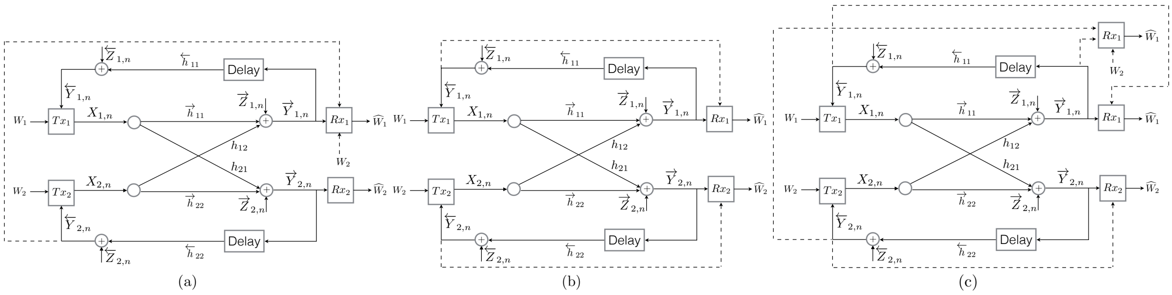


Figure G.1.: Genie-Aided GIC-NOF models for channel use n. (a) Model used to calculate the outer bound on R_1; (b) Model used to calculate the outer bound on R_1+R_2; and (c) Model used to calculate the outer bound on $2R_1+R_2$

uniform distribution over the set $\mathcal{W}_i$, the following holds for any $k \in \{1, 2, \ldots, N\}$:

$$\begin{aligned}
NR_i =& H\left(W_i\right) \\
=& H\left(W_i | W_j\right) \\
\overset{(a)}{\leqslant}& I\left(W_i; \overrightarrow{\boldsymbol{Y}}_i, \overleftarrow{\boldsymbol{Y}}_j | W_j\right) + N\delta(N) \\
\leqslant& \sum_{n=1}^{N} \left[h\Big(\overrightarrow{Y}_{i,n}, \overleftarrow{Y}_{j,n} | W_j, \overrightarrow{\boldsymbol{Y}}_{i,(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}, X_{j,n}\Big) - h\left(\overrightarrow{Z}_{i,n}\right) - h\left(\overleftarrow{Z}_{j,n}\right) \right] \\
& + N\delta(N) \\
\leqslant& \sum_{n=1}^{N} \left[h\left(\overrightarrow{Y}_{i,n}, \overleftarrow{Y}_{j,n} | X_{j,n}\right) - h\left(\overrightarrow{Z}_{i,n}\right) - h\left(\overleftarrow{Z}_{j,n}\right) \right] + N\delta(N) \\
=& N\left[h\left(\overrightarrow{Y}_{i,k}, \overleftarrow{Y}_{j,k} | X_{j,k}\right) - \log\left(2\pi e\right) \right] + N\delta(N),
\end{aligned} \tag{G.5}$$

where (a) follows from Fano's inequality with $\delta: \mathbb{N} \rightarrow \mathbb{R}_+$ a positive monotonically decreasing function (Lemma 58) (see Figure G.1a).

From (G.5), the following holds in the asymptotic block-length regime:

$$\begin{aligned}
R_i \leqslant& h\left(\overrightarrow{Y}_{i,k}, \overleftarrow{Y}_{j,k} | X_{j,k}\right) - \log\left(2\pi e\right) \\
\leqslant& \frac{1}{2}\log\Big(b_{3,i}+1\Big) + \frac{1}{2}\log\left(\frac{\Big(b_{3,i} + b_{4,j}(\rho) + 1\Big)\overleftarrow{\mathrm{SNR}}_j}{\Big(b_{1,j}(\rho)+1\Big)\Big(b_{3,i} + (1-\rho^2)\Big)} + 1 \right).
\end{aligned} \tag{G.6}$$

This completes the proof of (5.9b).

Proof of (5.9d):

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

$$\begin{aligned}
N\Big(R_1 + R_2\Big) =& H\left(W_1\right) + H\left(W_2\right) \\
\overset{(a)}{\leqslant}& I\left(W_1; \overrightarrow{\boldsymbol{Y}}_1, \overleftarrow{\boldsymbol{Y}}_1\right) + I\left(W_2; \overrightarrow{\boldsymbol{Y}}_2, \overleftarrow{\boldsymbol{Y}}_2\right) + N\delta(N)
\end{aligned}$$

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

$$
\begin{aligned}
\leqslant&\sum_{n=1}^{N}\Big[h\left(\overrightarrow{Y}_{1,n}|X_{1,C,n},X_{2,U,n}\right)+h\left(X_{1,U,n}|X_{2,C,n}\right)+h\left(\overleftarrow{Y}_{2,n}|X_{2,n},X_{1,U,n}\right)\\
&+h\left(\overrightarrow{Y}_{2,n}|X_{2,C,n},X_{1,U,n}\right)+h\left(X_{2,U,n}|X_{1,C,n}\right)+h\left(\overleftarrow{Y}_{1,n}|X_{1,n},X_{2,U,n}\right)-3\log\left(2\pi e\right)\Big]\\
&+N\log\left(2\pi e\right)+N\delta(N)\\
=&N\Big[h\left(\overrightarrow{Y}_{1,k}|X_{1,C,k},X_{2,U,k}\right)+h\left(X_{1,U,k}|X_{2,C,k}\right)+h\left(\overleftarrow{Y}_{2,k}|X_{2,k},X_{1,U,k}\right)\\
&+h\left(\overrightarrow{Y}_{2,k}|X_{2,C,k},X_{1,U,k}\right)+h\left(X_{2,U,k}|X_{1,C,k}\right)+h\left(\overleftarrow{Y}_{1,k}|X_{1,k},X_{2,U,k}\right)-3\log\left(2\pi e\right)\Big]\\
&+N\log\left(2\pi e\right)+N\delta(N), \qquad \text{(G.7)}
\end{aligned}
$$

where (a) follows from Fano's inequality with $\delta:\mathbb{N}\rightarrow\mathbb{R}_{+}$ a positive monotonically decreasing function (Lemma 58) (see Figure G.1b); (b) follows from the fact that $h\left(\overrightarrow{\boldsymbol{Y}}_{i}\right)-h\left(\boldsymbol{X}_{i,C},\boldsymbol{X}_{j,U}\right)+h\left(\boldsymbol{X}_{i,C},\boldsymbol{X}_{j,U}|\overrightarrow{\boldsymbol{Y}}_{i}\right)=h\left(\overrightarrow{\boldsymbol{Y}}_{i}|\boldsymbol{X}_{i,C},\boldsymbol{X}_{j,U}\right)$; (c) follows from the fact that $h\left(\overrightarrow{\boldsymbol{Z}}_{i}|W_{j},\overleftarrow{\boldsymbol{Y}}_{j},\boldsymbol{X}_{j},\boldsymbol{X}_{i,C}\right)-h\left(\overrightarrow{\boldsymbol{Z}}_{i},\overrightarrow{\boldsymbol{Z}}_{j}|\overleftarrow{\boldsymbol{Y}}_{j},\boldsymbol{X}_{i},\boldsymbol{X}_{j}\right)\leqslant 0$; and (d) follows from Lemma 21.

From (G.7), the following holds in the asymptotic block-length regime for any $k\in\{1,2,\ldots,N\}$:

$$
\begin{aligned}
R_1+R_2\leqslant&h\left(\overrightarrow{Y}_{1,k}|X_{1,C,k},X_{2,U,k}\right)+h\left(X_{1,U,k}|X_{2,C,k}\right)+h\left(\overleftarrow{Y}_{2,k}|X_{2,k},X_{1,U,k}\right)\\
&+h\left(\overrightarrow{Y}_{2,k}|X_{2,C,k},X_{1,U,k}\right)+h\left(X_{2,U,k}|X_{1,C,k}\right)+h\left(\overleftarrow{Y}_{1,k}|X_{1,k},X_{2,U,k}\right)-2\log\left(2\pi e\right)\\
\leqslant&\frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(\overrightarrow{Y}_{1,k},X_{1,C,k},X_{2,U,k}\right)\right)\right)-\frac{1}{2}\log\left(\mathrm{INR}_{12}+1\right)\\
&+\frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(\overleftarrow{Y}_{2,k},X_{2,k},X_{1,U,k}\right)\right)\right)-\frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(X_{2,k},X_{1,U,k}\right)\right)\right)\\
&+\frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(\overrightarrow{Y}_{2,k},X_{2,C,k},X_{1,U,k}\right)\right)\right)-\frac{1}{2}\log\left(\mathrm{INR}_{21}+1\right)\\
&+\frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(\overleftarrow{Y}_{1,k},X_{1,k},X_{2,U,k}\right)\right)\right)-\frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(X_{1,k},X_{2,U,k}\right)\right)\right)\\
&+\log\left(2\pi e\right), \qquad \text{(G.8)}
\end{aligned}
$$

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

$$
\begin{aligned}
\det\left(\mathrm{Var}\left(\overrightarrow{Y}_{j,k},X_{j,C,k},X_{i,U,k}\right)\right)=&\overrightarrow{\mathrm{SNR}}_j+\mathrm{INR}_{ji}+h_{ji,U}^2-2h_{ji,U}\sqrt{\mathrm{INR}_{ji}} \qquad \text{(G.9a)}\\
&+\left(1-\rho^2\right)\left(\mathrm{INR}_{ij}\mathrm{INR}_{ji}+h_{ji,U}^2\left(\overrightarrow{\mathrm{SNR}}_j+\mathrm{INR}_{ij}\right)-2h_{ji,U}\mathrm{INR}_{ij}\sqrt{\mathrm{INR}_{ji}}\right)\\
&+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_j}\left(\sqrt{\mathrm{INR}_{ji}}-h_{ji,U}\right),
\end{aligned}
$$

$$
\begin{aligned}
\det\left(\mathrm{Var}\left(\overleftarrow{Y}_{j,k},X_{j,k},X_{i,U,k}\right)\right)=&1+h_{ji,U}^2\left(1-\rho^2\right) \qquad \text{(G.9b)}\\
&+\frac{\overleftarrow{\mathrm{SNR}}_j\left(1-\rho^2\right)\left(h_{ji,U}^2-2h_{ji,U}\sqrt{\mathrm{INR}_{ji}}+\mathrm{INR}_{ji}\right)}{\left(\overrightarrow{\mathrm{SNR}}_j+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_j\mathrm{INR}_{ji}}+\mathrm{INR}_{ji}+1\right)}, \text{ and}\\
\det\left(\mathrm{Var}\left(X_{j,k},X_{i,U,k}\right)\right)=&1+\left(1-\rho^2\right)h_{ji,U}^2. \qquad \text{(G.9c)}
\end{aligned}
$$

The expressions in (G.9) depend on $S_{1,i}$, $S_{2,i}$, $S_{3,i}$, $S_{4,i}$, and $S_{5,i}$ via the parameter $h_{ji,U}$ in

(G.1). Hence, the following cases are identified:

Case 1: $(S_{1,2} \vee S_{2,2} \vee S_{5,2}) \wedge (S_{1,1} \vee S_{2,1} \vee S_{5,1})$. From (G.1), it follows that $h_{12,U} = 0$ and $h_{21,U} = 0$. Therefore, plugging the expression (G.9) into (G.8) yields (5.6a).

Case 2: $(S_{1,2} \vee S_{2,2} \vee S_{5,2}) \wedge (S_{3,1} \vee S_{4,1})$. From (G.1), it follows that $h_{12,U} = 0$ and $h_{21,U} = \sqrt{\frac{\mathrm{INR}_{12}\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_2}}$. Therefore, plugging the expression (G.9) into (G.8) yields (5.6b).

Case 3: $(S_{3,2} \vee S_{4,2}) \wedge (S_{1,1} \vee S_{2,1} \vee S_{5,1})$. From (G.1), it follows that $h_{12,U} = \sqrt{\frac{\mathrm{INR}_{12}\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_1}}$ and $h_{21,U} = 0$. Therefore, plugging the expression (G.9) into (G.8) yields (5.6c).

Case 4: $(S_{3,2} \vee S_{4,2}) \wedge (S_{3,1} \vee S_{4,1})$. From (G.1), it follows that $h_{12,U} = \sqrt{\frac{\mathrm{INR}_{12}\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_1}}$ and $h_{21,U} = \sqrt{\frac{\mathrm{INR}_{12}\mathrm{INR}_{21}}{\overrightarrow{\mathrm{SNR}}_2}}$. Therefore, plugging the expression (G.9) into (G.8) yields (5.6d).

This completes the proof of (5.9d).

Proof of (5.9e): From the assumption that the message indices W_i and W_j are i.i.d. following a uniform distribution over the sets $\mathcal{W}_i$ and $\mathcal{W}_j$ respectively, for all $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$, the following holds for any $k \in \{1,2,\ldots,N\}$:

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

$$
\begin{aligned}
&\overset{(e)}{\leqslant} h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) - h\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}|W_i, \overleftarrow{\boldsymbol{Y}}_i\right) + h\left(\overrightarrow{\boldsymbol{Z}}_j|W_i, \overleftarrow{\boldsymbol{Y}}_i, \boldsymbol{X}_i, \boldsymbol{X}_{j,C}\right)\\
&\quad + h\left(\overrightarrow{\boldsymbol{Y}}_i|W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right) - h\left(\overrightarrow{\boldsymbol{Y}}_i, \boldsymbol{X}_{i,C}|W_i, W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_i, \boldsymbol{X}_j\right) + h\left(\overrightarrow{\boldsymbol{Y}}_j\right) + N\log(2\pi e)\\
&\quad + N\delta(N)\\
&= h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) - h\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}|W_i, \overleftarrow{\boldsymbol{Y}}_i\right) + h\left(\overrightarrow{\boldsymbol{Z}}_j|W_i, \overleftarrow{\boldsymbol{Y}}_i, \boldsymbol{X}_i, \boldsymbol{X}_{j,C}\right)\\
&\quad + h\left(\overrightarrow{\boldsymbol{Y}}_i|W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right) - h\left(\overrightarrow{\boldsymbol{Z}}_i, \overrightarrow{\boldsymbol{Z}}_j|W_i, W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_i, \boldsymbol{X}_j\right) + h\left(\overrightarrow{\boldsymbol{Y}}_j\right)\\
&\quad + N\log(2\pi e) + N\delta(N)\\
&\overset{(f)}{\leqslant} h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) - h\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}|W_i, \overleftarrow{\boldsymbol{Y}}_i\right) + h\left(\overrightarrow{\boldsymbol{Y}}_i|W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right) + h\left(\overrightarrow{\boldsymbol{Y}}_j\right)\\
&\quad + N\log(2\pi e) + N\delta(N)\\
&\leqslant h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) + I\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}; W_i, \overleftarrow{\boldsymbol{Y}}_i\right) - h\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) + h\left(\overrightarrow{\boldsymbol{Y}}_i|W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right)\\
&\quad + h\left(\overrightarrow{\boldsymbol{Y}}_j\right) + h\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}|\overrightarrow{\boldsymbol{Y}}_j\right) + N\log(2\pi e) + N\delta(N)\\
&\overset{(g)}{=} h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) + I\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}; W_i, \overleftarrow{\boldsymbol{Y}}_i\right) + h\left(\overrightarrow{\boldsymbol{Y}}_i|W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right)\\
&\quad + h\left(\overrightarrow{\boldsymbol{Y}}_j|\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) + N\log(2\pi e) + N\delta(N)\\
&\leqslant h\left(\overrightarrow{\boldsymbol{Y}}_i\right) - h\left(\overleftarrow{\boldsymbol{Y}}_i|W_i\right) + I\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}, W_j, \overleftarrow{\boldsymbol{Y}}_j; W_i, \overleftarrow{\boldsymbol{Y}}_i\right) + h\left(\overrightarrow{\boldsymbol{Y}}_i|W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right)\\
&\quad + h\left(\overrightarrow{\boldsymbol{Y}}_j|\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) + N\log(2\pi e) + N\delta(N)\\
&\overset{(h)}{\leqslant} h\left(\overrightarrow{\boldsymbol{Y}}_i\right) + \sum_{n=1}^{N}\left[h\left(X_{i,U,n}|X_{j,C,n}\right) + h\left(\overleftarrow{\boldsymbol{Y}}_{j,n}|X_{j,n}, X_{i,U,n}\right) - \frac{3}{2}\log(2\pi e)\right]\\
&\quad + h\left(\overrightarrow{\boldsymbol{Y}}_i|W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_{i,C}\right) + h\left(\overrightarrow{\boldsymbol{Y}}_j|\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) + N\log(2\pi e) + N\delta(N)\\
&\overset{(i)}{\leqslant} h\left(\overrightarrow{\boldsymbol{Y}}_i\right) + \sum_{n=1}^{N}\left[h\left(X_{i,U,n}|X_{j,C,n}\right) + h\left(\overleftarrow{\boldsymbol{Y}}_{j,n}|X_{j,n}, X_{i,U,n}\right) - \frac{3}{2}\log(2\pi e)\right] + h\left(\overrightarrow{\boldsymbol{Y}}_i|\boldsymbol{X}_{i,C}, \boldsymbol{X}_j\right)\\
&\quad + h\left(\overrightarrow{\boldsymbol{Y}}_j|\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) + N\log(2\pi e) + N\delta(N)\\
&\leqslant \sum_{n=1}^{N}\Big[h\left(\overrightarrow{Y}_{i,n}\right) + h\left(X_{i,U,n}|X_{j,C,n}\right) + h\left(\overleftarrow{Y}_{j,n}|X_{j,n}, X_{i,U,n}\right) - \frac{3}{2}\log(2\pi e)\\
&\quad + h\left(\overrightarrow{Y}_{i,n}|X_{i,C,n}, X_{j,n}\right) + h\left(\overrightarrow{Y}_{j,n}|X_{j,C,n}, X_{i,U,n}\right)\Big] + N\log(2\pi e) + N\delta(N)\\
&= N\Big[h\left(\overrightarrow{Y}_{i,k}\right) + h\left(X_{i,U,k}|X_{j,C,k}\right) + h\left(\overleftarrow{Y}_{j,k}|X_{j,k}, X_{i,U,j}\right) - \frac{5}{2}\log(2\pi e) + h\left(\overrightarrow{Y}_{i,k}|X_{i,C,k}, X_{j,k}\right)\\
&\quad + h\left(\overrightarrow{Y}_{j,k}|X_{j,C,k}, X_{i,U,k}\right) + 2\log(2\pi e) + \delta(N)\Big],
\end{aligned}
\tag{G.10}
$$

where, (a) follows from the fact that W_1 and W_2 are mutually independent; (b) follows from Fano's inequality with $\delta : \mathbb{N} \to \mathbb{R}_+$ a positive monotonically decreasing function (Lemma 58) (see Figure G.1c); (c) follows from (2.5) and (G.2); (d) follows from (G.3); (e) follows from (2.1) and the fact that conditioning does not increase entropy (Lemma 40); (f) follows from the fact that $h\left(\overrightarrow{\boldsymbol{Z}}_j|W_j, \overleftarrow{\boldsymbol{Y}}_i, \boldsymbol{X}_i, \boldsymbol{X}_{j,C}\right) - h\left(\overrightarrow{\boldsymbol{Z}}_i, \overrightarrow{\boldsymbol{Z}}_j|W_i, W_j, \overleftarrow{\boldsymbol{Y}}_j, \boldsymbol{X}_i, \boldsymbol{X}_j\right) \leqslant 0$; (g) follows from the fact that $h\left(\overrightarrow{\boldsymbol{Y}}_j\right) - h\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right) + h\left(\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}|\overrightarrow{\boldsymbol{Y}}_j\right) = h\left(\overrightarrow{\boldsymbol{Y}}_j|\boldsymbol{X}_{j,C}, \boldsymbol{X}_{i,U}\right)$; (h) follows from Lemma 21; and (i) follows from the fact that conditioning does not increase entropy (Lemma 40).

From (G.10), the following holds in the asymptotic block-length regime for any $k \in \{1, 2, \ldots N\}$:

$$
\begin{aligned}
2R_i + R_j \leqslant & h\left(\overrightarrow{Y}_{i,k}\right) + h\left(X_{i,U,k}|X_{j,C,k}\right) + h\left(\overleftarrow{Y}_{j,k}|X_{j,k}, X_{i,U,k}\right) + h\left(\overrightarrow{Y}_{i,k}|X_{i,C,k}, X_{j,k}\right) \\
& + h\left(\overrightarrow{Y}_{j,k}|X_{j,C,k}, X_{i,U,k}\right) - \frac{1}{2}\log\left(2\pi e\right) \\
\leqslant & \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_i + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_i \mathrm{INR}_{ij}} + \mathrm{INR}_{ij} + 1\right) - \frac{1}{2}\log\left(\mathrm{INR}_{ij} + 1\right) \\
& + \frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(\overleftarrow{Y}_{j,k}, X_{j,k}, X_{i,U,k}\right)\right)\right) - \frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(X_{j,k}, X_{i,U,k}\right)\right)\right) \\
& + \frac{1}{2}\log\left(1 + \left(1 - \rho^2\right)\left(\overrightarrow{\mathrm{SNR}}_i + \mathrm{INR}_{ji}\right)\right) - \frac{1}{2}\log\left(1 + \left(1 - \rho^2\right)\mathrm{INR}_{ji}\right) \\
& + \frac{1}{2}\log\left(\det\left(\mathrm{Var}\left(\overrightarrow{Y}_{j,k}, X_{j,C,k}, X_{i,U,k}\right)\right)\right) + 2\log\left(2\pi e\right). \qquad \text{(G.11)}
\end{aligned}
$$

The outer bound on (G.11) depends on $S_{1,i}$, $S_{2,i}$, $S_{3,i}$, $S_{4,i}$, and $S_{5,i}$ via the parameter $h_{ji,U}$ in (G.1). Hence, as in the previous part, the following cases are identified:

Case 1: $(S_{1,i} \vee S_{2,i} \vee S_{5,i})$. From (G.1), it follows that $h_{ji,U} = 0$. Therefore, plugging the expressions (G.9) into (G.11) yields (5.7a).

Case 2: $(S_{3,i} \vee S_{4,i})$. From (G.1), it follows that $h_{ji,U} = \sqrt{\frac{\mathrm{INR}_{ij}\mathrm{INR}_{ji}}{\overrightarrow{\mathrm{SNR}}_j}}$. Therefore, plugging the expressions (G.9) into (G.11) yields (5.7b).

This completes the proof of (5.9e) and the proof of Theorem 8.

— H —

Proof of Theorem 9

THIS appendix presents a proof of the Theorem 9. The gap, denoted by δ, between the sets $\overline{\mathcal{C}}$ and $\underline{\mathcal{C}}$ (Definition 6) is approximated as follows:

$$\delta = \max\left(\delta_{R_1}, \delta_{R_2}, \frac{\delta_{2R}}{2}, \frac{\delta_{3R_1}}{3}, \frac{\delta_{3R_2}}{3}\right), \tag{H.1}$$

where,

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

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

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

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

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

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

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

It is important to highlight that, as suggested in [31, 53, 88], the gap between $\underline{\mathcal{C}}$ and $\overline{\mathcal{C}}$ can be calculated more precisely. However, the choice in (H.1) eases the calculations at the expense of less precision. Note also that whether or not the bounds are active (achievable or converse) in either of the equalities in (H.2) depend on the exact values of $\overrightarrow{\mathrm{SNR}}_1$, $\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, INR_{21}, $\overleftarrow{\mathrm{SNR}}_1$, and $\overleftarrow{\mathrm{SNR}}_2$. Hence, a key point in order to find the gap between the achievable region and the converse region is to choose a convenient coding scheme parameters for the achievable region, *i.e.*, the values of ρ, μ_1, and μ_2, according to the definitions in (H.2) for all $i \in \{1,2\}$. These particular coding scheme parameters are chosen such that the expressions in (H.2) become simpler to obtain an upper bound at the expense of a looser outer bound. These particular coding scheme parameters are different for each interference regime. The following describes all the key cases and the corresponding coding scheme parameters.

Case 1: $\mathrm{INR}_{12} > \overrightarrow{\mathrm{SNR}}_1$ and $\mathrm{INR}_{21} > \overrightarrow{\mathrm{SNR}}_2$. This case corresponds to the scenario in which both transmitter-receiver pairs are in HIR. Three subcases follow considering the SNR in the feedback links.

Case 1.1: $\overleftarrow{\mathrm{SNR}}_2 \leqslant \overrightarrow{\mathrm{SNR}}_1$ and $\overleftarrow{\mathrm{SNR}}_1 \leqslant \overrightarrow{\mathrm{SNR}}_2$. In this case the coding scheme has parameters: $\rho = 0$, $\mu_1 = 0$, and $\mu_2 = 0$.

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

Case 1.3: $\overleftarrow{\mathrm{SNR}}_2 \leqslant \overrightarrow{\mathrm{SNR}}_1$ and $\overleftarrow{\mathrm{SNR}}_1 > \overrightarrow{\mathrm{SNR}}_2$. In this case the coding scheme has parameters: $\rho = 0$, $\mu_1 = 0$, and $\mu_2 = 1$.

Case 2: $\mathrm{INR}_{12} \leqslant \overrightarrow{\mathrm{SNR}}_1$ and $\mathrm{INR}_{21} \leqslant \overrightarrow{\mathrm{SNR}}_2$. This case corresponds to the scenario in which both transmitter-receiver pairs are in LIR. There are twelve subcases that must be studied separately.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

1. Calculation of δ_{R_1}.

   From (H.2a) and considering the corresponding coding scheme parameters for the achievable region ($\rho = 0$, $\mu_1 = 0$ and $\mu_2 = 0$), it follows that

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

   where the exact value of ρ' is chosen to provide at least an outer bound for (H.3).

Note that in this case:

$$
\begin{aligned}
\kappa_{1,1}(\rho') &= \frac{1}{2}\log\left(b_{1,1}(\rho')+1\right)\\
&\overset{(a)}{\leqslant} \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}} + \mathrm{INR}_{12}+1\right)\\
&\overset{(b)}{\leqslant} \frac{1}{2}\log\left(2\overrightarrow{\mathrm{SNR}}_1 + 2\mathrm{INR}_{12}+1\right)\\
&\leqslant \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12}+1\right)+\frac{1}{2}, \qquad \text{(H.4a)}\\
\kappa_{2,1}(\rho') &= \frac{1}{2}\log\left(1+b_{4,1}(\rho')+b_{5,2}(\rho')\right)\\
&\leqslant \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{21}+1\right), \qquad \text{(H.4b)}\\
\kappa_{3,1}(\rho') &= \frac{1}{2}\log\left(b_{4,1}(\rho')+1\right)+\frac{1}{2}\log\left(\frac{\overleftarrow{\mathrm{SNR}}_2\left(b_{4,1}(\rho')+b_{5,2}(\rho')+1\right)}{\left(b_{1,2}(1)+1\right)\left(b_{4,1}(\rho')+1\right)}+1\right)\\
&\overset{(c)}{\leqslant} \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1+1\right)+\frac{1}{2}\log\left(\frac{\overleftarrow{\mathrm{SNR}}_2\left(\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{21}+1\right)}{\left(\overrightarrow{\mathrm{SNR}}_2+\mathrm{INR}_{21}+1\right)\left(\overrightarrow{\mathrm{SNR}}_1+1\right)}+1\right)\\
&= \frac{1}{2}\log\left(\frac{\overleftarrow{\mathrm{SNR}}_2\left(\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{21}+1\right)}{\overrightarrow{\mathrm{SNR}}_2+\mathrm{INR}_{21}+1}+\overrightarrow{\mathrm{SNR}}_1+1\right), \qquad \text{(H.4c)}
\end{aligned}
$$

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

$$
\left(\sqrt{\overrightarrow{\mathrm{SNR}}_1}-\sqrt{\mathrm{INR}_{12}}\right)^2 \geqslant 0; \qquad \text{(H.5)}
$$

and (c) follows from the fact that $\kappa_{3,1}(\rho')$ is a monotonically decreasing function of ρ'. Note also that the achievable bound $a_{1,1}+a_{4,2}(0,0)$ admits the following lower-bound:

$$
\begin{aligned}
a_{1,1}+a_{4,2}(0,0) &= \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}}+2\right)+\frac{1}{2}\log\left(\mathrm{INR}_{21}+1\right)-1\\
&\geqslant \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}}+2\right)+\frac{1}{2}\log\left(\mathrm{INR}_{21}\right)-1\\
&= \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1+2\mathrm{INR}_{21}\right)-1\\
&= \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{21}+\mathrm{INR}_{21}\right)-1\\
&\geqslant \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{21}+1\right)-1. \qquad \text{(H.6)}
\end{aligned}
$$

From (H.3), (H.4) and (H.6), assuming that $a_{1,1}+a_{4,2}(0,0) < a_{6,1}(0,0)$, it follows that

$$
\begin{aligned}
\delta_{R_1} &\leqslant \min\left(\kappa_{1,1}(\rho'),\kappa_{2,1}(\rho'),\kappa_{3,1}(\rho)\right)-\left(a_{1,1}+a_{4,2}(0,0)\right)\\
&\leqslant \kappa_{2,1}(\rho')-\left(a_{1,1}+a_{4,2}(0,0)\right)\\
&\leqslant 1. \qquad \text{(H.7)}
\end{aligned}
$$

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

$$\delta_{R_1} \leqslant \min\Big(\kappa_{1,1}(\rho'), \kappa_{2,1}(\rho'), \kappa_{3,1}(\rho)\Big) - a_{6,1}(0,0). \tag{H.8}$$

To calculate an upper-bound for (H.8), the following cases are considered:

<u>Case 1.1.1</u>: $\overrightarrow{\mathrm{SNR}}_1 \geqslant \mathrm{INR}_{21} \wedge \overrightarrow{\mathrm{SNR}}_2 < \mathrm{INR}_{12}$;

<u>Case 1.1.2</u>: $\overrightarrow{\mathrm{SNR}}_1 < \mathrm{INR}_{21} \wedge \overrightarrow{\mathrm{SNR}}_2 \geqslant \mathrm{INR}_{12}$; and

<u>Case 1.1.3</u>: $\overrightarrow{\mathrm{SNR}}_1 < \mathrm{INR}_{21} \wedge \overrightarrow{\mathrm{SNR}}_2 < \mathrm{INR}_{12}$.

In Case 1.1.1, from (H.4) and (H.8), it follows that

$$\begin{aligned}
\delta_{R_1} \leqslant& \kappa_{2,1}(\rho') - a_{6,1}(0,0)\\
\leqslant& \frac{1}{2}\log\Big(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{21} + 1\Big) - \frac{1}{2}\log\Big(\overrightarrow{\mathrm{SNR}}_1 + 2\Big) + \frac{1}{2}\\
\leqslant& \frac{1}{2}\log\Big(\overrightarrow{\mathrm{SNR}}_1 + \overrightarrow{\mathrm{SNR}}_1 + 1\Big) - \frac{1}{2}\log\Big(\overrightarrow{\mathrm{SNR}}_1 + 2\Big) + \frac{1}{2}\\
\leqslant& 1.
\end{aligned} \tag{H.9}$$

In Case 1.1.2, from (H.4) and (H.8), it follows that

$$\begin{aligned}
\delta_{R_1} \leqslant& \kappa_{3,1}(\rho') - a_{6,1}(0,0)\\
\leqslant& \frac{1}{2}\log\left(\frac{\overleftarrow{\mathrm{SNR}}_2\Big(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{21} + 1\Big)}{\overrightarrow{\mathrm{SNR}}_2 + \mathrm{INR}_{21} + 1} + \overrightarrow{\mathrm{SNR}}_1 + 1\right) - \frac{1}{2}\log\Big(\overrightarrow{\mathrm{SNR}}_1 + 2\Big) + \frac{1}{2}\\
\leqslant& \frac{1}{2}\log\Big(\overleftarrow{\mathrm{SNR}}_2 + \overrightarrow{\mathrm{SNR}}_1 + 1\Big) - \frac{1}{2}\log\Big(\overrightarrow{\mathrm{SNR}}_1 + 2\Big) + \frac{1}{2}\\
\leqslant& \frac{1}{2}\log\Big(\overrightarrow{\mathrm{SNR}}_1 + \overrightarrow{\mathrm{SNR}}_1 + 1\Big) - \frac{1}{2}\log\Big(\overrightarrow{\mathrm{SNR}}_1 + 2\Big) + \frac{1}{2}\\
\leqslant& 1.
\end{aligned} \tag{H.10}$$

In Case 1.1.3 two additional cases are considered:

<u>Case 1.1.3.1</u>: $\overrightarrow{\mathrm{SNR}}_1 \geqslant \overrightarrow{\mathrm{SNR}}_2$ and

<u>Case 1.1.3.2</u>: $\overrightarrow{\mathrm{SNR}}_1 < \overrightarrow{\mathrm{SNR}}_2$.

In Case 1.1.3.1, from (H.4) and (H.8), it follows that

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

$$\leqslant \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_1\left(\mathrm{INR}_{21}+\mathrm{INR}_{21}+\mathrm{INR}_{21}\right)}{\mathrm{INR}_{21}\overrightarrow{\mathrm{SNR}}_1}+1\right)+\frac{1}{2}$$
$$=\frac{3}{2}. \tag{H.11}$$

In Case 1.1.3.2, from (H.4) and (H.8), it follows that

$$\begin{aligned}
\delta_{R_1} \leqslant & \kappa_{3,1}(\rho') - a_{6,1}(0,0)\\
\leqslant & \frac{1}{2}\log\left(\frac{\overleftarrow{\mathrm{SNR}}_2\left(\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{21}+1\right)}{\overrightarrow{\mathrm{SNR}}_2+\mathrm{INR}_{21}+1}+\overrightarrow{\mathrm{SNR}}_1+1\right)-\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1+2\right)+\frac{1}{2}\\
\leqslant & \frac{1}{2}\log\left(\overleftarrow{\mathrm{SNR}}_2+\overrightarrow{\mathrm{SNR}}_1+1\right)-\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1+2\right)+\frac{1}{2}\\
\leqslant & \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1+\overrightarrow{\mathrm{SNR}}_1+1\right)-\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1+2\right)+\frac{1}{2}\\
\leqslant & 1.
\end{aligned} \tag{H.12}$$

Then, from (H.7), (H.9), (H.10), (H.11), and (H.12), it follows that in Case 1.1:

$$\delta_{R_1} \leqslant \frac{3}{2}. \tag{H.13}$$

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

$$\delta_{R_2} \leqslant \frac{3}{2}. \tag{H.14}$$

2. Calculation of δ_{2R}. From (H.2c) and considering the corresponding coding scheme parameters for the achievable region ($\rho = 0$, $\mu_1 = 0$ and $\mu_2 = 0$), it follows that

$$\begin{aligned}
\delta_{2R} \leqslant & \min\left(\kappa_4(\rho'),\kappa_5(\rho'),\kappa_6(\rho')\right)-\min\left(a_{2,1}(0)+a_{1,2}, a_{1,1}+a_{2,2}(0), a_{5,1}(0,0)+a_{5,2}(0,0)\right)\\
\leqslant & \min\left(\kappa_4(\rho'),\kappa_5(\rho')\right)-\min\left(a_{2,1}(0)+a_{1,2}, a_{1,1}+a_{2,2}(0), a_{5,1}(0,0)+a_{5,2}(0,0)\right).
\end{aligned} \tag{H.15}$$

Note that

$$\begin{aligned}
\kappa_4(\rho') = & \frac{1}{2}\log\left(1+\frac{b_{4,1}(\rho')}{1+b_{5,2}(\rho')}\right)+\frac{1}{2}\log\left(b_{1,2}(\rho')+1\right)\\
\leqslant & \frac{1}{2}\log\left(1+\frac{b_{4,1}(\rho')}{b_{5,2}(\rho')}\right)+\frac{1}{2}\log\left(b_{1,2}(\rho')+1\right)\\
= & \frac{1}{2}\log\left(1+\frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}}\right)+\frac{1}{2}\log\left(b_{1,2}(\rho')+1\right)\\
\overset{(d)}{\leqslant} & \frac{1}{2}\log\left(1+\frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}}\right)+\frac{1}{2}\log\left(2\overrightarrow{\mathrm{SNR}}_2+2\mathrm{INR}_{21}+1\right)\\
\leqslant & \frac{1}{2}\log\left(1+\frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}}\right)+\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_2+\mathrm{INR}_{21}+1\right)+\frac{1}{2}
\end{aligned}$$

$$\leqslant \frac{1}{2}\log\left(2+\frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}}\right)+\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_2+\mathrm{INR}_{21}+1\right)+\frac{1}{2} \text{ and} \tag{H.16a}$$

$$\begin{aligned}
\kappa_5(\rho')=&\frac{1}{2}\log\left(1+\frac{b_{4,2}(\rho')}{1+b_{5,1}(\rho')}\right)+\frac{1}{2}\log\left(b_{1,1}(\rho')+1\right)\\
\leqslant&\frac{1}{2}\log\left(1+\frac{b_{4,2}(\rho')}{b_{5,1}(\rho')}\right)+\frac{1}{2}\log\left(b_{1,1}(\rho')+1\right)\\
=&\frac{1}{2}\log\left(1+\frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}\right)+\frac{1}{2}\log\left(b_{1,1}(\rho')+1\right)\\
\overset{(e)}{\leqslant}&\frac{1}{2}\log\left(1+\frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}\right)+\frac{1}{2}\log\left(2\overrightarrow{\mathrm{SNR}}_1+2\mathrm{INR}_{12}+1\right)\\
\leqslant&\frac{1}{2}\log\left(1+\frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}\right)+\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{12}+1\right)+\frac{1}{2}\\
\leqslant&\frac{1}{2}\log\left(2+\frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}\right)+\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{12}+1\right)+\frac{1}{2},
\end{aligned} \tag{H.16b}$$

where, (d) follows from the fact that

$$\left(\sqrt{\overrightarrow{\mathrm{SNR}}_2}-\sqrt{\mathrm{INR}_{21}}\right)^2\geqslant 0; \tag{H.17}$$

and (e) follows from the fact that

$$\left(\sqrt{\overrightarrow{\mathrm{SNR}}_1}-\sqrt{\mathrm{INR}_{12}}\right)^2\geqslant 0. \tag{H.18}$$

From (H.15) and (H.16), assuming that $a_{2,1}(0)+a_{1,2}<\min\Big(a_{1,1}+a_{2,2}(0), a_{5,1}(0,0)+a_{5,2}(0,0)\Big)$, it follows that

$$\begin{aligned}
\delta_{2R}\leqslant&\min\left(\kappa_4(\rho'),\kappa_5(\rho')\right)-\left(a_{2,1}(0)+a_{1,2}\right)\\
\leqslant&\kappa_5(\rho')-\left(a_{2,1}(0)+a_{1,2}\right)\\
\leqslant&\frac{1}{2}\log\left(2+\frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}\right)+\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{12}+1\right)+\frac{1}{2}-\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1+\mathrm{INR}_{12}+1\right)\\
&-\frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}+2\right)+1\\
=&\frac{3}{2}.
\end{aligned} \tag{H.19}$$

From (H.15) and (H.16), assuming that $a_{1,1}+a_{2,2}(0)<\min\Big(a_{2,1}(0)+a_{1,2}, a_{5,1}(0,0)+a_{5,2}(0,0)\Big)$, it follows that

$$\delta_{2R}\leqslant\min\left(\kappa_4(\rho'),\kappa_5(\rho')\right)-\left(a_{1,1}+a_{2,2}(0)\right)$$

$$\begin{aligned}
&\leqslant \kappa_4(\rho') - \Big(a_{1,1} + a_{2,2}(0)\Big)\\
&\leqslant \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}}\right) + \frac{1}{2}\log\Big(\overrightarrow{\mathrm{SNR}}_2 + \mathrm{INR}_{21} + 1\Big) + \frac{1}{2} - \frac{1}{2}\log\Big(\overrightarrow{\mathrm{SNR}}_2 + \mathrm{INR}_{21} + 1\Big)\\
&\quad -\frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}} + 2\right) + 1\\
&= \frac{3}{2}. \qquad \text{(H.20)}
\end{aligned}$$

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

$$\delta_{2R} \leqslant \min\Big(\kappa_4(\rho'), \kappa_5(\rho')\Big) - \Big(a_{5,1}(0,0) + a_{5,2}(0,0)\Big). \qquad \text{(H.21)}$$

To calculate an upper-bound for (H.21), the cases 1.1.1 - 1.1.3 defined above are analyzed hereunder.

In Case 1.1.1, $a_{5,1}(0,0) + a_{5,2}(0,0)$ admits the following lower-bound:

$$\begin{aligned}
a_{5,1}(0,0) + a_{5,2}(0,0) &= \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}} + \mathrm{INR}_{12} + 1\right) + \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}} + \mathrm{INR}_{21} + 1\right) - 1\\
&\geqslant \frac{1}{2}\log\left(\mathrm{INR}_{12} + 1\right) - 1. \qquad \text{(H.22)}
\end{aligned}$$

From (H.16), (H.21), and (H.22), it follows that

$$\begin{aligned}
\delta_{2R} &\leqslant \min\Big(\kappa_4(\rho'), \kappa_5(\rho')\Big) - \Big(a_{5,1}(0,0) + a_{5,2}(0,0)\Big)\\
&\leqslant \kappa_5(\rho') - \Big(a_{5,1}(0,0) + a_{5,2}(0,0)\Big)\\
&\leqslant \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}\right) + \frac{1}{2}\log\Big(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1\Big) + \frac{1}{2} - \frac{1}{2}\log\left(\mathrm{INR}_{12} + 1\right) + 1\\
&\leqslant \frac{1}{2}\log\left(2 + 1\right) + \frac{1}{2}\log\left(\mathrm{INR}_{12} + \mathrm{INR}_{12} + 1\right) - \frac{1}{2}\log\left(\mathrm{INR}_{12} + 1\right) + \frac{3}{2}\\
&\leqslant \frac{1}{2}\log\left(3\right) + 2. \qquad \text{(H.23)}
\end{aligned}$$

In Case 1.1.2, $a_{5,1}(0,0) + a_{5,2}(0,0)$ admits the following lower-bound:

$$\begin{aligned}
a_{5,1}(0,0) + a_{5,2}(0,0) &= \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}} + \mathrm{INR}_{12} + 1\right) + \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}} + \mathrm{INR}_{21} + 1\right) - 1\\
&\geqslant \frac{1}{2}\log\left(\mathrm{INR}_{21} + 1\right) - 1. \qquad \text{(H.24)}
\end{aligned}$$

From (H.16), (H.21), and (H.24), it follows that

$$\begin{aligned}
\delta_{2R} \leqslant & \min\Big(\kappa_4(\rho'), \kappa_5(\rho')\Big) - \Big(a_{5,1}(0,0) + a_{5,2}(0,0)\Big) \\
\leqslant & \kappa_4(\rho') - \Big(a_{5,1}(0,0) + a_{5,2}(0,0)\Big) \\
\leqslant & \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}}\right) + \frac{1}{2}\log\Big(\overrightarrow{\mathrm{SNR}}_2 + \mathrm{INR}_{21} + 1\Big) + \frac{1}{2} - \frac{1}{2}\log\left(\mathrm{INR}_{21} + 1\right) + 1 \\
\leqslant & \frac{1}{2}\log\left(2 + 1\right) + \frac{1}{2}\log\left(\mathrm{INR}_{21} + \mathrm{INR}_{21} + 1\right) - \frac{1}{2}\log\left(\mathrm{INR}_{21} + 1\right) + \frac{3}{2} \\
\leqslant & \frac{1}{2}\log\left(3\right) + 2.
\end{aligned} \tag{H.25}$$

In Case 1.1.3, from (H.16), (H.21), and (H.22), it follows that

$$\begin{aligned}
\delta_{2R} \leqslant & \min\Big(\kappa_4(\rho'), \kappa_5(\rho')\Big) - \Big(a_{5,1}(0,0) + a_{5,2}(0,0)\Big) \\
\leqslant & \kappa_5(\rho') - \Big(a_{5,1}(0,0) + a_{5,2}(0,0)\Big) \\
\leqslant & \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}\right) + \frac{1}{2}\log\Big(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1\Big) + \frac{1}{2} - \frac{1}{2}\log\left(\mathrm{INR}_{12} + 1\right) + 1 \\
\leqslant & \frac{1}{2}\log\left(2 + 1\right) + \frac{1}{2}\log\left(\mathrm{INR}_{12} + \mathrm{INR}_{12} + 1\right) \\
& -\frac{1}{2}\log\left(\mathrm{INR}_{12} + 1\right) + \frac{3}{2} \\
\leqslant & \frac{1}{2}\log\left(3\right) + 2.
\end{aligned} \tag{H.26}$$

Then, from (H.19), (H.20), (H.23), (H.25), and (H.26), it follows that in Case 1.1:

$$\delta_{2R} \leqslant 2 + \frac{1}{2}\log\left(3\right). \tag{H.27}$$

3. Calculation of δ_{3R_1}. From (H.2d) and considering the corresponding coding scheme parameters for the achievable region ($\rho = 0$, $\mu_1 = 0$ and $\mu_2 = 0$), it follows that

$$\delta_{3R_1} \leqslant \kappa_{7,1}(\rho') - \Big(a_{1,1} + a_{7,1}(0,0,0) + a_{5,2}(0,0)\Big). \tag{H.28}$$

The sum $a_{1,1} + a_{7,1}(0,0,0) + a_{5,2}(0,0)$ admits the following lower-bound:

$$\begin{aligned}
a_{1,1} + a_{7,1}(0,0,0) + a_{5,2}(0,0) = & \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}} + 2\right) + \frac{1}{2}\log\Big(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1\Big) \\
& + \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}} + \mathrm{INR}_{21} + 1\right) - \frac{3}{2} \\
\geqslant & \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}} + 2\right) + \frac{1}{2}\log\Big(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1\Big) \\
& + \frac{1}{2}\log\left(\mathrm{INR}_{21} + 1\right) - \frac{3}{2}.
\end{aligned} \tag{H.29}$$

If the term $\kappa_{7,1}(\rho')$ is active in the converse region, it admits the sum $\kappa_{1,1}(\rho') + \kappa_4(\rho')$ as an upper bound, which corresponds to the sum of the single rate and sum-rate outer bounds respectively. The following upper bound holds on this sum:

$$\begin{aligned}
\kappa_{7,1}(\rho') \leqslant & \kappa_{1,1}(\rho') + \kappa_4(\rho') \\
\leqslant & \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1\right) + \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}}\right) + \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_2 + \mathrm{INR}_{21} + 1\right) + 1 \\
\leqslant & \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1\right) + \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}}\right) + \frac{1}{2}\log\left(\mathrm{INR}_{21} + \mathrm{INR}_{21} + 1\right) + 1 \\
\leqslant & \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1\right) + \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}}\right) + \frac{1}{2}\log\left(\mathrm{INR}_{21} + 1\right) + \frac{3}{2}. \qquad \text{(H.30)}
\end{aligned}$$

From (H.28), (H.29) and (H.30), it follows that in Case 1.1:

$$\begin{aligned}
\delta_{3R_1} \leqslant & \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1\right) + \frac{1}{2}\log\left(2 + \frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}}\right) + \frac{1}{2}\log\left(\mathrm{INR}_{21} + 1\right) + \frac{3}{2} \\
& - \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_1}{\mathrm{INR}_{21}} + 2\right) - \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1\right) - \frac{1}{2}\log\left(\mathrm{INR}_{21} + 1\right) + \frac{3}{2} \\
= & 3. \qquad \text{(H.31)}
\end{aligned}$$

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

$$\delta_{3R_2} \leqslant 3. \qquad \text{(H.32)}$$

Therefore, in Case 1.1, from (H.1), (H.14), (H.13), (H.27), (H.31) and (H.32) it follows that

$$\delta = \max\left(\delta_{R_1}, \delta_{R_2}, \frac{\delta_{2R}}{2}, \frac{\delta_{3R_1}}{3}, \frac{\delta_{3R_2}}{3}\right) \leqslant \frac{3}{2}. \qquad \text{(H.33)}$$

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

— I —

Proof of Theorem 10

To prove Theorem 10, the first step is to show that a rate pair $(R_1, R_2) \in \mathbb{R}_+^2$, with $R_i < L_i$ or $R_i > U_i$ for at least one $i \in \{1, 2\}$, is not achievable at an η-NE for all $\eta > 0$. That is,

$$\mathcal{N}_\eta \subseteq \mathcal{C} \cap \mathcal{B}_\eta. \tag{I.1}$$

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

$$\mathcal{N}_\eta \supseteq \mathcal{C} \cap \mathcal{B}_\eta. \tag{I.2}$$

This proves Theorem 10.

Proof of (I.1): The proof of (I.1) is completed by the following lemmas:

Lemma 22. *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-NE for all $\eta > 0$.*

Proof: Let $(s_1^*, s_2^*) \in \mathcal{A}_1 \times \mathcal{A}_2$ be an η-NE transmit-receive configuration pair such that $u_1(s_1^*, s_2^*) = R_1$ and $u_2(s_1^*, s_2^*) = R_2$, respectively. Assume, without loss of generality, that $R_1 < L_1$. 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.*, $R_1(s_1', s_2^*) - R_1 > \eta$, 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-NE, which contradicts the initial assumption. This proves that if (s_1^*, s_2^*) is an η-NE, then $R_1 \geqslant L_1$ and $R_2 \geqslant L_2$. This completes the proof. ∎

Lemma 23. *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-NE for all $\eta > 0$.*

Proof: Let $(s_1^*, s_2^*) \in \mathcal{A}_1 \times \mathcal{A}_2$ be an η-NE transmit-receive configuration pair such that $u_1(s_1^*, s_2^*) = R_1$ and $u_2(s_1^*, s_2^*) = R_2$, respectively. Hence, the following holds for transmitter-receiver i:

$$\begin{aligned} N\,R_i &= H(W_i) \\ &\overset{(a)}{=} H(W_i|\Omega_i) \\ &\overset{(b)}{\leqslant} I\left(W_i; \overrightarrow{\boldsymbol{Y}}_i|\Omega_i\right) + N\delta_i(N), \end{aligned} \tag{I.3}$$

where, (a) follows from the independence between the indices W_i and Ω_i; and (b) follows from Fano's inequality, since the rate R_i is achievable from the assumptions of the lemma, with $\delta_i : \mathbb{N} \to \mathbb{R}_+$ a positive monotonically decreasing function for all $i \in \{1, 2\}$ (Lemma 58). In particular, for transmitter-receiver pair 1 in (I.3), the following holds:

$$N\,R_1 \overset{(c)}{\leqslant} N\,\max(\overrightarrow{n}_{11}, n_{12}) - \sum_{n=1}^{N} H\left(\overrightarrow{\boldsymbol{Y}}_{1,n}|\Omega_1, W_1, \overrightarrow{\boldsymbol{Y}}_{1,(1:n-1)}\right) + N\delta_1(N), \tag{I.4}$$

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 encoding function in (2.1). Moreover, for all $n \in \{1, 2, \ldots, N\}$, the channel input $\boldsymbol{X}_{i,n}$ can be written as

$$\boldsymbol{X}_{i,n} = \left(\boldsymbol{X}_{i,C,n}, \boldsymbol{X}_{i,D,n}, \boldsymbol{X}_{i,P,n}, \boldsymbol{X}_{i,Q,n}\right), \tag{I.5}$$

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.*,

$$\dim \boldsymbol{X}_{i,C,n} = \min(\overrightarrow{n}_{ii}, n_{ji}); \tag{I.6}$$

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

$$\dim \boldsymbol{X}_{i,P,n} = (\overrightarrow{n}_{ii} - n_{ji})^+; \tag{I.7}$$

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

$$\dim \boldsymbol{X}_{i,D,n} = (n_{ji} - \overrightarrow{n}_{ii})^+; \tag{I.8}$$

finally, $\boldsymbol{X}_{i,Q,n} = (0, \ldots, 0)^{\mathsf{T}}$ is included for dimensional matching of the model in (2.32), *i.e.*,

$$\dim \boldsymbol{X}_{i,Q,n} = q - \max(\overrightarrow{n}_{ii}, n_{ji}). \tag{I.9}$$

Using this notation, the following holds from (I.4):

$$R_1 \leqslant \max(\overrightarrow{n}_{11}, n_{12}) - \frac{1}{N}\sum_{n=1}^{N} H\left(\boldsymbol{X}_{2,C,n}, \boldsymbol{X}_{2,D,n}|\Omega_1, W_1, \overrightarrow{\boldsymbol{Y}}_{1,(1:n-1)}\right) + \delta_1(N) \tag{I.10}$$

$$= \max(\overrightarrow{n}_{11}, n_{12}) - H\left(\widetilde{\boldsymbol{X}}_{2,C,n}, \widetilde{\boldsymbol{X}}_{2,D,n}\right) + \delta_1(N), \text{for any } n \in \{1, 2, \ldots, N\} \tag{I.11}$$

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 inequality in (I.11) follows from the signal construction in (2.31).

For all $i \in \{1, 2\}$, let $\boldsymbol{X}_{i,C,n} = \left(\boldsymbol{X}_{i,C1,n}^{\mathsf{T}}, \boldsymbol{X}_{i,C2,n}^{\mathsf{T}}\right)^{\mathsf{T}}$ be such that $\boldsymbol{X}_{i,C1,n}$ satisfies:

$$\dim \boldsymbol{X}_{i,C1,n} = \left(\min\left((\overrightarrow{n}_{ii} - n_{ij})^+, n_{ji}\right) - \left(\min\left((\overrightarrow{n}_{ii} - n_{ji})^+, n_{ij}\right) - (\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii})^+\right)^+\right)^+. \tag{I.12}$$

The dimension of $\boldsymbol{X}_{i,C1,n}$ is chosen as the non-negative difference between 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.*, $\left(\min\left((\overrightarrow{n}_{ii} - n_{ji})^+, n_{ij}\right) - (\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii})^+\right)^+$. 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,

$$\dim \boldsymbol{X}_{i,C2,n} = \min(\overrightarrow{n}_{ii}, n_{ji}) - \dim \boldsymbol{X}_{i,C1,n}. \tag{I.13}$$

For $i = 2$, $\boldsymbol{X}_{i,C1,n}$ satisfies:

$$\dim \boldsymbol{X}_{2,C1,n} = \left(\min\left((\overrightarrow{n}_{22} - n_{21})^+, n_{12}\right) - \left(\min\left((\overrightarrow{n}_{22} - n_{12})^+, n_{21}\right) - (\max(\overrightarrow{n}_{22}, n_{21}) - \overleftarrow{n}_{22})^+\right)^+\right)^+. \tag{I.14}$$

Given a fixed 5-tuple $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{22})$, consider the following events (Boolean variables):

$$C_{1,2}: \quad n_{21} < \overrightarrow{n}_{22} \leqslant n_{12}, \tag{I.15a}$$
$$C_{2,2}: \quad \max(n_{12}, n_{21}, \overleftarrow{n}_{22}) < \overrightarrow{n}_{22} < n_{12} + n_{21}, \tag{I.15b}$$
$$C_{3,2}: \quad \overleftarrow{n}_{22} \leqslant n_{12}, \tag{I.15c}$$
$$C_{4,2}: \quad n_{21} < n_{12} < \overrightarrow{n}_{22} \leqslant \overleftarrow{n}_{22}, \tag{I.15d}$$
$$C_{5,2}: \quad \overrightarrow{n}_{22} > \max(n_{12}, n_{21}, \overleftarrow{n}_{22}), \tag{I.15e}$$
$$C_{6,2}: \quad \overrightarrow{n}_{22} \geqslant \max(n_{12} + n_{21}, \overleftarrow{n}_{22} + n_{21}), \tag{I.15f}$$
$$C_{7,2}: \quad \max(n_{12}, n_{21}, \overleftarrow{n}_{22}, \overleftarrow{n}_{22} + n_{21} - n_{12}) < \overrightarrow{n}_{22} < \overleftarrow{n}_{22} + n_{21} \leqslant \overleftarrow{n}_{22} + \overrightarrow{n}_{22} - n_{12}, \tag{I.15g}$$
$$C_{8,2}: \quad \max(n_{12}, n_{21}, \overleftarrow{n}_{22}, \overleftarrow{n}_{22} + n_{21} - n_{12}) < \overrightarrow{n}_{22} < n_{12} + n_{21} < \overleftarrow{n}_{22} + n_{21}. \tag{I.15h}$$

Then, the following holds:

$$\dim \boldsymbol{X}_{2,C1,n} = \begin{cases} \overrightarrow{n}_{22} - n_{21} & \text{if} \quad C_{1,2} \vee \left(C_{2,2} \wedge C_{3,2}\right) \text{ holds true} \\ n_{12} - n_{21} & \text{if} \quad C_{4,2} \text{ holds true} \\ n_{12} & \text{if} \quad C_{5,2} \wedge C_{6,2} \text{ holds true} \\ \overrightarrow{n}_{22} + n_{12} - \overleftarrow{n}_{22} - n_{21} & \text{if} \quad C_{7,2} \vee C_{8,2} \text{ holds true} \\ 0 & \qquad \text{otherwise} \end{cases}. \tag{I.16}$$

The following step establishes a lower-bound on $H\left(\widetilde{\boldsymbol{X}}_{2,C,n}, \widetilde{\boldsymbol{X}}_{2,D,n}\right)$ at an η-NE. From (I.3), the following inequality holds for transmitter-receiver pair 2:

$$\begin{aligned}
N\,R_2 \leqslant & I\left(W_2; \overrightarrow{\boldsymbol{Y}}_2|\Omega_2\right) + N\delta_2(N) \\
= & \sum_{n=1}^{N}\left(H\left(\overrightarrow{\boldsymbol{Y}}_{2,n}|\Omega_1, \overrightarrow{\boldsymbol{Y}}_{2,(1:n-1)}\right) - H\left(\overrightarrow{\boldsymbol{Y}}_{2,n}|W_2, \Omega_1, \overrightarrow{\boldsymbol{Y}}_{2,(1:n-1)}\right)\right) + N\delta_2(N) \\
\leqslant & \sum_{n=1}^{N}\left(H\left(\overrightarrow{\boldsymbol{Y}}_{2,n}\right) - H\left(\overrightarrow{\boldsymbol{Y}}_{2,n}|W_2, \Omega_2, \overrightarrow{\boldsymbol{Y}}_{2,(1:n-1)}\right)\right) + N\delta_2(N) \\
\overset{(d)}{=} & \sum_{n=1}^{N}\left(H\left(\overrightarrow{\boldsymbol{Y}}_{2,n}\right) - H\left(\overrightarrow{\boldsymbol{Y}}_{2,n}|\boldsymbol{X}_{2,n}\right)\right) + N\delta_2(N) \\
= & \sum_{n=1}^{N} I\left(\overrightarrow{\boldsymbol{Y}}_{2,n}; \boldsymbol{X}_{2,n}\right) + N\delta_2(N) \\
= & \sum_{n=1}^{N} I\left(\overrightarrow{\boldsymbol{Y}}_{2,n}; \boldsymbol{X}_{2,C1,n}, \boldsymbol{X}_{2,C2,n}, \boldsymbol{X}_{2,P,n}, \boldsymbol{X}_{2,D,n}\right) + N\delta_2(N) \\
= & \sum_{n=1}^{N}\Bigg(I\left(\overrightarrow{\boldsymbol{Y}}_{2,n}; \boldsymbol{X}_{2,C2,n}, \boldsymbol{X}_{2,P,n}, \boldsymbol{X}_{2,D,n}\right) \\
& +I\left(\overrightarrow{\boldsymbol{Y}}_{2,n}; \boldsymbol{X}_{2,C1,n}|\boldsymbol{X}_{2,C2,n}, \boldsymbol{X}_{2,P,n}, \boldsymbol{X}_{2,D,n}\right)\Bigg) + N\delta_2(N) \\
= & \sum_{n=1}^{N}\left(I\left(\overrightarrow{\boldsymbol{Y}}_{2,n}; \boldsymbol{X}_{2,C2,n}, \boldsymbol{X}_{2,P,n}, \boldsymbol{X}_{2,D,n}\right) + H\left(\boldsymbol{X}_{2,C1,n}|\boldsymbol{X}_{2,C2,n}, \boldsymbol{X}_{2,P,n}, \boldsymbol{X}_{2,D,n}\right)\right) \\
& +N\delta_2(N) \\
= & \sum_{n=1}^{N}\left(I\left(\overrightarrow{\boldsymbol{Y}}_{2,n}; \boldsymbol{X}_{2,C2,n}, \boldsymbol{X}_{2,P,n}, \boldsymbol{X}_{2,D,n}\right) + H\left(\boldsymbol{X}_{2,C1,n}\right)\right) + N\delta_2(N),
\end{aligned} \tag{I.17}$$

where, (d) follows from the fact that $\boldsymbol{X}_{i,n} = f_{i,n}^{(N)}\left(W_i, \Omega_i, \overleftarrow{\boldsymbol{Y}}_{i,(1:n-1)}\right)$ from the definition of the encoding function and $\left(W_i, \Omega_i, \overleftarrow{\boldsymbol{Y}}_{i,(1:n-1)}\right) \rightarrow X_{i,n} \rightarrow \overrightarrow{Y}_{i,n}$.

Let $\varphi : \mathrm{N} \rightarrow \mathrm{R}^+$ be a monotonically decreasing function such that for all $n \in \{1, 2, \ldots, N\}$ in (I.17), the following holds:

$$R_2 = I\left(\overrightarrow{\boldsymbol{Y}}_{2,n}; \boldsymbol{X}_{2,C2,n}, \boldsymbol{X}_{2,P,n}, \boldsymbol{X}_{2,D,n}\right) + H\left(\boldsymbol{X}_{2,C1,n}\right) + \varphi(N). \tag{I.18}$$

Assume now that there exists another transmit-receive configuration for transmitter-receiver 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 $\boldsymbol{X}_{2,C2,n}$, $\boldsymbol{X}_{2,P,n}$, and $\boldsymbol{X}_{2,D,n}$ 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 symbol previously 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.*,

$$R_2(s^*_1, s'_2)\leqslant I\left(\boldsymbol{X}_{2,C2,n}, \boldsymbol{X}_{2,P,n}; \overrightarrow{\boldsymbol{Y}}_{2,n}\right) + \dim \boldsymbol{X}_{2,C1,n} + \delta'_2(N), \tag{I.19}$$

with $\delta'_2 : \mathbb{N} \to \mathbb{R}_+$ a positive monotonically decreasing function. From Definition 4, it follows that $R_2(s^*_1, s'_2) - R_2 \leqslant \eta$. Hence, from (I.18) and (I.19), it follows that

$$H\left(\boldsymbol{X}_{2,C1,n}\right)\geqslant\dim \boldsymbol{X}_{2,C1,n} - \eta - \varphi(N) + \delta'_2(N). \tag{I.20}$$

Note that for a finite N, (I.20) can be satisfied only if $\eta > -\varphi(N) + \delta'_2(N)$. This suggests that in the asymptotic block-length regime and given $\eta > 0$ arbitrarily small at an η-NE, the bits $\boldsymbol{X}_{2,C1,n}$ are used at maximum entropy. Note also that from the definition of $\widetilde{\boldsymbol{X}}_{2,C,n}$, it follows that the bits $\boldsymbol{X}_{2,C1,n}$ are contained into $\widetilde{\boldsymbol{X}}_{2,C,n}$ and thus, it follows that

$$\begin{aligned} H\left(\widetilde{\boldsymbol{X}}_{2,C,n}, \widetilde{\boldsymbol{X}}_{2,D,n}\right)&\geqslant H\left(\widetilde{\boldsymbol{X}}_{2,C,n}\right) \\ &\geqslant H\left(\boldsymbol{X}_{2,C1,n}\right) \\ &\geqslant\dim \boldsymbol{X}_{2,C1,n} - \eta - \varphi(N) + \delta'_2(N). \end{aligned} \tag{I.21}$$

Plugging (I.12) and (I.21) into (I.11), it follows that at an η-NE,

$$\begin{aligned} R_1\leqslant&\max\left(\overrightarrow{n}_{11}, n_{12}\right) \\ &-\left(\min\left((\overrightarrow{n}_{22}-n_{21})^+,n_{12}\right)-\left(\min\left((\overrightarrow{n}_{22}-n_{12})^+,n_{21}\right)-(\max\left(\overrightarrow{n}_{22},n_{21}\right)-\overleftarrow{n}_{22})^+\right)^+\right)^+ \\ &+\eta + \varphi(N) - \delta'_2(N), \end{aligned} \tag{I.22}$$

which proves, in the asymptotic block-length regime and given $\eta > 0$ arbitrarily small, that

$$\begin{aligned} U_1\leqslant&\max\left(\overrightarrow{n}_{11}, n_{12}\right) \\ &-\left(\min\left((\overrightarrow{n}_{22}-n_{21})^+,n_{12}\right)-\left(\min\left((\overrightarrow{n}_{22}-n_{12})^+,n_{21}\right)-(\max\left(\overrightarrow{n}_{22},n_{21}\right)-\overleftarrow{n}_{22})^+\right)^+\right)^+ + \eta, \end{aligned}$$

and this completes the proof of Lemma 23. ■

Proof of (I.2)**:** To continue with the second part of the proof of Theorem 10, consider a modification of the coding scheme with noisy feedback presented in the centralized part (Part II). The novelty consists in allowing users to introduce common randomness as suggested in [16, 66].

Consider without any loss of generality that $N = N_1 = N_2$. Let $W_i^{(t)} \in \{1, 2, \ldots, 2^{NR_i}\}$ and $\Omega_i^{(t)} \in \{1, 2, \ldots, 2^{NR_{i,R}}\}$ denote the message index and the random message index sent by transmitter i during the t-th block, with $t \in \{1, 2, \ldots, T\}$, respectively. 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 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 $2^{N\left(R_{1,C1}+R_{2,C1}+R_{1,R1}+R_{2,R1}\right)}$ 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 used by transmitter i is selected using $\left(W_{i,C1}^{(t)}, \Omega_{i,R1}^{(t)}\right)$ from the second code-layer, which is a sub-codebook of $2^{N\left(R_{i,C1}+R_{i,R1}\right)}$ 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 used by transmitter i is selected using $\left(W_{i,C2}^{(t)}, \Omega_{i,R2}^{(t)}\right)$ from the third code-layer, which is a sub-codebook of $2^{N\left(R_{i,C2}+R_{i,R2}\right)}$ 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 used by transmitter i is selected 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 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 M.

The proof of (I.2) uses the following results:

Lemma 24 proves that the RHK-NOF achieves all the rate pairs $(R_1, R_2) \in \mathcal{C}$; Lemma 25 provides the maximum rate improvement that a transmitter-receiver pair can obtain when it deviates from the RHK-NOF coding scheme; Lemma 26 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 for all $\eta > 0$; and Lemma 27 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 (I.2).

Lemma 24. *The achievable region of the randomized Han-Kobayashi coding scheme for the two-user D-LDIC-NOF is the set of rates* $\left(R_{1,C1},\ R_{1,R1},\ R_{1,C2},\ R_{1,R2},\ R_{1,P},\ R_{2,C1},\ R_{2,R1},\ R_{2,C2},\ R_{2,R2},\ R_{2,P}\right) \in \mathbb{R}_+^{10}$ *that satisfy, for all* $i \in \{1,2\}$ *and* $j \in \{1,2\} \setminus \{i\}$*, the following conditions:*

$$R_{j,C1} + R_{j,R1} \leqslant \theta_{1,i}, \tag{I.23a}$$
$$R_i + R_{j,C} + R_{j,R} \leqslant \theta_{2,i}, \tag{I.23b}$$
$$R_{j,C2} + R_{j,R2} \leqslant \theta_{3,i}, \tag{I.23c}$$
$$R_{i,P} \leqslant \theta_{4,i}, \tag{I.23d}$$
$$R_{i,P} + R_{j,C2} + R_{j,R2} \leqslant \theta_{5,i}, \tag{I.23e}$$
$$R_{i,C2} + R_{i,P} \leqslant \theta_{6,i}, \text{ and} \tag{I.23f}$$
$$R_{i,C2} + R_{i,P} + R_{j,C2} + R_{j,R2} \leqslant \theta_{7,i}, \tag{I.23g}$$

where,

$$\theta_{1,i} = \left(n_{ij} - \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^+\right)^+, \tag{I.24a}$$
$$\theta_{2,i} = \max\left(\overrightarrow{n}_{ii}, n_{ij}\right), \tag{I.24b}$$
$$\theta_{3,i} = \min\left(n_{ij}, \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^+\right), \tag{I.24c}$$
$$\theta_{4,i} = \left(\overrightarrow{n}_{ii} - n_{ji}\right)^+, \tag{I.24d}$$
$$\theta_{5,i} = \max\left(\left(\overrightarrow{n}_{ii} - n_{ji}\right)^+, \min\left(n_{ij}, \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^+\right) \right), \tag{I.24e}$$
$$\begin{aligned}\theta_{6,i} = {} & \min\left(n_{ji}, \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+\right) - \min\left(\left(n_{ji} - \overrightarrow{n}_{ii}\right)^+, \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+ \right) \\ & + \left(\overrightarrow{n}_{ii} - n_{ji}\right)^+, \text{ and}\end{aligned} \tag{I.24f}$$
$$\begin{aligned}\theta_{7,i} = \max\Big(& \min\left(n_{ij}, \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^+ \right), \min\left(n_{ji}, \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+ \right) \\ & - \min\left(\left(n_{ji} - \overrightarrow{n}_{ii}\right)^+, \left(\max\left(\overrightarrow{n}_{jj}, n_{ji}\right) - \overleftarrow{n}_{jj}\right)^+ \right) + \left(\overrightarrow{n}_{ii} - n_{ji}\right)^+ \Big).\end{aligned} \tag{I.24g}$$

Proof: The proof of Lemma 24 is presented in Appendix M. ■

The set of inequalities in (I.23) 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 (I.23) in terms of R_1 and R_2, corresponds to the region $\mathcal{C}$ (Theorem 1). Therefore, the relevance of Lemma 24 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 (I.23).

The following lemma shows that when both transmitter-receiver links use the RHK-NOF and one of them unilaterally changes its coding scheme, his rate improvement is always limited.

Lemma 25. *Let the rate* 6*-tuple* $\boldsymbol{R} = (R_{1,C}, R_{1,R}, R_{1,P}, R_{2,C}, R_{2,R}, R_{2,P})$ *be achievable with the RHK-NOF such that* $R_1 = R_{1,P} + R_{1,C}$ *and* $R_2 = R_{2,P} + R_{2,C}$*. Let* $i \in \{1,2\}$ *and* $j \in \{1,2\} \setminus \{i\}$*. Then, any unilateral deviation of transmitter-receiver pair* i *by using any* *other coding scheme leads to a transmission rate R_i' that satisfies:*

$$R_i' \leqslant \max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \left(R_{j,C} + R_{j,R}\right). \tag{I.25}$$

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 random symbols and it might possibly have a different block-length $N_1' \neq N_1$. Let $\overrightarrow{\boldsymbol{Y}}'_1 = \left(\overrightarrow{\boldsymbol{Y}}'^{\mathsf{T}}_{1,1}, \overrightarrow{\boldsymbol{Y}}'^{\mathsf{T}}_{1,2}, \ldots, \overrightarrow{\boldsymbol{Y}}'^{\mathsf{T}}_{1,N}\right)^{\mathsf{T}}$ be the super vector of channel outputs at receiver 1 during $N = \max(N_1', N_2)$ consecutive channel uses in the model in (2.31). Hence, an upper-bound on R_1' is obtained from the following inequalities:

$$\begin{aligned}
NR_1' &= H\left(W_1\right)\\
&= H\left(W_1 | \Omega_1\right)\\
&= I\left(W_1; \overrightarrow{\boldsymbol{Y}}'_1 | \Omega_1\right) + H\left(W_1 | \overrightarrow{\boldsymbol{Y}}'_1, \Omega_1\right)\\
&\overset{(a)}{\leqslant} I\left(W_1; \overrightarrow{\boldsymbol{Y}}'_1 | \Omega_1\right) + N\delta_1(N)\\
&= H\left(\overrightarrow{\boldsymbol{Y}}'_1 | \Omega_1\right) - H\left(\overrightarrow{\boldsymbol{Y}}'_1 | W_1, \Omega_1\right) + N\delta_1(N)\\
&\overset{(b)}{\leqslant} N \max\left(\overrightarrow{n}_{11}, n_{12}\right) - H\left(\overrightarrow{\boldsymbol{Y}}'_1 | W_1, \Omega_1\right) + N\delta_1(N),
\end{aligned} \tag{I.26}$$

where, (a) follows from Fano's inequality, since the rate R_1' is achievable as the indice W_1 can be reliably decoded by receiver 1 using the signals $\overrightarrow{\boldsymbol{Y}}'_1$ and Ω_1 from the assumptions of the lemma, with $\delta_1 : \mathbb{N} \rightarrow \mathbb{R}_+$ a positive monotonically decreasing function (Lemma 58); and (b) follows from $H(\overrightarrow{\boldsymbol{Y}}'_1 | \Omega_1) \leqslant N \dim \overrightarrow{\boldsymbol{Y}}'_{1,n} = N \max\left(\overrightarrow{n}_{11}, n_{12}\right)$, for all $n \in \{1, 2, \ldots, N\}$. To refine this upper bound, note that the term $H\left(\overrightarrow{\boldsymbol{Y}}'_1 | W_1, \Omega_1\right)$ in (I.26) satisfies the following lower bound:

$$\begin{aligned}
N\left(R_{2,C} + R_{2,R}\right) &= H(W_{2,C}, \Omega_2)\\
&\overset{(c)}{=} H(W_{2,C}, \Omega_2 | W_1, \Omega_1)\\
&= I(W_{2,C}, \Omega_2; \overrightarrow{\boldsymbol{Y}}'_1 | W_1, \Omega_1) + H(W_{2,C}, \Omega_2 | W_1, \Omega_1, \overrightarrow{\boldsymbol{Y}}'_1)\\
&\overset{(d)}{\leqslant} I(W_{2,C}, \Omega_2; \overrightarrow{\boldsymbol{Y}}'_1 | W_1, \Omega_1) + N\delta_2(N)\\
&= H(\overrightarrow{\boldsymbol{Y}}'_1 | W_1, \Omega_1) - H(\overrightarrow{\boldsymbol{Y}}'_1 | W_1, \Omega_1, W_{2,C}, \Omega_2) + N\delta_2(N)\\
&\leqslant H(\overrightarrow{\boldsymbol{Y}}'_1 | W_1, \Omega_1) + N\delta_2(N),
\end{aligned} \tag{I.27}$$

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, since the indices $W_{2,C}$ and Ω_2 can be reliably decoded by receiver 1 using the signals $\overrightarrow{\boldsymbol{Y}}'_1$ from the assumptions of the lemma with $\delta_2 : \mathbb{N} \rightarrow \mathbb{R}_+$ a positive monotonically decreasing function (Lemma 58). Hence, it follows from (I.27) that

$$H(\overrightarrow{\boldsymbol{Y}}'_1 | W_1, \Omega_1) \geqslant N\left(R_{2,C} + R_{2,R}\right) - N\delta_2(N). \tag{I.28}$$

Finally, plugging (I.28) into (I.26) yields, in the asymptotic block-length regime, the following upper-bound:

$$R_1' \leqslant \max\left(\overrightarrow{n}_{11}, n_{12}\right) - (R_{2,C} + R_{2,R}). \tag{I.29}$$

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

Lemma 25 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 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 6-tuple $\boldsymbol{R} = (R_{1,C}, R_{1,R}, R_{1,P}, R_{2,C}, R_{2,R}, R_{2,P})$, with $R_i = R_{i,P} + R_{i,C}$. Denote by $R'_{i,\max} = \max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - (R_{j,C} + R_{j,R})$ 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 26. *Let $\eta > 0$ be fixed and let the rate 6-tuple $\boldsymbol{R} = (R_{1,C}, R_{1,R}, R_{1,P}, R_{2,C}, R_{2,R}, R_{2,P})$ be achievable with the RHK-NOF and satisfy for all $i \in \{1,2\}$,*

$$R_{i,C} + R_{i,P} + R_{j,C} + R_{j,R} = \max(\overrightarrow{n}_{ii}, n_{ij}) - \eta. \tag{I.30}$$

Then, the rate pair $(R_1, R_2) \in \mathbb{R}_+^2$, with $R_i = R_{i,C} + R_{i,P}$ is achievable at an η-NE.

Proof: Let $(s_1^*, s_2^*) \in \mathcal{A}_1 \times \mathcal{A}_2$ be a transmit-receive configuration pair, in which the configuration s_i^* is a RHK-NOF satisfying condition (I.30). From the assumptions of the lemma, it follows that (s_1^*, s_2^*) is an η-NE at which $u_1(s_1^*, s_2^*) = R_{1,C} + R_{1,P}$ and $u_2(s_1^*, s_2^*) = R_{2,C} + R_{2,P}$. Consider that such a transmit-receive configuration pair (s_1^*, s_2^*) is not an η-NE. Then, from Definition 4, there exists at least one $i \in \{1,2\}$ and at least one configuration $s_i \in \mathcal{A}_i$ 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,

$$u_1(s_1, s_2^*) = R_1' \geqslant u_1(s_1^*, s_2^*) + \eta = R_{1,C} + R_{1,P} + \eta. \tag{I.31}$$

However, from Lemma 25, it follows that

$$R_1' \leqslant \max\left(\overrightarrow{n}_{11}, n_{12}\right) - (R_{2,C} + R_{2,R}), \tag{I.32}$$

and from the assumption in (I.30), with $i = 1$, the following holds:

$$R_{2,C} + R_{2,R} = \max(\overrightarrow{n}_{11}, n_{12}) - (R_{1,C} + R_{1,P}) - \eta, \tag{I.33}$$

it follows that

$$R_1' \leqslant R_{1,C} + R_{1,P} + \eta. \tag{I.34}$$

The result in (I.34) contradicts condition (I.31) for any $\eta > 0$ and shows that there exists no other coding scheme that brings an individual utility improvement greater 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 $(R_1, R_2) \in \mathcal{C} \cap \mathcal{B}_\eta$ are achievable by the RHK-NOF coding scheme at an η-NE, for all $\eta > 0$.

Lemma 27. *Let $\eta > 0$ be fixed. 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 26, it follows that the configuration pair (s_1^*, s_2^*) in which each player's transmit-receive configuration is the RHK-NOF satisfying condition (I.30) is an η-NE. Thus, from the conditions in (I.23) and (I.30), the following holds:

$$\begin{aligned}
R_{j,C1} + R_{j,R1} &\leqslant \theta_{1,i}, \\
R_i + R_{j,C} + R_{j,R} &\leqslant \theta_{2,i}, \\
R_i + R_{j,C} + R_{j,R} &\geqslant \theta_{2,i} - \eta, \\
R_{j,C2} + R_{j,R2} &\leqslant \theta_{3,i}, \\
R_{i,P} &\leqslant \theta_{4,i}, \\
R_{i,P} + R_{j,C2} + R_{j,R2} &\leqslant \theta_{5,i}, \\
R_{i,C2} + R_{i,P} &\leqslant \theta_{6,i}, \text{ and} \\
R_{i,C2} + R_{i,P} + R_{j,C2} + R_{j,R2} &\leqslant \theta_{7,i}.
\end{aligned} \tag{I.35}$$

The region characterized by (I.35) 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:

$$\begin{aligned}
R_1 &\geqslant (\theta_{2,1} - \theta_{1,1} - \theta_{3,1} - \eta)^+, \\
R_1 &\leqslant \min\left(\theta_{6,1} + \theta_{1,2}, \theta_{2,1} + \theta_{1,2} + \theta_{5,2} - \theta_{2,2} + \eta, \theta_{2,1}\right), \\
R_2 &\geqslant (\theta_{2,2} - \theta_{1,2} - \theta_{3,2} - \eta)^+, \\
R_2 &\leqslant \min\left(\theta_{1,1} + \theta_{6,2}, \theta_{2,2}, \theta_{1,1} + \theta_{5,1} + \theta_{2,2} - \theta_{2,1} + \eta\right), \\
R_1 + R_2 &\leqslant \min\left(\theta_{4,1} + \theta_{2,2} - \eta, \theta_{2,1} + \theta_{4,2}, \theta_{1,1} + \theta_{5,1} + \theta_{1,2} + \theta_{5,2}\right), \\
R_1 + 2R_2 &\leqslant \min\left(\theta_{1,1} + \theta_{5,1} + \theta_{2,2} + \theta_{4,2}, \theta_{1,1} + \theta_{2,1} + \theta_{4,2} + \theta_{6,2}\right), \\
2R_1 + R_2 &\leqslant \min\left(\theta_{4,1} + \theta_{6,1} + \theta_{1,2} + \theta_{2,2}, \theta_{2,1} + \theta_{4,1} + \theta_{1,2} + \theta_{5,2}\right).
\end{aligned} \tag{I.36}$$

The region described by (I.36) is identical to $\mathcal{C} \cap \mathcal{B}_\eta$. This completes the proof.

∎

— J —

Proof of Theorem 14

The proof of Theorem 14 consists of constructing a coding scheme that satisfies Definition 4. The coding scheme is a generalization to continuous channel inputs of the coding scheme introduced in Appendix I 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 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 28 describes all the rate pairs $(R_1, R_2) \in \mathbb{R}^2_+$ that can be achieved with the RHK-NOF scheme.

Lemma 28. *The RHK-NOF scheme achieves the set of rates* $\Big(R_{1,C1},\ R_{1,R1},\ R_{1,C2},\ R_{1,R2},\ R_{1,P},\ R_{2,C1},\ R_{2,R1},\ R_{2,C2},\ R_{2,R2},\ R_{2,P}\Big) \in \mathbb{R}^{10}_+$ *that satisfy the following conditions:*

$$R_{i,P} \leqslant a_{1,i}, \tag{J.1a}$$

$$R_i + R_{j,C} + R_{j,R} \leqslant a_{2,i}(\rho), \tag{J.1b}$$

$$R_{j,C1} + R_{j,R1} \leqslant a_{3,i}(\rho, \mu_j), \tag{J.1c}$$

$$R_{j,C2} + R_{j,R2} \leqslant a_{4,i}(\rho, \mu_j), \tag{J.1d}$$

$$R_{i,P} + R_{j,C2} + R_{j,R2} \leqslant a_{5,i}(\rho, \mu_j), \tag{J.1e}$$

$$R_{i,C2} + R_{i,P} \leqslant a_{6,i}(\rho, \mu_i), \quad \textit{and} \tag{J.1f}$$

$$R_{i,C2} + R_{i,P} + R_{j,C2} + R_{j,R2} \leqslant a_{7,i}(\rho, \mu_1, \mu_2), \tag{J.1g}$$

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 28 is presented in Appendix M and Appendix N. ■

The set of inequalities in (J.1) 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 $\underline{\mathcal{C}}$ (Theorem 7). Therefore, the relevance of Lemma 28 relies on the implication that any rate pair $(R_1, R_2) \in \underline{\mathcal{C}}$ 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 29 provides the maximum rate improvement that a given transmitter-receiver pair can achieve by unilateral deviation from the R-KH-NOF coding scheme.

Lemma 29. *Assume that the rate* 10*-tuple* $\boldsymbol{R} = (R_{1,C1}, R_{1,R1}, R_{1,C2}, R_{1,R2}, R_{1,P}, R_{2,C1}, R_{2,R1}, R_{2,C2}, R_{2,R2}, R_{2,P}) \in \mathbb{R}_+^{10}$ *is achievable with the RHK-NOF. Let* $i \in \{1,2\}$ *and* $j \in \{1,2\} \setminus \{i\}$*. Then, any unilateral deviation of transmitter-receiver pair* i *by using any other coding scheme leads to a transmission rate* $R_i' \in \mathbb{R}_+$ *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: Assume that both transmitters achieve the rates $\boldsymbol{R}$ by using the RHK-NOF coding scheme following the code construction in Appendix N.

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 by $\boldsymbol{X}_1' = \left(X_{1,1}', X_{1,2}', \ldots, X_{1,N}'\right)$ and $\overrightarrow{\boldsymbol{Y}}'_1 = \left(\overrightarrow{Y}'_{1,1}, \overrightarrow{Y}'_{1,2}, \ldots, \overrightarrow{Y}'_{1,N}\right)$ respectively the vector of channel outputs of transmitter 1 and channel inputs to receiver 1, with $N = \max(N_1', N_2)$. Hence, an upper-bound for R_1' is obtained from the following inequalities:

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

where, (a) follows from Fano's inequality, since the rate R_1' is achievable from the assumptions of the lemma with $\delta_1 : \mathbb{N} \to \mathbb{R}_+$ a positive monotonically decreasing function (Lemma 58), and (b) follows from the fact that 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, note that the term $h\left(\overrightarrow{\boldsymbol{Y}}'_1|W_1,\Omega_1\right)$ in (J.2) satisfies the following lower bound:

$$\begin{aligned}
N_2(R_{2,C} + R_{2,R}) &= H(W_{2,C}, \Omega_2)\\
&\overset{(d)}{=} H(W_{2,C}, \Omega_2|W_1, \Omega_1)\\
&= I(W_{2,C}, \Omega_2; \overrightarrow{\boldsymbol{Y}}'_1|W_1, \Omega_1) + H\left(W_{2,C}, \Omega_2|\overrightarrow{\boldsymbol{Y}}'_1, W_1, \Omega_1\right)
\end{aligned}$$

$$\begin{aligned}
&\overset{(e)}{\leqslant} I(W_{2,C},\Omega_2;\overrightarrow{\boldsymbol{Y}}'_1|W_1,\Omega_1)+N\delta_2(N)\\
&= h(\overrightarrow{\boldsymbol{Y}}'_1|W_1,\Omega_1)-h(\overrightarrow{\boldsymbol{Y}}'_1|W_1,\Omega_1,W_{2,C},\Omega_2)+N\delta_2(N)\\
&\overset{(f)}{\leqslant} h(\overrightarrow{\boldsymbol{Y}}'_1|W_1,\Omega_1)+N\left(\delta_2(N)-\frac{1}{2}\log(2\pi e)\right), \qquad \text{(J.3)}
\end{aligned}$$

where, (d) follows from the independence of the indices W_1, Ω_1, W_2, and Ω_2; (e) follows from Fano's inequality, since 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 from the assumptions of the lemma with $\delta_2 : \mathbb{N} \rightarrow \mathbb{R}_+$ a positive monotonically decreasing function (Lemma 58); 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 (J.3) into (J.2), it follows that

$$\begin{aligned}
R'_1 &\leqslant \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1+2\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}}+\mathrm{INR}_{12}+1\right)-(R_{2,C}+R_{2,R})-\frac{1}{2}\log(2\pi e)+\delta(N)\\
&\leqslant \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1+2\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}}+\mathrm{INR}_{12}+1\right)-(R_{2,C}+R_{2,R})+\delta(N). \qquad \text{(J.4)}
\end{aligned}$$

Note that $\delta(N) = \delta_1(N) + \delta_2(N)$ is a monotonically decreasing function of N. Hence, in the asymptotic block-length 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 10-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 $(R_1, R_2) \in \mathbb{R}_+^2$, 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 30. *Let $\eta \geqslant 1$ and let the rate 10-tuple $\boldsymbol{R} = (R_{1,C1}, R_{1,R1}, R_{1,C2}, R_{1,R2}, R_{1,P}, R_{2,C1}, R_{2,R1}, R_{2,C2}, R_{2,R2}, R_{2,P}) \in \mathbb{R}_+^{10}$ be achievable with the RHK-NOF scheme. Let also $\rho \in [0,1]$ and for all $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$,*

$$R_{i,C}+R_{i,P}+R_{j,C}+R_{j,R} = \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_i+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_i\mathrm{INR}_{ij}}+\mathrm{INR}_{ij}+1\right)-\frac{1}{2}. \qquad \text{(J.5)}$$

Then, the rate pair $(R_1, R_2) \in \mathbb{R}_+^2$, 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 30 follows the same steps as in the proof of Lemma 26.

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 (J.5), with $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$. From the assumptions of the lemma, the configuration pair (s_1^*, s_2^*) is an η-NE and

$$\begin{aligned} u_i(s_1^*, s_2^*) =& R_i \\ =& R_{i,C} + R_{i,P} \\ =& \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_i + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_i \mathrm{INR}_{ij}} + \mathrm{INR}_{ij} + 1\right) - (R_{j,C} + R_{j,R}) - \frac{1}{2}, \end{aligned} \quad (J.6)$$

where the last equality holds from (J.5). Then, from Definition 4, 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 cannot exceed η, that is,

$$u_i(s_i, s_j^*) - u_i(s_i^*, s_j^*) \leqslant \eta. \quad (J.7)$$

Without loss of generality, let $i = 1$ be the deviating transmitter-receiver pair and assume it achieves the highest improvement (Lemma 29), that is,

$$u_1(s_1, s_2^*) = \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) - (R_{2,C} + R_{2,R}). \quad (J.8)$$

Hence, plugging (J.6) and (J.8) into (J.7) yields:

$$\begin{aligned} u_1(s_1, s_2^*) - u_1(s_1^*, s_2^*) =& \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) \quad (J.9) \\ & - \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) + \frac{1}{2} \\ \stackrel{(a)}{\leqslant}& 1 \\ \leqslant& \eta, \end{aligned}$$

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:

$$\begin{aligned} \Delta =& \frac{1}{2}\log\left(1 + \frac{2(1-\rho)\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}}}{\overrightarrow{\mathrm{SNR}}_1 + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1}\right) + \frac{1}{2} \\ \leqslant& \frac{1}{2}\log\left(1 + \frac{2\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}}}{\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1}\right) + \frac{1}{2} \\ \leqslant& \frac{1}{2}\log\left(1 + \frac{\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12}}{\overrightarrow{\mathrm{SNR}}_1 + \mathrm{INR}_{12} + 1}\right) + \frac{1}{2} \\ \leqslant& \frac{1}{2}\log(2) + \frac{1}{2} \\ =& 1 \\ \leqslant& \eta. \end{aligned} \quad (J.10)$$

This verifies that any rate improvement by unilateral deviation of the transmit-receive configuration (s_1^*, s_2^*) cannot be greater than η, with η arbitrarily close to 1. The same can be proved for the other transmitter-receiver pair and this completes the proof. ■

Finally, Lemma 31 characterizes the achievable η-NE region $\underline{\mathcal{N}}_\eta$.

Lemma 31. *For all rate pairs* $(R_1, R_2) \in \underline{\mathcal{N}}_\eta$*, there always exists at least one* η*-NE 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 10-tuple $(R_{1,C1}, R_{1,R1}, R_{1,C2}, R_{1,R2}, R_{1,P}, R_{2,C1}, R_{2,R1}, R_{2,C2}, R_{2,R2}, R_{2,P}) \in \mathbb{R}_+^{10}$ that is achievable with the RHK-NOF coding scheme satisfies the inequalities in (J.1). Additionally, any rate 10-tuple $(R_{1,C1}, R_{1,R1}, R_{1,C2}, R_{1,R2}, R_{1,P}, R_{2,C1}, R_{2,R1}, R_{2,C2}, R_{2,R2}, R_{2,P}) \in \mathbb{R}_+^{10}$ that satisfies (J.1) and (J.5) is an η-NE (Lemma 30). A Fourier-Motzkin elimination applied on inequalities (J.1) and (J.5) leads to a region described in terms of the rates R_1 and R_2, as follows:

$$
\begin{aligned}
R_1 \geqslant & \Big(a_{2,1}(\rho) - a_{3,1}(\rho,\mu_2) - a_{4,1}(\rho,\mu_2) - \eta\Big)^+, \\
R_1 \leqslant & \min\Big(a_{2,1}(\rho), a_{6,1}(\rho,\mu_1) + a_{3,2}(\rho,\mu_1), a_{1,1} + a_{3,2}(\rho,\mu_1) + a_{4,2}(\rho,\mu_1), \\
& 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) + \eta, \\
& 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) + 2\eta, \\
& a_{2,1}(\rho) + a_{3,2}(\rho,\mu_1) + a_{5,2}(\rho,\mu_1) - a_{2,2}(\rho) + \eta\Big), \\
R_2 \geqslant & \Big(a_{2,2}(\rho) - a_{3,2}(\rho,\mu_1) - a_{4,2}(\rho,\mu_1) - \eta\Big)^+, \\
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_{4,1}(\rho,\mu_2) + a_{1,2}, \\
& 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) + 2\eta, \\
& 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) + \eta, \\
& a_{3,1}(\rho,\mu_2) + a_{5,1}(\rho,\mu_2) + a_{2,2}(\rho) - a_{2,1}(\rho) + \eta\Big), \\
R_1 + R_2 \leqslant & \min\Big(a_{1,1} + a_{2,2}(\rho), a_{1,2} + 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_{7,1}(\rho,\mu_1,\mu_2) + a_{2,2}(\rho) + a_{3,2}(\rho,\mu_1) - a_{2,1}(\rho) + \eta, \\
& 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) + \eta\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), \\
& 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), \\
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)\Big). \qquad \text{(J.11)}
\end{aligned}
$$

The region (J.11) corresponds to the achievable η-NE region for the two-user D-GIC-NOF, *i.e.*, $\underline{\mathcal{N}}_\eta$. Finally, the achievable η-NE region in (J.11) can be presented in terms of the achievable region $\underline{\mathcal{C}}$ (Theorem 7) and the bounds in (J.11) that are not in the achievable region $\underline{\mathcal{C}}$. This completes the proof. ■

— K —

Proof of Theorem 15

Given an $\eta \geqslant 1$, it is shown that $R_1 > L_1$, $R_2 > L_2$, $R_1 < U_1$, and $R_2 < U_2$ are necessary conditions for the rate pair (R_1, R_2) to be an η-NE. This shows that if any rate pair (R_1, R_2) is an η-NE, then $(R_1, R_2) \in \overline{\mathcal{C}} \cap \overline{\mathcal{B}}_\eta$. This proof is completed by Lemma 32 and Lemma 33.

Lemma 32. *A rate pair $(R_1, R_2) \in \mathcal{C}$, with either $R_1 < \left(\frac{1}{2}\log(1 + \frac{\overrightarrow{\mathrm{SNR}}_1}{1+\mathrm{INR}_{12}}) - \eta\right)^+$ or $R_2 < \left(\frac{1}{2}\log(1 + \frac{\overrightarrow{\mathrm{SNR}}_2}{1+\mathrm{INR}_{21}}) - \eta\right)^+$ is not an η-NE, for any given $\eta \geqslant 0$.*

Proof: Let (s_1^*, s_2^*) be an η-NE transmit-receive configuration pair such that $u_1(s_1^*, s_2^*) = R_1$ and $u_2(s_1^*, s_2^*) = R_2$, respectively. Hence, from Definition 4, it holds that any rate improvement of a transmitter-receiver pair that unilaterally deviates from (s_1^*, s_2^*) is always smaller than η. Without loss of generality, let $R_1 < \left(\frac{1}{2}\log(1 + \frac{\overrightarrow{\mathrm{SNR}}_1}{1+\mathrm{INR}_{12}}) - \eta\right)^+$. Then, note that independently of the transmit-receive configuration of transmitter-receiver pair 2, transmitter-receiver pair 1 can always use a transmit-receive configuration s_1' in which transmitter 1 saturates the average power constraint (2.7) and interference is treated as noise at receiver 1. Thus, transmitter-receiver pair 1 is always able to achieve the rate $R(s_1', s_2^*) = \frac{1}{2}\log\left(1 + \frac{\overrightarrow{\mathrm{SNR}}_1}{1+\mathrm{INR}_{12}}\right)$, which implies that a utility improvement $R(s_1', s_2^*) - R(s_1^*, s_2^*) > \eta$ is always possible. Thus, from Definition 4, the assumption that the rate pair (R_1, R_2) is an η-NE does not hold. This completes the proof. ■

Lemma 33. *A rate pair $(R_1, R_2) \in \overline{\mathcal{C}}$, with either $R_1 > U_1$ or $R_2 > U_2$ is not an η-NE, for any given $\eta \geqslant 1$.*

Proof: Let (s_1^*, s_2^*) be an η-NE transmit-receive configuration pair such that $u_1(s_1^*, s_2^*) = R_1$ and $u_2(s_1^*, s_2^*) = R_2$, respectively. Hence, from Definition 4, it holds that any rate improvement of a transmitter-receiver pair that unilaterally deviates from (s_1^*, s_2^*) is always smaller than η.

Without loss of generality, the focus is on user 1 to show the upper-bound on R_1. Then, the following holds:

$$
\begin{aligned}
N\,R_1 =& H\left(W_1\right)\\
\overset{(a)}{=}& H\left(W_1|\Omega_1\right)\\
=& I\left(W_1; \overrightarrow{\boldsymbol{Y}}_1|\Omega_1\right) + H\left(W_1|\Omega_1, \overrightarrow{\boldsymbol{Y}}_1\right)\\
\overset{(b)}{\leqslant}& I\left(W_1; \overrightarrow{\boldsymbol{Y}}_1|\Omega_1\right) + N\delta_1(N)\\
=& h\left(\overrightarrow{\boldsymbol{Y}}_1|\Omega_1\right) - h\left(\overrightarrow{\boldsymbol{Y}}_1|W_1, \Omega_1\right) + N\delta_1(N)\\
=& \sum_{n=1}^{N}\left(h\left(\overrightarrow{Y}_{1,n}|\Omega_1, \overrightarrow{Y}_{1,(1:n-1)}\right) - h\left(\overrightarrow{Y}_{1,n}|W_1, \Omega_1, \overrightarrow{Y}_{1,(1:n-1)}\right)\right) + N\delta_1(N)\\
\leqslant& \sum_{n=1}^{N}\left(h\left(\overrightarrow{Y}_{1,n}\right) - h\left(\overrightarrow{Y}_{1,n}|W_1, \Omega_1, \overrightarrow{Y}_{1,(1:n-1)}\right)\right) + N\delta_1(N)\\
\overset{(c)}{=}& \sum_{n=1}^{N}\left(h\left(\overrightarrow{Y}_{1,n}\right) - h\left(\overrightarrow{Y}_{1,n}|X_{1,n}\right)\right) + N\delta_1(N)\\
=& \sum_{n=1}^{N}\left(h\left(\overrightarrow{Y}_{1,n}\right) - h\left(\overrightarrow{Z}_{1,n}\right) - h\left(\overrightarrow{Y}_{1,n}|X_{1,n}\right) + h\left(\overrightarrow{Z}_{1,n}\right)\right) + N\delta_1(N)\\
\leqslant& \frac{N}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) - \sum_{n=1}^{N}\left(h\left(\overrightarrow{Y}_{1,n}|X_{1,n}\right)\right.\\
&\left. - h\left(\overrightarrow{Y}_{1,n}|X_{1,n}, X_{2,n}\right)\right) + N\delta_1(N)\\
=& \frac{N}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) - \sum_{n=1}^{N} I\left(\overrightarrow{Y}_{1,n}; X_{2,n}|X_{1,n}\right)\\
&+ N\delta_1(N),
\end{aligned}
\tag{K.1}
$$

where, (a) follows from the independence between the indices W_i and Ω_i; (b) follows from Fano's inequality, since the rate R_i is achievable from the assumptions of the lemma with $\delta_1 : \mathbb{N} \to \mathbb{R}_+$ a positive monotonically decreasing function (Lemma 58); and (c) follows from the fact that $\boldsymbol{X}_{i,n} = f_{i,n}^{(N)}\left(W_i, \Omega_i, \overleftarrow{Y}_{i,(1:n-1)}\right)$ from the definition of the encoding function in (2.1) and $\left(W_i, \Omega_i, \overleftarrow{\boldsymbol{Y}}_{i,(1:n-1)}\right) \to X_{i,n} \to \overrightarrow{Y}_{i,n}$.

Then, the following holds:

$$
R_1 \leqslant \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) - \frac{1}{N}\sum_{n=1}^{N} I\left(\overrightarrow{Y}_{1,n}; X_{2,n}|X_{1,n}\right) + \delta_1(N).
\tag{K.2}
$$

The term $\frac{1}{N}\sum_{n=1}^{N} I\left(\overrightarrow{Y}_{1,n}; X_{2,n}|X_{1,n}\right)$ in (K.2) plays the same role as $\frac{1}{N}\sum_{n=1}^{N} H\left(\boldsymbol{X}_{2,C,n}, \boldsymbol{X}_{2,D,n}|\Omega_1,\right.$ $\left.W_1, \overrightarrow{\boldsymbol{Y}}_{1,(1:n-1)}\right)$ in (I.10) in the linear deterministic case. The remainder of the proof consists in finding a lower bound on $\frac{1}{N}\sum_{n=1}^{N} I\left(\overrightarrow{Y}_{1,n}; X_{2,n}|X_{1,n}\right)$. Consider a set of events (Boolean variables) that are determined by the parameters $\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, INR_{21}, and $\overleftarrow{\mathrm{SNR}}_2$. Given a fixed 4-tuple ($\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, $\mathrm{INR}_{21}, \overleftarrow{\mathrm{SNR}}_2$), the events are defined below:

$$C_{1,2}: \quad \mathrm{INR}_{21} < \overrightarrow{\mathrm{SNR}}_2 \leqslant \mathrm{INR}_{12}, \tag{K.3a}$$

$$C_{2,2}: \quad \max\left(\mathrm{INR}_{12}, \mathrm{INR}_{21}, \overleftarrow{\mathrm{SNR}}_2\right) < \overrightarrow{\mathrm{SNR}}_2 < \mathrm{INR}_{12}\mathrm{INR}_{21}, \tag{K.3b}$$

$$C_{3,2}: \quad \overleftarrow{\mathrm{SNR}}_2 \leqslant \mathrm{INR}_{12}, \tag{K.3c}$$

$$C_{4,2}: \quad \mathrm{INR}_{21} < \mathrm{INR}_{12} < \overrightarrow{\mathrm{SNR}}_2 \leqslant \overleftarrow{\mathrm{SNR}}_2, \tag{K.3d}$$

$$C_{5,2}: \quad \overrightarrow{\mathrm{SNR}}_2 > \max\left(\mathrm{INR}_{12}, \mathrm{INR}_{21}, \overleftarrow{\mathrm{SNR}}_2\right), \tag{K.3e}$$

$$C_{6,2}: \quad \overrightarrow{\mathrm{SNR}}_2 \geqslant \max\left(\mathrm{INR}_{12}\mathrm{INR}_{21}, \overleftarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}\right), \tag{K.3f}$$

$$C_{7,2}: \quad \max\left(\mathrm{INR}_{12}, \mathrm{INR}_{21}, \overleftarrow{\mathrm{SNR}}_2, \frac{\overleftarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}{\mathrm{INR}_{12}}\right) < \overrightarrow{\mathrm{SNR}}_2 < \overleftarrow{\mathrm{SNR}}_2\mathrm{INR}_{21} \leqslant \frac{\overleftarrow{\mathrm{SNR}}_2\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}}, \tag{K.3g}$$

$$C_{8,2}: \quad \max\left(\mathrm{INR}_{12}, \mathrm{INR}_{21}, \overleftarrow{\mathrm{SNR}}_2, \frac{\overleftarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}{\mathrm{INR}_{12}}\right) < \overrightarrow{\mathrm{SNR}}_2 < \mathrm{INR}_{12}\mathrm{INR}_{21} < \overleftarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}. \tag{K.3h}$$

Let $0 \leqslant \gamma_2 \leqslant 1$, be a fixed positive real defined as follows:

$$\gamma_2 = \begin{cases} \min\left(\frac{\overrightarrow{\mathrm{SNR}}_2}{\mathrm{INR}_{12}\mathrm{INR}_{21}}, \frac{1}{\mathrm{INR}_{21}}\right) & \text{if} \quad C_{1,2} \vee \left(C_{2,2} \wedge C_{3,2}\right) \text{ holds true} \\ \min\left(\frac{1}{\mathrm{INR}_{21}}, \frac{\mathrm{INR}_{12}}{\mathrm{INR}_{21}\overleftarrow{\mathrm{SNR}}_2}\right) & \text{if} \quad C_{4,2} \text{ holds true} \\ \min\left(1, \frac{\mathrm{INR}_{12}}{\overrightarrow{\mathrm{SNR}}_2}\right) & \text{if} \quad C_{5,2} \wedge C_{6,2} \text{ holds true} \\ \min\left(\frac{\overrightarrow{\mathrm{SNR}}_2}{\overleftarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}, \frac{\mathrm{INR}_{12}}{\overleftarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}, 1\right) & \text{if} \quad C_{7,2} \vee C_{8,2} \text{ holds true} \\ 0 & \text{otherwise} \end{cases}. \tag{K.4}$$

For all $n \in \{1, 2, \ldots, N\}$, let $U_{2,n}$ and $V_{2,n}$ be two independent random variables with zero mean and variances $1 - \gamma_2$ and γ_2, respectively, with γ_2 defined as in (K.4), such that $X_{2,n} = U_{2,n} + V_{2,n}$. Using this notation, the inequality in (K.2) can be written as follows:

$$\begin{aligned} R_1 &\overset{(d)}{\leqslant} \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) - \frac{1}{N}\sum_{n=1}^{N} I\left(\overrightarrow{Y}_{1,n}; U_{2,n}, V_{2,n}|X_{1,n}\right) \\ &\quad + \delta_1(N) \\ &= \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1\mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) - \frac{1}{N}\sum_{n=1}^{N}\Bigg(I\left(\overrightarrow{Y}_{1,n}; V_{2,n}|X_{1,n}\right) \\ &\quad + I\left(\overrightarrow{Y}_{1,n}; U_{2,n}, |X_{1,n}, V_{2,n}\right)\Bigg) + \delta_1(N) \end{aligned}$$

$$\leqslant \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right) - \frac{1}{N}\sum_{n=1}^{N} I\left(\overrightarrow{Y}_{1,n}; V_{2,n}|X_{1,n}\right) + \delta_1(N), \tag{K.5}$$

where, (d) follows from the fact that $I\left(\overrightarrow{Y}_{1,n}; X_{2,n}|X_{1,n}\right) = I\left(\overrightarrow{Y}_{1,n}; U_{2,n}, V_{2,n}|X_{1,n}\right)$ for all $n \in \{1, 2, \ldots N\}$.

The remainder of the proof consists in finding a lower bound on $\frac{1}{N}\sum_{n=1}^{N} I\left(\overrightarrow{Y}_{1,n}; V_{2,n}|X_{1,n}\right)$.

Following the same steps as in (K.1), the following holds:

$$\begin{aligned}
N\,R_2 &= H\left(W_2\right)\\
&= H\left(W_2|\Omega_2\right)\\
&= I\left(W_2; \overrightarrow{\boldsymbol{Y}}_2|\Omega_2\right) + H\left(W_2|\Omega_2, \overrightarrow{\boldsymbol{Y}}_2\right)\\
&\leqslant I\left(W_2; \overrightarrow{\boldsymbol{Y}}_2|\Omega_2\right) + N\delta_2(N)\\
&= h\left(\overrightarrow{\boldsymbol{Y}}_2|\Omega_2\right) - h\left(\overrightarrow{\boldsymbol{Y}}_2|W_2, \Omega_2\right) + N\delta_2(N)\\
&= \sum_{n=1}^{N}\left(h\left(\overrightarrow{Y}_{2,n}|\Omega_2, \overrightarrow{Y}_{2,(1:n-1)}\right) - h\left(\overrightarrow{Y}_{2,n}|W_2, \Omega_2, \overrightarrow{Y}_{2,(1:n-1)}\right)\right) + N\delta_2(N)\\
&\leqslant \sum_{n=1}^{N}\left(h\left(\overrightarrow{Y}_{2,n}\right) - h\left(\overrightarrow{Y}_{2,n}|W_2, \Omega_2, \overrightarrow{Y}_{2,(1:n-1)}\right)\right) + N\delta_2(N)\\
&= \sum_{n=1}^{N}\left(h\left(\overrightarrow{Y}_{2,n}\right) - h\left(\overrightarrow{Y}_{2,n}|X_{2,n}\right)\right) + N\delta_2(N)\\
&= \sum_{n=1}^{N} I\left(\overrightarrow{Y}_{2,n}; X_{2,n}\right) + N\delta_2(N)\\
&\overset{(e)}{=} \sum_{n=1}^{N}\left(I\left(\overrightarrow{Y}_{2,n}; V_{2,n}\right) + h\left(\overrightarrow{Y}_{2,n}|V_{2,n}\right) - h\left(\overrightarrow{Y}_{2,n}|X_{2,n}\right)\right) + N\delta_2(N),
\end{aligned} \tag{K.6}$$

where, (e) follows from the fact that $I\left(\overrightarrow{Y}_{2,n};X_{2,n},V_{2,n}\right) = I\left(\overrightarrow{Y}_{2,n};X_{2,n}\right) + I\left(\overrightarrow{Y}_{2,n};V_{2,n}|X_{2,n}\right) = I\left(\overrightarrow{Y}_{2,n}; V_{2,n}\right) + I\left(\overrightarrow{Y}_{2,n}; X_{2,n}|V_{2,n}\right)$.

Let $\varphi : \mathbb{N} \to \mathbb{R}^{+}$ be a monotonically decreasing function such that (K.6) holds with equality, i.e.,

$$R_2 = \frac{1}{N}\sum_{n=1}^{N}\left(I\left(\overrightarrow{Y}_{2,n}; V_{2,n}\right) + h\left(\overrightarrow{Y}_{2,n}|V_{2,n}\right) - h\left(\overrightarrow{Y}_{2,n}|X_{2,n}\right)\right) + \varphi(N). \tag{K.7}$$

Consider that player 2 implements an alternative strategy s_2' that induces channel inputs $\boldsymbol{X}_2' = \left(X_{2,1}', X_{2,2}', \ldots, X_{2,N_2'}'\right)$ where $X_{2,n}' = U_{2,n} + V_{2,n}'$, with $V_{2,n}'$ a random variable with variance γ_2 and independent of any symbol transmitted by either transmitter until channel use n. Note that $U_{2,n}$ continues to be the same as with the strategy s_2^*. Then, the channel-output at receiver 2 after the deviation from s_2^* to s_2', denoted by $\overrightarrow{Y}_{2,n}'$, is:

$$\overrightarrow{Y}'_{2,n} = \overrightarrow{h}_{22}\left(U_{2,n} + V'_{2,n}\right) + h_{21}X_{1,n} + \overrightarrow{Z}_{2,n}. \tag{K.8}$$

Let $\rho = \mathbb{E}\left[X_{1,n}X_{2,n}\right]$, $\mathbb{E}\left[X_{1,n}U_{2,n}\right] = \sqrt{1-\gamma_2}\rho_{X_1U_2}$, $\mathbb{E}\left[X_{1,n}V_{2,n}\right] = \sqrt{\gamma_2}\rho_{X_1V_2}$, and $\rho = \sqrt{1-\gamma_2}\rho_{X_1U_2} + \sqrt{\gamma_2}\rho_{X_1V_2}$. Then, following the same steps as in (K.6), it follows that:

$$R_2(s_1^*, s_2') \leqslant \frac{1}{N}\sum_{n=1}^{N}\left(I\left(\overrightarrow{Y}'_{2,n}; V'_{2,n}\right) + h\left(\overrightarrow{Y}'_{2,n}|V'_{2,n}\right) - h\left(\overrightarrow{Y}'_{2,n}|X'_{2,n}\right)\right) + \delta'_2(N), \tag{K.9}$$

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

$$\begin{aligned}
I\left(\overrightarrow{Y}'_{2,n}; V'_{2,n}\right) =& h\left(\overrightarrow{Y}'_{2,n}\right) - h\left(\overrightarrow{Y}'_{2,n}|V'_{2,n}\right)\\
\geqslant& \frac{1}{2}\log\left(2\pi e\left(\overrightarrow{\mathrm{SNR}}_2 + 2\left(\rho - \rho_{X_1V_2}\sqrt{\gamma_2}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}} + \mathrm{INR}_{21} + 1\right)\right)\\
&-\frac{1}{2}\log\left(2\pi e\left(\left(1-\gamma_2\right)\overrightarrow{\mathrm{SNR}}_2 + 2\left(\rho - \rho_{X_1V_2}\sqrt{\gamma_2}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}} + \mathrm{INR}_{21} + 1\right)\right)\\
=& \frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_2 + 2\left(\rho - \rho_{X_1V_2}\sqrt{\gamma_2}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}} + \mathrm{INR}_{21} + 1}{\left(1-\gamma_2\right)\overrightarrow{\mathrm{SNR}}_2 + 2\left(\rho - \rho_{X_1V_2}\sqrt{\gamma_2}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}} + \mathrm{INR}_{21} + 1}\right).
\end{aligned} \tag{K.10}$$

The inequality in (K.10) follows from using a worst-case noise argument.

From Definition 4, it follows that $R_2(s_1^*, s_2') \leqslant R_2 + \eta$, with $\eta > 0$. Hence, the following holds:

$$\begin{aligned}
&\frac{1}{N}\sum_{n=1}^{N}\left(I\left(\overrightarrow{Y}'_{2,n}; V'_{2,n}\right) + h\left(\overrightarrow{Y}'_{2,n}|V'_{2,n}\right) - h\left(\overrightarrow{Y}'_{2,n}|X'_{2,n}\right)\right) + \delta'_2(N) \leqslant\\
&\frac{1}{N}\sum_{n=1}^{N}\left(I\left(\overrightarrow{Y}_{2,n}; V_{2,n}\right) + h\left(\overrightarrow{Y}_{2,n}|V_{2,n}\right) - h\left(\overrightarrow{Y}_{2,n}|X_{2,n}\right)\right) + \varphi(N) + \eta.
\end{aligned} \tag{K.11}$$

From (K.11) the following holds:

$$\begin{aligned}
\frac{1}{N}\sum_{n=1}^{N} I\left(\overrightarrow{Y}_{2,n}; V_{2,n}\right) \geqslant& \frac{1}{N}\sum_{n=1}^{N}\left(I\left(\overrightarrow{Y}'_{2,n}; V'_{2,n}\right) + h\left(\overrightarrow{Y}'_{2,n}|V'_{2,n}\right) - h\left(\overrightarrow{Y}'_{2,n}|X'_{2,n}\right) - h\left(\overrightarrow{Y}_{2,n}|V_{2,n}\right)\right.\\
&\left. + h\left(\overrightarrow{Y}_{2,n}|X_{2,n}\right)\right) + \delta'_2(N) - \varphi(N) - \eta\\
=& \frac{1}{N}\sum_{n=1}^{N}\left(I\left(\overrightarrow{Y}'_{2,n}; V'_{2,n}\right) + h\left(\overrightarrow{h}_{22}U_{2,n} + h_{21}X_{1,n} + \overrightarrow{Z}_{2,n}\right)\right.\\
&- h\left(h_{21}X_{1,n} + \overrightarrow{Z}_{2,n}|X'_{2,n}\right) - h\left(\overrightarrow{h}_{22}U_{2,n} + h_{21}X_{1,n} + \overrightarrow{Z}_{2,n}|V_{2,n}\right)\\
&\left. + h\left(h_{21}X_{1,n} + \overrightarrow{Z}_{2,n}|X_{2,n}\right)\right) + \delta'_2(N) - \varphi(N) - \eta
\end{aligned}$$

$$\geqslant \frac{1}{N}\sum_{n=1}^{N}\left(I\left(\overrightarrow{Y}'_{2,n};V'_{2,n}\right)+h\left(h_{21}X_{1,n}+\overrightarrow{Z}_{2,n}|X_{2,n}\right)-h\left(h_{21}X_{1,n}+\overrightarrow{Z}_{2,n}|X'_{2,n}\right)\right)+\delta'_2(N)$$
$$-\varphi(N)-\eta$$
$$\stackrel{(f)}{\geqslant}\frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_2+2\left(\rho-\rho_{X_1V_2}\sqrt{\gamma_2}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}+\mathrm{INR}_{21}+1}{\left(1-\gamma_2\right)\overrightarrow{\mathrm{SNR}}_2+2\left(\rho-\rho_{X_1V_2}\sqrt{\gamma_2}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}+\mathrm{INR}_{21}+1}\right)$$
$$+\frac{1}{2}\log\left(\mathrm{INR}_{21}\left(1-\rho^2\right)+1\right)-\frac{1}{2}\log\left(\mathrm{INR}_{21}\left(1-\left(\rho-\rho_{X1V2}\sqrt{\gamma_2}\right)^2\right)+1\right)$$
$$+\delta'_2(N)-\varphi(N)-\eta, \tag{K.12}$$

where (f) follows from (K.10).

The lower-bound on $\frac{1}{N}\sum_{n=1}^{N} I\left(\overrightarrow{Y}_{1,n};V_{2,n}|X_{1,n}\right)$ can be obtained as follows:

$$\frac{1}{N}\sum_{n=1}^{N} I\left(\overrightarrow{Y}_{1,n};V_{2,n}|X_{1,n}\right)\stackrel{(g)}{=}\frac{1}{N}\sum_{n=1}^{N}\left(I\left(\overrightarrow{Y}_{2,n};V_{2,n}\right)+I\left(\overrightarrow{Y}_{1,n},X_{1,n};V_{2,n}|\overrightarrow{Y}_{2,n}\right)-I\left(X_{1,n};V_{2,n}\right)\right.$$
$$\left.-I\left(\overrightarrow{Y}_{2,n};V_{2,n}|X_{1,n},\overrightarrow{Y}_{1,n}\right)\right)$$
$$=\frac{1}{N}\sum_{n=1}^{N}\left(I\left(\overrightarrow{Y}_{2,n};V_{2,n}\right)+h\left(V_{2,n}|\overrightarrow{Y}_{2,n}\right)-h\left(V_{2,n}|\overrightarrow{Y}_{2,n},\overrightarrow{Y}_{1,n},X_{1,n}\right)\right.$$
$$\left.-I\left(X_{1,n};V_{2,n}\right)-h\left(V_{2,n}|X_{1,n},\overrightarrow{Y}_{1,n}\right)+h\left(V_{2,n}|X_{1,n},\overrightarrow{Y}_{1,n},\overrightarrow{Y}_{2,n}\right)\right)$$
$$=\frac{1}{N}\sum_{n=1}^{N}\left(I\left(\overrightarrow{Y}_{2,n};V_{2,n}\right)+h\left(V_{2,n}|\overrightarrow{Y}_{2,n}\right)-I\left(X_{1,n};V_{2,n}\right)\right.$$
$$\left.-h\left(V_{2,n}|X_{1,n},\overrightarrow{Y}_{1,n}\right)\right)$$
$$\stackrel{(h)}{=}\frac{1}{N}\sum_{n=1}^{N}\left(I\left(\overrightarrow{Y}_{2,n};V_{2,n}\right)+h\left(V_{2,n}|\overrightarrow{Y}_{2,n}\right)-h\left(V_{2,n}|X_{1,n},\overrightarrow{Y}_{1,n}\right)\right)$$
$$+\frac{1}{2}\log\left(1-\rho^2_{X_1V_2}\right)$$
$$\stackrel{(i)}{=}\frac{1}{N}\sum_{n=1}^{N} I\left(\overrightarrow{Y}_{2,n};V_{2,n}\right)-\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_2+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}+\mathrm{INR}_{21}+1\right)$$
$$+\frac{1}{2}\log\left((2\pi e)\left(\overrightarrow{\mathrm{SNR}}_2\left(\gamma_2-\gamma_2^2\right)+2\gamma_2\left(\rho-\rho_{X_1V_2}\sqrt{\gamma_2}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}\right.\right.$$
$$\left.\left.+\gamma_2\mathrm{INR}_{21}\left(1-\rho^2_{X_1V_2}\right)+\gamma_2\right)\right)-\frac{1}{N}\sum_{n=1}^{N} h\left(V_{2,n}|X_{1,n},\overrightarrow{Y}_{1,n}\right)$$
$$+\frac{1}{2}\log\left(1-\rho^2_{X_1V_2}\right)$$

$$
\begin{aligned}
&\stackrel{(j)}{=}\frac{1}{N}\sum_{n=1}^{N} I\left(\overrightarrow{Y}_{2,n};V_{2,n}\right)\\
&-\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_2+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}+\mathrm{INR}_{21}+1\right)\\
&+\frac{1}{2}\log\Bigg((2\pi e)\Bigg(\overrightarrow{\mathrm{SNR}}_2\left(\gamma_2-\gamma_2^2\right)+2\gamma_2\left(\rho-\rho_{X_1V_2}\sqrt{\gamma_2}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}+\gamma_2\mathrm{INR}_{21}\left(1-\rho^2_{X_1V_2}\right)\\
&+\gamma_2\Bigg)\Bigg)+\frac{1}{2}\log\left(1-\rho^2_{X_1V_2}\right)+\frac{1}{2}\log\left(\mathrm{INR}_{12}\left(1-\rho^2\right)+1\right)\\
&-\frac{1}{2}\log\Bigg((2\pi e)\Bigg(\gamma_2\left(\mathrm{INR}_{12}\left(1-\rho^2\right)+1\right)-\rho^2_{X_1V_2}\gamma_2\left(\mathrm{INR}_{12}+1\right)+\gamma_2\mathrm{INR}_{12}\big(2\rho_{X_1V_2}\rho\sqrt{\gamma_2}\\
&-\gamma_2\big)\Bigg)\Bigg)\\
&\stackrel{(k)}{\geqslant}\frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_2+2\left(\rho-\rho_{X_1V_2}\sqrt{\gamma_2}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}+\mathrm{INR}_{21}+1}{\left(1-\gamma_2\right)\overrightarrow{\mathrm{SNR}}_2+2\left(\rho-\rho_{X_1V_2}\sqrt{\gamma_2}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}+\mathrm{INR}_{21}+1}\right)\\
&+\frac{1}{2}\log\left(\mathrm{INR}_{21}\left(1-\rho^2\right)+1\right)-\frac{1}{2}\log\left(\mathrm{INR}_{21}\left(1-\left(\rho-\rho_{X1V2}\sqrt{\gamma_2}\right)^2\right)+1\right)\\
&-\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_2+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}+\mathrm{INR}_{21}+1\right)\\
&+\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_2\left(\gamma_2-\gamma_2^2\right)+2\gamma_2\left(\rho-\rho_{X_1V_2}\sqrt{\gamma_2}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}+\gamma_2\mathrm{INR}_{21}\left(1-\rho^2_{X_1V_2}\right)+\gamma_2\right)\\
&+\frac{1}{2}\log\left(1-\rho^2_{X_1V_2}\right)+\frac{1}{2}\log\left(\mathrm{INR}_{12}\left(1-\rho^2\right)+1\right)\\
&-\frac{1}{2}\log\left(\gamma_2\left(\mathrm{INR}_{12}\left(1-\rho^2\right)+1\right)-\rho^2_{X_1V_2}\gamma_2\left(\mathrm{INR}_{12}+1\right)+\gamma_2\mathrm{INR}_{12}\left(2\rho_{X_1V_2}\rho\sqrt{\gamma_2}-\gamma_2\right)\right)\\
&+\delta_2'(N)-\varphi(N)-\eta, \qquad\qquad (K.13)
\end{aligned}
$$

where, (g) follows from the fact that $I\left(\overrightarrow{Y}_{1,n},\overrightarrow{Y}_{2,n},X_{1,n};V_{2,n}\right)=I\left(\overrightarrow{Y}_{2,n};V_{2,n}\right)+I\Big(\overrightarrow{Y}_{1,n},X_{1,n};V_{2,n}|\overrightarrow{Y}_{2,n}\Big)=I\left(X_{1,n};V_{2,n}\right)+I\left(\overrightarrow{Y}_{1,n};V_{2,n}|X_{1,n}\right)+I\left(\overrightarrow{Y}_{2,n};V_{2,n}|X_{1,n},\overrightarrow{Y}_{1,n}\right)$; (h) follows from the fact that $I\left(X_{1,n};V_{2,n}\right)=-\frac{1}{2}\log\left(1-\rho^2_{X_1V_2}\right)$; (i) follows from the fact that $h\left(V_{2,n}|\overrightarrow{Y}_{2,n}\right)=\frac{1}{2}\log\Bigg((2\pi e)\Bigg(\overrightarrow{\mathrm{SNR}}_2\left(\gamma_2-\gamma_2^2\right)+2\gamma_2\left(\rho-\rho_{X_1V_2}\sqrt{\gamma_2}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}+\gamma_2\mathrm{INR}_{21}\left(1-\rho^2_{X_1V_2}\right)+\gamma_2\Bigg)\Bigg)-\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_2+2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_2\mathrm{INR}_{21}}+\mathrm{INR}_{21}+1\right)$; (j) follows from the fact that $h\left(V_{2,n}|X_{1,n},\overrightarrow{Y}_{1,n}\right)=\frac{1}{2}\log\Bigg((2\pi e)\Bigg(\gamma_2\left(\mathrm{INR}_{12}\left(1-\rho^2\right)+1\right)-\rho^2_{X_1V_2}\gamma_2\left(\mathrm{INR}_{12}+1\right)+\gamma_2\mathrm{INR}_{12}\left(2\rho_{X_1V_2}\rho\sqrt{\gamma_2}-\gamma_2\right)\Bigg)\Bigg)-\frac{1}{2}\log\left(\mathrm{INR}_{12}\left(1-\rho^2\right)+1\right)$; and (k) follows from (K.12).

Plugging (K.13) into (K.5), and in the asymptotic block-length regime, the following holds:

$$
\begin{aligned}
R_1(s_1^*, s_2^*) \leqslant\ & \frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_1 + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_1 \mathrm{INR}_{12}} + \mathrm{INR}_{12} + 1\right)\\
& -\frac{1}{2}\log\left(\frac{\overrightarrow{\mathrm{SNR}}_2 + 2\left(\rho - \rho_{X_1V_2}\sqrt{\gamma_2}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_2 \mathrm{INR}_{21}} + \mathrm{INR}_{21} + 1}{(1-\gamma_2)\,\overrightarrow{\mathrm{SNR}}_2 + 2\left(\rho - \rho_{X_1V_2}\sqrt{\gamma_2}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_2 \mathrm{INR}_{21}} + \mathrm{INR}_{21} + 1}\right)\\
& -\frac{1}{2}\log\left(\mathrm{INR}_{21}\left(1-\rho^2\right)+1\right) + \frac{1}{2}\log\left(\mathrm{INR}_{21}\left(1-\left(\rho - \rho_{X1V2}\sqrt{\gamma_2}\right)^2\right)+1\right)\\
& +\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_2 + 2\rho\sqrt{\overrightarrow{\mathrm{SNR}}_2 \mathrm{INR}_{21}} + \mathrm{INR}_{21} + 1\right)\\
& -\frac{1}{2}\log\left(\overrightarrow{\mathrm{SNR}}_2\left(\gamma_2 - \gamma_2^2\right) + 2\gamma_2\left(\rho - \rho_{X_1V_2}\sqrt{\gamma_2}\right)\sqrt{\overrightarrow{\mathrm{SNR}}_2 \mathrm{INR}_{21}} + \gamma_2 \mathrm{INR}_{21}\left(1-\rho_{X_1V_2}^2\right) + \gamma_2\right)\\
& -\frac{1}{2}\log\left(1-\rho_{X_1V_2}^2\right) - \frac{1}{2}\log\left(\mathrm{INR}_{12}\left(1-\rho^2\right)+1\right)\\
& +\frac{1}{2}\log\left(\gamma_2\left(\mathrm{INR}_{12}\left(1-\rho^2\right)+1\right) - \rho_{X_1V_2}^2\gamma_2\left(\mathrm{INR}_{12}+1\right) + \gamma_2 \mathrm{INR}_{12}\left(2\rho_{X_1V_2}\rho\sqrt{\gamma_2} - \gamma_2\right)\right)\\
& +\eta. \qquad \text{(K.14)}
\end{aligned}
$$

The same procedure can be applied for the other user and this completes the proof. ■

— L —

Proof of Lemma 21

Lemma 21 is proved as follows:

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

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

where (a) follows from the fact that $\overleftarrow{Z}_{i,n}$ and $\overleftarrow{Z}_{j,n}$ are independent of W_i, W_j, $\boldsymbol{X}_{i,C,(1:n-1)}$, $\boldsymbol{X}_{j,U,(1:n-1)}$, $\overleftarrow{\boldsymbol{Y}}_{i,(1:n-1)}$, $\overleftarrow{\boldsymbol{Y}}_{j,(1:n-1)}$, $X_{i,n}$, $X_{j,n}$, $\overrightarrow{Z}_{j,n}$, and $\overrightarrow{Z}_{i,n}$. This completes the proof of Lemma 21.

— M —

Proof of Lemma 24

THIS appendix provides a description of the RHK-NOF and a proof of Lemma 24. This scheme is based on a three-part message splitting, superposition coding, common randomness and backward decoding.

Codebook Generation: Fix a strictly positive joint probability distribution

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

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

Let $R_{1,C1}$, $R_{1,R1}$, $R_{1,C2}$, $R_{1,R2}$, $R_{2,C1}$, $R_{2,R1}$, $R_{2,C2}$, $R_{1,R2}$, $R_{1,P}$, and $R_{2,P}$ be non-negative real numbers. Let $R_{1,C} = R_{1,C1} + R_{1,C2}$, $R_{2,C} = R_{2,C1} + R_{2,C2}$, $R_{1,R} = R_{1,R1} + R_{1,R2}$, $R_{2,R} = R_{2,R1} + R_{2,R2}$. Define 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 $R_{i,R}$, which 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 the product distribution

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

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

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 the conditional distribution

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

with $k \in \{1, 2, \ldots, 2^{N\left(R_{1,C1}+R_{1,R1}\right)}\}$.

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 the conditional distribution

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

with $l \in \{1, 2, \ldots, 2^{N\left(R_{1,C2}+R_{1,R2}\right)}\}$.

For each triplet 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 the conditional distribution

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

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 the conditional distribution

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

with $j \in \{1, 2, \ldots, 2^{N\left(R_{2,C1}+R_{2,R1}\right)}\}$.

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. N-length 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 the conditional distribution

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

with $m \in \{1, 2, \ldots, 2^{N\left(R_{2,C2}+R_{2,R2}\right)}\}$.

For each triplet 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 $\boldsymbol{x}_{2,P}(s,r,j,m,b) = \big(x_{2,P,1}(s,r,j,m,b), x_{2,P,2}(s,r,j,m,b), \ldots, x_{2,P,N}(s,r,j,m,b)\big)$ according to

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

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

Encoding: Denote by $W_i^{(t)} \in \mathcal{W}_i = \left\{1, 2, \ldots, 2^{N\left(R_{i,C}+R_{i,P}\right)}\right\}$ and $\Omega_i^{(t)} \in \mathcal{W}_{i,R} = \{1, 2, \ldots, 2^{NR_{i,R}}\}$ the message index and the random message index of transmitter i during block $t \in \{1, 2, \ldots, T\}$, respectively, with $T \in \mathbb{N}$ the total number of blocks. Let $W_i^{(t)}$ be decomposed into the message index $W_{i,C}^{(t)} \in \mathcal{W}_{i,C} = \{1, 2, \ldots, 2^{NR_{i,C}}\}$ and the message index $W_{i,P}^{(t)} \in \mathcal{W}_{i,P}=\{1,\ 2,\ \ldots, 2^{NR_{i,P}}\}$. 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 decomposed into the message indices $W_{i,C1}^{(t)} \in \mathcal{W}_{i,C1} = \{1, 2, \ldots, 2^{NR_{i,C1}}\}$ and $W_{i,C2}^{(t)} \in \mathcal{W}_{i,C2} = \{1, 2, \ldots, 2^{NR_{i,C2}}\}$. That is, $W_{i,C}^{(t)} = \Big(W_{i,C1}^{(t)}, W_{i,C2}^{(t)}\Big)$. Let $\Omega_i^{(t)}$ be decomposed into the message indices $\Omega_{i,R1}^{(t)} \in \mathcal{W}_{i,R1} = \{1, 2, \ldots, 2^{NR_{i,R1}}\}$ and $\Omega_{i,R2}^{(t)} \in \mathcal{W}_{i,R2} = \{1, 2, \ldots, 2^{NR_{i,R2}}\}$. That is, $\Omega_i^{(t)} = \left(\Omega_{i,R1}^{(t)},\ \Omega_{i,R2}^{(t)}\right)$. The blueindices $\left(W_{i,C1}^{(t)}, \Omega_{i,R1}^{(t)}\right)$ must be reliably decoded by transmitter j, with $j \in \{1, 2\} \setminus \{i\}$ (via feedback), and by both receivers. The indices $\left(W_{i,C2}^{(t)}, \Omega_{i,R2}^{(t)}\right)$ must be reliably decoded by both receivers but not by transmitter j.

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

$$
\begin{aligned}
\boldsymbol{x}_1^{(t)}=\Theta_1\Bigg(&\boldsymbol{u}\Big(\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}\right), \left(W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right)\Big), \\
&\boldsymbol{u}_1\Big(\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}\right), \left(W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right), \left(W_{1,C1}^{(t)}, \Omega_{1,R1}^{(t)}\right)\Big), \\
&\boldsymbol{v}_1\Big(\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}\right), \left(W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right), \left(W_{1,C1}^{(t)}, \Omega_{1,R1}^{(t)}\right), \left(W_{1,C2}^{(t)}, \Omega_{1,R2}^{(t)}\right)\Big), \\
&\boldsymbol{x}_{1,P}\Big(\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}\right), \left(W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right), \left(W_{1,C1}^{(t)}, \Omega_{1,R1}^{(t)}\right), \left(W_{1,C2}^{(t)}, \Omega_{1,R2}^{(t)}\right), W_{1,P}^{(t)}\Big)\Bigg),
\end{aligned}
\tag{M.2}
$$

where $\Theta_1 : (\mathcal{X}_1 \cap \mathcal{X}_2)^N \times \mathcal{X}_1^{3N} \rightarrow \mathcal{X}_1^N$ is a function that transforms the codewords $\boldsymbol{u}\Big(\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}\right), \left(W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right)\Big)$, $\boldsymbol{u}_1\Big(\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}\right), \left(W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right), \left(W_{1,C1}^{(t)}, \Omega_{1,R1}^{(t)}\right)\Big)$, $\boldsymbol{v}_1\Big(\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}\right), \left(W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right), \left(W_{1,C1}^{(t)}, \Omega_{1,R1}^{(t)}\right), \left(W_{1,C2}^{(t)}, \Omega_{1,R2}^{(t)}\right)\Big)$, and $\boldsymbol{x}_{1,P}\Big(\left(W_{1,C1}^{(t-1)}, \Omega_{1,R1}^{(t-1)}\right), \left(W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right), \left(W_{1,C1}^{(t)}, \Omega_{1,R1}^{(t)}\right), \left(W_{1,C2}^{(t)}, \Omega_{1,R2}^{(t)}\right), W_{1,P}^{(t)}\Big)$ into the N-dimensional vector $\boldsymbol{x}_1^{(t)}$ of channel inputs. The indices $\left(W_{1,C1}^{(0)}, \Omega_{1,R1}^{(0)}\right) = \left(W_{1,C1}^{(T)}, \Omega_{1,R1}^{(T)}\right) = s^*$ and $\left(W_{2,C1}^{(0)} = \Omega_{2,R1}^{(0)}\right) = \left(W_{2,C1}^{(T)}, \Omega_{2,R1}^{(T)}\right) = r^*$, and the pair $(s^*, r^*) \in \{1, 2, \ldots, 2^{N\left(R_{1,C1}+R_{1,R1}\right)}\} \times \{1, 2, \ldots, 2^{N\left(R_{2,C1}+R_{2,R1}\right)}\}$ are pre-defined and known by both receivers and transmitters. It is worth noting that the indices $\left(W_{2,C1}^{(t-1)}, \Omega_{2,R1}^{(t-1)}\right)$ 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 $\Big(\left(\widehat{W}_{1,C1}^{(T-t)}, \Omega_{1,R1}^{(T-t)}\right), \left(\widehat{W}_{2,C1}^{(T-t)}, \widehat{\Omega}_{2,R1}^{(T-t)}\right), \left(\widehat{W}_{1,C2}^{(T-(t-1))}, \Omega_{1,R2}^{(T-(t-1))}\right), \widehat{W}_{1,P}^{(T-(t-1))}, \left(\widehat{W}_{2,C2}^{(T-(t-1))}, \widehat{\Omega}_{2,R2}^{(T-(t-1))}\right)\Big) \in \mathcal{W}_{1,C1} \times \mathcal{W}_{1,R1} \times \mathcal{W}_{2,C1} \times \mathcal{W}_{2,R1} \times \mathcal{W}_{1,C2} \times \mathcal{W}_{1,R2} \times \mathcal{W}_{1,P} \times \mathcal{W}_{2,C2} \times \mathcal{W}_{2,R2}$ from the channel output $\overrightarrow{\boldsymbol{y}}_1^{(T-(t-1))}$. The 5-tuple $\Big(\left(\widehat{W}_{1,C1}^{(T-t)}, \Omega_{1,R1}^{(T-t)}\right), \left(\widehat{W}_{2,C1}^{(T-t)}, \widehat{\Omega}_{2,R1}^{(T-t)}\right), \left(\widehat{W}_{1,C2}^{(T-(t-1))}, \Omega_{1,R2}^{(T-(t-1))}\right), \widehat{W}_{1,P}^{(T-(t-1))}, \left(\widehat{W}_{2,C2}^{(T-(t-1))}, \widehat{\Omega}_{2,R2}^{(T-(t-1))}\right)\Big)$ is the unique 5-tuple that satisfies:

$$\Big(\boldsymbol{u}\left(\left(\widehat{W}_{1,C1}^{(T-t)},\Omega_{1,R1}^{(T-t)}\right),\left(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\right)\right),\boldsymbol{u}_1\Big(\left(\widehat{W}_{1,C1}^{(T-t)},\Omega_{1,R1}^{(T-t)}\right),\left(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\right),$$
$$\left(W_{1,C1}^{(T-(t-1))},\Omega_{1,R1}^{(T-(t-1))}\right)\Big),\boldsymbol{v}_1\Big(\left(\widehat{W}_{1,C1}^{(T-t)},\Omega_{1,R1}^{(T-t)}\right),\left(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\right),$$
$$\left(W_{1,C1}^{(T-(t-1))},\Omega_{1,R1}^{(T-(t-1))}\right),\left(\widehat{W}_{1,C2}^{(T-(t-1))},\Omega_{1,R2}^{(T-(t-1))}\right)\Big),\boldsymbol{x}_{1,P}\Big(\left(\widehat{W}_{1,C1}^{(T-t)},\Omega_{1,R1}^{(T-t)}\right),$$
$$\left(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\right),\left(W_{1,C1}^{(T-(t-1))},\Omega_{1,R1}^{(T-(t-1))}\right),\left(\widehat{W}_{1,C2}^{(T-(t-1))},\Omega_{1,R2}^{(T-(t-1))}\right),\widehat{W}_{1,P}^{(T-(t-1))}\Big),$$
$$\boldsymbol{u}_2\Big(\left(\widehat{W}_{1,C1}^{(T-t)},\Omega_{1,R1}^{(T-t)}\right),\left(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\right),\left(W_{2,C1}^{(T-(t-1))},\Omega_{2,R1}^{(T-(t-1))}\right)\Big),$$
$$\boldsymbol{v}_2\Big(\left(\widehat{W}_{1,C1}^{(T-t)},\Omega_{1,R1}^{(T-t)}\right),\left(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\right),\left(W_{2,C1}^{(T-(t-1))},\Omega_{2,R1}^{(T-(t-1))}\right),$$
$$\left(\widehat{W}_{2,C2}^{(T-(t-1))},\widehat{\Omega}_{2,R2}^{(T-(t-1))}\right)\Big),\overrightarrow{\boldsymbol{y}}_1^{(T-(t-1))}\Big)\in\mathcal{T}^{(N,\epsilon)}_{\left[U\ U_1\ V_1\ X_{1,P}\ U_2\ V_2\ \overrightarrow{Y}_1\right]},\tag{M.3}$$

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

Error Probability Analysis: An error might occur during encoding step t if the indices $\left(W_{2,C1}^{(t-1)},\Omega_{2,R1}^{(t-1)}\right)$ are not correctly decoded at transmitter 1 at the end of the step $t-1$. From the AEP [28], it follows that the blueindices $\left(W_{2,C1}^{(t-1)},\Omega_{2,R1}^{(t-1)}\right)$ can be reliably decoded at transmitter 1 during encoding step t, under the condition:

$$\begin{aligned}R_{2,C1}+R_{2,R1}\leqslant& I\left(\overleftarrow{Y}_1;U_2|U,U_1,V_1,X_1\right)\\=&I\left(\overleftarrow{Y}_1;U_2|U,X_1\right).\end{aligned}\tag{M.4}$$

An error might occur during the (backward) decoding step t if the indices $W_{1,C1}^{(T-t)}$, $\Big(W_{2,C1}^{(T-t)}$, $\Omega_{2,R1}^{(T-t)}\Big)$, $W_{1,C2}^{(T-(t-1))}$, $W_{1,P}^{(T-(t-1))}$, and $\left(W_{2,C2}^{(T-(t-1))},\Omega_{2,R2}^{(T-(t-1))}\right)$ are not decoded correctly given that the indices $W_{1,C1}^{(T-(t-1))}$ 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 5-tuple $\Big(\widehat{W}_{1,C1}^{(T-t)}$, $\left(\widehat{W}_{2,C1}^{(T-t)},\widehat{\Omega}_{2,R1}^{(T-t)}\right)$, $\widehat{W}_{1,C2}^{(T-(t-1))}$, $\widehat{W}_{1,P}^{(T-(t-1))}$, $\left(\widehat{W}_{2,C2}^{(T-(t-1))},\widehat{\Omega}_{2,R2}^{(T-(t-1))}\right)\Big)$ that satisfies (M.3), or (ii) there exist several 5-tuples $\Big(\widehat{W}_{1,C1}^{(T-t)}$, $\left(\widehat{W}_{2,C1}^{(T-t)}, \widehat{\Omega}_{2,R1}^{(T-t)}\right)$, $\widehat{W}_{1,C2}^{(T-(t-1))}$, $\widehat{W}_{1,P}^{(T-(t-1))}$, $\left(\widehat{W}_{2,C2}^{(T-(t-1))}, \widehat{\Omega}_{2,R2}^{(T-(t-1))}\right)\Big)$ that simultaneously satisfy (M.3). From the AEP [28], 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 $E_{(s,r,l,q,m)}^{(t)}$ that describes the case in which the codewords $\Big(\boldsymbol{u}(s,r)$, $\boldsymbol{u}_1\Big(s,r,\left(W_{1,C1}^{(T-(t-1))},\Omega_{1,R1}^{(T-(t-1))}\right)\Big)$, $\boldsymbol{v}_1\Big(s,r,\left(W_{1,C1}^{(T-(t-1))},\Omega_{1,R1}^{(T-(t-1))}\right),l\Big)$, $\boldsymbol{x}_{1,P}\Big(s,r,\left(W_{1,C1}^{(T-(t-1))},\Omega_{1,R1}^{(T-(t-1))}\right),l,q\Big)$, $\boldsymbol{u}_2\Big(s,r,\left(W_{2,C1}^{(T-(t-1))},\Omega_{2,R1}^{(T-(t-1))}\right)\Big)$, and $\boldsymbol{v}_2\Big(s,r,$

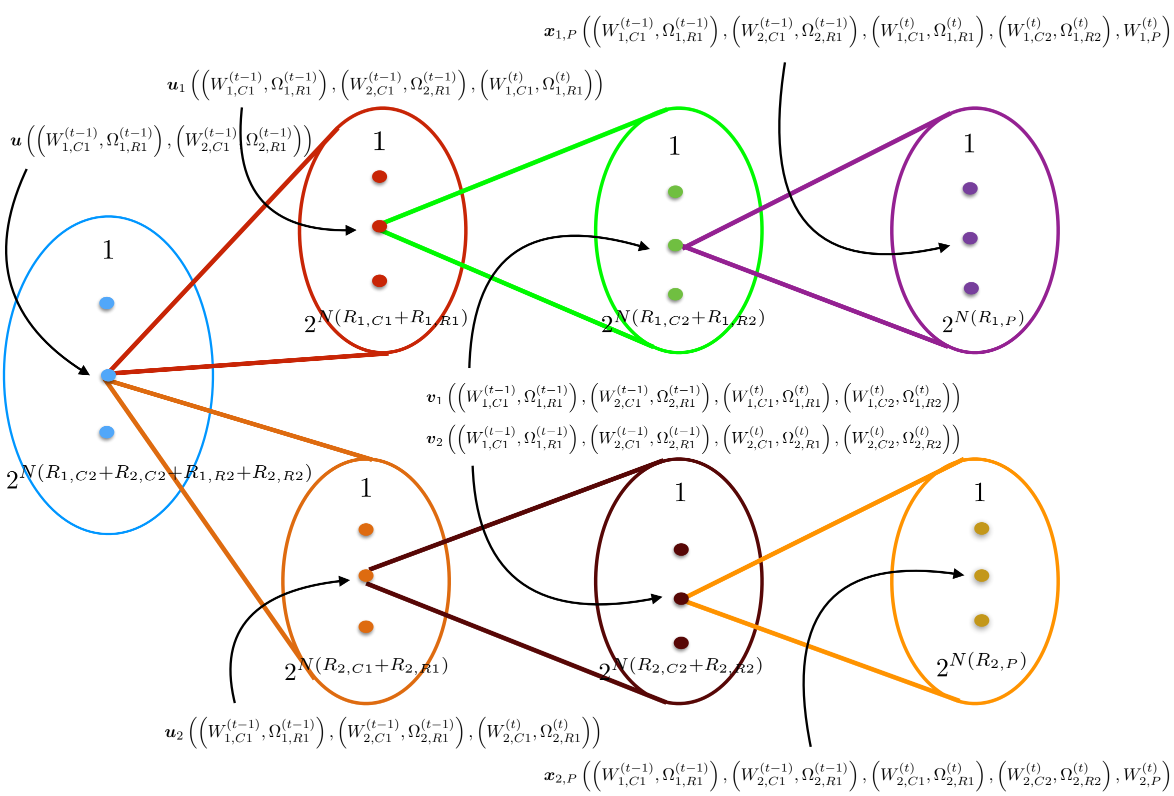


Figure M.1.: Structure of the superposition code. The codewords corresponding to the message indices $W^{(t-1)}_{1,C1}, W^{(t-1)}_{2,C1}, W^{(t)}_{i,C1}, W^{(t)}_{i,C2}, W^{(t)}_{i,P}$ with $i \in \{1,2\}$ as well as the block index t are both highlighted. The (approximate) number of codewords for each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

$\left(W^{(T-(t-1))}_{2,C1}, \Omega^{(T-(t-1))}_{2,R1}\right), m\Big)\Big)$ are jointly typical with $\overrightarrow{\boldsymbol{y}}^{(T-(t-1))}_1$ during decoding step t. Without loss of generality assume that the codeword to be decoded at decoding step t corresponds to the indices $(s,r,l,q,m) = (1,1,1,1,1)$ due to the symmetry of the code. Then, applying Boole's inequality to the probablility of error due to (ii) during decoding step t yields:

$$\begin{aligned} P^{(t)}_1(s_1,s_2) =& \Pr\left[\bigcup_{(s,r,l,q,m)\neq(1,1,1,1,1)} E^{(t)}_{(s,r,l,q,m)}\right] \\ \leqslant& \sum_{(s,r,l,q,m)\in\mathcal{T}} \Pr\left[E^{(t)}_{(s,r,l,q,m)}\right], \end{aligned} \tag{M.5}$$

with $\mathcal{T} = \left\{\mathcal{W}_{1,C1}\times\mathcal{W}_{1,R1}\times\mathcal{W}_{2,C1}\times\mathcal{W}_{2,R1}\times\mathcal{W}_{1,C2}\times\mathcal{W}_{1,R2}\times\mathcal{W}_{1,P}\times\mathcal{W}_{2,C2}\times\mathcal{W}_{2,R2}\right\} \setminus \{(1,1,1,1,1)\}$. Therefore,

$$\begin{aligned}
P_1^{(t)}(s_1,s_2) \leqslant & \sum_{s=1,r=1,l=1,q=1,m\neq 1} \Pr\left[E_{(s,r,l,q,m)}\right] + \sum_{s=1,r=1,l=1,q\neq 1,m=1} \Pr\left[E_{(s,r,l,q,m)}\right] \\
& + \sum_{s=1,r=1,l=1,q\neq 1,m\neq 1} \Pr\left[E_{(s,r,l,q,m)}\right] + \sum_{s=1,r=1,l\neq 1,q=1,m=1} \Pr\left[E_{(s,r,l,q,m)}\right] \\
& + \sum_{s=1,r=1,l\neq 1,q=1,m\neq 1} \Pr\left[E_{(s,r,l,q,m)}\right] + \sum_{s=1,r=1,l\neq 1,q\neq 1,m=1} \Pr\left[E_{(s,r,l,q,m)}\right] \\
& + \sum_{s=1,r=1,l\neq 1,q\neq 1,m\neq 1} \Pr\left[E_{(s,r,l,q,m)}\right] + \sum_{s=1,r\neq 1,l=1,q=1,m=1} \Pr\left[E_{(s,r,l,q,m)}\right] \\
& + \sum_{s=1,r\neq 1,l=1,q=1,m\neq 1} \Pr\left[E_{(s,r,l,q,m)}\right] + \sum_{s=1,r\neq 1,l=1,q\neq 1,m=1} \Pr\left[E_{(s,r,l,q,m)}\right] \\
& + \sum_{s=1,r\neq 1,l=1,q\neq 1,m\neq 1} \Pr\left[E_{(s,r,l,q,m)}\right] + \sum_{s=1,r\neq 1,l\neq 1,q=1,m=1} \Pr\left[E_{(s,r,l,q,m)}\right] \\
& + \sum_{s=1,r\neq 1,l\neq 1,q=1,m\neq 1} \Pr\left[E_{(s,r,l,q,m)}\right] + \sum_{s=1,r\neq 1,l\neq 1,q\neq 1,m=1} \Pr\left[E_{(s,r,l,q,m)}\right] \\
& + \sum_{s=1,r\neq 1,l\neq 1,q\neq 1,m\neq 1} \Pr\left[E_{(s,r,l,q,m)}\right] + \sum_{s\neq 1,r=1,l=1,q=1,m=1} \Pr\left[E_{(s,r,l,q,m)}\right] \\
& + \sum_{s\neq 1,r=1,l=1,q=1,m\neq 1} \Pr\left[E_{(s,r,l,q,m)}\right] + \sum_{s\neq 1,r=1,l=1,q\neq 1,m=1} \Pr\left[E_{(s,r,l,q,m)}\right] \\
& + \sum_{s\neq 1,r=1,l=1,q\neq 1,m\neq 1} \Pr\left[E_{(s,r,l,q,m)}\right] + \sum_{s\neq 1,r=1,l\neq 1,q=1,m=1} \Pr\left[E_{(s,r,l,q,m)}\right] \\
& + \sum_{s\neq 1,r=1,l\neq 1,q=1,m\neq 1} \Pr\left[E_{(s,r,l,q,m)}\right] + \sum_{s\neq 1,r=1,l\neq 1,q\neq 1,m=1} \Pr\left[E_{(s,r,l,q,m)}\right] \\
& + \sum_{s\neq 1,r=1,l\neq 1,q\neq 1,m\neq 1} \Pr\left[E_{(s,r,l,q,m)}\right] + \sum_{s\neq 1,r\neq 1,l=1,q=1,m=1} \Pr\left[E_{(s,r,l,q,m)}\right] \\
& + \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 \text{(M.6)}
\end{aligned}$$

$$\begin{aligned}
P_1^{(t)}(s_1,s_2) \leqslant & 2^{N(R_{2,C2}+R_{2,R2}-I(\overrightarrow{Y}_1;V_2|U,U_1,U_2,V_1,X_1)+2\epsilon)} \\
& + 2^{N(R_{1,P}-I(\overrightarrow{Y}_1;X_1|U,U_1,U_2,V_1,V_2)+2\epsilon)} \\
& + 2^{N(R_{2,C2}+R_{2,R2}+R_{1,P}-I(\overrightarrow{Y}_1;V_2,X_1|U,U_1,U_2,V_1)+2\epsilon)} \\
& + 2^{N(R_{1,C2}-I(\overrightarrow{Y}_1;V_1,X_1|U,U_1,U_2,V_2)+2\epsilon)} \\
& + 2^{N(R_{1,C2}+R_{2,C2}+R_{2,R2}-I(\overrightarrow{Y}_1;V_1,V_2,X_1|U,U_1,U_2)+2\epsilon)} \\
& + 2^{N(R_{1,C2}+R_{1,P}-I(\overrightarrow{Y}_1;V_1,X_1|U,U_1,U_2,V_2)+2\epsilon)} \\
& + 2^{N(R_{1,C2}+R_{1,P}+R_{2,C2}+R_{2,R2}-I(\overrightarrow{Y}_1;V_1,V_2,X_1|U,U_1,U_2)+2\epsilon)} \\
& + 2^{N(R_{2,C1}+R_{2,R1}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)} \\
& + 2^{N(R_{2,C}+R_{2,R}-I(\overrightarrow{Y}_1;U,U_1,U_2,V_1,V_2,X_1)+2\epsilon)}
\end{aligned}$$

$$
\begin{aligned}
&+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)}\\
&+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)}. \qquad \text{(M.7)}
\end{aligned}
$$

The same analysis of the probability of error holds for transmitter-receiver pair 2. From the AEP [28], and from (M.4) and (M.7), reliable decoding holds under the following conditions for transmitter $i \in \{1,2\}$, with $j \in \{1,2\} \setminus \{i\}$:

$$
\begin{aligned}
R_{j,C1} + R_{j,R1} \leqslant & \theta_{1,i}\\
& \triangleq I\left(\overleftarrow{Y}_i; U_j | U, U_i, V_i, X_i\right)\\
& = I\left(\overleftarrow{Y}_i; U_j | U, X_i\right), \qquad \text{(M.8a)}\\
R_i + R_{j,C} + R_{j,R} \leqslant & \theta_{2,i}\\
& \triangleq I(\overrightarrow{Y}_i; U, U_i, U_j, V_i, V_j, X_i)\\
& = I(\overrightarrow{Y}_i; U, U_j, V_j, X_i), \qquad \text{(M.8b)}\\
R_{j,C2} + R_{j,R2} \leqslant & \theta_{3,i}\\
& \triangleq I(\overrightarrow{Y}_i; V_j | U, U_i, U_j, V_i, X_i)\\
& = I(\overrightarrow{Y}_i; V_j | U, U_j, X_i), \qquad \text{(M.8c)}\\
R_{i,P} \leqslant & \theta_{4,i}\\
& \triangleq I(\overrightarrow{Y}_i; X_i | U, U_i, U_j, V_i, V_j), \qquad \text{(M.8d)}
\end{aligned}
$$

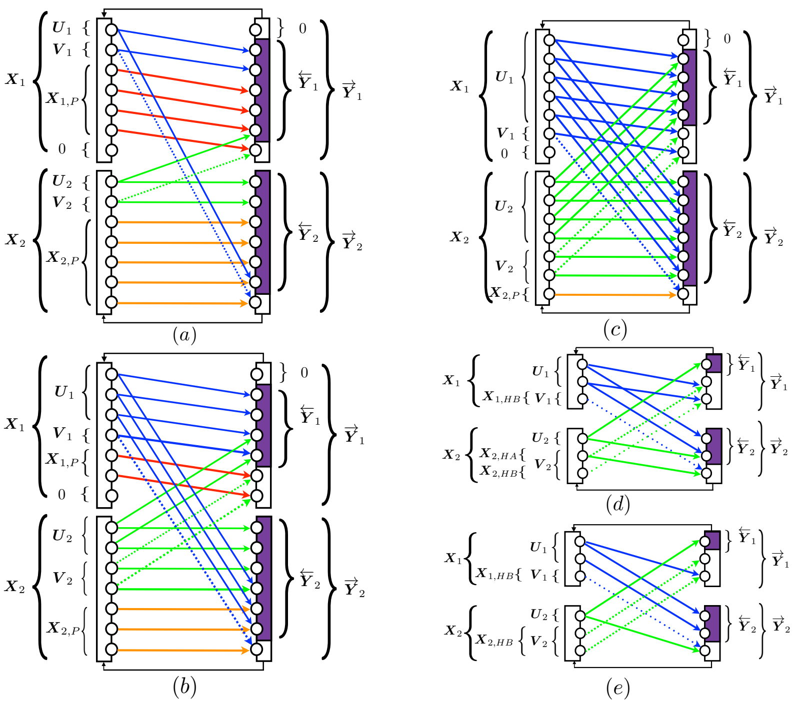


Figure M.2.: The auxiliary random variables and their relation with signals when channel-output feedback is considered in (a) very weak interference regime, (b) weak interference regime, (c) moderate interference regime, (d) strong interference regime and (e) very strong interference regime.

$$\begin{aligned} R_{i,P} + R_{j,C2} + R_{j,R2} \leqslant & \theta_{5,i} \\ \triangleq & I(\overrightarrow{Y}_i; V_j, X_i | U, U_i, U_j, V_i), \end{aligned} \tag{M.8e}$$

$$\begin{aligned} R_{i,C2} + R_{i,P} \leqslant & \theta_{6,i} \\ \triangleq & I(\overrightarrow{Y}_i; V_i, X_i | U, U_i, U_j, V_j) \\ = & I(\overrightarrow{Y}_i; X_i | U, U_i, U_j, V_j), \text{ and} \end{aligned} \tag{M.8f}$$

$$\begin{aligned} R_{i,C2} + R_{i,P} + R_{j,C2} + R_{j,R2} \leqslant & \theta_{7,i} \\ \triangleq & I(\overrightarrow{Y}_i; V_i, V_j, X_i | U, U_i, U_j) \\ = & I(\overrightarrow{Y}_i; V_j, X_i | U, U_i, U_j). \end{aligned} \tag{M.8g}$$

From the probability of error analysis, it follows that the rate-pairs achievable with the proposed randomized coding scheme with NOF are those simultaneously satisfying conditions (M.8) with $i \in \{1, 2\}$ and $j \in \{1, 2\} \backslash \{i\}$. Indeed, when $R_{1,R} = R_{2,R} = 0$, the coding scheme described above reduces to the coding scheme presented in Appendix A.

In the two-user LDIC-NOF model, the channel input of transmitter i at each channel use is a q-dimensional binary vector $\boldsymbol{X}_i \in \mathcal{X}_i = \{0, 1\}^q$ with $i \in \{1, 2\}$ and q as defined in (2.29). Following this observation, the random variables U, U_i, V_i, and $X_{i,P}$ described in (M.1) in the codebook generation are also vectors, and thus, they are denoted by $\boldsymbol{U}$, $\boldsymbol{U}_i$, $\boldsymbol{V}_i$ and $\boldsymbol{X}_{i,P}$, respectively. The random variables $\boldsymbol{U}_i$, $\boldsymbol{V}_i$, and $\boldsymbol{X}_{i,P}$ are assumed to be mutually independent and uniformly distributed over the sets $\{0,1\}^{\left(n_{ji}-\left(\max\left(\overrightarrow{n}_{jj},n_{ji}\right)-\overleftarrow{n}_{jj}\right)^{+}\right)^{+}}$, $\{0,1\}^{\left(\min\left(n_{ji},\left(\max\left(\overrightarrow{n}_{jj},n_{ji}\right)-\overleftarrow{n}_{jj}\right)^{+}\right)\right)}$, and $\{0,1\}^{\left(\overrightarrow{n}_{ii}-n_{ji}\right)^{+}}$, respectively. Note that the random variables $\boldsymbol{U}_i$, $\boldsymbol{V}_i$, and $\boldsymbol{X}_{i,P}$ have the dimensions indicated in (A.16a), (A.16b), and (A.16c), respectively.

Note that the random variable $\boldsymbol{U}$ in (M.1) is not used, and therefore, is a constant. The input symbol of transmitter i during channel use n is $\boldsymbol{X}_i = \left(\boldsymbol{U}_i^{\mathsf{T}}, \boldsymbol{V}_i^{\mathsf{T}}, \boldsymbol{X}_{i,P}^{\mathsf{T}}, (0,\ldots,0)\right)^{\mathsf{T}}$, where $(0,\ldots,0)$ is put to meet the dimension constraint $\dim \boldsymbol{X}_i = q$. Hence, during block $t \in \{1,2,\ldots,T\}$, the codeword $\boldsymbol{X}_i^{(t)}$ in the two-user LDIC-NOF is a $q \times N$ matrix, *i.e.*, $\boldsymbol{X}_i^{(t)} = (\boldsymbol{X}_{i,1}, \boldsymbol{X}_{i,2}, \ldots, \boldsymbol{X}_{i,N}) \in \{0,1\}^{q\times N}$.

The intuition behind this choice is based on the following observations: (a) the vector $\boldsymbol{U}_i$ represents the bits in $\boldsymbol{X}_i$ that can be observed by transmitter j via feedback but no necessarily by receiver i; (b) the vector $\boldsymbol{V}_i$ represents the bits in $\boldsymbol{X}_i$ that can be observed by receiver j but no necessarily by receiver i; and finally, (c) the vector $\boldsymbol{X}_{i,P}$ is a notational artefact to denote the bits of $\boldsymbol{X}_i$ that are neither in $\boldsymbol{U}_i$ nor $\boldsymbol{V}_i$. In particular, the bits in $\boldsymbol{X}_{i,P}$ are only observed by receiver i, as shown in Figure M.2. This intuition justifies the dimensions described in (A.16).

Considering this particular code structure, the terms $\theta_{l,i}$, with $(l,i) \in \{1,\ldots,7\} \times \{1,2\}$ in (M.8), are defined in (I.24). This completes the proof of Lemma 24.

— N —

Proof of Lemma 28

THIS appendix provides a proof of Lemma 28. For the two-user D-GIC-NOF model, consider that transmitter i uses the following random variable:

$$X_i = U + U_i + V_i + X_{i,P}, \tag{N.1}$$

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

$$U \sim \mathcal{N}(0, \rho), \tag{N.2a}$$
$$U_i \sim \mathcal{N}(0, \mu_i \lambda_{i,C}), \tag{N.2b}$$
$$V_i \sim \mathcal{N}(0, (1-\mu_i)\lambda_{i,C}), \tag{N.2c}$$
$$X_{i,P} \sim \mathcal{N}(0, \lambda_{i,P}), \tag{N.2d}$$

with

$$\rho + \lambda_{i,C} + \lambda_{i,P} = 1 \text{ and} \tag{N.3a}$$

$$\lambda_{i,P} = \min\left(\frac{1}{\mathrm{INR}_{ji}}, 1\right), \tag{N.3b}$$

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

The random variables U, U_1, U_2, V_1, V_2, $X_{1,P}$, and $X_{2,P}$ can be interpreted as components of the signals X_1 and X_2. The random variable U, which is used in this case, represents the common component of the channel inputs of transmitter 1 and transmitter 2.

The parameters ρ, μ_i, and $\lambda_{i,P}$ define a particular coding scheme for transmitter i. The assignment in (N.3b) is based on the intuition obtained from the linear deterministic model, in which the power of the signal $X_{i,P}$ from transmitter i to receiver j must be observed at the noise level. From (2.5), (2.6), and (N.1), the right-hand-side of the inequalities in (A.14) can be written in terms of $\overrightarrow{\mathrm{SNR}}_1$, $\overrightarrow{\mathrm{SNR}}_2$, INR_{12}, INR_{21}, $\overleftarrow{\mathrm{SNR}}_1$, $\overleftarrow{\mathrm{SNR}}_2$, ρ, μ_1, and μ_2. Then, the following holds in (A.14) for the two-user GIC-NOF:

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

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 (5.1). This completes the proof of Lemma 28.

— O —

Price of Anarchy and Maximum and Minimum Sum-Rates

THIS appendix presents the maximum sum-rate in the centralized case and the maximum and minimum sum-rate in the decentralized case. Denote by $\overline{\Sigma}_C$ the maximum sum-rate in the centralized case, which is the solution to the optimization problem in the numerator of (6.8) and the numerator of (6.16). A closed-form expression of $\overline{\Sigma}_C$ is given by the Lemma 34.

Lemma 34 (Maximum sum-rate in the capacity region). *For all* $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \in \mathbb{N}^6$, $\overline{\Sigma}_{\mathcal{C}}$ *satisfies the following equality:*

$$\begin{aligned}\overline{\Sigma}_{\mathcal{C}(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})}=\min\Bigg(& \max\left(\overrightarrow{n}_{11}, n_{12}\right)+\max\left(\overrightarrow{n}_{22}, n_{21}\right), \\ & \max\left(\overrightarrow{n}_{11}, n_{12}\right)+\max\left(\overrightarrow{n}_{22}, \overleftarrow{n}_{11}-\left(\overrightarrow{n}_{11}-n_{12}\right)^+\right), \\ & \max\left(\overrightarrow{n}_{11}, \overleftarrow{n}_{22}-\left(\overrightarrow{n}_{22}-n_{21}\right)^+\right)+\max\left(\overrightarrow{n}_{22}, n_{21}\right), \\ & \max\left(\overrightarrow{n}_{11}, \overleftarrow{n}_{22}-\left(\overrightarrow{n}_{22}-n_{21}\right)^+\right)+\max\left(\overrightarrow{n}_{22}, \overleftarrow{n}_{11}-\left(\overrightarrow{n}_{11}-n_{12}\right)^+\right), \\ & \max\left(\overrightarrow{n}_{22}, n_{12}\right)+\left(\overrightarrow{n}_{11}-n_{12}\right)^+, \max\left(\overrightarrow{n}_{11}, n_{21}\right)+\left(\overrightarrow{n}_{22}-n_{21}\right)^+, \\ & \max\left(\left(\overrightarrow{n}_{11}-n_{12}\right)^+, n_{21}, \overrightarrow{n}_{11}-\left(\max\left(\overrightarrow{n}_{11}, n_{12}\right)-\overleftarrow{n}_{11}\right)^+\right) \\ & +\max\left(\left(\overrightarrow{n}_{22}-n_{21}\right)^+, n_{12}, \overrightarrow{n}_{22}-\left(\max\left(\overrightarrow{n}_{22}, n_{21}\right)-\overleftarrow{n}_{22}\right)^+\right),\end{aligned}$$

$$\max\left(\overrightarrow{n}_{11}, n_{21}\right) + \left(\overrightarrow{n}_{11} - n_{12}\right)^{+}$$
$$+ \max\Big(\left(\overrightarrow{n}_{22} - n_{21}\right)^{+}, n_{12}, \overrightarrow{n}_{22} - \left(\max\left(\overrightarrow{n}_{22}, n_{21}\right) - \overleftarrow{n}_{22}\right)^{+} \Big),$$
$$\max\left(\overrightarrow{n}_{22}, n_{12}\right) + \left(\overrightarrow{n}_{22} - n_{21}\right)^{+}$$
$$+ \max\Big(\left(\overrightarrow{n}_{11} - n_{12}\right)^{+}, n_{21}, \overrightarrow{n}_{11} - \left(\max\left(\overrightarrow{n}_{11}, n_{12}\right) - \overleftarrow{n}_{11}\right)^{+} \Big)\Bigg). \tag{O.1}$$

Proof: The proof of Lemma 34 is obtained by combining the minimum between the sum-rate bounds (4.1c)-(4.1c), the weighted sum-rate bounds (4.1d), and the sum of single rate bounds (4.1a)-(4.1b) on the capacity region of the two-user LDIC-NOF (Theorem 1) for all $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$. ■

Denote by $\overline{\Sigma}_{\mathcal{N}_\eta}$ the maximum sum-rate in the decentralized case, which is the solution to the optimization problem in the denominator of (6.16). A closed-form expression of $\overline{\Sigma}_{\mathcal{N}_\eta}$ is given by the Lemma 35.

Lemma 35 (Maximum Sum-Rate at an η-NE)**.** *For all* $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \in \mathbb{N}^6$, $\overline{\Sigma}_{\mathcal{N}_\eta(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})}$ *satisfies the following equality:*

$$\overline{\Sigma}_{\mathcal{N}_\eta(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})} = \min\Bigg(\overline{\Sigma}_{\mathcal{C}(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})}, \max\left(\overrightarrow{n}_{11}, n_{21}\right) + \max\left(\overrightarrow{n}_{22}, n_{21}\right)$$
$$- \Bigg(\min\Big(\left(\overrightarrow{n}_{11} - n_{12}\right)^{+}, n_{21}\Big) - \Big(\min\Big(\left(\overrightarrow{n}_{11} - n_{21}\right)^{+}, n_{12}\Big) - \Big(\max\left(\overrightarrow{n}_{11}, n_{12}\right) - \overleftarrow{n}_{11}\Big)^{+}\Big)^{+}\Bigg)^{+} + \eta,$$
$$\max\Big(\overrightarrow{n}_{11}, \overleftarrow{n}_{22} - \left(\overrightarrow{n}_{22} - n_{21}\right)^{+}\Big) + \max\left(\overrightarrow{n}_{22}, n_{21}\right) + \eta$$
$$- \Bigg(\min\Big(\left(\overrightarrow{n}_{11} - n_{12}\right)^{+}, n_{21}\Big) - \Big(\min\Big(\left(\overrightarrow{n}_{11} - n_{21}\right)^{+}, n_{12}\Big) - \Big(\max\left(\overrightarrow{n}_{11}, n_{12}\right) - \overleftarrow{n}_{11}\Big)^{+}\Big)^{+}\Bigg)^{+},$$
$$\max\left(\overrightarrow{n}_{11}, n_{12}\right) + \max\left(\overrightarrow{n}_{22}, n_{21}\right) + \eta$$
$$- \Bigg(\min\Big(\left(\overrightarrow{n}_{11} - n_{12}\right)^{+}, n_{21}\Big) - \Big(\min\Big(\left(\overrightarrow{n}_{11} - n_{21}\right)^{+}, n_{12}\Big) - \Big(\max\left(\overrightarrow{n}_{11}, n_{12}\right) - \overleftarrow{n}_{11}\Big)^{+}\Big)^{+}\Bigg)^{+},$$
$$\max\left(\overrightarrow{n}_{11}, n_{12}\right) + \max\left(\overrightarrow{n}_{22}, n_{21}\right) + \eta$$
$$- \Bigg(\min\Big(\left(\overrightarrow{n}_{22} - n_{21}\right)^{+}, n_{12}\Big) - \Big(\min\Big(\left(\overrightarrow{n}_{22} - n_{12}\right)^{+}, n_{21}\Big) - \Big(\max\left(\overrightarrow{n}_{22}, n_{21}\right) - \overleftarrow{n}_{22}\Big)^{+}\Big)^{+}\Bigg)^{+},$$
$$\max\left(\overrightarrow{n}_{11}, n_{12}\right) + \max\left(\overrightarrow{n}_{22}, n_{12}\right) + \eta$$
$$- \Bigg(\min\Big(\left(\overrightarrow{n}_{22} - n_{21}\right)^{+}, n_{12}\Big) - \Big(\min\Big(\left(\overrightarrow{n}_{22} - n_{12}\right)^{+}, n_{21}\Big) - \Big(\max\left(\overrightarrow{n}_{22}, n_{21}\right) - \overleftarrow{n}_{22}\Big)^{+}\Big)^{+}\Bigg)^{+},$$
$$\max\left(\overrightarrow{n}_{11}, n_{12}\right) + \max\Big(\overrightarrow{n}_{22}, \overleftarrow{n}_{11} - \left(\overrightarrow{n}_{11} - n_{12}\right)^{+}\Big) + \eta$$
$$- \Bigg(\min\Big(\left(\overrightarrow{n}_{22} - n_{21}\right)^{+}, n_{12}\Big) - \Big(\min\Big(\left(\overrightarrow{n}_{22} - n_{12}\right)^{+}, n_{21}\Big) - \Big(\max\left(\overrightarrow{n}_{22}, n_{21}\right) - \overleftarrow{n}_{22}\Big)^{+}\Big)^{+}\Bigg)^{+},$$

$$\max\left(\overrightarrow{n}_{11}, n_{12}\right) + \max\left(\overrightarrow{n}_{22}, n_{21}\right) + 2\eta$$
$$-\left(\min\left(\left(\overrightarrow{n}_{22} - n_{21}\right)^{+}, n_{12}\right) - \left(\min\left(\left(\overrightarrow{n}_{22} - n_{12}\right)^{+}, n_{21}\right) - \left(\max\left(\overrightarrow{n}_{22}, n_{21}\right) - \overleftarrow{n}_{22}\right)^{+}\right)^{+}\right)^{+}$$
$$-\left(\min\left(\left(\overrightarrow{n}_{11} - n_{12}\right)^{+}, n_{21}\right) - \left(\min\left(\left(\overrightarrow{n}_{11} - n_{21}\right)^{+}, n_{12}\right) - \left(\max\left(\overrightarrow{n}_{11}, n_{12}\right) - \overleftarrow{n}_{11}\right)^{+}\right)^{+}\right)^{+},$$
$$\max\left(\overrightarrow{n}_{11}, n_{21}\right) + \left(\overrightarrow{n}_{11} - n_{12}\right)^{+} + \max\left(\left(\overrightarrow{n}_{22} - n_{21}\right)^{+}, n_{12}, \overrightarrow{n}_{22} - \left(\max\left(\overrightarrow{n}_{22}, n_{21}\right) - \overleftarrow{n}_{22}\right)^{+}\right)$$
$$-\left(\left(\overrightarrow{n}_{11} - n_{12}\right)^{+} - \eta\right)^{+}, \max\left(\overrightarrow{n}_{11}, n_{21}\right) + \left(\overrightarrow{n}_{11} - n_{12}\right)^{+} - \left(\left(\overrightarrow{n}_{22} - n_{21}\right)^{+} - \eta\right)^{+}$$
$$+\max\left(\left(\overrightarrow{n}_{11} - n_{12}\right)^{+}, n_{21}, \overrightarrow{n}_{11} - \left(\max\left(\overrightarrow{n}_{11}, n_{12}\right) - \overleftarrow{n}_{11}\right)^{+}\right)\Bigg), \quad \text{(O.2)}$$

where, $\overline{\Sigma}_{\mathcal{C}(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})}$ *is defined in Lemma 34.*

Proof: The proof of Lemma 35 follows from obtaining the maximum sum-rate for the η-NE region of the two-user LDIC-NOF (Theorem 10). It corresponds to the minimum between the sum-rate upper-bounds (4.1c)-(4.1c), the difference between the upper bound on the weighted sum-rate (4.1d) and the lower-bound on the single rate (6.2b), *i.e.*, U_i, and the sum of upper-bounds on single rates (4.1a)-(4.1b) in Theorem 1 and (6.2b), for all $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$. ■

Denote by $\underline{\Sigma}_{\mathcal{N}_\eta}$ the minimum sum-rate in the decentralized case, which is the solution to the optimization problem in the denominator of (6.8). A closed-form expression of $\underline{\Sigma}_{\mathcal{N}_\eta}$ is given by the Lemma 36.

Lemma 36 (Minimum Sum-Rate at an η-NE)**.** *For all* $(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}) \in \mathbb{N}^6$, $\underline{\Sigma}_{\mathcal{N}_\eta(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})}$ *satisfies the following equality:*

$$\underline{\Sigma}_{\mathcal{N}_\eta(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})} = \left(\left(\overrightarrow{n}_{11} - n_{12}\right)^{+} - \eta\right)^{+} + \left(\left(\overrightarrow{n}_{22} - n_{21}\right)^{+} - \eta\right)^{+}. \quad \text{(O.3)}$$

Proof: The proof of Lemma 36 follows from obtaining the minimum sum-rate for the η-NE region of the two-user LDIC-NOF (Theorem 10) and it is obtained as the sum of the lower-bounds on the single rates in (6.2a), *i.e.*, L_i, for all $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$. ■

O.1. PoA when both Transmitter-Receiver Pairs are in the Low-Interference Regime

When both transmitter-receiver pairs are in LIR, *i.e.*, $\overrightarrow{n}_{11} > n_{12}$ and $\overrightarrow{n}_{22} > n_{21}$, and assuming that $\overleftarrow{n}_{ii} \leqslant \max\left(\overrightarrow{n}_{ii}, n_{ij}\right)$ for all $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$, the following holds:

$$\overline{\Sigma}_{\mathcal{C}(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})} = \min\Bigg(\max\left(\overrightarrow{n}_{22}, n_{12}\right) + \overrightarrow{n}_{11} - n_{12}, \max\left(\overrightarrow{n}_{11}, n_{21}\right) + \overrightarrow{n}_{22} - n_{21},$$
$$\max\left(\overrightarrow{n}_{11} - n_{12}, n_{21}, \overleftarrow{n}_{11}\right) + \max\left(\overrightarrow{n}_{22} - n_{21}, n_{12}, \overleftarrow{n}_{22}\right),$$

$$\max\left(\overrightarrow{n}_{11}, n_{21}\right) + \overrightarrow{n}_{11} - n_{12} + \max\left(\overrightarrow{n}_{22} - n_{21}, n_{12}, \overleftarrow{n}_{22}\right),$$
$$\max\left(\overrightarrow{n}_{22}, n_{12}\right) + \overrightarrow{n}_{22} - n_{21} + \max\left(\overrightarrow{n}_{11} - n_{12}, n_{21}, \overleftarrow{n}_{11}\right)\Big) \tag{O.4}$$

and

$$\begin{aligned}\underline{\Sigma}_{\mathcal{N}_\eta(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})} =& \overrightarrow{n}_{11} - n_{12} + \overrightarrow{n}_{22} - n_{21} - 2\eta \\ =& \underline{\Sigma}_{\mathcal{N}1}.\end{aligned} \tag{O.5}$$

Then, the PoA when both transmitter-receiver pairs are in LIR can be calculated using (O.4) and (O.5).

If B_1 in (6.7c) holds true, the following holds:

$$\begin{aligned}\overline{\Sigma}_{\mathcal{C}(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})} =& \min\Big(\overrightarrow{n}_{22} + \overrightarrow{n}_{11} - n_{12}, \overrightarrow{n}_{11} + \overrightarrow{n}_{22} - n_{21}, \\ & \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right) + \max\left(\overrightarrow{n}_{22} - n_{21}, \overleftarrow{n}_{22}\right), \\ & 2\overrightarrow{n}_{11} - n_{12} + \max\left(\overrightarrow{n}_{22} - n_{21}, \overleftarrow{n}_{22}\right), \\ & 2\overrightarrow{n}_{22} - n_{21} + \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right)\Big) \\ =& \overline{\Sigma}_{\mathcal{C}1},\end{aligned} \tag{O.6}$$

and this proves (6.10a).

If $B_{2,i}$ in (6.7d) with $i = 1$ and $j = 2$ holds true, the following holds:

$$\begin{aligned}\overline{\Sigma}_{\mathcal{C}(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})} =& \min\Big(\overrightarrow{n}_{22} + \overrightarrow{n}_{11} - n_{12}, \overrightarrow{n}_{11} + \overrightarrow{n}_{22} - n_{21}, \\ & \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right) + \max\left(n_{12}, \overleftarrow{n}_{22}\right), \\ & 2\overrightarrow{n}_{11} - n_{12} + \max\left(n_{12}, \overleftarrow{n}_{22}\right), \\ & 2\overrightarrow{n}_{22} - n_{21} + \max\left(\overrightarrow{n}_{11} - n_{12}, \overleftarrow{n}_{11}\right)\Big) \\ =& \overline{\Sigma}_{\mathcal{C}2,1},\end{aligned} \tag{O.7}$$

and this proves (6.10b) with $i = 1$ and $j = 2$. The same procedure can be followed when $B_{2,i}$ in (6.7d) with $i = 2$ and $j = 1$ holds true.

If $B_{3,i}$ in (6.7e) or $B_{5,i}$ in (6.7g) with $i = 1$ and $j = 2$ holds true, the following holds:

$$\overline{\Sigma}_{\mathcal{C}(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})} = \overrightarrow{n}_{11}. \tag{O.8}$$

The same procedure can be followed when $B_{3,i}$ in (6.7e) or $B_{5,i}$ in (6.7g) with $i = 2$ and $j = 1$ holds true.

If B_4 in (6.7f) holds true, the following holds:

$$\begin{aligned}\overline{\Sigma}_{\mathcal{C}(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})}=&\min\Big(\overrightarrow{n}_{22}+\overrightarrow{n}_{11}-n_{12},\overrightarrow{n}_{11}+\overrightarrow{n}_{22}-n_{21},\\&\max\left(n_{21},\overleftarrow{n}_{11}\right)+\max\left(n_{12},\overleftarrow{n}_{22}\right),\\&2\overrightarrow{n}_{11}-n_{12}+\max\left(n_{12},\overleftarrow{n}_{22}\right),\\&2\overrightarrow{n}_{22}-n_{21}+\max\left(n_{21},\overleftarrow{n}_{11}\right)\Big)\\=&\overline{\Sigma}_{\mathcal{C}3},\end{aligned} \tag{O.9}$$

and this proves (6.10c). Plugging (O.5), (O.6), (O.7), (O.8), and (O.9) into (6.8) yields (6.9), and this completes the proof of Theorem 11.

O.2. PoA when Transmitter-Receiver Pair 1 is in the Low-Interference Regime and Transmitter-Receiver Pair 2 is in the High-Interference Regime

When transmitter-receiver pair 1 is in LIR, *i.e.*, $\overrightarrow{n}_{11} > n_{12}$, and transmitter-receiver pair 2 is in HIR, i.e, $\overrightarrow{n}_{22} \leqslant n_{21}$, and assuming that $\overleftarrow{n}_{ii} \leqslant \max\left(\overrightarrow{n}_{ii}, n_{ij}\right)$ for all $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$, the following holds:

$$\begin{aligned}\overline{\Sigma}_{\mathcal{C}(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})}=&\min\Big(\overrightarrow{n}_{11}+\max\left(\overrightarrow{n}_{22},\overleftarrow{n}_{11}-\overrightarrow{n}_{11}+n_{12}\right),\\&\max\left(\overrightarrow{n}_{22},n_{12}\right)+\overrightarrow{n}_{11}-n_{12},\max\left(\overrightarrow{n}_{11},n_{21}\right),\\&\max\left(\overrightarrow{n}_{11}-n_{12},n_{21},\overleftarrow{n}_{11}\right)+\max\left(n_{12},\overrightarrow{n}_{22}-n_{21}+\overleftarrow{n}_{22}\right)\Big)\end{aligned} \tag{O.10}$$

and

$$\underline{\Sigma}_{\mathcal{N}_{\eta}(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})}=\overrightarrow{n}_{11}-n_{12}-\eta. \tag{O.11}$$

If B_7 in (6.7i), or B_8 in (6.7j), or B_{10} in (6.7l) holds true, the following holds:

$$\overline{\Sigma}_{\mathcal{C}(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})}=\overrightarrow{n}_{11}. \tag{O.12}$$

If B_9 in (6.7k) holds true, the following holds:

$$\overline{\Sigma}_{\mathcal{C}(\overrightarrow{n}_{11},\overrightarrow{n}_{22},n_{12},n_{21},\overleftarrow{n}_{11},\overleftarrow{n}_{22})}=\min\left(\overrightarrow{n}_{22}+\overrightarrow{n}_{11}-n_{12},n_{21}\right). \tag{O.13}$$

Plugging (O.11), (O.12), and (O.13) into (6.8) yields (6.14), and this completes the proof of Theorem 12.

— P —

Information Measures

This chapter introduces some information measures that are used along this thesis. These fundamental notions correspond to Shannon's original measures [85], which build the foundations of information theory. Shannon's information measures include entropy, joint entropy, conditional entropy, mutual information, and conditional mutual information.

P.1. Discrete Random Variables

P.1.1. Entropy

The entropy $H(X)$ of a discrete random variable X is a functional of the pmf P_X, which measures the average amount of information contained into X.

Definition 9 (Entropy)**.** *Let $\mathcal{X}$ be a countable set and let also X be a random variable with pmf $P_X : \mathcal{X} \to [0,1]$. Then, the entropy of X, denoted by $H(X)$, is:*

$$H(X) = -\sum_{x \in \text{supp}(P_X)} P_X(x) \log P_X(x). \tag{P.1}$$

This entropy is measured in bits given that the base of the logarithm is two. Note that $H(X)$ depends only on P_X and not on the elements of $\mathcal{X}$.

The entropy of a random variable X can also be written as follows:

$$H(X) = -\mathbb{E}_X \left[\log P_X(X)\right]. \tag{P.2}$$

For each $x \in \text{supp}(P_X)$, define $\imath(x) = -\log P_X(x)$. Then, $\imath$ is a new random variable, and $H(X)$ is its average. The function $\imath(x)$ can be interpreted as the amount of information provided by the event $X = x$ (See Figure P.1) [32]. According to this interpretation, an unlikely event provides a very large amount of information and an event that occurs with probability close to one does not provide information [55].

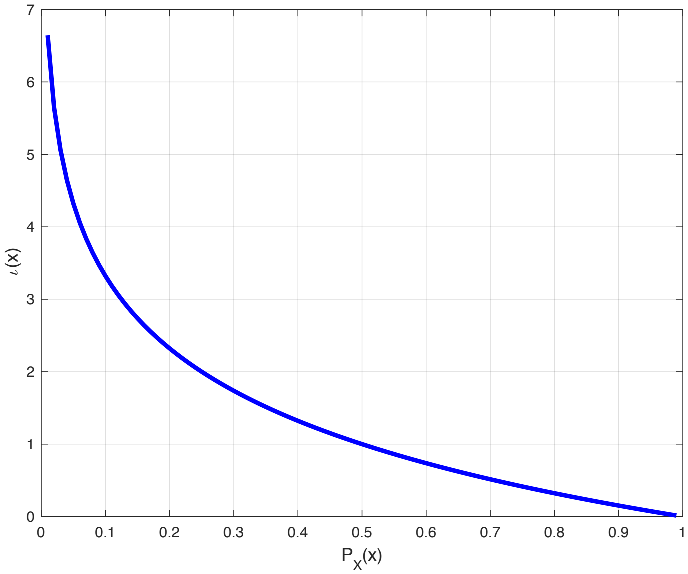


Figure P.1.: The function $\imath(x)$.

The following corollary presents the entropy of a binary random variable.

Corollary 11. *Let X be a binary random variable following the distribution P_X such that $P_X(0) = 1 - P_X(1) = p$ and $0 \leqslant p \leqslant 1$. Then,*

$$H(X) = \begin{cases} 0 & \text{if } p = 0 \\ -p \log p - (1-p) \log(1-p) & \text{otherwise} \end{cases}. \tag{P.3}$$

The binary entropy function in (P.3) is plotted in Figure P.2 as a function of p. It is worth noting that the binary entropy function is a non-negative function with a maximum equal to one when $P_X(0) = 1 - P_X(1) = \frac{1}{2}$ (uniform distribution).

In general, the entropy takes non-negative values, *i.e.*, $H(X) \geqslant 0$, with equality if and only if X is non-random. The entropy takes its maximum value when all the events have the same probability, *i.e.*, $H(X) = \log(|\mathcal{X}|)$, as stated by the following lemma:

Lemma 37. *Let $\mathcal{X}$ be a countable set and let also X be a random variable with pmf $P_X : \mathcal{X} \rightarrow [0, 1]$. Then,*

$$0 \leqslant H(X) \leqslant \log |\mathcal{X}|. \tag{P.4}$$

Proof: The lower-bound on the entropy of a random variable X is obtained from the fact that for all $x \in \text{supp}(P_X)$, $0 < P_X(x) \leqslant 1$, then $\frac{1}{P_X(x)} \geqslant 1$ and $\log\left(\frac{1}{P_X(x)}\right) \geqslant 0$. Thus, $H(X) \geqslant 0$.

The upper-bound on the entropy of the random variable X is also obtained from (P.2) as

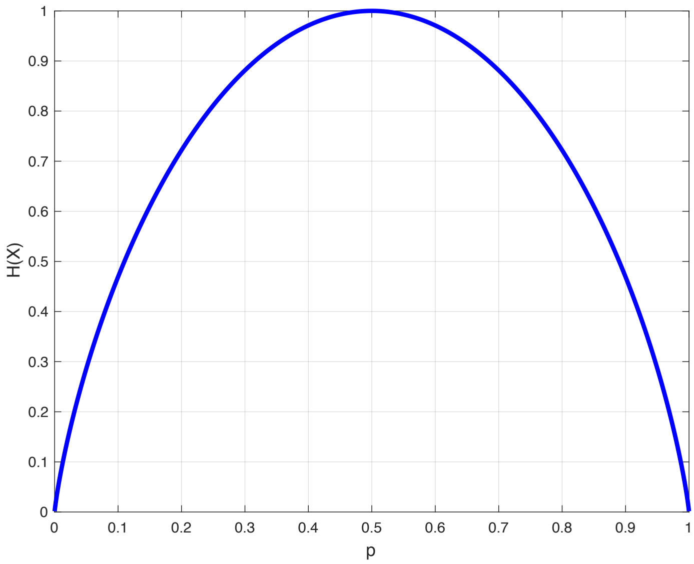


Figure P.2.: Entropy of a binary random variable.

follows:

$$H(X) = \mathbb{E}_X \left[\log \frac{1}{P_X(X)} \right] \tag{P.5a}$$

$$\leqslant \log \mathbb{E}_X \left[\frac{1}{P_X(X)} \right] \tag{P.5b}$$

$$= \log \sum_{x \in \mathrm{supp}(P_X)} 1 \tag{P.5c}$$

$$= \log |\mathcal{X}|, \tag{P.5d}$$

where, (P.5b) follows from Jensen's inequality. Thus, the maximum value of the entropy of a random variable X is obtained when it is uniformly distributed, *i.e.*, $P_X(x) = \frac{1}{|\mathcal{X}|}$ for all $x \in \mathrm{supp}(P_X)$. This completes the proof of Lemma 37. ■

P.1.2. Joint Entropy

The joint entropy $H(X, Y)$ of the discrete random variables X and Y is a functional of the pmf $P_{X,Y}$, which measures the average amount of information simultaneously contained into X and Y. It is a measure of the uncertainty about the simultaneous outcome of the random variables X and Y.

Definition 10 (Joint Entropy). *Let $\mathcal{X}$ and $\mathcal{Y}$ be two countable sets and let also X and Y be two random variables with joint pmf $P_{XY} : \mathcal{X} \times \mathcal{Y} \to [0, 1]$. Then, the joint entropy of X and Y, denoted by $H(X, Y)$, is:*

$$H(X, Y) = - \sum_{(x,y) \in \mathrm{supp}(P_{XY})} P_{XY}(x, y) \log P_{XY}(x, y). \tag{P.6}$$

The joint entropy of the random variables X and Y can also be written as follows:

$$H(X,Y) = -\mathbb{E}_{XY}\left[\log P_{XY}(X,Y)\right]. \tag{P.7}$$

The joint entropy between two random variables is less than or equal to the sum of the entropy of each random variable, as stated by the following lemma.

Lemma 38. *Let $\mathcal{X}$ and $\mathcal{Y}$ be two countable sets and let also X and Y be two random variables with joint pmf $P_{XY} : \mathcal{X} \times \mathcal{Y} \rightarrow [0,1]$. Then,*

$$H(X,Y) \leqslant H(X) + H(Y), \tag{P.8}$$

with equality if and only if the random variables X and Y are independent.

Proof: From (P.7), the following holds:

$$\begin{aligned}
H(X,Y) &= -\mathbb{E}_{XY}\left[\log\left(\frac{P_X(X)P_Y(Y)P_{XY}(X,Y)}{P_X(X)P_Y(Y)}\right)\right]. && \text{(P.9a)}\\
&= -\mathbb{E}_X\left[\log P_X(X)\right] - \mathbb{E}_Y\left[\log P_Y(Y)\right] - \mathbb{E}_{XY}\left[\log\left(\frac{P_{XY}(X,Y)}{P_X(X)P_Y(Y)}\right)\right] && \text{(P.9b)}\\
&= H(X) + H(Y) + \mathbb{E}_{XY}\left[\log\left(\frac{P_X(X)P_Y(Y)}{P_{X,Y}(XY)}\right)\right] && \text{(P.9c)}\\
&\leqslant H(X) + H(Y) + \log\left(\mathbb{E}_{XY}\left[\left(\frac{P_X(X)P_Y(Y)}{P_{XY}(X,Y)}\right)\right]\right) && \text{(P.9d)}\\
&= H(X) + H(Y) + \log\left(\sum_{(x,y)\in\mathrm{supp}(P_{XY})} P_X(X)P_Y(Y)\right) && \text{(P.9e)}\\
&= H(X) + H(Y), && \text{(P.9f)}
\end{aligned}$$

where, (P.9d) follows from Jensen's inequality.

If the random variables X and Y are independent, from (P.9c) the following holds:

$$\begin{aligned}
H(X,Y) &= H(X) + H(Y) + \mathbb{E}_{XY}\left[\log\left(\frac{P_X(X)P_Y(Y)}{P_X(X)P_Y(Y)}\right)\right] && \text{(P.10a)}\\
&= H(X) + H(Y), && \text{(P.10b)}
\end{aligned}$$

and this completes the proof of Lemma 38. ■

Definition 11 generalizes Definition 10.

Definition 11. *Let $\mathcal{X}_1$, $\mathcal{X}_2$, ..., $\mathcal{X}_N$ be $N \in \mathbb{N}$ countable sets and let also $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)^{\mathsf{T}}$ be a vector of N random variables with joint pmf $P_{\boldsymbol{X}} : \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_N \rightarrow [0,1]$. Then, the joint entropy of $\boldsymbol{X}$, denoted by $H(\boldsymbol{X})$, is:*

$$H(\boldsymbol{X}) = -\sum_{\boldsymbol{x}\in\mathrm{supp}(P_{\boldsymbol{X}})} P_{\boldsymbol{X}}(\boldsymbol{x}) \log P_{\boldsymbol{X}}(\boldsymbol{x}). \tag{P.11}$$

The joint entropy of a vector of discrete random variables $\boldsymbol{X}$ can also be written as follows:

$$H(\boldsymbol{X}) = -\mathbb{E}_{\boldsymbol{X}}\left[\log P_{\boldsymbol{X}}(\boldsymbol{X})\right]. \tag{P.12}$$

If $X_1, X_2, \ldots, X_N$ are mutually independent, the following holds:

$$H(\boldsymbol{X}) = \sum_{n=1}^{N} H(X_n). \tag{P.13}$$

P.1.3. Conditional Entropy

The conditional entropy $H(Y|X)$ of Y given X is a measure of the average amount of information necessary to identify the random variable Y given the observation of the random variable X.

Definition 12 (Conditional Entropy). *Let $\mathcal{X}$ and $\mathcal{Y}$ be two countable sets and let also X and Y be two random variables with joint pmf $P_{XY} : \mathcal{X} \times \mathcal{Y} \to [0,1]$. Then, the entropy of Y conditioning on X, denoted by $H(Y|X)$, is:*

$$H(Y|X) = -\sum_{(x,y)\in\text{supp}(P_{XY})} P_{XY}(x,y) \log P_{Y|X}(y|x). \tag{P.14}$$

The entropy of the random variable Y conditioning on the random variable X can also be written as follows:

$$H(Y|X) = -\mathbb{E}_{XY}\left[\log P_{Y|X}(Y|X)\right]. \tag{P.15}$$

Note also that the conditional entropy in (P.14) can be written as follows:

$$\begin{aligned} H(Y|X) &= \sum_{x\in\text{supp}(P_X)} P_X(x) \left[- \sum_{y\in\text{supp}\left(P_{Y|X=x}\right)} P_{Y|X}(y|x) \log P_{Y|X}(y|x) \right] \\ &= \sum_{x\in\text{supp}(P_X)} P_X(x) H(Y|X=x), \end{aligned} \tag{P.16}$$

where, $H(Y|X=x) = -\sum_{y\in\text{supp}(P_Y)} P_{Y|X}(y|x) \log P_{Y|X}(y|x)$ is the entropy of Y conditioning on a fixed $X = x$.

The following lemma presents a generalization of the chain rule for entropy and the conditional entropy.

Lemma 39 (Chain rule for entropy and chain rule for conditional entropy). *Let $\mathcal{X}_1$, $\mathcal{X}_2$, …, $\mathcal{X}_N$ and $\mathcal{Y}$ be $N+1$ countable sets, let $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)^{\mathsf{T}}$ be a vector of N random variables, and let also Y be a random variable with joint pmfs $P_{\boldsymbol{X}} : \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_N \to [0,1]$ and $P_{\boldsymbol{X}Y} : \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_N \times \mathcal{Y} \to [0,1]$. Then,*

$$H(X_1, \ldots, X_N) = H(X_1) + H(X_2|X_1) + \sum_{n=3}^{N} H(X_n|X_1, \ldots, X_{n-1}), \text{ and} \tag{P.17}$$

$$H(X_1, \ldots, X_N|Y) = H(X_1|Y) + H(X_2|Y, X_1) + \sum_{n=3}^{N} H(X_n|Y, X_1, \ldots, X_{n-1}). \tag{P.18}$$

Proof:

Proof of (P.17): From (P.12), the following holds:

$$H(\boldsymbol{X}) = -\mathbb{E}_{\boldsymbol{X}}\left[\log\left(P_{X1}(X_1)P_{X_2|X_1}(X_2|X_1)\dots P_{X_N|X_1X_2\dots X_{N-1}}(X_N|X_1,X_2,\dots,X_{N-1})\right)\right] \tag{P.19a}$$

$$= -\mathbb{E}_{X_1}\left[\log P_{X1}(X_1)\right] - \mathbb{E}_{X_1X_2}\left[\log P_{X_2|X_1}(X_2|X_1)\right] - \dots - \mathbb{E}_{\boldsymbol{X}}\left[\log P_{X_N|X_1,X_2,\dots,X_{N-1}}\right] \tag{P.19b}$$

$$= H(X_1) + H(X_2|X_1) + \dots + H(X_N|X_1,X_2,\dots,X_{N-1}), \tag{P.19c}$$

and this completes the proof of (P.17).

Proof of (P.18): From (P.15), the following holds:

$$H(\boldsymbol{X}|Y) = -\mathbb{E}_{\boldsymbol{X}Y}\left[\log P_{\boldsymbol{X}|Y}(\boldsymbol{X}|Y)\right] \tag{P.20a}$$

$$-\mathbb{E}_{\boldsymbol{X}Y}\Big[\log\Big(P_{X1|Y}(X_1|Y)P_{X_2|X_1Y}(X_2|Y,X_1)\dots P_{X_N|X_1X_2\dots X_{N-1}Y}(X_N|Y,X_1,X_2,\dots,X_{N-1})\Big)\Big] \tag{P.20b}$$

$$= -\mathbb{E}_{X_1Y}\left[\log P_{X1|Y}(X_1|Y)\right] - \mathbb{E}_{X_1X_2Y}\left[\log P_{X_2|X_1Y}(X_2|Y,X_1)\right] - \dots - \mathbb{E}_{\boldsymbol{X}Y}\left[\log P_{X_N|X_1,X_2,\dots,X_{N-1}Y}\right] \tag{P.20c}$$

$$= H(X_1|Y) + H(X_2|Y,X_1) + \dots + H(X_N|Y,X_1,X_2,\dots,X_{N-1}), \tag{P.20d}$$

and this completes the proof of (P.18). This completes the proof of Lemma 39. ■

Conditioning a random variable on another one cannot increase the a priori uncertainty on its realization. Thus, conditioning does not increase entropy as formalized in the following lemma.

Lemma 40 (Conditioning does not increase entropy)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be two countable sets and let also X and Y be two random variables with joint pmf $P_{XY} : \mathcal{X} \times \mathcal{Y} \to [0,1]$. Then,*

$$H(Y|X) \leqslant H(Y), \tag{P.21}$$

with equality if and only if the random variables X and Y are independent.

Proof: From (P.15), the following holds:

$$H(Y|X) = -\mathbb{E}_{XY}\left[\log\left(\frac{P_Y(Y)P_{X|Y}(X|Y)}{P_X(X)}\right)\right] \tag{P.22a}$$

$$= -\mathbb{E}_Y\left[\log P_Y(Y)\right] - \mathbb{E}_{XY}\left[\log\left(\frac{P_{X|Y}(X|Y)}{P_X(X)}\right)\right] \tag{P.22b}$$

$$= H(Y) - \mathbb{E}_{XY}\left[\log\left(\frac{P_{XY}(X,Y)}{P_X(X)P_Y(Y)}\right)\right] \tag{P.22c}$$

$$= H(Y) + \mathbb{E}_{XY}\left[\log\left(\frac{P_X(X)P_Y(Y)}{P_{XY}(X,Y)}\right)\right] \tag{P.22d}$$

$$\leqslant H(Y) + \log\left(\mathbb{E}_{XY}\left[\left(\frac{P_X(X)P_Y(Y)}{P_{XY}(X,Y)}\right)\right]\right) \tag{P.22e}$$

$$=H(Y)+\log\left(\sum_{(x,y)\in\text{supp}(P_{XY})} P_X(X)P_Y(Y)\right) \quad \text{(P.22f)}$$

$$=H(Y), \quad \text{(P.22g)}$$

where (P.22e) follows from Jensen's inequality.

If the random variables X and Y are independent, from (P.22d) the following holds:

$$H(Y|X)=H(Y)+\mathbb{E}_{XY}\left[\log\left(\frac{P_X(X)P_Y(Y)}{P_X(X)P_Y(Y)}\right)\right] \quad \text{(P.23a)}$$

$$=H(Y), \quad \text{(P.23b)}$$

and this completes the proof of Lemma 40. ■

The joint entropy of a vector of random variables is less than or equal to the sum of the entropy of each random variable, as stated by the following lemma.

Lemma 41. *Let $\mathcal{X}_1$, $\mathcal{X}_2$, …, $\mathcal{X}_N$ be $N \in \mathbb{N}$ countable sets and let also $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)^T$ be a vector of N random variables with joint pmf $P_{\boldsymbol{X}} : \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_N \to [0, 1]$. Then,*

$$H(X_1, \ldots, X_N) \leqslant \sum_{n=1}^{N} H(X_N), \quad \text{(P.24)}$$

with equality if and only if the random variables $X_1, X_2, \ldots, X_N$ are mutually independent.

Proof: The proof of Lemma 41 follows by combining Lemma 39 and Lemma 40. ■

The entropy of a deterministic function of the random variable X is less than or equal to the entropy of the random variable X, with equality only when the function is an injective function. This is stated in Lemma 42.

Lemma 42 (Entropy of a function)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be countable sets, let X be a random variable with pmf $P_X : \mathcal{X} \to [0, 1]$, and let also $f : \mathcal{X} \to \mathcal{Y}$ be a deterministic function of X. Then,*

$$H(X) \geqslant H(f(X)). \quad \text{(P.25)}$$

Proof: Let Y be a random variable with $Y = f(X)$ and $f : \mathcal{X} \to \mathcal{Y}$. From (P.2), the following holds:

$$H(Y)=-\mathbb{E}_Y\left[\log P_Y(Y)\right] \quad \text{(P.26a)}$$

$$\leqslant-\mathbb{E}_X\left[\log P_X(X)\right] \quad \text{(P.26b)}$$

$$=H(X), \quad \text{(P.26c)}$$

where (P.26b) follows from the fact that $P_Y(y) = \sum_{x\in\text{supp}(P_X), y=f(x)} P_X(x)$, which implies that $P_Y(y) \geqslant P_X(x)$ and $-\log P_Y(y) \leqslant -\log P_X(x)$. If f is an injective function $P_Y(y) = P_X(x)$, then $H(Y) = H(X)$. This completes the proof of Lemma 42. ■

P.1.4. Mutual Information

The mutual information $I(X; Y)$ between the random variables X and Y is the average amount of information about one of the random variables provided by the occurrence of the other random variable.

Definition 13 (Mutual Information)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be two countable sets and let also X and Y be two random variables with joint pmf $P_{XY}:\mathcal{X}\times\mathcal{Y}\rightarrow[0,1]$. Then, the mutual information between X and Y, denoted by $I(X;Y)$, is:*

$$I(X;Y) = -\sum_{(x,y)\in\text{supp}(P_{XY})} P_{XY}(x,y)\log\left(\frac{P_{XY}(x,y)}{P_X(x)P_Y(y)}\right). \tag{P.27}$$

The mutual information between the random variables X and Y can also be written as follows:

$$I(X;Y)=\mathbb{E}_{XY}\left[\log\left(\frac{P_{XY}(X,Y)}{P_X(X)P_Y(Y)}\right)\right] \tag{P.28a}$$

$$=\mathbb{E}_{XY}\left[\log\left(\frac{P_{Y|X}(Y|X)}{P_Y(Y)}\right)\right] \tag{P.28b}$$

$$=\mathbb{E}_{XY}\left[\log\left(\frac{P_{X|Y}(X|Y)}{P_X(X)}\right)\right]. \tag{P.28c}$$

The following lemma presents some useful properties of the mutual information.

Lemma 43. *Let $\mathcal{X}$ and $\mathcal{Y}$ be two countable sets and let also X and Y be two random variables with joint pmf $P_{XY}:\mathcal{X}\times\mathcal{Y}\rightarrow[0,1]$. Then,*

$$I(X;Y)=I(Y;X), \tag{P.29}$$

$$I(X;Y)=H(X)-H(X|Y), \tag{P.30}$$

$$I(X;Y)=H(Y)-H(Y|X), \tag{P.31}$$

$$I(X;Y)\geqslant 0, \tag{P.32}$$

$$I(X;Y)=H(X)+H(Y)-H(X,Y), \tag{P.33}$$

$$I(X;X)=H(X). \tag{P.34}$$

Proof:

Proof of (P.29): This follows directly from Definition 13.

Proof of (P.30): From (P.28c), the following holds:

$$I(X;Y)=-\mathbb{E}_X\left[\log P_X(X)\right]+\mathbb{E}_{XY}\left[\log P_{X|Y}(X|Y)\right] \tag{P.35a}$$

$$=H(X)-H(X|Y), \tag{P.35b}$$

and this completes the proof of (P.30).

Proof of (P.31): From (P.28b), the following holds:

$$I(X;Y)=\mathbb{E}_{XY}\left[\log\left(\frac{P_{Y|X}(Y|X)}{P_Y(Y)}\right)\right] \tag{P.36a}$$

$$=-\mathbb{E}_Y\left[\log P_Y(Y)\right]+\mathbb{E}_{XY}\left[\log P_{Y|X}(Y|X)\right] \tag{P.36b}$$

$$=H(Y)-H(Y|X), \tag{P.36c}$$

and this completes the proof of (P.31).

Proof of (P.32): From (P.30) and (P.31), the following holds:

$$I(X;Y) \geqslant H(X) - H(X) \tag{P.37a}$$
$$= 0, \tag{P.37b}$$

where, (P.37a) follows from Lemma 40. This completes the proof of (P.32).

Proof of (P.33): From (P.28a), the following holds:

$$I(X;Y) = -\mathbb{E}_X\left[\log P_X(X)\right] - \mathbb{E}_Y\left[\log P_Y(Y)\right] + \mathbb{E}_{XY}\left[\log P_{XY}(X,Y)\right] \tag{P.38a}$$
$$= H(X) + H(Y) - H(X,Y), \tag{P.38b}$$

and this completes the proof of (P.33).

Proof of (P.34): Let Y be a random variable identical to the random variable X, *i.e.*, $Y = X$. From (P.28a), the following holds:

$$I(X;X) = \mathbb{E}_{XY}\left[\log\left(\frac{P_{XY}(X,Y)}{P_X(X)P_Y(Y)}\right)\right] \tag{P.39a}$$
$$= \mathbb{E}_X\left[\log\left(\frac{P_X(X)}{P_X(X)P_X(X)}\right)\right] \tag{P.39b}$$
$$= \mathbb{E}_X\left[\log\left(\frac{1}{P_X(X)}\right)\right] \tag{P.39c}$$
$$= -\mathbb{E}_X\left[\log P_X(X)\right] \tag{P.39d}$$
$$= H(X), \tag{P.39e}$$

and this completes the proof of (P.34). This completes the proof of Lemma 43. ∎

The mutual information between two independent random variables is equal to zero. This means that the occurrence of one random variable does not provide information about the occurrence of the other random variable. This is stated by the following lemma.

Lemma 44 (Mutual information of independent random variables)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be two countable sets and let also X and Y be two independent random variables with pmfs $P_X : \mathcal{X} \rightarrow [0,1]$ and $P_Y : \mathcal{Y} \rightarrow [0,1]$. Then,*

$$I(X;Y) = 0, \tag{P.40}$$

Proof: From the assumption of the lemma, it follows that $P_{XY}(x,y) = P_X(x)P_Y(y)$ for all $(x,y) \in \mathcal{X} \times \mathcal{Y}$. Hence, from (P.28a) the following holds:

$$I(X;Y) = \mathbb{E}_{XY}\left[\log\left(\frac{P_X(X)P_Y(Y)}{P_X(X)P_Y(Y)}\right)\right] \tag{P.41a}$$
$$= \mathbb{E}_{XY}\left[\log 1\right] \tag{P.41b}$$
$$= 0, \tag{P.41c}$$

and this completes the proof. ∎

The mutual information between a random variable X and two random variables Y and Z is bigger than or equal to the mutual information between the random variable X and one of the random variables Y and Z. This is stated by the following lemma.

Lemma 45. *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be three countable sets and let also X, Y, and Z be three random variables with joint pmf $P_{XYZ} : \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \to [0,1]$. Then,*

$$I(X;Y,Z) \geqslant I(X;Y), \tag{P.42}$$

with equality if and only if $X \to Y \to Z$.

Proof: From (P.30), the following holds:

$$\begin{aligned} I(X;Y,Z) &= H(Y,Z) - H(Y,Z|X) && \text{(P.43a)}\\ &= H(Y) + H(Z|Y) - H(Y|X) - H(Z|X,Y) && \text{(P.43b)}\\ &= I(X;Y) + H(Z|Y) - H(Z|X,Y) && \text{(P.43c)}\\ &\geqslant I(X;Y), && \text{(P.43d)} \end{aligned}$$

where, (P.43d) follows from the fact the fact that $H(Z|Y) - H(Z|X,Y) \geqslant 0$ given that conditioning does not increase entropy (Lemma 40). Note that the equality holds if $H(Z|Y) - H(Z|X,Y) = H(Z|Y) - H(Z|Y) = 0$. This means that the random variables X and Z are independent conditioning on the random variable Y, *i.e.*, $X \to Y \to Z$. This completes the proof of Lemma 45. ■

P.1.5. Conditional Mutual Information

Definition 14 (Conditional Mutual Information)**.** *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be three countable sets and let X, Y and Z be three random variables with joint pmf $P_{XYZ} : \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \to [0,1]$. Then, the mutual information between X and Y conditioning on Z, denoted by $I(X;Y|Z)$, is:*

$$I(X;Y|Z) = -\sum_{(x,y,z)\in \text{supp}(P_{XYZ})} P_{XYZ}(x,y,z) \log\left(\frac{P_{XY|Z}(x,y|z)}{P_{X|Z}(x|z)P_{Y|Z}(y|z)}\right). \tag{P.44}$$

The mutual information between the random variables X and Y conditioning on the random variable Z can also be written as follows:

$$\begin{aligned} I(X;Y|Z) &= \mathbb{E}_{XYZ}\left[\log\left(\frac{P_{XY|Z}(X,Y|Z)}{P_{X|Z}(X|Z)P_{Y|Z}(Y|Z)}\right)\right] && \text{(P.45a)}\\ &= \mathbb{E}_{XYZ}\left[\log\left(\frac{P_{Y|XZ}(Y|X,Z)}{P_{Y|Z}(Y|Z)}\right)\right] && \text{(P.45b)}\\ &= \mathbb{E}_{XYZ}\left[\log\left(\frac{P_{X|YZ}(X|Y,Z)}{P_{X|Z}(X|Z)}\right)\right]. && \text{(P.45c)} \end{aligned}$$

Note also that the conditional mutual information in (P.44) can be written as follows:

$$\begin{aligned} I(X;Y|Z) &= \sum_{z\in \text{supp}(P_Z)} P_Z(z) \left[-\sum_{(x,y)\in \text{supp}\left(P_{XY|Z=z}\right)} P_{X,Y|Z}(x,y|z) \log\left(\frac{P_{XY|Z}(x,y|z)}{P_{X|Z}(x|z)P_{Y|Z}(y|z)}\right)\right]\\ &= \sum_{z\in \text{supp}(P_Z)} P_Z(z) I(X;Y|Z=z), && \text{(P.46)} \end{aligned}$$

where, $I(X;Y|Z=z) = -\sum_{(x,y)\in\text{supp}\left(P_{XY|Z=z}\right)} P_{XY|Z}(x,y|z)\log\left(\frac{P_{XY|Z}(x,y|z)}{P_{X|Z}(x|z)P_{Y|Z}(y|z)}\right)$ is the mutual information between X and Y conditioning on a fixed $Z=z$.

The following lemma presents some useful properties of the mutual information and conditional mutual information.

Lemma 46. *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be three countable sets and let X, Y and Z be three random variables with joint pmf $P_{XYZ}: \mathcal{X}\times\mathcal{Y}\times\mathcal{Z}\rightarrow[0,1]$. Then,*

$$I(X;Y|Z)=H(Y|Z)-H(Y|X,Z) \tag{P.47}$$
$$=H(X|Z)-H(X|Y,Z) \text{ and} \tag{P.48}$$
$$I(X,Y;Z)=I(X;Z)+I(Y;Z|X) \tag{P.49}$$
$$=I(Y;Z)+I(X;Z|Y). \tag{P.50}$$

Proof:

Proof of (P.47): From (P.45b), the following holds:

$$I(X;Y|Z)=\mathbb{E}_{XYZ}\left[\log P_{Y|XZ}(Y|X,Z)\right]-\mathbb{E}_{YZ}\left[\log P_{Y|Z}(Y|Z)\right] \tag{P.51a}$$
$$=H(Y|Z)-H(Y|X,Z), \tag{P.51b}$$

and this completes the proof of (P.47).

Proof of (P.48): From (P.45c), the following holds:

$$I(X;Y|Z)=\mathbb{E}_{XYZ}\left[\log P_{X|YZ}(X|Y,Z)\right]-\mathbb{E}_{XZ}\left[\log P_{X|Z}(X|Z)\right] \tag{P.52a}$$
$$=H(X|Z)-H(X|Y,Z), \tag{P.52b}$$

and this completes the proof of (P.48).

Proof of (P.49): From (P.30), the following holds:

$$I(X,Y;Z)=H(X,Y)-H(X,Y|Z) \tag{P.53a}$$
$$=H(X)+H(Y|X)-H(X|Z)-H(Y|X,Z) \tag{P.53b}$$
$$=I(X;Z)+I(Y;Z|X), \tag{P.53c}$$

and this completes the proof of (P.49).

Proof of (P.50): From (P.30), the following holds:

$$I(X,Y;Z)=H(X,Y)-H(X,Y|Z) \tag{P.54a}$$
$$=H(Y)+H(X|Y)-H(Y|Z)-H(X|Y,Z) \tag{P.54b}$$
$$=I(Y;Z)+I(X;Z|Y), \tag{P.54c}$$

and this completes the proof of (P.50). This completes the proof of Lemma 46.

■

The mutual information between the random variables X and Y conditioning on the random variable Z is equal to zero if X and Y are independent conditioning on Z, *i.e.*, $X\rightarrow Z\rightarrow Y$, as stated by the following lemma.

Lemma 47. *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be three countable sets and let also X, Y and Z be three random variables with joint pmf $P_{XYZ} : \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \rightarrow [0, 1]$ such that $X \rightarrow Z \rightarrow Y$. Then,*

$$I(X; Y|Z) = 0. \tag{P.55}$$

Proof: From (P.45c), the following holds:

$$\begin{aligned} I(X; Y|Z) &= \mathbb{E}_{XYZ}\left[\log\left(\frac{P_{X|Z}(x|z)}{P_{X|Z}(x|z)}\right)\right] && \text{(P.56a)} \\ &= \mathbb{E}_{XYZ}\left[\log 1\right] && \text{(P.56b)} \\ &= 0. && \text{(P.56c)} \end{aligned}$$

where, (P.56a) follows from the fact that the random variables X and Y are mutually independent conditioning on the random variable Z, *i.e.*, $X \rightarrow Z \rightarrow Y$. This completes the proof of Lemma 47. ■

The following lemma presents some additional useful properties of the mutual information and conditional mutual information.

Lemma 48 (Chain rule for mutual information and chain rule for conditional mutual information)**.** *Let $\mathcal{X}_1$, $\mathcal{X}_2$, ..., $\mathcal{X}_N$, $\mathcal{Y}$ and $\mathcal{Z}$ be $N + 2$ countable sets. Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)^{\mathsf{T}}$ be a vector of N random variables and let also Y and Z be two random variables with joint pmfs $P_{\boldsymbol{X}Y} : \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_N \times \mathcal{Y} \rightarrow [0, 1]$ and $P_{\boldsymbol{X}YZ} : \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_N \times \mathcal{Y} \times \mathcal{Z} \rightarrow [0, 1]$. Then,*

$$\begin{aligned} I(X_1, X_2, \ldots, X_N; Y) &= I(X_1; Y) + I(X_2; Y|X_1) + \sum_{n=3}^{N} I(X_n; Y|X_1, X_2, \ldots, X_{n-1}), && \text{(P.57a)} \\ I(X_1, X_2, \ldots, X_N; Y) &\geqslant 0, \ \ \textit{and} && \text{(P.57b)} \end{aligned}$$

$$I(X_1, X_2, \ldots, X_N; Y|Z) = I(X_1; Y|Z) + I(X_2; Y|Z, X_1) + \sum_{n=3}^{N} I(X_n; Y|Z, X_1, X_2, \ldots, X_{n-1}). \tag{P.57c}$$

Proof:

Proof of (P.57a): From (P.28a), the following holds:

$$I(\boldsymbol{X};Y) = \mathbb{E}_{\boldsymbol{X}Y}\left[\log\left(\frac{P_{\boldsymbol{X}Y}(\boldsymbol{X},Y)}{P_{\boldsymbol{X}}(\boldsymbol{X})P_Y(Y)}\right)\right] \tag{P.58a}$$

$$= \mathbb{E}_{\boldsymbol{X}Y}\Bigg[\log\Bigg(\frac{P_{X_1Y}(X_1,Y)}{P_{X_1}(X_1)P_Y(Y)}\frac{P_{X_2|X_1Y}(X_2|X_1,Y)}{P_{X_2|X_1}(X_2|X_1)}\frac{P_{X_3|X_1X_2Y}(X_3|X_1,X_2,Y)}{P_{X_3|X_1X_2}(X_3|X_1,X_2)}\cdots$$
$$\frac{P_{X_N|X_1X_2\ldots X_{N-1}Y}(X_N|X_1,X_2,\ldots,X_{N-1},Y)}{P_{X_N|X_1X_2\ldots X_{N-1}}(X_3|X_1,X_2,\ldots,X_N-1)}\Bigg)\Bigg] \tag{P.58b}$$

$$= \mathbb{E}_{X_1Y}\left[\log\frac{P_{X_1Y}(X_1,Y)}{P_{X_1}(X_1)P_Y(Y)}\right] + \mathbb{E}_{X_1X_2Y}\left[\log\frac{P_{X_2|X_1Y}(X_2|X_1,Y)}{P_{X_2|X_1}(X_2|X_1)}\right]$$
$$+\mathbb{E}_{X_1X_2X_3Y}\left[\log\frac{P_{X_3|X_1X_2Y}(X_3|X_1,X_2,Y)}{P_{X_3|X_1X_2}(X_3|X_1,X_2)}\right] + \ldots$$
$$+\mathbb{E}_{\boldsymbol{X}Y}\left[\log\frac{P_{X_N|X_1X_2\ldots X_{N-1}Y}(X_N|X_1,X_2,\ldots,X_{N-1},Y)}{P_{X_N|X_1X_2\ldots X_{N-1}}(X_3|X_1,X_2,\ldots,X_N-1)}\right] \tag{P.58c}$$

$$= I(X_1;Y) + I(X_2;Y|X_1) + I(X_3;Y|X_1,X_2) + \ldots + I(X_N;Y|X_1,X_2,\ldots,X_{N-1}), \tag{P.58d}$$

where, (P.58d) follows from (P.28a) and (P.45b). This completes the proof of (P.57a).

Proof of (P.57b): From (P.57a), the following holds:

$$I(X_1,X_2,\ldots,X_N;Y) = I(X_1;Y) + I(X_2;Y|X_1) + \sum_{n=3}^{N} I(X_n;Y|X_1,X_2,\ldots,X_{n-1}) \tag{P.59a}$$

$$= H(Y) - H(Y|X_1) + H(Y|X_1) - H(Y|X_1,X_2)$$
$$+\sum_{n=3}^{N}\left(H(Y|X_1,X_2,\ldots,X_{n-1}) - H(Y|X_1,X_2,\ldots,X_{n-1},X_n)\right) \tag{P.59b}$$

$$\geqslant 0, \tag{P.59c}$$

where (P.59c) follows from Lemma 40 and the fact that $H(Y) \geqslant H(Y|X_1)$, $H(Y|X_1) \geqslant H(Y|X_1,X_2)$, …, $H(Y|X_1,X_2,\ldots,X_{N-1}) \geqslant H(Y|X_1,X_2,\ldots,X_{N-1},X_N)$. This completes the proof of of (P.57b).

Proof of (P.57c): From (P.45c), the following holds:

$$I(\boldsymbol{X};Y|Z) = \mathbb{E}_{\boldsymbol{X}YZ}\left[\log\left(\frac{P_{\boldsymbol{X}|YZ}(\boldsymbol{X}|Y,Z)}{P_{\boldsymbol{X}|Z}(\boldsymbol{X})}\right)\right] \tag{P.60a}$$

$$
\begin{aligned}
=&\mathbb{E}_{\boldsymbol{X}YZ}\left[\log\left(\frac{P_{X_1|YZ}(X_1|YZ)}{P_{X_1|Z}(X_1|Z)}\frac{P_{X_2|X_1YZ}(X_2|X_1,Y,Z)}{P_{X_2|X_1Z}(X_2|X_1,Z)}\right.\right.\\
&\frac{P_{X_3|X_1X_2YZ}(X_3|X_1,X_2,Y,Z)}{P_{X_3|X_1X_2Z}(X_3|X_1,X_2,Z)}\dots\\
&\left.\left.\frac{P_{X_N|X_1X_2\dots X_{N-1}YZ}(X_N|X_1,X_2,\dots,X_{N-1},Y,Z)}{P_{X_N|X_1X_2\dots X_{N-1},Z}(X_3|X_1,X_2,\dots,X_N-1,Z)}\right)\right] \quad \text{(P.60b)}\\
=&\mathbb{E}_{X_1YZ}\left[\log\frac{P_{X_1|YZ}(X_1|Y,Z)}{P_{X_1|Z}(X_1|Z)}\right]+\mathbb{E}_{X_1X_2YZ}\left[\log\frac{P_{X_2|X_1YZ}(X_2|X_1,Y,Z)}{P_{X_2|X_1Z}(X_2|X_1,Z)}\right]\\
&+\mathbb{E}_{X_1X_2X_3YZ}\left[\log\frac{P_{X_3|X_1X_2YZ}(X_3|X_1,X_2,Y,Z)}{P_{X_3|X_1X_2Z}(X_3|X_1,X_2,Z)}\right]+\dots\\
&+\mathbb{E}_{\boldsymbol{X}YZ}\left[\log\frac{P_{X_N|X_1X_2\dots X_{N-1}YZ}(X_N|X_1,X_2,\dots,X_{N-1},Y,Z)}{P_{X_N|X_1X_2\dots X_{N-1}Z}(X_3|X_1,X_2,\dots,X_N-1,Z)}\right] \quad \text{(P.60c)}\\
=&I(X_1;Y|Z)+I(X_2;Y|X_1,Z)+I(X_3;Y|X_1,X_2,Z)+\dots\\
&+I(X_N;Y|X_1,X_2,\dots,X_{N-1},Z), \quad \text{(P.60d)}
\end{aligned}
$$

where (P.60d) follows from (P.45c). This completes the proof of (P.57c). This completes the proof of Lemma 48.

■

The mutual information between the random variables X and Z is less than or equal to the mutual information between the random variables X and Y, or between the random variables Y and Z, if the random variables X and Z are independent conditioning on the random variable Y, *i.e.*, $X \to Y \to Z$. This is stated in the following lemma.

Lemma 49 (Data Processing Inequality [28, Theorem 2.8.1])**.** *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be three countable sets and let X, Y and Z be three random variables with joint pmf $P_{XYZ}:\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}\to[0,1]$ such that $X\to Y\to Z$. Then,*

$$I(X;Z)\leqslant I(X;Y) \text{ and} \quad \text{(P.61a)}$$
$$I(X;Z)\leqslant I(Y;Z). \quad \text{(P.61b)}$$

If $Z=g(Y)$, then

$$I(X;g(Y))\leqslant I(X;Y). \quad \text{(P.61c)}$$

Proof:
Proof of (P.61a): From (P.57a), the following holds:

$$
\begin{aligned}
I(X;Y,Z)=&I(X;Z)+I(X;Y|Z) && \text{(P.62a)}\\
\geqslant&I(X;Z) && \text{(P.62b)}
\end{aligned}
$$

and

$$
\begin{aligned}
I(X;Y,Z)=&I(X;Y)+I(X;Z|Y) && \text{(P.62c)}\\
=&I(X;Y), && \text{(P.62d)}
\end{aligned}
$$

where (P.62d) follows from the fact that the random variables X and Z are mutually independent conditioning on the random variable Y, *i.e.*, $X \to Y \to Z$. From (P.62b) and (P.62d), the following holds:

$$I(X;Z) \leqslant I(X;Y), \tag{P.62e}$$

and this completes the proof of (P.61a).

Proof of (P.61b): From (P.57a), the following holds:

$$\begin{aligned} I(X,Y;Z) &= I(Y;Z) + I(X;Z|Y) && \text{(P.63a)}\\ &= I(Y;Z) && \text{(P.63b)} \end{aligned}$$

and

$$\begin{aligned} I(X,Y;Z) &= I(X;Z) + I(Y;Z|X) && \text{(P.63c)}\\ &\geqslant I(X;Z), && \text{(P.63d)} \end{aligned}$$

where (P.63b) follows from the fact that the random variables X and Z are mutually independent conditioning on the random variable Y, *i.e.*, $X \to Y \to Z$. From (P.63b) and (P.63d), the following holds:

$$I(X;Z) \leqslant I(Y;Z), \tag{P.63e}$$

and this completes the proof of (P.61b).

Proof of (P.61c): Plugging $Z = g(Y)$ into (P.62e), yields:

$$I\left(X; g(Y)\right) \leqslant I\left(X;Y\right), \tag{P.64}$$

and this completes the proof of (P.61c). This completes the proof of Lemma 49. ■

The following lemma presents some useful properties of the conditional mutual information if the random variables X and Z are independent conditioning on the random variable Y, *i.e.*, $X \to Y \to Z$.

Lemma 50 (Corollary [28, Theorem 2.8.1])**.** *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be three countable sets and let X, Y and Z be three random variables with joint pmf $P_{XYZ} : \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \to [0,1]$ such that $X \to Y \to Z$. Then,*

$$\begin{aligned} &I(X;Y|Z) \leqslant I(X;Y) \text{ and} && \text{(P.65a)}\\ &I(Y;Z|X) \leqslant I(Y;Z). && \text{(P.65b)} \end{aligned}$$

Proof:

Proof of (P.65a): From (P.57a), the following holds:

$$\begin{aligned} I(X;Y,Z) &= I(X;Z) + I(X;Y|Z) && \text{(P.66a)}\\ &\geqslant I(X;Y|Z). && \text{(P.66b)} \end{aligned}$$

From (P.62d) and (P.66b), the following holds:

$$I(X;Y|Z) \leqslant I(X;Y), \tag{P.66c}$$

and this completes the proof of (P.65a).

Proof of (P.65b): From (P.57a), the following holds:

$$\begin{aligned} I(X,Y;Z) &= I(X;Z) + I(Y;Z|X) \\ &\geqslant I(Y;Z|X). \end{aligned} \tag{P.67a}$$

From (P.63b) and (P.67a), the following holds:

$$I(Y;Z|X) \leqslant I(Y;Z), \tag{P.67b}$$

and this completes the proof of (P.65b). This completes the proof of Lemma 50. ■

The following lemma presents a property of the conditional mutual information when the random variables X, Y, and Z do not form a Markov Chain, which is contrary in result to the stated in Lemma 50.

Lemma 51 (Corollary [28, Theorem 2.8.1])**.** *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be three countable sets and let X, Y and Z be three random variables with joint pmf $P_{XYZ} : \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \to [0,1]$ such that $P_{XYZ}(x,y,z) = P_X(x)P_Y(y)P_{Z|XY}(z|x,y)$. Then,*

$$I(X;Y|Z) \geqslant I(X;Y). \tag{P.68}$$

Proof: From the assumption of the lemma, X and Y are two independent random variables, then $I(X;Y) = 0$. Hence, the inequality can be obtained from the non-negativity of mutual information. ■

The following two lemmas present some useful properties of the mutual information between two N-dimensional vectors of random variables. These two lemmas are considering that the components of the N-dimensional vector of random variables $\boldsymbol{Y}$ correspond to the channel-outputs generated by the components of the N-dimensional vector of random variables $\boldsymbol{X}$ as channel-inputs in a given channel. In the first lemma, the components of the N-dimensional vector of random variables $\boldsymbol{X}$ are assumed be mutually independent. In the second lemma, the channel is assumed to be memoryless.

Lemma 52 ([55, Theorem 1.8])**.** *Let $\mathcal{X}_1$, $\mathcal{X}_2$, ..., $\mathcal{X}_N$, and $\mathcal{Y}_1$, $\mathcal{Y}_2$, ..., $\mathcal{Y}_N$ be $2N$ countable sets, let $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)^{\mathsf{T}}$ be an N- dimensional vector of independent random variables, and let also $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_N)^{\mathsf{T}}$ be an N-dimensional vector of random variables such that the joint pmf is $P_{\boldsymbol{XY}} : \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_N \times \mathcal{Y}_1 \times \mathcal{Y}_2 \ldots \times \mathcal{Y}_N \to [0,1]$. Then,*

$$I(\boldsymbol{X};\boldsymbol{Y}) \geqslant \sum_{n=1}^{N} I(X_n;Y_n). \tag{P.69}$$

Proof: From (P.28c), the following holds:

$$I(\boldsymbol{X};\boldsymbol{Y}) = \mathbb{E}_{\boldsymbol{XY}}\left[\log\left(\frac{P_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{X}|\boldsymbol{Y})}{P_{\boldsymbol{X}}(\boldsymbol{X})}\right)\right] \tag{P.70a}$$

$$= \mathbb{E}_{\boldsymbol{XY}}\left[\log\left(\frac{P_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{X}|\boldsymbol{Y})}{P_{X_1}(X_1)P_{X_2}(X_2)\ldots P_{X_N}(X_N)}\right)\right], \tag{P.70b}$$

where, (P.70b) follows from the fact that $X_1, X_2, \ldots, X_N$ are mutually independent. On the other hand,

$$\sum_{n=1}^{N} I\left(X_n;Y_n\right)=\sum_{n=1}^{N} \mathbb{E}_{X_nY_n}\left[\log\left(\frac{P_{X_n|Y_n}(X_n|Y_n)}{P_{X_n}(X_n)}\right)\right] \tag{P.70c}$$

$$=\mathbb{E}_{\boldsymbol{XY}}\left[\log\left(\frac{P_{X_1|Y_1}(X_1|Y_1)P_{X_2|Y_2}(X_2|Y_2)\ldots P_{X_N|Y_N}(X_N|Y_N)}{P_{X_1}(X_1)P_{X_2}(X_2)\ldots P_{X_N}(X_N)}\right)\right]. \tag{P.70d}$$

Hence,

$$\sum_{n=1}^{N} I\left(X_n;Y_n\right) - I\left(\boldsymbol{X};\boldsymbol{Y}\right) = \mathbb{E}_{\boldsymbol{XY}}\left[\log\left(\frac{P_{X_1|Y_1}(X_1|Y_1)P_{X_2|Y_2}(X_2|Y_2)\ldots P_{X_N|Y_N}(X_N|Y_N)}{P_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{X}|\boldsymbol{Y})}\right)\right] \tag{P.70e}$$

$$\leqslant\log\left(\mathbb{E}_{\boldsymbol{XY}}\left[\left(\frac{P_{X_1|Y_1}(X_1|Y_1)P_{X_2|Y_2}(X_2|Y_2)\ldots P_{X_N|Y_N}(X_N|Y_N)}{P_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{X}|\boldsymbol{Y})}\right)\right]\right) \tag{P.70f}$$

$$=\log\left(\sum_{\boldsymbol{x}\in\mathcal{X}^N \boldsymbol{y}\in\mathcal{Y}^N}\left(P_{\boldsymbol{Y}}(\boldsymbol{y})P_{X_1|Y_1}(x_1|y_1)P_{X_2|Y_2}(x_2|y_2)\ldots P_{X_N|Y_N}(x_N|y_N)\right)\right) \tag{P.70g}$$

$$=\log\left(\sum_{\boldsymbol{y}\in\mathcal{Y}^N} P_{\boldsymbol{Y}}(\boldsymbol{y})\sum_{\boldsymbol{x}\in\mathcal{X}^N}\left(P_{X_1|Y_1}(x_1|y_1)P_{X_2|Y_2}(x_2|y_2)\ldots P_{X_N|Y_N}(x_N|y_N)\right)\right) \tag{P.70h}$$

$$=\log\left(\sum_{\boldsymbol{y}\in\mathcal{Y}^N} P_{\boldsymbol{Y}}(\boldsymbol{y})\right) \tag{P.70i}$$

$$=\log 1 \tag{P.70j}$$

$$=0, \tag{P.70k}$$

where (P.70f) follows from Jensen's inequality. Then,

$$I\left(\boldsymbol{X};\boldsymbol{Y}\right) \geqslant \sum_{n=1}^{N} I\left(X_n;Y_n\right), \tag{P.70l}$$

and this completes the proof of 52. ∎

Lemma 53 ([55, Theorem 1.9])**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be two countable sets. Let also $X_1, X_2, \ldots, X_N$, $Y_1, Y_2, \ldots, Y_N$ be $2N$ random variables with joint pmf $P_{\boldsymbol{XY}} : \mathcal{X}^N \times \mathcal{Y}^N \to [0,1]$ such that for all $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^N \times \mathcal{Y}^N$ it holds that $P_{\boldsymbol{XY}}(\boldsymbol{x},\boldsymbol{y}) = P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})P_{\boldsymbol{X}}(\boldsymbol{x})$, with $P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}) = \prod_{n=1}^{N} P_{Y_n|X_n}(y_n|x_n)$. Then,*

$$I\left(\boldsymbol{X};\boldsymbol{Y}\right) \leqslant \sum_{n=1}^{N} I\left(X_n;Y_n\right). \tag{P.71}$$

Proof: From (P.28b), the following holds:

$$I\left(\boldsymbol{X};\boldsymbol{Y}\right)=\mathbb{E}_{\boldsymbol{XY}}\left[\log\left(\frac{P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{Y}|\boldsymbol{X})}{P_{\boldsymbol{Y}}(\boldsymbol{Y})}\right)\right] \tag{P.72a}$$

$$=\mathbb{E}_{\boldsymbol{XY}}\left[\log\left(\frac{P_{Y_1|X_1}(Y_1|X_1)P_{Y_2|X_2}(Y_2|X_2)\dots P_{Y_N|X_N}(Y_N|X_N)}{P_{\boldsymbol{Y}}(\boldsymbol{Y})}\right)\right], \tag{P.72b}$$

where (P.72b) follows from the fact that $P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}) = \prod_{n=1}^{N} P_{Y_n|X_n}(y_n|x_n)$. On the other hand,

$$\sum_{n=1}^{N} I(X_n;Y_n)=\sum_{n=1}^{N}\mathbb{E}_{X_nY_n}\left[\log\left(\frac{P_{Y_n|X_n}(Y_n|X_n)}{P_{Y_n}(Y_n)}\right)\right] \tag{P.72c}$$

$$=\mathbb{E}_{\boldsymbol{XY}}\left[\log\left(\frac{P_{Y_1|X_1}(Y_1|X_1)P_{Y_2|X_2}(Y_2|X_2)\dots P_{Y_N|X_N}(Y_N|X_N)}{P_{Y_1}(Y_1)P_{Y_2}(Y_2)\dots P_{Y_N}(Y_N)}\right)\right]. \tag{P.72d}$$

Hence,

$$\sum_{n=1}^{N} I(\boldsymbol{X};\boldsymbol{Y}) - I(X_n;Y_n)=\mathbb{E}_{\boldsymbol{Y}}\left[\log\left(\frac{P_{Y_1}(Y_1)P_{Y_2}(Y_2)\dots P_{Y_N}(Y_N)}{P_{\boldsymbol{Y}}(\boldsymbol{Y})}\right)\right] \tag{P.72e}$$

$$\leqslant\log\left(\mathbb{E}_{\boldsymbol{Y}}\left[\left(\frac{P_{Y_1}(Y_1)P_{Y_2}(Y_2)\dots P_{Y_N}(Y_N)}{P_{\boldsymbol{Y}}(\boldsymbol{Y})}\right)\right]\right) \tag{P.72f}$$

$$=\log\left(\sum_{\boldsymbol{y}\in\mathcal{Y}^N}\left(P_{Y_1}(y_1)P_{Y_2}(y_2)\dots P_{Y_N}(y_N)\right)\right) \tag{P.72g}$$

$$=\log 1 \tag{P.72h}$$

$$=0, \tag{P.72i}$$

where (P.72f) follows from Jensen's inequality. Then,

$$I(\boldsymbol{X};\boldsymbol{Y}) \leqslant \sum_{n=1}^{N} I(X_n;Y_n), \tag{P.72j}$$

and this completes the proof of Lemma 53.

■

P.2. Real-Valued Random Variables

Shannon formalized the information measures on discrete random variables and these notions were extended to real-valued random variables. The differential entropy (the entropy of a real-valued random variable) does not have the same meaning as the entropy for the discrete case. Nonetheless, the real importance of the differential entropy is in the calculation of the mutual information between two real-valued random variables, which allows to compare two probability distributions and to keep the same meaning as in the discrete case.

P.2.1. Differential Entropy

The differential entropy $h(X)$ of a real-valued random variable X is a functional of the pdf f_X. Although entropy and differential entropy have similar mathematical forms, the differential entropy does not serve as a measure of the average amount of information contained in a real-valued random variable. In fact, a real-valued random variable generally contains an infinite amount of information [104].

Definition 15 (Differential Entropy)**.** *Let X be a random variable with pdf $f_X : \mathbb{R} \rightarrow [0, \infty)$. Then, the differential entropy of X, denoted by $h(X)$, is:*

$$h(X) = -\int_{-\infty}^{\infty} f_X(x) \log f_X(x)\, \mathrm{d}x. \tag{P.73}$$

Note that $h(X)$ depends only on f_X and not in the values in $\mathbb{R}$. The differential entropy of a random variable X can also be written as follows:

$$h(X) = -\mathbb{E}_X \left[\log f_X(X)\right]. \tag{P.74}$$

Corollary 12. *Let X be a random variable uniformly distributed on $[0, a]$. Then,*

$$h(X) = -\int_0^a \frac{1}{a} \log \frac{1}{a}\, \mathrm{d}x = \log a. \tag{P.75}$$

Proof: The proof of Corollary 12 follows directly from Definition 15. ■

Note that in this case $h(X) < 0$ if $a < 1$.

Corollary 13. *Let X be a Gaussian random variable with zero mean and variance σ^2,* i.e.*, $X \sim \mathcal{N}(0, \sigma^2)$. Then,*

$$h(X) = \frac{1}{2} \log \left(2\pi e \sigma^2\right). \tag{P.76}$$

Proof:

From Definition 15, with $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$ and e the base of the logarithm, the following holds:

$$\begin{aligned}
h(X) &= -\int_{-\infty}^{\infty} f_X(x) \ln f_X(x)\, \mathrm{d}x && \text{(P.77a)}\\
&= -\int_{-\infty}^{\infty} f_X(x) \left(-\frac{x^2}{2\sigma^2} - \ln\sqrt{2\pi\sigma^2}\right) \mathrm{d}x && \text{(P.77b)}\\
&= \frac{1}{2\sigma^2} \int_{-\infty}^{\infty} x^2 f_X(x)\, \mathrm{d}x + \ln\sqrt{2\pi\sigma^2} \int_{-\infty}^{\infty} f_X(x)\, \mathrm{d}x && \text{(P.77c)}\\
&= \frac{\mathbb{E}_X\left[X^2\right]}{2\sigma^2} + \ln\sqrt{2\pi\sigma^2} && \text{(P.77d)}\\
&= \frac{1}{2} + \frac{1}{2}\ln\left(2\pi\sigma^2\right) && \text{(P.77e)}\\
&= \frac{1}{2}\ln e + \frac{1}{2}\ln\left(2\pi\sigma^2\right) && \text{(P.77f)}\\
&= \frac{1}{2}\ln\left(2\pi e\sigma^2\right), && \text{(P.77g)}
\end{aligned}$$

in nats, where (P.77e) follows from the fact that $\mathbb{E}_X\left[X^2\right] = \mathrm{Var}_X\left[X\right] + \left(\mathbb{E}_X\left[X\right]\right)^2 = \sigma^2$. Changing the base of the logarithm to two completes the proof of Corollary 13. ■

Lemma 54. *Let X be a random variable with pdf $f_X : \mathbb{R} \to [0, \infty)$, zero mean, and variance σ^2. The maximum value of the differential entropy of the random variable X is obtained when the random variable X has a Gaussian distribution with zero mean and variance σ^2. Then,*

$$h(X) \leqslant \frac{1}{2} \log \left(2\pi e \sigma^2\right). \tag{P.78}$$

Proof: Let $\phi_X : \mathbb{R} \to [0, \infty)$ be a Gaussian pdf on the random variable X with zero mean and variance σ^2, *i.e.*, $\phi_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$. Then, the following holds:

$$\begin{aligned}
h(X) + \int_{-\infty}^{\infty} f_X(x) \log \phi_X(x)\, \mathrm{d}x &= - \int_{-\infty}^{\infty} f_X(x) \log f_X(x)\, \mathrm{d}x + \int_{-\infty}^{\infty} f_X(x) \log \phi_X(x)\, \mathrm{d}x && \text{(P.79a)} \\
&= \int_{-\infty}^{\infty} f_X(x) \log \frac{\phi_X(x)}{f_X(x)}\, \mathrm{d}x && \text{(P.79b)} \\
&= \mathbb{E}_X \left[\log \frac{\phi_X(X)}{f_X(X)} \right] && \text{(P.79c)} \\
&\leqslant \log \left(\mathbb{E}_X \left[\frac{\phi_X(X)}{f_X(X)} \right] \right) && \text{(P.79d)} \\
&= \log \left(\int_{-\infty}^{\infty} f_X(x) \frac{\phi_X(x)}{f_X(x)}\, \mathrm{d}x \right) && \text{(P.79e)} \\
&= \log \left(\int_{-\infty}^{\infty} \phi_X(x)\, \mathrm{d}x \right) && \text{(P.79f)} \\
&= \log 1 && \text{(P.79g)} \\
&= 0, && \text{(P.79h)}
\end{aligned}$$

where (P.79d) follows from Jensen's inequality.

Then,

$$\begin{aligned}
h(X) &\leqslant - \int_{-\infty}^{\infty} f_X(x) \log \phi_X(x)\, \mathrm{d}x && \text{(P.79i)} \\
&= \frac{1}{2} \log \left(2\pi e \sigma^2\right), && \text{(P.79j)}
\end{aligned}$$

and equality holds if $f_X(x) = \phi_X(x)$. This completes the proof of Lemma 54. ■

P.2.2. Joint Differential Entropy

The joint differential entropy can be understood as the extension of the joint entropy for discrete random variables to real-valued random variables. Although joint entropy and joint differential entropy have similar mathematical forms, the joint differential entropy does not serve as a measure of the average amount of information simultaneously contained into the considered real-valued random variables.

Definition 16 (Joint Differential Entropy)**.** *Let X and Y be two random variables with joint pdf $f_{XY} : \mathbb{R}^2 \to [0, \infty)$. Then, the joint differential entropy of the random variables X and Y,* *denoted by $h(X,Y)$, is:*

$$h(X,Y) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\log f_{XY}(x,y)\,\mathrm{d}x\,\mathrm{d}y. \tag{P.80}$$

The joint differential entropy of the random variables X and Y can also be written as follows:

$$h(X,Y) = -\mathbb{E}_{XY}\left[\log f_{XY}(X,Y)\right]. \tag{P.81}$$

Lemma 38 and Definition 11 can be extended to real-valued random variables.

Lemma 55 (Differential Entropy of a Bivariate Gaussian Distribution)**.** *Let X and Y be two Gaussian random variables with covariance matrix $\mathbf{K} = \mathbb{E}_{XY}\left[\begin{bmatrix}X\\Y\end{bmatrix}\begin{bmatrix}X & Y\end{bmatrix}\right] = \begin{bmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y\\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{bmatrix}$, where $\rho = \frac{\mathbb{E}_{XY}[XY]}{\sigma_X\sigma_Y}$ is the Pearson correlation coefficient. The joint differential entropy of the random variables X and Y is:*

$$h(X,Y) = \frac{1}{2}\log\left((2\pi e)^2\left|\mathbf{K}\right|\right), \tag{P.82}$$

where $|\mathbf{K}|$ is the determinant of the covariance matrix $\mathbf{K}$, i.e., $|\mathbf{K}| = \det(\mathbf{K})$.

Proof: From (P.81), the following holds:

$$h(X,Y) = -\mathbb{E}_{XY}\left[\log f_{XY}(X,Y)\right]. \tag{P.83a}$$

For all $(x,y) \in \mathbb{R}^2$, the following holds:

$$f_{XY}(x,y) = \frac{1}{\left(\sqrt{2\pi}\right)^2|\mathbf{K}|^{\frac{1}{2}}} e^{-\frac{1}{2}\left([x\,y]\,\mathbf{K}^{-1}\begin{bmatrix}x\\y\end{bmatrix}\right)}, \tag{P.83b}$$

where $\mathbf{K}^{-1}$ is the inverse of the covariance matrix. The determinant of the covariance matrix $\mathbf{K}$ is:

$$|\mathbf{K}| = \sigma_X^2\sigma_Y^2\left(1-\rho^2\right), \tag{P.83c}$$

and the inverse of the covariance matrix $\mathbf{K}$ is:

$$\mathbf{K}^{-1} = \frac{1}{|\mathbf{K}|}\begin{bmatrix}\sigma_Y^2 & -\rho\sigma_X\sigma_Y\\ -\rho\sigma_X\sigma_Y & \sigma_X^2\end{bmatrix}. \tag{P.83d}$$

Plugging (P.83d) into (P.83b) the following holds:

$$f_{XY}(x,y) = \frac{1}{\left(\sqrt{2\pi}\right)^2|\mathbf{K}|^{\frac{1}{2}}} e^{-\frac{1}{2}\left(\frac{1}{|\mathbf{K}|}[x\,y]\begin{bmatrix}\sigma_Y^2 & -\rho\sigma_X\sigma_Y\\ -\rho\sigma_X\sigma_Y & \sigma_X^2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}\right)} \tag{P.83e}$$

$$=\frac{1}{\left(\sqrt{2\pi}\right)^2|\mathbf{K}|^{\frac{1}{2}}}e^{-\frac{1}{2}\left(\frac{1}{|\mathbf{K}|}[x\,y]\begin{bmatrix}x\sigma_Y^2-y\rho\sigma_X\sigma_Y\\-x\rho\sigma_X\sigma_Y+y\sigma_X^2\end{bmatrix}\right)} \tag{P.83f}$$

$$=\frac{1}{\left(\sqrt{2\pi}\right)^2|\mathbf{K}|^{\frac{1}{2}}}e^{-\frac{1}{2}\left(\frac{1}{|\mathbf{K}|}\left(x^2\sigma_Y^2-2xy\rho\sigma_X\sigma_Y+y^2\sigma_X^2\right)\right)}. \tag{P.83g}$$

Plugging (P.83g) into (P.83a) and considering the logarithm of base e, the following holds:

$$h(X,Y)=-\mathbb{E}_{XY}\left[\ln\left(\frac{1}{\left(\sqrt{2\pi}\right)^2|\mathbf{K}|^{\frac{1}{2}}}e^{-\frac{1}{2}\left(\frac{1}{|\mathbf{K}|}\left(X^2\sigma_Y^2-2XY\rho\sigma_X\sigma_Y+Y^2\sigma_X^2\right)\right)}\right)\right] \tag{P.83h}$$

$$=-\mathbb{E}_{XY}\left[\ln\left(\frac{1}{\left(\sqrt{2\pi}\right)^2|\mathbf{K}|^{\frac{1}{2}}}\right)-\frac{1}{2}\left(\frac{1}{|\mathbf{K}|}\left(X^2\sigma_Y^2-2XY\rho\sigma_X\sigma_Y+Y^2\sigma_X^2\right)\right)\right] \tag{P.83i}$$

$$=\ln\left(\left(\sqrt{2\pi}\right)^2|\mathbf{K}|^{\frac{1}{2}}\right)+\frac{1}{2}\left(\frac{1}{|\mathbf{K}|}\left(\sigma_Y^2\mathbb{E}_X\left[X^2\right]-2\rho\sigma_X\sigma_Y\mathbb{E}_{XY}\left[XY\right]+\sigma_X^2\mathbb{E}_Y\left[Y^2\right]\right)\right) \tag{P.83j}$$

$$=\ln\left(\left(\sqrt{2\pi}\right)^2|\mathbf{K}|^{\frac{1}{2}}\right)+\frac{1}{2}\left(\frac{1}{|\mathbf{K}|}\left(\sigma_Y^2\sigma_X^2-2\rho^2\sigma_X^2\sigma_Y^2+\sigma_X^2\sigma_Y^2\right)\right) \tag{P.83k}$$

$$=\ln\left(\left(\sqrt{2\pi}\right)^2|\mathbf{K}|^{\frac{1}{2}}\right)+\left(\frac{1}{|\mathbf{K}|}\left(\sigma_X^2\sigma_Y^2\left(1-\rho^2\right)\right)\right) \tag{P.83l}$$

$$=\ln\left(\left(\sqrt{2\pi}\right)^2|\mathbf{K}|^{\frac{1}{2}}\right)+1 \tag{P.83m}$$

$$=\ln\left(\left(\sqrt{2\pi e}\right)^2|\mathbf{K}|^{\frac{1}{2}}\right) \tag{P.83n}$$

$$=\frac{1}{2}\ln\left((2\pi e)^2|\mathbf{K}|\right), \tag{P.83o}$$

in nats, where (P.83m) follows from (P.83c). Changing the base of the logarithm to two completes the proof of Lemma 55. ■

Lemma 56 generalizes Lemma 55.

Lemma 56. *Let* $\boldsymbol{X}=(X_1,X_2,\ldots,X_N)^{\mathsf{T}}\in\mathbb{R}^N$ *be a vector of* N *random variables with joint Gaussian pdf* $f_{\boldsymbol{X}}:\mathcal{X}_1\times\mathcal{X}_2\times\ldots\times\mathcal{X}_N\to[0,\infty)$*, i.e.,* $\boldsymbol{X}\sim\mathcal{N}(0,\mathbf{K})$*. Then, the joint entropy of* $\boldsymbol{X}$*, denoted by* $h(\boldsymbol{X})$*, is:*

$$h(\boldsymbol{X})=\frac{1}{2}\log\left((2\pi e)^N|\mathbf{K}|\right). \tag{P.84}$$

P.2.3. Conditional Differential Entropy

Definition 17 (Conditional Differential Entropy)**.** *Let* X *and* Y *be two random variables with joint pdf* $f_{XY}:\mathbb{R}^2\to[0,\infty)$*. Then, the differential entropy of* Y *conditioning on* X*, denoted* *by* $h(Y|X)$, *is:*

$$h(Y|X) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\log f_{Y|X}(y|x)\,\mathrm{d}x\,\mathrm{d}y. \tag{P.85}$$

The differential entropy of the random variable Y conditioning on the random variable X can be written as follows:

$$h(Y|X) = -\mathbb{E}_{XY}\left[\log f_{Y|X}(Y|X)\right]. \tag{P.86}$$

Note also that the conditional differential entropy in (P.85) can be written as follows:

$$\begin{aligned} h(Y|X) =& \int_{-\infty}^{\infty} f_X(x)\left[-\int_{-\infty}^{\infty} f_{Y|X}(y|x)\log f_{Y|X}(y|x)\right]\mathrm{d}y\,\mathrm{d}x \\ =& \int_{-\infty}^{\infty} f_X(x) h(Y|X=x)\,\mathrm{d}x, \end{aligned} \tag{P.87}$$

where $h(Y|X=x) = -\int_{-\infty}^{\infty} f_{Y|X}(y|x)\log f_{Y|X}(y|x)\,\mathrm{d}y$, the differental entropy of Y conditioning on a fixed $X = x$.

Lemmas 39-42 can be extended to real-valued random variables.

P.2.4. Mutual Information

Definition 18 (Mutual Information)**.** *Let X and Y be two random variables with joint pdf $f_{XY}: \mathcal{X}\times\mathcal{Y}\to[0,\infty)$. Then, the mutual information between X and Y, denoted by $I(X;Y)$, is:*

$$I(X;Y) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\log\left(\frac{f_{XY}(x,y)}{f_X(x)f_Y(y)}\right)\mathrm{d}x\,\mathrm{d}y. \tag{P.88}$$

The mutual information between the real-valued random variables X and Y can also be written as follows:

$$I(X;Y) = \mathbb{E}_{XY}\left[\log\left(\frac{f_{XY}(X,Y)}{f_X(X)f_Y(Y)}\right)\right] \tag{P.89a}$$

$$= \mathbb{E}_{XY}\left[\log\left(\frac{f_{Y|X}(Y|X)}{f_Y(Y)}\right)\right] \tag{P.89b}$$

$$= \mathbb{E}_{XY}\left[\log\left(\frac{f_{X|Y}(X|Y)}{f_X(X)}\right)\right]. \tag{P.89c}$$

Lemmas 43-45 can be extended to real-valued random variables.

Lemma 57 (Mutual information between two Gaussian random variables)**.** *Let X and Y be two Gaussian random variables with zero means, correlation ρ, and variances σ_X^2 and σ_Y^2, respectively,* i.e., $(X,Y)^{\mathsf{T}} \sim \mathcal{N}\left(\begin{bmatrix} X \\ Y\end{bmatrix}, \begin{bmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{bmatrix}\right)$. *The mutual information between the random variables X and Y is:*

$$I(X;Y) = -\frac{1}{2}\log\left(1-\rho^2\right). \tag{P.90}$$

Proof: From Lemma P.33, the following holds:

$$I(X;Y) = h(X) + h(Y) - h(X,Y). \tag{P.91a}$$

Plugging (P.76) and (P.82) into (P.91a), the following holds:

$$I(X;Y)=\frac{1}{2}\log\left(2\pi e\sigma_X^2\right)+\frac{1}{2}\log\left(2\pi e\sigma_Y^2\right)-\frac{1}{2}\log\left((2\pi e)^2\left|\mathbf{K}\right|\right) \tag{P.91b}$$
$$=\frac{1}{2}\log\left(\frac{\sigma_X^2\sigma_Y^2}{|\mathbf{K}|}\right) \tag{P.91c}$$
$$=-\frac{1}{2}\log\left(1-\rho^2\right), \tag{P.91d}$$

where (P.91d) follows from the fact that $|\mathbf{K}| = \det(\mathbf{K}) = \sigma_X^2\sigma_Y^2\left(1-\rho^2\right)$. This completes the proof. ■

Note that if $\rho = \pm 1$ (perfect correlation) then $I(X;Y)$ is infinite.

P.2.5. Conditional Mutual Information

Definition 19 (Conditional Mutual Information)**.** *Let X, Y, and Z be three random variables with joint pdf $f_{XYZ} : \mathbb{R}^3 \to [0,\infty)$. Then, the mutual information between X and Y conditioning on Z, denoted by $I(X;Y|Z)$, is:*

$$I(X;Y|Z) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XYZ}(x,y,z)\log\left(\frac{f_{XY|Z}(x,y|z)}{f_{X|Z}(x|z)f_{Y|Z}(y|z)}\right)\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{P.92}$$

The mutual information between the real-valued random variables X and Y conditioning on the real-valued random variable Z can also be written as follows:

$$I(X;Y|Z)=\mathbb{E}_{XYZ}\left[\log\left(\frac{f_{XY|Z}(X,Y|Z)}{f_{X|Z}(X|Z)f_{Y|Z}(Y|Z)}\right)\right] \tag{P.93}$$
$$=\mathbb{E}_{XYZ}\left[\log\left(\frac{f_{Y|XZ}(Y|X,Z)}{f_{Y|Z}(Y|Z)}\right)\right] \tag{P.94}$$
$$=\mathbb{E}_{XYZ}\left[\log\left(\frac{f_{X|YZ}(X|YZ)}{f_{X|Z}(X|)}\right)\right]. \tag{P.95}$$

Lemmas 46-53 can be extended to real-valued random variables.

— Q —

Fano's Inequality

In data transmission, Fano's inequality establishes a connection between a traditional practical measure, the probability of error, and an information measure of the effect of the channel noise, the equivocation or conditional entropy. This inequality gives a lower-bound on the probability of error or an upper-bound on the equivocation. Fano's inequality is of paramount importance due to its versatile use in establishing fundamental limits. This result is used in all converse proofs in this thesis.

Lemma 58 (Fano's Inequality). *Let $\mathcal{X}$ be a countable set and let X and $\widehat{X}$ be two random variables with joint pmf $P_{X\widehat{X}} : \mathcal{X}^2 \to (0,1]$ such that for all $(x,\widehat{x}) \in \mathcal{X}^2$, $P_{X\widehat{X}}(x,\widehat{x}) = P_{X|\widehat{X}}(x|\widehat{x})P_{\widehat{X}}(\widehat{x})$. Let also $E = \mathbb{1}_{\{X \neq \hat{X}\}}$ be a binary random variable with pmf $P_E : \{0,1\} \to [0,1]$ such that $p = P_E(1) = 1 - P_E(0)$. Then,*

$$H\left(X|\widehat{X}\right) \leqslant H\left(E\right) + p\log\left(|\mathcal{X}| - 1\right). \tag{Q.1}$$

Proof:

$$\begin{aligned}
H(X|\widehat{X}) &= H(X|\widehat{X}) + H(E|X,\widehat{X}) && \text{(Q.2a)}\\
&= H(E,X|\widehat{X}) && \text{(Q.2b)}\\
&= H(E|\widehat{X}) + H(X|E,\widehat{X}) && \text{(Q.2c)}\\
&\leqslant H(E) + H(X|E,\widehat{X}) && \text{(Q.2d)}\\
&= H(E) + \sum_{\widehat{x}\in \mathrm{supp}\left(P_{\widehat{X}}\right)} \Big(P_{E,\widehat{X}}(0,\widehat{x})H(X|E=0,\widehat{X}=\widehat{x}) && \text{(Q.2e)}\\
&\quad + P_{E,\widehat{X}}(1,\widehat{x})H(X|E=1,\widehat{X}=\widehat{x})\Big) && \text{(Q.2f)}\\
&= H(E) + \sum_{\widehat{x}\in \mathrm{supp}\left(P_{\widehat{X}}\right)} P_{E,\widehat{X}}(1,\widehat{x})H(X|E=1,\widehat{X}=\widehat{x}) && \text{(Q.2g)}
\end{aligned}$$

$$\leqslant H(E) + \sum_{\widehat{x} \in \text{supp}\left(P_{\widehat{X}}\right)} P_{E,\widehat{X}}(1,\widehat{x}) \log\left(|\mathcal{X}|-1\right) \tag{Q.2h}$$

$$= H(E) + \log\left(|\mathcal{X}|-1\right) \sum_{\widehat{x} \in \text{supp}\left(P_{\widehat{X}}\right)} P_{E,\widehat{X}}(1,\widehat{x}) \tag{Q.2i}$$

$$= H(E) + P_E(1) \log\left(|\mathcal{X}|-1\right) \tag{Q.2j}$$

$$= H(E) + p \log\left(|\mathcal{X}|-1\right), \tag{Q.2k}$$

where, (Q.2a) follows from the fact that the value of the random variable E is known given the knowledge of the random variables X and $\widehat{X}$, *i.e.*, $H(E|X,\widehat{X}) = 0$; (Q.2d) follows from the fact that conditioning does not increase entropy (Lemma 40); (Q.2e) follows from Definition 12 (equation P.16) and $P_{E|\widehat{X}} : \{0,1\} \times \mathcal{X} \to (0,1]$; (Q.2g) follows from the fact that if $E = 0$ the value of the random variable X is known given the knowledge of the random variable $\widehat{X}$, i.e, $H(X|E=0,\widehat{X}=\widehat{x}) = 0$; (Q.2h) follows from the fact that given $E = 1$ and $\widehat{X} = \widehat{x}$, X can take any of the $\mathcal{X} - 1$ values and an upper-bound can be obtained on the entropy assuming that X is uniformly distributed, *i.e.*, $H(X|E=1,\widehat{X}=\widehat{x}) \leqslant \log\left(|\mathcal{X}|-1\right)$ (Lemma 37); and (Q.2k) follows from the fact that $p = \Pr\left[X \neq \widehat{X}\right] = P_E(1)$. This completes the proof of Lemma 58. ■

Fano's inequality corresponds to a model of communication in which a message selected from a set $\mathcal{X}$ is encoded into an input signal for transmission through a noisy channel, and the resulting output signal is decoded into a message of the same set. The conditional entropy $H\left(X|\widehat{X}\right)$ or equivocation represents the remaining uncertainty on the random variable X. It can also be seen as the average number of bits needed to transmit such that the receiver can identify X with the knowledge of $\widehat{X}$. In other words, it is the average information loss in a noisy channel. If $H\left(X|\widehat{X}\right) = 0$, then, the probability of error p is equal to zero.

Consider the following Markov chain: $X \to Y \to \widehat{X}$, with $\widehat{X} = g(Y)$, where g is a deterministic function. Then, from Lemma 49, the following holds:

$$I(X;Y) \geqslant I(X;g(Y)) \tag{Q.3}$$

$$= I(X;\widehat{X}). \tag{Q.4}$$

From (Q.4), the following holds:

$$H(X|\widehat{X}) \geqslant H(X|Y) \tag{Q.5}$$

and

$$H(X|Y) \leqslant H\left(X|\widehat{X}\right) \leqslant H\left(E\right) + P_E \log\left(|\mathcal{X}|-1\right). \tag{Q.6}$$

A loose bound on the equivocation can be obtained as follows:

$$H\left(X|\widehat{X}\right) \leqslant 1 + P_E \log|\mathcal{X}|. \tag{Q.7}$$

A lower-bound on the probability of error can be obtained from (Q.7), as follows:

$$P_E \geqslant \frac{H\left(X|\widehat{X}\right) - 1}{\log|\mathcal{X}|}. \tag{Q.8}$$

If the probability of error P_E is small, then $H\left(X|\widehat{X}\right)$ should be also small. Note also that $H\left(X|\widehat{X}\right) = H(X) - I\left(X;\widehat{X}\right)$ in which a high equivocation implies a low mutual information, and this also implies a high probability of error. A low probability of error implies a high mutual information, and this implies a low equivocation.

— R —

Weak Typicality

The AEP is a direct consequence of the weak law of large numbers (WLLN). It states that a sequence of N independent and identically distributed (i.i.d.) random variables $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)$ with N sufficiently large, almost surely belongs to the subset of all possible sequences $\mathcal{X}^N$ having only $2^{NH(X)}$ elements, each having a probability close to $2^{-NH(X)}$ [54], where X is a random variable representing any of the random variables in the long sequence. This divides the set of all sequences into two sets: the typical set and the nontypical set. All of the sequences in the typical set, the set with a probability measure close to one, have roughly the same probability of occurrence. Thus, the sequences in the typical set are almost uniformly distributed.

R.1. Discrete Random Variables

Let $\mathcal{X} = \{0, 1\}$ and let also $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)^{\mathsf{T}}$ be an N-dimensional vector of i.i.d binary random variables with joint pmf $P_{\boldsymbol{X}} : \{0, 1\}^N \to [0, 1]$. The probability of a binary sequence that contains r ones and $N - r$ zeros is:

$$P_{\boldsymbol{X}}(\boldsymbol{x}) = p_1^r (1 - p_1)^{N-r}, \tag{R.1}$$

where, $p_1 = P_X(1)$. The total number of binary sequences that contain r ones in a binary sequence of N symbols is:

$$n = \binom{N}{r}. \tag{R.2}$$

Let $R \in \mathbb{N}$ be a random variable that represents the number of ones, r, in a binary sequence of N symbols. Then, the probability of all binary sequences of N symbols that contain r ones is:

$$P_R(r) = \binom{N}{r} p_1^r (1 - p_1)^{N-r}, \tag{R.3}$$

where $\mathbb{E}_R[R] = Np_1$ and $\mathrm{Var}_R[R] = Np_1(1-p_1)$ (standard deviation equal to $\sqrt{Np_1(1-p_1)}$). The number of binary sequences with r ones will be approximately equal to $Np_1 \pm \sqrt{Np_1(1-p_1)}$. As N increases, the probability distribution of the random variable R becomes more concentrated, in the sense that its expected value increases as N and the standard deviation increases only as $\sqrt{N}$. It implies that the binary sequence $\boldsymbol{x}$ is most likely to fall in a small subset of sequences that is called the typical set [54].

Now, let $\mathcal{X}$ be a countable set and let also $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)^\mathsf{T}$ be an N-dimensional vector of i.i.d random variables with joint pmf $P_{\boldsymbol{X}} : \mathcal{X}^N \to [0,1]$. A long typical sequence of N symbols contains approximately p_1N occurrences of the first symbol, p_2N occurrences of the second symbol, ..., $p_\ell N$ occurrences of the ℓ-th symbol, where $p_1 = P_X(1)$, $p_2 = P_X(2)$, ..., $p_\ell = P_X(\ell)$. When the probability distribution $(p_1, p_2, \ldots, p_\ell)$ is conmesurable with denominator $n \in \mathbb{N}$, the probability of that typical sequence is:

$$P_{\boldsymbol{X}}(\boldsymbol{x}) = P_X(x_1)P_X(x_2)\ldots P_X(x_N) \tag{R.4a}$$
$$= p_1^{(Np_1)} p_2^{(Np_2)} \ldots p_\ell^{(Np_\ell)}, \tag{R.4b}$$

where, ℓ is the cardinality of the set $\mathcal{X}$, *i.e.*, $\ell = |\mathcal{X}|$. The amount of information provided by the typical sequence $\boldsymbol{x}$ is:

$$\imath(\boldsymbol{x}) = -\log P_{\boldsymbol{X}}(\boldsymbol{x}) \tag{R.5a}$$
$$= -N \sum_{x \in \mathrm{supp}(P_X)} P_X(x) \log P_X(x) \tag{R.5b}$$
$$= NH(X). \tag{R.5c}$$

Then, the amount of information provided by the typical sequence $\boldsymbol{x}$ is equal to $NH(X)$, even when the distribution $(p_1, p_2, \ldots, p_\ell)$ are non conmesurable values. Here, $H(X) = -\frac{1}{N} \log P_{\boldsymbol{X}}(\boldsymbol{x})$ is called the empirical entropy of a typical sequence.

The following lemma is instrumental in the understanding of typical sequences.

Lemma 59 (Chebyshev Inequality)**.** *Let X be a random variable with finite expected value μ,* i.e.*, $\mathbb{E}_X[X] = \mu < \infty$, and variance σ^2,* i.e.*, $\mathrm{Var}_X[X] = \sigma^2$. Then, for any $a > 0$, the following holds:*

$$\Pr[|X - \mu| \geqslant a] \leqslant \frac{\sigma^2}{a^2}. \tag{R.6}$$

R.1.1. Weak Typicality

Lemma 60 ([104, Theorem 5.1])**.** *Let X be a random variable defined on a countable set $\mathcal{X}$ with pmf $P_X : \mathcal{X} \to [0,1]$. Let also $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)^\mathsf{T} \in \mathcal{X}^N$ be an N-dimensional vector of random variables whose joint pmf is:*

$$P_{\boldsymbol{X}}(x_1, x_2 \ldots, x_N) = \prod_{n=1}^{N} P_X(x_n), \tag{R.7}$$

for all $(x_1, x_2 \ldots, x_N) \in \mathcal{X}^N$*. Then, for any* $\epsilon > 0$ *arbitrarily small, there always exists an* N *sufficiently large such that* $\boldsymbol{X}$ *satisfies:*

$$\Pr\left[\left|-\frac{1}{N}\log P_{\boldsymbol{X}}\left(\boldsymbol{X}\right) - H(X)\right| < \epsilon\right] \geqslant 1 - \epsilon. \tag{R.8}$$

Proof: Let the discrete random variable Y be defined by:

$$Y = -\frac{1}{N}\log P_{\boldsymbol{X}}\left(\boldsymbol{X}\right). \tag{R.9}$$

Note that

$$\mathbb{E}_Y\left[Y\right] = \mathbb{E}_{\boldsymbol{X}}\left[-\frac{1}{N}\log P_{\boldsymbol{X}}\left(\boldsymbol{X}\right)\right] \tag{R.10a}$$

$$= -\frac{1}{N}\sum_{\boldsymbol{x}\in\operatorname{supp} P_{\boldsymbol{X}}} P_{\boldsymbol{X}}\left(\boldsymbol{x}\right)\log P_{\boldsymbol{X}}\left(\boldsymbol{x}\right) \tag{R.10b}$$

$$= -\frac{1}{N}\sum_{n=1}^{N}\sum_{\boldsymbol{x}\in\operatorname{supp} P_{\boldsymbol{X}}} P_{\boldsymbol{X}}\left(\boldsymbol{x}\right)\log P_X\left(x_n\right) \tag{R.10c}$$

$$= -\frac{1}{N}\sum_{n=1}^{N}\sum_{x_1\in\operatorname{supp} P_X}\sum_{x_2\in\operatorname{supp} P_X}\ldots\sum_{x_N\in\operatorname{supp} P_X} P_X\left(x_1\right)P_X\left(x_2\right)\ldots P_X\left(x_N\right)\log P_X\left(x_n\right) \tag{R.10d}$$

$$= -\frac{1}{N}\sum_{n=1}^{N}\sum_{x_n\in\operatorname{supp} P_X} P_X\left(x_n\right)\log P_X\left(x_n\right) \tag{R.10e}$$

$$= \frac{1}{N}\sum_{n=1}^{N} H(X) \tag{R.10f}$$

$$= H(X) \tag{R.10g}$$

and

$$\operatorname{Var}_Y\left[Y\right] = \operatorname{Var}_{\boldsymbol{X}}\left[-\frac{1}{N}\log P_{\boldsymbol{X}}\left(\boldsymbol{X}\right)\right] \tag{R.11a}$$

$$= \frac{1}{N^2}\operatorname{Var}_{\boldsymbol{X}}\left[\log P_{\boldsymbol{X}}\left(\boldsymbol{X}\right)\right] \tag{R.11b}$$

$$= \frac{1}{N^2}\sum_{n=1}^{N}\operatorname{Var}_{X_n}\left[\log P_{X_n}\left(X_n\right)\right] \tag{R.11c}$$

$$= \frac{1}{N}\operatorname{Var}_X\left[\log P_X\left(X\right)\right], \tag{R.11d}$$

where (R.10c) and (R.11c) follow from the fact that all the random variables in the vector of random variables are independent (R.7).

From Chebyshev inequality (Lemma 59), it holds for any $a > 0$ that:

$$\Pr\left[\left|Y - \mathbb{E}_Y\left[Y\right]\right| \geqslant a\right] \leqslant \frac{\operatorname{Var}_Y\left[Y\right]}{a^2}. \tag{R.12}$$

That is,

$$\Pr\left[\left|-\frac{1}{N}\log P_{\boldsymbol{X}}\left(\boldsymbol{X}\right)-H(X)\right|\geqslant a\right]\leqslant\frac{1}{a^2N}\mathrm{Var}_X\left[\log P_X\left(x\right)\right]. \tag{R.13}$$

Note that since the random variable X has finite expected value and a finite variance, it follows that $\frac{1}{a^2}\mathrm{Var}_X\left[\log P_X\left(x\right)\right]$ is always finite.

Thus, for all $\epsilon' > 0$, there always exists an N sufficiently large, such that

$$\Pr\left[\left|-\frac{1}{N}\log P_{\boldsymbol{X}}\left(\boldsymbol{X}\right)-H(X)\right|\geqslant a\right]\leqslant\epsilon'. \tag{R.14}$$

Finally, note that

$$\Pr\left[\left|-\frac{1}{N}\log P_{\boldsymbol{X}}\left(\boldsymbol{X}\right)-H(X)\right|< a\right]=1-\Pr\left[\left|-\frac{1}{N}\log P_{\boldsymbol{X}}\left(\boldsymbol{X}\right)-H(X)\right|\geqslant a\right] \tag{R.15a}$$

$$\geqslant 1-\epsilon'. \tag{R.15b}$$

Therefore, for all $\epsilon > 0$, there always exists an N sufficiently large such that

$$\Pr\left[\left|-\frac{1}{N}\log P_{\boldsymbol{X}}\left(\boldsymbol{X}\right)-H(X)\right|<\epsilon\right]\geqslant 1-\epsilon. \tag{R.16}$$

This completes the proof.

■

Remark 1. *Since the probability space is discrete and finite, it follows from Vitali convergence theorem [77] that the convergence in probability of* $-\frac{1}{N}\log P_{\boldsymbol{X}}\left(\boldsymbol{X}\right)$ *to* $H(X)$, i.e., $-\frac{1}{N}\log P_{\boldsymbol{X}}\left(\boldsymbol{X}\right)\xrightarrow{p}H(X)$ *established in Lemma 60 implies the* $\mathscr{L}^1$ *convergence of* $-\frac{1}{N}\log P_{\boldsymbol{X}}\left(\boldsymbol{X}\right)$ *to* $H(X)$, i.e., $-\frac{1}{N}\log P_{\boldsymbol{X}}\left(\boldsymbol{X}\right)\xrightarrow{\mathscr{L}^1}H(X)$.

Definition 20 (Weakly Typical Set)**.** *Consider a random variable* $X\in\mathcal{X}$ *distributed according to* P_X *and the joint pmf of the* N*-dimensional vector of random variables* $\boldsymbol{X}$ *in* (R.7)*. For any* $\epsilon>0$ *arbitrarily small, the set of weakly typical sequences with respect to* P_X *is the set of sequences* $\boldsymbol{x}=(x_1,x_2,\ldots,x_N)\in\mathcal{X}^N$*, denoted by* $\mathcal{T}_X^{(N,\epsilon)}$*, such that:*

$$\mathcal{T}_X^{(N,\epsilon)}=\left\{\boldsymbol{x}\in\mathcal{X}^N:\left|-\frac{1}{N}\log P_{\boldsymbol{X}}\left(\boldsymbol{x}\right)-H(X)\right|<\epsilon\right\}. \tag{R.17}$$

The expression $-\frac{1}{N}\log P_{\boldsymbol{X}}\left(\boldsymbol{x}\right)$ is called the empirical entropy of a weakly typical sequence. The typical sequences are those sequences that have probability close to $2^{-NH(X)}$. Note that the most probable sequence and the least probable sequence are not necessarily typical sequences. Nonetheless, the set formed by the typical sequences has a probability measure close to one as N increases. Note also that $\mathcal{T}_X$ depends only on N, ϵ, and the distribution P_X. Figure R.1 shows that the empirical entropy of a binary sequence approaches to the entropy of a binary random variable for $N\in\mathbb{N}$ sufficiently large.

Lemma 61 (Weak AEP)**.** *Let* $\mathcal{T}_X^{(N,\epsilon)}$ *be the set of weakly typical sequences with respect to* P_X

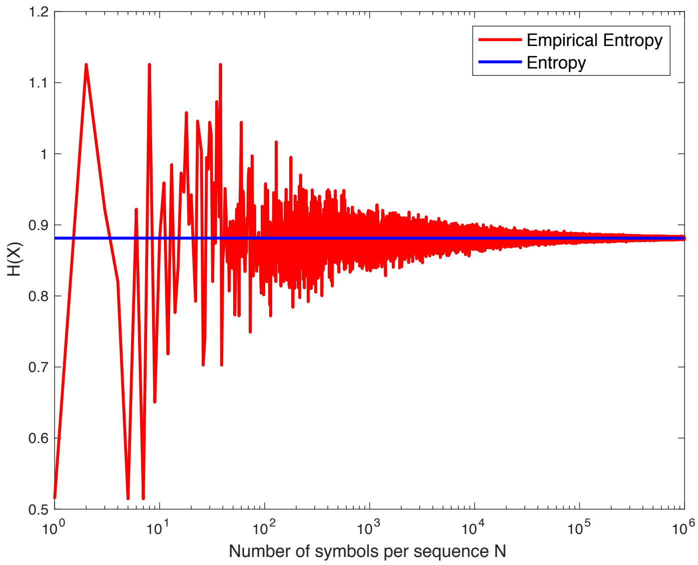


Figure R.1.: Empirical entropy of random binary sequence with $P_X(0) = 1 - P_X(1) = 0.3$.

and with $\epsilon > 0$*. Then, for* $N \in \mathbb{N}$ *sufficiently large and for all* $\boldsymbol{x} \in \mathcal{T}_X^{(N,\epsilon)}$*, the following holds:*

$$2^{-N(H(X)+\epsilon)} < P_{\boldsymbol{X}}(\boldsymbol{x}) < 2^{-N(H(X)-\epsilon)}, \tag{R.18a}$$

$$\sum_{\boldsymbol{x} \in \mathcal{T}_X^{(N,\epsilon)}} P_{\boldsymbol{X}}(\boldsymbol{x}) \geqslant 1 - \epsilon, \ \textit{and} \tag{R.18b}$$

$$(1-\epsilon)2^{N(H(X)-\epsilon)} < \left|\mathcal{T}_X^{(N,\epsilon)}\right| < 2^{N(H(X)+\epsilon)}. \tag{R.18c}$$

Proof:

Proof of (R.18a): This is obtained directly from Definition 20.

Proof of (R.18b): From (R.8), the following holds:

$$\sum_{\boldsymbol{x} \in \mathcal{T}_X^{(N,\epsilon)}} P_{\boldsymbol{X}}(\boldsymbol{x}) \geqslant 1 - \epsilon, \tag{R.19}$$

with $\epsilon > 0$, and this completes the proof of (R.18b).

Proof of (R.18c): From (R.18a) and (R.18b), the following holds:

$$1 = \sum_{\boldsymbol{x} \in \mathcal{X}^N} P_{\boldsymbol{X}}(\boldsymbol{x}) \tag{R.20a}$$

$$\geqslant \sum_{\boldsymbol{x} \in \mathcal{T}_X^{(N,\epsilon)}} P_{\boldsymbol{X}(\boldsymbol{x})} \tag{R.20b}$$

$$> \sum_{\boldsymbol{x} \in \mathcal{T}_X^{(N,\epsilon)}} 2^{-N(H(X)+\epsilon)} \tag{R.20c}$$

$$= \left|\mathcal{T}_X^{(N,\epsilon)}\right| 2^{-N(H(X)+\epsilon)}, \tag{R.20d}$$

for N sufficiently large, which implies:

$$\left|\mathcal{T}_X^{(N,\epsilon)}\right| < 2^{N(H(X)+\epsilon)}, \tag{R.20e}$$

and

$$1-\epsilon \leqslant \sum_{\boldsymbol{x}\in\mathcal{T}_X^{(N,\epsilon)}} P_{\boldsymbol{X}}(\boldsymbol{x}) \tag{R.20f}$$
$$< \sum_{\boldsymbol{x}\in\mathcal{T}_X^{(N,\epsilon)}} 2^{-N(H(X)-\epsilon)} \tag{R.20g}$$
$$= \left|\mathcal{T}_X^{(N,\epsilon)}\right| 2^{-N(H(X)-\epsilon)}, \tag{R.20h}$$

for N sufficiently large, which implies:

$$\left|\mathcal{T}_X^{(N,\epsilon)}\right| > (1-\epsilon)\, 2^{N(H(X)-\epsilon)}, \tag{R.20i}$$

and this completes the proof of (R.18c). This completes the proof of Lemma 61. ■

R.1.2. Weak Joint Typicality

The notion of typicality can be extended to multiple vectors of random variables.

Lemma 62. *Let $\mathcal{X}$ and $\mathcal{Y}$ be two countable sets and let also X and Y be two random variables with joint pmf $P_{XY}: \mathcal{X}\times\mathcal{Y} \to [0,1]$. Let also $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)^{\mathsf{T}}$ and $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_N)^{\mathsf{T}}$ be two N-dimensional vectors of random variables whose joint pmf is:*

$$P_{\boldsymbol{XY}}(\boldsymbol{x},\boldsymbol{y}) = \prod_{n=1}^{N} P_{XY}(x_n, y_n), \tag{R.21}$$

for all $(x_1, x_2 \ldots, x_N) \in \mathcal{X}^N$ and $(y_1, y_2 \ldots, y_N) \in \mathcal{Y}^N$. Then, for any $\epsilon > 0$ arbitrarily small, there always exists an N sufficiently large such that $\boldsymbol{X}$ and $\boldsymbol{Y}$ satisfies:

$$\Pr\left[\left|-\frac{1}{N}\log P_{\boldsymbol{XY}}(\boldsymbol{X},\boldsymbol{Y}) - H(X,Y)\right| < \epsilon\right] \geqslant 1-\epsilon. \tag{R.22}$$

Proof: This proof follows along the same lines as the proof of Lemma 60. Then, let the discrete random variable Z be defined by:

$$Z = -\frac{1}{N}\log P_{\boldsymbol{XY}}(\boldsymbol{XY}). \tag{R.23}$$

Note that

$$\mathbb{E}_Z[Z] = H(X,Y) \text{ and} \tag{R.24a}$$
$$\mathrm{Var}_Z[Z] = \frac{1}{N}\mathrm{Var}_{XY}[\log P_{XY}(X,Y)]. \tag{R.24b}$$

From Chebyshev inequality (Lemma 59), it holds for any $a > 0$ that:

$$\Pr[|Z - \mathbb{E}_Z[Z]| \geqslant a] \leqslant \frac{\mathrm{Var}_Z[Z]}{a^2}. \tag{R.25}$$

That is,

$$\Pr\left[\left|-\frac{1}{N}\log P_{\boldsymbol{XY}}\left(\boldsymbol{XY}\right)-H(X,Y)\right|\geqslant a\right]\leqslant\frac{1}{a^2N}\mathrm{Var}_{XY}\left[\log P_{XY}\left(X,Y\right)\right]. \tag{R.26}$$

Note that since the random variables X and Y have a finite joint expected value and a finite joint variance, it follows that $\frac{1}{a^2}\mathrm{Var}_{XY}\left[\log P_{XY}\left(X,Y\right)\right]$ is always finite.

Thus, for all $\epsilon' > 0$, there always exists an N sufficiently large, such that

$$\Pr\left[\left|-\frac{1}{N}\log P_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)-H(X,Y)\right|\geqslant a\right]\leqslant\epsilon'. \tag{R.27}$$

Finally, note that

$$\Pr\left[\left|-\frac{1}{N}\log P_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)-H(X,Y)\right|< a\right]=1-\Pr\left[\left|-\frac{1}{N}\log P_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)-H(X,Y)\right|\geqslant a\right] \tag{R.28a}$$

$$\geqslant 1-\epsilon'. \tag{R.28b}$$

Therefore, for all $\epsilon > 0$, there always exists an N sufficiently large such that

$$\Pr\left[\left|-\frac{1}{N}\log P_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)-H(X,Y)\right|<\epsilon\right]\geqslant 1-\epsilon. \tag{R.29}$$

This completes the proof.

■

Remark 2. *Since the probability space is discrete and finite, it follows from Vitali convergence theorem [77] that the convergence in probability of* $-\frac{1}{N}\log P_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)$ *to* $H(X,Y)$, i.e., $-\frac{1}{N}\log P_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)\xrightarrow{p}H(X,Y)$ *established in Lemma 62 implies the* $\mathscr{L}^1$ *convergence of* $-\frac{1}{N}\log P_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)$ *to* $H(X,Y)$*, i.e.,* $-\frac{1}{N}\log P_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)\xrightarrow{\mathscr{L}^1}H(X,Y)$.

Definition 21 (Weakly Joint Typical Set)**.** *Consider two random variables* $X\in\mathcal{X}$ *and* $Y\in\mathcal{Y}$ *distributed according to* P_{XY}*, and the pmfs and joint pmf of the* N*-dimensional vectors of random variables* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *according to* (R.7), $P_{\boldsymbol{Y}}(y_1,y_2\ldots,y_N)=\prod_{n=1}^{N}P_Y(y_n)$, *and* (R.21). *For any* $\epsilon>0$ *arbitrarily small, the set of weakly joint typical sequences with respect to* P_{XY} *is the set of sequences* $((x_1,y_1),(x_2,y_2),\ldots,(x_N,y_N))\in(\mathcal{X}\times\mathcal{Y})^N$*, denoted by* $\mathcal{T}_{XY}^{(N,\epsilon)}$*, such that:*

$$\begin{aligned}\mathcal{T}_{XY}^{(N,\epsilon)}=\Bigg\{(\boldsymbol{x},\boldsymbol{y})\in(\mathcal{X}\times\mathcal{Y})^N:&\left|-\frac{1}{N}\log\left(P_{\boldsymbol{X}}\left(\boldsymbol{x}\right)\right)-H(X)\right|<\epsilon,\\ &\left|-\frac{1}{N}\log\left(P_{\boldsymbol{Y}}\left(\boldsymbol{y}\right)\right)-H(Y)\right|<\epsilon,\ \textit{and}\\ &\left|-\frac{1}{N}\log\left(P_{\boldsymbol{XY}}\left(\boldsymbol{x},\boldsymbol{y}\right)\right)-H(X,Y)\right|<\epsilon\Bigg\}.\end{aligned} \tag{R.30}$$

Note that if $(\boldsymbol{x},\boldsymbol{y})\in\mathcal{T}_{XY}^{(N,\epsilon)}$ then $\boldsymbol{x}\in\mathcal{T}_{X}^{(N,\epsilon)}$ and $\boldsymbol{y}\in\mathcal{T}_{Y}^{(N,\epsilon)}$.

Lemma 63. *Let $\mathcal{T}_{X\,Y}^{(N,\epsilon)}$ be the set of weakly joint typical sequences with respect to P_{XY} and with $\epsilon > 0$. Then, for N sufficiently large and for all $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{T}_{XY}^{(N,\epsilon)}$, the following holds:*

$$2^{-N(H(X,Y)+\epsilon)} < P_{\boldsymbol{XY}}(\boldsymbol{x}, \boldsymbol{y}) < 2^{-N(H(X,Y)-\epsilon)}, \tag{R.31a}$$

$$2^{-N(H(X|Y)+2\epsilon)} < P_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) < 2^{-N(H(X|Y)-2\epsilon)}, \tag{R.31b}$$

$$\sum_{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{T}_{XY}^{(N,\epsilon)}} P_{\boldsymbol{XY}}(\boldsymbol{x}, \boldsymbol{y}) \geqslant 1 - \epsilon, \text{ and} \tag{R.31c}$$

$$(1-\epsilon) 2^{N(H(X,Y)-\epsilon)} < |\mathcal{T}_{XY}^{(N,\epsilon)}| < 2^{N(H(X,Y)+\epsilon)}. \tag{R.31d}$$

Proof:

Proof of (R.31a): This is obtained directly from Definition 21.

Proof of (R.31b): From the assumptions of the lemma, following along the steps of the proof of (R.18a) for all $\boldsymbol{x} \in \mathcal{T}_X^{(N,\epsilon)}$ and for all $\boldsymbol{y} \in \mathcal{T}_Y^{(N,\epsilon)}$ yields:

$$2^{-N(H(X)+\epsilon)} < P_{\boldsymbol{X}}(\boldsymbol{x}) < 2^{-N(H(X)-\epsilon)} \text{ and} \tag{R.32a}$$

$$2^{-N(H(Y)+\epsilon)} < P_{\boldsymbol{Y}}(\boldsymbol{y}) < 2^{-N(H(Y)-\epsilon)}. \tag{R.32b}$$

From (R.21) and (R.32b), the following holds:

$$2^{-N(H(X|Y)+2\epsilon)} < P_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) < 2^{-N(H(X|Y)-2\epsilon)}, \tag{R.32c}$$

and this completes the proof of (R.31b).

Proof of (R.31c): From Lemma 62, the following holds:

$$\sum_{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{T}_{XY}^{(N,\epsilon)}} P_{\boldsymbol{XY}}(\boldsymbol{x}, \boldsymbol{y}) \geqslant 1 - \epsilon, \tag{R.33}$$

with $\epsilon > 0$, and this completes the proof of (R.31d).

Proof of (R.31d): From (R.31a) and (R.31c), the following holds:

$$1 = \sum_{(\boldsymbol{x},\boldsymbol{y}) \in (\mathcal{X} \times \mathcal{Y})^N} P_{\boldsymbol{XY}}(\boldsymbol{x}, \boldsymbol{y}) \tag{R.34a}$$

$$\geqslant \sum_{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{T}_{XY}^{(N,\epsilon)}} P_{\boldsymbol{XY}}(\boldsymbol{x}, \boldsymbol{y}) \tag{R.34b}$$

$$> \sum_{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{T}_{XY}^{(N,\epsilon)}} 2^{-N(H(X,Y)+\epsilon)} \tag{R.34c}$$

$$= \left| \mathcal{T}_{XY}^{(N,\epsilon)} \right| 2^{-N(H(X,Y)+\epsilon)}, \tag{R.34d}$$

for N sufficiently large, which implies:

$$\left| \mathcal{T}_{XY}^{(N,\epsilon)} \right| < 2^{N(H(X,Y)+\epsilon)}, \tag{R.34e}$$

and

$$1 - \epsilon \leqslant \sum_{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{T}_{XY}^{(N,\epsilon)}} P_{\boldsymbol{XY}}(\boldsymbol{x},\boldsymbol{y}) \tag{R.34f}$$

$$< \sum_{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{T}_{XY}^{(N,\epsilon)}} 2^{-N(H(X,Y)-\epsilon)} \tag{R.34g}$$

$$= \left|\mathcal{T}_{XY}^{(N,\epsilon)}\right| 2^{-N(H(X,Y)-\epsilon)}, \tag{R.34h}$$

for N sufficiently large, which implies:

$$\left|\mathcal{T}_{\boldsymbol{X}\ \boldsymbol{Y}}^{(N,\epsilon)}\right| > (1-\epsilon)\, 2^{N(H(X,Y)-\epsilon)}, \tag{R.34i}$$

and this completes the proof. ∎

R.1.3. Weak Conditional Typicality

Definition 22 (Weakly Typical Set Subject to Conditioning)**.** *Consider two random variables $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$ distributed according to P_{XY}, and the conditional pmf of the N-dimensional vectors of random variables $\boldsymbol{X}$ and $\boldsymbol{Y}$ according to $P_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) = \prod_{n=1}^{N} P_{X|Y}(x_n|y_n)$ for all $\boldsymbol{x} \in \mathcal{X}^n$ and $\boldsymbol{y} \in \mathcal{Y}^n$. Let $\boldsymbol{y} = (y_1, \ldots, y_N)^{\mathsf{T}}$ be a sequence such that $\boldsymbol{y} \in \mathcal{T}_Y^{(N,\epsilon)}$, with $\epsilon > 0$ and N sufficiently large. Then, the set of weakly typical sequences with respect to P_X conditioning on the sequence $\boldsymbol{y}$ is the set of sequences $\boldsymbol{x} = (x_1, x_2, \ldots, x_N) \in \mathcal{X}^N$, denoted by $\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})$, such that:*

$$\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y}) = \Bigg\{\boldsymbol{x} \in \mathcal{X}^N : \left|-\frac{1}{N}\log\left(P_{\boldsymbol{X}}\left(\boldsymbol{x}\right)\right) - H(X)\right| < \epsilon, \text{ and } \left|-\frac{1}{N}\log\left(P_{\boldsymbol{XY}}\left(\boldsymbol{x},\boldsymbol{y}\right)\right) - H(X,Y)\right| < \epsilon\Bigg\}. \tag{R.35}$$

Lemma 64. *For any $\boldsymbol{y}$, let $\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})$ be the set of weakly typical sequences with respect to P_{XY} conditioning on $\boldsymbol{Y} = \boldsymbol{y}$ and with $\epsilon > 0$. Then, for $N \in \mathbb{N}$ sufficiently large and for all $\boldsymbol{x} \in \mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})$, the following holds:*

$$\left|\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})\right| < 2^{N(H(X|Y)+2\epsilon)} \tag{R.36a}$$

$$\sum_{\boldsymbol{y} \in \mathcal{Y}^N} P_{\boldsymbol{Y}}(\boldsymbol{y}) \left|\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})\right| > (1-\epsilon) 2^{N(H(X|Y)-2\epsilon)}. \tag{R.36b}$$

Proof:

Proof of (R.36a):

$$1= \sum_{(\boldsymbol{x},\boldsymbol{y})\in(\mathcal{X}\times\mathcal{Y})^N} P_{\boldsymbol{XY}}(\boldsymbol{x},\boldsymbol{y}) \tag{R.37a}$$

$$= \sum_{\boldsymbol{y}\in\mathcal{Y}^N} P_{\boldsymbol{Y}}(\boldsymbol{y}) \sum_{\boldsymbol{x}\in\mathcal{X}^N} P_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) \tag{R.37b}$$

$$> \sum_{\boldsymbol{y}\in\mathcal{Y}^N} P_{\boldsymbol{Y}}(\boldsymbol{y}) \sum_{\boldsymbol{x}\in\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})} P_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{Y}=\boldsymbol{y}) \tag{R.37c}$$

$$> \sum_{\boldsymbol{y}\in\mathcal{Y}^N} P_{\boldsymbol{Y}}(\boldsymbol{y}) \sum_{\boldsymbol{x}\in\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})} 2^{-N(H(X|Y)+2\epsilon)} \tag{R.37d}$$

$$= \sum_{\boldsymbol{y}\in\mathcal{Y}^N} P_{\boldsymbol{Y}}(\boldsymbol{y}) \left|\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})\right| 2^{-N(H(X|Y)+2\epsilon)} \tag{R.37e}$$

$$= \left|\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})\right| 2^{-N(H(X|Y)+2\epsilon)} \sum_{\boldsymbol{y}\in\mathcal{Y}^N} P_{\boldsymbol{Y}}(\boldsymbol{y}) \tag{R.37f}$$

$$= \left|\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})\right| 2^{-N(H(X|Y)+2\epsilon)}, \tag{R.37g}$$

for N sufficiently large, which implies:

$$\left|\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})\right| < 2^{N(H(X|Y)+2\epsilon)}, \tag{R.37h}$$

and this completes the proof of (R.36a).

Proof of (R.36b):

$$1-\epsilon \leqslant \sum_{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{T}_{XY}^{(N,\epsilon)}} P_{\boldsymbol{XY}}(\boldsymbol{x},\boldsymbol{y}) \tag{R.38a}$$

$$= \sum_{\boldsymbol{y}\in\mathcal{T}_{Y}^{(N,\epsilon)}} P_{\boldsymbol{Y}}(\boldsymbol{y}) \sum_{\boldsymbol{x}\in\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})} P_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) \tag{R.38b}$$

$$\leqslant \sum_{\boldsymbol{y}\in\mathcal{Y}^N} P_{\boldsymbol{Y}}(\boldsymbol{y}) \sum_{\boldsymbol{x}\in\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})} P_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) \tag{R.38c}$$

$$< \sum_{\boldsymbol{y}\in\mathcal{Y}^N} P_{\boldsymbol{Y}}(\boldsymbol{y}) \sum_{\boldsymbol{x}\in\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})} 2^{-N(H(X|Y)-2\epsilon)} \tag{R.38d}$$

$$= \sum_{\boldsymbol{y}\in\mathcal{Y}^N} P_{\boldsymbol{Y}}(\boldsymbol{y}) \left|\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})\right| 2^{-N(H(X|Y)-2\epsilon)} \tag{R.38e}$$

$$= 2^{-N(H(X|Y)-2\epsilon)} \sum_{\boldsymbol{y}\in\mathcal{Y}^N} P_{\boldsymbol{Y}}(\boldsymbol{y}) \left|\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})\right|, \tag{R.38f}$$

for N sufficiently large, which implies:

$$\sum_{\boldsymbol{y}\in\mathcal{Y}^N} P_{\boldsymbol{Y}}(\boldsymbol{y}) \left|\mathcal{T}_{X|Y}^{(N,\epsilon)}(\boldsymbol{y})\right| > (1-\epsilon) 2^{N(H(X|Y)-2\epsilon)}, \tag{R.38g}$$

and this completes the proof of (R.36b). This completes the proof of Lemma 64. ■

R.2. Real-Valued Random Variables

R.2.1. Weak Typicality

Lemma 65 ([104, Theorem 10.35])**.** *Let X be a random variable X with pdf $f_X : \mathbb{R} \to [0, \infty)$. Let also $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)^{\mathsf{T}} \in \mathcal{X}^N$ be an N-dimensional vector of random variables whose joint pdf is:*

$$f_{\boldsymbol{X}}(x_1, x_2 \ldots, x_N) = \prod_{n=1}^{N} f_X(x_n), \tag{R.39}$$

for all $(x_1, x_2 \ldots, x_N) \in \mathbb{R}^N$. Then, for any $\epsilon > 0$ arbitrarily small, there always exists an N sufficiently large such that $\boldsymbol{X}$ satisfies:

$$\Pr\left[\left|-\frac{1}{N}\log f_{\boldsymbol{X}}\left(\boldsymbol{X}\right) - h(X)\right| < \epsilon\right] \geqslant 1 - \epsilon. \tag{R.40}$$

Proof: Let the real-valued random variable Y be defined by:

$$Y = -\frac{1}{N}\log f_{\boldsymbol{X}}\left(\boldsymbol{X}\right). \tag{R.41}$$

Note that

$$\mathbb{E}_Y\left[Y\right]=\mathbb{E}_{\boldsymbol{X}}\left[-\frac{1}{N}\log f_{\boldsymbol{X}}\left(\boldsymbol{X}\right)\right] \tag{R.42a}$$

$$=-\frac{1}{N}\int_{-\infty}^{\infty} f_{\boldsymbol{X}}\left(\boldsymbol{x}\right)\log f_{\boldsymbol{X}}\left(\boldsymbol{x}\right)\mathrm{d}\boldsymbol{x} \tag{R.42b}$$

$$=-\frac{1}{N}\sum_{n=1}^{N}\int_{-\infty}^{\infty} f_{\boldsymbol{X}}\left(\boldsymbol{x}\right)\log f_X\left(x_n\right)\mathrm{d}x_n \tag{R.42c}$$

$$=-\frac{1}{N}\sum_{n=1}^{N}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} f_X\left(x_1\right)f_X\left(x_2\right)\ldots f_X\left(x_N\right)\log f_X\left(x_n\right)\mathrm{d}x_1\mathrm{d}x_2\ldots\mathrm{d}x_N \tag{R.42d}$$

$$=-\frac{1}{N}\sum_{n=1}^{N}\int_{-\infty}^{\infty} f_X\left(x_n\right)\log f_X\left(x_n\right) \tag{R.42e}$$

$$=\frac{1}{N}\sum_{n=1}^{N} h(X) \tag{R.42f}$$

$$=h(X) \tag{R.42g}$$

and

$$\mathrm{Var}_Y\left[Y\right]=\mathrm{Var}_{\boldsymbol{X}}\left[-\frac{1}{N}\log f_{\boldsymbol{X}}\left(\boldsymbol{X}\right)\right] \tag{R.43a}$$

$$=\frac{1}{N^2}\mathrm{Var}_{\boldsymbol{X}}\left[\log f_{\boldsymbol{X}}\left(\boldsymbol{X}\right)\right] \tag{R.43b}$$

$$=\frac{1}{N^2}\sum_{n=1}^{N}\mathrm{Var}_{X_n}\left[\log f_{X_n}\left(X_n\right)\right] \tag{R.43c}$$

$$=\frac{1}{N}\mathrm{Var}_X\left[\log f_X\left(X\right)\right], \tag{R.43d}$$

where (R.42c) and (R.43c) follow from the fact that all the random variables in the vector of random variables are independent (R.39).

From Chebyshev inequality (Lemma 59), it holds for any $a > 0$ that:

$$\Pr\left[|Y - \mathbb{E}_Y[Y]| \geqslant a\right] \leqslant \frac{\mathrm{Var}_Y[Y]}{a^2}. \tag{R.44}$$

That is,

$$\Pr\left[\left|-\frac{1}{N}\log f_{\boldsymbol{X}}(\boldsymbol{X}) - h(X)\right| \geqslant a\right] \leqslant \frac{1}{a^2 N}\mathrm{Var}_X\left[\log f_X(x)\right]. \tag{R.45}$$

Note that since the random variable X has finite expected value and a finite variance, it follows that $\frac{1}{a^2}\mathrm{Var}_X\left[\log f_X(x)\right]$ is always finite.

Thus, for all $\epsilon' > 0$, there always exists an N sufficiently large, such that

$$\Pr\left[\left|-\frac{1}{N}\log f_{\boldsymbol{X}}(\boldsymbol{X}) - h(X)\right| \geqslant a\right] \leqslant \epsilon'. \tag{R.46}$$

Finally, note that

$$\Pr\left[\left|-\frac{1}{N}\log f_{\boldsymbol{X}}(\boldsymbol{X}) - h(X)\right| < a\right] = 1 - \Pr\left[\left|-\frac{1}{N}\log f_{\boldsymbol{X}}(\boldsymbol{X}) - h(X)\right| \geqslant a\right] \tag{R.47a}$$

$$\geqslant 1 - \epsilon'. \tag{R.47b}$$

Therefore, for all $\epsilon > 0$, there always exists an N sufficiently large such that

$$\Pr\left[\left|-\frac{1}{N}\log f_{\boldsymbol{X}}(\boldsymbol{X}) - h(X)\right| < \epsilon\right] \geqslant 1 - \epsilon. \tag{R.48}$$

This completes the proof.

Remark 3. *Since the probability space is real-valued, it follows from Vitali convergence theorem [77] that the convergence in probability of* $-\frac{1}{N}\log f_{\boldsymbol{X}}(\boldsymbol{X})$ *to* $H(X)$, i.e., $-\frac{1}{N}\log f_{\boldsymbol{X}}(\boldsymbol{X}) \xrightarrow{p} H(X)$ *established in Lemma 65 implies the* $\mathscr{L}^1$ *convergence of* $-\frac{1}{N}\log f_{\boldsymbol{X}}(\boldsymbol{X})$ *to* $h(X)$, i.e., $-\frac{1}{N}\log f_{\boldsymbol{X}}(\boldsymbol{X}) \xrightarrow{\mathscr{L}^1} h(X)$.

■

Definition 23 (Weakly Typical Set)**.** *Consider a random variable* $X \in \mathbb{R}$ *distributed according to* f_X *and the joint pdf of the* N*-dimensional vector of random variables* $\boldsymbol{X}$ *in* (R.39)*. For any* $\epsilon > 0$ *arbitrarily small, the set of weakly typical sequences with respect to* f_X *is the set of sequences* $\boldsymbol{x} = (x_1, x_2, \ldots, x_N) \in \mathcal{X}^N$*, denoted by* $\mathcal{T}_X^{(N,\epsilon)}$*, such that:*

$$\mathcal{T}_X^{(N,\epsilon)} = \left\{\boldsymbol{x} \in \mathcal{X}^N : \left|-\frac{1}{N}\log f_{\boldsymbol{X}}(\boldsymbol{x}) - h(X)\right| < \epsilon\right\}, \tag{R.49}$$

where, ϵ *is an arbitrarily small positive real number and* $-\frac{1}{N}\log f_{\boldsymbol{X}}(\boldsymbol{x})$ *is called the empirical differential entropy of a weakly typical sequence.*

The expression $-\frac{1}{N}\log f_{\boldsymbol{X}}(\boldsymbol{x})$ is called the empirical differential entropy of a weakly typical sequence. Note also that $\mathcal{T}_X$ depends only on N, ϵ, and the distribution f_X.

Definition 24. *The volume of a set $\mathcal{A} \subseteq \mathbb{R}^N$, denoted by* Vol($\mathcal{A}$)*, is:*

$$\mathrm{Vol}(\mathcal{A}) = \int_{\mathcal{A}\subseteq\mathbb{R}^N} d\boldsymbol{x}. \tag{R.50}$$

Lemma 66 (Weak AEP [104, Theorem 10.38])**.** *Let $\mathcal{T}_X^{(N,\epsilon)}$ be the set of weakly typical sequences with respect to f_X, with $\epsilon > 0$. Then, for $N \in \mathbb{N}$ sufficiently large and for all $\boldsymbol{x} \in \mathcal{T}_X^{(N,\epsilon)}$, the following holds:*

$$2^{-N(h(X)+\epsilon)} < f_{\boldsymbol{X}}(\boldsymbol{x}) < 2^{-N(h(X)-\epsilon)}, \tag{R.51a}$$

$$\int_{\boldsymbol{x}\in\mathcal{T}_X^{(N,\epsilon)}} f_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x} \geqslant 1-\epsilon, \text{ and} \tag{R.51b}$$

$$(1-\epsilon)2^{N(h(X)-\epsilon)} < \mathrm{Vol}\left(\mathcal{T}_X^{(N,\epsilon)}\right) < 2^{N(h(X)+\epsilon)}. \tag{R.51c}$$

Proof:

Proof of (R.51a): This is obtained directly from Definition 23.

Proof of (R.51b): From (R.40), the following holds:

$$\int_{\boldsymbol{x}\in\mathcal{T}_X^{(N,\epsilon)}} f_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x} \geqslant 1-\epsilon, \tag{R.52}$$

with $\epsilon > 0$ and this completes the proof of (R.51b).

Proof of (R.51c): From (R.51a) and (R.51b), the following holds:

$$1 = \int_{\boldsymbol{x}\in\mathcal{X}^N} f_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x} \tag{R.53a}$$

$$\geqslant \int_{\boldsymbol{x}\in\mathcal{T}_X^{(N,\epsilon)}} f_{\boldsymbol{X}(\boldsymbol{x})}\, d\boldsymbol{x} \tag{R.53b}$$

$$> \int_{\boldsymbol{x}\in\mathcal{T}_X^{(N,\epsilon)}} 2^{-N(h(X)+\epsilon)}\, d\boldsymbol{x} \tag{R.53c}$$

$$= 2^{-N(h(X)+\epsilon)} \int_{\boldsymbol{x}\in\mathcal{T}_X^{(N,\epsilon)}} d\boldsymbol{x} \tag{R.53d}$$

$$= 2^{-N(h(X)+\epsilon)}\mathrm{Vol}\left(\mathcal{T}_X^{(N,\epsilon)}\right), \tag{R.53e}$$

for N sufficiently large, which implies:

$$\mathrm{Vol}\left(\mathcal{T}_X^{(N,\epsilon)}\right) < 2^{N(h(X)+\epsilon)}, \tag{R.53f}$$

and

$$
\begin{aligned}
1-\epsilon \leqslant & \int_{\boldsymbol{x} \in \mathcal{T}_X^{(N,\epsilon)}} f_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x} && \text{(R.53g)}\\
< & \int_{\boldsymbol{x} \in \mathcal{T}_X^{(N,\epsilon)}} 2^{-N(h(X)-\epsilon)}\, d\boldsymbol{x} && \text{(R.53h)}\\
= & 2^{-N(h(X)-\epsilon)} \int_{\boldsymbol{x} \in \mathcal{T}_X^{(N,\epsilon)}} d\boldsymbol{x} && \text{(R.53i)}\\
= & 2^{-N(h(X)-\epsilon)} \mathrm{Vol}\left(\mathcal{T}_X^{(N,\epsilon)}\right), && \text{(R.53j)}
\end{aligned}
$$

for N sufficiently large, which implies:

$$
\mathrm{Vol}\left(\mathcal{T}_X^{(N,\epsilon)}\right) > (1-\epsilon)\, 2^{N(h(X)-\epsilon)}, \tag{R.53k}
$$

and this completes the proof of (R.51c). This completes the proof of Lemma 66. ■

R.2.2. Weak Joint Typicality

Lemma 67. *Let X and Y be two random variables with joint pdf $f_{XY} : \mathbb{R}^2 \to [0,\infty)$. Let also $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)^\mathsf{T}$ and $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_N)^\mathsf{T}$ be two N-dimensional vectors of random variables whose joint pdf is:*

$$
f_{\boldsymbol{XY}}(\boldsymbol{x},\boldsymbol{y}) = \prod_{n=1}^{N} f_{XY}(x_n, y_n), \tag{R.54}
$$

for all $(x_1, x_2 \ldots, x_N) \in \mathbb{R}^N$ and $(y_1, y_2 \ldots, y_N) \in \mathbb{R}^N$. Then, for any $\epsilon > 0$ arbitrarily small, there always exists an N sufficiently large such that $\boldsymbol{X}$ and $\boldsymbol{Y}$ satisfies:

$$
\Pr\left[\left|-\frac{1}{N}\log f_{\boldsymbol{XY}}(\boldsymbol{X},\boldsymbol{Y}) - h(X,Y)\right| < \epsilon\right] > 1-\epsilon. \tag{R.55}
$$

Proof:

This proof follows along the same lines as the proof of Lemma 65. Then, let the real-valued random variable Z be defined by:

$$
Z = -\frac{1}{N}\log f_{\boldsymbol{XY}}(\boldsymbol{XY}). \tag{R.56}
$$

Note that

$$
\begin{aligned}
\mathbb{E}_Z[Z] = & h(X,Y) \text{ and} && \text{(R.57a)}\\
\mathrm{Var}_Z[Z] = & \frac{1}{N}\mathrm{Var}_{XY}[\log f_{XY}(X,Y)]. && \text{(R.57b)}
\end{aligned}
$$

From Chebyshev inequality (Lemma 59), it holds for any $a > 0$ that:

$$
\Pr[|Z - \mathbb{E}_Z[Z]| \geqslant a] \leqslant \frac{\mathrm{Var}_Z[Z]}{a^2}. \tag{R.58}
$$

That is,

$$\Pr\left[\left|-\frac{1}{N}\log f_{\boldsymbol{XY}}\left(\boldsymbol{XY}\right)-h(X,Y)\right|\geqslant a\right]\leqslant\frac{1}{a^2N}\mathrm{Var}_{XY}\left[\log f_{XY}\left(X,Y\right)\right].\tag{R.59}$$

Note that since the random variables X and Y have a finite joint expected value and a finite joint variance, it follows that $\frac{1}{a^2}\mathrm{Var}_{XY}\left[\log f_{XY}\left(X,Y\right)\right]$ is always finite. Thus, for all $\epsilon'>0$, there always exists an N sufficiently large, such that

$$\Pr\left[\left|-\frac{1}{N}\log f_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)-h(X,Y)\right|\geqslant a\right]\leqslant\epsilon'.\tag{R.60}$$

Finally, note that

$$\Pr\left[\left|-\frac{1}{N}\log f_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)-h(X,Y)\right|<a\right]=1-\Pr\left[\left|-\frac{1}{N}\log f_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)-h(X,Y)\right|\geqslant a\right]\tag{R.61a}$$

$$\geqslant 1-\epsilon'.\tag{R.61b}$$

Therefore, for all $\epsilon>0$, there always exists an N sufficiently large such that

$$\Pr\left[\left|-\frac{1}{N}\log f_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)-h(X,Y)\right|<\epsilon\right]\geqslant 1-\epsilon.\tag{R.62}$$

This completes the proof.

■

Remark 4. *Since the probability space is continuous, it follows from Vitali convergence theorem [77] that the convergence in probability of* $-\frac{1}{N}\log f_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)$ *to* $h(X,Y)$*,* i.e.*,* $-\frac{1}{N}\log f_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)\xrightarrow{p}h(X,Y)$ *established in Lemma 67 implies the* $\mathscr{L}^1$ *convergence of* $-\frac{1}{N}\log f_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)$ *to* $h(X,Y)$*,* i.e.*,* $-\frac{1}{N}\log f_{\boldsymbol{XY}}\left(\boldsymbol{X},\boldsymbol{Y}\right)\xrightarrow{\mathscr{L}^1}h(X,Y)$.

Definition 25 (Weakly Joint Typical Set)**.** *Consider two random variables* $X\in\mathbb{R}$ *and* $Y\in\mathbb{R}$ *distributed according to* f_{XY}*, and the pdfs and joint pdf of the* N*-dimensional vectors of random variables* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *according to* (R.7)*,* $P_{\boldsymbol{Y}}(y_1,y_2\dots,y_N)=\prod_{n=1}^{N}P_Y(y_n)$*, and* (R.54)*. For any* $\epsilon>0$ *arbitrarily small, the set of weakly joint typical sequences with respect to* P_{XY} *is the set of sequences* $((x_1,y_1),(x_2,y_2),\dots,(x_N,y_N))\in(\mathcal{X}\times\mathcal{Y})^N$*, denoted by* $\mathcal{T}_{XY}^{(N,\epsilon)}$*, such that:*

$$\begin{aligned}\mathcal{T}_{XY}^{(N,\epsilon)}=\Bigg\{(\boldsymbol{x},\boldsymbol{y})\in(\mathcal{X}\times\mathcal{Y})^N:&\left|-\frac{1}{N}\log\left(f_{\boldsymbol{X}}\left(\boldsymbol{x}\right)\right)-h(X)\right|<\epsilon,\\&\left|-\frac{1}{N}\log\left(f_{\boldsymbol{Y}}\left(\boldsymbol{y}\right)\right)-h(Y)\right|<\epsilon,\ \text{and}\\&\left|-\frac{1}{N}\log\left(f_{\boldsymbol{XY}}\left(\boldsymbol{x},\boldsymbol{y}\right)\right)-h(X,Y)\right|<\epsilon\Bigg\}.\end{aligned}\tag{R.63}$$

Note that if $(\boldsymbol{x},\boldsymbol{y})\in\mathcal{T}_{XY}^{(N,\epsilon)}$ then $\boldsymbol{x}\in\mathcal{T}_{X}^{(N,\epsilon)}$ and $\boldsymbol{y}\in\mathcal{T}_{Y}^{(N,\epsilon)}$.

Lemma 68. *Let $\mathcal{T}_{X\,Y}^{(N,\epsilon)}$ be the set of weakly joint typical sequences with respect to f_{XY} and with $\epsilon > 0$. Then, for $N \in \mathbb{N}$ sufficiently large and for all $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{T}_{XY}^{(N,\epsilon)}$, the following holds:*

$$2^{-N(h(X,Y)+\epsilon)} < f_{\boldsymbol{XY}}(\boldsymbol{x}, \boldsymbol{y}) < 2^{-N(h(X,Y)-\epsilon)}, \tag{R.64a}$$

$$2^{-N(h(X|Y)+2\epsilon)} < f_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) < 2^{-N(h(X|Y)-2\epsilon)}, \tag{R.64b}$$

$$\int_{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{T}_{XY}^{(N,\epsilon)}} f_{\boldsymbol{XY}}(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{x}\, d\boldsymbol{y} \geqslant 1 - \epsilon, \text{ and} \tag{R.64c}$$

$$(1-\epsilon)2^{N(h(X,Y)-\epsilon)} < \mathrm{Vol}\left(\mathcal{T}_{XY}^{(N,\epsilon)}\right) < 2^{N(h(X,Y)+\epsilon)}. \tag{R.64d}$$

Proof:

Proof of (R.64a): This is obtained directly from Definition 25.

Proof of (R.64b): From the assumptions of the lemma, it follows that

$$2^{-N(h(X)+\epsilon)} < f_{\boldsymbol{X}}(\boldsymbol{x}) < 2^{-N(h(X)-\epsilon)} \text{ and }, \tag{R.65a}$$

$$2^{-N(h(Y)+\epsilon)} < f_{\boldsymbol{Y}}(\boldsymbol{y}) < 2^{-N(h(Y)-\epsilon)}. \tag{R.65b}$$

From (R.64a) and (R.65b), the following holds:

$$2^{-N(h(X|Y)+2\epsilon)} < f_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) < 2^{-N(h(X|Y)-2\epsilon)}, \tag{R.65c}$$

and this completes the proof of (R.64b).

Proof of (R.64c): From Lemma 67, the following holds:

$$\int_{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{T}_{XY}^{(N,\epsilon)}} f_{\boldsymbol{XY}}(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{x}\, d\boldsymbol{y} \geqslant 1 - \epsilon, \tag{R.66}$$

with $\epsilon > 0$ and this completes the proof of (R.64c).

Proof of (R.64d)**:** From (R.64a) and (R.64c), the following holds:

$$1 = \int_{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^N\times\mathcal{Y}^N} f_{\boldsymbol{XY}}(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{x}\, d\boldsymbol{y} \tag{R.67a}$$

$$\geqslant \int_{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{T}_{XY}^{(N,\epsilon)}} f_{\boldsymbol{XY}}(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{x}\, d\boldsymbol{y} \tag{R.67b}$$

$$> \int_{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{T}_{XY}^{(N,\epsilon)}} 2^{-N(h(X,Y)+\epsilon)}\, d\boldsymbol{x}\, d\boldsymbol{y} \tag{R.67c}$$

$$= 2^{-N(h(X,Y)+\epsilon)} \int_{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{T}_{XY}^{(N,\epsilon)}} d\boldsymbol{x}\, d\boldsymbol{y} \tag{R.67d}$$

$$= 2^{-N(h(X,Y)+\epsilon)} \mathrm{Vol}\left(\mathcal{T}_{XY}^{(N,\epsilon)}\right), \tag{R.67e}$$

for N sufficiently large, which implies:

$$\mathrm{Vol}\left(\mathcal{T}_{XY}^{(N,\epsilon)}\right) < 2^{N(h(X,Y)+\epsilon)}, \tag{R.67f}$$

and

$$1 - \epsilon \leqslant \int_{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{T}_{XY}^{(N,\epsilon)}} f_{\boldsymbol{XY}}(\boldsymbol{x},\boldsymbol{y})\, d\boldsymbol{x}\, d\boldsymbol{y} \tag{R.67g}$$

$$< \int_{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{T}_{XY}^{(N,\epsilon)}} 2^{-N(h(X,Y)-\epsilon)}\, d\boldsymbol{x}\, d\boldsymbol{y} \tag{R.67h}$$

$$= 2^{-N(h(X,Y)-\epsilon)} \int_{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{T}_{XY}^{(N,\epsilon)}} d\boldsymbol{x}\, d\boldsymbol{y} \tag{R.67i}$$

$$= 2^{-N(h(X,Y)-\epsilon)} \mathrm{Vol}\left(\mathcal{T}_{XY}^{(N,\epsilon)}\right), \tag{R.67j}$$

for N sufficiently large, which implies:

$$\mathrm{Vol}\left(\mathcal{T}_{XY}^{(N,\epsilon)}\right) > (1-\epsilon)\, 2^{N(h(X,Y)-\epsilon)}, \tag{R.67k}$$

and this completes the proof of (R.64d). This completes the proof of Lemma 68. ■

Bibliography

[1] Ahlswede, R. "Multi-way communication channels". In: *Proc. 2nd International Symposium on Information Theory.* Tsahkadsor, Armenian S.S.R. Hungarian Academy of Sciences, Budapest, Sept. 1971, pp. 23–52 (Cited in page 12).

[2] Ahlswede, R. "The capacity region of a channel with two senders and two receivers". In: *The Annals of Probability* 2.5 (Oct. 1974), pp. 805–814 (Cited in page 12).

[3] Annapureddy, V. S. and Veeravalli, V. V. "Gaussian interference networks: Sum-capacity in the low-interference regime and new outer bounds on the capacity region". In: *IEEE Trans. Inf. Theory* 55.7 (July 2009), pp. 3032–3050 (Cited in page 16).

[4] Anshelevich, E., Dasgupta, A., Tardos, E., and Wexler, T. "Near-optimal network design with selfish agents". In: *Proc. Annual ACM Symposium on Theory of Computing.* San Diego, CA, USA, June 2003 (Cited in pages l, 80).

[5] Ardestanizadeh, E., Franceschetti, M., Javidi, T., and Kim, Y.-H. "Wiretap channel with secure rate-limited feedback". In: *IEEE Trans. Inf. Theory* 55.12 (Dec. 2009), pp. 5353–5361 (Cited in page 2).

[6] Ashraphijuo, M., Aggarwal, V., and Wang, X. "On the symmetric K-user linear deterministic interference channels with limited feedback". In: *Proc. 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton).* Oct. 2014 (Cited in pages 22, 23).

[7] Ashraphijuo, M., Aggarwal, V., and Wang, X. "On the Symmetric K-user Interference Channels with Limited Feedback". In: *IEEE Trans. Inf. Theory* 62.12 (Oct. 2016), pp. 6969–6985 (Cited in pages 22, 23).

[8] Avestimehr, S., Diggavi, S., and Tse, D. N. C. "Wireless network information flow: A deterministic approach". In: *IEEE Trans. Inf. Theory* 57.4 (Apr. 2011), pp. 1872–1905 (Cited in pages 22, 25).

[9] Belhadj Amor, S. and Perlaza, S. M. "Decentralized K-user Gaussian multiple access channels". In: *Proc. International conference on NETwork Games, COntrol and OPtimization (Netgcoop 2016).* Avignon, France, Nov. 2016 (Cited in page 2).

[10] Belhadj Amor, S. and Perlaza, S. M. "Decentralized simultaneous energy and information transmission in multiple access channels". In: *Proc. 2016 Annual Conference on Information Science and Systems (CISS).* Princeton, NJ, USA, Mar. 2016 (Cited in page 2).

[11] Belhadj Amor, S., Perlaza, S. M., Krikidis, I., and Poor, H. V. "Feedback enhances simultaneous energy and information transmission in multiple access channels". In: *Proc. IEEE International Symposium on Information Theory (ISIT).* Barcelona, Spain, July 2016, pp. 1974–1978 (Cited in page 2).

[12] Belhadj Amor, S., Perlaza, S. M., Krikidis, I., and Poor, H. V. “Feedback enhances simultaneous wireless information and energy transmission in multiple access channels”. In: *IEEE Trans. Inf. Theory* 63.8 (Aug. 2017), pp. 5244–5265 (Cited in page 2).

[13] Belhadj Amor, S., Steinberg, Y., and Wigger, M. “MIMO MAC-BC duality with linear-feedback Coding Schemes”. In: *IEEE Trans. Inf. Theory* 61.11 (Nov. 2015), pp. 5976–5998 (Cited in page 2).

[14] Bergmans, P. “Random Coding Theorems for Broadcast Channels with Degraded Components”. In: *IEEE Trans. Inf. Theory* 19.2 (Mar. 1973), pp. 197–207 (Cited in pages 2, 3, 17, 19, 22, 42).

[15] Berry, R. and Tse, D. N. C. “Information theoretic games on interference channels”. In: *Proc. of the IEEE International Symposium on Information Theory (ISIT).* Toronto,Canada, July 2008 (Cited in pages xxvi, xxxii, 4, 32).

[16] Berry, R. and Tse, D. N. C. “Shannon meets Nash on the interference channel”. In: *IEEE Trans. Inf. Theory* 57.5 (May 2011), pp. 2821–2836 (Cited in pages xxvi, 2, 4, 33, 34, 71, 74, 75, 85, 153).

[17] Bhaskaran, S. R. “Gaussian broadcast channel with feedback”. In: *IEEE Trans. Inf. Theory* 54.11 (Nov. 2008), pp. 5252–5257 (Cited in page 2).

[18] Bracher, A. and Wigger, M. “Feedback and partial message side-information on the semideterministic broadcast channel”. In: *IEEE Trans. Inf. Theory* 63.8 (Aug. 2017), pp. 505–5073 (Cited in page 2).

[19] Braess, D. “Über ein Paradoxon aus der Verkehrsplanung”. In: *Unternehmensforschung* 24.5 (May 1969), pp. 258–268 (Cited in page 74).

[20] Bresler, G. and Tse, D. N. C. “The two user Gaussian interference channel: A deterministic view”. In: *European Transactions on Telecommunications* 19.4 (Apr. 2008), pp. 333–354 (Cited in pages xxv, 3, 22, 28, 36, 40, 43, 71).

[21] Bross, S. I. and Wigger, M. “On the relay channel with receiver–transmitter feedback”. In: *IEEE Trans. Inf. Theory* 55.1 (Jan. 2009), pp. 275–291 (Cited in page 2).

[22] Carleial, A. “A case where Interference does not reduce capacity”. In: *IEEE Trans. Inf. Theory* 21.5 (Sept. 1975), pp. 569–570 (Cited in pages 1, 12, 16).

[23] Carleial, A. “Interference channels”. In: *IEEE Trans. Inf. Theory* 24.1 (Sept. 1978), pp. 60–70 (Cited in pages 1–3, 12–14, 17, 19, 22, 42).

[24] Chong, H.-F., Motani, M., Garg, H. K., and El Gamal, H. “On the Han-Kobayashi region for the interference channel”. In: *IEEE Trans. Inf. Theory* 54.7 (July 2008), pp. 3188–3195 (Cited in pages 1, 13–16, 22).

[25] Cover, T. “Broadcast channels”. In: *IEEE Trans. Inf. Theory* 18.1 (Jan. 1972), pp. 2–14 (Cited in page 15).

[26] Cover, T. and Gamal, A. E. “Capacity theorems for the relay channel”. In: *IEEE Trans. Inf. Theory* 25.5 (Sept. 1979), pp. 572–584 (Cited in page 2).

[27] Cover, T. M. and Leung, C. S. K. “An achievable rate region for the multiple-access channel with feedback”. In: *IEEE Trans. Inf. Theory* 27.3 (May 1981), pp. 292–298 (Cited in pages 1–3, 14, 17, 19, 22, 42).

[28] Cover, T. M. and Thomas, J. A. *Elements of information theory*. Hoboken, NJ, USA: Wiley-Interscience, 1991 (Cited in pages 102, 103, 178, 181, 206–208).

[29] Dueck, G. "Partial feedback for two-way and broadcast channels". In: *Information and Control* 46.1 (July 1980), pp. 1–15 (Cited in page 2).

[30] El Gamal, A. and Costa, M. "The capacity region of a class of deterministic interference channels". In: *IEEE Trans. Inf. Theory* 28.2 (Mar. 1982), pp. 343–346 (Cited in pages 22, 25).

[31] Etkin, R. H., Tse, D. N. C., and Hua, W. "Gaussian interference channel capacity to within one bit". In: *IEEE Trans. Inf. Theory* 54.12 (Dec. 2008), pp. 5534–5562 (Cited in pages xxvi, xxxiii, xxxix, 1, 4, 16, 17, 22, 24, 34, 44, 53, 140).

[32] Fano, R. *Transmission of information - a statistical theory of communication*. MIT Press, Mar. 1961 (Cited in page 193).

[33] Gaarder, N. T. and Wolf, J. K. "The capacity region of a mulitple-access discrete memoryless channel can increase with feedback". In: *IEEE Trans. Inf. Theory* 21.1 (Jan. 1975), pp. 100–102 (Cited in page 2).

[34] Gabbai, Y. and Bross, S. I. "Achievable rates for the discrete memoryless relay channel with partial feedback configurations". In: *IEEE Trans. Inf. Theory* 52.11 (Nov. 2006), pp. 4989–5007 (Cited in page 2).

[35] Gastpar, M. and Kramer, G. "On noisy feedback for interference channels". In: *Asilomar Conference on Signals, Systems and Computers*. Oct. 2006 (Cited in page 19).

[36] Gastpar, M., Lapidoth, A., Steinberg, Y., and Wigger, M. "Coding schemes and asymptotic capacity for the Gaussian broadcast and interference channels with feedback". In: *IEEE Trans. Inf. Theory* 60.1 (Jan. 2014), pp. 54–71 (Cited in page 2).

[37] Han, T. S. and Kobayashi, K. "A new achievable rate region for the interference channel". In: *IEEE Trans. Inf. Theory* 27.1 (1981), pp. 49–60 (Cited in pages 1–3, 12–17, 19, 22, 42).

[38] Han, Z., Niyato, D., Saad, W., and Hjørungnes, A. *Game theory in wireless and communication networks: Theory, models, and applications*. New York, NY.: Cambridge University Press, 2012 (Cited in page 2).

[39] Hekstra, A. P. and Willems, F. M. J. "Dependence balance bounds for single-output two-way channels". In: *IEEE Trans. Inf. Theory* 35.1 (Jan. 1989), pp. 44–53 (Cited in pages 19, 21).

[40] Jafar, S. A. "Interference alignment: A new look at signal dimensions in a communication network". In: *Foundations and Trends in Communications and Information Theory* 7.1 (2010), pp. 1–134 (Cited in pages 3, 24).

[41] Karakus, C., Wang, I.-H., and Diggavi, S. "Gaussian Interference Channel with Intermittent Feedback". In: *IEEE Trans. Inf. Theory* 61.9 (Sept. 2015), pp. 4663–4699 (Cited in page 23).

[42] Khalfet, N. and Perlaza, S. M. "Simultaneous information and energy transmission in Gaussian interference channels with feedback". In: *Proc. 55th Annual Allerton Conference on Communications, Control, and Computing*. Allerton, USA., Oct. 2017 (Cited in page 2).

[43] Khalfet, N. and Perlaza, S. M. *Simultaneous information and energy transmission in the interference channel.* Tech. rep. 9102. Lyon, France: IINRIA, July 2017 (Cited in pages 2, 15).

[44] Kobayashi, K. and Han, T. S. "A further consideration on the HK and the CMG regions for the interference channel". In: *Information Theory and Applications Workshop (ITA).* San Diego, CA, USA, Feb. 2007 (Cited in page 15).

[45] Koutsoupias, E. and Papadimitriou, C. "Worst-case equilibria". In: *Proc. 16th Annual Symposium on Theoretical Aspects of Computer Science.* Trier, Germany, Mar. 1999 (Cited in pages xlviii, 76).

[46] Kramer, G. "Outer bounds on the capacity of Gaussian interference channels". In: *IEEE Trans. Inf. Theory* 50.3 (Mar. 2004), pp. 581–586 (Cited in page 16).

[47] Kramer, G. "Correction to "Feedback strategies for white Gaussian interference networks"". In: *IEEE Trans. Inf. Theory* 50.6 (June 2004), pp. 1373–1374 (Cited in page 19).

[48] Kramer, G. "Feedback strategies for white Gaussian interference networks". In: *IEEE Trans. Inf. Theory* 48.6 (June 2002), pp. 1423–1438 (Cited in page 19).

[49] Kramer, G., Marić, I., and Yates, R. D. *Cooperative communications.* Now Publishers, 2007 (Cited in page 1).

[50] Lapidoth, A. and Wigger, M. "On the AWGN MAC with imperfect feedback". In: *IEEE Trans. Inf. Theory* 56.11 (Nov. 2010), pp. 5432–5476 (Cited in page 2).

[51] Lasaulce, S. and Tembine, H. *Game theory and learning in wireless networks: Fundamentals and applications.* Waltham, MA, USA: Elsevier Academic Press, 2011 (Cited in page 2).

[52] Le, S.-Q., Tandon, R., Motani, M., and Poor, H. V. "The Capacity Region of the Symmetric Linear Deterministic Interference Channel with Partial Feedback". In: *Proc. 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton).* Oct. 2012 (Cited in page 19).

[53] Le, S.-Q., Tandon, R., Motani, M., and Poor, H. V. "Approximate capacity region for the symmetric Gaussian interference channel with noisy feedback". In: *IEEE Trans. Inf. Theory* 61.7 (July 2015), pp. 3737–3762 (Cited in pages xxv, xxvi, xxxix, 2–4, 19, 20, 22, 28, 30, 40, 43, 52, 140).

[54] Mackay, D. J. C. *Information theory, inference, and learning algorithms.* Cambridge, UK: Cambridge University Press, 2003 (Cited in pages 221, 222).

[55] McEliece, R. *The theory of information and coding.* Cambridge, UK: Cambridge University Press, 2004 (Cited in pages 193, 208, 209).

[56] Mertikopoulos, P., Belmega, E. V., Moustakas, A. L., and Lasaulce, S. "Distributed learning policies for power allocation in multiple access channels". In: *IEEE Journal on Selected Areas in Communications* 30.1 (Jan. 2012), pp. 96–106 (Cited in page 2).

[57] Mohajer, S., Tandon, R., and Poor, H. V. "On the feedback capacity of the fully connected-user interference channel". In: *IEEE Trans. Inf. Theory* 59.5 (May 2013), pp. 2863–2881 (Cited in pages 3, 18, 22).

[58] Motahari, A. S. and Khandani, A. K. “Capacity bounds for the Gaussian interference channel”. In: *IEEE Trans. Inf. Theory* 55.2 (Feb. 2009), pp. 620–643 (Cited in page 16).

[59] Nash, J. F. “Equilibrium points in n-person games”. In: *Proc. National Academy of Sciences of the United States of America* 36.1 (Jan. 1950), pp. 48–49 (Cited in pages xxxii, 33).

[60] Nisan, N., Roughgarden, T., Tardos, É., and Vazirani, V. V. *Algorithmic game theory.* Cambridge, UK: Cambridge University Press, 2007 (Cited in page 33).

[61] Ozarow, L. H. “The capacity of the white Gaussian multiple access channel with feedback”. In: *IEEE Trans. Inf. Theory* 30.4 (July 1984), pp. 623–629 (Cited in page 2).

[62] Ozarow, L. H. and Leung-Yan-Cheong, S. “An achievable region and outer bound for the Gaussian broadcast channel with feedback”. In: *IEEE Trans. Inf. Theory* IT-30.4 (July 1984), pp. 667–671 (Cited in page 2).

[63] Pabhakaran, V. M. and Viswanath, P. “Interference channel with source cooperation”. In: *IEEE Trans. Inf. Theory* 57.1 (Jan. 2011), pp. 156–186 (Cited in pages 19, 22, 29, 30).

[64] Perlaza, S. M., Lasaulce, S., and Debbah, M. “Equilibria of channel selection games in parallel multiple access ahannels”. In: *EURASIP Journal in Wireless Communications Networks* 2013.15 (Jan. 2013), pp. 1–23 (Cited in page 2).

[65] Perlaza, S. M., Tandon, R., Poor, H. V., and Han, Z. “The Nash equilibrium region of the linear deterministic interference channel with feedback”. In: *Proc. 50th Annual Allerton Conference on Communications, Control, and Computing.* Monticello, IL, Oct. 2012 (Cited in page 2).

[66] Perlaza, S. M., Tandon, R., Poor, H. V., and Han, Z. “Perfect output feedback in the two-user decentralized interference channel”. In: *IEEE Trans. Inf. Theory* 61.10 (Oct. 2015), pp. 5441–5462 (Cited in pages xxvi, 2–4, 24, 34, 35, 71, 74, 75, 86, 153).

[67] Perlaza, S. M., Tandon, R., and Poor, H. V. “Decentralized interference channels with noisy feedback possess Pareto optimal Nash equilibria”. In: *Proc. of the 6th International Symposium on Communications, Control, and Signal Processing (ISCCSP 2014).* Athens, Greece, May 2014 (Cited in pages 2, 75).

[68] Perlaza, S. M., Tandon, R., and Poor, H. V. “Symmetric decentralized interference channels with noisy feedback”. In: *Proc. IEEE Intl. Symposium on Information Theory (ISIT).* Honolulu, HI, USA, June 2014 (Cited in pages xxvi, 2, 4, 19, 35, 71, 74, 75).

[69] Quintero, V., Perlaza, S. M., Esnaola, I., and Gorce, J.-M. “Approximate capacity of the Gaussian interference channel with noisy channel-output feedback”. In: *Proc. IEEE Information Theory Workshop (ITW).* Cambridge, UK, Sept. 2016 (Cited in pages xxv, 3).

[70] Quintero, V., Perlaza, S. M., Esnaola, I., and Gorce, J.-M. “Approximate capacity region of the two-user Gaussian interference channel with noisy channel-output feedback”. In: (Submitted to) *IEEE Trans. Inf. Theory* (Nov. 2016) (Cited in pages xxv, 3).

[71] Quintero, V., Perlaza, S. M., Esnaola, I., and Gorce, J.-M. “When does output feedback enlarge the capacity of the interference channel?” In: *IEEE Trans. Commun.* 66.2 (Feb. 2018), pp. 615–628 (Cited in pages xxvi, 4, 24).

[72] Quintero, V., Perlaza, S. M., and Gorce, J.-M. “Noisy channel-output feedback capacity of the linear deterministic interference channel”. In: *IEEE Information Theory Workshop.* Jeju Island, Korea, Oct. 2015 (Cited in pages xxv, 3).

[73] Quintero, V., Perlaza, S. M., Gorce, J.-M., and Poor, H. V. *Decentralized interference channels with noisy output feedback.* Tech. rep. 9011. Lyon, France: INRIA, Jan. 2017 (Cited in pages xxvi, 4).

[74] Quintero, V., Perlaza, S. M., Gorce, J.-M., and Poor, H. V. “Nash region of the linear deterministic interference channel with noisy output feedback”. In: *Proc. IEEE International Symposium on Information Theory (ISIT).* Aachen, Germany, June 2017 (Cited in pages xxvi, 4, 19).

[75] Ram, Z. “Lattice Coding for Signals and Networks”. In: University Printing House, Cambridge CB2 8BS, UK: Cambridge University Press, 2014 (Cited in page 22).

[76] Rose, L., Perlaza, S. M., and Debbah, M. “On the Nash equilibria in decentralized parallel interference channels”. In: *Proc. IEEE Intl. Conference on Communications (ICC).* Kyoto, Japan, June 2011 (Cited in page 2).

[77] Royden, H. and Fitzpatrick, P. *Real analysis.* Boston, MA, USA: Prentice Hall, 2014 (Cited in pages 224, 227, 232, 235).

[78] Saha, S. and Berry, R. “On information theoretic games for interference networks”. In: *Proc. 44th Asilomar Conference on Signals, Systems and Computers.* Pacific Grove, CA, USA, Nov. 2010 (Cited in page 2).

[79] Sahai, A., Aggarwal, V., Yüksel, M., and Sabharwal, A. “On Channel Output Feedback in Deterministic Interference Channels”. In: *Proc. IEEE Information Theory Workshop.* Taormina, Italy, Oct. 2009 (Cited in page 24).

[80] Sahai, A., Aggarwal, V., Yuksel, M., and Sabharwal, A. “Capacity of all nine models of channel output feedback for the two-user interference channel”. In: *IEEE Trans. Inf. Theory* 59.11 (2013), pp. 6957–6979 (Cited in pages xxv, 3, 19, 24, 29, 40).

[81] Sato, H. “The capacity of the Gaussian interference channel under strong interference”. In: *IEEE Trans. Inf. Theory* 27.6 (Nov. 1981), pp. 786–788 (Cited in pages 1, 12, 16).

[82] Sato, H. “Two-user communication channels”. In: *IEEE Trans. Inf. Theory* 23.3 (1977), pp. 295–304 (Cited in page 16).

[83] Shang, X., Kramer, G., and Chen, B. “A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels”. In: *IEEE Trans. Inf. Theory* 55.2 (Feb. 2009), pp. 689–699 (Cited in page 16).

[84] Shannon, C. E. “The zero-error capacity of a noisy channel”. In: *IRE Transactions on Information Theory* 2.3 (Sept. 1956), pp. 8–19 (Cited in page 111).

[85] Shannon, C. E. “A mathematical theory of communication”. In: *The Bell System Technical Journal* 27.3 (July 1948), pp. 379–423 (Cited in page 193).

[86] Shannon, C. E. “Channels with side information at the transmitter”. In: *IBM Journal of Research and Development* 2.4 (Oct. 1958), pp. 289–293 (Cited in pages xxxvi, 47).

[87] Shayevitz, O. and Wigger, M. “On the capacity of the discrete memoryless broadcast channel with feedback”. In: *IEEE Trans. Inf. Theory* To appear (2013) (Cited in page 2).

[88] Suh, C. and Tse, D. N. C. "Feedback capacity of the Gaussian interference channel to within 2 bits". In: *IEEE Trans. Inf. Theory* 57.5 (May 2011), pp. 2667–2685 (Cited in pages xxv, xxvi, xxxix, 2–4, 17, 18, 22, 24, 28, 40, 43, 53, 60, 71, 111, 131, 140).

[89] Tandon, R. and Ulukus, S. "Dependence balance based outer bounds for Gaussian networks with cooperation and feedback". In: *IEEE Trans. Inf. Theory* 57.7 (July 2011), pp. 4063–4086 (Cited in page 22).

[90] Tandon, R., Mohajer, S., and Poor, H. V. "On the symmetric feedback capacity of the K-user cyclic Z-interference channel". In: *IEEE Trans. Inf. Theory* 59.5 (May 2013), pp. 2713–2734 (Cited in pages 3, 18, 22).

[91] Thomas, J. A. "Feedback can at most double Gaussian multiple access channel capacity". In: *IEEE Trans. Inf. Theory* 33.5 (Sept. 1987), pp. 711–716 (Cited in page 2).

[92] Tuninetti, D. "An outer bound for the memoryless two-user interference channel with general cooperation". In: *IEEE Information Theory Workshop (ITW)*. Sept. 2012, pp. 217–221 (Cited in page 19).

[93] Tuninetti, D. "An outer bound region for interference channels with generalized feedback". In: *IEEE Information Theory and Applications Workshop (ITA)*. Feb. 2010, pp. 1–5 (Cited in page 19).

[94] Tuninetti, D. "On interference channel with generalized feedback (IFC-GF)". In: *Proc. of International Symposium on Information Theory (ISIT)*. Nice, France, June 2007, pp. 2661–2665 (Cited in pages xxv, 2, 3, 19, 43).

[95] Vahid, A., Suh, C., and Avestimehr, A. S. "Interference channels with rate-limited feedback". In: *IEEE Trans. Inf. Theory* 58.5 (May 2012), pp. 2788–2812 (Cited in page 22).

[96] Venkataramanan, R. and Pradhan, S. S. "An achievable rate region for the broadcast channel with feedback". In: *IEEE Trans. Inf. Theory* 59.10 (June 2013), pp. 6175–6191 (Cited in page 2).

[97] Willems, F. "The feedback capacity region of a class of discrete memoryless multiple access channels". In: *IEEE Trans. Inf. Theory* 28.1 (Jan. 1982), pp. 93–95 (Cited in page 2).

[98] Willems, F. M. J. "Information theoretical results for multiple access channels". PhD thesis. Leuven, Belgium: Katholieke Universiteit, Oct. 1982 (Cited in pages 2, 3, 17, 19, 22, 42).

[99] Willems, F. M. J. and Van Der Meulen, E. C. "The discrete memoryless multiple-access channel with cribbing encoders". In: *IEEE Trans. Inf. Theory* IT-31.3 (May 1985), pp. 313–327 (Cited in pages 2, 3, 17, 19, 22, 42).

[100] Wu, Y. and Wigger, M. "Any positive feedback rate increases the capacity of strictly less-noisy broadcast channels". In: *Information Theory Workshop*. IEEE, Sept. 2013, pp. 1–5 (Cited in page 2).

[101] Wu, Y. and Wigger, M. "Coding schemes with rate-limited feedback that improve over the no feedback capacity for a large class of broadcast channels". In: *IEEE Trans. Inf. Theory* 62.4 (Apr. 2016), pp. 2009–2033 (Cited in page 2).

[102] Yang, S. and Tuninetti, D. “Interference channel with generalized feedback (a.k.a. with source cooperation): Part I: Achievable Region”. In: *IEEE Trans. Inf. Theory* 5.57 (May 2011), pp. 2686–2710 (Cited in pages xxv, 2, 3, 19).

[103] Yates, R. D., Tse, D., and Li, Z. “Secret communication on interference channels”. In: *Proc. of the IEEE International Symposium on Information Theory (ISIT)*. Toronto, Canada, July 2008 (Cited in pages xxxii, 32).

[104] Yeung, R. W. *Information theory and network coding*. Springer, 2008 (Cited in pages 211, 222, 231, 233).

[105] Zhou, L. and Yu, W. “On the capacity of the K-user cyclic Gaussian interference channel”. In: *IEEE Trans. Inf. Theory* 59.1 (Jan. 2013), pp. 154–165 (Cited in page 19).