Noisy Channel-Output Feedback Capacity of the Linear Deterministic Interference Channel

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

Noisy Channel-Output Feedback Capacity of the Linear Deterministic Interference Channel

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

TECHNICAL REPORT
N° 456
January 2015
Project-Team Socrate

ISSN 0249-0803 ISRN INRIA/RT-456--FR+ENG

Noisy Channel-Output Feedback Capacity of the Linear Deterministic Interference Channel

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

Project-Team Socrate

Technical Report n° 456 — January 2015 — 41 pages

Abstract: In this technical report, the capacity region of the two-user linear deterministic (LD) interference channel with noisy output feedback (IC-NOF) is fully characterized. This result allows the identification of several asymmetric scenarios in which implementing channel-output feedback in only one of the transmitter-receiver pairs is as beneficial as implementing it in both links, in terms of achievable individual rate and sum-rate improvements w.r.t. the case without feedback. In other scenarios, the use of channel-output feedback in any of the transmitter-receiver pairs benefits only one of the two pairs in terms of achievable individual rate improvements or simply, it turns out to be useless, i.e., the capacity regions with and without feedback turn out to be identical even in the full absence of noise in the feedback links.

Key-words: Capacity, Linear Deterministic Interference Channel, Noisy Output-Feedback

The work of Victor Quintero is supported in part by Universidad del Cauca, Popayán, Colombia and the Administrative Department of Science, Technology and Innovation of Colombia (Colciencias), fellowship number 617-2013

RESEARCH CENTRE GRENOBLE – RHÔNE-ALPES

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

Capacité du Canal Linéaire Déterministe à Interférences avec Rétroalimentation Degradée par Bruit Additif.

Résumé : Dans ce rapport, la région de capacité du canal linéaire déterministe à interférences avec rétroalimentation degradée entre les récepteurs et leurs émetteurs correspondants est caractérisée. Ce résultat permet l'identification de plusieurs scenarios asymétriques dans lesquels la rétroalimentation dans un seul couple récepteur-émetteur montre autant de bénéfices que des rétroalimentations dans les deux couples récepteurs-émetteurs. Ces bénéfices sont mis en évidence par l'amélioration des taux de transmission individuels et de leur somme par rapport aux cas où il n'y a aucune rétroalimentation. D'autres scenarios montrent qu'une rétroalimentation dans un des couple émetteur-récepteur améliore le taux individuel d'un des deux couples émetteurs-récepteurs. D'ailleurs, il existe d'autres scenarios où l'utilisation d'un ou plusieurs liens de rétroalimentation ne montre aucun bénéfice ni pour les taux individuels ni pour leur somme. Dans ces scenarios, cela montre que les régions de capacité avec et sans rétroalimentation sont identiques.

Mots-clés : Région de Capacité, Modèle linéaire déterministe, canal à interférences, rétroalimentation degradée.

Contents

1 Introduction

Perfect channel-output feedback (POF) has been shown to dramatically enlarge the capacity region of the two-user interference channel (IC) [1, 2, 3]. The same observation has been made in other types of ICs with larger number of transmitter-receiver pairs [4]. In general, when a transmitter observes the channel-output at its intended receiver, it obtains a noisy version of the sum of its own transmitted signal and the interfering signals from other transmitters. This implies that, subject to a finite delay, transmitters know at least partially the information transmitted by other transmitters in the network. Hence, channel-output feedback allows using the interference as side-information. Unfortunately, the benefits of feedback are less well understood in a more realistic case in which the channel-output feedback links are impaired by additive noise. The capacity region of the LD-IC with noisy channel-output feedback (NOF) is known only in the two user-symmetric case, see [5].

In this technical report, the capacity region of the two-user LD-IC-NOF is fully characterized. This result allows the exact identification of the asymmetric interference regimes in which the capacity region of the LD-IC is enlarged thanks to the action of channel-output feedback. At the same time, it reveals that there exist configurations in which channel-output feedback is absolutely useless in terms of capacity region improvement. The achievability scheme, which is optimized for the IC-NOF using a three-part message splitting, superposition coding and backward decoding, as first suggested by other authors in [1, 3, 6], and the outer bound region (converse) are essential part of this technical report and are included in the annexes.

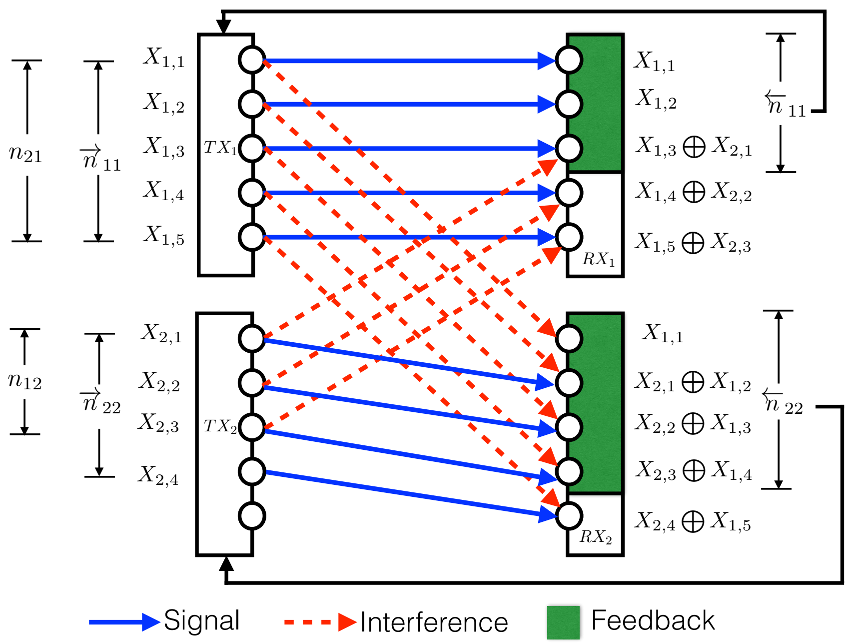

Figure 1: Two-user linear deterministic interference channel with noisy channel-output feedback (LD-IC-NOF).

2 Linear Deterministic Interference Channel with Noisy-Channel Output Feedback

Consider the two-user LD-IC-NOF, with parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$ described in Fig. 1. $\overrightarrow{n}_{ii}$, $i \in \{1,2\}$, is a non-negative integer used to represent the signal-noise ratio (SNR) in receiver i; n_{ij}, $i \in \{1,2\}$ and $j \in \{1,2\} \setminus \{i\}$, is a non-negative integer used to represent the interference-noise ratio (INR) in receiver i from transmitter j; and $\overleftarrow{n}_{ii}$, $i \in \{1,2\}$, is a non-negative integer used to represent the signal-noise ratio (SNR) in transmittter i in the feedback link from receiver i. At transmitter i, with $i \in \{1,2\}$, the channel-input $\boldsymbol{X}_i^{(n)}$ at channel use n, with $n \in \{1,\ldots,N\}$, is a q-dimensional binary vector $\boldsymbol{X}_i^{(n)} = \left(X_{i,1}^{(n)},\ldots,X_{i,q}^{(n)}\right)^{\mathsf{T}}$, with

$$q = \max\left(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}\right) \tag{1}$$

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

$$\overrightarrow{\boldsymbol{Y}}_i^{(n)}=\boldsymbol{S}^{q-\overrightarrow{n}_{ii}}\boldsymbol{X}_i^{(n)} + \boldsymbol{S}^{q-n_{ij}}\boldsymbol{X}_j^{(n)}, \tag{2}$$

and the feedback signal available at transmitter i at the end of channel use n is

$$\overleftarrow{\boldsymbol{Y}}_i^{(n)}=\boldsymbol{S}^{(q-\overleftarrow{n}_{ii})^+}\overrightarrow{\boldsymbol{Y}}_i^{(n-d)}, \tag{3}$$

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

$$\boldsymbol{S} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}. \tag{4}$$

Transmitter i sends M_i information bits $b_{i,1},\ldots,b_{i,M_i}$ by sending the codeword $\left(\boldsymbol{X}_i^{(1)},\ldots,\boldsymbol{X}_i^{(N)}\right)$. The encoder of transmitter i can be modeled as a set of deterministic mappings $f_i^{(1)},\ldots,f_i^{(N)}$, with $f_i^{(1)} : \{0,1\}^{M_i} \rightarrow \{0,1\}^q$ and $\forall n \in \{2,\ldots,N\}$, $f_i^{(n)} : \{0,1\}^{M_i} \times \{0,1\}^{q(n-1)} \rightarrow \{0,1\}^q$, such that

$$\boldsymbol{X}_i^{(1)}=f_i^{(1)}\left(b_{i,1},\ldots,b_{i,M_i}\right) \text{ and } \forall t \in \{2,\ldots,N\}, \tag{5}$$

$$\boldsymbol{X}_i^{(n)}=f_i^{(n)}\left(b_{i,1},\ldots,b_{i,M_i},\overleftarrow{\boldsymbol{Y}}_i^{(1)},\ldots,\overleftarrow{\boldsymbol{Y}}_i^{(n-1)}\right). \tag{6}$$

At the end of the block, receiver i uses the sequence $\boldsymbol{Y}_i^{(1)},\ldots,\boldsymbol{Y}_i^{(N)}$ to generate the estimates $\hat{b}_{i,1},\ldots,\hat{b}_{i,M_i}$. The average bit error probability at receiver i, denoted by p_i, is calculated as follows

$$p_i = \frac{1}{M_i}\sum_{\ell=1}^{M_i} \mathbb{1}_{\left\{\hat{b}_{i,\ell} \neq b_{i,\ell}\right\}}. \tag{7}$$

A rate pair $(R_1,R_2) \in \mathbb{R}_+^2$ is said to be achievable if it satisfies the following definition.

Definition 1 (Achievable Rate Pairs) *The rate pair* $(R_1, R_2) \in \mathbb{R}_+^2$ *is achievable if there exists at least one pair of codebooks* $\mathcal{X}_1$ *and* $\mathcal{X}_2$ *with codewords of length* N_1 *and* N_2*, respectively, with the corresponding encoding functions* $f_1^{(1)}, \ldots, f_1^{(N_1)}$ *and* $f_2^{(1)}, \ldots, f_2^{(N_2)}$ *such that the average bit error probability can be made arbitrarily small by letting the block lengths* N_1 *and* N_2 *grow to infinity.*

The following section determines the set of all the rate pairs (R_1, R_2) that are achievable in the LD-IC-NOF with parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$.

3 Main Results

Denote by $\mathcal{C}(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ the capacity region of the LD-IC-NOF with parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$. Theorem 1 fully characterizes the capacity region $\mathcal{C}(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$.

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

$$\begin{aligned}
R_i &\leqslant \min\left(\max\left(\overrightarrow{n}_{ii}, n_{ji}\right), \max\left(\overrightarrow{n}_{ii}, n_{ij}\right)\right), &(8)\\
R_i &\leqslant \min\left(\left(n_{ji} - \overrightarrow{n}_{ii}\right)^+, \left(\overleftarrow{n}_{jj} - \min\left(\overrightarrow{n}_{ii}, n_{ji}\right) - \min\left(\left(\overrightarrow{n}_{jj} - n_{ji}\right)^+, n_{ij}\right)\right)^+\right)\\
&\quad + \overrightarrow{n}_{ii}, &(9)\\
R_1 + R_2 &\leqslant \min\left(\max\left(\overrightarrow{n}_{11}, n_{12}\right) + \left(\overrightarrow{n}_{22} - n_{12}\right)^+, \max\left(\overrightarrow{n}_{22}, n_{21}\right) + \left(\overrightarrow{n}_{11} - n_{21}\right)^+\right), &(10)\\
R_1 + R_2 &\leqslant \left(\overrightarrow{n}_{11} - n_{21}\right)^+ + \left(\overrightarrow{n}_{22} - n_{12}\right)^+ + \left(\overleftarrow{n}_{11} - \left(\overrightarrow{n}_{11} - n_{12}\right)^+\right)^+\\
&\quad + \left(\overleftarrow{n}_{22} - \left(\overrightarrow{n}_{22} - n_{21}\right)^+\right)^+, &(11)\\
2R_i + R_j &\leqslant \max\left(\overrightarrow{n}_{ii}, n_{ij}\right) + \left(\overrightarrow{n}_{ii} - n_{ji}\right)^+ + \left(\overrightarrow{n}_{jj} - n_{ij}\right)^+ + \left(\overleftarrow{n}_{jj} - \left(\overrightarrow{n}_{jj} - n_{ji}\right)^+\right)^+. &(12)
\end{aligned}$$

3.1 Proofs

The proof of Theorem 1 is presented in two parts: the achievability scheme is formally described in Appendix A and an outer bound region (converse) is described in Appendix B.

3.2 Discussion

This section provides a set of examples in which particular scenarios are highlighted to show that channel-output feedback can be strongly beneficial for enlarging the capacity region of the two-user LD-IC. At the same time, it also highlights other examples in which channel-output feedback does not bring any benefit in terms of the capacity region. These benefits are given in terms of the following metrics: (a) individual rate improvements Δ_1 and Δ_2; and (b) sum-rate improvement Σ. In all cases, these improvements are measured with respect to the case without feedback.

In order to formally define Δ_1, Δ_2 and Σ, consider an LD-IC-NOF with parameters $\overrightarrow{n}_{11}$, $\overrightarrow{n}_{22}$, n_{12}, n_{21}, $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$. The maximum improvement $\Delta_i(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ of the individual rate R_i due to the effect of channel-output feedback with respect to the case without feedback is

$$\Delta_i(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}) = \max_{R_j > 0} \sup_{\substack{(R_i, R_j) \in \mathcal{C}_1 \\ (R_i^\dagger, R_j) \in \mathcal{C}_2}} R_i - R_i^\dagger, \tag{13}$$

and the maximum sum rate improvement $\Sigma(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ with respect to the case without feedback is

$$\Sigma(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22}) = \sup_{\substack{(R_1, R_2) \in \mathcal{C}_1 \\ (R_1^\dagger, R_2^\dagger) \in \mathcal{C}_2}} R_1 + R_2 - (R_1^\dagger + R_2^\dagger), \tag{14}$$

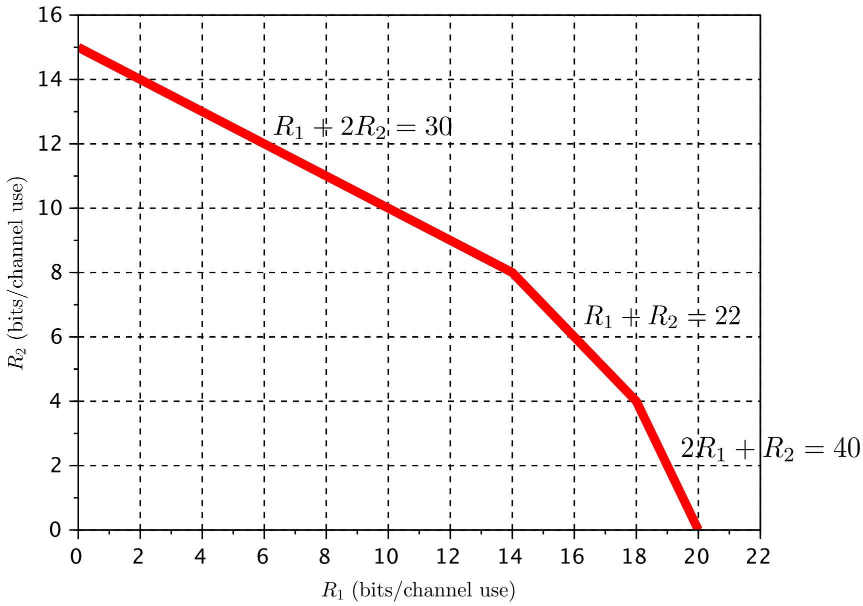

Figure 2: Capacity region $\mathcal{C}(20, 15, 12, 13, 0, 0)$ of the example in Sec. 3.2.1 without feedback

where $\mathcal{C}_1 = \mathcal{C}(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ and $\mathcal{C}_2 = \mathcal{C}(\overrightarrow{n}_{11}, \overrightarrow{n}_{22}, n_{12}, n_{21}, 0, 0)$ are the capacity region with noisy channel-output feedback and without feedback, respectively. The following describes particular scenarios that highlight some interesting observations.

3.2.1 Example 1: only one channel-output feedback link allows simultaneous maximum improvement of both individual rates

Consider the case in which transmitter-receiver pairs 1 and 2 are in weak and moderate interference regimes, with $\overrightarrow{n}_{11} = 20$, $\overrightarrow{n}_{22} = 15$, $n_{12} = 12$, $n_{21} = 13$. In Fig. 2, Fig. 3 and Fig. 4 the capacity region is plotted without channel-output feedback, without channel-output feedback and with noisy channel-output feedback, and without channel-output feedback and perfect channel-output feedback respectively. In Fig. 5, $\Delta_i(20, 15, 12, 13, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ is plotted for both $i = 1$ and $i = 2$ as a function of $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$. Therein, it is shown that: (a) Increasing the parameter $\overleftarrow{n}_{11}$, beyond a given threshold $\overleftarrow{n}^*_{11}$, allows simultaneous improvement of both individual rates. Note that in the case of perfect channel-output feedback, i.e., $\overleftarrow{n}_{11} = \max(\overrightarrow{n}_{11}, n_{12})$, the maximum improvement of both individual rates is simultaneously achieved. (b) Increasing the parameter $\overleftarrow{n}_{22}$, beyond a given threshold $\overleftarrow{n}^*_{22}$, also provides simultaneous improvement of both individual rates, however, even in the case of perfect channel-output feedback, i.e., $\overleftarrow{n}_{22} = \max(\overrightarrow{n}_{22}, n_{21})$, it does not achieve the maximum improvement, for any value of $\overrightarrow{n}_{11}$. (c) Finally, the sum rate does not increase by using channel-output feedback.

3.2.2 Example 2: only one channel-output feedback link allows maximum improvement of one individual rate and the sum-rate

Consider the case in which transmitter-receiver pairs 1 and 2 are in very weak and moderate interference regimes, with $\overrightarrow{n}_{11} = 10$, $\overrightarrow{n}_{22} = 10$, $n_{12} = 3$, $n_{21} = 8$. In Fig. 6, Fig. 7 and Fig. 8 the

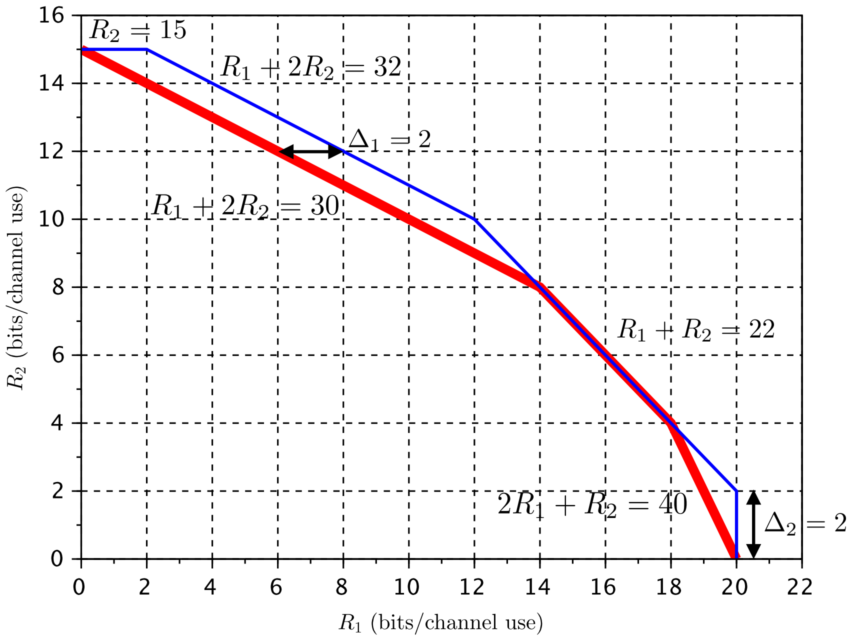

Figure 3: Capacity region $\mathcal{C}(20, 15, 12, 13, 0, 0)$ without feedback (thick red line) and $\mathcal{C}(20, 15, 12, 13, 15, 14)$ with noisy channel-output feedback (thin blue line) of the example in Sec. 3.2.1. Note that $\Delta_1(20, 15, 12, 13, 15, 14) = 2$ bits/ch.use, $\Delta_2(20, 15, 12, 13, 15, 14) = 2$ bits/ch.use and $\Sigma(20, 15, 12, 13, 15, 14) = 0$ bits/ch.use.

capacity region is plotted without channel-output feedback, without channel-output feedback and with noisy channel-output feedback, and without channel-output feedback and perfect channel-output feedback respectively. In Fig. 9, $\Delta_i(10, 10, 3, 8, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ is plotted for both $i = 1$ and $i = 2$ as a function of $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$. Therein, it is shown that: (a) For all $i \in \{1, 2\}$, increasing $\overleftarrow{n}_{ii}$, beyond a given threshold $\overleftarrow{n}_{ii}^*$, allows simultaneous improvement of both individual rates. Nonetheless, maximum improvement is achieved only for R_i. (b) Increasing either $\overleftarrow{n}_{11}$ or $\overleftarrow{n}_{22}$, beyond the thresholds $\overleftarrow{n}_{11}^*$ and $\overleftarrow{n}_{22}^*$, allows maximum improvement of the sum rate (see Fig. 9).

3.2.3 Example 3: at least one channel-output feedback link does not have any effect over the capacity region

Consider the case in which transmitter-receiver pairs 1 and 2 are in the weak interference regime, with $\overrightarrow{n}_{11} = 10$, $\overrightarrow{n}_{22} = 20$, $n_{12} = 6$, $n_{21} = 12$. In Fig. 10, Fig. 11 and Fig. 12 the capacity region is plotted without channel-output feedback, without channel-output feedback and with noisy channel-output feedback, and without channel-output feedback and perfect channel-output feedback respectively. In Fig. 13, $\Delta_i(10, 20, 6, 12, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ is plotted for both $i = 1$ and $i = 2$ as a function of $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$. Therein, it is shown that: (a) Increasing the parameter $\overleftarrow{n}_{11}$ does not enlarge the capacity region, independently of the value of $\overleftarrow{n}_{22}$. (b) Increasing the parameter $\overleftarrow{n}_{22}$, beyond a threshold $\overleftarrow{n}_{22}^*$, allows simultaneous improvement of both individual rates. (c) Finally, none of the parameters $\overleftarrow{n}_{11}$ or $\overleftarrow{n}_{22}$ increases the sum-rate.

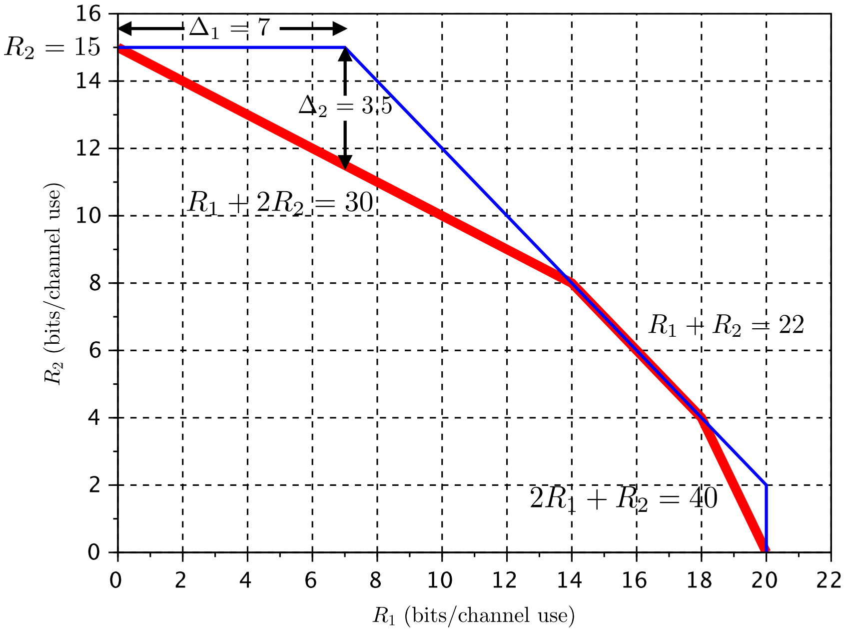

Figure 4: Capacity region $\mathcal{C}(20, 15, 12, 13, 0, 0)$ without feedback (thick red line) and $\mathcal{C}(20, 15, 12, 13, 20, 15)$ with perfect channel-output feedback (thin blue line) of the example in Sec. 3.2.1. Note that $\Delta_1(20, 15, 12, 13, 20, 15) = 7$ bits/ch.use, $\Delta_2(20, 15, 12, 13, 20, 15) = 3.5$ bits/ch.use and $\Sigma(20, 15, 12, 13, 20, 15) = 0$ bits/ch.use.

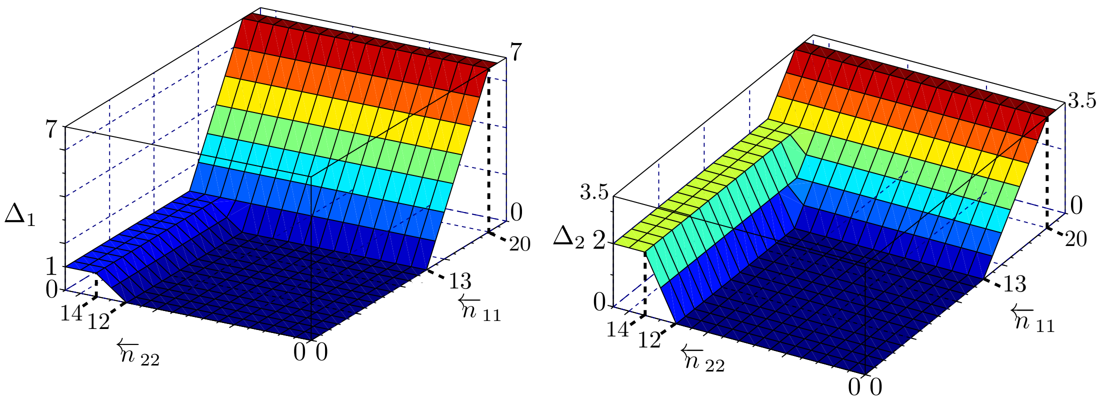

Figure 5: Maximum improvements $\Delta_1(20, 15, 12, 13, \cdot, \cdot)$ and $\Delta_2(20, 15, 12, 13, \cdot, \cdot)$ of individual rates of the example in Sec. 3.2.1

3.2.4 Example 4: the channel-output feedback of link i exclusively improves R_j

Consider the case in which transmitter-receiver pairs 1 and 2 are in the strong interference regime, with $\overrightarrow{n}_{11} = 7$, $\overrightarrow{n}_{22} = 8$, $n_{12} = 15$, $n_{21} = 13$. In Fig. 14, Fig. 15 and Fig. 16 the capacity region is plotted without channel-output feedback, without channel-output feedback and with

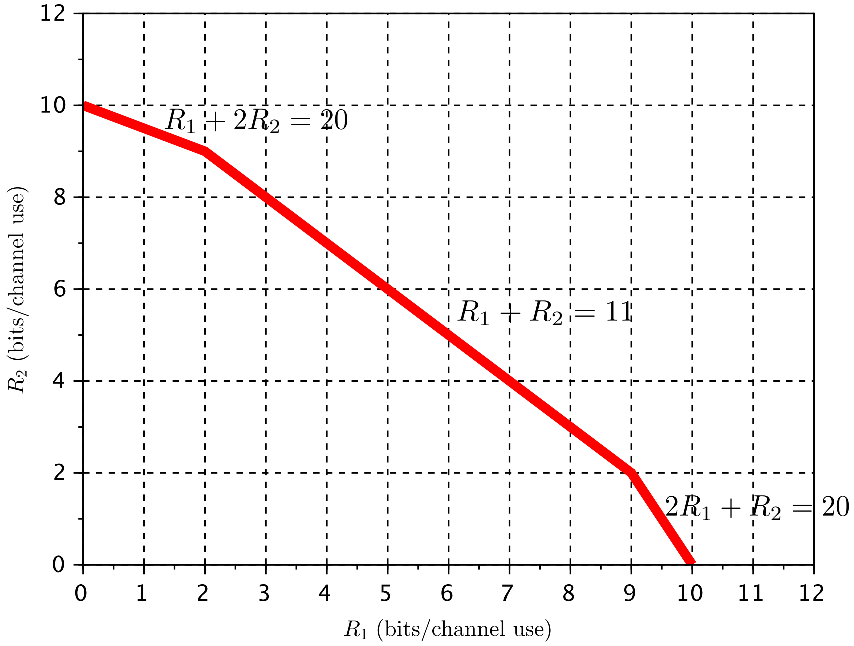

Figure 6: Capacity region of the example in Sec. 3.2.2 without feedback $\mathcal{C}(10, 10, 3, 8, 0, 0)$

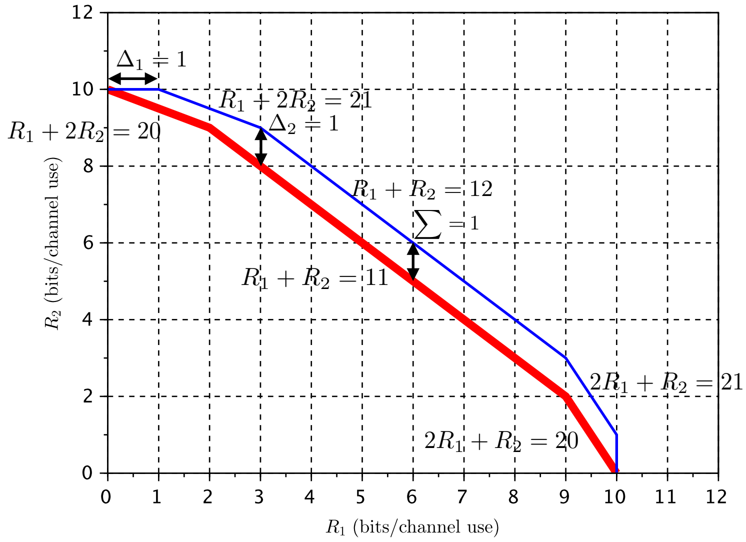

Figure 7: Capacity region $\mathcal{C}(10, 10, 3, 8, 0, 0)$ without feedback (thick red line) and $\mathcal{C}(10, 10, 3, 8, 9, 4)$ with noisy channel-output feedback (thin blue line) of the example in Sec. 3.2.2. Note that $\Delta_1(10, 10, 3, 8, 9, 4) = 1$ bit/ch.use, $\Delta_2(10, 10, 3, 8, 9, 4) = 1$ bit/ch.use and $\Sigma(10, 10, 3, 8, 9, 4) = 1$ bit/ch.use.

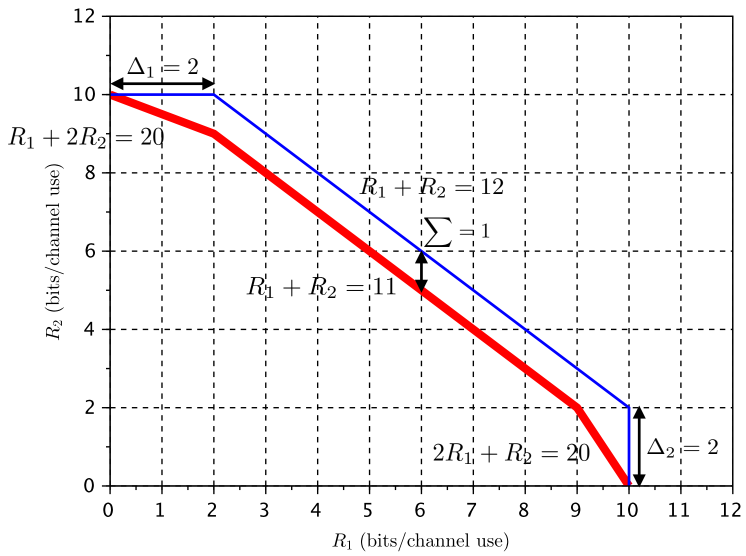

Figure 8: Capacity region $\mathcal{C}(10, 10, 3, 8, 0, 0)$ without feedback (thick red line) and $\mathcal{C}(10, 10, 3, 8, 10, 10)$ with perfect channel-output feedback (thin blue line) of the example in Sec. 3.2.2. Note that $\Delta_1(10, 10, 3, 8, 10, 10) = 2$ bits/ch.use, $\Delta_2(10, 10, 3, 8, 10, 10) = 2$ bits/ch.use and $\Sigma(10, 10, 3, 8, 10, 10) = 1$ bit/ch.use.

noisy channel-output feedback, and without channel-output feedback and perfect channel-output feedback respectively. In Fig. 17, $\Delta_i(7, 8, 15, 13, \overleftarrow{n}_{11}, \overleftarrow{n}_{22})$ is plotted for both $i = 1$ and $i = 2$ as a function of $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$. Therein, it is shown that: (a) Increasing the parameter $\overleftarrow{n}_{ii}$, beyond a threshold $\overleftarrow{n}^*_{ii}$, exclusively improves R_j. (b) None of the parameters $\overleftarrow{n}_{11}$ or $\overleftarrow{n}_{22}$ has an impact over the sum rate. Note that these observations are in line with the interpretation of channel-output feedback as an altruistic technique, as in [7, 8]. This is basically because the link implementing channel-output feedback provides an alternative path to the information sent by the other link, as first suggested in [1].

3.2.5 Example 5: none of the channel-output feedback links has any effect over the capacity region

Consider the case in which transmitter-receiver pairs 1 and 2 are in the very weak and strong interference regimes, with $\overrightarrow{n}_{11} = 10$, $\overrightarrow{n}_{22} = 9$, $n_{12} = 2$, $n_{21} = 15$. In Fig. 18 the capacity region is plotted without channel-output feedback and perfect channel-output feedback. Note that the capacity region of the LD-IC with and without channel-output feedback are identical. That is, none of the parameters either $\overleftarrow{n}_{11}$ or $\overleftarrow{n}_{22}$ enlarges the capacity region.

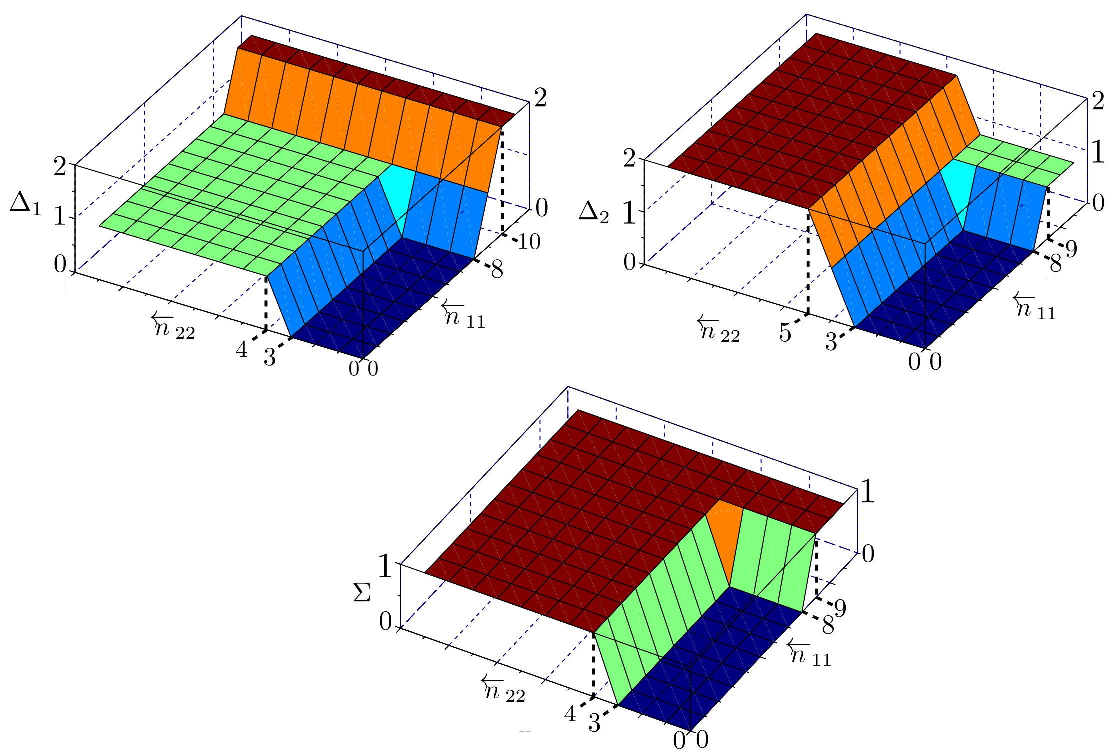

Figure 9: Maximum improvements $\Delta_1(10, 10, 3, 8, \cdot, \cdot)$ and $\Delta_2(10, 10, 3, 8, \cdot, \cdot)$ of one individual rate and $\Sigma(10, 10, 3, 8, \cdot, \cdot)$ of the sum rate of the example in Sec. 3.2.2.

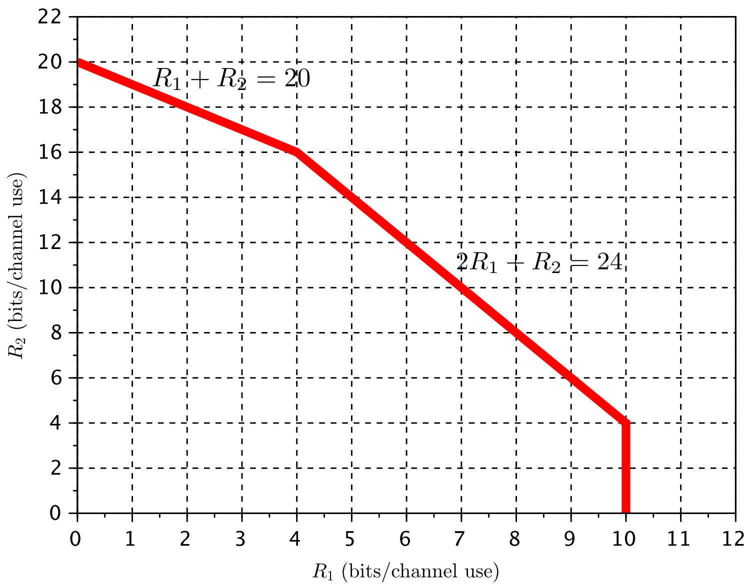

Figure 10: Capacity region of the example in Sec. 3.2.3 without feedback $\mathcal{C}(10, 20, 6, 12, 0, 0)$

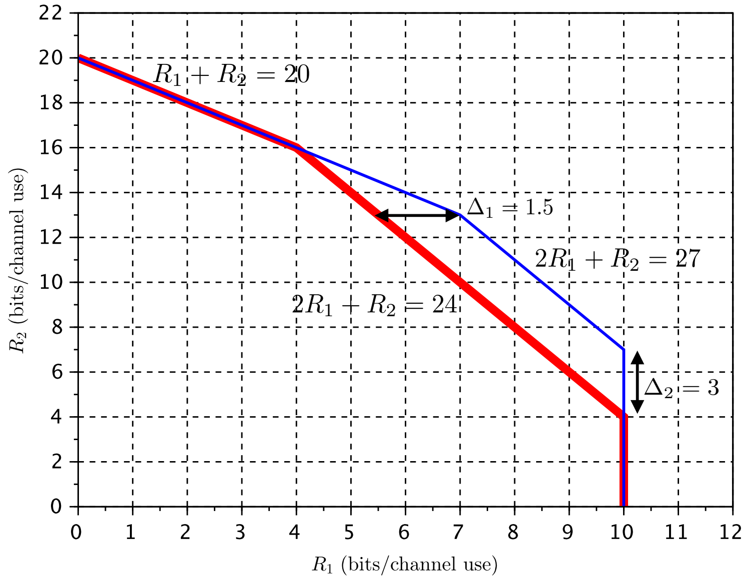

Figure 11: Capacity region $\mathcal{C}(10, 20, 6, 12, 0, 0)$ without feedback (thick red line) and $\mathcal{C}(10, 20, 6, 12, 10, 11)$ with noisy channel-output feedback (thin blue line) of the example in Sec. 3.2.3. Note that $\Delta_1(10, 20, 6, 12, 10, 11) = 1.5$ bits/ch.use, $\Delta_2(10, 20, 6, 12, 10, 11) = 2$ bits/ch.use and $\Sigma(10, 20, 6, 12, 10, 11) = 0$ bits/ch.use.

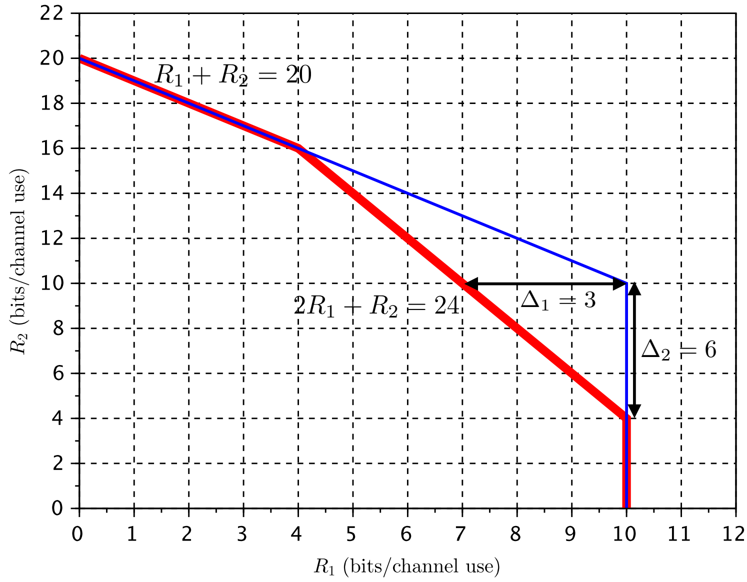

Figure 12: Capacity region $\mathcal{C}(10, 20, 6, 12, 0, 0)$ without feedback (thick red line) and $\mathcal{C}(10, 20, 6, 12, 10, 20)$ with perfect channel-output feedback (thin blue line) of the example in Sec. 3.2.3. Note that $\Delta_1(10, 20, 6, 12, 10, 20) = 3$ bits/ch.use, $\Delta_2(10, 20, 6, 12, 10, 20) = 6$ bits/ch.use and $\Sigma(10, 20, 6, 12, 10, 20) = 0$ bits/ch.use.

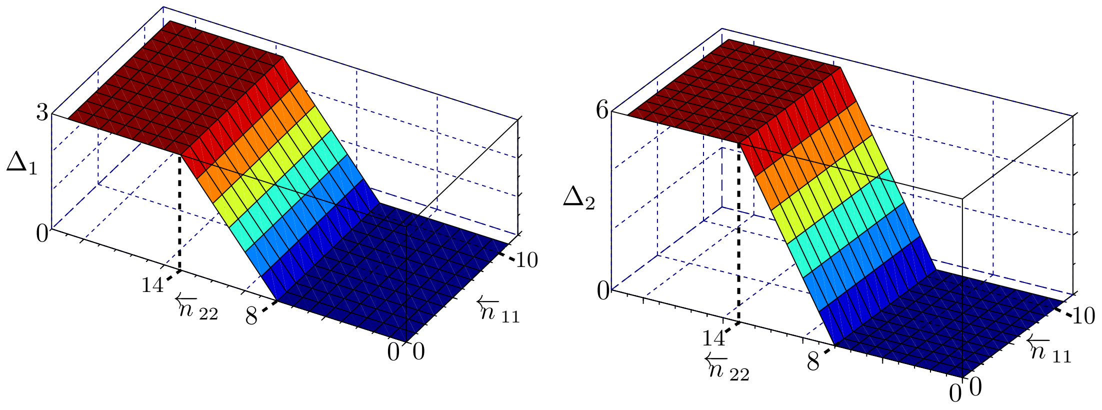

Figure 13: Maximum improvement $\Delta_1(10, 20, 6, 12, \cdot, \cdot)$ and $\Delta_2(10, 20, 6, 12, \cdot, \cdot)$ of one individual rate of the example in Sec. 3.2.3.

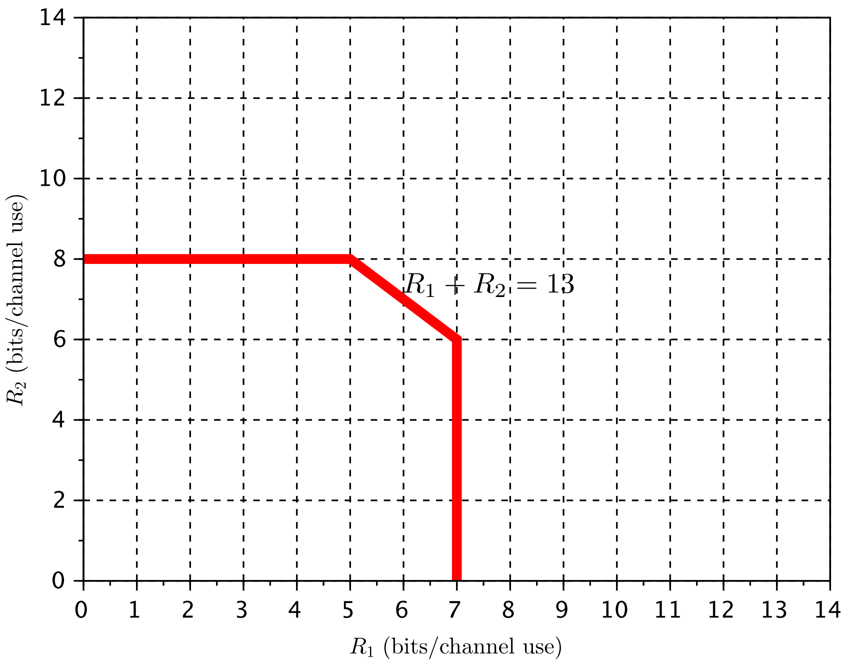

Figure 14: Capacity region of the example in Sec. 3.2.4 without feedback $\mathcal{C}(7, 8, 15, 13, 0, 0)$

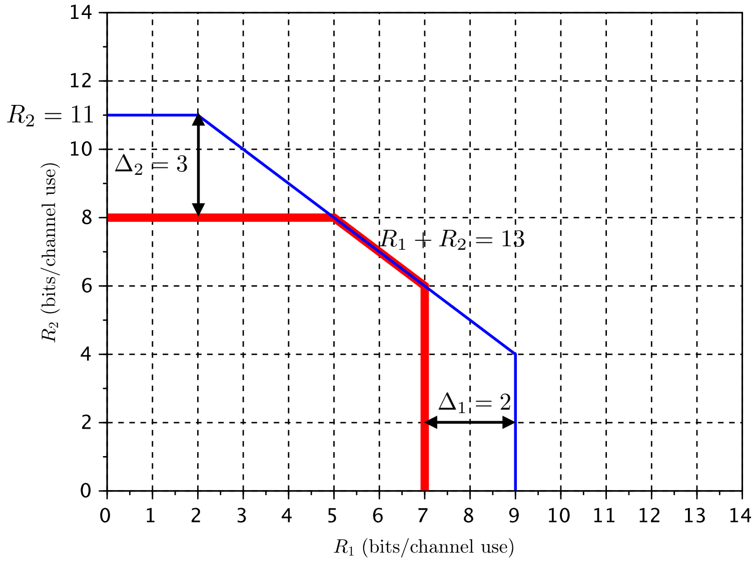

Figure 15: Capacity region $\mathcal{C}(7, 8, 15, 13, 0, 0)$ without feedback (thick red line) and $\mathcal{C}(7, 8, 15, 13, 11, 9)$ with noisy channel-output feedback (thin blue line) of the example in Sec. 3.2.4. Note that $\Delta_1(7, 8, 15, 13, 11, 9) = 2$ bits/ch.use, $\Delta_2(7, 8, 15, 13, 11, 9) = 3$ bits/ch.use and $\Sigma(7, 8, 15, 13, 11, 9) = 0$ bits/ch.use.

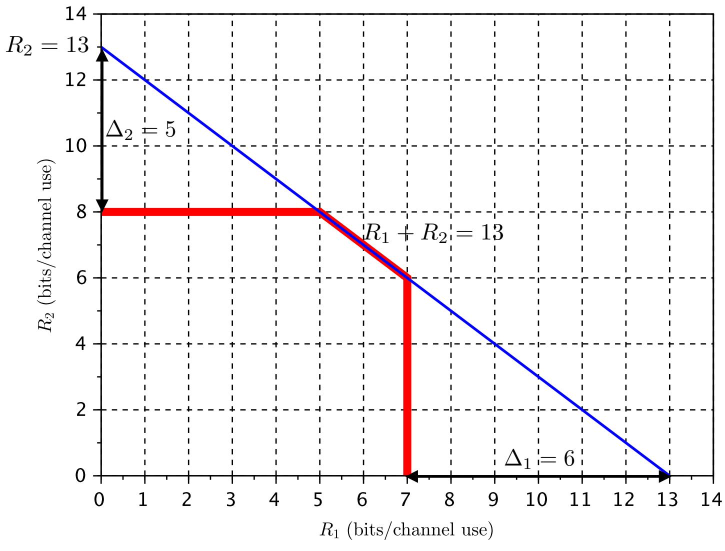

Figure 16: Capacity region of the example in Sec. 3.2.4 without feedback $\mathcal{C}(7, 8, 15, 13, 0, 0)$ (thick red line) and with perfect channel-output feedback $\mathcal{C}(7, 8, 15, 13, 15, 13)$ (thin blue line). Note that $\Delta_1(7, 8, 15, 13, 15, 13) = 6$ bits/ch.use, $\Delta_2(7, 8, 15, 13, 15, 13) = 5$ bits/ch.use and $\Sigma(7, 8, 15, 13, 15, 13) = 0$ bits/ch.use.

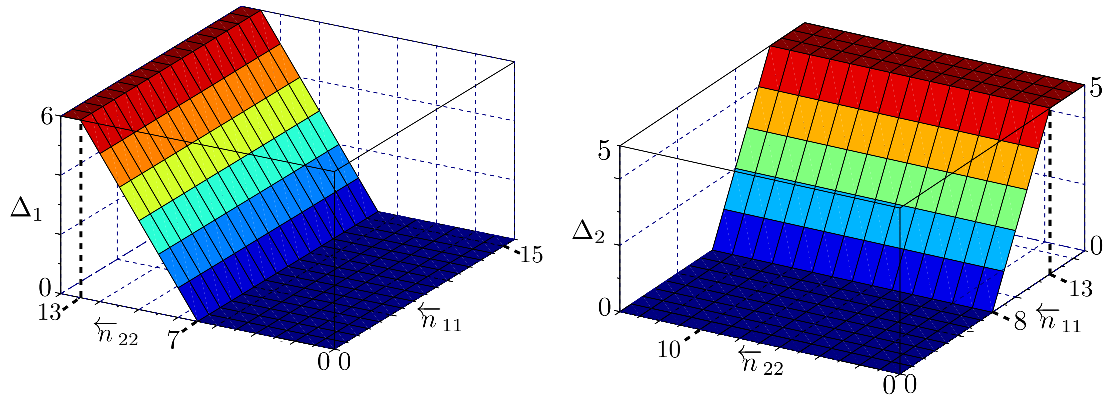

Figure 17: Maximum improvement $\Delta_1(7, 8, 15, 13, \cdot, \cdot)$ and $\Delta_1(7, 8, 15, 13, \cdot, \cdot)$ of one individual rate of the example in Sec. 3.2.4.

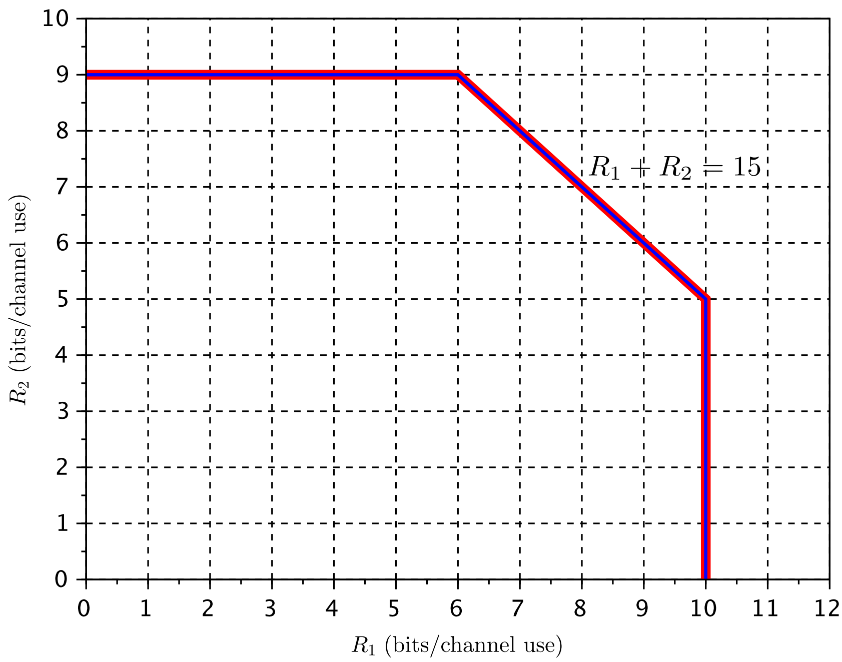

Figure 18: Capacity region $\mathcal{C}(10, 9, 2, 15, 0, 0)$ without feedback (thick red line) and $\mathcal{C}(10, 9, 2, 15, 10, 15)$ with perfect channel-output feedback (thin blue line) of the example in Sec. 3.2.5. Note that $\mathcal{C}(10, 9, 2, 15, 0, 0) = \mathcal{C}(10, 9, 2, 15, 10, 15)$.

4 Conclusions

In this technical report, the noisy channel-output feedback capacity of the linear deterministic interference channel has been fully characterized. Based on specific asymmetric examples, it is highlighted that even in the presence of noise, the benefits of channel-output feedback can be significantly relevant in terms of achievable individual rate and sum-rate improvements with respect to the case without feedback. Unfortunately, there also exist scenarios in which these benefits are totally inexistent.

Appendices

A Achievability Scheme

This appendix provides a description of the proposed achievability scheme, which is based on a three-part message splitting, superposition coding and backward decoding, as first suggested in [1, 3, 6]. The coding scheme is general and thus, it holds for other IC-NOF, i.e., Gaussian IC-NOF. However, the scope of this technical report is exclusively the case of the linear deterministic approximation.

Codebook Generation: fix a joint probability distribution $p(U, U_1, U_2, V_1, V_2, X_1, X_2) = p(U)$ $p(U_1|U)p(U_2|U)p(V_1|U, U_1)p(V_2|U, U_2)p(X_1|U, U_1, V_1)p(X_2|U, U_2, V_2)$.

Generate $2^{N(R_{1,C1}+R_{2,C1})}$ i.i.d. N-length codewords $\boldsymbol{u}(s, r) = \big(u_1(s, r), \ldots, u_N(s, r)\big)$ according to $p\big(\boldsymbol{u}(s, r)\big) = \prod_{i=1}^{N} p(u_i(s, r))$, with $s \in \{1, \ldots, 2^{NR_{1,C1}}\}$ and $r \in \{1, \ldots, 2^{NR_{2,C1}}\}$.

For encoder 1, generate for each codeword $\boldsymbol{u}(s, r)$, $2^{NR_{1,C1}}$ i.i.d. N-length codewords $\boldsymbol{u}_1(s, r, k) = \big(u_{1,1}(s, r, k), \ldots, u_{1,N}(s, r, k)\big)$ according to $p\big(\boldsymbol{u}_1(s, r, k)|\boldsymbol{u}(s, r)\big) = \prod_{i=1}^{N} p\big(u_{1,i}(s, r, k)|u_i(s, r)\big)$, with $k \in \{1, \ldots, 2^{NR_{1,C1}}\}$. 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 $\boldsymbol{v}_1(s, r, k, l) = \big(v_{1,1}(s, r, k, l), \ldots, v_{1,N}(s, r, k, l)\big)$ according to $p\big(\boldsymbol{v}_1(s, r, k, l)$ $\mid \boldsymbol{u}(s, r), \boldsymbol{u}_1(s, r, k)\big) = \prod_{i=1}^{N} p\big(v_{1,i}(s, r, k, l)|u_i(s, r), u_{1,i}(s, r, k)\big)$, with $l \in \{1, \ldots, 2^{NR_{1,C2}}\}$. For each tuple of codewords $\big(\boldsymbol{u}(s, r),\ \boldsymbol{u}_1(s, r, k),\ \boldsymbol{v}_1(s, r, k, l)\big)$, generate $2^{NR_{1,P}}$ i.i.d. N-length codewords $\boldsymbol{x}_1(s, r, k, l, q) = \big(x_{1,1}(s, r, k, l, q), \ldots, x_{1,N}(s, r, k, l, q)\big)$ according to $p\big(\boldsymbol{x}_1(s, r, k, l, q)$ $\mid \boldsymbol{u}(s, r), \boldsymbol{u}_1(s, r, k), \boldsymbol{v}_1(s, r, k, l)\big) = \prod_{i=1}^{N} p\big(x_{1,i}(s, r, k, l, q)|u_i(s, r), u_{1,i}(s, r, k), v_{1,i}(s, r, k, l)\big)$, with $q \in \{1, \ldots, 2^{NR_{1,P}}\}$.

For encoder 2, generate for each codeword $\boldsymbol{u}(s, r)$, $2^{NR_{2,C1}}$ i.i.d. N-length codewords $\boldsymbol{u}_2(s, r, j) = \big(u_{2,1}(s, r, j), \ldots, u_{2,N}(s, r, j)\big)$ according to $p\big(\boldsymbol{u}_2(s, r, j)|\boldsymbol{u}(s, r)\big) = \prod_{i=1}^{N} p\big(u_{2,i}(s, r, j)|u_i(s, r)\big)$, with $j \in \{1, \ldots, 2^{NR_{2,C1}}\}$. 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. length-N codewords $\boldsymbol{v}_2(s, r, j, m) = \big(v_{2,1}(s, r, j, m), \ldots, v_{2,N}(s, r, j, m)\big)$ according to $p\big(\boldsymbol{v}_2(s, r, j, m)$ $\mid \boldsymbol{u}(s, r),\ \boldsymbol{u}_2(s, r, j)\big) = \prod_{i=1}^{N} p(v_{2,i}(s, r, j, m) \mid u_i(s, r),\ u_{2,i}(s, r, j))$, with $m \in \{1, \ldots, 2^{NR_{2,C2}}\}$. For each tuple of codewords $\big(\boldsymbol{u}(s, r),\ \boldsymbol{u}_2(s, r, j), \boldsymbol{v}_2(s, r, j, m)\big)$, generate $2^{NR_{2,P}}$ i.i.d. N-length codewords $\boldsymbol{x}_2(s, r, j, m, b) = \big(x_{2,1}(s, r, j, m, b), \ldots, x_{2,N}(s, r, j, m, b)\big)$ according to $p\big(\boldsymbol{x}_2(s, r, j, m, b) \mid$ $\boldsymbol{u}(s, r),\ \boldsymbol{u}_2(s, r, j),\ \boldsymbol{v}_2(s, r, j, m)\big) = \prod_{i=1}^{N} p(x_{2,i}(s, r, j, m, b) \mid u_i(s, r),\ u_{2,i}(s, r, j),\ v_{2,i}(s, r, j, m, b))$, with $b \in \{1, \ldots, 2^{NR_{2,P}}\}$.

Encoding: denote by $W_i^{(t)} \in \{1, \ldots, 2^{N(R_{i,C}+R_{i,P})}\}$ the message index of transmitter i during block t, respectively. Let $W_i^{(t)} = (W_{i,C}^{(t)}, W_{i,P}^{(t)})$ be the message index composed by the message index $W_{i,C}^{(t)} \in \{1, \ldots, 2^{NR_{i,C}}\}$ and private message index $W_{i,P}^{(t)} \in \{1, \ldots, 2^{NR_{i,P}}\}$. Let also $W_{i,C}^{(t)} = (W_{i,C1}^{(t)}, W_{i,C2}^{(t)})$ be the message index composed by the message indexes $W_{i,C1}^{(t)} \in \{1, \ldots, 2^{NR_{i,C1}}\}$ and $W_{i,C2}^{(t)} \in \{1, \ldots, 2^{NR_{i,C2}}\}$. $W_{i,C1}^{(t)}$ is the message index that can be reliably decoded at least at one receiver and transmitter j (via feedback). $W_{i,C2}^{(t)}$ is the message index that can be reliably decoded at least at one receiver.

Consider Markov encoding with a length of T blocks. At encoding step t, with $t \in \{1, \ldots, T\}$, transmitter i sends the codeword $\boldsymbol{x}_1^{(t)} = \boldsymbol{x}_1\left(W_{1,C1}^{(t-1)}, W_{2,C1}^{(t-1)}, W_{1,C1}^{(t)}, W_{1,C2}^{(t)}, W_{1,P}^{(t)}\right)$, where $W_{1,C1}^{(0)} = W_{1,C1}^{(T)} = s^*$ and $W_{2,C1}^{(0)} = W_{2,C1}^{(T)} = r^*$. The pair $(s^*, r^*) \in \{1, \ldots, 2^{N\,R_{1,C1}}\} \times \{1, \ldots, 2^{NR_{2,C1}}\}$ is pre-defined and known at 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 messages indexes at the end of block T in a backward decoding fashion. At each decoding step t, with $t \in \{1, \ldots, T\}$, receiver 1 obtains the message indexes $\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, \widehat{W}_{1,C2}^{(T-(t-1))}, \widehat{W}_{1,P}^{(T-(t-1))}, \widehat{W}_{2,C2}^{(T-(t-1))}\right) \in \{1, \ldots, 2^{NR_{1,C1}}\} \times \{1, \ldots, 2^{NR_{2,C1}}\} \times \{1, \ldots, 2^{NR_{1,C2}}\} \times \{1, \ldots, 2^{NR_{1,P}}\} \times \{1, \ldots, 2^{NR_{2,C2}}\}$. The tuple $\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, \widehat{W}_{1,C2}^{(T-(t-1))}, \widehat{W}_{1,P}^{(T-(t-1))}, \widehat{W}_{2,C2}^{(T-(t-1))}\right)$ is the unique tuple that satisfies

$$\begin{aligned}
&\Big(\boldsymbol{u}\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}\right), \boldsymbol{u}_1\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, W_{1,C1}^{(T-(t-1))}\right), \\
&\boldsymbol{v}_1\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, W_{1,C1}^{(T-(t-1))}, \widehat{W}_{1,C2}^{(T-(t-1))}\right), \\
&\boldsymbol{x}_1\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, W_{1,C1}^{(T-(t-1))}, \widehat{W}_{1,C2}^{(T-(t-1))}, \widehat{W}_{1,P}^{(T-(t-1))}\right), \\
&\boldsymbol{u}_2\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, W_{2,C1}^{(T-(t-1))}\right), \\
&\boldsymbol{v}_2\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, W_{2,C1}^{(T-(t-1))}, \widehat{W}_{2,C2}^{(T-(t-1))}\right), \overrightarrow{\boldsymbol{y}}_1^{(T-(t-1))}\Big) \in \mathcal{A}_e^{(n)},
\end{aligned} \tag{15}$$

where $W_{1,C1}^{(T-(t-1))}$ and $W_{2,C1}^{(T-(t-1))}$ are assumed to be perfectly decoded in the previous decoding step $t-1$. The set $\mathcal{A}_e^{(n)}$ represents the set of jointly typical sequences. Finally, receiver 2 follows a similar decoding scheme.

Probability of Error Analysis: an error might occur during the coding phase at the beginning of block t if the common message index $W_{2,C1}^{(t-1)}$ is not correctly decoded at transmitter 1. From the asymptotic equipartion property (AEP) [9], it follows that the message index $W_{2,C1}^{(t-1)}$ can be reliably decoded at transmitter 1 during encoding step t, under the condition:

$$R_{2,C1} \leqslant I\left(\overleftarrow{Y}_1; U_2 | U, X_1\right). \tag{16}$$

An error might occur during the (backward) decoding step t if the messages indexes $W_{1,C1}^{(T-t)}$, $W_{2,C1}^{(T-t)}$, $W_{1,C2}^{(T-(t+1))}, W_{1,P}^{(T-(t+1))}$ and $W_{2,C2}^{(T-(t+1))}$ are not decoded correctly given that the message indexes $W_{1,C1}^{(T-(t+1))}$ and $W_{2,C1}^{(T-(t+1))}$ were correctly decoded in the previous decoding step $t-1$. These errors might arise for two reasons: (i) there does not exist a tuple $\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, \widehat{W}_{1,C2}^{(T-(t+1))}, \widehat{W}_{1,P}^{(T-(t+1))}, \widehat{W}_{2,C2}^{(T-(t+1))}\right)$ that satisfies (15), or (ii) there exist several tuples $\left(\widehat{W}_{1,C1}^{(T-t)}, \widehat{W}_{2,C1}^{(T-t)}, \widehat{W}_{1,C2}^{(T-(t+1))}, \widehat{W}_{1,P}^{(T-(t+1))}, \widehat{W}_{2,C2}^{(T-(t+1))}\right)$ that simultaneously satisfy (15). From the asymptotic equipartion property (AEP) [9], 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 $\big(\boldsymbol{u}(s,r)$, $\boldsymbol{u}_1(s,r,W_{1,C1}^{(T-(t+1))})$, $\boldsymbol{v}_1(s,r,W_{1,C1}^{(T-(t+1))},l)$, $\boldsymbol{x}_1(s,r,W_{1,C1}^{(T-(t+1))},l,q)$, $\boldsymbol{u}_2(s,r,W_{2,C1}^{(T-(t+1))})$ and $\boldsymbol{v}_2(s,r,W_{2,C1}^{(T-(t+1))},m)\big)$ are jointly typical with

$\overrightarrow{\boldsymbol{y}}_1^{(T-(t+1))}$ during decoding step t. Assume now that the codeword to be decoded at decoding step t corresponds to the indexes $(s,r,l,q,m) = (1,1,1,1,1)$. This is without loss of generality due to the symmetry of the code. Then, the probability of error due to (ii) during decoding step t, $p_e^{(t)}$, can be bounded as follows

$$
\begin{aligned}
p_e^{(t)} =& \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=1,r=1,l=1,q=1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) + \sum_{s=1,r=1,l=1,q\neq1,m=1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) \\
&+ \sum_{s=1,r=1,l=1,q\neq1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) + \sum_{s=1,r=1,l\neq1,q=1,m=1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) \\
&+ \sum_{s=1,r=1,l\neq1,q=1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) + \sum_{s=1,r=1,l\neq1,q\neq1,m=1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) \\
&+ \sum_{s=1,r=1,l\neq1,q\neq1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) + \sum_{s=1,r\neq1,l=1,q=1,m=1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) \\
&+ \sum_{s=1,r\neq1,l=1,q=1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) + \sum_{s=1,r\neq1,l=1,q\neq1,m=1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) \\
&+ \sum_{s=1,r\neq1,l=1,q\neq1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) + \sum_{s=1,r\neq1,l\neq1,q=1,m=1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) \\
&+ \sum_{s=1,r\neq1,l\neq1,q=1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) + \sum_{s=1,r\neq1,l\neq1,q\neq1,m=1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) \\
&+ \sum_{s=1,r\neq1,l\neq1,q\neq1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) + \sum_{s\neq1,r=1,l=1,q=1,m=1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) \\
&+ \sum_{s\neq1,r=1,l=1,q=1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) + \sum_{s\neq1,r=1,l=1,q\neq1,m=1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) \\
&+ \sum_{s\neq1,r=1,l=1,q\neq1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) + \sum_{s\neq1,r=1,l\neq1,q=1,m=1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) \\
&+ \sum_{s\neq1,r=1,l\neq1,q=1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) + \sum_{s\neq1,r=1,l\neq1,q\neq1,m=1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) \\
&+ \sum_{s\neq1,r=1,l\neq1,q\neq1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) + \sum_{s\neq1,r\neq1,l=1,q=1,m=1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) \\
&+ \sum_{s\neq1,r\neq1,l=1,q=1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) + \sum_{s\neq1,r\neq1,l=1,q\neq1,m=1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) \\
&+ \sum_{s\neq1,r\neq1,l=1,q\neq1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) + \sum_{s\neq1,r\neq1,l\neq1,q=1,m=1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) \\
&+ \sum_{s\neq1,r\neq1,l\neq1,q=1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) + \sum_{s\neq1,r\neq1,l\neq1,q\neq1,m=1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right) \\
&+ \sum_{s\neq1,r\neq1,l\neq1,q\neq1,m\neq1} \Pr\left(E^{(t)}_{(s,r,l,q,m)}\right). \qquad (17)
\end{aligned}
$$

From the asymptotic equipartion property (AEP) [9], it follows that:

$$
\begin{aligned}
p_e^{(t)} \leqslant\, & 2^{N\left(R_{2,C2}-I(\overrightarrow{Y}_1;V_2|U,U_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,P}-I(\overrightarrow{Y}_1;X_1|U,V_1,V_2)+2\epsilon\right)}\\
&+2^{N\left(R_{2,C2}+R_{1,P}-I(\overrightarrow{Y}_1;V_2,X_1|U,U_2,V_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C2}-I(\overrightarrow{Y}_1;X_1|U,U_1,V_2)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C2}+R_{2,C2}-I(\overrightarrow{Y}_1;V_2,X_1|U,U_1,U_2)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C2}+R_{1,P}-I(\overrightarrow{Y}_1;X_1|U,U_1,V_2)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C2}+R_{1,P}+R_{2,C2}-I(\overrightarrow{Y}_1;V_2,X_1|U,U_1,U_2)+2\epsilon\right)}\\
&+2^{N\left(R_{2,C1}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{2,C1}+R_{2,C2}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{2,C1}+R_{1,P}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{2,C1}+R_{1,P}+R_{2,C2}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{2,C1}+R_{1,C2}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{2,C1}+R_{1,C2}+R_{2,C2}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{2,C1}+R_{1,C2}+R_{1,P}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{2,C1}+R_{1,C2}+R_{1,P}+R_{2,C2}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}+R_{2,C2}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}+R_{1,P}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}+R_{1,P}+R_{2,C2}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}+R_{1,C2}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}+R_{1,C2}+R_{2,C2}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}+R_{1,C2}+R_{1,P}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}+R_{1,C2}+R_{1,P}+R_{2,C2}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}+R_{2,C1}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}+R_{2,C1}+R_{2,C2}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}+R_{2,C1}+R_{1,P}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}+R_{2,C1}+R_{1,P}+R_{2,C2}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}+R_{2,C1}+R_{1,C2}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}+R_{2,C1}+R_{1,C2}+R_{2,C2}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}+R_{2,C1}+R_{1,C2}+R_{1,P}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}\\
&+2^{N\left(R_{1,C1}+R_{2,C1}+R_{1,C2}+R_{1,P}+R_{2,C2}-I(\overrightarrow{Y}_1;U,V_2,X_1)+2\epsilon\right)}. \qquad (18)
\end{aligned}
$$

The same analysis of the probability of error holds for transmitter-receiver pair 2. Hence, in general, from (16) and (18), reliable decoding holds under the following conditions for transmitter

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

$$
\begin{aligned}
&R_{j,C1} &&\leqslant I\left(\overleftarrow{Y}_i; U_j | U, X_i\right) && (19)\\
&R_{i,C1}+R_{i,C2}+R_{i,P}+R_{j,C1}+R_{j,C2} &&\leqslant I(\overrightarrow{Y}_i; U, V_j, X_i) && (20)\\
&R_{j,C2} &&\leqslant I(\overrightarrow{Y}_i; V_j | U, U_j, X_i) && (21)\\
&R_{i,P} &&\leqslant I(\overrightarrow{Y}_i; X_i | U, V_i, V_j) && (22)\\
&R_{i,P}+R_{j,C2} &&\leqslant I(\overrightarrow{Y}_i; V_j, X_i | U, U_j, V_i) && (23)\\
&R_{i,C2}+R_{i,P} &&\leqslant I(\overrightarrow{Y}_i; X_i | U, U_i, V_j) && (24)\\
&R_{i,C2}+R_{i,P}+R_{j,C2} &&\leqslant I(\overrightarrow{Y}_i; V_j, X_i | U, U_i, U_j). && (25)
\end{aligned}
$$

In terms of the parameters of the linear deterministic model, the following holds

$$
\begin{aligned}
&I\left(\overleftarrow{Y}_i; U_j | U, X_i\right) &&\leqslant \left(n_{ij} - (\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii})^+\right)^+, && (26)\\
&I(\overrightarrow{Y}_i; U, V_j, X_i) &&\leqslant \max(\overrightarrow{n}_{ii}, n_{ij}), && (27)\\
&I(\overrightarrow{Y}_i; V_j | U, U_j, X_i) &&\leqslant \min\left(n_{ij}, (\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii})^+\right), && (28)\\
&I(\overrightarrow{Y}_i; X_i | U, V_i, V_j) &&\leqslant (\overrightarrow{n}_{ii} - n_{ji})^+, && (29)\\
&I(\overrightarrow{Y}_i; V_j, X_i | U, U_j, V_i) &&\leqslant \max\left((\overrightarrow{n}_{ii} - n_{ji})^+, \min\left(n_{ij}, (\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii})^+\right)\right), && (30)\\
&I(\overrightarrow{Y}_i; X_i | U, U_i, V_j) &&\leqslant \min\left(n_{ji}, (\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+\right) - \min\left((n_{ji} - \overrightarrow{n}_{ii})^+,\right. && \\
& &&\left.(\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})\right) + (\overrightarrow{n}_{ii} - n_{ji})^+, && (31)\\
&I(\overrightarrow{Y}_i; V_j, X_i | U, U_i, U_j) &&\leqslant \max\left(\min\left(n_{ij}, (\max(\overrightarrow{n}_{ii}, n_{ij}) - \overleftarrow{n}_{ii})^+\right), (\overrightarrow{n}_{ii} - n_{ji})^+ +\right. && \\
& &&\min\left(n_{ji}, (\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+\right) - \min\left((n_{ji} - \overrightarrow{n}_{ii})^+,\right. && \\
& &&\left.\left.(\max(\overrightarrow{n}_{jj}, n_{ji}) - \overleftarrow{n}_{jj})^+\right)\right). && (32)
\end{aligned}
$$

Finally, taking into account that $R_i = R_{i,C1} + R_{i,C2} + R_{i,P}$, a Fourier-Motzkin elimination process in (19) - (25) yields the system of inequalities in Theorem 1.

B An Outer Bound Region

This appendix provides a proof for the following Lemma.

Lemma 1 (Converse) 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 linear deterministic interference channel with noisy channel-output feedback is included in the set of non-negative rate pairs (R_1, R_2) that satisfies the following conditions for all $i \in \{1,2\}$ and for all $j \in \{1,2\} \setminus \{i\}$:

$$
\begin{aligned}
R_i &\leqslant \min\left(\max\left(\overrightarrow{n}_{ii}, n_{ji}\right), \max\left(\overrightarrow{n}_{ii}, n_{ij}\right)\right), && (33)\\
R_i &\leqslant \min\left(\left(n_{ji} - \overrightarrow{n}_{ii}\right)^+, \left(\overleftarrow{n}_{jj} - \min\left(\overrightarrow{n}_{ii}, n_{ji}\right) - \min\left(\left(\overrightarrow{n}_{jj} - n_{ji}\right)^+, n_{ij}\right)\right)^+\right)\\
&\quad + \overrightarrow{n}_{ii}, && (34)\\
R_1 + R_2 &\leqslant \min\left(\max\left(\overrightarrow{n}_{11}, n_{12}\right) + \left(\overrightarrow{n}_{22} - n_{12}\right)^+, \max\left(\overrightarrow{n}_{22}, n_{21}\right) + \left(\overrightarrow{n}_{11} - n_{21}\right)^+\right), && (35)\\
R_1 + R_2 &\leqslant \left(\overrightarrow{n}_{11} - n_{21}\right)^+ + \left(\overrightarrow{n}_{22} - n_{12}\right)^+ + \left(\overleftarrow{n}_{11} - \left(\overrightarrow{n}_{11} - n_{12}\right)^+\right)^+\\
&\quad + \left(\overleftarrow{n}_{22} - \left(\overrightarrow{n}_{22} - n_{21}\right)^+\right)^+, && (36)\\
2R_i + R_j &\leqslant \max\left(\overrightarrow{n}_{ii}, n_{ij}\right) + \left(\overrightarrow{n}_{ii} - n_{ji}\right)^+ + \left(\overrightarrow{n}_{jj} - n_{ij}\right)^+ + \left(\overleftarrow{n}_{jj} - \left(\overrightarrow{n}_{jj} - n_{ji}\right)^+\right)^+. && (37)
\end{aligned}
$$

Proof of (33) **and** (35)**:** inequalities (33) and (35) corresponds to the minimum cut-set bound [10] and the sum-rate bounds of the case of the two-user interference channel with perfect channel-output feedback [1]. ■

The rest of the proof of Lemma 1 is presented using particular notation. For all $i \in \{1,2\}$, the channel input $\boldsymbol{X}_i^{(n)}$ of the LD-IC-NOF in (2) for any channel use $n \in \{1,\ldots,N\}$ is a q-dimensional vector, with q in (1), 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., $\boldsymbol{X}_i^{(n)} = \left(\boldsymbol{X}_{i,C}^{(n)}, \boldsymbol{X}_{i,P}^{(n)}, \boldsymbol{X}_{i,D}^{(n)}, \boldsymbol{X}_{i,Q}^{(n)}\right)$, as shown in Fig. 19. Note that this notation is independent of the feedback parameters $\overleftarrow{n}_{11}$ and $\overleftarrow{n}_{22}$ and it holds for all $n \in \{1,\ldots,N\}$. More specifically,

- $\boldsymbol{X}_{i,C}^{(n)}$ contains the signal levels at transmitter i that are observed at both receivers and thus,
$$\dim \boldsymbol{X}_{i,C}^{(n)} = \min\left(\overrightarrow{n}_{ii}, n_{ji}\right); \tag{38}$$
- $\boldsymbol{X}_{i,P}^{(n)}$ contains the signal levels at transmitter i that are observed only at receiver i and thus,
$$\dim \boldsymbol{X}_{i,P}^{(n)} = \left(\overrightarrow{n}_{ii} - n_{ji}\right)^+; \tag{39}$$
- $\boldsymbol{X}_{i,D}^{(n)}$ contains the signal levels at transmitter i that are observed only at receiver j and thus,
$$\dim \boldsymbol{X}_{i,D}^{(n)} = \left(n_{ji} - \overrightarrow{n}_{ii}\right)^+; \text{ and} \tag{40}$$
- $\boldsymbol{X}_{i,Q}^{(n)}$ is included for dimensional matching of the model in (3). Then,
$$\dim \boldsymbol{X}_{i,Q}^{(n)} = q - \max\left(\overrightarrow{n}_{ii}, n_{ji}\right). \tag{41}$$

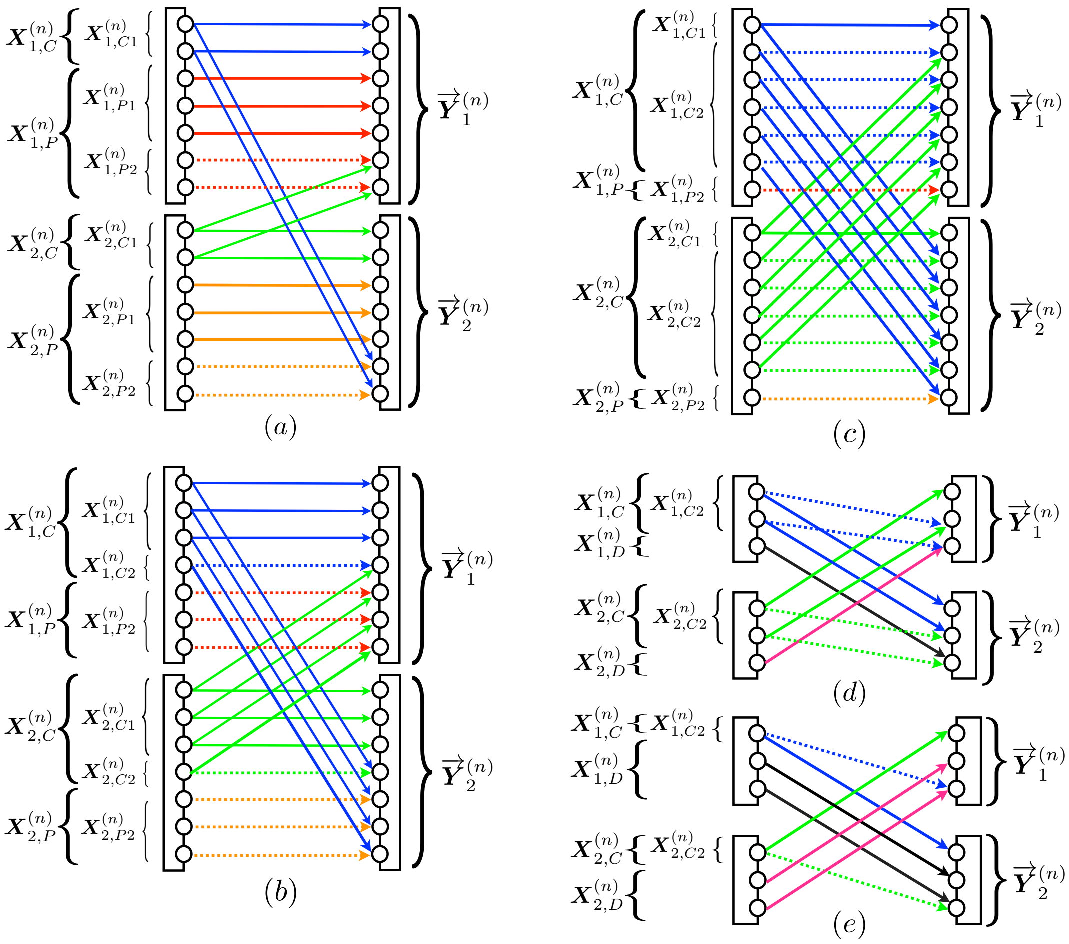

Figure 19: Examples of the channel inputs written in terms of $\boldsymbol{X}_{i,C}^{(n)}$, $\boldsymbol{X}_{i,P}^{(n)}$ and $\boldsymbol{X}_{i,D}^{(n)}$ in (a) very weak interference regime, (b) weak interference regime, (c) moderate interference regime, (d) strong interference regime and (e) very strong interference regime

These levels $\boldsymbol{X}_{i,Q}^{(n)}$ are not used for signal transmission by transmitter i and thus,

$$\begin{aligned}
H\left(\boldsymbol{X}_{i}^{(n)}\right) &= H\left(\boldsymbol{X}_{i,C}^{(n)}, \boldsymbol{X}_{i,P}^{(n)}, \boldsymbol{X}_{i,D}^{(n)}, \boldsymbol{X}_{i,Q}^{(n)}\right)\\
&= H\left(\boldsymbol{X}_{i,C}^{(n)}, \boldsymbol{X}_{i,P}^{(n)}, \boldsymbol{X}_{i,D}^{(n)}\right)\\
&\leqslant \dim \boldsymbol{X}_{i,C}^{(n)} + \dim \boldsymbol{X}_{i,P}^{(n)} + \dim \boldsymbol{X}_{i,D}^{(n)}.
\end{aligned} \tag{42}$$

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

$$\begin{aligned}
\dim \boldsymbol{X}_{i}^{(n)} &= \dim \boldsymbol{X}_{i,C}^{(n)} + \dim \boldsymbol{X}_{i,P}^{(n)} + \dim \boldsymbol{X}_{i,D}^{(n)} + \dim \boldsymbol{X}_{i,Q}^{(n)}\\
&= \min\left(\overrightarrow{n}_{ii}, n_{ji}\right) + \left(\overrightarrow{n}_{ii} - n_{ji}\right)^{+} + \left(n_{ji} - \overrightarrow{n}_{ii}\right)^{+} + q - \max\left(\overrightarrow{n}_{ii}, n_{ji}\right)\\
&= \max\left(\overrightarrow{n}_{ii}, n_{ji}\right) + q - \max\left(\overrightarrow{n}_{ii}, n_{ji}\right)\\
&= q.
\end{aligned} \tag{43}$$

Vector $\boldsymbol{X}_{i,C}^{(n)}$ can be written as the concatenation of two vectors: $\boldsymbol{X}_{i,C1}^{(n)}$ and $\boldsymbol{X}_{i,C2}^{(n)}$, i.e., $\boldsymbol{X}_{i,C}^{(n)} = \left(\boldsymbol{X}_{i,C1}^{(n)}, \boldsymbol{X}_{i,C2}^{(n)}\right)$. Vector $\boldsymbol{X}_{i,C1}^{(n)}$ (resp. $\boldsymbol{X}_{i,C2}^{(n)}$) contains the levels of $\boldsymbol{X}_{i,C}^{(n)}$ that are seen at receiver i without any interference (resp. with interference of transmitter j). Hence,

$$\dim \boldsymbol{X}_{i,C1}^{(n)} = \min\left(\left(\overrightarrow{n}_{ii} - n_{ij}\right)^+, n_{ji}\right), \text{ and} \tag{44}$$

$$\dim \boldsymbol{X}_{i,C2}^{(n)} = \min\left(\overrightarrow{n}_{ii}, n_{ji}\right) - \min\left(\left(\overrightarrow{n}_{ii} - n_{ij}\right)^+, n_{ji}\right). \tag{45}$$

Vector $\boldsymbol{X}_{i,P}^{(n)}$ can also be written as the concatenation of two vectors: $\boldsymbol{X}_{i,P1}^{(n)}$ and $\boldsymbol{X}_{i,P2}^{(n)}$, i.e., $\boldsymbol{X}_{i,P}^{(n)} = \left(\boldsymbol{X}_{i,P1}^{(n)}, \boldsymbol{X}_{i,P2}^{(n)}\right)$. Vector $\boldsymbol{X}_{i,P1}^{(n)}$ (resp. $\boldsymbol{X}_{i,P2}^{(n)}$) contains the levels of $\boldsymbol{X}_{i,P}^{(n)}$ that are seen at receiver i without any interference (resp. with interference of transmitter j). Hence,

$$\dim \boldsymbol{X}_{i,P1}^{(n)} = \left(\left(\overrightarrow{n}_{ii} - n_{ji}\right)^+ - n_{ij}\right)^+ \text{ and} \tag{46}$$

$$\dim \boldsymbol{X}_{i,P2}^{(n)} = \min\left(\left(\overrightarrow{n}_{ii} - n_{ji}\right)^+, n_{ij}\right)^+. \tag{47}$$

When feedback is taken into account, an alternative notation is needed. Let $\boldsymbol{X}_{i,C}^{(n)}$ be written in terms of $\boldsymbol{X}_{i,CF_j}^{(n)}$, $\boldsymbol{X}_{i,CG_j}^{(n)}$, with $j \in \{1,2\}$, i.e., $\boldsymbol{X}_{i,C}^{(n)} = \left(\boldsymbol{X}_{i,CF_j}^{(n)}, \boldsymbol{X}_{i,CG_j}^{(n)}\right)$. The vector $\boldsymbol{X}_{i,CF_j}^{(n)}$ represents the signal levels of $\boldsymbol{X}_{i,C}^{(n)}$ that can be fed back from receiver j to transmitter j; and $\boldsymbol{X}_{i,CG_j}^{(n)}$ represents the signal levels of $\boldsymbol{X}_{i,C}^{(n)}$ that can not be fed back from receiver j to transmitter j, as shown in Fig. 21. Let $\boldsymbol{X}_{i,D}^{(n)}$ be written in terms of $\boldsymbol{X}_{i,DF}^{(n)}$ and $\boldsymbol{X}_{i,DG}^{(n)}$, i.e., $\boldsymbol{X}_{i,D}^{(n)} = \left(\boldsymbol{X}_{i,DF}^{(n)}, \boldsymbol{X}_{i,DG}^{(n)}\right)$. The vector $\boldsymbol{X}_{i,DF}^{(n)}$ represents the signal levels of $\boldsymbol{X}_{i,D}^{(n)}$ that can be fed back from receiver j to transmitter j, with $j \in \{1,2\} \setminus \{i\}$. The dimension of vector $\boldsymbol{X}_{i,DF}^{(n)}$ is function of the dimensions of vectors $\boldsymbol{X}_{i,C}^{(n)}$, $\boldsymbol{X}_{i,D}^{(n)}$, $\boldsymbol{X}_{j,C1}^{(n)}$, and $\boldsymbol{X}_{j,P1}^{(n)}$, which were defined in (38), (41), (44) and (46) respectively. The vector $\boldsymbol{X}_{i,DG}^{(n)}$ represents the signal levels that are in $\boldsymbol{X}_{i,D}^{(n)}$ but not $\boldsymbol{X}_{i,DF}^{(n)}$, as shown in Fig. 21. The dimension of vectors $\boldsymbol{X}_{i,DF}^{(n)}$ and $\boldsymbol{X}_{i,DG}^{(n)}$ are defined as follows:

$$\begin{aligned}
\dim \boldsymbol{X}_{i,DF}^{(n)} &= \min\left(\dim \boldsymbol{X}_{i,D}^{(n)}, \left(\overleftarrow{n}_{jj} - \dim \boldsymbol{X}_{i,C}^{(n)} - \dim \boldsymbol{X}_{j,C1}^{(n)} - \dim \boldsymbol{X}_{j,P1}^{(n)}\right)^+\right), \\
&= \min\Big(\left(n_{ji} - \overrightarrow{n}_{ii}\right)^+, \Big(\overleftarrow{n}_{jj} - \min\left(\overrightarrow{n}_{ii}, n_{ji}\right) - \min\left(\left(\overrightarrow{n}_{jj} - n_{ji}\right)^+, n_{ij}\right) \\
&\quad - \left(\left(\overrightarrow{n}_{jj} - n_{ij}\right)^+ - n_{ji}\right)^+\Big)^+\Big), \\
&\overset{(a)}{=} \min\Big(\left(n_{ji} - \overrightarrow{n}_{ii}\right)^+, \Big(\overleftarrow{n}_{jj} - \overrightarrow{n}_{ii} - \min\left(\left(\overrightarrow{n}_{jj} - n_{ji}\right)^+, n_{ij}\right) \\
&\quad - \left(\left(\overrightarrow{n}_{jj} - n_{ij}\right)^+ - n_{ji}\right)^+\Big)^+\Big); \text{ and}
\end{aligned} \tag{48}$$

$$\dim \boldsymbol{X}_{i,DG}^{(n)} = \dim \boldsymbol{X}_{i,D}^{(n)} - \dim \boldsymbol{X}_{i,DF}^{(n)}, \tag{49}$$

where (a) follows from the fact that vector $\boldsymbol{X}_{i,D}^{(n)}$ only exists if $n_{ji} > \overrightarrow{n}_{ii}$. Note that it is not necessary to include a subindex related to the receiver that implements the feedback, this is mainly because the signal levels $\boldsymbol{X}_{i,D}^{(n)}$ are seen only at receiver j.

More generally, when needed, the vector $\boldsymbol{X}_{iF_k}^{(n)}$ is used to represent the signal levels of $\boldsymbol{X}_{i}^{(n)}$ that can be fed back from receiver k to transmitter k, with $k \in \{1,2\}$. The vector $\boldsymbol{X}_{iG_k}^{(n)}$ is used to represent the signal levels of $\boldsymbol{X}_{i}^{(n)}$ that can not be fed back from receiver k to transmitter k.

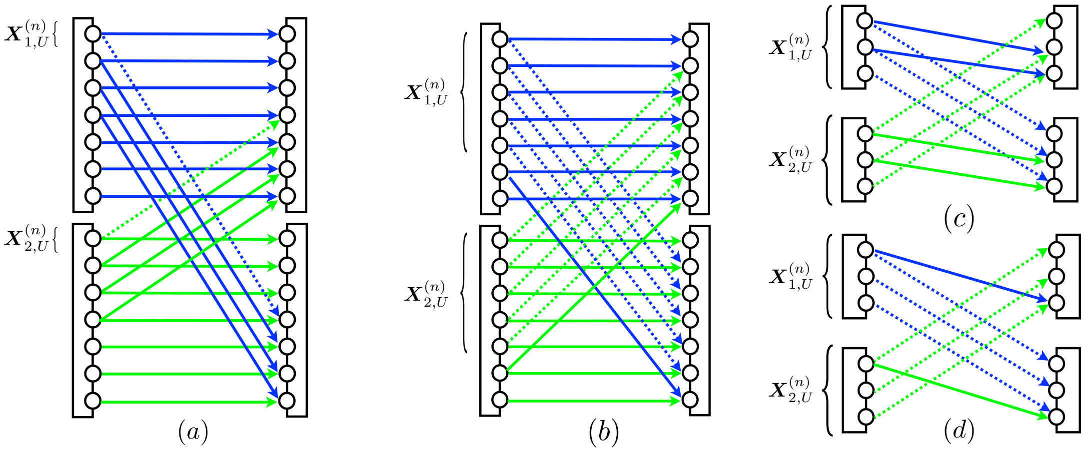

Figure 20: Vector $\boldsymbol{X}_{i,U}^{(n)}$ in (a) weak interference regime, (b) moderate interference regime, (c) strong interference regime and (d) very strong interference regime. Vector $\boldsymbol{X}_{i,U}^{(n)}$ does not exist in the very weak interference regime.

In the proofs of (36) and (37), the vector $\boldsymbol{X}_{i,U}^{(n)}$ is used to represent the signal levels of vector $\boldsymbol{X}_{i}^{(n)}$ that interfere with signal levels of $\boldsymbol{X}_{j,C}^{(n)}$ at receiver j and those signal levels of $\boldsymbol{X}_{i}^{(n)}$ that are seen at receiver j and are not used by transmitter j. An example is shown in Fig. 20. Vector $\boldsymbol{X}_{i,U}^{(n)}$ was used for the first time in the proof of the converse in the symmetric case in [5].

Finally, for all $i \in \{1, 2\}$, the channel output $\overrightarrow{\boldsymbol{Y}}_{i}^{(n)}$ of the LD-IC-NOF in (2) for any channel use $n \in \{1, \ldots, N\}$ is a q-dimensional vector, with q in (1), that can be written as the concatenation of three vectors: $\overleftarrow{\boldsymbol{Y}}_{i}^{(n)}$, $\overrightarrow{\boldsymbol{Y}}_{i,G}^{(n)}$ and $\overrightarrow{\boldsymbol{Y}}_{i,Q}^{(n)}$, i.e., $\overrightarrow{\boldsymbol{Y}}_{i}^{(n)} = \left(\overleftarrow{\boldsymbol{Y}}_{i}^{(n)}, \overrightarrow{\boldsymbol{Y}}_{i,G}^{(n)}, \overrightarrow{\boldsymbol{Y}}_{i,Q}^{(n)}\right)$, as shown in Fig. 21. More specifically,

- $\overleftarrow{\boldsymbol{Y}}_{i}^{(n)}$ contains the signal levels at receiver i that are fed back to transmitter i and thus,

$$\dim \overleftarrow{\boldsymbol{Y}}_{i}^{(n)} = \min\left(\overleftarrow{n}_{ii}, \max\left(\overrightarrow{n}_{ii}, n_{ij}\right)\right); \tag{50}$$

- $\overrightarrow{\boldsymbol{Y}}_{i,G}^{(n)}$ contains the signal levels at receiver i that are not fed back to transmitter i and thus,

$$\dim \overrightarrow{\boldsymbol{Y}}_{i,G}^{(n)} = \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^{+}; \tag{51}$$

- $\overrightarrow{\boldsymbol{Y}}_{i,Q}^{(n)}$ is included for dimensional matching of the model in (3). Then,

$$\dim \overrightarrow{\boldsymbol{Y}}_{i,Q}^{(n)} = q - \max\left(\overrightarrow{n}_{ii}, n_{ij}\right); \tag{52}$$

These levels $\overrightarrow{\boldsymbol{Y}}_{i,Q}^{(n)}$ do not represent an output of the channel and thus,

$$\begin{aligned} H\left(\overrightarrow{\boldsymbol{Y}}_{i}^{(n)}\right) &= H\left(\overleftarrow{\boldsymbol{Y}}_{i}^{(n)}, \overrightarrow{\boldsymbol{Y}}_{i,G}^{(n)}, \overrightarrow{\boldsymbol{Y}}_{i,Q}^{(n)}\right), \\ &= H\left(\overleftarrow{\boldsymbol{Y}}_{i}^{(n)}, \overrightarrow{\boldsymbol{Y}}_{i,G}^{(n)}\right), \\ &\leqslant \dim \overleftarrow{\boldsymbol{Y}}_{i}^{(n)} + \dim \overrightarrow{\boldsymbol{Y}}_{i,G}^{(n)}. \end{aligned} \tag{53}$$

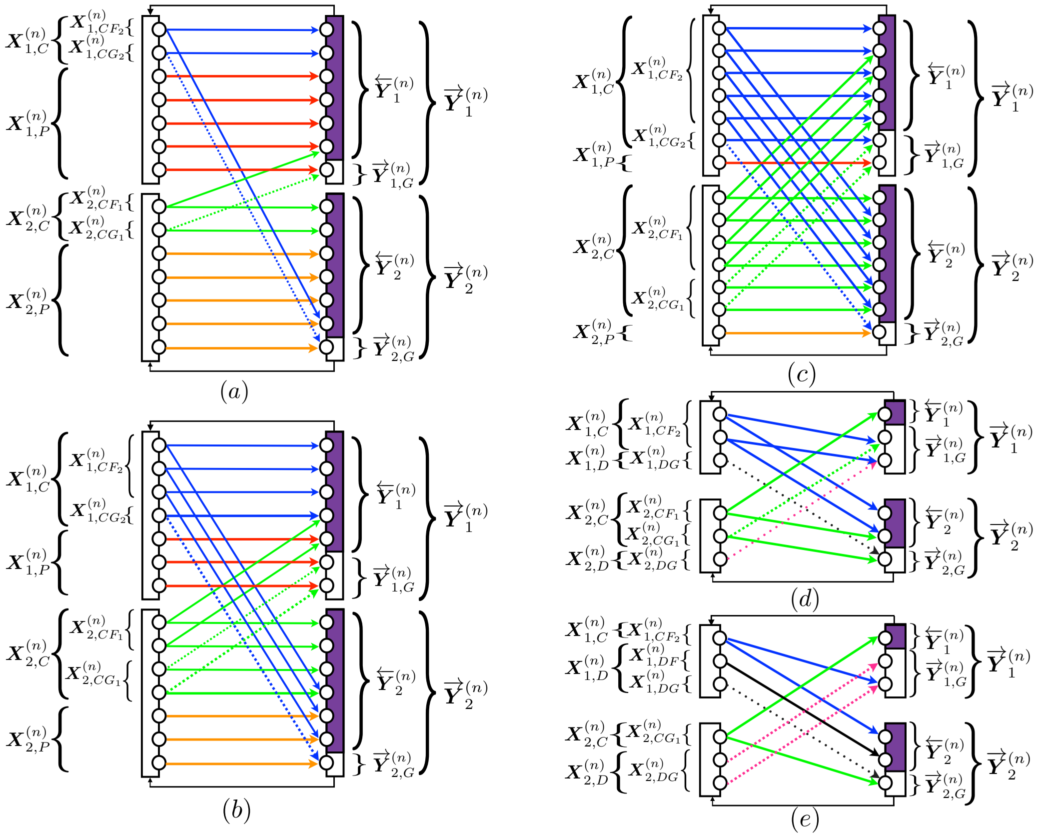

Figure 21: Different components of codewords when channel-output feedback is considered in (a) very weak interference regime, (b) weak interference regime, (c) moderate interference regime, (d) strong interference regime and (e) very strong interference regime.

The dimension of $\overrightarrow{\boldsymbol{Y}}_i^{(n)}$ satisfies

$$
\begin{aligned}
\dim \overrightarrow{\boldsymbol{Y}}_i^{(n)} =& \dim \overleftarrow{\boldsymbol{Y}}_i^{(n)} + \overrightarrow{\boldsymbol{Y}}_{i,G}^{(n)} + \overrightarrow{\boldsymbol{Y}}_{i,Q}^{(n)} \\
=& \min\left(\overleftarrow{n}_{ii}, \max\left(\overrightarrow{n}_{ii}, n_{ij}\right)\right) + \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^+ + q - \max\left(\overrightarrow{n}_{ii}, n_{ij}\right) \\
=& \min\left(\overleftarrow{n}_{ii}, \max\left(\overrightarrow{n}_{ii}, n_{ij}\right)\right) + \max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \min\left(\overleftarrow{n}_{ii}, \max\left(\overrightarrow{n}_{ii}, n_{ij}\right)\right) + q \\
& - \max\left(\overrightarrow{n}_{ii}, n_{ij}\right) \\
=& q.
\end{aligned} \tag{54}
$$

In the proofs of (36) and (37), the vector $\left(\boldsymbol{X}_{i,CF_j}^{(n)}, \boldsymbol{X}_{i,DF}^{(n)}\right)$, $i \in \{1,2\}$ is used to represent the signal levels of vector $\boldsymbol{X}_i^{(n)}$ that are seen at the receiver j and are included into the feedback from the receiver j to transmitter j.
The dimension of vector $\left(\boldsymbol{X}_{i,CF_j}^{(n)}, \boldsymbol{X}_{i,DF}^{(n)}\right)$ is

$$
\dim\left(\boldsymbol{X}_{i,CF_j}^{(n)}, \boldsymbol{X}_{i,DF}^{(n)}\right) = \left(\overleftarrow{n}_{jj} - \left(\overrightarrow{n}_{jj} - n_{ji}\right)^+\right)^+; \tag{55}
$$

Proof of (34)**:** first, consider $n_{ji} \leqslant \overrightarrow{n}_{ii}$, i.e., vector $\boldsymbol{X}_{i,P}^{(n)}$ exists and vector $\boldsymbol{X}_{i,D}^{(n)}$ does not exist. Then, the following holds

$$
\begin{aligned}
NR_i &= H\left(W_i\right), \\
&\stackrel{(a)}{=} H\left(W_i|W_j\right), \\
&= I\left(W_i; \boldsymbol{X}_i^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right) + H\left(W_i|W_j, \boldsymbol{X}_i^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right), \\
&\stackrel{(b)}{=} I\left(W_i; \boldsymbol{X}_i^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right) + H\left(W_i|W_j, \boldsymbol{X}_i^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}, \boldsymbol{X}_j^{(1:N)}\right), \\
&\stackrel{(c)}{=} I\left(W_i; \boldsymbol{X}_i^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right) + H\left(W_i|W_j, \boldsymbol{X}_i^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}, \boldsymbol{X}_j^{(1:N)}, \overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right), \\
&\stackrel{(d)}{\leqslant} I\left(W_i; \boldsymbol{X}_i^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right) + N\delta(N), \\
&= H\left(\boldsymbol{X}_i^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right) - H\left(\boldsymbol{X}_i^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_i, W_j\right) + N\delta(N), \\
&= \sum_{n=1}^{N} \left[H\left(\boldsymbol{X}_i^{(n)}, \overleftarrow{\boldsymbol{Y}}_j^{(n)}|W_j, \boldsymbol{X}_i^{(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:n-1)}\right) - H\left(\boldsymbol{X}_i^{(n)}, \overleftarrow{\boldsymbol{Y}}_j^{(n)}|W_i, W_j, \boldsymbol{X}_i^{(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:n-1)}\right)\right] \\
&\quad + N\delta(N), \\
&\stackrel{(e)}{=} \sum_{n=1}^{N} \left[H\left(\boldsymbol{X}_i^{(n)}, \overleftarrow{\boldsymbol{Y}}_j^{(n)}|W_j, \boldsymbol{X}_i^{(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:n-1)}, \boldsymbol{X}_j^{(1:n)}\right)\right. \\
&\quad \left. - H\left(\boldsymbol{X}_i^{(n)}, \overleftarrow{\boldsymbol{Y}}_j^{(n)}|W_i, W_j, \boldsymbol{X}_i^{(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:n-1)}, \boldsymbol{X}_j^{(1:n)}\right)\right] + N\delta(N), \\
&\stackrel{(f)}{\leqslant} \sum_{n=1}^{N} \left[H\left(\boldsymbol{X}_i^{(n)}, \overleftarrow{\boldsymbol{Y}}_j^{(n)}|\boldsymbol{X}_j^{(1:n)}\right) - H\left(\boldsymbol{X}_i^{(n)}, \overleftarrow{\boldsymbol{Y}}_j^{(n)}|W_i, W_j, \boldsymbol{X}_i^{(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:n-1)}, \boldsymbol{X}_j^{(1:n)}\right)\right] \\
&\quad + N\delta(N), \\
&\stackrel{(g)}{=} \sum_{n=1}^{N} \left[H\left(\boldsymbol{X}_i^{(n)}, \boldsymbol{X}_{i,CF_j}^{(n)}|\boldsymbol{X}_j^{(1:n)}\right)\right. \\
&\quad \left. - H\left(\boldsymbol{X}_i^{(n)}, \boldsymbol{X}_{i,CF_j}^{(n)}|W_i, W_j, \boldsymbol{X}_i^{(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:n-1)}, \boldsymbol{X}_j^{(1:n)}\right)\right] + N\delta(N), \\
&\stackrel{(h)}{=} \sum_{n=1}^{N} \left[H\left(\boldsymbol{X}_i^{(n)}|\boldsymbol{X}_j^{(1:n)}\right)\right. \\
&\quad \left. - H\left(\boldsymbol{X}_i^{(n)}|W_i, W_j, \boldsymbol{X}_i^{(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:n-1)}, \boldsymbol{X}_j^{(1:n)}\right)\right] + N\delta(N), \\
&\stackrel{(i)}{=} \sum_{n=1}^{N} \left[H\left(\boldsymbol{X}_i^{(n)}|\boldsymbol{X}_j^{(1:n)}\right)\right. \\
&\quad \left. - H\left(\boldsymbol{X}_i^{(n)}|W_i, W_j, \boldsymbol{X}_i^{(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:n-1)}, \boldsymbol{X}_j^{(1:n)}, \overrightarrow{\boldsymbol{Y}}_i^{(1:n-1)}\right)\right] + N\delta(N), \\
&\stackrel{(j)}{=} \sum_{n=1}^{N} \left[H\left(\boldsymbol{X}_i^{(n)}|\boldsymbol{X}_j^{(1:n)}\right)\right. \\
&\quad \left. - H\left(\boldsymbol{X}_i^{(n)}|W_i, W_j, \boldsymbol{X}_i^{(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:n-1)}, \boldsymbol{X}_j^{(1:n)}, \overrightarrow{\boldsymbol{Y}}_i^{(1:n-1)}, \boldsymbol{X}_i^{(1:n)}\right)\right] + N\delta(N),
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{n=1}^{N} H\left(\boldsymbol{X}_i^{(n)} | \boldsymbol{X}_j^{(1:n)}\right) + N\delta(N),\\
&= NH\left(\boldsymbol{X}_i^{(n)} | \boldsymbol{X}_j^{(n)}\right) + N\delta(N), \text{ for any } n \in \{1,\dots,N\},\\
&\overset{(k)}{\leqslant} NH\left(\boldsymbol{X}_i^{(n)}\right) + N\delta(N),
\end{aligned} \tag{56}
$$

where
(a) follows from the fact that W_i and W_j are independent, $i \in \{1,2\}$ and $j \in \{1,2\}\backslash\{i\}$;
(b) follows from the fact that $\boldsymbol{X}_j^{(1:N)} = f\left(W_j, \overleftarrow{\boldsymbol{Y}}_j^{(1:N-1)}\right)$;
(c) follows from the fact that $\overrightarrow{\boldsymbol{Y}}_i^{(1:N)} = f\left(\boldsymbol{X}_i^{(1:N)}, \boldsymbol{X}_j^{(1:N)}\right)$;
(d) follows from Fano's inequality;
(e) follows from the fact that $\boldsymbol{X}_j^{(1:n)} = f\left(W_j, \overleftarrow{\boldsymbol{Y}}_j^{(1:n-1)}\right)$;
(f) follows from the fact that conditioning reduces the entropy;
(g) follows from the fact that $\overleftarrow{\boldsymbol{Y}}_j^{(n)} = f\left(\boldsymbol{X}_{i,CF_j}^{(n)}, \boldsymbol{X}_{j,F_j}^{(n)}\right)$ where $\boldsymbol{X}_{j,F_j}^{(n)}$ is contained into $\boldsymbol{X}_j^{(n)}$;
(h) follows from the fact that $\boldsymbol{X}_{i,CF_j}^{(n)}$ and $\boldsymbol{X}_{i,DF}^{(n)}$ are contained into $\boldsymbol{X}_i^{(n)}$;
(i) follows from the fact that $\overrightarrow{\boldsymbol{Y}}_i^{(1:n-1)} = f\left(\boldsymbol{X}_i^{(1:n-1)}, \boldsymbol{X}_j^{(1:n-1)}\right)$;
(j) follows from the fact that that $\boldsymbol{X}_i^{(1:n)} = f\left(W_i, \overleftarrow{\boldsymbol{Y}}_i^{(1:n-1)}\right)$ and $\overleftarrow{\boldsymbol{Y}}_i^{(1:n-1)}$ is included into $\overrightarrow{\boldsymbol{Y}}_i^{(1:n-1)}$;
(k) follows from the fact that conditioning reduces the entropy.

From (56), in the asymptotic regime, the following holds

$$
\begin{aligned}
R_i \leqslant & H\left(\boldsymbol{X}_i^{(n)}\right)\\
=& H\left(\boldsymbol{X}_{i,C}^{(n)}\right) + H\left(\boldsymbol{X}_{i,P}^{(n)}\right)\\
\leqslant & \dim \boldsymbol{X}_{i,C}^{(n)} + \dim \boldsymbol{X}_{i,P}^{(n)}.
\end{aligned} \tag{57}
$$

A tighter bound can be obtained for the case in which $n_{ji} > \overrightarrow{n}_{ii}$. In this case the vector $\boldsymbol{X}_{i,P}^{(n)}$ and the vector $\boldsymbol{X}_{i,D}^{(n)}$ exists. Hence, the following holds

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

$$
\begin{aligned}
&\overset{(e)}{=}\sum_{n=1}^{N}\Big[H\left(\boldsymbol{X}_{i,C}^{(n)},\overleftarrow{\boldsymbol{Y}}_{j}^{(n)}|W_j,\boldsymbol{X}_{i,C}^{(1:n-1)},\overleftarrow{\boldsymbol{Y}}_{j}^{(1:n-1)},\boldsymbol{X}_{j}^{(1:n)}\right)\\
&\quad -H\left(\boldsymbol{X}_{i,C}^{(n)},\overleftarrow{\boldsymbol{Y}}_{j}^{(n)}|W_i,W_j,\boldsymbol{X}_{i,C}^{(1:n-1)},\overleftarrow{\boldsymbol{Y}}_{j}^{(1:n-1)},\boldsymbol{X}_{j}^{(1:n)}\right)\Big]+N\delta(N),\\
&\overset{(f)}{\leqslant}\sum_{n=1}^{N}\Big[H\left(\boldsymbol{X}_{i,C}^{(n)},\overleftarrow{\boldsymbol{Y}}_{j}^{(n)}|\boldsymbol{X}_{j}^{(1:n)}\right)-H\left(\boldsymbol{X}_{i,C}^{(n)},\overleftarrow{\boldsymbol{Y}}_{j}^{(n)}|W_i,W_j,\boldsymbol{X}_{i,C}^{(1:n-1)},\overleftarrow{\boldsymbol{Y}}_{j}^{(1:n-1)},\boldsymbol{X}_{j}^{(1:n)}\right)\Big]\\
&\quad +N\delta(N),\\
&\overset{(g)}{=}\sum_{n=1}^{N}\Big[H\left(\boldsymbol{X}_{i,C}^{(n)},\boldsymbol{X}_{i,CF_j}^{(n)},\boldsymbol{X}_{i,DF}^{(n)}|\boldsymbol{X}_{j}^{(1:n)}\right)\\
&\quad -H\left(\boldsymbol{X}_{i,C}^{(n)},\boldsymbol{X}_{i,CF_j}^{(n)},\boldsymbol{X}_{i,DF}^{(n)}|W_i,W_j,\boldsymbol{X}_{i,C}^{(1:n-1)},\overleftarrow{\boldsymbol{Y}}_{j}^{(1:n-1)},\boldsymbol{X}_{j}^{(1:n)}\right)\Big]+N\delta(N),\\
&\overset{(h)}{=}\sum_{n=1}^{N}\Big[H\left(\boldsymbol{X}_{i,C}^{(n)},\boldsymbol{X}_{i,DF}^{(n)}|\boldsymbol{X}_{j}^{(1:n)}\right)\\
&\quad -H\left(\boldsymbol{X}_{i,C}^{(n)},\boldsymbol{X}_{i,DF}^{(n)}|W_i,W_j,\boldsymbol{X}_{i,C}^{(1:n-1)},\overleftarrow{\boldsymbol{Y}}_{j}^{(1:n-1)},\boldsymbol{X}_{j}^{(1:n)}\right)\Big]+N\delta(N),\\
&\overset{(i)}{=}\sum_{n=1}^{N}\Big[H\left(\boldsymbol{X}_{i,C}^{(n)},\boldsymbol{X}_{i,DF}^{(n)}|\boldsymbol{X}_{j}^{(1:n)}\right)\\
&\quad -H\left(\boldsymbol{X}_{i,C}^{(n)},\boldsymbol{X}_{i,DF}^{(n)}|W_i,W_j,\boldsymbol{X}_{i,C}^{(1:n-1)},\overleftarrow{\boldsymbol{Y}}_{j}^{(1:n-1)},\boldsymbol{X}_{j}^{(1:n)},\overrightarrow{\boldsymbol{Y}}_{i}^{(1:n-1)}\right)\Big]+N\delta(N),\\
&\overset{(j)}{=}\sum_{n=1}^{N}\Big[H\left(\boldsymbol{X}_{i,C}^{(n)},\boldsymbol{X}_{i,DF}^{(n)}|\boldsymbol{X}_{j}^{(1:n)}\right)\\
&\quad -H\left(\boldsymbol{X}_{i,C}^{(n)},\boldsymbol{X}_{i,DF}^{(n)}|W_i,W_j,\boldsymbol{X}_{i,C}^{(1:n-1)},\overleftarrow{\boldsymbol{Y}}_{j}^{(1:n-1)},\boldsymbol{X}_{j}^{(1:n)},\overrightarrow{\boldsymbol{Y}}_{i}^{(1:n-1)},\boldsymbol{X}_{i}^{(1:n)}\right)\Big]+N\delta(N),\\
&=\sum_{n=1}^{N}H\left(\boldsymbol{X}_{i,C}^{(n)},\boldsymbol{X}_{i,DF}^{(n)}|\boldsymbol{X}_{j}^{(1:n)}\right)+N\delta(N),\\
&=NH\left(\boldsymbol{X}_{i,C}^{(n)},\boldsymbol{X}_{i,DF}^{(n)}|\boldsymbol{X}_{j}^{(n)}\right)+N\delta(N),\ \text{for any } n\in\{1,\ldots,N\},\\
&\overset{(k)}{\leqslant}NH\left(\boldsymbol{X}_{i,C}^{(n)},\boldsymbol{X}_{i,DF}^{(n)}\right)+N\delta(N),
\end{aligned}
\tag{58}
$$

$$
\begin{aligned}
&=NH\left(\boldsymbol{X}_{i,C}^{(n)}\right)+NH\left(\boldsymbol{X}_{i,DF}^{(n)}|\boldsymbol{X}_{i,C}^{(n)}\right)+N\delta(N),\\
&\leqslant NH\left(\boldsymbol{X}_{i,C}^{(n)}\right)+NH\left(\boldsymbol{X}_{i,DF}^{(n)}\right)+N\delta(N),
\end{aligned}
\tag{59}
$$

where
(a) follows from the fact that W_i and W_j are independent, $i \in \{1,2\}$ and $j \in \{1,2\}\backslash\{i\}$;
(b) follows from the fact that $\boldsymbol{X}_j^{(1:N)} = f\left(W_j, \overleftarrow{\boldsymbol{Y}}_j^{(1:N-1)}\right)$;
(c) follows from the fact that $\overrightarrow{\boldsymbol{Y}}_i^{(1:N)} = f\left(\boldsymbol{X}_{i,C}^{(1:N)}, \boldsymbol{X}_j^{(1:N)}\right)$ when $n_{ji} > \overrightarrow{n}_{ii}$;
(d) follows from Fano's inequality;
(e) follows from the fact that $\boldsymbol{X}_j^{(1:n)} = f\left(W_j, \overleftarrow{\boldsymbol{Y}}_j^{(1:n-1)}\right)$;
(f) follows from the fact that conditioning reduces the entropy;

(g) follows from the fact that $\overleftarrow{\boldsymbol{Y}}_j^{(n)} = f\left(\boldsymbol{X}_{i,CF_j}^{(n)}, \boldsymbol{X}_{i,DF_j}^{(n)}, \boldsymbol{X}_{j,F_j}^{(n)}\right)$ where $\boldsymbol{X}_{j,F_j}^{(n)}$ is contained into $\boldsymbol{X}_j^{(n)}$;
(h) follows from the fact that $\boldsymbol{X}_{i,CF_j}^{(n)}$ is contained into $\boldsymbol{X}_{i,C}^{(n)}$;
(i) follows from the fact that $\overrightarrow{\boldsymbol{Y}}_i^{(1:n-1)} = f\left(\boldsymbol{X}_{i,C}^{(1:n-1)}, \boldsymbol{X}_j^{(1:n-1)}\right)$ when $n_{ji} > \overrightarrow{n}_{ii}$;
(j) follows from the fact that that $\boldsymbol{X}_i^{(1:n)} = f\left(W_i, \overleftarrow{\boldsymbol{Y}}_i^{(1:n-1)}\right)$ and $\overleftarrow{\boldsymbol{Y}}_i^{(1:n-1)}$ is included into $\overrightarrow{\boldsymbol{Y}}_i^{(1:n-1)}$;
(k) follows from the fact that conditioning reduces the entropy.

From (59), in the asymptotic regime, it follows that:

$$\begin{aligned} R_i \leqslant & H\left(\boldsymbol{X}_{i,C}^{(n)}\right) + H\left(\boldsymbol{X}_{i,DF}^{(n)}\right) \\ \leqslant & \dim \boldsymbol{X}_{i,C}^{(n)} + \dim \boldsymbol{X}_{i,DF}^{(n)}. \end{aligned} \tag{60}$$

Inequality (57) is valid when $n_{ji} \leqslant \overrightarrow{n}_{ii}$. Inequality (60) is valid when $n_{ji} > \overrightarrow{n}_{ii}$. It is worth noting that in the case in which $n_{ji} \leqslant \overrightarrow{n}_{ii}$, $H\left(\boldsymbol{X}_{i,D}^{(n)}\right) = 0$; and in the other case, i.e., $n_{ji} > \overrightarrow{n}_{ii}$, $H\left(\boldsymbol{X}_{i,P}^{(n)}\right) = 0$. Then, inequalities (57) and (60) can be expressed as one inequality as follows

$$\begin{aligned} R_i \leqslant & H\left(\boldsymbol{X}_{i,C}^{(n)}\right) + H\left(\boldsymbol{X}_{i,P}^{(n)}\right) + H\left(\boldsymbol{X}_{i,DF}^{(n)}\right), \\ \leqslant & \dim \boldsymbol{X}_{i,C}^{(n)} + \dim \boldsymbol{X}_{i,P}^{(n)} + \dim \boldsymbol{X}_{i,DF}^{(n)}. \end{aligned} \tag{61}$$

Plugging (38), (39) and (48) in (61), it yields

$$\begin{aligned} R_i \leqslant & \min\left(\overrightarrow{n}_{ii}, n_{ji}\right) + \left(\overrightarrow{n}_{ii} - n_{ji}\right)^+ + \min\Big(\left(n_{ji} - \overrightarrow{n}_{ii}\right)^+, \Big(\overleftarrow{n}_{jj} - \overrightarrow{n}_{ii} - \min\left(\left(\overrightarrow{n}_{jj} - n_{ji}\right)^+, n_{ij}\right) \\ & - \left(\left(\overrightarrow{n}_{jj} - n_{ij}\right)^+ - n_{ji}\right)^+\Big)^+\Big), \\ = & \overrightarrow{n}_{ii} + \min\Big(\left(n_{ji} - \overrightarrow{n}_{ii}\right)^+, \Big(\overleftarrow{n}_{jj} - \overrightarrow{n}_{ii} - \min\left(\left(\overrightarrow{n}_{jj} - n_{ji}\right)^+, n_{ij}\right) \\ & - \left(\left(\overrightarrow{n}_{jj} - n_{ij}\right)^+ - n_{ji}\right)^+\Big)^+\Big), \end{aligned} \tag{62}$$

and this completes the proof of (34). ■

Proof of (36)**:**

$$
\begin{aligned}
N\left(R_1+R_2\right) &= H\left(W_1\right)+H\left(W_2\right),\\
&= I\left(W_1;\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}\right)+H\left(W_1|\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}\right)+I\left(W_2;\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}\right)+H\left(W_2|\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}\right),\\
&\overset{(a)}{\leqslant} I\left(W_1;\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}\right)+I\left(W_2;\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}\right)+N\delta_1(N)+N\delta_2(N),\\
&\overset{(b)}{=} I\left(W_1;\overrightarrow{\boldsymbol{Y}}_1^{(1:N)},\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}\right)+I\left(W_2;\overrightarrow{\boldsymbol{Y}}_2^{(1:N)},\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}\right)+N\delta(N),\\
&= H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)},\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}\right)-H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)},\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)},\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}\right)\\
&\quad -H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)},\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)+N\delta(N),\\
&= H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}\right)+H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}\right)-H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)-H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}\right)\\
&\quad +H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}\right)+H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}\right)-H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)-H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}\right)\\
&\quad +N\delta(N),\\
&\overset{(b)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}\right)-H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)-H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}\right)\\
&\quad -H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)-H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}\right)+N\delta(N),\\
&\overset{(c)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}\right)-H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)-H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)},\boldsymbol{X}_1^{(1:N)}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}\right)\\
&\quad -H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)-H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)},\boldsymbol{X}_2^{(1:N)}\right)+N\delta(N),\\
&= H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}\right)-H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)-H\left(\boldsymbol{X}_{2,C}^{(1:N)}|W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)},\boldsymbol{X}_1^{(1:N)}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}\right)\\
&\quad -H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)-H\left(\boldsymbol{X}_{1,C}^{(1:N)}|W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)},\boldsymbol{X}_2^{(1:N)}\right)+N\delta(N),\\
&\overset{(d)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}\right)-H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)-H\left(\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)}|W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)},\boldsymbol{X}_1^{(1:N)}\right)\\
&\quad +H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}\right)-H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)-H\left(\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)}|W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)},\boldsymbol{X}_2^{(1:N)}\right)\\
&\quad +N\delta(N),\\
&= H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}\right)-H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)-H\left(\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)}|W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}\right)\\
&\quad -H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)-H\left(\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)}|W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}\right)+N\delta(N),\\
&= H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}\right)+\left[I\left(\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)};W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}\right)-H\left(\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)}\right)\right]\\
&\quad +H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}\right)+\left[I\left(\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)};W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}\right)-H\left(\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)}\right)\right]\\
&\quad -H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)-H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)+N\delta(N),\\
&\overset{(e)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}|\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)}\right)-H\left(\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)}|\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}\right)\\
&\quad +H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}|\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)}\right)-H\left(\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)}|\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}\right)\\
&\quad +I\left(\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)};W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}\right)+I\left(\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)};W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}\right)\\
&\quad -H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)-H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)+N\delta(N),\\
&\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}|\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}|\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)}\right)\\
&\quad +I\left(\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)};W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}\right)+I\left(\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)};W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}\right)\\
&\quad -H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)-H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)+N\delta(N),
\end{aligned}
$$

$$
\begin{aligned}
&\overset{(f)}{\leqslant} H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}|\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}|\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)}\right)\\
&\quad+I\left(\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)},W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)};W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}\right)+I\left(\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)},W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)};W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}\right)\\
&\quad-H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)-H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)+N\delta(N),\\
&=H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}|\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}|\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)}\right)+I\left(W_2;W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}\right)\\
&\quad+I\left(\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)},\overleftarrow{\boldsymbol{Y}}_2^{(1:N)};W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_2\right)+I\left(W_1;W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}\right)\\
&\quad+I\left(\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)},\overleftarrow{\boldsymbol{Y}}_1^{(1:N)};W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_1\right)-H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)-H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)+N\delta(N),\\
&=H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}|\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}|\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)}\right)+H\left(W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}\right)\\
&\quad-H\left(W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_2\right)+H\left(\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)},\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)\\
&\quad-H\left(\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)},\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2,W_1,\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}\right)+H\left(W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}\right)-H\left(W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_1\right)\\
&\quad+H\left(\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)},\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)-H\left(\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)},\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1,W_2,\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}\right)\\
&\quad-H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)-H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)+N\delta(N),\\
&\overset{(g)}{=}H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}|\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}|\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)}\right)+H\left(W_1\right)+H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)\\
&\quad-H\left(W_1|W_2\right)-H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_2,W_1\right)+H\left(\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)},\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)\\
&\quad+H\left(W_2\right)+H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)-H\left(W_2|W_1\right)-H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_1,W_2\right)\\
&\quad+H\left(\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)},\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)-H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)-H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)+N\delta(N),\\
&\overset{(h)}{=}H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}|\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}|\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)}\right)-H\left(\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_2,W_1\right)\\
&\quad+H\left(\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)},\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)-H\left(\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_1,W_2\right)+H\left(\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)},\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)\\
&\quad+N\delta(N),\\
&\overset{(i)}{\leqslant}H\left(\overrightarrow{\boldsymbol{Y}}_1^{(1:N)}|\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2^{(1:N)}|\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)}\right)+H\left(\boldsymbol{X}_{2,C}^{(1:N)},\boldsymbol{X}_{1,U}^{(1:N)},\overleftarrow{\boldsymbol{Y}}_2^{(1:N)}|W_2\right)\\
&\quad+H\left(\boldsymbol{X}_{1,C}^{(1:N)},\boldsymbol{X}_{2,U}^{(1:N)},\overleftarrow{\boldsymbol{Y}}_1^{(1:N)}|W_1\right)+N\delta(N),\\
&=\sum_{n=1}^{N}\Big[H\left(\overrightarrow{\boldsymbol{Y}}_1^{(n)}|\boldsymbol{X}_{1,C}^{(n)},\boldsymbol{X}_{2,U}^{(n)},\overrightarrow{\boldsymbol{Y}}_1^{(1:n-1)}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2^{(n)}|\boldsymbol{X}_{2,C}^{(n)},\boldsymbol{X}_{1,U}^{(n)},\overrightarrow{\boldsymbol{Y}}_2^{(1:n-1)}\right)\\
&\quad+H\left(\boldsymbol{X}_{2,C}^{(n)},\boldsymbol{X}_{1,U}^{(n)},\overleftarrow{\boldsymbol{Y}}_2^{(n)}|W_2,\boldsymbol{X}_{2,C}^{(1:n-1)},\boldsymbol{X}_{1,U}^{(1:n-1)},\overleftarrow{\boldsymbol{Y}}_2^{(1:n-1)}\right)\\
&\quad+H\left(\boldsymbol{X}_{1,C}^{(n)},\boldsymbol{X}_{2,U}^{(n)},\overleftarrow{\boldsymbol{Y}}_1^{(n)}|W_1,\boldsymbol{X}_{1,C}^{(1:n-1)},\boldsymbol{X}_{2,U}^{(1:n-1)},\overleftarrow{\boldsymbol{Y}}_1^{(1:n-1)}\right)\Big]+N\delta(N),\\
&\overset{(j)}{=}\sum_{n=1}^{N}\Big[H\left(\overrightarrow{\boldsymbol{Y}}_1^{(n)}|\boldsymbol{X}_{1,C}^{(n)},\boldsymbol{X}_{2,U}^{(n)},\overrightarrow{\boldsymbol{Y}}_1^{(1:n-1)}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2^{(n)}|\boldsymbol{X}_{2,C}^{(n)},\boldsymbol{X}_{1,U}^{(n)},\overrightarrow{\boldsymbol{Y}}_2^{(1:n-1)}\right)\\
&\quad+H\left(\boldsymbol{X}_{2,C}^{(n)},\boldsymbol{X}_{1,U}^{(n)},\overleftarrow{\boldsymbol{Y}}_2^{(n)}|W_2,\boldsymbol{X}_{2,C}^{(1:n-1)},\boldsymbol{X}_{1,U}^{(1:n-1)},\overleftarrow{\boldsymbol{Y}}_2^{(1:n-1)},\boldsymbol{X}_2^{(1:n)}\right)\\
&\quad+H\left(\boldsymbol{X}_{1,C}^{(n)},\boldsymbol{X}_{2,U}^{(n)},\overleftarrow{\boldsymbol{Y}}_1^{(n)}|W_1,\boldsymbol{X}_{1,C}^{(1:n-1)},\boldsymbol{X}_{2,U}^{(1:n-1)},\overleftarrow{\boldsymbol{Y}}_1^{(1:n-1)},\boldsymbol{X}_1^{(1:n)}\right)\Big]+N\delta(N),\\
&=\sum_{n=1}^{N}\Big[H\left(\overrightarrow{\boldsymbol{Y}}_1^{(n)}|\boldsymbol{X}_{1,C}^{(n)},\boldsymbol{X}_{2,U}^{(n)},\overrightarrow{\boldsymbol{Y}}_1^{(1:n-1)}\right)+H\left(\overrightarrow{\boldsymbol{Y}}_2^{(n)}|\boldsymbol{X}_{2,C}^{(n)},\boldsymbol{X}_{1,U}^{(n)},\overrightarrow{\boldsymbol{Y}}_2^{(1:n-1)}\right)\\
&\quad+H\left(\boldsymbol{X}_{1,U}^{(n)},\overleftarrow{\boldsymbol{Y}}_2^{(n)}|\boldsymbol{X}_{1,U}^{(1:n-1)},\boldsymbol{X}_2^{(1:n)}\right)+H\left(\boldsymbol{X}_{2,U}^{(n)},\overleftarrow{\boldsymbol{Y}}_1^{(n)}|\boldsymbol{X}_{2,U}^{(1:n-1)},\boldsymbol{X}_1^{(1:n)}\right)\Big]+N\delta(N),
\end{aligned}
$$

$$
\begin{aligned}
&\overset{(k)}{\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^{(1:n)}\right)\\
&\quad+H\left(\boldsymbol{X}_{2,U}^{(n)},\overleftarrow{\boldsymbol{Y}}_1^{(n)}|\boldsymbol{X}_1^{(1:n)}\right)\Big]+N\delta(N),\\
&=\sum_{n=1}^{N}\Big[H\left(\boldsymbol{X}_{1,P}^{(n)}|\boldsymbol{X}_{1,C}^{(n)},\boldsymbol{X}_{2,U}^{(n)}\right)+H\left(\boldsymbol{X}_{2,P}^{(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^{(1:n)}\right)\\
&\quad+H\left(\boldsymbol{X}_{2,U}^{(n)},\overleftarrow{\boldsymbol{Y}}_1^{(n)}|\boldsymbol{X}_1^{(1:n)}\right)\Big]+N\delta(N),\\
&\overset{(k)}{\leqslant}\sum_{n=1}^{N}\Big[H\left(\boldsymbol{X}_{1,P}^{(n)}\right)+H\left(\boldsymbol{X}_{2,P}^{(n)}\right)+H\left(\boldsymbol{X}_{1,U}^{(n)},\overleftarrow{\boldsymbol{Y}}_2^{(n)}|\boldsymbol{X}_2^{(1:n)}\right)+H\left(\boldsymbol{X}_{2,U}^{(n)},\overleftarrow{\boldsymbol{Y}}_1^{(n)}|\boldsymbol{X}_1^{(1:n)}\right)\Big]\\
&\quad+N\delta(N),\\
&=N\Big[H\left(\boldsymbol{X}_{1,P}^{(n)}\right)+H\left(\boldsymbol{X}_{2,P}^{(n)}\right)+H\left(\boldsymbol{X}_{1,U}^{(n)},\overleftarrow{\boldsymbol{Y}}_2^{(n)}|\boldsymbol{X}_2^{(n)}\right)+H\left(\boldsymbol{X}_{2,U}^{(n)},\overleftarrow{\boldsymbol{Y}}_1^{(n)}|\boldsymbol{X}_1^{(n)}\right)\Big]\\
&\quad+N\delta(N),\ \text{for any } n\in\{1,\ldots,N\},\\
&=N\Big[H\left(\boldsymbol{X}_{1,P}^{(n)}\right)+H\left(\boldsymbol{X}_{2,P}^{(n)}\right)+H\left(\boldsymbol{X}_{1,U}^{(n)}|\boldsymbol{X}_2^{(n)}\right)+H\left(\overleftarrow{\boldsymbol{Y}}_2^{(n)}|\boldsymbol{X}_2^{(n)},\boldsymbol{X}_{1,U}^{(n)}\right)+H\left(\boldsymbol{X}_{2,U}^{(n)}|\boldsymbol{X}_1^{(n)}\right)\\
&\quad+H\left(\overleftarrow{\boldsymbol{Y}}_1^{(n)}|\boldsymbol{X}_1^{(n)},\boldsymbol{X}_{2,U}^{(n)}\right)\Big]+N\delta(N),\\
&\overset{(k)}{\leqslant}N\Big[H\left(\boldsymbol{X}_{1,P}^{(n)}\right)+H\left(\boldsymbol{X}_{2,P}^{(n)}\right)+H\left(\boldsymbol{X}_{1,U}^{(n)}\right)+H\left(\overleftarrow{\boldsymbol{Y}}_2^{(n)}|\boldsymbol{X}_2^{(n)},\boldsymbol{X}_{1,U}^{(n)}\right)+H\left(\boldsymbol{X}_{2,U}^{(n)}\right)\\
&\quad+H\left(\overleftarrow{\boldsymbol{Y}}_1^{(n)}|\boldsymbol{X}_1^{(n)},\boldsymbol{X}_{2,U}^{(n)}\right)\Big]+N\delta(N),\\
&=N\Big[H\left(\boldsymbol{X}_{1,P}^{(n)}\right)+H\left(\boldsymbol{X}_{2,P}^{(n)}\right)+H\left(\boldsymbol{X}_{1,U}^{(n)}\right)+H\left(\boldsymbol{X}_{1,CF_2}^{(n)},\boldsymbol{X}_{1,DF}^{(n)}|\boldsymbol{X}_2^{(n)},\boldsymbol{X}_{1,U}^{(n)}\right)+H\left(\boldsymbol{X}_{2,U}^{(n)}\right)\\
&\quad+H\left(\boldsymbol{X}_{2,CF_1}^{(n)},\boldsymbol{X}_{2,DF}^{(n)}|\boldsymbol{X}_1^{(n)},\boldsymbol{X}_{2,U}^{(n)}\right)\Big]+N\delta(N),\\
&=N\Big[H\left(\boldsymbol{X}_{1,P}^{(n)}\right)+H\left(\boldsymbol{X}_{2,P}^{(n)}\right)+H\left(\boldsymbol{X}_{1,U}^{(n)}\right)+H\left(\boldsymbol{X}_{1,CF_2}^{(n)},\boldsymbol{X}_{1,DF}^{(n)}|\boldsymbol{X}_{1,U}^{(n)}\right)+H\left(\boldsymbol{X}_{2,U}^{(n)}\right)\\
&\quad+H\left(\boldsymbol{X}_{2,CF_1}^{(n)},\boldsymbol{X}_{2,DF}^{(n)}|\boldsymbol{X}_{2,U}^{(n)}\right)\Big]+N\delta(N),
\end{aligned}\tag{63}
$$

where
(a) follows from Fano's inequality;
(b) follows from the fact that $\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}$ is included into $\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}$;
(c) follows from the fact that $\boldsymbol{X}_i^{(1:N)}=f\left(W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N-1)}\right)$;
(d) follows from the fact that $\boldsymbol{X}_{i,U}^{(1:N)}$ is included into $\boldsymbol{X}_i^{(1:N)}$;
(e) follows from the fact that $H(Y)-H(X)=H(Y|X)-H(X|Y)$;
(f) follows from the fact that injecting information increases the mutual information;
(g) follows from the fact that $H\left(\boldsymbol{X}_{i,C}^{(1:N)},\boldsymbol{X}_{j,U}^{(1:N)},\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i,W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right)=0$;
(h) follows from the fact that W_i is independent of W_j, then $H(W_i|W_j)=H(W_i)$;
(i) follows from the fact that conditioning reduces the entropy;
(j) follows from the fact that $\boldsymbol{X}_i^{(1:n)}=f\left(W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:n-1)}\right)$;
(k) follows from the fact that conditioning reduces the entropy.

From (63), in the asymptotic regime, it follows that:

$$
\begin{aligned}
R_1+R_2\leqslant & H\left(\boldsymbol{X}_{1,P}^{(n)}\right)+H\left(\boldsymbol{X}_{2,P}^{(n)}\right)+H\left(\boldsymbol{X}_{1,U}^{(n)}\right)+H\left(\boldsymbol{X}_{1,CF_2}^{(n)},\boldsymbol{X}_{1,DF}^{(n)}|\boldsymbol{X}_{1,U}^{(n)}\right)+H\left(\boldsymbol{X}_{2,U}^{(n)}\right)\\
&+H\left(\boldsymbol{X}_{2,CF_1}^{(n)},\boldsymbol{X}_{2,DF}^{(n)}|\boldsymbol{X}_{2,U}^{(n)}\right),\\
=&\dim \boldsymbol{X}_{1,P}^{(n)}+\dim \boldsymbol{X}_{2,P}^{(n)}+\dim \boldsymbol{X}_{1,U}^{(n)}+\dim\left(\boldsymbol{X}_{1,CF_2}^{(n)},\boldsymbol{X}_{1,DF}^{(n)}\right)-\dim \boldsymbol{X}_{1,U}^{(n)}+\dim \boldsymbol{X}_{2,U}^{(n)}\\
&+\dim\left(\boldsymbol{X}_{2,CF_1}^{(n)},\boldsymbol{X}_{2,DF}^{(n)}\right)-\dim \boldsymbol{X}_{2,U}^{(n)},\\
=&\dim \boldsymbol{X}_{1,P}^{(n)}+\dim \boldsymbol{X}_{2,P}^{(n)}+\dim\left(\boldsymbol{X}_{1,CF_2}^{(n)},\boldsymbol{X}_{1,DF}^{(n)}\right)+\dim\left(\boldsymbol{X}_{2,CF_1}^{(n)},\boldsymbol{X}_{2,DF}^{(n)}\right). \qquad (64)
\end{aligned}
$$

Plugging (39), (55) in (64), it yields

$$
\begin{aligned}
R_1+R_2\leqslant&(\overrightarrow{n}_{11}-n_{21})^++(\overrightarrow{n}_{22}-n_{12})^++\left(\overleftarrow{n}_{11}-(\overrightarrow{n}_{11}-n_{12})^+\right)^+\\
&+\left(\overleftarrow{n}_{22}-(\overrightarrow{n}_{22}-n_{21})^+\right)^+, \qquad (65)
\end{aligned}
$$

and this completes the proof of (36). ■

Proof of (37)**:** for all $i\in\{1,2\}$ and $j\in\{1,2\}\setminus\{i\}$, it follows that

$$
\begin{aligned}
N\left(2R_i+R_j\right)\leqslant&\, 2H\left(W_i\right)+H\left(W_j\right),\\
=&\,2I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right)+2H\left(W_i|\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right)+I\left(W_j;\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right)+H\left(W_j|\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right),\\
\overset{(a)}{\leqslant}&\,2I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right)+I\left(W_j;\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right)+2N\delta_i(N)+N\delta_j(N),\\
=&\,2I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right)+I\left(W_j;\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right)+N\delta(N),\\
=&\,I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right)+I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right)+I\left(W_j;\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right)+N\delta(N),\\
=&\,I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i^{(1:N)},\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right)+I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right)+I\left(W_j;\overrightarrow{\boldsymbol{Y}}_j^{(1:N)},\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right)\\
&+N\delta(N),\\
\overset{(b)}{\leqslant}&\,I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i^{(1:N)},\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right)+I\left(W_i;\overrightarrow{\boldsymbol{Y}}_i^{(1:N)},\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right)\\
&+I\left(W_j;\overrightarrow{\boldsymbol{Y}}_j^{(1:N)},\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right)+N\delta(N),\\
=&\,H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)},\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right)-H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)},\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right)+H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)},\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right)\\
&-H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)},\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_i,W_j\right)+H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)},\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right)-H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)},\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right)\\
&+N\delta(N),\\
=&\,H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right)+H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right)-H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right)-H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right)\\
&+H\left(\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right)+H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right)-H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)},\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_i,W_j\right)\\
&+H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right)+H\left(\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right)-H\left(\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right)\\
&-H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right)+N\delta(N),\\
=&\,H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right)+H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right)-H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right)-H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right)\\
&+H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right)-H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)},\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_i,W_j\right)+H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right)\\
&+H\left(\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right)-H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right)+N\delta(N),
\end{aligned}
$$

$$\begin{aligned}
&\overset{(c)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right) - H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right)\\
&\quad - H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)},\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_i,W_j\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right) - H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right)\\
&\quad + N\delta(N),\\
&\overset{(d)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right) - H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right)\\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right) - H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right) + N\delta(N)\\
&\overset{(e)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right) - H\left(\boldsymbol{X}_{j,C}^{(1:N)}|W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right)\\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right) - H\left(\boldsymbol{X}_{i,C}^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right) + N\delta(N),\\
&\overset{(f)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right) - H\left(\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)}|W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right)\\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right) - H\left(\boldsymbol{X}_{i,C}^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right) + N\delta(N),\\
&\overset{(g)}{\leqslant} H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right) - H\left(\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)}|W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right)\\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)},\boldsymbol{X}_{i,C}^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right) - H\left(\boldsymbol{X}_{i,C}^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right) + N\delta(N),\\
&= H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right) + \left[I\left(\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)};W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)}\right)\right]\\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)},\boldsymbol{X}_{i,C}^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right) - H\left(\boldsymbol{X}_{i,C}^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right) + N\delta(N),\\
&= H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right) + \left[I\left(\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)};W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)}\right)\right]\\
&\quad + H\left(\boldsymbol{X}_{i,C}^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)},\boldsymbol{X}_{i,C}^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right)\\
&\quad - H\left(\boldsymbol{X}_{i,C}^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)}\right) + N\delta(N),\\
&= H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right) + \left[I\left(\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)};W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)}\right)\right]\\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)},\boldsymbol{X}_{i,C}^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}\right) + N\delta(N),\\
&\overset{(g)}{\leqslant} H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right) + \left[I\left(\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)};W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)}\right)\right]\\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)},\boldsymbol{X}_{i,C}^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)},\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)}\right) + N\delta(N),\\
&\overset{(h)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right) + I\left(\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)};W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right)\\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)},\boldsymbol{X}_{i,C}^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}|\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)}\right) + N\delta(N),\\
&\overset{(i)}{\leqslant} H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right) + I\left(\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)},W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)};W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right)\\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)},\boldsymbol{X}_{i,C}^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}|\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)}\right) + N\delta(N),\\
&= H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right) + I\left(W_j;W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right)\\
&\quad + I\left(\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)},\overleftarrow{\boldsymbol{Y}}_j^{(1:N)};W_i,\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j,\overleftarrow{\boldsymbol{Y}}_j^{(1:N)},\boldsymbol{X}_{i,C}^{(1:N)}\right)\\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}|\boldsymbol{X}_{j,C}^{(1:N)},\boldsymbol{X}_{i,U}^{(1:N)}\right) + N\delta(N),
\end{aligned}$$

$$
\begin{aligned}
&= H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_i\right) + H\left(W_i, \overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(W_i, \overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j\right) \\
&\quad + H\left(\boldsymbol{X}_{j,C}^{(1:N)}, \boldsymbol{X}_{i,U}^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right) - H\left(\boldsymbol{X}_{j,C}^{(1:N)}, \boldsymbol{X}_{i,U}^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j, W_i, \overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}, \boldsymbol{X}_{i,C}^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}|\boldsymbol{X}_{j,C}^{(1:N)}, \boldsymbol{X}_{i,U}^{(1:N)}\right) + N\delta(N), \\
&= H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) + H\left(W_i\right) - H\left(W_i|W_j\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j, W_i, \right) \\
&\quad + H\left(\boldsymbol{X}_{j,C}^{(1:N)}, \boldsymbol{X}_{i,U}^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right) - H\left(\boldsymbol{X}_{j,C}^{(1:N)}, \boldsymbol{X}_{i,U}^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j, W_i, \overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}, \boldsymbol{X}_{i,C}^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}|\boldsymbol{X}_{j,C}^{(1:N)}, \boldsymbol{X}_{i,U}^{(1:N)}\right) + N\delta(N), \\
&\overset{(j)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j, W_i, \right) + H\left(\boldsymbol{X}_{j,C}^{(1:N)}, \boldsymbol{X}_{i,U}^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right) \\
&\quad - H\left(\boldsymbol{X}_{j,C}^{(1:N)}, \boldsymbol{X}_{i,U}^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j, W_i, \overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}, \boldsymbol{X}_{i,C}^{(1:N)}\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}|\boldsymbol{X}_{j,C}^{(1:N)}, \boldsymbol{X}_{i,U}^{(1:N)}\right) + N\delta(N), \\
&\overset{(k)}{=} H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) - H\left(\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j, W_i, \right) + H\left(\boldsymbol{X}_{j,C}^{(1:N)}, \boldsymbol{X}_{i,U}^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}, \boldsymbol{X}_{i,C}^{(1:N)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}|\boldsymbol{X}_{j,C}^{(1:N)}, \boldsymbol{X}_{i,U}^{(1:N)}\right) + N\delta(N), \\
&\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) + H\left(\boldsymbol{X}_{j,C}^{(1:N)}, \boldsymbol{X}_{i,U}^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}|W_j, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}, \boldsymbol{X}_{i,C}^{(1:N)}\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(1:N)}|\boldsymbol{X}_{j,C}^{(1:N)}, \boldsymbol{X}_{i,U}^{(1:N)}\right) + N\delta(N), \\
&= \sum_{n=1}^{N} \left[H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}|\overrightarrow{\boldsymbol{Y}}_i^{(1:n-1)}\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)\right. \\
&\quad \left. + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}|W_j, \overleftarrow{\boldsymbol{Y}}_j^{(1:n)}, \boldsymbol{X}_{i,C}^{(1:n)}, \overrightarrow{\boldsymbol{Y}}_i^{(1:n-1)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(n)}|\boldsymbol{X}_{j,C}^{(1:n)}, \boldsymbol{X}_{i,U}^{(1:n)}, \overrightarrow{\boldsymbol{Y}}_j^{(1:n-1)}\right)\right] \\
&\quad + N\delta(N), \\
&\overset{(l)}{\leqslant} \sum_{n=1}^{N} \left[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)\right. \\
&\quad \left. + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}|W_j, \overleftarrow{\boldsymbol{Y}}_j^{(1:n)}, \boldsymbol{X}_{i,C}^{(1:n)}, \overrightarrow{\boldsymbol{Y}}_i^{(1:n-1)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(n)}|\boldsymbol{X}_{j,C}^{(1:n)}, \boldsymbol{X}_{i,U}^{(1:n)}, \overrightarrow{\boldsymbol{Y}}_j^{(1:n-1)}\right)\right] \\
&\quad + N\delta(N), \\
&\overset{(m)}{=} \sum_{n=1}^{N} \left[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^{(1:n)}\right)\right. \\
&\quad \left. + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}|W_j, \overleftarrow{\boldsymbol{Y}}_j^{(1:n)}, \boldsymbol{X}_{i,C}^{(1:n)}, \overrightarrow{\boldsymbol{Y}}_i^{(1:n-1)}, \boldsymbol{X}_j^{(1:n)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(n)}|\boldsymbol{X}_{j,C}^{(1:n)}, \boldsymbol{X}_{i,U}^{(1:n)}, \overrightarrow{\boldsymbol{Y}}_j^{(1:n-1)}\right)\right] \\
&\quad + N\delta(N), \\
&= \sum_{n=1}^{N} \left[H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}\right) + H\left(\boldsymbol{X}_{i,U}^{(n)}, \overleftarrow{\boldsymbol{Y}}_j^{(n)}|W_j, \boldsymbol{X}_{i,U}^{(1:n-1)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:n-1)}, \boldsymbol{X}_j^{(1:n)}\right)\right. \\
&\quad \left. + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}|W_j, \overleftarrow{\boldsymbol{Y}}_j^{(1:n)}, \boldsymbol{X}_{i,C}^{(1:n)}, \overrightarrow{\boldsymbol{Y}}_i^{(1:n-1)}, \boldsymbol{X}_j^{(1:n)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(n)}|\boldsymbol{X}_{j,C}^{(1:n)}, \boldsymbol{X}_{i,U}^{(1:n)}, \overrightarrow{\boldsymbol{Y}}_j^{(1:n-1)}\right)\right] \\
&\quad + N\delta(N), \\
&= \sum_{n=1}^{N} \left[H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}\right) + H\left(\boldsymbol{X}_{i,U}^{(n)}, \overleftarrow{\boldsymbol{Y}}_j^{(n)}|\boldsymbol{X}_{i,U}^{(1:n-1)}, \boldsymbol{X}_j^{(1:n)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}|\boldsymbol{X}_{i,C}^{(1:n)}, \overrightarrow{\boldsymbol{Y}}_i^{(1:n-1)}, \boldsymbol{X}_j^{(1:n)}\right)\right. \\
&\quad \left. + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(n)}|\boldsymbol{X}_{j,C}^{(1:n)}, \boldsymbol{X}_{i,U}^{(1:n)}, \overrightarrow{\boldsymbol{Y}}_j^{(1:n-1)}\right)\right] + N\delta(N), \\
&\overset{(n)}{\leqslant} \sum_{n=1}^{N} \left[H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}\right) + H\left(\boldsymbol{X}_{i,U}^{(n)}, \overleftarrow{\boldsymbol{Y}}_j^{(n)}|\boldsymbol{X}_j^{(1:n)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}|\boldsymbol{X}_{i,C}^{(1:n)}, \boldsymbol{X}_j^{(1:n)}\right)\right. \\
&\quad \left. + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(n)}|\boldsymbol{X}_{j,C}^{(1:n)}, \boldsymbol{X}_{i,U}^{(1:n)}\right)\right] + N\delta(N),
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{n=1}^{N} \Big[H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}\right) + H\left(\boldsymbol{X}_{i,U}^{(n)}|\boldsymbol{X}_j^{(1:n)}\right) + H\left(\overleftarrow{\boldsymbol{Y}}_j^{(n)}|\boldsymbol{X}_j^{(1:n)}, \boldsymbol{X}_{i,U}^{(n)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}|\boldsymbol{X}_{i,C}^{(1:n)}, \boldsymbol{X}_j^{(1:n)}\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(n)}|\boldsymbol{X}_{j,C}^{(1:n)}, \boldsymbol{X}_{i,U}^{(1:n)}\right) \Big] + N\delta(N), \\
&\overset{(n)}{\leqslant} \sum_{n=1}^{N} \Big[H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}\right) + H\left(\boldsymbol{X}_{i,U}^{(n)}\right) + H\left(\overleftarrow{\boldsymbol{Y}}_j^{(n)}|\boldsymbol{X}_j^{(1:n)}, \boldsymbol{X}_{i,U}^{(n)}\right) + H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}|\boldsymbol{X}_{i,C}^{(1:n)}, \boldsymbol{X}_j^{(1:n)}\right) \\
&\quad + H\left(\overrightarrow{\boldsymbol{Y}}_j^{(n)}|\boldsymbol{X}_{j,C}^{(1:n)}, \boldsymbol{X}_{i,U}^{(1:n)}\right) \Big] + N\delta(N), \\
&= N\Big[H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}\right) + H\left(\boldsymbol{X}_{i,U}^{(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), \text{ for any } n \in \{1, \ldots, N\}, \\
&= N\Big[H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}\right) + H\left(\boldsymbol{X}_{i,U}^{(n)}\right) + H\left(\overleftarrow{\boldsymbol{Y}}_j^{(n)}|\boldsymbol{X}_j^{(n)}, \boldsymbol{X}_{i,U}^{(n)}\right) + H\left(\boldsymbol{X}_{i,P}^{(n)}|\boldsymbol{X}_{i,C}^{(n)}, \boldsymbol{X}_j^{(n)}\right) \\
&\quad + H\left(\boldsymbol{X}_{j,P}^{(n)}|\boldsymbol{X}_{j,C}^{(n)}, \boldsymbol{X}_{i,U}^{(n)}\right) \Big] + N\delta(N), \\
&\overset{(n)}{\leqslant} N\Big[H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}\right) + H\left(\boldsymbol{X}_{i,U}^{(n)}\right) + H\left(\overleftarrow{\boldsymbol{Y}}_j^{(n)}|\boldsymbol{X}_j^{(n)}, \boldsymbol{X}_{i,U}^{(n)}\right) + H\left(\boldsymbol{X}_{i,P}^{(n)}\right) + H\left(\boldsymbol{X}_{j,P}^{(n)}\right) \Big] \\
&\quad + N\delta(N), \\
&= N\Big[H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}\right) + H\left(\boldsymbol{X}_{i,U}^{(n)}\right) + H\left(\boldsymbol{X}_{i,CF_j}^{(n)}, \boldsymbol{X}_{i,DF}^{(n)}|\boldsymbol{X}_j^{(n)}, \boldsymbol{X}_{i,U}^{(n)}\right) + H\left(\boldsymbol{X}_{i,P}^{(n)}\right) + H\left(\boldsymbol{X}_{j,P}^{(n)}\right) \Big] \\
&\quad + N\delta(N), \\
&= N\Big[H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}\right) + H\left(\boldsymbol{X}_{i,U}^{(n)}\right) + H\left(\boldsymbol{X}_{i,CF_j}^{(n)}, \boldsymbol{X}_{i,DF}^{(n)}|\boldsymbol{X}_{i,U}^{(n)}\right) + H\left(\boldsymbol{X}_{i,P}^{(n)}\right) + H\left(\boldsymbol{X}_{j,P}^{(n)}\right) \Big] \\
&\quad + N\delta(N),
\end{aligned} \tag{66}
$$

where
(a) follows from Fano's inequality;
(b) follows from the fact that $I\left(W_i; \overrightarrow{\boldsymbol{Y}}_i^{(1:N)}\right) \leqslant I\left(W_i; \overrightarrow{\boldsymbol{Y}}_i^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j\right)$;
(c) follows from the fact that $\overleftarrow{\boldsymbol{Y}}_i^{(1:N)}$ is included into $\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}$;
(d) follows from the fact that $H\left(\overrightarrow{\boldsymbol{Y}}_i^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_i, W_j\right) = 0$;
(e) follows from the fact that $\boldsymbol{X}_i^{(1:N)} = f\left(W_i, \overleftarrow{\boldsymbol{Y}}_i^{(1:N-1)}\right)$;
(f) follows from the fact that $\boldsymbol{X}_{i,U}^{(1:N)}$ is included into $\boldsymbol{X}_i^{(1:N)}$;
(g) follows from the fact that including another random variable increases the mutual information;
(h) follows from the fact that $H(Y|X) = H(X,Y) - H(X)$;
(i) follows from the fact that injecting information increases the mutual information;
(j) follows from the fact that W_i is independent of W_j, then $H(W_i|W_j) = H(W_i)$;
(k) follows from the fact that $H\left(\boldsymbol{X}_{j,C}^{(1:N)}, \boldsymbol{X}_{i,U}^{(1:N)}, \overleftarrow{\boldsymbol{Y}}_j^{(1:N)}|W_j, W_i, \overleftarrow{\boldsymbol{Y}}_i^{(1:N)}\right) = 0$;
(l) follows from the fact that conditioning reduces the entropy;
(m) follows from the fact that $\boldsymbol{X}_j^{(1:n)} = f\left(W_j, \overleftarrow{\boldsymbol{Y}}_j^{(1:n-1)}\right)$;
(n) follows from the fact that conditioning reduces the entropy.

From (66), in the asymptotic regime, it holds that:

$$
\begin{aligned}
2R_i + R_j &\leqslant H\left(\overrightarrow{\boldsymbol{Y}}_i^{(n)}\right) + H\left(\boldsymbol{X}_{i,U}^{(n)}\right) + H\left(\boldsymbol{X}_{i,CF_j}^{(n)}, \boldsymbol{X}_{i,DF}^{(n)}|\boldsymbol{X}_{i,U}^{(n)}\right) + H\left(\boldsymbol{X}_{i,P}^{(n)}\right) + H\left(\boldsymbol{X}_{j,P}^{(n)}\right), \\
&\overset{(o)}{\leqslant} \dim \overleftarrow{\boldsymbol{Y}}_i^{(n)} + \dim \overrightarrow{\boldsymbol{Y}}_{i,G}^{(n)} + \dim \boldsymbol{X}_{i,U}^{(n)} + \dim\left(\boldsymbol{X}_{i,CF_j}^{(n)}, \boldsymbol{X}_{i,DF}^{(n)}\right) - \dim \boldsymbol{X}_{i,U}^{(n)} + \dim \boldsymbol{X}_{i,P}^{(n)} + \dim \boldsymbol{X}_{j,P}^{(n)}, \\
&= \dim \overleftarrow{\boldsymbol{Y}}_i^{(n)} + \dim \overrightarrow{\boldsymbol{Y}}_{i,G}^{(n)} + \dim\left(\boldsymbol{X}_{i,CF_j}^{(n)}, \boldsymbol{X}_{i,DF}^{(n)}\right) + \dim \boldsymbol{X}_{i,P}^{(n)} + \dim \boldsymbol{X}_{j,P}^{(n)},
\end{aligned} \tag{67}
$$

where
(o) follows from (53).
Plugging (50), (52), (55), and (39) in (67), it yields

$$\begin{aligned}2R_i + R_j \leqslant & \min\left(\overleftarrow{n}_{ii}, \max\left(\overrightarrow{n}_{ii}, n_{ij}\right)\right) + \left(\max\left(\overrightarrow{n}_{ii}, n_{ij}\right) - \overleftarrow{n}_{ii}\right)^+ + \left(\overleftarrow{n}_{jj} - \left(\overrightarrow{n}_{jj} - n_{ji}\right)^+\right)^+ \\ & + \left(\overrightarrow{n}_{ii} - n_{ji}\right)^+ + \left(\overrightarrow{n}_{jj} - n_{ij}\right)^+, \\ = & \max\left(\overrightarrow{n}_{ii}, n_{ij}\right) + \left(\overleftarrow{n}_{jj} - \left(\overrightarrow{n}_{jj} - n_{ji}\right)^+\right)^+ + \left(\overrightarrow{n}_{ii} - n_{ji}\right)^+ + \left(\overrightarrow{n}_{jj} - n_{ij}\right)^+, \end{aligned} \tag{68}$$

and this completes the proof of (37). ■

References

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

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

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

[4] S. Mohajer, R. Tandon, and H. V. Poor, "On the feedback capacity of the fully connected-user interference channel," *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 2863–2881, May 2013.

[5] S.-Q. Le, R. Tandon, M. Motani, and H. V. Poor, "Approximate capacity region for the symmetric Gaussian interference channel with noisy feedback," *(Submitted to) IEEE Transactions on Information Theory*, Dec. 2012. http://www.ece.vt.edu/tandonr/IC-Partial-Feedback.pdf.

[6] D. Tuninetti, "On interference channel with generalized feedback (IFC-GF)," in *Proc. of International Symposium on Information Theory (ISIT)*, Nice, France, Jun. 2007, pp. 2661–2665.

[7] S. M. Perlaza, R. Tandon, H. V. Poor, and Z. Han, "Perfect output feedback in the two-user decentralized interference channel," (To appear in) *IEEE Transactions on Information Theory*, Jun. 2013. http://arxiv.org/abs/1306.2878.

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

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

[10] C. E. Shannon, "The zero-error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 8–19, Sep. 1956.

RESEARCH CENTRE
GRENOBLE – RHÔNE-ALPES

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

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

ISSN 0249-0803