



HAL
open science

Simultaneous Information and Energy Transmission

Nizar Khalfet

► **To cite this version:**

Nizar Khalfet. Simultaneous Information and Energy Transmission. Computer Science [cs]. INSA LYON, 2020. English. NNT: . tel-03919602

HAL Id: tel-03919602

<https://hal.science/tel-03919602>

Submitted on 3 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



INSA

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

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

Spécialité / discipline de doctorat :
Génie Électrique

Soutenue publiquement le 13/02/2020, par :
Nizar Khalfet

Transmission Simultanée d'Information et d'Énergie

Devant le jury composé de :

Rapporteurs :

LASAULCE Samson	Directeur de Recherche	CNRS	France
KRIKIDIS Ioannis	Maître de Conférences	University of Cyprus	Chypre

Examineurs :

KOBAYASHI Mari	Professeur des Universités	CentraleSupélec	France
----------------	----------------------------	-----------------	--------

Directeur de Thèse :

GORCE Jean-Marie	Professeur des Universités	INSA Lyon	France
------------------	----------------------------	-----------	--------

Co-encadrant :

PERLAZA Samir	Chargé de Recherche	INRIA	France
---------------	---------------------	-------	--------

Abstract

IN this thesis, the fundamental limits of simultaneous information and energy transmission (SIET) are studied from two perspectives: the asymptotic and non-asymptotic block-length regimes. In the asymptotic block-length regime, the fundamental limits on SIET in the two-user Gaussian interference channel (G-IC) with and without feedback are characterized. More specifically, an achievable and converse regions 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 on information transmission rates, as well as a new outer bound on the energy transmission rate. For the finite block-length regime, the case of a transmitter simultaneously sending information to a receiver and energy to an energy harvester through the binary symmetric channel has been studied. Given a finite number of channel uses (latency constraint) as well as tolerable average decoding error probability and energy shortage probability (reliability constraints), two sets of information and energy transmission rates are presented. One consists in rate pairs for which the existence of at least one code achieving such rates under the latency and reliability constraints is proved (achievable region). The second one consists in a set whose complement contains the rate pairs for which there does not exist a code capable of achieving such rates (converse region). These two sets approximate the information-energy capacity region, which allows analyzing the trade-offs among performance, latency, and reliability in SIET systems.

Résumé

Dans cette thèse, les limites fondamentales de la transmission simultanée d'information et d'énergie sont étudiées sous deux perspectives : le régime asymptotique et le régime non-asymptotique. Pour le régime asymptotique, 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. Un 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 de faisabilité basé sur division de puissance, des variable aléatoire en commun, division de débit, codage superposé en block du Markov, et décodage en arrière est présenté. Finalement, une 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. Deuxièmement, le problème d'introduction d'un récepteur d'énergie dans le canal binaire symétrique est présenté. Pour ce problème, une borne de faisabilité et une borne d'impossibilité sont obtenues dans le régime non-asymptotique. Ces deux bornes caractérisent le nombre maximal de bits d'information et d'énergie qui peuvent être introduits de manière furtive pour le canal binaire symétrique.

Contents

Notation	x
Présentation des contributions majeures	xi
1. Synthèse des contributions	1
1.1. Canal Guassien à interférence avec récolte d'énergie	1
1.1.1. Transmission de l'information	2
1.1.2. Transmission d'énergie	3
1.1.3. Transmission simultanée d'informations et d'énergie	4
1.1.4. Résultats Principaux	5
1.1.5. Cas sans voie de retour	5
1.1.6. Cas avec une rétroaction parfaite entre le canal et la sortie	7
1.1.7. Amélioration maximale du débit d'énergie avec la voie de retour	9
1.2. Récolte d'énergie avec des contraintes de fiabilité et de latence	9
1.2.1. Transmission de l'information	10
1.2.2. Transmission d'énergie	11
1.2.3. Limites du débit de transmission de l'énergie	11
1.2.4. Limites du débit de transmission de l'information	12
1.2.5. Région de capacité de l'information et de l'énergie	12
2. Introduction	15
2.1. Energy Harvesting	15
2.1.1. Solar Energy Harvesting	15
2.1.2. Wind Energy Harvesting	16
2.1.3. Thermal-Based Energy Harvesting	16
2.1.4. Mechanical Energy Harvesting	16
2.1.5. Biochemical Energy Harvesting	17
2.1.6. RF Energy Harvesting	17
2.2. Wireless Energy Transmission (WET)	18
2.2.1. WET Technologies	18
2.2.2. History and Evolution	18
2.3. Simultaneous Information and Energy Transmission	19
2.3.1. Time Switching (TS) Architectures	20

2.3.2. Power Splitting (PS) Architectures	20
2.3.3. Antenna Division Architecture	21
3. State of the Art	23
3.1. Point-to-Point Channels	23
3.1.1. Discrete Memoryless Channels	23
3.1.2. The additive white Gaussian noise (AWGN) channels	26
3.2. Multiple Access Channels	27
3.2.1. System Model	27
3.2.2. Case Without Feedback	29
3.2.3. Case With Feedback	30
4. Gaussian Interference Channel	33
4.0.1. Case without Feedback	34
4.0.2. Case with Perfect Channel-Output Feedback	35
5. Energy Harvesting in the Gaussian Interference Channel	37
5.1. Introduction	37
5.2. Gaussian Interference Channel with Energy Harvesting	38
5.2.1. Information Transmission Task	38
5.2.2. Energy Transmission Task	39
5.2.3. Simultaneous Information and Energy Transmission	41
5.3. Main Result	42
5.3.1. Case without Channel-Output Feedback	42
5.3.2. Case with Perfect Channel-Output Feedback	44
5.3.3. Maximum Energy Rate Improvement with Feedback	46
5.4. Numerical Analysis	47
6. Energy Harvesting with Reliability and Latency Constraints	51
6.1. Introduction	51
6.2. System Model	52
6.2.1. Information Transmission Task	53
6.2.2. Energy Transmission Task	53
6.3. Fundamental Limits	55
6.3.1. Bounds on the Energy Transmission Rate	55
6.3.2. Bounds on the Information Transmission Rate	59
6.3.3. Information-Energy Capacity Region	60
7. Conclusion	63
A. Proof of Theorem 12	65
A.0.1. Proof of (5.14a)-(5.14g)	65
A.0.2. Proof of (5.14h)	74
B. Proof of Theorem 13	77
C. Proof of Theorem 15	87
C.0.1. Proof of (5.17a)-(5.17f)	87

C.0.2. Proof of (5.17g)	96
D. Proof of Theorem 16	97
E. Preliminary Results	105
F. Proof of Proposition 5	107
G. Proof of Proposition 6	109
H. Proof of Proposition 7	113
I. Proof of Lemma 6	115
I.1. Proof of Proposition 8	116
J. Proof of Theorem 18	119
K. Proof of Theorem 19	123
Bibliography	127

List of Figures

1.1. Canal Guassien à interférence avec un collecteur d'énergie non-colocalisé à l'utilisation du canal n	3
1.2. Canal Guassien à interférence avec un collecteur d'énergie non-colocalisé à l'utilisation du canal n avec voie de retour.	4
2.1. Electrical characteristics of a solar cell model. Taken from [1].	16
2.2. RF scheme model for harvesting Energy. Taken from [2]	17
2.3. Time switching technique with one transmitter and one receiver. Taken from [3].	20
2.4. Power splitting technique with one transmitter and one receiver. Taken from [3].	21
3.1. Two-user memoryless Gaussian MAC with feedback and energy harvester. Taken from [4].	28
3.2. Intersection of the planes $B = b_0$ and $B = b_1$ with the information-energy capacity region of the symmetric G-MAC without feedback. Taken from [4] .	31
4.1. Gaussian interference channel with noisy channel-output at channel use n . .	33
5.1. Two-user Gaussian interference channel with a non-colocated energy harvester at channel use n	40
5.2. Two-user Gaussian interference channel with channel-output feedback and a non-colocated energy harvester during channel use n	40
5.3. Superposition of $\underline{\mathcal{E}}$ and $\bar{\mathcal{E}}$, with parameters $\text{SNR}_1 = \text{SNR}_2 = 20\text{dB}$, $\text{INR}_1 = \text{INR}_2 = \text{SNR}_{31} = \text{SNR}_{32} = 10\text{dB}$, $\sigma_3^2 = 1$	48
5.4. 3-D superposition of $\underline{\mathcal{E}}^{\text{F}}$ and $\bar{\mathcal{E}}^{\text{F}}$, with parameters $\text{SNR}_1 = \text{SNR}_2 = 20\text{ dB}$, $\text{INR}_1 = \text{INR}_2 = \text{SNR}_{31} = \text{SNR}_{32} = 10\text{ dB}$ and $\sigma_3^2 = 1$	49
5.5. Convex hull of pairs (R_2, B) that are in the sets $\{(R_2, B) \in \mathbb{R}_+^2 : (r_1, R_2, B) \in \bar{\mathcal{E}}\}$ (solid line) and $\{(R_2, B) \in \mathbb{R}_+^2 : (r_1, R_2, B) \in \bar{\mathcal{E}}^{\text{F}}\}$ (dashed line), with $r_1 \in \{0, 3\}$. Parameters $\text{SNR}_1 = \text{SNR}_2 = 20\text{ dB}$, $\text{INR}_1 = \text{INR}_2 = \text{SNR}_{31} = \text{SNR}_{32} = 10\text{ dB}$, and $\sigma_3^2 = 1$	49
5.6. The ratio $\frac{\sup \bar{B}_{\text{F}}}{\sup B}$ for different ratios of SNR at the EH, i.e., $\text{SNR}_{31} = \text{SNR}_{32} = \text{SNR}$; $\frac{\text{SNR}_{31}}{2} = \text{SNR}_{32} = \text{SNR}$; and $\frac{\text{SNR}_{31}}{10} = \text{SNR}_{32} = \text{SNR}$, respectively. Parameters $\text{SNR}_1 = \text{SNR}_2 = 20\text{ dB}$, $\text{INR}_1 = \text{INR}_2 = \text{SNR}_{31} = \text{SNR}_{32} = 10\text{ dB}$, and $\sigma_3^2 = 1$	50

- 5.7. The ratio $\frac{\max \bar{\mathcal{B}}_F}{\max \mathcal{B}}$ for different ratios of SNR at the EH, i.e., $\text{SNR}_{31} = \text{SNR}_{32} = \text{SNR}$; $\frac{\text{SNR}_{31}}{2} = \text{SNR}_{32} = \text{SNR}$; and $\frac{\text{SNR}_{31}}{10} = \text{SNR}_{32} = \text{SNR}$, respectively. Parameters $\text{SNR}_1 = \text{SNR}_2 = 20$ dB, $\text{INR}_1 = \text{INR}_2 = \text{SNR}_{31} = \text{SNR}_{32} = 10$ dB, and $\sigma_3^2 = 1$ 50

Notation

THROUGHOUT this thesis, sets are denoted with uppercase calligraphic letters, i.e., \mathcal{X} . Random variables are denoted by uppercase letters, e.g., X . Denote by x and \mathcal{X} the realization and the set of events from which the random variable X takes value, respectively. Denote by P_X the probability distribution of X over the set \mathcal{X} . Denote by P_{XY} and $P_{Y|X}$ the joint probability distribution of (X, Y) and the conditional probability distribution of Y given X , respectively. Let N be a fixed natural number. Denote by $\mathbf{X} = (X_1, X_2, \dots, X_N)^\top$ an N -dimensional vector of random variables and a corresponding realization is denoted by $\mathbf{x} = (x_1, x_2, \dots, x_N)^\top$. Given $\mathbf{X} = (X_1, X_2, \dots, X_N)^\top$ and $(a, b) \in \mathbb{N}^2$, with $a < b \leq N$, denote by $\mathbf{X}_{(a,b)} = (X_a, X_{a+1}, \dots, X_b)^\top$ the $(b - a + 1)$ -dimensional vector of random variables formed by the components a to b of X . Denote by $(\cdot)^+ = \max(\cdot, 0)$ the positive part operator. Denote by $\mathbb{E}_X[\cdot]$ the expectation with respect to the distribution of the random variable X . The logarithm function is assumed to be base 2. Let \mathbf{x} be a binary vector, hence denote by $N(0|\mathbf{x})$ and $N(1|\mathbf{x})$ the number of zeros and ones in \mathbf{x} , respectively. Given a binary vector \mathbf{y} , denote by $d(\mathbf{x}, \mathbf{y})$ the Hamming distance between \mathbf{x} and \mathbf{y} , which is given by

$$d(\mathbf{x}, \mathbf{y}) = \sum_{t=1}^n \mathbb{1}_{\{x_t \neq y_t\}}. \quad (1)$$

The complementary cumulative distribution function $Q : \mathbb{R} \rightarrow [0, 1]$ of the standard Gaussian distribution is

$$Q(t) = \frac{1}{\sqrt{2\pi}} \int_t^\infty \exp\left(-\frac{x^2}{2}\right) dx, \quad (2)$$

and the functional inverse of Q is $Q^{-1} : [0, 1] \rightarrow \mathbb{R}$. Given two integers n and t , the coefficient of the term x^t in the expansion of the binomial power $(1 + x)^n$ is denoted by $\binom{n}{t}$. Therefore, for all $t < 0$ or $t > n$, it is assumed that $\binom{n}{t} = 0$.

Présentation des contributions majeures

Cette thèse a pour objet d'étude les limites fondamentales de la transmission simultanée d'information et d'énergie sont étudiées sous deux perspectives : le régime asymptotique et le régime non-asymptotique. Pour le régime asymptotique, 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. Un 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 de faisabilité basé sur division de puissance, des variable aléatoire en commun, division de débit, codage superposé en block du Markov , et décodage en arrière est présenté. Finalement, une 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. Deuxièmement, le problème d'introduction d'un récepteur d'énergie dans le canal binaire symétrique est présenté. Pour ce problème, une borne de faisabilité et une borne d'impossibilité sont obtenues dans le régime non-asymptotique. Ces deux bornes caractérisent le nombre maximal de bits d'information et d'énergie qui peuvent être introduits de manière furtive pour le canal binaire symétrique.

Publications

Journals

“*Simultaneous Information and Energy Transmission in the Two-User Gaussian Interference Channel*”. Nizar Khalfet and Samir M. Perlaza. IEEE Journal on Selected Areas in Communications, Special Issue on Wireless Transmission of Information and Power, Sep., 2018.

Conferences

1. “*On Ultra-Reliable and Low Latency Simultaneous Information and Energy Transmission Systems*”. Nizar Khalfet, Samir M. Perlaza, Ali Tajer, and H. Vincent Poor. IEEE International Workshop on Signal Processing Advances in Wireless Communications, Cannes, France, Jul., 2019.
2. “*On the Maximum Energy Transmission Rate in Ultra-Reliable and Low Latency SIET*”. Nizar Khalfet and Samir M. Perlaza. Third International Balkan Conference on Communications and Networking, Skopje, North Macedonia, Jun., 2019.
3. “*Simultaneous Information and Energy Transmission in Gaussian Interference Channels*”. Nizar Khalfet and Samir M. Perlaza. The 2018 International Zurich Seminar on Information and Communications, Zurich, Switzerland, Feb., 2018.
4. “*Simultaneous Information and Energy Transmission in Gaussian Interference Channels with Feedback*”. Nizar Khalfet and Samir M. Perlaza. The 55th Annual Allerton Conference on Communications, Control, and Computing, Monticello, IL, US, Oct., 2017.

INRIA Technical Reports

1. “*On Ultra-Reliable and Low Latency Simultaneous Information and Energy Transmission Systems*”. Nizar Khalfet, Samir M. Perlaza, Ali Tajer, and H. Vincent Poor. Technical Report, INRIA, No. RR-9261, Lyon, France, May., 2019.
2. “*Simultaneous Information and Energy Transmission in the Interference Channel*”. Nizar Khalfet and Samir M. Perlaza. Technical Report, INRIA, No. RR-9102, Lyon, France, Nov., 2017.

Synthèse des contributions

Cette these a pour objectif d'étudier 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 en considérant le cas du regime asymptotique. Pour le régime non asymptotique, la région atteignable et la région d'impossibilité sont caractérisées pour le canal binaire symétrique. Voici les contributions principales de cette these :

- Pour le canal Gaussien a interference G-IC avec et sans voie de retour, les limites fondamentales de la transmission simultanée d'information et d'énergie sont identifiées. Pour les deux cas, un schéma d'atteignabilité basé sur la division de puissance, exploitant les techniques d'en aléatoire commun, de la division du débit , de block-Markov codage par superposition, et de décodage en sens inverse est présenté. Finalement, une région d'impossibilité pour les deux cas est obtenue en utilisant des techniques de majoration de la littérature pour les débits d'information et aussi une borne supérieure pour le débit d'énergie en utilisant la loi des grands nombres.
- Pour le canal binaire symétrique, compte tenu d'un nombre fini d'utilisations du canal (contrainte de latence), de la contrainte sur la probabilité moyenne d'erreur de décodage sur la probabilité du seuil d'énergie tolérables (contrainte de fiabilité), deux ensembles du débit de transmission d'information et d'énergie sont présentés. Le premier consiste en des paires de débits pour lesquels l'existence d'au moins un code permettant d'atteindre de tels débits sous les contraintes de latence est présenté (région atteignabilité). La seconde consiste en un ensemble les paires de débits pour lesquelles il n'existe pas de code capable d'atteindre de tels débits (région d'impossibilité) . Ces deux ensembles permettent d'approximer la région de capacité d'information et d'énergie, ce qui permet d'analyser le compromis entre performance, latence et fiabilité dans les systèmes SIET.

1.1 Canal Gaussien à interférence avec récolte d'énergie

Considérons un canal d'interférence gaussien (G-IC) avec un collecteur d'énergie non colocalisé représenté sur la figure 5.1 sans voie de retour et sur la figure 5.2 avec une voie de retour canal-sortie parfaite. L'émetteur i , avec $i \in \{1, 2\}$, vise à exécuter deux tâches : (a) transmission d'informations (b) transmission d'énergie.

1.1.1 Transmission de l'information Du point de vue de la transmission d'information, le but de l'émetteur i est de transmettre un indice de message $W_i \in \mathcal{W}_i = \{1, 2, \dots, 2^{NR_i}\}$ au récepteur i en utilisant N symboles d'entrée de canal $X_{i,1}, X_{i,2}, \dots, X_{i,N}$. L'indice W_i est supposé être uniformément distribué sur \mathcal{W} . Le coefficient de canal entre l'émetteur k et le récepteur i , avec $k \in \{1, 2\}$, est noté par $h_{i,k} \in \mathbb{R}_+$. Au récepteur i , pendant l'utilisation du canal n , le symbole d'entrée $X_{i,n}$ est observé avec l'interférence produite par le symbole $X_{j,n}$ envoyé par l'émetteur j , avec $j \in \{1, 2\} \setminus \{i\}$, et un bruit Gaussien additif réel $Z_{i,n}$ de moyenne nulle et de variance σ_i^2 . Par conséquent, la sortie du canal au niveau du récepteur i pendant l'utilisation du canal n , dénote par $Y_{i,n}$, est :

$$Y_{i,n} = h_{i,i}X_{i,n} + h_{i,j}X_{j,n} + Z_{i,n}. \quad (1.1)$$

Dans le cas sans voie de retour, à chaque utilisation du canal n , le symbole $X_{i,n}$ envoyé par l'émetteur i dépend de l'indice de message W_i et d'un indice généré aléatoirement $\Omega \in \mathbb{N}$ indépendant de W_1 et W_2 . Soit $f_{i,n}^{(N)} : \mathcal{W}_i \times \mathbb{N} \rightarrow \mathbb{R}$ la fonction de codage à l'utilisation du canal n , telle que pour tout $n \in \{1, 2, \dots, N\}$ est donné par :

$$X_{i,n} = f_{i,n}^{(N)}(W_i, \Omega). \quad (1.2)$$

Dans le cas avec voie de retour, le symbole $X_{i,n}$ envoyé par l'émetteur i dépend des indices W_i et Ω et aussi de toutes les sorties de canal précédentes $Y_{i,1}, Y_{i,2}, \dots, Y_{i,n-d}$, avec $d \in \mathbb{N}$ le délai de rétroaction. De plus, on suppose que d est égal à l'utilisation d'un canal. Plus précisément, pour tout $n \in \{2, 3, \dots, N\}$, $f_{i,1}^{(N)} : \mathcal{W}_i \times \mathbb{N} \rightarrow \mathbb{R}$ et $f_{i,n}^{(N)} : \mathcal{W}_i \times \mathbb{N} \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ sont des fonctions de codage telles que :

$$X_{i,1} = f_{i,1}^{(N)}(W_i, \Omega) \quad \text{et} \quad (1.3a)$$

$$X_{i,n} = f_{i,n}^{(N)}(W_i, \Omega, Y_{i,1}, Y_{i,2}, \dots, Y_{i,n-1}). \quad (1.3b)$$

Dans les deux cas, avec et sans voie de retour, l'indice aléatoire Ω est supposé être connu par tous les émetteurs et récepteurs. De plus, les symboles d'entrée du canal $X_{i,1}, X_{i,2}, \dots, X_{i,N}$ sont soumis à une contrainte de puissance moyenne de la forme

$$\frac{1}{N} \sum_{n=1}^N \mathbb{E}_{X_{i,n}} [X_{i,n}^2] \leq P_i, \quad (1.4)$$

où P_i désigne la puissance moyenne de l'émetteur i en unités d'énergie par utilisation de canal.

Le récepteur i observe les sorties du canal $Y_{i,1}, Y_{i,2}, \dots, Y_{i,N}$ et utilise une fonction de décodage $\phi_i^{(N)} : \mathbb{N} \times \mathbb{R}^N \rightarrow \mathcal{W}_i$, pour obtenir une estimation de l'indice des messages :

$$\widehat{W}_i = \phi_i^{(N)}(\Omega, Y_{i,1}, Y_{i,2}, \dots, Y_{i,N}), \quad (1.5)$$

où \widehat{W}_i est une estimation de l'indice du message W_i . La probabilité d'erreur de décodage d'un dictionnaire de longueur de bloc N , désignée par $P_{DE}^{(N)}$, est donnée par

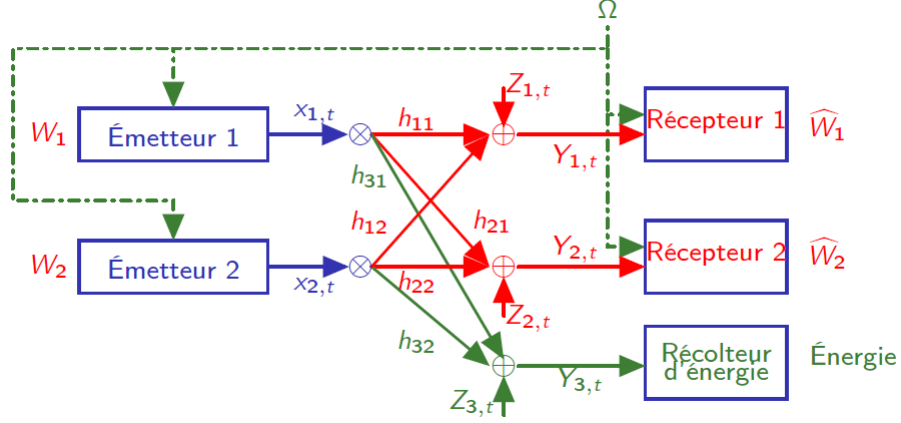


FIGURE 1.1. : Canal Guassien à interférence avec un collecteur d'énergie non-cocalisé à l'utilisation du canal n .

$$P_{\text{DE}}^{(N)} = \max \left[\Pr \left(\widehat{W}_1 \neq W_1 \right), \Pr \left(\widehat{W}_2 \neq W_2 \right) \right]. \quad (1.6)$$

Le rapport signal/bruit (SNR) au niveau du récepteur i est défini comme suit

$$\text{SNR}_i \triangleq \frac{|h_{i,i}|^2 P_i}{\sigma_i^2}. \quad (1.7a)$$

De plus, Le rapport interférence/bruit (INR) au récepteur i est défini comme suit

$$\text{INR}_i \triangleq \frac{|h_{i,j}|^2 P_j}{\sigma_i^2}, \text{ avec } j \neq i. \quad (1.7b)$$

1.1.2 Transmission d'énergie Soit $h_{3i} \in \mathbb{R}_+$ est le coefficient de canal de l'émetteur i au collecteur d'énergie (EH). Les symboles envoyés par les émetteurs pendant l'utilisation du canal n sont observés par le EH sous réserve d'un bruit Gaussien additif $Z_{3,n}$ de moyenne nulle et de variance σ_3^2 . Plus précisément, la sortie du canal au niveau de l'EH pendant l'utilisation du canal n , désignée par $Y_{3,n}$, est :

$$Y_{3,n} = h_{3,1}X_{1,n} + h_{3,2}X_{2,n} + Z_{3,n}. \quad (1.8)$$

Du point de vue de la transmission d'énergie, l'objectif des deux émetteurs est de assurer conjointement un débit suffisent d'énergie moyen à l'EH. Soit $B^{(N)} : \mathbb{R}^N \rightarrow \mathbb{R}_+$ est une fonction qui détermine l'énergie obtenue à partir des sorties du canal $Y_{3,1}, Y_{3,2}, \dots, Y_{3,N}$. Dans la suite, cette fonction est choisie comme étant le débit d'énergie moyen (en unités d'énergie par utilisation de canal) à la fin de N utilisations de canal. C'est-à-dire ,

$$B^{(N)}(Y_{3,1}, Y_{3,2}, \dots, Y_{3,N}) \triangleq \frac{1}{N} \sum_{n=1}^N Y_{3,n}^2, \quad (1.9)$$

Le SNR de l'émetteur i à l'EH est noté par

$$\limsup_{N \rightarrow \infty} P_{\text{ES}}^{(N)} = 0. \quad (1.13b)$$

En utilisant la définition 8, les limites fondamentales de la transmission simultanée d'informations et d'énergie dans le G-IC est décrites par la région de capacité de l'information et de l'énergie, définie comme suit.

Definition 2 (Région de capacité d'information et d'énergie). *La région de capacité d'information et d'énergie, désignée par \mathcal{E}^F dans le cas avec rétroaction et \mathcal{E} dans le cas sans rétroaction, correspond à la fermeture de tous les triplets de débit d'information et d'énergie réalisables (R_1, R_2, B) .*

1.1.4 Résultats Principaux Le résultat principal consiste en une description des régions de capacité d'information et d'énergie avec rétroaction \mathcal{E}^F et sans rétroaction \mathcal{E} . Une telle description est présentée sous la forme d'une *approximation* au sens de la définition ci-dessous.

Definition 3 (Approximation d'un ensemble). *Soit $n \in \mathbb{N}$ fixé. Un ensemble $\mathcal{X} \subset \mathbb{R}_+^n$ est approximé par les ensembles $\underline{\mathcal{X}}$ et $\overline{\mathcal{X}}$ si $\underline{\mathcal{X}} \subseteq \mathcal{X} \subseteq \overline{\mathcal{X}}$ et $\forall \mathbf{x} = (x_1, \dots, x_n) \in \overline{\mathcal{X}}$ alors $\exists (\xi_1, \xi_2, \dots, \xi_n) \in \mathbb{R}_+^n$ tel que $((x_1 - \xi_1)^+, (x_2 - \xi_2)^+, \dots, (x_n - \xi_n)^+) \in \underline{\mathcal{X}}$*

1.1.5 Cas sans voie de retour La région de capacité d'énergie d'information E , est approximée par les régions $\underline{\mathcal{E}}$ (Théorème 12), qui représente une région réalisable d'information et d'énergie, et $\overline{\mathcal{E}}$ (Théorème 13), qui représente une région d'impossibilité d'information et d'énergie.

Région atteignable Le théorème suivant introduit une région d'énergie d'information réalisable.

Theorem 1. *La région de capacité énergétique d'information \mathcal{E} contient l'ensemble $\underline{\mathcal{E}} \subseteq \mathbb{R}_+^3$ de tous les tuples de débits (R_1, R_2, B) qui satisfont :*

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{(1 - \lambda_{1e})\text{SNR}_1}{1 + \lambda_{2p}\text{INR}_1} \right), \quad (1.14a)$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{(1 - \lambda_{2e})\text{SNR}_2}{1 + \lambda_{1p}\text{INR}_2} \right), \quad (1.14b)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{1e})\text{SNR}_1 + (1 - \lambda_{2e})\text{INR}_1}{1 + \lambda_{2p}\text{INR}_1} \right) + \frac{1}{2} \log \left(1 + \frac{\lambda_{2p}\text{SNR}_2}{1 + \lambda_{1p}\text{INR}_2} \right), \quad (1.14c)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{2e})\text{SNR}_2 + (1 - \lambda_{1e})\text{INR}_2}{1 + \lambda_{1p}\text{INR}_2} \right) + \frac{1}{2} \log \left(1 + \frac{\lambda_{1p}\text{SNR}_1}{1 + \lambda_{2p}\text{INR}_1} \right), \quad (1.14d)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(\frac{1 + \lambda_{1p}\text{SNR}_1 + (1 - \lambda_{2e})\text{INR}_1}{1 + \lambda_{2p}\text{INR}_1} \right) + \frac{1}{2} \log \left(\frac{1 + \lambda_{2p}\text{SNR}_2 + (1 - \lambda_{1e})\text{INR}_2}{1 + \lambda_{1p}\text{INR}_2} \right) \quad (1.14e)$$

$$2R_1 + R_2 \leq \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{1e})\text{SNR}_1 + (1 - \lambda_{2e})\text{INR}_1}{1 + \lambda_{2p}\text{INR}_1} \right) + \frac{1}{2} \log \left(\frac{1 + \lambda_{2p}\text{SNR}_2 + (1 - \lambda_{1e})\text{INR}_2}{1 + \lambda_{1p}\text{INR}_2} \right) + \frac{1}{2} \log \left(1 + \frac{\lambda_{1p}\text{SNR}_1}{1 + \lambda_{2p}\text{INR}_1} \right), \quad (1.14f)$$

$$R_1 + 2R_2 \leq \frac{1}{2} \log \left(\frac{1 + (1-\lambda_{2e})\text{SNR}_2 + (1-\lambda_{1e})\text{INR}_2}{1 + \lambda_{1p}\text{INR}_2} \right) + \frac{1}{2} \log \left(\frac{1 + \lambda_{1p}\text{SNR}_1 + (1-\lambda_{2e})\text{INR}_1}{1 + \lambda_{2p}\text{INR}_1} \right) + \frac{1}{2} \log \left(1 + \frac{\lambda_{2p}\text{SNR}_2}{1 + \lambda_{1p}\text{INR}_2} \right), \quad (1.14g)$$

$$B \leq \sigma_3^2 \left(1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}\sqrt{\lambda_{1e}\lambda_{2e}} \right), \quad (1.14h)$$

avec $(\lambda_{ip}, \lambda_{ie}) \in [0, 1]^2$ tel que $\lambda_{ip} + \lambda_{ie} \leq 1$, pour tout $i \in \{1, 2\}$.

Région d'impossibilité Le théorème suivant introduit la region d'impossibilite d'information et d'énergie.

Theorem 2. *La région de capacité d'information et d'énergie \mathcal{E} contient l'ensemble $\bar{\mathcal{E}} \subseteq \mathbb{R}_+^3$ de tous les tuples de débits (R_1, R_2, B) qui satisfont :*

$$R_1 \leq \frac{1}{2} \log(1 + \beta_1 \text{SNR}_1), \quad (1.15a)$$

$$R_2 \leq \frac{1}{2} \log(1 + \beta_2 \text{SNR}_2), \quad (1.15b)$$

$$R_1 + R_2 \leq \frac{1}{2} \log(1 + \beta_1 \text{SNR}_1 + \beta_2 \text{INR}_1) + \frac{1}{2} \log \left(1 + \frac{\beta_2 \text{SNR}_2}{1 + \beta_2 \text{INR}_1} \right), \quad (1.15c)$$

$$R_1 + R_2 \leq \frac{1}{2} \log(1 + \beta_2 \text{SNR}_2 + \beta_1 \text{INR}_2) + \frac{1}{2} \log \left(1 + \frac{\beta_1 \text{SNR}_1}{1 + \beta_1 \text{INR}_2} \right), \quad (1.15d)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{\beta_1 \text{SNR}_1 + \beta_2 \text{INR}_1 + \beta_1 \beta_2 \text{INR}_1 \text{INR}_2}{1 + \beta_1 \text{INR}_2} \right) + \frac{1}{2} \log \left(1 + \frac{\beta_2 \text{SNR}_2 + \beta_1 \text{INR}_2 + \beta_1 \beta_2 \text{INR}_1 \text{INR}_2}{1 + \beta_2 \text{INR}_1} \right), \quad (1.15e)$$

$$2R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{\beta_1 \text{SNR}_1}{1 + \beta_1 \text{INR}_2} \right) + \frac{1}{2} \log(1 + \beta_1 \text{SNR}_1 + \beta_2 \text{INR}_1) + \frac{1}{2} \log \left(1 + \frac{\beta_2 \text{SNR}_2 + \beta_1 \text{INR}_2 + \beta_1 \beta_2 \text{INR}_1 \text{INR}_2}{1 + \beta_2 \text{INR}_1} \right), \quad (1.15f)$$

$$R_1 + 2R_2 \leq \frac{1}{2} \log \left(1 + \frac{\beta_2 \text{SNR}_2}{1 + \beta_2 \text{INR}_1} \right) + \frac{1}{2} \log(1 + \beta_2 \text{SNR}_2 + \beta_1 \text{INR}_2) + \frac{1}{2} \log \left(1 + \frac{\beta_1 \text{SNR}_1 + \beta_2 \text{INR}_1 + \beta_1 \beta_2 \text{INR}_1 \text{INR}_2}{1 + \beta_1 \text{INR}_2} \right), \quad (1.15g)$$

$$B \leq \sigma_3^2 \left(1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}\sqrt{(1-\beta_1)(1-\beta_2)} \right), \quad (1.15h)$$

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

Le dictionnaire de l'émetteur i , avec $i \in \{1, 2\}$, est généré par la superposition de trois couches de code différentes. La première couche de code est un sous-code généré dans le but de la transmission d'énergie. Notez que cette couche de code peut être choisie comme étant la même pour les deux émetteurs. Pour chaque mot de code de la première couche, un nouveau sous-code est généré. Cet ensemble de sous-code est appelé la deuxième couche de code et il est construit pour diffuser des informations aux deux récepteurs. Cependant, même si ce n'est pas l'objectif principale, ces mots de code transportent naturellement de l'énergie vers l'EH également. Enfin, pour chaque mot de code de la deuxième couche, un nouveau sous dictionnaire est généré. Cet ensemble de sous-code est appelé la troisième couche du

dictionnaire et il est construit dans l'objectif de transmettre exclusivement de l'information au récepteur i .

Theorem 3 (Approximation de \mathcal{E}). *Soit $\underline{\mathcal{E}} \subset \mathbb{R}_+^3$ et $\bar{\mathcal{E}} \subset \mathbb{R}_+^3$ sont les ensembles de tuples (R_1, R_2, B) décrits par le théorème 12 et le théorème 13, respectivement, comme le suit :*

$$\underline{\mathcal{E}} \subset \mathcal{E} \subset \bar{\mathcal{E}}, \quad (1.16)$$

et pour tout $(R_1, R_2, B) \in \bar{\mathcal{E}}$ il s'ensuit que $\left((R_1 - 1/2)^+, (R_2 - 1/2)^+, \left(B - \frac{B_{\max}}{2} \right)^+ \right) \in \underline{\mathcal{E}}$.

$$\begin{aligned} \frac{B - B'}{B_{\max}} &= \frac{2h_{3,1}h_{3,2}\sqrt{P_1P_2} \left(\sqrt{(1-\beta_1)(1-\beta_2)} - \sqrt{\lambda_{1e}\lambda_{2e}} \right)}{\sigma_3^2 + h_{3,1}^2P_1 + h_{3,2}^2P_2 + 2|h_{3,1}||h_{3,2}|\sqrt{P_1P_2}} \\ &\leq \frac{2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}}{1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}} \\ &\leq \frac{2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}}{1 + 4\sqrt{\text{SNR}_{31}\text{SNR}_{32}}} \\ &\leq \frac{1}{2}. \end{aligned}$$

1.1.6 Cas avec une rétroaction parfaite entre le canal et la sortie La région de capacité d'information et d'énergie \mathcal{E}^F contient l'ensemble $\mathcal{E}^F \subseteq \mathbb{R}_+^3$ de tous les tuples de débit (R_1, R_2, B) qui satisfont

Theorem 4.

$$R_1 \leq \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{1e})\text{SNR}_1 + (1 - \lambda_{2e})\text{INR}_1 + 2\rho\sqrt{\text{SNR}_1\text{INR}_1}}{1 + \lambda_{2p}\text{INR}_1} \right), \quad (1.17a)$$

$$R_1 \leq \frac{1}{2} \log \left(\frac{1 + (1 - (\rho + \lambda_{1e}))\text{INR}_2}{1 + \lambda_{1p}\text{INR}_2} \right) + \frac{1}{2} \log \left(\frac{1 + \lambda_{1p}\text{SNR}_1 + \lambda_{2p}\text{INR}_1}{1 + \lambda_{2p}\text{INR}_1} \right), \quad (1.17b)$$

$$R_2 \leq \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{2e})\text{SNR}_2 + (1 - \lambda_{1e})\text{INR}_2 + 2\rho\sqrt{\text{SNR}_2\text{INR}_2}}{1 + \lambda_{1p}\text{INR}_2} \right), \quad (1.17c)$$

$$R_2 \leq \frac{1}{2} \log \left(\frac{1 + (1 - (\rho + \lambda_{2e}))\text{INR}_1}{1 + \lambda_{2p}\text{INR}_1} \right) + \frac{1}{2} \log \left(\frac{1 + \lambda_{2p}\text{SNR}_1 + \lambda_{1p}\text{INR}_1}{1 + \lambda_{1p}\text{INR}_1} \right) \quad (1.17d)$$

$$\begin{aligned} R_1 + R_2 &\leq \frac{1}{2} \log \left(\frac{1 + \lambda_{1p}\text{SNR}_1 + \lambda_{2p}\text{INR}_1}{1 + \lambda_{2p}\text{INR}_1} \right) \\ &\quad + \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{2e})\text{SNR}_2 + (1 - \lambda_{1e})\text{INR}_2 + 2\rho\sqrt{\text{SNR}_2\text{INR}_2}}{1 + \lambda_{1p}\text{INR}_2} \right), \quad (1.17e) \end{aligned}$$

$$\begin{aligned} R_1 + R_2 &\leq \frac{1}{2} \log \left(\frac{1 + \lambda_{2p}\text{SNR}_1 + \lambda_{1p}\text{INR}_1}{1 + \lambda_{1p}\text{INR}_1} \right) \\ &\quad + \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{1e})\text{SNR}_1 + (1 - \lambda_{2e})\text{INR}_1 + 2\rho\sqrt{\text{SNR}_1\text{INR}_1}}{1 + \lambda_{2p}\text{INR}_1} \right), \quad (1.17f) \end{aligned}$$

$$B \leq \sigma_3^2 \left(1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}(\rho + \sqrt{\lambda_{1e}\lambda_{2e}}) \right), \quad (1.17g)$$

où $(\rho, \lambda_{ip}, \lambda_{ie}) \in [0, 1]^3$ et $\rho + \lambda_{ip} + \lambda_{ie} \leq 1$, pour tout $i \in \{1, 2\}$.

Le dictionnaire de l'émetteur i , avec $i \in \{1, 2\}$, est généré par la superposition de quatre sous code des codes différents. Cela contraste avec le dictionnaire à trois couches utilisé dans le cas sans rétroaction. Cependant, les deux dictionnaires partagent de profondes similitudes. La première couche dans le cas avec et sans voie de retour est identique et joue le même rôle. La deuxième couche du dictionnaire avec voie de retour est obtenue par l'union des deuxièmes couches des deux émetteurs dans le cas sans voie de retour. Les troisième et quatrième couches du dictionnaire avec voie de retour sont identiques aux deuxième et troisième couches du dictionnaire sans rétroaction. Les rôles de ces deux couches sont identiques dans le cas avec et sans voie de retour.

Région d'impossibilité Le théorème suivant décrit une région converse désignée par $\bar{\mathcal{E}}^F$.

Theorem 5. *La région de capacité d'information et d'énergie \mathcal{E}^F est contenue dans l'ensemble $\bar{\mathcal{E}}^F \in \mathbb{R}_+^3$, qui contient tous les tuples de débit (R_1, R_2, B) qui satisfont :*

$$R_1 \leq \frac{1}{2} \log \left(1 + \beta_1 \text{SNR}_1 + \beta_2 \text{INR}_1 + 2\rho \sqrt{\beta_1 \text{SNR}_1 \beta_2 \text{INR}_1} \right), \quad (1.18a)$$

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{\beta_1(1 - \rho^2) \text{SNR}_1}{1 + \beta_1(1 - \rho^2) \text{INR}_2} \right) + \frac{1}{2} \log \left(1 + \beta_1(1 - \rho^2) \text{INR}_2 \right), \quad (1.18b)$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \beta_2 \text{SNR}_2 + \beta_1 \text{INR}_2 + 2\rho \sqrt{\beta_2 \text{SNR}_2 \beta_1 \text{INR}_2} \right), \quad (1.18c)$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{\beta_2(1 - \rho^2) \text{SNR}_2}{1 + \beta_2(1 - \rho^2) \text{INR}_1} \right) + \frac{1}{2} \log \left(1 + \beta_2(1 - \rho^2) \text{INR}_1 \right), \quad (1.18d)$$

$$\begin{aligned} R_1 + R_2 &\leq \frac{1}{2} \log \left(1 + \frac{\beta_1(1 - \rho^2) \text{SNR}_1}{1 + \beta_1(1 - \rho^2) \text{INR}_2} \right) \\ &\quad + \frac{1}{2} \log \left(1 + \beta_2 \text{SNR}_2 + \beta_1 \text{INR}_2 + 2\rho \sqrt{\beta_2 \text{SNR}_2 \beta_1 \text{INR}_2} \right), \end{aligned} \quad (1.18e)$$

$$\begin{aligned} R_1 + R_2 &\leq \frac{1}{2} \log \left(1 + \frac{\beta_2(1 - \rho^2) \text{SNR}_2}{1 + \beta_2(1 - \rho^2) \text{INR}_1} \right) \\ &\quad + \frac{1}{2} \log \left(1 + \beta_1 \text{SNR}_1 + \beta_2 \text{INR}_1 + 2\rho \sqrt{\beta_1 \text{SNR}_1 \beta_2 \text{INR}_1} \right), \end{aligned} \quad (1.18f)$$

$$B \leq \sigma_3^2 \left(1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31} \text{SNR}_{32}} (\rho \sqrt{\beta_1 \beta_2} + \sqrt{(1 - \beta_1)(1 - \beta_2)}) \right), \quad (1.18g)$$

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

La preuve repose sur le fait que le dictionnaire n'est pas supposés être formés par des mots de code de moyenne nulle. En d'autres termes, le mot de code considéré peut avoir une moyenne non nulle. Les limites supérieures des débit d'information s'appuient fortement sur les limites des ensembles de coupures [5], l'inégalité de Fano [6] et les modèles assistés par génie. La limite supérieure du débit de transmission d'énergie est une conséquence immédiate de l'inégalité de Markov [7].

Approximation de la région de capacité d'information et d'énergie En utilisant la région intérieure $\underline{\mathcal{E}}^F$ et la région extérieure $\overline{\mathcal{E}}^F$, décrites respectivement par le théorème 15 et le théorème 16, la région de capacité d'énergie et d'information \mathcal{E}^F peut être approximée au sens de la définition 10.

Theorem 6 (Approximation de \mathcal{E}^F). *Soit $\underline{\mathcal{E}}^F \subset \mathbb{R}_+^3$ et $\overline{\mathcal{E}}^F \subset \mathbb{R}_+^3$ sont les ensembles de tuples (R_1, R_2, B) décrits par le théorème 15 et le théorème 16, respectivement. Alors,*

$$\underline{\mathcal{E}}^F \subset \mathcal{E}^F \subset \overline{\mathcal{E}}^F, \quad (1.19)$$

et pour tout $(R_1, R_2, B) \in \overline{\mathcal{E}}^F$ il s'ensuit que $\left((R_1 - 1)^+, (R_2 - 1)^+, \left(B - \frac{B_{\max}}{2} \right)^+ \right) \in \underline{\mathcal{E}}_b^F$.

1.1.7 Amélioration maximale du débit d'énergie avec la voie de retour Considérons les ensembles suivants de débits d'énergie : $\underline{\mathcal{B}} = \{b \in \mathbb{R}_+ : (R_1, R_2, b) \in \underline{\mathcal{E}}\}$, $\overline{\mathcal{B}} = \{b \in \mathbb{R}_+ : (R_1, R_2, b) \in \overline{\mathcal{E}}\}$, $\underline{\mathcal{B}}_F = \{b \in \mathbb{R}_+ : (R_1, R_2, b) \in \underline{\mathcal{E}}^F\}$, et $\overline{\mathcal{B}}_F = \{b \in \mathbb{R}_+ : (R_1, R_2, b) \in \overline{\mathcal{E}}^F\}$.

On peut montrer que l'amélioration maximale que l'on peut obtenir sur le débit d'énergie grâce à la voie de retour est double que celle dans le cas sans voie de retour. La proposition suivante montre ce résultat en majorant les limites supérieures des ratios

$$\frac{\max \underline{\mathcal{B}}_F}{\max \underline{\mathcal{B}}} \text{ et } \frac{\max \overline{\mathcal{B}}_F}{\max \overline{\mathcal{B}}}.$$

Proposition 1 (Amélioration du débit avec la voie de retour). Le débit d'énergie réalisable dans le G-IC à deux utilisateurs avec une rétroaction parfaite du canal de sortie peut être le double du débit d'énergie réalisable dans le G-IC à deux utilisateurs sans rétroaction. C'est-à-dire

$$1 < \frac{\max \underline{\mathcal{B}}_F}{\max \underline{\mathcal{B}}} \leq 2. \quad (1.20)$$

Toute amélioration au-delà d'un facteur deux n'est pas faisable, comme le montre l'inégalité suivante,

$$1 < \frac{\max \overline{\mathcal{B}}_F}{\max \overline{\mathcal{B}}} \leq 2. \quad (1.21)$$

1.2 Récolte d'énergie avec des contraintes de fiabilité et de latence

Considérons un système de communication à trois parties, où l'émetteur transmet simultanément de l'information à un récepteur d'information (RI) et de l'énergie à un récepteur EH par l'intermédiaire d'un canal symétrique binaire. Pour le cas non-asymptotique, la transformation aléatoire pour ce système est donnée par

$$(\{0, 1\}^n, \{0, 1\}^n \times \{0, 1\}^n, P_{\mathbf{Y}\mathbf{Z}|\mathbf{X}}), \quad (1.22)$$

où $n \in \mathbb{N}$ présente la longueur du bloc. Étant donné une entrée $\mathbf{x} \triangleq (x_1, x_2, \dots, x_n) \in \{0, 1\}^n$, les sorties $\mathbf{y} \triangleq (y_1, y_2, \dots, y_n) \in \{0, 1\}^n$ et $\mathbf{z} \triangleq (z_1, z_2, \dots, z_n) \in \{0, 1\}^n$ sont observés dans le RI et dans l'EH, respectivement, avec la probabilité suivante

$$P_{\mathbf{Y}\mathbf{Z}|\mathbf{X}}(\mathbf{y}, \mathbf{z}|\mathbf{x}) = \prod_{t=1}^n P_{Y|X}(y_t|x_t)P_{Z|X}(z_t|x_t), \quad (1.23)$$

où pour tous les $(x, y, z) \in \{0, 1\}^3$,

$$P_{Y|X}(y|x) = \alpha_1 \mathbb{1}_{\{x \neq y\}} + (1 - \alpha_1) \mathbb{1}_{\{x=y\}}, \quad (1.24)$$

$$P_{Z|X}(z|x) = \alpha_2 \mathbb{1}_{\{x \neq z\}} + (1 - \alpha_2) \mathbb{1}_{\{x=z\}}, \quad (1.25)$$

et $\alpha_1 \in [0, \frac{1}{2})$ et $\alpha_2 \in (0, \frac{1}{2})$. Dans ce contexte, deux tâches sont effectuées par l'émetteur : (a) la transmission d'information ; et (b) la transmission d'énergie.

1.2.1 Transmission de l'information L'objectif de cette tâche est d'envoyer un message de l'émetteur vers le receveur d'information (RI). L'indice du message est une réalisation d'une variable aléatoire uniformément distribuée dans $\{1, 2, \dots, M\}$, avec $M \in \mathbb{N}$. Pour transmettre l'information dans n utilisations de canaux, l'émetteur utilise un code (n, M) défini comme suit

Definition 4 ((n, M) -code). An (n, M) -code pour la transformation aléatoire dans (6.1) est un système

$$\{(\mathbf{u}(1), \mathcal{D}_1), (\mathbf{u}(2), \mathcal{D}_2), \dots, (\mathbf{u}(M), \mathcal{D}_M)\}, \quad (1.26)$$

où pour tous les $(i, j) \in \{1, 2, \dots, M\}^2$, avec $i \neq j$,

$$\mathbf{u}(i) \triangleq (u_1(i), u_2(i), \dots, u_n(i)) \in \{0, 1\}^n, \quad (1.27a)$$

$$\mathcal{D}_i \cap \mathcal{D}_j = \emptyset, \text{ and} \quad (1.27b)$$

$$\bigcup_{i=1}^M \mathcal{D}_i \subseteq \{0, 1\}^n. \quad (1.27c)$$

Pour tout $i \in \{1, 2, \dots, M\}$, l'émetteur encode l'indice de message i , dans un mot de code $\mathbf{u}(i) = (u_1(i), \dots, u_n(i))$. Au bout de n utilisations du canal, le RI observe la sortie $\mathbf{y} = (y_1, \dots, y_n)$ et décide que le symbole i a été transmis s'il satisfait

$$(y_1, y_2, \dots, y_n) \in \mathcal{D}_i. \quad (1.28)$$

Dénotons par $\lambda_i \in [0, 1]$ la probabilité d'erreur de décodage associée à la transmission du message d'indice i , qui est donnée par

$$\lambda_i \triangleq \Pr[\mathbf{Y} \in \mathcal{D}_i^c | \mathbf{X} = \mathbf{u}(i)], \quad (1.29)$$

où la probabilité est prise par rapport à la distribution marginal $P_{\mathbf{Y}|\mathbf{X}}$, et \mathcal{D}_i^c représente l'ensemble complémentaire de \mathcal{D}_i par rapport à $\{0, 1\}^n$. La probabilité moyenne d'erreur, notée λ est donnée par

$$\lambda \triangleq \frac{1}{M} \sum_{m=1}^M \lambda_m. \quad (1.30)$$

1.2.2 Transmission d'énergie Soit $g : \{0, 1\} \rightarrow \mathbb{R}_+$ est une fonction positive à valeur réelle qui détermine l'énergie récoltée à partir des symboles de sortie du canal. Soit

$$b_0 \triangleq g(0), \quad \text{et} \quad (1.31a)$$

$$b_1 \triangleq g(1) \quad (1.31b)$$

sont l'énergie récoltée lorsque les sorties du canal au niveau de l'EH sont respectivement de 0 et 1. À la fin de n utilisations de canaux, l'énergie moyenne délivrée à l'EH par les sorties du canal $\mathbf{z} = (z_1, z_2, \dots, z_n)$ est donnée par la fonction $B_n : \{0, 1\}^n \rightarrow \mathbb{R}_+$, avec

$$B_n(\mathbf{z}) \triangleq \frac{1}{n} \sum_{t=1}^n g(z_t) = (b_0 - b_1) \frac{N(0|\mathbf{z})}{n} + b_1. \quad (1.32)$$

Une transmission fiable de l'énergie se produit lorsque $B_n(\mathbf{z}) \geq b$, avec $b \geq 0$. Cependant, un événement de manque d'énergie se produit lorsque $B_n(\mathbf{z}) < b$ à la fin de la transmission. Pour le reste de la section, l'hypothèse suivante est adoptée sans perte de généralité

$$b_1 < b_0. \quad (1.33)$$

1.2.3 Limites du débit de transmission de l'énergie Cette section fournit quelques limites supérieures sur le débit de transmission d'énergie pour tout code $(n, M, \epsilon, \delta, b)$ donné avec des contraintes probabilité de manque d'énergie (ESP) maximales ou moyennes. Ces limites sont exprimées en fonction des paramètres du code et des distributions d'entrée empiriques induites par le code, défini comme suit

Definition 5 (Les distributions empiriques). *Soit un code (n, M) , pour tout $i \in \{1, 2, \dots, M\}$, la distribution de probabilité empirique des symboles d'entrée du canal induits par le mot de code $\mathbf{u}(i)$ est*

$$\bar{P}_X^{(i)}(0) \triangleq \frac{1}{n} N(0|\mathbf{u}(i)) = 1 - \bar{P}_X^{(i)}(1), \quad (1.34)$$

La distribution empirique des symboles induits conjointement par tous les mots de code est défini comme suit

$$\bar{P}_X(0) \triangleq \frac{1}{nM} \sum_{i=1}^M N(0|\mathbf{u}(i)) = 1 - \bar{P}_X(1). \quad (1.35)$$

Souvent, le vecteur $(\bar{P}_X^{(i)}(0), \bar{P}_X^{(i)}(1))$ est appelé le *type* du mot de code $\mathbf{u}(i)$ [8].

En utilisant les distributions empiriques de la définition 17, certaines limites supérieures du débit de transmission de l'énergie peuvent être décrites.

Proposition 2 (Approximation Gaussienne). *Considérons un code $(n, M, \epsilon, \delta, b)$ pour la transformation aléatoire dans (6.1) satisfaisant (6.12). Ensuite, soumis d'une contrainte ESP maximale, il s'avère que pour tout $i \in \{1, 2, \dots, M\}$,*

$$b \leq (b_0 - b_1) \left((1 - 2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 \right) + b_1 - \sqrt{\frac{(b_0 - b_1)^2 \alpha_2 (1 - \alpha_2)}{n}} Q^{-1} \left(\delta + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}} \right) \quad (1.36)$$

et sous d'une contrainte d'ESP moyen, le débit d'énergie b satisfait à

$$b \leq \hat{B}, \quad (1.37)$$

où \hat{B} est le plus grand réel positif qui satisfait

$$\frac{1}{M} \sum_{i=1}^M Q \left(\frac{n \left((1 - 2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 - \frac{\hat{B} - b_1}{b_0 - b_1} \right)}{\sqrt{n\alpha_2(1 - \alpha_2)}} \right) - \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}} < \delta. \quad (1.38)$$

Notez que la limite supérieure dans (6.30), est valide lorsque la condition suivante est satisfaite

$$0 < \delta + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)^2}} < 1, \quad (1.39)$$

étant donné que le domaine de la fonction Q^{-1} est $(0, 1)$.

1.2.4 Limites du débit de transmission de l'information Étant donné un code $(n, M, \epsilon, \delta, b)$, le lemme suivant décrit une borne supérieure sur M , qui ne dépend pas de la probabilité d'erreur de décodage ϵ . Cependant, elle joue un rôle important lorsque $\underline{b}(n, \delta) < b < \bar{b}(n, \delta)$.

Lemma 1. Étant donné un $(n, M, \epsilon, \delta, b)$ -code pour la transformation aléatoire dans (6.1) qui satisfait (6.12). Ensuite, en supposant à une contrainte de probabilité de manque d'énergie maximale, il s'avère que

$$M \leq \binom{n}{\lceil n\rho^*(b) \rceil} 2^{(n - \lceil n\rho^*(b) \rceil)}, \quad (1.40)$$

où $\rho^*(b)$ est défini par

$$\rho^*(b) \triangleq \min(1, \rho^+(b)), \quad (1.41)$$

avec $\rho^+ : [b_1, b_0] \rightarrow \mathbb{R}_+$ est défini comme le suit

$$\rho^+(b) \triangleq \frac{(b - b_1) - \alpha_2(b_0 - b_1)}{(b_0 - b_1)(1 - 2\alpha_2)} + \frac{\sqrt{\alpha_2(1 - \alpha_2)}}{\sqrt{n}(1 - 2\alpha_2)} Q^{-1} \left(\delta + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}} \right).$$

Notez que $\rho^*(b)$ est strictement croissante avec le débit d'énergie b . En fait, lorsque $\rho^*(b) \in (\frac{1}{2}, 1]$, le côté droit de (6.40) est monotone décroissant avec b . Cela met en évidence le compromis entre la transmission d'informations et la transmission d'énergie. C'est-à-dire que dans le régime dans lequel $\rho^*(b) \in (\frac{1}{2}, 1]$, l'augmentation du débit d'énergie impliquerait nécessairement la diminution du débit d'information.

1.2.5 Région de capacité de l'information et de l'énergie Étant donné une longueur de bloc fixe n et une paire $(\epsilon, \delta) \in [0, 1]^2$, la région de capacité d'information et d'énergie $\mathcal{C}(n, \epsilon, \delta)$ de la transformation aléatoire dans (6.1) soumise à (6.12) est approximée par un ensemble $\underline{\mathcal{C}}(n, \epsilon, \delta)$ qui est contenu dans $\mathcal{C}(n, \epsilon, \delta)$ (Théorème 18) et un autre ensemble $\bar{\mathcal{C}}(n, \epsilon, \delta)$ qui contient $\mathcal{C}(n, \epsilon, \delta)$ (Théorème 19). C'est-à-dire ,

$$\underline{\mathcal{C}}(n, \epsilon, \delta) \subseteq \mathcal{C}(n, \epsilon, \delta) \subseteq \bar{\mathcal{C}}(n, \epsilon, \delta). \quad (1.42)$$

Cette approximation est obtenue en considérant la contrainte de probabilité d'erreur de décodage (DEP) moyenne et une contrainte de DEP maximal. La notation suivante est utilisée pour décrire l'ensemble $\underline{\mathcal{C}}(n, \epsilon, \delta)$. Les fonctions $\phi : \mathbb{N} \times [0, 1] \rightarrow [0, 1]$, et $\chi : \mathbb{R}_+ \times [0, 1] \rightarrow [0, 1]$ sont définis comme suit,

$$\phi(m, \rho) \triangleq \min \left\{ 1, (m-1) \sum_{\ell_0=0}^n \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\min\{\ell_0, \ell_1\}} \sum_{\ell_3=0}^{\ell_0+\ell_1-2\ell_2} \sum_{\ell_4=0}^{\ell_3} \binom{n}{\ell_0} \binom{\ell_0}{\ell_2} \binom{n-\ell_0}{\ell_1-\ell_2} \binom{\ell_0}{\ell_4} \binom{n-\ell_0}{\ell_3-\ell_4} \cdot \alpha_1^{\ell_0+\ell_1-2\ell_2} (1-\alpha_1)^{n-\ell_0-\ell_1+2\ell_2} \rho^{\ell_3+\ell_0-2\ell_4+\ell_1} (1-\rho)^{2n-(\ell_3+\ell_0-2\ell_4+\ell_1)} \right\}, \quad (1.43)$$

et

$$\chi(s, \rho) \triangleq \sum_{t=0}^n \binom{n}{t} \rho^t (1-\rho)^{n-t} Q \left(\frac{(1-2\alpha_2)t + n \left(\alpha_2 - \frac{s-b_1}{b_0-b_1} \right)}{\sqrt{n\alpha_2(1-\alpha_2)}} \right) + \frac{(1-\alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}}. \quad (1.44)$$

Considérons les fonctions $M_1^* : [0, 1] \rightarrow \mathbb{N}$, et $B^* : [0, 1] \rightarrow \mathbb{R}_+$ être défini comme suit

$$M_1^*(\rho) \triangleq \begin{cases} \operatorname{argmax}_{m \in \mathbb{N}} \phi(m, \rho) \\ \text{s.t. } \phi(m, \rho) < \epsilon \end{cases}, \quad (1.45)$$

et

$$B^*(\rho) \triangleq \begin{cases} \operatorname{argmax}_{s \in \mathbb{R}_+} \chi(s, \rho) \\ \text{s.t. } \chi(s, \rho) < \delta \end{cases}. \quad (1.46)$$

En utilisant cette notation, étant donné une longueur de bloc fixe n et une paire $(\epsilon, \delta) \in [0, 1]^2$, le théorème suivant introduit l'ensemble $\underline{\mathcal{C}}(n, \epsilon, \delta)$, qui est contenu dans la région de capacité d'information et d'énergie $\mathcal{C}(n, \epsilon, \delta)$.

Theorem 7. *La région de capacité de l'information et de l'énergie $\mathcal{C}(n, \epsilon, \delta)$ de la transformation aléatoire dans (6.1) soumise à (6.12), contient l'ensemble*

$$\underline{\mathcal{C}}(n, \epsilon, \delta) \triangleq \left\{ (M, b) \in \mathbb{N} \times \mathbb{R}_+ : \exists \rho \in [0, 1], M < M_1^*(\rho) \text{ et } b < B^*(\rho) \right\}, \quad (1.47)$$

avec $M_1^* : [0, 1] \rightarrow \mathbb{N}$ est défini dans (6.45) et $B^* : [0, 1] \rightarrow \mathbb{R}_+$ est défini dans (6.46).

La description de l'ensemble $\overline{\mathcal{C}}(n, \epsilon, \delta)$ utilise la notation suivante. Considérons la fonction $\gamma : [0, 1]^2 \rightarrow [0, 1]$, telle que

$$\gamma(\rho, q) = \sum_{\ell_0=0}^n \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\min\{\ell_0, \ell_1\}} \binom{n}{\ell_0} \binom{\ell_0}{\ell_2} \binom{n-\ell_0}{\ell_1-\ell_2} \rho^{\ell_1} (1-\rho)^{n-\ell_1} q^{\ell_0} (1-q)^{n-\ell_0} \quad (1.48)$$

$$\left(\mathbb{1}_{\left\{ (\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} > L \right\}} + \lambda \mathbb{1}_{\left\{ (\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} = L \right\}} \right)$$

où $\lambda \in [0, 1]$ et $L \in \mathbb{R}$ sont choisis pour satisfaire

$$1 - \epsilon = \sum_{\ell_0=0}^n \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\min\{\ell_0, \ell_1\}} \binom{n}{\ell_0} \binom{\ell_0}{\ell_2} \binom{n-\ell_0}{\ell_1-\ell_2} (1-\alpha_1)^{n-(\ell_0+\ell_1-2\ell_2)} \alpha_1^{\ell_0+\ell_1-2\ell_2} \rho^{\ell_1} (1-\rho)^{n-\ell_1}$$

$$\cdot \left(\mathbb{1}_{\{(\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} > L\}} + \lambda \mathbb{1}_{\{(\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} = L\}} \right).$$

Considérons également la fonction $\Gamma : [b_1, b_0] \rightarrow \mathbb{R}_+$ définie comme suit

$$\Gamma(b) \triangleq \sup_{\rho > \rho^*(b)} \inf_{q \in [0, \frac{1}{2}]} \frac{1}{\gamma(\rho, q)}. \quad (1.49)$$

Enfin, laissons aussi $B^+ \in \mathbb{R}_+$ être défini par,

$$B^+ \triangleq (1 - \alpha_2)b_0 + \alpha_2 b_1 - \sqrt{\frac{(b_0 - b_1)^2 \alpha_2 (1 - \alpha_2)}{n}} \cdot Q^{-1} \left(\delta + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}} \right). \quad (1.50)$$

En utilisant cette notation, étant donné une longueur de bloc fixe n et une paire $(\epsilon, \delta) \in [0, 1]^2$, le théorème suivant introduit un ensemble, désigné par $\bar{\mathcal{C}}(n, \epsilon, \delta)$, qui contient la région de capacité de l'information et de l'énergie $\mathcal{C}(n, \epsilon, \delta)$.

Theorem 8. *La région de capacité de l'information et de l'énergie $\mathcal{C}(n, \epsilon, \delta)$ de la transformation aléatoire dans (6.1) soumise à (6.12), est contenue dans l'ensemble*

$$\bar{\mathcal{C}}(n, \epsilon, \delta) \triangleq \{(M, b) \in \mathbb{N} \times \mathbb{R}_+ : M < \Gamma(b) \text{ et } b < B^+\}, \quad (1.51)$$

avec $\Gamma : \mathbb{R}_+ \rightarrow [0, 1]$ est défini dans (6.49) et B^+ est défini dans (6.50).

Introduction

2.1 Energy Harvesting

Energy, in physics, is the capacity for doing work. Work is defined by a force applied to an object to change this object's physical properties. In general, this ability exists in different variations. Some of the variations are more abundant depending on the environment and can be available in the form of solar light, kinetic, wind, mechanical, gravitational, sound, nuclear or electromagnetic energy. The term "Energy Harvesting" refers to increasing the amount of energy through the adoption of an alternative energy resource. This approach would enable an energy-autonomous operation of low-energy consumption electronic devices. Therefore, it plays an important role in providing energy to the IoT (Internet of Things), in which billions of wireless sensor nodes are deployed. Accordingly, a transducer is used to harvest the ambient energy available at different forms within the environment. Usually, the amount of available energy is small and strictly depends on the environmental conditions. In general, energy harvesting can be categorized into some basic categories, namely light, wind, thermal, mechanical vibration, biochemical, and electromagnetic (RF) waves, among many others. The fundamentals of each category will be shortly discussed in the following sections to evaluate their feasibility as potential energy sources.

2.1.1 Solar Energy Harvesting Harvesting energy from solar light is the most common alternative energy source. Basically, it consists of transmitting the light into electricity using the advantage of the photovoltaic (PV) effect produced by semiconducting materials. A PV system uses a solar module, which contains several solar cells to convert the solar light into electrical energy. Due to their increased efficiency, decreasing cost and environmental concern, photovoltaic installations have increased dramatically in recent years. A better description of photovoltaic energy harvesters can be explained by examining a model of a solar cell. A PV cell is modeled as a current source in parallel with a diode as well as a shunt and a series of resistances to model non-ideality effects. Figure 2.1 shows the electrical characteristics curve. The red curve shows the relation between the current and voltage across the solar cell. It is clear that increasing illumination has a positive effect on the extracted power by increasing the short-circuit current. While the open-circuit voltage of a photovoltaic cell is not affected too much for illumination variations, it is affected greatly by temperature variations [9]. The

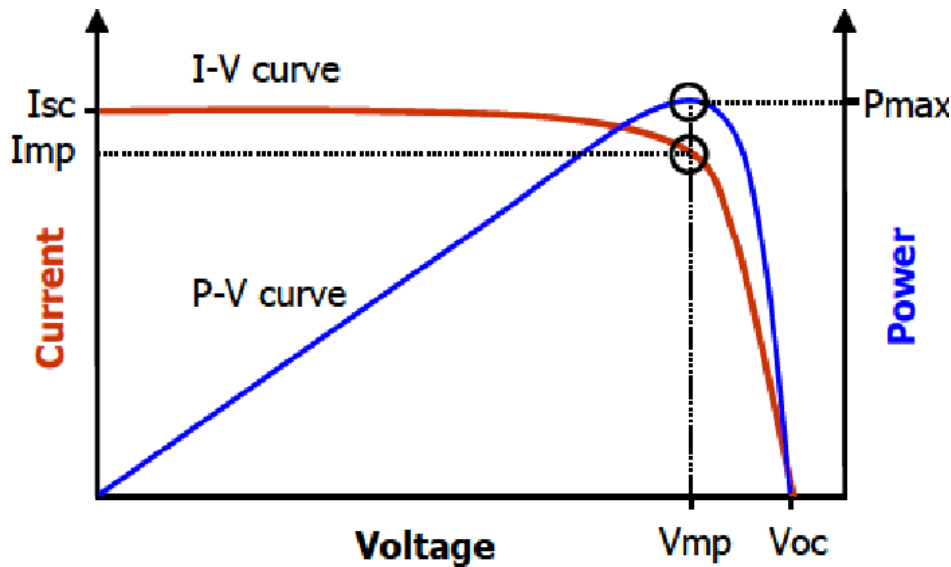


Figure 2.1.: Electrical characteristics of a solar cell model. Taken from [1].

power curve reaches a peak, known as the maximum power point at a particular voltage and current. Hence, designing an efficient solar energy harvester requires a good understanding of many factors such as illumination levels, daylight availability, photovoltaic cell properties, probability of shading on photovoltaic cells and power supply requirements [10].

2.1.2 Wind Energy Harvesting Wind energy is considered as a reliable source of free energy hundreds of years. Since the early 1980s, wind turbines played a significant role to convert the kinetic energy into mechanical energy. These mechanical sources can be used for particular purposes (such as or pumping water and pounding grain) and a generator can harvest this mechanical energy into electricity. For instance, the wind turbine-generator is an essential energy supplier for electrical grids and cellular base stations. Like solar energy, the wind is considered to be part of the renewable energy sources.

2.1.3 Thermal-Based Energy Harvesting Thermal energy harvesting is the process of scavenging heat which is either freely available in the environment or which is waste energy given off by engines, machines, and other sources and putting it to use. Reports shows that more than half of the energy produced from renewable and non-renewable energy sources in the world is rejected to the environment, mostly in the form of waste heat. Technically speaking based on the Seebeck effect, if two different metals joined in two junctions with a temperature difference between the junctions, an open circuit voltage would be established between them. Fundamentally, thermoelectric phenomena have been used for heating and cooling applications. By reversing the process, a temperature gradient in a conducting material results in heat flow which results in the diffusion of charge carriers.

2.1.4 Mechanical Energy Harvesting Mechanical energy harvesting is known also as waste energy harvesting, which is the process of using vibrations and convert it into electricity. This conversion can be performed in general using various mechanisms as electromagnetic,

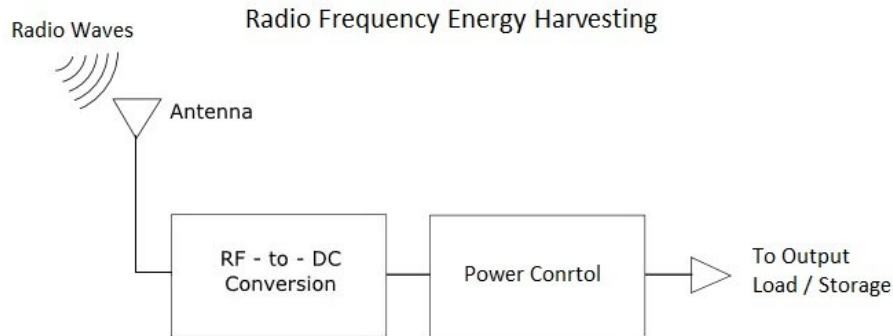


Figure 2.2.: RF scheme model for harvesting Energy. Taken from [2]

electrostatic and piezoelectric transductions. The most common one is the piezoelectric, because of its large power densities. This property allows the material to produce voltage when stimulated by mechanical stress. The main concept of an electromagnetic harvester is the use of an inductive spring-mass to convert the vibration into electricity. An electrostatic harvester, in general, takes place between the parallel plates of the capacitor on which electrical charge is stored. The idea is to oppose those plates where one is fixed and the other is initially charged and is moving, this allows the mechanical movement to be converted into electricity.

2.1.5 Biochemical Energy Harvesting This energy source includes also the energy that can be harvested from the human body. Due to its richness, there is a lot of chemical energy stored in the human body which can be converted into positive mechanical work. Harvesting energy could be realized outside the body from everyday human actions or intentionally generated by humans. Many devices could take advantage of the energy produced by the human body, to produce electricity, such as hand-crank generators, radios, and mobile phone chargers. However, devices could be implemented inside the body to harvest energy from movement.

2.1.6 RF Energy Harvesting Radio Frequency (RF) signals are an alternative available source for energy harvesting. Energy can be harvested from ubiquitous radio transmitters used for TV broadcasting and cell phone network coverage and is essentially a free-energy. But due to a factor like terrain, number of users, network congestion, there might be variations on the energy level. In addition to that, in some cases, it is hard to distinguish the useful RF signal and the waste RF signal. Through using a receiving antenna with a power conditioning circuit as shown in Figure 2.2, this free-flowing energy can be harvested and converted into DC voltage. Typical value for the energy transmitted through RF broadcasting reaches few tens of watts for frequencies in the order of GHz. However, the harvester device has to be close enough to the RF transmitter to harvest a useful amount of this power. Otherwise, the power density level would be very low. Power density levels of $0.1\text{-}1\text{ mW}/\text{m}^2$ would be expected at distances of 25-100m away from a GSM base station [11]. Intel demonstrated a prototype with an antenna of around 30 cm by 20 cm where $60\mu\text{W}$ has been harvested from a TV tower station away from the harvester by 4.1km [12].

2.2 Wireless Energy Transmission (WET)

Wireless energy transmission is the process of transmitting energy without any physical contact between the source and the load. Specifically, it eliminates the need to connecting the battery-depending devices to the power grid. Hence, it extends the battery lifetime and ensures sustainability of wireless networks such as sensor networks since their batteries are generally inaccessible and cannot be replaced. In this context, WET becomes an alternative to eliminate the need for battery recharging. Different WET technologies have been implemented and in the following sections, an overview of those technologies is given based on the work in [13].

2.2.1 WET Technologies

Inductive Coupling This technology is based on the principle of Ampere’s law and Faraday’s law of induction, in which energy is transmitted between two properly aligned transmitter/receiver coils by magnetic field. A time-varying magnetic field appears when an alternating current passes through a receiving coil. A consequence of Faraday’s low, it induces an alternating current in the receiving circuit, that could be served to create energy. An energy transmission efficiency up to 90 percent could be achieved using this technology [13].

Magnetic resonant coupling Another WET technology, is the magnetic resonant coupling that is based on the principle of resonant coupling. This principle states that two objects operating at the same frequency, increases the efficiency of their coupling [13]. There is a huge similarity between this technology and the previous one since they both use the magnetic field as a medium for WET. However, Magnetic resonant coupling is more suitable than inductive coupling over longer distances, since it attain higher energy efficiency.

Electromagnetic Radiation It is also known as radiative WET, and it has been especially used for wireless communication, and it provides a promising technology for WET. In the opposite, of the previous technologies, it is considered as a far-field wireless energy transmission technology, where the energy absorption by the receiver is not related to the radiation of the transmitter. In radiative WET, the main component is the rectifier, consists of a matching circuit, a diode, and a low-pass filter. The rectifier, capture the designated radio frequency signal and converted it into the usable direct current (DC).

Laser Power Beaming To achieve efficient energy delivery over long distances, this technology, uses highly concentrated laser light aiming at the energy receiver. Using the same approach, as solar energy, the receiver makes use of the photovoltaic cells to convert the received laser light into electricity. However, this technology has a limitation in practice, since laser beaming requires line-of-sight link, in addition to an accurate pointing to the photo-voltaic cell.

The table 2.1 summarizes the different WET technologies.

2.2.2 History and Evolution Since Faraday discovered the electromagnetic induction phenomenon in 1830 the physical wire is the main carrier of electricity transmission, but the traditional methods cannot meet the requirements of modern needs. People ask for convenience and safety. Besides, with the improvement of environmental protection consciousness, the

Table 2.1.: Comparison of the main technologies of WET. Taken from [13]

WPT technology	Main devices	Typical range	Typical frequency	Main advantages and limitations	Current and potential applications	Representative companies
Inductive coupling	Wire coils	Millimeters to centimeters	Hz to MHz	High efficiency, require precise tx/rx coil alignment, very short range, single receiver only	Electric tooth brush and razor battery charging, transcutaneous charging of bio-medical implants, electrical vehicle charging, cell phone charging, factory automation	Powermat, Delphi, GetPowerPad, WildCharge, Primove
Magnetic resonant coupling	Tuned wire coils, lumped element resonators	A few meters, typically 4 to 10 times the coil diameter	kHz to MHz	High efficiency, safe, mid-range, large tx/rx size	Consumer electronics (e.g., cell phones, laptops, household robots) charging, biomedical implants charging, electrical vehicles charging, RFID, smart cards, industrial applications	PowerbyProxi, WiTricity, WiPower, Intel (Wireless Resonant Energy Link)
EM radiation	Dish antenna, antenna array, rectenna	Several meters to hundreds of kilometers	MHz to dozens of GHz	Long range, small receiver form factors, flexible in deployment and movement, support power multicasting, potential for SWIPT, LoS link is not a must, low efficiency, safety and health issues	Wireless sensor charging, IoT, RFID, consumer electronics charging, wireless-powered aircrafts, solar power satellite	Intel (WISP), Energon (Wattup), PowerCast, Ossia (Cota)
Laser power beaming	Laser emitter, photovoltaic receiver	up to kilometers	THz	Compact size, high energy concentration, no interference to existing communication systems or electronics, laser radiation is hazardous, require LoS link and accurate receiver targeting, vulnerable to atmospheric absorption and scattering by clouds, fog, and rain	Laser-powered UAVs, laser-powered space elevator climbers, laser-based solar power satellite	LaserMotive

battery caused serious damage in the water and soil. These problems all contribute to the exploration and thinking of the new method. In the early 1900s, Heinrich Hertz demonstrated the existence of electromagnetic waves in free spaces, through his famous experience. He uses the equivalent of a dipole antenna, to generate and detect a high-frequency power at the receiver end. After, Nikola Tesla in 1899, came up with the idea of wireless transmitting energy. He tried to use the earth itself and the atmosphere layer as conductors to achieve wireless energy transmission. Nonetheless, he was limited by technical conditions at that time. His experiment had failed to achieve the goal [14]. In the late 1950s, Goubau and Schweing theoretically calculate the free space beam guided wave transmission efficiency can reach almost 100 percent and verified it on the reflected beam guided wave system [15]. Followed by these, Raytheon company, William Brown [16] and others as leaders, began to research this technology. In 1964, they finished driving helicopters experiment with 2.45 GHz microwave. In 1968, Peter Glaser imagined the construction of solar power stations in satellites [17]. In the late 1980s, the Canadian communication research center established a service platform for radio, television, communications and other applications, and used wireless methods to supply energy to the platform [14]. Until the 1990s, the wireless transmitting energy technology received widespread attention in close distance application, small and medium-sized power.

2.3 Simultaneous Information and Energy Transmission

The dual capability of RF signals for conveying both information and energy motivates the unified study of wireless information and energy transmission systems and it is called SIET systems [18]. The main interest of SIET systems is for low-power devices with no access for wired energy supplies and interested in not only receiving information but also energy. Existing technologies that already uses this principle are radio frequency identification (RFID) devices,

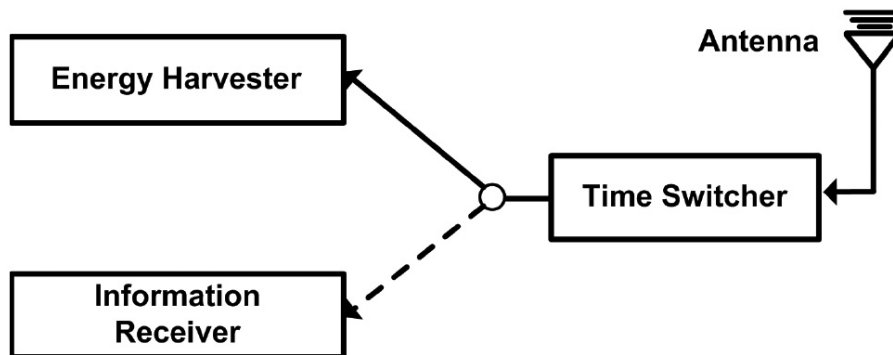


Figure 2.3.: Time switching technique with one transmitter and one receiver. Taken from [3].

power line communications, wireless sensor networks, and wireless body sensor networks. Beyond the existing applications, simultaneous energy and information transmission (SEIT) appears as a promising technology for a variety of emerging applications including low-power short-range communication systems, sensor networks, machine-to-machine networks, and body-area networks, among others [19].

Despite the recent interest in SIET, the practical implementation remains a challenging task, since the EH circuits in practical are not able yet to decode information. Hence, to operate SIET, some practical receivers are proposed such that: time switching, power splitting, and antenna switching.

2.3.1 Time Switching (TS) Architectures TS consists of dividing the transmission block into two orthogonal time slots, one for transferring power and the other for data [20]. In this case, signal splitting is performed in the time domain, and thus the entire signal received in one-time slot is used either for information decoding or power transfer, see Figure 2.3. The TS technique allows for a simple hardware implementation at the receiver but requires accurate time synchronization and information/energy scheduling. Especially, in [21], TS is employed in the broadcast channel with SIET, to characterize the achievable rate region under EH constraint at the receiver.

2.3.2 Power Splitting (PS) Architectures The power splitting (PS) technique consists of splitting the received signal into two streams, where one stream with PS ratio $0 \leq \rho \leq 1$ is used for EH, and the other with power ratio $1 - \rho$ is used for information decoding [20], see Figure 2.4. The PS technique entails higher receiver complexity compared to TS and requires the optimization of the PS factor ρ , however, it achieves instantaneous SIET, as the signal received in one-time slot is used for both information decoding and energy transmission. Therefore, it is more suitable for applications with critical information/energy or delay constraints and closer to the information-theoretical optimum. The PS design has been widely studied in the literature. In [22], a PS design adopted by the receiver, such that a fraction of the received signal is used for channel estimation or data detection, while the rest is used for energy harvesting. In [23], a water-filling algorithm is proposed for an effective PS ratio, to reduce the computational complexity and to achieve the optimal performance bound for SIET. There are some advantages and disadvantages in both these techniques, the power

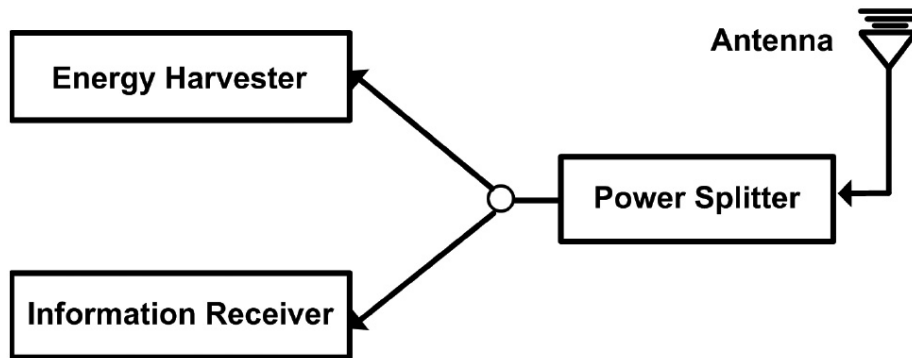


Figure 2.4.: Power splitting technique with one transmitter and one receiver. Taken from [3].

splitting method can achieve instantaneous SIET, therefore it is easier to manage applications with delay constraints and has better performance. But because power splitting requires high receiver complexity, it is more expensive and additional noise can be introduced in the process of power splitting. So, although power splitting technique may be better in terms of performance, time switching is often preferred because they are much cheaper and suitable for mass production, and the relatively simple circuitry involved reduces the chance of mechanical failure and unlike the power splitting circuit, it will introduce as much noise. However, both techniques have been used simultaneously, for coordinating information decoding and energy harvesting for OFDM systems with SIET [24].

2.3.3 Antenna Division Architecture The antenna switching (AS) technique dynamically switches each antenna element between decoding/rectifying to achieve SIET in the antenna domain. In the AS scheme, it assumes that each receiver acts as either an information decoding (ID) receiver or an energy harvesting (EH) receiver [20]. The AS technique requires the solution of an optimization problem in each communication frame in order to decide the optimal assignment of the antenna elements for information decoding and energy harvesting. In [25], a constrained energy efficiency optimization in MIMO system with SIET is proposed for joint antenna selection and spacial switching. Since AS, introduces a high computational complexity, a low-complexity AS techniques are proposed in [26] using the principle of generalized selection combining (GSC).

3

State of the Art

In this section, previous results on simultaneous information and energy transmission in point-to-point channels and multiple access channels (MAC) are presented. In both cases a trade-off between the information rate (bits per channel use) and the energy rate (energy units per channel use) is shown.

3.1 Point-to-Point Channels

In a point-to-point scenario, by choosing a given modulation as in [27] a trade-off between information rate (bits per channel use) and energy rate (energy units per channel use) is observed. For instance, consider a 4-PAM codeword with alphabet $\{-2, -1, 1, 2\}$. If the energy rate constraint is omitted, the transmitter can clearly convey up to 2 bits/ch.use by using all symbols of the constellation. However, imposing a maximum energy rate constraint at the EH receiver, the maximum information rate achieved is 1 bit/ch.use.

3.1.1 Discrete Memoryless Channels Consider a three-party communication system in which a transmitter aims at simultaneously sending information to an information receiver (IR) and energy to an energy harvester (EH) through a noisy communication medium. This system can be fully characterized by a random transformation $(\mathcal{X}, \mathcal{Y} \times \mathcal{Z}, P_{YZ|X})$, where \mathcal{X} is the input alphabet set, $\mathcal{Y} \times \mathcal{Z}$ is the output alphabet set for the information and energy receiver, and $P_{YZ|X}$ is the transition probability kernel. That is, given an input $x \in \mathcal{X}$, the output $(y, z) \in \mathcal{Y} \times \mathcal{Z}$ is observed with probability $P_{YZ|X}(y, z|x)$. A memoryless channel is a random transformation given by

$$(\mathcal{X}^n, \mathcal{Y}^n \times \mathcal{Z}^n, P_{YZ|X}), \quad (3.1)$$

where $n \in \{1, 2, \dots\}$ is the block length and $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n) \in \mathcal{Y}^n$, $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n) \in \mathcal{Z}^n$ and $\mathbf{X} = (X_1, X_2, \dots, X_n) \in \mathcal{X}^n$ are n -dimensional vectors of random variables, such that given an input $\mathbf{x} = (x_1, x_2, \dots, x_n)$, the outputs $\mathbf{y} = (y_1, y_2, \dots, y_n)$ and $\mathbf{z} = (z_1, z_2, \dots, z_n)$ are observed with probability:

$$P_{YZ|X}(\mathbf{y}, \mathbf{z}|\mathbf{x}) = \prod_{t=1}^n P_{YZ|X}(y_t, z_t|x_t). \quad (3.2)$$

Within this context, two tasks are carried out by the transmitter: (a) the information transmission task; and (b) the energy transmission task.

Information Transmission Task From the information transmission standpoint, the goal of the transmitter is to convey an independent message index $W \in \mathcal{W} = \{1, 2, \dots, 2^{NR}\}$ to the receiver using N channel input symbols X_1, X_2, \dots, X_N . The message index W is assumed to be uniformly distributed over \mathcal{W} . Let $f_n^{(N)} : \mathcal{W} \rightarrow \mathcal{X}$ be the encoding function at channel use n , such that for all $n \in \{1, 2, \dots, N\}$, the following holds:

$$X_n = f_n^{(N)}(W). \quad (3.3)$$

The decoder of receiver i observes the channel outputs Y_1, Y_2, \dots, Y_N and uses a decoding function $\phi^{(N)} : \mathbb{N} \times \mathcal{Y}^N \rightarrow \mathcal{W}$, to get an estimate of the message indices:

$$\widehat{W} = \phi^{(N)}(Y_1, Y_2, \dots, Y_N), \quad (3.4)$$

where \widehat{W} is an estimate of the message index W . The decoding error probability of a codebook of block-length N , denoted by $P_{\text{DE}}^{(N)}$, is given by

$$P_{\text{DE}}^{(N)} = \Pr(\widehat{W} \neq W). \quad (3.5)$$

Energy Transmission Task Let $B^{(N)} : \mathbb{R}^N \rightarrow \mathbb{R}_+$ be a function that determines the energy obtained from the channel outputs Z_1, Z_2, \dots, Z_N . In the following, this function is chosen to be the average energy rate (in energy-units per channel use) at the end of N channel uses. That is,

$$B^{(N)}(Z_1, Z_2, \dots, Z_N) \triangleq \frac{1}{N} \sum_{n=1}^N B^{(N)}(Z_n). \quad (3.6)$$

Hence, given an energy rate $B \in [0, B_{\max}]$. Here B_{\max} is the maximum feasible energy rate. The energy shortage probability, denoted by $P_{\text{ES}}^{(N)}(B)$, is:

$$P_{\text{ES}}^{(N)} \triangleq \Pr[B^{(N)}(\mathbf{Z}) < B]. \quad (3.7)$$

Simultaneous Information and Energy Transmission The system is said to operate at the information-energy rate couple $(R, B) \in \mathbb{R}_+^2$ when both transmitter-receiver pairs use an encoding-decoding scheme such that: (i) reliable information transmission at rate R is ensured; and (ii) reliable energy transmission at rate B is ensured. A formal definition is given below.

Definition 6 (Achievable Rates). The couple $(R, B) \in \mathbb{R}_+^2$ is achievable if for all $i \in \{1, 2\}$, there exists a sequence of encoding functions $f_{i,1}^{(N)}, f_{i,2}^{(N)}, \dots, f_{i,N}^{(N)}$ and a decoding function $\phi^{(N)}$ such that both the average decoding error probability $P_{\text{DE}}^{(N)}$ and the energy-shortage probability $P_{\text{ES}}^{(N)}$ tend to zero as the block-length N tends to infinity. That is,

$$\limsup_{N \rightarrow \infty} P_{\text{DE}}^{(N)} = 0 \text{ and} \quad (3.8a)$$

$$\limsup_{N \rightarrow \infty} P_{\text{ES}}^{(N)} = 0. \quad (3.8b)$$

The information and energy transmission task are often a conflicting tasks, this trade-off can be characterized by the notion of information-energy capacity region.

Definition 7 (Information-Energy Capacity Function). *Denote by $b \in [0, B_{\max}]$ the minimum energy rate that must be guaranteed at the input of the EH and denoted by n the blocklength. The information-energy capacity is defined as the following*

$$C^{(N)}(b) \triangleq \max_{\mathbf{X}: B^{(N)} \geq b} I(\mathbf{X}; \mathbf{Y}),$$

where $I(\mathbf{X}, \mathbf{Y})$ is the mutual information between \mathbf{X} and \mathbf{Y} .

Then, the information-energy capacity function is defined as the following:

$$C(b) \triangleq \limsup_{N \rightarrow \infty} \frac{1}{N} C^{(N)}(b). \quad (3.9)$$

Examples In the following, consider a three binary channels, in which the receiver and the EH observe the same output sequence, i.e., $Z_n = Y_n$, for all $n \in \{1, 2, \dots, N\}$. For the noiseless binary channel, using (3.9), the information-energy capacity region for a minimum energy rate b

$$C(b) = \begin{cases} 1, & \text{if } 0 \leq b \leq \frac{1}{2}, \\ H_2(b), & \text{if } \frac{1}{2} \leq b \leq 1, \end{cases}$$

where $H_2(\cdot)$ is the binary entropy function. For $0 \leq b \leq \frac{1}{2}$, the energy transmission does not impact the information transmission task, and the uniform distribution achieves the information capacity. In the second regime when, $\frac{1}{2} \leq b \leq 1$, the optimal distribution is characterized by a Bernoulli distribution with parameter b . Note that a trade-off between both tasks is observed by increasing the energy rate beyond $\frac{1}{2}$.

Consider now the general case, with a binary symmetric channel with cross-over probability α_1 . Thus, the information-energy capacity function with a minimum energy rate $b \in [0, 1]$, follows:

$$C(b) = \begin{cases} 1 - H_2(\alpha_1), & \text{if } 0 \leq b \leq \frac{1}{2}, \\ H_2(b) - H_2(\alpha_1), & \text{if } \frac{1}{2} \leq b \leq 1. \end{cases}$$

The same analysis follows in this case as well. For any $0 \leq b \leq \frac{1}{2}$, the uniform distribution maximizes the information rate and satisfies the energy constraint. However, increasing the energy demands beyond this point will force the transmitter to use the most energetic symbol and it reduces the information rate.

The third example is the Z -channel, with 1-to-0 cross-over probability ϵ , i.e, the binary DMC with $P_{Y|X} = \begin{bmatrix} 1 & 0 \\ \epsilon & 1 - \epsilon \end{bmatrix}$. The information capacity for a minimum energy rate b is

$$C(b) = \begin{cases} C(0), & \text{if } 0 \leq b \leq (1 - \epsilon)\pi^*, \\ H_2(b) - \frac{b}{1 - \epsilon} H_2(\epsilon), & \text{if } (1 - \epsilon)\pi^* \leq b \leq 1 - \epsilon, \end{cases}$$

where $C(0)$ denotes the capacity of this channel given by

$$C(0) = \log_2 \left(1 - \epsilon^{\frac{1}{1-\epsilon}} + \epsilon^{\frac{\epsilon}{1-\epsilon}} \right). \quad (3.10)$$

The capacity achieving input distribution is Bernoulli with parameter

$$\pi^* = \frac{\epsilon^{\frac{\epsilon}{1-\epsilon}}}{1 + (1-\epsilon)\epsilon^{\frac{\epsilon}{1-\epsilon}}}. \quad (3.11)$$

The previous examples show that by increasing the energy rate constraints, the transmitter choose the most energetic symbols at the expense of a lower information rate for the information receiver.

3.1.2 The additive white Gaussian noise (AWGN) channels Consider a three-party communication system, in which the transmitter aims to convey both information and energy through a memoryless Gaussian channel. At each channel use $t \in \{1, \dots, N\}$, the transmitter inputs a symbol X_t . The channel output observed by the information decoder is

$$Y_t = h_1 X_t + Z_t, \quad (3.12)$$

and the EH receiver observes

$$S_t = h_2 X_t + Q_t, \quad (3.13)$$

where h_1 and h_2 are the channel fading gains, which assumed to be constant and non-negative satisfying the \mathcal{L}_2 -norm condition: $\|\mathbf{h}\|^2 \leq 1$, with $\mathbf{h} \triangleq (h_1, h_2)^T$ in order to satisfy the principle of energy conservation. Z_t and Q_t are distributed according to a standard real Gaussian variables. Define $\omega(s) \triangleq s^2$ as the output energy function. the input sequence $\{X_t\}_{t=1}^N$ is subjected to an average *input power constraint* given by

$$\frac{1}{N} \sum_{t=1}^N \mathbb{E}[X_t^2] \leq P, \quad (3.14)$$

with P the average transmit power in energy-units per channel use. The signal to noise ratios (SNR_s) defined as $\text{SNR}_i \triangleq \frac{h_i^2 P}{\sigma_i^2}$, for $i \in \{1, 2\}$.

Note that without considering the energy rate constraint, the information capacity is $\mathcal{C}(0, P) = \frac{1}{2} \log_2(1 + \text{SNR}_1)$ and the maximum energy rate that can be achieved at the input of the EH is $B_{\max} \triangleq \sigma_2^2(1 + \text{SNR}_2)$.

For any $0 \leq b \leq \sigma_2^2(1 + \text{SNR}_2)$, the information capacity function is

$$\mathcal{C}(b, P) = \max_{X: \mathbb{E}[X^2] \leq P \text{ and } \mathbb{E}[S^2] \geq b} I(X; Y). \quad (3.15)$$

Using the same arguments as [28] and [29], it can be shown that

$$\mathcal{C}(b, P) = \frac{1}{2} \log_2(1 + \text{SNR}_1), \quad (3.16)$$

which is inline with expression for the case without energy constraints. Hence, for any feasible energy rate $0 \leq b \leq \sigma_2^2(1 + \text{SNR}_2)$ there is no trade-off between information and energy

transmission. Nevertheless, a trade-off is observed in the Gaussian channel with peak power constraint [29], and it was shown that the optimal input distribution is discrete with finite number of mass point.

3.2 Multiple Access Channels

In the case of multi-user network SIET, to satisfy the minimum energy rate required at the input of the EH, in some cases, the transmitters cooperate to increase the energy transmission. Consider a scenario in which the goal of the transmitter is to simultaneously transmit information to the receiver and energy to an EH. Two cases are presented:

- The requested energy rate is less than what is able to deliver by only transmitting information, in this case, no coordination between the transmitters is required and the optimal strategy of the transmitter is to use the available power to maximize the information transmission.
- The transmitter is required to deliver an energy rate that is higher than what it is able to deliver by only transmitting information. In this case the transmitters might use different policy to ensure a higher energy rate at the EH. For instance, a power splitting approach is used to decompose the transmitted signal into two signal components, one that is designed to transmit information and another that is designed to transmit energy. The former carries both energy and information, however the latter is assumed to be known by all devices and can be build such that the energy received at the EH is maximized.

Many scenarios of SIET in multi-user networks has been studied. An important case to look at is the multi-access channel (MAC) with an EH. From an information theoretic standpoint, the information-energy trade-off was characterized by Fouladgar et al. [30] in the discrete memoryless two-user MAC. Recently, Belhadj Amor et al. studied SIET in the centralized Gaussian MAC (G-MAC) with and without channel-output feedback [4], [31] as well as the decentralized G-MAC [32]. In the following section a review of the results obtained in the centralized case is presented.

3.2.1 System Model

Information Transmission From the information transmission standpoint, the goal of transmitter i , with $i \in \{1, 2\}$, is to convey a message index $W_i \in \mathcal{W}_i = \{1, 2, \dots, \lfloor 2^{NR_i} \rfloor\}$ to receiver i using N channel input symbols $X_{i,1}, X_{i,2}, \dots, X_{i,N}$. That is, information is transmitted at rate $R_i > 0$ bits per channel use. The channel coefficient from transmitter k to receiver i , with $k \in \{1, 2\}$, is denoted by $h_{i,k} \in \mathbb{R}_+$, where \mathbb{R}_+ denotes the positive reals. The channel output at the receiver during channel use n , is denoted by $Y_{1,n}$ and

$$Y_{1,n} = h_{1,1}X_{1,n} + h_{1,2}X_{2,n} + Z_n. \quad (3.17)$$

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 $f_{i,n}^{(N)} : \mathcal{W}_i \times \mathbb{N} \rightarrow \mathbb{R}$ be the encoding function at channel use n , such that for all $n \in \{1, 2, \dots, N\}$, the following holds:

$$X_{i,n} = f_{i,n}^{(N)}(W_i, \Omega). \quad (3.18)$$

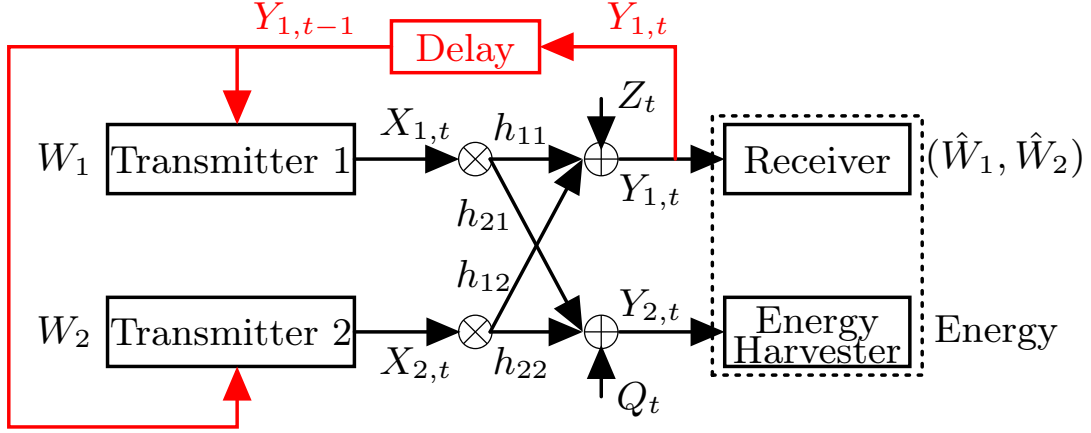


Figure 3.1.: Two-user memoryless Gaussian MAC with feedback and energy harvester. Taken from [4].

In the case with feedback as described in Figure 3.1, the symbol $X_{i,n}$ sent by transmitter i depends upon the indices W_i and Ω , but also upon all previous channel-outputs $Y_{1,1}, Y_{1,2}, \dots, Y_{1,n-d}$, 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. Thus, the first channel input symbol $X_{i,1}$ depends only on the message index W_i and Ω . More specifically, $f_{i,1}^{(N)} : \mathcal{W}_i \times \mathbb{N} \rightarrow \mathbb{R}$. Alternatively, for all $n \in \{2, 3, \dots, N\}$, the encoding functions are $f_{i,n}^{(N)} : \mathcal{W}_i \times \mathbb{N} \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}$. Essentially,

$$X_{i,1} = f_{i,1}^{(N)}(W_i, \Omega), \quad (3.19a)$$

and for all $n > 1$,

$$X_{i,n} = f_{i,n}^{(N)}(W_i, \Omega, Y_{1,1}, \dots, Y_{1,n-1}). \quad (3.20)$$

In both cases, with and without feedback, the random index Ω is assumed to be known by all transmitters and receivers. Moreover, channel input symbols $X_{i,1}, X_{i,2}, \dots, X_{i,N}$ are subject to an average power constraint of the form

$$\frac{1}{N} \sum_{n=1}^N \mathbb{E}_{X_{i,n}} [X_{i,n}^2] \leq P_i, \quad (3.21)$$

where P_i denotes the average transmit power of transmitter i in energy units per channel use. The decoder of receiver i observes the channel outputs $Y_{1,1}, Y_{1,2}, \dots, Y_{1,N}$ and uses a decoding function $\phi^{(N)} : \mathbb{N} \times \mathbb{R}^N \rightarrow \mathcal{W}_i$, to get an estimate of the message indices:

$$\widehat{W}_i = \phi^{(N)}(\Omega, Y_{1,1}, Y_{1,2}, \dots, Y_{1,N}), \quad (3.22)$$

where \widehat{W}_i is an estimate of the message index W_i . The decoding error probability of a codebook

of block-length N , denoted by $P_{\text{DE}}^{(N)}$, is given by

$$P_{\text{DE}}^{(N)} = \max \left[\Pr \left(\widehat{W}_1 \neq W_1 \right) \right]. \quad (3.23)$$

The G-MAC-F and G-MAC with an EH are fully described by the signal to noise ratios (SNRs): SNR_{ji} , with $\forall (i, j) \in \{1, 2\}^2$. These SNRs are defined as follows

$$\text{SNR}_{ji} \triangleq \frac{|h_{ji}|^2 P_i}{\sigma_j^2}. \quad (3.24)$$

Energy Transmission Task From the energy transmission standpoint, the goal of both transmitters is to jointly guarantee an average energy rate at the EH. Let $B^{(N)} : \mathbb{R}^N \rightarrow \mathbb{R}_+$ be a function that determines the energy obtained from the channel outputs $Y_{2,1}, Y_{2,2}, \dots, Y_{2,N}$. In the following, this function is chosen to be the average energy rate (in energy-units per channel use) at the end of N channel uses. That is,

$$B^{(N)}(Y_{2,1}, Y_{2,2}, \dots, Y_{2,N}) \triangleq \frac{1}{N} \sum_{n=1}^N Y_{2,n}^2, \quad (3.25)$$

which implies that the energy carried by a given channel output $Y_{2,t}$, with $t \in \{1, 2, \dots, N\}$, is $Y_{2,t}^2$. Note that the maximum average energy rate, denoted by B_{max} , is:

$$B_{\text{max}} = \sigma_2^2 \left(1 + \text{SNR}_{21} + \text{SNR}_{22} + 2\sqrt{\text{SNR}_{21}\text{SNR}_{22}} \right), \quad (3.26)$$

which can be achieved in the asymptotic block-length regime when both channel inputs exhibit a correlation coefficient equal to one. Hence, given an energy rate $B \in [0, B_{\text{max}}]$, the energy shortage probability, denoted by $P_{\text{ES}}^{(N)}(B)$, is:

$$P_{\text{ES}}^{(N)} \triangleq \Pr \left[B^{(N)}(\mathbf{Y}_2) < B \right]. \quad (3.27)$$

Simultaneous Information and Energy Transmission An information-energy rate triplet $(R_1, R_2, B) \in \mathbb{R}_+^3$ is said to be achievable, when both transmitter-receiver pairs use an encoding and decoding scheme such that: (i) reliable information transmission at rates R_1 and R_2 is satisfied; and (ii) reliable energy transmission at energy rate B is satisfied.

3.2.2 Case Without Feedback The information-energy capacity region of the Gaussian multiple access channel with EH (GMACB) is fully characterized by the following theorem.

Theorem 9 (Information-Energy Capacity Region of the GMACB [4]). *The information-energy capacity region $\mathcal{E}(\sigma_2^2, \text{SNR}_{11}, \text{SNR}_{12}, \text{SNR}_{21}, \text{SNR}_{22})$ of the GMACB is the set of all information-energy rate triplets (R_1, R_2, B) that satisfy*

$$0 \leq R_1 \leq \frac{1}{2} \log_2 (1 + \beta_1 \text{SNR}_{11}), \quad (3.28a)$$

$$0 \leq R_2 \leq \frac{1}{2} \log_2 (1 + \beta_2 \text{SNR}_{12}), \quad (3.28b)$$

$$0 \leq R_1 + R_2 \leq \frac{1}{2} \log_2 (1 + \beta_1 \text{SNR}_{11} + \beta_2 \text{SNR}_{12}), \quad (3.28c)$$

$$0 \leq B \leq \sigma_2^2 \left(1 + \text{SNR}_{21} + \text{SNR}_{22} + 2\sqrt{(1 - \beta_1)\text{SNR}_{21}(1 - \beta_2)\text{SNR}_{22}} \right), \quad (3.28d)$$

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

Comments: Note that β_i , with $i \in \{1, 2\}$, can be interpreted as the power splitting factor dedicated to information transmission and the remaining fraction $(1 - \beta_i)$ is dedicated to energy transmission task. The proof of the achievability of (3.28a)-(3.28c) is obtained by using the same arguments as [33] or [34]. On the other hand, (3.28d) is obtained by using a non-information carrying signals (common randomness) that is known to both transmitters and the receiver. Specifically, the common randomness does not carry any information and does not produce any interference to the information-carrying (IC) signals and used to correlate the transmitted signal. At the receiver, this knowledge is used for cancelling the interference of the no-information carrying component.

Example: Fig. 3.2 shows the intersection of the information energy capacity region $\mathcal{E}(1, 10, 10, 10, 10)$, in the Cartesian coordinates (R_1, R_2, B) , with a plane $B = b$. Three different regimes are observed based on the value of b :

- $b \in \left[0, \sigma_2^2 (1 + \text{SNR}_{21} + \text{SNR}_{22}) \right]$. In this case, the intersection of the plan $B = b$ with $\mathcal{E}(\sigma_2^2, \text{SNR}_{11}, \text{SNR}_{12}, \text{SNR}_{21}, \text{SNR}_{22})$ forms a set, which is identical to the capacity region of G-MAC without an EH. In this case, transmitting information is enough to satisfy the energy constraints. For instance, all the points Q_2, Q_3, Q_4 , and Q_5 are achievable.
- $b \in \left[\sigma_2^2 (1 + \text{SNR}_{21} + \text{SNR}_{22}), \sigma_2^2 (1 + \text{SNR}_{21} + \text{SNR}_{22} + 2\sqrt{\text{SNR}_{21}\text{SNR}_{22}}) \right]$. In this case the energy rate B required at the energy harvester is higher than what is strictly necessary to guarantee reliable communication at the information sum-rate. In this case, a trade-off is accruing between the two tasks: information and energy transmission. For instance, the intersection of the volume $\mathcal{E}(1, 10, 10, 10, 10)$ with the plane $B = b_1 \in (\sigma_2^2 (1 + \text{SNR}_{21} + \text{SNR}_{22}), \sigma_2^2 (1 + \text{SNR}_{21} + \text{SNR}_{22} + 2\sqrt{\text{SNR}_{21}\text{SNR}_{22}})$ in Fig. 3.2 form a subset of the information capacity region of the previous case.
- $b \in \left(\sigma_2^2 (1 + \text{SNR}_{21} + \text{SNR}_{22} + 2\sqrt{(1 - \beta_1)\text{SNR}_{21}(1 - \beta_2)\text{SNR}_{22}}), \infty \right)$. In this case, no positive information rate could be achieved (See for the instance the point Q_1 in Fig. 3.2).

3.2.3 Case With Feedback A complete characterization of the information-energy capacity region for the Gaussian multiple access channel with perfect channel output feedback feedback is given by the following theorem.

Theorem 10 (Information-Energy Capacity Region of the GMACFB [4]). *The information-energy capacity region is characterized by the set of information-energy rate triplets (R_1, R_2, B) that satisfy*

$$0 \leq R_1 \leq \frac{1}{2} \log_2 (1 + \beta_1 \text{SNR}_{11} (1 - \rho^2)), \quad (3.29a)$$

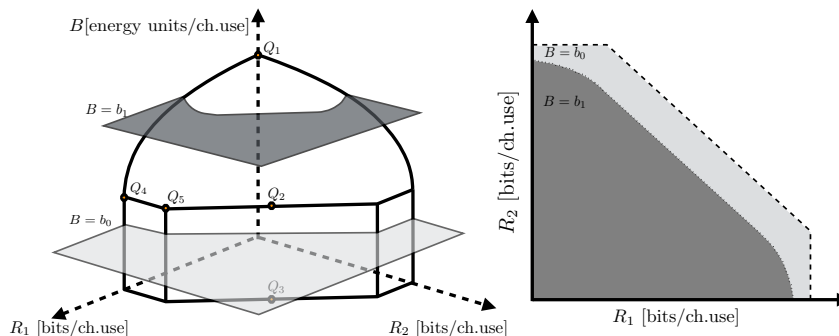


Figure 3.2.: Intersection of the planes $B = b_0$ and $B = b_1$ with the information-energy capacity region of the symmetric G-MAC without feedback. Taken from [4]

$$0 \leq R_2 \leq \frac{1}{2} \log_2 (1 + \beta_2 \text{SNR}_{12} (1 - \rho^2)), \quad (3.29b)$$

$$0 \leq R_1 + R_2 \leq \frac{1}{2} \log_2 (1 + \beta_1 \text{SNR}_{11} + \beta_2 \text{SNR}_{12} + 2\rho \sqrt{\beta_1 \text{SNR}_{11} \beta_2 \text{SNR}_{12}}), \quad (3.29c)$$

$$0 \leq B \leq \sigma_2^2 \left(1 + \text{SNR}_{21} + \text{SNR}_{22} + 2\rho \sqrt{\beta_1 \text{SNR}_{21} \beta_2 \text{SNR}_{22}} + 2\sqrt{(1 - \beta_1) \text{SNR}_{21} (1 - \beta_2) \text{SNR}_{22}} \right), \quad (3.29d)$$

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

Comments on the proof: The proof of the achievable information rates in Theorem 10 follows the same arguments with Ozarow's capacity achieving scheme [35]. The parameters β_1 and β_2 present the power splitting ratio as defined in Theorem 9 above. The key difference with respect to the case without feedback is an additive correlation, i.e., Pearson correlation coefficient between the information-carrying signals (IC) transmitted by both transmitters. The Pearson correlation coefficient between the no-information-carrying (NIC) signals doesn't appear in Theorem 10, since it is unity in order to maximize the transmission of energy. It is worth noting that both IC and NIC contribute to the total harvested energy in (5.9), where the IC signal convey at most $\sigma_2^2 (\beta_1 \text{SNR}_{21} + \beta_2 \text{SNR}_{22} + 2\rho \sqrt{\beta_1 \text{SNR}_{21} \beta_2 \text{SNR}_{22}})$ energy-units per channel use, while the common randomness signals is able to deliver at most $\sigma_2^2 ((1 - \beta_1) \text{SNR}_{21} + (1 - \beta_2) \text{SNR}_{22} + 2\sqrt{(1 - \beta_1) \text{SNR}_{21} (1 - \beta_2) \text{SNR}_{22}})$ energy-units per channel use.

The Role of Feedback In the previous sections, a full characterization of the information energy capacity region is given for the case with and without feedback. From the shape of both region, it can be seen that the feedback enhances the transmission of energy rate. By fixing the information sum-rate as the information sum-capacity without feedback, it can be shown analytically that feedback can at most double the transmission of energy by assuming a symmetric G-MAC. Specifically, B_{NF} and B_{F} denotes the maximum energy rate that can be guaranteed at the EH in the G-MAC without and with feedback respectively. The following theorem characterizes the energy rate enhancement by using a perfect channel feedback

Theorem 11 (Maximum Energy Rate enhancement with Feedback [4]). *The energy rate*

enhancement due to the feedback is given by

$$1 \leq \frac{B_F}{B_{NF}} \leq 2. \quad (3.30)$$

As shown in Theorem 11, feedback can at most double the energy transmission rate in the case of the Gaussian multiple access channel.

4

Gaussian Interference Channel

Consider a two-user Gaussian interference channel without channel output feedback. For this setup, there are two transmitter-receiver pairs and each transmitter i , $i \in \{1, 2\}$, convey a message index $W_i \in \mathcal{W}_i$, with $\mathcal{W}_i = \{1, 2, \dots, 2^{NR_i}\}$, where N denotes the number of channel uses and R_i is the transmission rate in bits per channel use. The channel coefficient from transmitter k to receiver i , with $k \in \{1, 2\}$, is denoted by $h_{i,k} \in \mathbb{R}_+$. 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 $j \neq i \in \{1, 2\}$, and a real additive Gaussian noise $Z_{i,n}$. Hence during channel use n , the channel output at receiver i is given by :

$$Y_{i,n} = h_{i,i}X_{i,n} + h_{i,j}X_{j,n} + Z_{i,n}, \quad (4.1)$$

where $Z_{i,n}$ is a standard real AWGN random variables. Let $f_i : \mathcal{W}_i \rightarrow \mathcal{X}_i$ and $\psi_i : \mathbb{R}_i^N \rightarrow \mathcal{W}_i$ represent respectively the encoding and decoding function and it is assumed deterministic in this section. Moreover, channel input symbols $\{X_{i,1}, \dots, X_{i,n}\}$ are subject to an average power constraint in the form

$$\frac{1}{n} \sum_{t=1}^n \mathbb{E}[X_{i,t}^2] \leq 1. \quad (4.2)$$

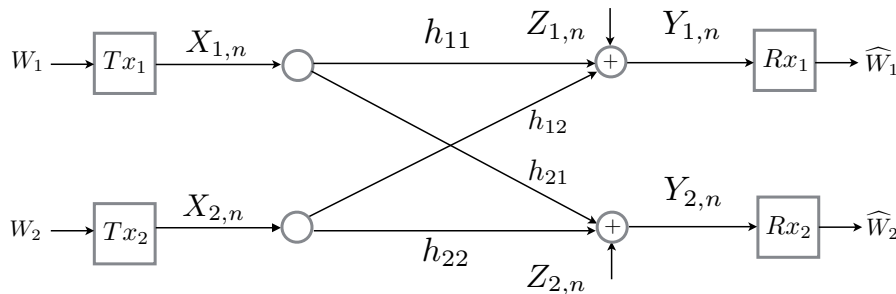


Figure 4.1.: Gaussian interference channel with noisy channel-output at channel use n

4.0.1 Case without Feedback Assessing the capacity region of the two-user Gaussian interference channel (GIC) is an interesting problem in network information theory. The capacity region is perfectly known in the very strong interference regime [36], which is the same capacity region of two non-interfering point to point links. In this case, the interference in both receivers is stronger than the intended signals and therefore the interference can be decoded and subtracted from the received signals to decode the intended signals in each receiver (successive interference cancellation, SIC). The capacity region of the GIC is also known in the case of strong interference regime and it was independently obtained by [37] and [38]. The capacity region of the GIC for the case of strong interference regime in [38] is obtained considering that each receiver must decode both messages. Thus, each transmitter with both receivers can be seen as a multiple access channel (MAC) and the capacity region of the GIC under strong interference can be obtained as the intersection of the capacity regions of the two MACs [39]. This capacity region was initially introduced in [40]. This approach considered the joint decoding instead of sequential decoding as in [36].

The best known achievable region for the two-GIC is given in [37]. This achievable region is simplified in [41]. The strategy in [37] uses rate-splitting [42], which implies that the transmitter possesses two indices to transmit: common index that can be decoded by receiver i and j and private index to be decoded only by the the corresponding receiver. This strategy also implies to arbitrarily split the user signal power into the common and private parts of the message. In reception, this strategy uses joint typical decoding. A more simplified achievable bound was presented in [41], which uses rate-splitting [42] and superposition coding technique for the broadcast channel [43]. The codebook of transmitter i , with $i \in \{1, 2\}$, is generated by two different code layers. The set of sub-code is designed to broadcast the common index denoted by $W_{i,C}$ to both receiver. For each codeword in the second layer, a new sub-codebook is generated. This set of sub-codebooks is referred to as the second layer of the codebook and it is designed for the exclusive transmitting the message index $W_{i,P}$ to receiver i . Finally, transmitter i sends the codeword $\mathbf{X}_i(W_{i,C}, W_{i,P})$ using the superposing technique described above. The proof of the simplified Han-Kobayashi achievable region [41] uses the observation that receiver i is not interested in decoding the common message $W_{i,P}$, with $j \neq i$.

Since the capacity region of the GIC is still known problem in the literature, hence several converse bounds are proposed in [44–48]. In [47], an outer bound of the capacity region of the GIC is characterized for the strong interference regime. An interesting case was presented in [46], where the interference is treated as noise and achieves the sum capacity of the GIC. In addition, an outer bounds that achieves the capacity within one bit per channel was proposed [44], which obtained by using genie-aided models. In the following, two lemmas are presented that shows two outer bounds for the capacity region of the two-user GIC for the weak and mixed interference regime

Lemma 2 (Outer bound for weak GIC [44, Theorem 3]). *The capacity region of the GIC is contained within the set of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy the following inequalities:*

$$R_1 \leq \frac{1}{2} \log(1 + \text{SNR}_1), \quad (4.3a)$$

$$R_2 \leq \frac{1}{2} \log(1 + \text{SNR}_2), \quad (4.3b)$$

$$R_1 + R_2 \leq \frac{1}{2} \log(1 + \text{SNR}_1) + \frac{1}{2} \log\left(1 + \frac{\text{SNR}_2}{1 + \text{INR}_{21}}\right), \quad (4.3c)$$

$$R_1 + R_2 \leq \frac{1}{2} \log(1 + \text{SNR}_2) + \frac{1}{2} \log\left(1 + \frac{\text{SNR}_1}{1 + \text{INR}_{12}}\right), \quad (4.3d)$$

$$R_1 + R_2 \leq \frac{1}{2} \log\left(1 + \text{INR}_{12} + \frac{\text{SNR}_1}{1 + \text{INR}_{21}}\right) + \frac{1}{2} \log\left(1 + \text{INR}_{21} + \frac{\text{SNR}_2}{1 + \text{INR}_{12}}\right), \quad (4.3e)$$

$$2R_1 + R_2 \leq \frac{1}{2} \log(1 + \text{SNR}_1 + \text{INR}_{12}) + \frac{1}{2} \log\left(1 + \text{INR}_{21} + \frac{\text{SNR}_2}{1 + \text{INR}_{12}}\right) + \frac{1}{2} \log\left(\frac{1 + \text{SNR}_1}{1 + \text{INR}_{21}}\right), \quad (4.3f)$$

$$R_1 + 2R_2 \leq \frac{1}{2} \log(1 + \text{SNR}_2 + \text{INR}_{21}) + \frac{1}{2} \log\left(1 + \text{INR}_{12} + \frac{\text{SNR}_1}{1 + \text{INR}_{21}}\right) + \frac{1}{2} \log\left(\frac{1 + \text{SNR}_2}{1 + \text{INR}_{12}}\right). \quad (4.3g)$$

Lemma 3 (Outer bound for mixed GIC [44, Theorem 4]). *Let $\mathcal{C} \subset \mathbb{R}_+^2$ denote the capacity region of the GIC. Then, \mathcal{C} is contained within the set of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy the following inequalities:*

$$R_1 \leq \frac{1}{2} \log(1 + \text{SNR}_1), \quad (4.4a)$$

$$R_2 \leq \frac{1}{2} \log(1 + \text{SNR}_2), \quad (4.4b)$$

$$R_1 + R_2 \leq \frac{1}{2} \log(1 + \text{SNR}_1) + \frac{1}{2} \log\left(1 + \frac{\text{SNR}_2}{1 + \text{INR}_{21}}\right), \quad (4.4c)$$

$$R_1 + R_2 \leq \frac{1}{2} \log(1 + \text{SNR}_1 + \text{INR}_{12}), \quad (4.4d)$$

$$R_1 + 2R_2 \leq \frac{1}{2} \log(1 + \text{SNR}_2 + \text{INR}_{21}) + \frac{1}{2} \log\left(1 + \text{INR}_{12} + \frac{\text{SNR}_1}{1 + \text{INR}_{21}}\right) \quad (4.4e)$$

$$+ \frac{1}{2} \log\left(1 + \frac{\text{SNR}_2}{1 + \text{INR}_{12}}\right). \quad (4.4f)$$

4.0.2 Case with Perfect Channel-Output Feedback An approximation within two bits per channel use for the case with perfect channel-output feedback is characterized in [49]. The proof of the achievable region in [49] is built using the following techniques: block Markov superposition coding [43], rate-splitting [37, 42], and backward decoding [50]. The converse region is obtained by using genie-aided models. For high SNR regime, it has been shown in [49] that feedback provides multiple gain for the symmetric case. The achievable and converse region for two-user GIC-F are presented in the following lemmas:

Lemma 4 (Achievable region two-user GIC-F [49, Theorem 2]). *The capacity region of the two-user GIC-F contains all the rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy the following inequalities:*

$$R_1 \leq \frac{1}{2} \log\left(1 + \text{SNR}_1 + \text{INR}_{12} + 2\rho\sqrt{\text{SNR}_1\text{INR}_{12}}\right) - \frac{1}{2}, \quad (4.5a)$$

$$R_1 \leq \frac{1}{2} \log(1 + (1 - \rho)\text{INR}_{21}) + \frac{1}{2} \log\left(2 + \frac{\text{SNR}_1}{\text{INR}_{21}}\right) - 1, \quad (4.5b)$$

$$\leq \frac{1}{2} \log\left(1 + \text{SNR}_2 + \text{INR}_{21} + 2\rho\sqrt{\text{SNR}_2\text{INR}_{21}}\right) - \frac{1}{2}, \quad (4.5c)$$

$$R_2 \leq \frac{1}{2} \log(1 + (1 - \rho) \text{INR}_{12}) + \frac{1}{2} \log\left(2 + \frac{\text{SNR}_2}{\text{INR}_{12}}\right) - 1, \quad (4.5d)$$

$$R_1 + R_2 \leq \frac{1}{2} \log\left(2 + \frac{\text{SNR}_1}{\text{INR}_{21}}\right) + \frac{1}{2} \log\left(1 + \text{SNR}_2 + \text{INR}_{21} + 2\rho\sqrt{\text{SNR}_2\text{INR}_{21}}\right) - 1, \quad (4.5e)$$

$$R_1 + R_2 \leq \frac{1}{2} \log\left(2 + \frac{\text{SNR}_2}{\text{INR}_{12}}\right) + \frac{1}{2} \log\left(1 + \text{SNR}_1 + \text{INR}_{12} + 2\rho\sqrt{\text{SNR}_1\text{INR}_{12}}\right) - 1, \quad (4.5f)$$

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

Lemma 5 (Converse Region two-user GIC-F, [49, Theorem 3]). *The capacity region of the GIC-F is contained within the set of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy the following inequalities:*

$$R_1 \leq \frac{1}{2} \log\left(1 + \text{SNR}_1 + \text{INR}_{12} + 2\rho\sqrt{\text{SNR}_1\text{INR}_{12}}\right), \quad (4.6a)$$

$$R_1 \leq \frac{1}{2} \log\left(1 + (1 - \rho) \text{INR}_{21}\right) + \frac{1}{2} \log\left(1 + \frac{(1 - \rho^2) \text{SNR}_1}{1 + (1 - \rho^2) \text{INR}_{21}}\right), \quad (4.6b)$$

$$R_2 \leq \frac{1}{2} \log\left(1 + \text{SNR}_2 + \text{INR}_{21} + 2\rho\sqrt{\text{SNR}_2\text{INR}_{21}}\right), \quad (4.6c)$$

$$R_2 \leq \frac{1}{2} \log\left(1 + (1 - \rho) \text{INR}_{12}\right) + \frac{1}{2} \log\left(1 + \frac{(1 - \rho^2) \text{SNR}_2}{1 + (1 - \rho^2) \text{INR}_{12}}\right), \quad (4.6d)$$

$$R_1 + R_2 \leq \frac{1}{2} \log\left(1 + \frac{(1 - \rho^2) \text{SNR}_1}{1 + (1 - \rho^2) \text{INR}_{21}}\right) + \frac{1}{2} \log\left(1 + \text{SNR}_2 + \text{INR}_{21} + 2\rho\sqrt{\text{SNR}_2\text{INR}_{21}}\right), \quad (4.6e)$$

$$R_1 + R_2 \leq \frac{1}{2} \log\left(1 + \frac{(1 - \rho^2) \text{SNR}_2}{1 + (1 - \rho^2) \text{INR}_{12}}\right) + \frac{1}{2} \log\left(1 + \text{SNR}_1 + \text{INR}_{12} + 2\rho\sqrt{\text{SNR}_1\text{INR}_{12}}\right), \quad (4.6f)$$

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

Energy Harvesting in the Gaussian Interference Channel

In this chapter, the fundamental limits of simultaneous information and energy transmission for the two-user Gaussian interference channel (G-IC) with and without feedback are presented. Specifically, an inner and an outer bounds in terms of information and energy transmission rates (in bits per channel use and energy-units per channel use, respectively) are fully characterized. 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.

5.1 Introduction

Battery dependency is a critical issue for modern communication system such as as low-power short-range communication systems, sensor networks, machine-type networks, and body-area networks. Harvesting energy from ambient sources presents a promising solution to power up those devices. Specifically, simultaneous information and energy transmission (SIET) is a technology that exploits the duality of the radio frequency (RF) signals, which can carry both information and energy through appropriate co-design and engineering has gained an upsurge of interest.

The idea of wireless energy transmission traces back to Tesla in the 20-th century [51]. However, only recently the idea of SIET has been formalized. For instance, the point-to-point channel has been studied from the perspective of SIET in the asymptotic block-length regime in [28, 52, 53] and [54]. In the non-asymptotic block-length regime, very little is known as pointed out in [55] and [56]. The Gaussian multiple access channel with one non-located energy harvester (EH) has been studied in [4]. The special case of a colocated EH was studied in [30]. The Gaussian interference channel (G-IC) has been studied in the context of SIET in [57, 58] and [59]. A closely related result is the analysis of energy cooperation at the transmitters [60]. Nonetheless, none of these works tackles the fundamental limits of SIET in the G-IC. Other multi-user channels are studied in [19, 27, 61–64] and references therein.

This chapter focuses on the case of the two-user G-IC under the following assumptions:

(a) A non-colocated energy harvester (EH) recollects energy from the signals sent by the transmitters; (b) a perfect channel-output feedback link might be available from each receiver to the corresponding transmitter. The main results in the case with and without feedback are twofold. First, a set of achievable information and energy transmission rates is presented. That is, a subset of the information-energy capacity region [4] is characterized. Second, a set that contains the information-energy capacity region is introduced. These two sets are shown to be a constant gap approximation to the information-energy capacity region when the energy rate is normalized by the maximum energy rate. Using these results, it is finally shown that perfect channel-output feedback can at most double the energy rate in the G-IC.

This work builds upon existing results whose foci are exclusively on information transmission, in particular [65–67] and [49]. For instance, in the proof of achievability with and without feedback, the novelty lies upon the use of power splitting to superimpose two signal components: One that is designed to transmit information; and another that is designed to transmit energy. Note that the former naturally carries both information and energy, whereas the latter exclusively carries energy. From this perspective, the information-carrying component is built up along the same lines of the achievability scheme presented in [66] and [49] in the case with and without feedback, respectively. The non-information carrying component is built along the lines of the scheme presented in [4] using the idea of common randomness. More specifically, both the transmitters and the receivers know in advance the realization of a given random variable. This realization is used by the transmitters to correlate their transmitted signals in order to benefit the transmission of energy. At the receivers, this knowledge is used for cancelling the interference of the no-information carrying component. The proof of the converse, in the case with and without feedback, follows along the lines of [65] and [49], respectively. The main difference relies on the fact that no constraint is imposed on the first moment of codewords of a given feasible coding scheme. Note for instance that a codeword with a strictly positive mean carries an amount of energy equal to its mean, i.e., direct current (DC) component. Finally, new inner and outer bounds on the energy transmission rate are obtained using Markov’s concentration inequality [68].

5.2 Gaussian Interference Channel with Energy Harvesting

Consider a Gaussian interference channel (G-IC) with a non-colocated energy harvester depicted in Figure 5.1 without feedback and in Figure 5.2 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.

5.2.1 Information Transmission Task From the information transmission standpoint, the goal of transmitter i is to convey an independent message index $W_i \in \mathcal{W}_i = \{1, 2, \dots, 2^{NR_i}\}$ to receiver i using N channel input symbols $X_{i,1}, X_{i,2}, \dots, X_{i,N}$. The message index W_i is assumed to be uniformly distributed over \mathcal{W} . The channel coefficient from transmitter k to receiver i , with $k \in \{1, 2\}$, is denoted by $h_{i,k} \in \mathbb{R}_+$. 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:

$$Y_{i,n} = h_{i,i}X_{i,n} + h_{i,j}X_{j,n} + Z_{i,n}. \quad (5.1)$$

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 $\Omega \in \mathbb{N}$ independent of W_1 and W_2 . Let $f_{i,n}^{(N)} : \mathcal{W}_i \times \mathbb{N} \rightarrow \mathbb{R}$ be the encoding function at channel use n , such that for all $n \in \{1, 2, \dots, N\}$, the following holds:

$$X_{i,n} = f_{i,n}^{(N)}(W_i, \Omega). \quad (5.2)$$

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 $Y_{i,1}, Y_{i,2}, \dots, Y_{i,n-d}$, with $d \in \mathbb{N}$ the feedback delay. In the following, without any loss of generality, it is assumed that d is equal to one channel use. More specifically, $f_{i,1}^{(N)} : \mathcal{W}_i \times \mathbb{N} \rightarrow \mathbb{R}$ and for all $n \in \{2, 3, \dots, N\}$, $f_{i,n}^{(N)} : \mathcal{W}_i \times \mathbb{N} \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ are encoding functions such that:

$$X_{i,1} = f_{i,1}^{(N)}(W_i, \Omega) \quad \text{and} \quad (5.3a)$$

$$X_{i,n} = f_{i,n}^{(N)}(W_i, \Omega, Y_{i,1}, Y_{i,2}, \dots, Y_{i,n-1}). \quad (5.3b)$$

In both cases, with and without feedback, the random index Ω is assumed to be known by all transmitters and receivers. Moreover, channel input symbols $X_{i,1}, X_{i,2}, \dots, X_{i,N}$ are subject to an average power constraint of the form

$$\frac{1}{N} \sum_{n=1}^N \mathbb{E}_{X_{i,n}} [X_{i,n}^2] \leq P_i, \quad (5.4)$$

where P_i denotes the average transmit power of transmitter i in energy units per channel use. The decoder of receiver i observes the channel outputs $Y_{i,1}, Y_{i,2}, \dots, Y_{i,N}$ and uses a decoding function $\phi_i^{(N)} : \mathbb{N} \times \mathbb{R}^N \rightarrow \mathcal{W}_i$, to get an estimate of the message indices:

$$\widehat{W}_i = \phi_i^{(N)}(\Omega, Y_{i,1}, Y_{i,2}, \dots, Y_{i,N}), \quad (5.5)$$

where \widehat{W}_i is an estimate of the message index W_i . The decoding error probability of a codebook of block-length N , denoted by $P_{\text{DE}}^{(N)}$, is given by

$$P_{\text{DE}}^{(N)} = \max \left[\Pr(\widehat{W}_1 \neq W_1), \Pr(\widehat{W}_2 \neq W_2) \right]. \quad (5.6)$$

The signal to noise ratio (SNR) at receiver i is defined as follows

$$\text{SNR}_i \triangleq \frac{|h_{i,i}|^2 P_i}{\sigma_i^2}. \quad (5.7a)$$

The interference to noise ratio (INR) at receiver i is defined as follows

$$\text{INR}_i \triangleq \frac{|h_{i,j}|^2 P_j}{\sigma_i^2}, \quad \text{with } j \neq i. \quad (5.7b)$$

5.2.2 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

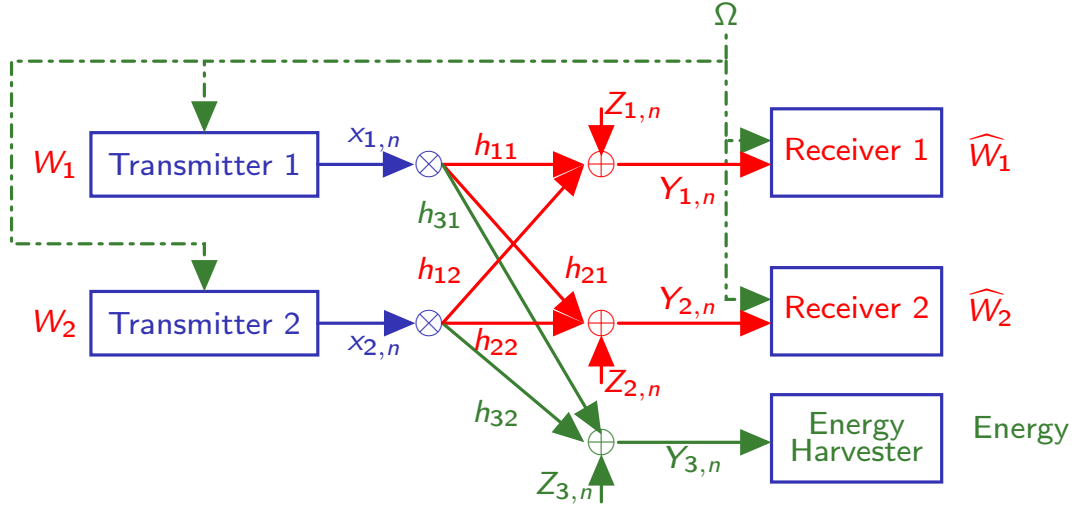


Figure 5.1.: Two-user Gaussian interference channel with a non-colocated energy harvester at channel use n .

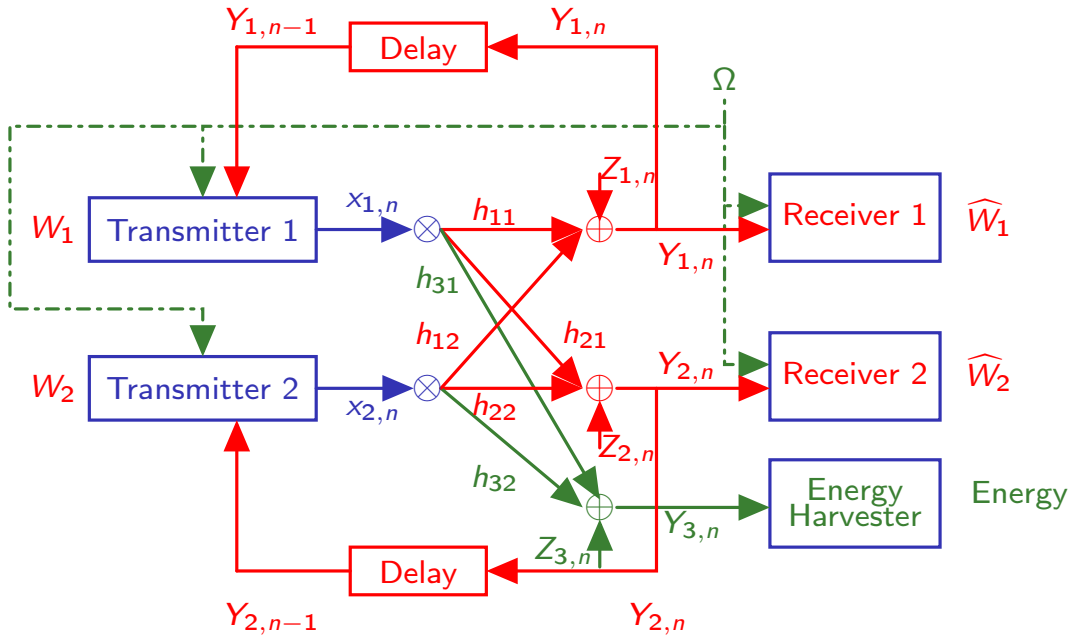


Figure 5.2.: Two-user Gaussian interference channel with channel-output feedback and a non-colocated energy harvester during channel use n .

use n are observed by the EH subject to an 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:

$$Y_{3,n} = h_{3,1}X_{1,n} + h_{3,2}X_{2,n} + Z_{3,n}. \quad (5.8)$$

From the energy transmission standpoint, the goal of both transmitters is to jointly guarantee an average energy rate at the EH. Let $B^{(N)} : \mathbb{R}^N \rightarrow \mathbb{R}_+$ be a function that determines the energy obtained from the channel outputs $Y_{3,1}, Y_{3,2}, \dots, Y_{3,N}$. In the following, this function is chosen to be the average energy rate (in energy-units per channel use) at the end of N channel uses. That is,

$$B^{(N)}(Y_{3,1}, Y_{3,2}, \dots, Y_{3,N}) \triangleq \frac{1}{N} \sum_{n=1}^N Y_{3,n}^2, \quad (5.9)$$

which implies that the energy carried by a given channel output $Y_{3,t}$, with $t \in \{1, 2, \dots, N\}$, is $Y_{3,t}^2$. This assumption is very optimistic, and it is considered for the simplicity of the linear model and it is used in many works in [28, 52, 69–76]. However the nonlinearity is a fundamental propriety of the rectifier given the dependency of the delivered DC power on higher order statistics of the channel input distribution [77–85]. Nonetheless, from the fundamental limits point of view, the assumption in (5.9) has the justification of the assumption of perfect channel output feedback, i.e., any more realistic model would induce fundamental limits that would be more pessimistic than the results presented in the following sections. The SNR of transmitter i at the EH is denoted by

$$\text{SNR}_{3i} = \frac{|h_{3,i}|^2 P_i}{\sigma_3^2}. \quad (5.10)$$

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

$$B_{\max} = \sigma_3^2 \left(1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}} \right). \quad (5.11)$$

which can be achieved in the asymptotic block-length regime when both channel inputs exhibit a correlation coefficient equal to one. Hence, given an energy rate $B \in [0, B_{\max}]$, the energy shortage probability, denoted by $P_{\text{ES}}^{(N)}$, is:

$$P_{\text{ES}}^{(N)} \triangleq \Pr \left[B^{(N)}(\mathbf{Y}_3) < B \right]. \quad (5.12)$$

5.2.3 Simultaneous Information and Energy Transmission An information-energy rate triplet $(R_1, R_2, B) \in \mathbb{R}_+^3$ is achievable when both transmitter-receiver pairs use an encoding-decoding scheme such that: (i) reliable information transmission 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 8 (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 $f_{i,1}^{(N)}, f_{i,2}^{(N)}, \dots, f_{i,N}^{(N)}$ and a decoding functions $\phi_1^{(N)}$ and $\phi_2^{(N)}$ such that both the average decoding error probability $P_{\text{DE}}^{(N)}$ and the energy-shortage probability $P_{\text{ES}}^{(N)}$ tend to zero as the block-length N tends to infinity. That is,*

$$\limsup_{N \rightarrow \infty} P_{\text{DE}}^{(N)} = 0 \text{ and} \quad (5.13a)$$

$$\limsup_{N \rightarrow \infty} P_{\text{ES}}^{(N)} = 0. \quad (5.13b)$$

Using Definition 8, the fundamental limits of simultaneous information and energy transmission in the G-IC can be described by the information-energy capacity region, defined as follows.

Definition 9 (Information-Energy Capacity Region). *La région de capacité information-énergie, désignée par \mathcal{E}^F dans le cas avec rétroaction et \mathcal{E} dans le cas sans rétroaction, correspond à la fermeture de tous les triplets du débit information-énergie réalisables (R_1, R_2, B) .*

5.3 Main Result

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

Definition 10 (Approximation of a Set). *Let $n \in \mathbb{N}$ be fixed. A set $\mathcal{X} \subset \mathbb{R}_+^n$ is approximated by the sets $\underline{\mathcal{X}}$ and $\bar{\mathcal{X}}$ if $\underline{\mathcal{X}} \subseteq \mathcal{X} \subseteq \bar{\mathcal{X}}$ and $\forall \mathbf{x} = (x_1, \dots, x_n) \in \bar{\mathcal{X}}$ then $\left((x_1 - \xi_1)^+, (x_2 - \xi_2)^+, \dots, (x_n - \xi_n)^+ \right) \in \underline{\mathcal{X}}$, for some $(\xi_1, \xi_2, \dots, \xi_n) \in \mathbb{R}_+^n$.*

5.3.1 Case without Channel-Output Feedback The information-energy capacity region \mathcal{E} , is approximated by the regions $\underline{\mathcal{E}}$ (Theorem 12), which represents an information-energy achievable region, and $\bar{\mathcal{E}}$ (Theorem 13), which represents an information-energy converse region.

An Achievable Region The following theorem introduces an achievable information-energy region.

Theorem 12. *The information-energy capacity region \mathcal{E} contains the set $\underline{\mathcal{E}} \subseteq \mathbb{R}_+^3$ of all rate tuples (R_1, R_2, B) that satisfy:*

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{(1 - \lambda_{1e})\text{SNR}_1}{1 + \lambda_{2p}\text{INR}_1} \right), \quad (5.14a)$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{(1 - \lambda_{2e})\text{SNR}_2}{1 + \lambda_{1p}\text{INR}_2} \right), \quad (5.14b)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{1e})\text{SNR}_1 + (1 - \lambda_{2e})\text{INR}_1}{1 + \lambda_{2p}\text{INR}_1} \right) + \frac{1}{2} \log \left(1 + \frac{\lambda_{2p}\text{SNR}_2}{1 + \lambda_{1p}\text{INR}_2} \right), \quad (5.14c)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{2e})\text{SNR}_2 + (1 - \lambda_{1e})\text{INR}_2}{1 + \lambda_{1p}\text{INR}_2} \right) + \frac{1}{2} \log \left(1 + \frac{\lambda_{1p}\text{SNR}_1}{1 + \lambda_{2p}\text{INR}_1} \right), \quad (5.14d)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(\frac{1 + \lambda_{1p}\text{SNR}_1 + (1 - \lambda_{2e})\text{INR}_1}{1 + \lambda_{2p}\text{INR}_1} \right) + \frac{1}{2} \log \left(\frac{1 + \lambda_{2p}\text{SNR}_2 + (1 - \lambda_{1e})\text{INR}_2}{1 + \lambda_{1p}\text{INR}_2} \right), \quad (5.14e)$$

$$2R_1 + R_2 \leq \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{1e})\text{SNR}_1 + (1 - \lambda_{2e})\text{INR}_1}{1 + \lambda_{2p}\text{INR}_1} \right) + \frac{1}{2} \log \left(\frac{1 + \lambda_{2p}\text{SNR}_2 + (1 - \lambda_{1e})\text{INR}_2}{1 + \lambda_{1p}\text{INR}_2} \right) + \frac{1}{2} \log \left(1 + \frac{\lambda_{1p}\text{SNR}_1}{1 + \lambda_{2p}\text{INR}_1} \right), \quad (5.14f)$$

$$R_1 + 2R_2 \leq \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{2e})\text{SNR}_2 + (1 - \lambda_{1e})\text{INR}_2}{1 + \lambda_{1p}\text{INR}_2} \right) + \frac{1}{2} \log \left(\frac{1 + \lambda_{1p}\text{SNR}_1 + (1 - \lambda_{2e})\text{INR}_1}{1 + \lambda_{2p}\text{INR}_1} \right) + \frac{1}{2} \log \left(1 + \frac{\lambda_{2p}\text{SNR}_2}{1 + \lambda_{1p}\text{INR}_2} \right), \quad (5.14g)$$

$$B \leq \sigma_3^2 \left(1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}\sqrt{\lambda_{1e}\lambda_{2e}} \right), \quad (5.14h)$$

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 12 is presented in Appendix A. ■

A Converse Region The following theorem introduces an information-energy converse region.

Theorem 13. *The information-energy capacity region \mathcal{E} contains the set $\bar{\mathcal{E}} \subseteq \mathbb{R}_+^3$ of all rate tuples (R_1, R_2, B) that satisfy:*

$$R_1 \leq \frac{1}{2} \log(1 + \beta_1 \text{SNR}_1), \quad (5.15a)$$

$$R_2 \leq \frac{1}{2} \log(1 + \beta_2 \text{SNR}_2), \quad (5.15b)$$

$$R_1 + R_2 \leq \frac{1}{2} \log(1 + \beta_1 \text{SNR}_1 + \beta_2 \text{INR}_1) + \frac{1}{2} \log\left(1 + \frac{\beta_2 \text{SNR}_2}{1 + \beta_2 \text{INR}_1}\right), \quad (5.15c)$$

$$R_1 + R_2 \leq \frac{1}{2} \log(1 + \beta_2 \text{SNR}_2 + \beta_1 \text{INR}_2) + \frac{1}{2} \log\left(1 + \frac{\beta_1 \text{SNR}_1}{1 + \beta_1 \text{INR}_2}\right), \quad (5.15d)$$

$$R_1 + R_2 \leq \frac{1}{2} \log\left(1 + \frac{\beta_1 \text{SNR}_1 + \beta_2 \text{INR}_1 + \beta_1 \beta_2 \text{INR}_1 \text{INR}_2}{1 + \beta_1 \text{INR}_2}\right) + \frac{1}{2} \log\left(1 + \frac{\beta_2 \text{SNR}_2 + \beta_1 \text{INR}_2 + \beta_1 \beta_2 \text{INR}_1 \text{INR}_2}{1 + \beta_2 \text{INR}_1}\right), \quad (5.15e)$$

$$2R_1 + R_2 \leq \frac{1}{2} \log\left(1 + \frac{\beta_1 \text{SNR}_1}{1 + \beta_1 \text{INR}_2}\right) + \frac{1}{2} \log(1 + \beta_1 \text{SNR}_1 + \beta_2 \text{INR}_1) + \frac{1}{2} \log\left(1 + \frac{\beta_2 \text{SNR}_2 + \beta_1 \text{INR}_2 + \beta_1 \beta_2 \text{INR}_1 \text{INR}_2}{1 + \beta_2 \text{INR}_1}\right), \quad (5.15f)$$

$$R_1 + 2R_2 \leq \frac{1}{2} \log\left(1 + \frac{\beta_2 \text{SNR}_2}{1 + \beta_2 \text{INR}_1}\right) + \frac{1}{2} \log(1 + \beta_2 \text{SNR}_2 + \beta_1 \text{INR}_2) + \frac{1}{2} \log\left(1 + \frac{\beta_1 \text{SNR}_1 + \beta_2 \text{INR}_1 + \beta_1 \beta_2 \text{INR}_1 \text{INR}_2}{1 + \beta_1 \text{INR}_2}\right), \quad (5.15g)$$

$$B \leq \sigma_3^2 \left(1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31} \text{SNR}_{32}} \sqrt{(1 - \beta_1)(1 - \beta_2)}\right), \quad (5.15h)$$

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

Proof: The proof of Theorem 13 is presented in Appendix B. ■

An Approximation to the Information-Energy Capacity Region Using the inner region $\underline{\mathcal{E}}$ and the outer region $\bar{\mathcal{E}}$, described respectively by Theorem 12 and Theorem 13, the information-energy capacity region \mathcal{E} can be approximated in the sense of Definition 10.

Theorem 14 (Approximation of \mathcal{E}). *Let $\underline{\mathcal{E}} \subset \mathbb{R}_+^3$ and $\bar{\mathcal{E}} \subset \mathbb{R}_+^3$ be the sets of tuples (R_1, R_2, B) described by Theorem 12 and Theorem 13, respectively. Then,*

$$\underline{\mathcal{E}} \subset \mathcal{E} \subset \bar{\mathcal{E}}, \quad (5.16)$$

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

Proof: et pour tout $(R_1, R_2, B) \in \bar{\mathcal{E}}$ il s'ensuit que $\left((R_1 - 1/2)^+, (R_2 - 1/2)^+, \left(B - \frac{B_{\max}}{2}\right)^+\right) \in \underline{\mathcal{E}}$

$$\left. \frac{B_{\max}}{2} \right)^+ \in \mathcal{E}.$$

$$\begin{aligned} \frac{B - B'}{B_{\max}} &= \frac{2h_{3,1}h_{3,2}\sqrt{P_1P_2} \left(\sqrt{(1-\beta_1)(1-\beta_2)} - \sqrt{\lambda_{1e}\lambda_{2e}} \right)}{\sigma_3^2 + h_{3,1}^2P_1 + h_{3,2}^2P_2 + 2|h_{3,1}||h_{3,2}|\sqrt{P_1P_2}} \\ &\leq \frac{2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}}{1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}} \\ &\leq \frac{2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}}{1 + 4\sqrt{\text{SNR}_{31}\text{SNR}_{32}}} \\ &\leq \frac{1}{2}. \end{aligned}$$

The proof that $R_1 - R'_1 \leq \frac{1}{2}$ and $R_2 - R'_2 \leq \frac{1}{2}$ is presented in [65]. This completes the proof. \blacksquare

5.3.2 Case with Perfect Channel-Output Feedback The information-energy capacity region \mathcal{E}^F , is approximated by the regions $\underline{\mathcal{E}}^F$ (Theorem 15) and $\underline{\mathcal{E}}^F$ (Theorem 16).

An Achievable Region The following theorem introduces an achievable information-energy region.

Theorem 15. *La région de capacité énergétique d'information \mathcal{E}^F contient l'ensemble $\mathcal{E}^F \subseteq \mathbb{R}_+^3$ de tous les tuples de taux (R_1, R_2, B) qui satisfont :*

$$R_1 \leq \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{1e})\text{SNR}_1 + (1 - \lambda_{2e})\text{INR}_1 + 2\rho\sqrt{\text{SNR}_1\text{INR}_1}}{1 + \lambda_{2p}\text{INR}_1} \right), \quad (5.17a)$$

$$R_1 \leq \frac{1}{2} \log \left(\frac{1 + (1 - (\rho + \lambda_{1e}))\text{INR}_2}{1 + \lambda_{1p}\text{INR}_2} \right) + \frac{1}{2} \log \left(\frac{1 + \lambda_{1p}\text{SNR}_1 + \lambda_{2p}\text{INR}_1}{1 + \lambda_{2p}\text{INR}_1} \right), \quad (5.17b)$$

$$R_2 \leq \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{2e})\text{SNR}_2 + (1 - \lambda_{1e})\text{INR}_2 + 2\rho\sqrt{\text{SNR}_2\text{INR}_2}}{1 + \lambda_{1p}\text{INR}_2} \right), \quad (5.17c)$$

$$R_2 \leq \frac{1}{2} \log \left(\frac{1 + (1 - (\rho + \lambda_{2e}))\text{INR}_1}{1 + \lambda_{2p}\text{INR}_1} \right) + \frac{1}{2} \log \left(\frac{1 + \lambda_{2p}\text{SNR}_1 + \lambda_{1p}\text{INR}_1}{1 + \lambda_{1p}\text{INR}_1} \right) \quad (5.17d)$$

$$\begin{aligned} R_1 + R_2 &\leq \frac{1}{2} \log \left(\frac{1 + \lambda_{1p}\text{SNR}_1 + \lambda_{2p}\text{INR}_1}{1 + \lambda_{2p}\text{INR}_1} \right) \\ &\quad + \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{2e})\text{SNR}_2 + (1 - \lambda_{1e})\text{INR}_2 + 2\rho\sqrt{\text{SNR}_2\text{INR}_2}}{1 + \lambda_{1p}\text{INR}_2} \right), \quad (5.17e) \end{aligned}$$

$$\begin{aligned} R_1 + R_2 &\leq \frac{1}{2} \log \left(\frac{1 + \lambda_{2p}\text{SNR}_1 + \lambda_{1p}\text{INR}_1}{1 + \lambda_{1p}\text{INR}_1} \right) \\ &\quad + \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{1e})\text{SNR}_1 + (1 - \lambda_{2e})\text{INR}_1 + 2\rho\sqrt{\text{SNR}_1\text{INR}_1}}{1 + \lambda_{2p}\text{INR}_1} \right), \quad (5.17f) \end{aligned}$$

$$B \leq \sigma_3^2 \left(1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}(\rho + \sqrt{\lambda_{1e}\lambda_{2e}}) \right), \quad (5.17g)$$

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 15 is presented in Appendix C. ■

A Converse Region The following theorem describes a converse region denoted by $\bar{\mathcal{E}}^F$.

Theorem 16. *The information-energy capacity region \mathcal{E}^F is contained into the set $\bar{\mathcal{E}}^F \in \mathbb{R}_+^3$, which contains all rate tuples (R_1, R_2, B) that satisfy:*

$$R_1 \leq \frac{1}{2} \log \left(1 + \beta_1 \text{SNR}_1 + \beta_2 \text{INR}_1 + 2\rho \sqrt{\beta_1 \text{SNR}_1 \beta_2 \text{INR}_1} \right), \quad (5.18a)$$

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{\beta_1(1-\rho^2)\text{SNR}_1}{1+\beta_1(1-\rho^2)\text{INR}_2} \right) + \frac{1}{2} \log \left(1 + \beta_1(1-\rho^2)\text{INR}_2 \right), \quad (5.18b)$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \beta_2 \text{SNR}_2 + \beta_1 \text{INR}_2 + 2\rho \sqrt{\beta_2 \text{SNR}_2 \beta_1 \text{INR}_2} \right), \quad (5.18c)$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{\beta_2(1-\rho^2)\text{SNR}_2}{1+\beta_2(1-\rho^2)\text{INR}_1} \right) + \frac{1}{2} \log \left(1 + \beta_2(1-\rho^2)\text{INR}_1 \right), \quad (5.18d)$$

$$\begin{aligned} R_1 + R_2 \leq & \frac{1}{2} \log \left(1 + \frac{\beta_1(1-\rho^2)\text{SNR}_1}{1+\beta_1(1-\rho^2)\text{INR}_2} \right) \\ & + \frac{1}{2} \log \left(1 + \beta_2 \text{SNR}_2 + \beta_1 \text{INR}_2 + 2\rho \sqrt{\beta_2 \text{SNR}_2 \beta_1 \text{INR}_2} \right), \end{aligned} \quad (5.18e)$$

$$\begin{aligned} R_1 + R_2 \leq & \frac{1}{2} \log \left(1 + \frac{\beta_2(1-\rho^2)\text{SNR}_2}{1+\beta_2(1-\rho^2)\text{INR}_1} \right) \\ & + \frac{1}{2} \log \left(1 + \beta_1 \text{SNR}_1 + \beta_2 \text{INR}_1 + 2\rho \sqrt{\beta_1 \text{SNR}_1 \beta_2 \text{INR}_1} \right), \end{aligned} \quad (5.18f)$$

$$B \leq \sigma_3^2 \left(1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}(\rho\sqrt{\beta_1\beta_2} + \sqrt{(1-\beta_1)(1-\beta_2)}) \right), \quad (5.18g)$$

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

Proof: The proof of Theorem 16 is presented in Appendix D. ■

An Approximation to the Information-Energy Capacity Region Using the inner region $\underline{\mathcal{E}}^F$ and the outer region $\bar{\mathcal{E}}^F$, described respectively by Theorem 15 and Theorem 16, the information-energy capacity region \mathcal{E}^F can be approximated in the sense of Definition 10.

Theorem 17 (Approximation of \mathcal{E}^F). *Let $\underline{\mathcal{E}}^F \subset \mathbb{R}_+^3$ and $\bar{\mathcal{E}}^F \subset \mathbb{R}_+^3$ be the sets of tuples (R_1, R_2, B) described by Theorem 15 and Theorem 16, respectively. Then,*

$$\underline{\mathcal{E}}^F \subset \mathcal{E}^F \subset \bar{\mathcal{E}}^F, \quad (5.19)$$

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

Proof: Note that for all $(R_1, R_2, B) \in \bar{\mathcal{E}}^F$ and for all $(R'_1, R'_2, B') \in \underline{\mathcal{E}}^F$, there always exists

a tuple $(\beta_1, \beta_2, \lambda_{1e}, \lambda_{2e}, \rho)$ such that:

$$\begin{aligned}
 \frac{B - B'}{B_{\max}} &\leq \frac{2h_{3,1}h_{3,2}\sqrt{P_1P_2} \left(\sqrt{(1-\beta_1)(1-\beta_2)} + \rho\sqrt{\beta_1\beta_2} - \rho' - \sqrt{\lambda_{1e}\lambda_{2e}} \right)}{\sigma_3^2 + h_{3,1}^2P_1 + h_{3,2}^2P_2 + 2h_{3,1}h_{3,2}\sqrt{P_1P_2}} \\
 &\leq \frac{2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}}{1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}} \\
 &\leq \frac{2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}}{1 + 4\sqrt{\text{SNR}_{31}\text{SNR}_{32}}} \\
 &\leq \frac{1}{2}.
 \end{aligned}$$

The proof of the inequalities $R_1 - R'_1 \leq 1$ and $R_2 - R'_2 \leq 1$ is presented in [49]. This completes the proof. \blacksquare

5.3.3 Maximum Energy Rate Improvement with Feedback Consider the following sets of energy rates: $\underline{\mathcal{B}} = \{b \in \mathbb{R}_+ : (R_1, R_2, b) \in \underline{\mathcal{E}}\}$, $\overline{\mathcal{B}} = \{b \in \mathbb{R}_+ : (R_1, R_2, b) \in \overline{\mathcal{E}}\}$, $\underline{\mathcal{B}}_F = \{b \in \mathbb{R}_+ : (R_1, R_2, b) \in \underline{\mathcal{E}}^F\}$, and $\overline{\mathcal{B}}_F = \{b \in \mathbb{R}_+ : (R_1, R_2, b) \in \overline{\mathcal{E}}^F\}$. The maximum improvement that can be achieved on the energy rate due to feedback can be shown to be at most a factor of two. The following proposition shows this by providing upper bounds on the ratios $\frac{\max \underline{\mathcal{B}}_F}{\max \underline{\mathcal{B}}}$ and $\frac{\max \overline{\mathcal{B}}_F}{\max \overline{\mathcal{B}}}$.

Proposition 3 (Rate improvement with Feedback). The energy rate achievable in the two-user G-IC with perfect channel-output feedback can be twice the energy rate achievable in the two-user G-IC without feedback. That is,

$$1 < \frac{\max \underline{\mathcal{B}}_F}{\max \underline{\mathcal{B}}} \leq 2. \quad (5.20)$$

Any improvement beyond a factor of two is not feasible. That is,

$$1 < \frac{\max \overline{\mathcal{B}}_F}{\max \overline{\mathcal{B}}} \leq 2. \quad (5.21)$$

Before presenting the proof of Proposition 3, the following auxiliary result is presented

Proposition 4. For all $(x, y) \in [0, 1]^2$, it holds that

$$\sqrt{xy} + \sqrt{(1-x)(1-y)} \leq 1. \quad (5.22)$$

Proof: Given that $(\sqrt{x} - \sqrt{y})^2 = x + y - 2\sqrt{xy} \geq 0$, it holds that

$$\sqrt{xy} \leq \frac{x + y}{2}. \quad (5.23)$$

Using (5.23), it holds that

$$\sqrt{xy} \leq \frac{x + y}{2}, \text{ and} \quad (5.24)$$

$$\sqrt{(1-x)(1-y)} \leq \frac{2 - (x + y)}{2}. \quad (5.25)$$

Adding (5.24) and (5.25) yields the inequality in (5.22). This completes the proof. \blacksquare

Proof of Proposition 3: The proof is divided into two parts. The first part presents the proof of (5.20). The second part presents the proof of (5.21)

1. Proof of (5.20): Let $\max \underline{\mathcal{B}}_F$ be written in terms of $(\rho, \lambda_{1p}, \lambda_{2p}, \lambda_{1e}, \lambda_{2e})$ and let also $\max \underline{\mathcal{B}}$ be written in terms of $(\rho', \lambda'_{1p}, \lambda'_{2p}, \lambda'_{1e}, \lambda'_{2e})$ such that:

$$\frac{\sup \underline{\mathcal{B}}_F}{\sup \underline{\mathcal{B}}} = \frac{1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}(\rho + \sqrt{\lambda_{1e}\lambda_{2e}})}{1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}\lambda'_{1e}\lambda'_{2e}} \quad (5.26)$$

$$\leq \frac{1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}(\rho + \max(\lambda_{1e}, \lambda_{2e}))}{1 + \text{SNR}_{31} + \text{SNR}_{32}} \quad (5.27)$$

$$\leq 1 + \frac{2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}}{1 + \text{SNR}_{31} + \text{SNR}_{32}} \quad (5.28)$$

$$\leq 1 + \frac{\text{SNR}_{31} + \text{SNR}_{32}}{1 + \text{SNR}_{31} + \text{SNR}_{32}} \quad (5.29)$$

$$\leq 2. \quad (5.30)$$

2. Proof of (5.21) : Let $\max \bar{\mathcal{B}}_F$ be written in terms of $(\rho, \beta_1, \beta_2) \in [0, 1]^3$ and let also $\max \bar{\mathcal{B}}$ be written in terms of $(\beta'_1, \beta'_2) \in [0, 1]^2$ such that:

$$\frac{\sup \bar{\mathcal{B}}_F}{\sup \bar{\mathcal{B}}} = \frac{1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}(\rho\sqrt{\beta_1\beta_2} + \sqrt{(1-\beta_1)(1-\beta_2)})}{1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}(1-\beta'_1)(1-\beta'_2)} \quad (5.31)$$

$$\leq 1 + \frac{2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}}{1 + \text{SNR}_{31} + \text{SNR}_{32} + 2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}(1-\beta_1)(1-\beta_2)} \quad (5.32)$$

$$\leq 1 + \frac{2\sqrt{\text{SNR}_{31}\text{SNR}_{32}}}{1 + \text{SNR}_{31} + \text{SNR}_{32}} \quad (5.33)$$

$$\leq 1 + \frac{\text{SNR}_{31} + \text{SNR}_{32}}{1 + \text{SNR}_{31} + \text{SNR}_{32}} \quad (5.34)$$

$$\leq 2, \quad (5.35)$$

where (5.32) follows from Proposition 4. This completes the proof. \blacksquare

5.4 Numerical Analysis

Consider the two-user G-ICs with and without channel-output feedback depicted in Figure 5.1 and Figure 5.2 with parameters $\text{SNR}_1 = \text{SNR}_2 = 20$ dB, $\text{INR}_1 = \text{INR}_2 = \text{SNR}_{31} = \text{SNR}_{32} = 10$ dB, and $\sigma_3^2 = 1$. The corresponding achievable region $\underline{\mathcal{E}}$ and converse region $\bar{\mathcal{E}}$ are shown in Figure 5.3. In the case with feedback, the corresponding achievable region $\underline{\mathcal{E}}^F$ and converse region $\bar{\mathcal{E}}^F$ are shown in Figure 5.4. Note the strict inclusions $\underline{\mathcal{E}} \subset \bar{\mathcal{E}}$ and $\underline{\mathcal{E}}^F \subset \bar{\mathcal{E}}^F$ (Definition 10). Note also that for all $B \leq 21$ energy units, the set of triplets $\underline{\mathcal{E}}^F$ and the set of triplets $\bar{\mathcal{E}}^F$ are prisms whose bases correspond to the inner and outer regions approximating the information capacity region presented in [37] and [65] respectively. For all $B > 21$, the trade-off between information transmission rates and the energy transmission rate becomes evident as both regions $\underline{\mathcal{E}}^F$ and $\bar{\mathcal{E}}^F$ monotonically shrink when B increases (Remark 3). The same observation can be made for the case without feedback.

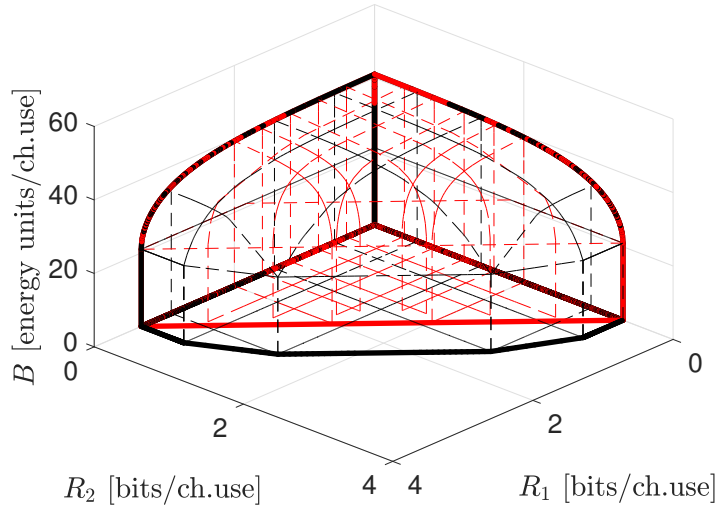


Figure 5.3.: Superposition of $\underline{\mathcal{E}}$ and $\bar{\mathcal{E}}$, with parameters $\text{SNR}_1 = \text{SNR}_2 = 20\text{dB}$, $\text{INR}_1 = \text{INR}_2 = \text{SNR}_{31} = \text{SNR}_{32} = 10\text{dB}$, $\sigma_3^2 = 1$.

Figure 5.5 shows the pairs (R_2, B) that are in the sets $\{(R_2, B) \in \mathbb{R}_+^2 : (r_1, R_2, B) \in \bar{\mathcal{E}}\}$ (solid line) and $\{(R_2, B) \in \mathbb{R}_+^2 : (r_1, R_2, B) \in \bar{\mathcal{E}}^F\}$ (dashed line), with $r_1 = 0$ and $r_1 = 3$. Note that thanks to feedback, the information rate R_2 can be increased one bit per channel use while keeping both the information rate R_1 and the energy rate B invariant.

Figure 5.6 shows the set of pairs (R_1, R_2) that are in the sets $\{(R_1, R_2) \in \mathbb{R}_+^2 : (R_1, R_2, b) \in \bar{\mathcal{E}}\}$ (solid line) and $\{(R_1, R_2) \in \mathbb{R}_+^2 : (R_1, R_2, b) \in \bar{\mathcal{E}}^F\}$ (dashed line), with $b = 21$ and $b = 35$. Note that thanks to feedback, both the information rates R_1 and R_2 can be increased more than half a bit per channel use while keeping the energy rate B constant.

Figure 5.7 shows the ratio $\frac{\max \bar{B}_F}{\max B}$ for different ratios of SNR at the EH, i.e., $\text{SNR}_{31} = \text{SNR}_{32} = \text{SNR}$; $\frac{\text{SNR}_{31}}{2} = \text{SNR}_{32} = \text{SNR}$; and $\frac{\text{SNR}_{31}}{10} = \text{SNR}_{32} = \text{SNR}$, respectively. Note that the upper bound in Proposition 3 is tight in the case in the symmetric case.

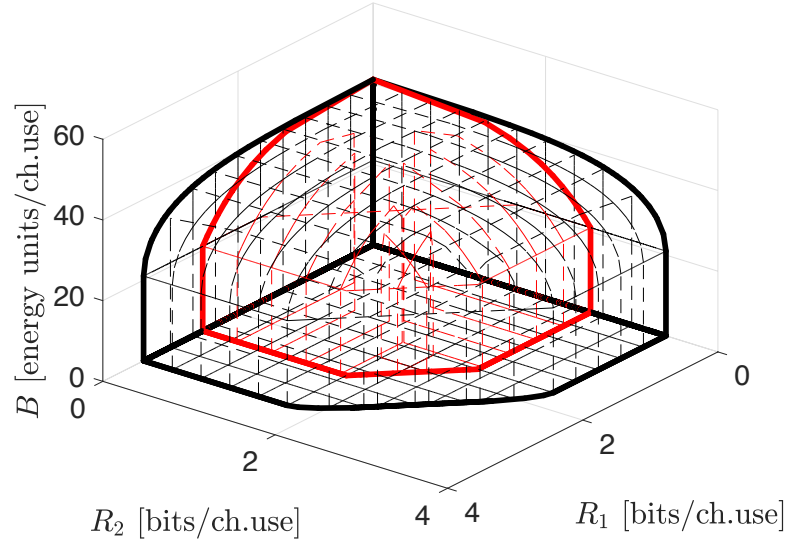


Figure 5.4.: 3-D superposition of $\underline{\mathcal{E}}^F$ and $\overline{\mathcal{E}}^F$, with parameters $\text{SNR}_1 = \text{SNR}_2 = 20$ dB, $\text{INR}_1 = \text{INR}_2 = \text{SNR}_{31} = \text{SNR}_{32} = 10$ dB and $\sigma_3^2 = 1$.

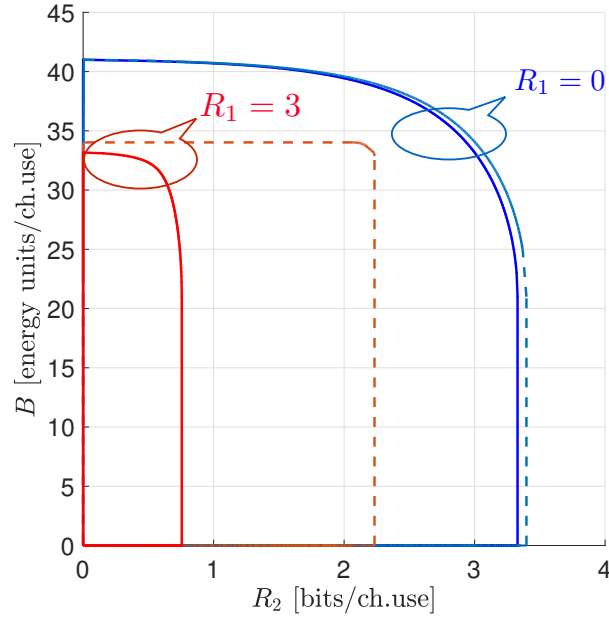


Figure 5.5.: Convex hull of pairs (R_2, B) that are in the sets $\{(R_2, B) \in \mathbb{R}_+^2 : (r_1, R_2, B) \in \overline{\mathcal{E}}\}$ (solid line) and $\{(R_2, B) \in \mathbb{R}_+^2 : (r_1, R_2, B) \in \underline{\mathcal{E}}^F\}$ (dashed line), with $r_1 \in \{0, 3\}$. Parameters $\text{SNR}_1 = \text{SNR}_2 = 20$ dB, $\text{INR}_1 = \text{INR}_2 = \text{SNR}_{31} = \text{SNR}_{32} = 10$ dB, and $\sigma_3^2 = 1$.

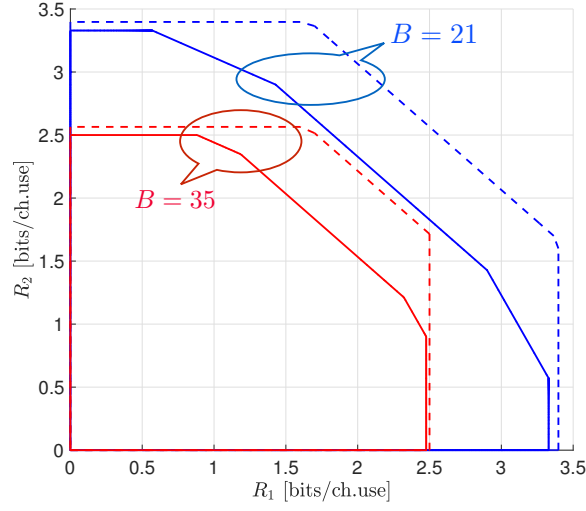


Figure 5.6.: The ratio $\frac{\sup \bar{\mathcal{B}}_F}{\sup \bar{\mathcal{B}}}$ for different ratios of SNR at the EH, i.e., $\text{SNR}_{31} = \text{SNR}_{32} = \text{SNR}$; $\frac{\text{SNR}_{31}}{2} = \text{SNR}_{32} = \text{SNR}$; and $\frac{\text{SNR}_{31}}{10} = \text{SNR}_{32} = \text{SNR}$, respectively. Parameters $\text{SNR}_1 = \text{SNR}_2 = 20$ dB, $\text{INR}_1 = \text{INR}_2 = \text{SNR}_{31} = \text{SNR}_{32} = 10$ dB, and $\sigma_3^2 = 1$.

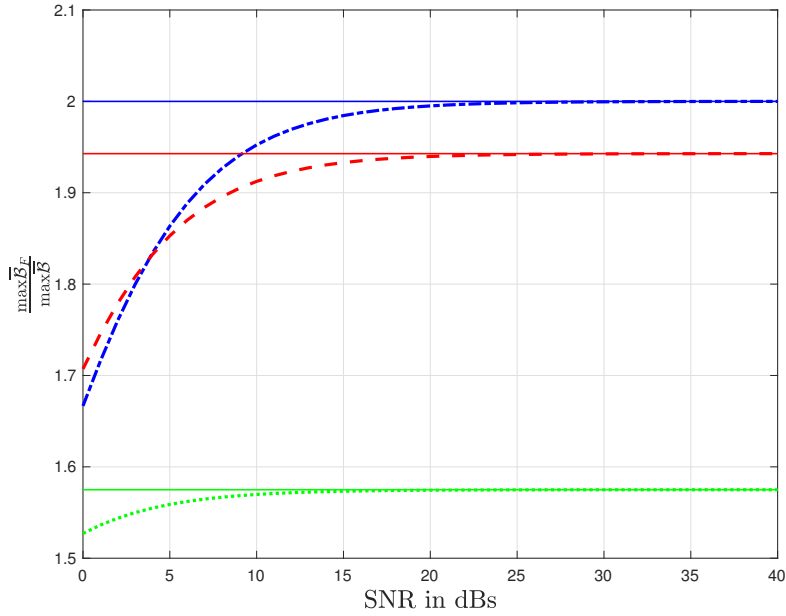


Figure 5.7.: The ratio $\frac{\max \bar{\mathcal{B}}_F}{\max \bar{\mathcal{B}}}$ for different ratios of SNR at the EH, i.e., $\text{SNR}_{31} = \text{SNR}_{32} = \text{SNR}$; $\frac{\text{SNR}_{31}}{2} = \text{SNR}_{32} = \text{SNR}$; and $\frac{\text{SNR}_{31}}{10} = \text{SNR}_{32} = \text{SNR}$, respectively. Parameters $\text{SNR}_1 = \text{SNR}_2 = 20$ dB, $\text{INR}_1 = \text{INR}_2 = \text{SNR}_{31} = \text{SNR}_{32} = 10$ dB, and $\sigma_3^2 = 1$.

6

Energy Harvesting with Reliability and Latency Constraints

In this chapter, the fundamental limits of simultaneous information and energy transmission (SIET) are fully characterized, when the system operates at the non-asymptotic block-length regime. The focus is on the case of a transmitter simultaneously sending information to a receiver and energy to an energy harvester through the binary symmetric channel. Given a finite number of channel uses (latency constraint) as well as tolerable average decoding error probability and energy shortage probability (reliability constraints), two sets of information and energy transmission rates are presented. One consists in rate pairs for which the existence of at least one code achieving such rates under the latency and reliability constraints is proved (achievable region). The second one consists in a set whose complement contains the rate pairs for which there does not exist a code capable of achieving such rates (converse region). These two sets approximate the information-energy capacity region, which allows analyzing the trade-offs among performance, latency, and reliability in SIET systems.

6.1 Introduction

Simultaneous information and energy transmission (SIET) refers to systems in which at least one transmitter aims to simultaneously send information to a set of information receivers (IRs) and energy to a set of energy harvesters (EHs). This idea traces back to Nikola Tesla, who proposed SIET in 1914 [14]. In modern communications systems, SIET is one of the central ideas for wirelessly powering up devices with low-energy consumption [86].

The fundamental limits of SIET are characterized by the *information-energy capacity region* [61]. This region consists in the set of all information and energy transmission rates that can be simultaneously achieved. In general, it can be characterized in two different regimes: (i) the asymptotic block-length regime; and (ii) the non-asymptotic block-length regime. The former refers to a case in which the block length is assumed to be infinitely long, while the decoding error probability (DEP) and the energy-shortage probability (ESP) are assumed to be arbitrarily close to zero. From this perspective, the asymptotic block-length regime does not capture the constraints on latency. Essentially, these limits apply only to the scenarios in which the duration of the transmission is arbitrarily long. The non-asymptotic regime, on the other hand, refers to the case in which the block length is assumed to be finite and both the

DEP and the ESP are bounded away from zero. In this case, the information-energy capacity region is parametrized by a finite block length, an upper bound on the DEP, and an upper bound on the ESP. This allows taking into account the constraints on latency in terms of channel uses, and reliability in terms of DEP and ESP.

The information energy capacity region in the asymptotic regime was characterized for point-to-point memoryless channels in [28, 61], and [52]. Alternatively, in multi-user channels, characterizations of the information-energy capacity region of multiple access channels were presented in [30] and [4]. A characterization of this region in the context of the interference channel was presented in [87]. In the non-asymptotic regime, however, the information-energy capacity region in point-to-point channels is not well-investigated. A first attempt to characterize it was made in [88], building upon the existing results on the fundamental limits on information transmission in the non-asymptotic block-length regime in [89] and [90]. In multi-user channels, a characterization of the information-energy capacity region is unknown.

The focus of this research is on a system in which a transmitter simultaneously sends information to an information receiver and energy to an energy harvester through binary symmetric channels. The main contribution is characterizing the information-energy capacity region. This characterization is achieved by providing a set that is confined by the information-energy capacity region and another set that contains it. The inner set contains the information and energy transmission rates for which there always exists at least one code achieving such rates (achievable region). The outer set is a set whose complement contains the information and energy transmission rates that cannot be achieved by any code (converse region).

This chapter is organized as follows. The system model and the notion of the information-energy capacity region in the non-asymptotic block-length regime are presented in Section 6.2. Section 6.3 presents the achievable and converse region for the binary symmetric channel in the non-asymptotic regime.

6.2 System Model

Consider a three part communication system, where a transmitter aims simultaneously convey information to an information receiver (IR) and energy to an EH receiver through a binary symmetric channel. For the non-asymptotic case, the random transformation for such system is given by

$$(\{0, 1\}^n, \{0, 1\}^n \times \{0, 1\}^n, P_{\mathbf{Y}\mathbf{Z}|\mathbf{X}}), \quad (6.1)$$

where $n \in \mathbb{N}$ presents the block length. Given an input $\mathbf{x} \triangleq (x_1, x_2, \dots, x_n) \in \{0, 1\}^n$, the outputs $\mathbf{y} \triangleq (y_1, y_2, \dots, y_n) \in \{0, 1\}^n$ and $\mathbf{z} \triangleq (z_1, z_2, \dots, z_n) \in \{0, 1\}^n$ are observed at the IR and at the EH, respectively, with probability

$$P_{\mathbf{Y}\mathbf{Z}|\mathbf{X}}(\mathbf{y}, \mathbf{z}|\mathbf{x}) = \prod_{t=1}^n P_{Y|X}(y_t|x_t)P_{Z|X}(z_t|x_t), \quad (6.2)$$

where for all $(x, y, z) \in \{0, 1\}^3$,

$$P_{Y|X}(y|x) = \alpha_1 \mathbb{1}_{\{x \neq y\}} + (1 - \alpha_1) \mathbb{1}_{\{x = y\}}, \quad (6.3)$$

$$P_{Z|X}(z|x) = \alpha_2 \mathbb{1}_{\{x \neq z\}} + (1 - \alpha_2) \mathbb{1}_{\{x = z\}}, \quad (6.4)$$

and $\alpha_1 \in [0, \frac{1}{2})$ and $\alpha_2 \in (0, \frac{1}{2})$. In this context, two tasks are carried out by the transmitter: (a) the information transmission task; and (b) the energy transmission task.

6.2.1 Information Transmission Task The purpose of this task is to send a message from the transmitter to the IR. The message index is a realization of a random variable uniformly distributed in $\{1, 2, \dots, M\}$, with $M \in \mathbb{N}$. To carry out this task within n channel uses, the transmitter uses an (n, M) -code.

Definition 11 ((n, M) -code). An (n, M) -code for the random transformation in (6.1) is a system

$$\{(\mathbf{u}(1), \mathcal{D}_1), (\mathbf{u}(2), \mathcal{D}_2), \dots, (\mathbf{u}(M), \mathcal{D}_M)\}, \quad (6.5)$$

where for all $(i, j) \in \{1, 2, \dots, M\}^2$, with $i \neq j$,

$$\mathbf{u}(i) \triangleq (u_1(i), u_2(i), \dots, u_n(i)) \in \{0, 1\}^n, \quad (6.6a)$$

$$\mathcal{D}_i \cap \mathcal{D}_j = \emptyset, \text{ and} \quad (6.6b)$$

$$\bigcup_{i=1}^M \mathcal{D}_i \subseteq \{0, 1\}^n. \quad (6.6c)$$

Given the system in (6.5), for all $i \in \{1, 2, \dots, M\}$, the transmitter encodes the message index i , into a codeword $\mathbf{u}(i) = (u_1(i), \dots, u_n(i))$. At the end of n channel uses, the IR observes the output $\mathbf{y} = (y_1, \dots, y_n)$ and decides that the symbol i was transmitted if it satisfies the rule

$$(y_1, y_2, \dots, y_n) \in \mathcal{D}_i. \quad (6.7)$$

Denote by $\lambda_i \in [0, 1]$ by the decoding error probability associated with the transmission of message index i , which is given by

$$\lambda_i \triangleq \Pr[\mathbf{Y} \in \mathcal{D}_i^c \mid \mathbf{X} = \mathbf{u}(i)], \quad (6.8)$$

where the probability is taken with respect to the marginal $P_{\mathbf{Y}|\mathbf{X}}$, and \mathcal{D}_i^c represents the complement set of \mathcal{D}_i with respect to $\{0, 1\}^n$. The average probability of error, denoted by λ , is

$$\lambda \triangleq \frac{1}{M} \sum_{m=1}^M \lambda_m. \quad (6.9)$$

Information transmission is said to be reliable if the average or maximum DEP is controlled. This leads to the following refinements of Definition 11.

Definition 12 ((n, M, ϵ) -code with maximum DEP). An (n, M) -code is said to be an (n, M, ϵ) -code with maximum DEP if $\lambda_i < \epsilon$, for all $i \in \{1, 2, \dots, M\}$, for a given $\epsilon \in [0, 1]$.

Definition 13 ((n, M, ϵ) -code with average DEP). An (n, M) -code is said to be an (n, M, ϵ) -code with average DEP if $\lambda < \epsilon$, for all $i \in \{1, 2, \dots, M\}$, for a given $\epsilon \in [0, 1]$.

Note that any (n, M, ϵ) -code with maximum DEP is also a (n, M, ϵ) -code with average DEP. Nonetheless, the converse is not necessarily true.

6.2.2 Energy Transmission Task Let $g : \{0, 1\} \rightarrow \mathbb{R}_+$ be a positive real-valued function that determines the energy harvested from the channel output symbols. Let

$$b_0 \triangleq g(0), \text{ and} \quad (6.10a)$$

$$b_1 \triangleq g(1) \tag{6.10b}$$

be the energy harvested when the channel outputs at the EH are 0 and 1, respectively. At the end of n channel uses, the average energy delivered to the EH by the channel outputs $\mathbf{z} = (z_1, z_2, \dots, z_n)$ is given by the function $B_n : \{0, 1\}^n \rightarrow \mathbb{R}_+$, with

$$B_n(\mathbf{z}) \triangleq \frac{1}{n} \sum_{t=1}^n g(z_t) = (b_0 - b_1) \frac{N(0|\mathbf{z})}{n} + b_1. \tag{6.11}$$

A reliable transmission of energy occurs when $B_n(\mathbf{z}) \geq b$, with $b \geq 0$. However, an energy-shortage event occurs when $B_n(\mathbf{z}) < b$ at the end of the transmission. For the rest of the section, the following assumption is adopted without loss of generality

$$b_1 < b_0. \tag{6.12}$$

Since for the case where $b_0 = b_1$ is trivial, it holds that $B_n(\mathbf{z}) = b_0 = b_1$. That is, the average energy rate at the input of the EH is independent of the codebook, and either an energy shortage is never observed if $b \geq b_0 = b_1$; or the the system is always under energy shortage if $b < b_0 = b_1$. Denote by θ_i the probability of energy-shortage when transmitting the message with index $i \in \{1, 2, \dots, M\}$, which defined as follows

$$\theta_i \triangleq \Pr [B_n(\mathbf{Z}) < b \mid \mathbf{X} = \mathbf{u}(i)], \tag{6.13}$$

where the probability is taken with respect to the marginal $P_{\mathbf{Z}|\mathbf{X}}$. Denote by θ , the average probability of energy-shortage, which is given by

$$\theta \triangleq \frac{1}{M} \sum_{i=1}^M \theta_i. \tag{6.14}$$

It is worth noting that for all $\mathbf{z} \in \mathcal{Z}^n$, the average energy rate $B_n(\mathbf{z})$ is bounded according to

$$b_1 \leq B_n(\mathbf{z}) \leq b_0. \tag{6.15}$$

The inequalities in (6.15) imply that there exists a case in which energy transmission might occur with zero (maximal or average) ESP for all energy transmission rates $b \leq b_1$. This is because the event $B_n(\mathbf{Z}) < b_1$ is observed with zero probability. Alternatively, any energy transmission rate $b > b_0$ cannot be achieved with an average or maximal energy-shortage probability strictly smaller than one.

Energy transmission is said to be reliable if the average or maximum ESP is controlled. This leads to the following refinements of Definition 11.

Definition 14 ($(n, M, \epsilon, \delta, b)$ -code with maximum ESP). *Let $\delta \in [0, 1]$ and $b \geq 0$ be fixed. An (n, M, ϵ) -code that satisfies $\theta_i < \delta$, for all $i \in \{1, 2, \dots, M\}$, is said to be an $(n, M, \epsilon, \delta, b)$ -code with maximum ESP.*

Definition 15 ($(n, M, \epsilon, \delta, b)$ -code with average ESP). *Let $\delta \in [0, 1]$ and $b \geq 0$ be fixed. An (n, M, ϵ) -code that satisfies $\theta < \delta$ is said to be an $(n, M, \epsilon, \delta, b)$ -code with average ESP.*

Note that any $(n, M, \epsilon, \delta, b)$ -code with maximum ESP is also a $(n, M, \epsilon, \delta, b)$ -code with average ESP. Nonetheless, the converse is not necessarily true.

6.3 Fundamental Limits

The non-asymptotic fundamental limits of the system described in Section 6.2 are described by the notion of information-energy capacity region. That is, the set of all information and energy transmission rates that are achievable within a given block length subject to an average or maximum DEP and an average or maximum ESP. Note that an average or maximum DEP constraint leads to different definitions of the information-energy capacity region, and so does an average or maximum ESP constraint.

Definition 16 (Information-Energy Capacity Region). *The information-energy capacity region $\mathcal{C}(n, \epsilon, \delta)$ with average or maximum DEP and average or maximum ESP of the random transformation in (6.1) is the set of all pairs $(\frac{\log_2(M)}{n}, b)$ for which there exists an $(n, M, \epsilon, \delta, b)$ -code with average or maximum DEP and average or maximum ESP, respectively.*

6.3.1 Bounds on the Energy Transmission Rate This section provides some upper bounds on the energy transmission rate for any given $(n, M, \epsilon, \delta, b)$ -code with either maximum or average ESP constraints. These bounds are expressed in terms of the parameters of the code (Definition 11), the parameters of the random transformation in (6.1) and the empirical input distributions induced by the code.

Definition 17 (Empirical Distributions). *Consider an (n, M) -code described by the system in (6.5). For all $i \in \{1, 2, \dots, M\}$, the empirical probability distribution of the channel input symbols induced by the codeword $\mathbf{u}(i)$ is*

$$\bar{P}_X^{(i)}(0) \triangleq \frac{1}{n} N(0|\mathbf{u}(i)) = 1 - \bar{P}_X^{(i)}(1), \quad (6.16)$$

The empirical distribution of the channel input symbols jointly induced by all codewords is denoted by

$$\bar{P}_X(0) \triangleq \frac{1}{nM} \sum_{i=1}^M N(0|\mathbf{u}(i)) = 1 - \bar{P}_X(1). \quad (6.17)$$

Often, the vector $(\bar{P}_X^{(i)}(0), \bar{P}_X^{(i)}(1))$ is referred to as the *type* of the codeword $\mathbf{u}(i)$ [8].

Using the empirical distributions in Definition 17, some upper bounds on the energy transmission rate can be described. These upper bounds are obtained from the analysis of the ESP in (6.13). The following proposition provides the exact value of the ESP for any given $(n, M, \epsilon, \delta, b)$ -code.

Proposition 5 (Ground-Truth ESP). *Consider an $(n, M, \epsilon, \delta, b)$ -code described by the system in (6.5) for the random transformation in (6.1) satisfying (6.12). Then, for all $i \in \{1, 2, \dots, M\}$, the ESP in (6.13) satisfies*

$$\theta_i = \sum_{k=0}^{\lfloor \frac{n(b-b_1)}{b_0-b_1} \rfloor} \sum_{s=0}^k \binom{N(0|\mathbf{u}(i))}{s} \binom{N(1|\mathbf{u}(i))}{k-s} (1 - \alpha_2)^{N(1|\mathbf{u}(i))-k+2s} \alpha_2^{N(0|\mathbf{u}(i))+k-2s}. \quad (6.18)$$

Proof: The proof of Proposition 5 is presented in Appendix F. ■

The equality in (6.18) together with Definition 14 and Definition 15 provide the first bounds on the energy transmission rate of any given code. The following corollary describes these bounds.

Corollary 1. *Consider an $(n, M, \epsilon, \delta, b)$ -code described by the system in (6.5) for the random transformation in (6.1) satisfying (6.12). Then, subject to a maximal ESP constraint, it holds that,*

$$b < \tilde{B} \quad (6.19)$$

where \tilde{B} is the largest real that satisfies for all $i \in \{1, 2, \dots, M\}$,

$$\sum_{k=0}^{\lfloor \frac{n(\tilde{B}-b_1)}{b_0-b_1} \rfloor} \sum_{s=0}^k \binom{N(0|\mathbf{u}(i))}{s} \binom{N(1|\mathbf{u}(i))}{k-s} (1-\alpha_2)^{N(1|\mathbf{u}(i))-k+2s} \alpha_2^{N(0|\mathbf{u}(i))+k-2s} < \delta, \quad (6.20)$$

and subject to an average ESP constraint, the energy rate b satisfies

$$b < \check{B} \quad (6.21)$$

where \check{B} is the biggest positive real that satisfies

$$\frac{1}{M} \sum_{i=1}^M \sum_{k=0}^{\lfloor \frac{n(\check{B}-b_1)}{b_0-b_1} \rfloor} \sum_{s=0}^k \binom{N(0|\mathbf{u}(i))}{s} \binom{N(1|\mathbf{u}(i))}{k-s} (1-\alpha_2)^{N(1|\mathbf{u}(i))-k+2s} \alpha_2^{N(0|\mathbf{u}(i))+k-2s} < \delta. \quad (6.22)$$

The bounds in Corollary 1 are not in closed-form and thus, are difficult to calculate. Moreover, they bring very little insight to obtain a bound on the energy rate b at which an $(n, M, \epsilon, \delta, b)$ code can transmit energy. Therefore, it would be desirable to approximate the individual ESP in order to obtain an upper bound on the energy transmission rate in a closed form expression, probably, at the expense of some precision. The following proposition provides some approximations on the ESP using tools from large deviations theory [91].

Proposition 6. *Consider an $(n, M, \epsilon, \delta, b)$ -code described by the system in (6.5) for the random transformation in (6.1) satisfying (6.12). Then, for all $i \in \{1, 2, \dots, M\}$, the ESP in (6.13) satisfies*

$$\theta_i > 1 - \exp \left(-n \frac{\left(\frac{b-b_1}{b_0-b_1} - \left((1-2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 \right) \right)^2}{\frac{b-b_1}{b_0-b_1} + \left((1-2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 \right)} \right), \quad (6.23)$$

and

$$\theta_i < \exp \left(-n \frac{\left(\frac{b-b_1}{b_0-b_1} - \left((1-2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 \right) \right)^2}{2 \left((1-2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 \right)} \right). \quad (6.24)$$

Proof: The proof of Proposition 6 is presented in Appendix G. ■

Using Proposition 6, the energy transmission rate of an $(n, M, \epsilon, \delta, b)$ -code can be upper-bounded.

Proposition 7 (Large Deviation Bound). *Consider an $(n, M, \epsilon, \delta, b)$ -code described by the system in (6.5) for the random transformation in (6.1) satisfying (6.12). Then, subject to a maximal ESP constraint, it holds that for all $i \in \{1, 2, \dots, M\}$,*

$$b < (b_0 - b_1) \left((1 - 2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 \right) + b_1 + \frac{b_0 - b_1}{\sqrt{n}} \sqrt{-2 \left((1 - 2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 \right) \log(1 - \delta)} - \frac{b_0 - b_1}{n} \log(1 - \delta). \quad (6.25)$$

and subject to an average ESP constraint, the energy rate b satisfies

$$b < \hat{B}, \quad (6.26)$$

where \hat{B} is the biggest positive real that satisfies

$$1 - \delta < \frac{1}{M} \sum_{i=1}^M \exp \left(-n \frac{\left(\frac{b-b_1}{b_0-b_1} - \left((1-2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 \right) \right)^2}{\frac{b-b_1}{b_0-b_1} + \left((1-2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 \right)} \right). \quad (6.27)$$

Proof: The proof of Proposition 7 is presented in Appendix H. ■

Another approximation to the ESP in (6.13) is obtained from the Berry-Esseen theorem (Theorem 20 in Appendix E). The following lemma presents this approximation.

Lemma 6. *Consider an $(n, M, \epsilon, \delta, b)$ -code described by the system in (6.5) for the random transformation in (6.1) satisfying (6.12). Then, for all $i \in \{1, 2, \dots, M\}$, the ESP in (6.13) satisfies*

$$\theta_{i \geq Q} \left(\frac{n \left((1 - 2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 - \frac{b-b_1}{b_0-b_1} \right)}{\sqrt{n\alpha_2(1-\alpha_2)}} \right) - \frac{n [\alpha_2(1-\alpha_2)^3 + (1-\alpha_2)\alpha_2^3]}{2(n\alpha_2(1-\alpha_2))^{3/2}} \quad (6.28)$$

and

$$\theta_{i \leq Q} \left(\frac{n \left((1 - 2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 - \frac{b-b_1}{b_0-b_1} \right)}{\sqrt{n\alpha_2(1-\alpha_2)}} \right) + \frac{n [\alpha_2(1-\alpha_2)^3 + (1-\alpha_2)\alpha_2^3]}{2(n\alpha_2(1-\alpha_2))^{3/2}}. \quad (6.29)$$

Proof: The proof of Lemma 6 is presented in Appendix I. ■

Using Lemma 6, the energy transmission rate of an $(n, M, \epsilon, \delta, b)$ -code can be upper-bounded.

Proposition 8 (Gaussian Approximation Bound). *Consider an $(n, M, \epsilon, \delta, b)$ -code described by the system in (6.5) for the random transformation in (6.1) satisfying (6.12). Then, subject to a maximal ESP constraint, it holds that for all $i \in \{1, 2, \dots, M\}$,*

$$b \leq (b_0 - b_1) \left((1 - 2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 \right) + b_1 - \sqrt{\frac{(b_0 - b_1)^2 \alpha_2 (1 - \alpha_2)}{n}} Q^{-1} \left(\delta + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}} \right) \quad (6.30)$$

and subject to an average ESP constraint, the energy rate b satisfies

$$b \leq \hat{B}, \quad (6.31)$$

where \hat{B} is the biggest positive real that satisfies

$$\frac{1}{M} \sum_{i=1}^M Q \left(\frac{n \left((1 - 2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 - \frac{\hat{B} - b_1}{b_0 - b_1} \right)}{\sqrt{n\alpha_2(1 - \alpha_2)}} \right) - \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}} < \delta. \quad (6.32)$$

Proof: The proof of Proposition 8 is presented in Appendix I.1. ■

Note that the upper bound in (6.30), is valid when the following condition is satisfied

$$0 < \delta + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)^2}} < 1, \quad (6.33)$$

given that the domain of the function Q^{-1} is $(0, 1)$. In general, the bounds presented in Proposition 8 are tighter than those presented in Proposition 7 for small values of the block length n . Nonetheless, the bounds in Proposition 7 are easier to calculate and perform equally well for large n .

It is worth noting that the uniform distribution [89], i.e., $\bar{P}_X^{(i)}(0) = 1 - \bar{P}_X^{(i)}(1) = \frac{1}{2}$ is optimal for information transmission, without any energy constraint. By using this observation, the following corollary provides an outer bound on the energy rate that can be transmitted by an (n, M, ϵ) -code that possesses an empirical input distribution that is uniform

Corollary 2. *Given an $(n, M, \epsilon, \delta, b)$ -code for the random transformation in (6.2) satisfying (6.12). Assume that such a code uses an information-rate optimal empirical distribution. Then, it follows that $b < \underline{b}(n, \delta)$, with $\underline{b} : \mathbb{N} \times [0, 1] \rightarrow \mathbb{R}$, such that*

$$\underline{b}(n, \delta) \triangleq \frac{b_0 + b_1}{2} - \sqrt{\frac{(b_0 - b_1)^2 \alpha_2 (1 - \alpha_2)}{n}} Q^{-1} \left(\delta + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}} \right). \quad (6.34)$$

Specifically, Corollary 2 characterizes a threshold on the energy rate b beyond which an $(n, M, \epsilon, \delta, b)$ -code, if it exists, exhibits a trade-off between the energy transmission task and the information transmission task. More specifically, if there exists an $(n, M, \epsilon, \delta, b)$ -code whose energy transmission rate b is beyond the threshold \underline{b} in Corollary 2, it exhibits an empirical input distribution for which $\bar{P}_X(0) > \bar{P}_X(1)$. This implies a zero is transmitted more often than a one, which is not information-rate optimal.

Proposition 8 also provides upper bounds on the largest energy rate that can be transmitted by any $(n, M, \epsilon, \delta, b)$ -code. Note that the largest energy-transmission rate is achieved by a zero information-rate code whose codewords contain only zeros, i.e., $\bar{P}_X(0) = 1 - \bar{P}_X(1) = 1$. The following corollary describes this observation.

Corollary 3. *Consider an $(n, M, \epsilon, \delta, b)$ -code for the random transformation in (6.1) satisfying (6.12). Then, it follows that: $b < \bar{b}(n, \delta)$, with $\bar{b} : \mathbb{N} \times [0, 1] \rightarrow \mathbb{R}$, such that*

$$\bar{b}(n, \delta) \triangleq (1 - \alpha_2)b_0 + \alpha_2 b_1 - \sqrt{\frac{(b_0 - b_1)^2 \alpha_2 (1 - \alpha_2)}{n}} Q^{-1} \left(\delta + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}} \right). \quad (6.35)$$

From Corollary 2 and Corollary 3, it is expected that a trade-off between the energy transmission task and the information transmission task is occurred.

Denote by $\rho^* : [b_1, b_0] \rightarrow [0, 1]$ as

$$\rho^*(b) \triangleq \min(1, \rho^+(b)), \quad (6.36)$$

with $\rho^+ : [b_1, b_0] \rightarrow \mathbb{R}_+$ such that

$$\rho^+(b) \triangleq \frac{(b - b_1) - \alpha_2(b_0 - b_1)}{(b_0 - b_1)(1 - 2\alpha_2)} + \frac{\sqrt{\alpha_2(1 - \alpha_2)}}{\sqrt{n}(1 - 2\alpha_2)} Q^{-1} \left(\delta + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}} \right).$$

Note that for a fixed $b > 0$, $\rho^+(b)$ describes the empirical input distribution that saturates the inequality in (6.30). The following corollary from Proposition 8 highlights this observation.

Corollary 4. *Given an $(n, M, \epsilon, \delta, b)$ -code described by the system in (6.5) for the random transformation in (6.1) satisfying (6.12). Then, by assuming a maximal energy-shortage probability constraint, for all $i \in \{1, 2, \dots, M\}$, the empirical input distribution $\bar{P}_X^{(i)}$ satisfies the following the inequality*

$$\bar{P}_X^{(i)}(0) \geq \rho^*(b), \quad (6.37)$$

where $\rho^*(b)$ is defined in (6.36).

Corollary 4 leads to an interesting observation by noticing that $\rho^*(b)$ is a lower bound on the fraction of zeros in each codeword (maximal energy-shortage probability) when energy is transmitted at an average energy rate b . This is inline with the assumption in (6.12), which implies that the symbol zero carries more energy than the symbol one.

An interesting class of codes is that of homogeneous codes. A formal definition of these codes is hereunder.

Definition 18 (Homogeneous Codes). *A code \mathcal{C} described by the system in (6.5) is said to be homogeneous if the following conditions hold:*

$$N(0|\mathbf{u}(1)) = N(0|\mathbf{u}(2)) = \dots = N(0|\mathbf{u}(M)) \text{ and} \quad (6.38)$$

$$N(1|\mathbf{u}(1)) = N(1|\mathbf{u}(2)) = \dots = N(1|\mathbf{u}(M)). \quad (6.39)$$

The interest in this class of codes stems from the fact that an average ESP constraint or a maximum ESP constraint leads to the same fundamental limits on the energy rate.

Corollary 5. *Given an $(n, M, \epsilon, \delta, b)$ -code described by the system in (6.5) for the random transformation in (6.1) satisfying (6.12), and assume it is a homogeneous code. Then, the bounds on the energy rate b subject to a maximum ESP and average ESP are identical. That is, the bound in (6.19) is identical to (6.21); the one in (6.25) is identical to (6.26); and the one in (6.30) is identical to (6.31).*

6.3.2 Bounds on the Information Transmission Rate Given an $(n, M, \epsilon, \delta, b)$ -code, the following lemma describes a loose bound on M , which does not depend on the decoding error probability ϵ . However, it plays an important role when $\underline{b}(n, \delta) < b < \bar{b}(n, \delta)$.

Lemma 7. Given an $(n, M, \epsilon, \delta, b)$ -code for the random transformation in (6.1) which satisfies (6.12). Then by assuming to a maximal energy-shortage probability constraint, it holds that

$$M \leq \binom{n}{\lceil n\rho^*(b) \rceil} 2^{(n - \lceil n\rho^*(b) \rceil)}, \quad (6.40)$$

where $\rho^*(b)$ is defined by (6.36).

Proof of Proposition 7: An approximation to the minimum number of zeros in each codeword in any given $(n, M, \epsilon, \delta, b)$ -code with maximal energy-shortage probability is characterized in Corollary 4. That is, for all $i \in \{1, 2, \dots, M\}$ it follows that

$$N(0|\mathbf{u}(i)) \geq \lceil n\rho^*(b) \rceil. \quad (6.41)$$

This is an immediate result that provides an upper-bound on M given that all codewords must contain at least $\lceil n\rho^*(b) \rceil$ zero symbols. Hence, the right-hand side of (6.40) is the maximum number of codewords of length n for which at least $\lceil n\rho^*(b) \rceil$ symbols are zeros. This completes the proof. \blacksquare

Note that $\rho^*(b)$ is monotonically increasing with the energy rate b . Interestingly, when $\rho^*(b) \in (\frac{1}{2}, 1]$, the right-hand side of (6.40) is monotonically decreasing with b . This highlights the existing trade-off between the information transmission task and the energy transmission task. That is, in the regime in which $\rho^*(b) \in (\frac{1}{2}, 1]$, increasing the energy rate would necessarily imply decreasing the information rate.

6.3.3 Information-Energy Capacity Region Given a fixed block length n and a pair $(\epsilon, \delta) \in [0, 1]^2$, the information-energy capacity region $\mathcal{C}(n, \epsilon, \delta)$ (Definition 16) of the random transformation in (6.1) subject to (6.12) is approximated by a set $\underline{\mathcal{C}}(n, \epsilon, \delta)$ that is contained in $\mathcal{C}(n, \epsilon, \delta)$ (Theorem 18) and another set $\overline{\mathcal{C}}(n, \epsilon, \delta)$ that contains $\mathcal{C}(n, \epsilon, \delta)$ (Theorem 19). That is,

$$\underline{\mathcal{C}}(n, \epsilon, \delta) \subseteq \mathcal{C}(n, \epsilon, \delta) \subseteq \overline{\mathcal{C}}(n, \epsilon, \delta). \quad (6.42)$$

This approximation is obtained by considering an average DEP constraint and a maximum ESP constraint. The following notation is used to describe the set $\underline{\mathcal{C}}(n, \epsilon, \delta)$. Let functions $\phi : \mathbb{N} \times [0, 1] \rightarrow [0, 1]$, and $\chi : \mathbb{R}_+ \times [0, 1] \rightarrow [0, 1]$ be defined as,

$$\phi(m, \rho) \triangleq \min \left\{ 1, (m-1) \sum_{\ell_0=0}^n \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\min\{\ell_0, \ell_1\}} \sum_{\ell_3=0}^{\ell_0 + \ell_1 - 2\ell_2} \sum_{\ell_4=0}^{\ell_3} \binom{n}{\ell_0} \binom{\ell_0}{\ell_2} \binom{n - \ell_0}{\ell_1 - \ell_2} \binom{\ell_0}{\ell_4} \binom{n - \ell_0}{\ell_3 - \ell_4} \cdot \alpha_1^{\ell_0 + \ell_1 - 2\ell_2} (1 - \alpha_1)^{n - \ell_0 - \ell_1 + 2\ell_2} \rho^{\ell_3 + \ell_0 - 2\ell_4 + \ell_1} (1 - \rho)^{2n - (\ell_3 + \ell_0 - 2\ell_4 + \ell_1)} \right\}, \quad (6.43)$$

and

$$\chi(s, \rho) \triangleq \sum_{t=0}^n \binom{n}{t} \rho^t (1 - \rho)^{n-t} Q \left(\frac{(1 - 2\alpha_2)t + n \left(\alpha_2 - \frac{s-b_1}{b_0-b_1} \right)}{\sqrt{n\alpha_2(1 - \alpha_2)}} \right) + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}} \quad (6.44)$$

Let the functions $M_1^* : [0, 1] \rightarrow \mathbb{N}$, and $B^* : [0, 1] \rightarrow \mathbb{R}_+$ be defined as

$$M_1^*(\rho) \triangleq \begin{cases} \operatorname{argmax}_{m \in \mathbb{N}} \phi(m, \rho) \\ \text{s.t. } \phi(m, \rho) < \epsilon \end{cases}, \quad (6.45)$$

and

$$B^*(\rho) \triangleq \begin{cases} \operatorname{argmax}_{s \in \mathbb{R}_+} \chi(s, \rho) \\ \text{s.t. } \chi(s, \rho) < \delta \end{cases}. \quad (6.46)$$

Note that functions ϕ in (6.43), χ in (6.44), M_1^* in (6.45), and B^* in (6.46) depend on the block length n , the parameters of the random transformation in (6.1), i.e., α_1 and α_2 , and the energy harvested from symbols 0 and 1, i.e., b_0 and b_1 in (6.10). Nonetheless, none of these parameters is put as an argument of these functions given that they remain constant during this analysis. Using this notation, given a fixed block length n and a pair $(\epsilon, \delta) \in [0, 1]^2$, the following theorem introduces the set $\underline{\mathcal{C}}(n, \epsilon, \delta)$, that is contained in the information-energy capacity region $\mathcal{C}(n, \epsilon, \delta)$.

Theorem 18. *The information-energy capacity region $\mathcal{C}(n, \epsilon, \delta)$ of the random transformation in (6.1) subject to (6.12), contains the set*

$$\underline{\mathcal{C}}(n, \epsilon, \delta) \triangleq \left\{ (M, b) \in \mathbb{N} \times \mathbb{R}_+ : \exists \rho \in [0, 1], M < M_1^*(\rho) \text{ and } b < B^*(\rho) \right\}, \quad (6.47)$$

where $M_1^* : [0, 1] \rightarrow \mathbb{N}$ is defined in (6.45) and $B^* : [0, 1] \rightarrow \mathbb{R}_+$ is defined in (6.46).

Proof of Theorem 18: The proof of Theorem 18 is presented in Appendix J. ■

The description of the set $\bar{\mathcal{C}}(n, \epsilon, \delta)$ uses the following notation. Consider the function $\gamma : [0, 1]^2 \rightarrow [0, 1]$, such that

$$\begin{aligned} \gamma(\rho, q) = & \sum_{\ell_0=0}^n \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\min\{\ell_0, \ell_1\}} \binom{n}{\ell_0} \binom{\ell_0}{\ell_2} \binom{n-\ell_0}{\ell_1-\ell_2} \rho^{\ell_1} (1-\rho)^{n-\ell_1} q^{\ell_0} (1-q)^{n-\ell_0} \\ & \left(\mathbb{1}_{\{(\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} > L\}} + \lambda \mathbb{1}_{\{(\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} = L\}} \right) \end{aligned} \quad (6.48)$$

where $\lambda \in [0, 1]$ and $L \in \mathbb{R}$ are chosen to satisfy

$$\begin{aligned} 1 - \epsilon = & \sum_{\ell_0=0}^n \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\min\{\ell_0, \ell_1\}} \binom{n}{\ell_0} \binom{\ell_0}{\ell_2} \binom{n-\ell_0}{\ell_1-\ell_2} (1-\alpha_1)^{n-(\ell_0+\ell_1-2\ell_2)} \alpha_1^{\ell_0+\ell_1-2\ell_2} \rho^{\ell_1} (1-\rho)^{n-\ell_1} \\ & \cdot \left(\mathbb{1}_{\{(\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} > L\}} + \lambda \mathbb{1}_{\{(\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} = L\}} \right). \end{aligned}$$

Consider also the function $\Gamma : [b_1, b_0] \rightarrow \mathbb{R}_+$ defined as

$$\Gamma(b) \triangleq \sup_{\rho > \rho^*(b)} \inf_{q \in [0, \frac{1}{2}]} \frac{1}{\gamma(\rho, q)}. \quad (6.49)$$

Finally, let also $B^+ \in \mathbb{R}_+$ be defined as,

$$B^+ \triangleq (1 - \alpha_2)b_0 + \alpha_2 b_1 - \sqrt{\frac{(b_0 - b_1)^2 \alpha_2 (1 - \alpha_2)}{n}} \cdot Q^{-1} \left(\delta + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}} \right). \quad (6.50)$$

Using this notation, given a fixed block length n and a pair $(\epsilon, \delta) \in [0, 1]^2$, the following theorem introduces a set, denoted by $\bar{\mathcal{C}}(n, \epsilon, \delta)$, that contains the information-energy capacity region $\mathcal{C}(n, \epsilon, \delta)$.

Theorem 19. *The information-energy capacity region $\mathcal{C}(n, \epsilon, \delta)$ of the random transformation in (6.1) subject to (6.12), is contained in the set*

$$\bar{\mathcal{C}}(n, \epsilon, \delta) \triangleq \{(M, b) \in \mathbb{N} \times \mathbb{R}_+ : M < \Gamma(b) \text{ and } b < B^+\}, \quad (6.51)$$

where $\Gamma : \mathbb{R}_+ \rightarrow [0, 1]$ is defined in (6.49) and B^+ is defined in (6.50).

Proof: The proof of Theorem 19 is presented in Appendix K. ■

Conclusion

IN THIS thesis two new problems have been introduced. First, the problem of approximating the information-energy capacity regions of the two-user Gaussian interference channel, with and without perfect output feedback by two regions, i.e., an achievable region and a converse region was introduced. When the energy transmission rate is normalized by the maximum energy rate, the approximation of these information-energy capacity regions is within a constant gap. In the proof of achievability, the key idea is the use of power-splitting between two signal components: an information-carrying component and a no-information component. Random coding arguments are used for the case of the information-carrying component, whereas a deterministic sequence known by all transmitters and receivers is used for the no-information component. The proof of the converse of the information rates follows along the same lines of the case in which only information is transmitted. The difference stems from lifting the constraints on the mean of the channel input signals. The proof of the converse of the energy rate uses Markov's concentration inequality. The results presented in Chapter 5 are a first step in the study of the fundamental limits of simultaneous information and energy transmission, nonetheless many questions are left open. On the one hand, there exists sufficient evidence that the use of multiple antennas at either the transmitters or the receivers enhances the energy rate [92]. However, very little is known from the perspectives of fundamental limits. On the other hand, an interesting question is about the degradation of energy rates due to noisy feedback or rate-limited feedback. Similarly, another interesting question is about the benefits of other topologies of feedback, i.e., feedback from the receivers to both transmitters.

Second, the fundamental limits of simultaneous information and energy transmission (SIET) are studied in the non-asymptotic block-length regime. The focus is on the case of a transmitter simultaneously sending information to a receiver and energy to an energy harvester through the binary symmetric channel. Given a finite number of channel uses (latency constraint) as well as tolerable average decoding error probability and energy shortage probability (reliability constraints), two sets of information and energy transmission rates are presented. One consists of rate pairs for which the existence of at least one code achieving such rates under the latency and reliability constraints is proved (achievable region). The second one consists of a set whose complement contains the rate pairs for which there does not exist a code capable of achieving such rates (converse region). These two sets approximate the information energy capacity

7. Conclusion

region, which allows analyzing the trade-offs among performance, latency, and reliability in SIET systems.

A

Proof of Theorem 12

This proof is divided into two parts. The first part consists of the proof of (5.14a)-(5.14g). The second part consists of the proof of (5.14h).

A.0.1 Proof of (5.14a)-(5.14g) 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 $(v, u_1, u_2, s_1, s_2) \in \mathcal{X}_1 \cap \mathcal{X}_2 \times (\mathcal{X}_1 \times \mathcal{X}_2)^2$. 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 $\mathbf{v}(\omega) = (v_1(\omega), \dots, v_N(\omega))$ according to

$$P_V(\mathbf{v}(\omega)) = \prod_{m=1}^N P_V(v_m(\omega)), \quad (\text{A.1})$$

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

$$P_{U_1|V}(\mathbf{u}_1(\omega, i)|\mathbf{v}(\omega)) = \prod_{m=1}^N P_{U_1|V}(u_{1,m}(\omega, i)|v_m(\omega)), \quad (\text{A.2})$$

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

$$P_{S_1|U_1V}(\mathbf{s}_1(\omega, i, j)|\mathbf{u}_1(\omega, i), \mathbf{v}(\omega)) = \prod_{m=1}^N P_{S_1|U_1V}(s_{1,m}(\omega, i, j)|u_{1,m}(\omega, i), v_m(\omega)), \quad (\text{A.3})$$

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

$$P_{U_2|V}(\mathbf{u}_2(\omega, k)|\mathbf{v}(\omega)) = \prod_{m=1}^N P_{U_2|V}(u_{2,m}(\omega, k)|v_m(\omega)), \quad (\text{A.4})$$

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

$$P_{\mathbf{S}_2|U_2V}(\mathbf{s}_2(\omega, k, l)|\mathbf{u}_2(\omega, k), \mathbf{v}(\omega)) = \prod_{m=1}^N P_{S_{2,m}|U_2V}(s_{2,m}(\omega, k, l)|u_{2,m}(\omega, k), v_m(\omega)), \quad (\text{A.5})$$

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

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

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

$$\left(\mathbf{v}(\Omega), \mathbf{u}_1(\Omega, \widehat{W}_{1,C}), \mathbf{s}_1(\Omega, \widehat{W}_{1,C}, \widehat{W}_{1,P}), \mathbf{u}_2(\Omega, \widehat{W}_{2,C}), \mathbf{y}_1 \right) \in \mathcal{T}_{VU_1S_1U_2Y_1}^{(N,\epsilon)}, \quad (\text{A.6})$$

where Ω are assumed to be perfectly known by both transmitters and receivers. The set $\mathcal{T}_{VU_1S_1U_2Y_1}^{(N,\epsilon)}$ 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 (A.6), 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 (A.6). From the asymptotic equipartition property (AEP) [93], 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 $\mathbf{v}(\Omega)$, $\mathbf{u}_1(i, \Omega)$, $\mathbf{s}_1(i, j, \Omega)$, and $\mathbf{u}_2(k, \Omega)$ are jointly typical with a given \mathbf{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. Note that no error is declared when codewords $\left(\mathbf{v}(\Omega), \mathbf{u}_1(\Omega, 1), \mathbf{s}_1(\Omega, 1, 1), \mathbf{u}_2(\Omega, \hat{k}) \right)$, where $\hat{k} \neq 1$, are the only jointly typical sequences with the received sequence \mathbf{y}_1 . This is due to the fact that receiver 1 is interested only on the indices i and j . Then, the probability of error P_e due to the event (ii), can be bounded as follows:

$$\begin{aligned} P_{\text{DE}}^{(N)} &= \Pr \left(\bigcup_{(i,j) \neq (1,1)} \bigcup_{k=1}^{2^{NR_{2,C}}} E_{ijk} \right) \\ &\leq \sum_{i \neq 1, j \neq 1, k \neq 1} \Pr(E_{ijk}) + \sum_{i=1, j \neq 1, k=1} \Pr(E_{ijk}) + \sum_{i \neq 1, j \neq 1, k=1} \Pr(E_{ijk}) \\ &\quad + \sum_{i \neq 1, j=1, k \neq 1} \Pr(E_{ijk}) + \sum_{i \neq 1, j=1, k=1} \Pr(E_{ijk}) + \sum_{i=1, j \neq 1, k \neq 1} \Pr(E_{ijk}). \end{aligned} \quad (\text{A.7})$$

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

$$\Pr[(E_{i11})]=\Pr\left[(\mathbf{V}, \mathbf{U}_1, \mathbf{S}_1, \mathbf{U}_2, \mathbf{Y}_1) \in \mathcal{T}_{VU_1S_1U_2Y_1}^{(N,\epsilon)}\right] \quad (\text{A.8a})$$

$$= \sum_{(\mathbf{v}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2, \mathbf{y}_1) \in \mathcal{T}_{VU_1S_1U_2Y_1}^{(N,\epsilon)}} P_{U_1S_1|V}(\mathbf{u}_1, \mathbf{s}_1|\mathbf{v})P_{U_2Y_1|V}(\mathbf{u}_2, \mathbf{y}_1|\mathbf{v})P_V(\mathbf{v}), \quad (\text{A.8b})$$

$$\leq \left| \mathcal{T}_{VU_1S_1U_2Y_1}^{(N,\epsilon)} \right| 2^{-N(H(U_1, S_1|V)+H(U_2, Y_1|V)+H(V)-3\epsilon)} \quad (\text{A.8c})$$

$$\leq 2^{-N(H(U_1, S_1|V)+H(U_2, Y_1|V)+H(V)-H(V, U_1, S_1, U_2, Y_1)-4\epsilon)} \quad (\text{A.8d})$$

$$= 2^{-N(H(U_1, S_1|V)+H(U_2, Y_1|V)-H(U_1, S_1, U_2, Y_1|V)-4\epsilon)} \quad (\text{A.8e})$$

$$= 2^{-N(I(U_1, S_1; U_2, Y_1|V)-4\epsilon)} \quad (\text{A.8f})$$

$$= 2^{-N(I(U_1, S_1; Y_1|U_2, V)+I(U_1, S_1; U_2|V)-4\epsilon)} \quad (\text{A.8g})$$

$$= 2^{-N(I(U_1, S_1; Y_1|U_2, V)-4\epsilon)} \quad (\text{A.8h})$$

$$= 2^{-N(I(X_1; Y_1|U_2, V)-4\epsilon)}, \quad (\text{A.8i})$$

where (A.8f) follows from the fact that, for any three random variables X , Y , and Z it holds that

$$I(X; Y|Z) = H(X|Z) + H(Y|Z) - H(X, Y|Z); \quad (\text{A.9})$$

and (A.8i) follows from the fact that $X_i = \theta_i(S_i, U_i, V)$. Note that the probability operator $\Pr[\cdot]$ in (A.8a) applies with a probability distribution $P_{VU_1S_1U_2Y_1}$ that factorizes as $P_V P_{U_1S_1|V} P_{U_2Y_1|V}$ given that all the codewords \mathbf{s}_1 and \mathbf{u}_1 are independent from the given channel output \mathbf{y}_1 .

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

$$\Pr(E_{1j1})=\Pr\left[(\mathbf{V}, \mathbf{U}_1, \mathbf{S}_1, \mathbf{U}_2, \mathbf{Y}_1) \in \mathcal{T}_{VU_1S_1U_2Y_1}^{(N,\epsilon)}\right] \quad (\text{A.10a})$$

$$= \sum_{(\mathbf{v}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{y}_1) \in \mathcal{T}_{VU_1S_1U_2Y_1}^{(N,\epsilon)}} P_{S_1|U_1V}(\mathbf{s}_1|\mathbf{u}_1, \mathbf{v})P_{Y_1|U_1U_2V}(\mathbf{y}_1|\mathbf{u}_1, \mathbf{u}_2, \mathbf{v})P_{U_1U_2V}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}), \quad (\text{A.10b})$$

$$\leq \left| \mathcal{T}_{VU_1S_1U_2Y_1}^{(N,\epsilon)} \right| 2^{-N(H(S_1|V, U_1)+H(Y_1|V, U_1, U_2)+H(V, U_1, U_2)-3\epsilon)} \quad (\text{A.10c})$$

$$\leq 2^{-N(H(S_1|V, U_1)+H(Y_1|V, U_1, U_2)+H(V, U_1, U_2)-H(V, U_1, S_1, U_2, Y_1)-4\epsilon)} \quad (\text{A.10d})$$

$$= 2^{-N(H(S_1|V, U_1, U_2)+H(Y_1|V, U_1, U_2)-H(S_1, Y_1|U_1, U_2, V)-4\epsilon)} \quad (\text{A.10e})$$

$$= 2^{-N(I(S_1; Y_1|U_1, U_2, V)-4\epsilon)} \quad (\text{A.10f})$$

$$= 2^{-N(I(X_1; Y_1|U_1, U_2, V)-4\epsilon)}, \quad (\text{A.10g})$$

where (A.10f) follows from (A.9); and (A.10g) follows from the fact that $X_i = \theta_i(S_i, U_i, V)$. Note that the probability operator $\Pr[\cdot]$ in (A.10a) applies with a probability distribution $P_{VU_1S_1U_2Y_1}$ that factorizes as $P_{S_1|U_1V} P_{Y_1|U_1U_2V} P_{U_1U_2V}$ given that the codeword \mathbf{s}_1 is independent from the given channel output \mathbf{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:

$$\Pr(E_{ij1})=\Pr\left[(\mathbf{V}, \mathbf{U}_1, \mathbf{S}_1, \mathbf{U}_2, \mathbf{Y}_1) \in \mathcal{T}_{VU_1S_1U_2Y_1}^{(N,\epsilon)}\right] \quad (\text{A.11a})$$

$$= \sum_{(\mathbf{v}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2, \mathbf{y}_1) \in \mathcal{T}_{VU_1S_1U_2Y_1}^{(N, \epsilon)}} P_{U_1S_1|V}(\mathbf{u}_1, \mathbf{s}_1|\mathbf{v}) P_{U_2Y_1|V}(\mathbf{u}_2, \mathbf{y}_1|\mathbf{v}) P_V(\mathbf{v}), \quad (\text{A.11b})$$

$$\leq 2^{-N(H(U_1, S_1|V) + H(U_2, Y_1|V) + H(V) - H(V, U_1, S_1, U_2, Y_1) - 4\epsilon)} \quad (\text{A.11c})$$

$$= 2^{-N(H(U_1, S_1|V) + H(U_2, Y_1|V) - H(U_1, S_1, U_2, Y_1|V) - 4\epsilon)} \quad (\text{A.11d})$$

$$= 2^{-N(I(U_1, S_1; U_2, Y_1|V) - 4\epsilon)} \quad (\text{A.11e})$$

$$= 2^{-N(I(U_1, S_1; Y_1|U_2, V) + I(U_1, S_1; U_2|V) - 4\epsilon)} \quad (\text{A.11f})$$

$$= 2^{-N(I(U_1, S_1; Y_1|U_2, V) - 4\epsilon)} \quad (\text{A.11g})$$

$$= 2^{-N(I(X_1; Y_1|U_2, V) - 4\epsilon)}, \quad (\text{A.11h})$$

where (A.11e) follows from (A.9); and (A.11h) follows from the fact that $X_i = \theta_i(S_i, U_i, V)$. Note that the probability operator $\Pr[\cdot]$ in (A.11a) applies with a probability distribution $P_{VU_1S_1U_2Y_1}$ that factorizes as $P_V P_{U_1S_1|V} P_{U_2Y_1|V}$ given that all the codewords \mathbf{s}_1 and \mathbf{u}_1 are independent from the output of the channel output \mathbf{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:

$$\Pr(E_{i1k}) = \Pr\left[(\mathbf{V}, \mathbf{U}_1, \mathbf{S}_1, \mathbf{U}_2, \mathbf{Y}_1) \in \mathcal{T}_{VU_1S_1U_2Y_1}^{(N, \epsilon)}\right] \quad (\text{A.12a})$$

$$= \sum_{(\mathbf{v}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2, \mathbf{y}_1) \in \mathcal{T}_{VU_1S_1U_2Y_1}^{(N, \epsilon)}} P_{U_1S_1U_2|V}(\mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2|\mathbf{v}) P_{Y_1|V}(\mathbf{y}_1|\mathbf{v}) P_V(\mathbf{v}), \quad (\text{A.12b})$$

$$\leq \left| \mathcal{T}_{VU_1S_1U_2Y_1}^{(N, \epsilon)} \right| 2^{-N(H(U_1, S_1, U_2|V) + H(Y_1|V) + H(V) - 3\epsilon)} \quad (\text{A.12c})$$

$$\leq 2^{-N(H(U_1, S_1, U_2|V) + H(Y_1|V) + H(V) - H(V, U_1, S_1, U_2, Y_1) - 4\epsilon)} \quad (\text{A.12d})$$

$$= 2^{-N(H(U_1, S_1, U_2|V) + H(Y_1|V) - H(Y_1, U_1, S_1, U_2|V) - 4\epsilon)} \quad (\text{A.12e})$$

$$= 2^{-N(I(S_1, U_1, U_2; Y_1|V) - 4\epsilon)} \quad (\text{A.12f})$$

$$= 2^{-N(I(X_1, U_2; Y_1|V) - 4\epsilon)}, \quad (\text{A.12g})$$

where (A.12f) follows from (A.9); and (A.12g) follows from the fact that $X_i = \theta_i(S_i, U_i, V)$. Note that the probability operator $\Pr[\cdot]$ in (A.12a) applies with a probability distribution $P_{VU_1S_1U_2Y_1}$ that factorizes as $P_{U_1S_1U_2|V} P_{Y_1|V} P_V$ given that all the codewords \mathbf{s}_1 , \mathbf{u}_1 and \mathbf{u}_2 are independent from the given channel output \mathbf{y}_1 .

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

$$\Pr(E_{ij1}) = \Pr\left[(\mathbf{V}, \mathbf{U}_1, \mathbf{S}_1, \mathbf{U}_2, \mathbf{Y}_1) \in \mathcal{T}_{VU_1S_1U_2Y_1}^{(N, \epsilon)}\right] \quad (\text{A.13a})$$

$$= \sum_{(\mathbf{v}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2, \mathbf{y}_1) \in \mathcal{T}_{VU_1S_1U_2Y_1}^{(N, \epsilon)}} P_{U_1S_1|V}(\mathbf{u}_1, \mathbf{s}_1|\mathbf{v}) P_{U_2Y_1|V}(\mathbf{u}_2, \mathbf{y}_1|\mathbf{v}) P_V(\mathbf{v}), \quad (\text{A.13b})$$

$$\leq 2^{-N(H(U_1, S_1|V) + H(U_2, Y_1|V) + H(V) - H(V, U_1, S_1, U_2, Y_1) - 4\epsilon)} \quad (\text{A.13c})$$

$$= 2^{-N(H(U_1, S_1|V) + H(U_2, Y_1|V) - H(U_1, S_1, U_2, Y_1|V) - 4\epsilon)} \quad (\text{A.13d})$$

$$= 2^{-N(I(U_1, S_1; U_2, Y_1|V) - 4\epsilon)} \quad (\text{A.13e})$$

$$= 2^{-N(I(U_1, S_1; Y_1|U_2, V) + I(U_1, S_1; U_2|V) - 4\epsilon)} \quad (\text{A.13f})$$

$$= 2^{-N(I(U_1, S_1; Y_1|U_2, V) - 4\epsilon)} \quad (\text{A.13g})$$

$$= 2^{-N(I(X_1; Y_1|U_2, V) - 4\epsilon)}, \quad (\text{A.13h})$$

where (A.13e) follows from (A.9); and (A.13h) follows from the fact that $X_i = \theta_i(S_i, U_i, V)$. Note that the probability operator $\Pr[\cdot]$ in (A.13a) applies with a probability distribution $P_{\mathbf{V}U_1S_1U_2Y_1}$ that factorizes as $P_{\mathbf{V}}P_{U_1S_1|V}P_{U_2Y_1|V}$ given that all the codewords \mathbf{s}_1 and \mathbf{u}_1 are independent from the output of the channel output \mathbf{y}_1 .

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

$$\Pr(E_{ijk}) = \Pr\left[(\mathbf{V}, U_1, \mathbf{S}_1, U_2, \mathbf{Y}_1) \in \mathcal{T}_{VU_1S_1U_2Y_1}^{(N,\epsilon)}\right] \quad (\text{A.14a})$$

$$= \sum_{(\mathbf{v}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2, \mathbf{y}_1) \in \mathcal{T}_{VU_1S_1U_2Y_1}^{(N,\epsilon)}} P_{U_1S_1U_2|V}(\mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2|\mathbf{v}) P_{Y_1|V}(\mathbf{y}_1|\mathbf{v}) P_{\mathbf{V}}(\mathbf{v}), \quad (\text{A.14b})$$

$$\stackrel{(b)}{\leq} \left| \mathcal{T}_{VU_1S_1U_2Y_1}^{(N,\epsilon)} \right| 2^{-n(H(U_1, S_1, U_2|V) + H(Y_1|V) + H(V) - 3\epsilon)} \quad (\text{A.14c})$$

$$\leq 2^{-N(H(U_1, S_1, U_2|V) + H(Y_1|V) + H(V) - H(V, U_1, S_1, U_2, Y_1) - 4\epsilon)} \quad (\text{A.14d})$$

$$= 2^{-N(H(U_1, S_1, U_2|V) + H(Y_1|V) - H(Y_1, U_1, S_1, U_2|V) - 4\epsilon)} \quad (\text{A.14e})$$

$$= 2^{-N(I(S_1, U_1, U_2; Y_1|V) - 4\epsilon)} \quad (\text{A.14f})$$

$$= 2^{-N(I(X_1, U_2; Y_1|V) - 4\epsilon)}, \quad (\text{A.14g})$$

where (A.14f) follows from (A.9); and (A.14g) follows from the fact that $X_i = \theta_i(S_i, U_i, V)$. Note that the probability operator $\Pr[\cdot]$ in (A.14a) applies with a probability distribution $P_{\mathbf{V}U_1S_1U_2Y_1}$ that factorizes as $P_{U_1S_1U_2|V}P_{Y_1|V}P_{\mathbf{V}}(\mathbf{v})$ given that all the codewords \mathbf{s}_1 , \mathbf{u}_1 and \mathbf{u}_2 are independent from the given channel output \mathbf{y}_1 .

Using (A.8), (A.10), (A.11), (A.12), (A.13), and (A.14) in (A.7), the following holds:

$$\begin{aligned} P_{\text{DE}}^{(N)} &\leq 2^{N(R_{1c} - I(X_1; Y_1|U_2, V) + 4\epsilon)} 2^{N(R_{1p} - I(X_1; Y_1|U_1, U_2, V) + 4\epsilon)} \\ &\quad + 2^{N(R_{1c} + R_{1p} - I(X_1; Y_1|U_2, V) + 4\epsilon)} + 2^{N(R_{1c} + R_{2c} - I(U_2, X_1; Y_1|V) + 4\epsilon)} \\ &\quad + 2^{N(R_{1p} + R_{2c} - I(U_2, X_1; Y_1|U_1, V) + 4\epsilon)} + 2^{N(R_{1p} + R_{1c} + R_{2c} - I(U_2, X_1; Y_1|V) + 4\epsilon)}. \end{aligned} \quad (\text{A.15})$$

The same analysis of the probability of error holds for transmitter-receiver pair 2. Hence in general, from (A.15), reliable decoding holds under the following conditions:

$$R_{1p} \leq a_1, \quad (\text{A.16a})$$

$$R_{1p} + R_{1c} \leq d_1, \quad (\text{A.16b})$$

$$R_{1p} + R_{2c} \leq e_1, \quad (\text{A.16c})$$

$$R_{1p} + R_{1c} + R_{2c} \leq g_1, \quad (\text{A.16d})$$

$$-R_{1p} \leq 0, \quad (\text{A.16e})$$

$$-R_{1c} \leq 0, \quad (\text{A.16f})$$

$$R_{2p} \leq a_2, \quad (\text{A.16g})$$

$$R_{2p} + R_{2c} \leq d_2, \quad (\text{A.16h})$$

$$R_{2p} + R_{1c} \leq e_2, \quad (\text{A.16i})$$

$$R_{1p} + R_{1c} + R_{2c} \leq g_2, \quad (\text{A.16j})$$

$$-R_{2p} \leq 0, \quad (\text{A.16k})$$

$$-R_{2c} \leq 0, \quad (\text{A.16l})$$

where

$$a_1 = I(Y_1; X_1 | U_1, U_2, V), \quad (\text{A.17a})$$

$$d_1 = I(Y_1; X_1 | U_2, V), \quad (\text{A.17b})$$

$$e_1 = I(Y_1; X_1, U_2 | U_1, V), \quad (\text{A.17c})$$

$$g_1 = I(Y_1; X_1, U_2 | V), \quad (\text{A.17d})$$

$$a_2 = I(Y_2; X_2 | U_1, U_2, V), \quad (\text{A.17e})$$

$$d_2 = I(Y_2; X_2 | U_1, V), \quad (\text{A.17f})$$

$$e_2 = I(Y_2; X_2, U_1 | U_2, V), \text{ and } \quad (\text{A.17g})$$

$$g_2 = I(Y_2; X_2, U_1 | V). \quad (\text{A.17h})$$

The proof continues by applying a Fourier-Motzkin elimination process on (A.16). Set $R_{1p} = R_1 - R_{1c}$, $R_{2p} = R_2 - R_{2c}$ and eliminate R_{1p} , R_{2p} from the set of inequalities (A.16) to obtain

$$R_1 - R_{1c} \leq a_1, \quad (\text{A.18a})$$

$$R_1 \leq d_1, \quad (\text{A.18b})$$

$$R_1 - R_{1c} + R_{2c} \leq e_1, \quad (\text{A.18c})$$

$$R_1 + R_{2c} \leq g_1, \quad (\text{A.18d})$$

$$-R_1 + R_{1c} \leq 0, \quad (\text{A.18e})$$

$$-R_{1c} \leq 0, \quad (\text{A.18f})$$

$$R_2 - R_{2c} \leq a_2, \quad (\text{A.18g})$$

$$R_2 \leq d_2, \quad (\text{A.18h})$$

$$R_2 - R_{2c} + R_{1c} \leq e_2, \quad (\text{A.18i})$$

$$R_2 + R_{1c} \leq g_2, \quad (\text{A.18j})$$

$$-R_2 + R_{2c} \leq 0, \quad (\text{A.18k})$$

$$-R_{2c} \leq 0. \quad (\text{A.18l})$$

Collect the inequalities in (A.18) that do not include R_{1c} to obtain:

$$R_1 \leq d_1, \quad (\text{A.19a})$$

$$R_1 + R_{2c} \leq g_1, \quad (\text{A.19b})$$

$$R_2 - R_{2c} \leq a_2, \quad (\text{A.19c})$$

$$R_2 \leq d_2, \quad (\text{A.19d})$$

$$-R_2 + R_{2c} \leq 0, \quad (\text{A.19e})$$

$$-R_{2c} \leq 0. \quad (\text{A.19f})$$

Also, collect the inequalities in (A.18) that include R_{1c} with positive coefficients to obtain:

$$R_2 - R_{2c} + R_{1c} \leq e_2, \quad (\text{A.20a})$$

$$R_2 + R_{1c} \leq g_2, \quad (\text{A.20b})$$

$$-R_1 + R_{1c} \leq 0. \quad (\text{A.20c})$$

Furthermore, collect the inequalities in (A.18) that include R_{1c} with negative coefficients to obtain:

$$-R_{1c} \leq 0, \quad (\text{A.21a})$$

$$R_1 - R_{1c} \leq a_1, \quad (\text{A.21b})$$

$$R_1 - R_{1c} + R_{2c} \leq e_1. \quad (\text{A.21c})$$

Eliminate R_{1c} by summing each inequality in (A.20) with (A.21) to obtain:

$$R_2 - R_{2c} \leq e_2, \quad (\text{A.22a})$$

$$R_2 \leq g_2, \quad (\text{A.22b})$$

$$-R_1 \leq 0, \quad (\text{A.22c})$$

$$R_1 + R_2 - R_{2c} \leq a_1 + e_2, \quad (\text{A.22d})$$

$$R_1 + R_2 \leq a_1 + g_2, \quad (\text{A.22e})$$

$$0 \leq a_1, \quad (\text{A.22f})$$

$$R_1 + R_2 \leq e_1 + e_2, \quad (\text{A.22g})$$

$$R_1 + R_2 + R_{2c} \leq e_1 + g_2, \quad (\text{A.22h})$$

$$R_{2c} \leq e_1. \quad (\text{A.22i})$$

Collect the inequalities in (A.19) and (A.22) that do not include R_{2c} to obtain:

$$0 \leq a_1, \quad (\text{A.23a})$$

$$R_1 \leq d_1, \quad (\text{A.23b})$$

$$R_2 \leq d_2, \quad (\text{A.23c})$$

$$R_2 \leq g_2, \quad (\text{A.23d})$$

$$-R_1 \leq 0, \quad (\text{A.23e})$$

$$R_1 + R_2 \leq a_1 + g_2, \quad (\text{A.23f})$$

$$R_1 + R_2 \leq e_1 + e_2. \quad (\text{A.23g})$$

Note that (A.23a) is redundant due to the positivity of mutual information. The inequality (A.23d) is redundant with respect to (A.23c), given that:

$$\begin{aligned} g_2 &= I(Y_2; X_2, U_1 | V) \\ &= I(Y_2; U_1 | V) + I(Y_2; X_2 | U_1, V) \\ &= d_2 + I(Y_2; U_1 | V) \\ &\geq d_2. \end{aligned} \quad (\text{A.24})$$

Collect the inequalities in (A.19) and (A.22) that include R_{2c} with positive coefficients to obtain:

$$R_{2c} \leq e_1, \quad (\text{A.25a})$$

$$R_1 + R_{2c} \leq g_1, \quad (\text{A.25b})$$

$$R_1 + R_2 + R_{2c} \leq e_1 + g_2, \quad (\text{A.25c})$$

$$-R_2 + R_{2c} \leq 0. \quad (\text{A.25d})$$

Collect the inequalities in (A.19) and (A.22) that include R_{2c} with negative coefficients to obtain:

$$-R_{2c} \leq 0, \quad (\text{A.26a})$$

$$R_2 - R_{2c} \leq a_2, \quad (\text{A.26b})$$

$$R_2 - R_{2c} \leq e_2, \quad (\text{A.26c})$$

$$R_1 + R_2 - R_{2c} \leq a_1 + e_2. \quad (\text{A.26d})$$

The inequality (A.26c) is redundant with respect to (A.26b), as shown hereunder:

$$\begin{aligned} e_2 &= I(Y_2; X_2, U_1 | U_2, V) \\ &= I(Y_2; U_1 | U_2, V) + I(Y_2; X_2 | U_2, U_1, V) \\ &= a_2 + I(Y_2; U_1 | U_2, V) \\ &\geq a_2. \end{aligned} \quad (\text{A.27})$$

Eliminate R_{2c} by adding each inequality from (A.25) to each inequality (A.26) to obtain inequalities not including R_{2c} :

$$0 \leq e_1, \quad (\text{A.28a})$$

$$R_1 \leq g_1, \quad (\text{A.28b})$$

$$R_1 + R_2 \leq e_1 + g_2, \quad (\text{A.28c})$$

$$-R_2 \leq 0, \quad (\text{A.28d})$$

$$R_2 \leq a_2 + e_1, \quad (\text{A.28e})$$

$$R_1 + R_2 \leq a_2 + g_1, \quad (\text{A.28f})$$

$$R_1 + 2R_2 \leq a_2 + e_1 + g_2, \quad (\text{A.28g})$$

$$0 \leq a_2, \quad (\text{A.28h})$$

$$R_1 + R_2 \leq a_1 + e_2 + e_1, \quad (\text{A.28i})$$

$$2R_1 + R_2 \leq a_1 + e_2 + g_1, \quad (\text{A.28j})$$

$$2R_1 + 2R_2 \leq a_1 + e_2 + e_1 + g_2, \quad (\text{A.28k})$$

$$R_1 \leq a_1 + e_2. \quad (\text{A.28l})$$

Note that (A.28a) and (A.28h) are redundant due to the positivity of mutual information. The inequality (A.28b) is redundant with respect to (A.23b), as shown hereunder:

$$\begin{aligned} g_1 &= I(Y_1; X_1, U_2 | V) \\ &= I(Y_1; U_2 | V) + I(Y_1; X_1 | U_2, V) \\ &= d_1 + I(Y_1; U_2 | V) \\ &\geq d_1. \end{aligned} \quad (\text{A.29})$$

The inequality (A.28c) is redundant with respect to (A.23f), since

$$\begin{aligned}
e_1 &= I(Y_1; X_1, U_2 | U_1, V) \\
&= I(Y_1; U_2 | U_1, V) + I(Y_1; X_1 | U_1, U_2, V) \\
&= a_1 + I(Y_1; U_2 | U_1, V) \\
&\geq a_1.
\end{aligned} \tag{A.30}$$

Note also that the inequality (A.28i) is redundant with respect to (A.23g). Furthermore, the inequality (A.28k) is redundant with respect to (A.23f) and (A.23g). Hence from (A.23) and (A.28), the system of inequalities in (A.16) can be simplified as follows:

$$R_1 \leq d_1, \tag{A.31a}$$

$$R_1 \leq a_1 + e_2, \tag{A.31b}$$

$$R_2 \leq d_2, \tag{A.31c}$$

$$R_2 \leq a_2 + e_1, \tag{A.31d}$$

$$R_1 + R_2 \leq a_1 + g_2, \tag{A.31e}$$

$$R_1 + R_2 \leq a_2 + g_1, \tag{A.31f}$$

$$R_1 + R_2 \leq e_1 + e_2, \tag{A.31g}$$

$$2R_1 + R_2 \leq a_1 + g_1 + e_2, \tag{A.31h}$$

$$R_1 + 2R_2 \leq a_2 + g_2 + e_1. \tag{A.31i}$$

Finally, following the result of [41], the inequalities (A.31b) and (A.31d) can be dropped. Let $k \in \{1, 2\}$ be fixed. Consider the following Gaussian input distribution for transmitter k :

$$V \sim \mathcal{N}(0, 1); U_k \sim \mathcal{N}(0, \lambda_{kc}); S_k \sim \mathcal{N}(0, \lambda_{kp}), \tag{A.32}$$

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 (5.14a), (5.14c), (5.14e) and (5.14f). The choice of the Gaussian input distribution in (A.32) yields:

$$\begin{aligned}
I(X_1; Y_1 | U_2, V) &= h(Y_1 | U_2, V) - h(Y_1 | X_1, U_2, V) \\
&= \frac{1}{2} \log \left(2\pi e \text{Var}[Y_1 | U_2, V] \right) - \frac{1}{2} \log \left(2\pi e \text{Var}[Y_1 | X_1, U_2, V] \right) \\
&= \frac{1}{2} \log \left(2\pi e \left(\sigma_1^2 + h_{1,1}^2 (1 - \lambda_{1e}) P_1 - h_{2,1}^2 \lambda_{2p} P_2 \right) \right) \\
&\quad - \frac{1}{2} \log \left(2\pi e \left(\sigma_1^2 + h_{2,1}^2 \lambda_{2p} P_2 \right) \right) \\
&= \frac{1}{2} \log \left(1 + \frac{(1 - \lambda_{1e}) \text{SNR}_1}{1 + \lambda_{2p} \text{INR}_1} \right),
\end{aligned} \tag{A.33a}$$

$$\begin{aligned}
I(X_1, U_2; Y_1 | V) &= h(Y_1 | V) - h(Y_1 | X_1, U_2, V) \\
&= \frac{1}{2} \log \left(2\pi e \text{Var}[Y_1 | V] \right) - \frac{1}{2} \log \left(2\pi e \text{Var}[Y_1 | X_1, U_2, V] \right) \\
&= \frac{1}{2} \log \left(2\pi e \left(\sigma_1^2 + h_{1,1}^2 (1 - \lambda_{1e}) P_1 + h_{2,2}^2 (1 - \lambda_{2e}) P_2 \right) \right)
\end{aligned}$$

$$\begin{aligned}
& -\frac{1}{2} \log \left(2\pi e(\sigma_1^2 + h_{2,1}^2 \lambda_{2p} P_2) \right) \\
& = \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{1e}) \text{SNR}_1 + (1 - \lambda_{2e}) \text{INR}_1}{1 + \lambda_{2p} \text{INR}_1} \right), \tag{A.33b}
\end{aligned}$$

$$\begin{aligned}
I(X_2; Y_2 | U_1, U_2, V) &= h(Y_2 | U_1, U_2, V) - h(Y_2 | U_1, U_2, X_2, V) \\
&= \frac{1}{2} \log \left(2\pi e \text{Var}[Y_2 | U_1, U_2, V] \right) - \frac{1}{2} \log \left(2\pi e \text{Var}[Y_2 | U_1, U_2, X_2, V] \right) \\
&= \frac{1}{2} \log \left(2\pi e(\sigma_2^2 + h_{2,2}^2 \lambda_{2p} P_2 + h_{2,1}^2 \lambda_{1p} P_1) \right) \\
&\quad - \frac{1}{2} \log \left(2\pi e(\sigma_2^2 + h_{2,1}^2 \lambda_{1p} P_1) \right) \\
&= \frac{1}{2} \log \left(1 + \frac{\lambda_{2p} \text{SNR}_2}{1 + \lambda_{1p} \text{INR}_2} \right), \tag{A.33c}
\end{aligned}$$

$$\begin{aligned}
I(X_1, U_2; Y_1 | U_1, V) &= h(Y_1 | U_1, V) - h(Y_1 | U_1, U_2, X_1, V) \\
&= \frac{1}{2} \log \left(2\pi e \text{Var}[Y_1 | U_1, V] \right) - \frac{1}{2} \log \left(2\pi e \text{Var}[Y_1 | U_1, U_2, X_1, V] \right) \\
&= \frac{1}{2} \log \left(2\pi e(\sigma_1^2 + h_{1,1}^2 \lambda_{1p} P_1 + h_{1,2}^2 (1 - \lambda_{2e}) P_2) \right) \\
&\quad - \frac{1}{2} \log \left(2\pi e(\sigma_1^2 + h_{1,2}^2 \lambda_{2p} P_2) \right) \\
&= \frac{1}{2} \log \left(\frac{1 + \lambda_{1p} \text{SNR}_1 + (1 - \lambda_{2e}) \text{INR}_1}{1 + \lambda_{2p} \text{INR}_1} \right), \tag{A.33d}
\end{aligned}$$

$$\begin{aligned}
I(X_2, U_1; Y_2 | U_2, V) &= h(Y_2 | U_2, V) - h(Y_2 | U_2, U_1, X_2, V) \\
&= \frac{1}{2} \log \left(2\pi e \text{Var}[Y_2 | U_2, V] \right) - \frac{1}{2} \log \left(2\pi e \text{Var}[Y_2 | U_2, U_1, X_2, V] \right) \\
&= \frac{1}{2} \log \left(2\pi e(\sigma_2^2 + h_{2,2}^2 \lambda_{2p} P_2 + h_{2,1}^2 (1 - \lambda_{1e}) P_1) \right) \\
&\quad - \frac{1}{2} \log \left(2\pi e(\sigma_2^2 + h_{2,1}^2 \lambda_{1p} P_1) \right) \\
&= \frac{1}{2} \log \left(\frac{1 + \lambda_{2p} \text{SNR}_2 + (1 - \lambda_{1e}) \text{INR}_2}{1 + \lambda_{1p} \text{INR}_1} \right), \tag{A.33e}
\end{aligned}$$

which proves (5.14a), (5.14c) and (5.14e). Finally, using (A.33c), (A.33d) and (A.33e), the proof of (5.14f) follows immediately .

A.0.2 Proof of (5.14h) The choice of the channel input in (A.32) guarantees that the random variables $Y_{3,1}, \dots, Y_{3,n}$ are independently and identically distributed. For all $n \in \{1, 2, \dots, N\}$, $Y_{3,n}$ follows a zero-mean Gaussian distribution with variance \bar{B} given by

$$\begin{aligned}
\bar{B} &= \mathbb{E} [Y_{3,n}^2] \\
&= \mathbb{E} [(h_{3,1} X_{1,n} + h_{3,2} X_{2,n} + Z_{3,n})^2] \\
&= h_{3,1}^2 \mathbb{E} [X_{1,n}^2] + h_{3,2}^2 \mathbb{E} [X_{2,n}^2] + 2h_{3,1} h_{3,2} \mathbb{E} [X_{1,n} X_{2,n}] + \sigma_3^2 \\
&= h_{3,1}^2 P_1 + h_{3,2}^2 P_2 + 2h_{3,1} h_{3,2} \mathbb{E} [X_{1,n} X_{2,n}] + \sigma_3^2
\end{aligned}$$

$$\leq h_{3,1}^2 P_1 + h_{3,2}^2 P_2 + 2h_{3,1}h_{3,2}\sqrt{\lambda_{1e}P_1\lambda_{2e}P_2} + \sigma_3^2.$$

By the weak law of large numbers, it holds that

$$\lim_{N \rightarrow \infty} \Pr(|B^{(N)} - \bar{B}| > 0) = 0. \quad (\text{A.34})$$

Consequently,

$$\lim_{N \rightarrow \infty} \Pr(B^{(N)} > \bar{B}) = 0 \quad \text{and} \quad (\text{A.35a})$$

$$\lim_{N \rightarrow \infty} \Pr(B^{(N)} < \bar{B}) = 0. \quad (\text{A.35b})$$

From (A.35b), it holds that for any energy rate B which satisfies $0 < B \leq \bar{B}$

$$\lim_{N \rightarrow \infty} \Pr(B^{(N)} < B) = 0. \quad (\text{A.36})$$

This proves (5.14h) and completes the proof of Theorem 12.

B

Proof of Theorem 13

Fix an information-energy rate triplet (R_1, R_2, B) achievable with a given coding scheme (Definition 8). Denote by \mathbf{X}_1 and \mathbf{X}_2 the channel inputs resulting from transmitting the independent message W_1 and W_2 using such a coding scheme. Denote by \mathbf{Y}_1 and \mathbf{Y}_2 the corresponding channel outputs. The bounds (5.15a) and (5.15b) 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:

$$NR_1 \leq \sum_{n=1}^N h(Y_{1,n}|X_{2,n}) - Nh(Z_1) + o(N) \text{ and} \quad (\text{B.1})$$

$$NR_2 \leq \sum_{n=1}^N h(Y_{2,n}|X_{1,n}) - Nh(Z_2) + o(N), \quad (\text{B.2})$$

where $\frac{o(N)}{N}$ tends to zero as N tends to infinity. Define the following random variables:

$$T_1 = h_{2,1}X_{1,n} + Z_{2,n}, \quad (\text{B.3})$$

$$T_2 = h_{1,2}X_{2,n} + Z_{1,n}, \quad (\text{B.4})$$

$$U_1 = h_{2,1}X_{1,n} + Z'_{2,n}, \text{ and} \quad (\text{B.5})$$

$$U_2 = h_{1,2}X_{2,n} + Z'_{1,n}, \quad (\text{B.6})$$

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) . Let \mathbf{T}_1 and \mathbf{T}_2 be the vectors:

$$\mathbf{T}_1 = (T_{1,1}, T_{1,2}, \dots, T_{1,N}) \text{ and} \quad (\text{B.7})$$

$$\mathbf{T}_2 = (T_{2,1}, T_{2,2}, \dots, T_{2,N}). \quad (\text{B.8})$$

The outer bound is established by using a genie aided argument. For all $j \in \{1, 2\}$, by Fano's inequality, it follows that

$$\begin{aligned} NR_j &= H(M_j) \\ &\leq I(M_j; \mathbf{Y}_j) + o(N) \end{aligned}$$

$$\leq I(\mathbf{X}_j; \mathbf{Y}_j) + o(N). \quad (\text{B.9})$$

Using the definition of mutual information, yields:

$$I(\mathbf{X}_1; \mathbf{Y}_1) = h(\mathbf{Y}_1) - h(\mathbf{Y}_1 | \mathbf{X}_1) \quad (\text{B.10a})$$

$$= h(\mathbf{Y}_1) - h(\mathbf{T}_2 | \mathbf{X}_1) \quad (\text{B.10b})$$

$$= h(\mathbf{Y}_1) - h(\mathbf{T}_2) \quad (\text{B.10c})$$

$$\leq \sum_{n=1}^N h(Y_{1,n}) - h(\mathbf{T}_2). \quad (\text{B.10d})$$

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

$$I(\mathbf{X}_1; \mathbf{Y}_1) \leq I(\mathbf{X}_1; \mathbf{Y}_1, \mathbf{U}_1, \mathbf{X}_2) \quad (\text{B.11a})$$

$$= I(\mathbf{X}_1; \mathbf{U}_1) + I(\mathbf{X}_1; \mathbf{X}_2 | \mathbf{U}_1) + I(\mathbf{X}_1; \mathbf{Y}_1 | \mathbf{U}_1, \mathbf{X}_2) \quad (\text{B.11b})$$

$$= h(\mathbf{U}_1) - h(\mathbf{U}_1 | \mathbf{X}_1) + h(\mathbf{Y}_1 | \mathbf{U}_1, \mathbf{X}_2) - h(\mathbf{Y}_1 | \mathbf{X}_1, \mathbf{U}_1, \mathbf{X}_2) \quad (\text{B.11c})$$

$$= h(\mathbf{T}_1) - h(\mathbf{U}_1 | \mathbf{X}_1) + h(\mathbf{Y}_1 | \mathbf{U}_1, \mathbf{X}_2) - h(\mathbf{Y}_1 | \mathbf{X}_1, \mathbf{U}_1, \mathbf{X}_2) \quad (\text{B.11d})$$

$$= h(\mathbf{T}_1) - h(\mathbf{U}_1 | \mathbf{X}_1) + h(\mathbf{Y}_1 | \mathbf{U}_1, \mathbf{X}_2) - h(\mathbf{T}_2 | \mathbf{X}_2) \quad (\text{B.11e})$$

$$\leq h(\mathbf{T}_1) - h(\mathbf{Z}'_2) + \sum_{n=1}^N h(Y_{1,n} | U_{1,n}, X_{2,n}) - h(\mathbf{Z}_1) \quad (\text{B.11f})$$

$$= h(\mathbf{T}_1) - Nh(\mathbf{Z}'_2) + \sum_{n=1}^N h(Y_{1,n} | U_{1,n}, X_{2,n}) - Nh(\mathbf{Z}_{1,n}), \quad (\text{B.11g})$$

where (B.11f) follows from the fact that conditioning does not increase the entropy.

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

$$I(\mathbf{X}_1; \mathbf{Y}_1) \leq I(\mathbf{X}_1; \mathbf{Y}_1, \mathbf{U}_1) \quad (\text{B.12a})$$

$$= I(\mathbf{X}_1; \mathbf{U}_1) + I(\mathbf{X}_1; \mathbf{Y}_1 | \mathbf{U}_1) \quad (\text{B.12b})$$

$$= h(\mathbf{U}_1) - h(\mathbf{U}_1 | \mathbf{X}_1) + h(\mathbf{Y}_1 | \mathbf{U}_1) - h(\mathbf{Y}_1 | \mathbf{X}_1, \mathbf{U}_1) \quad (\text{B.12c})$$

$$= h(\mathbf{T}_1) - h(\mathbf{U}_1 | \mathbf{X}_1) + h(\mathbf{Y}_1 | \mathbf{U}_1) - h(\mathbf{T}_2) \quad (\text{B.12d})$$

$$\leq h(\mathbf{T}_1) - h(\mathbf{T}_2) - h(\mathbf{Z}'_2) + \sum_{i=1}^N h(Y_{1,n} | U_{1,n}) \quad (\text{B.12e})$$

$$= h(\mathbf{T}_1) - h(\mathbf{T}_2) - Nh(\mathbf{Z}'_2) + \sum_{n=1}^N h(Y_{1,n} | U_{1,n}). \quad (\text{B.12f})$$

where (B.12e) follows from the fact that conditioning does not increase the entropy.

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

$$I(\mathbf{X}_1; \mathbf{Y}_1) \leq I(\mathbf{X}_1; \mathbf{Y}_1, \mathbf{X}_2) \quad (\text{B.13a})$$

$$= I(\mathbf{X}_1; \mathbf{X}_2) + I(\mathbf{X}_1; \mathbf{Y}_1 | \mathbf{X}_2) \quad (\text{B.13b})$$

$$= h(\mathbf{Y}_1 | \mathbf{X}_2) - h(\mathbf{Y}_1 | \mathbf{X}_1, \mathbf{X}_2) \quad (\text{B.13c})$$

$$=h(\mathbf{Y}_1|\mathbf{X}_2) - h(\mathbf{Z}_1) \quad (\text{B.13d})$$

$$\leq \sum_{n=1}^N h(Y_{1,n}|X_{2,n}) - Nh(\mathbf{Z}_1). \quad (\text{B.13e})$$

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

$$I(\mathbf{X}_2; \mathbf{Y}_2) \leq \sum_{n=1}^N h(Y_{2,n}) - h(\mathbf{T}_1), \quad (\text{B.14a})$$

$$I(\mathbf{X}_2; \mathbf{Y}_2) \leq h(\mathbf{T}_2) - Nh(\mathbf{Z}'_1) + \sum_{n=1}^N h(Y_{2,n}|U_{2,n}, X_{1,n}) - Nh(\mathbf{Z}_1), \quad (\text{B.14b})$$

$$I(\mathbf{X}_2; \mathbf{Y}_2) \leq h(\mathbf{T}_2) - h(\mathbf{T}_1) - Nh(\mathbf{Z}'_1) + \sum_{n=1}^N h(Y_{2,n}|U_{2,n}), \quad (\text{B.14c})$$

$$I(\mathbf{X}_2; \mathbf{Y}_2) \leq \sum_{n=1}^N h(Y_{2,n}|X_{2,n}) - Nh(\mathbf{Z}_1). \quad (\text{B.14d})$$

The key idea of the proof is to consider a linear combination of the inequalities in (B.10d), (B.11g), (B.12f), (B.13e), (B.14a), (B.14b), (B.14c), and (B.14d), where all the terms on $h(\mathbf{T}_1)$ and $h(\mathbf{T}_2)$ are removed. Adding (B.10d) and (B.14b) and plugging into (B.9) yields the first bound on sum rate:

$$N(R_1 + R_2) \leq \sum_{n=1}^N h(Y_{1,n}) + \sum_{n=1}^N h(Y_{2,n}|U_{2,n}, X_{1,n}) - Nh(\mathbf{Z}_2) - Nh(\mathbf{Z}'_1) + o(N). \quad (\text{B.15a})$$

Adding (B.11g) and (B.14a) and plugging into (B.9) yields the second bound on sum rate:

$$N(R_1 + R_2) \leq \sum_{n=1}^N h(Y_{2,n}) + \sum_{n=1}^N h(Y_{1,n}|U_{1,n}, X_{2,n}) - Nh(\mathbf{Z}_1) - Nh(\mathbf{Z}'_2) + o(N). \quad (\text{B.15b})$$

Adding (B.12f) and (B.14c) and plugging into (B.9) yields the third bound on sum rate:

$$N(R_1 + R_2) \leq \sum_{n=1}^N h(Y_{1,n}|U_{1,n}) + \sum_{n=1}^N h(Y_{2,n}|U_{2,n}) - Nh(\mathbf{Z}'_1) - Nh(\mathbf{Z}'_2) + o(N). \quad (\text{B.15c})$$

Adding (B.10d), (B.11g) and (B.14c) and plugging into (B.9) yields the first bound for the weighted sum rate:

$$\begin{aligned} N(2R_1 + R_2) &\leq \sum_{n=1}^N h(Y_{1,n}) + \sum_{n=1}^N h(Y_{1,n}|U_{1,n}, X_{2,n}) + \sum_{n=1}^N h(Y_{2,n}|U_{2,n}) \\ &\quad - N \left(h(\mathbf{Z}_1) + h(\mathbf{Z}_2) + h(\mathbf{Z}'_1) + h(\mathbf{Z}'_2) \right) + o(N). \end{aligned} \quad (\text{B.15d})$$

Adding (B.14a), (B.12f) and (B.14b) and plugging into (B.9) yields the second bound for the weighted sum rate :

$$N(R_1 + 2R_2) \leq \sum_{n=1}^N h(Y_{2,n}) + \sum_{n=1}^N h(Y_{2,n}|U_{2,n}, X_{1,n}) + \sum_{n=1}^N h(Y_{1,n}|U_{1,n}) - N(h(Z_1) + h(Z_2) + h(Z'_1) + h(Z'_2)) + o(N). \quad (\text{B.15e})$$

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

$$NR_1 \leq \sum_{n=1}^N h(Y_{1,n}|X_{2,n}) - Nh(Z_1) + o(N), \quad (\text{B.16a})$$

$$NR_2 \leq \sum_{n=1}^N h(Y_{2,n}|X_{1,n}) - Nh(Z_2) + o(N), \quad (\text{B.16b})$$

$$N(R_1 + R_2) \leq \sum_{n=1}^N h(Y_{1,n}) + \sum_{n=1}^N h(Y_{2,n}|U_{2,n}, X_{1,n}) - Nh(Z_2) - Nh(Z'_1) + o(N), \quad (\text{B.16c})$$

$$N(R_1 + R_2) \leq \sum_{n=1}^N h(Y_{2,n}) + \sum_{n=1}^N h(Y_{1,n}|U_{1,n}, X_{2,n}) - Nh(Z_1) - Nh(Z'_2) + o(N), \quad (\text{B.16d})$$

$$N(R_1 + R_2) \leq \sum_{n=1}^N h(Y_{1,n}|U_{1,n}) + \sum_{n=1}^N h(Y_{2,n}|U_{2,n}) - Nh(Z'_1) - Nh(Z'_2) + o(N), \quad (\text{B.16e})$$

$$N(2R_1 + R_2) \leq \sum_{n=1}^N h(Y_{1,n}) + \sum_{n=1}^N h(Y_{1,n}|U_{1,n}, X_{2,n}) + \sum_{n=1}^N h(Y_{2,n}|U_{2,n}) - N(h(Z_1) + h(Z_2) + h(Z'_1) + h(Z'_2)) + o(N), \quad (\text{B.16f})$$

$$N(R_1 + 2R_2) \leq \sum_{n=1}^N h(Y_{2,n}) + \sum_{n=1}^N h(Y_{2,n}|U_{2,n}, X_{1,n}) + \sum_{n=1}^N h(Y_{1,n}|U_{1,n}) - N(h(Z_1) + h(Z_2) + h(Z'_1) + h(Z'_2)) + o(N). \quad (\text{B.16g})$$

Using assumption (5.13b), for any $\eta > 0$ there exist $N_0(\eta)$ such that for any $N \geq N_0(\eta)$ it holds that

$$\Pr(B^{(N)} < B) < \eta. \quad (\text{B.17})$$

Equivalently,

$$\Pr(B^{(N)} \geq B) \geq 1 - \eta. \quad (\text{B.18})$$

From Markov's inequality [94], the following holds:

$$B\Pr(B^{(N)} \geq B) \leq \mathbb{E}[B^{(N)}]. \quad (\text{B.19})$$

Combining (B.18) and (B.19) yields

$$B(1 - \eta) \leq \mathbb{E}[B^{(N)}], \quad (\text{B.20})$$

which can be written as

$$(B - \delta_N) \leq \mathbb{E}[B^{(N)}], \quad (\text{B.21})$$

for some $\delta_N > 0$ (for sufficiently large N).

In the following, for all $n \in \mathbb{N}$, the bounds in (B.16) and (B.21) are evaluated assuming that the channel inputs $X_{1,n}$ and $X_{2,n}$ are arbitrary independent. Define the following parameters:

$$\mu_{i,n} \triangleq \mathbb{E}[X_{i,n}], \quad (\text{B.22})$$

$$\gamma_{i,n}^2 \triangleq \text{Var}[X_{i,n}], \quad (\text{B.23})$$

for $n \in \{1, \dots, N\}$ and for $i \in \{1, 2\}$. The input sequence must satisfy the input power constraint in (5.4), which can be written, for $i \in \{1, 2\}$, as

$$\frac{1}{n} \sum_{n=1}^N \mathbb{E}[X_{i,n}^2] = \left(\frac{1}{N} \sum_{n=1}^N \gamma_{i,n}^2 \right) + \left(\frac{1}{N} \sum_{n=1}^N \mu_{i,n}^2 \right) \leq P_i. \quad (\text{B.24})$$

Using these assumptions the following holds:

$$\begin{aligned} h(Y_{i,n}) &\leq \frac{1}{2} \log \left(2\pi e \text{Var}[Y_{i,n}] \right) \\ &= \frac{1}{2} \log \left(2\pi e (\sigma^2 + h_{i,i}^2 \gamma_{i,n}^2 + h_{i,j}^2 \gamma_{j,n}^2) \right), \end{aligned} \quad (\text{B.25a})$$

$$\begin{aligned} h(Y_{i,n}|U_{i,n}, X_{j,n}) &\leq \frac{1}{2} \log \left(2\pi e \text{Var}[Y_{i,n}|U_{i,n}, X_{j,n}] \right) \\ &\leq \frac{1}{2} \log \left(2\pi e |\text{Cov}[h_{i,i}X_{i,n} + Z_{i,n}, U_{i,n}]| \right) - \frac{1}{2} \log \left(2\pi e \text{Var}[U_{i,n}] \right) \\ &= \frac{1}{2} \log \left(\frac{\sigma_j^2 \gamma_{i,n}^2 h_{i,i}^2 + \sigma_i^2 \gamma_{i,n}^2 h_{j,i}^2 + \sigma_i^2 \sigma_j^2}{\gamma_{i,n}^2 h_{j,i}^2 + \sigma_j^2} \right) + \frac{1}{2} \log(2\pi e \sigma_i^2 \sigma_j^2) \\ &= \frac{1}{2} \log \left(1 + \frac{\frac{h_{i,i}^2 \gamma_{i,n}^2}{\sigma_i^2}}{1 + \frac{h_{j,i}^2 \gamma_{i,n}^2}{\sigma_j^2}} \right) + \frac{1}{2} \log(2\pi e \sigma_i^2 \sigma_j^2), \end{aligned} \quad (\text{B.25b})$$

$$\begin{aligned} h(Y_{i,n}|U_{i,n}) &\leq \frac{1}{2} \log \left(2\pi e \text{Var}[Y_{i,n}|U_{i,n}] \right) \\ &\leq \frac{1}{2} \log \left(2\pi e |\text{Cov}[Y_{i,n}, U_{i,n}]| \right) - \frac{1}{2} \log \left(2\pi e \text{Var}[U_{i,n}] \right) \\ &= \frac{1}{2} \log \left(\frac{\sigma_j^2 \gamma_i^2 h_{i,i}^2 + \sigma_i^2 \gamma_{i,n}^2 h_{j,i}^2 + h_{i,j}^2 h_{j,i}^2 \gamma_{i,n}^2 \gamma_{j,n}^2 + \sigma_i^2 \sigma_j^2}{\gamma_i^2 h_{j,i}^2 + \sigma_j^2} \right) + \frac{1}{2} \log(2\pi e \sigma_i^2 \sigma_j^2) \\ &\leq \frac{1}{2} \log \left(1 + \frac{\frac{h_{i,i}^2 \gamma_{i,n}^2}{\sigma_i^2} + \frac{h_{i,j}^2 \gamma_{j,n}^2}{\sigma_i^2} + \frac{\gamma_{i,n}^2 \gamma_{j,n}^2 h_{i,j}^2 h_{j,i}^2}{\sigma_i^2 \sigma_j^2}}{1 + \frac{\gamma_{i,n}^2 h_{j,i}^2}{\sigma_j^2}} \right) + \frac{1}{2} \log(2\pi e \sigma_i^2 \sigma_j^2), \end{aligned} \quad (\text{B.25c})$$

Finally, plugging (B.25) in (B.16), it yields:

$$NR_1 \leq \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\gamma_{1,n}^2 h_{1,1}^2}{\sigma_1^2} \right) + o(N), \quad (\text{B.26a})$$

$$NR_2 \leq \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\gamma_{2,n}^2 h_{2,2}^2}{\sigma_2^2} \right) + o(N), \quad (\text{B.26b})$$

$$N(R_1 + R_2) \leq \sum_{n=1}^N \log \left(1 + \frac{h_{1,1}^2 \gamma_{1,n}^2}{\sigma_1^2} + \frac{h_{1,2}^2 \gamma_{2,n}^2}{\sigma_1^2} \right) + \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\frac{h_{2,2}^2 \gamma_{2,n}^2}{\sigma_2^2}}{1 + \frac{h_{1,2}^2 \gamma_{2,n}^2}{\sigma_1^2}} \right) + o(N), \quad (\text{B.26c})$$

$$N(R_1 + R_2) \leq \sum_{n=1}^N \log \left(1 + \frac{h_{2,2}^2 \gamma_{2,n}^2}{\sigma_2^2} + \frac{h_{2,1}^2 \gamma_{1,n}^2}{\sigma_2^2} \right) + \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\frac{h_{1,1}^2 \gamma_{1,n}^2}{\sigma_1^2}}{1 + \frac{h_{2,1}^2 \gamma_{1,n}^2}{\sigma_2^2}} \right) + o(N), \quad (\text{B.26d})$$

$$N(R_1 + R_2) \leq \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\frac{h_{1,1}^2 \gamma_{1,n}^2}{\sigma_1^2} + \frac{h_{1,2}^2 \gamma_{2,n}^2}{\sigma_1^2} + \frac{\gamma_{1,n}^2 \gamma_{2,n}^2 h_{1,2}^2 h_{2,1}^2}{\sigma_1^2 \sigma_2^2}}{1 + \frac{\gamma_{1,n}^2 h_{2,1}^2}{\sigma_2^2}} \right) + \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\frac{h_{2,2}^2 \gamma_{2,n}^2}{\sigma_2^2} + \frac{h_{2,1}^2 \gamma_{1,n}^2}{\sigma_2^2} + \frac{\gamma_{1,n}^2 \gamma_{2,n}^2 h_{1,2}^2 h_{2,1}^2}{\sigma_1^2 \sigma_2^2}}{1 + \frac{\gamma_{2,n}^2 h_{1,2}^2}{\sigma_1^2}} \right) + o(N), \quad (\text{B.26e})$$

$$N(2R_1 + R_2) \leq \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\frac{h_{1,1}^2 \gamma_{1,n}^2}{\sigma_1^2}}{1 + \frac{h_{2,1}^2 \gamma_{1,n}^2}{\sigma_2^2}} \right) + \sum_{n=1}^N \log \left(1 + \frac{h_{1,1}^2 \gamma_{1,n}^2}{\sigma_1^2} + \frac{h_{1,2}^2 \gamma_{2,n}^2}{\sigma_1^2} \right) + \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\frac{h_{2,2}^2 \gamma_{2,n}^2}{\sigma_2^2} + \frac{h_{2,1}^2 \gamma_{1,n}^2}{\sigma_2^2} + \frac{\gamma_{1,n}^2 \gamma_{2,n}^2 h_{1,2}^2 h_{2,1}^2}{\sigma_1^2 \sigma_2^2}}{1 + \frac{\gamma_{2,n}^2 h_{1,2}^2}{\sigma_1^2}} \right) + o(N), \quad (\text{B.26f})$$

$$N(R_1 + 2R_2) \leq \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\frac{h_{2,2}^2 \gamma_{2,n}^2}{\sigma_2^2}}{1 + \frac{h_{1,2}^2 \gamma_{2,n}^2}{\sigma_1^2}} \right) + \sum_{n=1}^N \log \left(1 + \frac{h_{2,2}^2 \gamma_{2,n}^2}{\sigma_2^2} + \frac{h_{2,1}^2 \gamma_{1,n}^2}{\sigma_2^2} \right) + \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\frac{h_{1,1}^2 \gamma_{1,n}^2}{\sigma_1^2} + \frac{h_{1,2}^2 \gamma_{2,n}^2}{\sigma_1^2} + \frac{\gamma_{1,n}^2 \gamma_{2,n}^2 h_{1,2}^2 h_{2,1}^2}{\sigma_1^2 \sigma_2^2}}{1 + \frac{\gamma_{1,n}^2 h_{2,1}^2}{\sigma_2^2}} \right) + o(N). \quad (\text{B.26g})$$

The expectation of the average received energy rate is given by

$$\begin{aligned} \mathbb{E} [B^{(N)}] &= \mathbb{E} \left[\frac{1}{N} \sum_{n=1}^N Y_{3,n}^2 \right] \\ &= \sigma_3^2 + h_{3,1}^2 \left(\frac{1}{N} \sum_{n=1}^N (\gamma_{1,n}^2 + \mu_{1,n}^2) \right) + h_{3,2}^2 \left(\frac{1}{N} \sum_{n=1}^N (\gamma_{2,n}^2 + \mu_{2,n}^2) \right) \\ &\quad + 2h_{3,1}h_{3,2} \frac{1}{N} \sum_{n=1}^N \mu_{1,n}\mu_{2,n}. \end{aligned} \quad (\text{B.27})$$

Using Cauchy-Schwarz inequality, the energy rate in (B.27) can be upper-bounded as follows:

$$\begin{aligned} \mathbb{E} [B^{(N)}] &\leq \sigma_3^2 + h_{3,1}^2 \left(\frac{1}{N} \sum_{n=1}^N (\gamma_{1,n}^2 + \mu_{1,n}^2) \right) + h_{3,2}^2 \left(\frac{1}{N} \sum_{n=1}^N (\gamma_{2,n}^2 + \mu_{2,n}^2) \right) \\ &\quad + 2h_{3,1}h_{3,2} \left(\frac{1}{N} \sum_{n=1}^N \mu_{1,n}^2 \right)^{1/2} \left(\frac{1}{N} \sum_{n=1}^N \mu_{2,n}^2 \right)^{1/2}. \end{aligned} \quad (\text{B.28})$$

Combining (B.21) and (B.28) yields the following upper-bound on the energy rate B :

$$\begin{aligned} B &\leq \sigma_3^2 + h_{3,1}^2 \left(\frac{1}{N} \sum_{n=1}^N (\gamma_{1,n}^2 + \mu_{1,n}^2) \right) + h_{3,2}^2 \left(\frac{1}{N} \sum_{n=1}^N (\gamma_{2,n}^2 + \mu_{2,n}^2) \right) \\ &\quad + 2h_{3,1}h_{3,2} \left(\frac{1}{N} \sum_{n=1}^N \mu_{1,n}^2 \right)^{1/2} \left(\frac{1}{N} \sum_{n=1}^N \mu_{2,n}^2 \right)^{1/2} + \delta_N. \end{aligned} \quad (\text{B.29})$$

In order to obtain a single-letterization of the upper-bound given by constraints (B.26) and (B.29), define also

$$\mu_i^2 \triangleq \frac{1}{N} \sum_{n=1}^N \mu_{i,n}^2, \quad i \in \{1, 2\}, \quad (\text{B.30})$$

$$\gamma_i^2 \triangleq \frac{1}{N} \sum_{n=1}^N \gamma_{i,n}^2, \quad i \in \{1, 2\}, \quad (\text{B.31})$$

Using these notations, the input power constraint in (D.12) can be rewritten as

$$\gamma_i^2 + \mu_i^2 \leq P_i, \quad \text{with } i \in \{1, 2\}. \quad (\text{B.32})$$

By the concavity of the mutual information, applying Jensen's inequality to the bounds in (B.26) yields in the asymptotic blocklength regime:

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{\gamma_1^2 h_{1,1}^2}{\sigma_1^2} \right), \quad (\text{B.33a})$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{\gamma_2^2 h_{2,2}^2}{\sigma_2^2} \right), \quad (\text{B.33b})$$

$$R_1 + R_2 \leq \log \left(1 + \frac{h_{1,1}^2 \gamma_1^2}{\sigma_1^2} + \frac{h_{1,2}^2 \gamma_2^2}{\sigma_1^2} \right) + \frac{1}{2} \log \left(1 + \frac{\frac{h_{2,2}^2 \gamma_2^2}{\sigma_2^2}}{1 + \frac{h_{1,2}^2 \gamma_2^2}{\sigma_1^2}} \right), \quad (\text{B.33c})$$

$$R_1 + R_2 \leq \log \left(1 + \frac{h_{2,2}^2 \gamma_2^2}{\sigma_2^2} + \frac{h_{2,1}^2 \gamma_1^2}{\sigma_2^2} \right) + \frac{1}{2} \log \left(1 + \frac{\frac{h_{1,1}^2 \gamma_1^2}{\sigma_1^2}}{1 + \frac{h_{2,1}^2 \gamma_1^2}{\sigma_2^2}} \right), \quad (\text{B.33d})$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{\frac{h_{1,1}^2 \gamma_1^2}{\sigma_1^2} + \frac{h_{1,2}^2 \gamma_2^2}{\sigma_1^2} + \frac{\gamma_1^2 \gamma_2^2 h_{1,2}^2 h_{2,1}^2}{\sigma_1^2 \sigma_2^2}}{1 + \frac{\gamma_1^2 h_{2,1}^2}{\sigma_2^2}} \right) \quad (\text{B.33e})$$

$$\begin{aligned}
 & + \frac{1}{2} \log \left(1 + \frac{\frac{h_{2,2}^2 \gamma_2^2}{\sigma_2^2} + \frac{h_{2,1}^2 \gamma_1^2}{\sigma_2^2} + \frac{\gamma_1^2 \gamma_2^2 h_{1,2}^2 h_{2,1}^2}{\sigma_1^2 \sigma_2^2}}{1 + \frac{\gamma_2^2 h_{1,2}^2}{\sigma_1^2}} \right), \\
 2R_1 + R_2 & \leq \frac{1}{2} \log \left(1 + \frac{\frac{h_{1,1}^2 \gamma_1^2}{\sigma_1^2}}{1 + \frac{h_{2,1}^2 \gamma_1^2}{\sigma_2^2}} \right) + \log \left(1 + \frac{h_{1,1}^2 \gamma_1^2}{\sigma_1^2} + \frac{h_{1,2}^2 \gamma_2^2}{\sigma_1^2} \right) \quad (\text{B.33f})
 \end{aligned}$$

$$\begin{aligned}
 & + \frac{1}{2} \log \left(1 + \frac{\frac{h_{2,2}^2 \gamma_2^2}{\sigma_2^2} + \frac{h_{2,1}^2 \gamma_1^2}{\sigma_2^2} + \frac{\gamma_1^2 \gamma_2^2 h_{1,2}^2 h_{2,1}^2}{\sigma_1^2 \sigma_2^2}}{1 + \frac{\gamma_2^2 h_{1,2}^2}{\sigma_1^2}} \right), \\
 R_1 + 2R_2 & \leq \frac{1}{2} \log \left(1 + \frac{\frac{h_{2,2}^2 \gamma_2^2}{\sigma_2^2}}{1 + \frac{h_{1,2}^2 \gamma_2^2}{\sigma_1^2}} \right) + \log \left(1 + \frac{h_{2,2}^2 \gamma_2^2}{\sigma_2^2} + \frac{h_{2,1}^2 \gamma_1^2}{\sigma_2^2} \right) \quad (\text{B.33g}) \\
 & + \frac{1}{2} \log \left(1 + \frac{\frac{h_{1,1}^2 \gamma_1^2}{\sigma_1^2} + \frac{h_{1,2}^2 \gamma_2^2}{\sigma_1^2} + \frac{\gamma_1^2 \gamma_2^2 h_{1,2}^2 h_{2,1}^2}{\sigma_1^2 \sigma_2^2}}{1 + \frac{\gamma_1^2 h_{2,1}^2}{\sigma_2^2}} \right),
 \end{aligned}$$

and the upper-bound on the energy rate (B.29) yields

$$B \leq \sigma_3^2 + h_{3,1}^2(\gamma_1^2 + \mu_1^2) + h_{3,2}^2(\gamma_2^2 + \mu_2^2) + 2h_{3,1}h_{3,2}|\mu_1||\mu_2|. \quad (\text{B.33h})$$

To sum up, it has been shown so far that, in the limit when N tends to infinity, any information-energy rate triplet $(R_1, R_2, B) \in \mathcal{E}$ can be bounded by the constraints in (B.33) for some $\gamma_1, \gamma_2, \mu_1, \mu_2$ satisfying (B.32). Let $\mathcal{R}(\gamma_1, \gamma_2, \mu_1, \mu_2)$ denote the set of information-energy rate triplets satisfying (B.33) for some $\gamma_1, \gamma_2, \mu_1, \mu_2$ such that (B.32) is true. Thus, it holds that

$$\bar{\mathcal{E}} \subseteq \bigcup_{\substack{0 \leq \gamma_1^2 + \mu_1^2 \leq P_1 \\ 0 \leq \gamma_2^2 + \mu_2^2 \leq P_2}} \mathcal{R}(\gamma_1, \gamma_2, \mu_1, \mu_2). \quad (\text{B.34})$$

In this union, it suffices to consider $\mu_1 \geq 0, \mu_2 \geq 0$, and $\gamma_1, \gamma_2, \mu_1$, and μ_2 that saturate the input power constraint (i.e., (B.32) holds with equality). Thus,

$$\bar{\mathcal{E}} \subseteq \bigcup_{\substack{0 \leq \gamma_1^2 + \mu_1^2 \leq P_1 \\ 0 \leq \gamma_2^2 + \mu_2^2 \leq P_2}} \mathcal{R}(\gamma_1^2, \gamma_2^2, \mu_1, \mu_2) \subseteq \bigcup_{\substack{\sigma_1^2 + \mu_1^2 = P_1 \\ \gamma_2^2 + \mu_2^2 = P_2}} \mathcal{R}(\gamma_1^2, \gamma_2^2, \mu_1, \mu_2). \quad (\text{B.35})$$

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

$$\beta_i \triangleq \frac{\gamma_i^2}{P_i} = \frac{P_i - \mu_i^2}{P_i}, \quad i \in \{1, 2\}. \quad (\text{B.36})$$

With these notations, any region $\mathcal{R}(\gamma_1^2, \gamma_2^2, \mu_1, \mu_2)$ 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:

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{\beta_1 P_1 h_{1,1}^2}{\sigma_1^2} \right), \quad (\text{B.37a})$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{\beta_2 P_2 h_{2,2}^2}{\sigma_2^2} \right), \quad (\text{B.37b})$$

$$R_1 + R_2 \leq \log \left(1 + \frac{h_{1,1}^2 \beta_1 P_1}{\sigma_1^2} + \frac{h_{1,2}^2 \beta_2 P_2}{\sigma_1^2} \right) + \frac{1}{2} \log \left(1 + \frac{\frac{h_{2,2}^2 \beta_2 P_2}{\sigma_2^2}}{1 + \frac{h_{1,2}^2 \beta_2 P_2}{\sigma_1^2}} \right), \quad (\text{B.37c})$$

$$R_1 + R_2 \leq \log \left(1 + \frac{h_{2,2}^2 \beta_2 P_2}{\sigma_2^2} + \frac{h_{2,1}^2 \beta_1 P_1}{\sigma_2^2} \right) + \frac{1}{2} \log \left(1 + \frac{\frac{h_{1,1}^2 \beta_1 P_1}{\sigma_1^2}}{1 + \frac{h_{2,1}^2 \beta_1 P_1}{\sigma_2^2}} \right), \quad (\text{B.37d})$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{\frac{h_{1,1}^2 \beta_1 P_1}{\sigma_1^2} + \frac{h_{1,2}^2 \beta_2 P_2}{\sigma_1^2} + \frac{\beta_1 \beta_2 P_1 P_2 h_{1,2}^2 h_{2,1}^2}{\sigma_1^2 \sigma_2^2}}{1 + \frac{\beta_1 P_1 h_{2,1}^2}{\sigma_2^2}} \right) + \frac{1}{2} \log \left(1 + \frac{\frac{h_{2,2}^2 \beta_2 P_2}{\sigma_2^2} + \frac{h_{2,1}^2 \beta_1 P_1}{\sigma_2^2} + \frac{\beta_1 \beta_2 P_1 P_2 h_{1,2}^2 h_{2,1}^2}{\sigma_1^2 \sigma_2^2}}{1 + \frac{\beta_2 P_2 h_{1,2}^2}{\sigma_1^2}} \right), \quad (\text{B.37e})$$

$$2R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{\frac{h_{1,1}^2 \beta_1 P_1}{\sigma_1^2}}{1 + \frac{h_{2,1}^2 \beta_1 P_1}{\sigma_2^2}} \right) + \log \left(1 + \frac{h_{1,1}^2 \beta_1 P_1}{\sigma_1^2} + \frac{h_{1,2}^2 \beta_2 P_2}{\sigma_1^2} \right) + \frac{1}{2} \log \left(1 + \frac{\frac{h_{2,2}^2 \beta_2 P_2}{\sigma_2^2} + \frac{h_{2,1}^2 \beta_1 P_1}{\sigma_2^2} + \frac{\beta_1 \beta_2 P_1 P_2 h_{1,2}^2 h_{2,1}^2}{\sigma_1^2 \sigma_2^2}}{1 + \frac{\beta_2 P_2 h_{1,2}^2}{\sigma_1^2}} \right), \quad (\text{B.37f})$$

$$R_1 + 2R_2 \leq \frac{1}{2} \log \left(1 + \frac{\frac{h_{2,2}^2 \beta_2 P_2}{\sigma_2^2}}{1 + \frac{h_{1,2}^2 \beta_2 P_2}{\sigma_1^2}} \right) + \log \left(1 + \frac{h_{2,2}^2 \beta_2 P_2}{\sigma_2^2} + \frac{h_{2,1}^2 \beta_1 P_1}{\sigma_2^2} \right) + \frac{1}{2} \log \left(1 + \frac{\frac{h_{1,1}^2 \beta_1 P_1}{\sigma_1^2} + \frac{h_{1,2}^2 \beta_2 P_2}{\sigma_1^2} + \frac{\beta_1 \beta_2 P_1 P_2 h_{1,2}^2 h_{2,1}^2}{\sigma_1^2 \sigma_2^2}}{1 + \frac{\beta_1 P_1 h_{2,1}^2}{\sigma_2^2}} \right), \quad (\text{B.37g})$$

$$B \leq \sigma_3^2 + h_{3,1}^2 P_1 + h_{3,2}^2 P_2 + 2h_{3,1} h_{3,2} \sqrt{(1 - \beta_1) P_1 (1 - \beta_2) P_2}, \quad (\text{B.37h})$$

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 13 and this completes the proof of Theorem 13.

C

Proof of Theorem 15

This proof is divided into two parts. The first part consists of the proof of (5.17a)-(5.17f). The second part consists of the proof of (5.17g).

C.0.1 Proof of (5.17a)-(5.17f) Codebook Generation: Fix a strictly positive joint probability distribution:

$P_{VU_1U_2S_1S_2}(v, u, u_1, u_2, s_1, s_2) = P_V(v)P_{U|V}(u|v)P_{U_1|UV}(u_1|u, v)P_{U_2|U,V}(u_2|u, v)P_{S_1|UU_1V}(s_1|u, u_1, v)P_{S_2|UU_2V}(s_2|u, u_2, v)$, for all $(v, u, u_1, u_2, x_1, x_2) \in \mathcal{X}_1 \cap \mathcal{X}_2 \times (\mathcal{X}_1 \cap \mathcal{X}_2) \times (\mathcal{X}_1 \times \mathcal{X}_2)$. 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 $\mathbf{v}(\omega) = (v_1(\omega), \dots, v_N(\omega))$ according to

$$P_V(\mathbf{v}(\omega)) = \prod_{m=1}^N P_V(v_m(\omega)), \quad (\text{C.1})$$

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

$$P_{U|V}(\mathbf{u}(\omega, s, r)|\mathbf{v}(\omega)) = \prod_{m=1}^N P_{U|V}(u_m(\omega, s, r)|v_m(\omega)), \quad (\text{C.2})$$

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

$$P_{U_1|UV}(\mathbf{u}_1(\omega, s, r, k)|\mathbf{u}(s, r), \mathbf{v}(\omega)) = \prod_{m=1}^N P_{U_1|UV}(u_{1,m}(\omega, s, r, k)|u_m(\omega, s, r), v_m(\omega)), \quad (\text{C.3})$$

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

according to

$$P_{\mathbf{S}_1|U_1UV} \left(\mathbf{s}_1(\omega, s, r, k, l) | \mathbf{u}_1(\omega, s, r, k), \mathbf{u}(s, r), \mathbf{v}(\omega) \right) = \prod_{m=1}^N P_{\mathbf{S}_1|U_1UV} \left(s_{1,m}(\omega, s, r, k, l) | u_{1,m}(\omega, s, r, k), u_m(\omega, s, r), v_m(\omega) \right), \quad (\text{C.4})$$

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

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

$$P_{U_2|UV}(\mathbf{u}_2(\Omega, s, r, q) | \mathbf{u}(s, r), \mathbf{v}(\omega)) = \prod_{m=1}^N P_{U_2|UV} \left(u_{2,m}(\omega, s, r, q) | u_m(\omega, s, r), v_m(\omega) \right), \quad (\text{C.5})$$

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

$$P_{\mathbf{S}_2|U_2UV}(\mathbf{s}_2(\omega, s, r, q, z) | \mathbf{u}_2(\omega, s, r, q), \mathbf{u}(\omega, s, r), \mathbf{v}(\omega)) = \prod_{m=1}^N P_{\mathbf{S}_2|U_2UV} \left(s_{2,m}(\omega, s, r, q, z) | u_{2,m}(\omega, s, r, q), u_m(\omega, s, r), v_m(\omega) \right), \quad (\text{C.6})$$

with $z \in \{1, \dots, 2^{NR_{2,P}}\}$.

Encoding: Consider Markov encoding over T blocks. Let $W_i^{(t)}$ be the message index transmitted during block t . Let i also be represented by the message index $W_{i,C}^{(t)} \in \{1, 2, \dots, 2^{NR_{i,C}}\}$ and the message index $W_{i,P}^{(t)} \in \{1, 2, \dots, 2^{NR_{i,P}}\}$. The message index $\Omega^{(t)}$ is known by both transmitters and receivers. At encoding step t , with $t \in \{1, 2, \dots, T\}$, transmitter 1 sends the codeword $\mathbf{x}_1^{(t)} = \theta_1 \left(\mathbf{v}^{(t)}(\Omega^{(t)}), \mathbf{u} \left(\Omega^{(t)}, W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)} \right), \mathbf{u}_1 \left(\Omega^{(t)}, W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}, W_{1,C}^{(t)} \right), \right.$

$\left. \mathbf{s}_1 \left(\Omega^{(t)}, W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}, W_{1,C}^{(t)}, W_{1,P}^{(t)} \right) \right)$, where $\theta_1 : \mathcal{X}^N \times (\mathcal{X}_1 \cup \mathcal{X}_2)^N \times \mathcal{X}_1 \times \mathcal{X}_1^N \rightarrow \mathcal{X}_1^N$ is

a function that transforms the codewords $\mathbf{v}^{(t)}(\Omega)$, $\mathbf{u} \left(\Omega^{(t)}, W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)} \right)$, $\mathbf{u}_1 \left(\Omega^{(t)}, W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}, W_{1,C}^{(t)} \right)$, and $\mathbf{s}_1 \left(\Omega^{(t)}, W_{1,C}^{(t-1)}, W_{2,C}^{(t-1)}, W_{1,C}^{(t)}, W_{1,P}^{(t)} \right)$ into the N -dimensional vector $\mathbf{x}_1^{(t)}$.

The indices $W_{1,C}^{(0)} = W_{1,C}^{(T)} = s^*$ and $W_{2,C}^{(0)} = W_{2,C}^{(T)} = r^*$, and the pair $(s^*, r^*) \in \{1, 2, \dots, 2^{NR_{1,C}}\} \times \{1, 2, \dots, 2^{NR_{2,C}}\}$ 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 $t \in \{1, 2, \dots, T\}$, receiver 1 obtains the indices $\left(\widehat{W}_{1,C}^{(T-t)}, \widehat{W}_{2,C}^{(T-t)}, \widehat{W}_{1,P}^{(T-(t-1))} \right)$ from the given channel output \mathbf{y}_1 .

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

$$\left(\mathbf{v}(\Omega^{(t)}), \mathbf{u} \left(\Omega^{(t)}, \widehat{W}_{1,C}^{(T-t)}, \widehat{W}_{2,C}^{(T-t)} \right), \mathbf{u}_1 \left(\Omega^{(t)}, \widehat{W}_{1,C}^{(T-t)}, \widehat{W}_{2,C}^{(T-t)}, W_{1,C}^{(T-(t-1))} \right) \right),$$

$$\begin{aligned} & \mathbf{s}_1 \left(\Omega^{(t)}, \widehat{W}_{1,C}^{(T-t)}, \widehat{W}_{2,C}^{(T-t)}, W_{1,C}^{(T-(t-1))}, W_{1,P}^{(T-(t-1))} \right), \mathbf{u}_2 \left(\Omega^{(t)}, \widehat{W}_{1,C}^{(T-t)}, \widehat{W}_{2,C}^{(T-t)}, W_{2,C}^{(T-(t-1))} \right), \\ & \mathbf{y}_1^{(T-(t-1))} \Big) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}, \end{aligned} \quad (\text{C.7})$$

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 $t > 2$ at transmitter 1 if the index $\widehat{W}_{2,C}^{(t-1)}$ is not correctly decoded. Without any loss of generality, let $W_{2,C}^{(t-1)} = 1$ and $\widehat{W}_{2,C}^{(t-1)} = 1$. Define the event E_k that describes the case in which there exists another message index $k \neq 1$ that satisfies: $\left(\mathbf{v}(\Omega^{(t)}), \mathbf{u} \left(\Omega^{(t)}, W_{1,C}^{(t-2)}, W_{2,C}^{(t-2)} \right), \mathbf{u}_1 \left(\Omega^{(t)}, W_{1,C}^{(t-2)}, W_{2,C}^{(t-2)}, W_{1,C}^{(t-1)} \right), \mathbf{s}_1 \left(\Omega^{(t)}, W_{1,C}^{(t-2)}, W_{2,C}^{(t-2)}, W_{1,C}^{(t-1)}, W_{1,P}^{(t-1)} \right), \mathbf{u}_2 \left(\Omega^{(t)}, W_{1,C}^{(t-2)}, W_{2,C}^{(t-2)}, k \right) \right) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}$, with $t \in \{2, 3, \dots, T\}$ 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:

$$\Pr(E_k) = \Pr \left[(\mathbf{V}, \mathbf{U}, \mathbf{U}_1, \mathbf{S}_1, \mathbf{U}_2, \mathbf{Y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)} \right] \quad (\text{C.8a})$$

$$= \sum_{(\mathbf{v}, \mathbf{u}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}} P_{\mathbf{V}}(\mathbf{v}) P_{\mathbf{U}, \mathbf{U}_1, \mathbf{S}_1, \mathbf{Y}_1 | \mathbf{V}}(\mathbf{u}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{y}_1 | \mathbf{v}) P_{\mathbf{U}_2 | \mathbf{V}}(\mathbf{u}_2 | \mathbf{v}), \quad (\text{C.8b})$$

$$\leq \left| \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)} \right| 2^{-N(H(U, U_1, S_1, Y_1 | V) + H(U_2 | V) + H(V) - 4\epsilon)} \quad (\text{C.8c})$$

$$\leq 2^{-N(H(U, U_1, S_1, Y_1 | V) + H(U_2 | V) + H(V) - H(U, U_1, U_2, S_1, Y_1) - 4\epsilon)} \quad (\text{C.8d})$$

$$= 2^{-N(H(U, U_1, S_1, Y_1 | V) + H(U_2 | V) - H(U, U_1, U_2, S_1, Y_1 | V) - 4\epsilon)} \quad (\text{C.8e})$$

$$= 2^{-N(I(U, U_1, S_1, Y_1; U_2 | V) - 4\epsilon)} \quad (\text{C.8f})$$

$$= 2^{-N(I(U, X_1, Y_1; U_2 | V) - 4\epsilon)} \quad (\text{C.8g})$$

$$= 2^{-N(I(Y_1; U_2 | X_1, V, U) - 4\epsilon)}, \quad (\text{C.8h})$$

where (C.8f) follows from (A.9); and (C.8h) follows from the fact that $X_i = \theta_i(U, U_i, S_i, V)$. Note that the probability operator $\Pr[\cdot]$ in (C.8a) applies with a probability distribution $P_{VUU_1X_1U_2Y_1}$ that factorizes as $P_{\mathbf{V}}P_{\mathbf{U}, \mathbf{U}_1, \mathbf{X}_1, \mathbf{Y}_1 | \mathbf{V}}P_{\mathbf{U}_2 | \mathbf{V}}$ given that all the codewords \mathbf{u}_2 are independent from the given channel output \mathbf{y}_1 . The error probability becomes arbitrarily small (as N goes to infinity) if

$$R_{2c} \leq I(U_2; Y_1 | X_1, U, V). \quad (\text{C.9})$$

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 (C.7), 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 (C.7). From the

asymptotic equipartition property [93], 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 $\mathbf{v}(\Omega^{(t)})$, $\mathbf{u}(\Omega^{(t)}, s, r)$, $\mathbf{u}_1(\Omega^{(t)}, s, r, W_{1,C}^{(T-(t-1))})$, $\mathbf{s}_1(\Omega^{(t)}, s, r, W_{1,C}^{(T-(t-1))}, l)$, and $\mathbf{u}_2(\Omega^{(t)}, s, r, W_{2,C}^{(T-(t-1))})$ are jointly typical with $\mathbf{y}_1^{(T-(t-1))}$ 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:

$$\begin{aligned}
& \Pr\left(\bigcup_{(s,r,l) \neq (1,1,1)} E_{srl}\right) \\
& \leq \sum_{s \neq 1, r \neq 1, l \neq 1} \Pr(E_{srl}) + \sum_{s \neq 1, r \neq 1, l=1} \Pr(E_{srl}) + \sum_{s \neq 1, r=1, l \neq 1} \Pr(E_{srl}) \\
& \quad + \sum_{s \neq 1, r=1, l=1} \Pr(E_{srl}) + \sum_{s=1, r \neq 1, l \neq 1} \Pr(E_{srl}) + \sum_{s=1, r \neq 1, l=1} \Pr(E_{srl}) \\
& \quad + \sum_{s=1, r=1, l \neq 1} \Pr(E_{srl}), \tag{C.10}
\end{aligned}$$

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

$$\Pr[E_{s11}] = \Pr\left[(\mathbf{V}, \mathbf{U}, \mathbf{U}_1, \mathbf{S}_1, \mathbf{U}_2, \mathbf{Y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}\right] \tag{C.11a}$$

$$= \sum_{(\mathbf{v}, \mathbf{u}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2, \mathbf{y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}} P_V(\mathbf{v}) P_{U,U_1,S_1,U_2|V}(\mathbf{u}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2|\mathbf{v}) P_{Y_1|V}(\mathbf{y}_1|\mathbf{v}), \tag{C.11b}$$

$$\leq \left| \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)} \right| 2^{-N(H(U,U_1,U_2,S_1|V) + H(Y_1|V) + H(V) - 4\epsilon)} \tag{C.11c}$$

$$\leq 2^{-N(H(U,U_1,U_2,S_1|V) + H(Y_1|V) + H(V) - H(U,U_1,U_2,S_1,Y_1,V) - 4\epsilon)} \tag{C.11d}$$

$$= 2^{-N(H(U,U_1,U_2,S_1|V) + H(Y_1|V) - H(U,U_1,U_2,S_1,Y_1|V) - 4\epsilon)} \tag{C.11e}$$

$$= 2^{-N(I(U,U_1,U_2,S_1;Y_1|V) - 4\epsilon)} \tag{C.11f}$$

$$= 2^{-N(I(U,U_2,X_1;Y_1|V) - 4\epsilon)}, \tag{C.11g}$$

where (C.11f) follows from (A.9); and (C.11g) follows from the fact that $X_i = \theta_i(U, U_i, S_i, V)$. Note that the probability operator $\Pr[\cdot]$ in (C.11a) applies with a probability distribution $P_{VUU_1S_1U_2Y_1}$ that factorizes as $P_V P_{U,U_1,S_1,U_2|V} P_{Y_1|V}$ given that all the codewords $\mathbf{u}, \mathbf{u}_1, \mathbf{s}_1$, and \mathbf{u}_2 are independent from the output of the channel \mathbf{y}_1 . For all $r \in \{2, 3, \dots, 2^{NR_{2,C}}\}$, the following holds

$$\Pr[E_{1r1}] = \Pr\left[(\mathbf{V}, \mathbf{U}, \mathbf{U}_1, \mathbf{S}_1, \mathbf{U}_2, \mathbf{Y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}\right] \tag{C.12a}$$

$$= \sum_{(\mathbf{v}, \mathbf{u}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2, \mathbf{y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}} P_V(\mathbf{v}) P_{U,U_1,S_1,U_2|V}(\mathbf{u}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2|\mathbf{v}) P_{Y_1|V}(\mathbf{y}_1|\mathbf{v}), \tag{C.12b}$$

$$\leq \left| \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)} \right| 2^{-N(H(U,U_1,U_2,S_1|V) + H(Y_1|V) + H(V) - 4\epsilon)} \tag{C.12c}$$

$$\leq 2^{-N(H(U,U_1,U_2,S_1|V) + H(Y_1|V) + H(V) - H(U,U_1,U_2,S_1,Y_1,V) - 4\epsilon)} \tag{C.12d}$$

$$= 2^{-N(H(U,U_1,U_2,S_1|V) + H(Y_1|V) - H(U,U_1,U_2,S_1,Y_1|V) - 4\epsilon)} \tag{C.12e}$$

$$=2^{-N(I(U,U_1,U_2,S_1;Y_1|V)-4\epsilon)} \quad (\text{C.12f})$$

$$=2^{-N(I(U,U_2,X_1;Y_1|V)-4\epsilon)}. \quad (\text{C.12g})$$

where (C.12f) follows from (A.9); and (C.12g) follows from the fact that $X_i = \theta_i(U, U_i, S_i, V)$. Note that the probability operator $\Pr[\cdot]$ in (C.12a) applies with a probability distribution $P_{VUU_1S_1U_2Y_1}$ that factorizes as $P_V P_{U,U_1,S_1,U_2|V} P_{Y_1|V}$ given that all the codewords \mathbf{u} , \mathbf{u}_1 , \mathbf{s}_1 , and \mathbf{u}_2 are independent from the output of the channel output \mathbf{y}_1 . For all $s \in \{2, 3, \dots, 2^{NR_{1,C}}\}$ and $r \in \{2, 3, \dots, 2^{NR_{2,C}}\}$, the following holds

$$\Pr[E_{sr1}] = \Pr \left[(\mathbf{V}, \mathbf{U}, \mathbf{U}_1, \mathbf{S}_1, \mathbf{U}_2, \mathbf{Y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)} \right] \quad (\text{C.13a})$$

$$= \sum_{(\mathbf{v}, \mathbf{u}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2, \mathbf{y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}} P_V(\mathbf{v}) P_{U,U_1,S_1,U_2|V}(\mathbf{u}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2|\mathbf{v}) P_{Y_1|V}(\mathbf{y}_1|\mathbf{v}), \quad (\text{C.13b})$$

$$\leq \left| \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)} \right| 2^{-N(H(U,U_1,U_2,S_1|V)+H(Y_1|V)+H(V)-4\epsilon)} \quad (\text{C.13c})$$

$$\leq 2^{-N(H(U,U_1,U_2,S_1|V)+H(Y_1|V)+H(V)-H(U,U_1,U_2,S_1,Y_1,V)-4\epsilon)} \quad (\text{C.13d})$$

$$= 2^{-N(H(U,U_1,U_2,S_1|V)+H(Y_1|V)-H(U,U_1,U_2,S_1,Y_1|V)-4\epsilon)} \quad (\text{C.13e})$$

$$= 2^{-N(I(U,U_1,U_2,S_1;Y_1|V)-4\epsilon)} \quad (\text{C.13f})$$

$$= 2^{-N(I(U,U_2,X_1;Y_1|V)-4\epsilon)}, \quad (\text{C.13g})$$

where (C.13f) follows from (A.9); and (C.13g) follows from the fact that $X_i = \theta_i(U, U_i, S_i, V)$. Note that the probability operator $\Pr[\cdot]$ applies in (C.13a) with a probability distribution $P_{VUU_1S_1U_2Y_1}$ that factorizes as $P_V P_{U,U_1,S_1,U_2|V} P_{Y_1|V}$ given that all the codewords \mathbf{u} , \mathbf{u}_1 , \mathbf{s}_1 , and \mathbf{u}_2 are independent from the output of the channel output \mathbf{y}_1 . For all $s \in \{2, 3, \dots, 2^{NR_{1,C}}\}$ and $l \in \{2, 3, \dots, 2^{NR_{1,P}}\}$, the following holds

$$\Pr[E_{s1l}] = \Pr \left[(\mathbf{V}, \mathbf{U}, \mathbf{U}_1, \mathbf{S}_1, \mathbf{U}_2, \mathbf{Y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)} \right] \quad (\text{C.14a})$$

$$= \sum_{(\mathbf{v}, \mathbf{u}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2, \mathbf{y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}} P_V(\mathbf{v}) P_{U,U_1,S_1,U_2|V}(\mathbf{u}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2|\mathbf{v}) P_{Y_1|V}(\mathbf{y}_1|\mathbf{v}), \quad (\text{C.14b})$$

$$\leq \left| \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)} \right| 2^{-N(H(U,U_1,U_2,S_1|V)+H(Y_1|V)+H(V)-4\epsilon)} \quad (\text{C.14c})$$

$$\leq 2^{-N(H(U,U_1,U_2,S_1|V)+H(Y_1|V)+H(V)-H(U,U_1,U_2,S_1,Y_1,V)-4\epsilon)} \quad (\text{C.14d})$$

$$= 2^{-N(H(U,U_1,U_2,S_1|V)+H(Y_1|V)-H(U,U_1,U_2,S_1,Y_1|V)-4\epsilon)} \quad (\text{C.14e})$$

$$= 2^{-N(I(U,U_1,U_2,S_1;Y_1|V)-4\epsilon)} \quad (\text{C.14f})$$

$$= 2^{-N(I(U,U_2,X_1;Y_1|V)-4\epsilon)}, \quad (\text{C.14g})$$

where (C.14f) follows from (A.9); and (C.14g) follows from the fact that $X_i = \theta_i(U, U_i, S_i, V)$. Note that the probability operator $\Pr[\cdot]$ in (C.14a) applies with a probability distribution $P_{VUU_1S_1U_2Y_1}$ that factorizes as $P_V P_{U,U_1,S_1,U_2|V} P_{Y_1|V}$ given that all the codewords \mathbf{u} , \mathbf{u}_1 , \mathbf{s}_1 , and \mathbf{u}_2 are independent from the output of the channel output \mathbf{y}_1 .

For all $r \in \{2, 3, \dots, 2^{NR_{2,C}}\}$ and $l \in \{2, 3, \dots, 2^{NR_{1,P}}\}$, the following holds

$$\Pr[E_{1rl}] = \Pr \left[(\mathbf{V}, \mathbf{U}, \mathbf{U}_1, \mathbf{S}_1, \mathbf{U}_2, \mathbf{Y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)} \right] \quad (\text{C.15a})$$

$$= \sum_{(\mathbf{v}, \mathbf{u}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2, \mathbf{y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}} P_V(\mathbf{v}) P_{U, U_1, S_1, U_2 | V}(\mathbf{u}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2 | \mathbf{v}) P_{Y_1 | V}(\mathbf{y}_1 | \mathbf{v}), \quad (\text{C.15b})$$

$$\leq \left| \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)} \right| 2^{-N(H(U, U_1, U_2, S_1 | V) + H(Y_1 | V) + H(V) - 4\epsilon)} \quad (\text{C.15c})$$

$$\leq 2^{-N(H(U, U_1, U_2, S_1 | V) + H(Y_1 | V) + H(V) - H(U, U_1, U_2, S_1, Y_1, V) - 4\epsilon)} \quad (\text{C.15d})$$

$$= 2^{-N(H(U, U_1, U_2, S_1 | V) + H(Y_1 | V) - H(U, U_1, U_2, S_1, Y_1 | V) - 4\epsilon)} \quad (\text{C.15e})$$

$$= 2^{-N(I(U, U_1, U_2, S_1; Y_1 | V) - 4\epsilon)} \quad (\text{C.15f})$$

$$= 2^{-N(I(U, U_2, X_1; Y_1 | V) - 4\epsilon)}, \quad (\text{C.15g})$$

where (C.15f) follows from (A.9); and (C.15g) follows from the fact that $X_i = \theta_i(U, U_i, S_i, V)$. Note that the probability operator $\Pr[\cdot]$ in (C.15a) applies with a probability distribution $P_{VUU_1S_1U_2Y_1}$ that factorizes as $P_V P_{U, U_1, S_1, U_2 | V} P_{Y_1 | V}$ given that all the codewords \mathbf{u} , \mathbf{u}_1 , \mathbf{s}_1 , and \mathbf{u}_2 are independent from the output of the channel output \mathbf{y}_1 .

For all $s \in \{2, 3, \dots, 2^{NR_{1,C}}\}$, $r \in \{2, 3, \dots, 2^{NR_{2,C}}\}$ and $l \in \{2, 3, \dots, 2^{NR_{1,P}}\}$, the following holds

$$\Pr[E_{srl}] = \Pr \left[(\mathbf{V}, \mathbf{U}, \mathbf{U}_1, \mathbf{S}_1, \mathbf{U}_2, \mathbf{Y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)} \right] \quad (\text{C.16a})$$

$$= \sum_{(\mathbf{v}, \mathbf{u}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2, \mathbf{y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}} P_V(\mathbf{v}) P_{U, U_1, S_1, U_2 | V}(\mathbf{u}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2 | \mathbf{v}) P_{Y_1 | V}(\mathbf{y}_1 | \mathbf{v}), \quad (\text{C.16b})$$

$$\leq \left| \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)} \right| 2^{-N(H(U, U_1, U_2, S_1 | V) + H(Y_1 | V) + H(V) - 4\epsilon)} \quad (\text{C.16c})$$

$$\leq 2^{-N(H(U, U_1, U_2, S_1 | V) + H(Y_1 | V) + H(V) - H(U, U_1, U_2, S_1, Y_1, V) - 4\epsilon)} \quad (\text{C.16d})$$

$$= 2^{-N(H(U, U_1, U_2, S_1 | V) + H(Y_1 | V) - H(U, U_1, U_2, S_1, Y_1 | V) - 4\epsilon)} \quad (\text{C.16e})$$

$$= 2^{-N(I(U, U_1, U_2, S_1; Y_1 | V) - 4\epsilon)} \quad (\text{C.16f})$$

$$= 2^{-N(I(U, U_2, X_1; Y_1 | V) - 4\epsilon)}, \quad (\text{C.16g})$$

where (C.16f) follows from (A.9); and (C.16g) follows from the fact that $X_i = \theta_i(U, U_i, S_i, V)$. Note that the probability operator $\Pr[\cdot]$ in (C.16a) applies with a probability distribution $P_{VUU_1S_1U_2Y_1}$ that factorizes as $P_V P_{U, U_1, S_1, U_2 | V} P_{Y_1 | V}$ given that all the codewords \mathbf{u} , \mathbf{u}_1 , \mathbf{s}_1 , and \mathbf{u}_2 are independent from the output of the channel output \mathbf{y}_1 .

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

$$\Pr[E_{11l}] = \Pr \left[(\mathbf{V}, \mathbf{U}, \mathbf{U}_1, \mathbf{S}_1, \mathbf{U}_2, \mathbf{Y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)} \right] \quad (\text{C.17a})$$

$$= \sum_{(\mathbf{v}, \mathbf{u}, \mathbf{u}_1, \mathbf{s}_1, \mathbf{u}_2, \mathbf{y}_1) \in \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)}} P_V(\mathbf{v}) P_{U, U_1, U_2, Y_1 | V}(\mathbf{u}, \mathbf{u}_1, \mathbf{u}_2, \mathbf{y}_1 | \mathbf{v}) P_{S_1 | V}(\mathbf{s}_1 | \mathbf{v}), \quad (\text{C.17b})$$

$$\leq \left| \mathcal{T}_{VUU_1S_1U_2Y_1}^{(N,\epsilon)} \right| 2^{-N(H(U, U_1, U_2, Y_1 | V) + H(S_1 | V) + H(V) - 4\epsilon)} \quad (\text{C.17c})$$

$$\leq 2^{-N(H(U, U_1, U_2, Y_1 | V) + H(S_1 | V) + H(V) - H(U, U_1, U_2, S_1, Y_1, V) - 4\epsilon)} \quad (\text{C.17d})$$

$$= 2^{-N(I(S_1; U, U_1, U_2, Y_1 | V) - 4\epsilon)} \quad (\text{C.17e})$$

$$=2^{-N(I(X_1;Y_1|U,U_1,U_2,V)-4\epsilon)}, \quad (\text{C.17f})$$

where (C.17e) follows from (A.9); and (C.17f) follows from the fact that $X_i = \theta_i(U, U_i, S_i, V)$. Note that the probability operator $\Pr[\cdot]$ in (C.17a) applies with a probability distribution $P_{VUU_1S_1U_2Y_1}$ that factorizes as $P_V P_{U,U_1,U_2,Y_1|V} P_{S_1|V}$ given that the codewords \mathbf{s}_1 is independent from the output of the channel \mathbf{y}_1 . Plugging (C.11) into (C.10) yields:

$$\begin{aligned} P_e &\leq 2^{N(R_{1C}+R_{2C}+R_{2P}-I(U,X_1,U_2;Y_1|V)+4\epsilon)} + 2^{N(R_{1C}+R_{2C}-I(U,X_1,U_2;Y_1|V)+4\epsilon)} \\ &\quad + 2^{N(R_{1C}+R_{1P}-I(U,X_1,U_2;Y_1|V)+4\epsilon)} + 2^{N(R_{1C}-I(U,X_1,U_2;Y_1|V)+4\epsilon)} \\ &\quad + 2^{N(R_{2C}+R_{1P}-I(U,X_1,U_2;Y_1|V)+4\epsilon)} + 2^{N(R_{2C}-I(U,X_1,U_2;Y_1|V)+4\epsilon)} \\ &\quad + 2^{N(R_{1P}-I(X_1;Y_1|U,U_1,U_2,V)+4\epsilon)}. \end{aligned} \quad (\text{C.18})$$

The same analysis of the probability of error holds for transmitter-receiver pair 2. Hence in general, from (C.9) and (C.18), reliable decoding holds under the following conditions:

$$R_{2C} \leq I(U_2; Y_1 | X_1, U, V), \quad (\text{C.19a})$$

$$R_{1P} \leq I(X_1; Y_1 | U_1, U_2, U, V), \quad (\text{C.19b})$$

$$R_{1C} + R_{2C} + R_{1P} \leq I(U, X_1, U_2; Y_1 | V), \quad (\text{C.19c})$$

$$R_{1C} \leq I(U_1; Y_2 | X_2, U, V), \quad (\text{C.19d})$$

$$R_{2P} \leq I(X_2; Y_2 | U_1, U_2, U, V), \quad (\text{C.19e})$$

$$R_{1C} + R_{2C} + R_{2P} \leq I(U, X_2, U_1; Y_2 | V). \quad (\text{C.19f})$$

The proof continues by applying a Fourier-Motzkin elimination process on (C.19). Set $R_{1P} = R_1 - R_{1C}$, $R_{2P} = R_2 - R_{2C}$. Eliminate R_{1P} , R_{2P} from the inequalities in (C.19) to obtain:

$$R_{2C} \leq a_1, \quad (\text{C.20a})$$

$$R_1 - R_{1C} \leq a_2, \quad (\text{C.20b})$$

$$R_1 + R_{2C} \leq a_3, \quad (\text{C.20c})$$

$$R_{1C} \leq b_1, \quad (\text{C.20d})$$

$$R_2 - R_{2C} \leq b_2, \quad (\text{C.20e})$$

$$R_2 + R_{1C} \leq b_3, \quad (\text{C.20f})$$

$$-R_{1C} \leq 0, \quad (\text{C.20g})$$

$$-R_1 + R_{1C} \leq 0, \quad (\text{C.20h})$$

$$-R_{2C} \leq 0, \quad (\text{C.20i})$$

$$-R_2 + R_{2C} \leq 0. \quad (\text{C.20j})$$

Collect the inequalities in (C.20) that do not include R_{1C} among the above inequalities to obtain:

$$R_{2C} \leq a_1, \quad (\text{C.21a})$$

$$R_1 + R_{2C} \leq a_3, \quad (\text{C.21b})$$

$$R_2 - R_{2C} \leq b_2, \quad (\text{C.21c})$$

$$-R_{2C} \leq 0, \quad (\text{C.21d})$$

$$-R_2 + R_{2C} \leq 0. \quad (\text{C.21e})$$

Collect the inequalities in (C.20) that include R_{1C} with positive coefficients to obtain:

$$R_{1C} \leq b_1, \quad (\text{C.22a})$$

$$R_2 + R_{1C} \leq b_3, \quad (\text{C.22b})$$

$$-R_1 + R_{1C} \leq 0. \quad (\text{C.22c})$$

Collect the inequalities in (C.20) that include R_{1C} with negative coefficients to obtain:

$$R_1 - R_{1C} \leq a_2, \quad (\text{C.23a})$$

$$-R_{1C} \leq 0. \quad (\text{C.23b})$$

Eliminate R_{1C} by adding each inequality from (C.22) and each inequality from (C.23) to obtain:

$$R_1 \leq b_1 + a_2, \quad (\text{C.24a})$$

$$R_2 \leq b_3, \quad (\text{C.24b})$$

$$R_2 + R_1 \leq b_3 + a_2, \quad (\text{C.24c})$$

$$-R_1 \leq 0, \quad (\text{C.24d})$$

$$R_{2C} \leq a_1, \quad (\text{C.24e})$$

$$R_1 + R_{2C} \leq a_3, \quad (\text{C.24f})$$

$$-R_2 + R_{2C} \leq 0, \quad (\text{C.24g})$$

$$R_2 - R_{2C} \leq b_2, \quad (\text{C.24h})$$

$$-R_{2C} \leq 0. \quad (\text{C.24i})$$

Collect the inequalities in (C.24) that do not include R_{2C} to obtain:

$$R_1 \leq b_1 + a_2, \quad (\text{C.25a})$$

$$R_2 \leq b_3, \quad (\text{C.25b})$$

$$R_2 + R_1 \leq b_3 + a_2, \quad (\text{C.25c})$$

$$-R_1 \leq 0. \quad (\text{C.25d})$$

Collect the inequalities in (C.24) that include R_{2C} with positive coefficients to obtain:

$$R_{2C} \leq a_1, \quad (\text{C.26a})$$

$$-R_2 + R_{2C} \leq 0, \quad (\text{C.26b})$$

$$R_1 + R_{2C} \leq a_3. \quad (\text{C.26c})$$

Collect the inequalities in (C.24) that include R_{2C} with negative coefficients to obtain:

$$R_2 - R_{2C} \leq b_2, \quad (\text{C.27a})$$

$$-R_{2C} \leq 0. \quad (\text{C.27b})$$

Eliminate R_{2C} by adding each inequality in (C.26) with each inequality in (C.27) to obtain:

$$R_1 \leq a_3, \quad (\text{C.28a})$$

$$R_1 \leq b_1 + a_2, \quad (\text{C.28b})$$

$$R_2 \leq b_3, \quad (\text{C.28c})$$

$$R_2 \leq a_1 + b_2, \quad (\text{C.28d})$$

$$R_1 + R_2 \leq b_3, \quad (\text{C.28e})$$

$$R_1 + R_2 \leq a_3 + b_2. \quad (\text{C.28f})$$

The proof of Theorem 15 continues as follows, let $k \in \{0, 1\}$ be fixed and consider the following Gaussian input distribution for transmitter k :

$$V \sim \mathcal{N}(0, 1); U \sim \mathcal{N}(0, \rho); U_k \sim \mathcal{N}(0, \lambda_{ke}); \text{ and } S_k \sim \mathcal{N}(0, \lambda_{kp}), \quad (\text{C.29})$$

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 (5.17a), (5.17b) and (5.17e). The choice of the Gaussian input distribution in (C.28) yields:

$$\begin{aligned} I(U, X_1, U_2; Y_1|V) &= h(Y_1|V) - h(Y_1|U, X_1, U_2, V) \\ &= \frac{1}{2} \log \left(2\pi e \text{Var}[Y_1|V] \right) - \frac{1}{2} \log \left(2\pi e \text{Var}[Y_1|U, X_1, U_2, V] \right) \\ &= \frac{1}{2} \log \left(2\pi e (\sigma_1^2 + (1 - \lambda_{1e})P_1 h_{1,1}^2 + (1 - \lambda_{2e})P_2 h_{1,2}^2) \right) \\ &\quad - \frac{1}{2} \log \left(2\pi e (\sigma_1^2 + \lambda_{2p}P_2 h_{1,2}^2) \right) \\ &= \frac{1}{2} \log \left(\frac{1 + (1 - \lambda_{1e})\text{SNR}_1 + (1 - \lambda_{2e})\text{INR}_1 + 2\rho\sqrt{\text{SNR}_1\text{INR}_1}}{1 + \lambda_{2p}\text{INR}_1} \right) \end{aligned} \quad (\text{C.30a})$$

which proves (5.17a). With the same power setting in (C.29), the following holds

$$\begin{aligned} I(U_1; Y_2|U, X_2, V) &= h(Y_2|U, X_2, V) - h(Y_2|U, X_2, U_1, V) \\ &= \frac{1}{2} \log \left(2\pi e \text{Var}[Y_2|U, X_2, V] \right) - \frac{1}{2} \log \left(2\pi e \text{Var}[Y_2|U, X_2, U_1, V] \right) \\ &= \frac{1}{2} \log \left(2\pi e (\sigma_2^2 + (1 - (\rho + \lambda_{1e}))P_1 h_{2,1}^2) \right) \\ &\quad - \frac{1}{2} \log \left(2\pi e (\sigma_2^2 + \lambda_{1p}P_1 h_{2,1}^2) \right) \\ &= \frac{1}{2} \log \left(\frac{1 + (1 - (\rho + \lambda_{1e}))\text{INR}_2}{1 + \lambda_{1p}\text{INR}_2} \right), \text{ and} \end{aligned} \quad (\text{C.30b})$$

$$\begin{aligned} I(X_1; Y_1|U, U_1, U_2, V) &= h(Y_1|U, U_1, U_2, V) - h(Y_1|U, U_1, U_2, X_1, V) \\ &= \frac{1}{2} \log \left(2\pi e \text{Var}[Y_1|U, U_1, U_2, V] \right) - \frac{1}{2} \log \left(2\pi e \text{Var}[Y_1|U, U_1, U_2, X_1, V] \right), \\ &= \frac{1}{2} \log \left(2\pi e (\sigma_1^2 + \lambda_{1p}P_1 h_{1,1}^2 + \lambda_{2p}P_2 h_{1,2}^2) \right) \end{aligned}$$

$$\begin{aligned}
 & -\frac{1}{2} \log \left(2\pi e(\sigma_1^2 + \lambda_{2p} P_1 h_{2,1}^2) \right) \\
 & = \frac{1}{2} \log \left(\frac{1 + \lambda_{1p} \text{SNR}_1 + \lambda_{2p} \text{INR}_1}{1 + \lambda_{2p} \text{INR}_1} \right). \tag{C.30c}
 \end{aligned}$$

This proves (5.17b). Finally, using (C.30b) and (C.30c), yields the proof of (5.17e).

C.0.2 Proof of (5.17g) The choice of the channel input in (C.29) guarantee that the random variables $Y_{3,1}, \dots, Y_{3,n}$ are independently and identically distributed. For all $n \in \{1, 2, \dots, N\}$, $Y_{3,n}$ follows a zero-mean Gaussian distribution with variance \bar{B} given by

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

By the weak law of large numbers, it holds that

$$\lim_{n \rightarrow \infty} \Pr (B^{(N)} < \bar{B}) = 0. \tag{C.32}$$

From (C.32), it holds that for any energy B which satisfies $0 < B \leq \bar{B}$, it holds that

$$\lim_{n \rightarrow \infty} \Pr (B^{(N)} < B) = 0 \tag{C.33}$$

This proves (5.17g) and completes the proof of Theorem 15.

D

Proof of Theorem 16

Fix an information-energy rate triplet (R_1, R_2, B) achievable with a given coding scheme (Definition 8). Denote by \mathbf{X}_1 and \mathbf{X}_2 the channel inputs resulting from transmitting the independent messages (W_1, Ω) and (W_2, Ω) using such coding scheme. Denote by \mathbf{Y}_1 and \mathbf{Y}_2 the corresponding channel outputs. Define the following random variables:

$$S_1 = h_{2,1}X_1 + Z_2 \text{ and} \quad (\text{D.1})$$

$$S_2 = h_{1,2}X_2 + Z_1, \quad (\text{D.2})$$

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 (5.13a), considering Fano's inequality and following similar steps as in [49], it can be shown that the information rates R_1 and R_2 must satisfy the following inequalities

$$NR_1 \leq \sum_{n=1}^N [h(Y_{1,n}) - h(Z_{1,n})] + o(N), \quad (\text{D.3a})$$

$$NR_1 \leq \sum_{n=1}^N [h(Y_{2,n}|X_{2,n}) - h(Z_{2,n}) + h(Y_{1,n}|X_{2,n}, S_{1,n}) - h(Z_{1,n})] + o(N), \quad (\text{D.3b})$$

$$NR_2 \leq \sum_{n=1}^N [h(Y_{2,n}) - h(Z_{2,n})] + o(N), \quad (\text{D.3c})$$

$$NR_2 \leq \sum_{n=1}^N [h(Y_{1,n}|X_{1,n}) - h(Z_{1,n}) + h(Y_{2,n}|X_{1,n}, S_{2,n}) - h(Z_{2,n})] + o(N), \quad (\text{D.3d})$$

$$N(R_1 + R_2) \leq \sum_{n=1}^N [h(Y_{1,n}|S_{1,n}, X_{2,n}) - h(Z_{1,n}) + h(Y_{2,n}) - h(Z_{2,n})] + o(N), \quad (\text{D.3e})$$

$$N(R_1 + R_2) \leq \sum_{n=1}^N [h(Y_{2,n}|S_{2,n}, X_{1,n}) - h(Z_{2,n}) + h(Y_{1,n}) - h(Z_{1,n})] + o(N). \quad (\text{D.3f})$$

Using assumption (5.13b), for any $\eta > 0$ there exists $N_0(\eta)$ such that for any $n \geq N_0(\eta)$ it holds that

$$\Pr(B^{(N)} < B) < \eta. \quad (\text{D.4})$$

Equivalently,

$$\Pr(B^{(N)} \geq B) \geq 1 - \eta. \quad (\text{D.5})$$

Using Markov's inequality, the probability in (D.5) can be upper-bounded as follows:

$$B \Pr(B^{(N)} \geq B) \leq \mathbb{E}[B^{(N)}]. \quad (\text{D.6})$$

Combining (D.5) and (D.6) yields:

$$B(1 - \eta) \leq \mathbb{E}[B^{(N)}], \quad (\text{D.7})$$

which can be written as

$$(B - \delta^{(N)}) \leq \mathbb{E}[B^{(N)}], \quad (\text{D.8})$$

for some $\delta^{(N)} > 0$ (for sufficiently large N). The bounds in (D.3) and (D.8) are evaluated assuming that the channel inputs $X_{1,n}$ and $X_{2,n}$ are arbitrary correlated random variables with

$$\mu_{i,n} \triangleq \mathbb{E}[X_{i,n}], \quad (\text{D.9})$$

$$\gamma_{i,n}^2 \triangleq \text{Var}[X_{i,n}], \quad (\text{D.10})$$

$$\lambda_n \triangleq \text{Cov}[X_{1,n}X_{2,n}], \quad (\text{D.11})$$

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

$$\frac{1}{N} \sum_{n=1}^N \mathbb{E}[X_{i,n}^2] = \left(\frac{1}{N} \sum_{n=1}^N \gamma_{i,n}^2 \right) + \left(\frac{1}{N} \sum_{n=1}^N \mu_{i,n}^2 \right) \leq P_i. \quad (\text{D.12})$$

Using this assumptions, the following holds:

$$\begin{aligned} h(Y_{1,n}) &\leq \frac{1}{2} \log \left(2\pi e \text{Var}[Y_{1,n}] \right) \\ &= \frac{1}{2} \log \left(2\pi e (h_{1,1}^2 \gamma_{1,n}^2 + h_{1,2}^2 \gamma_{2,n}^2 + 2h_{1,1}h_{1,2}\lambda_n + \sigma_1^2) \right), \end{aligned} \quad (\text{D.13a})$$

$$\begin{aligned} h(Y_{2,n}|X_{2,n}) &\leq \frac{1}{2} \log \left(2\pi e \text{Var}[Y_{2,n}|X_{2,n}] \right) \\ &= \frac{1}{2} \log \left(2\pi e \left(\gamma_{2,n}^2 + \frac{h_{2,1}^2}{\sigma_2^2} (\gamma_{1,n}^2 \gamma_{2,n}^2 - \lambda_n^2) \right) \right) - \frac{1}{2} \log(2\pi e \gamma_{2,n}^2) + \frac{1}{2} \log(2\pi e \sigma_2^2) \\ &= \frac{1}{2} \log \left(1 + \gamma_{1,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2} \right) \frac{h_{2,1}^2}{\sigma_2^2} \right) + \frac{1}{2} \log(2\pi e \sigma_2^2), \end{aligned} \quad (\text{D.13b})$$

$$h(Y_{1,n}|X_{2,n}, S_{1,n}) \leq \frac{1}{2} \log(2\pi e \text{Var}[Y_{1,n}|X_{2,n}, S_{1,n}])$$

$$\begin{aligned}
&\leq \frac{1}{2} \log \left(1 + \frac{(\gamma_{1,n}^2 \gamma_{2,n}^2 - \lambda_n^2) \frac{h_{1,1}^2}{\sigma_1^2}}{\gamma_{2,n}^2 + (\gamma_{1,n}^2 \gamma_{2,n}^2 - \lambda_n^2) \frac{h_{2,1}^2}{\sigma_2^2}} \right) + \frac{1}{2} \log(2\pi e \sigma_1^2) \\
&= \frac{1}{2} \log \left(1 + \frac{\gamma_{1,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{1,1}^2}{\sigma_1^2}}{1 + \gamma_{1,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{2,1}^2}{\sigma_2^2}} \right) + \frac{1}{2} \log(2\pi e \sigma_1^2). \tag{D.13c}
\end{aligned}$$

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

$$\text{Var}[Y_{1,n}|X_{2,n}, S_{1,n}] \leq K_{Y_{1,n}} - K_{Y_{1,n}(X_{2,n}, S_{1,n})} K_{(X_{2,n}, S_{1,n})}^{-1} K_{Y_{1,n}(X_{2,n}, S_{1,n})}^\top, \tag{D.14}$$

where

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

Finally, the bounds in (D.3) can be rewritten as follows:

$$NR_1 \leq \sum_{n=1}^N \frac{1}{2} \log \left(\frac{h_{1,1}^2 \gamma_{1,n}^2}{\sigma_1^2} + \frac{h_{1,2}^2 \gamma_{2,n}^2}{\sigma_1^2} + \frac{2h_{1,1} h_{1,2} \lambda_n}{\sigma_1^2} + 1 \right) + o(N), \tag{D.15a}$$

$$\begin{aligned}
NR_1 &\leq \sum_{n=1}^N \frac{1}{2} \log \left(1 + \gamma_{1,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{2,1}^2}{\sigma_2^2} \right) \\
&\quad + \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\gamma_{1,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{1,1}^2}{\sigma_1^2}}{1 + \gamma_{1,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{2,1}^2}{\sigma_2^2}} \right) + o(N), \tag{D.15b}
\end{aligned}$$

$$NR_2 \leq \sum_{n=1}^N \frac{1}{2} \log \left(\frac{h_{2,2}^2 \gamma_{2,n}^2}{\sigma_2^2} + \frac{h_{2,1}^2 \gamma_{1,n}^2}{\sigma_2^2} + \frac{2h_{2,2} h_{2,1} \lambda_n}{\sigma_2^2} + 1 \right) + o(N), \tag{D.15c}$$

$$\begin{aligned}
NR_2 &\leq \sum_{n=1}^N \frac{1}{2} \log \left(1 + \gamma_{2,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{2,1}^2}{\sigma_1^2} \right) \\
&\quad + \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\gamma_{2,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{2,2}^2}{\sigma_2^2}}{1 + \gamma_{2,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{1,2}^2}{\sigma_1^2}} \right) + o(N), \tag{D.15d}
\end{aligned}$$

$$\begin{aligned}
N(R_1 + R_2) &\leq \sum_{n=1}^N \frac{1}{2} \log \left(\frac{h_{2,2}^2 \gamma_{2,n}^2}{\sigma_2^2} + \frac{h_{2,1}^2 \gamma_{1,n}^2}{\sigma_2^2} + \frac{2h_{2,2} h_{2,1} \lambda_n}{\sigma_2^2} + 1 \right) \\
&\quad + \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\gamma_{1,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{1,1}^2}{\sigma_1^2}}{1 + \gamma_{1,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2}\right) \frac{h_{2,1}^2}{\sigma_2^2}} \right) + o(N), \tag{D.15e}
\end{aligned}$$

$$\begin{aligned}
N(R_1 + R_2) &\leq \sum_{n=1}^N \frac{1}{2} \log \left(\frac{h_{1,1}^2 \gamma_{1,n}^2}{\sigma_1^2} + \frac{h_{1,2}^2 \gamma_{2,n}^2}{\sigma_1^2} + \frac{2h_{1,1}h_{1,2}\lambda_n}{\sigma_1^2} + 1 \right) \\
&\quad + \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\gamma_{2,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2} \right) \frac{h_{2,2}^2}{\sigma_2^2}}{1 + \gamma_{2,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{1,n}^2 \gamma_{2,n}^2} \right) \frac{h_{1,2}^2}{\sigma_1^2}} \right) + o(N). \quad (\text{D.15f})
\end{aligned}$$

The average received energy rate is given by

$$\begin{aligned}
\mathbb{E} \left[\frac{1}{N} \sum_{n=1}^N Y_{3,n}^2 \right] &= \sigma_3^2 + h_{3,1}^2 \left(\frac{1}{N} \sum_{n=1}^N (\gamma_{1,n}^2 + \mu_{1,n}^2) \right) + h_{3,2}^2 \left(\frac{1}{N} \sum_{n=1}^N (\gamma_{2,n}^2 + \mu_{2,n}^2) \right) \\
&\quad + 2h_{3,1}h_{3,2} \left(\frac{1}{N} \sum_{n=1}^N (\lambda_n + \mu_{1,n}\mu_{2,n}) \right). \quad (\text{D.16})
\end{aligned}$$

Using Cauchy-Schwarz inequality, the expected value on the energy rate in (D.16) can be upper-bounded as follows:

$$\begin{aligned}
\mathbb{E} \left[\frac{1}{N} \sum_{n=1}^N Y_{3,n}^2 \right] &\leq \sigma_3^2 + h_{3,1}^2 \left(\frac{1}{N} \sum_{n=1}^N (\gamma_{1,n}^2 + \mu_{1,n}^2) \right) + h_{3,2}^2 \left(\frac{1}{N} \sum_{n=1}^N (\gamma_{2,n}^2 + \mu_{2,n}^2) \right) \\
&\quad + 2h_{3,1}h_{3,2} \left(\left| \frac{1}{N} \sum_{n=1}^N \lambda_n \right| + \left(\frac{1}{N} \sum_{n=1}^N \mu_{1,n}^2 \right)^{1/2} \left(\frac{1}{N} \sum_{n=1}^N \mu_{2,n}^2 \right)^{1/2} \right) \quad (\text{D.17})
\end{aligned}$$

Combining (D.8) and (D.17) yields the following upper-bound on the energy rate B :

$$\begin{aligned}
B &\leq \sigma_3^2 + h_{3,1}^2 \left(\frac{1}{N} \sum_{n=1}^N (\gamma_{1,n}^2 + \mu_{1,n}^2) \right) + h_{3,2}^2 \left(\frac{1}{N} \sum_{n=1}^N (\gamma_{2,n}^2 + \mu_{2,n}^2) \right) \\
&\quad + 2h_{3,1}h_{3,2} \left(\left| \frac{1}{N} \sum_{n=1}^N \lambda_n \right| + \left(\frac{1}{N} \sum_{n=1}^N \mu_{1,n}^2 \right)^{1/2} \left(\frac{1}{N} \sum_{n=1}^N \mu_{2,n}^2 \right)^{1/2} \right) + \delta_N. \quad (\text{D.18})
\end{aligned}$$

In order to obtain a single-letterization of the upper-bound given by constraints (D.15) and (D.18), define also

$$\mu_i^2 \triangleq \frac{1}{N} \sum_{n=1}^N \mu_{i,n}^2, \quad i \in \{1, 2\}, \quad (\text{D.19})$$

$$\gamma_i^2 \triangleq \frac{1}{N} \sum_{n=1}^N \gamma_{i,n}^2, \quad i \in \{1, 2\}, \quad (\text{D.20})$$

$$\rho \triangleq \frac{\left(\frac{1}{N} \sum_{n=1}^N \lambda_n \right)}{|\gamma_1| |\gamma_2|}. \quad (\text{D.21})$$

With this notation, the input power constraint in (D.12) can be rewritten as

$$\gamma_i^2 + \mu_i^2 \leq P_i, \quad i \in \{1, 2\}. \quad (\text{D.22})$$

Note that

$$\frac{1}{N} \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\gamma_{i,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{i,n}^2 \gamma_{j,n}^2} \right) \frac{h_{i,i}^2}{\sigma_i^2}}{1 + \gamma_{i,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{i,n}^2 \gamma_{j,n}^2} \right) \frac{h_{j,i}^2}{\sigma_j^2}} \right) \leq \frac{1}{2} \log \left(1 + \frac{\gamma_i^2 (1 - \rho^2) \frac{h_{i,i}^2}{\sigma_i^2}}{1 + \gamma_i^2 (1 - \rho^2) \frac{h_{j,i}^2}{\sigma_j^2}} \right). \quad (\text{D.23})$$

The proof of (D.23) is obtained using the fact that for all $(\alpha, \beta) \in \mathbb{R}_+^2$, $\log \left(1 + \frac{\alpha x}{1 + \beta x} \right)$ is a concave function on x , then it follows that

$$\begin{aligned} & \frac{1}{N} \sum_{n=1}^N \frac{1}{2} \log \left(1 + \frac{\gamma_{i,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{i,n}^2 \gamma_{j,n}^2} \right) \frac{h_{i,i}^2}{\sigma_i^2}}{1 + \gamma_{i,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{i,n}^2 \gamma_{j,n}^2} \right) \frac{h_{j,i}^2}{\sigma_j^2}} \right) \\ & \leq \frac{1}{2} \log \left(1 + \frac{\frac{1}{N} \sum_{n=1}^N \gamma_{i,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{i,n}^2 \gamma_{j,n}^2} \right) \frac{h_{i,i}^2}{\sigma_i^2}}{1 + \frac{1}{N} \sum_{n=1}^N \gamma_{i,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{i,n}^2 \gamma_{j,n}^2} \right) \frac{h_{j,i}^2}{\sigma_j^2}} \right) \\ & \leq \frac{1}{2} \log \left(1 + \frac{\frac{1}{N} \sum_{n=1}^N \gamma_{i,n}^2 \left(1 - \frac{(\frac{1}{N} \sum_{n=1}^N \lambda_n)^2}{\frac{1}{N} \sum_{n=1}^N \gamma_{i,n}^2 \frac{1}{N} \sum_{n=1}^N \gamma_{j,n}^2} \right) \frac{h_{i,i}^2}{\sigma_i^2}}{1 + \frac{1}{N} \sum_{n=1}^N \gamma_{i,n}^2 \left(1 - \frac{(\frac{1}{N} \sum_{n=1}^N \lambda_n)^2}{\frac{1}{N} \sum_{n=1}^N \gamma_{i,n}^2 \frac{1}{N} \sum_{n=1}^N \gamma_{j,n}^2} \right) \frac{h_{j,i}^2}{\sigma_j^2}} \right) \quad (\text{D.24}) \\ & = \frac{1}{2} \log \left(1 + \frac{\gamma_i^2 (1 - \rho^2) \frac{h_{i,i}^2}{\sigma_i^2}}{1 + \gamma_i^2 (1 - \rho^2) \frac{h_{j,i}^2}{\sigma_j^2}} \right), \quad (\text{D.25}) \end{aligned}$$

where (D.24) follows from the fact that $n \in \{1, 2, \dots, N\}$, $\sigma_{i,n}^2 \left(1 - \frac{\lambda_n^2}{\gamma_{i,n}^2 \gamma_{j,n}^2} \right)$ is a concave function on $\gamma_{1,n}^2$, $\gamma_{2,n}^2$ and λ_n ; and for all $(\alpha, \beta) \in \mathbb{R}_+^2$, $\frac{\alpha x}{1 + \beta x}$ is an increasing function on x . By the concavity of the mutual information, applying Jensen's inequality and (D.23) in the bounds (D.15) yields, in the limit when $N \rightarrow \infty$,

$$R_1 \leq \frac{1}{2} \log \left(\frac{h_{1,1}^2 \gamma_1^2}{\sigma_1^2} + \frac{h_{1,2}^2 \gamma_2^2}{\sigma_1^2} + 2\rho \sqrt{\frac{h_{1,1}^2 h_{1,2}^2 \gamma_1^2 \gamma_2^2}{\sigma_1^4}} \right), \quad (\text{D.26a})$$

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{h_{2,1}^2}{\sigma_2^2} \gamma_1^2 (1 - \rho^2) \right) + \frac{1}{2} \log \left(1 + \frac{\gamma_1^2 (1 - \rho^2) \frac{h_{1,1}^2}{\sigma_1^2}}{1 + \gamma_1^2 (1 - \rho^2) \frac{h_{2,1}^2}{\sigma_2^2}} \right), \quad (\text{D.26b})$$

$$R_2 \leq \frac{1}{2} \log \left(\frac{h_{2,2}^2 \gamma_2^2}{\sigma_2^2} + \frac{h_{2,1}^2 \sigma_1^2}{\sigma_2^2} + 2\rho \sqrt{\frac{h_{2,2} h_{2,1} \gamma_1^2 \gamma_2^2}{\sigma_2^4}} \right), \quad (\text{D.26c})$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{h_{1,2}^2}{\sigma_1^2} \gamma_2^2 (1 - \rho^2) \right) + \frac{1}{2} \log \left(1 + \frac{\gamma_2^2 (1 - \rho^2) \frac{h_{2,2}^2}{\sigma_2^2}}{1 + \gamma_2^2 (1 - \rho^2) \frac{h_{1,2}^2}{\sigma_1^2}} \right), \quad (\text{D.26d})$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(\frac{h_{2,2}^2 \gamma_2^2}{\sigma_2^2} + \frac{h_{2,1}^2 \gamma_1^2}{\sigma_2^2} + 2\rho \sqrt{\frac{h_{2,2}^2 h_{2,1}^2 \gamma_1^2 \gamma_2^2}{\sigma_2^4}} \right) + \frac{1}{2} \log \left(1 + \frac{\gamma_1^2 (1 - \rho^2) \frac{h_{1,1}^2}{\sigma_1^2}}{1 + \gamma_1^2 (1 - \rho^2) \frac{h_{2,1}^2}{\sigma_2^2}} \right) \quad (\text{D.26e})$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(\frac{h_{1,1}^2 \gamma_1^2}{\sigma_1^2} + \frac{h_{1,2}^2 \gamma_2^2}{\sigma_1^2} + 2\rho \sqrt{\frac{h_{1,1}^2 h_{1,2}^2 \gamma_1^2 \gamma_2^2}{\sigma_1^4}} \right) + \frac{1}{2} \log \left(1 + \frac{\gamma_2^2 (1 - \rho^2) \frac{h_{2,2}^2}{\sigma_2^2}}{1 + \gamma_2^2 (1 - \rho^2) \frac{h_{1,2}^2}{\sigma_1^2}} \right) \quad (\text{D.26f})$$

and the upper-bound on the energy rate (B.29) :

$$B \leq \sigma_3^2 + h_{3,1}^2 (\gamma_1^2 + \mu_1^2) + h_{3,2}^2 (\gamma_2^2 + \mu_2^2) + 2h_{2,1} h_{2,2} (|\rho| |\gamma_1| |\gamma_2| + |\mu_1| |\mu_2|). \quad (\text{D.26g})$$

In the limit when N tends to infinity, any information-energy rate triplet such that $(R_1, R_2, B) \in \mathcal{E}^F$ satisfies (D.26) for some $\gamma_1, \gamma_2, \mu_1, \mu_2$ satisfying (D.22) and for some $\rho \in [-1, 1]$. Let $\mathcal{R}^F(\gamma_1, \gamma_2, \mu_1, \mu_2, \rho)$ denote the set of information-energy rate triplets satisfying (D.26), for some $\gamma_1, \gamma_2, \mu_1, \mu_2$ such that (D.22) is true and for some $\rho \in [-1, 1]$. Thus, it holds that

$$\mathcal{E}^F \subseteq \bigcup_{\substack{0 \leq \gamma_1^2 + \mu_1^2 \leq P_1 \\ 0 \leq \gamma_2^2 + \mu_2^2 \leq P_2 \\ -1 \leq \rho \leq 1}} \mathcal{R}^F(\gamma_1^2, \gamma_2^2, \mu_1, \mu_2, \rho). \quad (\text{D.27})$$

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

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

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

$$\beta_i \triangleq \frac{\gamma_i^2}{P_i} = \frac{P_i - \mu_i^2}{P_i}, \quad i \in \{1, 2\}. \quad (\text{D.28})$$

Using (D.28), any region $\mathcal{R}(\gamma_1, \gamma_2, \mu_1, \mu_2, \rho)$ 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 $0 \leq \rho \leq 1$, can be rewritten as follows:

$$R_1 \leq \frac{1}{2} \log \left(\frac{h_{1,1}^2 \beta_1 P_1}{\sigma_1^2} + \frac{h_{1,2}^2 \beta_2 P_2}{\sigma_1^2} + 2\rho \sqrt{\frac{h_{1,1}^2 h_{1,2}^2 \beta_1 \beta_2 P_1 P_2}{\sigma_1^4}} \right), \quad (\text{D.29a})$$

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{h_{2,1}^2 \beta_1 P_1 (1 - \rho^2)}{\sigma_2^2} \right) + \frac{1}{2} \log \left(1 + \frac{\beta_1 P_1 (1 - \rho^2) \frac{h_{1,1}^2}{\sigma_1^2}}{1 + \beta_1 P_1 (1 - \rho^2) \frac{h_{2,1}^2}{\sigma_2^2}} \right), \quad (\text{D.29b})$$

$$R_2 \leq \frac{1}{2} \log \left(\frac{h_{2,2}^2 \beta_2 P_2}{\sigma_2^2} + \frac{h_{2,1}^2 \beta_1 P_1}{\sigma_2^2} + 2\rho \sqrt{\frac{h_{2,2}^2 h_{2,1}^2 \beta_1 \beta_2 P_1 P_2}{\sigma_2^4}} \right), \quad (\text{D.29c})$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{h_{1,2}^2 \beta_2 P_2 (1 - \rho^2)}{\sigma_1^2} \right) + \frac{1}{2} \log \left(1 + \frac{\beta_2 P_2 (1 - \rho^2) \frac{h_{2,2}^2}{\sigma_2^2}}{1 + \beta_2 P_2 (1 - \rho^2) \frac{h_{1,2}^2}{\sigma_1^2}} \right), \quad (\text{D.29d})$$

$$\begin{aligned}
R_1 + R_2 &\leq \frac{1}{2} \log \left(\frac{h_{2,2}^2 \beta_2 P_2}{\sigma_2^2} + \frac{h_{2,1}^2 \beta_1 P_1}{\sigma_2^2} + 2\rho \sqrt{\frac{h_{2,2}^2 h_{2,1}^2 \beta_1 \beta_2 P_1 P_2}{\sigma_2^4}} \right) \\
&\quad + \frac{1}{2} \log \left(1 + \frac{\beta_1 P_1 (1 - \rho^2) \frac{h_{1,1}^2}{\sigma_1^2}}{1 + \beta_1 P_1 (1 - \rho^2) \frac{h_{2,1}^2}{\sigma_2^2}} \right), \tag{D.29e}
\end{aligned}$$

$$\begin{aligned}
R_1 + R_2 &\leq \frac{1}{2} \log \left(\frac{h_{1,1}^2 \beta_1 P_1}{\sigma_1^2} + \frac{h_{1,2}^2 \beta_2 P_2}{\sigma_1^2} + 2\rho \sqrt{\frac{h_{1,1}^2 h_{1,2}^2 \beta_1 \beta_2 P_1 P_2}{\sigma_1^4}} \right) \\
&\quad + \frac{1}{2} \log \left(1 + \frac{\beta_2 P_2 (1 - \rho^2) \frac{h_{2,2}^2}{\sigma_2^2}}{1 + \beta_2 P_2 (1 - \rho^2) \frac{h_{1,2}^2}{\sigma_1^2}} \right), \tag{D.29f}
\end{aligned}$$

$$B \leq \sigma_3^2 + h_{3,1}^2 P_1 + h_{3,2}^2 P_2 + 2h_{3,1} h_{3,2} (|\rho| \sqrt{\beta_1 P_1 \beta_2 P_2} + \sqrt{(1 - \beta_1)(1 - \beta_2) P_1 P_2}), \tag{D.29g}$$

for some $(\beta_1, \beta_2) \in [0, 1]^2$ and $\rho \in [0, 1]$. Hence, using the definitions in (5.7) and (5.10), the region (D.29) contains all information-energy rate triplets (R_1, R_2, B) satisfying constraints (5.18), which completes the proof of Theorem 16.

E

Preliminary Results

This section introduces some auxiliary results that play a key role in the following appendices.

Definition 19 (Moment Generating Function). *Given a random variable X , its moment generating function is denoted by $\phi_X : \mathbb{R} \rightarrow \mathbb{R}$ and*

$$\phi_X(\lambda) = \mathbb{E}_X [e^{\lambda X}]. \quad (\text{E.1})$$

The following lemmas highlight some properties of the moment generating function.

Lemma 8. *Let $Z = \sum_{i=1}^n X_i$ be a random variable formed by the sum of n independent random variables X_1, X_2, \dots, X_n . Then, for all $\lambda \in \mathbb{R}$,*

$$\phi_Z(\lambda) = \prod_{t=1}^n \phi_{X_t}(\lambda). \quad (\text{E.2})$$

Proof:

$$\phi_Z(\lambda) = \mathbb{E}_Z [e^{\lambda Z}] \quad (\text{E.3})$$

$$= \mathbb{E}_{X_1, X_2, \dots, X_n} [e^{\lambda \sum_{t=1}^n X_t}] \quad (\text{E.4})$$

$$= \mathbb{E}_{X_1, X_2, \dots, X_n} \prod_{t=1}^n e^{\lambda X_t} \quad (\text{E.5})$$

$$= \prod_{t=1}^n \mathbb{E}_{X_t} [e^{\lambda X_t}] \quad (\text{E.6})$$

$$= \prod_{t=1}^n \phi_{X_t}(\lambda). \quad (\text{E.7})$$

■

Lemma 9. *Let X be a Bernoulli random variable with $P_X(1) = 1 - P_X(0) = \rho$. Then, for all $\lambda \in \mathbb{R}$*

$$\phi_X(\lambda) = 1 + \rho (e^\lambda - 1). \quad (\text{E.8})$$

Proof:

$$\mathbb{E}_X [e^{\lambda X}] = P_X(1)e^\lambda + P_X(0) \quad (\text{E.9})$$

$$= \rho e^\lambda + (1 - \rho) \quad (\text{E.10})$$

$$= 1 + \rho(e^\lambda - 1). \quad (\text{E.11})$$

■

Theorem 20 (Berry-Esseen Theorem, [7]). *Let X_1, X_2, \dots, X_n be independent random variables such that for all $t \in \{1, 2, \dots, n\}$,*

$$\mu_t = \mathbb{E}_{X_t} [X_t], \quad (\text{E.12})$$

$$\sigma_t^2 = \mathbb{E}_{X_t} [X_t^2] - \mu_t^2, \quad (\text{E.13})$$

$$\phi_t = \mathbb{E}_{X_t} [|X_t - \mu_t|^3]. \quad (\text{E.14})$$

Then, it holds for all $\lambda \in \mathbb{R}$ that

$$\left| \Pr \left(\sum_{t=1}^n X_t - \mu_t \geq \sigma \lambda \right) - Q(\lambda) \right| \leq \frac{c_0 \phi}{\sigma^3}, \quad (\text{E.15})$$

where

$$\mu = \sum_{t=1}^n \mu_t, \quad \sigma^2 = \sum_{t=1}^n \sigma_t^2, \quad \text{and} \quad \phi = \sum_{t=1}^n \phi_t. \quad (\text{E.16})$$

The best value of the constant c_0 is $c_0 = 0.4748$.

F

Proof of Proposition 5

Consider the definition of the ESP in (6.13). Hence, for all $i \in \{1, 2, \dots, M\}$,

$$\theta_i = \Pr \left[\sum_{t=1}^n \mathbb{1}_{\{Z_t=0\}} < \left(\frac{n(b-b_1)}{b_0-b_1} \right) \middle| \mathbf{X} = \mathbf{u}(i) \right]. \quad (\text{F.1})$$

Assume that the Transmitter uses an $(n, M, \epsilon, \delta, b)$ -code and it aims at sending the message index $i \in \{1, 2, \dots, M\}$. Then, for all $t \in \{1, 2, \dots, n\}$, the random variable $\mathbb{1}_{\{Z_t=0\}}$ in (F.1) follows a Bernoulli distribution and the probability of a “one” is

$$P_{Z|X}(0|u_t(i)) = \begin{cases} \alpha_2 & \text{if } u_t(i) = 1 \\ 1 - \alpha_2 & \text{if } u_t(i) = 0 \end{cases}. \quad (\text{F.2})$$

This implies that the random variable $\sum_{t=1}^n \mathbb{1}_{\{Z_t=0\}}$ can be expressed as the sum of two random variables with binomial distributions $\mathcal{B}(N(0|\mathbf{u}(i)), 1 - \alpha_2)$ and $\mathcal{B}(N(1|\mathbf{u}(i)), \alpha_2)$, respectively. That is,

$$\sum_{t=1}^n \mathbb{1}_{\{Z_t=0\}} = \sum_{t \in \{m: u_m(i)=0\}} \mathbb{1}_{\{Z_t=0\}} + \sum_{t \in \{m: u_m(i)=1\}} \mathbb{1}_{\{Z_t=0\}}. \quad (\text{F.3})$$

Hence,

$$\begin{aligned} \theta_i &= \sum_{m=0}^{\lfloor \frac{n(b-b_1)}{b_0-b_1} \rfloor} \Pr \left[\sum_{t \in \{m: u_m(i)=0\}} \mathbb{1}_{\{Z_t=0\}} + \sum_{t \in \{m: u_m(i)=1\}} \mathbb{1}_{\{Z_t=0\}} = m \middle| \mathbf{X} = \mathbf{u}(i) \right] \\ &= \sum_{k=0}^{\lfloor \frac{n(b-b_1)}{b_0-b_1} \rfloor} \sum_{s=0}^k \Pr \left[\sum_{t \in \{m: u_m(i)=0\}} \mathbb{1}_{\{Z_t=0\}} = s \middle| \mathbf{X} = \mathbf{u}(i) \right] \\ &\quad \Pr \left[\sum_{t \in \{m: u_m(i)=1\}} \mathbb{1}_{\{Z_t=0\}} = k - s \middle| \mathbf{X} = \mathbf{u}(i) \right] \\ &= \sum_{k=0}^{\lfloor \frac{n(b-b_1)}{b_0-b_1} \rfloor} \sum_{s=0}^k \binom{N(0|\mathbf{u}(i))}{s} (1 - \alpha_2)^s \alpha_2^{N(0|\mathbf{u}(i))-s} \binom{N(1|\mathbf{u}(i))}{k-s} (1 - \alpha_2)^{N(1|\mathbf{u}(i))-k+s} \alpha_2^{k-s} \end{aligned}$$

$$= \sum_{k=0}^{\lfloor \frac{n(b-b_1)}{b_0-b_1} \rfloor} \sum_{s=0}^k \binom{N(0|\mathbf{u}(i))}{s} \binom{N(1|\mathbf{u}(i))}{k-s} (1-\alpha_2)^{N(1|\mathbf{u}(i))-k+2s} \alpha_2^{N(0|\mathbf{u}(i))+k-2s}, \quad (\text{F.4})$$

which completes the proof.

G

Proof of Proposition 6

For all $i \in \{1, 2, \dots, M\}$, let V_i be the following random variable:

$$V_i = \sum_{t \in \{m: u_m(i)=0\}} \mathbb{1}_{\{Z_t=0\}} + \sum_{t \in \{m: u_m(i)=1\}} \mathbb{1}_{\{Z_t=0\}}. \quad (\text{G.1})$$

From Lemma 8 and Lemma 9, the following holds,

$$\phi_{V_i}(\lambda) = \prod_{t \in \{m: u_m(i)=0\}}^n 1 + (1 - \alpha_2)(e^\lambda - 1) \prod_{t \in \{m: u_m(i)=1\}}^n 1 + \alpha_2(e^\lambda - 1) \quad (\text{G.2})$$

$$\leq \prod_{t \in \{m: u_m(i)=0\}}^n \exp\left((1 - \alpha_2)(e^\lambda - 1)\right) \prod_{t \in \{m: u_m(i)=1\}}^n \exp\left(\alpha_2(e^\lambda - 1)\right) \quad (\text{G.3})$$

$$\leq \exp\left(N(0|\mathbf{u}(i))(1 - \alpha_2)(e^\lambda - 1) + N(1|\mathbf{u}(i))\alpha_2(e^\lambda - 1)\right) \quad (\text{G.4})$$

$$= \exp\left((e^\lambda - 1)(N(0|\mathbf{u}(i))(1 - \alpha_2) + N(1|\mathbf{u}(i))\alpha_2)\right) \quad (\text{G.5})$$

$$= \exp\left(n(e^\lambda - 1)\left(\bar{P}_X^{(i)}(0)(1 - \alpha_2) + (1 - \bar{P}_X^{(i)}(0))\alpha_2\right)\right) \quad (\text{G.6})$$

$$= \exp\left(n(e^\lambda - 1)(\bar{P}_X^{(i)}(0)(1 - 2\alpha_2) + \alpha_2)\right), \quad (\text{G.7})$$

where (G.3) follows from the fact that for all $x \in \mathbb{R}$, $1 + x \leq e^x$. In order to ease the notation, let μ be

$$\mu \triangleq n\left(\bar{P}_X^{(i)}(0)(1 - 2\alpha_2) + \alpha_2\right). \quad (\text{G.8})$$

From Markov's inequality, it holds that for all $(\lambda, \gamma) \in \mathbb{R}^2$,

$$\Pr(V_i > (1 + \gamma)\mu) = \Pr\left(e^{\lambda V_i} > e^{\lambda(1+\gamma)\mu}\right) \quad (\text{G.9})$$

$$\leq \frac{\mathbb{E}_{V_i}\left[e^{\lambda V_i}\right]}{e^{\lambda(1+\gamma)\mu}} \quad (\text{G.10})$$

$$= \frac{\phi_{V_i}(\lambda)}{e^{\lambda(1+\gamma)\mu}} \quad (\text{G.11})$$

$$\leq \frac{\exp\left((e^\lambda - 1)\mu\right)}{e^{\lambda(1+\gamma)\mu}} \quad (\text{G.12})$$

$$= \exp\left((e^\lambda - 1)\mu - \lambda(1 + \gamma)\mu\right) \quad (\text{G.13})$$

$$= \exp\left(\mu(e^\lambda - 1 - \lambda(1 + \gamma))\right). \quad (\text{G.14})$$

Note that the choice of λ can be improved to tight the bound in (G.14). Note that,

$$\frac{d}{d\lambda} e^\lambda - (1 + \gamma)\lambda = e^\lambda - (1 + \gamma). \quad (\text{G.15})$$

Then, the optimal λ is the solution to $e^\lambda - (1 + \gamma) = 0$. That is, $\lambda = \log(1 + \gamma)$. This implies:

$$\Pr(V_i > (1 + \gamma)\mu) \leq \exp\left(\mu(1 + \gamma - 1 - (1 + \gamma)\log(1 + \gamma))\right) \quad (\text{G.16})$$

$$= \exp\left(\mu(\gamma - (1 + \gamma)\log(1 + \gamma))\right) \quad (\text{G.17})$$

$$= \left(\frac{e^\lambda}{(1 + \gamma)^{(1 + \gamma)}}\right)^\mu \quad (\text{G.18})$$

$$< \exp\left(-\frac{\gamma^2}{2 + \gamma}\mu\right), \quad (\text{G.19})$$

where (G.19) follows from the fact that

$$\log\left(\left(\frac{e^\lambda}{(1 + \gamma)^{(1 + \gamma)}}\right)^\mu\right) = \mu \log(\exp \gamma) - \mu(1 + \gamma) \log(1 + \gamma) \quad (\text{G.20})$$

$$\leq \mu \left(\gamma - \frac{2(1 + \gamma)\gamma}{2 + \gamma}\right) \quad (\text{G.21})$$

$$= \mu \left(-\frac{\gamma^2}{2 + \gamma}\right), \quad (\text{G.22})$$

and the inequality in (G.21) is due to the fact that for all $x > 0$, $\log(1 + x) \geq \frac{x}{1 + \frac{x}{2}}$. Hence, it follows that

$$\begin{aligned} & \Pr\left(\sum_{t=1}^n \mathbb{1}_{\{Z_t=0\}} > \frac{n(b - b_1)}{b_0 - b_1}\right) \\ &= \Pr\left[\sum_{t=1}^n \mathbb{1}_{\{Z_t=0\}} > \left(1 + \left(\frac{b - b_1}{(\bar{P}_X^{(i)}(0)(1 - 2\alpha_2) + \alpha_2)(b_0 - b_1)}\right) - 1\right) \right. \\ & \quad \left. n(\bar{P}_X^{(i)}(0)(1 - 2\alpha_2) + \alpha_2)(b_0 - b_1)\right] \end{aligned} \quad (\text{G.23})$$

$$\leq \exp\left(\frac{-n\left(\frac{b - b_1}{(\bar{P}_X^{(i)}(0)(1 - 2\alpha_2) + \alpha_2)(b_0 - b_1)} - 1\right)^2 (\bar{P}_X^{(i)}(0)(1 - 2\alpha_2) + \alpha_2)}{\frac{b - b_1}{(\bar{P}_X^{(i)}(0)(1 - 2\alpha_2) + \alpha_2)(b_0 - b_1)}}\right) \quad (\text{G.24})$$

$$= \exp\left(-n \frac{\left(\frac{b - b_1}{b_0 - b_1} - (P_X^{(i)}(0)(1 - 2\alpha_2) + \alpha_2)\right)^2}{\left(\frac{b - b_1}{b_0 - b_1} + (P_X^{(i)}(0)(1 - 2\alpha_2) + \alpha_2)\right)}\right). \quad (\text{G.25})$$

This completes the proof of (6.23).

On the other hand, from Markov's inequality, it holds that for all $(\lambda, \gamma) \in \mathbb{R}^2$,

$$\Pr(V_i < (1 - \gamma)\mu) = \Pr(e^{-\lambda V_i} > e^{-\lambda(1-\gamma)\mu}) \quad (\text{G.26})$$

$$\leq \frac{\mathbb{E}_{V_i}[e^{-\lambda V_i}]}{e^{\lambda(1-\gamma)\mu}} \quad (\text{G.27})$$

$$= \frac{\phi_{V_i}(-\lambda)}{e^{-\lambda(1-\gamma)\mu}} \quad (\text{G.28})$$

$$\leq \frac{\exp((e^{-\lambda} - 1)\mu)}{e^{-\lambda(1-\gamma)\mu}} \quad (\text{G.29})$$

$$= \exp(\mu(e^{-\lambda} - 1 + \lambda(1 - \gamma))). \quad (\text{G.30})$$

Note that the choice of λ can be improved to tight the bound in (G.30). Note that,

$$\frac{d}{d\lambda} (e^{-\lambda} + (1 - \gamma)\lambda) = -e^{-\lambda} + (1 - \gamma). \quad (\text{G.31})$$

Then, the optimal λ is the solution to $-e^{-\lambda} + (1 - \gamma) = 0$. That is, $\lambda = -\log(1 - \gamma)$. This implies:

$$\Pr(V_i < (1 - \gamma)\mu) \leq \exp(\mu(e^{-\lambda} - 1 + \lambda(1 - \gamma))) \quad (\text{G.32})$$

$$\leq \exp\left(-\mu\left(\gamma + (1 - \gamma)\left(-\gamma + \frac{\gamma^2}{2}\right)\right)\right) \quad (\text{G.33})$$

$$= \exp\left(-\mu\left(\frac{\gamma^2}{2} + \gamma^2\left(1 - \frac{\gamma^3}{2}\right)\right)\right) \quad (\text{G.34})$$

$$\leq \exp\left(-\frac{\mu\gamma^2}{2}\right) \quad (\text{G.35})$$

Then,

$$\begin{aligned} & \Pr\left(\sum_{t=1}^n \mathbf{1}_{\{Z_t=0\}} < \frac{n(b - b_1)}{b_0 - b_1}\right) \\ &= \Pr\left[\sum_{t=1}^n \mathbf{1}_{\{Z_t=0\}} < \left(1 - \left(1 - \frac{b - b_1}{(\bar{P}_X^{(i)}(0)(1 - 2\alpha_2) + \alpha_2)(b_0 - b_1)}\right)\right)\right. \\ & \quad \left. n(\bar{P}_X^{(i)}(0)(1 - 2\alpha_2) + \alpha_2)(b_0 - b_1)\right] \\ &\leq \exp\left(-n \frac{(\bar{P}_X^{(i)}(0)(1 - 2\alpha_2) + \alpha_2)(b_0 - b_1)}{2} \left(1 - \frac{b - b_1}{(\bar{P}_X^{(i)}(0)(1 - 2\alpha_2) + \alpha_2)(b_0 - b_1)}\right)^2\right) \\ &= \exp\left(-n \frac{\left(\frac{b - b_1}{b_0 - b_1} - (\bar{P}_X^{(i)}(0)(1 - 2\alpha_2) + \alpha_2)\right)^2}{2(\bar{P}_X^{(i)}(0)(1 - 2\alpha_2) + \alpha_2)}\right). \end{aligned} \quad (\text{G.36})$$

This completes the proof of (6.24).

H

Proof of Proposition 7

From Proposition 6, it follows that

$$\Pr \left(\sum_{t=1}^n \mathbb{1}_{\{Z_t=0\}} > \frac{n(b-b_1)}{b_0-b_1} \right) < \exp \left(-n \frac{\left(\frac{b-b_1}{b_0-b_1} - (P_X^{(i)}(0)(1-2\alpha_2) + \alpha_2) \right)^2}{\left(\frac{b-b_1}{b_0-b_1} + (P_X^{(i)}(0)(1-2\alpha_2) + \alpha_2) \right)} \right), \quad (\text{H.1})$$

and thus subject to a maximal ESP constraint, it holds that for all $i \in \{1, 2, \dots, M\}$,

$$\delta > \theta_i > 1 - \exp \left(-n \frac{\left(\frac{b-b_1}{b_0-b_1} - (P_X^{(i)}(0)(1-2\alpha_2) + \alpha_2) \right)^2}{\left(\frac{b-b_1}{b_0-b_1} + (P_X^{(i)}(0)(1-2\alpha_2) + \alpha_2) \right)} \right). \quad (\text{H.2})$$

Hence, it follows that

$$\exp \left(-n \frac{\left(\frac{b-b_1}{b_0-b_1} - (P_X^{(i)}(0)(1-2\alpha_2) + \alpha_2) \right)^2}{\left(\frac{b-b_1}{b_0-b_1} + (P_X^{(i)}(0)(1-2\alpha_2) + \alpha_2) \right)} \right) > 1 - \delta, \quad (\text{H.3})$$

which implies

$$\begin{aligned} & \left(\frac{b-b_1}{b_0-b_1} \right)^2 - \left(2(P_X^{(i)}(0)(1-2\alpha_2) + \alpha_2) - \frac{1}{n} \log(1-\delta) \right) \left(\frac{b-b_1}{b_0-b_1} \right) + (P_X^{(i)}(0)(1-2\alpha_2) + \alpha_2) \\ & \cdot \left((P_X^{(i)}(0)(1-2\alpha_2) + \alpha_2) + \frac{1}{n} \log(1-\delta) \right) < 0. \end{aligned} \quad (\text{H.4})$$

Denote Γ_1 and Γ_2 the roots of the quadratic function in the left-hand side of (H.4). Then,

$$\begin{aligned} \Gamma_1 &= \frac{1}{2} \left(2(P_X^{(i)}(0)(1-2\alpha_2) + \alpha_2) - \frac{1}{n} \log(1-\delta) \right) \\ & \quad + \frac{1}{2} \sqrt{-\frac{8}{n} (P_X^{(i)}(0)(1-2\alpha_2) + \alpha_2) \log(1-\delta) + \frac{1}{n^2} \log(1-\delta)^2} \end{aligned} \quad (\text{H.5})$$

$$= (P_X^{(i)}(0)(1-2\alpha_2) + \alpha_2) - \frac{1}{2n} \log(1-\delta) \quad (\text{H.6})$$

$$+\sqrt{\frac{-\log(1-\delta)}{n}}\sqrt{2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)-\frac{1}{4n}\log(1-\delta)} \quad (\text{H.7})$$

$$\Gamma_2 = \frac{1}{2}\left(2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)-\frac{1}{n}\log(1-\delta)\right) \quad (\text{H.8})$$

$$-\frac{1}{2}\sqrt{-\frac{8}{n}\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)\log(1-\delta)+\frac{1}{n^2}\log(1-\delta)^2} \quad (\text{H.9})$$

$$= \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)-\frac{1}{2n}\log(1-\delta) \quad (\text{H.10})$$

$$-\sqrt{\frac{-\log(1-\delta)}{n}}\sqrt{2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)-\frac{1}{4n}\log(1-\delta)} \quad (\text{H.11})$$

From (H.4), it follows that $\left(\frac{b-b_1}{b_0-b_1}\right)$ must satisfy:

$$\Gamma_2 < \left(\frac{b-b_1}{b_0-b_1}\right) < \Gamma_1. \quad (\text{H.12})$$

This implies that:

$$\left(\frac{b-b_1}{b_0-b_1}\right) \leq \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)-\frac{1}{2n}\log(1-\delta) \quad (\text{H.13})$$

$$+\sqrt{\frac{-\log(1-\delta)}{n}}\sqrt{2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)-\frac{1}{4n}\log(1-\delta)} \quad (\text{H.14})$$

Thus,

$$\begin{aligned} b &\leq \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)+b_1-\frac{(b_0-b_1)}{2n}\log(1-\delta) \\ &\quad +\sqrt{\frac{-(b_0-b_1)\log(1-\delta)}{n}}\sqrt{2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)-\frac{(b_0-b_1)}{4n}\log(1-\delta)} \\ &\leq \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)+b_1-\frac{(b_0-b_1)}{2n}\log(1-\delta) \\ &\quad +\sqrt{\frac{2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)^2\log\left(\frac{1}{1-\delta}\right)}{n}}+\sqrt{\frac{(b_0-b_1)^2\left(\log\left(\frac{1}{1-\delta}\right)\right)^2}{4n^2}} \end{aligned} \quad (\text{H.15})$$

$$\begin{aligned} &= \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)+b_1-\frac{(b_0-b_1)}{n}\log(1-\delta) \\ &\quad +\frac{(b_0-b_1)}{\sqrt{n}}\sqrt{-2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)\log(1-\delta)}, \end{aligned} \quad (\text{H.16})$$

and this completes the proof of (6.25). The proof of (6.26) follows immediately from (6.14) and (H.16)

Proof of Lemma 6

For all $t \in \{1, 2, \dots, n\}$, consider the first moment, the second moment, and the third absolute moment of the random variable $\mathbf{1}_{\{Z_t=0\}}$ given that the channel input is $\mathbf{u}(i)$, with $i \in \{1, 2, \dots, M\}$:

$$\mathbb{E}_{Z|X=u_t(i)} [\mathbf{1}_{\{Z_t=0\}}] = P_{Z|X}(0|u_t(i)) \quad (\text{I.1a})$$

$$\begin{aligned} \mathbb{V}_{Z|X=u_t(i)} [\mathbf{1}_{\{Z_t=0\}}] &= \mathbb{E}_{Z|X=u_t(i)} [\mathbf{1}_{\{Z_t=0\}}^2] - \mathbb{E}_{Z|X=u_t(i)} [\mathbf{1}_{\{Z_t=0\}}]^2 \\ &= \mathbb{E}_{Z|X=u_t(i)} [\mathbf{1}_{\{Z_t=0\}}] - \mathbb{E}_{Z|X=u_t(i)} [\mathbf{1}_{\{Z_t=0\}}]^2 \\ &= P_{Z|X}(0|u_t(i)) - P_{Z|X}(0|u_t(i))^2 \\ &= P_{Z|X}(0|u_t(i)) (1 - P_{Z|X}(0|u_t(i))) \\ &= P_{Z|X}(0|u_t(i)) P_{Z|X}(1|u_t(i)), \text{ and} \end{aligned} \quad (\text{I.1b})$$

$$\begin{aligned} \mathbb{S}_{Z|X=u_t(i)} [\mathbf{1}_{\{Z_t=0\}}] &= \mathbb{E}_{Z|X=u_t(i)} [|\mathbf{1}_{\{Z_t=0\}} - P_{Z|X}(0|u_t(i))|^3] \\ &= P_{Z|X}(0|u_t(i)) |1 - P_{Z|X}(0|u_t(i))|^3 + P_{Z|X}(1|u_t(i)) P_{Z|X}(0|u_t(i))^3 \\ &= P_{Z|X}(0|u_t(i)) P_{Z|X}(1|u_t(i))^3 + P_{Z|X}(1|u_t(i)) P_{Z|X}(0|u_t(i))^3. \end{aligned} \quad (\text{I.1c})$$

Using the equalities above, the following holds,

$$\begin{aligned} \sum_{t=1}^n \mathbb{E}_{Z|X} [\mathbf{1}_{\{Z_t=0\}}] &= n P_{Z|X}(0|0) \bar{P}_X^{(i)}(0) + n P_{Z|X}(0|1) \bar{P}_X^{(i)}(1) \\ &= n \left((1 - 2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 \right), \end{aligned} \quad (\text{I.2a})$$

$$\begin{aligned} \sum_{t=1}^n \mathbb{V}_{Z|X=u_t(i)} [\mathbf{1}_{\{Z_t=0\}}] &= \sum_{t=1}^n P_{Z|X}(0|u_t(i)) P_{Z|X}(1|u_t(i)) \\ &= N(0|\mathbf{u}(i)) P_{Z|X}(0|0) P_{Z|X}(1|0) + N(1|\mathbf{u}(i)) P_{Z|X}(0|1) P_{Z|X}(1|1) \\ &= N(0|\mathbf{u}(i)) (1 - \alpha_2) \alpha_2 + N(1|\mathbf{u}(i)) \alpha_2 (1 - \alpha_2) \\ &= n (\alpha_2 (1 - \alpha_2)), \text{ and} \end{aligned} \quad (\text{I.2b})$$

$$\begin{aligned} \sum_{k=1}^n \mathbb{S}_{Z|X=u_t(i)} [\mathbf{1}_{\{Z_t=0\}}] &= \sum_{t=1}^n P_{Z|X}(0|u_t(i)) P_{Z|X}(1|u_t(i))^3 + P_{Z|X}(1|u_t(i)) P_{Z|X}(0|u_t(i))^3 \\ &= N(0|\mathbf{u}(i)) (P_{Z|X}(0|0) P_{Z|X}(1|0))^3 + P_{Z|X}(1|0) P_{Z|X}(0|0)^3 \end{aligned}$$

$$\begin{aligned}
& +N(1|\mathbf{u}(i)) (P_{Z|X}(0|1)P_{Z|X}(1|1)^3 + P_{Z|X}(1|1)P_{Z|X}(0|1)^3) \\
& =N(0|\mathbf{u}(i)) (\alpha_2^3(1 - \alpha_2) + \alpha_2(1 - \alpha_2)^3) \\
& \quad +N(1|\mathbf{u}(i)) (\alpha_2^3(1 - \alpha_2) + \alpha_2(1 - \alpha_2)^3) \\
& =n (\alpha_2^3(1 - \alpha_2) + \alpha_2(1 - \alpha_2)^3). \tag{I.2c}
\end{aligned}$$

Using (I.2), it follows that for all $i \in \{1, 2, \dots, M\}$,

$$\begin{aligned}
\theta_i & =\Pr [B_n(\mathbf{Z}) < b | \mathbf{X} = \mathbf{u}(i)] \\
& =\Pr \left[\sum_{k=1}^n \mathbf{1} \{Z_k = 0\} < \frac{n(b - b_1)}{b_0 - b_1} \middle| \mathbf{X} = \mathbf{u}(i) \right] \\
& =\Pr \left[\sum_{t=1}^n (\mathbf{1} \{Z_t = 0\} - P_{Z|X}(0|u_t(i))) \right. \\
& \quad \left. \leq \sqrt{n\alpha_2(1 - \alpha_2)} \left(\frac{\frac{n(b-b_1)}{b_0-b_1} - n \left((1 - 2\alpha_2)\bar{P}_X^{(i)}(0) + \alpha_2 \right)}{\sqrt{n\alpha_2(1 - \alpha_2)}} \right) \middle| \mathbf{X} = \mathbf{u}(i) \right].
\end{aligned}$$

From the Berry-Esseen theorem (Theorem 20 in Appendix E), it follows that

$$\begin{aligned}
\theta_i & \geq Q \left(\frac{n \left((1 - 2\alpha_2)\bar{P}_X^{(i)}(0) + \alpha_2 - \frac{b-b_1}{b_0-b_1} \right)}{\sqrt{n\alpha_2(1 - \alpha_2)}} \right) - \frac{n [\alpha_2(1 - \alpha_2)^3 + (1 - \alpha_2)\alpha_2^3]}{2(n\alpha_2(1 - \alpha_2))^{3/2}} \\
& =Q \left(\frac{n \left((1 - 2\alpha_2)\bar{P}_X^{(i)}(0) + \alpha_2 - \frac{b-b_1}{b_0-b_1} \right)}{\sqrt{n\alpha_2(1 - \alpha_2)}} \right) - \frac{(\alpha_2(1 - \alpha_2)) \left((1 - \alpha_2)^2 + \alpha_2^2 \right)}{2\sqrt{n} (\alpha_2(1 - \alpha_2))^{3/2}} \tag{I.3}
\end{aligned}$$

$$=Q \left(\frac{n \left((1 - 2\alpha_2)\bar{P}_X^{(i)}(0) + \alpha_2 - \frac{b-b_1}{b_0-b_1} \right)}{\sqrt{n\alpha_2(1 - \alpha_2)}} \right) - \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}}, \tag{I.4}$$

and

$$\theta_i \leq Q \left(\frac{n \left((1 - 2\alpha_2)\bar{P}_X^{(i)}(0) + \alpha_2 - \frac{b-b_1}{b_0-b_1} \right)}{\sqrt{n\alpha_2(1 - \alpha_2)}} \right) + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}}.$$

This completes the proof.

I.1 Proof of Proposition 8

Using the fact that $Q^{-1}(\cdot)$ is a decreasing function, it follows from Lemma 6 that

$$\frac{n \left((1 - 2\alpha_2)\bar{P}_X^{(i)}(0) + \alpha_2 - \frac{b-b_1}{b_0-b_1} \right)}{\sqrt{n\alpha_2(1 - \alpha_2)}} \geq Q^{-1} \left(\delta + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}} \right). \tag{I.5}$$

From (I.5), it holds that

$$\frac{n(b - b_1)}{b_0 - b_1} \leq n \left((1 - 2\alpha_2)\bar{P}_X^{(i)}(0) + \alpha_2 \right) - \sqrt{n\alpha_2(1 - \alpha_2)} Q^{-1} \left(\delta + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}} \right). \tag{I.6}$$

Finally, the energy rate b is upper bounded as the following

$$b \leq (b_0 - b_1) \left((1 - 2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 \right) + b_1 - \sqrt{\frac{(b_0 - b_1)^2 \alpha_2 (1 - \alpha_2)}{n}} Q^{-1} \left(\delta + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}} \right),$$

which ends the proof of the energy bound (6.25). The proof of (6.26) follows immediately by following the definition of θ in (6.14) and using (I.4). This completes the proof. ■

J

Proof of Theorem 18

The proof of Theorem 18 is based on the random coding arguments.

Codebook Generation: Let $\rho \in [0, 1]$ be a fixed parameter. Consider a probability distribution P_X that satisfies

$$P_X(0) = 1 - P_X(1) = \rho. \quad (\text{J.1})$$

Let also $M \in \mathbb{N}$ and $b \in \mathbb{R}_+$ be fixed parameters. An (n, M) -code is randomly generated as follows: first, the codewords $\mathbf{u}(1), \mathbf{u}(2), \dots, \mathbf{u}(M)$ are realizations of a random variable \mathbf{X} following a distribution P_X such that for all $\mathbf{x} \in \mathcal{X}^n$,

$$P_X(\mathbf{x}) = \rho^{N(0|\mathbf{x})} (1 - \rho)^{n - N(0|\mathbf{x})}. \quad (\text{J.2})$$

Second, the decoding sets $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_M$ are defined using the information density function $\imath : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$, such that

$$\imath(\mathbf{x}, \mathbf{y}) = \log \left(\frac{P_{Y|X}(\mathbf{y}|\mathbf{x})}{P_Y(\mathbf{y})} \right). \quad (\text{J.3})$$

Using this notation, for all $\ell \in \{1, 2, \dots, M - 1\}$:

$$\mathcal{D}_\ell \triangleq \left\{ \mathbf{y} \in \mathcal{Y}^n : \ell \in \arg \max_{k \in \{1, 2, \dots, M\}} \imath(\mathbf{u}(k), \mathbf{y}) \right\} \setminus \bigcup_{j=1}^{\ell-1} \mathcal{D}_j, \text{ and} \quad (\text{J.4a})$$

$$\mathcal{D}_M \triangleq \mathcal{Y}^n \setminus \bigcup_{j=1}^{M-1} \mathcal{D}_j. \quad (\text{J.4b})$$

Given a generated (n, M) -code, the transmitter inputs the symbol $u_t(i)$ at channel use t , with $t \in \{1, 2, \dots, n\}$ to transmit the message index $i \in \{1, 2, \dots, M\}$. After n channel uses, the IR observes an n -dimensional channel output vector \mathbf{y} . The IR decides upon the index i following the rule in (6.7).

Decoding Error Probability Analysis: Let $\bar{\lambda}$ be the average over all possible codebooks of the DEP λ in (6.9). An immediate consequence of the random coding union bound [89] is

that the average DEP $\bar{\lambda}$ is upper bounded by:

$$\begin{aligned} \bar{\lambda} &< \mathbb{E}_{\mathbf{X}\mathbf{Y}} \left[\min \left\{ 1, (M-1) \Pr \left(\iota(\bar{\mathbf{X}}; \mathbf{Y}) \geq \iota(\mathbf{X}; \mathbf{Y}) \right) \right\} \right] \\ &= \sum_{\mathbf{x} \in \mathcal{X}^n} \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{\mathbf{X}}(\mathbf{x}) P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) \min \left\{ 1, (M-1) \Pr \left(\iota(\bar{\mathbf{X}}; \mathbf{y}) \geq \iota(\mathbf{x}; \mathbf{y}) \right) \right\}, \end{aligned} \quad (\text{J.5})$$

where the probability in (J.5) is with respect to the random variable $\bar{\mathbf{X}}$, whose probability mass function is $P_{\mathbf{X}}$ in (J.2). Note that for all $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, the following holds

$$\begin{aligned} \mathbb{1}_{\{\iota(\bar{\mathbf{x}}; \mathbf{y}) \geq \iota(\mathbf{x}; \mathbf{y})\}} &= \mathbb{1}_{\left\{ \log \frac{P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\bar{\mathbf{x}})}{P_{\mathbf{Y}}(\mathbf{y})} \geq \log \frac{P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})}{P_{\mathbf{Y}}(\mathbf{y})} \right\}} \\ &= \mathbb{1}_{\{d(\bar{\mathbf{x}}, \mathbf{y}) \log \alpha_1 + (n-d(\bar{\mathbf{x}}, \mathbf{y})) \log(1-\alpha_1) \geq d(\mathbf{x}, \mathbf{y}) \log \alpha_1 + (n-d(\mathbf{x}, \mathbf{y})) \log(1-\alpha_1)\}} \\ &= \mathbb{1}_{\{(d(\bar{\mathbf{x}}, \mathbf{y}) - d(\mathbf{x}, \mathbf{y})) \log \alpha_1 - (d(\bar{\mathbf{x}}, \mathbf{y}) - d(\mathbf{x}, \mathbf{y})) \log(1-\alpha_1)\}} \\ &= \mathbb{1}_{\left\{ (d(\bar{\mathbf{x}}, \mathbf{y}) - d(\mathbf{x}, \mathbf{y})) \log \left(\frac{\alpha_1}{1-\alpha_1} \right) \geq 0 \right\}} \end{aligned} \quad (\text{J.6})$$

$$= \mathbb{1}_{\{d(\bar{\mathbf{x}}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{y})\}}, \quad (\text{J.7})$$

where (J.6) follows from the fact that $\log \left(\frac{\alpha_1}{1-\alpha_1} \right) < 0$ for all $\alpha_1 \in [0, \frac{1}{2}]$. Now, from (J.7) it holds that for all $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$:

$$\begin{aligned} \Pr \left(\iota(\bar{\mathbf{X}}; \mathbf{y}) \geq \iota(\mathbf{x}; \mathbf{y}) \right) &= \sum_{\bar{\mathbf{x}} \in \mathcal{X}^n} \mathbb{1}_{\{d(\bar{\mathbf{x}}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{y})\}} P_{\mathbf{X}}(\bar{\mathbf{x}}) \\ &= \Pr \left(d(\bar{\mathbf{X}}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{y}) \right) \\ &= \sum_{\ell=0}^{d(\mathbf{x}, \mathbf{y})} \Pr \left(d(\bar{\mathbf{X}}, \mathbf{y}) = \ell \right), \end{aligned} \quad (\text{J.8})$$

where,

$$\begin{aligned} \Pr \left(d(\bar{\mathbf{X}}, \mathbf{y}) = \ell \right) &= \Pr \left(\sum_{t=0}^n \mathbb{1}_{\{\bar{X}_t \neq y_t\}} = \ell \right) \\ &= \Pr \left(\sum_{t \in \{m: y_m=0\}} \mathbb{1}_{\{\bar{X}_t=1\}} + \sum_{t \in \{m: y_m=1\}} \mathbb{1}_{\{\bar{X}_t=0\}} = \ell \right), \end{aligned} \quad (\text{J.10})$$

with $\ell \in \{0, 1, \dots, n\}$. Note that for all $\mathbf{y} \in \{0, 1\}^n$,

$$\sum_{t \in \{m: y_m=0\}} \mathbb{1}_{\{\bar{X}_t=1\}} \sim \text{Binomial}(N(0|\mathbf{y}), 1-\rho), \quad \text{and} \quad (\text{J.11})$$

$$\sum_{t \in \{m: y_m=1\}} \mathbb{1}_{\{\bar{X}_t=0\}} \sim \text{Binomial}(n - N(0|\mathbf{y}), \rho). \quad (\text{J.12})$$

Therefore, it holds that

$$\Pr \left(d(\bar{\mathbf{X}}, \mathbf{y}) = \ell \right) = \sum_{s=0}^{\ell} \Pr \left(\sum_{t \in \{m: y_m=0\}} \mathbb{1}_{\{\bar{X}_t=1\}} = s \right) \Pr \left(\sum_{t \in \{m: y_m=1\}} \mathbb{1}_{\{\bar{X}_t=0\}} = \ell - s \right)$$

$$\begin{aligned}
&= \sum_{s=0}^{\ell} \binom{N(0|\mathbf{y})}{s} (1-\rho)^s \rho^{N(0|\mathbf{y})-s} \binom{n-N(0|\mathbf{y})}{\ell-s} \rho^{\ell-s} (1-\rho)^{n-N(0|\mathbf{y})-\ell+s} \\
&= \sum_{s=0}^{\ell} \binom{N(0|\mathbf{y})}{s} \binom{n-N(0|\mathbf{y})}{\ell-s} \rho^{\ell+N(0|\mathbf{y})-2s} (1-\rho)^{n-(\ell+N(0|\mathbf{y})-2s)}. \quad (\text{J.13})
\end{aligned}$$

Plugging (J.13) in (J.9) yields

$$\begin{aligned}
\Pr(\iota(\bar{\mathbf{X}}; \mathbf{y}) \geq \iota(\mathbf{x}; \mathbf{y})) &= \sum_{\ell_1=0}^{d(\mathbf{x}, \mathbf{y})} \Pr(d(\bar{\mathbf{X}}, \mathbf{y}) = \ell_1) \\
&= \sum_{\ell_1=0}^{d(\mathbf{x}, \mathbf{y})} \sum_{\ell_2=0}^{\ell_1} \binom{N(0|\mathbf{y})}{\ell_2} \binom{n-N(0|\mathbf{y})}{\ell_1-\ell_2} \rho^{\ell_1+N(0|\mathbf{y})-2\ell_2} (1-\rho)^{n-(\ell_1+N(0|\mathbf{y})-2\ell_2)}. \quad (\text{J.14})
\end{aligned}$$

The term $\mathbb{E}_{\mathbf{X}\mathbf{Y}} \left[\Pr(\iota(\bar{\mathbf{X}}; \mathbf{Y}) \geq \iota(\mathbf{X}; \mathbf{Y})) \right]$ in (J.5) can be calculated as follows:

$$\begin{aligned}
&\mathbb{E}_{\mathbf{X}\mathbf{Y}} \left[\Pr(\iota(\bar{\mathbf{X}}; \mathbf{Y}) \geq \iota(\mathbf{X}; \mathbf{Y})) \right] \quad (\text{J.15}) \\
&= \sum_{\mathbf{y} \in \mathcal{Y}^n} \sum_{\mathbf{x} \in \mathcal{X}^n} P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) P_{\mathbf{X}}(\mathbf{x}) \Pr(\iota(\bar{\mathbf{X}}; \mathbf{y}) \geq \iota(\mathbf{x}; \mathbf{y})) \\
&= \sum_{\mathbf{y} \in \mathcal{Y}^n} \sum_{\mathbf{x} \in \mathcal{X}^n} (1-\alpha_1)^{d(\mathbf{x}, \mathbf{y})} \alpha_1^{n-d(\mathbf{x}, \mathbf{y})} \rho^{N(0|\mathbf{x})} (1-\rho)^{n-N(0|\mathbf{x})} \Pr(d(\bar{\mathbf{X}}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{y})).
\end{aligned}$$

There are 2^n vectors in \mathcal{X}^n and such a number can be written as follows for all $\mathbf{y} \in \mathcal{Y}^n$:

$$2^n = \sum_{\ell=0}^n \binom{n}{\ell} = \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\ell_1} \binom{N(0|\mathbf{y})}{\ell_2} \binom{n-N(0|\mathbf{y})}{\ell_1-\ell_2}, \quad (\text{J.16})$$

where the second equality follows from the Vandermonde's inequality. Hence, given an $\ell \leq n$, $\binom{n}{\ell}$ can be interpreted as the number of vectors in \mathcal{X}^n with ℓ zeros. Then, $\sum_{\ell=0}^n \binom{n}{\ell}$ is the number of all vectors in \mathcal{X}^n . To continue with this analysis, given a vector $\mathbf{y} \in \mathcal{Y}^n$ and a vector $\mathbf{x} \in \mathcal{X}^n$, let ℓ_1 be the number of zeros in \mathbf{x} ; and let also ℓ_2 be the number of zeros in \mathbf{x} that are in the same components in which the vector \mathbf{y} also has zeros. Hence, the number $\binom{N(0|\mathbf{y})}{\ell_2} \binom{n-N(0|\mathbf{y})}{\ell_1-\ell_2}$ can be interpreted as the number of vectors \mathbf{x} that contain exactly ℓ_1 zeros among which ℓ_2 zeros are placed on components in which \mathbf{y} also contains zeros. Therefore, $\sum_{\ell_2=0}^{\ell_1} \binom{N(0|\mathbf{y})}{\ell_2} \binom{n-N(0|\mathbf{y})}{\ell_1-\ell_2} = \binom{n}{\ell_1}$ is the number of vectors in \mathcal{X}^n with exactly ℓ_1 zeros. Using these interpretations, it holds that

$$N(0|\mathbf{x}) = \ell_1 \quad \text{and} \quad (\text{J.17})$$

$$d(\mathbf{x}, \mathbf{y}) = N(0|\mathbf{y}) - \ell_2 + \ell_1 - \ell_2 \quad (\text{J.18})$$

$$= N(0|\mathbf{y}) + \ell_1 - 2\ell_2, \quad (\text{J.19})$$

and

$$\mathbb{E}_{\mathbf{X}\mathbf{Y}} \left[\Pr(\iota(\bar{\mathbf{X}}; \mathbf{Y}) \geq \iota(\mathbf{X}; \mathbf{Y})) \right] \quad (\text{J.20})$$

$$\begin{aligned}
&= \sum_{\mathbf{y} \in \mathcal{Y}^n} \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\min\{N(0|\mathbf{y}), \ell_1\}} \binom{N(0|\mathbf{y})}{\ell_2} \binom{n-N(0|\mathbf{y})}{\ell_1-\ell_2} \alpha_1^{N(0|\mathbf{y})+\ell_1-2\ell_2} (1-\alpha_1)^{n-N(0|\mathbf{y})-\ell_1+2\ell_2} \\
&\cdot \rho^{\ell_1} (1-\rho)^{n-\ell_1} \Pr(d(\bar{\mathbf{X}}, \mathbf{y}) \leq N(0|\mathbf{y}) + \ell_1 - 2\ell_2) \\
&= \sum_{\mathbf{y} \in \mathcal{Y}^n} \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\min\{N(0|\mathbf{y}), \ell_1\}} \binom{N(0|\mathbf{y})}{\ell_2} \binom{n-N(0|\mathbf{y})}{\ell_1-\ell_2} \alpha_1^{N(0|\mathbf{y})+\ell_1-2\ell_2} (1-\alpha_1)^{n-N(0|\mathbf{y})-\ell_1+2\ell_2} \\
&\cdot \rho^{\ell_1} (1-\rho)^{n-\ell_1} \sum_{\ell_3=0}^{N(0|\mathbf{y})+\ell_1-2\ell_2} \sum_{\ell_4=0}^{\ell_3} \binom{N(0|\mathbf{y})}{\ell_4} \binom{n-N(0|\mathbf{y})}{\ell_3-\ell_4} \rho^{\ell_3+N(0|\mathbf{y})-2\ell_4} (1-\rho)^{n-(\ell_3+N(0|\mathbf{y})-2\ell_4)} \\
&= \sum_{\ell_0=0}^n \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\min\{\ell_0, \ell_1\}} \binom{n}{\ell_0} \binom{\ell_0}{\ell_2} \binom{n-\ell_0}{\ell_1-\ell_2} \alpha_1^{\ell_0+\ell_1-2\ell_2} (1-\alpha_1)^{n-\ell_0-\ell_1+2\ell_2} \rho^{\ell_1} (1-\rho)^{n-\ell_1} \\
&\cdot \sum_{\ell_3=0}^{\ell_0+\ell_1-2\ell_2} \sum_{\ell_4=0}^{\ell_3} \binom{\ell_0}{\ell_4} \binom{n-\ell_0}{\ell_3-\ell_4} \rho^{\ell_3+\ell_0-2\ell_4} (1-\rho)^{n-(\ell_3+\ell_0-2\ell_4)} \tag{J.21}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{\ell_0=0}^n \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\min\{\ell_0, \ell_1\}} \sum_{\ell_3=0}^{\ell_0+\ell_1-2\ell_2} \sum_{\ell_4=0}^{\ell_3} \binom{n}{\ell_0} \binom{\ell_0}{\ell_2} \binom{n-\ell_0}{\ell_1-\ell_2} \binom{\ell_0}{\ell_4} \binom{n-\ell_0}{\ell_3-\ell_4} \\
&\cdot \alpha_1^{\ell_0+\ell_1-2\ell_2} (1-\alpha_1)^{n-\ell_0-\ell_1+2\ell_2} \rho^{\ell_3+\ell_0-2\ell_4+\ell_1} (1-\rho)^{2n-(\ell_3+\ell_0-2\ell_4+\ell_1)}. \tag{J.22}
\end{aligned}$$

Replacing (J.22) into (J.5), leads to $\bar{\lambda} < \phi(m, \rho)$. This completes the proof of the bound on the information rate. The proof continues with the proof of the bound on the energy rate.

Energy-Shortage Probability Analysis: Consider an (n, M) -code described by the system in (6.5) generated using the probability mass function in (J.1). Hence, at an energy transmission rate b , it follows from Lemma 6 that for all $i \in \{1, 2, \dots, M\}$,

$$\theta_i \leq Q \left(\frac{(1-2\alpha_2)N(0|\mathbf{u}(i)) + n \left(\alpha_2 - \frac{b-b_1}{b_0-b_1} \right)}{\sqrt{n\alpha_2(1-\alpha_2)}} \right) + \frac{(1-\alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}}.$$

Let $\bar{\theta}_i$ be the average over all possible codebooks of the energy-shortage probability θ_i in (6.13) while transmitting at an energy rate b . Hence, the following holds:

$$\begin{aligned}
\bar{\theta}_i &= \sum_{\mathbf{x} \in \mathcal{X}^n} P_{\mathbf{X}}(\mathbf{x}) \Pr[B_n(\mathbf{Z}) < b | \mathbf{X} = \mathbf{x}] \\
&\leq \sum_{\mathbf{x} \in \mathcal{X}^n} P_{\mathbf{X}}(\mathbf{x}) Q \left(\frac{(1-2\alpha_2)N(0|\mathbf{x}) + n \left(\alpha_2 - \frac{b-b_1}{b_0-b_1} \right)}{\sqrt{n\alpha_2(1-\alpha_2)}} \right) + \frac{(1-\alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}} \\
&= \sum_{t=0}^n \binom{n}{t} \rho^t (1-\rho)^{n-t} Q \left(\frac{(1-2\alpha_2)t + n \left(\alpha_2 - \frac{b-b_1}{b_0-b_1} \right)}{\sqrt{n\alpha_2(1-\alpha_2)}} \right) + \frac{(1-\alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}} \\
&= \chi(b, \rho). \tag{J.23}
\end{aligned}$$

Hence from (J.23), it follows that for all $i \in \{1, 2, \dots, M\}$, $\bar{\theta} < \chi(b, \rho)$. This completes the proof. \blacksquare

K

Proof of Theorem 19

Consider an $(n, M, \epsilon, \delta, b)$ -code with maximum ESP described by the system in (6.5) and empirical input distribution

$$\bar{P}_X(0) = 1 - \bar{P}_X(1) \triangleq \rho. \quad (\text{K.1})$$

The proof is based on the notion of the *meta converse* introduced in [89]. Consider the following hypotheses:

$$\mathcal{H}_0 : (\mathbf{X}, \mathbf{Y}) \sim \bar{P}_X Q_Y \text{ and} \quad (\text{K.2a})$$

$$\mathcal{H}_1 : (\mathbf{X}, \mathbf{Y}) \sim \bar{P}_X P_{Y|X}, \quad (\text{K.2b})$$

where for all $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$

$$Q_Y(\mathbf{y}) = q^{N(0;\mathbf{y})} (1 - q)^{n - N(0;\mathbf{y})}, \quad (\text{K.3})$$

$$\bar{P}_X(\mathbf{x}) = \rho^{N(0|\mathbf{x})} (1 - \rho)^{n - N(0|\mathbf{x})}, \text{ and} \quad (\text{K.4})$$

$$P_{Y|X}(\mathbf{y}|\mathbf{x}) = \alpha_1^{d(\mathbf{x},\mathbf{y})} (1 - \alpha_1)^{n - d(\mathbf{x},\mathbf{y})}. \quad (\text{K.5})$$

The goal of the binary hypothesis test in (K.2) is to determine, based on the observation of $\mathbf{x} \in \mathcal{X}^n$ and $\mathbf{y} \in \mathcal{Y}^n$, whether these vectors are realizations of the random variables in hypothesis \mathcal{H}_0 or \mathcal{H}_1 . Consider a random transformation $P_{T|XY}$ from $\mathcal{X}^n \times \mathcal{Y}^n \rightarrow \{0, 1\}$. Note that this transformation can be a randomized test for the hypothesis test in (K.2). More specifically, $P_{T|X,Y}(1|\mathbf{x}, \mathbf{y}) = 1 - P_{T|X,Y}(0|\mathbf{x}, \mathbf{y})$ is the probability with which \mathcal{H}_1 is accepted given \mathbf{x} and \mathbf{y} . Define the function $\beta_{1-\epsilon} : \Delta(\mathcal{X}^n \times \mathcal{Y}^n)^2 \rightarrow [0, 1]$ by

$$\begin{aligned} \beta_{1-\epsilon}(P_{XY}, \bar{P}_X Q_Y) = \\ P_{T|XY}: \sum_{\mathbf{y} \in \mathcal{Y}^n} \sum_{\mathbf{x} \in \mathcal{X}^n} P_{T|XY}(1|\mathbf{x}, \mathbf{y}) P_{Y|X}(\mathbf{y}|\mathbf{x}) \bar{P}_X(\mathbf{x}) \geq 1 - \epsilon \left[\sum_{\mathbf{y} \in \mathcal{Y}^n} \sum_{\mathbf{x} \in \mathcal{X}^n} P_{T|XY}(1|\mathbf{x}, \mathbf{y}) \bar{P}_X(\mathbf{x}) Q_Y(\mathbf{y}) \right], \end{aligned} \quad (\text{K.6})$$

that is, the minimum probability of falsely rejecting \mathcal{H}_0 given that the probability of successfully accepting \mathcal{H}_1 is lower bounded by $1 - \epsilon$. Note that the corresponding log-likelihood ratio for

these hypotheses is for all $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$,

$$\begin{aligned} \log \frac{P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})}{Q_{\mathbf{Y}}(\mathbf{y})} &= \log \frac{(1 - \alpha_1)^{n-d(\mathbf{x}, \mathbf{y})} \alpha_1^{d(\mathbf{x}, \mathbf{y})}}{q^{N(0, \mathbf{y})} (1 - q)^{n-N(0, \mathbf{y})}} \\ &= (n - d(\mathbf{x}, \mathbf{y})) \log(1 - \alpha_1) + d(\mathbf{x}, \mathbf{y}) \log \alpha_1 - N(0|\mathbf{y}) \log(q) - (n - N(0|\mathbf{y})) \log(1 - q) \\ &= n \log \left(\frac{1 - \alpha_1}{1 - q} \right) + d(\mathbf{x}, \mathbf{y}) \log \frac{\alpha_1}{1 - \alpha_1} + N(0|\mathbf{y}) \log \frac{1 - q}{q}. \end{aligned}$$

Hence, from the Neyman-Pearson lemma, it follows that the optimal test is of the form

$$P_{T|\mathbf{X}\mathbf{Y}}^*(1|\mathbf{x}, \mathbf{y}) = \begin{cases} 0 & \text{if } d(\mathbf{x}, \mathbf{y}) \log \frac{\alpha_1}{1 - \alpha_1} + N(0|\mathbf{y}) \log \frac{1 - q}{q} < L \\ \lambda & \text{if } d(\mathbf{x}, \mathbf{y}) \log \frac{\alpha_1}{1 - \alpha_1} + N(0|\mathbf{y}) \log \frac{1 - q}{q} = L \\ 1 & \text{if } d(\mathbf{x}, \mathbf{y}) \log \frac{\alpha_1}{1 - \alpha_1} + N(0|\mathbf{y}) \log \frac{1 - q}{q} > L, \end{cases} \quad (\text{K.7})$$

where the constants $\lambda \in [0, 1]$, and $L \in \mathbb{R}$ are chosen to satisfy:

$$\begin{aligned} 1 - \epsilon &= \sum_{\mathbf{y} \in \mathcal{Y}^n} \sum_{\mathbf{x} \in \mathcal{X}^n} P_{T|\mathbf{X}\mathbf{Y}}^*(1|\mathbf{x}, \mathbf{y}) P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) \bar{P}_{\mathbf{X}}(\mathbf{x}) \\ &= \sum_{\mathbf{y} \in \mathcal{Y}^n} \sum_{\mathbf{x} \in \mathcal{X}^n} (1 - \alpha_1)^{n-d(\mathbf{x}, \mathbf{y})} \alpha_1^{d(\mathbf{x}, \mathbf{y})} \rho^{N(0|\mathbf{x})} (1 - \rho)^{N(0|\mathbf{x})} \\ &\quad \cdot \left(\mathbb{1}_{\left\{d(\mathbf{x}, \mathbf{y}) \log \frac{\alpha_1}{1 - \alpha_1} + N(0|\mathbf{y}) \log \frac{1 - q}{q} > L\right\}} + \lambda \mathbb{1}_{\left\{d(\mathbf{x}, \mathbf{y}) \log \frac{\alpha_1}{1 - \alpha_1} + N(0|\mathbf{y}) \log \frac{1 - q}{q} = L\right\}} \right). \end{aligned} \quad (\text{K.8})$$

The equality in (K.8) can be rewritten by noting that there are 2^n vectors in \mathcal{X}^n and for all $\mathbf{y} \in \mathcal{Y}^n$:

$$2^n = \sum_{\ell=0}^n \binom{n}{\ell} = \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\ell_1} \binom{N(0|\mathbf{y})}{\ell_2} \binom{n - N(0|\mathbf{y})}{\ell_1 - \ell_2}, \quad (\text{K.9})$$

where the second equality follows from the Vandermonde's inequality. Given a vector $\mathbf{y} \in \mathcal{Y}^n$ and a vector $\mathbf{x} \in \mathcal{X}^n$, let ℓ_1 be the number of zeros in \mathbf{x} ; and let also ℓ_2 be the number of zeros in \mathbf{x} that are in the same components in which the vector \mathbf{y} also has zeros. Hence, the number $\binom{N(0|\mathbf{y})}{\ell_2} \binom{n - N(0|\mathbf{y})}{\ell_1 - \ell_2}$ can be interpreted as the number of vectors \mathbf{x} that contain exactly ℓ_1 zeros among which ℓ_2 zeros are placed on components in which \mathbf{y} also contains zeros. Therefore, $\sum_{\ell_2=0}^{\ell_1} \binom{N(0|\mathbf{y})}{\ell_2} \binom{n - N(0|\mathbf{y})}{\ell_1 - \ell_2} = \binom{n}{\ell_1}$ is the number of vectors in \mathcal{X}^n with exactly ℓ_1 zeros. Using these interpretations, it holds that

$$N(0|\mathbf{x}) = \ell_1 \quad \text{and} \quad (\text{K.10})$$

$$d(\mathbf{x}, \mathbf{y}) = N(0|\mathbf{y}) - \ell_2 + \ell_1 - \ell_2 \quad (\text{K.11})$$

$$= N(0|\mathbf{y}) + \ell_1 - 2\ell_2, \quad (\text{K.12})$$

and thus, the equality in (K.8) can be rewritten as follows:

$$\begin{aligned} 1 - \epsilon &= \sum_{\mathbf{y} \in \mathcal{Y}^n} \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\ell_1} \binom{N(0|\mathbf{y})}{\ell_2} \binom{n - N(0|\mathbf{y})}{\ell_1 - \ell_2} (1 - \alpha_1)^{n - (N(0|\mathbf{y}) + \ell_1 - 2\ell_2)} \alpha_1^{N(0|\mathbf{y}) + \ell_1 - 2\ell_2} \rho^{\ell_1} (1 - \rho)^{n - \ell_1} \\ &\quad \cdot \left(\mathbb{1}_{\left\{(N(0|\mathbf{y}) + \ell_1 - 2\ell_2) \log \frac{\alpha_1}{1 - \alpha_1} + N(0|\mathbf{y}) \log \frac{1 - q}{q} > L\right\}} + \lambda \mathbb{1}_{\left\{(N(0|\mathbf{y}) + \ell_1 - 2\ell_2) \log \frac{\alpha_1}{1 - \alpha_1} + N(0|\mathbf{y}) \log \frac{1 - q}{q} = L\right\}} \right) \end{aligned}$$

$$\begin{aligned}
&= \sum_{\ell_0=0}^n \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\ell_1} \binom{n}{\ell_0} \binom{\ell_0}{\ell_2} \binom{n-\ell_0}{\ell_1-\ell_2} (1-\alpha_1)^{n-(\ell_0+\ell_1-2\ell_2)} \alpha_1^{\ell_0+\ell_1-2\ell_2} \rho^{\ell_1} (1-\rho)^{n-\ell_1} \\
&\quad \cdot \left(\mathbb{1}_{\{(\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} > L\}} + \lambda \mathbb{1}_{\{(\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} = L\}} \right).
\end{aligned}$$

Plugging (K.7) into (K.6) yields

$$\begin{aligned}
\beta_{1-\epsilon}(P_{\mathbf{X}\mathbf{Y}}, \bar{P}_{\mathbf{X}} Q_{\mathbf{Y}}) &= \sum_{\mathbf{y} \in \mathcal{Y}^n} \sum_{\mathbf{x} \in \mathcal{X}^n} P_{T|\mathbf{X}\mathbf{Y}}^*(1|\mathbf{x}, \mathbf{y}) Q_{\mathbf{Y}}(\mathbf{y}) \bar{P}_{\mathbf{X}}(\mathbf{x}) \\
&= \sum_{\mathbf{y} \in \mathcal{Y}^n} \sum_{\mathbf{x} \in \mathcal{X}^n} \rho^{N(0|\mathbf{x})} (1-\rho)^{n-N(0|\mathbf{x})} q^{N(0|\mathbf{y})} (1-q)^{n-N(0|\mathbf{y})} \\
&\quad \left(\mathbb{1}_{\{d(\mathbf{x}, \mathbf{y}) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\mathbf{y}) \log \frac{1-q}{q} > L\}} + \lambda \mathbb{1}_{\{d(\mathbf{x}, \mathbf{y}) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\mathbf{y}) \log \frac{1-q}{q} = L\}} \right) \\
&= \sum_{\mathbf{y} \in \mathcal{Y}^n} \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\ell_1} \binom{N(0|\mathbf{y})}{\ell_2} \binom{n-N(0|\mathbf{y})}{\ell_1-\ell_2} \rho^{\ell_1} (1-\rho)^{n-\ell_1} q^{N(0|\mathbf{y})} (1-q)^{n-N(0|\mathbf{y})} \\
&\quad \left(\mathbb{1}_{\{(N(0|\mathbf{y})+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\mathbf{y}) \log \frac{1-q}{q} > L\}} \right. \\
&\quad \left. + \lambda \mathbb{1}_{\{(N(0|\mathbf{y})+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\mathbf{y}) \log \frac{1-q}{q} = L\}} \right) \\
&= \sum_{\ell_0=0}^n \sum_{\ell_1=0}^n \sum_{\ell_2=0}^{\ell_1} \binom{n}{\ell_0} \binom{\ell_0}{\ell_2} \binom{n-\ell_0}{\ell_1-\ell_2} \rho^{\ell_1} (1-\rho)^{n-\ell_1} q^{\ell_0} (1-q)^{n-\ell_0} \\
&\quad \left(\mathbb{1}_{\{(\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} > L\}} + \lambda \mathbb{1}_{\{(\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} = L\}} \right) \\
&\triangleq \gamma(\rho, q).
\end{aligned}$$

Finally, from Theorem 29 in [89], it follows that

$$M < \Gamma(b), \quad (\text{K.13})$$

where the function Γ is defined in (6.49) and the optimization domain over ρ is due to Corollary 4, which requires that $\rho > \rho^*(b)$. This completes the proof of the information bound.

The proof continues with the proof of the bound on the energy rate. From Proposition 8, subject to a maximum ESP constraint, it follows that for any $(n, M, \epsilon, \delta, b)$ -code described by the system in (6.5) for the random transformation in (6.1) satisfying (6.12), it holds that for all $i \in \{1, 2, \dots, M\}$:

$$b \leq (b_0 - b_1) \left((1-2\alpha_2) \bar{P}_X^{(i)}(0) + \alpha_2 \right) + b_1 - \sqrt{\frac{(b_0 - b_1)^2 \alpha_2 (1-\alpha_2)}{n}} Q^{-1} \left(\delta + \frac{(1-\alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}} \right) \quad (\text{K.14})$$

$$= (1-\alpha_2)b_0 + \alpha_2 b_1 - \sqrt{\frac{(b_0 - b_1)^2 \alpha_2 (1-\alpha_2)}{n}} \cdot Q^{-1} \left(\delta + \frac{(1-\alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}} \right) \quad (\text{K.15})$$

$$= B^+. \quad (\text{K.16})$$

From (K.13) and (K.16), it follows that the information rate $\frac{\log M}{n}$ and the energy rate b of any $(n, M, \epsilon, \delta, b)$ -code for the random transformation in (6.1) satisfying (6.12) subject to a

K. Proof of Theorem 19

maximum ESP constraint are such that, $M < \Gamma(b)$ and $b < B^+$. This completes the proof of Theorem 19. ■

Bibliography

- [1] Sam Davis. Solar system efficiency: Maximum power point tracking is key, Fact Sheet N°282, 2015. <https://www.powerelectronics.com/solar/solar-system-efficiency-maximum-power-point-tracking-key>, Last accessed on 2015-12-15. *2 citations in pages viii and 16*
- [2] Mark Hughes . Wireless rf energy harvesting: Rf-to-dc conversion and a look at powercast hardware, Fact Sheet N°282, 2018. <https://www.allaboutcircuits.com/technical-articles/wireless-rf-energy-harvesting-rf-to-dc-conversion-powercast-hardware/>, Last accessed on 2018-08-03. *2 citations in pages viii and 17*
- [3] X. Lu, I. Flint, D. Niyato, N. Privault, and P. Wang. Self-sustainable communications with rf energy harvesting: Ginibre point process modeling and analysis. *IEEE Journal on Selected Areas in Communications*, 34(5):1518–1535, May 2016. *3 citations in pages viii, 20, and 21*
- [4] 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*, 63(8):5244–5265, Aug. 2017. *9 citations in pages viii, 27, 28, 29, 30, 31, 37, 38, and 52*
- [5] C. E. Shannon. The zero-error capacity of a noisy channel. *IRE Transactions on Information Theory*, 2(3):8–19, Sep. 1956. *Cited in page 8*
- [6] Robert Fano. *Transmission of information - a statistical theory of communication*. MIT Press, Mar. 1961. *Cited in page 8*
- [7] W. Feller. *An Introduction to Probability Theory and Its Application*, volume 2. John Wiley and Sons, New York, NY, 2nd edition, 1971. *2 citations in pages 8 and 106*
- [8] I. Csiszár and J. Körner. *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic Press, 2nd edition, 1981. *2 citations in pages 11 and 55*
- [9] J.D. Kraus. *Electromagnetics*. McGraw-Hill, 1992. *Cited in page 15*

- [10] B. Thide. *Electromagnetic field theory*. Upsilon Books Uppsala, 2004. *Cited in page 16*
- [11] G.A. Kendir, W. Liu, and M. Ghovanloo. Design and optimization of a 3-coil inductive link for efficient wireless power transmission. *IEEE Transactions on Biomedical Circuits and Systems*, 5(6):579–591, 2011. *Cited in page 17*
- [12] G. A. Kendir, W. Liu, G. Wang, M. Sivaprakasam, R. Bashirullah, M. S. Humayun, and J. D. Weiland. An optimal design methodology for inductive power link with class-e amplifier. *IEEE Trans. Circuits Syst*, 52(5):857–866, 2005. *Cited in page 17*
- [13] Y. Zeng, B. Clerckx, and R. Zhang. Communications and signals design for wireless power transmission. *IEEE Transactions on Communications*, 65(5):2264–2290, May 2017. *2 citations in pages 18 and 19*
- [14] N. Tesla. Apparatus for transmitting electrical energy., December 1 1914. US Patent 1,119,732. *2 citations in pages 19 and 51*
- [15] G. Goubau and F. Schwing. On the guided propagation of electromagnetic wave beams. *IRE Transactions on Antennas and Propagation*, 9(3):248–256, 1961. *Cited in page 19*
- [16] L. Chen, J. Wang, J. Zhou, D. W. K. Ng, R. Schober, , and C. Zhao. Distributed user-centric scheduling for visible light communication networks. *Optics Express*, 24:15570–15589, 2016. *Cited in page 19*
- [17] P. Glaser. An overview of the solar power satellite option. *IEEE Transactions on Microwave Theory and Techniques*, 40(6):1230–1238, 1992. *Cited in page 19*
- [18] S. Bi, C. K. Ho, and R. Zhang. Wireless powered communication: opportunities and challenges. *IEEE Communications Magazine*, 53(4):117–125, Apr. 2015. *Cited in page 19*
- [19] I. Krikidis, S. Timotheou, S. Nikolaou, G. Zheng, D. W. K. Ng, and R. Schober. Simultaneous wireless information and power transfer in modern communication systems. *IEEE Communications Magazine*, 52(11):104–110, Nov. 2014. *2 citations in pages 20 and 37*
- [20] R. Zhang and C. K. Ho. MIMO broadcasting for simultaneous wireless information and power transfer. *IEEE Transactions on Wireless Communications*, 12(5):1989–2001, May 2013. *2 citations in pages 20 and 21*
- [21] H. Lee, K. Lee, H. Kim, and I. Lee. Joint transceiver optimization for miso swipt systems with time switching. *IEEE Transactions on Wireless Communications*, 17(5):3298–3312, May 2018. *Cited in page 20*

-
- [22] X. Zhou. Training-based swipt: Optimal power splitting at the receiver. *IEEE Transactions on Vehicular Technology*, 64(9):4377–4382, Sep. 2015.
Cited in page 20
- [23] K. Lee and J. Ko. Wireless information and power transfer: Probability-based power allocation and splitting with low complexity. *IEEE Systems Journal*, 12(1):1060–1064, March 2018.
Cited in page 20
- [24] D. Xu and H. Zhu. Outage minimized resource allocation for multiuser ofdm systems with swipt. *IEEE Access*, 7:79714–79725, 2019.
Cited in page 21
- [25] J. Tang, D. K. C. So, A. Shojaeifard, K. Wong, and J. Wen. Joint antenna selection and spatial switching for energy efficient mimo swipt system. *IEEE Transactions on Wireless Communications*, 16(7):4754–4769, July 2017.
Cited in page 21
- [26] I. Krikidis, S. Sasaki, S. Timotheou, and Z. Ding. A low complexity antenna switching for joint wireless information and energy transfer in mimo relay channels. *IEEE Transactions on Communications*, 62(5):1577–1587, May 2014.
Cited in page 21
- [27] P. Popovski, A. M. Fouladgar, and O. Simeone. Interactive joint transfer of energy and information. *IEEE Transactions on Communications*, 61(5):2086–2097, May 2013.
2 citations in pages 23 and 37
- [28] L. R. Varshney. Transporting information and energy simultaneously. In *Proc. IEEE International Symposium on Information Theory*, pages 1612–1616, Toronto, ON, Canada, Jul. 2008.
4 citations in pages 26, 37, 41, and 52
- [29] Joel G. Smith. The information capacity of amplitude- and variance-constrained scalar Gaussian channels. *Information and Control*, 18(3):203–219, 1971.
2 citations in pages 26 and 27
- [30] Ali Mohammad Fouladgar; and Osvaldo Simeone. On the transfer of information and energy in multi-user systems. *IEEE Communications Letters*, 16(11):1733–1736, Nov. 2012.
3 citations in pages 27, 37, and 52
- [31] Selma Belhadj Amor, Samir M. Perlaza, and Ioannis Krikidis. Simultaneous energy and information transmission in Gaussian multiple access channels. In *Proc. 5th International Conference on Communications and Networking (ComNet)*, Hammamet, Tunisia, Nov. 2015.
Cited in page 27
- [32] S. Belhadj Amor, S. M. Perlaza, and H. V. Poor. Decentralized simultaneous energy and information transmission in multiple access channels. Technical report, INRIA, Lyon, France, Apr. 2016.
Cited in page 27

- [33] T. M. Cover. Some advances in broadcast channels. In A. Viterbi, editor, *Advances in Communication Systems, Vol. 4, Theory and Applications*, chapter 4. Academic Press, New York, 1975.
Cited in page 30
- [34] A. D. Wyner. Recent results in the Shannon theory. *IEEE Transactions on Information Theory*, 20(1):2–10, 1974.
Cited in page 30
- [35] L. H. Ozarow. The capacity of the white Gaussian multiple access channel with feedback. *IEEE Transactions on Information Theory*, 30(4):623–629, Jul. 1984.
Cited in page 31
- [36] A. Carleial. A case where interference does not reduce capacity. *IEEE Trans. Inf. Theory*, 21(5):569–570, Sep. 1975.
Cited in page 34
- [37] T. S. Han and K. Kobayashi. A new achievable rate region for the interference channel. *IEEE Trans. Inf. Theory*, 27(1):49–60, 1981. *3 citations in pages 34, 35, and 47*
- [38] H. Sato. The capacity of the Gaussian interference channel under strong interference. *IEEE Trans. Inf. Theory*, 27(6):786–788, Nov. 1981.
Cited in page 34
- [39] R. Ahlswede. Multi-way communication channels. In *Proc. 2nd International Symposium on Information Theory*, pages 23–52, Tsahkadsor, Armenian S.S.R. Hungarian Academy of Sciences, Budapest, Sep. 1971.
Cited in page 34
- [40] R. Ahlswede. The capacity region of a channel with two senders and two receivers. *The Annals of Probability*, 2(5):805–814, Oct. 1974.
Cited in page 34
- [41] 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*, 54(7):3188–3195, Jul. 2008.
2 citations in pages 34 and 73
- [42] A. Carleial. Interference channels. *IEEE Trans. Inf. Theory*, 24(1):60–70, Sep. 1978.
2 citations in pages 34 and 35
- [43] Thomas M. Cover and Cyril S. K. Leung. An achievable rate region for the multiple-access channel with feedback. *IEEE Transactions on Information Theory*, 27(3):292–298, May 1981.
2 citations in pages 34 and 35
- [44] R. H. Etkin, D. N. C. Tse, and Wang Hua. Gaussian interference channel capacity to within one bit. *IEEE Trans. Inf. Theory*, 54(12):5534–5562, Dec. 2008.
2 citations in pages 34 and 35
- [45] G. Kramer. Outer bounds on the capacity of Gaussian interference channels. *IEEE Trans. Inf. Theory*, 50(3):581–586, Mar. 2004.
Cited in page 34

-
- [46] Abolfazl Seyed Motahari and Amir Keyvan Khandani. Capacity bounds for the Gaussian interference channel. *IEEE Trans. Inf. Theory*, 55(2):620–643, Feb. 2009.
Cited in page 34
- [47] H. Sato. Two-user communication channels. *IEEE Trans. Inf. Theory*, 23(3):295–304, 1977.
Cited in page 34
- [48] Xiaohu Shang, Gerhard Kramer, and Biao Chen. A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels. *IEEE Trans. Inf. Theory*, 55(2):689–699, Feb. 2009.
Cited in page 34
- [49] C. Suh and D. N. C. Tse. Feedback capacity of the Gaussian interference channel to within 2 bits. *IEEE Trans. Inf. Theory*, 57(5):2667–2685, May 2011.
5 citations in pages 35, 36, 38, 46, and 97
- [50] Frans M. J. Willems and Edward C. Van Der Meulen. The discrete memoryless multiple-access channel with cribbing encoders. *IEEE Trans. Inf. Theory*, IT-31(3):313–327, May. 1985.
Cited in page 35
- [51] Nikola Tesla. Apparatus for transmitting electrical energy, Dec. 1914.
Cited in page 37
- [52] P. Grover and A. Sahai. Shannon meets Tesla: Wireless information and power transfer. In *Proc. IEEE International Symposium on Information Theory*, pages 2363–2367, Austin, TX, USA, Jun. 2010.
3 citations in pages 37, 41, and 52
- [53] Morteza Varasteh, Borzoo Rassouli, and Bruno Clerckx. Wireless information and power transfer over an AWGN channel: Nonlinearity and asymmetric Gaussian signaling. In *Proc. IEEE Information Theory Workshop*, Kaohsiung, Taiwan, Nov. 2017.
Cited in page 37
- [54] L. R. Varshney. On energy/information cross-layer architectures. In *Proc. IEEE International Symposium on Information Theory*, pages 1356–1360, Cambridge, MA, USA, Jul. 2012.
Cited in page 37
- [55] Samir M. Perlaza, Ali Tajer, and H. Vincent Poor. Simultaneous energy and information transmission: A finite block-length analysis. In *Proc. 19th IEEE International Workshop on Signal Processing Advances in Wireless Communication*, Kalamata, Greece, Jun. 2018.
Cited in page 37
- [56] K Gautam Shenoy and Vinod Sharma. Finite blocklength achievable rates for energy harvesting AWGN channels with infinite buffer. In *Proc. IEEE International Symposium on Information Theory (ISIT)*, Barcelona, Spain, Jul. 2016.
Cited in page 37
- [57] Shen Chao, Wei-Chiang Li, and Tsung-Hui Chang. Simultaneous information and energy transfer: A two-user MISO interference channel case. In *Proc. Signal Processing for*

- Communications Symposium*, pages 3862–3867, Anaheim, CA, USA, Dec. 2012.
Cited in page 37
- [58] Shen Chao, Wei-Chiang Li, and Tsung-Hui Chang. Wireless information and energy transfer in multi-antenna interference channel. *IEEE Transactions on Signal Processing*, 62(23):6249–6264, Dec. 2014.
Cited in page 37
- [59] J. Park and B. Clerckx. Joint wireless information and energy transfer in a two-user MIMO interference channel. *IEEE Trans. Wireless Commun.*, 12(8):4210–4221, Aug. 2013.
Cited in page 37
- [60] D. Sacristan-Murga and D. Guenduez. Joint energy and information transmission in a two-user symmetric interference channel. In *European Wireless Conference*, pages 1–6, 2014.
Cited in page 37
- [61] S. Belhadj Amor and Samir M. Perlaza. Fundamental limits of simultaneous energy and information transmission. In *Proc. International Symposium on Telecommunications*, Thessaloniki, Greece, May 2016. *3 citations in pages 37, 51, and 52*
- [62] Z. Ding, S. M. Perlaza, I. Esnaola, and H. V. Poor. Power allocation strategies in energy harvesting wireless cooperative networks. *IEEE Transactions on Wireless Communications*, 13(2):846–860, Feb. 2014.
Cited in page 37
- [63] M. Gastpar. Gaussian multiple-access channels under received-power constraints. In *Proc. IEEE Information Theory Workshop*, pages 452–457, 2004.
Cited in page 37
- [64] S. Ulukus, A. Yener, E. Erkip, O. Simeone, M. Zorzi, P. Grover, and K. Huang. Energy harvesting wireless communications: A review of recent advances. *IEEE Journal on Selected Areas in Communications*, 33(3):360–381, Mar. 2015.
Cited in page 37
- [65] R. H. Etkin, D. N. C. Tse, and H. Wang. Gaussian interference channel capacity to within one bit. *IEEE Trans. Inf. Theory*, 54(12):5534–5562, Dec. 2008. *3 citations in pages 38, 44, and 47*
- [66] T. S. Han and K. Kobayashi. A new achievable rate region for the interference channel. *IEEE Trans. Inf. Theory*, 27:49–60, Jan. 1981.
Cited in page 38
- [67] Victor Quintero, Samir M. Perlaza, Iñaki Esnaola, and Jean-Marie Gorce. Approximate capacity region of the two-user Gaussian interference channel with noisy channel-output feedback. *IEEE Trans. Inf. Theory*, 64(7):5326 – 5358, Jul. 2018.
Cited in page 38
- [68] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration Inequalities : A Nonasymptotic Theory of Independence*. Oxford University Press, New York, USA, Mar.

2013.

Cited in page 38

- [69] J. Park and B. Clerckx. Joint wireless information and energy transfer in a two-user MIMO interference channel. *IEEE Transactions on Wireless Communications*, 12(8):4210–4221, Aug. 2013.

Cited in page 41

- [70] K. Huang and E. Larsson. Simultaneous information and power transfer for broadband wireless systems. *IEEE Transactions on Signal Processing*, 61(23):5972–5986, Dec 2013.

Cited in page 41

- [71] D. W. K. Ng, E. S. Lo, and R. Schober. Wireless information and power transfer: Energy efficiency optimization in ofdma systems. *IEEE Transactions on Wireless Communications*, 12(12):6352–6370, December 2013.

Cited in page 41

- [72] F. Gholami, H. Meghdadi, and A. Shahzadi. Throughput analysis for decode-and-forward relaying protocol with wireless energy harvesting and information processing. In *2018 International Conference on Internet of Things, Embedded Systems and Communications (IINTEC)*, pages 132–135, Dec 2018.

Cited in page 41

- [73] Y. Huang and B. Clerckx. Joint wireless information and power transfer for an autonomous multiple antenna relay system. *IEEE Communications Letters*, 19(7):1113–1116, July 2015.

Cited in page 41

- [74] H. Lee, K. Lee, H. Kim, B. Clerckx, and I. Lee. Resource allocation techniques for wireless powered communication networks with energy storage constraint. *IEEE Transactions on Wireless Communications*, 15(4):2619–2628, April 2016.

Cited in page 41

- [75] C. Boyer and S. Roy. Backscatter communication and rfid: Coding, energy, and mimo analysis. *IEEE Transactions on Communications*, 62(3):770–785, March 2014.

Cited in page 41

- [76] G. Yang, C. K. Ho, and Y. L. Guan. Multi-antenna wireless energy transfer for backscatter communication systems. *IEEE Journal on Selected Areas in Communications*, 33(12):2974–2987, Dec 2015.

Cited in page 41

- [77] M. S. Trotter, J. D. Griffin, and G. D. Durgin. Power-optimized waveforms for improving the range and reliability of rfid systems. In *2009 IEEE International Conference on RFID*, pages 80–87, April 2009.

Cited in page 41

- [78] B. Clerckx. Waveform optimization for swipt with nonlinear energy harvester modeling. In *WSA 2016; 20th International ITG Workshop on Smart Antennas*, pages 1–5, March 2016.

Cited in page 41

- [79] B. Clerckx, R. Zhang, R. Schober, D. W. K. Ng, D. I. Kim, and H. V. Poor. Fundamentals of wireless information and power transfer: From rf energy harvester models to signal and system designs. *IEEE Journal on Selected Areas in Communications*, 37(1):4–33, Jan 2019.
Cited in page 41
- [80] D. I. Kim, J. H. Moon, and J. J. Park. New swipt using papr: How it works. *IEEE Wireless Communications Letters*, 5(6):672–675, Dec 2016.
Cited in page 41
- [81] M. Varasteh, B. Rassouli, and B. Clerckx. Wireless information and power transfer over an awgn channel: Nonlinearity and asymmetric gaussian signaling. In *2017 IEEE Information Theory Workshop (ITW)*, pages 181–185, Nov 2017.
Cited in page 41
- [82] R. Morsi, V. Jamali, D. W. K. Ng, and R. Schober. On the capacity of swipt systems with a nonlinear energy harvesting circuit. In *2018 IEEE International Conference on Communications (ICC)*, pages 1–7, May 2018.
Cited in page 41
- [83] J. Kang, I. Kim, and D. I. Kim. Wireless information and power transfer: Rate-energy tradeoff for nonlinear energy harvesting. *IEEE Transactions on Wireless Communications*, 17(3):1966–1981, March 2018.
Cited in page 41
- [84] M. Varasteh, B. Rassouli, H. Joudeh, and B. Clerckx. Swipt signalling over complex awgn channels with two nonlinear energy harvester models. In *2018 IEEE International Symposium on Information Theory (ISIT)*, pages 866–870, June 2018.
Cited in page 41
- [85] M. Varasteh, B. Rassouli, and B. Clerckx. Wireless information and power transfer over an awgn channel: Nonlinearity and asymmetric gaussian signaling. In *2017 IEEE Information Theory Workshop (ITW)*, pages 181–185, Nov 2017.
Cited in page 41
- [86] Faisal Tariq, Muhammad R. A. Khandaker, Kai-Kit Wong, Muhammad Imran, Mehdi Bennis, and Mérouane Debbah. A speculative study on 6G. *online: <https://arxiv.org/pdf/1902.06700.pdf>*, Feb. 2019.
Cited in page 51
- [87] N. Khalfet and S. M. Perlaza. Simultaneous Information and Energy Transmission in the Two-User Gaussian Interference Channel. *IEEE Journal on Selected Areas in Communications*, 37(1):156–170, Jan. 2019.
Cited in page 52
- [88] S. M. Perlaza, A. Tajer, and H. V. Poor. Simultaneous energy and information transmission: A finite block-length analysis. In *Proc. of the 19th IEEE International Workshop on Signal Processing Advances in Wireless Communication*, Kalamata, Greece, Jun. 2018.
Cited in page 52

- [89] Yury Polyanskiy. *Channel Coding: Non-Asymptotic Fundamental Limits*. PhD thesis, Electrical Engineering, Princeton University, Princeton, NJ, USA, Nov. 2010.
5 citations in pages 52, 58, 119, 123, and 125
- [90] Y. Polyanskiy, H. V. Poor, and S. Verdú. Channel coding rate in the finite blocklength regime. *IEEE Trans. Inf. Theory*, 56(5):2307–2359, May 2010.
Cited in page 52
- [91] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*. New York: Springer, 2nd edition, 1998.
Cited in page 56
- [92] R. Zhang and C. Keong. Ho. MIMO broadcasting for simultaneous wireless information and power transfer. *IEEE Transactions on Wireless Communications*, 12(5):1989–2001, May. 2013.
Cited in page 63
- [93] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience Publication, Hoboken, NJ, USA, 2005.
2 citations in pages 66 and 90
- [94] Rick Durrett. *Probability Theory and Examples*. Cambridge University Press, New York, NY, 2010.
Cited in page 80

