Simultaneous Information and Energy Transmission in the Interference Channel

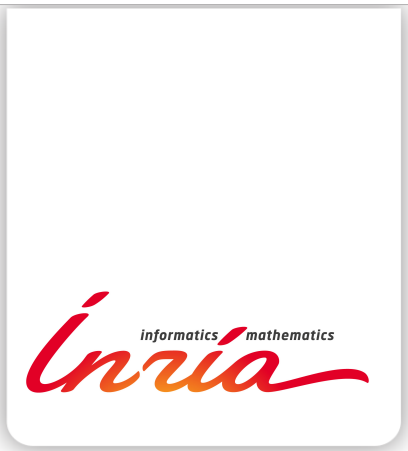
Nizar Khalfet, Samir Perlaza

Simultaneous Information and Energy Transmission in the Interference Channel

Nizar Khalfet and Samir M. Perlaza

RESEARCH REPORT
N° 9102
October 2017
Project-Teams Socrate

ISSN 0249-6399 ISRN INRIA/RR--9102--FR+ENG



Simultaneous Information and Energy Transmission in the Interference Channel

Nizar Khalfet and Samir M. Perlaza

Project-Teams Socrate

Research Report n° 9102 — version 1 — initial version October 2017 — revised version Novembre 2017 — 51 pages

Abstract: In this report, the fundamental limits of simultaneous information and energy transmission in the two-user Gaussian interference channel (G-IC) with and without feedback are fully characterized. More specifically, an achievable and converse region in terms of information and energy transmission rates (in bits per channel use and energy-units per channel use, respectively) are identified. In both cases, with and without feedback, an achievability scheme based on power-splitting, common randomness, rate splitting, block-Markov superposition coding, and backward decoding is presented. Finally, converse regions for both cases are obtained using some of the existing outer bounds for information transmission rates, as well as a new outer bound for the energy transmission rate.

Key-words: Two-user Gaussian interference channel, simultaneous information and energy transmission, feedback, information-energy capacity region, achievability, converse.

Nizar Khalfet and Samir M. Perlaza are with the Laboratoire CITI (a joint laboratory between the Université de Lyon, INSA de Lyon, and INRIA). 6 Avenue des Arts, F-69621, Villeurbanne, France. ({nizar.khalfet, samir.perlaza}@inria.fr).
Samir M. Perlaza is also with the Department of Electrical Engineering at Princeton University, Princeton, NJ.
This research was supported in part by the European Commission under Marie Skłodowska-Curie Individual Fellowship No. 659316 (CYBERNETS) and in part by the Euro-Mediterranean Cooperation ERA-NET project COM-MED.

RESEARCH CENTRE
GRENOBLE – RHÔNE-ALPES

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

Résumé : Dans ce rapport, les limites fondamentales de la transmission simultanée d'information et d'énergie dans le canal Gaussien à interférence (G-IC) avec et sans voie de retour sont déterminées. L'ensemble des débits atteignables de transmission d'information et d'énergie (en bits par utilisation du canal et en unités d'énergie par utilisation du canal respectivement) est identifié. Pour les deux cas, un schéma d'atteignabilité est basé sur power-splitting, common randomness, rate splitting, block-Markov superposition coding, et backward decoding est présenté. Finalement, la région converse pour les deux cas est obtenu en utilisant des techniques de majoration dans la littérature pour les débits d'information et aussi un majorant pour le débit d'énergie en utilisant la loi des grands nombres.

Mots-clés : Canal à interference, voie de retour, transmission simultanée d'information et d'énergie

Contents

1 Notation

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

2 Gaussian Interference Channel with Energy Harvesting

Consider the Gaussian interference channel (G-IC) with a non-colocated energy harvester depicted in figure 1 without feedback and in figure 2 with perfect channel-output feedback. Transmitter i, with $i \in \{1,2\}$, aims to execute two tasks: (a) an information transmission task and (b) an energy transmission task.

2.1 Information Transmission Task

From the information transmission standpoint, the goal of transmitter i is to convey an independent message index $W_i \in \mathcal{W}_i = \{1, 2, \ldots, 2^{NR_i}\}$ to receiver i using N channel input symbols $X_{i,1}, X_{i,2}, \ldots, X_{i,N}$. The channel coefficient from transmitter k to receiver i, with $k \in \{1,2\}$, is denoted by $h_{ik} \in \mathbb{R}_+$. At receiver i, during channel use n, input symbol $X_{i,n}$ is observed at receiver i subject to the interference produced by the symbol $X_{j,n}$ sent by transmitter j, with $j \in \{1,2\} \setminus \{i\}$, and a real additive Gaussian noise $Z_{i,n}$ with zero mean and variance σ_i^2. Hence, the channel output at receiver i during channel use n, denoted by $Y_{i,n}$, is:

$$Y_{i,n} = h_{ii}X_{i,n} + h_{ij}X_{j,n} + Z_{i,n}. \tag{1}$$

In the case without feedback, at each channel use n, the symbol $X_{i,n}$ sent by transmitter i depends upon the message index W_i and a randomly generated index $\Omega \in \mathbb{N}$. Let $f_{i,n}^{(N)} : \mathcal{W}_i \times \mathbb{N} \to \mathbb{R}$ be the encoding function at channel use n, such that for all $n \in \{1, 2, \ldots, N\}$, the following holds:

$$X_{i,n}{=}f_{i,n}^{(N)}(W_i, \Omega). \tag{2}$$

In the case with feedback, the symbol $X_{i,n}$ sent by transmitter i depends upon the indices W_i and Ω, but also upon all previous channel-outputs $Y_{i,1}, Y_{i,2}, \ldots, Y_{i,n-d}$, with $d \in \mathbb{N}$ the feedback delay. In the following, it is assumed that d is equal to one channel use, without any loss of generality. The first channel input symbol $X_{i,1}$ depends only on the message index W_i and Ω. More specifically, $f_{i,1}^{(N)} : \mathcal{W}_i \times \mathbb{N} \to \mathbb{R}$ and for all $n \in \{2, 3, \ldots, N\}$, $f_{i,n}^{(N)} : \mathcal{W}_i \times \mathbb{N} \times \mathbb{R}^{n-1} \to \mathbb{R}$ are the encoding functions such that:

$$X_{i,1}{=}f_{i,1}^{(N)}(W_i, \Omega) \quad \text{and} \tag{3a}$$

$$X_{i,n}{=}f_{i,n}^{(N)}(W_i, \Omega, Y_{i,1}, Y_{i,2}, \ldots, Y_{i,n-1}) \text{ for all } n > 1. \tag{3b}$$

In both cases, the random index Ω is assumed to be independent of both W_1 and W_2 and known by all transmitters and receivers. Moreover, channel input symbols $X_{i,1}, X_{i,2}, \ldots, X_{i,N}$ are subject to an average power constraint of the form

$$\frac{1}{N}\sum_{n=1}^{N} \mathrm{E}[X_{i,n}^2] \leq P_i, \tag{4}$$

where P_i denotes the average transmit power of transmitter i in energy units per channel use. The expectation in (4) is taken with respect to W_i and Ω in the case without feedback. In the case with feedback, the expectation is taken over the joint distribution of the message indices W_1 and W_2, the random index Ω, and the noise terms, i.e., Z_1 and Z_2. The dependence of $X_{i,n}$ on W_1, W_2, Ω, and the previously observed noise realizations is due to the effect of feedback as shown in (3). Note that W_i and Ω are assumed to be independent and uniformly distributed over their corresponding sets.

Let $T \in \mathbb{N}$ be fixed and assume that during a given communication, T blocks of N channel uses

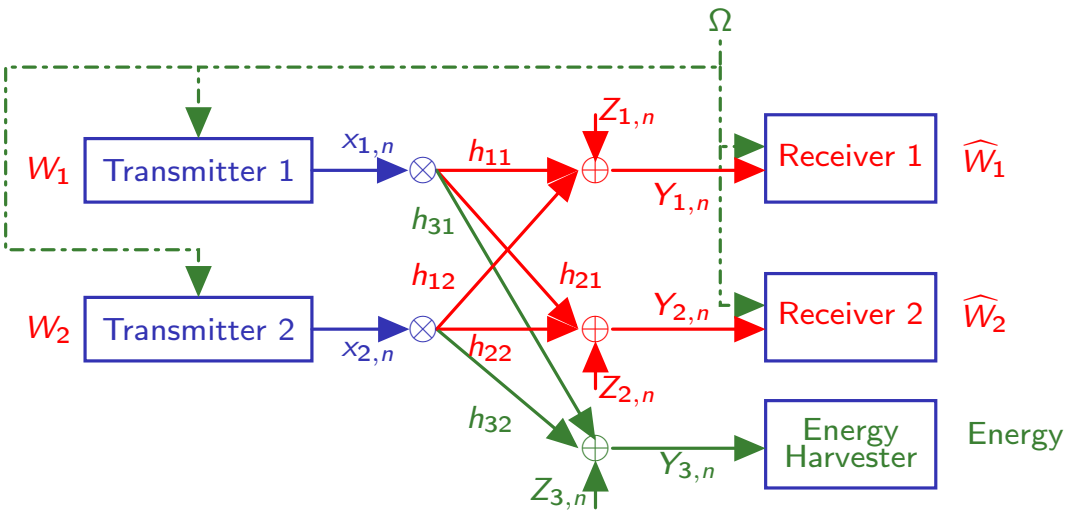


Figure 1: Two-user Gaussian interference channel with a non-colocated energy harvester at channel use n.

are transmitted. The decoder of receiver i observes the channel outputs $Y_{i,1}, Y_{i,2}, \ldots, Y_{i,NT}$ and uses a decoding function $\phi_i \colon \mathbb{N} \times \mathbb{R}^{NT} \to \mathcal{W}_i^T$, to get an estimate of the message indices:

$$\left(\widehat{W}_i^{(1)}, \widehat{W}_i^{(2)}, \ldots, \widehat{W}_i^{(T)}\right) = \phi_i\left(\Omega, Y_{i,1}, Y_{i,2}, \ldots, Y_{i,NT}\right), \tag{5}$$

where $\widehat{W}_i^{(t)}$ is an estimate of the message index $W_i^{(t)}$ sent during block $t \in \{1, 2, \ldots, T\}$. The decoding error probability during block t of a codebook of block-length N, denoted by $P_{\mathrm{e}}^{(t)}(N)$, is given by

$$P_{\mathrm{e}}^{(t)}(N) = \max\left(\Pr\left[\widehat{W_1}^{(t)} \neq W_1^{(t)}\right], \Pr\left[\widehat{W_2}^{(t)} \neq W_2^{(t)}\right]\right). \tag{6}$$

The signal to noise ratio (SNR) at receiver i is denoted by

$$\mathrm{SNR}_i = \frac{|h_{ii}|^2 P_i}{\sigma_i^2}. \tag{7a}$$

The interference to noise ratio (INR) at receiver i is denoted by

$$\mathrm{INR}_i = \frac{|h_{ij}|^2 P_j}{\sigma_i^2}, \text{ with } j \neq i. \tag{7b}$$

2.2 Energy Transmission Task

Let $h_{3i} \in \mathbb{R}_+$ be the channel coefficient from transmitter i to the energy harvester (EH). The symbols sent by the transmitters during channel use n are observed by the EH subject to an

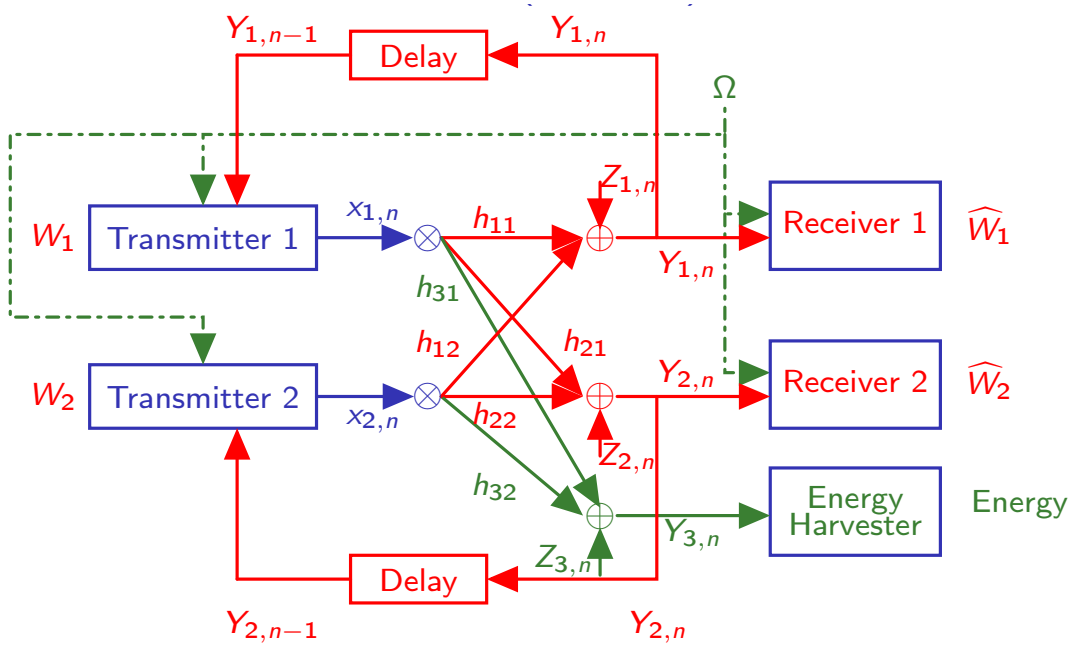


Figure 2: Two-user Gaussian interference channel with channel-output feedback and a non-collocated energy harvester during channel use n.

additive Gaussian noise $Z_{3,n}$ with zero mean and variance σ_3^2. More specifically, the channel output at the EH during channel use n, denoted by $Y_{3,n}$, is:

$$Y_{3,n} = h_{31}X_{1,n} + h_{32}X_{2,n} + Z_{3,n}. \tag{8}$$

From the energy transmission standpoint, the goal of both transmitters is to jointly guarantee an average energy rate at the EH. Let $b \geqslant 0$ denote the minimum average energy rate that must be guaranteed at the input of the EH. Let also $B^{(N)}$ be the average energy rate (in energy-units per channel use) at the end of N channel uses. That is,

$$B^{(N)} \triangleq \frac{1}{N}\sum_{n=1}^{N} Y_{3,n}^2. \tag{9}$$

The SNR of transmitter i at the EH is denoted by

$$\mathrm{SNR}_{3i} = \frac{|h_{3i}|^2 P_i}{\sigma_3^2}. \tag{10}$$

Note that the maximum average energy rate, denoted by $B_{\max}$, is:

$$B_{\max} = \sigma_3^2\left(1 + \mathrm{SNR}_{31} + \mathrm{SNR}_{32} + 2\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}\right). \tag{11}$$

The probability of energy outage, given an average energy rate B, is defined as follows:

$$P_{\text{outage}}^{(N,\epsilon)}(B) \triangleq \Pr\left[B^{(N)} < B - \epsilon\right], \tag{12}$$

for all $B > b$ and some $\epsilon > 0$.

2.3 Simultaneous Information and Energy Transmission

Given a minimum energy rate b to be satisfied at the EH, the system is said to be operating at the information-energy rate triplet $(R_1, R_2, B) \in \mathbb{R}_+^3$ when both transmitter-receiver pairs use a transmit-receive configurations such that: (i) reliable communication at information rates R_1 and R_2 is ensured; and (ii) reliable energy transmission at energy rate B is ensured. A formal definition is given below.

Definition 1 (Achievable Rates) The triplet $(R_1, R_2, B) \in \mathbb{R}_+^3$ is achievable if for all $i \in \{1, 2\}$, there exists a sequence of encoding functions $f_{i,1}^{(N)}, f_{i,2}^{(N)}, \ldots, f_{i,N}^{(N)}$ and a decoding function ϕ_i such that both the average error probability $P_{\mathrm{e}}^{(t)}(N)$, for all $t \in \{1, 2, \ldots, T\}$, and the energy-outage probability $P_{\mathrm{outage}}^{(N,\epsilon)}(B)$ tend to zero as the block-length N tends to infinity. That is,

$$\limsup_{N\to\infty} P_{\mathrm{e}}^{(t)}(N) = 0 \text{ and} \tag{13a}$$

$$\limsup_{N\to\infty} P_{\mathrm{outage}}^{(N,\epsilon)} = 0. \tag{13b}$$

Using Definition 1, the fundamental limits of simultaneous information and energy transmission in the Gaussian interference channel can be described by the information-energy capacity region, defined as follows.

Definition 2 (Information-Energy Capacity Region) The information-energy capacity region given a minimum energy rate b, denoted by $\mathcal{E}_b^{\mathrm{F}}$ in the case with feedback and $\mathcal{E}_b$ in the case without feedback, corresponds to the closure of all achievable information-energy rate triplets (R_1, R_2, B).

3 Main result

The main result consists of a description of the information-energy capacity regions with feedback $\mathcal{E}_b^{\mathrm{F}}$ and without feedback $\mathcal{E}_b$, for a given $b \geqslant 0$. Such a description is presented in the form of an *approximation* in the sense of the definition hereunder.

Definition 3 (Approximation of a Set) Let $n \in \mathbb{N}$ be fixed. A closed and convex region $\mathcal{X} \subset \mathbb{R}_+^n$ is approximated by the sets $\underline{\mathcal{X}}$ and $\overline{\mathcal{X}}$ if $\underline{\mathcal{X}} \subseteq \mathcal{X} \subseteq \overline{\mathcal{X}}$ and $\forall \boldsymbol{x} = (x_1, \ldots, x_n) \in \overline{\mathcal{X}}$ then $((x_1 - \xi_1)^+, (x_1 - \xi_2)^+, \ldots, (x_n - \xi_n)^+) \in \underline{\mathcal{X}}$, for some $(\xi_1, \xi_2, \ldots, \xi_n) \in \mathbb{R}_+^n$.

3.1 Case without Channel-Output Feedback

The information-energy capacity region $\mathcal{E}_b$, with b any positive real number, is approximated by the regions $\underline{\mathcal{E}}_b$ (Theorem 1), which represents an information-energy achievable region, and $\overline{\mathcal{E}}_b$ (Theorem 2), which represents an information-energy converse region.

3.1.1 An Achievable Region

The following theorem introduces an achievable information-energy region.

Theorem 1 *Let b be a fixed positive real. Then, the information-energy capacity region $\mathcal{E}_b$ contains all the rate tuples (R_1, R_2, B) that satisfy:*

$$R_1 \leq \frac{1}{2}\log\left(1+\frac{(1-\lambda_{1e})\mathrm{SNR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right), \tag{14a}$$

$$R_2 \leq \frac{1}{2}\log\left(1+\frac{(1-\lambda_{2e})\mathrm{SNR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right), \tag{14b}$$

$$R_1+R_2 \leq \frac{1}{2}\log\left(\frac{1+(1-\lambda_{1e})\mathrm{SNR}_1+(1-\lambda_{2e})\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right)+\frac{1}{2}\log\left(1+\frac{\lambda_{2p}\mathrm{SNR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right), \tag{14c}$$

$$R_1+R_2 \leq \frac{1}{2}\log\left(\frac{1+(1-\lambda_{2e})\mathrm{SNR}_2+(1-\lambda_{1e})\mathrm{INR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right)+\frac{1}{2}\log\left(1+\frac{\lambda_{1p}\mathrm{SNR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right), \tag{14d}$$

$$R_1+R_2 \leq \frac{1}{2}\log\left(\frac{1+\lambda_{1p}\mathrm{SNR}_1+(1-\lambda_{2e})\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right)+\frac{1}{2}\log\left(\frac{1+\lambda_{2p}\mathrm{SNR}_2+(1-\lambda_{1e})\mathrm{INR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right), \tag{14e}$$

$$\begin{aligned}2R_1+R_2 \leq{}& \frac{1}{2}\log\left(\frac{1+(1-\lambda_{1e})\mathrm{SNR}_1+(1-\lambda_{2e})\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right)+\frac{1}{2}\log\left(\frac{1+\lambda_{2p}\mathrm{SNR}_2+(1-\lambda_{1e})\mathrm{INR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right)\\ &+\frac{1}{2}\log\left(1+\frac{\lambda_{1p}\mathrm{SNR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right),\end{aligned} \tag{14f}$$

$$\begin{aligned}R_1+2R_2 \leq{}& \frac{1}{2}\log\left(\frac{1+(1-\lambda_{2e})\mathrm{SNR}_2+(1-\lambda_{1e})\mathrm{INR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right)+\frac{1}{2}\log\left(\frac{1+\lambda_{1p}\mathrm{SNR}_1+(1-\lambda_{2e})\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right)\\ &+\frac{1}{2}\log\left(1+\frac{\lambda_{2p}\mathrm{SNR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right),\end{aligned} \tag{14g}$$

$$b \leq B \leq \sigma_3^2\left(1+\mathrm{SNR}_{31}+\mathrm{SNR}_{32}+2\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}\sqrt{\lambda_{1e}\lambda_{2e}}\right), \tag{14h}$$

with $(\lambda_{ip}, \lambda_{ie}) \in [0,1]^2$ such that $\lambda_{ip}+\lambda_{ie} \leq 1$, for all $i \in \{1,2\}$.

Proof: The proof of Theorem 1 is presented in Appendix A. ∎

The achievability scheme used to obtain Theorem 1 is built upon random coding arguments using rate-splitting [1], superposition coding [2], common randomness, and power-spliting [3]. Following a rate-splitting argument, the index W_i is divided into two sub-indices $W_{i,P} \in \{1, 2\ldots, 2^{NR_{i,P}}\}$ and $W_{i,C} \in \{1, 2\ldots, 2^{NR_{i,C}}\}$, where $R_{i,C}+R_{i,P}=R_i$. Note that the block index (t) in $W_i^{(t)}$ and $\Omega^{(t)}$ are dropped as the encoding and decoding are identical at each block. The message index $W_{i,C}$ must be decoded at both receivers, whereas the index $W_{i,P}$ must be decoded only at the intended receiver. This rate-splitting is reminiscent of the Han-Kobayashi scheme in [1].

The codebook generation at transmitter i follows a three-level superposition coding scheme. The first code-layer is a sub-codebook of 2^{NR_E} codewords. Denote by $\boldsymbol{v}(\Omega)$ the corresponding codeword of the first code-layer. Note that both transmitters know Ω, hence they are able to choose the same codeword $\boldsymbol{v}(\Omega)$ from the first-layer codebook. The index Ω as well as the codeword $\boldsymbol{v}(\Omega)$ are also known at the receivers, which highlights that the role of this codebook is not information transmission but energy transmission. The second codeword used by transmitter i is selected using $W_{i,C}$ from the second code-layer, which is a sub-codebook of $2^{NR_{i,C}}$ codewords associated to $\boldsymbol{v}(\Omega)$. Denote by $\boldsymbol{u}_i(\Omega, W_{i,C})$ the corresponding codeword in the second code-layer. The third codeword used by transmitter i is selected using $W_{i,P}^{(t)}$ from the second code-layer, which is a sub-codebook of $2^{NR_{i,P}}$ codewords associated to $\boldsymbol{u}_i(\Omega, W_{i,C})$. Denote by $\boldsymbol{s}_i(\Omega, W_{i,C}, W_{i,P})$ the corresponding codeword in the third code-layer. Finally, for transmitting the triplet $(\Omega, W_{i,C}, W_{i,P})$, the channel input symbol $X_{i,n}$ at channel use $n \in \{1,2,\ldots,N\}$ is a deterministic function of the n-th components of the codewords $\boldsymbol{v}(\Omega)$, $\boldsymbol{u}_i(\Omega, W_{i,C})$ and $\boldsymbol{s}_i(\Omega, W_{i,C}, W_{i,P})$.

3.1.2 A Converse Region

The following Theorem introduces an information-energy converse region.

Theorem 2 *Let b be a fixed positive real. Then, the information-energy capacity region $\mathcal{E}_b$ is contained into the set of all the rate tuples (R_1, R_2, B) that satisfy:*

$$R_1 \leq \frac{1}{2}\log(1+\beta_1\mathrm{SNR}_1), \tag{15a}$$

$$R_2 \leq \frac{1}{2}\log(1+\beta_2\mathrm{SNR}_2), \tag{15b}$$

$$R_1+R_2 \leq \frac{1}{2}\log(1+\beta_1\mathrm{SNR}_1+\beta_2\mathrm{INR}_1)+\frac{1}{2}\log\left(1+\frac{\beta_2\mathrm{SNR}_2}{1+\beta_2\mathrm{INR}_1}\right), \tag{15c}$$

$$R_1+R_2 \leq \frac{1}{2}\log(1+\beta_2\mathrm{SNR}_2+\beta_1\mathrm{INR}_2)+\frac{1}{2}\log\left(1+\frac{\beta_1\mathrm{SNR}_1}{1+\beta_1\mathrm{INR}_2}\right), \tag{15d}$$

$$\begin{aligned} R_1+R_2 \leq & \frac{1}{2}\log\left(1+\frac{\beta_1\mathrm{SNR}_1+\beta_2\mathrm{INR}_1+\beta_1\beta_2\mathrm{INR}_1\mathrm{INR}_2}{1+\beta_1\mathrm{INR}_2}\right) \\ & +\frac{1}{2}\log\left(1+\frac{\beta_2\mathrm{SNR}_2+\beta_1\mathrm{INR}_2+\beta_1\beta_2\mathrm{INR}_1\mathrm{INR}_2}{1+\beta_2\mathrm{INR}_1}\right), \end{aligned} \tag{15e}$$

$$\begin{aligned} 2R_1+R_2 \leq & \frac{1}{2}\log\left(1+\frac{\beta_1\mathrm{SNR}_1}{1+\beta_1\mathrm{INR}_2}\right)+\frac{1}{2}\log(1+\beta_1\mathrm{SNR}_1+\beta_2\mathrm{INR}_1) \\ & +\frac{1}{2}\log\left(1+\frac{\beta_2\mathrm{SNR}_2+\beta_1\mathrm{INR}_2+\beta_1\beta_2\mathrm{INR}_1\mathrm{INR}_2}{1+\beta_2\mathrm{INR}_1}\right), \end{aligned} \tag{15f}$$

$$\begin{aligned} R_1+2R_2 \leq & \frac{1}{2}\log\left(1+\frac{\beta_2\mathrm{SNR}_2}{1+\beta_2\mathrm{INR}_1}\right)+\frac{1}{2}\log(1+\beta_2\mathrm{SNR}_2+\beta_1\mathrm{INR}_2) \\ & +\frac{1}{2}\log\left(1+\frac{\beta_1\mathrm{SNR}_1+\beta_2\mathrm{INR}_1+\beta_1\beta_2\mathrm{INR}_1\mathrm{INR}_2}{1+\beta_1\mathrm{INR}_2}\right), \end{aligned} \tag{15g}$$

$$b \leq B \leq \sigma_3^2\left(1+\mathrm{SNR}_{31}+\mathrm{SNR}_{32}+2\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}\sqrt{(1-\beta_1)(1-\beta_2)}\right), \tag{15h}$$

with $(\beta_1, \beta_2) \in [0,1]^2$.

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

3.1.3 An Approximation to the Information-Energy Capacity Region

Using the inner region $\underline{\mathcal{E}}_b$ and the outer region $\overline{\mathcal{E}}_b$, described respectively by Theorem 1 and Theorem 2, the information-energy capacity region $\mathcal{E}_b$ can be approximated in the sense of Definition 3.

Theorem 3 (Approximation of $\mathcal{E}_b$) Let $\underline{\mathcal{E}}_b \subset \mathbb{R}_+^3$ and $\overline{\mathcal{E}}_b \subset \mathbb{R}_+^3$ be the sets of tuples (R_1, R_2, B) described by Theorem 1 and Theorem 2, respectively. Then,

$$\underline{\mathcal{E}}_b \subset \mathcal{E}_b \subset \overline{\mathcal{E}}_b, \tag{16}$$

and for all $(R_1, R_2, B) \in \overline{\mathcal{E}}_b$ it follows that $\left((R_1-1/2)^+, (R_2-1/2)^+, \left(B-\frac{B_{\max}}{2}\right)^+\right) \in \underline{\mathcal{E}}_b$.

Proof: Following similar steps as in [4], it can be shown that for all $(R_1, R_2, 0) \in \overline{\mathcal{E}}_b$ it follows that $((R_1-1/2)^+, (R_2-1/2)^+, 0) \in \underline{\mathcal{E}}_b$. Note also that for all $(R_1, R_2, B) \in \overline{\mathcal{E}}_b$ and for all $(R_1, R_2, B') \in \underline{\mathcal{E}}_b$, there always exists a tuple $(\beta_1, \beta_2, \lambda_{1e}, \lambda_{2e})$ such that:

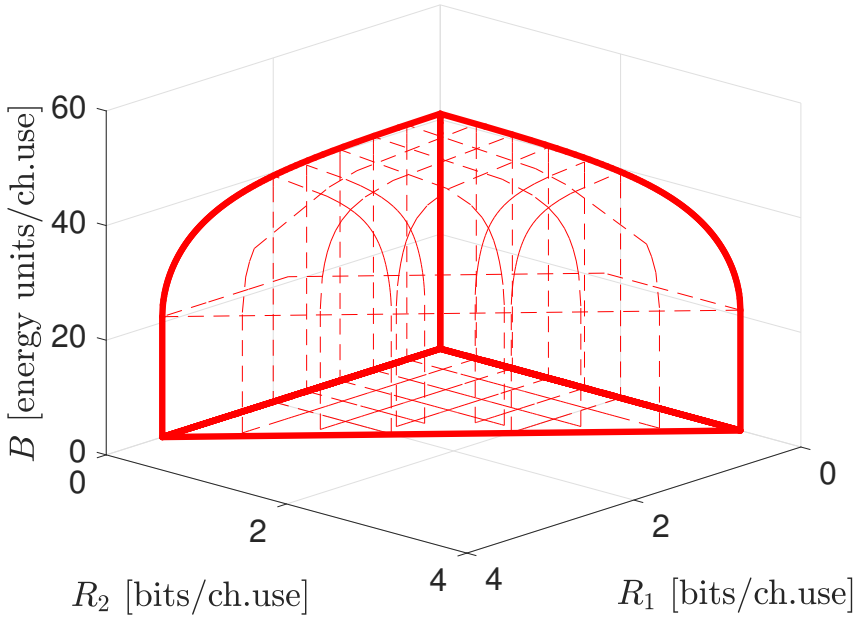


Figure 3: 3-D representation of $\underline{\mathcal{E}}_b$, with parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB, $\sigma_3^2 = 1$, and $b = 0$.

$$\begin{aligned}\frac{B - B'}{B_{\max}} &= \frac{2h_{31}h_{32}\sqrt{P_1P_2}\left(\sqrt{(1-\beta_1)(1-\beta_2)} - \sqrt{\lambda_{1e}\lambda_{2e}}\right)}{\sigma_3^2 + h_{31}^2P_1 + h_{32}^2P_2 + 2|h_{31}||h_{32}|\sqrt{P_1P_2}} \\ &\leq \frac{2\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}}{1 + \mathrm{SNR}_{31} + \mathrm{SNR}_{32} + 2\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}} \\ &\leq \frac{2\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}}{1 + 4\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}} \\ &\leq \frac{1}{2},\end{aligned}$$

which completes the proof. ■

3.1.4 Examples

Consider a Gaussian interference channel with an external EH with parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB and $\sigma_3^2 = 1$. Figure 3 and Figure 4 show $\underline{\mathcal{E}}_b$ and $\overline{\mathcal{E}}_b$, respectively, with $b = 0$. Figure 5 shows both $\underline{\mathcal{E}}_b$ and $\overline{\mathcal{E}}_b$ in the same axes.

Note that for all $B \in [0, 1 + \mathrm{SNR}_{31} + \mathrm{SNR}_{32}]$, transmitting information with independent codewords is enough to satisfy the energy rate constraints. This implies that $\beta_1 = \beta_2 = 1$ is optimal in this regime. Alternatively, for all $B \in [1 + \mathrm{SNR}_{31} + \mathrm{SNR}_{32}, B_{\max}]$, transmitters deal with trade-off between the information and energy rate. Increasing B reduces the information region and and shrinks the information-energy capacity region.

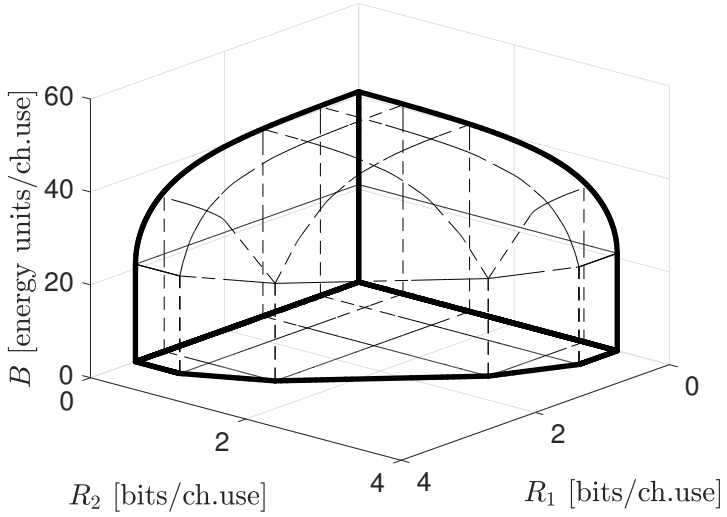


Figure 4: 3-D representation of $\overline{\mathcal{E}}_b$, with parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB, $\sigma_3^2 = 1$, and $b = 0$.

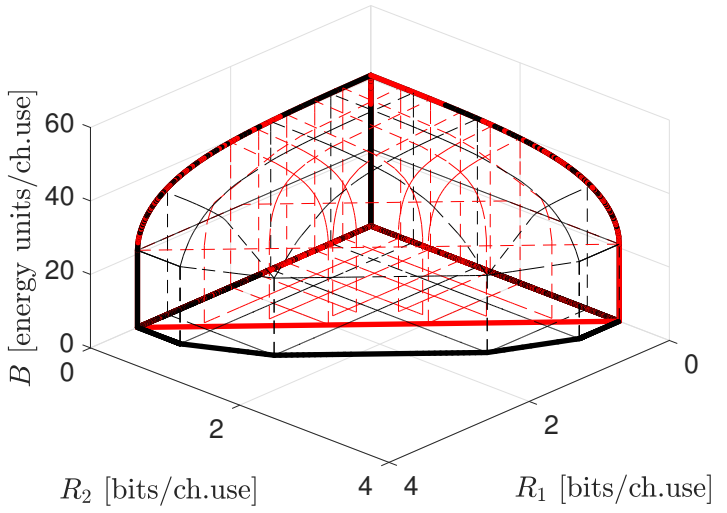


Figure 5: 3-D superposition of $\underline{\mathcal{E}}_b$ and $\overline{\mathcal{E}}_b$, with parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB, $\sigma_3^2 = 1$, and $b = 0$.

3.2 Case with Perfect Channel-Output Feedback

The information-energy capacity region $\mathcal{E}_b^{\mathrm{F}}$, with b any positive real number, is approximated by the regions $\underline{\mathcal{E}}_b^{\mathrm{F}}$ (Theorem 4) and $\overline{\mathcal{E}}_b^{\mathrm{F}}$ (Theorem 5).

3.2.1 An Achievable Region

The following theorem introduces an achievable information-energy region.

Theorem 4 *Let b be a fixed positive real. Then, the information-energy capacity region $\mathcal{E}_b^{\mathrm{F}}$ contains all the rate tuples (R_1, R_2, B) that satisfy:*

$$R_1 \leq \frac{1}{2}\log\left(\frac{1+(1-\lambda_{1e})\mathrm{SNR}_1+(1-\lambda_{2e})\mathrm{INR}_1+2\rho\sqrt{\mathrm{SNR}_1\mathrm{INR}_1}}{1+\lambda_{2p}\mathrm{INR}_1}\right), \tag{17a}$$

$$R_1 \leq \frac{1}{2}\log\left(\frac{1+(1-(\rho+\lambda_{1e}))\mathrm{INR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\right)+\frac{1}{2}\log\left(\frac{1+\lambda_{1p}\mathrm{SNR}_1+\lambda_{2p}\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right), \tag{17b}$$

$$R_2 \leq \frac{1}{2}\log\left(\frac{1+(1-\lambda_{2e})\mathrm{SNR}_2+(1-\lambda_{1e})\mathrm{INR}_2+2\rho\sqrt{\mathrm{SNR}_2\mathrm{INR}_2}}{1+\lambda_{1p}\mathrm{INR}_2}\right), \tag{17c}$$

$$R_2 \leq \frac{1}{2}\log\left(\frac{1+(1-(\rho+\lambda_{2e}))\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right)+\frac{1}{2}\log\left(\frac{1+\lambda_{2p}\mathrm{SNR}_1+\lambda_{1p}\mathrm{INR}_1}{1+\lambda_{1p}\mathrm{INR}_1}\right), \tag{17d}$$

$$\begin{aligned} R_1+R_2 \leq{}& \frac{1}{2}\log\left(\frac{1+\lambda_{1p}\mathrm{SNR}_1+\lambda_{2p}\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\right) \\ &+\frac{1}{2}\log\left(\frac{1+(1-\lambda_{2e})\mathrm{SNR}_2+(1-\lambda_{1e})\mathrm{INR}_2+2\rho\sqrt{\mathrm{SNR}_2\mathrm{INR}_2}}{1+\lambda_{1p}\mathrm{INR}_2}\right), \end{aligned} \tag{17e}$$

$$\begin{aligned} R_1+R_2 \leq{}& \frac{1}{2}\log\left(\frac{1+\lambda_{2p}\mathrm{SNR}_1+\lambda_{1p}\mathrm{INR}_1}{1+\lambda_{1p}\mathrm{INR}_1}\right) \\ &+\frac{1}{2}\log\left(\frac{1+(1-\lambda_{1e})\mathrm{SNR}_1+(1-\lambda_{2e})\mathrm{INR}_1+2\rho\sqrt{\mathrm{SNR}_1\mathrm{INR}_1}}{1+\lambda_{2p}\mathrm{INR}_1}\right), \end{aligned} \tag{17f}$$

$$b \leq B \leq \sigma_3^2\left(1+\mathrm{SNR}_{31}+\mathrm{SNR}_{32}+2\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}(\rho+\sqrt{\lambda_{1e}\lambda_{2e}})\right), \tag{17g}$$

where $(\rho, \lambda_{ip}, \lambda_{ie}) \in [0,1]^3$ and $\rho+\lambda_{ip}+\lambda_{ie} \leq 1$, for all $i \in \{1,2\}$.

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

The achievability scheme used to obtain Theorem 4 is built upon random coding arguments using rate-splitting [1], block-Markov superposition coding, backward decoding [5], common randomness, and power-spliting [3]. Let $W_i^{(t)} \in \{1, 2\ldots, 2^{NR_i}\}$ and $\Omega \in \{1, 2\ldots, 2^{NR_E}\}$ be the message index and the common random index at transmitter i during the t-th block. Following a rate-splitting argument, the index $W_i^{(t)}$ is divided into two sub-indices $W_{i,P}^{(t)} \in \{1, 2\ldots, 2^{NR_{i,P}}\}$ and $W_{i,C}^{(t)} \in \{1, 2\ldots, 2^{NR_{i,C}}\}$, where $R_{i,C}+R_{i,P}=R_i$. At the end of block t, the message indices $W_{i,C}^{(t)}$ and $W_{i,P}^{(t)}$ must be decoded by receiver i, whereas $W_{i,C}^{(t)}$ must be decoded by receiver j, and by transmitter j via feedback. Therefore at the beginning of block t, each transmitter possesses the knowledge of the indices $W_{1,C}^{(t-1)}$ and $W_{2,C}^{(t-1)}$. In the case of the first block $t=1$, the indices $W_{1,C}^{(0)}$ and $W_{2,C}^{(0)}$ correspond to two indices assumed to be known by all transmitters and receivers. The codebook generation at transmitter i follows a four-level superposition coding scheme. The first code-layer is a sub-codebook of 2^{NR_E} codewords. Denote by $\boldsymbol{v}(\Omega)$ the corresponding codeword of the first code-layer. Note that both transmitters know Ω, hence they are able to choose the same codeword $\boldsymbol{v}(\Omega)$ from the first-layer codebook. The index Ω as well as the codeword $\boldsymbol{v}(\Omega)$ are also known at the receivers, which highlights that the role of this codebook is not information transmission but energy transmission. The second code-layer is a sub-codebook of $2^{N(R_{1,C}+R_{2,C})}$ codewords. Denote by $\boldsymbol{u}\left(\Omega, W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}\right)$ the corresponding codeword in the second code-layer. The third codeword used by transmitter i is selected using $W_{i,C}^{(t)}$ from the third code-layer,

which is a sub-codebook of $2^{NR_{i,C}}$ codewords associated to $\boldsymbol{u}\left(\Omega, W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}\right)$. Denote by $\boldsymbol{u}_i\left(\Omega, W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}, W_{i,C}^{(t)}\right)$ the corresponding codeword in the third code-layer. The fourth codeword used by transmitter i is selected using $W_{i,P}^{(t)}$ from the fourth code-layer, which is a sub-codebook of $2^{NR_{i,P}}$ codewords associated to $\boldsymbol{s}_i\left(\Omega, W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}, W_{i,C}^{(t)}, W_{i,P}^{(t)}\right)$ the corresponding in the fourth code-layer. Finally, For transmitting the triplet $(\Omega, W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}, W_{i,C}^{(t)}, W_{i,P}^{(t)})$, the channel input-symbol $X_{i,n}$ at channel use $n \in \{1, 2, \ldots, N\}$ is a deterministic function of the n-th components of the codewords $\boldsymbol{v}(\Omega)$, $\boldsymbol{u}\left(\Omega, W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}\right)$, $\boldsymbol{u}_i\left(\Omega, W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}, W_{i,C}^{(t)}\right)$ and $\boldsymbol{s}_i\left(\Omega, W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}, W_{i,C}^{(t)}, W_{i,P}^{(t)}\right)$.

3.2.2 A Converse Region

The following theorem describes a converse region denoted by $\overline{\mathcal{E}}_b^{\mathrm{F}}$.

Theorem 5 *Let b be a fixed positive real. Then, the information-energy capacity region $\mathcal{E}_b^{\mathrm{F}}$ is contained into the set of all the rate tuples (R_1, R_2, B) that satisfy:*

$$R_1 \leq \frac{1}{2}\log\left(1+\beta_1\mathrm{SNR}_1+\beta_2\mathrm{INR}_1+2\rho\sqrt{\beta_1\mathrm{SNR}_1\beta_2\mathrm{INR}_1}\right), \tag{18a}$$

$$R_1 \leq \frac{1}{2}\log\left(1+\frac{\beta_1(1-\rho^2)\mathrm{SNR}_1}{1+\beta_1(1-\rho^2)\mathrm{INR}_2}\right)+\frac{1}{2}\log\left(1+\beta_1(1-\rho^2)\mathrm{INR}_2\right), \tag{18b}$$

$$R_2 \leq \frac{1}{2}\log\left(1+\beta_2\mathrm{SNR}_2+\beta_1\mathrm{INR}_2+2\rho\sqrt{\beta_2\mathrm{SNR}_2\beta_1\mathrm{INR}_2}\right), \tag{18c}$$

$$R_2 \leq \frac{1}{2}\log\left(1+\frac{\beta_2(1-\rho^2)\mathrm{SNR}_2}{1+\beta_2(1-\rho^2)\mathrm{INR}_1}\right)+\frac{1}{2}\log\left(1+\beta_2(1-\rho^2)\mathrm{INR}_1\right), \tag{18d}$$

$$\begin{aligned} R_1+R_2 \leq & \frac{1}{2}\log\left(1+\frac{\beta_1(1-\rho^2)\mathrm{SNR}_1}{1+\beta_1(1-\rho^2)\mathrm{INR}_2}\right) \\ & +\frac{1}{2}\log\left(1+\beta_2\mathrm{SNR}_2+\beta_1\mathrm{INR}_2+2\rho\sqrt{\beta_2\mathrm{SNR}_2\beta_1\mathrm{INR}_2}\right), \end{aligned} \tag{18e}$$

$$\begin{aligned} R_1+R_2 \leq & \frac{1}{2}\log\left(1+\frac{\beta_2(1-\rho^2)\mathrm{SNR}_2}{1+\beta_2(1-\rho^2)\mathrm{INR}_1}\right) \\ & +\frac{1}{2}\log\left(1+\beta_1\mathrm{SNR}_1+\beta_2\mathrm{INR}_1+2\rho\sqrt{\beta_1\mathrm{SNR}_1\beta_2\mathrm{INR}_1}\right), \end{aligned} \tag{18f}$$

$$\begin{aligned} b \leq B \quad \leq & \sigma_3^2\Big(1+\mathrm{SNR}_{31}+\mathrm{SNR}_{32}+2\rho\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}(\rho\sqrt{\beta_1\beta_2} \\ & +2\sqrt{(1-\beta_1)(1-\beta_2)})\Big), \end{aligned} \tag{18g}$$

with $(\beta_1, \beta_2, \rho) \in [0,1]^3$.

Proof: The proof of Theorem 5 is presented in Appendix D. ∎

3.2.3 An Approximation to the Information-Energy Capacity Region

Using the inner region $\underline{\mathcal{E}}_b^{\mathrm{F}}$ and the outer region $\overline{\mathcal{E}}_b^{\mathrm{F}}$, described respectively by Theorem 4 and Theorem 5, the information-energy capacity region $\mathcal{E}_b^{\mathrm{F}}$ can be approximated in the sense of Definition 3.

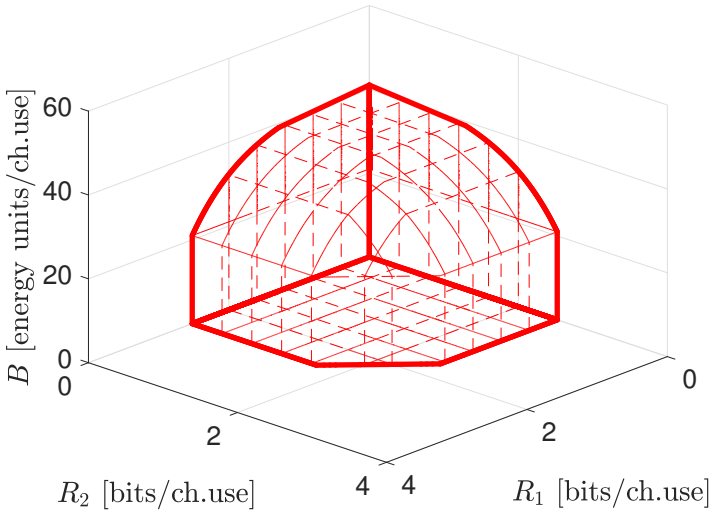


Figure 6: 3-D representation of $\underline{\mathcal{E}}_b^{\mathrm{F}}$, with parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB, $\sigma_3^2 = 1$, and $b = 0$.

Theorem 6 (Approximation of $\mathcal{E}_b^{\mathrm{F}}$) Let $\underline{\mathcal{E}}_b^{\mathrm{F}} \subset \mathbb{R}_+^3$ and $\overline{\mathcal{E}}_b^{\mathrm{F}} \subset \mathbb{R}_+^3$ be the sets of tuples (R_1, R_2, B) described by Theorem 4 and Theorem 5, respectively. Then,

$$\underline{\mathcal{E}}_b^{\mathrm{F}} \subset \mathcal{E}_b^{\mathrm{F}} \subset \overline{\mathcal{E}}_b^{\mathrm{F}}, \tag{19}$$

and for all $(R_1, R_2, B) \in \overline{\mathcal{E}}_b^{\mathrm{F}}$ it follows that $\left((R_1-1)^+, (R_2-1)^+, \left(B - \frac{B_{\max}}{2}\right)^+\right) \in \underline{\mathcal{E}}_b^{\mathrm{F}}$.

Proof: Following similar steps as in [5], it can be shown that for all $(R_1, R_2, 0) \in \overline{\mathcal{E}}_b^{\mathrm{F}}$ it follows that $((R_1-1)^+, (R_2-1)^+, 0) \in \underline{\mathcal{E}}_b^{\mathrm{F}}$. Note also that for all $(R_1, R_2, B) \in \overline{\mathcal{E}}_b^{\mathrm{F}}$ and for all $(R_1, R_2, B') \in \underline{\mathcal{E}}_b^{\mathrm{F}}$, there always exists a tuple $(\beta_1, \beta_2, \lambda_{1e}, \lambda_{2e}, \rho)$ such that:

$$\begin{aligned}
\frac{B - B'}{B_{\max}} &\leq \frac{2h_{31}h_{32}\sqrt{P_1P_2}\left(\sqrt{(1-\beta_1)(1-\beta_2)} + \rho\sqrt{\beta_1\beta_2}\right)}{\sigma_3^2 + h_{31}^2P_1 + h_{32}^2P_2 + 2h_{31}h_{32}\sqrt{P_1P_2}} \\
&\leq \frac{2\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}}{1 + \mathrm{SNR}_{31} + \mathrm{SNR}_{32} + 2\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}} \\
&\leq \frac{2\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}}{1 + 4\sqrt{\mathrm{SNR}_{31}\mathrm{SNR}_{32}}} \\
&\leq \frac{1}{2},
\end{aligned}$$

which completes the proof. ■

3.2.4 Example

Consider a Gaussian interference channel with feedback and an external EH with parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB and $\sigma_3^2 = 1$. Figure 6 and

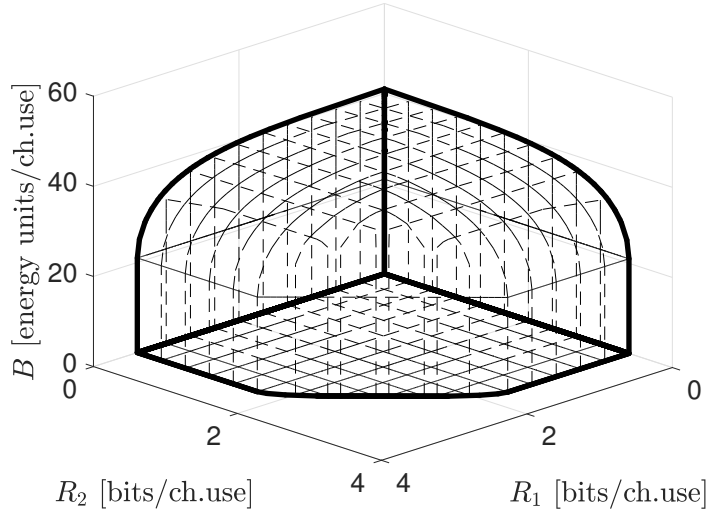


Figure 7: 3-D representation of $\overline{\mathcal{E}}_b^{\mathrm{F}}$, with parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB,$\sigma_3^2 = 1$,and $b = 0$

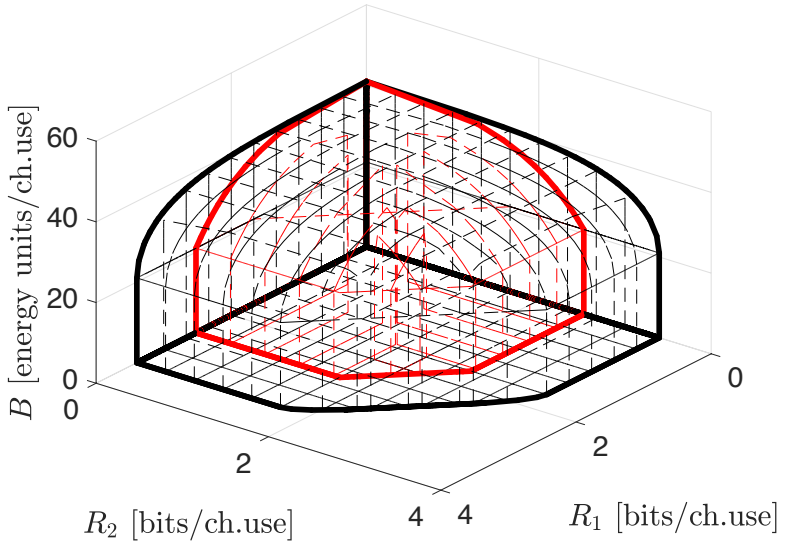


Figure 8: 3-D superposition of $\underline{\mathcal{E}}_b^{\mathrm{F}}$ and $\overline{\mathcal{E}}_b^{\mathrm{F}}$, with parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB and $\sigma_3^2 = 1$.

Figure 7 show $\underline{\mathcal{E}}_b$ and $\overline{\mathcal{E}}_b$, respectively, with $b = 0$. Figure 8 shows both $\underline{\mathcal{E}}_b$ and $\overline{\mathcal{E}}_b$ in the same axes. Note that for all $B \in [0, 1 + \mathrm{SNR}_{31} + \mathrm{SNR}_{32}]$, transmitting information with independent codewords is enough to satisfy the energy rate constraints. This implies that $\beta_1 = \beta_2 = 1$ is optimal in this regime. Alternatively, for all $B \in [1 + \mathrm{SNR}_{31} + \mathrm{SNR}_{32}, B_{\max}]$, transmitters deal with a trade-off between the information and energy rate. More specifically, increasing B reduces

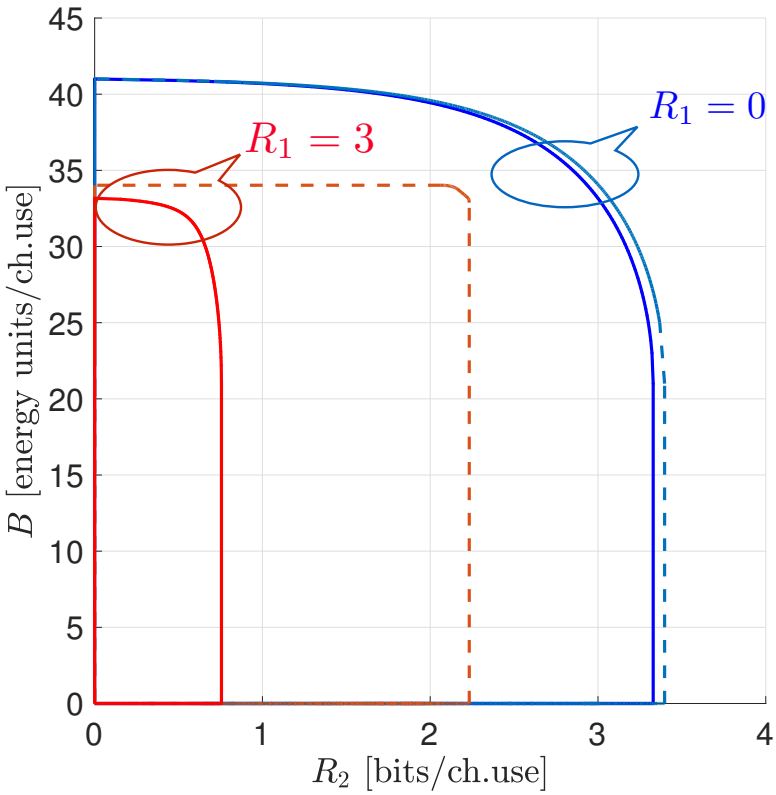


Figure 9: Energy rate B of the symmetric two-user Gaussian interference channel with feedback (dashed line) and without feedback (solid line), with co-located receiver EH, as a function of R_2 with two different values of R_1 and parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB, $\sigma_3^2 = 1$, and $b = 0$.

the information region and shrinks the information-energy capacity region.

3.2.5 Comparison with the Case of no Feedback

In this section, the impact of the feedback in terms of information and energy transmission under the assumption of the symmetric interference channel is quantified through an example. From the energy transmission standpoint, figure 9 shows the enhancement of the energy transmission due to the use of feedback when the information rate R_1 corresponding to transmitter-receiver pair 1 is fixed.

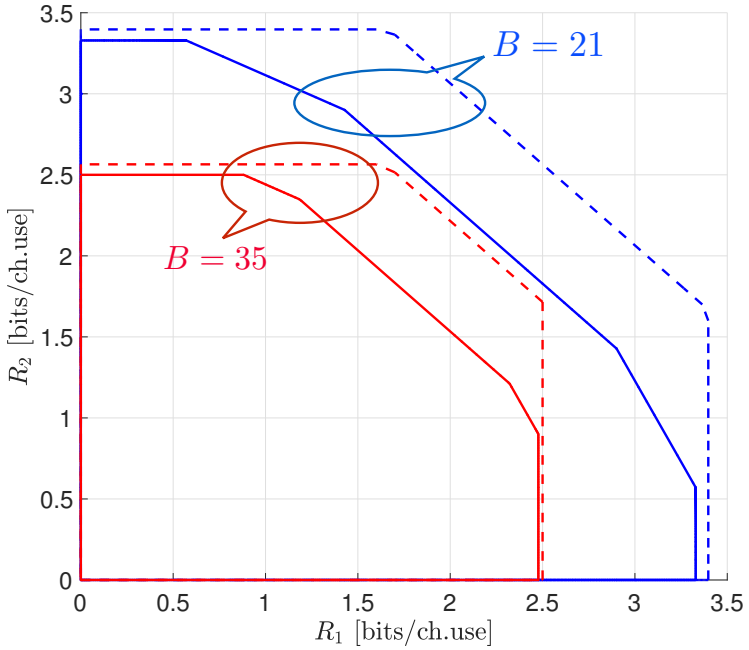


Figure 10: The Information capacity region of the symmetric two-user Gaussian interference channel with feedback (dashed line) and without feedback (solid line), with co-located receiver EH, with two different values of B and parameters $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 20$ dB, $\mathrm{INR}_1 = \mathrm{INR}_2 = \mathrm{SNR}_{31} = \mathrm{SNR}_{32} = 10$ dB, $\sigma_3^2 = 1$, and $b = 0$

Appendices

A Proof of Theorem 1

This proof is divided into two parts. The first part consists of the proof of (14a)-(14g). The second part consists of the proof of (14h).

A.1 Proof of (14a)-(14g)

Lemma 1 *Let $B \in \mathbb{R}^+$ be fixed, then for all $(R_1, R_2, B) \in \underline{\mathcal{E}}_b$, the following holds:*

$$R_1 \leq I(X_1; Y_1|U_2, V), \tag{20a}$$
$$R_2 \leq I(X_2; Y_2|U_1, V), \tag{20b}$$
$$R_1 + R_2 \leq I(X_1, U_2; Y_1|V) + I(X_2; Y_2|U_1, U_2, V), \tag{20c}$$
$$R_1 + R_2 \leq I(X_2, U_1; Y_2|V) + I(X_1; Y_1|U_1, U_2, V), \tag{20d}$$
$$R_1 + R_2 \leq I(X_1, U_2; Y_1|U_1, V) + I(X_2, U_1; Y_2|U_2, V), \tag{20e}$$
$$2R_1 + R_2 \leq I(X_1, U_2; Y_1|V) + I(X_1; Y_1|U_1, U_2, V) + (X_2U_1; Y_2|U_2V), \text{ and} \tag{20f}$$
$$R_1 + 2R_2 \leq I(X_1, U_2; Y_1|V) + I(X_1; Y_1|U_1, U_2, V) + I(X_2, U_1; Y_2|U_2, V), \tag{20g}$$

for all joint distributions

$$P_{VU_1U_2S_1S_2}(v,u_1,u_2,s_1,s_2) = P_V(v)P_{U_1|V}(u_1|v)P_{U_2|V}(u_2|v)P_{S_1|U_1V}(s_1|u_1,v)P_{S_2|U_2V}(s_2|u_2,v). \tag{21}$$

Proof: **Codebook Generation**: Fix a strictly positive joint probability distribution: $P_{VU_1U_2S_1S_2}(v,u_1,u_2,s_1,s_2) = P_V(v)P_{U_1|V}(u_1|v)P_{U_2|V}(u_2|v)P_{S_1|U_1V}(s_1|u_1,v)P_{S_2|U_2V}(s_2|u_2,v)$, for all $(v,u_1,u_2,s_1,s_2) \in \mathcal{X}_1 \cap \mathcal{X}_2 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_1 \times \mathcal{X}_2$. Let R_E, $R_{1,C}$, $R_{2,C}$, $R_{1,P}$ and $R_{2,P}$ be non negative real numbers. Generate 2^{NR_E} i.i.d N-length codewords $\boldsymbol{v}(\omega) = (v_1(\omega),\ldots,v_N(\omega))$ according to

$$P_{\boldsymbol{V}}(\boldsymbol{v}(\omega)) = \prod_{m=1}^{N} P_V(v_m(\omega)), \tag{22}$$

with $\omega \in \{1,2,\ldots,2^{NR_E}\}$. For encoder 1, for each codeword $\boldsymbol{v}(\omega)$, generate $2^{NR_{1,C}}$ i.i.d. N-length codewords $\boldsymbol{u}_1(\omega,i) = (u_{1,1}(\omega,i),\ldots,u_{1,N}(\omega,i))$ according to

$$P_{\boldsymbol{U_1}|\boldsymbol{V}}(\boldsymbol{u}_1(\omega,i)|\boldsymbol{v}(\omega)) = \prod_{m=1}^{N} P_{U_1|V}(u_{1,m}(\omega,i)|v_m(\omega)), \tag{23}$$

with $i \in \{1,\ldots,2^{NR_{1,C}}\}$. For each pair of codewords $\boldsymbol{u}_1(\omega,i)$ and $\boldsymbol{v}(\omega)$, generate $2^{NR_{1,P}}$ i.i.d. N-length codewords $\boldsymbol{s}_1(\omega,i,j) = (s_{1,1}(\omega,i,j),\ldots,s_{1,N}(\omega,i,j))$ according to

$$P_{\boldsymbol{S_1}|\boldsymbol{U_1V}}(\boldsymbol{s}_1(\omega,i,j)|\boldsymbol{u}_1(\omega,i),\boldsymbol{v}(\omega)) = \prod_{m=1}^{N} P_{S_1|U_1V}(s_{1,m}(\omega,i,j)|u_{1,m}(\omega,i),v_m(\omega)), \tag{24}$$

with $j \in \{1,\ldots,2^{NR_{1,P}}\}$. For encoder 2, for each codeword $\boldsymbol{v}(\omega)$, with $\omega \in \{1,2,\ldots,2^{NR_E}\}$ generate $2^{NR_{2,C}}$ i.i.d. N-length codewords $\boldsymbol{u}_2(\omega,k) = (u_{2,1}(\omega,k),\ldots,u_{2,N}(\omega,k))$ according to

$$P_{\boldsymbol{U_2}|\boldsymbol{V}}(\boldsymbol{u}_2(\omega,k)|\boldsymbol{v}(\omega)) = \prod_{m=1}^{N} P_{U_2|V}(u_{2,m}(\omega,k)|v_m(\omega)), \tag{25}$$

with $k \in \{1,\ldots,2^{NR_{2,C}}\}$. For each pair of codewords $\boldsymbol{u}_2(\omega,k)$ and $\boldsymbol{v}(\omega)$, generate $2^{NR_{2,P}}$ i.i.d. N-length codewords $\boldsymbol{s}_2(\omega,k,l) = (s_{2,1}(\omega,k,l),\ldots,s_{2,N}(\omega,k,l))$ according to

$$P_{\boldsymbol{S_2}|\boldsymbol{U_2V}}(\boldsymbol{s}_2(\omega,k,l)|\boldsymbol{u}_2(\omega,k),\boldsymbol{v}(\omega)) = \prod_{m=1}^{N} P_{S_2|U_2V}(s_{2,m}(\omega,k,l)|u_{2,m}(\omega,k),v_m(\omega)), \tag{26}$$

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

Encoding: Denote by $(W_i,\Omega) \in \{1,2,\ldots,2^{N(R_{i,C}+R_{i,P}}\} \times \{1,2,\ldots,2^{NR_E}\}$ the message index and the random message index of transmitter i. Let W_i be represented by the message index $W_{i,C} \in \{1,2,\ldots,2^{NR_{i,C}}\}$ and the message index $W_{i,P} \in \{1,2,\ldots,2^{NR_{i,P}}\}$. Transmitter i sends the codeword $\boldsymbol{x}_i = \theta_i\Big(\boldsymbol{v}(\Omega),\boldsymbol{u}_i(W_{i,C},\Omega),\boldsymbol{s}_i(W_{i,C},W_{i,P},\Omega)\Big)$, where $\theta_i : (\mathcal{X}_1 \cap \mathcal{X}_2)^N \times \mathcal{X}_i^N \times \mathcal{X}_i^N \to \mathcal{X}_i^N$ is a function that transforms the codewords $\boldsymbol{v}(\Omega)$, $\boldsymbol{u}_i(W_{i,C},\Omega)$, and $\boldsymbol{s}_i(W_{i,C},W_{i,P},\Omega)$ into the N-dimensional vector $\boldsymbol{x}_i$.

Decoding: Given the channel output $\boldsymbol{y}_1$, receiver 1 estimates the unique tuple $(\Omega,\widehat{W}_{1,C},\widehat{W}_{1,P})$ that satisfies:

$$\Big(\boldsymbol{v}(\Omega),\boldsymbol{u}_1(\widehat{W}_{1,C},\Omega),\boldsymbol{s}_1(\widehat{W}_{1,C},\widehat{W}_{1,P},\Omega),\boldsymbol{u}_2(\widehat{W}_{2,C},\Omega),\boldsymbol{y}_1\Big) \in \mathcal{T}^{(N,\epsilon)}_{VU_1S_1U_2Y_1}, \tag{27}$$

where Ω are assumed to be perfectly known by both transmitters and receivers. The set $\mathcal{T}^{(N,\epsilon)}_{VU_1S_1U_2Y_1}$ represents the set of jointly typical sequences of the random variables V, U_1, S_1, U_2, and Y_1, with $\epsilon > 0$. Finally, receiver 2 follows a similar decoding scheme.

Error Probability Analysis: an error might occur at receiver 1 if the indices $W_{1,C}$ and $W_{1,P}$ are not decoded correctly given that Ω is known by both transmitters and receivers. These errors might arise for two reasons: (i) there does not exist a tuple $(\Omega, \widehat{W}_{1,C}, \widehat{W}_{1,P})$, for at least one $\widehat{W}_{2,C}$ that satisfy (27), or (ii) there exist several tuples $(\Omega, \widehat{W}_{1,C}, \widehat{W}_{1,P})$, for at least one $\widehat{W}_{2,C}$ that simultaneously satisfy (27). From the asymptotic equipartion property (AEP) [6], 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_{ijk} that describes the case in which the codewords $\boldsymbol{v}(\Omega), \boldsymbol{u}_1(i,\Omega), \boldsymbol{s}_1(i,j,\Omega)$, and $\boldsymbol{u}_2(k,\Omega)$ are jointly typical with $\boldsymbol{y}_1$. Assume now that the codeword to be decoded corresponds to the indices $(i,j,k) = (1,1,1)$, this is without loss of generality due to the symmetry of the code. No error is declared when codewords $\Big(\boldsymbol{v}(\Omega), \boldsymbol{u}_1(1,\Omega), \boldsymbol{s}_1(1,1,\Omega), \boldsymbol{u}_2(\hat{k},\Omega)\Big)$, where $\hat{k} \neq 1$, are the only jointly typical sequences with the received sequence $\boldsymbol{y}_1$. Then, the probability of error P_e due to (ii), can be bounded as follows:

$$\begin{aligned}
P_e =& \Pr\left[\bigcup_{(i,j)\neq(1,1)} \bigcup_{k=1}^{2^{NR_{2,C}}} E_{ijk}\right] \\
\leq& \sum_{i\neq1,j\neq1,k\neq1} \Pr\left[E_{ijk}\right] + \sum_{i=1,j\neq1,k=1} \Pr\left[E_{ijk}\right] + \sum_{i\neq1,j\neq1,k=1} \Pr\left[E_{ijk}\right] \\
&+ \sum_{i\neq1,j=1,k\neq1} \Pr\left[E_{ijk}\right] + \sum_{i\neq1,j=1,k=1} \Pr\left[E_{ijk}\right] + \sum_{i=1,j\neq1,k\neq1} \Pr\left[E_{ijk}\right]. \qquad (28)
\end{aligned}$$

For all $i \in \{2, 3, \ldots, 2^{NR_{1,C}}\}$, the following holds

$$\begin{aligned}
\Pr[(E_{i11})] \overset{(a)}{=} & \Pr\left[(\boldsymbol{V}, \boldsymbol{U}_1, \boldsymbol{X}_1, \boldsymbol{U}_2, \boldsymbol{Y}_1) \in \mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}\right] \\
&= \sum_{(\boldsymbol{v},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{u}_2,\boldsymbol{y}_1)\in\mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}} P_{\boldsymbol{U}_1\boldsymbol{X}_1|\boldsymbol{V}}(\boldsymbol{u}_1,\boldsymbol{x}_1|\boldsymbol{v}) P_{\boldsymbol{U}_2\boldsymbol{Y}_1|\boldsymbol{V}}(\boldsymbol{u}_2,\boldsymbol{y}_1|\boldsymbol{v}) P_{\boldsymbol{V}}(\boldsymbol{v}), \\
&\leq \left|\mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}\right| 2^{-N(H(U_1,X_1|V)+H(U_2,Y_1|V)+H(V)-3\epsilon)} \\
&\leq 2^{-N(H(U_1,X_1|V)+H(U_2,Y_1|V)+H(V)-H(V,U_1,X_1,U_2,Y_1)-4\epsilon)} \\
&= 2^{-N(I(X_1;Y_1|U_2,V)-4\epsilon)}, \qquad (29a)
\end{aligned}$$

where the probability operator $\Pr[.]$ in (a) applies with a probability distribution $P_{\boldsymbol{V}\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2\boldsymbol{Y}_1}$ that factorizes as $P_{\boldsymbol{V}} P_{\boldsymbol{U}_1\boldsymbol{X}_1|\boldsymbol{V}} P_{\boldsymbol{U}_2\boldsymbol{Y}_1|\boldsymbol{V}}$ given that all the codewords $\boldsymbol{x}_1$ and $\boldsymbol{u}_1$ are independent from the output of the channel $\boldsymbol{y}_1$.

For all $j \in \{2,3,\dots,2^{NR_{1,P}}\}$, the following holds:

$$\begin{aligned}
\Pr(E_{1j1}) &\overset{(a)}{=} \Pr\left[(\boldsymbol{V},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,\boldsymbol{Y}_1) \in \mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}\right] \\
&= \sum_{(\boldsymbol{v},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{y}_1)\in\mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}} P_{\boldsymbol{X}_1|\boldsymbol{U}_1\boldsymbol{V}}(\boldsymbol{x}_1|\boldsymbol{u}_1,\boldsymbol{v})P_{\boldsymbol{Y}_1|\boldsymbol{U}_1\boldsymbol{U}_2\boldsymbol{V}}(\boldsymbol{y}_1|\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{v})P_{\boldsymbol{U}_1\boldsymbol{U}_2\boldsymbol{V}}(\boldsymbol{u}_1\boldsymbol{u}_2\boldsymbol{v}), \\
&\leq \left|\mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}\right| 2^{-N(H(X_1|V,U_1)+H(Y_1|V,U_1,U_2)+H(V,U_1,U_2)-3\epsilon)} \\
&\leq 2^{-N(H(X_1|V,U_1)+H(Y_1|V,U_1,U_2)+H(V,U_1U_2)-H(V,U_1,X_1,U_2,Y_1)-4\epsilon)} \\
&= 2^{-N(I(X_1;Y_1|U_1,U_2,V)-4\epsilon)},
\end{aligned} \tag{29b}$$

where the probability operator Pr[.] in (a) applies with a probability distribution $P_{\boldsymbol{V}\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2\boldsymbol{Y}_1}$ that factorizes as $P_{\boldsymbol{X}_1|\boldsymbol{U}_1\boldsymbol{V}}P_{\boldsymbol{Y}_1|\boldsymbol{U}_1\boldsymbol{U}_2\boldsymbol{V}}P_{\boldsymbol{U}_1\boldsymbol{U}_2\boldsymbol{V}}$ given that the codeword $\boldsymbol{x}_1$ is independent from the output of the channel $\boldsymbol{y_1}$.
For all $i \in \{2,3,\dots,2^{NR_{1,C}}\}$ and $j \in \{2,3,\dots,2^{NR_{1,P}}\}$ the following holds:

$$\begin{aligned}
\Pr(E_{ij1}) &\overset{(a)}{=} \Pr\left[(\boldsymbol{V},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,\boldsymbol{Y}_1) \in \mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}\right] \\
&= \sum_{(\boldsymbol{v},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{u}_2,\boldsymbol{y}_1)\in\mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}} P_{\boldsymbol{U}_1\boldsymbol{X}_1|\boldsymbol{V}}(\boldsymbol{u}_1,\boldsymbol{x}_1|\boldsymbol{v})P_{\boldsymbol{U}_2\boldsymbol{Y}_1|\boldsymbol{V}}(\boldsymbol{u}_2,\boldsymbol{y}_1|\boldsymbol{v})P_{\boldsymbol{V}}(\boldsymbol{v}), \\
&\leq \left|\mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}\right| 2^{-N(H(U_1,X_1|V)+H(U_2,Y_1|V)+H(V)-3\epsilon)} \\
&\leq 2^{-N(H(U_1,X_1|V)+H(U_2,Y_1|V)+H(V)-H(V,U_1,X_1,U_2,Y_1)-4\epsilon)} \\
&= 2^{-N(I(X_1;Y_1|U_2,V)-4\epsilon)},
\end{aligned} \tag{29c}$$

where the probability operator Pr[.] in (a) applies with a probability distribution $P_{\boldsymbol{V}\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2\boldsymbol{Y}_1}$ that factorizes as $P_{\boldsymbol{V}}P_{\boldsymbol{U}_1\boldsymbol{X}_1|\boldsymbol{V}}P_{\boldsymbol{U}_2\boldsymbol{Y}_1|\boldsymbol{V}}$ given that all the codewords $\boldsymbol{x}_1$ and $\boldsymbol{u}_1$ are independent from the output of the channel $\boldsymbol{y}_1$.
For all $i \in \{2,3,\dots,2^{NR_{1,C}}\}$ and $k \in \{2,3,\dots,2^{NR_{2,C}}\}$ the following holds:

$$\begin{aligned}
\Pr(E_{i1k}) &\overset{(a)}{=} \Pr\left[(\boldsymbol{V},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,\boldsymbol{Y}_1) \in \mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}\right] \\
&= \sum_{(\boldsymbol{v},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{u}_2,\boldsymbol{y}_1)\in\mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}} P_{\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2|\boldsymbol{V}}(\boldsymbol{u}_1\boldsymbol{x}_1\boldsymbol{u}_2|\boldsymbol{v})P_{\boldsymbol{Y}_1|\boldsymbol{V}}(\boldsymbol{y}_1|\boldsymbol{v})P_{\boldsymbol{V}}(\boldsymbol{v}), \\
&\leq \left|\mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}\right| 2^{-N(H(U_1,X_1,U_2|V)+H(Y_1|V)+H(V)-3\epsilon)} \\
&\leq 2^{-N(H(U_1,X_1,U_2|V)+H(Y_1|V)+H(V)-H(V,U_1,X_1,U_2,Y_1)-4\epsilon)} \\
&= 2^{-N(I(X_1,U_2;Y_1|V)-4\epsilon)},
\end{aligned} \tag{29d}$$

where the probability operator Pr[.] in (a) applies with a probability distribution $P_{\boldsymbol{V}\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2\boldsymbol{Y}_1}$ that factorizes as $P_{\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2|\boldsymbol{V}}P_{\boldsymbol{Y}_1|\boldsymbol{V}}P_{\boldsymbol{V}}$ given that all the codewords $\boldsymbol{x}_1$, $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are independent from the output of the channel $\boldsymbol{y}_1$.

For all $i \in \{2,3,\ldots,2^{NR_{1,P}}\}$ and $j \in \{2,3,\ldots,2^{NR_{2,C}}\}$ the following holds:

$$\begin{aligned}
\Pr(E_{ij1}) &\overset{(a)}{=} \Pr\left[(\boldsymbol{V},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,\boldsymbol{Y}_1) \in \mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}\right] \\
&= \sum_{(\boldsymbol{v},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{u}_2,\boldsymbol{y}_1)\in\mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}} P_{\boldsymbol{X}_1\boldsymbol{U}_2|\boldsymbol{U}_1\boldsymbol{V}}(\boldsymbol{x}_1\boldsymbol{u}_2|\boldsymbol{u}_1\boldsymbol{v})P_{\boldsymbol{Y}_1|\boldsymbol{U}_1\boldsymbol{V}}(\boldsymbol{y}_1|\boldsymbol{u}_1\boldsymbol{v})P_{\boldsymbol{U}_1\boldsymbol{V}}(\boldsymbol{u}_1\boldsymbol{v}), \\
&\leq \left|\mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}\right| 2^{-n(H(U_1,X_1,U_2|V)+H(Y_1|V)+H(V)-3\epsilon)} \\
&\leq 2^{-N(H(U_1,X_1,U_2|V)+H(Y_1|V)+H(V)-H(V,U_1,X_1,U_2,Y_1)-4\epsilon)} \\
&= 2^{-N(I(X_1,U_2;Y_1|V)-4\epsilon)},
\end{aligned} \tag{29e}$$

where $(a)$the probability operator Pr[.] applies with a probability distribution $P_{\boldsymbol{V}\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2\boldsymbol{Y}_1}$ that factorizes as $P_{\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2|\boldsymbol{V}}P_{\boldsymbol{Y}_1|\boldsymbol{V}}P_{\boldsymbol{V}}$ given that all the codewords $\boldsymbol{x}_1$ and $\boldsymbol{u}_2$ are independent from the output of the channel $\boldsymbol{y}_1$.
For all $i \in \{2,3,\ldots,2^{NR_{1,C}}\}$, $j \in \{2,3,\ldots,2^{NR_{1,P}}\}$ and $k \in \{2,3,\ldots,2^{NR_{2,C}}\}$ the following holds

$$\begin{aligned}
\Pr(E_{ijk}) &\overset{(a)}{=} \Pr\left[(\boldsymbol{V},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,\boldsymbol{Y}_1) \in \mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}\right] \\
&= \sum_{(\boldsymbol{v},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{u}_2,\boldsymbol{y}_1)\in\mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}} P_{\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2|\boldsymbol{V}}(\boldsymbol{u}_1\boldsymbol{x}_1\boldsymbol{u}_2|\boldsymbol{v})P_{\boldsymbol{Y}_1|\boldsymbol{V}}(\boldsymbol{y}_1|\boldsymbol{v})P_{\boldsymbol{V}}(\boldsymbol{v}), \\
&\overset{(b)}{\leq} \left|\mathcal{T}^{(N,\epsilon)}_{VU_1X_1U_2Y_1}\right| 2^{-n(H(U_1,X_1,U_2|V)+H(Y_1|V)+H(V)-3\epsilon)} \\
&\leq 2^{-N(H(U_1,X_1,U_2|V)+H(Y_1|V)+H(V)-H(V,U_1,X_1,U_2,Y_1)-4\epsilon)} \\
&= 2^{-N(I(X_1,U_2;Y_1|V)-4\epsilon)},
\end{aligned} \tag{29f}$$

where the probability operator Pr[.] in (a) applies with a probability distribution $P_{\boldsymbol{V}\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2\boldsymbol{Y}_1}$ that factorizes as $P_{\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2|\boldsymbol{V}}P_{\boldsymbol{Y}_1|\boldsymbol{V}}P_{\boldsymbol{V}}(\boldsymbol{v})$ given that all the codewords $\boldsymbol{x}_1$, $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are independent from the output of the channel $\boldsymbol{y}_1$.
Using (29) in (28), the following holds:

$$\begin{aligned}
P_{\mathrm{e}} \leq\ & 2^{N(R_{1c}-I(X_1;Y_1|U_2,V)+4\epsilon)}2^{N(R_{1p}-I(X_1;Y_1|U_1,U_2,V)+4\epsilon)} \\
& + 2^{N(R_{1c}+R_{1p}-I(X_1;Y_1|U_2,V)+4\epsilon)} + 2^{N(R_{1c}+R_{2c}-I(U_2,X_1;Y_1|V)+4\epsilon)} \\
& + 2^{N(R_{1p}+R_{2c}-I(U_2,X_1;Y_1|U_1,V)+4\epsilon)} + 2^{N(R_{1p}+R_{1c}+R_{2c}-I(U_2,X_1;Y_1|V)+4\epsilon)}.
\end{aligned} \tag{30}$$

The same analysis of the probability of error holds for transmitter-receiver pair 2. Hence in general, from (30), reliable decoding holds under the following conditions:

$$R_{1p} \leq a_1, \tag{31a}$$
$$R_{1p} + R_{1c} \leq d_1, \tag{31b}$$
$$R_{1p} + R_{2c} \leq e_1, \tag{31c}$$
$$R_{1p} + R_{1c} + R_{2c} \leq g_1, \tag{31d}$$
$$-R_{1p} \leq 0, \tag{31e}$$
$$-R_{1c} \leq 0, \tag{31f}$$

$$R_{2p} \leq a_2, \tag{31g}$$
$$R_{2p} + R_{2c} \leq d_2, \tag{31h}$$
$$R_{2p} + R_{1c} \leq e_2, \tag{31i}$$
$$R_{1p} + R_{1c} + R_{2c} \leq g_2, \tag{31j}$$
$$-R_{2p} \leq 0, \tag{31k}$$
$$-R_{2c} \leq 0, \tag{31l}$$

where

$$a_1{=}I(Y_1; X_1|U_1, U_2, V), \tag{32a}$$
$$d_1{=}I(Y_1; X_1|U_2, V), \tag{32b}$$
$$e_1{=}I(Y_1; X_1, U_2|U_1, V), \tag{32c}$$
$$g_1{=}I(Y_1; X_1, U_2|V), \tag{32d}$$
$$a_2{=}I(Y_2; X_2|U_1, U_2, V), \tag{32e}$$
$$d_2{=}I(Y_2; X_2|U_1, V), \tag{32f}$$
$$e_2{=}I(Y_2; X_2, U_1|U_2, V), \tag{32g}$$
$$g_2{=}I(Y_2; X_2, U_1|V). \tag{32h}$$

The proof continues by applying a Fourrier-Motzkin elimination process on (31).
Set $R_{1p} = R_1 - R_{1c}$, $R_{2p} = R_2 - R_{2c}$ and eliminate R_{1p}, R_{2p} from the set of inequalities (31) to obtain

$$R_1 - R_{1c} \leq a_1, \tag{33a}$$
$$R_1 \leq d_1, \tag{33b}$$
$$R_1 - R_{1c} + R_{2c} \leq e_1, \tag{33c}$$
$$R_1 + R_{2c} \leq g_1, \tag{33d}$$
$$-R_1 + R_{1c} \leq 0, \tag{33e}$$
$$-R_{1c} \leq 0, \tag{33f}$$
$$R_2 - R_{2c} \leq a_2, \tag{33g}$$
$$R_2 \leq d_2, \tag{33h}$$
$$R_2 - R_{2c} + R_{1c} \leq e_2, \tag{33i}$$
$$R_2 + R_{1c} \leq g_2, \tag{33j}$$
$$-R_2 + R_{2c} \leq 0, \tag{33k}$$
$$-R_{2c} \leq 0. \tag{33l}$$

Collect the inequalities in (33) that do not include R_{1c} to obtain:

$$R_1 \leq d_1, \tag{34a}$$
$$R_1 + R_{2c} \leq g_1, \tag{34b}$$
$$R_2 - R_{2c} \leq a_2, \tag{34c}$$
$$R_2 \leq d_2, \tag{34d}$$
$$-R_2 + T_2 \leq 0, \tag{34e}$$
$$-R_{2c} \leq 0. \tag{34f}$$

Also, collect the inequalities in (33) that include R_{1c} with positive coefficients to obtain:

$$R_2 - R_{2c} + R_{1c} \leq e_2, \tag{35a}$$
$$R_2 + R_{1c} \leq g_2, \tag{35b}$$
$$-R_1 + R_{1c} \leq 0. \tag{35c}$$

Furthermore, collect the inequalities in (33) that include $-R_{1c}$ to obtain:

$$-R_{1c} \leq 0, \tag{36a}$$
$$R_1 - R_{1c} \leq a_1, \tag{36b}$$
$$R_1 - R_{1c} + R_{2c} \leq e_1. \tag{36c}$$

Next, eliminate R_{1c} by summing each inequality in (35) with (36) to obtain:

$$R_2 - R_{2c} \leq e_2, \tag{37a}$$
$$R_2 \leq g_2, \tag{37b}$$
$$-R_1 \leq 0, \tag{37c}$$
$$R_1 + R_2 - R_{2c} \leq a_1 + e_2, \tag{37d}$$
$$R_1 + R_2 \leq a_1 + g_2, \tag{37e}$$
$$0 \leq a_1, \tag{37f}$$
$$R_1 + R_2 \leq e_1 + e_2, \tag{37g}$$
$$R_1 + R_2 + R_{2c} \leq e_1 + g_2, \tag{37h}$$
$$R_{2c} \leq e_1. \tag{37i}$$

Collect the inequalities in (33) and (37) that do not include R_{2c} to obtain:

$$0 \leq a_1, \tag{38a}$$
$$R_1 \leq d_1, \tag{38b}$$
$$R_2 \leq d_2, \tag{38c}$$
$$R_2 \leq g_2, \tag{38d}$$
$$-R_1 \leq 0, \tag{38e}$$
$$R_1 + R_2 \leq a_1 + g_2, \tag{38f}$$
$$R_1 + R_2 \leq e_1 + e_2. \tag{38g}$$

Collect the inequalities in (38) that include R_{2c} to obtain:

$$R_{2c} \leq e_1, \tag{39a}$$
$$R_1 + R_{2c} \leq g_1, \tag{39b}$$
$$R_1 + R_2 + R_{2c} \leq e_1 + g_2, \tag{39c}$$
$$-R_2 + R_{2c} \leq 0. \tag{39d}$$

Collect the inequalities in (38) that include $-R_{2c}$ to obtain:

$$-R_{2c} \leq 0, \tag{40a}$$
$$R_2 - R_{2c} \leq a_2, \tag{40b}$$
$$R_2 - R_{2c} \leq e_2, \tag{40c}$$
$$R_1 + R_2 - R_{2c} \leq a_1 + e_2. \tag{40d}$$

Note that (38a) is redundant due to the positivity of mutual information. The inequality (38d) is redundant with respect to (38c), given that:

$$\begin{aligned} g_2 &= I(Y_2; X_2, U_1|V) \\ &= I(Y_2; U_1|V) + I(Y_2; X_2|U_1, V) \\ &= d_2 + I(Y_2; U_1|V) \\ &\geq d_2. \end{aligned} \tag{41}$$

The inequality (40c) is redundant with respect to (40b), as shown hereunder:

$$\begin{aligned} e_2 &= I(Y_2; X_2, U_1|U_2, V) \\ &= I(Y_2; U_1|U_2, V) + I(Y_2; X_2|U_2, U_1, V) \\ &= a_2 + I(Y_2; U_1|U_2, V) \\ &\geq a_2. \end{aligned} \tag{42}$$

Eliminate R_{2c} by adding each inequality from (39) to each inequality (40) to obtain inequalities not including R_{2c}:

$$0 \leq e_1, \tag{43a}$$
$$R_1 \leq g_1, \tag{43b}$$
$$R_1 + R_2 \leq e_1 + g_2, \tag{43c}$$
$$-R_2 \leq 0, \tag{43d}$$
$$R_2 \leq a_2 + e_1, \tag{43e}$$
$$R_1 + R_2 \leq a_2 + g_1, \tag{43f}$$
$$R_1 + 2R_2 \leq a_2 + e_1 + g_2, \tag{43g}$$
$$0 \leq a_2, \tag{43h}$$
$$R_1 + R_2 \leq a_1 + e_2 + e_1, \tag{43i}$$
$$2R_1 + R_2 \leq a_1 + e_2 + g_1, \tag{43j}$$
$$2R_1 + 2R_2 \leq a_1 + e_2 + e_1 + g_2, \tag{43k}$$
$$R_1 \leq a_1 + e_2. \tag{43l}$$

Note that (43a) and (43h) are redundant due to the positivity of mutual information. The inequality (43b) is redundant with respect to (38b), as shown hereunder:

$$\begin{aligned} g_1 &= I(Y_1; X_1, U_2|V) \\ &= I(Y_1; U_2|V) + I(Y_1; X_1|U_2, V) \\ &= d_1 + I(Y_1; U_2|V) \\ &\geq d_1. \end{aligned} \tag{44}$$

The inequality (43c) is redundant with respect to (38f), since

$$\begin{aligned} e_1 &= I(Y_1; X_1, U_2|U_1, V) \\ &= I(Y_1; U_2|U_1, V) + I(Y_1; X_1|U_1, U_2, V) \\ &= a_1 + I(Y_1; U_2|U_1, V) \\ &\geq a_1. \end{aligned} \tag{45}$$

Note also that the inequality (43i) is redundant with respect to (38g). Furthermore, the inequality (43k) is redundant with respect to (38f) and (38g). Hence, the system of inequalities in (31) can be simplified as folllows:

$$R_1 \leq d_1, \tag{46a}$$
$$R_1 \leq a_1 + e_2, \tag{46b}$$
$$R_2 \leq d_2, \tag{46c}$$
$$R_2 \leq a_2 + e_1, \tag{46d}$$
$$R_1 + R_2 \leq a_1 + g_2, \tag{46e}$$
$$R_1 + R_2 \leq a_2 + g_1, \tag{46f}$$
$$R_1 + R_2 \leq e_1 + e_2, \tag{46g}$$
$$2R_1 + R_2 \leq a_1 + g_1 + e_2, \tag{46h}$$
$$R_1 + 2R_2 \leq a_2 + g_2 + e_1. \tag{46i}$$

Finally, following the result of [2], the inequalities (46b) and (46d) can be dropped and this completes the proof of Lemma 1. ■

The proof of Theorem 1 continues as follows. Let $k \in \{0, 1\}$ be fixed. Consider the following Gaussian input distribution for transmitter k:

$$V \sim \mathcal{N}(0,1);\ U_k \sim \mathcal{N}(0, \lambda_{kc});\ S_k \sim \mathcal{N}(0, \lambda_{kp}), \tag{47}$$

where $X_k = \sqrt{P_k}S_k + \sqrt{P_k}U_k + \sqrt{\lambda_{ke}P_k}V$; and $(\lambda_{kp}, \lambda_{kc}, \lambda_{ke}) \in [0,1]^3$ and $\lambda_{kp} + \lambda_{kc} + \lambda_{ke} \leq 1$. By symmetry, it suffices to prove (14a), (14c), (14e) and (14f). The choice of the Gaussian input distribution in (47) yields:

$$\begin{aligned}
I(X_1; Y_1|U_2, V) =& h(Y_1|U_2, V) - h(Y_1|X_1, U_2, V) \\
=& \frac{1}{2}\log\Big(2\pi e \mathrm{Var}[Y_1|U_2, V]\Big) - \frac{1}{2}\log\Big(2\pi e \mathrm{Var}[Y_1|X_1, U_2, V]\Big) \\
=& \frac{1}{2}\log\Big(2\pi e\left(\sigma_1^2 + h_{11}^2(1-\lambda_{1e})P_1 - h_{21}^2\lambda_{2p}P_2\right)\Big) \\
& - \frac{1}{2}\log\Big(2\pi e(\sigma_1^2 + h_{21}^2\lambda_{2p}P_2)\Big) \\
=& \frac{1}{2}\log\Big(1 + \frac{(1-\lambda_{1e})\mathrm{SNR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\Big),
\end{aligned} \tag{48a}$$

$$\begin{aligned}
I(X_1, U_2; Y_1|V) =& h(Y_1|V) - h(Y_1|X_1, U_2, V) \\
=& \frac{1}{2}\log\Big(2\pi e \mathrm{Var}[Y_1|V]\Big) - \frac{1}{2}\log\Big(2\pi e \mathrm{Var}[Y_1|X_1, U_2, V]\Big) \\
=& \frac{1}{2}\log\Big(2\pi e(\sigma_1^2 + h_{11}^2(1-\lambda_{1e})P_1 + h_{22}^2(1-\lambda_{2e}P_2)\Big) \\
& - \frac{1}{2}\log\Big(2\pi e(\sigma_1^2 + h_{21}^2\lambda_{2p}P_2)\Big) \\
=& \frac{1}{2}\log\Big(\frac{1 + (1-\lambda_{1e})\mathrm{SNR}_1 + (1-\lambda_{2e})\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\Big),
\end{aligned} \tag{48b}$$

$$
\begin{aligned}
I(X_2;Y_2|U_1,U_2,V)=&h(Y_2|U_1,U_2,V)-h(Y_2|U_1,U_2,X_2,V)\\
=&\frac{1}{2}\log\Big(2\pi e\mathrm{Var}[Y_2|U_1,U_2,V]\Big)-\frac{1}{2}\log\Big(2\pi e\mathrm{Var}[Y_2|U_1,U_2,X_2,V]\Big)\\
=&\frac{1}{2}\log\Big(2\pi e(\sigma_2^2+h_{22}^2\lambda_{2p}P_2+h_{21}^2\lambda_{1p}P_1)\Big)\\
&-\frac{1}{2}\log\Big(2\pi e(\sigma_2^2+h_{21}^2\lambda_{1p}P_1)\Big)\\
=&\frac{1}{2}\log\Big(1+\frac{\lambda_{2p}\mathrm{SNR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\Big),
\end{aligned}
\tag{48c}
$$

$$
\begin{aligned}
I(X_1,U_2;Y_1|U_1,V)=&h(Y_1|U_1,V)-h(Y_1|U_1,U_2,X_1,V)\\
=&\frac{1}{2}\log\Big(2\pi e\mathrm{Var}[Y_1|U_1,V]\Big)-\frac{1}{2}\log\Big(2\pi e\mathrm{Var}[Y_1|U_1,U_2,X_1,V]\Big)\\
=&\frac{1}{2}\log\Big(2\pi e(\sigma_1^2+h_{11}^2\lambda_{1p}P_1+h_{12}^2(1-\lambda_{2e})P_2)\Big)\\
&-\frac{1}{2}\log\Big(2\pi e(\sigma_1^2+h_{12}^2\lambda_{2p}P_2)\Big)\\
=&\frac{1}{2}\log\Big(\frac{1+\lambda_{1p}\mathrm{SNR}_1+(1-\lambda_{2e})\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\Big),
\end{aligned}
\tag{48d}
$$

$$
\begin{aligned}
I(X_2,U_1;Y_2|U_2,V)=&h(Y_2|U_2,V)-h(Y_2|U_2,U_1,X_2,V)\\
=&\frac{1}{2}\log\Big(2\pi e\mathrm{Var}[Y_2|U_2,V]\Big)-\frac{1}{2}\log\Big(2\pi e\mathrm{Var}[Y_2|U_2,U_1,X_2,V]\Big)\\
=&\frac{1}{2}\log\Big(2\pi e(\sigma_1^2+h_{22}^2\lambda_{2p}P_2+h_{21}^2(1-\lambda_{1e})P_1)\Big)\\
&-\frac{1}{2}\log\Big(2\pi e(\sigma_2^2+h_{21}^2\lambda_{1p}P_1)\Big)\\
=&\frac{1}{2}\log\Big(\frac{1+\lambda_{2p}\mathrm{SNR}_2+(1-\lambda_{1e})\mathrm{INR}_2}{1+\lambda_{1p}\mathrm{INR}_1}\Big),
\end{aligned}
\tag{48e}
$$

which proves (14a), (14c) and (14e). Finally, using (48c), (48d) and (48e), the proof of (14f) follows immediately .

A.2 Proof of (14h)

The choice of the channel input in (47) guarantee that the random variables $Y_{3,1},\ldots,Y_{3,n}$ are independently and identically distributed (i.i.d.). For all $n\in\{1,2,\ldots,N\}$, $Y_{3,n}$ follows a zero-mean Gaussian distribution with variance $\bar{B}$ given by

$$
\begin{aligned}
\bar{B}=&E[Y_{3,n}^2]\\
=&E[(h_{31}X_{1,n}+h_{32}X_{2,n}+Z_{3,n})^2]\\
=&h_{31}^2E[X_{1,n}^2]+h_{32}^2E[X_{2,n}^2]+2h_{31}h_{32}E[X_{1,n}X_{2,n}]+\sigma_3^2\\
=&h_{31}^2P_1+h_{32}^2P_2+2h_{31}h_{32}E[X_{1,n}X_{2,n}]+\sigma_3^2\\
\leq&h_{31}^2P_1+h_{32}^2P_2+2h_{31}h_{32}\sqrt{\lambda_{1e}P_1\lambda_{2e}P_2}+\sigma_3^2.
\end{aligned}
$$

By the weak law of large numbers, it holds that $\forall\epsilon>0$,

$$
\lim_{N\to\infty}\Pr\Big(|B^{(N)}-\bar{B}|>\epsilon\Big)=0. \tag{49}
$$

Consequently,

$$\lim_{N\to\infty} \Pr\left(B^{(N)} > \bar{B} + \epsilon\right) = 0 \quad \text{and} \tag{50a}$$

$$\lim_{N\to\infty} \Pr\left(B^{(N)} < \bar{B} - \epsilon\right) = 0. \tag{50b}$$

From (50b), it holds that for any energy rate B which satisfies $0 < B \leqslant \bar{B}$, it holds that

$$\lim_{N\to\infty} \Pr\left(B^{(N)} < B - \epsilon\right) = 0. \tag{51}$$

This proves (14h) and completes the proof of Theorem 1.

B Proof of Theorem 2

Fix an information-energy rate triplet (R_1, R_2, B) achievable with a given coding scheme (Definition 1). Denote by $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ the channel inputs resulting from transmitting the independent message W_1 and W_2 using such coding scheme. Denote by $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ the corresponding channel outputs. The bounds (15a) and (15b) on R_1 and R_2 are trivial and can be obtained by removing the interference from the other user and calculating the point-to-point capacity:

$$NR_1 \leqslant \sum_{n=1}^{N} h(Y_{1,n}|X_{2,n}) - Nh(Z_1) + \mathrm{o}(N) \text{ and} \tag{52}$$

$$NR_2 \leqslant \sum_{n=1}^{N} h(Y_{2,n}|X_{1,n}) - Nh(Z_2) + \mathrm{o}(N). \tag{53}$$

Define the following random variables:

$$T_1 = h_{21}X_1 + Z_2, \tag{54}$$

$$T_2 = h_{12}X_2 + Z_1, \tag{55}$$

$$U_1 = h_{21}X_1 + Z_2^{'}, \text{ and} \tag{56}$$

$$U_2 = h_{12}X_2 + Z_1^{'}, \tag{57}$$

where, $Z_1^{'}$ and $Z_2^{'}$ are real Gaussian random variables with zero mean and variances σ_1^2 and σ_2^2, respectively, independent of each other and of (X_1, X_2, Z_1, Z_2). The outer bound is established by using a genie aided argument. For all $j \in \{1, 2\}$, by Fano's inequality, it follows that

$$\begin{aligned} NR_j &= H(M_j) \\ &\leq I(M_j; \boldsymbol{Y_j}) + \mathrm{o}(N) \\ &\leq I(\boldsymbol{X}_j; \boldsymbol{Y}_j) + \mathrm{o}(N), \end{aligned} \tag{58}$$

where $\frac{\mathrm{o}(N)}{N}$ tends to zero as N tends to infinity. Using the definition of mutual information, yields:

$$\begin{aligned} I(\boldsymbol{X}_1; \boldsymbol{Y}_1) &= h(\boldsymbol{Y}_1) - h(\boldsymbol{Y}_1|\boldsymbol{X}_1) \\ &= h(\boldsymbol{Y}_1) - h(\boldsymbol{T_2}|\boldsymbol{X}_1) \\ &= h(\boldsymbol{Y}_1) - h(\boldsymbol{T}_2) \\ &\leq \sum_{n=1}^{N} h(Y_{1,n}) - h(\boldsymbol{T}_2). \end{aligned} \tag{59a}$$

Consider the genie-aided channel in which a genie provides $\boldsymbol{U}_1$ and $\boldsymbol{X}_2$ to receiver 1, then $I(\boldsymbol{X}_1;\boldsymbol{Y}_1)$ can be upper bounded as follows:

$$\begin{aligned} I(\boldsymbol{X}_1;\boldsymbol{Y}_1) &\leq I(\boldsymbol{X}_1;\boldsymbol{Y}_1,\boldsymbol{U}_1,\boldsymbol{X}_2) \\ &= I(\boldsymbol{X}_1,\boldsymbol{U}_1) + I(\boldsymbol{X}_1,\boldsymbol{X}_2|\boldsymbol{U}_1) + I(\boldsymbol{X}_1;\boldsymbol{Y}_1|\boldsymbol{U}_1,\boldsymbol{X}_2) \\ &= h(\boldsymbol{T}_1) - h(\boldsymbol{U}_1|\boldsymbol{X}_1) + h(\boldsymbol{Y}_1|\boldsymbol{U}_1,\boldsymbol{X}_2) - h(\boldsymbol{Y}_1|\boldsymbol{X}_1,\boldsymbol{U}_1,\boldsymbol{X}_2) \\ &= h(\boldsymbol{T}_1) - h(\boldsymbol{U}_1|\boldsymbol{X}_1) + h(\boldsymbol{Y}_1|\boldsymbol{U}_1,\boldsymbol{X}_2) - h(\boldsymbol{T}_2|\boldsymbol{X}_2) \\ &\overset{(a)}{\leq} h(\boldsymbol{T}_1) - h(\boldsymbol{Z}_2^{'}) + \sum_{n=1}^{N} h(Y_{1,n}|U_{1,n},X_{2,n}) - h(\boldsymbol{Z}_2) \\ &= h(\boldsymbol{T}_1) - Nh(Z_2^{'}) + \sum_{n=1}^{N} h(Y_{1,n}|U_{1,n},X_{2,n}) - Nh(Z_2), \end{aligned} \tag{59b}$$

where (a) follows from the fact that conditioning reduces the entropy. Consider the genie-aided channel in which a genie provides $\boldsymbol{U}_1$ to receiver 1, then $I(\boldsymbol{X}_1;\boldsymbol{Y}_1)$ can be upper bounded as follows:

$$\begin{aligned} I(\boldsymbol{X}_1;\boldsymbol{Y}_1) &\leq I(\boldsymbol{X}_1;\boldsymbol{Y}_1,\boldsymbol{U}_1) \\ &= I(\boldsymbol{X}_1;\boldsymbol{U}_1) + I(\boldsymbol{X}_1;\boldsymbol{Y}_1|\boldsymbol{U}_1) \\ &= h(\boldsymbol{U}_1) - h(\boldsymbol{U}_1|\boldsymbol{X}_1) + h(\boldsymbol{Y}_1|\boldsymbol{U}_1) - h(\boldsymbol{Y}_1|\boldsymbol{X}_1,\boldsymbol{U}_1) \\ &= h(\boldsymbol{T}_1) - h(\boldsymbol{U}_1|\boldsymbol{X}_1) + h(\boldsymbol{Y}_1|\boldsymbol{U}_1) - h(\boldsymbol{T}_2) \\ &\overset{(a)}{\leq} h(\boldsymbol{T}_1) - h(\boldsymbol{T}_2) - h(\boldsymbol{Z}_2^{'}) + \sum_{i=1}^{N} h(Y_{1,n}|U_{1,n}) \\ &= h(\boldsymbol{T}_1) - h(\boldsymbol{T}_2) - Nh(Z_2^{'}) + \sum_{n=1}^{N} h(Y_{1,n}|U_{1,n}). \end{aligned} \tag{59c}$$

Consider the genie-aided channel in which a genie provides $\boldsymbol{X}_2$ to receiver 1, then $I(\boldsymbol{X}_1;\boldsymbol{Y}_1)$ can be upper bounded as follows:

$$\begin{aligned} I(\boldsymbol{X}_1;\boldsymbol{Y}_1) &\leq I(\boldsymbol{X}_1;\boldsymbol{Y}_1,\boldsymbol{X}_2) \\ &= I(\boldsymbol{X}_1;\boldsymbol{X}_2) + I(\boldsymbol{X}_1;\boldsymbol{Y}_1|\boldsymbol{X}_2) \\ &= h(\boldsymbol{Y}_1|\boldsymbol{X}_2) - h(\boldsymbol{Y}_1|\boldsymbol{X}_1,\boldsymbol{X}_2) \\ &= h(\boldsymbol{Y}_1|\boldsymbol{X}_2) - h(\boldsymbol{T}_2|\boldsymbol{X}_2) \\ &\overset{(a)}{\leq} \sum_{n=1}^{N} h(Y_{1,n}|X_{2,n}) - h(\boldsymbol{Z}_1) \\ &= \sum_{n=1}^{N} h(Y_{1,n}|X_{2,n}) - Nh(Z_1). \end{aligned} \tag{59d}$$

By symmetry, similar bounds can be established for $I(\boldsymbol{X}_2;\boldsymbol{Y}_2)$, namely,

$$I(\boldsymbol{X_2};\boldsymbol{Y_2}) \leq \sum_{n=1}^{N} h(Y_{2,n}) - h(\boldsymbol{T}_1), \tag{59e}$$

$$I(\boldsymbol{X_2};\boldsymbol{Y_2}) \leq h(\boldsymbol{T}_2) - Nh(Z_1^{'}) + \sum_{n=1}^{N} h(Y_{2,n}|U_{2,n},X_{1,n}) - Nh(Z_1), \tag{59f}$$

$$I(\boldsymbol{X_2};\boldsymbol{Y}_2) \leq h(\boldsymbol{T}_2) - h(\boldsymbol{T}_1) - Nh(Z_1^{'}) + \sum_{n=1}^{N} h(Y_{2,n}|U_{2,n}), \tag{59g}$$

$$I(\boldsymbol{X}_2;\boldsymbol{Y}_2) \leq \sum_{n=1}^{N} h(Y_{2,n}|X_{2,n}) - Nh(Z_1). \tag{59h}$$

The key idea of the proof is to consider a linear combination of the inequalities in (59), where all the terms on $h(\boldsymbol{T}_1)$ and $h(\boldsymbol{T}_2)$ are removed. Adding (59f) and (59a) and plugging into (58) yields the first bound on sum rate:

$$N(R_1+R_2) \leqslant \sum_{n=1}^{N} h(Y_{1,n}) + \sum_{n=1}^{N} h(Y_{2,n}|U_{2,n},X_{1,n}) - Nh(Z_2) - Nh(Z_1^{'}) + \mathrm{o}(N). \tag{60a}$$

Adding (59b) and (59e) and plugging into (58) yields the second bound on sum rate:

$$N(R_1+R_2) \leqslant \sum_{n=1}^{N} h(Y_{2,n}) + \sum_{n=1}^{N} h(Y_{1,n}|U_{1,n},X_{2,n}) - Nh(Z_1) - Nh(Z_2^{'}) + \mathrm{o}(N). \tag{60b}$$

Adding (59c) and (59g) and plugging into (58) yields the third bound on sum rate:

$$N(R_1+R_2) \leqslant \sum_{n=1}^{N} h(Y_{1,n}|U_{1,n}) + \sum_{n=1}^{N} h(Y_{2,n}|U_{2,n}) - Nh(Z_1^{'}) - Nh(Z_2^{'}) + \mathrm{o}(N). \tag{60c}$$

Adding (59a), (59b) and (59g) and plugging into (58) yields the first bound for the weighted sum rate:

$$\begin{aligned} N(2R_1+R_2) \leqslant & \sum_{n=1}^{N} h(Y_{1,n}) + \sum_{n=1}^{N} h(Y_{1,n}|U_{1,n},X_{2,n}) + \sum_{n=1}^{N} h(Y_{2,n}|U_{2,n}) \\ & - N\left(h(Z_1)+h(Z_2)+h(Z_1^{'})+h(Z_2^{'})\right) + \mathrm{o}(N). \end{aligned} \tag{60d}$$

Adding (59e), (59f) and (59c) and plugging into (58) yields the second bound for the weighted sum rate :

$$\begin{aligned} N(R_1+2R_2) \leqslant & \sum_{n=1}^{N} h(Y_{2,n}) + \sum_{n=1}^{N} h(Y_{2,n}|U_{2,n},X_{1,n}) + \sum_{n=1}^{N} h(Y_{1,n}|U_{1,n}) \\ & - N\left(h(Z_1)+h(Z_2)+h(Z_1^{'})+h(Z_2^{'})\right) + \mathrm{o}(N). \end{aligned} \tag{60e}$$

Hence, the information rates can be upper bounded as follows:

$$NR_1 \leqslant \sum_{n=1}^{N} h(Y_{1,n}|X_{2,n}) - Nh(Z_1) + \mathrm{o}(N), \tag{61a}$$

$$NR_2 \leqslant \sum_{n=1}^{N} h(Y_{2,n}|X_{1,n}) - Nh(Z_2) + \mathrm{o}(N), \tag{61b}$$

$$N(R_1+R_2) \leqslant \sum_{n=1}^{N} h(Y_{1,n}) + \sum_{n=1}^{N} h(Y_{2,n}|U_{2,n},X_{1,n}) - Nh(Z_2) - Nh(Z_1^{'}) + \mathrm{o}(N), \tag{61c}$$

$$N(R_1+R_2) \leqslant \sum_{n=1}^{N} h(Y_{2,n}) + \sum_{n=1}^{N} h(Y_{1,n}|U_{1,n},X_{2,n}) - Nh(Z_1) - Nh(Z_2^{'}) + \mathrm{o}(N), \tag{61d}$$

$$N(R_1+R_2) \leqslant \sum_{n=1}^{N} h(Y_{1,n}|U_{1,n}) + \sum_{n=1}^{N} h(Y_{2,n}|U_{2,n}) - Nh(Z_1^{'}) - Nh(Z_2^{'}) + \mathrm{o}(N), \tag{61e}$$

$$\begin{aligned} N(2R_1+R_2) \leqslant & \sum_{n=1}^{N} h(Y_{1,n}) + \sum_{n=1}^{N} h(Y_{1,n}|U_{1,n},X_{2,n}) + \sum_{n=1}^{N} h(Y_{2,n}|U_{2,n}) \\ & - N\left(h(Z_1)+h(Z_2)+h(Z_1^{'})+h(Z_2^{'})\right) + \mathrm{o}(N), \end{aligned} \tag{61f}$$

$$\begin{aligned} N(R_1+2R_2) \leqslant & \sum_{n=1}^{N} h(Y_{2,n}) + \sum_{n=1}^{N} h(Y_{2,n}|U_{2,n},X_{1,n}) + \sum_{n=1}^{N} h(Y_{1,n}|U_{1,n}) \\ & - N\left(h(Z_1)+h(Z_2)+h(Z_1^{'})+h(Z_2^{'})\right) + \mathrm{o}(N). \end{aligned} \tag{61g}$$

Using assumption (13b), for a given $\epsilon_N > 0$, for any $\eta > 0$ there exist $N_0(\eta)$ such that for any $N \geq N_0(\eta)$ it holds that

$$\Pr\left[B^{(N)} < B - \epsilon_N\right] < \eta. \tag{62}$$

Equivalently,

$$\Pr\left[B^{(N)} \geq B - \epsilon_N\right] \geq 1 - \eta. \tag{63}$$

From Markov's inequality [7], the following holds:

$$(B-\epsilon_N)\Pr\left[B^{(N)} \geq B - \epsilon_N\right] \leq E[B^{(N)}]. \tag{64}$$

Combining (63) and (64) yields

$$(B-\epsilon_N)(1-\eta) \leq E[B^{(N)}], \tag{65}$$

which can be written as

$$(B-\delta_N) \leq E[B^{(N)}], \tag{66}$$

for some $\delta_N > \epsilon_N$ (for sufficiently large N).
In the following, for all $n \in \mathbb{N}$, the bounds in (71) and (66) are evaluated assuming that the channel inputs $X_{1,n}$ and $X_{2,n}$ are arbitrary independent with

$$\mu_{i,n} \triangleq \mathrm{E}[X_{i,n}], \tag{67}$$

$$\gamma_{i,n}^2 \triangleq \mathrm{Var}[X_{i,n}], \tag{68}$$

for $n \in \{1, \dots, N\}$ and for $i \in \{1, 2\}$. The input sequence must satisfy the input power constraint in (4), which can be written, for $i \in \{1, 2\}$, as

$$\frac{1}{n}\sum_{n=1}^{N} \mathrm{E}[X_{i,n}^2] = \left(\frac{1}{N}\sum_{n=1}^{N} \gamma_{i,n}^2\right) + \left(\frac{1}{N}\sum_{n=1}^{N} \mu_{i,n}^2\right) \leqslant P_i. \tag{69}$$

Using these assumptions the following holds:

$$\begin{aligned}
h(Y_{i,n}) &\leqslant \frac{1}{2}\log\Big(2\pi e \mathrm{Var}[Y_{i,n}]\Big) \\
&= \frac{1}{2}\log\Big(2\pi e(\sigma^2 + h_{ii}^2\gamma_{i,n}^2 + h_{ij}^2\gamma_{j,n}^2)\Big), && \text{(70a)} \\
h(Y_{i,n}|U_{i,n}, X_{j,n}) &\leqslant \frac{1}{2}\log\Big(2\pi e \mathrm{Var}[Y_{i,n}|U_{i,n}, X_{j,n}]\Big) \\
&\leqslant \frac{1}{2}\log\Big(2\pi e \mathrm{Var}[h_{ii}X_{i,n} + Z_{i,n}, U_{i,n}]\Big) - \frac{1}{2}\log\Big(2\pi e \mathrm{Var}[U_{i,n}]\Big) \\
&= \frac{1}{2}\log\left(\frac{\sigma_j^2\gamma_i^2 h_{ii}^2 + \sigma_1^2\gamma_i^2 h_{ji}^2 + \sigma_i^2\sigma_j^2}{\gamma_i^2 h_{ji}^2 + \sigma_j^2}\right) \\
&= \frac{1}{2}\log\left(1 + \frac{\frac{h_{ii}^2\gamma_{i,n}^2}{\sigma_i^2}}{1 + \frac{h_{ji}^2\gamma_{i,n}^2}{\sigma_j^2}}\right) + \frac{1}{2}\log(2\pi e\sigma_i^2\sigma_j^2), && \text{(70b)} \\
h(Y_{i,n}|U_{i,n}) &\leqslant \frac{1}{2}\log\Big(2\pi e \mathrm{Var}[Y_{i,n}|U_{i,n}]\Big) \\
&\leqslant \frac{1}{2}\log\Big(2\pi e \mathrm{Var}\left[Y_{i,n}, U_{i,n}\right]\Big) - \frac{1}{2}\log\Big(2\pi e \mathrm{Var}[U_{i,n}]\Big) \\
&= \frac{1}{2}\log\left(\frac{\sigma_j^2\gamma_i^2 h_{ii}^2 + \sigma_i^2\gamma_{i,n}^2 h_{ji}^2 + h_{ij}^2 h_{ji}^2\gamma_{i,n}^2\gamma_{j,n}^2 + \sigma_i^2\sigma_j^2}{\gamma_i^2 h_{ji}^2 + \sigma_j^2}\right) \\
&\leqslant \frac{1}{2}\log\left(1 + \frac{\frac{h_{ii}^2\gamma_{i,n}^2}{\sigma_i^2} + \frac{h_{ij}^2\gamma_{j,n}^2}{\sigma_i^2} + \frac{\gamma_{i,n}^2\gamma_{j,n}^2 h_{ij}^2 h_{ji}^2}{\sigma_i^2\sigma_j^2}}{1 + \frac{\gamma_{i,n}^2 h_{ji}^2}{\sigma_j^2}}\right) + \frac{1}{2}\log(2\pi e\sigma_i^2\sigma_j^2), && \text{(70c)}
\end{aligned}$$

Finally, plugging (70) in (61), it yields:

$$\begin{aligned}
NR_1 &\leqslant \sum_{n=1}^{N}\frac{1}{2}\log\left(1 + \frac{\gamma_{1,n}^2 h_{11}^2}{\sigma_1^2}\right) + \mathrm{o}(N), && \text{(71a)} \\
NR_2 &\leqslant \sum_{n=1}^{N}\frac{1}{2}\log\left(1 + \frac{\gamma_{2,n}^2 h_{22}^2}{\sigma_2^2}\right) + \mathrm{o}(N), && \text{(71b)} \\
N(R_1 + R_2) &\leqslant \sum_{n=1}^{N}\log\left(1 + \frac{h_{11}^2\gamma_{1,n}^2}{\sigma_1^2} + \frac{h_{12}^2\gamma_{2,n}^2}{\sigma_1^2}\right) + \sum_{n=1}^{N}\frac{1}{2}\log\left(1 + \frac{\frac{h_{22}^2\gamma_{2,n}^2}{\sigma_2^2}}{1 + \frac{h_{12}^2\gamma_{2,n}^2}{\sigma_1^2}}\right) + \mathrm{o}(N), && \text{(71c)} \\
N(R_1 + R_2) &\leqslant \sum_{n=1}^{N}\log\left(1 + \frac{h_{22}^2\gamma_{2,n}^2}{\sigma_2^2} + \frac{h_{21}^2\gamma_{1,n}^2}{\sigma_2^2}\right) + \sum_{n=1}^{N}\frac{1}{2}\log\left(1 + \frac{\frac{h_{11}^2\gamma_{1,n}^2}{\sigma_1^2}}{1 + \frac{h_{21}^2\gamma_{1,n}^2}{\sigma_2^2}}\right) + \mathrm{o}(N), && \text{(71d)}
\end{aligned}$$

$$N(R_1+R_2) \leqslant \sum_{n=1}^{N}\frac{1}{2}\log\left(1+\frac{\frac{h_{11}^2\gamma_{1,n}^2}{\sigma_1^2}+\frac{h_{12}^2\gamma_{2,n}^2}{\sigma_1^2}+\frac{\gamma_{1,n}^2\gamma_{2,n}^2h_{12}^2h_{21}^2}{\sigma_1^2\sigma_2^2}}{1+\frac{\gamma_{1,n}^2h_{21}^2}{\sigma_2^2}}\right) \tag{71e}$$
$$+\sum_{n=1}^{N}\frac{1}{2}\log\left(1+\frac{\frac{h_{22}^2\gamma_{2,n}^2}{\sigma_2^2}+\frac{h_{21}^2\gamma_{1,n}^2}{\sigma_2^2}+\frac{\gamma_{1,n}^2\gamma_{2,n}^2h_{12}^2h_{21}^2}{\sigma_1^2\sigma_2^2}}{1+\frac{\gamma_{2,n}^2h_{12}^2}{\sigma_1^2}}\right)+\mathrm{o}(N),$$

$$N(2R_1+R_2) \leqslant \sum_{n=1}^{N}\frac{1}{2}\log\left(1+\frac{\frac{h_{11}^2\gamma_{1,n}^2}{\sigma_1^2}}{1+\frac{h_{21}^2\gamma_{1,n}^2}{\sigma_2^2}}\right)+\sum_{n=1}^{N}\log\left(1+\frac{h_{11}^2\gamma_{1,n}^2}{\sigma_1^2}+\frac{h_{12}^2\gamma_{2,n}^2}{\sigma_1^2}\right) \tag{71f}$$
$$+\sum_{n=1}^{N}\frac{1}{2}\log\left(1+\frac{\frac{h_{22}^2\gamma_{2,n}^2}{\sigma_2^2}+\frac{h_{21}^2\gamma_{1,n}^2}{\sigma_2^2}+\frac{\gamma_{1,n}^2\gamma_{2,n}^2h_{12}^2h_{21}^2}{\sigma_1^2\sigma_2^2}}{1+\frac{\gamma_{2,n}^2h_{12}^2}{\sigma_1^2}}\right)+\mathrm{o}(N),$$

$$N(R_1+2R_2) \leqslant \sum_{n=1}^{N}\frac{1}{2}\log\left(1+\frac{\frac{h_{22}^2\gamma_{2,n}^2}{\sigma_2^2}}{1+\frac{h_{12}^2\gamma_{2,n}^2}{\sigma_1^2}}\right)+\sum_{n=1}^{N}\log\left(1+\frac{h_{22}^2\gamma_{2,n}^2}{\sigma_2^2}+\frac{h_{21}^2\gamma_{1,n}^2}{\sigma_2^2}\right) \tag{71g}$$
$$+\sum_{n=1}^{N}\frac{1}{2}\log\left(1+\frac{\frac{h_{11}^2\gamma_{1,n}^2}{\sigma_1^2}+\frac{h_{12}^2\gamma_{2,n}^2}{\sigma_1^2}+\frac{\gamma_{1,n}^2\gamma_{2,n}^2h_{12}^2h_{21}^2}{\sigma_1^2\sigma_2^2}}{1+\frac{\gamma_{1,n}^2h_{21}^2}{\sigma_2^2}}\right)+\mathrm{o}(N).$$

The expectation of the average received energy rate is given by

$$\begin{aligned}\mathrm{E}\left[B^{(N)}\right] &= \mathrm{E}\left[\frac{1}{N}\sum_{n=1}^{N}Y_{3,n}^2\right]\\ &= \sigma_3^2+h_{31}^2\left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{1,n}^2+\mu_{1,n}^2)\right)+h_{32}^2\left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{2,n}^2+\mu_{2,n}^2)\right)\\ &\quad+2h_{31}h_{32}\frac{1}{N}\sum_{n=1}^{N}\mu_{1,n}\mu_{2,n}.\end{aligned} \tag{72}$$

Using Cauchy-Schwarz inequality, the energy rate in (72) can be upper-bounded as follows:

$$\begin{aligned}\mathrm{E}\left[B^{(N)}\right] &\leqslant \sigma_3^2+h_{31}^2\left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{1,n}^2+\mu_{1,n}^2)\right)+h_{32}^2\left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{2,n}^2+\mu_{2,n}^2)\right)\\ &\quad+2h_{31}h_{32}\left(\frac{1}{N}\sum_{n=1}^{N}\mu_{1,n}^2\right)^{1/2}\left(\frac{1}{N}\sum_{n=1}^{N}\mu_{2,n}^2\right)^{1/2}.\end{aligned} \tag{73}$$

Combining (66) and (73) yields the following upper-bound on the energy rate B:

$$\begin{aligned}B &\leqslant \sigma_3^2+h_{31}^2\left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{1,n}^2+\mu_{1,n}^2)\right)+h_{32}^2\left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{2,n}^2+\mu_{2,n}^2)\right)\\ &\quad+2h_{31}h_{32}\left(\frac{1}{N}\sum_{n=1}^{N}\mu_{1,n}^2\right)^{1/2}\left(\frac{1}{N}\sum_{n=1}^{N}\mu_{2,n}^2\right)^{1/2}+\delta_N.\end{aligned} \tag{74}$$

In order to obtain a single-letterization of the upper-bound given by constraints (71) and (74), define also

$$\mu_i^2 \triangleq \frac{1}{N}\sum_{n=1}^{N}\mu_{i,t}^2, \quad i \in \{1,2\}, \tag{75}$$

$$\gamma_i^2 \triangleq \frac{1}{N}\sum_{n=1}^{N}\gamma_{i,n}^2, \quad i \in \{1,2\}, \tag{76}$$

Using these notations, the input power constraint in (125) can be rewritten as

$$\gamma_i^2 + \mu_i^2 \leqslant P_i, \text{ with } \quad i \in \{1,2\}. \tag{77}$$

By the concavity of the mutual information, applying Jensen's inequality to the bounds in (71) yields in the asymptotic blocklength regime:

$$R_1 \leqslant \frac{1}{2}\log\left(1+\frac{\gamma_1^2 h_{11}^2}{\sigma_1^2}\right), \tag{78a}$$

$$R_2 \leqslant \frac{1}{2}\log\left(1+\frac{\gamma_2^2 h_{22}^2}{\sigma_2^2}\right), \tag{78b}$$

$$R_1+R_2 \leqslant \log\left(1+\frac{h_{11}^2\gamma_1^2}{\sigma_1^2}+\frac{h_{12}^2\gamma_2^2}{\sigma_1^2}\right)+\frac{1}{2}\log\left(1+\frac{\frac{h_{22}^2\gamma_2^2}{\sigma_2^2}}{1+\frac{h_{12}^2\gamma_2^2}{\sigma_1^2}}\right), \tag{78c}$$

$$R_1+R_2 \leqslant \log\left(1+\frac{h_{22}^2\gamma_2^2}{\sigma_2^2}+\frac{h_{21}^2\gamma_1^2}{\sigma_2^2}\right)+\frac{1}{2}\log\left(1+\frac{\frac{h_{11}^2\gamma_1^2}{\sigma_1^2}}{1+\frac{h_{21}^2\gamma_1^2}{\sigma_2^2}}\right), \tag{78d}$$

$$\begin{aligned} R_1+R_2 \leqslant{}& \frac{1}{2}\log\left(1+\frac{\frac{h_{11}^2\gamma_1^2}{\sigma_1^2}+\frac{h_{12}^2\gamma_2^2}{\sigma_1^2}+\frac{\gamma_1^2\gamma_2^2h_{12}^2h_{21}^2}{\sigma_1^2\sigma_2^2}}{1+\frac{\gamma_1^2h_{21}^2}{\sigma_2^2}}\right) \\ &+\frac{1}{2}\log\left(1+\frac{\frac{h_{22}^2\gamma_2^2}{\sigma_2^2}+\frac{h_{21}^2\gamma_1^2}{\sigma_2^2}+\frac{\gamma_1^2\gamma_2^2h_{12}^2h_{21}^2}{\sigma_1^2\sigma_2^2}}{1+\frac{\gamma_{2,n}^2h_{12}^2}{\sigma_1^2}}\right), \end{aligned} \tag{78e}$$

$$\begin{aligned} 2R_1+R_2 \leqslant{}& \frac{1}{2}\log\left(1+\frac{\frac{h_{11}^2\gamma_1^2}{\sigma_1^2}}{1+\frac{h_{21}^2\gamma_1^2}{\sigma_2^2}}\right)+\log\left(1+\frac{h_{11}^2\gamma_1^2}{\sigma_1^2}+\frac{h_{12}^2\gamma_2^2}{\sigma_1^2}\right) \\ &+\frac{1}{2}\log\left(1+\frac{\frac{h_{22}^2\gamma_2^2}{\sigma_2^2}+\frac{h_{21}^2\gamma_1^2}{\sigma_2^2}+\frac{\gamma_1^2\gamma_2^2h_{12}^2h_{21}^2}{\sigma_1^2\sigma_2^2}}{1+\frac{\gamma_2^2h_{12}^2}{\sigma_1^2}}\right), \end{aligned} \tag{78f}$$

$$\begin{aligned} R_1+2R_2 \leqslant{}& \frac{1}{2}\log\left(1+\frac{\frac{h_{22}^2\gamma_2^2}{\sigma_2^2}}{1+\frac{h_{12}^2\gamma_2^2}{\sigma_1^2}}\right)+\log\left(1+\frac{h_{22}^2\gamma_2^2}{\sigma_2^2}+\frac{h_{21}^2\gamma_1^2}{\sigma_2^2}\right) \\ &+\frac{1}{2}\log\left(1+\frac{\frac{h_{11}^2\gamma_1^2}{\sigma_1^2}+\frac{h_{12}^2\gamma_2^2}{\sigma_1^2}+\frac{\gamma_1^2\gamma_2^2h_{12}^2h_{21}^2}{\sigma_1^2\sigma_2^2}}{1+\frac{\gamma_1^2h_{21}^2}{\sigma_2^2}}\right), \end{aligned} \tag{78g}$$

and the upper-bound on the energy rate (74) yields

$$B \leqslant \sigma_3^2 + h_{31}^2(\gamma_1^2+\mu_1^2)+h_{32}^2(\gamma_2^2+\mu_2^2)+2h_{31}h_{32}|\mu_1||\mu_2|. \tag{78h}$$

To sum up, it has been shown so far that, in the limit when N tends to infinity, any information-energy rate triplet
$(R_1, R_2, B) \in \mathcal{E}_b^{\mathrm{NFB}}$ can be bounded by the constraints in (78) for some γ_1, γ_2, μ_1, μ_2 satisfying (77). Let $\mathcal{R}_b(\gamma_1, \gamma_2, \mu_1, \mu_2)$ denote the set of information-energy rate triplets satisfying (78) for some γ_1, γ_2, μ_1, μ_2 such that (77) is true. Thus, it holds that

$$\overline{\mathcal{E}}_b \subseteq \bigcup_{\substack{0 \leqslant \gamma_1^2 + \mu_1^2 \leqslant P_1 \\ 0 \leqslant \gamma_2^2 + \mu_2^2 \leqslant P_2}} \mathcal{R}_b(\gamma_1, \gamma_2, \mu_1, \mu_2). \tag{79}$$

In this union, it suffices to consider $\mu_1 \geqslant 0$, $\mu_2 \geqslant 0$, and γ_1, γ_2, μ_1, and μ_2 that saturate the input power constraint (i.e., (77) holds with equality). Thus,

$$\overline{\mathcal{E}}_b \subseteq \bigcup_{\substack{0 \leqslant \gamma_1^2 + \mu_1^2 \leqslant P_1 \\ 0 \leqslant \gamma_2^2 + \mu_2^2 \leqslant P_2}} \mathcal{R}_b(\gamma_1^2, \gamma_2^2, \mu_1, \mu_2) \subseteq \bigcup_{\substack{\sigma_1^2 + \mu_1^2 = P_1 \\ \gamma_2^2 + \mu_2^2 = P_2}} \mathcal{R}_b(\gamma_1^2, \gamma_2^2, \mu_1, \mu_2). \tag{80}$$

Let $\beta_i \in [0, 1]$ be defined as follows:

$$\beta_i \triangleq \frac{\gamma_i^2}{P_i} = \frac{P_i - \mu_i^2}{P_i}, \quad i \in \{1, 2\}. \tag{81}$$

With these notations, any region $\mathcal{R}_b(\gamma_1^2, \gamma_2^2, \mu_1, \mu_2)$ in the union over all $(\mu_1, \mu_2, \gamma_1, \gamma_2)$ that satisfy $\gamma_1^2 + \mu_1^2 = P_1$ and $\gamma_2^2 + \mu_2^2 = P_2$, can be rewritten as follows:

$$R_1 \leqslant \frac{1}{2} \log\left(1 + \frac{\beta_1 P_1 h_{11}^2}{\sigma_1^2}\right), \tag{82a}$$

$$R_2 \leqslant \frac{1}{2} \log\left(1 + \frac{\beta_2 P_2 h_{22}^2}{\sigma_2^2}\right), \tag{82b}$$

$$R_1 + R_2 \leqslant \log\left(1 + \frac{h_{11}^2 \beta_1 P_1}{\sigma_1^2} + \frac{h_{12}^2 \beta_2 P_2}{\sigma_1^2}\right) + \frac{1}{2} \log\left(1 + \frac{\frac{h_{22}^2 \beta_2 P_2}{\sigma_2^2}}{1 + \frac{h_{12}^2 \beta_2 P_2}{\sigma_1^2}}\right), \tag{82c}$$

$$R_1 + R_2 \leqslant \log\left(1 + \frac{h_{22}^2 \beta_2 P_2}{\sigma_2^2} + \frac{h_{21}^2 \beta_1 P_1}{\sigma_2^2}\right) + \frac{1}{2} \log\left(1 + \frac{\frac{h_{11}^2 \beta_1 P_1}{\sigma_1^2}}{1 + \frac{h_{21}^2 \beta_1 P_1}{\sigma_2^2}}\right), \tag{82d}$$

$$\begin{aligned} R_1 + R_2 \leqslant{}& \frac{1}{2} \log\left(1 + \frac{\frac{h_{11}^2 \beta_1 P_1}{\sigma_1^2} + \frac{h_{12}^2 \beta_2 P_2}{\sigma_1^2} + \frac{\beta_1 \beta_2 P_1 P_2 h_{12}^2 h_{21}^2}{\sigma_1^2 \sigma_2^2}}{1 + \frac{\beta_1 P_1 h_{21}^2}{\sigma_2^2}}\right) \\ &+ \frac{1}{2} \log\left(1 + \frac{\frac{h_{22}^2 \beta_2 P_2}{\sigma_2^2} + \frac{h_{21}^2 \beta_1 P_1}{\sigma_2^2} + \frac{\beta_1 \beta_2 P_1 P_2 h_{12}^2 h_{21}^2}{\sigma_1^2 \sigma_2^2}}{1 + \frac{\beta_2 P_2 h_{12}^2}{\sigma_1^2}}\right), \end{aligned} \tag{82e}$$

$$\begin{aligned} 2R_1 + R_2 \leqslant{}& \frac{1}{2} \log\left(1 + \frac{\frac{h_{11}^2 \beta_1 P_1}{\sigma_1^2}}{1 + \frac{h_{21}^2 \beta_1 P_1}{\sigma_2^2}}\right) + \log\left(1 + \frac{h_{11}^2 \beta_1 P_1}{\sigma_1^2} + \frac{h_{12}^2 \beta_2 P_2}{\sigma_1^2}\right) \\ &+ \frac{1}{2} \log\left(1 + \frac{\frac{h_{22}^2 \beta_2 P_2}{\sigma_2^2} + \frac{h_{21}^2 \beta_1 P_1}{\sigma_2^2} + \frac{\beta_1 \beta_2 P_1 P_2 h_{12}^2 h_{21}^2}{\sigma_1^2 \sigma_2^2}}{1 + \frac{\beta_2 P_2 h_{12}^2}{\sigma_1^2}}\right), \end{aligned} \tag{82f}$$

$$R_1+2R_2 \leqslant \frac{1}{2}\log\left(1+\frac{\frac{h_{22}^2\beta_2P_2}{\sigma_2^2}}{1+\frac{h_{12}^2\beta_2P_2}{\sigma_1^2}}\right)+\log\left(1+\frac{h_{22}^2\beta_2P_2}{\sigma_2^2}+\frac{h_{21}^2\beta_1P_1}{\sigma_2^2}\right) \tag{82g}$$
$$+\frac{1}{2}\log\left(1+\frac{\frac{h_{11}^2\beta_1P_1}{\sigma_1^2}+\frac{h_{12}^2\beta_2P_2}{\sigma_1^2}+\frac{\beta_1\beta_2P_1P_2h_{12}^2h_{21}^2}{\sigma_1^2\sigma_2^2}}{1+\frac{\beta_1P_1h_{21}^2}{\sigma_2^2}}\right),$$
$$B \leqslant \sigma_3^2+h_{31}^2P_1+h_{32}^2P_2+2h_{31}h_{32}\sqrt{(1-\beta_1)P_1(1-\beta_2)P_2}, \tag{82h}$$
$$B \geqslant b, \tag{82i}$$

for some $(\beta_1,\beta_2)\in[0,1]^2$. Hence, such a region contains all information-energy rate triplets (R_1,R_2,B) satisfying the constraints of Theorem 2 and this completes the proof of Theorem 2.

C Proof of Theorem 3

This proof is divided into two parts. The first part consists of the proof of (17a)-(17f). The second part consists of the proof of (17g).

C.1 Proof of (17a)-(17f)

Lemma 2 *For all $(R_1,R_2,B)\in\underline{\mathcal{E}}_b$, the following holds:*

$$R_1 \leq I(U,X_1,U_2;Y_1|V), \tag{83}$$
$$R_1 \leq I(U_1;Y_2|U,X_2,V)+I(X_1;Y_1|U_1,U_2,U,V), \tag{84}$$
$$R_2 \leq I(U,X_2,U_1;Y_2|V), \tag{85}$$
$$R_2 \leq I(U_2;Y_1|U,X_1,V)+I(X_2;Y_2|U_1,U_2,U,V), \tag{86}$$
$$R_1+R_2 \leq I(X_1;Y_1|U_1,U_2,U,V)+I(V,U_2,X_1;Y_1), \tag{87}$$
$$R_1+R_2 \leq I(X_2;Y_2|U_1,U_2,U,V)+I(V,U_2,X_1;Y_1), \tag{88}$$

over all joint distribution: $P_{VUU_1U_2S_1S_2}(v,u,u_1,u_2,s_1,s_2)=$
$P_V(v)P_{U|V}(u|v)P_{U_1|UV}(u_1|u,v)P_{U_2|UV}(u_2|u,v)P_{S_1|U_1UV}(s_1|u_1,u,v)P_{S_2|U_2UV}(s_2|u_2,u,v)$.

Proof: **Codebook Generation**: Fix a strictly positive joint probability distribution: $P_{VUU_1U_2S_1S_2}(v,u,u_1,u_2,s_1,s_2)=P_V(v)P_{U|V}(u|v)P_{U_1|UV}(u_1|u,v)P_{U_2|U,V}(u_2|u,v)$ $P_{S_1|UU_1V}(x_1|u,u_1,v)P_{S_2|UU_2V}(s_2|u,u_2,v)$, for all $(v,u,u_1,u_2,x_1,x_2)\in\mathcal{X}_1\cap\mathcal{X}_2\times(\mathcal{X}_1\cap\mathcal{X}_2)\times\mathcal{X}_1\times\mathcal{X}_2\times\mathcal{X}_1\times\mathcal{X}_2$. Let $R_E,R_{1,C},R_{2,C},R_{1,P}$ and $R_{2,P}$ be non-negative real numbers. For transmitter 1, generate 2^{NR_E} i.i.d N-length codewords $\boldsymbol{v}(\omega)=(v_1(\omega),\ldots,v_N(\omega))$ according to

$$P_{\boldsymbol{V}}(\boldsymbol{v}(\omega))=\prod_{m=1}^{N}P_V(v_m(\omega)), \tag{89}$$

with $\omega\in\{1,2,\ldots,2^{NR_E}\}$. For each codeword $\boldsymbol{v}(\omega)$, generate $2^{N(R_{1,C}+R_{2,C})}$ i.i.d. N-length codewords $\boldsymbol{u}(\omega,s,r)=(u_1(\omega,s,r),\ldots,u_N(\omega,s,r))$ according to

$$P_{\boldsymbol{U}|\boldsymbol{V}}(\boldsymbol{u}(s,r,\omega)|\boldsymbol{v}(\omega))=\prod_{m=1}^{N}P_{U|V}(u_m(s,r,\omega)|v_m(\omega)), \tag{90}$$

with $s \in \{1, \ldots, 2^{NR_{1,C}}\}$ and $r \in \{1, \ldots, 2^{NR_{2,C}}\}$. For transmitter 1, for each codeword $\boldsymbol{u}(\omega, s, r)$, generate $2^{NR_{1,C}}$ i.i.d. N-length codewords $\boldsymbol{u}_1(\omega, s, r, k) = (u_{1,1}(\omega, s, r, k), \ldots, u_{1,N}(\omega, s, r, k))$ according to

$$P_{\boldsymbol{U_1}|\boldsymbol{UV}}\Big(\boldsymbol{u}_1(s, r, k, \omega)|\boldsymbol{u}(s, r), \boldsymbol{v}(\omega)\Big) = \prod_{m=1}^{N} P_{U_1|UV}\Big(u_{1,m}(s, r, k, \omega)|u_m(s, r, \omega), v_m(\omega)\Big), \quad (91)$$

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

$$\begin{aligned} &P_{\boldsymbol{S_1}|\boldsymbol{U_1UV}}\Big(\boldsymbol{s}_1(\omega, s, r, k, l)|\boldsymbol{u}_1(\omega, s, r, k), \boldsymbol{u}(s, r), \boldsymbol{v}(\omega)\Big) = \\ &\prod_{m=1}^{N} P_{\boldsymbol{S}_1|\boldsymbol{U}_1\boldsymbol{UV}}\Big(s_{1,m}(\omega, s, r, k, l)|u_{1,m}(\omega, s, r, k), u_m(\omega, s, r), v_m(\omega)\Big), \end{aligned} \quad (92)$$

with $l \in \{1, \ldots, 2^{NR_{1,P}}\}$.
For encoder 2, for each codeword $\boldsymbol{u}(\omega, s, r)$, generate $2^{NR_{2,C}}$ i.i.d. N-length codewords $\boldsymbol{u}_2(\omega, s, r, q) = (u_{2,1}(\omega, s, r, q), \ldots, u_{2,1}(\omega, s, r, q))$ according to

$$P_{\boldsymbol{U_2}|\boldsymbol{UV}}(\boldsymbol{u}_2(s, r, q, \omega)|\boldsymbol{u}(s, r), \boldsymbol{v}(\omega)) = \prod_{m=1}^{N} P_{U_2|UV}\Big(u_{2,m}(\omega, s, r, q)|u_m(\omega, s, r), v_m(\omega)\Big), \quad (93)$$

with $q \in \{1, \ldots, 2^{NR_{2,C}}\}$. For each tuple of codewords $(\boldsymbol{v}(\omega), \boldsymbol{u}(\omega, s, r), \boldsymbol{u}_2(\omega, s, r, q))$, generate $2^{NR_{2,P}}$ i.i.d. N-length codewords $\boldsymbol{s}_2(\omega, s, r, q, z) = (s_{2,1}(\omega, s, r, q, z), \ldots, s_{2,N}(\omega, s, r, q, z))$ according to

$$\begin{aligned} &P_{\boldsymbol{S}_2|\boldsymbol{U}_2\boldsymbol{UV}}(\boldsymbol{s}_2(\omega, s, r, q, z)|\boldsymbol{u}_2(\omega, s, r, q), \boldsymbol{u}(\omega, s, r), \boldsymbol{v}(\omega)) \\ &= \prod_{m=1}^{N} P_{S_2|U_2UV}(s_{2,m}\Big(\omega, s, r, q, z)|u_{2,m}(\omega, s, r, q), u_m(\omega, s, r), v_m(\omega)\Big), \end{aligned} \quad (94)$$

with $z \in \{1, \ldots, 2^{NR_{2,P}}\}$.
Encoding: Let $W_i^{(t)}$ be represented by the message index $W_{i,C}^{(t)} \in \{1, 2, \ldots, 2^{NR_{i,C}}\}$ and the message index $W_{i,P}^{(t)} \in \{1, 2, \ldots, 2^{NR_{i,P}}\}$. The message index $W_{i,P}^{(t)}$ must be reliably decoded at receiver i and the message index $\Omega^{(t)}$ is known by both transmitters and receivers. The index $W_{i,C}^{(t-1)}$ must be reliably decoded by transmitter j (via feedback) but not necessarily by receiver i.
Consider Markov encoding over T blocks. At encoding step t, with $t \in \{1, 2, \ldots, T\}$, transmitter 1 sends the codeword $\boldsymbol{x}_1^{(t)} = \theta_1\Big((\boldsymbol{v}^{(t)}(\Omega), \boldsymbol{u}\left(W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}, \Omega\right), \boldsymbol{u}_1\left(W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}, W_{1,C}^{(t)}, \Omega\right),$ $\boldsymbol{s}_1\left(W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}, W_{1,C}^{(t)}, W_{1,P}^{(t)}, \Omega\right)\Big)$, where $\theta_1 : \mathcal{X}^N \times (\mathcal{X}_1 \cup \mathcal{X}_2)^N \times \mathcal{X}_1 \times \mathcal{X}_1^N \to \mathcal{X}_1^N$ is a function that transforms the codewords $\boldsymbol{v}^{(t)}(\Omega), \boldsymbol{u}\left(W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}, \Omega\right), \boldsymbol{u}_1\left(W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}, W_{1,C}^{(t)}, \Omega\right)$, and $\boldsymbol{s}_1\left(W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}, W_{1,C}^{(t)}, W_{1,P}^{(t)}, \Omega\right)$ into the N-dimensional vector $\boldsymbol{x}_1^{(t)}$. The indices $W_{1,C}^{(0)} = W_{1,C}^{(T)} = s^*$ and $W_{2,C}^{(0)} = W_{2,C}^{(T)} = r^*$, and the pair $(s^*, r^*) \in \{1, 2, \ldots, 2^{NR_{1,C}}\} \times \{1, 2, \ldots, 2^{NR_{2,C}}\}$ are pre-defined and known by both receivers and transmitters. Transmitter 2 follows a similar encoding scheme.

Decoding: Both receivers decode their message indices at the end of block T in a backward decoding fashion. At each decoding step t, with $t \in \{1, 2, \ldots, T\}$, receiver 1 obtains the indices $\left(\widehat{W}_{1,C}^{(T-t)}, \widehat{W}_{2,C}^{(T-t)}, W_{1,P}^{(T-(t-1))}\right)$ from the channel output $\boldsymbol{y}_1$.

The tuple $\left(\widehat{W}_{1,C}^{(T-t)}, \widehat{W}_{2,C}^{(T-t)}, W_{1,P}^{(T-(t-1))}\right)$ is the unique tuple that satisfy:

$$\Big(\big(\boldsymbol{v}(\Omega^{(t)}), \boldsymbol{u}\left(\Omega^{(t)}, \widehat{W}_{1,C}^{(T-t)}, \widehat{W}_{2,C}^{(T-t)}\right), \boldsymbol{u}_1\left(\Omega^{(t)}, \widehat{W}_{1,C}^{(T-t)}, \widehat{W}_{2,C}^{(T-t)}, W_{1,C}^{(T-(t-1))}\right),$$
$$\boldsymbol{s}_1\left(\Omega^{(t)}, \widehat{W}_{1,C}^{(T-t)}, \widehat{W}_{2,C}^{(T-t)}, W_{1,C}^{(T-(t-1))}, W_{1,P}^{(T-(t-1))}\right), \boldsymbol{u}_2\left(\Omega^{(t)}, \widehat{W}_{1,C}^{(T-t)}, \widehat{W}_{2,C}^{(T-t)}, W_{2,C}^{(T-(t-1))}\right),$$
$$\boldsymbol{y}_1^{(T-(t-1))}\Big) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}, \tag{95}$$

where $W_{1,C}^{(T-(t-1))}$ and $W_{2,C}^{(T-(t-1))}$ are assumed to be perfectly decoded in the previous decoding step $t-1$. The set $\mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}$ represent the set of jointly typical sequences of the random variables V, U, U_1, S_1, U_2, and Y_1, with $\epsilon > 0$. Finally, receiver 2 follows a similar decoding scheme.

Probability of Error Analysis: An error might occur during encoding step t at transmitter 1 if the index $\widehat{W}_{2,C}^{(t-1)}$ is not correctly decoded. Define the event E_k that describes the case in which there exist another message index k that satisfy: $\Big(\boldsymbol{v}(\Omega^{(t)}), \boldsymbol{u}\left(\Omega^{(t)}, W_{1,C}^{(t-2)}, W_{2,C}^{(t-2)}\right)$, $\boldsymbol{u}_1\left(\Omega^{(t)}, W_{1,C}^{(t-2)}, W_{2,C}^{(t-2)}, W_{1,C}^{(t-1)}, \Omega\right)$, $\boldsymbol{s}_1\left(\Omega^{(t)}, W_{1,C}^{(t-2)}, W_{2,C}^{(t-2)}, W_{1,C}^{(t-1)}, W_{1,P}^{(t-1)}\right)$, $\boldsymbol{u}_2\left(\Omega^{(t)}, W_{1,C}^{(t-2)}, W_{2,C}^{(t-2)}, k\right)\Big) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}$, with $t \in \{2, 3, \ldots, T\}$ and $W_{2,C}^{(t-2)}$ is assumed to be perfectly decoded in the previous block $t-1$. Then, the probability of event E_k can be bounded as follows:

$$\begin{aligned}
\Pr(E_k) &\overset{(a)}{=} \Pr\left[(\boldsymbol{V}, \boldsymbol{U}, \boldsymbol{U}_1, \boldsymbol{X}_1, \boldsymbol{U}_2, \boldsymbol{Y}_1) \in \mathcal{T}_{VUU_1X_1U_2Y_1}^{(N,\epsilon)}\right] \\
&= \sum_{(\boldsymbol{v},\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{y}_1)\in\mathcal{T}_{VUU_1X_1U_2Y_1}^{(N,\epsilon)}} P_{\boldsymbol{V}}(\boldsymbol{v}) P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{Y}_1|\boldsymbol{V}}(\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x_1},\boldsymbol{y}_2|\boldsymbol{v}) P_{\boldsymbol{U_2}|\boldsymbol{V}}(\boldsymbol{u}_2|\boldsymbol{v}), \\
&\leq \left|\mathcal{T}_{VUU_1X_1U_2Y_1}^{(N,\epsilon)}\right| 2^{-N(H(U,U_1,X_1,Y_1|V)+H(U_2|V)+H(V)-4\epsilon)} \\
&\leq 2^{-N(H(U,U_1,X_1,Y_1|V)+H(U_2|V)+H(V)-H(U,U_1,U_2,X_1,Y_1)-4\epsilon)} \\
&\leq 2^{-N(H(U,U_1,X_1,Y_1|V)+H(U_2|V)-H(U,U_1,U_2,X_1,Y_1|V)-4\epsilon)} \\
&= 2^{-N(I(U,U_1,X_1,Y_1;U_2|V)-4\epsilon)} \\
&= 2^{-N(I(Y_1;U_2|X_1,V)-4\epsilon)},
\end{aligned} \tag{96}$$

where the probability operator Pr[.] in (a) applies with a probability distribution $P_{\boldsymbol{V}\boldsymbol{U}\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2\boldsymbol{Y}_1}$ that factorizes as $P_{\boldsymbol{V}} P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{Y}_1} P_{\boldsymbol{U}_2|\boldsymbol{V}}$ given that all the codewords $\boldsymbol{u}_2$ are independent from the output of the channel $\boldsymbol{y}_1$. The error probability becomes arbitrarily small (as N goes to infinity) if

$$R_{2c} \leq I(U_2, ; Y_1|X_1, U, V). \tag{97}$$

An error might occur during the (backward) decoding step t if the indices $\widehat{W}_{1,C}^{(T-t)}, \widehat{W}_{2,C}^{(T-t)}$ or $\widehat{W}_{1,P}^{(T-(t-1)}$ are not decoded correctly given that the indices $W_{1,C}^{(T-(t-1))}$ and $W_{2,C}^{(T-(t-1))}$ were correctly decoded in the previous decoding step $t-1$. These errors might arise for two reasons: (i) there does not exist a tuple $\left(\widehat{W}_{1,C}^{(T-t)}, \widehat{W}_{2,C}^{(T-t)}, W_{1,P}^{(T-(t-1)}\right)$ that satisfies (95), or (ii) there exist

several tuples $\left(\widehat{W}_{1,C}^{(T-t)},\widehat{W}_{2,C}^{(T-t)},W_{1,P}^{(T-(t-1)}\right)$ that simultaneously satisfy (95). From the asymptotic equipartition property [6], the probability of error due to (i) tends to zero when N grows to infinity. Consider the error due to (ii) and define the event E_{srl} that describes the case in which the codewords $\boldsymbol{v}(\Omega^{(t)}), \boldsymbol{u}\left(\Omega^{(t)},s,r\right), \boldsymbol{u}_1\left(\Omega^{(t)},s,r,W_{1,C}^{(T-(t-1))}\right), \boldsymbol{s}_1\left(\Omega^{(t)},s,r,W_{1,C}^{(T-(t-1))},l\right)$, and $\boldsymbol{u}_2\left(\Omega^{(t)},s,r,W_{2,C}^{(T-(t-1))}\right)$ are jointly typical with $\boldsymbol{y}_1^{(T-(t-1))}$ during decoding step t. Assume now that the codeword to be decoded at decoding step t corresponds to the indices $(s,r,l)=(1,1,1)$. This is without loss of generality due to the symmetry of the code. Then, the probability of error due to (ii) during decoding step t, can be bounded as follows:

$$\begin{aligned}
&\Pr\left[\bigcup_{(s,r,l)\neq(1,1,1)} E_{srl}\right]\\
&\leq \sum_{s\neq1,r\neq1,l\neq1}\Pr\left[E_{srl}\right]+\sum_{s\neq1,r\neq1,l=1}\Pr\left[E_{srl}\right]+\sum_{s\neq1,r=1,l\neq1}\Pr\left[E_{srl}\right]\\
&+\sum_{s\neq1,r=1,l=1}\Pr\left[E_{srl}\right]+\sum_{s=1,r\neq1,l\neq1}\Pr\left[E_{srl}\right]+\sum_{s=1,r\neq1,l=1}\Pr\left[E_{srl}\right]\\
&+\sum_{s=1,r=1,l\neq1}\Pr\left[E_{srl}\right],
\end{aligned}\tag{98}$$

For all $s\in\{2,3,\ldots,2^{NR_{1,C}}\}$, the following holds

$$\begin{aligned}
\Pr[E_{s11}]&\overset{(a)}{=}\Pr\left[(\boldsymbol{V},\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,\boldsymbol{Y}_1)\in\mathcal{T}_{VUU_1X_1U_2Y_1}^{(N,\epsilon)}\right]\\
&=\sum_{(\boldsymbol{v},\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{u}_2,\boldsymbol{y}_1)\in\mathcal{T}_{VUU_1X_1U_2Y_1}^{(N,\epsilon)}} P_{\boldsymbol{V}}(\boldsymbol{v})P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,|\boldsymbol{V}}(\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x_1},\boldsymbol{u}_2|\boldsymbol{v})P_{\boldsymbol{Y_1}|\boldsymbol{V}}(\boldsymbol{y}_1|\boldsymbol{v}),\\
&\leq\left|\mathcal{T}_{VUU_1X_1U_2Y_1}^{(N,\epsilon)}\right|2^{-N(H(U,U_1,U_2,X_1|V)+H(Y_1|V)+H(V)-4\epsilon)}\\
&\leq 2^{-N(H(U,U_1,U_2,X_1|V)+H(Y_1|V)+H(V)-H(U,U_1,U_2,X_1,Y_1)-4\epsilon)}\\
&=2^{-N(I(U,U_1,U_2,X_1;Y_1|V)-4\epsilon)}\\
&=2^{-N(I(U,U_2,X_1;Y_1|V)-4\epsilon)},
\end{aligned}\tag{99a}$$

where the probability operator Pr[.] in (a) applies with a probability distribution $P_{\boldsymbol{V}\boldsymbol{U}\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2\boldsymbol{Y}_1}$ that factorizes as $P_{\boldsymbol{V}}P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,|\boldsymbol{V}}P_{\boldsymbol{Y_1}|\boldsymbol{V}}$ given that all the codewords $\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1$, and $\boldsymbol{u}_2$ are independent from the output of the channel $\boldsymbol{y}_1$. For all $r\in\{2,3,\ldots,2^{NR_{2,C}}\}$, the following holds

$$\begin{aligned}
\Pr[E_{1r1}]&\overset{(a)}{=}\Pr\left[(\boldsymbol{V},\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,\boldsymbol{Y}_1)\in\mathcal{T}_{VUU_1X_1U_2Y_1}^{(N,\epsilon)}\right]\\
&=\sum_{(\boldsymbol{v},\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{u}_2,\boldsymbol{y}_1)\in\mathcal{T}_{VUU_1X_1U_2Y_1}^{(N,\epsilon)}} P_{\boldsymbol{V}}(\boldsymbol{v})P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,|\boldsymbol{V}}(\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x_1},\boldsymbol{u}_2|\boldsymbol{v})P_{\boldsymbol{Y_1}|\boldsymbol{V}}(\boldsymbol{y}_1|\boldsymbol{v}),\\
&\leq\left|\mathcal{T}_{VUU_1X_1U_2Y_1}^{(N,\epsilon)}\right|2^{-N(H(U,U_1,U_2,X_1|V)+H(Y_1|V)+H(V)-4\epsilon)}\\
&\leq 2^{-N(H(U,U_1,U_2,X_1|V)+H(Y_1|V)+H(V)-H(U,U_1,U_2,X_1,Y_1)-4\epsilon)}\\
&=2^{-N(I(U,U_1,U_2,X_1;Y_1|V)-4\epsilon)}\\
&=2^{-N(I(U,U_2,X_1;Y_1|V)-4\epsilon)}.
\end{aligned}\tag{99b}$$

where the probability operator Pr[.] in (a) applies with a probability distribution $P_{\boldsymbol{V}\boldsymbol{U}\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2\boldsymbol{Y}_1}$ that factorizes as $P_{\boldsymbol{V}}P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,|\boldsymbol{V}}P_{\boldsymbol{Y_1}|\boldsymbol{V}}$ given that all the codewords $\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1$, and $\boldsymbol{u}_2$ are

independent from the output of the channel $\boldsymbol{y}_1$. For all $s \in \{2,3,\ldots,2^{NR_{1,C}}\}$ and $r \in \{2,3,\ldots,2^{NR_{2,C}}\}$, the following holds

$$\begin{aligned}
\Pr[E_{sr1}] &\overset{(a)}{=} \Pr\left[(\boldsymbol{V},\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,\boldsymbol{Y}_1) \in \mathcal{T}^{(N,\epsilon)}_{VUU_1X_1U_2Y_1}\right] \\
&= \sum_{(\boldsymbol{v},\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{u}_2,\boldsymbol{y}_1)\in\mathcal{T}^{(N,\epsilon)}_{VUU_1X_1U_2Y_1}} P_{\boldsymbol{V}}(\boldsymbol{v})P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,|\boldsymbol{V}}(\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{u}_2|\boldsymbol{v})P_{\boldsymbol{Y_1}|\boldsymbol{V}}(\boldsymbol{y}_1|\boldsymbol{v}), \\
&\leq \left|\mathcal{T}^{(N,\epsilon)}_{VUU_1X_1U_2Y_1}\right| 2^{-N(H(U,U_1,U_2,X_1|V)+H(Y_1|V)+H(V)-4\epsilon)} \\
&\leq 2^{-N(H(U,U_1,U_2,X_1|V)+H(Y_1|V)+H(V)-H(U,U_1,U_2,X_1,Y_1)-4\epsilon)} \\
&= 2^{-N(I(U,U_1,U_2,X_1;Y_1|V)-4\epsilon)} \\
&= 2^{-N(I(U,U_2,X_1;Y_1|V)-4\epsilon)},
\end{aligned} \tag{99c}$$

where the probability operator Pr[.] applies in (a) with a probability distribution $P_{\boldsymbol{V}\boldsymbol{U}\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2\boldsymbol{Y}_1}$ that factorizes as $P_{\boldsymbol{V}}P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,|\boldsymbol{V}}P_{\boldsymbol{Y_1}|\boldsymbol{V}}$ given that all the codewords $\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1$, and $\boldsymbol{u}_2$ are independent from the output of the channel $\boldsymbol{y}_1$. For all $s \in \{2,3,\ldots,2^{NR_{1,C}}\}$ and $l \in \{2,3,\ldots,2^{NR_{1,P}}\}$, the following holds

$$\begin{aligned}
\Pr[E_{s1l}] &\overset{(a)}{=} \Pr\left[(\boldsymbol{V},\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,\boldsymbol{Y}_1) \in \mathcal{T}^{(N,\epsilon)}_{VUU_1X_1U_2Y_1}\right] \\
&= \sum_{(\boldsymbol{v},\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{u}_2,\boldsymbol{y}_1)\in\mathcal{T}^{(N,\epsilon)}_{VUU_1X_1U_2Y_1}} P_{\boldsymbol{V}}(\boldsymbol{v})P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,|\boldsymbol{V}}(\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{u}_2|\boldsymbol{v})P_{\boldsymbol{Y_1}|\boldsymbol{V}}(\boldsymbol{y}_1|\boldsymbol{v}), \\
&\leq \left|\mathcal{T}^{(N,\epsilon)}_{VUU_1X_1U_2Y_1}\right| 2^{-N(H(U,U_1,U_2,X_1|V)+H(Y_1|V)+H(V)-4\epsilon)} \\
&\leq 2^{-N(H(U,U_1,U_2,X_1|V)+H(Y_1|V)+H(V)-H(U,U_1,U_2,X_1,Y_1)-4\epsilon)} \\
&= 2^{-N(I(U,U_1,U_2,X_1;Y_1|V)-4\epsilon)} \\
&= 2^{-N(I(U,U_2,X_1;Y_1|V)-4\epsilon)},
\end{aligned} \tag{99d}$$

where the probability operator Pr[.] in (a) applies with a probability distribution $P_{\boldsymbol{V}\boldsymbol{U}\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2\boldsymbol{Y}_1}$ that factorizes as $P_{\boldsymbol{V}}P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,|\boldsymbol{V}}P_{\boldsymbol{Y_1}|\boldsymbol{V}}$ given that all the codewords $\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1$, and $\boldsymbol{u}_2$ are independent from the output of the channel $\boldsymbol{y}_1$.
For all $r \in \{2,3,\ldots,2^{NR_{2,C}}\}$ and $l \in \{2,3,\ldots,2^{NR_{1,P}}\}$, the following holds

$$\begin{aligned}
\Pr[E_{1rl}] &\overset{(a)}{=} \Pr\left[(\boldsymbol{V},\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,\boldsymbol{Y}_1) \in \mathcal{T}^{(N,\epsilon)}_{VUU_1X_1U_2Y_1}\right] \\
&= \sum_{(\boldsymbol{v},\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{u}_2,\boldsymbol{y}_1)\in\mathcal{T}^{(N,\epsilon)}_{VUU_1X_1U_2Y_1}} P_{\boldsymbol{V}}(\boldsymbol{v})P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,|\boldsymbol{V}}(\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{u}_2|\boldsymbol{v})P_{\boldsymbol{Y_1}|\boldsymbol{V}}(\boldsymbol{y}_1|\boldsymbol{v}), \\
&\leq \left|\mathcal{T}^{(N,\epsilon)}_{VUU_1X_1U_2Y_1}\right| 2^{-N(H(U,U_1,U_2,X_1|V)+H(Y_1|V)+H(V)-4\epsilon)} \\
&\leq 2^{-N(H(U,U_1,U_2,X_1|V)+H(Y_1|V)+H(V)-H(U,U_1,U_2,X_1,Y_1)-4\epsilon)} \\
&= 2^{-N(I(U,U_1,U_2,X_1;Y_1|V)-4\epsilon)} \\
&= 2^{-N(I(U,U_2,X_1;Y_1|V)-4\epsilon)},
\end{aligned} \tag{99e}$$

where the probability operator Pr[.] in (a) applies with a probability distribution $P_{\boldsymbol{V}\boldsymbol{U}\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2\boldsymbol{Y}_1}$ that factorizes as $P_{\boldsymbol{V}}P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,|\boldsymbol{V}}P_{\boldsymbol{Y_1}|\boldsymbol{V}}$ given that all the codewords $\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1$, and $\boldsymbol{u}_2$ are independent from the output of the channel $\boldsymbol{y}_1$.
For all $s \in \{2,3,\ldots,2^{NR_{1,C}}\}$, $r \in \{2,3,\ldots,2^{NR_{2,C}}\}$ and $l \in \{2,3,\ldots,2^{NR_{1,P}}\}$, the following

holds

$$\begin{aligned}
\Pr[E_{srl}] &\overset{(a)}{=} \Pr\Big[(\boldsymbol{V},\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,\boldsymbol{Y}_1) \in \mathcal{T}^{(N,\epsilon)}_{VUU_1X_1U_2Y_1}\Big] \\
&= \sum_{(\boldsymbol{v},\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{u}_2,\boldsymbol{y}_1)\in\mathcal{T}^{(N,\epsilon)}_{VUU_1X_1U_2Y_1}} P_{\boldsymbol{V}}(\boldsymbol{v})P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,|\boldsymbol{V}}(\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x_1},\boldsymbol{u}_2|\boldsymbol{v})P_{\boldsymbol{Y_1}|\boldsymbol{V}}(\boldsymbol{y}_1|\boldsymbol{v}), \\
&\leq \left|\mathcal{T}^{(N,\epsilon)}_{VUU_1X_1U_2Y_1}\right| 2^{-N(H(U,U_1,U_2,X_1|V)+H(Y_1|V)+H(V)-4\epsilon)} \\
&\leq 2^{-N(H(U,U_1,U_2,X_1|V)+H(Y_1|V)+H(V)-H(U,U_1,U_2,X_1,Y_1)-4\epsilon)} \\
&= 2^{-N(I(U,U_1,U_2,X_1;Y_1|V)-4\epsilon)} \\
&= 2^{-N(I(U,U_2,X_1;Y_1|V)-4\epsilon)},
\end{aligned} \tag{99f}$$

where the probability operator Pr[.] in (a) applies with a probability distribution $P_{\boldsymbol{V}\boldsymbol{U}\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2\boldsymbol{Y}_1}$ that factorizes as $P_{\boldsymbol{V}}P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,|\boldsymbol{V}}P_{\boldsymbol{Y_1}|\boldsymbol{V}}$ given that all the codewords $\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1$, and $\boldsymbol{u}_2$ are independent from the output of the channel $\boldsymbol{y}_1$.
For all $l \in \{2,3,\ldots,2^{NR_{1,P}}\}$, the following holds

$$\begin{aligned}
\Pr[E_{11l}] &\overset{(a)}{=} \Pr\Big[(\boldsymbol{V},\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{X}_1,\boldsymbol{U}_2,\boldsymbol{Y}_1) \in \mathcal{T}^{(N,\epsilon)}_{VUU_1X_1U_2Y_1}\Big] \\
&= \sum_{(\boldsymbol{v},\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{x}_1,\boldsymbol{u}_2,\boldsymbol{y}_1)\in\mathcal{T}^{(N,\epsilon)}_{VUU_1X_1U_2Y_1}} P_{\boldsymbol{V}}(\boldsymbol{v})P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{U}_2,\boldsymbol{Y}_1|\boldsymbol{V}}(\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{y}_1|\boldsymbol{v})P_{\boldsymbol{X_1}|\boldsymbol{V}}(\boldsymbol{x}_1|\boldsymbol{v}), \\
&\leq \left|\mathcal{T}^{(N,\epsilon)}_{VUU_1X_1U_2Y_1}\right| 2^{-N(H(U,U_1,U_2,Y_1|V)+H(X_1|V)+H(V)-4\epsilon)} \\
&\leq 2^{-N(H(U,U_1,U_2,Y_1|V)+H(X_1|V)+H(V)-H(U,U_1,U_2,X_1,Y_1)-4\epsilon)} \\
&= 2^{-N(I(X_1;U,U_1,U_2,Y_1|V)-4\epsilon)} \\
&= 2^{-N(I(X_1;Y_1|U,U_1,U_2,V)-4\epsilon)},
\end{aligned} \tag{99g}$$

where the probability operator Pr[.] in (a) applies with a probability distribution $P_{\boldsymbol{V}\boldsymbol{U}\boldsymbol{U}_1\boldsymbol{X}_1\boldsymbol{U}_2\boldsymbol{Y}_1}$ that factorizes as $P_{\boldsymbol{V}}P_{\boldsymbol{U},\boldsymbol{U}_1,\boldsymbol{U}_2,\boldsymbol{Y}_1|\boldsymbol{V}}P_{\boldsymbol{X_1}|\boldsymbol{V}}$ given that the codewords $\boldsymbol{x}_1$ is independent from the output of the channel $\boldsymbol{y}_1$. Plugging (99) into (98) yields:

$$\begin{aligned}
P_{\mathrm{e}} \leq\ & 2^{N(R_{1C}+R_{2C}+R_{2P}-I(U,X_1,U_2;Y_1|V)+4\epsilon)} + 2^{N(R_{1C}+R_{2C}-I(U,X_1,U_2;Y_1|V)+4\epsilon)} \\
&+ 2^{N(R_{1C}+R_{1P}-I(U,X_1,U_2;Y_1V)+4\epsilon)} + 2^{N(R_{1C}-I(U,X_1,U_2;Y_1|V)+4\epsilon)} \\
&+ 2^{N(R_{2C}+R_{1P}-I(U,X_1,U_2;Y_1|V)+4\epsilon)} + 2^{N(R_{2C}-I(U,X_1,U_2;Y_1|V)+4\epsilon)} \\
&+ 2^{N(R_{1P}-I(X_1;Y_1|U,U_1,U_2,V)+4\epsilon)}.
\end{aligned} \tag{100}$$

The same analysis of the probability of error holds for transmitter-receiver pair 2. Hence in general, from (97) and (100), reliable decoding holds under the following conditions:

$$R_{2C} \leq I(U_2;Y_1|X_1,U,V), \tag{101a}$$
$$R_{1P} \leq I(X_1;Y_1|U_1,U_2,,U,V), \tag{101b}$$
$$R_{1C}+R_{2C}+R_{1P} \leq I(U,X_1,U_2;Y_1|V), \tag{101c}$$
$$R_{1C} \leq I(U_1;Y_2|X_2,U,V), \tag{101d}$$
$$R_{2P} \leq I(X_2;Y_2|U_1,U_2,U,V), \tag{101e}$$
$$R_{1C}+R_{2C}+R_{2P} \leq I(U,X_2,U_1;Y_2|V). \tag{101f}$$

The proof continues by applying a Fourrier-Motzkin elimination process on (101). Set $R_{1P} = R_1 - R_{1C}$, $R_{2P} = R_2 - R_{2C}$. Eliminate R_{1P}, R_{2P} from the inequalities in (101) to obtain:

$$R_{2C} \leq a_1, \tag{102a}$$
$$R_1 - R_{1C} \leq a_2, \tag{102b}$$
$$R_1 + R_{2C} \leq a_3, \tag{102c}$$
$$R_{1C} \leq b_1, \tag{102d}$$
$$R_2 - R_{2C} \leq b_2, \tag{102e}$$
$$R_2 + R_{1C} \leq b_3, \tag{102f}$$
$$-R_{1C} \leq 0, \tag{102g}$$
$$-R_1 + R_{1C} \leq 0, \tag{102h}$$
$$-R_{2C} \leq 0, \tag{102i}$$
$$-R_2 + R_{2C} \leq 0. \tag{102j}$$

Collect the inequalities in (102) that do not include R_{1C} among the above inequalities to obtain:

$$R_{2C} \leq a_1, \tag{103a}$$
$$R_1 + R_{2C} \leq a_3, \tag{103b}$$
$$R_2 - R_{2C} \leq b_2, \tag{103c}$$
$$-R_{2C} \leq 0, \tag{103d}$$
$$-R_2 + R_{2C} \leq 0. \tag{103e}$$

Collect the inequalities in (102) that include R_{1C} with positive coefficients to obtain:

$$R_{1C} \leq b_1, \tag{104a}$$
$$R_2 + R_{1C} \leq b_3, \tag{104b}$$
$$-R_1 + R_{1C} \leq 0. \tag{104c}$$

Collect the inequalities in (102) that include R_{1C} with negative coefficients to obtain:

$$R_1 - R_{1C} \leq a_2, \tag{105a}$$
$$-R_{1C} \leq 0. \tag{105b}$$

Eliminate R_{1C} by adding each inequality from (104) and each inequality from (105) to obtain:

$$R_1 \leq b_1 + a_2, \tag{106a}$$
$$R_2 \leq b_3, \tag{106b}$$
$$R_2 + R_1 \leq b_3 + a_2, \tag{106c}$$
$$-R_1 \leq 0, \tag{106d}$$
$$R_{2C} \leq a_1, \tag{106e}$$
$$R_1 + R_{\ 2C} \leq a_3, \tag{106f}$$
$$-R_2 + R_{2C} \leq 0, \tag{106g}$$
$$R_2 - R_{2C} \leq b_2, \tag{106h}$$
$$-R_{2C} \leq 0. \tag{106i}$$

Collect the inequalities in (106) that do not include R_{2C} to obtain:

$$R_1 \leq b_1 + a_2, \tag{107a}$$
$$R_2 \leq b_3, \tag{107b}$$
$$R_2 + R_1 \leq b_3 + a_2, \tag{107c}$$
$$-R_1 \leq 0. \tag{107d}$$

Collect the inequalities in (106) that include R_{2C} with positive coefficients to obtain:

$$R_{2C} \leq a_1, \tag{108a}$$
$$-R_2 + R_{2C} \leq 0, \tag{108b}$$
$$R_1 + R_{\,2C} \leq a_3. \tag{108c}$$

Collect the inequalities in (106) that include R_{2C} with negative coefficients to obtain:

$$R_2 - R_{2C} \leq b_2, \tag{109a}$$
$$-R_{2C} \leq 0. \tag{109b}$$

Eliminate R_{2C} by adding each inequality in (108) with each inequality in (109) to obtain:

$$R_1 \leq a_3, \tag{110a}$$
$$R_1 \leq b_1 + a_2, \tag{110b}$$
$$R_2 \leq b_3, \tag{110c}$$
$$R_2 \leq a_1 + b_2, \tag{110d}$$
$$R_1 + R_2 \leq b_3, \tag{110e}$$
$$R_1 + R_2 \leq a_3 + b_2. \tag{110f}$$

This completes the proof of Lemma 2. ■

The proof of Theorem 4 continues as follows, let $k \in \{0, 1\}$ be fixed and consider the following Gaussian input distribution for transmitter k:

$$V \sim \mathcal{N}(0,1);\ U \sim \mathcal{N}(0,\rho);\ U_k \sim \mathcal{N}(0,\lambda_{kc});\ \text{and}\ S_k \sim \mathcal{N}(0,\lambda_{kp}), \tag{111}$$

where $X_k = \sqrt{P_k}U + \sqrt{P_k}X_{kp} + \sqrt{P_k}U_k + \sqrt{\lambda_{ke}P_k}V$; and $(\rho, \lambda_{kp}, \lambda_{kc}, \lambda_{ke}) \in [0,1]^4$ and $\rho + \lambda_{kp} + \lambda_{kc} + \lambda_{ke} \leq 1$. By symmetry, it suffices to prove (17a), (17b) and (17e). The choice of the Gaussian input distribution in Lemma 2 yields:

$$\begin{aligned} I(U, X_1, U_2; Y_1|V) &= h(Y_1|V) - h(Y_1|U, X_1, U_2, V) \\ &= \frac{1}{2}\log\Big(2\pi e\mathrm{Var}[Y_1|V]\Big) - \frac{1}{2}\log\Big(2\pi e\mathrm{Var}[Y_1|U, X_1, U_2, V]\Big) \\ &= \frac{1}{2}\log\Big(2\pi e(\sigma_1^2 + (1-\lambda_{1e})P_1h_{11}^2 + (1-\lambda_{2e})P_2h_{12}^2)\Big) \\ &\quad -\frac{1}{2}\log\Big(2\pi e(\sigma_1^2 + \lambda_{2p}P_2h_{12}^2)\Big) \\ &= \frac{1}{2}\log\Big(\frac{1 + (1-\lambda_{1e})\mathrm{SNR}_1 + (1-\lambda_{2e})\mathrm{INR}_1 + 2\rho\sqrt{\mathrm{SNR}_1\mathrm{INR}_1}}{1 + \lambda_{2p}\mathrm{INR}_1}\Big), \end{aligned} \tag{112a}$$

which proves (17a). With the same power setting in (111), the following holds

$$\begin{aligned}
I(U_1;Y_2|U,X_2,V) &= h(Y_2|U,X_2,V) - h(Y_2|U,X_2,U_1,V) \\
&= \frac{1}{2}\log\Big(2\pi e\mathrm{Var}[Y_2|U,X_2,V]\Big) - \frac{1}{2}\log\Big(2\pi e\mathrm{Var}[Y_2|U,X_2,U_1,V]\Big) \\
&= \frac{1}{2}\log\Big(2\pi e\left(\sigma_2^2 + (1-(\rho+\lambda_{1e}))P_1h_{21}^2\right)\Big) \\
- &\quad \frac{1}{2}\log\Big(2\pi e(\sigma_2^2+\lambda_{1p}P_1h_{21}^2)\Big) \\
&= \frac{1}{2}\log\Big(\frac{1+(1-(\rho+\lambda_{1e}))\mathrm{INR}_2}{1+\lambda_{1p}\mathrm{INR}_2}\Big), \text{and} \qquad (112b) \\
I(X_1;Y_1|U,U_1,U_2,V) &= h(Y_1|U,U_1,U_2,V) - h(Y_1|U,U_1,U_2,X_1,V) \\
&= \frac{1}{2}\log\Big(2\pi e\mathrm{Var}[Y_1|U,U_1,U_2,V]\Big) - \frac{1}{2}\log\Big(2\pi e\mathrm{Var}[Y_1|U,U_1,U_2,X_1,V]\Big), \\
&= \frac{1}{2}\log\Big(2\pi e(\sigma_1^2+\lambda 1pP_1h_{11}^2+\lambda_{2p}P_2h_{12}^2)\Big) \\
- &\quad \frac{1}{2}\log\Big(2\pi e(\sigma_1^2+\lambda_{2p}P_1h_{21}^2)\Big) \\
&= \frac{1}{2}\log\Big(\frac{1+\lambda_{1p}\mathrm{SNR}_1+\lambda_{2p}\mathrm{INR}_1}{1+\lambda_{2p}\mathrm{INR}_1}\Big). \qquad (112c)
\end{aligned}$$

This proves (17b). Finally, using (112b) and (112c), yields the proof of (17e).

C.2 Proof of 17g

The choice of the channel input in (111) guarantee that the random variables $Y_{3,1},\ldots,Y_{3,n}$ are independently and identically distributed (i.i.d.). For all $n\in\{1,2,\ldots,N\}$, $Y_{3,n}$ follows a zero-mean Gaussian distribution with variance $\bar{B}$ given by

$$\begin{aligned}
\bar{B} &= E[Y_{3,n}^2] \\
&= E[(h_{31}X_{1,n}+h_{32}X_{2,n}+Z_{3,n})^2] \\
&= h_{31}^2E[X_{1,n}^2]+h_{32}^2E[X_{2,n}^2]+2h_{31}h_{32}E[X_{1,n}X_{2,n}]+\sigma_3^2 \\
&= h_{31}^2P_1+h_{32}^2P_2+2h_{31}h_{32}E[X_{1,n}X_{2,n}]+\sigma_3^2 \\
&\le h_{31}^2P_1+h_{32}^2P_2+2h_{31}h_{32}\sqrt{P_1P_2}(\rho+\sqrt{\lambda_{1e}\lambda_{2e}})+\sigma_3^2, \qquad (113)
\end{aligned}$$

By the weak law of large numbers, it holds that $\forall\epsilon>0$

$$\lim_{n\to\infty}\Pr\left[B^{(N)}<\bar{B}-\epsilon\right]=0 \qquad (114)$$

From (114), it holds that for any energy B which satisfies $0<B\le\bar{B}$, it holds that

$$\lim_{n\to\infty}\Pr\left[B^{(N)}<B-\epsilon\right]=0 \qquad (115)$$

This proves (17g) and completes the proof of Theorem 3.

D Proof of Theorem 4

Fix an information-energy rate triplet (R_1, R_2, B) achievable with a given coding scheme (Definition 1). Denote by $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ the channel inputs resulting from transmitting the independent messages (W_1, Ω) and (W_2, Ω) using such coding scheme. Denote by $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ the corresponding channel outputs. Define the following random variables:

$$S_1 = h_{21}X_1 + Z_2 \text{ and} \tag{116}$$
$$S_2 = h_{12}X_2 + Z_1, \tag{117}$$

where, Z_1 and Z_2 are real Gaussian random variables independent of each other with zero means and variances σ_1^2 and σ_2^2, respectively. Using assumption (13a) and Fano's inequality and following similar steps as in [5], it can be shown that the information rates R_1 and R_2 must satisfy the following inequalities

$$NR_1 \leq \sum_{n=1}^{N}[h(Y_{1,n}) - h(Z_{1,n})] + \mathrm{o}(N), \tag{118a}$$
$$NR_1 \leq \sum_{n=1}^{N}[h(Y_{2,n})|X_{2,n}) - h(Z_{2,n}) + h(Y_{1,n}|X_{2,n}, S_{1,n}) - h(Z_{1,n})] + \mathrm{o}(N), \tag{118b}$$
$$NR_2 \leq \sum_{n=1}^{N}[h(Y_{2,n}) - h(Z_{2,n})] + \mathrm{o}(N), \tag{118c}$$
$$NR_2 \leq \sum_{n=1}^{N}[h(Y_{1,n})|X_{1,n}) - h(Z_{1,n}) + h(Y_{2,n}|X_{1,n}, S_{2,n}) - h(Z_{2,n})] + \mathrm{o}(N), \tag{118d}$$
$$N(R_1 + R_2) \leq \sum_{n=1}^{N}[h(Y_{1,n}|S_{1,n}, X_{2,n}) - h(Z_{1,n}) + h(Y_{2,n}) - h(Z_{2,n})] + \mathrm{o}(N), \tag{118e}$$
$$N(R_1 + R_2) \leq \sum_{n=1}^{N}[h(Y_{2,n}|S_{2,n}, X_{1,n}) - h(Z_{2,n}) + h(Y_{1,n}) - h(Z_{1,n})] + \mathrm{o}(N). \tag{118f}$$

Using assumptions (13b), for a given $\epsilon^{(N)} > 0$, for any $\eta > 0$ there exists $N_0(\eta)$ such that for any $n \geq N_0(\eta)$ it holds that

$$\Pr\left(B^{(N)} < B - \epsilon^{(N)}\right) < \eta. \tag{119}$$

Equivalently,

$$\Pr\left(B^{(N)} \geqslant B - \epsilon^{(N)}\right) \geqslant 1 - \eta. \tag{120}$$

Using Markov's inequality, the probability in (118) can be upper-bounded as follows:

$$(B - \epsilon^{(N)}) \Pr\left(B^{(N)} \geqslant B - \epsilon^{(N)}\right) \leqslant \mathrm{E}[B^{(N)}]. \tag{121}$$

Combining (118) and (119) yields:

$$(B - \epsilon^{(N)})(1 - \eta) \leqslant \mathrm{E}[B^{(N)}], \tag{122}$$

which can be written as

$$(B - \delta^{(N)}) \leqslant \mathrm{E}[B^{(N)}], \tag{123}$$

for some $\delta^{(N)} > \epsilon^{(N)}$ (for sufficiently large N). The bounds in (116) and (121) are evaluated assuming that the channel inputs $X_{1,n}$ and $X_{2,n}$ are arbitrary correlated random variables with

$$\mu_{i,n} \triangleq \mathrm{E}[X_{i,n}], \tag{124}$$

$$\gamma_{i,n}^2 \triangleq \mathrm{Var}[X_{i,n}], \tag{125}$$

$$\lambda_n \triangleq \mathrm{Cov}[X_{1,n}X_{2,n}], \tag{126}$$

for $n \in \{1,\ldots,N\}$ and for $i \in \{1,2\}$. The input sequence must satisfy the input power constraint which can be written, for $i \in \{1,2\}$, as follows:

$$\frac{1}{N}\sum_{n=1}^{N}\mathrm{E}[X_{i,n}^2] = \left(\frac{1}{N}\sum_{n=1}^{N}\gamma_{i,n}^2\right) + \left(\frac{1}{N}\sum_{n=1}^{N}\mu_{i,n}^2\right) \leqslant P_i. \tag{127}$$

Using these asumptions, the following holds:

$$\begin{aligned} h(Y_{1,n}) &\leqslant \frac{1}{2}\log\Big(2\pi e\mathrm{Var}[Y_{1,n}]\Big) \\ &= \frac{1}{2}\log\Big(2\pi e(h_{11}^2\gamma_{1,n}^2 + h_{12}^2\gamma_{2,n}^2 + 2h_{11}h_{12}\lambda_n + \sigma_1^2)\Big), \qquad (128a) \\ h(Y_{2,n}|X_{2,n}) &\leqslant \frac{1}{2}\log\Big(2\pi e\mathrm{Var}[Y_{2,n}|X_{2,n}]\Big) \\ &= \frac{1}{2}\log\left(2\pi e\left(\gamma_{2,n}^2 + \frac{h_{21}^2}{\sigma_2^2}(\gamma_{1,n}^2\gamma_{2,n}^2 - \lambda_n^2)\right)\right) - \frac{1}{2}\log\left(2\pi e\gamma_{2,n}^2\right) + \frac{1}{2}\log(2\pi e\sigma_2^2) \\ &= \frac{1}{2}\log\left(1 + \gamma_{1,n}^2\left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2\gamma_{2,n}^2}\right)\frac{h_{21}^2}{\sigma_2^2}\right) + \frac{1}{2}\log(2\pi e\sigma_2^2), \qquad (128b) \\ h(Y_{1,n}|X_{2,n},S_{1,n}) &\leqslant \frac{1}{2}\log\left(2\pi e\mathrm{Var}[Y_{1,n}|X_{2,n},S_{1,n}]\right) \\ &\leqslant \frac{1}{2}\log\left(1 + \frac{\left(\gamma_{1,n}^2\gamma_{2,n}^2 - \lambda_n^2\right)\frac{h_{11}^2}{\sigma_1^2}}{\gamma_{2,n}^2 + \left(\gamma_{1,n}^2\gamma_{2,n}^2 - \lambda_n^2\right)\frac{h_{21}^2}{\sigma_2^2}}\right) + \frac{1}{2}\log(2\pi e\sigma_1^2) \\ &= \frac{1}{2}\log\left(1 + \frac{\gamma_{1,n}^2\left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2\gamma_{2,n}^2}\right)\frac{h_{11}^2}{\sigma_1^2}}{1 + \gamma_{1,n}^2\left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2\gamma_{2,n}^2}\right)\frac{h_{21}^2}{\sigma_2^2}}\right) + \frac{1}{2}\log(2\pi e\sigma_1^2). \qquad (128c) \end{aligned}$$

Given $(X_{2,n}, S_{1,n})$, the variance of $Y_{1,n}$ is upper-bounded by

$$\mathrm{Var}[Y_{1,n}|X_{2,n},S_{1,n}] \leqslant K_{Y_{1,n}} - K_{Y_{1,n}(X_{2,n},S_{1,n})}K_{(X_{2,n},S_{1,n})}^{-1}K_{Y_{1,n}(X_{2,n},S_{1,n})}^{\mathsf{T}}, \tag{129}$$

where

$$\begin{aligned} K_{Y_{1,n}} &= \sigma_{1,n}^2h_{11}^2 + \sigma_{2,n}^2h_{12}^2 + 2\lambda_nh_{11}h_{12} + \sigma_1^2, \\ K_{Y_{1,n}(X_{2,n},S_{1,n})} &= [\lambda_nh_{11} + h_{12}\gamma_{2,n}^2, h_{21}h_{11}\gamma_{1,n}^2 + \lambda_nh_{21}h_{12}], \text{ and} \\ K_{(X_{2,n},S_{1,n})} &= \begin{bmatrix}\gamma_{2,n}^2 & \lambda_nh_{21} \\ \lambda_nh_{21} & \gamma_{1,n}^2h_{21}^2 + 1\end{bmatrix}, \end{aligned}$$

Finally, the bounds in (116) can be rewritten as follows:

$$NR_1 \leqslant \sum_{n=1}^{N} \frac{1}{2} \log \left(\frac{h_{11}^2 \gamma_{1,n}^2}{\sigma_1^2} + \frac{h_{12}^2 \gamma_{2,n}^2}{\sigma_1^2} + \frac{2h_{11}h_{12}\lambda_n}{\sigma_1^2} + 1 \right) + \mathrm{o}(N), \tag{130a}$$

$$\begin{aligned} NR_1 \leqslant & \sum_{n=1}^{N} \frac{1}{2} \log \left(1 + \gamma_{1,n}^2 (1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}) \frac{h_{21}^2}{\sigma_2^2} \right) \\ & + \sum_{n=1}^{N} \frac{1}{2} \log \left(1 + \frac{\gamma_{1,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{11}^2}{\sigma_1^2}}{1 + \gamma_{1,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{21}^2}{\sigma_2^2}} \right) + \mathrm{o}(N), \end{aligned} \tag{130b}$$

$$NR_2 \leqslant \sum_{n=1}^{N} \frac{1}{2} \log \left(\frac{h_{22}^2 \gamma_{2,n}^2}{\sigma_2^2} + \frac{h_{21}^2 \gamma_{1,n}^2}{\sigma_2^2} + \frac{2h_{22}h_{21}\lambda_n}{\sigma_2^2} + 1 \right) + \mathrm{o}(N), \tag{130c}$$

$$\begin{aligned} NR_2 \leqslant & \sum_{n=1}^{N} \frac{1}{2} \log \left(1 + \gamma_{2,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{21}^2}{\sigma_1^2} \right) \\ & + \sum_{n=1}^{N} \frac{1}{2} \log \left(1 + \frac{\gamma_{2,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{22}^2}{\sigma_2^2}}{1 + \gamma_{2,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{12}^2}{\sigma_1^2}} \right) + \mathrm{o}(N), \end{aligned} \tag{130d}$$

$$\begin{aligned} N(R_1 + R_2) \leqslant & \sum_{n=1}^{N} \frac{1}{2} \log \left(\frac{h_{22}^2 \gamma_{2,n}^2}{\sigma_2^2} + \frac{h_{21}^2 \gamma_{1,n}^2}{\sigma_2^2} + \frac{2h_{22}h_{21}\lambda_n}{\sigma_2^2} + 1 \right) \\ & + \sum_{n=1}^{N} \frac{1}{2} \log \left(1 + \frac{\gamma_{1,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{11}^2}{\sigma_1^2}}{1 + \gamma_{1,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{21}^2}{\sigma_2^2}} \right) + \mathrm{o}(N), \end{aligned} \tag{130e}$$

$$\begin{aligned} N(R_1 + R_2) \leqslant & \sum_{n=1}^{N} \frac{1}{2} \log \left(\frac{h_{11}^2 \gamma_{1,n}^2}{\sigma_1^2} + \frac{h_{12}^2 \gamma_{2,n}^2}{\sigma_1^2} + \frac{2h_{11}h_{12}\lambda_n}{\sigma_1^2} + 1 \right) \\ & + \sum_{n=1}^{N} \frac{1}{2} \log \left(1 + \frac{\gamma_{2,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{22}^2}{\sigma_2^2}}{1 + \gamma_{2,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{12}^2}{\sigma_1^2}} \right) + \mathrm{o}(N). \end{aligned} \tag{130f}$$

The expectation of the average received energy rate is given by

$$\begin{aligned} \mathrm{E}\left[B^{(N)}\right] &= \mathrm{E}\left[\frac{1}{N}\sum_{n=1}^{N} Y_{3,n}^2\right] \\ &= \sigma_3^2 + h_{31}^2 \left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{1,n}^2 + \mu_{1,n}^2)\right) + h_{32}^2 \left(\frac{1}{N}\sum_{n=N}^{n}(\gamma_{2,n}^2 + \mu_{2,n}^2)\right) \\ &\quad + 2h_{31}h_{32}\left(\frac{1}{N}\sum_{n=1}^{N}(\lambda_n + \mu_{1,n}\mu_{2,n})\right). \end{aligned} \tag{131}$$

Using Cauchy-Schwarz inequality, the expected value on the energy rate in (129) can be upper-

bounded as follows:

$$\mathrm{E}\left[B^{(N)}\right] \leqslant \sigma_3^2 + h_{31}^2\left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{1,n}^2+\mu_{1,n}^2)\right) + h_{32}^2\left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{2,n}^2+\mu_{2,n}^2)\right) + 2h_{31}h_{32}\left(\left|\frac{1}{N}\sum_{n=1}^{N}\lambda_n\right| + \left(\frac{1}{N}\sum_{n=1}^{N}\mu_{1,n}^2\right)^{1/2}\left(\frac{1}{N}\sum_{n=1}^{N}\mu_{2,n}^2\right)^{1/2}\right). \quad (132)$$

Combining (121) and (130) yields the following upper-bound on the energy rate B:

$$B \leqslant \sigma_3^2 + h_{31}^2\left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{1,n}^2+\mu_{1,n}^2)\right) + h_{32}^2\left(\frac{1}{N}\sum_{n=1}^{N}(\gamma_{2,n}^2+\mu_{2,n}^2)\right) + 2h_{31}h_{32}\left(\left|\frac{1}{N}\sum_{n=1}^{N}\lambda_n\right| + \left(\frac{1}{N}\sum_{n=1}^{N}\mu_{1,n}^2\right)^{1/2}\left(\frac{1}{N}\sum_{n=1}^{N}\mu_{2,n}^2\right)^{1/2}\right) + \delta_N. \quad (133)$$

In order to obtain a single-letterization of the upper-bound given by constraints (128) and (131), define also

$$\mu_i^2 \triangleq \frac{1}{N}\sum_{n=1}^{N}\mu_{i,n}^2, \quad i \in \{1,2\}, \quad (134)$$

$$\gamma_i^2 \triangleq \frac{1}{N}\sum_{n=1}^{N}\gamma_{i,n}^2, \quad i \in \{1,2\}, \quad (135)$$

$$\rho \triangleq \frac{\left(\frac{1}{N}\sum_{n=1}^{N}\lambda_n\right)}{|\gamma_1||\gamma_2|}. \quad (136)$$

With these notations, the input power constraint in (125) can be rewritten as

$$\gamma_i^2 + \mu_i^2 \leqslant P_i, \quad i \in \{1,2\}. \quad (137)$$

Lemma 3

$$\frac{1}{N}\sum_{n=1}^{N}\frac{1}{2}\log\left(1+\frac{\gamma_{i,n}^2\left(1-\frac{\lambda_n^2}{\gamma_{i,n}^2\gamma_{j,n}^2}\right)\frac{h_{ii}^2}{\sigma_i^2}}{1+\gamma_{i,n}^2\left(1-\frac{\lambda_n^2}{\gamma_{i,n}^2\gamma_{j,n}^2}\right)\frac{h_{ji}^2}{\sigma_j^2}}\right) \leqslant \frac{1}{2}\log\left(1+\frac{\gamma_i^2\left(1-\rho^2\right)\frac{h_{ii}^2}{\sigma_i^2}}{1+\gamma_i^2\left(1-\rho^2\right)\frac{h_{ji}^2}{\sigma_j^2}}\right). \quad (138)$$

Proof: Using the fact that $H_{\alpha,\beta}(x) = \log\left(1+\frac{\alpha x}{1+\beta x}\right)$ is a concave function on x with

positive α and β, it follows that

$$\begin{aligned}
&\frac{1}{N}\sum_{n=1}^{N}\ \frac{1}{2}\ \log\left(1+\frac{\gamma_{i,n}^2\left(1-\frac{\lambda_n^2}{\gamma_{i,n}^2\gamma_{j,n}^2}\right)\frac{h_{ii}^2}{\sigma_i^2}}{1+\gamma_{i,n}^2\left(1-\frac{\lambda_n^2}{\gamma_{i,n}^2\gamma_{j,n}^2}\right)\frac{h_{ji}^2}{\sigma_j^2}}\right)\\
&\qquad\leqslant\ \frac{1}{2}\log\left(1+\frac{\frac{1}{N}\sum_{n=1}^{N}\gamma_{i,n}^2\left(1-\frac{\lambda_n^2}{\gamma_{i,n}^2\gamma_{j,n}^2}\right)\frac{h_{ii}^2}{\sigma_i^2}}{1+\frac{1}{N}\sum_{n=1}^{N}\gamma_{i,n}^2\left(1-\frac{\lambda_n^2}{\gamma_{i,n}^2\gamma_{j,n}^2}\right)\frac{h_{ji}^2}{\sigma_j^2}}\right)\\
&\qquad\overset{(a)}{\leqslant}\ \frac{1}{2}\log\left(1+\frac{\frac{1}{N}\sum_{n=1}^{N}\gamma_{i,n}^2\left(1-\frac{(\frac{1}{N}\sum_{n=1}^{N}\lambda_n)^2}{\frac{1}{N}\sum_{n=1}^{N}\gamma_{i,n}^2\frac{1}{N}\sum_{n=1}^{N}\gamma_{j,n}^2}\right)\frac{h_{ii}^2}{\sigma_i^2}}{1+\frac{1}{N}\sum_{n=1}^{N}\gamma_{i,n}^2\left(1-\frac{(\frac{1}{N}\sum_{n=1}^{N}\lambda_n)^2}{\frac{1}{N}\sum_{n=1}^{N}\gamma_{i,n}^2\frac{1}{N}\sum_{n=1}^{N}\gamma_{j,n}^2}\right)\frac{h_{ji}^2}{\sigma_j^2}}\right)\\
&\qquad=\ \frac{1}{2}\log\left(1+\frac{\gamma_i^2\left(1-\rho^2\right)\frac{h_{ii}^2}{\sigma_i^2}}{1+\gamma_i^2\left(1-\rho^2\right)\frac{h_{ji}^2}{\sigma_j^2}}\right),
\end{aligned}\tag{139}$$

where (a) follows from the fact that $G(\gamma_{i,n}^2,\gamma_{j,n}^2,\lambda_n)=\sigma_{i,n}^2\left(1-\frac{\lambda_n^2}{\gamma_{i,n}^2\gamma_{j,n}^2}\right)$ is a concave function on $\gamma_{i,n}^2$, $\gamma_{j,n}^2$ and λ_n; and $K_{\alpha,\beta}(x)=\frac{\alpha x}{1+\beta x}$ is an increasing function on x. ■

By the concavity of the mutual information, applying Jensen's inequality and Lemma 3 in the bounds (128) yields, in the limit when $N\to\infty$,

$$R_1 \leqslant \frac{1}{2}\log\left(\frac{h_{11}^2\gamma_1^2}{\sigma_1^2}+\frac{h_{12}^2\gamma_2^2}{\sigma_1^2}+2\rho\sqrt{\frac{h_{11}^2h_{12}^2\gamma_1^2\gamma_2^2}{\sigma_1^4}}\right),\tag{140a}$$

$$R_1 \leqslant \frac{1}{2}\log\left(1+\frac{h_{21}^2}{\sigma_2^2}\gamma_1^2\left(1-\rho^2\right)\right)+\frac{1}{2}\log\left(1+\frac{\gamma_1^2(1-\rho^2)\frac{h_{11}^2}{\sigma_1^2}}{1+\gamma_1^2(1-\rho^2)\frac{h_{21}^2}{\sigma_2^2}}\right),\tag{140b}$$

$$R_2 \leqslant \frac{1}{2}\log\left(\frac{h_{22}^2\gamma_2^2}{\sigma_2^2}+\frac{h_{21}^2\sigma_1^2}{\sigma_2^2}+2\rho\sqrt{\frac{h_{22}h_{21}\gamma_1^2\gamma_2^2}{\sigma_2^4}}\right),\tag{140c}$$

$$R_2 \leqslant \frac{1}{2}\log\left(1+\frac{h_{12}^2}{\sigma_1^2}\gamma_2^2\left(1-\rho^2\right)\right)+\frac{1}{2}\log\left(1+\frac{\gamma_2^2\left(1-\rho^2\right)\frac{h_{22}^2}{\sigma_2^2}}{1+\gamma_2^2\left(1-\rho^2\right)\frac{h_{12}^2}{\sigma_1^2}}\right),\tag{140d}$$

$$R_1+R_2 \leqslant \frac{1}{2}\log\left(\frac{h_{22}^2\gamma_2^2}{\sigma_2^2}+\frac{h_{21}^2\gamma_1^2}{\sigma_2^2}+2\rho\sqrt{\frac{h_{22}^2h_{21}^2\gamma_1^2\gamma_2^2}{\sigma_2^4}}\right)+\frac{1}{2}\log\left(1+\frac{\gamma_1^2(1-\rho^2)\frac{h_{11}^2}{\sigma_1^2}}{1+\gamma_1^2(1-\rho^2)\frac{h_{21}^2}{\sigma_2^2}}\right),\tag{140e}$$

$$R_1+R_2 \leqslant \frac{1}{2}\log\left(\frac{h_{11}^2\gamma_1^2}{\sigma_1^2}+\frac{h_{12}^2\gamma_2^2}{\sigma_1^2}+2\rho\sqrt{\frac{h_{11}^2h_{12}^2\gamma_1^2\gamma_2^2}{\sigma_1^4}}\right)+\frac{1}{2}\log\left(1+\frac{\gamma_2^2(1-\rho^2)\frac{h_{22}^2}{\sigma_2^2}}{1+\gamma_2^2(1-\rho^2)\frac{h_{12}^2}{\sigma_1^2}}\right),\tag{140f}$$

and the upper-bound on the energy rate (74) :

$$B \leqslant \sigma_3^2+h_{31}^2(\gamma_1^2+\mu_1^2)+h_{32}^2(\gamma_2^2+\mu_2^2)+2h_{21}h_{22}(|\rho|\ |\gamma_1||\gamma_2|+|\mu_1||\mu_2|).\tag{140g}$$

To sum up, it has been shown so far that, in the limit when N tends to infinity, any information-energy rate triplet $(R_1,R_2,B)\in\mathcal{E}_b^{\mathrm{FB}}$ can be bounded by the constraints in (138) for some γ_1^2,

γ_2^2, μ_1, μ_2 satisfying (135) and for some $\rho \in [-1,1]$. Let $\mathcal{R}_b^{\mathrm{FB}}(\gamma_1,\gamma_2,\mu_1,\mu_2,\rho)$ denote the set of information-energy rate triplets satisfying (138), for some γ_1, γ_2, μ_1, μ_2 such that (135) is true and for some $\rho \in [-1,1]$. Thus, it holds that

$$\mathcal{E}_b^{\mathrm{FB}} \subseteq \bigcup_{\substack{0\leqslant\gamma_1^2+\mu_1^2\leqslant P_1\\ 0\leqslant\gamma_2^2+\mu_2^2\leqslant P_2\\ -1\leqslant\rho\leqslant 1}} \mathcal{R}_b^{\mathrm{FB}}(\gamma_1^2,\gamma_2^2,\mu_1,\mu_2,\rho). \tag{141}$$

In this union, it suffices to consider $0 \leqslant \rho \leqslant 1$ because for any $-1 \leqslant \rho \leqslant 1$, $\mathcal{R}_b^{\mathrm{FB}}(\gamma_1^2,\gamma_2^2,\mu_1^2,\mu_2^2,\rho) \subseteq \mathcal{R}_b^{\mathrm{FB}}(\gamma_1^2,\gamma_2^2,\mu_1^2,\mu_2^2,|\rho|)$. Thus,

$$\mathcal{E}_b^{\mathrm{FB}} \subseteq \bigcup_{\substack{0\leqslant\gamma_1^2+\mu_1^2\leqslant P_1\\ 0\leqslant\gamma_2^2+\mu_2^2\leqslant P_2\\ -1\leqslant\rho\leqslant 1}} \mathcal{R}_b^{\mathrm{FB}}(\gamma_1^2,\gamma_2^2,\mu_1,\mu_2,\rho) \subseteq \bigcup_{\substack{\gamma_1^2+\mu_1^2=P_1\\ \gamma_2^2+\mu_2^2=P_2\\ 0\leqslant\rho\leqslant 1}} \mathcal{R}_b^{\mathrm{FB}}(\gamma_1^2,\gamma_2^2,\mu_1,\mu_2,\rho).$$

Let $\beta_i \in [0,1]$ be defined as follows:

$$\beta_i \stackrel{\triangle}{=} \frac{\gamma_i^2}{P_i} = \frac{P_i-\mu_i^2}{P_i}, \quad i \in \{1,2\}. \tag{142}$$

Using (140), any region $\mathcal{R}_b(\gamma_1,\gamma_2,\mu_1,\mu_2,\rho)$ in the union over all $(\mu_1,\mu_2,\gamma_1,\gamma_2)$ that satisfy $\gamma_1^2+\mu_1^2=P_1$, $\gamma_2^2+\mu_2^2=P_2$ and $0 \leqslant \rho \leqslant 1$, can be rewritten as follows:

$$R_1 \leqslant \frac{1}{2}\log\left(\frac{h_{11}^2\beta_1P_1}{\sigma_1^2}+\frac{h_{12}^2\beta_2P_2}{\sigma_1^2}+2\rho\sqrt{\frac{h_{11}^2h_{12}^2\beta_1\beta_2P_1P_2}{\sigma_1^4}}\right), \tag{143a}$$

$$R_1 \leqslant \frac{1}{2}\log\left(1+\frac{h_{21}^2}{\sigma_2^2}\beta_1P_1\left(1-\rho^2\right)\right)+\frac{1}{2}\log\left(1+\frac{\beta_1P_1(1-\rho^2)\frac{h_{11}^2}{\sigma_1^2}}{1+\beta_1P_1(1-\rho^2)\frac{h_{21}^2}{\sigma_2^2}}\right), \tag{143b}$$

$$R_2 \leqslant \frac{1}{2}\log\left(\frac{h_{22}^2\beta_2P_2}{\sigma_2^2}+\frac{h_{21}^2\beta_1P_1}{\sigma_2^2}+2\rho\sqrt{\frac{h_{22}h_{21}\beta_1\beta_2P_1P_2}{\sigma_2^4}}\right), \tag{143c}$$

$$R_2 \leqslant \frac{1}{2}\log\left(1+\frac{h_{12}^2}{\sigma_1^2}\beta_2P_2\left(1-\rho^2\right)\right)+\frac{1}{2}\log\left(1+\frac{\beta_2P_2\left(1-\rho^2\right)\frac{h_{22}^2}{\sigma_2^2}}{1+\beta_2P_2\left(1-\rho^2\right)\frac{h_{12}^2}{\sigma_1^2}}\right), \tag{143d}$$

$$\begin{aligned}R_1+R_2 \leqslant{}& \frac{1}{2}\log\left(\frac{h_{22}^2\beta_2P_2}{\sigma_2^2}+\frac{h_{21}^2\beta_1P_1}{\sigma_2^2}+2\rho\sqrt{\frac{h_{22}^2h_{21}^2\beta_1\beta_2P_1P_2}{\sigma_2^4}}\right)\\ &+\frac{1}{2}\log\left(1+\frac{\beta_1P_1\left(1-\rho^2\right)\frac{h_{11}^2}{\sigma_1^2}}{1+\beta_1P_1\left(1-\rho^2\right)\frac{h_{21}^2}{\sigma_2^2}}\right),\end{aligned} \tag{143e}$$

$$\begin{aligned}R_1+R_2 \leqslant{}& \frac{1}{2}\log\left(\frac{h_{11}^2\beta_1P_1}{\sigma_1^2}+\frac{h_{12}^2\beta_2P_2}{\sigma_1^2}+2\rho\sqrt{\frac{h_{11}^2h_{12}^2\beta_1\beta_2P_1P_2}{\sigma_1^4}}\right)\\ &+\frac{1}{2}\log\left(1+\frac{\beta_2P_2\left(1-\rho^2\right)\frac{h_{22}^2}{\sigma_2^2}}{1+\beta_2P_2\left(1-\rho^2\right)\frac{h_{12}^2}{\sigma_1^2}}\right),\end{aligned} \tag{143f}$$

$$B \leqslant \sigma_3^2+h_{31}^2P_1+h_{32}^2P_2+2h_{31}h_{32}(|\rho|\sqrt{\beta_1P_1\beta_2P_2}+\sqrt{(1-\beta_1)(1-\beta_2)P_1P_2}), \tag{143g}$$

$$B \geqslant b, \tag{143h}$$

for some $(\beta_1, \beta_2) \in [0,1]^2$ and $\rho \in [0,1]$. Hence, using the definitions in (7) and (10), the region (141) contains all information-energy rate triplets (R_1, R_2, B) satisfying constraints (18), which completes the proof of Theorem 4.

References

[1] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, 1981.

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

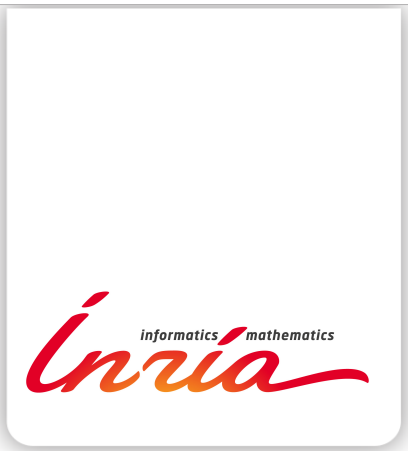
[3] S. Belhadj Amor, S. M. Perlaza, I. Krikidis, and H. V. Poor, "Feedback enhances simultaneous wireless information and energy transmission in multiple access channels," *IEEE Trans. Inf. Theory*, vol. 63, no. 8, pp. 5244–5265, Aug. 2017.

[4] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.

[5] 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.

[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* Hoboken, NJ, USA: Wiley-Interscience Publication, 2005.

[7] R. Durret, *Probability Theory and Examples.* New York, NY: Cambridge University Press, 2010.



RESEARCH CENTRE
GRENOBLE – RHÔNE-ALPES

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

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

ISSN 0249-6399