Réseaux coopératifs avec incertitude du canal

Arash Behboodi

N° d'ordre : 2012-08-TH

THÈSE DE DOCTORAT

DOMAINE : STIC

SPECIALITE : Télécommunications

Ecole Doctorale « Sciences et Technologies de l'Information des Télécommunications et des Systèmes »

Présentée par :

Arash BEHBOODI

Sujet :

Réseaux Coopératifs avec Incertitude du Canal

Cooperative Networks with Channel Uncertainty

Soutenue le le 13 Juin 2012 devant les membres du jury :

Prof. Giuseppe CAIRE	University of Southern California	Rapporteur
Prof. Michael C. GASTPAR	École Polytechnique Fédérale de Lausanne	Rapporteur
Prof. Jean Claude BELFIORE	Institut Télécom/Télécom ParisTech//CNRS LTCI	Examinateur
Prof. Bartek BLASZCZYSZYN	INRIA- TREC/ ENS de Paris	Examinateur
Prof. Pierre DUHAMEL	L2S CNRS/SUPELEC/Univ. Paris-Sud 11	Examinateur
Prof. Dirk SLOCK	EURECOM	Examinateur
Dr. Pablo PIANTANIDA	SUPELEC	Encadrant
Dr. Antoine BERTHET	SUPELEC	Invité (Directeur de thèse)

To those waiting *before the Law*

Abstract

In this thesis, we focus on cooperative networks where the transmitter is uncertain about the channel in operation. The main contributions are organized in three chapters.

In the first chapter, cooperative strategies are developed for *simultaneous relay channels* (SRC) which consist of a set of two single relay channels out of which the channel in operation is chosen. Provided that the channel uncertainty involves a limited number of possibilities, this problem is recognized as being equivalent to that of sending common and private information to several destinations in presence of helper relays where each channel outcome becomes a branch of the *broadcast relay channel* (BRC). For instance, the source can design a code with three messages (W_0, W_1, W_2) such that (W_0, W_1) are decoded when the relay uses DF scheme and (W_0, W_2) when CF scheme is used. Inner bounds on the capacity region of the general BRC with two helper relays are derived for the cases where two relays use different variations of coding strategies. First, both relays use partially *Decode-and-Forward* (DF) scheme –DF-DF region–, secondly both relays use *Compress-and-Forward* (CF) scheme –CF-CF region– and finally one uses DF scheme while the other one CF scheme –DF-CF region–. An outer bound on the capacity region of the general BRC is also derived. Capacity results are obtained for specific cases of semi-degraded and degraded Gaussian simultaneous relay channels.

In the second chapter, the composite relay channel is considered where the channel is randomly drawn from a set (non necessarily finite) of conditional distributions with index $\theta \in \Theta$, which represents the vector of channel parameters with a distribution $\mathbb{P}_\theta$ characterizing the probability that each channel is in operation. The specific draw θ is assumed to be unknown at the source, fully known at the destination and only partly known at the relay. In this setting, the transmission rate is fixed regardless of the current channel index and the asymptotic error probability is characterized. A novel selective coding strategy (SCS) is introduced which enables the relay to select –based on its channel measurement– the best coding scheme between CF and DF. Indeed, provided that the channel source-to-relay is good enough for decoding the message, the relay decides on DF and otherwise it may switch to CF. We derive bounds on the asymptotic average error probability of the memoryless relay channel. This result is later extended to the case of unicast composite networks with multiple relays. Generalized *Noisy Network Coding* theorems are shown for the case of unicast general networks where the relays are divided between those using DF scheme and the others using CF scheme. It is also shown that the

DF relays can exploit the help of CF relays using offset coding. An application example to the case of fading Gaussian relay channel is also investigated where it is demonstrated that SCS clearly outperforms well-known DF and CF schemes.

In the third chapter, the asymptotic behavior of error probability is studied for composite multiterminal networks. Here instead of finding the maximum achievable rate subject to a small error probability (EP), we look at the behavior of error probability (not necessarily zero) for a given rate. It can be seen that, as in case of composite binary symmetric averaged channel, the common notion of outage probability is not enough precise to characterize the error probability. Instead, various notions are introduced as a measure of performance among which the asymptotic spectrum of error probability is introduced as a novel performance measure for composite networks. It is shown that the behavior of EP for composite networks is directly related to their ϵ-capacity. Indeed, the notion of asymptotic spectrum of EP yields a more general measure for the performance of composite networks. The asymptotic spectrum of EP is upper bounded by using available achievable rate regions and lower bounded by a new region referred to as *full error region*. It is shown that every code with a rate belonging to this region, which is outer bounded by the cutset bound, yields EP equal to one. In this sense, for the networks satisfying strong converse condition, the asymptotic spectrum of EP coincides with the outage probability.

Publications

Journals

1. **Arash Behboodi**, Pablo Piantanida, "Selective Coding for Composite Unicast Networks", in preparation for IEEE Trans. on Information Theory.
2. **Arash Behboodi**, Pablo Piantanida, "On the Asymptotic Error Probability of Composite Networks", to be submitted to IEEE Trans. on Information Theory, May 2012.
3. **Arash Behboodi**, Pablo Piantanida, "Cooperative Strategies for Simultaneous and Broadcast Relay Channels", IEEE Trans. on Information Theory, March 2011 (under revision).

Conferences

1. **Arash Behboodi**, Pablo Piantanida, "The Asymptotic Error Probability of Composite Networks", Information Theory Workshop 2012 (ITW). September 3-7, 2012. Lausanne - Switzerland (invited paper).
2. **Arash Behboodi**, Pablo Piantanida, "Selective Coding Strategy for Unicast Composite Networks", IEEE International Symposium on Information Theory , ISIT 2012, July 2012 Cambridge, MA, USA.
3. **Arash Behboodi**, Pablo Piantanida, "Selective Coding Strategy for Composite Relay Channels", 5th International Symposium on Communication, Control, and signal processing, ISCCSP 2012, May 2012 Rome, Italy (invited paper).
4. **Arash Behboodi**, Pablo Piantanida, "Cooperative Strategies for Simultaneous Relay Channels", 23rd Symposium on Signal and Image Processing, GRETSI, September 2011 Bordeaux, France.
5. **Arash Behboodi**, Pablo Piantanida, "On the Asymptotic Error Probability of Composite Relay Channel", IEEE International Symposium on Information Theory, ISIT 2011, August 2011 St Petersburg, Russia.
6. **Arash Behboodi**, Pablo Piantanida, "Broadcasting over the Relay Channel with Oblivious Cooperative Strategy", Forty-Eighth Annual Allerton Conference on Communication, Control, and Computing, September 29 - October 1, 2010 Allerton Retreat Center, Monticello, Illinois.

7. **Arash Behboodi**, Pablo Piantanida, "Capacity of a Class of Broadcast Relay Channels", IEEE International Symposium on Information Theory, ISIT 2010, June 13-18, 2010 Austin, Texas.
8. **Arash Behboodi**, Pablo Piantanida, "On the Simultaneous Relay Channel with Collocated Relay and Destination Nodes", WNC3 2010, International Workshop on Wireless Networks: Communication, Cooperation and Competition, May 31, 2010, Avignon, France.
9. **Arash Behboodi**, Pablo Piantanida, "On the Simultaneous Relay Channel with Informed Receivers", IEEE International Symposium on Information Theory, ISIT 2009, June 28-July 3, 2009 Seoul, Korea.

Contents

Abstract iii

Publications v

Contents xi

List of Figures xiii

List of Tables xiv

List of Abbreviations xv

I Réseaux Coopératifs avec Incertitude du Canal 1

1 Introduction 3

1.1 Motivation . . . 10
1.2 Résumé des Contributions . . . 12

2 Stratégies de Coopération pour les Canaux à Relais Simultané 15

2.1 Formulation de problème . . . 16
2.2 Région Atteignable pour la Stratégie DF-DF . . . 19
2.3 Région Atteignable pour la Stratégie CF-DF . . . 23
2.4 Région Atteignable pour la Stratégie CF-CF . . . 26
2.5 Bornes extérieures sur la région de capacité de BRC général . . . 27
2.6 BRC à relais commun dégradé et semi-dégradé . . . 30
2.7 BRC Gaussien dégradé avec relais commun . . . 32
2.8 BRC Gaussien Dégradé avec Coopération Partielle . . . 34

2.9 Résultats Numériques 35
2.9.1 Source est ignorante de la stratégie coopérative adopté par le relais . 35
2.9.1.1 RC Compound 35
2.9.1.2 RC Composite 37

3 Stratégie de Codage Sélectif pour les Réseaux Unicasts Composites 41
3.1 Définition de Problème 42
3.2 Bornes sur la probabilité d'erreur moyenne 44
3.2.1 Décoder-et-Transmettre (DF) et Comprimer-et-Transmettre (CF) . 44
3.2.2 Stratégie de Codage Sélectif (SCS) 45
3.3 Réseaux à Multiple Relais Composite 49
3.3.1 Réseaux à Multiple Relais 51
3.4 Réseaux à Multiples Relais Composite 54
3.4.1 Stratégies de Codage Non-Sélective 55
3.4.2 Stratégie de Codage Sélectif 56
3.5 Canaux à deux relais Gaussiens avec évanouissement 59

4 Spectre Asymptotique de EP pour les réseaux composites 65
4.1 Notation et Arrière-Plan 67
4.2 Définition de Réseau Multi-terminaux Composite (CMN) 67
4.3 Fonction de Fiabilité pour Réseaux Composites 72
4.4 Spectre Asymptotique de Probabilité d'erreur 75
4.5 Résultats Principaux 77

II Cooperative Networks with Channel Uncertainty 87

5 Introduction 89
5.1 Related Works 92
5.1.1 Cooperation and Multiterminal Networks 92
5.1.2 Compound and Composite Channels 95
5.1.3 Cooperative Networks with Channel Uncertainty 97
5.2 Motivation 98
5.3 Summary of Contributions 99

6 Cooperative Strategies for Simultaneous Relay Channels **103**
6.1 Introduction 103
6.2 Main Definitions and Achievable Regions 107
6.2.1 Problem Statement 108
6.2.2 Achievable region based on DF-DF strategy 110
6.2.3 Achievable region based on CF-DF strategy 114
6.2.4 Achievable region based on CF-CF strategy 117
6.3 Outer Bounds and Capacity Results 118
6.3.1 Outer bounds on the capacity region of general BRC 118
6.3.2 Degraded and semi-degraded BRC with common relay 121
6.3.3 Degraded Gaussian BRC with common relay 124
6.3.4 Degraded Gaussian BRC with partial cooperation 125
6.4 Gaussian Simultaneous and Broadcast Relay Channels 126
6.4.1 DF-DF region for Gaussian BRC 127
6.4.2 DF-CF region for Gaussian BRC 129
6.4.3 CF-CF region for Gaussian BRC 132
6.4.4 Source is oblivious to the cooperative strategy adopted by the relay 133
6.4.4.1 Compound RC 133
6.4.4.2 Composite RC 135
6.4.5 Source is oblivious to the presence of relay 137

7 Selective Coding Strategy for Composite Unicast Networks **139**
7.1 Introduction 139
7.2 Problem Definition 142
7.3 Bounds on the Average Error Probability 143
7.3.1 Decode-and-Forward (DF) and Compress-and-Forward (CF) Schemes143
7.3.2 Selective Coding Strategy (SCS) 144
7.4 Composite Multiple Relay Networks 148
7.4.1 The Multiple Relay Networks 149
7.4.2 The Composite Multiple Relay Networks 153
7.4.3 Non-Selective Coding Strategy 154
7.4.4 Selective Coding Strategy 154
7.5 Application Example and Discussions 158
7.5.1 Gaussian Fading Multiple Relay Channels 158

7.5.2 Gaussian Fading Single Relay Channels 160
7.5.3 Gaussian Fading Two Relay Channels 163

8 The Asymptotic Spectrum of the EP for Composite Networks 169
8.1 Introduction 169
8.2 Main Definitions and Background 172
8.2.1 Notation and Background 172
8.2.2 Definition of Composite Multiterminal Network (CMN) 173
8.2.3 Reliability Function for Composite Networks 178
8.2.4 Asymptotic Spectrum of Error Probability 181
8.3 Main Results 183

9 Conclusion and Future Work 193
9.1 Summary and Conclusion 193
9.2 Future Work 195

A Appendix chapter 2 199
A.1 Proof of Theorem 24 199
A.2 Proof of Theorem 25 203
A.3 Proof of Theorem 26 206
A.4 Proof of Theorem 27 208
A.5 Proof of Theorem 30 220
A.6 Proof of Theorem 31 223
A.7 Proof of Theorem 33 224
A.8 Proof of Theorem 34 227

B Appendix chapter 3 231
B.1 Proof of Theorem 35 and Corollary 36 231
B.1.1 Proof of Theorem 35 231
B.1.2 Proof of Corollary 36 236
B.2 Proof of Theorem 37 239
B.3 Proof of Theorem 38 244
B.4 Proof of the theorem 39 248
B.5 Proof of the Corollaries 17 and 18 252

C Appendix chapter 4 **257**
C.1 Proof of Theorem 257
C.1.1 First Step 261
C.1.2 Second Step 262
C.1.3 Final Step 264

Bibliography **267**

List of Figures

1.1 Le canal à relais sans mémoire . . . 4
1.2 Canal à relais AWGN d'évanouissement lent . . . 10

2.1 Canaux de diffusion à relais et à relais simultané . . . 16
2.2 Canal de diffusion à relais (BRC) . . . 16
2.3 Diagramme de variables aléatoires auxiliaires . . . 20
2.4 Ré-configuration des message . . . 21
2.5 BRC Gaussien Dégradé (DG-BRC) . . . 33
2.6 Débit commun de BRC Gaussien avec stratégies DF-CF . . . 37
2.7 Débit espéré pour le canal à relais Gaussien composite . . . 39

3.1 Réseau à deux relais . . . 46
3.2 Réseau Composite à Multiple Relais . . . 55
3.3 Canal à deux relais Gaussien avec évanouissement . . . 60
3.4 La probabilités d'erreur asymptotique $\bar{\epsilon}(r)$ vs. débit de codage r. . . . 64

4.1 Réseau Multi-terminaux Composite (CMN) . . . 68
4.2 Réseau Multi-terminal Composite (CMN) avec $\theta = (\theta_r, \theta_d)$ . . . 71
4.3 Réseaux Multiterminaux satisfaisant la converse forte . . . 81

5.1 The memoryless relay channel . . . 91
5.2 A slow-fading AWGN relay channel . . . 98

6.1 Simultaneous and broadcast relay channels . . . 107
6.2 Broadcast relay channel (BRC) . . . 108
6.3 Diagram of auxiliary random variables . . . 112
6.4 The message reconfiguration . . . 112
6.5 Degraded Gaussian BRC . . . 124

6.6 Gaussian BRC 127
6.7 Common-rate of the Gaussian BRC with DF-CF strategies 134
6.8 Expected rate of the composite Gaussian relay channel 136
6.9 Inner bound on the capacity of the Gaussian BRC. 138

7.1 Two-relay network. 145
7.2 Composite Multiple Relay Network 153
7.3 Gaussian Multiple Relay Channels 158
7.4 Fading Gaussian relay channel 161
7.5 The asymptotic error probability $\bar{\epsilon}(r)$ vs. the coding rate r. for single relay channel162
7.6 Fading Gaussian two relay channel 163
7.7 Asymptotic error probability $\bar{\epsilon}(r)$ vs. coding rate r. for two relay channels . 167

8.1 Composite Multiterminal Network (CMN) 174
8.2 Composite Multiterminal Network (CMN) with $\theta = (\theta_r, \theta_d)$ 177
8.3 Multiterminal Network with Strong Capacity 187

9.1 Broadcast relay channel with common relay 196

List of Tables

2.1 Stratégie DF avec $b = \{1, 2\}$ 21

6.1 DF strategy with $b = \{1, 2\}$ 113

B.1 Coding for the two-relay channel 232

B.2 Coding for CMNNC 245

List of Abbreviations

ASEP	Asymptotic Spectrum of Error Probability
AWGN	Additive White Gaussian Noise
BRC	Broadcast Relay Channels
BRC-CR	Broadcast Relay Channels with Common Relay
CF	Compress and Forward
CMN	Composite Multiterminal Network
CMNNC	Cooperative Mixed Noisy Network Coding
DF	Decode and Forward
EP	Error Probability
MNNC	Mixed Noisy Network Coding
NNC	Noisy Network Coding
PD	Probability Distribution
RC	Relay Channel
SRC	Simultaneous Relay Channel
SCS	Selective Coding Strategy

Première partie

Réseaux Coopératifs avec Incertitude du Canal

Chapitre 1

Introduction

Les réseaux sans fil occupent une place incontestable dans l'industrie de télécommunication et il est important d'analyser leurs facettes différentes. Ils possèdent quelques spécificités propre à eux. Tout d'abord, ils sont composé de beaucoup de récepteurs et de émetteurs et dans ce sens ils sont un cas particulier de réseaux multiterminaux. Ils sont composé de beaucoup de noeuds qui peuvent être récepteur, émetteur ou tous les deux en même temps et il y a l'ensemble des messages destinés pour quelques noeuds venant d'autres noeuds. Deuxièmement, les noeuds peuvent s'aider, par ex. en envoyant le message aux autres utilisateurs à travers du réseau. Cela signifie que chaque noeud peut choisir le code transmis comme fonction de son propre message et l'observation précédente du canal. Cela ouvre la possibilité de coopération dans un réseau. Finalement, le canal dans les réseaux sans fil est soumis aux changements en raison de l'évanouissement [1] et la mobilité d'utilisateurs, qui nécessite pour considérer l'incertitude inhérente dans la structure de ces réseaux. C'est autour de ces trois axes, c'est-à-dire réseaux multiterminaux, coopération et incertitude que cette thèse est organisée. Une quantité considérable de recherche a été consacré à la théorie d'information de réseau, les réseaux coopératifs et la communication avec l'incertitude de canal. L'essence de coopération est l'opération de relayer. Comme on peut noter dans la figure 1.1, le canal à relais se compose de l'entrée de canal $X \in \mathscr{X}$ et l'entrée de relais $X_1 \in \mathscr{X}_1$, la sortie de canal $Y_1 \in \mathscr{Y}_1$ et la sortie de relais $Z_1 \in \mathscr{Z}_1$. Le canal est caractérisé par $\mathbb{W}(y_1, z_1|x, x_1)$ et il est supposé d'être sans mémoire :

$$\mathbb{W}(\underline{y}_1, \underline{z}_1|\underline{x}, \underline{x}_1) = \prod_{i=1}^{n} \mathbb{W}(y_{1i}, z_{1i}|x_i, x_{1i})$$

1. Fading

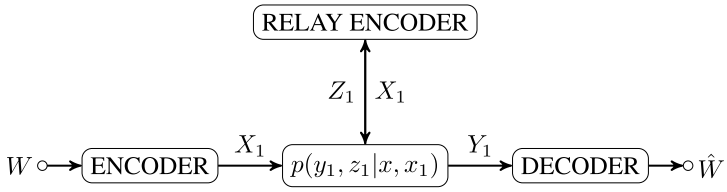


FIGURE 1.1 – Le canal à relais sans mémoire

pour $\underline{x} = (x_1, x_2..., x_n)$, où x_i signifie l'entrée de canal au temps i. L'entrée de relais X_1 au temps i est une fonction des sorties précédentes de relais Z_1, à savoir $X_{1i} = f_i(Z_1^{i-1})$. La difficulté centrale dans ce problème réside dans la découverte de la bonne fonction de relais.

La contribution originale pour ce canal est due de Cover et El Gamal [1]. Ils ont développé les stratégies coopératives principales pour les canaux de relais, à savoir Décoder-et-Transmettre [2] (DF) et Comprimer-et-Transmettre [3] (CF). Dans le codage de DF, comme présenté par Cover et El Gamal, les messages de source sont distribués dans les boîtes [4] indexées. Le relais décode le message de source et transmet ensuite son index de boîte. Le débit atteignable [5] pour le schéma DF est donné par

$$R_{\mathrm{DF}} = \max_{p(x,x_1)} \min \left\{ I(X; Z_1|X_1), I(XX_1; Y_1) \right\}.$$

Le débit précédent est effectivement la combinaison de deux conditions. Le relais doit décoder le message dans ce schéma et la première condition correspond à la condition de décodage réussi au relais à savoir $R \leq I(X; Z_1|X_1)$. La condition suivante est la condition du décodage réussi à la destination $R \leq I(XX_1; Y_1)$. C'est intéressant de voir qu'intuitivement la destination observe un canal d'accès multiple avec deux entrées X, X_1 et ce débit corresponde à ce canal. D'autre part, quand CF est utilisé, le relais trouve une version comprimée de son observation de sa sortie, à savoir $\hat{Z}_1$ et en utilisant la technique de binning, la version comprimée est alors transmise. Le débit atteignable pou CF est comme suit :

$$R_{\mathrm{CF}} = \max_{p(x)p(x_1)p(\hat{z}_1|z_1,x_1)} I(X; Y_1 \hat{Z}_1 | X_1)$$

2. Decode-and-Forward
3. Compress-and-Forward
4. Bins
5. Achievable rate

à condition de

$$I(X_1;Y_1) \geq I(Z_1;\hat{Z}_1|X_1,Y_1).$$

En fait, les schéma DF et CF sont les stratégies coopératives fondamentales développées pour le canal à relais. Une autre région de débit atteignable a été obtenue par Cover-El Gamal en combinant des schéma de DF et CF, où le relais utilise tant DF que CF. Dans un cas spécial de ce résultat, le relais utilise DF pour décoder et envoyer qu'une partie du message de source et le reste du message est directement transmis à la destination. Cela est appelé le décodage partiel DF schéma de DF. En fait, les régions de DF et CF peuvent être obtenus à travers des méthodes différentes. Par exemple la région de DF peut être obtenue, en utilisant des méthodes développées par Willems et Carleial dans [2, 3], où au lieu d'utiliser technique de binning, la source et le relais utilise les livres de code [6] avec la même grandeur, ce qui est appelé l'encodage régulier [7]. Willems a développé le décodage en arrière [8] et Carleial a utilisé le décodage avec la fenêtre-glissante [9] pour décoder le message à la destination. El Gamal-Mohseni-Zahedi dans [4] ont développé un schéma alternatif pour CF où le débit atteignable se révèle être équivalent au débit de CF de Cover-El Gamal.

Bien qu'en général les bornes précédentes ne soient pas serrées, il a été montré que le schéma DF accomplit la capacité des canaux à relais *physiquement dégradé* et *contrairement dégradé* [10] . Le canal à relais dégradé est défini avec la chaîne de Markov suivante $X \ominus (X_1, Z_1) \ominus Y_1$. Y_1 étant dégradé par rapport Z_1 implique intuitivement que Z_1 est en général mieux que Y_1. La notion apparaît aussi dans autres canaux, .e.g. canaux de diffusion. Particulièrement, quand il y a un feedback sans bruit de la destination au relais, le canal à relais peut être considéré comme physiquement dégradé et la capacité est accomplie en utilisant le schéma de DF. D'autre part, le schéma DF partiel résulte la capacité de canaux à relais semi-déterministes [11], comme il a été montré par Aref-El Gamal [5].

Un autre élément important de réseaux est le canal de diffusion [12](BC), où un ensemble de messages communs et privés est destiné à plusieurs destinations. Particulièrement BC sans mémoire à deux utilisateurs, caractérisé par $\mathbb{W}(y_1, y_2|x)$, a été profondément étudié.

6. livre de codes
7. Regular Encoding
8. Backward decoding
9. Sliding Window decoding
10. *physically degraded* and *reversely degraded* relay channels
11. semideterministic
12. Broadcast Channel

La région de capacité de BC dégradé a été trouvée par Bergmans, Gallager et Ahlswede et Korner [6–9]. Korner et Marton ont établi la capacité du BC avec l'ensemble de message dégradés [13] [10]. Ils ont présenté les notions de BC moins-bruyant et plus-capable [14] [11] et ils ont démontré la capacité de BC moins-bruyant. El Gamal a prouvé la capacité de BCs plus capable dans [4]. La borne intérieur le plus connu pour BC général est attribué à Marton [12]. C'est fondé sur l'idée de *binning* où une démonstration alternative a été aussi annoncée par El Gamal et Van der Meulen dans [13]. La région suivante est appelé la région de Marton.

$$
\begin{aligned}
\mathscr{R}_{\mathrm{BC}} = co\Big\{ (R_1, R_2) : R_1, R_2 \geq 0, \\
R_1 &\leq I(U_0 U_1; Y_1) \\
R_2 &\leq I(U_0 U_2; Y_2) \\
R_1 + R_2 &\leq \min\{I(U_0; Y_1), I(U_0; Y_1)\} \\
&\quad + I(U_1; Y_1|U_0) + I(U_2; Y_2|U_0) - I(U_2; U_1|U_0) \text{ pour tous } \mathbb{P}_{U_0 U_1 U_2 X} \in \mathcal{P} \Big\},
\end{aligned}
$$

où $\mathcal{P}$ est l'ensemble de toutes les distributions de probabilité (PDs) $\mathbb{P}_{U_0 U_1 U_2 X}$.

Une recherche vaste a été faite pendant des années pour étudier la région de capacité de réseaux plus général en combinant des canaux à relais simples, des canaux de diffusion et des canaux d'accès multiple. Par exemple, les canaux de diffusion à relais et les canaux d'accès multiple à relais , avec les réseaux généraux [14] ont été étudiés. Les régions de débit atteignables ont été dérivées en combinant des techniques de codage comme : schémas de (partiel) DF et CF, codage de Marton, superposition, bloc-Markov, etc. Pourtant comme la région de capacité n'est pas connue pour la plupart des réseaux fondamentaux comme le canal à relais et le canal de diffusion, les débit atteignables obtenus ne sont pas serrés en général.

La recherche sur les réseaux généraux a passionné les chercheurs à partir du début de la théorie d'information. Elias-Feinstein-Shannon aux alentours de 1956 ont exposé une borne supérieure sur la capacité de réseaux multi-terminaux [15]. Théorème (Elias-Feinstein-Shannon '56) : le flot maximum possible de gauche à droit à travers un réseau est égal à la valeur minimale parmi tous les coupes simples. La preuve de ce théorème a été aussi donnée par Ford-Fulkerson dans [16] et Dantzig-Fulkerson [17]. De plus les auteurs

13. Degraded Message Set
14. Less noisy and more capable channel

ont clairement déclaré que ce n'est d'aucune façon "évident" si cette région pour les réseaux généraux peut être accomplie. Supposons maintenant un réseau avec N utilisateurs de paire (X_i, Y_i), pour $i \in \mathcal{N} = \{1, 2..., N\}$ et le canal $\mathbb{W}(y_1, y_2..., y_N|x_1, x_2..., x_N)$. Alors la borne correspondant à ce coupe est comme suit $(R(S) = \sum_{k\in S} R_k)$:

$$\mathscr{R}_{\text{CB}} = co \bigcup_{P\in\mathcal{P}} \Big\{(R(S) \geq 0) : R(S) < I(X(S); Y(S)|X(S^c)) \text{ pour tous } S \subseteq \mathcal{N}\Big\}.$$

En fait, la région mentionné auparavant n'est pas en général atteignable. En outre, il est difficile de généraliser le codage pour les canaux de diffusion et les canaux à relais aux réseaux arbitraires. Pourtant, dans le travail récent [18] par Lim-Kim-El Gamal-Chung, le codage de réseau bruyant [15](NNC) a été présentée pour les réseaux généraux. Cela est fondé sur le schéma CF généralisé. Ce schéma est capable d'atteindre la borne de flot-max/coupe-min [16] à une distance constante prés.

Theorem 1 *(Lim-Kim-El Gamal-Chung, 2011) Une borne intérieure sur la région de capacité des réseaux sans mémoire avec N utilisateurs et ensemble de destinations D est donnée par*

$$\begin{aligned}\mathscr{R}_{NNC} =& co \bigcup_{P\in\mathscr{Q}} \Big\{(R(S) \geq 0) : \\ & R(S) < \min_{d\in S^c\cap\mathcal{D}} I(X(S); \hat{Y}(S^c), Y_d|X(S^c), Q) - I(Y(S); \hat{Y}(S)|X^N, \hat{Y}(S^c), Y_d, Q)\Big\}\end{aligned}$$

pour tous les coupes $S \subset \mathcal{N}$ avec $S^c \cap \mathcal{D} = \emptyset$.

NNC est fondé sur la transmission du même message dans tous les blocs - l'encodage répétitif- et le décodage non-unique d'index de compression où la compression ne profite pas de technique binning. Le schéma NNC accomplit la capacité de quelques réseaux, par exemple les réseaux déterministes linéaires finis de terrain [17] [19].

Le problème de la communication avec l'incertitude de canal a été étudiée via les modèles différents pourtant l'hypothèse principale est que le canal est inconnu aux terminaux. Soit le canal change arbitrairement pendant chaque rond de transmission, soit il reste fixe pendant le cours de transmission. Dans le premier cas nous sommes en face des canaux avec les états [18] pendant que dans le deuxième cas, nous arrivons au problème du canal

15. Noisy Network Coding
16. Max Flow Min Cut
17. Finite field deterministic networks
18. Channels with state

compound. Ces modèles correspondent grossièrement aux cas des canaux de communications sans fil avec l'évanouissement vite et lent. Dans le cadre d'évanouissement vite, la longueur de code est considérablement plus grande que le temps de cohérence du canal [19] et on peut considérer la capacité ergodique. Pour le cas d'évanouissement lent, la longueur de code est dans l'ordre du temps de cohérence. Bien que il y a des stratégies différentes développées pour ces scénarios, on se concentre sur la capacité compound. Le canal compound se compose à un ensemble de canaux indexé par θ :

$$\mathcal{W}_\Theta = \{\mathbb{W}_\theta(y|x) : \mathcal{X} \longmapsto \mathcal{Y}\}_{\theta \in \Theta} .$$

Il est important de noter qu'il n'y a aucune distribution supposée sur Θ. De plus pour avoir un débit atteignable pour le canal compound, le code devrait avoir la probabilité d'erreur petite pour chaque θ. La capacité du canal compound est donnée par [20–22]

$$C_{\mathrm{CC}} = \max_{p(x)} \inf_{\theta \in \Theta} I(X; Y_\theta),$$

où Y_θ est la sortie du canal avec la distribution $\mathbb{W}_\theta(y|x)$. Pourtant en cas d'un canal AWGN d'évanouissement lent, on ne peut pas garantir de petite probabilité d'erreur pour tous les canaux possibles parce que finalement on ne peut garantir que le débit zéro. Supposons maintenant qu'au lieu d'un ensemble simple des messages, on permet à l'encodeur de transmettre plusieurs ensemble des messages - le codage du canal de débit variable [23] - et ensuite la destination, selon l'index θ, décode autant que possible des messages. La connexion existante entre la diffusion et les canaux compound a été d'abord remarquée par Cover dans [24, la Section IX], où il a suggéré que le problème de canaux compound peut être étudié de ce point de vue de la diffusion. Cette idée a été complètement développée par Shamai dans [25] et appelée l'approche de la diffusion.

Considérons le canal AWGN d'évanouissement lent défini comme

$$Y = hX + \mathcal{N},$$

où $\mathcal{N}$ est le bruit AWGN et h est le coefficient d'évanouissement. Le canal est un canal d'évanouissement lent, qui signifie que h est choisi aléatoirement au préalable et reste constant pendant la communication. Ici l'incertitude vient du coefficient d'évanouissement h et pour chaque triage de h, il y a un canal qui peut être en opération. L'ensemble de tous les canaux possibles peut être considéré comme étant indexé par h, c'est-à-dire $\theta = h$.

19. Coherence Time

Maintenant l'émetteur considère le canal d'évanouissement comme un canal de diffusion Gaussien dégradé avec un continuum de récepteurs chacun avec un différent rapport signal sur bruit spécifié par uSNR où u est l'index continu. Shamai a construit une codage à plusieurs-couches, une couche pour chaque triage de h, tel que pour chaque tirage de h, toutes les couches avec $u = |h'|^2 \leq v = |h|^2$ peuvent être décodées et le reste de couches apparaissent comme l'interférence. L'allocation de pouvoir pour v est $\mathrm{SNR}(v)dv \geq 0$. Le débit pour ce canal est une fonction de v et suit comme

$$R(v) = \int_0^v \frac{-udy(u)}{1 + uy(u)},$$

où $y(u) = \int_v^\infty \mathrm{SNR}(v)dv$. L'idée principale derrière la stratégie de la diffusion est d'envoyer de différents messages pour que la destination puisse choisir combien entre eux peuvent être décodés selon le canal en opération. Dans la stratégie de la diffusion, le code transmis garantie des débit variables pour chacun des canaux possibles dans l'ensemble.

Il y a d'autres approches afin de s'occuper de l'incertitude dans les réseaux. Dans les cadres compound, il n'y a aucune distribution de probabilité présentée sur θ. Pour tenir compte de la distribution de θ, la notion de capacité de panne [20] a été proposée dans [26] pour les canaux d'évanouissement. Pour la probabilité de panne [21] désirées p, la capacité de panne est définie comme le débit maximum qui peut être transmis avec la probabilité $1 - p$. Au contraire, la capacité ergodique est le débit maximum d'information pour lequel la probabilité d'erreur diminue exponentiellement avec la longueur de code. À la différence de la stratégie de la diffusion, le code transmis envoie un débit fixe pour tous les canaux possibles dans l'ensemble. Effros-Goldsmith-Liang ont présenté le canal composite [27]. "Un canal composite se compose d'une collection de différents canaux avec une distribution caractérisant la probabilité que chaque canal est en opération." Donc le canal composite est défini comme l'ensemble des canaux $\mathcal{W}_\Theta$ comme auparavant, mais avec PD associé $\mathbb{P}_\theta$ sur l'index θ de canal. Les modèles composites à la différence des modèles compound, tiennent compte de l'incertitude de canal en présentant un PD $\mathbb{P}_\theta$ sur l'ensemble. Les auteurs dans [27] élargissent la définition de capacité pour permettre de panne en partie. Effectivement, la notion de *la capacité de panne* est définie comme le plus haut débit asymptotiquement atteignable avec une probabilité données de panne connue à décodeur.

20. Outage Capacity
21. Outage probability

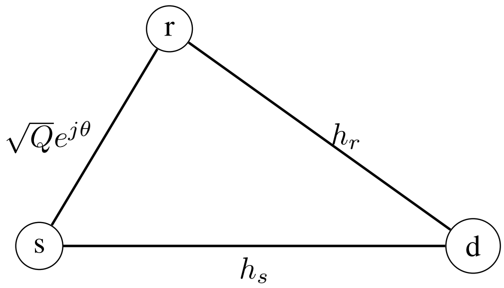


FIGURE 1.2 – Canal à relais AWGN d'évanouissement lent

L'incertitude dans les réseaux généraux peut être tant en raison de la mobilité d'utilisateur que l'évanouissement. Supposons maintenant que le relais peut être également présent ou absent et la source est inconscient de ce fait. De plus la topologie de réseau et du canal lui-même reste fixe pendant le cours de communication. Alors la source devrait être en mesure de concevoir un code pour préserver la performance malgré l'absence du relais. Katz-Shamai ont étudié ce problème à [28] avec un relais proche occasionnel, comme montré dans Fig. 1.2. Par le relais proche, on veut indiquer que le schéma de DF exécute mieux que CF. Ils ont utilisé la notion de débit espéré [22] pour mesurer la performance. Il a été montré que la superposition et le décodage en arrière de codage permettent à la destination de décoder le message sans perte de performance, même si le relais n'est pas présent. Autrement dit, si le relais n'est pas présent le débit espéré reste le même.

Les auteurs ont présenté la notion de *la coopération ignorant* [23] pour faire allusion aux protocoles coopératifs qui améliorent la performance quand le relais est présent et ne la dégradent pas quand le relais est absent, même si la source est non-informée de la topologie réelle.

1.1 Motivation

Il est bien reconnu comment les réseaux sans fil sont soumis aux changements statistiques, surtout en raison de la mobilité d'utilisateur et l'évanouissement. Dans quelques scénarios, la longueur de code est de façon significative plus petite que l'intervalle de temps de cohérence et donc le canal reste fixe pendant la communication. Considérons par

22. Expected rate
23. Oblivious cooperation

exemple un canal à relais simple où les canaux source-relais et relais-destination peuvent changer aléatoirement. Si le canal source-relais est assez bonne alors le relais serait en mesure de décoder le message de source et l'envoyer à la destination. Il semble mieux d'exécuter le schéma DF dans ce cas-là. Pourtant, si le canal source-relais est trié aléatoirement, la qualité de canal peut être de façon significative détériorée pour quelques cas et il n'est pas garanti que le relais peut décoder le message avec succès. Dans ces cas, une erreur est déclarée et le décodage ne peut pas être réalisé avec succès. De plus la source ne peut rien faire parce qu'il n'est pas conscient de la réalisation de canal. Dans le cas des canaux d'AWGN d'évanouissement lent, la qualité de canal source-relais, qui est fixe pendant la transmission, peut être pauvre pendant le cours entier de la communication et cela produit le décodage erroné à la destination. Une situation semblable peut être considérée si la source emploie CF. Quand le canal relais-destination est pauvre, l'utilisation de CF n'est pas adéquate et ainsi le décodage peut être erroné de nouveau. En effet, même si le canal de source-relais est assez bon pour permettre au relais de décoder le massage, il ne peut pas complètement exploiter le schéma DF parce que le code de source est conçu pour CF et alors indépendant du code de relais. Remarquons que dans ces exemples le relais peut avoir accès, au moins partiellement, aux informations d'état de canal (CSI) [24] parce qu'il a un récepteur et donc il peut avoir une estimation du canal. Mais parce que la source doit fixer le codage *a priori*, une stratégie coopérative est imposée au relais et donc ce n'est pas capable de profiter de CSI disponible.

Les susdits problèmes mentionnés sont centraux dans les réseaux multiterminaux avec l'incertitude de canal. Le problème principal est que le codage nécessaire (par ex. la stratégie coopérative) dépend de la qualité de canal source-relais et relais-destination. Donc il est désirable d'explorer comment le codage opportuniste et-ou adaptable pour la coopération est possible, même si la source est ignorante des états de canal. Autrement dit, comment les utilisateurs qui sont partiellement ou complètement conscients de CSI peuvent exploiter leurs informations disponibles pour fournir une meilleure performance de coopération. Nous ferons allusion à ces stratégies comme les stratégies ignorants, qui signifie que la source est inconscient de la stratégie du code déployée dans les autres terminus. Un exemple d'une stratégie ignorant a été donné auparavant quand la source ne sait pas si le relais est présent ou pas, mais sait que si le relais est présent alors il est proche d'elle. Il a été discuté que si le codage de superposition est permis à la source, donc la destination peut décoder le sujet de message à la contrainte de la capacité de canal direct

24. Channel State Information

à utilisateur simple, même si le relais n'est pas présent. Donc on peut le dire que le codage de superposition est un code ignorant en ce qui concerne la présence du relais.

Dans cette thèse, nous enquêtons sur les stratégies coopératives avec l'incertitude de canal. En particulier, nous nous intéressons à deux cas. D'abord, le cas de canaux à relais simultanés qui se compose d'un ensemble de canaux à relais et deuxièmement, le cas de modèles composites où le canal en opération est choisi parmi l'ensemble des canaux indexés avec θ suivant d'un PD $\mathbb{P}_\theta$. Dans ce cadre, la source (ou les sources) est ignorante du canal en opération indexée par θ. Les autres terminus sont partiellement ou complètement conscients du canal en opération. Comme nous avons mentionné, l'incertitude de canal peut être étudiée basée sur les modèles compound et via l'approche de la diffusion, ou sur la notion de la capacité de panne. Dans la même direction, nous verrons comment ces approches peuvent nous aider à comprendre mieux des bornes fondamentales et des schémas de codage originaux pour les réseaux coopératifs avec l'incertitude de canal.

1.2 Résumé des Contributions

La contribution de cette thèse est organisée dans trois chapitres :

- Les stratégies de coopération pour les canaux à relais simultané et les canaux de diffusion à relais,
- La stratégie de codage sélective pour les réseaux unicasts composites,
- Sur le Spectre Asymptotique [25] de la probabilité d'erreur de réseaux composites.

Dans le premier chapitre, les stratégies coopératives sont développées pour *le canal à relais simultané* (SRC), qui se compose d'un ensemble de canaux à relais simples parmi lesquels le canal en opération est choisi. L'approche de la diffusion est adoptée pour ce canal où la source veut transmettre des informations communes et privées à chacun de canaux possibles. Cela ouvre la possibilité d'utiliser l'approche de la diffusion afin d'envoyer des messages à chaque canal. Par exemple, supposons que le relais utilise DF ou CF mais il est toujours présent. Maintenant bien que la source puisse être ignorante de la stratégie du code au relais, il sait que cette stratégie est soit DF soit CF qui produit deux possibilités. Alors la source peut concevoir un code avec trois messages (W_0, W_1, W_2) tel que (W_0, W_1) est décodé quand le relais utilise DF et (W_0, W_2) quand CF est permis. Donc, ce problème est reconnu comme étant équivalent au problème d'envoyer des informations communes et privées à plusieurs destinations en présence de relais où chaque canal possible devient

25. Asymptotic Spectrum

une branche du *canal de diffusion à relais* (BRC). Les schémas coopératifs et la région de capacité pour un ensemble de deux canaux à relais sont enquêtés. Les stratégies de codage proposées doivent être capables de transmettre des informations simultanément à toutes les destinations dans un tel ensemble. Les bornes intérieures sur la région de capacité de général BRC sont dérivées pour trois cas d'intérêt particulier :

– Les canaux source-relais des deux destinations sont supposés plus forts que les autres et alors la coopération est fondé sur la stratégie DF pour les deux utilisateurs (la région DF-DF),
– Les canaux relais-destination des deux destinations sont supposés plus forts que les autres et alors la coopération est fondé sur la stratégie CF pour les deux utilisateurs (la région CF-CF),
– Le canal source-relais d'une destination est supposé plus fort que les autres pendant que pour l'autre est le canal relais-destination et alors la coopération est fondé sur la stratégie DF pour une destination et CF pour une autre (la région DF-CF).

Les techniques utilisé pour obtenir les bornes intérieures comptent sur la recombinaison de bits de messages et les stratégies de codage efficaces différentes pour le relais et les canaux de diffusion. Ces résultats peuvent être vus comme une généralisation et alors l'unification de travail précédent dans ce thème. Une borne extérieure sur la région de capacité de général BRC est aussi dérivé. Les résultats de capacité sont obtenus pour les cas spécifiques des canaux à relais simultanés semi-dégradé et Gaussien dégradé. Les débits sont calculées pour les modèles AWGN.

Dans le deuxième chapitre, le canal à relais composite est considéré où le canal est tiré aléatoirement d'un ensemble de distributions conditionnelles avec l'index $\theta \in \Theta$, qui représente le vecteur de paramètres de canal avec PD $\mathbb{P}_\theta$ caractérisant la probabilité que chaque canal est en opération. Le tirage spécifique θ est supposée inconnue à la source, complètement connue à la destination et seulement partiellement connue au relais. Dans ce cadre, le débit de transmission est fixe sans tenir compte de l'index de canal actuel. Alors l'encodeur ne peut pas nécessairement garantir de probabilité d'erreur arbitrairement petite pour tous les canaux. la probabilité d'erreur asymptotique sont utilisées comme métriques pour caractériser la performance. Dans ce cadre, la stratégie du code est communément choisie sans tenir compte de la mesure de canal au relais. Nous présentons une nouvel codage qui permet au relais de choisir - basé sur sa mesure de canal - le meilleur schéma de codage entre les schémas CF et DF. Effectivement, à condition que le canal source-relais est assez bonne pour décoder le message, le relais se décide pour DF et autrement il peut changer à CF. La stratégie de codage sélective proposée (SCS) est fondée sur le codage de superposition, DF et CF, le décodage en arrière et collectif à la destination. Nous dérivons des bornes sur la probabilité d'erreur asymptotique du canal à relais sans mémoire. Ce résultat est plus tard prolongé au cas de réseaux composites unicast avec les multiple relais. Comme conséquence de cela, nous généralisons le théorème NNC pour le cas de réseaux unicast où les relais sont divisés entre ceux qui utilisent le schéma DF et ceux qui utilisent CF. Il est aussi montré que les relais en utilisant le schéma de DF et codage en offset peut exploiter l'aide des relais qui utilisent CF. Un exemple d'application au cas du canal d'évanouissement à relais Gaussien est aussi enquêté où il est démontré que SCS remporte clairement les schémas célèbre DF et CF.

Le troisième chapitre est consacré à quelques considérations théoriques des réseaux multiterminaux composites. Comme nous avons déjà mentionné auparavant, une probabilité d'erreur arbitrairement petites ne peuvent pas être garanties pour tous les canaux dans l'ensemble. Ici au lieu de trouver le débit atteignable maximum subit aux petites probabilités d'erreur (EP), nous regardons la conduite de probabilités d'erreur (pas nécessairement zéro) pour un débit donné. La notion commune comme la mesure de performance de réseau composite est la notion de probabilité de panne. Mais il est vu en cas du canal symétrique binaire en moyenne composite [26], pour lequel ϵ-capacité est connue, la probabilité de panne n'est pas assez précise comme la mesure de performance. Au lieu de cela les notions différentes de performance sont discutées parmi lequel le spectre asymptotique de probabilités d'erreur est présenté comme une mesure de performance originale pour les réseaux composites. On montre que la conduite d'EP est directement liée à leur ϵ-capacité. Par exemple, la notion de spectre asymptotique d'EP est la mesure plus générale pour la performance de ces réseaux. Le spectre asymptotique d'EP peut être borné en utilisant des régions de débit atteignables disponibles et une nouvelle région appelé *la région d'erreur complète*. Chaque code avec un débit appartenant à cette région produit EP égal à un. Dans ce sens, pour les réseaux satisfaisant la condition de converse forte [27], le spectre asymptotique d'EP coïncide avec la probabilité de panne. À ce but il est montré que la borne flot-max coupe-min fournit une borne supérieure à la région d'erreur complète.

26. Composite binary symmetric averaged channel
27. Strong Converse Condition

Chapitre 2

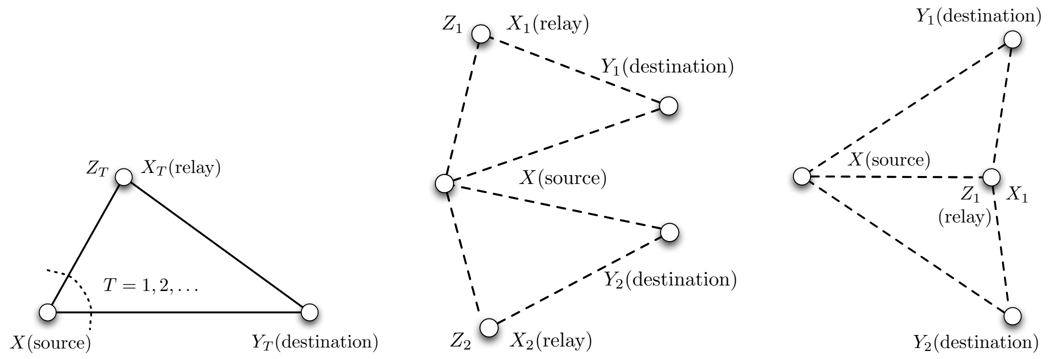
Stratégies de Coopération pour les Canaux à Relais Simultané

Le canal à relais simultané (SRC) est défini par un ensemble de canaux à relais, où la source veut communiquer des informations communes et privées à chacune des destinations dans l'ensemble. Pour envoyer des informations communes sans tenir compte du canal en opération, la source doit simultanément considérer tous les canaux comme décrit dans Fig. 2.1(a). Le scénario décrit offre une perspective d'applications pratiques, quant à l'exemple, la liaison descendante [1] en communication des réseaux cellulaires où la station de base (la source) peut être aidée par les relais, ou dans les réseaux ad hoc où la source peut ne pas prendre conscience de la présence d'un relais proche (par ex. coopération opportuniste).

Le problème du canal à relais simultané est équivalent à ce du canal de diffusion à relais (BRC), avec les chaînes de Markov supplémentaires. La source envoie des informations communes et privées à plusieurs destinations qui sont aidées par leurs propres relais. Donc le problème de SRC peut être étudié en utilisant le problème de BRC.

Dans cette section, nous étudions différentes stratégies de codage et région de capacité pour le cas de BRC général avec deux relais et destinations, aussi montrées dans Fig. 2.1(b), en tant que modèle équivalent pour SRC avec deux canaux à relais sans mémoire. Notons que chaque modèle présenté pour BRC peut être considéré comme un modèle équivalent pour le SRC en ajoutant des chaînes de Markov nécessaires pourtant pour le reste nous n'affirmons pas explicitement les chaînes de Markov. Dans la section suivante, nous formalisons d'abord le problème du canal à relais simultané et ensuite nous présentons

1. Downlink



(a) canal à relais simultané (SRC) (b) BRC avec deux relais (c) BRC avec relais commun

FIGURE 2.1 – Canaux de diffusion à relais et à relais simultané

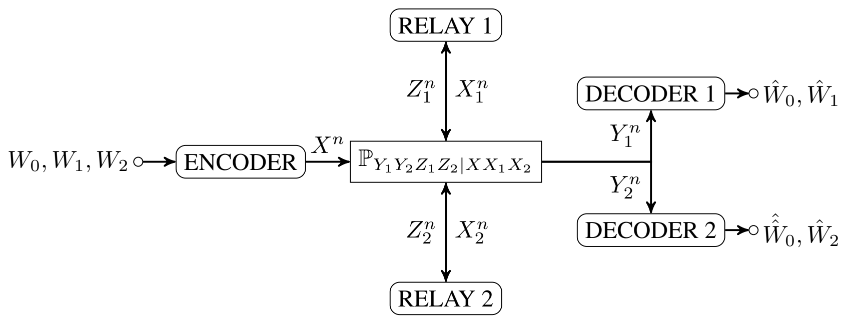


FIGURE 2.2 – Canal de diffusion à relais (BRC)

des régions de débit atteignables pour de différents cas de stratégie DF-DF , CF-CF et DF-CF. Des variables aléatoires sont présentées par les lettres majuscules X, Y. Les lettres gras $\mathbf{X}, \mathbf{Y}$ présentent la suite de n variables aléatoires, c'est-à-dire X^n, Y^n. D'autre part la chaîne de Markov entre trois variables aléatoires A, B et C est présenté en utilisant la notation suivante :

$$A \multimap B \multimap C.$$

2.1 Formulation de problème

Le canal à relais simultané [29] avec les entrées de la source discrète et du relais $x \in \mathscr{X}$, $x_T \in \mathscr{X}_T$, les sorties de relais et canal discrète $y_T \in \mathscr{Y}_T$, $z_T \in \mathscr{Z}_T$, est caractérisé par un ensemble de canaux à relais, chacun d'entre eux défini par une distribution de probabilités conditionnelle (PD)

$$\mathscr{P}_{SRC} = \left\{ P_{Y_T Z_T|X X_T} : \mathscr{X} \times \mathscr{X}_T \longmapsto \mathscr{Y}_T \times \mathscr{Z}_T \right\},$$

où T dénote l'index de canal. Le SRC modèle la situation dans laquelle seulement un canal simple est présent immédiatement et il ne change pas pendant la communication. Pourtant l'émetteur (la source) n'est pas instruit de la réalisation de T qui gouverne la communication. Dans ce cadre, T est supposé pour être connu à la destination et au relais. La transition PD de l'extension n-ième sans mémoire avec les entrées $(\mathbf{x}, \mathbf{x}_T)$ et les sorties $(\mathbf{y}_T, \mathbf{z}_T)$ est donnée par

$$P^n_{Y_T Z_T|X X_T}(\mathbf{y}_T, \mathbf{z}_T|\mathbf{x}, \mathbf{x}_T) = \prod_{i=1}^{n} W_T(y_{T,i}, z_{T,i}|x_i, x_{T,i}).$$

Ici nous nous concentrons sur le cas où $T \in \{1, 2\}$, autrement dit il y a deux canaux à relais dans l'ensemble.

Definition 1 (Code) *Un code pour SRC se compose de*

– *Une fonction d'encodeur* $\{\varphi : \mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}_2 \longmapsto \mathscr{X}^n\}$,

– *Deux fonctions de décodeur* $\{\psi_T : \mathscr{Y}_T^n \longmapsto \mathcal{W}_0 \times \mathcal{W}_T\}$,

– *Un ensemble des fonctions de relais* $\{f_{T,i}\}_{i=1}^n$ *telle que* $\{f_{T,i} : \mathscr{Z}_T^{i-1} \longmapsto \mathscr{X}_T^n\}_{i=1}^n$,

pour $T = \{1, 2\}$ *et pour quelques ensembles finis des entiers* $\mathcal{W}_T = \left\{1, \dots, M_T\right\}_{T=\{0,1,2\}}$. *Les débits de tel code sont* $n^{-1} \log M_T$ *et la probabilité d'erreur maximum correspondantes sont définies comme*

$$T = \{1, 2\} : \quad P^{(n)}_{e,T} = \max_{(w_0, w_T) \in \mathcal{W}_0 \times \mathcal{W}_T} \Pr\left\{\psi(\mathbf{Y}_T) \neq (w_0, w_T)\right\}.$$

Notons que le canal à relais compound a effectivement la même définition que le canal à relais simultané pourtant nous gardons les deux termes pour indiquer la différence dans les codes pour chacun. D'une part un code garantie un débit commun pour tous les canaux, c'est-à-dire tous T et le débit privé pour chaque canal, c'est-à-dire chacun de T, comme le code défini ci-dessus. Nous appelons ce cas comme le canal à relais simultané. D'autre part un autre code garantie seulement un débit commun et envoie le message commun w_0 à tous les canaux, c'est-à-dire tous T. En utilisant le canal à relais compound nous voulons référer à ce cas.

Definition 2 (Capacité et débit atteignable) *Pour chaque* $0 < \epsilon, \gamma < 1$, *un triplet des nombres non-négatifs* (R_0, R_1, R_2) *est atteignable pour SRC si pour tous* n *suffisam-*

ment large, il y a un code de longueur n dont la probabilité d'erreur satisfait :

$$P_{e,T}^{(n)}\left(\varphi,\psi,\{f_{T,i}\}_{i=1}^{n}\right)\leq\epsilon$$

pour chaque $T=\{1,2\}$ *et les débits*

$$\frac{1}{n}\log M_T\geq R_T-\gamma,$$

pour $T=\{0,1,2\}$. *On appelle l'ensemble de tous les débit atteignables* $\mathscr{C}_{BRC}$ *la région de capacité du SRC. Nous insistons qu'aucune distribution préalable sur* T *n'est supposée et ainsi l'encodeur doit construire un code qui produit de petites probabilités d'erreur pour chaque* $T=\{1,2\}$.

Une définition semblable peut être offerte pour SRC à message commun avec un ensemble de message simple $\mathcal{W}_0$, $n^{-1}\log M_0$ *et le débit* R_0. *SRC à message commun est équivalent au canal à relais compound et ainsi le débit atteignable pour le canal à relais compound est défini de la même façon.*

Remark 1 *Remarquons que, puisque le relais et le récepteur sont supposés instruits de la réalisation de* T, *le problème du codage de SRC peut être transformé au problème du canal de diffusion à relais (BRC) [29]. Parce que la source est incertaine du canal réel, il doit compter sur la présence de chaque canal et donc supposer la présence de tous les deux canaux simultanément. Cela résulte à un modèle équivalent de diffusion qui se compose de deux branches à relais, où chacun correspond à un canal à relais avec* $T=\{1,2\}$, *comme illustré dans Fig. 2.1(b) et 2.2. L'encodeur envoie les messages commun et privé* (W_0,W_T) *à la destination* T *au débit* (R_0,R_T). *BRC général est défini par le PD*

$$\mathscr{P}_{BRC}=\left\{P_{Y_1Z_1Y_2Z_2|XX_1X_2}:\mathscr{X}\times\mathscr{X}_1\times\mathscr{X}_2\longmapsto\mathscr{Y}_1\times\mathscr{Z}_1\times\mathscr{Y}_2\times\mathscr{Z}_2\right\},$$

avec entrées de relais et de canal (X,X_1,X_2) *et les sorties de relais et de canal* (Y_1,Z_1,Y_2,Z_2). *Les notions de débit atteignable pour* (R_0,R_1,R_2) *et la capacité restent la même que BCs conventionnel (voir [24], [14] et [30]). Semblable au cas de canaux de diffusion, la région de capacité du BRC dans Fig. 2.1(b) dépend seulement du PDs marginal suivant* $\{P_{Y_1|XX_1X_2Z_1Z_2},P_{Y_2|XX_1X_2Z_1Z_2},P_{Z_1Z_2|XX_1X_2}\}$.

Remark 2 *Nous insistons que la définition de canaux de diffusion à relais n'écarte pas la possibilité de dépendance de la première (également deuxième) destination* Y_1 *sur le deuxième (également le premier) relais* X_2 *et donc c'est plus général que les canaux à* *relais simultanés. Autrement dit, la définition actuelle de BRC correspond à SRC avec les contraintes supplémentaires pour garantir que* (Y_T, Z_T) *conditionné par* (X, X_T) *pour* $T = \{1, 2\}$ *est indépendant d'autres variables aléatoires. En dépit du fait que cette condition n'est pas nécessaire jusqu'aux preuves converses, la région atteignable développée cidessous est plus adaptée au canal à relais simultané. Pourtant les régions de débit atteignables n'ont pas besoin d'hypothèse supplémentaire et sont alors valides pour BRC général.*

2.2 Région Atteignable pour la Stratégie DF-DF

Considérons la situation où les canaux source–relais sont plus forts que les autres canaux. Dans ce cas-là, la stratégie de codage la plus efficace pour les deux relais se révèle d'être Décode-et-Transmettre (DF). La source doit transmettre les informations aux destinations basées sur un code de diffusion combiné avec le schéma DF. La codage derrière cette idée est comme suit. Les informations communes sont aidées par la partie commune des deux relais pendant que les informations privées sont envoyées en utilisant la division de débit dans deux parties, une partie par l'aide du relais correspondant et une autre partie par la transmission directe de la source à la destination correspondante. Le théorème suivant présente la région de débit atteignable générale.

Theorem 2 *(région DF-DF) Une borne intérieur sur la région de capacité* $\mathscr{R}_{DF\text{-}DF} \subseteq \mathscr{C}_{BRC}$ *de canal de diffusion à relais est donnée par*

$$
\begin{aligned}
\mathscr{R}_{DF\text{-}DF} = co \bigcup_{P \in \mathscr{Q}} \Big\{ & (R_0 \geq 0, R_1 \geq 0, R_2 \geq 0) : \\
R_0 + R_1 &\leq I_1 - I(U_0, U_1; X_2 | X_1, V_0), \\
R_0 + R_2 &\leq I_2 - I(U_0, U_2; X_1 | X_2, V_0), \\
R_0 + R_1 + R_2 &\leq I_1 + J_2 - I(U_0, U_1; X_2 | X_1, V_0) - I(U_1, X_1; U_2 | X_2, U_0, V_0) - I_M \\
R_0 + R_1 + R_2 &\leq J_1 + I_2 - I(U_0, U_2; X_1 | X_2, V_0) - I(U_1; U_2, X_2 | X_1, U_0, V_0) - I_M \\
2R_0 + R_1 + R_2 &\leq I_1 + I_2 - I(U_0, U_1; X_2 | X_1, V_0) - I(U_0, U_2; X_1 | X_2, V_0) \\
& \quad - I(U_1; U_2 | X_1, X_2, U_0, V_0) - I_M \Big\},
\end{aligned}
$$

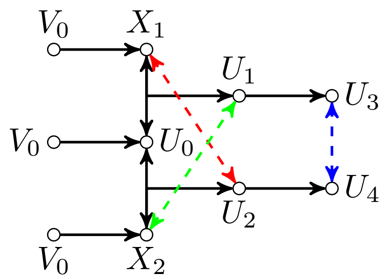


FIGURE 2.3 – Diagramme de variables aléatoires auxiliaires

où (I_i, J_i, I_M) *avec* $i = \{1, 2\}$ *est comme suit*

$$\begin{aligned}
I_i &= \min \left\{ I(U_0, U_i; Z_i|V_0, X_i) + I(U_{i+2}; Y_i|U_0, V_0, X_i, U_i), I(U_0, V_0, U_i, U_{i+2}, X_i; Y_i) \right\}, \\
J_i &= \min \left\{ I(U_i; Z_i|U_0, V_0, X_i) + I(U_{i+2}; Y_i|U_0, V_0, X_i, U_i), I(U_{i+2}, U_i, X_i; Y_i|U_0, V_0) \right\}, \\
I_M &= I(U_3; U_4|U_1, U_2, X_1, X_2, U_0, V_0),
\end{aligned}$$

co$\{\cdot\}$ *signifie l'enveloppe convexe et la union est sur toutes PDs* $P_{U_0V_0U_1U_2U_3U_4X_1X_2X} \in \mathscr{Q}$ *telle que*

$$\mathscr{Q} = \big\{ P_{U_0V_0U_1U_2U_3U_4X_1X_2X} = P_{U_3U_4X|U_1U_2} P_{U_1U_2|U_0X_1X_2} P_{U_0|X_1X_2V_0} P_{X_2|V_0} P_{X_1|V_0} P_{V_0} \\
\textit{satisfaisant} \quad (U_0, V_0, U_1, U_2, U_3, U_4) \ominus (X_1, X_2, X) \ominus (Y_1, Z_1, Y_2, Z_2) \big\}.$$

Proof Ici nous présentons un aperçu de la preuve. D'abord, les messages originaux sont réorganisés via le fendage de débit [2] dans nouveaux messages, comme montré dans Fig. 2.4, où nous ajoutons une partie des messages privés ensemble avec le message commun dans nouveaux messages (de la même façon à [14]). L'idée de codage générale de la preuve est représentée dans Fig.2.3. Le VA V_0 représente la partie commune pour (X_1, X_2) (les informations envoyées par les relais), qui est destiné pour aider les informations communes encodées dans U_0. Les informations privées sont envoyé en deux étapes, en utilisant d'abord l'aide des relais par (U_1, U_2) et basées sur la stratégie DF. Alors le lien direct entre la source et les destinations est utilisé pour décoder(U_3, U_4). Le codage de Marton est utilisée pour permettre la corrélation entre le VAs dénoté par les flèches dans Fig. 2.3. Faire une variable aléatoire était simultanément en corrélation avec VAs multiple, nous avons utilisé le codage de Marton à plusieurs niveaux. Pour ce but, nous commençons avec un ensemble donné de

2. Rate splitting

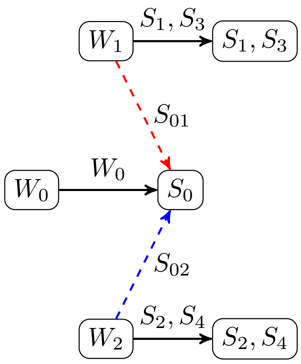


Figure 2.4 – Ré-configuration des message

VAs produit dans une façon i.i.d. et ensuite dans chaque étape nous avons choisi un sous-ensemble tel que tous leurs membres sont conjointement typiques[3] avec fixe VA. Alors dans chaque étape nous cherchons un tel sous-ensemble à l'intérieur du précédent. La table 2.1

Table 2.1 – Stratégie DF avec $b = \{1, 2\}$

$\underline{v}_0(t_{0(i-1)})$	$\underline{v}_0(t_{0(i)})$
$\underline{u}_0(t_{0(i-1)}, t_{0i})$	$\underline{u}_0(t_{0i}, t_{0(i+1)})$
$\underline{x}_b(t_{0(i-1)}, t_{b(i-1)})$	$\underline{x}_b(t_{0i}, t_{bi})$
$\underline{u}_b(t_{0(i-1)}, t_{0i}, t_{b(i-1)}, t_{bi})$	$\underline{u}_b(t_{0i}, t_{0(i+1)}, t_{bi}, t_{b(i+1)})$
$\underline{u}_{b+2}(t_{0(i-1)}, t_{0i}, t_{b(i-1)}, t_{bi}, t_{(b+2)i})$	$\underline{u}_{b+2}(t_{0i}, t_{0(i+1)}, t_{bi}, t_{b(i+1)}, t_{(b+2)(i+1)})$
$\underline{y}_{bi}$	$\underline{y}_{b(i+1)}$

montre des détails pour la transmission dans le temps. Les deux relais connaissant $\underline{v}_0, \underline{x}_b$ décodent $\underline{u}_0, \underline{u}_b$ dans le même bloc. Alors chaque destination en utilisant le décodage en arrière décode tout le livre de codes dans le dernier bloc. La région finale est une combinaison de toutes les contraintes de Marton, le codage et le décodage qui simplifiera à la région en utilisant l'élimination de Fourier-Motzkin.

Remark 3 *On fait les remarques suivantes :*

- *Les deux débit dans Théorème 2 coïncident avec le débit conventionnel basé sur partiel DF [1],*
- *Il est facile de vérifier que, en mettant $(X_1, X_2, V_0) = \emptyset$, $U_3 = U_1, U_4 = U_2$ $Z_1 = Y_1$ et $Z_2 = Y_2$, la région de débit dans Théorème 2 est équivalente à la région de*

3. Jointly Typical

Marton [12],

– *la région précédente améliore celui tiré pour le BRC dans [29] et pour le BRC avec le relais commun comme représenté dans Fig. 2.1(c). En choisissant $X_1 = X_2 = V_0$ et $U_1 = U_2 = U_0$, la région de débit dans Théorème 2 peut être montrée pour être équivalente à la borne intérieure de Kramer et al. dans [14]. Pourtant le corollaire suivant montre que la région de débit dans Théorème 2 est strictement mieux que celle de Kramer et al..*

Le corollaire suivant fournit une borne intérieure plus pointu sur la région de capacité du BRC avec le relais commun (BRC-CR). Dans la région suivante, le relais aide aussi les informations privées pour la première destination en divisant l'aide de relais en deux parties V_0 et X_1 pourtant le relais dans la région de Kramer *et al.* aide seulement des informations communes. Par exemple quand $Y_2 = \emptyset$ et la première destination est la version dégradée du relais, intuitivement quand le canal de deuxième destination est tellement faible que nous pouvons l'ignorer, alors la région de Kramer *et al.* ne peut pas accomplir la capacité du premier canal à relais parce que le relais peut seulement aider les informations communes. Pourtant ce n'est pas le cas dans la région suivante.

Corollary 1 (BRC à relais commun) *Une borne intérieure sur la région de capacité de BRC-CR $\mathscr{R}_{BRC\text{-}CR} \subseteq \mathscr{C}_{BRC\text{-}CR}$ est donnée par*

$$
\begin{aligned}
\mathscr{R}_{BRC\text{-}CR} = co \bigcup_{P_{V_0U_0U_1U_3U_4X_1X} \in \mathscr{Q}} \Big\{ & (R_0 \geq 0, R_1 \geq 0, R_2 \geq 0) : \\
R_0 + R_1 &\leq \min\{I_1 + I_{1p}, I_3 + I_{3p}\} + I(U_3; Y_1|U_1, U_0, X_1, V_0), \\
R_0 + R_2 &\leq I(U_0, V_0, U_4; Y_2) - I(U_0; X_1|V_0), \\
R_0 + R_1 + R_2 &\leq \min\{I_2, I_3\} + I_{3p} + I(U_3; Y_1|U_1, U_0, X_1, V_0) \\
&\quad + I(U_4; Y_2|U_0, V_0) - I(U_0; X_1|V_0) - I_M, \\
R_0 + R_1 + R_2 &\leq \min\{I_1, I_3\} + I_{1p} + I(U_3; Y_1|U_1, U_0, X_1, V_0) \\
&\quad + I(U_4; Y_2|U_0, V_0) - I(U_0; X_1|V_0) - I_M, \\
2R_0 + R_1 + R_2 &\leq I(U_3; Y_1|U_1, U_0, X_1, V_0) + I(U_4; Y_2|U_0, V_0) + I_2 \\
&\quad + \min\{I_1 + I_{1p}, I_3 + I_{3p}\} - I(U_0; X_1|V_0) - I_M \Big\}
\end{aligned}
$$

avec

$$
\begin{aligned}
I_1 &= I(U_0, V_0; Y_1),\\
I_2 &= I(U_0, V_0; Y_2),\\
I_3 &= I(U_0; Z_1|X_1, V_0),\\
I_{1p} &= I(U_1 X_1; Y_1|U_0, V_0),\\
I_{3p} &= I(U_1; Z_1|U_0, V_0, X_1),\\
I_M &= I(U_3; U_4|X_1, U_1, U_0, V_0),
\end{aligned}
$$

$co\{\cdot\}$ *signifie l'enveloppe convexe et* $\mathscr{Q}$ *est l'ensemble de toutes PDs* $P_{V_0U_0U_1U_3U_4X_1X}$ *satisfaisant*

$$(V_0, U_0, U_1, U_3, U_4) \ominus (X_1, X) \ominus (Y_1, Z_1, Y_2).$$

L'idée centrale est que le relais doit ici aider des informations privées et communes pour un utilisateur au moins. Il sera montré dans la section suivante qu'un cas spécial de ce corollaire atteint la capacité de BRC-CR Gaussien dégradé et BRC-CR semi-dégradé.

2.3 Région Atteignable pour la Stratégie CF-DF

Considérons maintenant un canal de diffusion à relais où le canal source-relais est plus fort que le canal relais-destination pour une branche et plus faible pour l'autre branche. La stratégie coopérative est mieux d'être fondé sur DF pour une branche et CF pour l'autre. La source doit transmettre les informations aux destinations basées sur un code de diffusion combiné avec les schémas CF et DF. Ce scénario peut survenir quand l'encodeur ne sait pas (par ex. en raison de la mobilité d'utilisateur et évanouissement) si le canal source-relais est mieux ou pas que le canal relais-destination. Le théorème suivant présente la région de débit atteignable générale pour le cas où le premier relais emploie DF et le deuxième relais emploie CF pour aider des informations communes et privées.

Theorem 3 (Région CF-DF) *Une borne intérieur sur la région de capacité de BRC*

$\mathscr{R}_{DF\text{-}CF} \subseteq \mathscr{C}_{BRC}$ *avec des stratégies coopératives hétérogènes est donnée par*

$$\begin{aligned}
\mathscr{R}_{CF\text{-}DF} = co \bigcup_{P\in\mathscr{Q}} \Big\{ (R_0 \geq 0, & R_1 \geq 0, R_2 \geq 0) : \\
R_0 + R_1 &\leq I_1, \\
R_0 + R_2 &\leq I_2 - I(U_2; X_1|U_0, V_0), \\
R_0 + R_1 + R_2 &\leq I_1 + J_2 - I(U_1, X_1; U_2|U_0, V_0), \\
R_0 + R_1 + R_2 &\leq J_1 + I_2 - I(U_1, X_1; U_2|U_0, V_0), \\
2R_0 + R_1 + R_2 &\leq I_1 + I_2 - I(U_1, X_1; U_2|U_0, V_0) \Big\},
\end{aligned}$$

où la quantité (I_i, J_i, Δ_0) *avec* $i = \{1, 2\}$ *est comme suit*

$$\begin{aligned}
I_1 &= \min\big\{ I(U_0, U_1; Z_1|X_1, V_0), I(U_1, U_0, X_1, V_0; Y_1) \big\}, \\
I_2 &= I(U_2, U_0, V_0; \hat{Z}_2, Y_2|X_2), \\
J_1 &= \min\big\{ I(U_1; Z_1|X_1, U_0, V_0), I(U_1, X_1; Y_1|U_0, V_0) \big\}, \\
J_2 &= I(U_2; \hat{Z}_2, Y_2|X_2, U_0, V_0),
\end{aligned}$$

$co\{\cdot\}$ *signifie l'enveloppe convexe et l'ensemble de toutes admissibles PDs* $\mathscr{Q}$ *est défini comme*

$$\begin{aligned}
\mathscr{Q} = \Big\{ & P_{V_0U_0U_1U_2X_1X_2XY_1Y_2Z_1Z_2\hat{Z}_2} = P_{V_0}P_{X_2}P_{X_1|V_0}P_{U_0|V_0}P_{U_2U_1|X_1U_0}P_{X|U_2U_1} \times \\
& P_{Y_1Y_2Z_1Z_2|XX_1X_2}P_{\hat{Z}_2|X_2Z_2}, \textit{satisfaisant} \\
& I(X_2; Y_2) \geq I(Z_2; \hat{Z}_2|X_2Y_2), \;\; \textit{and} \\
& (V_0, U_0, U_1, U_2) \ominus (X_1, X_2, X) \ominus (Y_1, Z_1, Y_2, Z_2) \Big\}.
\end{aligned}$$

C'est possible pour les canaux de diffusion à relais généraux de changer la stratégie du code dans le premier et deuxième relais, de DF à CF et vice versa et obtenir une autre région. Finalement une plus grande région peut être obtenue en prenant l'union de deux régions. Pour transmettre les informations communes et en même temps exploiter l'aide du relais pour la destination DF, le codage régulière est utilisée avec le codage de bloc-Markov. En fait, V_0 est la partie de X_1 pour aider la transmission de U_0. Mais la deuxième destination utilise CF où l'entré de relais et l'entré de canal sont surtout indépendantes.

Bien qu'il semble, au premier coup d'œil, que ce codage de bloc-Markov avec le codage de superposition ne sont pas compatible avec CF. La source utilise l'encodage régulier et surimpose [4] le code du bloc actuel sur le code du bloc précédent. Quand le relais utilise DF, il transmet le code du bloc précédent et alors la destination peut exploiter cette aide pour décoder tous les codes. Mais quand le relais utilise CF, la destination semble être en face de deux codes surimposés qui doivent être décodé. Parce que le code de centre porte un message factice [5] dans le premier bloc, la destination peut décoder le nuage en sachant le centre. Alors dans le bloc suivant en utilisant la même idée, elle continue à décoder en enlevant le code de centre. Mais cela cause la perte de performance parce qu'une partie du code transmis est effectivement négligée. Donc il semble que le codage de superposition n'est pas propre pour CF. Pourtant il peut être montré que ce n'est pas le cas. En utilisant le décodage en arrière, le code peut être aussi exploité pour CF aussi, sans perte de performance. Effectivement en CF, la destination prend V_0 pas comme le code à relais, mais comme le code de source sur lequel U_0 est superposé. Alors au dernier bloc U_0 porte le message factice, mais superposé sur V_0 qui porte le message du dernier bloc. Alors la destination peut conjointement décoder (U_0, V_0) et exploiter ainsi les deux codes sans perte de performance en ce qui concerne ordinaire CF.

Considérons maintenant le canal à relais compound où le canal en opération est choisi d'un ensemble de canaux à relais. Pour la simplicité supposons que l'ensemble inclut seulement deux canaux tel que la stratégie DF en comparaison de CF résulte à un meilleur débit pour le premier canal et à un plus mauvais débit pour le deuxième. Le but est d'émettre un débit avec arbitrairement petite erreur pour les deux canaux. Ensuite en utilisant l'encodage régulier, on peut voir que la meilleure stratégie coopérative peut être choisie pour chaque canal parce que le relais emploie DF dans le premier canal et CF dans le deuxième canal sans aucun problème. Le corollaire suivant s'ensuit directement de cette observation.

Corollary 2 (information-commun) *Une borne intérieure sur la capacité de canal à relais compound (ou BRC à message commun) est donnée par*

$$R_0 \leq \max_{P_{X_1 X_2 X} \in \mathscr{Q}} \min \left\{ I(X; Z_1 | X_1), I(X, X_1; Y_1), I(X; \hat{Z}_2, Y_2 | X_2) \right\}.$$

4. Superpose
5. Dummy message

Corollaire 2 suit de Théorème 3 avec le choix $U_1 = U_2 = U_0 = X$, $V_0 = X_1$. tandis que le corollaire suivant suit en mettant $U_0 = V_0 = \emptyset$.

Corollary 3 (information privée) *Une borne intérieure sur la région de capacité de BRC avec les stratégies coopératives hétérogènes est donnée par l'enveloppe convexe de l'ensemble de débits* (R_1, R_2) *satisfaisant*

$$R_1 \leq \min\big\{I(U_1; Z_1|X_1), I(U_1, X_1; Y_1)\big\}, \tag{2.1}$$

$$R_2 \leq I(U_2; \hat{Z}_2, Y_2|X_2) - I(U_2; X_1), \tag{2.2}$$

$$R_1 + R_2 \leq \min\big\{I(U_1; Z_1|X_1), I(U_1, X_1; Y_1)\big\} + I(U_2; \hat{Z}_2, Y_2|X_2) - I(U_1, X_1; U_2), \tag{2.3}$$

pour toutes PDs $P_{U_1U_2X_1X_2XY_1Y_2Z_1Z_2\hat{Z}_2} \in \mathscr{Q}$.

Remark 4 *La région dans Théorème 3 est équivalente à la région de Marton [12] avec* $(X_1, X_2, V_0) = \emptyset$, $Z_1 = Y_1$ *et* $Z_2 = Y_2$. *Remarquons que le débit conformément au schéma DF qui apparaît dans Théorème 3 coïncide avec le débit DF conventionnel, alors que le débit CF apparaît avec une petite différence. En fait,* X *est décomposé à* (U, X_1) *qui le remplacent dans le débit conformément à CF.*

2.4 Région Atteignable pour la Stratégie CF-CF

Nous considérons maintenant un autre scénario où les canaux relais-destination sont plus forts que les autres et alors la stratégie de codage efficace tourne pour être CF pour les deux utilisateurs. La borne intérieure basé sur cette stratégie est donné par le théorème suivant.

Theorem 4 (région CF-CF) *Une borne intérieure sur la région de capacité de BRC* $\mathscr{R}_{CF\text{-}CF} \subseteq \mathscr{C}_{BRC}$ *est donnée par*

$$\begin{aligned}
\mathscr{R}_{CF\text{-}CF} = co \bigcup_{P \in \mathscr{Q}} \Big\{ &(R_0 \geq 0, R_1 \geq 0, R_2 \geq 0) : \\
R_0 + R_1 &\leq I(U_0, U_1; Y_1, \hat{Z}_1|X_1), \\
R_0 + R_2 &\leq I(U_0, U_2; Y_2, \hat{Z}_2|X_2), \\
R_0 + R_1 + R_2 &\leq I_0 + I(U_1; Y_1, \hat{Z}_1|X_1, U_0) + I(U_2; Y_2, \hat{Z}_2|X_2, U_0) - I(U_1; U_2|U_0), \\
2R_0 + R_1 + R_2 &\leq I(U_0, U_1; Y_1, \hat{Z}_1|X_1) + I(U_0, U_2; Y_2, \hat{Z}_2|X_2) - I(U_1; U_2|U_0) \Big\},
\end{aligned}$$

où I_0 est défini par

$$I_0 = \min\left\{I(U_0; Y_1, \hat{Z}_1|X_1), I(U_0; Y_2, \hat{Z}_2|X_2)\right\},$$

co{·} signifie l'enveloppe convexe et l'ensemble de toute PDs admissibles $\mathscr{Q}$ est défini comme

$$\begin{aligned}\mathscr{Q} = \big\{P_{U_0U_1U_2X_1X_2XY_1Y_2Z_1Z_2\hat{Z}_1\hat{Z}_2} = P_{X_2}P_{X_1}P_{U_0}P_{U_2U_1|U_0}P_{X|U_2U_1}\times \\ P_{Y_1Y_2Z_1Z_2|XX_1X_2}P_{\hat{Z}_1|X_1Z_1}P_{\hat{Z}_2|X_2Z_2}, \\ I(X_2; Y_2) \geq I(Z_2; \hat{Z}_2|X_2, Y_2), \\ I(X_1; Y_1) \geq I(Z_1; \hat{Z}_1|X_1, Y_1), \\ (U_0, U_1, U_2) \ominus (X_1, X_2, X) \ominus (Y_1, Z_1, Y_2, Z_2)\big\}.\end{aligned}$$

Notons que cette région est équivalente à la région de Marton [12] en mettant $(X_1, X_2) = \emptyset$, $Z_1 = Y_1$ et $Z_2 = Y_2$.

Remark 5 *Une région de débit atteignable générale suit en utilisant le partage de temps*[6] *entre toutes les régions précédentes exposées dans les Théorèmes 2, 3 et 4.*

2.5 Bornes extérieures sur la région de capacité de BRC général

Les théorèmes suivants fournissent des bornes extérieures générales sur les régions de capacité du BRC et du BRC-CR où $X_1 = X_2$ et $Z_1 = Z_2$, respectivement.

Theorem 5 (borne extérieur de BRC) *La région de capacité $\mathscr{C}_{BRC}$ de BRC (voir*

6. Time-sharing

Fig. 2.2) est inclue dans l'ensemble $\mathscr{C}_{BRC}^{out}$ *de tous les débits* (R_0, R_1, R_2) *satisfaisant*

$$\begin{aligned}
\mathscr{C}_{BRC}^{out} = co \bigcup_{P_{VV_1U_1U_2X_1X}\in\mathscr{Q}} \Big\{(R_0 \geq & 0, R_1 \geq 0, R_2 \geq 0) : \\
R_0 \leq & \min\big\{I(V;Y_2), I(V;Y_1)\big\}, \\
R_0 + R_1 \leq & \min\big\{I(V;Y_1), I(V;Y_2)\big\} + I(U_1;Y_1|V), \\
R_0 + R_2 \leq & \min\big\{I(V;Y_1), I(V;Y_2)\big\} + I(U_2;Y_2|V), \\
R_0 + R_1 \leq & \min\big\{I(V,V_1;Y_1,Z_1|X_1), I(V,V_1;Y_2,Z_2)\big\} \\
& + I(U_1;Y_1,Z_1|V,V_1,X_1), \\
R_0 + R_2 \leq & \min\big\{I(V,V_1;Y_1,Z_1|X_1), I(V,V_1;Y_2,Z_2)\big\} \\
& + I(U_2;Y_2,Z_2|V,V_1,X_1), \\
R_0 + R_1 + R_2 \leq & I(V;Y_1) + I(U_2;Y_2|V) + I(U_1;Y_1|U_2,V), \\
R_0 + R_1 + R_2 \leq & I(V;Y_2) + I(U_1;Y_1|V) + I(U_2;Y_2|U_1,V), \\
R_0 + R_1 + R_2 \leq & I(V,V_1;Y_1,Z_1|X_1) + I(U_2;Y_2,Z_2|V,V_1,X_1) \\
& + I(U_1;Y_1,Z_1|X_1,U_2,V,V_1), \\
R_0 + R_1 + R_2 \leq & I(V,V_1;Y_2,Z_2) + I(U_1;Y_1,Z_1|V,V_1,X_1) \\
& + I(U_2;Y_2,Z_2|X_1,U_1,V,V_1)\Big\},
\end{aligned}$$

où $co\{\cdot\}$ *signifie l'enveloppe convexe et* $\mathscr{Q}$ *est l'ensemble de toutes PDs* $P_{VV_1U_1U_2X_1X_2X}$ *satisfaisant* $X_1 \ominus V_1 \ominus (V, U_1, U_2, X)$*. La cardinalité de VAs auxiliaires sont conditionné par* $\|\mathscr{V}\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\|\|\mathscr{X}_2\|\|\mathscr{Z}_1\|\|\mathscr{Z}_2\| + 25$, $\|\mathscr{V}_1\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\|\|\mathscr{X}_2\|\|\mathscr{Z}_1\|\|\mathscr{Z}_2\| + 17$ *and* $\|\mathscr{U}_1\|, \|\mathscr{U}_2\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\|\|\mathscr{X}_2\|\|\mathscr{Z}_1\|\|\mathscr{Z}_2\| + 8$.

Remark 6 *On peut voir de la preuve que* V_1 *est une variable aléatoire composant des parties causales et non-causales du relais. Ainsi* V_1 *peut être intuitivement considéré comme l'aide de relais pour* V*. Il peut aussi être déduit de la forme de la borne supérieures que* V *et* U_1, U_2 *représentent respectivement les informations communes et privées.*

Remark 7 *On fais les observations suivantes*

- *La borne extérieur est valide pour général BRC, c'est-à-dire pour des canaux de diffusion de 2 relais de 2 récepteurs. Pourtant dans notre cas, la paire de* Y, Y_b *dépend seulement de* X, X_b *pour* $b = 1, 2$*. L'utilisation de ces relations de Markov,*$I(U_b; Y_b, Z_b|X_b, T)$ *et* $I(U_b; Y_b|T)$ *peut être borné par* $I(X; Y_b, Z_b|X_b, T)$ *et* $I(X, X_b; Y_b|T)$ *pour la variable aléatoire* $T \in \{V, V_1, U_1, U_2\}$*. Cela simplifiera la région précédente.*

- *De plus nous pouvons voir que la région dans Théorème 5 n'est pas complètement symétrique. Donc une autre borne supérieure peut être obtenu en remplaçant l'indices 1 et 2, c'est-à-dire en présentant V_2 et X_2 au lieu V_1 et X_1. La borne finale sera l'intersection de ces deux régions.*
- *Si les relais ne sont pas présents, c'est-à-dire, $Z_1 = Z_2 = X_1 = X_2 = V_1 = \emptyset$, il n'est pas difficile de voir que la précédente borne réduit à la borne extérieur pour les canaux de diffusion généraux fait allusion à comme UVW-extérieur a borne [31]. En outre, il a été récemment montré qu'un tel relié est au moins aussi bon que toutes les bornes extérieures actuellement développées pour la région de capacité de canaux de diffusion [32].*

Le théorème suivant présente une borne supérieure sur la capacité du message commun BRC. La borne supérieure est utile pour l'évaluation de la capacité dans le canal à relais compound.

Theorem 6 (borne extérieure à information commune) *Une borne supérieure sur la capacité de BRC à message commun est donnée par*

$$R_0 \leq \max_{P_{X_1 X_2 X} \in \mathscr{Q}} \min \big\{ I(X;Z_1 Y_1|X_1), I(X,X_1;Y_1), I(X;Z_2,Y_2|X_2), I(X,X_2;Y_2) \big\}.$$

Le théorème suivant présente une borne extérieure sur la région de capacité du BRC avec le relais commun. Dans ce cas-là, en raison du fait que $Z_1 = Z_2$ et $X_1 = X_2$, nous pouvons choisir $V_1 = V_2$ à cause de la définition de V_b. Donc la borne extérieure du Théorème 5 avec la borne extérieure symétrique mentionné au-dessus, qui profite X_2, V_2, résulte à la borne suivante.

Theorem 7 (borne extérieure BRC-CR) *la région de capacité $\mathscr{C}_{BRC\text{-}CR}$ de BRC-CR* *est inclue dans l'ensemble $\mathscr{C}^{out}_{BRC\text{-}CR}$ de tous débits (R_0, R_1, R_2) satisfaisant*

$$\begin{aligned}
\mathscr{C}^{out}_{BRC\text{-}CR} = co \bigcup_{P_{VV_1U_1U_2X_1X}\in\mathscr{Q}} \Big\{ &(R_0 \geq 0, R_1 \geq 0, R_2 \geq 0) : \\
R_0 &\leq \min\big\{I(V;Y_2), I(V;Y_1)\big\}, \\
R_0 + R_1 &\leq \min\big\{I(V;Y_1), I(V;Y_2)\big\} + I(U_1;Y_1|V), \\
R_0 + R_2 &\leq \min\big\{I(V;Y_1), I(V;Y_2)\big\} + I(U_2;Y_2|V), \\
R_0 + R_1 &\leq \min\big\{I(V,V_1;Y_1,Z_1|X_1), I(V,V_1;Y_2,Z_1|X_1)\big\} \\
&\quad + I(U_1;Y_1,Z_1|V,V_1,X_1), \\
R_0 + R_2 &\leq \min\big\{I(V,V_1;Y_1,Z_1|X_1), I(V,V_1;Y_2,Z_1|X_1)\big\} \\
&\quad + I(U_2;Y_2,Z_1|V,V_1,X_1), \\
R_0 + R_1 + R_2 &\leq I(V;Y_1) + I(U_2;Y_2|V) + I(U_1;Y_1|U_2,V), \\
R_0 + R_1 + R_2 &\leq I(V;Y_2) + I(U_1;Y_1|V) + I(U_2;Y_2|U_1,V), \\
R_0 + R_1 + R_2 &\leq I(V,V_1;Y_1,Z_1|X_1) + I(U_2;Y_2,Z_1|V,V_1,X_1) \\
&\quad + I(U_1;Y_1,Z_1|X_1,U_2,V,V_1), \\
R_0 + R_1 + R_2 &\leq I(V,V_1;Y_2,Z_1|X_1) + I(U_1;Y_1,Z_1|V,V_1,X_1) \\
&\quad + I(U_2;Y_2,Z_1|X_1,U_1,V,V_1)\Big\},
\end{aligned}$$

où $co\{\cdot\}$ signifie l'enveloppe convexe et $\mathscr{Q}$ est l'ensemble de toutes PDs $P_{VV_1U_1U_2X_1X}$ vérifiant $(X_1) \multimap V_1 \multimap (V, U_1, U_2, X)$ où la cardinalité de VAs auxiliaires satisfait $\|\mathscr{V}\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\|\|\mathscr{Z}_1\| + 19$, $\|\mathscr{V}_1\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\|\|\mathscr{Z}_1\| + 11$ et $\|\mathscr{U}_1\|, \|\mathscr{U}_2\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\|\|\mathscr{Z}_1\| + 8$.

2.6 BRC à relais commun dégradé et semi-dégradé

Nous présentons maintenant des bornes intérieures et extérieures et des résultats de capacité pour une classe spéciale de BRC-CR. On définit d'abord deux classes de BRC-CRs.

Definition 3 (BRC-CR dégradé) *On dit qu'un canal de diffusion à relais avec le relais commun (BRC-CR) (montré dans Fig. 2.3), qui signifie $Z_1 = Z_2$ et $X_1 = X_2$, est dégradé (respectivement semi-dégradé) si PD $\big\{P_{Y_1Z_1Y_2|XX_1} : \mathscr{X} \times \mathscr{X}_1 \longmapsto \mathscr{Y}_1 \times \mathscr{Z}_1 \times \mathscr{Y}_2\big\}$ satisfait les chaînes de Markov pour un des cas suivants :*

I $X \multimap (X_1, Z_1) \multimap (Y_1, Y_2)$ *et* $(X, X_1) \multimap Y_1 \multimap Y_2$,

II $X \ominus (X_1, Z_1) \ominus Y_2$ *et* $X \ominus (Y_1, X_1) \ominus Z_1$,

où la condition (I) est considérée comme BRC-CR dégradé la condition (II) est considérée comme BRC-CR semi-dégradé .

Remarquons que BRC-CR dégradé peut être vu comme la combinaison d'un canal à relais dégradé avec un canal de diffusion dégradé. D'autre part, le cas semi-dégradé peut être vu comme la combinaison d'un canal de diffusion dégradé avec un canal à relais contrairement dégradé. La région de capacité de BRC-CR semi-dégradé est exposée dans le théorème suivant.

Theorem 8 (semi-degraded BRC-CR) *la région de capacité de BRC-CR semi-dégradé est donnée par la région suivante :*

$$\mathscr{C}_{BRC\text{-}CR} = \bigcup_{P_{UX_1X} \in \mathscr{Q}} \Big\{ (R_1 \geq 0, R_2 \geq 0) :$$

$$R_2 \leq \min\{I(U, X_1; Y_2), I(U; Z_1|X_1)\},$$

$$R_1 + R_2 \leq \min\{I(U, X_1; Y_2), I(U; Z_1|X_1)\} + I(X; Y_1|X_1, U) \Big\},$$

où $\mathscr{Q}$ *est l'ensemble de toutes PDs* P_{UX_1X} *satisfaisant* $U \ominus (X_1, X) \ominus (Y_1, Z_1, Y_2)$ *où la cardinalité de VA auxiliaire* U *satisfait* $\|\mathscr{U}\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\| + 2$.

Les théorèmes suivants fournissent des bornes extérieures et intérieures sur la région de capacité de BRC-CR dégradé.

Theorem 9 (BRC-CR dégradé) *la région de capacité* $\mathscr{C}_{BRC\text{-}CR}$ *de BRC-CR dégradé est inclue dans l'ensemble de débits* (R_0, R_1) *satisfaisant*

$$\mathscr{C}^{out}_{BRC\text{-}CR} = \bigcup_{P_{UX_1X} \in \mathscr{Q}} \Big\{ (R_0 \geq 0, R_1 \geq 0) :$$

$$R_0 \leq I(U; Y_2),$$

$$R_1 \leq \min\big\{I(X; Z_1|X_1, U), I(X, X_1; Y_1|U)\big\},$$

$$R_0 + R_1 \leq \min\big\{I(X; Z_1|X_1), I(X, X_1; Y_1)\big\} \Big\},$$

où $\mathscr{Q}$ *est l'ensemble de toutes PDs* P_{UX_1X} *satisfaisant* $U \ominus (X_1, X) \ominus (Y_1, Z_1, Y_2)$ *où la cardinalité VA auxiliaire* U *satisfait* $\|\mathscr{U}\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\| + 2$.

Il n'est pas difficile de voir que, en appliquant la condition dégradée, la borne supérieure de Théorème 9 est inclus dans celle de Théorème 7.

Theorem 10 (BRC-CR dégradé) *La borne intérieure sur la région de capacité* $\mathscr{R}_{BRC\text{-}CR} \subseteq \mathscr{C}_{BRC\text{-}CR}$ *de BRC-CR est donnée par l'ensemble de débits* (R_0, R_1) *satisfaisant*

$$\mathscr{R}_{BRC\text{-}CR} = co \bigcup_{P_{UVX_1X} \in \mathscr{Q}} \Big\{ (R_0 \geq 0, R_1 \geq 0) :$$

$$\begin{aligned} R_0 &\leq I(U,V;Y_2) - I(U;X_1|V), \\ R_0 + R_1 &\leq \min\big\{I(X;Z_1|X_1,V), I(X,X_1;Y_1)\big\}, \\ R_0 + R_1 &\leq \min\big\{I(X;Z_1|X_1,U,V), I(X,X_1;Y_1|U,V)\big\} \\ &\quad + I(U,V;Y_2) - I(U;X_1|V)\Big\}, \end{aligned}$$

où $co\{\cdot\}$ *signifie l'enveloppe convexe pour toutes PDs dans* $\mathscr{Q}$ *vérifiant*

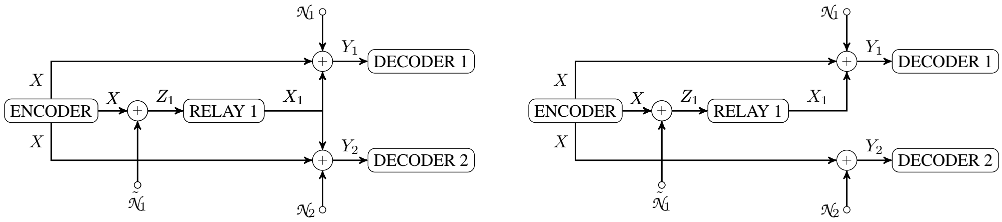
$$P_{UVX_1X} = P_{X|UX_1} P_{X_1U|V} P_V$$

avec $(U,V) \ominus (X_1,X) \ominus (Y_1,Z_1,Y_2)$.

Remark 8 *Dans la borne précédente V peut être intuitivement pris comme l'aide de relais pour* R_0*. La partie délicate est comment partager l'aide de relais entre les informations communes et privées. D'une part, le choix de* $V = \emptyset$ *enlèverait l'aide de relais pour les informations communes et alors pour le cas de* $Y_1 = Y_2$ *il impliquerait que l'aide de relais n'est pas exploitée et ainsi la région sera sous-optimale. Alors que le choix de* $V = X_1$ *causera un problème semblable quand* $Y_2 = \emptyset$*. Le code pour les informations communes ne peut pas être superposé sur le code de relais entier parce qu'il limite l'aide de relais pour les informations privées. Une solution est de super-imposer le code commun d'information à une variable aléatoire supplémentaire V qui joue le rôle de l'aide de relais pour les informations communes. Pourtant cela provoque un autre problème. Maintenant que U n'est pas superposé sur* X_1*, ces variables n'ont plus de dépendance complète et alors la borne extérieure ne tient pas pour le canal. Pour résumer, le codage de Marton enlève le problème de la corrélation avec le prix de déviation de la borne extérieur, c'est-à-dire les termes négatifs dans les bornes intérieures. C'est la raison principale pourquoi les bornes ne sont pas serrées pour BRC dégradé avec le relais commun.*

2.7 BRC Gaussien dégradé avec relais commun

D'une façon intéressante, les bornes intérieures et extérieures données par Théorèmes 10 et 9 arrivent à coïncider pour le cas de Gaussien dégradé BRC-CR, Fig. 2.5(a). La



(a) DG-BRC avec relais commun

(b) DG-BRC avec coopération partielle

FIGURE 2.5 – BRC Gaussien Dégradé (DG-BRC)

capacité de ce canal a été d'abord dérivée via une différente approche dans [33]. On définit BRC-CR Gaussien dégradé par les sorties de canal suivantes :

$$\begin{aligned} Y_1 &= X + X_1 + \mathcal{N}_1, \\ Y_2 &= X + X_1 + \mathcal{N}_2, \\ Z_1 &= X + \tilde{\mathcal{N}}_1 \end{aligned}$$

où la source et le relais ont des contraintes de pouvoir P, P_1 et $\mathcal{N}_1, \mathcal{N}_2, \tilde{\mathcal{N}}_1$ sont des bruits Gaussien indépendants avec les variances $N_1, N_2, \tilde{N}_1$, respectivement, tel que les bruits $\mathcal{N}_1, \mathcal{N}_2, \tilde{\mathcal{N}}_1$ satisfont les conditions de Markov nécessaires dans la définition 3. Notons qu'il est suffisant de supposer les récepteurs comme versions physiquement dégradé [7] du relais et prendre un récepteur comme seulement une version dégradé stochastique [8] d'autre récepteur. Cela signifie qu'il y a $\mathcal{N}, \mathcal{N}'$ tel que :

$$\begin{aligned} \mathcal{N}_1 &= \tilde{\mathcal{N}}_1 + \mathcal{N}, \\ \mathcal{N}_2 &= \tilde{\mathcal{N}}_1 + \mathcal{N}'. \end{aligned}$$

et aussi $N_1 < N_2$. Le théorème suivant tient comme un cas spécial des Théorèmes 9 et 10.

Theorem 11 (BRC-CR Gaussien dégradé) *La région de capacité de BRC-CR Gaus-*

7. Physically degraded
8. Stochastically degraded

sien dégradé est donnée par

$$\mathscr{C}_{BRC\text{-}CR} = \bigcup_{0\leq\beta,\alpha\leq 1} \Big\{(R_0 \geq 0, R_1 \geq 0) :$$

$$R_0 \leq C\left(\frac{\alpha(P + P_1 + 2\sqrt{\overline{\beta}PP_1})}{\overline{\alpha}(P + P_1 + 2\sqrt{\overline{\beta}PP_1}) + N_2}\right),$$

$$R_1 \leq C\left(\frac{\overline{\alpha}(P + P_1 + 2\sqrt{\overline{\beta}PP_1})}{N_1}\right),$$

$$R_0 + R_1 \leq C\left(\frac{\beta P}{\tilde{N}_1}\right)\Big\},$$

où $C(x) = 1/2\log(1 + x)$.

α et β peuvent être respectivement interprété comme l'allocation de pouvoir à la source pour deux destinations et le coefficient de corrélation entre le code à relais et la source.

2.8 BRC Gaussien Dégradé avec Coopération Partielle

Nous présentons maintenant une autre région de capacité pour BRC dégradé Gaussien avec coopération partielle (BRC-PC), Fig. 2.5(b), où il n'y a aucune coopération de destination-relais pour la deuxième destination et la première destination est la version dégradée de l'observation à relais. De plus la première destination est la version dégradée stochastique de l'observation de relais.

Les relations des entrées et sorties sont comme suit :

$$Y_1 = X + X_1 + \mathcal{N}_1,$$
$$Y_2 = X + \mathcal{N}_2,$$
$$Z_1 = X + \tilde{\mathcal{N}}_1.$$

La source et le relais ont des contraintes de pouvoir P, P_1 et $\mathcal{N}_1, \mathcal{N}_2, \tilde{\mathcal{N}}_1$ sont des bruits Gaussien indépendants avec les variances $N_1, N_2, \tilde{N}_1$ et il y a $\mathcal{N}$ tel que $\mathcal{N}_1 = \tilde{\mathcal{N}}_1 + \mathcal{N}$ qui signifie que Y_1 est physiquement dégradé par rapport Z_1. Nous supposons aussi $N_2 < \tilde{N}_1$ entre Y_2 et Z_1. Pour ce canal le théorème suivant tient.

Theorem 12 (BRC-PC Gaussien Dégradé) *la région de capacité de BRC-PC Gaus-*

sien dégradé est donnée par :

$$\mathscr{C}_{BRC\text{-}PC} = \bigcup_{0\leq\beta,\alpha\leq 1} \Big\{ (R_1 \geq 0, R_2 \geq 0) :$$

$$R_1 \leq \max_{\beta\in[0,1]} \min \left\{ C\left(\frac{\alpha\beta P}{\overline{\alpha}P + \widehat{N}_1}\right), C\left(\frac{\alpha P + P_1 + 2\sqrt{\overline{\beta}\alpha P P_1}}{\overline{\alpha}P + N_1}\right) \right\},$$

$$R_2 \leq C\left(\frac{\overline{\alpha}P}{N_2}\right) \Big\},$$

où $C(x) = 1/2\log(1+x)$.

α et β sont comme auparavant. Effectivement la source désigne le pouvoir αP afin de porter le message de Y_1 et $\overline{\alpha}P$ pour Y_2. Le théorème est en effet semblable au Théorème 8 sur la capacité de BRC semi-dégradé. Y_2 est le meilleur récepteur donc il peut décoder le message destiné pour Y_1 même après être aidé par le relais. Cela signifie que la première destination et le relais apparaissent tous ensemble comme dégradé à la deuxième destination. Donc la deuxième destination peut correctement décoder l'interférence d'autres utilisateurs et exploiter complètement le pouvoir alloué à cela $\overline{\alpha}P$ comme on peut voir dans la dernière condition de Théorème 12. Notons pourtant que Z_1 n'est pas nécessairement physiquement dégradé par rapport Y_2 qui fait un plus fort résultat que celui de Théorème 8.

2.9 Résultats Numériques

2.9.1 Source est ignorante de la stratégie coopérative adopté par le relais

2.9.1.1 RC Compound

Considérons d'abord des bornes intérieures et supérieures sur le débit commun pour la région DF-CF. La définition des canaux reste le même. Nous mettons $X = U + \sqrt{\frac{\overline{\beta}P}{P_1}}X_1$ et évaluons Corollaire 2. Le but est d'envoyer des informations communes au débit R_0. Il est facile de vérifier que les deux débit DF sont :

$$R_{DF} \leq \min \left\{ C\left(\frac{\beta P}{d_{z_1}^{\delta}\tilde{N}_1}\right), C\left(\frac{\frac{P}{d_{y_1}^{\delta}} + \frac{P_1}{d_{z_1y_1}^{\delta}} + 2\sqrt{\frac{\overline{\beta}PP_1}{d_{y_1}^{\delta}d_{z_1y_1}^{\delta}}}}{N_1}\right) \right\}, \tag{2.4}$$

R_{DF} est le débit atteignable pour la destination Y_1. Pour la destination Y_2, le débit de CF $I(X;Y_2,\hat{Z}_2|X_2)$ suit comme

$$R_{CF} \leq C\left(\frac{P}{d_{y_2}^{\delta} N_2} + \frac{P}{d_{z_2}^{\delta}(\widehat{N}_2 + \tilde{N}_2)}\right). \tag{2.5}$$

La borne supérieure du Théorème 6 se transforme en débit suivant

$$C = \max_{0\leq\beta_1,\beta_2\leq 1} \min\Bigg\{ C\left(\beta_1 P\left[\frac{1}{d_{z_1}^{\delta}\tilde{N}_1} + \frac{1}{d_{y_1}^{\delta}N_1}\right]\right), C\left(\frac{\frac{P}{d_{y_1}^{\delta}} + \frac{P_1}{d_{z_1y_1}^{\delta}} + 2\sqrt{\frac{\overline{\beta}_1 P P_1}{d_{y_1}^{\delta}d_{z_1y_1}^{\delta}}}}{N_1}\right),$$
$$C\left(\beta_2 P\left[\frac{1}{d_{z_2}^{\delta}\tilde{N}_2} + \frac{1}{d_{y_2}^{\delta}N_2}\right]\right), C\left(\frac{\frac{P}{d_{y_2}^{\delta}} + \frac{P_2}{d_{z_2y_2}^{\delta}} + 2\sqrt{\frac{\overline{\beta}_2 P P_2}{d_{y_2}^{\delta}d_{z_2y_2}^{\delta}}}}{N_2}\right)\Bigg\}. \tag{2.6}$$

Remarquons que le débit (2.5) est exactement le même comme le débit CF Gaussien [14]. Cela signifie que l'encodage régulier DF peut aussi être décodé avec la stratégie CF, aussi pour le cas avec le relais proche du récepteur (semblable à [34]). En utilisant le codage proposée c'est possible d'envoyer des informations communes au débit minimal entre les schémas CF et DF $R_0 = \min\{R_{DF}, R_{CF}\}$ (c'est-à-dire les expressions (2.4) à (2.5)). Pour le cas d'informations privées, tous paires de débit ($R_{DF} \leq R_1^*, R_{CF} \leq R_2^*$) sont admissibles, où

$$R_1^* = \max_{0\leq\beta,\lambda\leq 1} \min\left\{R_{11}^{(\beta,\lambda)}, R_{12}^{(\beta,\lambda)}\right\}. \tag{2.7}$$

$$R_2^* = C\left(\frac{\overline{\alpha}P}{d_{y_2}^{\delta}N_2 + \beta\alpha P} + \frac{\overline{\alpha}P}{d_{z_2}^{\delta}(\widehat{N}_2 + \tilde{N}_2) + \beta\alpha P}\right). \tag{2.8}$$

Alors (R_{DF}, R_{CF}) peuvent être simultanément transmis.

Fig. 2.6 montre l'évaluation numérique de R_0 pour le cas de débit commun. Tous les bruits de canal sont mis au variance d'unité et $P = P_1 = P_2 = 10$. La distance entre X et (Y_1, Y_2) est 1, pendant que $d_{z_1} = d_1$, $d_{z_1y_1} = 1 - d_1$, $d_{z_2} = d_2$, $d_{z_2y_2} = 1 - d_2$. le relais 1 bouge avec $d_1 \in [-1, 1]$ et Fig. 2.6 présente des débit en fonction de d_1. Mais la position du relais 2 est supposée fixe à $d_2 = 0.7$ ainsi R_{CF} qui ne dépend pas de d_1, est une fonction constante de d_1. D'autre part R_{DF} dépend de d_1. le débit CF pour Y_1 est aussi tracé qui correspond au cas où le premier relais utilise CF. Ce cadre sert pour comparer

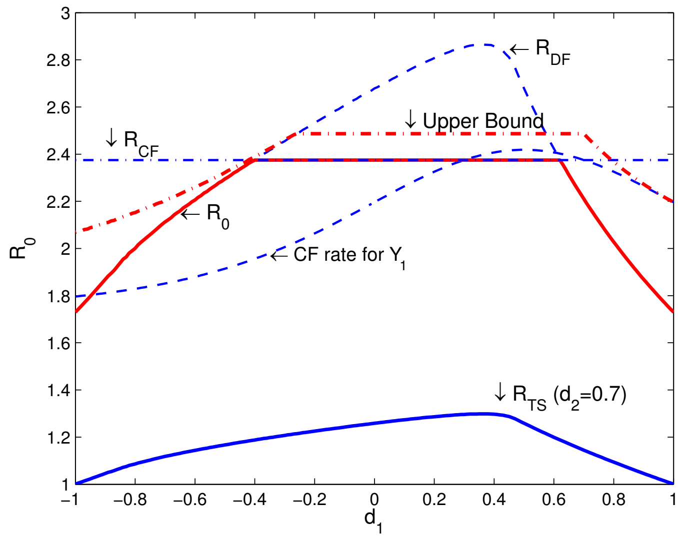


Figure 2.6 – Débit commun de BRC Gaussien avec stratégies DF-CF

les performances de nos schémas de codage quant à la position du relais. On peut voir que l'on peut accomplir le minimum entre les deux débits possibles CF et DF. Ces débit sont aussi comparés à une stratégie de partage de temps naïve qui se compose à l'utilisation du schéma de DF $\tau\%$ du temps et le schéma CF $(1-\tau)\%$ du temps [9]. Le partage de temps produit le débit atteignable

$$R_{TS} = \max_{0\leq\tau\leq1} \min\{\tau R_{DF}, (1-\tau)R_{CF}\}.$$

Remarquons qu'avec le schéma de codage proposé les augmentations significatives peuvent être accomplies quand le relais est près de la source (c'est-à-dire. Le schéma de DF est plus convenable), comparé au plus mauvais cas.

2.9.1.2 RC Composite

Considérons maintenant un modèle composite où le relais est proche de la source avec la probabilité p (appelez-le comme le premier canal) et proche de la destination avec la probabilité $1-p$ (le deuxième canal). Donc le schéma de DF est la stratégie convenable pour le premier canal pendant que le schéma CF joue mieux sur le deuxième. Pour chacun

9. Il ne faudrait pas confondre le partage de temps dans les cadres compound avec le partage de temps conventionnel qui produit la combinaison convexe de débit.

triplet du débit (R_0, R_1, R_2) nous définissons le débit espéré comme

$$R_{av} = R_0 + pR_1 + (1-p)R_2.$$

Le débit espéré basé sur la stratégie de codage proposée est comparé aux stratégies conventionnelles. Les schémas de codage alternatifs pour ce scénario sont possibles où l'encodeur peut simplement investir sur un schéma de codage DF ou CF., qui est utile quand probabilités sont hautes. Il y a de différentes façons de procéder :

– En envoyant des informations via le schéma DF au meilleur débit possible entre les deux canaux. Alors le plus mauvais canal ne peut pas décoder et ainsi le débit espéré devient $p_{DF}^{\max} R_{DF}^{\max}$, où $R_{DF}^{\max}$ est le débit DF accompli sur le meilleur canal et $p_{DF}^{\max}$ est sa probabilité.

– En envoyant des informations via le schéma DF au débit du plus mauvais (deuxième) canal et alors les deux utilisateurs peuvent décoder les informations au débit $R_{DF}^{\min}$. Finalement le débit espéré suivant est atteignable en investissant sur seulement un schéma du code

$$R_{av}^{DF} = \max\left\{p_{DF}^{\max} R_{DF}^{\max}, R_{DF}^{\min}\right\}.$$

– En investissant sur le schéma CF avec les mêmes arguments qu'avant le débit espéré écrit comme

$$R_{av}^{CF} = \max\left\{p_{CF}^{\max} R_{CF}^{\max}, R_{CF}^{\min}\right\},$$

avec les définitions de $(R_{CF}^{\min}, R_{CF}^{\max}, p_{CF}^{\max})$ semblable à auparavant.

Fig.2.7 montre l'évaluation numérique du débit moyen. Tous les bruits de canal sont mis au variance d'unité et $P = P_1 = P_2 = 10$. La distance entre X et (Y_1, Y_2) est $(3, 1)$, pendant que $d_{z_1} = 1$, $d_{z_1 y_1} = 2$, $d_{z_2} = 0.9$, $d_{z_2 y_2} = 0.1$. Comme on peut voir, la stratégie de débit commune fournit un débit tout le temps qui est toujours mieux que le plus mauvais cas. Pourtant dans un coin les investissements complets sur un débit est mieux puisque la haute probabilité d'un canal réduit l'effet de l'autre. Basé sur le schéma de codage proposé, c'est-à-dire l'utilisation du codage privé et du codage commun en même temps, on peut couvrir les points de coin et toujours faire mieux que les deux stratégies d'investissements complètes. Il vaut pour noter que dans cette région de coin, seulement les informations privées d'un canal sont nécessaires.

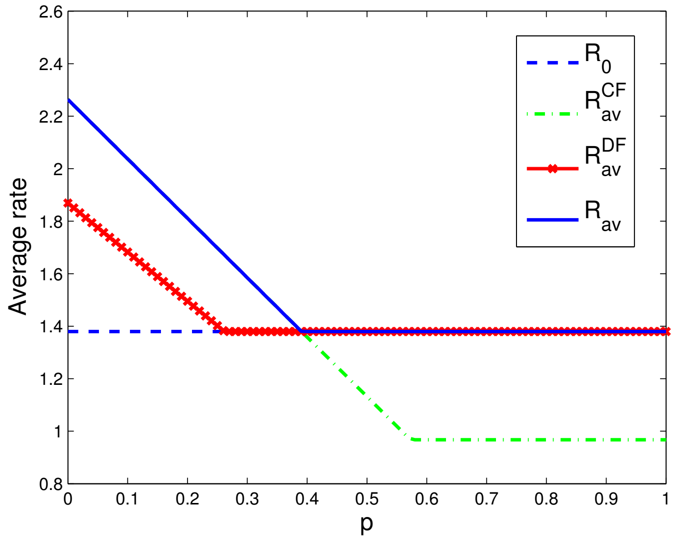


Figure 2.7 – Débit espéré pour le canal à relais Gaussien composite

Chapitre 3

Stratégie de Codage Sélectif pour les Réseaux Unicasts Composites

Comme mentionné auparavant, la nature variable avec temps de canaux sans fil, par ex. en raison d'évanouissement et mobilité d'utilisateur, ne permet pas aux terminus d'avoir une connaissance complète de tous les paramètres de canal impliqués dans la communication. Particulièrement sans feedback les informations d'état de canal (CSI) ne peuvent pas être disponible pour les encodeurs. Pendant des années, un ensemble de travaux a été fait en adressant des aspects tant théoriques que pratiques de problèmes de communication en présence de l'incertitude de canal. Peut-être, de point de vue de la théorie d'information, le canal compound d'abord présenté par Wolfowitz [21] est le modèle le plus important pour s'occuper de l'incertitude de canal, qui continue à attirer une grande partie d'attention des chercheurs (voir [35] et les références là). Les modèles composites sont plus adaptés pour s'occuper des scénarios sans fil car ils adressent l'incertitude de canal en présentant un PD $\mathbb{P}_\theta$ sur les canaux. Ces modèles se composent d'un ensemble de PDs conditionnel avec l'index θ de canal actuel - le vecteur de paramètres - tiré selon $\mathbb{P}_\theta$ et fixe pendant la communication. La capacité pour cette classe de canaux a été largement étudiée au préalable (voir [27] et les références là), pour les scénarios sans fil via la notion célèbre de capacité de panne (voir [36] et des références là) et la coopération ignorant des canaux de Gaussien évanouissement dans [28, 37, 38].

Ici nous enquêtons sur le canal à relais composite où l'index $\theta \in \Theta$ de canal est tiré au hasard selon $\mathbb{P}_\theta$. Le tirage de canal $\theta = (\theta_r, \theta_d)$ reste fixe pendant la communication, pourtant il est inconnu à la source, complètement connu à la destination et partiellement connu θ_r au relais. Bien qu'une approche compound puisse garantir une probabilité d'erreur asymptotiquement zéro sans tenir compte de θ, il serait pas un choix adéquat pour la plupart de scénarios sans fil où l'index du canal le plus mauvais possible résulte à un débit non-positifs. Une différente approche à ce problème qui est surtout préféré quand on s'occupe des modèles sans fil consiste au choix du débit de code r sans tenir compte de l'index de canal actuel. Alors l'encodeur ne peut pas nécessairement garantir - peu importe la valeur de r - une probabilité d'erreur arbitrairement petite. En fait, la probabilité d'erreur asymptotique deviennent la mesure caractérisant la fonction de fiabilité [39]. De plus, il se trouve que selon le tirage de paramètres de canal, il peut ne pas y avoir une fonction à relais unique - entre les stratégies coopératives les plus connues - qui minimise la probabilité d'erreur. Pourtant, puisque CSI n'est pas disponible à tous les noeuds la fonction de relais devrait être faite indépendamment de sa mesure de canal et cela deviendra la limite dans la performance de code. Nous présentons une stratégie de codage originale où le relais peut choisir, basé sur sa mesure de canal θ_r, la stratégie de codage adéquate.

À ce but, les débits atteignables sont d'abord tirés pour le réseau à deux relais avec la stratégie de codage mélangée. Cette région améliore la région atteignable pour les réseaux à deux relais avec les stratégies mélangées dans [14]. En fait, il est montré que le même code pour ce réseau à deux relais marche aussi pour le canal à relais composite où on permet que le relais choisisse DF ou CF. Ici la source envoie l'information sans tenir compte de la fonction à relais. Plus spécialement, nous montrons que le récent schéma CF de [40] peut simultanément être utilisé avec le schéma DF. En outre, seulement CSI du canal de source-à-relais est nécessaire pour décider - à relais - de la fonction de relais adéquate. Donc le relais n'a pas besoin de savoir CSI complète pour décider de la stratégie efficace. Cette idée peut être prolongée aux réseaux composites généraux avec les multiple relais. À ce but, un codage semblable devrait être développée tel qu'il peut être choisi si un relais dans le réseau utilise DF ou CF. La région atteignable est présentée laquelle généralise NNC au cas de stratégie de codage mélangée, c'est-à-dire tant DF que CF. Il est aussi montré que les relais DF peuvent exploiter l'aide des relais CF en utilisant codage de offset dans les relais.

3.1 Définition de Problème

Le canal à relais composite se compose d'un ensemble de canaux à relais $\left\{P_{Y_{1\theta}^n Z_{1\theta_r}^n | X^n X_{1\theta_r}^n}\right\}_{n=1}^{\infty}$ indexé par les vecteurs de paramètres $\theta = (\theta_d, \theta_r)$ avec $(\theta_d, \theta_r) \in \Theta$, où $\theta_r \in \Theta_r$ dénote tous les paramètres affectant la sortie de relais et $\theta_d \in \Theta_d$ sont les restes des paramètres impliqués dans la communication. Prenons $\mathbb{P}_\theta$ une mesure de probabilités collective sur $\Theta = \Theta_d \times \Theta_r$ et définissons chaque canal par PD conditionnelle $\{P_{Y_{1\theta}Z_{1\theta_r}|XX_1} : \mathcal{X}\times\mathcal{X}_1 \longmapsto \mathcal{Y}_1 \times \mathcal{Z}_1\}$. Nous supposons un canal à relais sans mémoire qui implique la décomposition suivante

$$P_{Y_{1\theta}^n Z_{1\theta_r}^n|X^n X_{1\theta_r}^n}(\underline{y}_1, \underline{z}_1|\underline{x}, \underline{x}_1) = \prod_{i=1}^{n} P_{Y_{1\theta}Z_{1\theta_r}|XX_{1\theta_r}}(y_1, z_1|x, x_1),$$

où l'entré de canal est dénotée par $\underline{x} = (x_1, \ldots, x_n) \in \mathcal{X}^n$, l'entré de relais par $\underline{x_1} = (x_{1,1}, \ldots, x_{1,n}) \in \mathcal{X}_1^n$, les observations de relais par $\underline{z}_1 = (z_{1,1}, \ldots, z_{1,n}) \in \mathcal{Z}_1^n$ et les sorties de canal par $\underline{y}_1 = (y_{1,1}, \ldots, y_{1,n}) \in \mathcal{Y}_1^n$. Les paramètres de canal affectant le relais et les sorties de destination $\theta = (\theta_r, \theta_d)$ sont tirées selon la PD collective $\mathbb{P}_\theta$ et restent fixent pendant la communication. Pourtant, le tirage spécifique de θ est supposée inconnue à la source, complètement connue à la destination et partiellement connu θ_r au relais. Remarquons que θ_r est suffisant pour savoir $P_{Z_{1\theta_r}^n|X^n X_{1\theta_r}^n}$ et alors le relais connaît [1] tous les paramètres à son propre canal.

Definition 4 (code et débit atteignable) *Un code-$\mathcal{C}(n, M_n, r)$ pour le canal à relais composite se compose de :*

- *Une fonction d'encodeur $\{\varphi : \mathcal{M}_n \longmapsto \mathcal{X}^n\}$,*
- *Une fonction de décodeur $\{\phi_\theta : \mathcal{Y}_1^n \longmapsto \mathcal{M}_n\}$,*
- *Un ensemble des fonctions de relais $\{f_{i,\theta_r} : \mathcal{Z}_1^{i-1} \longmapsto \mathcal{X}_1\}_{i=1}^n$, pour un ensemble de messages uniformément distribués $W \in \mathcal{M}_n = \{1, \ldots, M_n\}$. Notons que seulement CSI partielle au relais est supposé (désigné par θ_r) laquelle est en principe basé sur le lien source-relais.*

La probabilité d'erreur $0 \leq \epsilon < 1$ est r-atteignable, s'il y a un code-$\mathcal{C}(n, M_n, r)$ avec un débit satisfaisant

$$\liminf_{n\to\infty} \frac{1}{n} \log M_n \geq r \tag{3.1}$$

et la probabilité d'erreur moyenne

$$\limsup_{n\to\infty} \mathbb{E}_\theta \left[\Pr\left\{\phi_\theta(Y_{1\theta}^n) \neq W \middle| \theta\right\}\right] \leq \epsilon. \tag{3.2}$$

L'infimume de toutes la probabilité d'erreur r-atteignable $\bar{\epsilon}(r)$ est définis comme

$$\bar{\epsilon}(r) = \inf\{0 \leq \epsilon < 1 : \epsilon \textit{ is } r\textit{-achievable}\}. \tag{3.3}$$

1. Nous insistons qu'il n'y a aucune perte de généralité en le supposant parce que l'index θ est fixe pendant la communication et ainsi chaque destination peut tout à fait savoir ses propres paramètres de canal.

Nous insistons que pour les canaux satisfaisant la propriété de converse forte la quantité (3.3) *coïncide avec la définition ordinaire des probabilité de panne, qui tournent pour être (asymptotique) une probabilité d'erreur moyennes.*

3.2 Bornes sur la probabilité d'erreur moyenne

Dans le cadre présent, nous supposons que la source n'est pas consciente du tirage spécifique $\theta \sim \mathbb{P}_\theta$ et alors, le débit du code r et la stratégie du code - par ex. DF ou CF - doivent être choisis indépendamment du tirage de canal. En outre, tous les deux restent fixes pendant la communication sans tenir compte de la mesure de canal au relais. Nous voulons caractériser la probabilité d'erreur moyenne la plus petite possible définie par (3.2), en tant que fonction du débit du code r. Dans le travail récent [38], il a été montré que la probabilités d'erreur moyenne $\bar{\epsilon}(r)$ peut être borné comme suit

$$\mathbb{P}_\theta(\underline{r} \in \mathcal{S}_\theta) \leq \bar{\epsilon}(r) \leq \inf_{\phi} \mathbb{P}_\theta(\underline{r} \notin \mathcal{R}_\theta(\phi)), \tag{3.4}$$

où S_θ est la borne flot-max coupe-min

$$S_\theta = \min\left\{I(X; Z_{1\theta_r} Y_{1\theta} | X_{1\theta_r}), I(X X_{1\theta_r}; Y_{1\theta}\right\}$$

et $\mathcal{R}_\theta$ est un débit atteignable pour le canal à relais pour θ donné et ϕ est l'ensemble de toutes les fonctions d'encodage φ.

3.2.1 Décoder-et-Transmettre (DF) et Comprimer-et-Transmettre (CF)

Supposons d'abord que le schéma DF est choisi et ainsi la probabilité de panne deviennent une borne supérieure en la probabilité d'erreur moyenne, qui sont données par

$$P_{\text{out}}^{\text{DF}}(r) = \min_{p(x,x_1)} \mathbb{P}_\theta\big[r > \min\{I(X; Z_{1\theta_r}|X_1), I(XX_1; Y_{1\theta})\}\big]. \tag{3.5}$$

Remarquons que puisque la source ignore $\theta = (\theta_r, \theta_d)$, et $p(x, x_1)$ doit être connu tant au relais qu'à la source, alors $p(x_1)$ ne peut pas être de façon indépendante optimisé sur θ_r au relais pour minimiser la probabilité de panne.

Considérons maintenant le cas de CF, pour lequel la source n'a pas besoin de savoir $p(x_1)$. Donc le relais peut choisir $p(x_1)$ pour minimiser la probabilité de panne conditionnée par chaque θ_r. Ce processus se poursuit en deux étapes d'optimisation , et en utilisant le codage de réseau bruyant [18], la probabilité de panne de schéma CF s'écrit comme

$$P_{\text{out}}^{\text{CF}}(r) = \min_{p(x,q)} \mathbb{E}_{\theta_r}\Big[\min_{p(x_1|q)p(\hat{z}_1|x_1,z_1,q)} \mathbb{P}_{\theta|\theta_r}\big[r > \min\{I(X;\hat{Z}_1 Y_{1\theta}|X_{1\theta_r}Q), \\ I(XX_{1\theta_r};Y_{1\theta}|Q) - I(Z_{1\theta_r};\hat{Z}_1|XX_{1\theta_r}Y_{1\theta}Q)\}\big|\theta_r\big]\Big]. \tag{3.6}$$

Alors il est facile de voir que la stratégie minimisante la probabilité de panne fournira la borne supérieure la plus serrée sur la probabilité d'erreur moyenne. De (3.4), il tient pour tout le débit r que

$$\bar{\epsilon}(r) \leq \min\left\{P_{\text{out}}^{\text{DF}}(r), P_{\text{out}}^{\text{CF}}(r)\right\}. \tag{3.7}$$

La question centrale qui survient ici est de voir si la probabilité d'erreur (3.7) peut être améliorée selon quelque stratégie de codage élégante. Spécialement, le relais choisirait la meilleure stratégie instantanément selon sa mesure de canal θ_r. À ce but, le code de source devrait être capable d'être utilisé simultanément avec DF et CF. Dans la section suivante, nous prouvons d'abord un débit atteignable pour le réseau à deux relais où un noeud de relais emploie le codage de DF tandis que l'autre utilise CF et ensuite en comptant sur cette codage nous présentons la Stratégie de Codage Sélectif (SCS).

3.2.2 Stratégie de Codage Sélectif (SCS)

Considérons le canal à deux relais comme montré dans Fig. 3.1. Ce canal est caractérisé par PD conditionnel $\left\{P_{Y_1Z_1Z_2|XX_1X_2} : \mathcal{X} \times \mathcal{X}_1 \times \mathcal{X}_2 \longmapsto \mathcal{Y}_1 \times \mathcal{Z}_1 \times \mathcal{Z}_2\right\}$ pour toute entré de source $x \in \mathcal{X}$, l'entré de relais $x_1 \in \mathcal{X}_1$ et $x_2 \in \mathcal{X}_2$, la sortie de canal $y_1 \in \mathcal{Y}_1$, les sorties de relais $z_1 \in \mathcal{Z}_1$ et $z_2 \in \mathcal{Z}_2$. Le débit atteignable est défini comme d'habitude pour ce canal. Le théorème suivant fournit une borne intérieure sur la capacité de ce canal où une fonction à relais utilise le schéma DF tandis que l'autre exécute CF.

Theorem 13 (réseau à deux relais) *Une borne intérieure à la capacité de réseau à deux relais est donnée par tous les débits satisfaisant*

$$R \leq \max_{P\in\mathcal{P}} \min \Big\{ I(X;Z_1|X_1Q), \max\Big\{ I(XX_1;Y_1|Q), \\ \min\Big[I(XX_1;\hat{Z}_2Y_1|X_2Q), I(XX_1X_2;Y_1|Q) - I(Z_2;\hat{Z}_2|Y_1XX_1X_2Q)\Big]\Big\}\Big\} \tag{3.8}$$

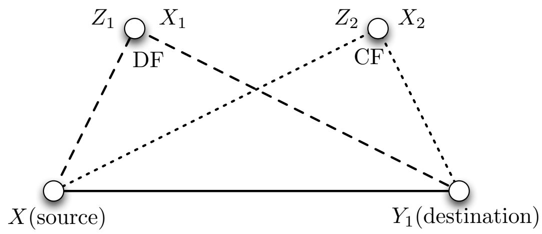


Figure 3.1 – Réseau à deux relais

et l'ensemble de toutes PDs admissible $\mathcal{P}$ est défini comme

$$\mathcal{P} = \left\{ P_{QX_2X_1XY_1Z_1Z_2\hat{Z}_2} = P_Q P_{X_2|Q} P_{XX_1|Q} P_{Y_1Z_1Z_2|XX_1X_2Q} P_{\hat{Z}_2|X_2Z_2Q} \right\}.$$

La preuve de Théorème 13 est fondée sur le codage de superposition, les schémas DF et CF, le décodage collectif et en arrière à la destination.

Le deuxième maximum dans (3.8) détermine si le deuxième relais, qui utilise CF, augmente vraiment le débit ou il est mieux de l'ignorer. On peut voir dans la preuve que le relais 2 augmente le débit si la condition suivante tient :

$$I(X_2; Y_1|XX_1Q) \geq I(Z_2; \hat{Z}_2|Y_1XX_1X_2Q) \tag{3.9}$$

Notons que la région dans Théorème 13 améliore la région dans [14] en n'utilisant pas CF au relais quand son lien est trop bruyant et en utilisant le codage de réseau bruyant qui améliore la contrainte de CF à (3.9).

Remark 9 *Le code de source X est superposé sur le code X_1 de relais DF. Observons les deux derniers termes dans* (3.8). *Ces expressions sont la condition de décodage réussi à la destination tandis que le premier terme est la condition de décodage réussi de X_1 pour le premier relais. Si on compare les deux derniers termes avec le débit ordinaire de CF, on peut voir qu'ils ont la forme semblable avec la différence mineure que le code de source X_1 a été remplacé par le paire (X, X_1).*

Considérons ensuite un scénario pour lequel le premier relais X_1 n'est pas présent mais la source utilise toujours le codage de superposition. Dans ce cas-là, nous nous occupons d'un canal à relais simple, le schéma CF à relais et la superposition du code à la source. Suivant presque les mêmes lignes que la preuve de Théorème 13, les débit atteignables peuvent être développés pour ce cas, mais il n'y a aucun besoin pour la condition de décodage réussi de X_1 au relais 1 parce qu'un tel relais n'est pas présent. Alors nous recevrions seulement les deux derniers termes dans (3.8) comme le débit final. Cette observation signifie que la superposition du code via (X, X_1) accomplit le même débit d'habitude de CF avec le paire (X, X_1) au lieu X. Après l'application des chaînes de Markov nécessaires, il se trouve que la région finale reste intacte. Autrement dit, le code de DF peut simultanément être utilisé avec le schéma CF sans n'importe quelle perte de performance. Basé sur cette remarque, le résultat suivant pour le canal à relais composite peut être tiré (voir la définition 4). Nous ferons allusion à cela comme *la Stratégie de Codage Sélectif (SCS)* et nous verrons qu'il peut davantage améliorer la borne supérieure sur la probabilité d'erreur moyenne.

Corollary 4 (SCS avec CSI partielle à relais) *La probabilité d'erreur moyenne de réseau à relais composite avec CSI partielθ_r à relais peut être borné par*

$$\bar{\epsilon}(r) \leq \min_{p(x,x_1,q)} \inf_{\mathcal{D}_{DF}\subseteq\Theta_r} \mathbb{E}_{\theta_r}\Big\{\mathbb{P}_{\theta|\theta_r}\big[r > I_{DF}, \theta_r \in \mathcal{D}_{DF}\big|\theta_r\big] + \min_{p(x_2|q)p(\hat{z}_2|x_2,z_1,q)} \mathbb{P}_{\theta|\theta_r}\big[r > I_{CF}, \theta_r \notin \mathcal{D}_{DF}\big|\theta_r\big]\Big\} \quad (3.10)$$

où (X_1, X_2) *réfère à l'entrée correspondante de relais choisit comme suit*

$$X_{1\theta_r} = \begin{cases} X_1 & si\ \theta_r \in \mathcal{D}_{DF} \\ X_2 & si\ \theta_r \notin \mathcal{D}_{DF} \end{cases} \quad denote$$

et I_{DF}, I_{CF} *sont définis par*

$$I_{DF} = \min\big\{I(X; Z_{1\theta_r}|X_1Q), I(XX_1; Y_{1\theta}|Q)\big\}, \quad (3.11)$$

$$I_{CF} = \max\Big\{\min\big[I(X; \hat{Z}_2Y_{1\theta}|X_2Q), I(XX_2; Y_{1\theta}) - I(Z_{1\theta_r}; \hat{Z}_2|Y_{1\theta}XX_2Q)\big], I(X; Y_{1\theta})\Big\}. \quad (3.12)$$

De plus, la région de décision optimale dans (3.10) *est donnée par l'ensemble*

$$\mathcal{D}^{\star}_{DF} = \big\{\theta_r \in \Theta_r : I(X; Z_{1\theta_r}|X_1Q) > r\big\}. \quad (3.13)$$

La preuve de Corollaire 4 est une conséquence directe de la preuve de débit atteignable dans Théorème 13 et quelques subtilités. Tout d'abord, nous insistons que le même code proposé dans Théorème 13 peut être utilisé pour le cadre de canal à relais composite. Fondamentalement, le relais a deux ensemble de livre de codes X_1 et X_2 et il envoie $X_{1\theta_r} = X_1$ quand la condition $\theta_r \in \mathcal{D}_{\mathrm{DF}}$ tient et autrement il envoie $X_{1\theta_r} = X_2$. Évidemment, X_1

correspond à DF pendant que X_2 correspond CF. Donc, puisque la probabilité d'erreur peut être rendues arbitrairement petites pour chaque fonction à relais, comme montré dans Fig. 3.1, alors la source n'a pas besoin de savoir la fonction à relais spécifique exécutée. Avec cette technique, le relais peut choisir la stratégie du code selon sa mesure de canal θ_r, qui doit améliorer la probabilité d'erreur générale.

Deuxièmement, remarquons que pour le cas de CF il peut y avoir la condition supplémentaire (3.9) pour le décodage. Pourtant, puisque la destination est savant de θ et par conséquent de la stratégie du code, il peut prendre conscience si la condition (3.9) est satisfaite ou non. Dans le cas où il échoue, la destination traitera la contribution de relais comme le bruit - sans exécuter son décodage - et ensuite la condition pour le décodage erroné simple devient $\{r > I(X; Y_{1\theta})\}$.

Nous remarquons que SCS est au moins aussi bon que les schémas DF ou CF tout seul. D'autre part, il peut être montré qu'en général le meilleur choix pour $\mathcal{D}_{\mathrm{DF}}$ est la région pour laquelle le relais peut décoder le message de source. Parce que $I(XX_2; Y_{1\theta})$ est la borne de flot-max coupe-min et c'est plus grand que I_{CF}, ainsi si r est plus grand que $I(XX_2; Y_{1\theta})$ alors ce sera plus grand que I_{CF} aussi. Donc quand le décodage au relais est réussi, le schéma CF ne peut pas faire mieux que DF et, alors le choix optimal est le schéma DF. En fait, CSI complète au relais n'est pas nécessaire de se décider pour le meilleur schéma coopératif et CSI sur le lien de source-à-relais est suffisante à ce but. CSI complète quand même améliore davantage la description de codage de source que le relais envoie à la destination. Cela produit le résultat suivant qui est une extension de Corollaire 5.

Corollary 5 (SCS avec CSI complète à relais) *La probabilité d'erreur moyenne de réseau à relais composite avec CSI complète $\theta = (\theta_r, \theta_d)$ au relais peut être borné par*

$$\bar{\epsilon}(r) \leq \min_{p(x,x_1,q)} \inf_{\mathcal{D}_{DF} \subseteq \Theta_r} \left\{ \mathbb{P}_\theta\left[r > I_{DF}, \theta_r \in \mathcal{D}_{DF}\right] + \mathbb{P}_\theta\left[r > I_{CF}, \theta_r \notin \mathcal{D}_{DF}\right] \right\} \tag{3.14}$$

où (X_1, X_2) réfère à l'entrée correspondant de relais choisit comme suit

$$X_{1\theta_r} = \begin{cases} X_1 & si\ \theta_r \in \mathcal{D}_{DF} \\ X_2 & si\ \theta_r \notin \mathcal{D}_{DF} \end{cases}$$

et I_{DF}, I_{CF} sont définis comme

$$I_{DF} = \min\big\{I(X;Z_{1\theta_r}|X_1Q), I(XX_1;Y_{1\theta}|Q)\big\}, \quad (3.15)$$

$$I_{CF} = \max_{p(x_2|q)p(\hat{z}_2|x_2,z_1,q)} \min\big\{I(X;\hat{Z}_2Y_{1\theta}|X_2Q), I(XX_2;Y_{1\theta}|Q) - I(Z_{1\theta_r};\hat{Z}_2|Y_{1\theta}XX_2Q)\big\}. \quad (3.16)$$

De plus, la région de décision optimale dans (3.14) *est donnée par l'ensemble*

$$\mathcal{D}^{\star}_{DF} = \big\{\theta_r \in \Theta_r : \, I(X;Z_{1\theta_r}|X_1Q) > r\big\}. \quad (3.17)$$

La preuve de Corollaire 5 découle facilement des mêmes lignes que Corollaire 4. Quand même, il vaudrait pour dire en passant ici que puisque CSI complète est disponible au relais, l'entré de relais peut être optimisée sur $\theta = (\theta_r, \theta_d)$ et ensuite I_{CF} ne peut jamais être moins que la capacité du lien direct de source-à-destination.

3.3 Réseaux à Multiple Relais Composite

Dans cette section, d'abord une région atteignable est présentée pour les réseaux à multiple relais avec une destination et une source simples où une partie des relais utilise DF et les restes utilisent CF. Le résultat est utilisé pour prouver la stratégie de codage sélective pour les canaux de multiple relais composites.

Le canal à multiple relais composite se compose d'un ensemble de canaux à multiple relais dénoté comme suit :

$$\big\{\mathbb{W}^n_\theta = P_{Y^n_{1,\theta}Z^n_{1,\theta_r}Z^n_{2,\theta_r}...Z^n_{N,\theta_r}|X^nX^n_{1,\theta_r}X^n_{2,\theta_r}X^n_{N,\theta_r}}\big\}_{n=1}^{\infty}.$$

Semblable au cas de canaux à relais simples, ils sont indexés par les vecteurs de paramètres $\theta = (\theta_d, \theta_r)$ avec $\theta_d, \theta_r \in \Theta$, où θ_r dénote tous les paramètres affectant la sortie des relais et θ_d sont les autres paramètres impliqués dans la communication. Supposons $\mathbb{P}_\theta$ d'être une mesure de probabilité sur Θ et définissons chaque canal par PD conditionnel

$$\big\{P_{Y_{1,\theta}Z_{1,\theta_r}...Z_{N,\theta_r}|XX_{1,\theta_r}...X_{N,\theta_r}} : \mathcal{X} \times \mathcal{X}_1 \times ... \times \mathcal{X}_N \longmapsto \mathcal{Y}_1 \times \mathcal{Z}_1 \times ... \times \mathcal{Z}_N\big\}.$$

Nous supposons de nouveau un canal à multiple relais sans mémoire.

Les paramètres de canal affectant le relais et les sorties de destination $\theta = (\theta_r, \theta_d)$ sont tirés selon la PD collectif $\mathbb{P}_\theta$ et restent fixe pendant la communication. Pourtant, le tirage spécifique de θ est supposée pour être inconnue à la source, complètement connue à la destination et partiellement connu (θ_r) au relais.

Definition 5 (code et débit atteignable) *Un code-$\mathcal{C}(n, M_n, r)$ pour les réseaux à multiple relais composite se compose de :*

- *Une fonction d'encodeur* $\{\varphi : \mathcal{M}_n \longmapsto \mathcal{X}^n\}$,
- *Une fonction de décodeur* $\{\phi_\theta : \mathcal{Y}_1^n \longmapsto \mathcal{M}_n\}$,
- *Un ensemble de fonctions de relais* $\left\{f_{i,\theta_r}^{(k)} : \mathcal{Z}_k^{i-1} \longmapsto \mathcal{X}_k\right\}_{i=1}^{n}$ *for* $k \in \mathcal{N}$. *Seulement CSI partielle dans le relais sont supposé (désigné par* θ_r*) laquelle est basé sur le lien source-relais.*

La probabilité d'erreur $0 \leq \epsilon < 1$ *est dite r-atteignable, si il y a un code-*$\mathcal{C}(n, M_n, r)$ *avec un débit satisfaisant*

$$\liminf_{n\to\infty} \frac{1}{n} \log M_n \geq r \tag{3.18}$$

et la probabilité d'erreur moyenne

$$\limsup_{n\to\infty} \mathbb{E}_\theta \left[\Pr\left\{\phi_\theta(Y_{1\theta}^n) \neq W \middle| \theta\right\}\right] \leq \epsilon. \tag{3.19}$$

L'infimum de toutes probabilités d'erreur r-atteignable $\bar{\epsilon}(r)$ *est défini comme*

$$\bar{\epsilon}(r) = \inf\left\{0 \leq \epsilon < 1 \,:\, \epsilon \textit{ is } r\textit{-achievable}\right\}. \tag{3.20}$$

Nous insistons que pour les canaux satisfaisant la propriété de converse forte[2] *, (3.20) coïncide avec la définition des probabilité de panne, qui sont (asymptotiquement) une probabilité d'erreur moyenne.*

Supposons que $\mathcal{N} = \{1, ..., N\}$ et pour chaque $\mathcal{S} \subseteq \mathcal{N}$, $X_{\mathcal{S}} = \{X_i, i \in \mathcal{S}\}$. Dans une manière similaire, $\bar{\epsilon}(r)$ peut être borné par :

$$\mathbb{P}_\theta(\underline{r} \in \mathcal{S}_\theta) \leq \bar{\epsilon}(r) \leq \inf_{\phi} \mathbb{P}_\theta(\underline{r} \notin \mathcal{R}_\theta(\phi)), \tag{3.21}$$

où $\mathcal{R}_\theta$ est un débit atteignable pour le réseau unicast pour θ donné, et $\mathcal{S}_\theta$ est l'infimume de tous les débits tel que chaque code avec tel débit produit la probabilité d'erreur tendant vers un, et ϕ est l'ensemble de toutes les fonctions d'encodage φ. On peut montrer que S_θ peut être remplacé par la borne flot-max coupe-min.

2. Strong Converse Property

3.3.1 Réseaux à Multiple Relais

Avant borner l'erreur espérée pour le modèle composite, nous regardons le canal à multiple relais caractérisé par $\{P_{Y_1^n Z_1^n Z_2^n \dots Z_N^n | X^n X_1^n X_2^n \dots X_N^n}\}_{n=1}^{\infty}$. Ce canal n'est pas composite et donc connu à chaque utilisateur. Le théorème suivant présente une borne intérieure sur la capacité de canaux à multiple relais qui généralise la borne précédente de NNC sur les réseaux à multiple relais.

Theorem 14 (Codage de Réseaux Bruyant Mixé) *Pour le réseaux à multiple relais, le débit suivant est atteignable :*

$$R \leq \max_{P \in \mathcal{P}} \max_{\mathcal{V} \subseteq \mathcal{N}, \mathcal{T} \in \Upsilon(\mathcal{V})} \min \left\{ \min_{\mathcal{S} \subseteq \mathcal{T}} R_{\mathcal{T}}(\mathcal{S}), \min_{i \in \mathcal{V}^c} I(X; Z_i | X_{\mathcal{V}^c} Q) \right\} \tag{3.22}$$

où

$$R_{\mathcal{T}}(\mathcal{S}) = I(X X_{\mathcal{V}^c} X_{\mathcal{S}}; \hat{Z}_{\mathcal{S}^c} Y_1 | X_{\mathcal{S}^c} Q) - I(Z_{\mathcal{S}}; \hat{Z}_{\mathcal{S}} | X X_{\mathcal{T} \cup \mathcal{V}^c} \hat{Z}_{\mathcal{S}^c} Y_1 Q),$$

pour $\mathcal{T} \subseteq \mathcal{V} \subseteq \mathcal{N}$ et $\mathcal{V}^c = \mathcal{N} - \mathcal{V}$, $\mathcal{S}^c = \mathcal{T} - \mathcal{S}$. Comme convention prenons $\min_{\emptyset} = +\infty$. *En plus $\Upsilon(\mathcal{V})$ est défini comme :*

$$\Upsilon(\mathcal{V}) = \{\mathcal{T} \subseteq \mathcal{V} : \textit{pour tous} \mathcal{S} \subseteq \mathcal{T}, Q_{\mathcal{T}}(\mathcal{S}) \geq 0\}, \tag{3.23}$$

où $Q_{\mathcal{T}}(\mathcal{S})$ est défini comme :

$$Q_{\mathcal{T}}(\mathcal{S}) = I(X_{\mathcal{S}}; \hat{Z}_{\mathcal{S}^c} Y_1 | X X_{\mathcal{S}^c \cup \mathcal{V}^c} Q) - I(Z_{\mathcal{S}}; \hat{Z}_{\mathcal{S}} | X X_{\mathcal{T} \cup \mathcal{V}^c} \hat{Z}_{\mathcal{S}^c} Y_1 Q).$$

L'ensemble de toutes les distributions admissibles $\mathcal{P}$ est défini comme :

$$\mathcal{P} = \{P_{Q X X_{\mathcal{N}} Z_{\mathcal{N}} \hat{Z}_{\mathcal{N}} Y_1} = P_Q P_{X X_{\mathcal{V}^c} | Q} P_{Y_1 Z_{\mathcal{N}} | X X_{\mathcal{N}} Q} \prod_{j \in \mathcal{V}} P_{X_j | Q} P_{\hat{Z}_j | X_j Z_j Q}\}.$$

Remark 10 *Il peut être montré en utilisant la même technique que [41] que l'optimisation dans* (3.22) *peut être refaite sur $\mathcal{T} \subseteq \mathcal{V}$ au lieu $\mathcal{T} \in \Upsilon(\mathcal{V})$. Ainsi* (3.22) *peut être réécrit comme suit :*

$$R \leq \max_{P \in \mathcal{P}} \max_{\mathcal{T} \subseteq \mathcal{V} \subseteq \mathcal{N}} \min \left\{ \min_{\mathcal{S} \subseteq \mathcal{T}} R_{\mathcal{T}}(\mathcal{S}), \min_{i \in \mathcal{V}^c} I(X; Z_i | X_{\mathcal{V}^c} Q) \right\}. \tag{3.24}$$

Pour démontrer cela, Il est suffisant de prouver que (3.24) *est inclus dans* (3.22)*. Autrement dit il est suffisant de montrer que chaque $\mathcal{T} \subseteq \mathcal{V}$ dans $\Upsilon(\mathcal{V})^c$ n'affecte pas le* *maximum dans* (3.24)*. D'abord l'égalité suivante peut être vérifiée, en utilisant la même idée que [41], pour* $\mathcal{A} \subseteq \mathcal{S} \subseteq \mathcal{T}$ *:*

$$R_{\mathcal{T}}(\mathcal{S}) = R_{\mathcal{T}\cap\mathcal{A}^c}(\mathcal{S}\cap\mathcal{A}^c) + Q_{\mathcal{T}}(\mathcal{A}). \tag{3.25}$$

Maintenant pour chaque $\mathcal{T} \subseteq \mathcal{V}$ *et* $\mathcal{T} \in \Upsilon(\mathcal{V})^c$*, selon la définition, il y a* $\mathcal{A} \subseteq \mathcal{T}$ *tel que* $Q_{\mathcal{T}}(\mathcal{A}) < 0$ *. De* (3.25)*, on peut voir que pour chaque* $\mathcal{S} \subseteq \mathcal{T}\mathcal{A}^c$*,* $R_{\mathcal{T}}(\mathcal{S}\cup\mathcal{A}) < R_{\mathcal{T}\cap\mathcal{A}^c}(\mathcal{S})$*, qui signifie qu'en remplaçant* $\mathcal{T}$ *avec* $\mathcal{T}\cap\mathcal{A}^c$ *on augmente le débit final. Autrement dit pour chaque* $\mathcal{T} \subseteq \mathcal{V}$ *et* $\mathcal{T} \in \Upsilon(\mathcal{V})^c$*, il y a* $\mathcal{T}' \subset \mathcal{T} \subseteq \mathcal{V}$*, pas nécessairement en* $\Upsilon(\mathcal{V})^c$ *tel que la région en ce qui concerne* $\mathcal{T}'$ *est augmentée ; ce qui finit la preuve.*

La conséquence de cette observation est que pour chaque $\mathcal{T}$ *et* $\mathcal{A} \subseteq \mathcal{T}$*, tel que* $Q_{\mathcal{T}}(\mathcal{A}) < 0$ *il est suffisant d'ignorer ces relais dans* $\mathcal{A}$ *et ne pas utiliser leur compression. La région* (3.24) *est plus facile d'être traitée particulièrement dans le cadre composite.*

Dans le théorème précédent en choisissant $\mathcal{V} = \mathcal{N}$, le théorème réduit à la région SNNC comme dans [41,42] qui est équivalent à la région NNC [18]. Donc le théorème 14 généralise et inclut le schéma NNC précédent et il fournit une région potentiellement plus grande. Par exemple pour le canal à relais dégradé simple il accomplit la capacité qui n'est pas le cas pour NNC. En fait les relais sont divisés en deux groupes. Les premiers groupes dans $\mathcal{V}^c$ utilisent DF et ceux dans $\mathcal{V}$ utilise CF. Pourtant un ensemble $\mathcal{T}$ de relais dans $\mathcal{V}^c$ peut être utile et augmenter le débit seulement s'ils satisfont conjointement (3.23). Sinon il est mieux de les considérer comme bruit.

La preuve est en général inspirée par la preuve du théorème 13 dans le sens qu'au lieu X_1, X_2, nous avons $X_{\mathcal{V}^c}, X_{\mathcal{V}}$.

Dans le théorème précédent, il n'y a aucune coopération entre les relais de DF et CF. Plus particulièrement les relais qui utilisent DF, ceux dans $\mathcal{V}^c$ décodent le message de source seul et sans n'importe quelle aide d'autres relais comme on peut voir dans la région. Pourtant il est possible que les relais dans $\mathcal{V}^c$ utilisent l'aide d'entre ceux dans $\mathcal{V}$ en décodant les indices de la compression transmise. Cela signifie que chaque relais dans $\mathcal{V}^c$ agit comme une destination potentielle et utilise un schéma NNC semblable pour décoder le message de source. Le théorème suivant prouve ce résultat pour ce réseau.

Theorem 15 (Codage de Réseau Bruyant Mélangé Coopérative) *Pour les réseaux à multiple relais , le débit suivant est atteignable :*

$$R \leq \max_{P\in\mathcal{P}} \max_{\mathcal{V}\subseteq\mathcal{N}} \min\left(\max_{\mathcal{T}\in\Upsilon(\mathcal{V})} \min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S}), \min_{k\in\mathcal{V}^c} \max_{\mathcal{T}_k\in\Upsilon_k(\mathcal{V})} \min_{\mathcal{S}\subseteq\mathcal{T}_k} R^{(k)}_{\mathcal{T}_k}(\mathcal{S}) \right). \tag{3.26}$$

où

$$R_{\mathcal{T}}(\mathcal{S}) = I(XX_{\mathcal{V}^c}X_{\mathcal{S}};\hat{Z}_{\mathcal{S}^c}Y_1|X_{\mathcal{S}^c}Q) - I(Z_{\mathcal{S}};\hat{Z}_{\mathcal{S}}|XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{S}^c}Y_1Q),$$

avec $(\mathcal{S}^c = \mathcal{T} - \mathcal{S})$*, et*

$$R^{(k)}_{\mathcal{T}_k}(\mathcal{S}) = I(X;\hat{Z}_{\mathcal{T}_k}Z_k|X_{\mathcal{V}^c}X_{\mathcal{T}_k}Q) + I(X_{\mathcal{S}};Z_k|X_{\mathcal{V}^c\cup\mathcal{S}^c}Q) - I(\hat{Z}_{\mathcal{S}};Z_{\mathcal{S}}|X_{\mathcal{V}^c\cup\mathcal{T}_k}\hat{Z}_{\mathcal{S}^c}Z_kQ)$$

avec $(\mathcal{S}^c = \mathcal{T}_k - \mathcal{S})$*, pour* $\mathcal{T},\mathcal{T}_k \subseteq \mathcal{V} \subseteq \mathcal{N}$ *et* $\mathcal{V}^c = \mathcal{N} - \mathcal{V}$*. En plus* $\Upsilon(\mathcal{V})$ *et* $\Upsilon_k(\mathcal{V})$ *sont définis comme suit :*

$$\begin{aligned}\Upsilon(\mathcal{V}) &= \{\mathcal{T} \subseteq \mathcal{V} : \text{ pour tous } \mathcal{S} \subseteq \mathcal{T}, Q_{\mathcal{T}}(\mathcal{S}) \geq 0\},\\ \Upsilon_k(\mathcal{V}) &= \{\mathcal{T} \subseteq \mathcal{V} : \text{ pour tous}\mathcal{S} \subseteq \mathcal{T}, Q^{(k)}_{\mathcal{T}}(\mathcal{S}) \geq 0\}\end{aligned} \tag{3.27}$$

où $Q_{\mathcal{T}}(\mathcal{S})$ *et* $Q^{(k)}_{\mathcal{T}}(\mathcal{S})$ *sont définis comme suit :*

$$\begin{aligned}Q_{\mathcal{T}}(\mathcal{S}) &= I(X_{\mathcal{S}};\hat{Z}_{\mathcal{S}^c}Y_1|XX_{\mathcal{S}^c\cup\mathcal{V}^c}Q) - I(Z_{\mathcal{S}};\hat{Z}_{\mathcal{S}}|XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{S}^c}Y_1Q),\\ Q^{(k)}_{\mathcal{T}}(\mathcal{S}) &= I(X_{\mathcal{S}};Z_k|X_{\mathcal{V}^c\cup\mathcal{S}^c}Q) - I(\hat{Z}_{\mathcal{S}};Z_{\mathcal{S}}|XX_{\mathcal{V}^c\cup\mathcal{T}}\hat{Z}_{\mathcal{S}^c}Z_kQ).\end{aligned}$$

L'ensemble de toutes les distributions admissibles $\mathcal{P}$ *sont définis comme auparavant.*

La seule différence en ce qui concerne le théorème 14 est que les relais DF utilisent le codage de réseau bruyant pour décoder le message de source. À ce but les relais de CF n'utilisent pas la technique de binning et les relais de DF utilise le décodage simultané de message et index de compression en utilisant le décodage en avant [3]. Pour le reste, nous faisons quelques observations sur le théorème précédent.

Notons que $R^{(k)}_{\mathcal{T}_k}(\mathcal{S})$ est strictement moins que la situation où NNC est utilisé dans le relais en utilisant le décodage en arrière. La raison comme discuté dans [42] est que l'augmentation dans NNC est accomplie en retardant le décodage. Dans les relais qui utilisent DF, retardement du décodage n'est pas possible et c'est la raison derrière la perte dans le débit.

De plus après l'argument semblable dans la remarque 10, l'optimisation dans (3.26) peut être faite sur les sous-ensembles de $\mathcal{V}$ au lieu $\Upsilon(\mathcal{V})$. Cela donnera la région suivante :

$$R \leq \max_{P\in\mathcal{P}} \max_{\mathcal{V}\subseteq\mathcal{N}} \min\left(\max_{\mathcal{T}\subseteq\mathcal{V}} \min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S}), \min_{k\in\mathcal{V}^c} \max_{\mathcal{T}_k\in\Upsilon_k(\mathcal{V})} \min_{\mathcal{S}\subseteq\mathcal{T}_k} R^{(k)}_{\mathcal{T}_k}(\mathcal{S})\right). \tag{3.28}$$

3. Forward Decoding

Prenons finalement un réseau où tous les relais utilisent CF. Pourtant la destination utilise le décodage en avant au lieu du décodage en arrière qui est la même méthode de décodage que la méthode de décodage des relais DF dans le théorème 15. Comme conséquence du théorème 15, nous pouvons obtenir une région atteignable avec le décodage en avant pour ce réseau qui est le corollaire suivant.

Corollary 6 *La région suivante présente une borne intérieure sur la capacité des réseaux à multiple relais :*

$$R \leq \max_{P \in \mathcal{P}} \max_{\mathcal{T} \subseteq \Upsilon(\mathcal{N})} \min_{\mathcal{S} \subseteq \mathcal{T}} R_{\mathcal{T}}^{FD}(\mathcal{S}), \tag{3.29}$$

où

$$R_{\mathcal{T}}^{FD}(\mathcal{S}) = I(X; \hat{Z}_{\mathcal{T}} Y_1 | X_{\mathcal{T}} Q) + I(X_{\mathcal{S}}; Y_1 | X_{\mathcal{S}^c} Q) - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | X_{\mathcal{T}} \hat{Z}_{\mathcal{S}^c} Y_1 Q) \quad (\mathcal{S}^c = \mathcal{T} - \mathcal{S}).$$

En plus $\Upsilon(\mathcal{N})$ est défini comme suit :

$$\Upsilon(\mathcal{N}) = \{\mathcal{T} \subseteq \mathcal{N} : \text{ pour tous} \mathcal{S} \subseteq \mathcal{T}, Q_{\mathcal{T}}(\mathcal{S}) \geq 0\}, \tag{3.30}$$

où $Q_{\mathcal{T}}(\mathcal{S})$ est défini comme :

$$Q_{\mathcal{T}}(\mathcal{S}) = I(X_{\mathcal{S}}; Y_1 | X_{\mathcal{S}^c} Q) - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | X X_{\mathcal{T}} \hat{Z}_{\mathcal{S}^c} Y_1 Q).$$

La région précédente n'est pas aussi bonne que la Codage de Réseau Bruyante, à cause du décodage en avant mais pourtant c'est potentiellement mieux que l'utilisation de binning et d'autres techniques de décodage en avant [42]. Particulièrement la condition qui détermine la région d'optimisation dans [42], c'est-à-dire $Q_{\mathcal{T}}^*(\mathcal{S}) \geq 0$ est définie comme suit :

$$Q_{\mathcal{T}}^*(\mathcal{S}) = I(X_{\mathcal{S}}; Y_1 | X_{\mathcal{S}^c} Q) - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | X_{\mathcal{T}} \hat{Z}_{\mathcal{S}^c} Y_1 Q).$$

Cela créera une plus petite région d'optimisation que $\Upsilon(\mathcal{N})$ parce que :

$$I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | X_{\mathcal{T}} \hat{Z}_{\mathcal{S}^c} Y_1 Q) \geq I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | X X_{\mathcal{T}} \hat{Z}_{\mathcal{S}^c} Y_1 Q).$$

Donc le décodage en avant et simultané sans binning agit potentiellement mieux que le décodage en avant simultané avec binning.

3.4 Réseaux à Multiples Relais Composite

Après trouver les débit atteignables pour les réseaux à multiple relais, nous bougeons aux réseaux composites. Supposons que les paramètres de canal affectant des relais et des sorties de destination sont $\theta = (\theta_r, \theta_d)$. Les relais sont indexés par $\mathcal{N} = \{1, ..., N\}$.

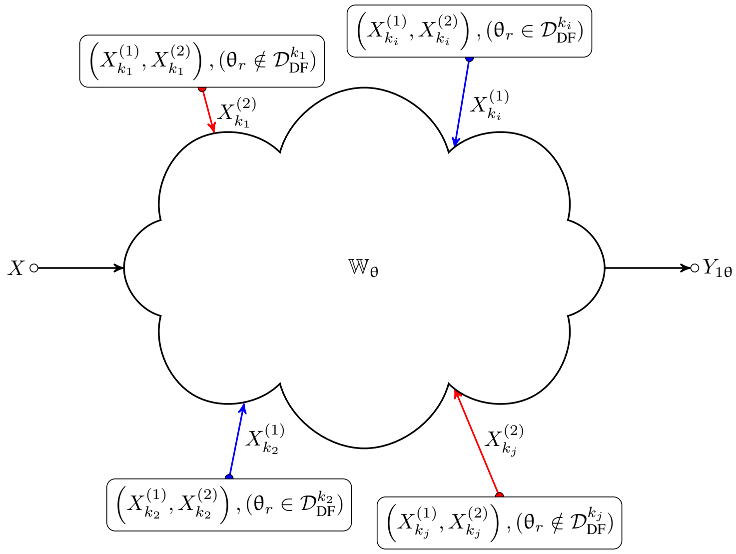


Figure 3.2 – Réseau Composite à Multiple Relais

3.4.1 Stratégies de Codage Non-Sélective

L'option commune est que chaque relais fixe sa stratégie de codage (Décoder-et-Transmettre ou Comprimer-et-Transmettre) sans tenir compte de θ. Autrement dit $\mathcal{V}$ est choisi au préalable. Cela aboutit à la borne suivante sur la probabilité d'erreur espérée pour les canaux à multiple relais composites avec CSI partielθ_r aux relais :

$$\bar{\epsilon}(r) \leq \inf_{\mathcal{V}\subseteq\mathcal{N}} \min_{p(x,x_{\mathcal{V}^c},q)} \mathbb{E}_{\theta_r}\left\{ \min_{\prod_{j\in\mathcal{V}} p(x_j|q)p(\hat{z}_j|x_j z_j q)} \mathbb{P}_{\theta|\theta_r}\left[r > I_{\text{CMNNC}}(\mathcal{V})\middle|\theta_r\right]\right\} \tag{3.31}$$

et $I_{\text{MNNC}}(\mathcal{V})$ est définis comme suit ($\theta = (\theta_r, \theta_d)$) :

$$I_{\text{CMNNC}}(\mathcal{V}) = \min\left(\max_{\mathcal{T}\subseteq\mathcal{V}}\min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S},\theta), \min_{k\in\mathcal{V}^c}\max_{\mathcal{T}_k\in\Upsilon_k(\mathcal{V})}\min_{\mathcal{S}\subseteq\mathcal{T}_k} R_{\mathcal{T}_k}^{(k)}(\mathcal{S},\theta)\right). \tag{3.32}$$

avec

$$\begin{aligned} R_{\mathcal{T}}(\mathcal{S},\theta) &= I(XX_{\mathcal{V}^c}X_{\mathcal{S}};\hat{Z}_{\mathcal{S}^c}Y_{1\theta}|X_{\mathcal{S}^c}Q) - I(Z_{\mathcal{S}\theta_r};\hat{Z}_{\mathcal{S}}|XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{S}^c}Y_{1\theta}Q) \quad (\mathcal{S}^c = \mathcal{T}-\mathcal{S}),\\ R_{\mathcal{T}_k}^{(k)}(\mathcal{S},\theta) &= I(X;\hat{Z}_{\mathcal{T}_k}Z_{k\theta_r}|X_{\mathcal{V}^c}X_{\mathcal{T}_k}Q) + I(X_{\mathcal{S}};Z_{k\theta_r}|X_{\mathcal{V}^c\cup\mathcal{S}^c}Q)\\ &\quad - I(\hat{Z}_{\mathcal{S}};Z_{\mathcal{S}\theta_r}|X_{\mathcal{V}^c\cup\mathcal{T}_k}\hat{Z}_{\mathcal{S}^c}Z_{k\theta_r}Q) \qquad (\mathcal{S}^c = \mathcal{T}_k-\mathcal{S}). \end{aligned}$$

pour $\mathcal{T},\mathcal{T}_k \subseteq \mathcal{V} \subseteq \mathcal{N}$ et $\mathcal{V}^c = \mathcal{N}-\mathcal{V}$. Similairement $\Upsilon_k(\mathcal{V})$ est définis comme suit :

$$\Upsilon_k(\mathcal{V}) = \{\mathcal{T}\subseteq\mathcal{V}:\ \text{pour tous } \mathcal{S}\subseteq\mathcal{T}, Q_{\mathcal{T}}^{(k)}(\mathcal{S},\theta_r) \geq 0\} \tag{3.33}$$

où $Q_{\mathcal{T}}^{(k)}(\mathcal{S},\theta_r)$ est définis comme suit :

$$Q_{\mathcal{T}}^{(k)}(\mathcal{S},\theta_r) = I(X_{\mathcal{S}}; Z_{k\theta_r}|X_{\mathcal{V}^c\cup\mathcal{S}^c}Q) - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}\theta_r}|XX_{\mathcal{V}^c\cup\mathcal{T}}\hat{Z}_{\mathcal{S}^c}Z_{k\theta_r}Q).$$

On peut voir que seulement les relais de CF peuvent adapter la distribution de probabilité de la version comprimée de la sortie du canal et les autres relais doivent fixer tant leur codage que sa distribution sans tenir compte de CSI disponible pour eux. Comme nous avons vu en cas de canaux à deux relais un codage peut être développée qui rend possible pour les relais de choisir leur codage selon CSI disponible.

3.4.2 Stratégie de Codage Sélectif

Semblable au cas de canaux à deux relais, chaque relais a beaucoup de livre de codes, un pour le cas qu'il utilise Décoder-et-Transmettre et les autres pour le cas de Comprimer-et-Transmettre. Chaque relais selon son CSI (θ_r) échange entre son code pour DF et CF. Nous supposons que θ_r est connu à tous les relais. Chaque relais k a une région de décision $\mathcal{D}_{\text{DF}}^{(k)}$ tel que pour tous $\theta_r \in \mathcal{D}_{\text{DF}}^{(k)}$, le relais k utilise Décoder-et-Transmettre et autrement il utilise Comprimer-et-Transmettre. Maintenant pour chaque $\mathcal{V} \subseteq \mathcal{N}$ définissons $\mathcal{D}_{\mathcal{V}}$ comme suit :

$$\mathcal{D}_{\mathcal{V}} = \left(\bigcap_{k\in\mathcal{V}^c} \mathcal{D}_{\text{DF}}^{(k)}\right) \bigcap \left(\bigcap_{k\in\mathcal{V}} \mathcal{D}_{\text{DF}}^{(k)\,c}\right)$$

Si $\theta_r \in \mathcal{D}_{\mathcal{V}}$, donc $\theta_r \notin \mathcal{D}_{\text{DF}}^{(k)}$ pour tous $k \in \mathcal{V}$ et $\theta_r \in \mathcal{D}_{\text{DF}}^{(k)}$ pour tous $k \notin \mathcal{V}$. Ainsi le relais k, pour chaque $k \in \mathcal{V}$ utilise CF et le relais k' pour $k' \in \mathcal{V}^c$ utilise DF. L'ensemble de région de la décision de relais fournira ainsi les régions $\mathcal{D}_{\mathcal{V}}$ qui sont mutuellement disjoints et forment une partition sur l'ensemble Θ_r. Maintenant si $\theta_r \in \mathcal{D}_{\mathcal{V}}$, nous avons un réseau à multiple relais où les relais dans $\mathcal{V}$ utilisent CF. Le débit atteignable conforme à ce cas est connu du théorème 14.

Regardons Fig. 3.2. Chaque relais a deux livre de codes principaux, $X_{(k)}^{(1)}$ et $X_{(k)}^{(2)}$. Le premier code ($X_{(k)}^{(1)}$) est transmis quand $\theta_r \in \mathcal{D}_{\text{DF}}^{(k)}$. Ce code est fondé sur la stratégie DF donc le relais k décode le message de source et le transmet à la destination. Pourtant la source, ne sachant pas si le relais k envoie $X_{(k)}^{(1)}$ ou pas, utilise le codage de superposition et super-impose son code sur $X_{(k)}^{(1)}$. Si le relais k envoie $X_{(k)}^{(1)}$ alors cela deviendra la stratégie de DF. Pourtant dans ce cas-là la source doit choisir la distribution mutuelle entre le relais k et lui-même a priori sans savoir θ_r. Donc la corrélation optimale entre la source et le relais ne peut pas être trouvée dans une façon dynamique basé sur θ_r.

D'autre part si $\theta_r \notin \mathcal{D}_{\mathrm{DF}}^{(k)}$, la stratégie de CF est utilisée. Notons qu'à la différence de DF, le code qui est utilisé pour CF, $X_{(k)}^{(2)}$, est indépendant du code source et donc sa distribution de probabilités peut être choisie dans une façon dynamique basée sur θ_r. Le choix optimal pour $\mathcal{D}_{\mathcal{V}}$ donnera potentiellement la meilleure probabilité de panne que le cas où chaque relais utilise une codage fixe pour tous θ_r. Cette idée résulte au théorème suivant.

Theorem 16 (SCS avec CSI partielle) *La probabilité d'erreur moyenne de réseaux à multiple relais composite avec CSI partielθ_r dans les relais peut être borné par*

$$\bar{\epsilon}(r) \leq \min_{p(x,x_{\mathcal{N}}^{(1)},q)} \inf_{\{\mathcal{D}_{\mathcal{V}},\mathcal{V}\subseteq\mathcal{N}\}\in\Pi(\Theta_r,N)} \sum_{\mathcal{V}\subseteq\mathcal{N}} \mathbb{E}_{\theta_r}\Big\{ \min_{\prod_{j\in\mathcal{V}} p(x_j^{(2)}|q)p(\hat{z}_j|x_j^{(2)}z_jq)} \mathbb{P}_{\theta|\theta_r}\big[r > I_{MNNC}(\mathcal{V}), \theta_r \in \mathcal{D}_{\mathcal{V}}\big|\theta_r\big]\Big\} \tag{3.34}$$

$\Pi\left(\Theta_r, N\right)$ est l'ensemble de toute partition de Θ_r à au maximum 2^N ensembles disjonctifs. $I_{MNNC}(\mathcal{V})$ est défini comme $\theta = (\theta_r, \theta_d)$:

$$I_{MNNC}(\mathcal{V}) = \max_{\mathcal{T}\subseteq\mathcal{V}} \min\Big\{ \min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S},\theta), \min_{i\in\mathcal{V}^c} I(X;Z_{i\theta_r}|X_{\mathcal{N}}^{(1)}Q)\Big\} \tag{3.35}$$

avec

$$R_{\mathcal{T}}(\mathcal{S},\theta) = I(XX_{\mathcal{V}^c}^{(1)}X_{\mathcal{S}}^{(2)};\hat{Z}_{\mathcal{S}^c}Y_{1\theta}|X_{\mathcal{S}^c}^{(2)}Q) - I(Z_{\mathcal{S}\theta_r};\hat{Z}_{\mathcal{S}}|XX_{\mathcal{T}}^{(2)}X_{\mathcal{V}^c}^{(1)}\hat{Z}_{\mathcal{S}^c}Y_{1\theta}Q),$$

où $(X_k^{(1)}, X_k^{(2)})$ réfère à l'entrée correspondante de relais choisit comme suit

$$X_{k\theta_r} = \begin{cases} X_k^{(1)} & \text{if } \theta_r \in \mathcal{D}_{DF}^k \\ X_k^{(2)} & \text{if } \theta_r \notin \mathcal{D}_{DF}^k \end{cases} \quad (\mathcal{D}_{DF}^k = \bigcup_{\mathcal{V}\subset\mathcal{N},k\notin\mathcal{V}} \mathcal{D}_{\mathcal{V}})$$

Autrement dit pour $\theta_r \in \mathcal{D}_{\mathcal{V}}$, la chaîne de Markov suivante tient :

$$X_{\mathcal{V}}^{(1)}X_{\mathcal{V}^c}^{(2)} \ominus XX_{\mathcal{V}^c}^{(1)}X_{\mathcal{V}}^{(2)} \ominus Y_{1\theta}Z_{\mathcal{N}\theta_r}.$$

On peut voir que l'utilisation du code de superposition ne change pas $R_{\mathcal{T}}(\mathcal{S},\theta)$, mais à la différence du cas de canal à relais simple, la condition de décodage correct aux relais DF est changée de $I(X;Z_{i\theta_r}|X_{\mathcal{V}^c}^{(1)}Q)$ à $I(X;Z_{i\theta_r}|X_{\mathcal{N}}^{(1)}Q)$.

Le corollaire suivant présente le cas CSI complète.

Corollary 7 (SCS avec CSI complète) *de La probabilité d'erreur moyenne de réseaux à multiple relais composite avec CSI complète dans les relais peut être borné par*

$$\bar{\epsilon}(r) \leq \min_{p(x,x_{\mathcal{N}}^{(1)},q)} \inf_{\{\mathcal{D}_{\mathcal{V}},\mathcal{V}\subseteq\mathcal{N}\}\in\Pi(\Theta_r,N)} \sum_{\mathcal{V}\subseteq\mathcal{N}} \mathbb{P}_{\theta}\big[r > I_{MNNC}(\mathcal{V}), \theta_r \in \mathcal{D}_{\mathcal{V}}\big] \tag{3.36}$$

$\Pi\left(\Theta_r, N\right)$ est l'ensemble de tous les partitions sur Θ_r à au maximum 2^N ensembles disjonctifs. $I_{MNNC}(\mathcal{V})$ est définies comme suit $\theta = (\theta_r, \theta_d)$:

$$I_{MNNC}(\mathcal{V}) = \max_{\prod_{j\in\mathcal{V}} p(x_j^{(2)}|q)p(\hat{z}_j|x_j^{(2)}z_jq)} \max_{\mathcal{T}\subseteq\mathcal{V}} \min\left\{\min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S},\theta), \min_{i\in\mathcal{V}^c} I(X; Z_{i\theta_r}|X_{\mathcal{N}}^{(1)}Q)\right\} \tag{3.37}$$

et

$$R_{\mathcal{T}}(\mathcal{S},\theta) = I(XX_{\mathcal{V}^c}^{(1)}X_{\mathcal{S}}^{(2)}; \hat{Z}_{\mathcal{S}^c}Y_{1\theta}|X_{\mathcal{S}^c}^{(2)}Q) - I(Z_{\mathcal{S}\theta_r}; \hat{Z}_{\mathcal{S}}|XX_{\mathcal{T}}^{(2)}X_{\mathcal{N}}^{(1)}\hat{Z}_{\mathcal{S}^c}Y_{1\theta}Q),$$

où $(X_k^{(1)}, X_k^{(2)})$ réfère à l'entrée correspondante de relais choisit comme suit

$$X_{k\theta_r} = \begin{cases} X_k^{(1)} & \text{if } \theta_r \in \mathcal{D}_{DF}^k \\ X_k^{(2)} & \text{if } \theta_r \notin \mathcal{D}_{DF}^k \end{cases} \quad (\mathcal{D}_{DF}^k = \bigcup_{\mathcal{V}\subset\mathcal{N},k\notin\mathcal{V}} \mathcal{D}_{\mathcal{V}})$$

La preuve du corollaire reste le même comme le théorème 16 avec cette différence que les relais qui utilisent CF, peuvent choisir la distribution de probabilités optimale pour la compression en sachant tant θ_r que θ_d.

Le théorème 16 et le corollaire 7 peuvent être changé au cas qu'au lieu du schéma atteignable du théorème 14, le codage de réseau bruyante mélangé coopératif du théorème 15 est utilisé. Nous exposons le théorème suivant sans preuve qui sera utilisée dans la section suivante.

Theorem 17 (SCS avec CSI partielle-Relais Coopératifs) *La probabilité d'erreur moyenne de réseaux à multiple relais composites avec CSI partielθ_r dans les relais peut être borné par*

$$\bar{\epsilon}(r) \leq \min_{p(x,x_{\mathcal{N}}^{(1)},q)} \inf_{\{\mathcal{D}_{\mathcal{V}},\mathcal{V}\subseteq\mathcal{N}\}\in\Pi(\Theta_r,N)} \tag{3.38}$$

$$\sum_{\mathcal{V}\subseteq\mathcal{N}} \mathbb{E}_{\theta_r}\left\{\min_{\prod_{j\in\mathcal{V}} p(x_j^{(2)}|q)p(\hat{z}_j|x_j^{(2)}z_jq)} \mathbb{P}_{\theta|\theta_r}\big[r > I_{CMNNC}(\mathcal{V}), \theta_r \in \mathcal{D}_{\mathcal{V}}\big|\theta_r\big]\right\} \tag{3.39}$$

$\Pi(\Theta_r, N)$ est l'ensemble de toutes les partitions sur Θ_r à au maximum 2^N ensembles disjonctifs. $I_{MNNC}(\mathcal{V})$ est défini comme suit $\theta = (\theta_r, \theta_d)$:

$$I_{CMNNC}(\mathcal{V}) = \min\left(\max_{\mathcal{T}\subseteq\mathcal{V}} \min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S},\theta), \min_{k\in\mathcal{V}^c} \max_{\mathcal{T}_k\in\Upsilon_k(\mathcal{V})} \min_{\mathcal{S}\subseteq\mathcal{T}_k} R^{(k)}_{\mathcal{T}_k}(\mathcal{S},\theta) \right). \quad (3.40)$$

avec

$$R_{\mathcal{T}}(\mathcal{S},\theta) = I(XX_{\mathcal{V}^c}X_{\mathcal{S}};\hat{Z}_{\mathcal{S}^c}Y_{1\theta}|X_{\mathcal{S}^c}Q) - I(Z_{\mathcal{S}\theta_r};\hat{Z}_{\mathcal{S}}|XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{S}^c}Y_{1\theta}Q) \quad (\mathcal{S}^c = \mathcal{T}-\mathcal{S}),$$
$$R^{(k)}_{\mathcal{T}_k}(\mathcal{S},\theta) = I(X;\hat{Z}_{\mathcal{T}_k}Z_{k\theta_r}|X_{\mathcal{V}^c}X_{\mathcal{T}_k}Q) + I(X_{\mathcal{S}};Z_{k\theta_r}|X_{\mathcal{V}^c\cup\mathcal{S}^c}Q) - I(\hat{Z}_{\mathcal{S}};Z_{\mathcal{S}\theta_r}|X_{\mathcal{V}^c\cup\mathcal{T}_k}\hat{Z}_{\mathcal{S}^c}Z_{k\theta_r}Q) \quad (\mathcal{S}^c = \mathcal{T}_k-\mathcal{S}).$$

pour $\mathcal{T},\mathcal{T}_k \subseteq \mathcal{V} \subseteq \mathcal{N}$ et $\mathcal{V}^c = \mathcal{N}-\mathcal{V}$. En plus $\Upsilon_k(\mathcal{V})$ est défini comme suit :

$$\Upsilon_k(\mathcal{V}) = \{\mathcal{T}\subseteq\mathcal{V} : \text{ pour tous } \mathcal{S}\subseteq\mathcal{T}, Q^{(k)}_{\mathcal{T}}(\mathcal{S},\theta_r) \geq 0\} \quad (3.41)$$

où $Q^{(k)}_{\mathcal{T}}(\mathcal{S},\theta_r)$ est défini comme suit :

$$Q^{(k)}_{\mathcal{T}}(\mathcal{S},\theta_r) = I(X_{\mathcal{S}};Z_{k\theta_r}|X_{\mathcal{V}^c\cup\mathcal{S}^c}Q) - I(\hat{Z}_{\mathcal{S}};Z_{\mathcal{S}\theta_r}|XX_{\mathcal{V}^c\cup\mathcal{T}}\hat{Z}_{\mathcal{S}^c}Z_{k\theta_r}Q).$$

$(X_k^{(1)}, X_k^{(2)})$ réfère aux entrées correspondantes de relais choisit similaire au théorème 16.

3.5 Canaux à deux relais Gaussiens avec évanouissement

Nous bougeons à un autre exemple qui est le canal à deux relais Gaussien avec évanouissement, Fig. 3.3 défini par les relations suivantes :

$$Z_1 = h_{01}X + h_{21}X_2 + \mathcal{N}_1, \quad (3.42)$$
$$Z_2 = h_{02}X + h_{12}X_1 + \mathcal{N}_2, \quad (3.43)$$
$$Y_1 = h_{03}X + h_{13}X_1 + h_{23}X_2 + \mathcal{N}_3, \quad (3.44)$$

De la même façon $\mathcal{N}_i$'s sont les bruits additifs, VAs i.i.d. complexe circulairement symétrique Gaussien avec moyen-zéro. h_{ij}'s sont VAs complexe moyen circulairement symétrique zéro indépendant Gaussien qui font la matrice évanouissement **H**. Le pouvoir moyen de source X, X_1 et X_2 ne doit pas excéder des pouvoirs P, P_1 et P_2 respectivement. Il est supposé que la source n'est pas consciente des coefficients évanouissement et des relais savent tous les coefficients évanouissement sauf h_{i3}'s et la destination est complètement consciente de tous les coefficients évanouissement. Il y a les possibilités suivantes pour choisir la stratégie coopérative nécessaire

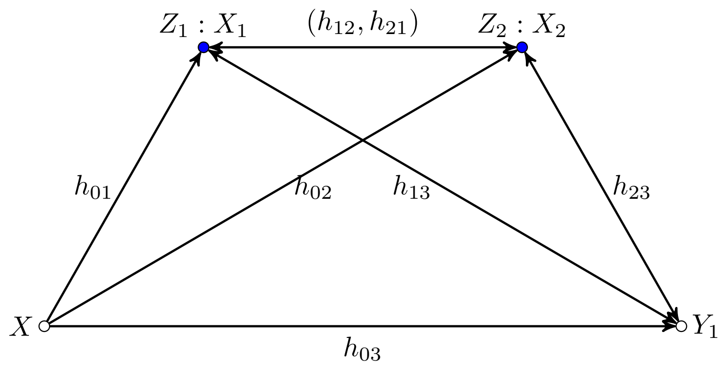


Figure 3.3 – Canal à deux relais Gaussien avec évanouissement

1. Codage de DF (DF) : les deux relais utilisent Décoder-et-Transmettre pour transmettre les informations.
2. Codage CF (CF) : les deux relais utilisent Comprimer-et-Transmettre pour transmettre les informations. Pourtant ici la destination peut considérer un ou deux relais comme le bruit pour prévenir la dégradation de performance.
3. Codage Mélangée : Un relais utilise DF et les autres utilisent CF. Pourtant il y a deux possibilités ici :
   (a) Codage Mélangée Non-coopératif (MC) : le relais de DF décode le message de source de façon indépendante et sans l'aide d'autres relais.
   (b) Codage Mélangée Coopératif (CMC) : le relais de DF exploite la version comprimée de la sortie de relais CF pour décoder le message de source.

Nous présentons d'abord les débit atteignables pour chacun de ces cas pour un triage donné de coefficients d'évanouissement en utilisant la formule générale présentée dans la section précédente. Pour le premier cas nous avons le débit suivant comme la borne intérieure de la capacité ($P_0 = P$) :

$$I_{\mathrm{DF}}(\mathbf{H}) = \min\left\{ \frac{1}{2}\log\left(1 + \frac{|h_{03}|^2 P + |h_{13}|^2 P_1 + |h_{23}|^2 P_2 + \sum_{0\leq i<j\leq 2} 2\rho_{ij}\sqrt{P_i P_j}\,\mathbb{R}e\{h_{i3}h_{j3}^{\star}\}}{N_3}\right),\right.$$
$$\left. \frac{1}{2}\log\left(1+\frac{|h_{01}|^2(1-\beta)P}{N_1}\right), \frac{1}{2}\log\left(1+\frac{|h_{02}|^2(1-\beta)P}{N_2}\right)\right\}, \quad (3.45)$$

où β est défini comme suit :

$$\beta = \frac{\rho_{01}^2 + \rho_{02}^2 - 2\rho_{01}\rho_{02}\rho_{12}}{1-\rho_{12}^2}.$$

En ce qui concerne le second cas, où tous les relais utilisent le codage de CF, le débit suivant peut être atteignable :

$$\begin{aligned}
I_{\mathrm{CF}}(\mathbf{H}) = \min\Bigg\{ & \frac{1}{2}\log\left(1 + P\left(\frac{|h_{01}|^2}{N_1+\hat{N}_1} + \frac{|h_{02}|^2}{N_2+\hat{N}_2} + \frac{|h_{03}|^2}{N_3}\right)\right), \\
& \frac{1}{2}\log\left(1+\frac{|h_{03}|^2P+|h_{13}|^2P_1+|h_{23}|^2P_2}{N_3}\right) - \frac{1}{2}\log\left(1+\frac{N_1}{\hat{N}_1}\right) - \frac{1}{2}\log\left(1+\frac{N_2}{\hat{N}_2}\right), \\
& \frac{1}{2}\log\left(1+\frac{|h_{03}|^2P+|h_{13}|^2P_1}{N_3} + \frac{|h_{02}|^2P+|h_{12}|^2P_1}{N_2+\hat{N}_2} + \frac{PP_1|h_{02}h_{13}-h_{03}h_{12}|^2}{N_3(N_2+\hat{N}_2)}\right) \\
& \qquad - \frac{1}{2}\log\left(1+\frac{N_1}{\hat{N}_1}\right), \\
& \frac{1}{2}\log\left(1+\frac{|h_{03}|^2P+|h_{23}|^2P_2}{N_3} + \frac{|h_{01}|^2P+|h_{21}|^2P_2}{N_1+\hat{N}_1} + \frac{PP_2|h_{01}h_{23}-h_{03}h_{21}|^2}{N_3(N_1+\hat{N}_1)}\right) \\
& \qquad - \frac{1}{2}\log\left(1+\frac{N_2}{\hat{N}_2}\right)\Bigg\}.
\end{aligned} \tag{3.46}$$

Notons que dans le débit précédent, la destination décode toute les indices de compression. Dénotons par $I^i_{\mathrm{CF.}}(\mathbf{H})$ le débit quand seulement la version comprimée du relais i est utilisée. Dans ce cas-là l'autre entré de relais apparaît comme le bruit pour la destination et aussi pour l'autre relais pourtant ce soi-disant bruit total au relais et à la destination est corrélé en raison de la présence de la contribution de l'autre relais. Pour ce cas le débit est comme suit :

$$\begin{aligned}
I^{(1)}_{\mathrm{CF}}(\mathbf{H}) = \min\Bigg\{ & \frac{1}{2}\log\left(1+\frac{\left(|h_{03}|^2(N_1+\hat{N}_1)+|h_{01}|^2N_3\right)P + PP_2|h_{01}h_{23}-h_{03}h_{21}|^2}{N_3(N_1+\hat{N}_1)+P_2\left(N_3|h_{21}|^2+(N_1+\hat{N}_1)|h_{23}|^2\right)}\right), \\
& \frac{1}{2}\log\left(1+\frac{|h_{03}|^2P+|h_{13}|^2P_1}{|h_{23}|^2P_2+N_3}\right) - \frac{1}{2}\log\left(1+\frac{N_1}{\hat{N}_1}+\frac{N_3|h_{21}|^2P_2}{\hat{N}_1(|h_{23}|^2P_2+N_3)}\right)\Bigg\}.
\end{aligned} \tag{3.47}$$

Quant à l'étape suivante, nous bougeons au cas de codage mélangée non-coopératif. Le débit suivant est atteignable quand X_1 utilise DF :

$$
\begin{aligned}
I_{\mathrm{MC}}(\mathbf{H}) = \min \Bigg\{ & \frac{1}{2}\log\left(1+\frac{|h_{01}|^2(1-\rho_{01}^2)P}{N_1+P_2|h_{21}|^2}\right), \\
& \frac{1}{2}\log\left(1+\frac{|h_{03}|^2P+|h_{13}|^2P_1+|h_{23}|^2P_2+2\rho_{01}\sqrt{PP_1}\mathbb{R}e\{h_{03}h_{13}^\star\}}{N_3}\right)-\frac{1}{2}\log\left(1+\frac{N_2}{\hat{N}_2}\right), \\
& \frac{1}{2}\log\Bigg(1+\frac{|h_{03}|^2P+|h_{13}|^2P_1+2\rho_{01}\sqrt{PP_1}\mathbb{R}e\{h_{03}h_{13}^\star\}}{N_3} \\
& \qquad +\frac{|h_{02}|^2P+|h_{12}|^2P_1+2\rho_{01}\sqrt{PP_1}\mathbb{R}e\{h_{02}h_{12}^\star\}}{N_2+\hat{N}_2} \\
& \qquad +\frac{PP_1(1-\rho_{01}^2)|h_{02}h_{13}-h_{03}h_{12}|^2}{N_3(N_2+\hat{N}_2)}+\frac{2\rho_{01}\sqrt{PP_1}\alpha}{N_3(N_2+\hat{N}_2)}\Bigg)\Bigg\},
\end{aligned} \tag{3.48}
$$

où α est :

$$
\alpha = (1-\mathbb{R}e\{h_{13}h_{03}^\star\})(|h_{02}|^2P+|h_{12}|^2P_1)+(1-\mathbb{R}e\{h_{12}h_{02}^\star\})(|h_{03}|^2P+|h_{13}|^2P_1).
$$

Finalement pour le codage mélangée coopératif seulement le premier terme change (par $(a)^-$ nous voulons dire la partie négative de a) :

$$
\begin{aligned}
I_{\mathrm{CMC}}(\mathbf{H}) = \min \Bigg\{ & \frac{1}{2}\log\left(1+P(1-\rho_{01}^2)\left(\frac{|h_{01}|^2}{N_1}+\frac{|h_{02}|^2}{N_2+\hat{N}_2}\right)\right)+ \\
& \left(\frac{1}{2}\log\left(1+\frac{|h_{21}|^2P_2}{N_1+|h_{01}|^2(1-\rho_{01}^2)P}\right)-\frac{1}{2}\log\left(1+\frac{N_2}{\hat{N}_2}+\frac{N_1|h_{02}|^2(1-\rho_{01}^2)P}{\hat{N}_2(|h_{01}|^2(1-\rho_{01}^2)P+N_1)}\right)\right)^-, \\
& \frac{1}{2}\log\left(1+\frac{|h_{03}|^2P+|h_{13}|^2P_1+|h_{23}|^2P_2+2\rho_{01}\sqrt{PP_1}\mathbb{R}e\{h_{03}h_{13}^\star\}}{N_3}\right)-\frac{1}{2}\log\left(1+\frac{N_2}{\hat{N}_2}\right), \\
& \frac{1}{2}\log\Bigg(1+\frac{|h_{03}|^2P+|h_{13}|^2P_1+2\rho_{01}\sqrt{PP_1}\mathbb{R}e\{h_{03}h_{13}^\star\}}{N_3} \\
& \qquad +\frac{|h_{02}|^2P+|h_{12}|^2P_1+2\rho_{01}\sqrt{PP_1}\mathbb{R}e\{h_{02}h_{12}^\star\}}{N_2+\hat{N}_2} \\
& \qquad +\frac{PP_1(1-\rho_{01}^2)|h_{02}h_{13}-h_{03}h_{12}|^2}{N_3(N_2+\hat{N}_2)}+\frac{2\rho_{01}\sqrt{PP_1}\alpha}{N_3(N_2+\hat{N}_2)}\Bigg)\Bigg\}.
\end{aligned} \tag{3.49}
$$

(3.45) à (3.49) peuvent être utilisé pour calculer les limites sur l'erreur espéré. Comme un exemple considérons le réseau à deux relais Gaussien d'évanouissement, représenté dans Fig. 3.3, qui est définie par les relations suivantes :

$$
Z_1 = \frac{h_{01}}{d^\alpha}X + h_{21}X_2 + \mathcal{N}_1, \tag{3.50}
$$

$$
Z_2 = h_{02}X + h_{12}X_1 + \mathcal{N}_2, \tag{3.51}
$$

$$
Y_1 = h_{03}X + h_{13}X_1 + h_{23}X_2 + \mathcal{N}_3. \tag{3.52}
$$

Définissons $\mathcal{N}_i$'s des bruits additifs, VAs i.i.d. complexe circulairement symétrique Gaussien avec moyen-zéro et la variance d'unité ; prenons h_{ij}'s VAs complexe moyen-zéro circulairement symétrique indépendant Gaussien de nouveau avec la variance d'unité. Mettons $\mathbf{H}$ la matrice d'évanouissement. d est l'affaiblissement de propagation [4] aléatoire. Le pouvoir moyen de la source et des entrées de relais X, X_1 et X_2 ne doit pas excéder des pouvoirs P, P_1 et P_2, respectivement. La compression est obtenue en ajoutant un bruit additif $\hat{Z}_1 = Z_1\hat{\mathcal{N}}_1$, $\hat{Z}_2 = Z_2\hat{\mathcal{N}}_2$. Il est supposé que la source n'est pas consciente des coefficients d'évanouissement, les relais savent tous les coefficients d'évanouissement sauf h_{i3}'s et la destination est complètement consciente de tout. Les pouvoirs de source et relais sont respectivement 1 et 10.

Les possibilités de choisir la stratégie coopérative propre sont comme suit. Premièrement les deux relais utilisent DF pour transmettre les informations, à savoir le cas de DF total. Deuxièmement les deux relais utilisent CF pour transmettre les informations, appelez-le CF total. Ici la destination peut considérer un ou deux relais comme le bruit pour prévenir la dégradation de performance. Alors un autre cas est quand un relais utilise DF et les autres utilisent CF, à savoir Codage Mélangée. Finalement les relais peuvent choisir leur stratégie de codage basée sur les paramètres de canal, à savoir Codage Sélectif.

La figure 3.4 présente l'analyse numérique de ces stratégies. d est choisi avec la distribution uniforme entre 0 et 0.1 qui signifie que le premier relais est toujours autour de la source. Étant donné cette hypothèse, nous supposons que le premier relais utilise DF en cas du codage mélangée et c'est l'autre relais qui utilise CF. On peut voir qu'aucune des stratégies non-sélectives comme DF total, CF total et Codage Mélangée n'est en général le meilleur sans tenir compte des coefficients d'évanouissement. Pourtant si on permet au relais de choisir leur stratégie selon les coefficients d'évanouissement, ce codage sélectif causera l'amélioration significative comparée au borne flot max et coup min relatif aux autres stratégies.

4. Path Loss

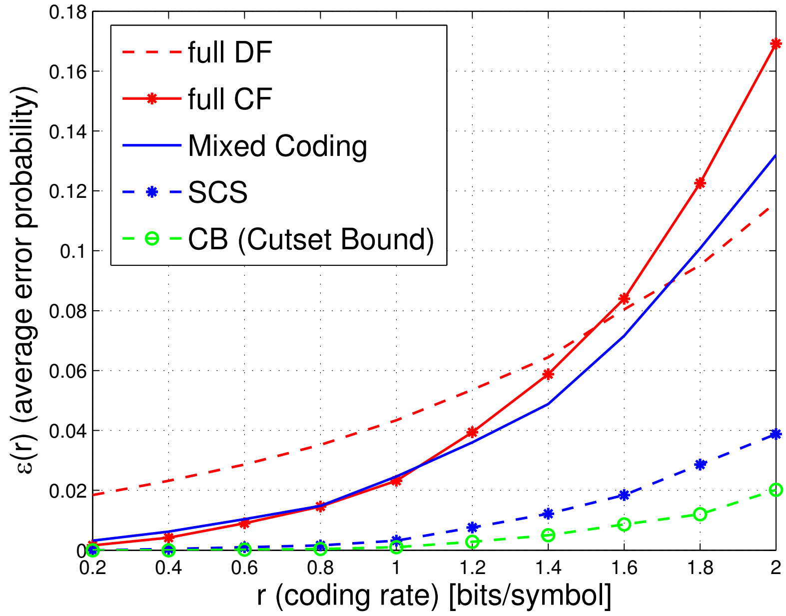


Figure 3.4 – La probabilités d'erreur asymptotique $\bar{\epsilon}(r)$ vs. débit de codage r.

Chapitre 4

Spectre Asymptotique de EP pour les réseaux composites

La plupart du temps, une probabilités d'erreur (EP) arbitrairement petite en pratique ne peuvent pas être garanties pour tous les canaux dans l'ensemble. Alors la contrainte de EP arbitrairement petite peut produire un débit atteignable nul pour quelques canaux composites. Dans ces cas la formulation conventionnelle de capacité n'est pas adéquate d'évaluer la performance du canal.

Ici au lieu de trouver les débit atteignables avec EP arbitrairement petite, nous fixons le débit r et regardons les caractéristiques d'EP pour le débit fixe r dans les réseaux composites. Le débit est fixe et ensuite nous regardons EP pour chaque tirage de canal. Dans ce cas-là EP est considéré comme une variable aléatoire qui est la fonction du paramètre de canal. La notion de probabilité de panne, en signifiant la probabilité qu'un code de débit r ne peut pas être correctement décodé, a été abondamment utilisée comme une mesure de performance pour les scénarios d'évanouissement [34]. Par exemple, il est fréquent que l'encodeur sans informations publiques envoient leurs messages en utilisant des codes de débit fixe et ensuite la probabilité de panne est utilisée pour mesurer la performance. La relation de capacité et de probabilité de panne a été discutée par Effros *et al.* pour les canaux généraux [27]. Pourtant on peut voir que la notion de probabilité de panne n' est pas assez précise pour caractériser l'EP pour les canaux ne satisfaisant pas converse forte. Les notions différentes sont présentées pour étudier EP. La notion de *spectre asymptotique d'EP* pour (r, ϵ) est défini comme la probabilité asymptotique que l'EP tombe au dessus ϵ. On montre que cette notion implique que d'autres notions disponibles utilisé pour mesurer la performance de réseaux composites.

Pourtant la ϵ-capacité n'est pas connue en général pour les réseaux généraux. On montre que le spectre asymptotique d'EP peut être borné en utilisant des régions de débit atteignables disponibles. Particulièrement le concept de *la région d'erreur complète* [1] est définie pour les réseaux généraux où chaque code avec les débit de transmission dans cette région produit EP égal à un. Pour les réseaux composites cette région est aussi une région aléatoire. Le spectre asymptotique d'EP au débit r est borné par la probabilité que la région d'erreur complète inclut les débit de transmission. Pour réseaux point à point [2] général, la région d'erreur complète est réduite à une valeur appelé *la capacité d'erreur complète*. En outre, il se trouve que pour les canaux satisfaisant la propriété de converse forte [43] l'EP coïncide avec la probabilité de panne. Dans ce sens la performance de réseaux composites peut être étudiée en utilisant des régions de débit atteignables disponibles et des régions d'erreur complètes. Spécialement les canaux satisfaisants converse fort sont de l'énorme intérêt parce que le spectre asymptotique d'EP peut coïncider avec la probabilité de panne pour eux.

Ici nous prouvons que la borne de flot-max coupe-min pour les réseaux sans mémoire discrète réside dans la région d'erreur complète. Autrement dit pour chaque code avec les débit de transmission pas satisfaisant la borne flot-max coupe-min, la probabilité d'erreur tend vers un. Cela fournira une borne sur le spectre asymptotique d'EP pour les réseaux sans mémoire discrète composite. Par conséquent les réseaux multi-terminaux qui ont la borne flot-max coupe-min comme la capacité sont satisfaisants la converse fort et ainsi le spectre asymptotique d'EP peut coïncider avec la probabilité de panne si la distribution de probabilité accomplissant la capacité pour chaque indice est la même. Pour prouver ce résultat la méthode de spectre d'information est utilisée.

1. Full error region
2. Point to point network

4.1 Notation et Arrière-Plan

La densité d'information[3] est défini par [44][4]

$$i(\mathbf{M}_n;\underline{\mathrm{Y}}) \triangleq \log \frac{\mathbb{P}_{Y^n|\mathbf{M}_n}(\underline{\mathrm{Y}}|\mathbf{M}_n)}{\mathbb{P}_{Y^n}(\underline{\mathrm{Y}})},$$

pour une suite arbitraire à n-dimension des sorties $\underline{\mathrm{Y}} = (Y_1, \dots, Y_n) \in \mathscr{Y}^n$ et $\mathbf{M}_n$ une VA uniforme sur l'ensemble des index $\mathcal{M}_n = \{1, \dots, M_n\}$. On utilise $\limsup$*en probabilité* des suites aléatoires Z_n, lequel est défini comme

$$\text{p-}\limsup_{n\to\infty} Z_n \triangleq \inf\left\{\beta : \lim_{n\to\infty} \Pr\{Z_n > \beta\} = 0\right\},$$

et également $\liminf$*en probabilité* de la suite aléatoire Z_n, définie comme

$$\text{p-}\liminf_{n\to\infty} Z_n \triangleq \sup\left\{\alpha : \lim_{n\to\infty} \Pr\{Z_n < \alpha\} = 0\right\}.$$

4.2 Définition de Réseau Multi-terminaux Composite (CMN)

Nous commençons par la description du Réseau Multi-terminal Composite (CMN) avec m-noeuds, qui est caractérisé par un ensemble de distributions de probabilités conditionnelles (PDs) :

$$\mathcal{W}_\theta = \left\{\mathbb{P}_{Y_{1\theta}^n \cdots Y_{m\theta}^n | X_{1\theta}^n \cdots X_{m\theta}^n} : \mathscr{X}_1^n \times \cdots \times \mathscr{X}_m^n \longmapsto \mathscr{Y}_1^n \times \cdots \times \mathscr{Y}_m^n\right\}_{n=1}^{\infty}$$

indexé avec n'importe quel vecteur de paramètres $\theta \in \Theta$ et où chaque noeud $i = \{1, \dots, m\}$ est équipé avec un émetteur $\underline{\mathrm{X}}_{i\theta} \in \mathscr{X}_i^n$ et un récepteur $\underline{\mathrm{Y}}_{i\theta} \in \mathscr{Y}_i^n$, comme décrit dans Fig. 4.1. Les entrées du matrices $\mathbb{P}_{Y_{1\theta}^n \cdots Y_{m\theta}^n | X_{1\theta}^n \cdots X_{m\theta}^n}$ seront souvent écrites comme $\mathbb{W}_\theta^n$. En plus, on dit que le réseau soit sans mémoire non-stationnaire si le PD collectif du canal multi-terminal se décompose comme

$$\mathbb{P}_{Y_{1\theta}^n \cdots Y_{m\theta}^n | X_{1\theta}^n \cdots X_{m\theta}^n}(\underline{\mathrm{y}}_1, \cdots, \underline{\mathrm{y}}_m | \underline{\mathrm{x}}_1, \cdots, \underline{\mathrm{x}}_m) = \prod_{t=1}^{n} \mathbb{W}_{\theta,t}(y_{1t}, \cdots, y_{mt} | x_{1t}, \cdots, x_{mt})$$

avec l'entré source $\underline{\mathrm{x}}_k = (x_{k1}, \dots, x_{kn}) \in \mathscr{X}_k^n$ et les sorties de canal $\underline{\mathrm{y}}_k = (y_{k1}, \dots, y_{kn}) \in \mathscr{Y}_k^n$, pour tout $k = \{1, \dots, m\}$, où $\left\{\mathbb{W}_{\theta,t}\right\}_{t=1}^{\infty}$ est une suite de canaux multi-terminaux à

3. Information Density

4. Laissons $\mathbb{P}_{Y^n|\mathbf{M}_n} = \dfrac{\mathbb{P}_{Y^n|\mathbf{M}_n}(d\underline{\mathrm{Y}}|\mathbf{M}_n)}{\mathbb{P}_{Y^n}(d\underline{\mathrm{Y}})}$ dénote *le dérivé de Radon-Nikodym* entre deux mesures de probabilités sur $\mathscr{Y}^n$ avec les valeurs sur un ensemble singulier supposé conventionnellement pour être $+\infty$. Alors, $\dfrac{\mathbb{P}_{Y^n|\mathbf{M}_n}(d\underline{\mathrm{Y}}|\mathbf{M}_n)}{\mathbb{P}_{Y^n}(d\underline{\mathrm{Y}})}$ est défini pour être $\mathbb{P}_{Y^n|\mathbf{M}_n}$, qui est évidemment une variable aléatoire (VA).

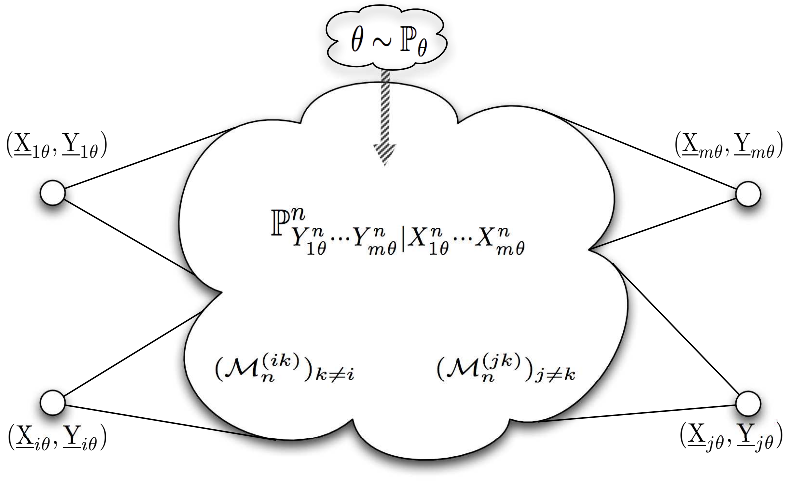


FIGURE 4.1 – Réseau Multi-terminaux Composite (CMN)

lettre simple [5]. De même on dit que le canal soit stationnaire et sans mémoire si $\mathbb{W}_{\theta,t} = \mathbb{W}_\theta$ pour tous $t = \{1, 2, \ldots, n\}$.

Laissons $\mathbb{P}_\theta$ dénoter n'importe quel PD arbitraire sur l'ensemble des paramètres de réseau (ou indices du canal) Θ. Avant que la communication commence, l'index $\theta \in \Theta$ de canal est supposé d'être tiré de $\mathbb{P}_\theta$ et reste fixe pendant la transmission entière.

L'ensemble $\mathcal{M}_n^{(ki)} \triangleq \{1, \ldots, M_n^{(ki)}\}$ représente l'ensemble des messages possibles à envoyer (dans n utilisations de canal) de la source k à la destination i avec $i \in \{1, \ldots, m\}\backslash\{k\}$. S'il n'y a aucun message destiné au noeud i du noeud k nous mettrons $\mathcal{M}_n^{(ki)} = \emptyset$.

Definition 6 (code et probabilité d'erreur) $\big(n, M_n^{(kj)}, (\epsilon_{n,\theta})_{\theta\in\Theta}\big)$*-code pour CMN se compose de :*

– *Un ensemble de fonction d'encodage pour $t = \{1, \ldots, n\}$ à chaque nœud $k \in \{1, \ldots, m\}$,*

$$\varphi_{t,\theta}^{(k)} : \bigotimes_{i=\{1,\ldots,m\}\backslash\{k\}} \mathcal{M}_n^{(ki)} \otimes \mathscr{Y}_k^{t-1} \longmapsto \mathscr{X}_k$$

où $\mathcal{M}_n^{(ki)}$ est l'ensemble de message de la source k destiné à la destination i, pour chaque $i = \{1, \ldots, m\}\backslash\{k\}$. Des symboles transmis $x_{kt} = \varphi_{t,\theta}^{(k)}(\underline{w}, y_k^{t-1})$ sont fonctions de symboles reçus passé y_k^{t-1} et tous les messages à envoyer de nœud k

$$\underline{w} \in \bigotimes_{i=\{1,\ldots,m\}\backslash\{k\}} \mathcal{M}_n^{(ki)}.$$

5. Single Letter

– *Une fonction de décodeur dans chaque noeud* $k \in \{1, \dots, m\}$,

$$\phi_{n,\theta}^{(jk)} : \mathscr{Y}_k^n \otimes \bigotimes_{i \in \{1,\dots,m\} \setminus \{k\}} \mathcal{M}_n^{(ki)} \longmapsto \mathcal{M}_n^{(jk)}$$

pour tout le noeud de source $j \neq k \in \{1, \dots, m\}$. *Cette fonction de décodage est par rapport le message destiné pour le noeud de destination* k *du noeud de source* j. *L'ensemble de décodage conformément à chaque fonction de décodage est défini par* $\mathcal{D}_{l,\theta}^{(jk)} \triangleq \phi_{n,\theta}^{(jk)^{-1}}(l)$ *pour tous les messages* $l \in \mathcal{M}_n^{(jk)}$, *qui correspond aux ensembles de décodage pour les messages* l *destiné au noeud* k *du noeud* j.

– *l'événement d'erreur* $\mathcal{E}_\theta^{(jk)}(l) \triangleq \{Y_{k\theta}^n \notin \mathcal{D}_{l,\theta}^{(jk)}\}$ *pour toutes les paires* $j \neq k \in \{1..., lem\}$ *et chaque* $l \in \mathcal{M}_n^{(jk)}$ *est défini comme l'événement que le message* l *du noeud* j *ne peut pas être correctement décodé à la destination* k. *Alors la probabilité d'erreur correspondante, basée sur chaque ensemble de décodage, sont définies par*

$$\mathbf{e}_{n,\theta}^{(jk)}(l) \triangleq \Pr\left(Y_{k\theta}^n \notin \mathcal{D}_{l,\theta}^{(jk)} | \mathbf{M}_n^{(jk)} = l\right), \tag{4.1}$$

où $\mathbf{M}_n^{(jk)}$ *dénote le VA correspondant au message transmis. En supposant un PD uniforme sur les ensembles de message, la probabilité d'erreur moyennes (EP) sont définies comme*

$$\epsilon_{n,\theta}^{(jk)} \triangleq \frac{1}{M_n^{(jk)}} \sum_{l=1}^{M_n^{(jk)}} \mathbf{e}_{n,\theta}^{(jk)}(l) \tag{4.2}$$

et l'EP maximum comme

$$\epsilon_{\max,n,\theta}^{(jk)} \triangleq \max_{l \in \mathcal{M}_n^{(jk)}} \mathbf{e}_{n,\theta}^{(jk)}(l) \geq \epsilon_{n,\theta}^{(jk)}. \tag{4.3}$$

Donc l'événement d'erreur pour le réseau est l'union de tous les événements d'erreur $\mathcal{E}_\theta^{(jk)}$ *sur toutes les sources* j *et les destinations* k *avec les messages correspondants* l. *L'EP du réseau s'écrit comme*

$$\epsilon_{n,\theta} \triangleq \Pr\left(\bigcup_{j \neq k \in \{1,\dots,m\}} \bigcup_{l \in \mathcal{M}_n^{(jk)}} \left\{Y_{k\theta}^n \notin \mathcal{D}_{l,\theta}^{(jk)}, \mathbf{M}_n^{(jk)} = l\right\}\right) \tag{4.4}$$

où c'est facile de voire

$$\begin{aligned} \epsilon_{n,\theta} &\leq \sum_{j \neq k \in \{1,\dots,m\}} \epsilon_{n,\theta}^{(jk)} \\ &\leq \sum_{j \neq k \in \{1,\dots,m\}} \epsilon_{\max,n,\theta}^{(jk)}. \end{aligned} \tag{4.5}$$

A travers ce chapitre EP moyenne sera utilisé. Remarquons que dans cadre de CMN, les probabilités d'erreur $\epsilon^{(jk)}_{\max,n,\theta}$, $\epsilon^{(jk)}_{n,\theta}$ et $\epsilon_{n,\theta}$ sont VAs, effectivement elles sont des fonctions de l'index aléatoire θ de canal. Par exemple, pour chaque fixe $\theta = \theta$ le CMN est réduit à un réseau multi-terminal conventionnel. Pour le reste, nous utilisons aussi la notion de C pour désigner un code.

Une chose est importante de noter ici. En général il est raisonnable de supposer si $\mathcal{M}^{(ki)}_n \neq \emptyset$ pour tout k, i, autrement dit si chaque noeud envoie quelque chose et chaque noeud reçoit quelque chose alors chaque noeud peut être instruit de θ parce que l'encodeur et les fonctions de décodeur peuvent maintenant être choisis basés sur θ, c'est-à-dire CSIT et CSIR sont tous les deux disponibles. Dans ce cas-là, si chacun sait le canal qui est en opération, alors le problème est réduit au réseau multi-terminal ordinaire et il n'y a aucun besoin pour l'étude de plus du sujet au-delà des méthodes ordinaires.

Pourtant dans les réseaux réels, pas chaque noeud transmet et reçoit dans le même temps et de plus pas chaque noeud est la source et la destination. Si le noeud k est seulement l'émetteur qui signifie que $\mathcal{Y}_k = \emptyset$, donc il n'y a aucune voie pour cela pour avoir des informations du canal en opération et donc il est effectivement ignorant. D'autre part si le noeud k est seulement le récepteur qui signifie que $\mathcal{X}_k = \emptyset$, alors il n'y a aucune voie pour cela pour envoyer les informations de son observation du canal aux autres utilisateurs et cela signifie que les autres utilisateurs sont nécessairement inconscient de CSI de cet utilisateur. D'autre part il y a des utilisateurs qui servent des relais qui signifie que $\mathcal{M}^{(ik)}_n = \emptyset$ et $\mathcal{M}^{(ki)}_n = \emptyset$ pour tous $i \neq k$. Ces utilisateurs peuvent seulement partiellement savoir CSI et ils ne sont pas naturellement instruits de CSI des utilisateurs sans entré de canal.

Le modèle qui semble être plus adapté au scénario pratique se compose de trois types de noeuds comme dans Fig. 4.2. Ces noeuds avec l'index appartenant à l'ensemble $\mathcal{T}$ sont seulement des émetteurs et des sources. Le noeud i où $i \in \mathcal{T}$ émet $\underline{X}_i$ indépendant de θ parce qu'il ne peut pas y avoir accès. Ces noeuds avec l'index dans $\mathcal{D}$ sont seulement des destinations. Ces noeuds ne peuvent pas transmettre leur propre observation aux autres noeuds pourtant ils peuvent avoir accès à CSI de tous les émetteurs. Finalement ces noeuds avec l'index dans $\mathcal{R}$ transmettent simultanément et reçoivent des informations. Les relais font partie essentiellement de ces noeuds. Ils ne peuvent pas avoir l'accès à CSI complète à cause de la présence de noeuds dans $\mathcal{D}$ et peuvent seulement être partiellement conscients de CSI. Nous pouvons supposer que θ est effectivement composé des deux parties θ_d et θ_r où les noeuds dans $\mathcal{R}$ peuvent seulement être instruits de θ_r. Pour le reste il est toujours présupposé que nous nous occupons de la sorte de réseau où CSI complète n'est

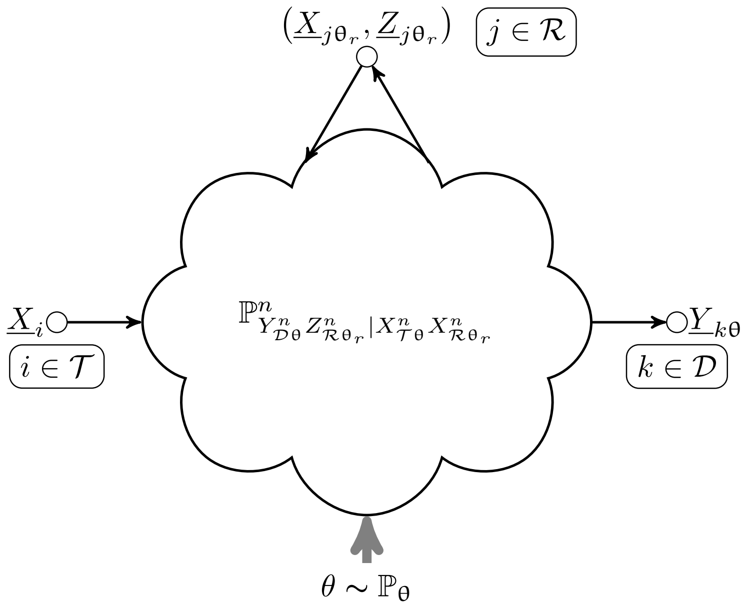


Figure 4.2 – Réseau Multi-terminal Composite (CMN) avec $\theta = (\theta_r, \theta_d)$

pas disponible à chaque noeud. Nous définissons ensuite la notion de débit atteignable et la région de capacité pour le CMN.

Definition 7 (capacité et débit atteignable) *Une probabilité d'erreur* $0 \leq \epsilon < 1$ *et un vecteur des débits* $\underline{r} = (r_{jk})_{j \neq k \in \{1,\ldots,m\}}$ *sont dit atteignable pour les réseaux multi-terminaux composite, s'il y a un* $\big(n, M_n^{(jk)}, (\epsilon_{n,\theta})_{\theta \in \Theta}\big)$*-code tel que pour tous les paires* $j \neq k \in \{1, \ldots, m\}$

$$\liminf_{n\to\infty} \frac{1}{n} \log M_n^{(jk)} \geq r_{jk} \tag{4.6}$$

et

$$\limsup_{n\to\infty} \sup_{\theta\in\Theta} \epsilon_{n,\theta} \leq \epsilon. \tag{4.7}$$

Alors la région de ϵ*-capacité* $\mathcal{C}_\epsilon$ *est définie par la région comprenant tous les débit atteignables satisfaisant* (4.7)*. De même* $(\epsilon, \epsilon_{n,\theta})$ *dans la définition* (4.7) *peut être remplacée par* $(\epsilon^{(jk)}, \epsilon^{(jk)}_{n,\theta})$ *qui correspond à l'erreur tolérée du noeud source* j *au noeud de destination* k*. Alors la notion de région de* $\underline{\epsilon}$*-capacité*

$$\mathcal{C} \triangleq \lim_{\underline{\epsilon}\to 0} \mathcal{C}_{\underline{\epsilon}}.$$

Remarquons que la fonction de fiabilité (4.7) a été choisie dans le sens le plus fort. En général, les définitions précédentes peuvent causer des débit atteignables nuls parce que chaque noeud doit fixer le débit tel que $\epsilon_{n,\theta}$ est moins que ϵ et l'index $\theta \in \Theta$ le plus mauvais possible peut avoir la ϵ-capacité zéro, pour n'importe quels $0 \leq \epsilon < 1$. En outre, dans les réseaux sans fil il est rare d'avoir le débit non-zéro pour les plus mauvaises tirages de canal mais il est désirable d'envoyer les informations et la performance de mesure d'une manière ou d'une autre. Comme on voit dans la section suivante, en comptant sur le PD $\mathbb{P}_\theta$, plusieurs différentes notions pour la fiabilité peuvent être suggérées.

4.3 Fonction de Fiabilité pour Réseaux Composites

Une approche alternative est l'étude de la conduite de probabilités d'erreur $\epsilon_{n,\theta}^{(jk)}$, $\epsilon_{n,\theta}$ comme n tend vers l'infinité pour les débit fixes. Pour le reste, nous nous concentrerons sur l'étude de la probabilité d'erreur. Nous supposons que $\epsilon_{n,\theta}$ converge à ϵ_θ presque partout pour faciliter du travail avec les bornes. Pourtant les résultats restent valides si nous le remplaçons avec la beaucoup plus faible hypothèse que $\epsilon_{n,\theta}$ converge dans la distribution à ϵ_θ. Parce que $\epsilon_{n,\theta}$ est uniformément intégrable, et alors le travail avec la Limite reste intact. [6]

Definition 8 (fonctions de fiabilité) *On dit que la valeur $0 \leq \epsilon < 1$ soit atteignable pour un tuple de débit $\underline{r} = (r_{jk})_{j \neq k \in \{1,\ldots,m\}}$ basé sur les fonctions de fiabilité suivantes, s'il y a un code de $\big(n, M_n^{(jk)}, (\epsilon_{n,\theta})_{\theta \in \Theta}\big)$ tel que pour toutes les paires $j \neq k \in \{1,\ldots,m\}$, les débit satisfaisants*

$$\liminf_{n\to\infty} \frac{1}{n} \log M_n^{(jk)} \geq r_{jk},$$

alors $\epsilon_{n,\theta}$ satisfait certaine condition de fiabilité, listé comme la dessous.

– *Si nous regardons la limite de $\epsilon_{n,\theta}$ quand $n \to \infty$ en utilisant la notion de Convergence presque partout (p.p), alors ϵ est dit être atteignable si la limite est moins ou égale à ϵ presque partout. Cela signifie que :*

$$\mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} \leq \epsilon) = 1. \tag{4.8}$$

Il garantie que pour tous les sous-ensembles dans Θ ayant une mesure non-zéro l'EP asymptotique sera pas plus grand que ϵ. Alors basé sur la convergence p.p.,

6. Il est toujours possible d'utiliser lim sup pour assurer la convergence pourtant les égalités ne sont pas tous valides et se transforment en inégalités pour quelques cas.

le désavantage provoqué par les événements de mesure zéro peut être enlevée et la fonction de fiabilité détendue. Pourtant on peut voir que :

$$\begin{aligned}\mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > \epsilon) &= \mathbb{E}(\mathbf{1}[\lim_{n\to\infty} \epsilon_{n,\theta} > \epsilon]) \\ &= \mathbb{E}(\lim_{n\to\infty} \mathbf{1}[\epsilon_{n,\theta} > \epsilon]) \\ &\overset{(a)}{=} \lim_{n\to\infty} \mathbb{E}(\mathbf{1}[\epsilon_{n,\theta} > \epsilon]) \\ &= \lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \epsilon)\end{aligned}$$

où (a) vient du théorème de convergence dominé de Lebesgue. Cela signifie que la notion de convergence presque partout est équivalente à la notion d'habitude plus lâche de Convergence dans la probabilité pour ce cas. Cela signifie aussi que ϵ est dit être atteignable si :

$$\lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \epsilon) = 0$$

EP atteignable peut aussi être caractérisé par

$$\epsilon_{\text{-}p}(\underline{r}, \mathsf{C}) = p\text{-}\limsup_{n\to\infty} \epsilon_{n,\theta} \tag{4.9}$$

où nous avons

$$p\text{-}\limsup_{n\to\infty} \epsilon_{n,\theta} = \inf\left\{\alpha : \lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \alpha) = 0\right\}. \tag{4.10}$$

Cela signifie que ϵ est atteignable s'il y a un code C tel que l'EP est plus grand que ou égal à $\epsilon_{\text{-}p}(\underline{r}, \mathsf{C})$ qui signifie qu'avec la probabilité non-zéro, $\epsilon_{n,\theta}$ peut excéder ϵ. Pourtant, le problème avec ces notions est que pour chaque $\epsilon < 1$ il ne peut être aucun code satisfaisant $\mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > \epsilon) = 0$ ou $\lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \epsilon) = 0$. Autrement dit, pour chaque $0 \leq \epsilon < 1$ et chaque code, il y a une probabilité non-zéro que l'erreur y excède (par ex. réseaux sans fil). Pourtant la condition dans (4.10) peut être détendue à

$$\epsilon_{\text{-}\delta}(\underline{r}, \mathsf{C}) = \delta\text{-}\limsup_{n\to\infty} \epsilon_{n,\theta} \tag{4.11}$$

où

$$\delta\text{-}\limsup_{n\to\infty} \epsilon_{n,\theta} = \inf\left\{\alpha : \lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \alpha) \leq \delta\right\}, \tag{4.12}$$

pour chaque constant $0 \leq \delta < 1$. Nous appelons cette notion EP $\delta-$atteignable. Cela signifie que ϵ est $\delta-$atteignable s'il y a un code C tel que $\epsilon_{\text{-}\delta}(\underline{r}, \mathsf{C})$ est moins que ϵ qui signifie qu'il y a un code tel que $\epsilon_{n,\theta}$ est moins que ϵ avec au moins la probabilités $1 - \delta$.

– *Aux mêmes lignes, la probabilité d'erreur moyenne peut être caractérisée pour le code* C *comme suit*

$$\bar{\epsilon}(\underline{r}, \mathsf{C}) = \lim_{n\to\infty} \mathbb{E}_\theta[\epsilon_{n,\theta}]. \tag{4.13}$$

Cette définition peut aboutir à la définition selon laquelle ϵ *est dit atteignable s'il y a un code* C *tel que* ϵ *est plus grand que l'erreur moyenne qui implique l'existence de codes avec EP moins que* ϵ *dans* $\mathcal{L}^1$*, mais pas partout, en signifiant que pour certains* $\theta \in \Theta$ *l'asymptotique EP peut tomber au dessus* $\bar{\epsilon}$*. Cela montre que l'erreur moyenne n'est pas assez précise pour caractériser la probabilité d'erreur, comme il sera montré plus tard. Il devrait être dit en passant ici qu'EP espéré est équivalent à la définition d'EP pour le canal moyenné*[7] *dans [45].*

– *Le débit EP*[8] *est défini pour un code* C *par*

$$\epsilon_T(\underline{r}, \mathsf{C}) = \sup_{0\le\alpha<1} \lim_{n\to\infty} \alpha\, \mathbb{P}_\theta(\epsilon_{n,\theta} > \alpha). \tag{4.14}$$

Le débit EP tient compte de la probabilité d'erreur désirée ϵ *et de la probabilité que l'erreur y excède. Ainsi si la probabilité d'erreur excédent sur un grand* ϵ *avec de petite probabilité, le débit EP le prend en compte.* ϵ *est dit atteignable en ce qui concerne cette mesure s'il y a un code* C *tel que* ϵ *est plus grand que* $\epsilon_T(\underline{r}, \mathsf{C})$*.*

Il est particulièrement intéressant de définir *le plus petit EP atteignable*, caractérisé par

$$\epsilon_{\text{-p}}(\underline{r}) = \inf_{\mathsf{C}} \epsilon_{\text{-p}}(\underline{r}, \mathsf{C}) = \inf_{\mathsf{C}} \text{p-}\limsup_{n\to\infty} \epsilon_{n,\theta} \tag{4.15}$$

où l'infimum est pris sur tous les codes. Cela signifie que pour ϵ plus petit que $\epsilon_{\text{-p}}(\underline{r})$, il y a un code tel que nous avons :

$$\lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \epsilon) > 0.$$

Le plus petit EP atteignable est un bon indicateur des canaux composites. Remarquons que le sens du plus petit EP atteignable ϵ est que les informations peuvent être envoyées au débit $\underline{r}$ avec une probabilité diminuant que l'EP tombe au dessus de ϵ.

Pourtant, le même problème avec cette notion est que pour quelques cas cette valeur n'est pas non-banale. Alors la même idée peut être utilisée ici pour détendre la condition précédente et définir *le* δ *plus petit EP atteignable* :

$$\epsilon_{\text{-}\delta}(\underline{r}) = \inf_{\mathsf{C}} \epsilon_{\text{-}\delta}(\underline{r}, \mathsf{C}) = \inf_{\mathsf{C}} \delta\text{-}\limsup_{n\to\infty} \epsilon_{n,\theta} \tag{4.16}$$

7. Averaged Channel
8. The throughput error probability

où infimum est pris de nouveau sur tous les codes. Cela signifie que pour ϵ plus grand que $\epsilon_{\text{-}\delta}(\underline{r})$, il y a un code tel que $\epsilon_{n,\theta}$ est moins que ϵ avec au moins probabilité de $1-\delta$.

Donc, d'une part l'EP espéré (4.13) peut ne pas toujours être la fonction de fiabilité adéquate pour le CMN, mais d'autre part (4.9) peut produire des débit très pessimistes. En outre, selon l'application prévue il peut y avoir de différentes fonctions de fiabilité d'intérêt pour les modèles composites. Alors la question est s'il y a une mesure universelle de fiabilité dont les autres peuvent en être tirés. La section suivante présente une telle quantité fondamentale, appelée le *le spectre asymptotique de probabilités d'erreur* (ASEP).

4.4 Spectre Asymptotique de Probabilité d'erreur

Dans la section précédente, nous avons discuté de différentes notions d'être atteignable pour une erreur. La plus petite erreur atteignable a été définie pour un fixe $\underline{r}$ en utilisant de différents critères. Maintenant nous enquêtons sur PD cumulatif asymptotique d'EP pour le vecteur fixe de débit de transmission $\underline{r} = (r_{jk})_{j\neq k\in\{1,\ldots,m\}}$, qui est donné par

$$\lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} \leq \epsilon),$$

pour chaque $0 \leq \epsilon \leq 1$.

Definition 9 (spectre asymptotique de EP) *Pour chaque $0 \leq \epsilon \leq 1$ et les débit de transmission $\underline{r} = (r_{jk})_{j\neq k\in\{1,\ldots,m\}}$, le spectre asymptotique d'EP pour le code donné C, $\mathcal{E}(\underline{r},\epsilon,\mathsf{C})$ est défini comme*

$$\mathcal{E}(\underline{r},\epsilon,\mathsf{C}) = \lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \epsilon). \tag{4.17}$$

Le spectre asymptotique d'EP pour CMN est défini comme :

$$\mathcal{E}(\underline{r},\epsilon) = \inf_{\mathsf{C}} \lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \epsilon) \tag{4.18}$$

où l'infimum est pris sur tous les codes $\left(n, M_n^{(jk)}, (\epsilon_{n,\theta})_{\theta\in\Theta}\right)$ avec les débits satisfaisant

$$\liminf_{n\to\infty} \frac{1}{n} \log M_n^{(jk)} \geq r_{jk},$$

for all pairs $j \neq k \in \{1,\ldots,m\}$.

La notion de $\mathcal{E}(\underline{r},\epsilon)$ indique intuitivement ce qui est la plus petite probabilité que l'erreur tombe au dessus de ϵ. Il sera montré que cette notion est la mesure la plus générale pour la performance de réseaux composites et implique toutes les autres notions.

Il est aussi intéressant de voir en particulier que pour un débit de transmission donné $\underline{r}$, quel sont les possibles probabilités d'erreur ; autrement dit, pour trouver si la valeur d'asymptotique de $\epsilon_{n,\theta}$ est moins qu'une valeur désirée ou non. Cette idée sous-tend l'idée d'erreur atteignable pour les réseaux multiterminaux composites.

Le théorème suivant fournit une relation entre le spectre asymptotique d'EP et les autres notions présentées auparavant.

Theorem 18 *Pour les réseaux composites avec le débit* $\underline{r}$*, le spectre asymptotique d'EP implique d'autres fonctions de fiabilité présentées auparavant. Le plus petit EP atteignable et* δ *plus petit EP atteignable peuvent être obtenus comme suit :*

$$\begin{aligned}\epsilon_{\text{-}p}(\underline{r}) &= \inf\left\{0 \leq \epsilon < 1 : \mathcal{E}(\underline{r}, \epsilon) = 0\right\} \\ \epsilon_{\text{-}\delta}(\underline{r}) &= \inf\left\{0 \leq \epsilon < 1 : \mathcal{E}(\underline{r}, \epsilon) \leq \delta\right\}.\end{aligned}$$

Le débit EP pour le code C *est obtenu comme suit*

$$\epsilon_T(\underline{r}, \mathsf{C}) = \sup_{0 \leq \epsilon < 1} \epsilon \mathcal{E}(\underline{r}, \epsilon, \mathsf{C}).$$

Finalement EP espéré pour le code C *est obtenu comme suit :*

$$\bar{\epsilon}(\underline{r}, \mathsf{C}) = \int_0^1 \mathcal{E}(\underline{r}, \epsilon, \mathsf{C}) d\epsilon.$$

Proof La preuve de trois premières égalités suit directement de la définition. Pour la dernière inégalité, en utilisant le fait que $\epsilon_{n,\theta}$ est positif et borné, nous avons :

$$\begin{aligned}\bar{\epsilon}(\underline{r}, \mathsf{C}) &= \lim_{n\to\infty} \mathbb{E}[\epsilon_{n,\theta}] \\ &\overset{(a)}{=} \mathbb{E}[\lim_{n\to\infty} \epsilon_{n,\theta}] \\ &= \int_0^{+\infty} \mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > t) dt \\ &= \int_0^1 \mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > t) dt \\ &= \int_0^1 \lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > t) dt\end{aligned}$$

où (a) vient du théorème de convergence dominé de Lebesgue. Et cela conclura la preuve.

Le théorème précédent déclare que l'EP espéré n'est pas nécessairement atteignable dans le sens strict pour un débit donné r. Donc, c'est possible en général que l'EP excède sur EP espéré. Cette observation montre que l'erreur espérée, bien qu'indicatif, ne serait pas toujours une mesure nécessaire pour EP.

Nous nous intéressons à la conduite une probabilité d'erreur $\epsilon_{n,\theta}$ qui est une variable aléatoire. En particulier la caractérisation du spectre asymptotique d'EP est du grand intérêt pour les réseaux au hasard. Cela nous donnera des meilleurs critères sur la probabilité d'erreur qui peuvent être accomplies dans un canal comparé aux probabilité de panne. Spécialement, ce sera utile pour les cas où l'émetteur fixe un débit r sans tenir compte du canal en opération. C'est d'habitude le cas en pratique où le débit est déterminé par les médias en utilisation. Pourtant il est intéressant de voir la relation entre la probabilité de panne et le spectre asymptotique d'EP. Dans la section suivante nous estimons comment caractériser le spectre asymptotique d'EP et la notion est mise en corrélation avec la probabilité de panne.

4.5 Résultats Principaux

Considérons un canal composite général. On peut remarquer que la distribution de probabilité de $\epsilon_{n,\theta}$ comme $n \to \infty$ est directement liée aux probabilités que le vecteur de débit $\underline{r}$ tombe dans la région de ϵ-capacité $\mathcal{C}_{\epsilon,\theta}$, où $\mathcal{C}_{\epsilon,\theta}$ est une région aléatoire avec θ comme le paramètre aléatoire. Supposons que la transmission est faite au débit $\underline{r}$ sur un canal non-composite. Alors si le code accomplit la probabilité d'erreur ϵ alors son débit devrait nécessairement appartenir à la région de ϵ-capacité. D'autre part si le débit appartient à la région de ϵ-capacité, donc il y a un code tel qu'il accomplit la probabilité d'erreur ϵ.

Pourtant en cas des réseaux composites, l'émetteur, en ignorant θ, a un code simple pour tous θ. Alors si le débit n'appartient pas à $\mathcal{C}_{\epsilon,\theta}$ pour un θ, donc $\epsilon_{n,\theta}$ excédera ϵ sûrement. Mais si le débit appartient à $\mathcal{C}_{\epsilon,\theta}$ pour un θ, donc il n'est pas garanti que $\epsilon_{n,\theta}$ n'excédera pas ϵ. Parce que bien qu'il y ait un code tel que l'EP est moins que ϵ, mais cela peut ne pas être le code utilisé par l'émetteur. Cela causera le théorème suivant.

Theorem 19 *Pour le réseau multiterminal composite avec le paramètre aléatoire θ nous avons :*

$$\mathbb{P}_\theta(\limsup_{n\to\infty} \epsilon_{n,\theta} > \epsilon) \geq \mathbb{P}_\theta(\underline{r} \notin \mathcal{C}_{\epsilon,\theta}), \tag{4.19}$$

où $\mathcal{C}_{\epsilon,\theta}$ est ϵ-capacité du réseau W_θ pour θ donné et $0 \leq \epsilon < 0$.

Proof Selon la définition, pour chaque θ, $\underline{r}$ est à l'intérieur de $\mathcal{C}_{\epsilon,\theta}$ si $\limsup_{n\to\infty} \epsilon_{n,\theta} \leq \epsilon$. Cela nous donne la preuve pour le théorème.

Dans le théorème 19 ϵ peut être remplacé avec $\epsilon^{(ij)}$ et respectivement $\epsilon_{n,\theta}$ à $\epsilon_{n,\theta}^{(ij)}$. Ce changement de la définition change aussi la définition de ϵ-capacité à la $\underline{\epsilon}$-capacité et ainsi le théorème reste valide sous le changement.

Supposons que les émetteurs qui ignorent le canal, fixent leur fonction encodage basée sur $\varphi_t^{(k)}$ et définissent ϕ comme l'ensemble de ces fonctions. Pour chacun θ et ϕ, définissons $\mathcal{R}_{\epsilon,\theta}(\phi)$ comme la région ϵ-atteignable tel que si le débit y appartient, alors l'EP est moins ou égal que ϵ pour le choix de ϕ. Maintenant nous avons :

$$\mathcal{E}(\underline{r}, \epsilon, \mathsf{C}) = \mathbb{P}_\theta(\underline{r} \notin \mathcal{R}_{\epsilon,\theta}(\phi)).$$

Cela présente une borne supérieure sur le spectre asymptotique d'EP. De plus, en prenant la limite à l'extérieur $\mathbb{P}(\lim_{n\to\infty} \epsilon_{n,\theta} < \epsilon)$, nous recevons le corollaire suivant.

Corollary 8 *Pour la probabilité d'erreur $\epsilon_{n,\theta}$ et la ϵ-capacité définie comme auparavant, le spectre asymptotique d'EP est comme suit :*

$$\inf_{\phi} \mathbb{P}_\theta(\underline{r} \notin \mathcal{R}_{\epsilon,\theta}(\phi)) \geq \mathcal{E}(\underline{r}, \epsilon) = \lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \epsilon) \geq \mathbb{P}_\theta(\underline{r} \notin \mathcal{C}_{\epsilon,\theta}). \tag{4.20}$$

Il y a des canaux composites comme les canaux symétriques binaires composites (CBSC) où un code unique, le code uniformément distribué pour CBSC, produit le meilleur code pour chaque canal. Dans ce cas-là, nous avons l'égalité suivante :

$$\mathcal{E}(\underline{r}, \epsilon) = \mathbb{P}_\theta(\underline{r} \notin \mathcal{C}_{\epsilon,\theta}). \tag{4.21}$$

Effectivement l'exemple suivant est un cas de ces canaux composites avec le meilleur code unique. Nous jetons un coup d'oeil plus proche à ces notions et à leur relation avec $\epsilon-$capacité.

l'Exemple (Canal Symétrique Binaire Moyenné Composite) [46] : un canal symétrique binaire moyenné avec trois paramètres est défini comme la moyenne de trois canaux symétriques binaires $(\mathbb{B}_1, \mathbb{B}_2, \mathbb{B}_3)$ avec les paramètres suivants :

$$p_1 < p_2 < p_3 \leq \frac{1}{2}$$

Les coefficients du fait en moyenne sont $\alpha_1, \alpha_2, \alpha_3$ tel que :

$$\alpha_1\alpha_2\alpha_3 = 1.$$

Le canal moyenné est alors défini comme $\mathbb{B} = \alpha_1\mathbb{B}_1 + \alpha_2\mathbb{B}_2 + \alpha_3\mathbb{B}_3$. La capacité de canal symétrique binaire avec le paramètre p est connue comme :

$$\mathcal{C}(p) = 1 - H(p).$$

Kieffer a calculé la capacité de canal symétrique binaire moyenné et a montré que le canal n'est pas satisfaisant la converse forte. De plus $\epsilon-$capacité de ce canal est caractérisé comme suit :

$$\mathcal{C}_\epsilon = \begin{cases} \mathcal{C}(p_3) & 0 < \epsilon < \alpha_3 \\ \mathcal{C}(\lambda(p_2, p_3)) & \epsilon = \alpha_3 \\ \mathcal{C}(p_2) & \alpha_3 < \epsilon < \alpha_3 + \alpha_2 \\ \mathcal{C}(\lambda(p_1, p_2)) & \epsilon = \alpha_3 + \alpha_2 \\ \mathcal{C}(p_1) & \alpha_3 + \alpha_2 < \epsilon < 1 \end{cases} \tag{4.22}$$

où $\lambda(p_1, p_2)$ est défini comme :

$$\lambda(p, q) = \frac{\log\left(\frac{1-p}{1-q}\right)}{\log\left(\frac{1-p}{1-q}\right) + \log\left(\frac{q}{p}\right)}.$$

Supposons maintenant qu'il y a un élément aléatoire associé à ce canal. Supposons par exemple que $\mathbf{p}_3$ prend sa valeur aléatoirement entre p_2 et $\frac{1}{2}$ avec la mesure $\mathbb{P}_{\mathbf{p}_3}$. Autrement dit, le paramètre de canal θ est $\mathbf{p}_3$. Alors le spectre asymptotique de probabilités d'erreur est comme suit :

$$\mathcal{E}(\underline{r}, \epsilon) = \begin{cases} \mathbb{P}_{\mathbf{p}_3}\big(r > \mathcal{C}(\mathbf{p}_3)\big) & 0 < \epsilon < \alpha_3 \\ \mathbb{P}_{\mathbf{p}_3}\big(r > \mathcal{C}(\lambda(p_2, \mathbf{p}_3))\big) & \epsilon = \alpha_3 \\ \mathbf{1}[r > \mathcal{C}(p_2)] & \alpha_3 < \epsilon < \alpha_3 + \alpha_2 \\ \mathbf{1}[r > \mathcal{C}(\lambda(p_1, p_2))] & \epsilon = \alpha_3 + \alpha_2 \\ \mathbf{1}[r > \mathcal{C}(p_1)] & \alpha_3 + \alpha_2 < \epsilon < 1 \end{cases} \tag{4.23}$$

Pour obtenir le plus petit EP atteignable, nous devons jeter un coup d'œil à la plus petite valeur de ϵ tel que la probabilité d'erreur asymptotique ne l'excède pas. Dans cet exemple, le seul élément aléatoire du canal est gouverné par $\mathbf{p}_3$ et la source prend conscience du fait que s'il transmet un code avec le débit $r > \mathcal{C}(p_2)$, alors le code ne sera pas décodé correctement. Ainsi pour le reste supposons que la source transmet un code avec $r \leq \mathcal{C}(p_2)$.

Maintenant les trois derniers termes dans le spectre asymptotique d'EP sont automatiquement zéro et nous avons :

$$\mathcal{E}(\underline{r},\epsilon) = \begin{cases} \mathbb{P}_{\mathbf{p}_3}\big(r > \mathcal{C}(\mathbf{p}_3)\big) & 0 < \epsilon < \alpha_3 \\ \mathbb{P}_{\mathbf{p}_3}\big(r > \mathcal{C}(\lambda(p_2,\mathbf{p}_3))\big) & \epsilon = \alpha_3 \\ 0 & \alpha_3 < \epsilon < 1 \end{cases} \tag{4.24}$$

Pour ce cas on peut voir que le plus petit EP atteignable est comme suit :

$$\epsilon_{\text{p-}}(\underline{r}) = \inf\{0 \le \epsilon < 1 : \mathcal{E}(\underline{r},\epsilon) = 0\} \le \alpha_3.$$

Dans d'autre mot pour ce canal, la probabilité d'erreur est au dessous α_3 avec la probabilité 1 pour $r < \mathcal{C}(p_2)$. D'autre part l'erreur espérée peut être calculée comme suit :

$$\bar{\epsilon}(\underline{r}) = \int_0^1 \mathcal{E}(\underline{r},\epsilon)d\epsilon = \int_0^{\alpha_3} \mathbb{P}_{\mathbf{p}_3}\big(r > \mathcal{C}(\mathbf{p}_3)\big)d\epsilon = \alpha_3 \times \mathbb{P}_{\mathbf{p}_3}\big(r > \mathcal{C}(\mathbf{p}_3)\big).$$

Il peut être directement vu que l'erreur espérée néglige l'information de l'erreur pour la probabilité d'erreur au point $\epsilon = \alpha_3$. Cela montre de nouveau que l'erreur espérée n'est pas assez général pour fournir des informations complètes sur l'erreur. Finalement le débit EP est calculé comme suit :

$$\bar{\epsilon}_T(\underline{r}) = \sup_{0\le\epsilon<1} \epsilon\mathcal{E}(\underline{r},\epsilon) = \alpha_3 \times \mathbb{P}_{\mathbf{p}_3}\big(r > \mathcal{C}(\lambda(p_2,\mathbf{p}_3))\big).$$

Ici les informations sur l'erreur pour ϵ moins que α_3 se sentent perdues dans la notion. Cet exemple montre clairement la relation entre toutes ces notions et comment le spectre asymptotique d'EP est la notion qui implique toutes les autres notions et inclut toutes les informations par rapport aux probabilités d'erreur dans les canaux composites.

Pourtant le problème principal est que la capacité n'est pas connue en général pour la plupart des réseaux multiterminaux et par conséquent non plus la ϵ-capacité. Donc nous devons chercher des façons de caractériser le spectre asymptotique d'EP d'autres façons.

Une option consiste à analyser la relation entre la notion de probabilité de panne et du spectre asymptotique d'EP. La probabilité de panne P_{out} est définie comme la probabilité qu'un code avec le débit $\underline{r}$, ne peut pas être correctement décodé qui signifie qu'il a l'erreur non-zéro. Les probabilité de panne sont alors égales à :

$$P_{out} = \mathbb{P}_\theta(\underline{r} \notin \mathcal{C}_\theta).$$

Supposons maintenant que chaque canal pour chaque θ satisfait la condition de converse forte. Cela signifie que pour chaque canal, chaque code avec le vecteur de débit à l'extérieur

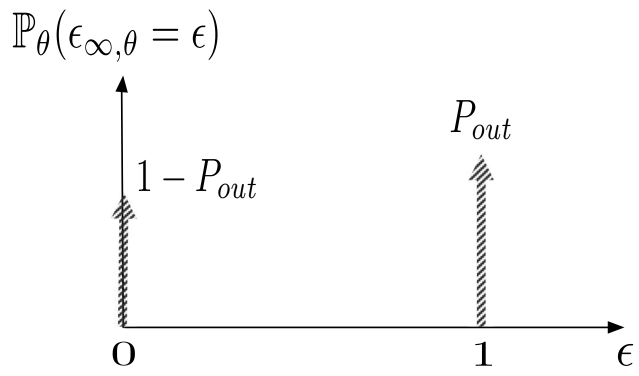


Figure 4.3 – Réseaux Multiterminaux satisfaisant la converse forte

de la région de capacité produit asymptotique la probabilité d'erreur 1. Cela signifie aussi

$$\mathcal{C}_\theta = \mathcal{C}_{\epsilon\theta} \ \ 0 \le \epsilon < 1.$$

Ainsi pour chacun θ, la probabilité d'erreur asymptotique, c'est-à-dire $\limsup_{n\to\infty} \epsilon_{n,\theta}$ prend comme la valeur zéro ou un. De plus s'il y a le meilleur code unique pour le canal composite alors de (4.21), il en suit que la probabilité d'erreur asymptotique peuvent être considérées comme un événement Bernoulli avec le paramètre P_{out} où P_{out} est la probabilité de panne (Fig. 4.3). Donc ces canaux satisfaisant la condition de converse forte sont de l'intérêt particulier parce que la notion de probabilité de panne dans ces cas coïncide avec la notion du spectre asymptotique d'EP.

Une autre option consiste à essayer de borner le spectre asymptotique d'EP. Il y a des bornes intérieures différentes et des bornes supérieures, des débit atteignables et des converses pour les réseaux multi-terminaux. Considérons un réseau multi-terminal composite avec le paramètre θ où une région atteignable est connue pour chaque θ et ϕ, défini comme auparavant, $\mathcal{R}_\theta(\phi)$. Maintenant si le débit $\underline{r}$ est à l'intérieur de la région alors la probabilité d'erreur tend vers zéro et sera moins que ϵ pour $0 < \epsilon < 1$. Pour le débit $\underline{r}$, le nombre de ces canaux avec la probabilité d'erreur plus grande que ϵ est moins ou égal au nombre de canaux avec la probabilité d'erreur non-zéro qui impliquent que le spectre asymptotique d'EP est essentiellement moins ou égal que la probabilité que le débit $\underline{r}$ n'est pas à l'intérieur de la région atteignable.

De la même façon pour le débit $\underline{r}$, le nombre de ces canaux avec la probabilité d'erreur plus grande que ϵ est moins ou égal au nombre de canaux avec la probabilité d'erreur égale à un. Pour un canal donné, il est intéressant de voir pour lesquelles valeurs de $\underline{r}$, la probabilité d'erreur tend vers un. Apparemment pour les canaux satisfaisants converse fort, les débit plus grands que la capacité produisent la probabilité d'erreur égale à 1. Cela aboutit à la définition suivante qui sera utile pour la caractérisation de l'asymptotique EP.

Definition 10 *Considérons un canal multiterminal $\mathbb{W}^n$ avec m sources et destinations. Le région d'erreur complète est une région $\mathcal{S} \subset \mathbb{R}_+^{m(m-1)}$ telle que pour tous les codes $\big(n, M_n^{(ij)}, \epsilon_n\big)$, si le vecteur de débit $\underline{r} = \big(\liminf_{n\to\infty} \frac{1}{n} \log M_n^{(ij)}\big)$ est à l'intérieur de la région $\mathcal{S}$ alors la probabilité d'erreur tend vers un :*

$$\lim_{n\to\infty} \epsilon_n = 1.$$

La définition précédente est simplement la définition que la probabilité d'erreur sera égale à 1 pour tous les noeuds si le débit des codes appartient à cette région. Dans cette définition, la notion de région d'erreur complète a été définie pour le réseau entier. Il peut être aussi défini particulièrement pour une communication de point-à-point. Dans ce cas-là la région est déterminée par une valeur simple appelée *la capacité d'erreur complète* $\mathcal{S}$ qui est défini comme l'infimume de tous les débit pour lesquels chaque code avec un tel débit produira une probabilité d'erreur asymptotique 1.

En utilisant cette définition, le théorème suivant fournit les bornes sur la distribution de probabilités de l'erreur.

Theorem 20 *Pour le réseau multiterminal composite avec le paramètre aléatoire θ, nous avons :*

$$\mathbb{P}_\theta(\underline{r} \in \mathcal{S}_\theta) \leq \mathcal{E}(\underline{r}, \epsilon) \leq \inf_{\phi} \mathbb{P}_\theta(\underline{r} \notin \mathcal{R}_\theta(\phi)), \tag{4.25}$$

où $\mathcal{R}_\theta$ est la région atteignable de réseau $\mathbb{W}_\theta$ pour θ, et $\mathcal{S}_\theta$ est la région d'erreur complète pour ce canal pour θ.

Proof Pour prouver le théorème nous commençons de la définition du spectre asymptotique d'EP et en utilisant la convergence d'EP :

$$\begin{aligned}
\mathcal{E}(\underline{r}, \epsilon) &= \mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > \epsilon) \\
&= \mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > \epsilon \text{ et } \underline{r} \in \mathcal{S}_\theta) + \mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > \epsilon \text{ et } \underline{r} \notin \mathcal{S}_\theta) \\
&= \mathbb{P}_\theta(1 > \epsilon \text{ et } \underline{r} \in \mathcal{S}_\theta) + \mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > \epsilon \text{ et } \underline{r} \notin \mathcal{S}_\theta) \\
&\geq \mathbb{P}_\theta(\underline{r} \in \mathcal{S}_\theta)
\end{aligned}$$

La preuve de la partie suivante est aussi conclue du corollaire 8 en utilisant le fait que $\mathcal{C}_\theta$ est inclus dans $\mathcal{C}_{\epsilon,\theta}$.

D'une façon intéressante $\mathbb{P}_\theta(\underline{r} \notin \mathcal{R}_\theta(\phi))$ est la probabilité de panne. De nouveau on peut voir si le canal satisfait la condition de converse forte, c'est-à-dire $\mathcal{S}_\theta = \mathcal{C}_\theta$ et il a un unique meilleur code alors la probabilité d'erreur asymptotique sera égale à la probabilité de panne, qui soutiennent le sens opérationnel de cette notion.

Il y a des régions atteignables différentes disponibles pour les réseaux multiterminaux [18], mais il n'y a pas beaucoup de la région d'erreur complète. Nous essayons de fournir quelques résultats dans cette direction pour le cas de canaux multiterminaux sans mémoire discrète avec $W_\theta = \mathbb{P}_{Y_\theta^{(1)} ..., Y_\theta^{(m)} | X_\theta^{(1)} ..., X_\theta^{(m)}}$. Les débit atteignables différents peuvent être trouvés pour ces canaux qui fournissent la borne intérieure sur la distribution de probabilités selon le corollaire 8.

D'autre part la borne supérieure bien connu pour le réseau multiterminal est la borne de flot-max coupe-min [15, 47]. Cela déclare que n'importe quel débit à l'extérieur de la région formée par borne flot-max coupe-min aura EP non-zéro. Dans le théorème suivant, nous prouvons que l'erreur est nécessairement un pour n'importe quel débit à l'intérieur de cette région. Ce résultat fournit une borne supérieure sur la région d'erreur complète.

Maintenant nous nous concentrons sans perte de généralité sur le cas qu'un groupe de noeuds source $S \subset \{1, 2..., m\}$ envoyant des informations aux noeuds de destination S^c avec le vecteur de débit $\underline{r} = (r^{ij})_{i \in S, j \in S^c}$. La définition de débit atteignable sera limitée au cas où $i \in S$ et $j \in S^c$.

Theorem 21 *Considérons un canal multiterminal sans mémoire discrète avec* m *les noeuds. Pour tous les codes* $\left(n, leM_n^{(ij)}, \epsilon_n\right)$, *Supposons que les débit du code,* $\underline{r} = \left(\liminf_{n\to\infty} \frac{1}{n} \log M_n^{(ij)}\right)$, *tombent à l'extérieur de la fermeture suivante pour tous* $S \subset \{1, 2..., m\}$

$$\mathscr{S}_{CB} = co \bigcup_{P \in \mathcal{P}} \left\{ (R(S) \geq 0) : R(S) < I(X_S; Y_{S^c} | X_{S^c}) \right\}$$

où

$$R(S) = \sum_{i \in S, j \in S^c} R_{ij}.$$

Autrement dit, supposons que $\underline{r} \notin \mathscr{S}_{CB}$. *alors* $\lim_{n\to\infty} \epsilon_n = 1$.

Ce théorème implique que la borne de flot-max coupe-min fournit aussi une borne sur la région d'erreur complète. Effectivement toutes ces bornes dans la théorie d'information de réseau qui sont obtenues en utilisant la borne flot-max coupe-min, dans le même temps fournissent une borne sur à la région d'erreur complète.

En utilisant la borne de flot-max coupe-min et les régions atteignables disponibles comme ceux qui sont développées dans [18], on peut obtenir des bornes sur le spectre asymptotique de probabilités d'erreur. Pourtant nous ne supposons pas que tous les utilisateurs sont dans les mêmes temps les récepteurs et les émetteurs. Le canal est supposé composé des sources, relais et destinations comme dans Fig. 4.2. Supposons que chaque source $i \in \mathcal{T}$ envoie le message aux destinations dans l'ensemble $\mathcal{D}$ et tous les autres utilisateurs dans $\mathcal{R}$ sont des relais. La source i envoie le message avec le débit R_i à toutes les destinations. la borne flot-max coupe-min pour ce canal est caractérisé par :

$$\mathscr{S}^*_{\text{CB}} = co \bigcup_{P \in \mathcal{P}} \Big\{ (R(S) \geq 0) : R(S) < \min_{S \subseteq \mathcal{T}} \min_{S' \subseteq \mathcal{R}} \min_{d \in \mathcal{D}} I(X_S X_{S'}; Z_{S'^c} Y_d | X_{S^c} X_{S'^c}) \Big\}$$

où $S^c = \mathcal{T} - S$, $S'^c = \mathcal{R} - S'$.

D'autre part un borne intérieure a été développé pour ce canal en utilisant Comprimer-et-Transmettre comme la stratégie coopérative dans [18]. Le théorème suivant est la réaffirmation du théorème de Codage de Réseau Bruyant pour ce canal.

Theorem 22 (Lim-Kim-El Gamal-Chung [18]) *Une borne intérieure sur la région de capacité de réseau de DM avec les sources dans $\mathcal{T}$, les relais dans $\mathcal{R}$ et les destinations dans D est donné par*

$$\mathscr{R}_{IB} = co \bigcup_{P \in \mathcal{P}} \mathscr{R}_{NNC} \tag{4.26}$$

où

$$\mathscr{R}_{NNC} = \Big\{ (R(S) \geq 0) : R(S) < \min_{S \subseteq \mathcal{T}} \min_{S' \subseteq \mathcal{R}} \min_{d \in \mathcal{D}} I(X_S X_{S'}; \hat{Z}_{S'^c} Y_d | X_{S^c} X_{S'^c} Q) - I(Z(S'); \hat{Z}(S') | X(\mathcal{R}) X(\mathcal{T}) \hat{Z}(S'^c) Y_d Q) \Big\}$$

où $S^c = \mathcal{T} - S$, $S'^c = \mathcal{R} - S'$ et $R(S) = \sum_{k \in S} R_k$.

Prenons maintenant le réseau multiterminal composite avec le paramètre θ. Et supposons que les sources utilisent le schéma du Codage de Réseau Bruyant précédent pour la communication. Pourtant, à la différence des cas non-composites, les sources ne peuvent pas choisir la distribution de probabilités du canal P de $\mathcal{P}$ pour maximiser la région parce qu'ils ne sont pas conscients de θ. Donc la distribution de probabilités devrait être ramassée au préalable pour ne pas minimiser la probabilité de panne.

Maintenant les régions $\mathscr{R}_{\mathrm{NNC}}, \mathscr{S}^*_{\mathrm{CB}}$ peut être paramétrisé en utilisant θ comme $\mathscr{R}_{\mathrm{NNC},\theta}$ et $\mathscr{S}^*_{\mathrm{CB},\theta}$. Ces régions peuvent être exploitées pour fournir la borne suivante sur le spectre asymptotique d'EP en utilisant le théorème 20.

Corollary 9 *Le spectre asymptotique d'EP pour le débit $\underline{r}$ et chaque ϵ satisfait les limites suivantes :*

$$\mathbb{P}_\theta(\underline{r} \in \mathscr{S}^*_{CB,\theta}) \leq \mathcal{E}(\underline{r}, \epsilon) \leq \min_{P \in \mathcal{P}} \mathbb{P}_\theta(\underline{r} \notin \mathscr{R}_{NNC,\theta}). \tag{4.27}$$

Notons que la distribution de probabilité est choisie tel qu'il minimise la probabilité de panne. Le codage de réseau bruyant est une borne serrée pour le groupe de canaux. Pour le cas de réseau déterministe sans interférence [48] ou le cas de réseaux déterministes linéaires de tribu finis[9] $Y_k = \sum_{i=1}^m g_{ik} X_i$ [19], si nous choisissons $\hat{Z}_k = Z_k$ pour $k \in \{1..., m\}$ alors on peut voir que les bornes de codage de réseau bruyante sont serrées et coïncident avec la borne flot-max coupe-min. Pourtant c'est seulement pour le réseau déterministe linéaire de tribu fini que la valeur optimale est obtenue par la distribution de probabilités indépendante et uniforme .

Considérons maintenant un réseau déterministe linéaire de tribu fini composite où le canal en opération est choisi de l'ensemble des réseaux déterministes linéaires finis de terrain, indexés par $\theta \sim \mathbb{P}_\theta$. Chaque canal satisfait la converse fort et il y a une unique meilleur fonction d'encodage, c'est-à-dire une distribution de probabilité optimale unique pour tous les canaux. Alors la probabilité de panne sont le spectre asymptotique d'EP dans ce réseau et le corollaire suivant peut être obtenu.

Corollary 10 *Pour le réseau déterministe linéaire de tribu fini composite, le spectre asymptotique d'EP pour le débit $\underline{r}$ et chacun ϵ est comme suit :*

$$\mathcal{E}(\underline{r}, \epsilon) = \mathbb{P}_\theta(\underline{r} \notin \mathscr{C}_{DN,\theta}). \tag{4.28}$$

où $\mathscr{C}_{DN,\theta}$ pour θ est défini :

$$\mathscr{C}_{DN,\theta} = \Big\{(R(S) \geq 0) : R(S) < \min_{S \subseteq \mathcal{R}} \min_{d \in \mathcal{D}} H(Z_{S^c\theta} Y_{d\theta} | X_{\mathcal{T}} X_{S^c\theta})\Big\},$$

où la distribution d'entrée est choisie à chaque source comme indépendante et uniformément distribué.

Il est intéressant de voir que le côté droit de (10) est indépendant de ϵ qui signifie que la probabilité de panne est une mesure suffisante pour la performance de ce réseau.

9. Finite field linear deterministic networks

Part II

Cooperative Networks with Channel Uncertainty

Chapter 5

Introduction

Consider two distant points such that at one point, say the source there is an information, a set of messages, desired at the other end, the destination. The goal of communication is to guarantee this information at the destination with specific reliability criteria. The two ends can communicate via a channel. The channel "is the medium used to transmit the signal from transmitter to receiver [49]." The channel is in general a conditional probability distribution $\mathbb{W}(y|x)$ determining the probability that the channel output is y, from the alphabets $\mathcal{Y}$, if the channel input is chosen as x from the alphabets $\mathcal{X}$. The communication is attained via a code. A code consists of an encoding function that maps the messages to the channel input and a decoding function that maps the channel output to the set of messages. The channel is particularly known to the transmitter so it can choose the probability distribution of the channel input in the encoding process in order to increase the rate. Shannon formulated the problems of channel coding and derived main conditions for the transmission of a discrete memoryless source over a channel. Shannon in his canonical paper showed that "it is possible to send information at the rate C –the capacity– through the channel with as small a frequency of errors or equivocation desired by proper encoding. This statement is not true for any rate greater than C [49]." This rather unexpected result at that period confirmed the possibility of point-to-point transmission of information through a noisy channel with non-zero rate. The model considered in the seminal work by Shannon consists of a source-encoder, a channel and a destination-decoder.

Later the research within the discipline of information theory was pursued in various directions where in particular we can spot two out of them. First direction was taken by

Shannon himself who generalized the original setting to the multiuser case where there are many receivers and transmitters in the network. He introduced two-way communication networks in [50] where the communication is carried out in both direction between two nodes. Shannon in the very same paper stipulated about the cases with more than two terminals. He declared that:

> *In another paper we will discuss the case of a channel with two or more terminals having inputs only and one terminal with an output only, a case for which a complete and simple solution of the capacity region has been found [50].*

This promised the capacity of multiple access channels. Unfortunately, Shannon did not publish anything afterward and "the simple solution" took almost a decade to be ready. Ahlswede [51] and Liao [52] around 1971 found the capacity of multiple access channels. The various models like broadcast channels [24] and relay channels were later introduced but it turned out that finding the capacity of general multiterminal networks is a laborious and difficult task. Cover in 1975 [53] conjectured that "it seems very likely that the capacity region of broadcast channel will be information theoretic and pretty". The task appeared to be much more challenging than once seemed to and Ahlswede in 1987 [54] qualified it as "one of the very challenging problems". The problem of capacity region of broadcast channel is still open.

In particular, there is a special multiterminal model which is very attracting because its practical application. Consider a three terminal network where one terminal has neither a message to transmit nor a message to receive. All this terminal can do is to find a best way to facilitate the communication between the other terminals. This channel was firstly introduced by Van der Meulen around 1971 in [55]. A special case of three terminal networks is the relay channel. If we add another node to the basic Shannon model for communication which is capable of transmitting and receiving the information we obtain the relay channel, see Fig. 5.1. This model can be considered as a basic model for cooperative communication, however the capacity region of this channel is still unknown in general. The various achievable rates and upper bounds have been developed which yields the capacity region for some special cases [1–3, 5, 14]. The cooperation between terminals opens up the possibility to increase the rate and reliability in multiterminal networks.

On the other hand, a second direction taken by researcher was to consider the uncertainty in the communication model. This aims to address scenarios where the terminals are

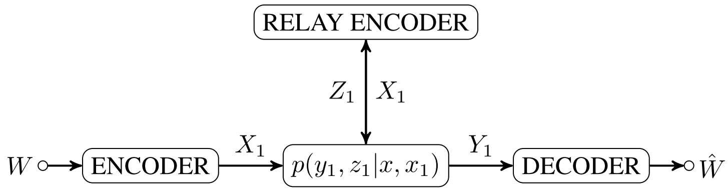


Figure 5.1: The memoryless relay channel

not aware of the exact probability distribution of the channel. Arbitrary varying channels (AVCs), time-varying channels with states and compound channels are the main models introduced to deal with uncertainty. The compound channel, also referred to as the simultaneous channel, is a model where the probability distribution (PD) of the channel in operation is chosen from a set of probability distributions Θ and remains fixed throughout the transmission. Both transmitter and receiver are assumed ignorant of the PD governing the transmission but they know the set of possible channel indices Θ. The compound channel and its capacity was derived in [20–22] in 1959.

What makes these two directions particularly interesting is the rise of wireless networks. Wireless networks occupy nowadays an undeniable place in the telecommunication industry and it is important to analyze their various facets. They possess some specificities. First of all, they are composed of many receivers and transmitters and in this sense they are a particular case of multiterminal networks. They are composed of many nodes which can be receivers, transmitters or both and there are set of messages destined for some nodes coming from other nodes. Secondly, the nodes can help each other, e.g. by forwarding the message of other users through the network. It means that each node can choose the transmitted code as a function of its own message and the previous observation of the channel. This opens up the possibility of cooperation within a network. Finally, the channel in wireless networks is subject to changes due to fading and mobility of users, which necessitates to consider the uncertainty inherent in the structure of these networks.

It is around these three axes, i.e. multiterminal networks, cooperation and uncertainty that this thesis is organized.

5.1 Related Works

A considerable amount of research was carried out on network information theory, cooperative networks and communication with channel uncertainty. We now review basic results and related work.

5.1.1 Cooperation and Multiterminal Networks

The essence of cooperation is relaying operation. As it can be seen from Fig. 5.1, the relay channel consists of the channel input $X \in \mathscr{X}$ and the relay input $X_1 \in \mathscr{X}_1$, the channel output $Y_1 \in \mathscr{Y}_1$ and the relay output $Z_1 \in \mathscr{Z}_1$. The channel is characterized by $\mathbb{W}(y_1, z_1|x, x_1)$ and it is assumed to be memoryless:

$$\mathbb{W}(\underline{y}_1, \underline{z}_1|\underline{x}, \underline{x}_1) = \prod_{i=1}^{n} \mathbb{W}(y_{1i}, z_{1i}|x_i, x_{1i})$$

for $\underline{x} = (x_1, x_2, ..., x_n)$, where x_i means the channel input at the time i. The relay input X_1 at the time i is a function of previous relay outputs Z_1, namely $X_{1i} = f_i(Z_1^{i-1})$. The central difficulty in this problem is to find a proper relay function.

The original contribution for this channel is due to Cover and El Gamal [1]. They developed the main cooperative strategies for relay channels, namely Decode-and-Forward (DF) scheme and Compress-and-Forward (CF) scheme. In DF coding, as first introduced by Cover-El Gamal, the source messages are distributed into indexed bins. The relay decodes the source message and then transmit its bin index. The achievable rate for DF scheme is given by

$$R_{\text{DF}} = \max_{p(x,x_1)} \min \left\{ I(X; Z_1|X_1), I(XX_1; Y_1) \right\}.$$

The previous rate is indeed the combination of two conditions. The relay has to decode the message in DF and the first condition corresponds to the condition of successful decoding at the relay $R \leq I(X; Z_1|X_1)$. The next condition is the condition of successful decoding at the destination $R \leq I(XX_1; Y_1)$. It is interesting to see that intuitively the destination sees a multiple access channel with two inputs X, X_1, jointly delivering a same message and not necessarily independent, and the rate is related to this setting.

On the other hand, when CF scheme is allowed, the relay finds a compressed version of its output observation, namely $\hat{Z}_1$, and using binning the compressed version is then transmitted. The achievable rate for CF is as follows

$$R_{\text{CF}} = \max_{p(x)p(x_1)p(\hat{z}_1|z_1,x_1)} I(X; Y_1\hat{Z}_1|X_1)$$

subject to the condition

$$I(X_1; Y_1) \geq I(Z_1; \hat{Z}_1|X_1, Y_1).$$

In fact, DF and CF schemes are the fundamental cooperative strategies developed for the relay channel. Another achievable rate region was obtained by Cover-El Gamal by combining DF-CF schemes, where the relay uses both DF and CF. As a special case, the relay use DF to decode and forward only part of the source message and the rest of the message is directly transmitted to the destination. It is referred to as partial decoding DF scheme. As a matter of fact, DF and CF regions can be obtained by different methods. For instance DF region can be obtained, using methods developed by Willems and Carleial in [2, 3], where instead of using binning, the source and the relay uses the codebooks with the same size. This is called regular encoding. Willems developed backward decoding and Carleial used sliding window to decode the message at the destination. El Gamal-Mohseni-Zahedi in [4] developed an alternative scheme for CF where the achievable rate turns out to be equivalent to final CF rate (for detailed survey on the results look at the introduction of the next chapter).

Although in general the previous bounds are not tight, it was shown that DF scheme achieves the capacity of *physically degraded* and *reversely degraded* relay channels. The degraded relay channel is defined with the following Markov chain $X \multimap (X_1, Z_1) \multimap Y_1$. The notion of degradedness of Y_1 with respect to Z_1 implies intuitively that Z_1 is in general better than Y_1. The notion appears also in other channels, .e.g. broadcast channels. Particularly, when there is noiseless feedback from the destination to the relay, the relay channel can be considered as physically degraded and the capacity is achieved using DF scheme. On the other hand, partial DF scheme yields the capacity of semi-deterministic relay channels, as it was shown by Aref-El Gamal [5].

Another important piece of networks is the broadcast channel (BC), where a set of common and private messages are intended to several destinations. In particular, the 2-user memoryless BC, characterized by $\mathbb{W}(y_1, y_2|x)$, has been extensively studied. The capacity region of the degraded BC was found by Bergmans, Gallager, and Ahlswede and Korner [6–9]. Korner and Marton established the capacity of the BC with degraded message sets [10]. They introduced the notions of less-noisy and more capable BCs [11] and showed the capacity of less-noisy BCs. El Gamal proved the capacity of more capable BCs in [4]. The best known inner bound for the general BC is due to Marton [12]. This is based on the idea of *binning* where an alternative proof was also reported by El Gamal and van der Meulen in [13]. The next region is referred to as the Marton region.

$$\begin{aligned}\mathscr{R}_{\text{BC}} = co\Big\{&(R_1, R_2) : R_1, R_2 \geq 0,\\ &R_1 \leq I(U_0U_1;Y_1)\\ &R_2 \leq I(U_0U_2;Y_2)\\ &R_1 + R_2 \leq \min\{I(U_0;Y_1), I(U_0;Y_1)\}\\ &\quad + I(U_1;Y_1|U_0) + I(U_2;Y_2|U_0) - I(U_2;U_1|U_0) \text{ for all } \mathbb{P}_{U_0U_1U_2X} \in \mathcal{P}\Big\},\end{aligned}$$

where $\mathcal{P}$ is the set of all PDs $\mathbb{P}_{U_0U_1U_2X}$.

Extensive research has been done during years in order to study the capacity region of more general networks by combining single relay channels, broadcast channels and multiple access channels. For instance, broadcast relay channels and multiple access relay channels, along with general networks [14] were studied. Achievable rate regions have been derived by combining coding techniques like: (partial) DF and CF schemes, Marton and superposition coding, block-Markov coding, etc. However as the capacity region is not known for most of the basic networks like the relay channel and the broadcast channel, the obtained achievable rates are not tight in general.

The research on general networks fascinated the researchers from the beginning. Elias-Feinstein-Shannon around 1956 stated a general upper bound on the capacity of multiterminal networks [15].

> *Theorem (Elias-Feinstein-Shannon'56):* The maximum possible flow from left to right through a network is equal to the minimum value among all simple cut-sets.

The proof of this theorem was also given by Ford-Fulkerson in [16] and Dantzig-Fulkerson [17]. Moreover the authors clearly stated that it is "by no means obvious" if this region for general networks can be achieved. Assume now a network with N pair users (X_i, Y_i), for $i \in \mathcal{N} = \{1, 2, ..., N\}$ and channel $\mathbb{W}(y_1, y_2, ..., y_N|x_1, x_2, ..., x_N)$. Then the cut-set bound is as follows ($R(S) = \sum_{k\in S} R_k$):

$$\mathscr{R}_{\text{CB}} = co \bigcup_{P\in\mathcal{P}} \Big\{(R(S) \geq 0) : R(S) < I(X(S); Y(S)|X(S^c)) \text{ for all } S \subseteq \mathcal{N}\Big\}.$$

As a matter of fact, the region as mentioned before is in general not achievable. Furthermore, it is difficult to generalize coding for broadcast and relay channels to arbitrary networks. However, in recent work [18] by Lim-Kim-El Gamal-Chung, noisy network coding (NNC) strategy was introduced for general networks. This is based on CF scheme, achieving a constant gap from the cut-set bound.

Theorem 23 *(Lim-Kim-El Gamal-Chung, 2011) An inner bound on the capacity region of DM network with N user and the set of destinations D is given by*

$$\begin{aligned}\mathscr{R}_{NNC} =& co \bigcup_{P \in \mathscr{Q}} \Big\{ (R(S) \geq 0) : \\ & R(S) < \min_{d \in S^c \cap \mathcal{D}} I(X(S); \hat{Y}(S^c), Y_d | X(S^c), Q) - I(Y(S); \hat{Y}(S) | X^N, \hat{Y}(S^c), Y_d, Q) \Big\}\end{aligned}$$

for all cutsets $S \subset \mathcal{N}$ with $S^c \cap \mathcal{D} = \emptyset$.

NNC is based on sending the same message in all blocks –repetitive encoding– and non-unique compression index decoding where compression does not rely on binning. The NNC scheme achieves the capacity of some networks, as for example finite field linear deterministic networks [19].

5.1.2 Compound and Composite Channels

The problem of communication with channel uncertainty has been studied via various models however the main assumption is that the channel is unknown to the terminals. Either the channel changes arbitrarily during each round of transmission or the channel remains fixed during the course of transmission. In the former case we face with channels with states while in the later with the compound channel. These models roughly correspond to the cases of fast fading and slow fading in wireless communications. In fast fading setting, the code-length is significantly larger than the coherence time interval of the channel and hence the channel can be assumed to be ergodic. This yields the Shannon capacity in this case is called ergodic capacity. For slow fading case, the code length is in the order of the coherence time for the corresponding interval. Although there exist various strategies developed for these scenarios, our focus is to look at the compound capacity.

The compound channel consists in a set of channels indexed by θ:

$$\mathcal{W}_{\Theta} = \{ \mathbb{W}_{\theta}(y|x) : \mathcal{X} \longmapsto \mathcal{Y} \}_{\theta \in \Theta} .$$

It is important to note that there is no probability distribution assumed on Θ. Moreover in order to have an achievable rate for the compound channel, the code should yield small error probability for every θ. The capacity of the compound channel (CC) is given by [20–22]

$$C_{\mathrm{CC}} = \max_{p(x)} \inf_{\theta\in\Theta} I(X;Y_\theta),$$

where Y_θ is the output of the channel with distribution $\mathbb{W}_\theta(y|x)$. However in the case of a slow-fading AWGN channel, one cannot guarantee small error probability for all possible channel gains because ultimately one gets zero rate. Assume now that instead of a single set of messages, the encoder is allowed to send several set of messages –variable-rate channel coding [23]– and then the destination, depending on the index θ, decodes as much as possible messages. The existent connection between broadcast and compound channels was first noticed by Cover in [24, Section IX], where he suggested "that the compound channels problem can be reinvestigated from this broadcasting point of view." This idea was fully developed by Shamai in [25] and later referred to as the broadcasting approach.

Consider the Gaussian fading AWGN channel defined as

$$Y = hX + \mathcal{N},$$

where $\mathcal{N}$ is AWGN noise and h is fading coefficient. The channel is assumed to be slowly fading, which means that h is chosen randomly beforehand and remains constant during the communication. Here the uncertainty comes from the fading coefficient h, and for each draw of h, there is a channel which can be possibly in operation. The set of all possible channels can be considered as being indexed by h, i.e. $\theta = h$. Now "the transmitter views the fading channel as a degraded Gaussian broadcast channel with a continuum of receivers each experiencing a different signal-to-noise ratio specified by uSNR where u is the continuous index". Shamai constructed a multi-layered coding, one layer for each fading draw h, such that for each fading draw h, all the layers with $u = |h'|^2 \leq v = |h|^2$ can be decoded and the rest of layers appear as interference. The power allocation for v is $\mathrm{SNR}(v)dv \geq 0$. The rate for this channel is a function of v and follows as

$$R(v) = \int_0^v \frac{-u dy(u)}{1 + uy(u)},$$

where $y(u) = \int_v^\infty \mathrm{SNR}(v)dv$. The main idea behind broadcasting strategy is to send different messages so that the destination can choose how many of them can be decoded depending on the channel in operation. In the broadcasting strategy, the transmitted code guarantees variable rates for each of the possible channels in the set.

There are other approaches to deal with uncertainty in networks. In compound settings, there is no probability distribution introduced over θ. To take into account the probability distribution of θ, the notion of outage capacity was proposed in [26] for fading channels. For a desired outage probability p, the outage capacity is defined as the maximum rate that can be transmitted with probability $1 - p$. In contrast, ergodic capacity is the maximum information rate for which error probability decays exponentially with the code length. Unlike broadcasting strategy, the transmitted code sends a fixed rate for all possible channels in the set. Effros-Goldsmith-Liang introduced the composite channel [27]. "A composite channel consists of a collection of different channels with a distribution characterizing the probability that each channel is in operation." So the composite channel is defined as the set of channels $\mathcal{W}_\Theta$ as before, but with an associated PD $\mathbb{P}_\theta$ on the channel index θ. Composite models unlike compound, take into account the channel uncertainty by introducing a PD $\mathbb{P}_\theta$ on the set. The authors in [27] broaden the definition of capacity to allow for some outage. Indeed, the notion of *capacity versus outage* is defined as "the highest rate asymptotically achievable with a given probability of decoder-recognized outage".

5.1.3 Cooperative Networks with Channel Uncertainty

The uncertainty in general networks can be both due to fading and user mobility. Now suppose that the relay can be present or absent and the source is oblivious to this fact. Moreover the topology of network and the channel itself remains fixed during the course of communication. Then the source should be able to design a code to sustain the performance in spite of the absence of the relay. Katz-Shamai studied this problem in [28] with an occasional nearby relay, as shown in Fig. 5.2. By nearby relay it is meant that DF scheme performs better than CF scheme. They used the notion of expected throughput to gauge the performance. It was shown that the superposition coding-backward decoding allows the destination to decode the message without performance loss, even if the relay is not present. In other words, if the relay is not present the expected throughput remains the same. The authors introduced the notion of *oblivious cooperation* to refer to cooperative protocols which improve performance when the relay is present while not degrading it when the relay is absent, even if the source is uninformed of the actual topology.

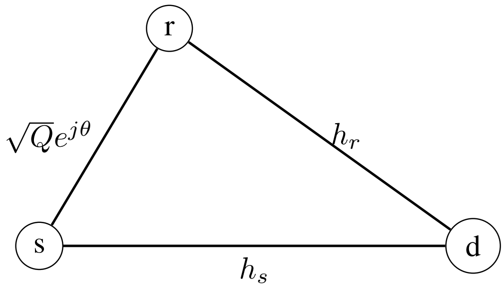


Figure 5.2: A slow-fading AWGN relay channel

5.2 Motivation

It is well recognized how wireless networks are subject to statistical changes, mainly due to fading and user mobility. In some scenarios, the code-length is significantly smaller than the coherence time interval so that the channel remains fixed during the communication. Let us consider for instance a single relay channel where the source-relay and relay-destination channel gains can randomly change. If the source-relay channel gain is good enough then the relay would be able to decode the source message and forward it to the destination. It may appear to be better to implement DF scheme in this case. However, if the channel is randomly drawn, the source-relay channel quality can be significantly deteriorated for some cases and it is not guaranteed that the relay can decode the message successfully. In these cases, an error is declared and decoding cannot be successfully carried out. Moreover the source cannot do anything because it is not aware of the channel realization. In slowly fading AWGN channels, the source-relay channel quality, which is fix during the transmission, can be poor through the whole course of the communication and this yields unsuccessful decoding at the destination. A similar situation can be considered if the source employs CF code. When the relay-destination channel is poor, the use of CF scheme is not adequate and thus decoding may be unsuccessful again. As a matter of fact, even if the source-relay channel is good enough for allowing decoding of the massage at the relay, it cannot fully exploit DF scheme because of the source code is designed for CF coding, and so independent of the relay code. Notice that in these examples the relay can have access, at least partly, to the channel state information (CSI) because it has a receiver and so it can have an estimation of the channel. But because the source has to fix the coding *a priori*, a cooperative strategy is imposed at relay and therefore it is not capable of benefiting from the available CSI.

The above mentioned problems are central in multiterminal networks with channel uncertainty. The main problem is that the proper coding (e.g. cooperative strategy) depends on the source-relay and the relay-destination channel quality. So it is desirable to explore how opportunistic and/or adaptive coding for cooperation is possible, even if the source is ignorant of the channel states. In other words, how the users that are partly or fully aware of CSI can exploit their available information to provide a better cooperation performance. We shall refer to these strategies to as oblivious strategies, which means that the source is oblivious to the coding strategy deployed in the other terminals. An example of an oblivious strategy was given before when the source does not know whether the relay is present or not but knows that if the relay is present then it is collocated. It was discussed that if superposition coding is allowed at the source, then the destination can decode the message subject to constraint of the single user channel capacity, even if the relay is not present. So it can be said that superposition coding is an oblivious code with respect to the presence of the relay.

In this thesis, we investigate cooperative strategies with channel uncertainty. In particular, we are interested in two cases. First, the case of simultaneous relay channels which consists of a set of relay channels and secondly, the case of composite models where the channel in operation is selected from the set of channels indexed with θ by following a PD $\mathbb{P}_\theta$. In this setting, the source (or sources) is ignorant of the channel in operation indexed by θ. The other terminals are partly or fully aware of the channel in operation. As we mentioned, channel uncertainty can be studied based on compound models and via the broadcasting approach, or by relaying on the notion of the outage capacity. In the same direction, we will see how these approaches can help us to understand better fundamental limits and novel coding schemes for cooperative networks with channel uncertainty.

5.3 Summary of Contributions

The contribution of this thesis is organized in three chapters:

– Cooperative Strategies for Simultaneous and Broadcast Relay Channels,

– Selective Coding Strategy for Composite Unicast Networks,

– On the Asymptotic Spectrum of the Error Probability of Composite Networks.

In the first chapter, cooperative strategies are developed for *simultaneous relay channel* (SRC), which consists of a set of single relay channels out of which the channel in operation is chosen. The broadcasting approach is adopted for this channel where the source wishes to transmit common and private information to each of possible channels. This opens up the possibility of using the broadcasting approach to send messages to each channel. For example, suppose that the relay uses either DF or CF scheme but it is always present. Now although the source may be ignorant of the coding strategy at the relay, it knows that it is either DF or CF scheme which yields two possibilities. Then the source can design a code with three messages (W_0, W_1, W_2) such that (W_0, W_1) is decoded when the relay uses DF and (W_0, W_2) when CF is allowed. Therefore, this problem is recognized as being equivalent to that of sending common and private information to several destinations in presence of helper relays where each channel outcome becomes a branch of the *broadcast relay channel* (BRC). Cooperative schemes and capacity region for a set of two relay channels are investigated. The proposed coding strategies must be capable of transmitting information simultaneously to all destinations in such set. Inner bounds on the capacity region of the general BRC are derived for three cases of particular interest:

- The channels from source-to-relays of both destinations are assumed to be stronger [1] than the others and hence cooperation is based on DF strategy for both users (referred to as DF-DF region),
- The channels from relay-to-destination of both destinations are assumed to be stronger than the others and hence cooperation is based on CF strategy for both users (referred to as CF-CF region),
- The channel from source-to-relay of one destination is assumed to be stronger than the others while for the other one is the channel from relay-to-destination and hence cooperation is based on DF strategy for one destination and CF for the other one (referred to as DF-CF region).

The techniques used to derive the inner bounds rely on recombination of message bits and various effective coding strategies for relay and broadcast channels. These results can be seen as a generalization and hence unification of previous work in this topic. An outer bound on the capacity region of the general BRC is also derived. Capacity results are obtained for specific cases of semi-degraded and degraded Gaussian simultaneous relay channels. Rate regions are computed for AWGN models.

In the second chapter, the composite relay channel is considered where the channel is randomly drawn from a set of conditional distributions with index $\theta \in \Theta$, which represents

1. As it will be mentioned later, the formal definition of the notion of *stronger channel* is not necessary until converse proofs.

the vector of channel parameters with PD $\mathbb{P}_\theta$ characterizing the probability that each channel is in operation. The specific draw θ is assumed to be unknown at the source, fully known at the destination and only partly known at the relay. In this setting, the transmission rate is fixed regardless of the current channel index. Hence the encoder cannot necessarily guarantee arbitrary small error probability for all channels. The asymptotic error probability is used as metric to characterize the reliability function. In this setting, the coding strategy is commonly chosen regardless of the channel measurement at the relay end. We introduce a novel coding which enables the relay to select –based on its channel measurement– the best coding scheme between CF and DF schemes. Indeed, provided that the channel source-to-relay is good enough for decoding the message, the relay decides on DF and otherwise it may switch to CF. The proposed selective coding strategy (SCS) is based on superposition coding, DF and CF coding schemes, backward and joint decoding at the destination. We derive bounds on the asymptotic average error probability of the memoryless relay channel. This result is later extended to the case of unicast composite networks with multiple relays. As a consequence of this, we generalize the NNC theorem for the case of unicast networks where the relays are divided between those who use DF scheme and those based on CF scheme. It is also shown that from offset coding the relays using DF scheme can exploit the help of those using CF scheme. Offset coding is the name for kind of Decode-and-Forward scheme where the relays in the block i of transmission are not necessarily transmitting the source message from the previous block and each can defer its retransmission. For instance the relay in block i, sends the message of the block $i-2$. An application example to the case of fading Gaussian relay channel is also investigated where it is demonstrated that SCS clearly outperforms the well-known DF and CF schemes.

The third chapter is dedicated to some theoretical considerations about the composite multiterminal networks. As we already mentioned before, arbitrary small error probability cannot be guaranteed for all channels in the set. Here instead of finding the maximum achievable rate subject to a small error probability (EP), we look at the behavior of error probability (not necessarily zero) for a given rate. The common notion as performance measure of composite network is the notion of outage probability. But it is seen in case of composite binary symmetric averaged channel, for which ϵ-capacity is known, the outage probability is not enough precise as performance measure. Instead, various notions of performance are discussed among which the asymptotic spectrum of error probability is introduced as a novel performance measure for composite networks. It is shown that the behavior of EP is directly related to their ϵ-capacity. For instance, the notion of asymptotic spectrum of EP yields a more general measure for the performance of these networks. The asymptotic spectrum of EP is bounded by using available achievable rate regions and a new region called *full error region*. Every code with a rate belonging to this region yields EP equal to one. In this sense, for the networks satisfying the strong converse condition, asymptotic spectrum of EP coincides with the outage probability. To this purpose it is shown that the cutset bound provides an upper bound for the full error region.

Chapter 6

Cooperative Strategies for Simultaneous Relay Channels

6.1 Introduction

The simultaneous relay channel (SRC) is defined by a set of relay channels, where the source wishes to communicate common and private information to each of the destinations in the set. In order to send common information regardless of the actual channel, the source must simultaneously consider all the channels as described in Fig. 6.1(a). The described scenario offers a perspective of practical applications, as for example, downlink communication on cellular networks where the base station (source) may be aided by relays, or on ad-hoc networks where the source may not be aware of the presence of a nearby relay (e.g. opportunistic cooperation).

Cooperative networks have been of huge interest during recent years between researchers as a possible candidate for future wireless networks [56–58]. Using the multiplicity of information in nodes, provided by the appropriate coding strategy, these networks can increase capacity and reliability. Diversity in cooperative networks has been assessed in [59–61] where multiple relays were introduced as an antenna array using distributed space-time coding. The advantage of cooperative MIMO over point-to-point multiple-antenna systems was analyzed in [62]. Also coded cooperation has been assessed in [63].

The simplest of cooperative networks is the relay channel. First introduced in [55], it consists of a sender-receiver pair whose communication is aided by a relay node. In other words, it consists of a channel input X, a relay input X_1, a channel output Y_1 and a relay output Z_1, where the relay input depends only on the past relay observations. The significant contribution was made by Cover and El Gamal [1], where the main strategies of Decode-and-Forward (DF) and Compress-and-Forward (CF), and a max-flow min-cut upper bound were developed for this channel. Moreover the capacity of the degraded and the reversely degraded relay channel were established by the authors. A general theorem that combines DF and CF in a single coding scheme was also presented. In general, the performance of DF and CF schemes are directly related to the noise condition between the relay and the destination. More precisely, it is well-known that DF scheme performs much better than CF when the source-to-relay channel is of high quality. Whereas, in contrast, CF is more suitable when the relay-to-destination channel is better. Furthermore, these schemes provide rates that are very close to the cut-set bound for Gaussian relay channels when the quality of one of these channels are good.

DF and CF inner bounds can be obtained using different coding and decoding techniques. Coding techniques can be classified [14] into *regular and irregular coding*. Irregular coding exploits the codebooks, which are involved between relay and source, with different sizes while regular coding requires the same size. Decoding techniques also can roughly be classified into *successive and simultaneous decoding*. Successive decoding method decodes the transmitted codebooks in a consecutive manner. In each block, it starts with a group of codebooks (e.g. relay codebook) and then afterward, it moves to the next group (e.g. source codebook). However the simultaneous decoding decodes all the codebooks in a given block at the same time.

Generally speaking, the latter provides the better results than the former. Cover and El Gamal [1] have proposed irregular coding with successive decoding. Regular coding with simultaneous decoding was first developed in [64]. It can be exploited for decoding the channel outputs of a single or various blocks. For instance, the author in [2] has exploited this issue by decoding with the channel outputs of two consecutive blocks which is called *sliding window decoding*. The notion of *backward decoding*, which was introduced in [3], consists of a decoder who waits until the last block and then starts to decode from the last to the first message. Backward coding is shown to provide better performances than other schemes based on simultaneous decoding [65, 66] such as sliding window. Backward decoding can use a single block like in [3] or multiple blocks as in [28] to decode the information. The best lower bound known for the relay channel was derived in [67] by using a generalized backward decoding strategy.

Based on these strategies, further work has been recently done on cooperative networks from different aspects. The capacity of semi-deterministic relay channels and the capacity of cascaded relay channels were found in [5,68]. A converse for the relay channel has been developed in [69]. Recently Compute-and-Forward strategy was proposed in [70] where Lattice coding is used to perform coding. It appears that the use of structured codes outperforms Decode-and-Forward strategy in some settings.

Multiple relay networks have been studied in [71] and practical scenarios have been also considered, like Gaussian relay channel [40,72,73], Gaussian parallel relay network [74–78], wireless relay channel and resource allocation [79–82]. The capacity of orthogonal relay channels was found in [83] while the relay channel with private messages was discussed in [84]. The capacity of a class of modulo-Sum relay channels was also found in [85]. The combination of relay channel with other networks has been studied in various papers, like multiple access relay, broadcast relay and multiple relays, fading relay channels. The multiple access relay channel (MARC) was analyzed in [86–88]. Offset decoding for MARC has been proposed in [89] to improve the sliding window rate while avoiding the problem of delay in the backward decoding. The relay-broadcast channel (RBC) where a user, which can be either the receiver or an distinct node, serves as a relay for transmitting the information to the receivers was also studied. An achievable rate region for the dedicated RBC was obtained in [14]. Preliminary works on the cooperative RBC were done in [90–92] and the capacity region of physically degraded cooperative RBC was found in [93]. Rate regions and upper bound for the cooperative RBC were developed further in [30, 94, 95]. The capacity of Gaussian dedicated RBC with degraded relay channel was presented in [33]. The simultaneous relay channel was also investigated through broadcast channels in [29, 96, 97].

An interesting relation between compound and broadcast channels was first mentioned in [24]. Indeed, the concept of broadcasting has been used as method for mitigating the channel uncertainty effect in numerous papers [25, 28, 98–100]. This strategy facilitates to adapt the rate to the actual channel in operation without having any feedback link to the transmitter. Extensive research has been done on compound channels [21, 101], including *Zero-Error* [102], side information [103], interference channels [104], MIMO [105], finite-states [106], multiple-access channel [107], feedback capacity [108], binary codes [109] and degraded MIMO broadcast channel [110]. The broadcast channel (BC) was introduced in [24] along with the capacity of binary symmetric, product, push-to- talk and orthogonal BCs. The capacity of the degraded BC was established in [6–9]. It was shown that feedback does not increase capacity of degraded BCs [111, 112] but it does for Gaussian BCs [113].

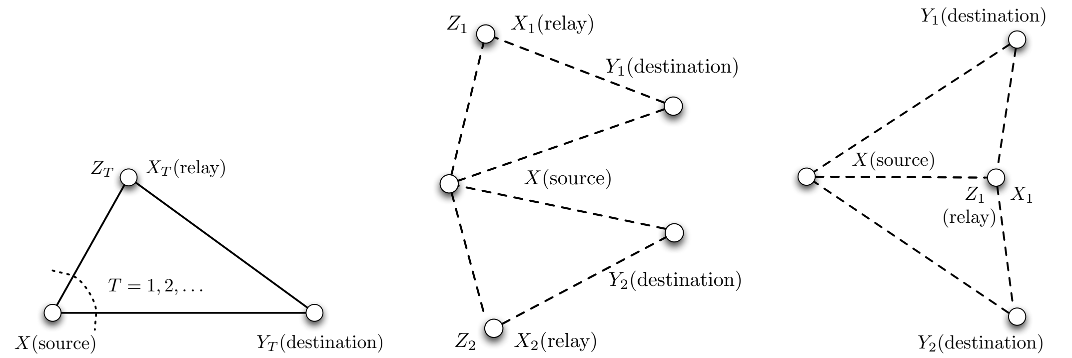
The capacity of the BC with degraded message sets was found in [10] while that of more capable and less-noisy were established in [4]. The best known inner bound for general BCs is due to Marton [12] and an alternative proof was given in [13] (see [114] and reference therein). Such bound is tight for channels with one deterministic component [115] and deterministic channels [116, 117]. Lately, another strategy called indirect decoding was introduced in [118, 119], which achieves the capacity of 3-receiver BC with two degraded message sets. A converse for the general BC was established in [12] and improved later in [31, 120].

The problem of the simultaneous relay channel is equivalent to that of the broadcast relay channel (BRC), with additional Markov chains. The source sends common and private information to several destinations which are aided by their own relays. So the problem of SRC can be studied using the problem of BRC.

In this chapter, we study different coding strategies and capacity region for the case of a general BRC with two relays and destinations, as shown in Fig. 6.1(b), as an equivalent model for SRC with two simultaneous memoryless relay channels. Note that each model introduced for BRC can be considered as an equivalent model for the SRC by adding proper Markov chains however we do not explicitly assert the Markov chains for the rest. The rest of the chapter is organized as follows. Section II presents main definitions and the problem statement. Inner bounds on the capacity region are derived for three cases of particular interest:

– The channels from source-to-relays are stronger [1] than the others and hence cooperation is based on DF strategy for both users (refer to as DF-DF region). This case corresponds to the SRC where DF is employed in both relay channels,
– The channels from relay-to-destination are stronger than the others and hence cooperation is based on CF strategy for both users (refer to as CF-CF region). This case corresponds to the SRC where CF is employed in both relay channels,
– The channel from source-to-relay of one destination is stronger than the others while for the other one is the channel from relay-to-destination and hence cooperation is based on DF strategy for one destination and CF for the other (refer to as DF-CF region). This case corresponds to the SRC where different coding, CF and DF, is

1. We shall not provide any formal definition to the notion of *stronger channel* since this is not necessary until converse proofs. However the operational meaning of this notion is that if channel A is assumed to stronger than channel B then the coding scheme will assume that decoder A can fully decode the information intended to decoder B.



(a) Simultaneous relay channel (SRC) (b) BRC with two relays (c) BRC with common relay

Figure 6.1: Simultaneous and broadcast relay channels

employed in each relay channel.

Section III examines general outer bounds and capacity results for several classes of BRCs. In particular, the case of the broadcast relay channel with common relay (BRC-CR) is investigated, as shown in Fig. 6.1(c). We show that the DF-DF region improves existent results on BRC with common relay, previously found in [14]. Capacity results are obtained for the specific cases of semi-degraded and degraded Gaussian simultaneous relay channels. In Section IV, rates are computed for the case of distant based additive white Gaussian noise (AWGN) relay channels. Achievability and converse proofs are relegated to the appendices. Finally, summarize and discussions are given in Section V.

6.2 Main Definitions and Achievable Regions

In this section, we first formalize the problem of the simultaneous relay channel and then the next three subsections present achievable rate regions for the cases of DF-DF strategy (DF-DF region), CF-CF strategy (CF-CF region) and DF-CF strategy (DF-CF region). We denote random variables by upper case letters X, Y and by bold letters $\mathbf{X}, \mathbf{Y}$ the sequence of n random variables, i.e. X^n, Y^n. On the other hand the Markov chain between three random variables A, B and C is presented using the following notation:

$$A \multimap B \multimap C.$$

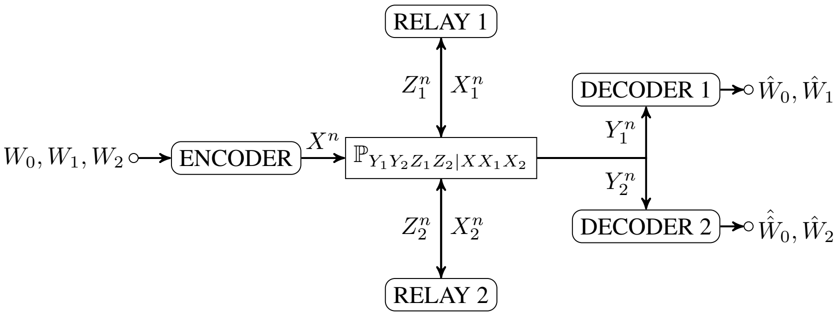


Figure 6.2: Broadcast relay channel (BRC)

6.2.1 Problem Statement

The simultaneous relay channel [29] with discrete source and relay inputs $x \in \mathscr{X}$, $x_T \in \mathscr{X}_T$, discrete channel and relay outputs $y_T \in \mathscr{Y}_T$, $z_T \in \mathscr{Z}_T$, is characterized by a set of relay channels, each of them defined by a conditional probability distribution (PD)

$$\mathscr{P}_{SRC} = \left\{ P_{Y_T Z_T | X X_T} : \mathscr{X} \times \mathscr{X}_T \longmapsto \mathscr{Y}_T \times \mathscr{Z}_T \right\},$$

where T denotes the channel index. The SRC models the situation in which only one single channel is present at once, and it does not change during the communication. However the transmitter (source) is not cognizant of the realization of T governing the communication. In this setting, T is assumed to be known at the destination and the relay ends. The transition PD of the n-memoryless extension with inputs $(\mathbf{x}, \mathbf{x}_T)$ and outputs $(\mathbf{y}_T, \mathbf{z}_T)$ is given by

$$P^n_{Y_T Z_T | X X_T}(\mathbf{y}_T, \mathbf{z}_T | \mathbf{x}, \mathbf{x}_T) = \prod_{i=1}^{n} W_T(y_{T,i}, z_{T,i} | x_i, x_{T,i}).$$

Here we focus on the case where $T \in \{1, 2\}$, in other words there are two relay channels in the set.

Definition 11 (Code) *A code for the SRC consists of*

– *An encoder mapping* $\{\varphi : \mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}_2 \longmapsto \mathscr{X}^n\}$,

– *Two decoder mappings* $\{\psi_T : \mathscr{Y}_T^n \longmapsto \mathcal{W}_0 \times \mathcal{W}_T\}$,

– *Two sets of relay functions* $\{f_{T,i}\}_{i=1}^n$ *such that* $\{f_{T,i} : \mathscr{Z}_T^{i-1} \longmapsto \mathscr{X}_T^n\}_{i=1}^n$,

for $T = \{1, 2\}$ *and some finite sets of integers* $\mathcal{W}_0 = \{1, \dots, M_0\}$ *and* $\mathcal{W}_T = \{1, \dots, M_T\}_{T=\{1,2\}}$. *The rates of such code are* $n^{-1} \log M_T$ *and the corresponding maximum error probabilities* *are defined as*

$$T = \{1,2\}: \quad P_{e,T}^{(n)} = \max_{(w_0,w_T)\in\mathcal{W}_0\times\mathcal{W}_T} \Pr\{\psi(\mathbf{Y}_T) \neq (w_0,w_T)\}.$$

Note that the compound relay channel has indeed the very same definition as the simultaneous relay channel however we keep both terms to indicate the difference in codes for each one. One code guarantees a common rate for all channels, i.e. all T and private rate for each channel, i.e. each T, as the code defined above. We refer to this case as the simultaneous relay channel. Another code guarantees only a common rate for, and sends a common message w_0 to all channels, i.e. all T. By using the compound relay channel we mean this case.

Definition 12 (Achievable rates and capacity) *For every $0 < \epsilon, \gamma < 1$, a triple of non-negative numbers (R_0, R_1, R_2) is achievable for the SRC if for every sufficiently large n there exists a n-length block code whose error probability satisfies*

$$P_{e,T}^{(n)}\big(\varphi, \psi, \{f_{T,i}\}_{i=1}^{n}\big) \leq \epsilon$$

for each $T = \{1,2\}$ and the rates

$$\frac{1}{n}\log M_T \geq R_T - \gamma,$$

for $T = \{0,1,2\}$. The set of all achievable rates $\mathscr{C}_{BRC}$ is called the capacity region of the SRC. We emphasize that no prior distribution on T is assumed and thus the encoder must exhibit a code that yields small error probability for every $T = \{1,2\}$.

A similar definition can be offered for the common-message SRC with a single message set $\mathcal{W}_0$, $n^{-1}\log M_0$ and rate R_0. The common-message SRC is equivalent to the compound relay channel and so the achievable rate for the compound relay channel is defined similarly.

Remark 11 *Notice that, since the relay and the receiver are assumed cognizant of the realization of T, the problem of coding for the SRC can be turned into that of the broadcast relay channel (BRC) [29]. Because the source is uncertain about the actual channel, it has to count on the presence of each one of them and therefore to assume the presence of both simultaneously. This leads to the equivalent broadcast model which consists of two relay branches, where each one corresponds to a relay channel with $T = \{1,2\}$, as illustrated in Fig. 6.1(b) and 6.2. The encoder sends common and private messages (W_0, W_T) to destination T at rates (R_0, R_T). The general BRC is defined by the PD*

$$\mathscr{P}_{BRC} = \left\{P_{Y_1Z_1Y_2Z_2|XX_1X_2} : \mathscr{X}\times\mathscr{X}_1\times\mathscr{X}_2 \longmapsto \mathscr{Y}_1\times\mathscr{Z}_1\times\mathscr{Y}_2\times\mathscr{Z}_2\right\},$$

with channel and relay inputs (X, X_1, X_2) *and channel and relay outputs* (Y_1, Z_1, Y_2, Z_2). *Notions of achievability for* (R_0, R_1, R_2) *and capacity remain the same as for conventional BCs (see [24], [14] and [30]). Similar to the case of broadcast channels, the capacity region of the BRC in Fig. 6.1(b) depends only on the following marginal PDs* $\{P_{Y_1|XX_1X_2Z_1Z_2}, P_{Y_2|XX_1X_2Z_1Z_2}, P_{Z_1Z_2|XX_1X_2}\}$.

Remark 12 *We emphasize that the definition of broadcast relay channels does not dismiss the possibility of dependence of the first (respectively the second) destination* Y_1 *on the second (respectively the first) relay* X_2 *and hence it is more general than the simultaneous relay channels. In other words, the current definition of BRC corresponds to that of SRC with the additional constraints to guarantee that* (Y_T, Z_T) *given* (X, X_T) *for* $T = \{1, 2\}$ *are independent of other random variables. Despite the fact that this condition is not necessary until converse proofs the achievable region developed below are more adapted to the simultaneous relay channel. However all the achievable rate regions do not need any additional assumption and hence are valid for the general BRC.*

The next subsections provide achievable rate regions for three different coding strategies.

6.2.2 Achievable region based on DF-DF strategy

Consider the situation where the channels from source-to-relay are stronger than the other channels. In this case, the best known coding strategy for both relays turns out to be Decode-and-Forward (DF). The source must broadcast the information to the destinations based on a broadcast code combined with DF scheme. The coding behind this idea is as follows. The common information is being helped by the common part of both relays while private information is sent by using rate-splitting in two parts, one part by the help of the corresponding relay and the other part by direct transmission from the source to the corresponding destination. The next theorem presents the general achievable rate region.

Theorem 24 *(DF-DF region) An inner bound on the capacity region* $\mathscr{R}_{DF\text{-}DF} \subseteq \mathscr{C}_{BRC}$ *of* *the broadcast relay channel is given by*

$$
\begin{aligned}
\mathscr{R}_{\textit{DF-DF}} = co \bigcup_{P \in \mathscr{Q}} \Big\{ &(R_0 \geq 0, R_1 \geq 0, R_2 \geq 0) : \\
R_0 + R_1 &\leq I_1 - I(U_0, U_1; X_2 | X_1, V_0), \\
R_0 + R_2 &\leq I_2 - I(U_0, U_2; X_1 | X_2, V_0), \\
R_0 + R_1 + R_2 &\leq I_1 + J_2 - I(U_0, U_1; X_2 | X_1, V_0) - I(U_1, X_1; U_2 | X_2, U_0, V_0) - I_M \\
R_0 + R_1 + R_2 &\leq J_1 + I_2 - I(U_0, U_2; X_1 | X_2, V_0) - I(U_1; U_2, X_2 | X_1, U_0, V_0) - I_M \\
2R_0 + R_1 + R_2 &\leq I_1 + I_2 - I(U_0, U_1; X_2 | X_1, V_0) - I(U_0, U_2; X_1 | X_2, V_0) \\
&\quad - I(U_1; U_2 | X_1, X_2, U_0, V_0) - I_M \Big\},
\end{aligned}
$$

where (I_i, J_i, I_M) *with* $i = \{1, 2\}$ *are as follows*

$$
\begin{aligned}
I_i &= \min \big\{ I(U_0, U_i; Z_i | V_0, X_i) + I(U_{i+2}; Y_i | U_0, V_0, X_i, U_i), I(U_0, V_0, U_i, U_{i+2}, X_i; Y_i) \big\}, \\
J_i &= \min \big\{ I(U_i; Z_i | U_0, V_0, X_i) + I(U_{i+2}; Y_i | U_0, V_0, X_i, U_i), I(U_{i+2}, U_i, X_i; Y_i | U_0, V_0) \big\}, \\
I_M &= I(U_3; U_4 | U_1, U_2, X_1, X_2, U_0, V_0),
\end{aligned}
$$

$co\{\cdot\}$ *denotes the convex hull and the union is over all joint PDs* $P_{U_0V_0U_1U_2U_3U_4X_1X_2X} \in \mathscr{Q}$ *such that*

$$
\begin{aligned}
\mathscr{Q} = \big\{ P_{U_0V_0U_1U_2U_3U_4X_1X_2X} &= P_{U_3U_4X|U_1U_2} \, P_{U_1U_2|U_0X_1X_2} \, P_{U_0|X_1X_2V_0} \, P_{X_2|V_0} \, P_{X_1|V_0} \, P_{V_0} \\
&\textit{satisfying} \quad (U_0, V_0, U_1, U_2, U_3, U_4) \multimap (X_1, X_2, X) \multimap (Y_1, Z_1, Y_2, Z_2) \big\}.
\end{aligned}
$$

Proof The complete proof of this theorem is relegated to Appendix A.1. Instead, here we provide an overview of it. First, the original messages are reorganized via rate-splitting into new messages, as shown in Fig. 6.4, where we add part of the private messages together with the common message into new messages (similarly to [14]). The general coding idea of the proof is depicted in Fig. 6.3. The RV V_0 represents the common part for the RVs (X_1, X_2) (the information sent by the relays), which is intended to help the common information encoded in U_0. Private information is sent in two steps, first using the relay help through (U_1, U_2) and based on DF strategy. Then the direct link between source and destinations is used to decode (U_3, U_4). Marton coding is used to allow correlation between the RVs denoted by arrows in Fig. 6.3. To make a random variable simultaneously correlated with multiple RVs, we used multi-level Marton coding. For this purpose, we start with a given set of i.i.d. generated RVs and then in each step

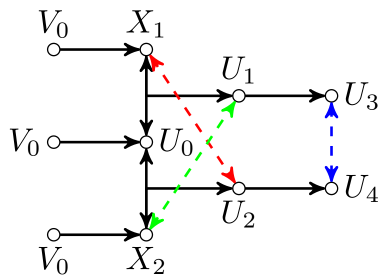


Figure 6.3: Diagram of auxiliary random variables

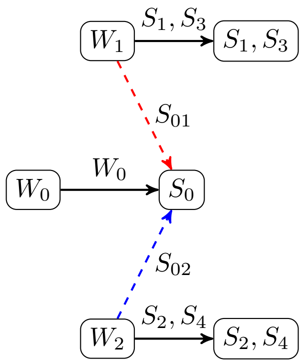


Figure 6.4: The message reconfiguration

we chose a subset such that all their members are jointly typical with a fix RV. Then in each step we look for such a subset inside the previous one. Full details for this process are explained in Appendix A.1.

Table 6.1 shows details for the transmission in time. Both relays knowing $\underline{v}_0, \underline{x}_b$ decode $\underline{u}_0, \underline{u}_b$ in the same block. Then each destination by using backward decoding decodes all the codebooks in the last block. The final region is a combination of all constraints from Marton coding and decoding which will simplify to the region by using Fourier-Motzkin elimination.

Remark 13 *We have the following observations:*

- *Both rates in Theorem 24 coincide with the conventional rate based on partially DF [1],*
- *It is easy to verify that, by setting $(X_1, X_2, V_0) = \emptyset$, $U_3 = U_1, U_4 = U_2$ $Z_1 = Y_1$ and $Z_2 = Y_2$, the rate region in Theorem 24 is equivalent to Marton's region [12],*

Table 6.1: DF strategy with $b = \{1, 2\}$

$\underline{v}_0(t_{0(i-1)})$	$\underline{v}_0(t_{0(i)})$
$\underline{u}_0(t_{0(i-1)}, t_{0i})$	$\underline{u}_0(t_{0i}, t_{0(i+1)})$
$\underline{x}_b(t_{0(i-1)}, t_{b(i-1)})$	$\underline{x}_b(t_{0i}, t_{bi})$
$\underline{u}_b(t_{0(i-1)}, t_{0i}, t_{b(i-1)}, t_{bi})$	$\underline{u}_b(t_{0i}, t_{0(i+1)}, t_{bi}, t_{b(i+1)})$
$\underline{u}_{b+2}(t_{0(i-1)}, t_{0i}, t_{b(i-1)}, t_{bi}, t_{(b+2)i})$	$\underline{u}_{b+2}(t_{0i}, t_{0(i+1)}, t_{bi}, t_{b(i+1)}, t_{(b+2)(i+1)})$
$\underline{y}_{bi}$	$\underline{y}_{b(i+1)}$

– *The previous region improves one derived for the BRC in [29] and for the BRC with common relay as depicted in Fig. 6.1(c). By choosing $X_1 = X_2 = V_0$ and $U_1 = U_2 = U_0$, the rate region in Theorem 24 can be shown to be equivalent to the inner bound by Kramer et al. in [14]. However the following corollary shows that the rate region in Theorem 24 is strictly better than that of Kramer et al..*

The following corollary provides a sharper inner bound on the capacity region of the BRC with common relay (BRC-CR). In the following region, the relay helps also the private information for the first destination by dividing relay help into two parts V_0 and X_1 however the relay in Kramer *et al.*'s region helps only common information. For instance when $Y_2 = \emptyset$ and the first destination is the degraded version of the relay, intuitively when the second destination channel is so weak that we can ignore it, then the region of Kramer *et al.* cannot achieve the capacity of the first relay channel because the relay can only help the common information. However this is not the case in the following region.

Corollary 11 (BRC with common relay) *An inner bound on the capacity region of* *the BRC-CR* $\mathscr{R}_{BRC\text{-}CR} \subseteq \mathscr{C}_{BRC\text{-}CR}$ *is given by*

$$
\begin{aligned}
\mathscr{R}_{BRC\text{-}CR} = co \bigcup_{P_{V_0U_0U_1U_3U_4X_1X} \in \mathscr{Q}} \Big\{ &(R_0 \geq 0, R_1 \geq 0, R_2 \geq 0) : \\
R_0 + R_1 &\leq \min\{I_1 + I_{1p}, I_3 + I_{3p}\} + I(U_3; Y_1|U_1, U_0, X_1, V_0), \\
R_0 + R_2 &\leq I(U_0, V_0, U_4; Y_2) - I(U_0; X_1|V_0), \\
R_0 + R_1 + R_2 &\leq \min\{I_2, I_3\} + I_{3p} + I(U_3; Y_1|U_1, U_0, X_1, V_0) \\
&\quad + I(U_4; Y_2|U_0, V_0) - I(U_0; X_1|V_0) - I_M, \\
R_0 + R_1 + R_2 &\leq \min\{I_1, I_3\} + I_{1p} + I(U_3; Y_1|U_1, U_0, X_1, V_0) \\
&\quad + I(U_4; Y_2|U_0, V_0) - I(U_0; X_1|V_0) - I_M, \\
2R_0 + R_1 + R_2 &\leq I(U_3; Y_1|U_1, U_0, X_1, V_0) + I(U_4; Y_2|U_0, V_0) + I_2 \\
&\quad + \min\{I_1 + I_{1p}, I_3 + I_{3p}\} - I(U_0; X_1|V_0) - I_M \Big\}
\end{aligned}
$$

with

$$
\begin{aligned}
I_1 &= I(U_0, V_0; Y_1), \\
I_2 &= I(U_0, V_0; Y_2), \\
I_3 &= I(U_0; Z_1|X_1, V_0), \\
I_{1p} &= I(U_1X_1; Y_1|U_0, V_0), \\
I_{3p} &= I(U_1; Z_1|U_0, V_0, X_1), \\
I_M &= I(U_3; U_4|X_1, U_1, U_0, V_0),
\end{aligned}
$$

$co\{\cdot\}$ *denotes the convex hull and* $\mathscr{Q}$ *is the set of all joint PDs* $P_{V_0U_0U_1U_3U_4X_1X}$ *satisfying*

$$(V_0, U_0, U_1, U_3, U_4) \multimap (X_1, X) \multimap (Y_1, Z_1, Y_2).$$

The central idea is that here the relay must help common information and private information for one user at least. It will be shown in the next section that a special case of this corollary reaches the capacity of the degraded Gaussian BRC-CR and semi-degraded BRC-CR.

6.2.3 Achievable region based on CF-DF strategy

Consider now a broadcast relay channel where the source-to-relay channel is stronger that the relay-to-destination channel for one branch and weaker for the other branch.

Hence cooperative strategy is better to be based on DF for one branch and CF for the other. The source must broadcast the information to the destinations based on a broadcast code combined with CF and DF schemes. This scenario may arise when the encoder does not know (e.g. due to user mobility and fading) whether the channel from source-to-relay is better or not than the channel from relay-to-destination. The next theorem presents the general achievable rate region for the case where the first relay employs DF and the second relay employs CF to help common and private information.

Theorem 25 (CF-DF region) *An inner bound on the capacity region of the BRC* $\mathscr{R}_{DF\text{-}CF} \subseteq \mathscr{C}_{BRC}$ *with heterogeneous cooperative strategies is given by*

$$\begin{aligned}\mathscr{R}_{CF\text{-}DF} = co \bigcup_{P\in\mathscr{Q}} \Big\{(R_0 \geq 0, R_1 \geq 0, R_2 \geq 0) :& \\ R_0 + R_1 &\leq I_1, \\ R_0 + R_2 &\leq I_2 - I(U_2; X_1|U_0, V_0), \\ R_0 + R_1 + R_2 &\leq I_1 + J_2 - I(U_1, X_1; U_2|U_0, V_0), \\ R_0 + R_1 + R_2 &\leq J_1 + I_2 - I(U_1, X_1; U_2|U_0, V_0), \\ 2R_0 + R_1 + R_2 &\leq I_1 + I_2 - I(U_1, X_1; U_2|U_0, V_0)\Big\},\end{aligned}$$

where the quantities (I_i, J_i) *with* $i = \{1, 2\}$ *are given by*

$$\begin{aligned} I_1 &= \min\big\{I(U_0, U_1; Z_1|X_1, V_0), I(U_1, U_0, X_1, V_0; Y_1)\big\}, \\ I_2 &= I(U_2, U_0, V_0; \hat{Z}_2, Y_2|X_2), \\ J_1 &= \min\big\{I(U_1; Z_1|X_1, U_0, V_0), I(U_1, X_1; Y_1|U_0, V_0)\big\}, \\ J_2 &= I(U_2; \hat{Z}_2, Y_2|X_2, U_0, V_0), \end{aligned}$$

$co\{\cdot\}$ *denotes the convex hull and the set of all admissible PDs* $\mathscr{Q}$ *is defined as*

$$\begin{aligned}\mathscr{Q} = \Big\{ & P_{V_0U_0U_1U_2X_1X_2XY_1Y_2Z_1Z_2\hat{Z}_2} = P_{V_0}P_{X_2}P_{X_1|V_0}P_{U_0|V_0}P_{U_2U_1|X_1U_0}P_{X|U_2U_1}\times \\ & P_{Y_1Y_2Z_1Z_2|XX_1X_2}P_{\hat{Z}_2|X_2Z_2}, \text{ satisfying} \\ & I(X_2;Y_2) \geq I(Z_2;\hat{Z}_2|X_2Y_2), \text{ and} \\ & (V_0, U_0, U_1, U_2) \ominus (X_1, X_2, X) \ominus (Y_1, Z_1, Y_2, Z_2)\Big\}.\end{aligned}$$

The proof is presented in Appendix A.2.

It is possible for general broadcast relay channels to change the coding strategy in first and second relay, from DF to CF and vice versa to obtain another region. Finally a bigger region can be obtained by taking the union of two regions. In order to transmit the common information and at the same time to exploit the help of the relay for the DF destination, the regular coding is used with block-Markov coding scheme. In fact, V_0 is the part of X_1 to help the transmission of U_0. But the second destination uses CF where the relay input and the channel input are mainly independent. Although it seems, at the first look, that block-Markov coding with superposition coding is not compatible with CF scheme. The source uses regular encoding and superimpose the code from the current block over the code from the previous block. When the relay uses DF, it transmits the code from the previous block and hence the destination can exploit this help to decode all codes. But when the relay uses CF, the destination seems to be faced with two superimposed codes which has to be decoded. Because the center codeword carries the dummy message in the first block, the destination can decode the cloud knowing the center. Then in the next block using the same idea, continues to decode by removing the center code. But this leads to performance loss because one part of the transmitted code is indeed thrown away. So it seems that the superposition coding is not proper for CF. However it can be shown that this is not the case. By using backward decoding, the code can be also exploited for CF scheme as well, without loss of performance. Indeed the CF destination takes V_0 not as the relay code but as the source code over which U_0 is superimposed. Then at the last block U_0 carries the dummy message but superimposed on V_0 which carries the message from the last block. Hence the destination can jointly decode (U_0, V_0) and thus exploiting both codes without performance loss with respect to usual CF.

Now consider the compound relay channel where the channel in operation is chosen from a set of relay channels. For simplicity suppose that the set includes only two channels such that DF strategy compared to CF yields a better rate for the first channel and a worse rate for the second one. The goal is to transmit with a rate with arbitrary small error for both channels. Then using the regular encoding, it can be seen that the best cooperative strategy can be picked up for each channel because the relay employs DF in the first channel and CF in the second channel without any problem. The next corollary results directly from this observation.

Corollary 12 (common-information) *A lower bound on the capacity of the compound* *(or common-message BRC) relay channel is given by*

$$R_0 \leq \max_{P_{X_1 X_2 X} \in \mathscr{Q}} \min \big\{ I(X; Z_1|X_1), I(X, X_1; Y_1), I(X; \hat{Z}_2, Y_2|X_2) \big\}.$$

Corollary 12 follows from Theorem 25 by choosing $U_1 = U_2 = U_0 = X$, $V_0 = X_1$. Whereas the following corollary follows by setting $U_0 = V_0 = \emptyset$.

Corollary 13 (private information) *An inner bound on the capacity region of the BRC with heterogeneous cooperative strategies is given by the convex hull of the set of rates* (R_1, R_2) *satisfying*

$$R_1 \leq \min \big\{ I(U_1; Z_1|X_1), I(U_1, X_1; Y_1) \big\}, \tag{6.1}$$

$$R_2 \leq I(U_2; \hat{Z}_2, Y_2|X_2) - I(U_2; X_1), \tag{6.2}$$

$$R_1 + R_2 \leq \min \big\{ I(U_1; Z_1|X_1), I(U_1, X_1; Y_1) \big\} + I(U_2; \hat{Z}_2, Y_2|X_2) - I(U_1, X_1; U_2), \tag{6.3}$$

for all joint PDs $P_{U_1 U_2 X_1 X_2 X Y_1 Y_2 Z_1 Z_2 \hat{Z}_2} \in \mathscr{Q}$.

Remark 14 *The region in Theorem 25 is equivalent to Marton's region [12] with* $(X_1, X_2, V_0) = \emptyset$, $Z_1 = Y_1$ *and* $Z_2 = Y_2$. *Observe that the rate corresponding to DF scheme that appears in Theorem 25 coincides with the conventional DF rate, whereas the CF rate appears with a little difference. In fact,* X *is decomposed into* (U, X_1) *which replace it in the rate corresponding to CF scheme.*

6.2.4 Achievable region based on CF-CF strategy

We consider now another scenario where the channels from relay-to-destination are stronger than the others and hence the efficient coding strategy turns to be CF for both users. The inner bound based on this strategy is given by the following theorem.

Theorem 26 (CF-CF region) *An inner bound on the capacity region of the BRC* $\mathscr{R}_{CF\text{-}CF} \subseteq \mathscr{C}_{BRC}$ *is given by*

$$\begin{aligned}
\mathscr{R}_{CF\text{-}CF} = co \bigcup_{P \in \mathscr{Q}} \Big\{ & (R_0 \geq 0, R_1 \geq 0, R_2 \geq 0) : \\
& R_0 + R_1 \leq I(U_0, U_1; Y_1, \hat{Z}_1|X_1), \\
& R_0 + R_2 \leq I(U_0, U_2; Y_2, \hat{Z}_2|X_2), \\
& R_0 + R_1 + R_2 \leq I_0 + I(U_1; Y_1, \hat{Z}_1|X_1, U_0) + I(U_2; Y_2, \hat{Z}_2|X_2, U_0) - I(U_1; U_2|U_0), \\
& 2R_0 + R_1 + R_2 \leq I(U_0, U_1; Y_1, \hat{Z}_1|X_1) + I(U_0, U_2; Y_2, \hat{Z}_2|X_2) - I(U_1; U_2|U_0) \Big\},
\end{aligned}$$

where the quantity I_0 is defined by

$$I_0 = \min\left\{I(U_0;Y_1,\hat{Z}_1|X_1), I(U_0;Y_2,\hat{Z}_2|X_2)\right\},$$

co$\{\cdot\}$ *denotes the convex hull and the set of all admissible PDs* $\mathscr{Q}$ *is defined as*

$$\begin{aligned}\mathscr{Q} = \big\{P_{U_0U_1U_2X_1X_2XY_1Y_2Z_1Z_2\hat{Z}_1\hat{Z}_2} = P_{X_2}P_{X_1}P_{U_0}P_{U_2U_1|U_0}P_{X|U_2U_1}\times \\ P_{Y_1Y_2Z_1Z_2|XX_1X_2}P_{\hat{Z}_1|X_1Z_1}P_{\hat{Z}_2|X_2Z_2}, \\ I(X_2;Y_2) \geq I(Z_2;\hat{Z}_2|X_2,Y_2), \\ I(X_1;Y_1) \geq I(Z_1;\hat{Z}_1|X_1,Y_1), \\ (U_0,U_1,U_2) \ominus (X_1,X_2,X) \ominus (Y_1,Z_1,Y_2,Z_2)\big\}.\end{aligned}$$

Proof The proof is presented in Appendix A.3.

Notice that this region is equivalent to Marton's region [12] by setting $(X_1, X_2) = \emptyset$, $Z_1 = Y_1$ and $Z_2 = Y_2$.

Remark 15 *A general achievable rate region follows by using time-sharing between all previous regions stated in Theorems 24, 25 and 26.*

6.3 Outer Bounds and Capacity Results

In this section, we first provide an outer bound on the capacity region of the general BRC. Then some capacity results for the cases of semi-degraded BRC with common relay (BRC-CR) and degraded Gaussian BRC-CR are stated.

6.3.1 Outer bounds on the capacity region of general BRC

The next theorems provide general outer bounds on the capacity regions of the BRC and the BRC-CR where $X_1 = X_2$ and $Z_1 = Z_2$, respectively.

Theorem 27 (outer bound BRC) *The capacity region* $\mathscr{C}_{BRC}$ *of the BRC (see Fig. 6.2)* *is included in the set* $\mathscr{C}_{BRC}^{out}$ *of all rates* (R_0, R_1, R_2) *satisfying*

$$\mathscr{C}_{BRC}^{out} = co \bigcup_{P_{VV_1U_1U_2X_1X} \in \mathscr{Q}} \Big\{ (R_0 \geq 0, R_1 \geq 0, R_2 \geq 0) :$$

$$R_0 \leq \min \big\{ I(V;Y_2), I(V;Y_1) \big\},$$

$$R_0 + R_1 \leq \min \big\{ I(V;Y_1), I(V;Y_2) \big\} + I(U_1;Y_1|V),$$

$$R_0 + R_2 \leq \min \big\{ I(V;Y_1), I(V;Y_2) \big\} + I(U_2;Y_2|V),$$

$$R_0 + R_1 \leq \min \big\{ I(V,V_1;Y_1,Z_1|X_1), I(V,V_1;Y_2,Z_2) \big\} + I(U_1;Y_1,Z_1|V,V_1,X_1),$$

$$R_0 + R_2 \leq \min \big\{ I(V,V_1;Y_1,Z_1|X_1), I(V,V_1;Y_2,Z_2) \big\} + I(U_2;Y_2,Z_2|V,V_1,X_1),$$

$$R_0 + R_1 + R_2 \leq I(V;Y_1) + I(U_2;Y_2|V) + I(U_1;Y_1|U_2,V),$$

$$R_0 + R_1 + R_2 \leq I(V;Y_2) + I(U_1;Y_1|V) + I(U_2;Y_2|U_1,V),$$

$$R_0 + R_1 + R_2 \leq I(V,V_1;Y_1,Z_1|X_1) + I(U_2;Y_2,Z_2|V,V_1,X_1) + I(U_1;Y_1,Z_1|X_1,U_2,V,V_1),$$

$$R_0 + R_1 + R_2 \leq I(V,V_1;Y_2,Z_2) + I(U_1;Y_1,Z_1|V,V_1,X_1) + I(U_2;Y_2,Z_2|X_1,U_1,V,V_1) \Big\},$$

where $co\{\cdot\}$ *denotes the convex hull and* $\mathscr{Q}$ *is the set of all joint PDs* $P_{VV_1U_1U_2X_1X_2X}$ *satisfying* $X_1 \multimap V_1 \multimap (V, U_1, U_2, X)$. *The cardinality of auxiliary RVs can be subject to satisfy* $\|\mathscr{V}\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\|\|\mathscr{X}_2\|\|\mathscr{Z}_1\|\|\mathscr{Z}_2\| + 25$, $\|\mathscr{V}_1\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\|\|\mathscr{X}_2\|\|\mathscr{Z}_1\|\|\mathscr{Z}_2\| + 17$ *and* $\|\mathscr{U}_1\|, \|\mathscr{U}_2\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\|\|\mathscr{X}_2\|\|\mathscr{Z}_1\|\|\mathscr{Z}_2\| + 8$.

Proof The proof is presented in Appendix A.4.

Remark 16 *It can be seen from the proof that* V_1 *is a random variable composed of causal and non-causal parts of the relay. So* V_1 *can be intuitively considered as the help of relays for* V. *It can also be inferred from the form of upper bound that* V *and* U_1, U_2 *represent respectively the common and private information.*

Remark 17 *We have the following observations:*

– *The outer bound is valid for the general BRC, i.e. for a 2-receiver 2-relay broadcast channels. However in our case, the pair of* Y, Y_b *depends only on* X, X_b *for* $b = 1, 2$. *Using these Markov relations,* $I(U_b; Y_b, Z_b|X_b, T)$ *and* $I(U_b; Y_b|T)$ *can be bounded by*

$I(X;Y_b,Z_b|X_b,T)$ and $I(X,X_b;Y_b|T)$ for the random variable $T \in \{V,V_1,U_1,U_2\}$. This will simplify the previous region.

- *Moreover we can see that the region in the Theorem 27 is not totally symmetric. So another upper bound can be obtained by replacing the indices 1 and 2, i.e. by introducing V_2 and X_2 instead of V_1 and X_1. The final bound will be the intersection of these two regions.*
- *If relays are not present, i.e., $Z_1 = Z_2 = X_1 = X_2 = V_1 = \emptyset$, it is not difficult to see that the previous bound reduces to the outer bound for general broadcast channels refers to as UVW-outer bound [31]. Furthermore, it was recently shown that such bound is at least as good as all the currently developed outer bounds for the capacity region of broadcast channels [32].*

The next theorem presents an upper bound on capacity of the common-message BRC. The upper bound is useful for evaluation of the capacity in the compound relay channel.

Theorem 28 (upper bound on common-information) *An upper bound on the capacity of the common-message BRC is given by*

$$R_0 \leq \max_{P_{X_1X_2X}\in\mathscr{Q}} \min\big\{I(X;Z_1Y_1|X_1), I(X,X_1;Y_1), I(X;Z_2,Y_2|X_2), I(X,X_2;Y_2)\big\}.$$

Proof The proof follows the conventional method. The common information W_0 is supposed to be decoded by all the users. The upper bound on the rate of each destination is obtained by using this fact and the same proof as [1]. Indeed the upper bound is the combination of the cut-set bound on each relay channel.

The next theorem presents an outer bound on the capacity region of the BRC with common relay. In this case, due to the fact that $Z_1 = Z_2$ and $X_1 = X_2$, we can choose $V_1 = V_2$ because of the definition of V_b (cf. Appendix A.4). Therefore the outer bound of Theorem 27 with the aforementioned symmetric outer bound, which makes use of X_2, V_2, yield the following bound.

Theorem 29 (outer bound BRC-CR) *The capacity region $\mathscr{C}_{BRC\text{-}CR}$ of the BRC-CR* *is included in the set* $\mathscr{C}^{out}_{BRC\text{-}CR}$ *of all rate pairs* (R_0, R_1, R_2) *satisfying*

$$
\begin{aligned}
\mathscr{C}^{out}_{BRC\text{-}CR} = co \bigcup_{P_{VV_1U_1U_2X_1X} \in \mathscr{Q}} \Big\{ (R_0 \geq 0, R_1 \geq 0, R_2 \geq 0) : \\
R_0 &\leq \min \big\{ I(V;Y_2), I(V;Y_1) \big\}, \\
R_0 + R_1 &\leq \min \big\{ I(V;Y_1), I(V;Y_2) \big\} + I(U_1;Y_1|V), \\
R_0 + R_2 &\leq \min \big\{ I(V;Y_1), I(V;Y_2) \big\} + I(U_2;Y_2|V), \\
R_0 + R_1 &\leq \min \big\{ I(V,V_1;Y_1,Z_1|X_1), I(V,V_1;Y_2,Z_1|X_1) \big\} \\
&\quad + I(U_1;Y_1,Z_1|V,V_1,X_1), \\
R_0 + R_2 &\leq \min \big\{ I(V,V_1;Y_1,Z_1|X_1), I(V,V_1;Y_2,Z_1|X_1) \big\} \\
&\quad + I(U_2;Y_2,Z_1|V,V_1,X_1), \\
R_0 + R_1 + R_2 &\leq I(V;Y_1) + I(U_2;Y_2|V) + I(U_1;Y_1|U_2,V), \\
R_0 + R_1 + R_2 &\leq I(V;Y_2) + I(U_1;Y_1|V) + I(U_2;Y_2|U_1,V), \\
R_0 + R_1 + R_2 &\leq I(V,V_1;Y_1,Z_1|X_1) + I(U_2;Y_2,Z_1|V,V_1,X_1) \\
&\quad + I(U_1;Y_1,Z_1|X_1,U_2,V,V_1), \\
R_0 + R_1 + R_2 &\leq I(V,V_1;Y_2,Z_1|X_1) + I(U_1;Y_1,Z_1|V,V_1,X_1) \\
&\quad + I(U_2;Y_2,Z_1|X_1,U_1,V,V_1) \Big\},
\end{aligned}
$$

where $co\{\cdot\}$ *denotes the convex hull and* $\mathscr{Q}$ *is the set of all joint PDs* $P_{VV_1U_1U_2X_1X}$ *verifying* $(X_1) \ominus V_1 \ominus (V, U_1, U_2, X)$ *where the cardinality of auxiliary RVs can be subject to satisfy* $\|\mathscr{V}\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\|\|\mathscr{Z}_1\| + 19$, $\|\mathscr{V}_1\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\|\|\mathscr{Z}_1\| + 11$ *and* $\|\mathscr{U}_1\|, \|\mathscr{U}_2\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\|\|\mathscr{Z}_1\| + 8$.

Proof It is enough to replace Z_2 with Z_1 in Theorem 27. Then the proof follows by taking the union with the symmetric region and using the fact that $I(V, V_1; Y_2, Z_1|X_1)$ is less than $I(V, V_1; Y_2, Z_1)$ due to Markov relationship between V_1 and X_1.

6.3.2 Degraded and semi-degraded BRC with common relay

We now present inner and outer bounds, and capacity results for a special class of BRC-CR. Let us first define two classes of BRC-CRs.

Definition 13 (degraded BRC-CR) *A broadcast relay channel with common relay (BRC-CR) (as is shown in Fig. 6.3), which means* $Z_1 = Z_2$ *and* $X_1 = X_2$*, is said to be de-*

graded (respectively semi-degraded) if the stochastic mapping $\big\{P_{Y_1Z_1Y_2|XX_1} : \mathscr{X} \times \mathscr{X}_1 \longmapsto \mathscr{Y}_1 \times \mathscr{Z}_1 \times \mathscr{Y}_2\big\}$ *satisfies the Markov chains for one of the following cases:*

(I) $X \multimap (X_1, Z_1) \multimap (Y_1, Y_2)$ *and* $(X, X_1) \multimap Y_1 \multimap Y_2$,

(II) $X \multimap (X_1, Z_1) \multimap Y_2$ *and* $X \multimap (Y_1, X_1) \multimap Z_1$,

where the condition (I) is referred to as degraded BRC-CR, and the condition (II) is referred to as semi-degraded BRC-CR.

Notice that the degraded BRC-CR can be seen as the combination of a degraded relay channel with a degraded broadcast channel. On the other hand, the semi-degraded case can be seen as the combination of a degraded broadcast channel with a reversely degraded relay channel. The capacity region of semi-degraded BRC-CR is stated in the following theorem.

Theorem 30 (semi-degraded BRC-CR) *The capacity region of the semi-degraded BRC-CR is given by the following rate region*

$$\mathscr{C}_{BRC\text{-}CR} = \bigcup_{P_{UX_1X} \in \mathscr{Q}} \Big\{(R_1 \geq 0, R_2 \geq 0) : \\ R_2 \leq \min\{I(U, X_1; Y_2), I(U; Z_1|X_1)\}, \\ R_1 + R_2 \leq \min\{I(U, X_1; Y_2), I(U; Z_1|X_1)\} + I(X; Y_1|X_1, U)\Big\},$$

where $\mathscr{Q}$ *is the set of all joint PDs* P_{UX_1X} *satisfying* $U \multimap (X_1, X) \multimap (Y_1, Z_1, Y_2)$ *where the alphabet of the auxiliary RV* U *can be subject to satisfy* $\|\mathscr{U}\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\| + 2$.

Proof It easy to show that the rate region stated in Theorem 30 directly follows from that of Theorem 24 by setting $X_1 = X_2 = V_0$, $Z_1 = Z_2$, $U_0 = U_2 = U_4 = U$, and $U_1 = U_3 = X$. Whereas the converse proof is presented in Appendix A.5.

The next theorems provide outer and inner bounds on the capacity region of the degraded BRC-CR.

Theorem 31 (degraded BRC-CR) *The capacity region* $\mathscr{C}_{BRC\text{-}CR}$ *of the degraded BRC-*

CR is included in the set of pair rates (R_0, R_1) *satisfying*

$$\mathscr{C}^{out}_{BRC\text{-}CR} = \bigcup_{P_{UX_1X}\in\mathscr{Q}} \Big\{(R_0 \geq 0, R_1 \geq 0) : \\ R_0 \leq I(U;Y_2), \\ R_1 \leq \min\big\{I(X;Z_1|X_1,U), I(X,X_1;Y_1|U)\big\}, \\ R_0 + R_1 \leq \min\big\{I(X;Z_1|X_1), I(X,X_1;Y_1)\big\}\Big\},$$

where $\mathscr{Q}$ *is the set of all joint PDs* P_{UX_1X} *satisfying* $U \ominus (X_1, X) \ominus (Y_1, Z_1, Y_2)$ *where the alphabet of the auxiliary RV* U *can be subject to satisfy* $\|\mathscr{U}\| \leq \|\mathscr{X}\|\|\mathscr{X}_1\| + 2$.

Proof The proof is presented in Appendix A.6.

It is not difficult to see that, by applying the degraded condition, the upper bound of Theorem 31 is included in that of Theorem 29.

Theorem 32 (degraded BRC-CR) *An inner bound on the capacity region* $\mathscr{R}_{BRC\text{-}CR} \subseteq \mathscr{C}_{BRC\text{-}CR}$ *of the BRC-CR is given by the set of rates* (R_0, R_1) *satisfying*

$$\mathscr{R}_{BRC\text{-}CR} = co \bigcup_{P_{UVX_1X}\in\mathscr{Q}} \Big\{(R_0 \geq 0, R_1 \geq 0) : \\ R_0 \leq I(U,V;Y_2) - I(U;X_1|V), \\ R_0 + R_1 \leq \min\big\{I(X;Z_1|X_1,V), I(X,X_1;Y_1)\big\}, \\ R_0 + R_1 \leq \min\big\{I(X;Z_1|X_1,U,V), I(X,X_1;Y_1|U,V)\big\} \\ + I(U,V;Y_2) - I(U;X_1|V)\Big\},$$

where $co\{\cdot\}$ *denotes the convex hull for all PDs in* $\mathscr{Q}$ *verifying*

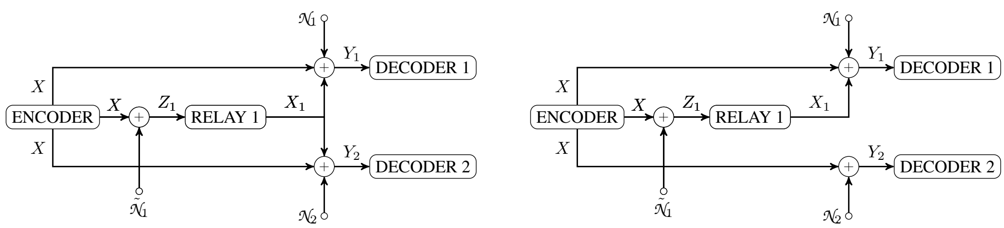
$$P_{UVX_1X} = P_{X|UX_1}P_{X_1U|V}P_V$$

with $(U,V) \ominus (X_1, X) \ominus (Y_1, Z_1, Y_2)$.

Proof The proof of this theorem easily follows by choosing $U_0 = U_2 = U_4 = U$, $V_0 = V$, $U_1 = U_3 = X$ in Corollary 11.

Remark 18 *In the previous bound* V *can be intuitively taken as the help of relay for* R_0. *The tricky part is how to share the help of relay between common and private information. On the one hand, the choice of* $V = \emptyset$ *would remove the help of relay for the common* *information and hence for the case of $Y_1 = Y_2$ it would imply that the help of relay is not exploited and thus the region will be suboptimal. Whereas the choice of $V = X_1$ will lead to a similar problem when $Y_2 = \emptyset$. The code for common information cannot be superimposed on the whole relay code because it limits the relay help for private information. The solution is to superimpose the common information code on an additional random variable V which plays the role of the relay help for common information. However this causes another problem. Now that U is not superimposed over X_1, these variables do not have full dependence anymore and hence the converse does not hold for the channel. To summarize, Marton coding remove the problem of correlation with the price of deviation from the outer bound, i.e. the negative terms in the inner bounds. This is the main reason why the bounds are not tight for the degraded BRC with common relay.*

6.3.3 Degraded Gaussian BRC with common relay



(a) DG-BRC with common relay (b) DG-BRC with partial cooperation

Figure 6.5: Degraded Gaussian BRC

Interestingly, the inner and the outer bounds given by Theorems 32 and 31 happen to coincide for the case of the degraded Gaussian BRC-CR, Fig. 6.5(a). The capacity of this channel was first derived via a different approach in [33]. Let us define the degraded Gaussian BRC-CR by the following channel outputs:

$$Y_1 = X + X_1 + \mathcal{N}_1,$$
$$Y_2 = X + X_1 + \mathcal{N}_2,$$
$$Z_1 = X + \tilde{\mathcal{N}}_1$$

where the source and the relay have power constraints P, P_1, and $\mathcal{N}_1, \mathcal{N}_2, \tilde{\mathcal{N}}_1$ are independent Gaussian noises with variances $N_1, N_2, \tilde{N}_1$, respectively, such that the noises $\mathcal{N}_1, \mathcal{N}_2, \tilde{\mathcal{N}}_1$ satisfy the necessary Markov conditions in definition 13. Note that it is enough to suppose the physical degradedness of receivers respect to the relay and the stochastic degradedness of one receiver respect to another. It means that there exist $\mathcal{N}, \mathcal{N}'$ such that:

$$\mathcal{N}_1 = \tilde{\mathcal{N}}_1 + \mathcal{N},$$
$$\mathcal{N}_2 = \tilde{\mathcal{N}}_1 + \mathcal{N}'.$$

and also $N_1 < N_2$. The following theorem holds as special case of Theorems 31 and 32.

Theorem 33 (degraded Gaussian BRC-CR) *The capacity region of the degraded Gaussian BRC-CR is given by*

$$\begin{aligned}\mathscr{C}_{BRC\text{-}CR} = \bigcup_{0\leq\beta,\alpha\leq 1} \Big\{(R_0 \geq 0, R_1 \geq 0) :& \\ R_0 &\leq C\left(\frac{\alpha(P + P_1 + 2\sqrt{\beta P P_1})}{\overline{\alpha}(P + P_1 + 2\sqrt{\beta P P_1}) + N_2}\right), \\ R_1 &\leq C\left(\frac{\overline{\alpha}(P + P_1 + 2\sqrt{\beta P P_1})}{N_1}\right), \\ R_0 + R_1 &\leq C\left(\frac{\beta P}{\tilde{N}_1}\right)\Big\},\end{aligned}$$

where $C(x) = 1/2 \log(1 + x)$.

Proof The proof is presented in the appendix A.7.

α and β can be respectively interpreted as the power allocation at the source for two destinations and the correlation coefficient between the source and relay code.

6.3.4 Degraded Gaussian BRC with partial cooperation

We now present another capacity region for the Gaussian degraded BRC with partial cooperation (BRC-PC), Fig. 6.5(b), where there is no relay-destination cooperation for the second the destination and the first destination is the degraded version of the relay observation. Moreover the first destination is (stochastically) degraded version of the relay observation.

The input and output relations are as follows:

$$Y_1 = X + X_1 + \mathcal{N}_1,$$
$$Y_2 = X + \mathcal{N}_2,$$
$$Z_1 = X + \tilde{\mathcal{N}}_1.$$

The source and the relay have power constraints P, P_1, and $\mathcal{N}_1, \mathcal{N}_2, \tilde{\mathcal{N}}_1$ are independent Gaussian noises with variances $N_1, N_2, \tilde{N}_1$ and there exists $\mathcal{N}$ such that $\mathcal{N}_1 = \tilde{\mathcal{N}}_1 + \mathcal{N}$ which means that Y_1 is physically degraded respect to Z_1. We also assume $N_2 < \tilde{N}_1$ between Y_2 and Z_1. For this channel the following theorem holds.

Theorem 34 (Gaussian degraded BRC-PC) *The capacity region of the Gaussian degraded BRC-PC is given by:*

$$\mathscr{C}_{BRC\text{-}PC} = \bigcup_{0\leq\beta,\alpha\leq 1} \Big\{(R_1 \geq 0, R_2 \geq 0) :$$

$$R_1 \leq \max_{\beta\in[0,1]} \min \left\{ C\left(\frac{\alpha\beta P}{\overline{\alpha}P + \widehat{N}_1}\right), C\left(\frac{\alpha P + P_1 + 2\sqrt{\overline{\beta}\alpha P P_1}}{\overline{\alpha}P + N_1}\right)\right\},$$

$$R_2 \leq C\left(\frac{\overline{\alpha}P}{N_2}\right)\Big\},$$

where $C(x) = 1/2\log(1+x)$.

Proof The proof is presented in the appendix A.8.

α and β are same as before. Indeed the source assigns the power αP to carry the message of Y_1 and $\overline{\alpha}P$ for Y_2. The theorem is indeed similar to Theorem 30 on the capacity of semi-degraded BRC. Y_2 is the best receiver so it can decode the message destined for Y_1 even after being helped by the relay. It means that the first destination and the relay appear all together as degraded to the second destination. So the second destination can correctly decode the interference of other users and exploit fully the power assigned to it $\overline{\alpha}P$ as it can be seen in the last condition of Theorem 34. However note that Z_1 is not necessarily physically degraded respect to Y_2 which fact makes it a stronger result than that of Theorem 30.

6.4 Gaussian Simultaneous and Broadcast Relay Channels

In this section, based on the achievable rate regions presented in Section 6.2, we compute achievable rate regions for the Gaussian BRC. The Gaussian BRC is modeled as follows:

$$\begin{aligned} Y_{1i} &= \frac{X_i}{\sqrt{d_{y_1}^{\delta}}} + \frac{X_{1i}}{\sqrt{d_{z_1y_1}^{\delta}}} + \mathcal{N}_{1i}, \quad \text{and} \quad Z_{1i} = \frac{X_i}{\sqrt{d_{z_1}^{\delta}}} + \tilde{\mathcal{N}}_{1i}, \\ Y_{2i} &= \frac{X_i}{\sqrt{d_{y_2}^{\delta}}} + \frac{X_{2i}}{\sqrt{d_{z_2y_2}^{\delta}}} + \mathcal{N}_{2i}, \quad \text{and} \quad Z_{2i} = \frac{X_i}{\sqrt{d_{z_2}^{\delta}}} + \tilde{\mathcal{N}}_{2i}. \end{aligned} \tag{6.4}$$

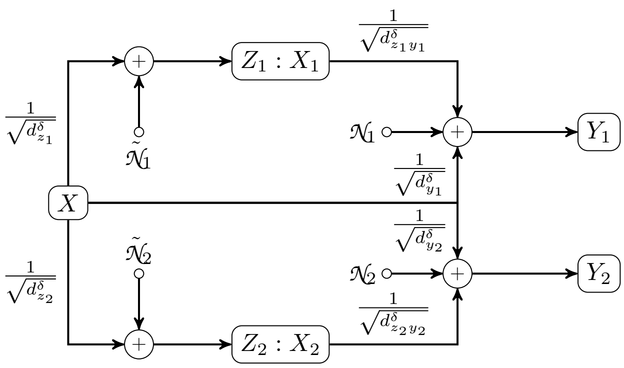


Figure 6.6: Gaussian BRC

The channel inputs $\{X_i\}$ and the relay inputs $\{X_{1i}\}$ and $\{X_{2i}\}$ must satisfy the power constraints

$$\sum_{i=1}^{n} X_i^2 \leq nP, \quad \text{and} \quad \sum_{i=1}^{n} X_{ki}^2 \leq nP_k, \;\; k = \{1,2\}. \tag{6.5}$$

The channel noises $\tilde{\mathcal{N}}_{1i}, \tilde{\mathcal{N}}_{2i}, \mathcal{N}_{1i}, \mathcal{N}_{2i}$ are independent zero-mean i.i.d. Gaussian RVs of variances $\tilde{N}_1, \tilde{N}_2, N_1, N_2$ independent of the channel and relay inputs. The distances (d_{y_1}, d_{y_2}) between source and destinations 1 and 2, respectively, are assumed to be fixed during the communication. Similarly for the distances between the relays and their destinations $(d_{z_1y_1}, d_{z_2y_2})$. Notice that, since (6.4) models the simultaneous Gaussian relay channel where a single pair relay-destination is present at once, no interference is allowed from the relay b to the destination $\overline{b} = \{1,2\} \setminus \{b\}$ for $b = \{1,2\}$. In the remainder of this section, we evaluate DF-DF, DF-CF, CF-CF regions and outer bounds for the channel model (6.4). As for the classical broadcast channel, by using superposition coding, we decompose X as a sum of two independent RVs such that $\mathbb{E}\left\{X_A^2\right\} = \alpha P$ and $\mathbb{E}\left\{X_B^2\right\} = \overline{\alpha} P$, where $\overline{\alpha} = 1 - \alpha$. The codewords (X_A, X_B) contain the information for user Y_1 and user Y_2, respectively.

6.4.1 DF-DF region for Gaussian BRC

We aim to evaluate the rate region in Theorem 24 for the presented Gaussian BRC. To this end, we rely on well-known coding schemes for broadcast and relay channels. A *Dirty-Paper Coding* (DPC) scheme is needed for destination Y_2 to cancel the interference coming from the relay code X_1. Similarly, a DPC scheme is needed for destination Y_1 to cancel the signal noise X_B coming from the code for the other user. The auxiliary RVs (U_1, U_2) are chosen as follow

$$\begin{aligned} U_1 &= X_A + \lambda\, X_B \text{ with } X_A = \tilde{X}_A + \sqrt{\frac{\overline{\beta_1}\alpha P}{P_1}} X_1, \\ U_2 &= X_B + \gamma X_1 \text{ with } X_B = \tilde{X}_B + \sqrt{\frac{\overline{\beta_2}\overline{\alpha} P}{P_1}} X_2, \end{aligned} \tag{6.6}$$

for some parameters $\beta_1, \beta_2, \alpha, \gamma, \lambda \in [0,1]$, where the encoder sends $X = X_A + X_B$. Now choose $V_0 = U_0 = \emptyset$, $U_1 = U_3$ and $U_4 = U_2$ in the theorem 24 in this evaluation. It can be seen that this choice leads to $I_M = 0$ and $I_i = J_i$ for $i = 1, 2$. Then if we choose $R_0 = 0$ and based on the chosen RVs, the following rates are achievable:

$$R_1 \leq \min\big\{I(U_1; Z_1|X_1), I(U_1, X_1; Y_1)\big\} - I(U_1; X_2, U_2|X_1), \tag{6.7}$$

$$R_2 \leq \min\big\{I(U_2; Z_2|X_2), I(U_2, X_2; Y_2)\big\} - I(X_1; U_2|X_2). \tag{6.8}$$

We try to evaluate these rates.

For destination 1, the achievable rate is the minimum of two mutual informations, where the first term is given by $R_{11} \leq I(U_1; Z_1|X_1) - I(U_1; X_2, U_2|X_1)$. The current problem appears as the conventional DPC with $\tilde{X}_A$ as the main message, X_B as the interference and $\tilde{N}_1$ as the noise. Hence the derived rate

$$R_{11}^{(\beta_1,\lambda)} = \frac{1}{2}\log\left[\frac{\alpha\beta_1 P(\alpha\beta_1 P + \overline{\alpha}P + d_{z_1}^{\delta}\tilde{N}_1)}{d_{z_1}^{\delta}\tilde{N}_1(\alpha\beta_1 P + \lambda^2\overline{\alpha}P) + (1-\lambda)^2\overline{\alpha}P\alpha\beta_1 P}\right]. \tag{6.9}$$

The second term is $R_{12} = I(U_1, X_1; Y_1) - I(U_1; X_2, U_2|X_1)$, where the first mutual information can be decomposed into two terms $I(X_1; Y_1)$ and $I(U_1; Y_1|X_1)$. Notice that regardless of the former, the rest of the terms in the expression of the rate R_{12} are similar to R_{11}. The main codeword is $\tilde{X}_A$, while X_B, $\mathcal{N}_1$ are the random state and the noise. After adding the term $I(X_1; Y_1)$ we have

$$R_{12}^{(\beta_1,\lambda)} = \frac{1}{2}\log\left[\frac{\alpha\beta_1 P d_{y_1}^{\delta}\left(\frac{P}{d_{y_1}^{\delta}} + \frac{P_1}{d_{z_1y_1}^{\delta}} + 2\sqrt{\frac{\overline{\beta_1}\alpha P P_1}{d_{y_1}^{\delta}d_{z_1y_1}^{\delta}}} + N_1\right)}{d_{y_1}^{\delta}N_1(\alpha\beta_1 P + \lambda^2\overline{\alpha}P) + (1-\lambda)^2\overline{\alpha}P\alpha\beta_1 P}\right]. \tag{6.10}$$

Based on expressions (6.10) and (6.9), the maximum achievable rate follows as

$$R_1^* = \max_{0 \leq \beta_1, \lambda \leq 1} \min \left\{ R_{11}^{(\beta_1,\lambda)}, R_{12}^{(\beta_1,\lambda)} \right\}.$$

For the second destination, the argument is similar to the one above with the difference that for the current DPC, where only X_1 can be canceled, the rest of X_A appears as noise for the destinations. So it becomes the conventional DPC with $\tilde{X}_B$ as the main message, X_1 as the interference and the $\tilde{\mathcal{N}}_1$ and $\tilde{X}_A$ as noises. The rate writes as

$$R_{21}^{(\beta_1,\beta_2,\gamma)} = \frac{1}{2} \log \left[\frac{\overline{\alpha}\beta_2 P(\overline{\alpha}\beta_2 P + \alpha P + d_{z_2}^{\delta} \tilde{N}_2)}{(d_{z_2}^{\delta} \tilde{N}_2 + \alpha\beta_1 P)(\overline{\alpha}\beta_2 P + \gamma^2 \overline{\beta_1} \alpha P) + (1-\gamma)^2 \overline{\alpha}\beta_2 P \alpha \overline{\beta_1} P} \right], \tag{6.11}$$

and for the other one

$$R_{22}^{(\beta_1,\beta_2,\gamma)} = \frac{1}{2} \log \left[\frac{\overline{\alpha}\beta_2 P d_{y_2}^{\delta} \left(\frac{P}{d_{y_2}^{\delta}} + \frac{P_2}{d_{z_2 y_2}^{\delta}} + 2\sqrt{\frac{\overline{\beta_2}\overline{\alpha} P P_2}{d_{y_2}^{\delta} d_{z_2 y_2}^{\delta}}} + N_2 \right)}{(d_{y_2}^{\delta} N_2 + \alpha\beta_1 P)(\overline{\alpha}\beta_2 P + \gamma^2 \overline{\beta_1} \alpha P) + (1-\gamma)^2 \overline{\alpha}\beta_2 P \alpha \overline{\beta_1} P} \right]. \tag{6.12}$$

And finally the maximum achievable rate follows as

$$R_2^* = \max_{0 \leq \beta_2, \gamma \leq 1} \min \left\{ R_{21}^{(\beta_1,\beta_2,\gamma)}, R_{22}^{(\beta_1,\beta_2,\gamma)} \right\}.$$

6.4.2 DF-CF region for Gaussian BRC

As for the conventional broadcast channel, by using superposition coding, we decompose $X = X_A + X_B$ as a sum of two independent RVs such that $\mathbb{E}\left\{X_A^2\right\} = \alpha P$ and $\mathbb{E}\left\{X_B^2\right\} = \overline{\alpha} P$, where $\overline{\alpha} = 1 - \alpha$. The codewords (X_A, X_B) contain the information intended to receivers Y_1 and Y_2. First, we identify two different cases for which DPC schemes are derived. This is due two asymmetry between two channels. In the first case the code is such that the CF decoder can remove the interference caused by DF code. In the second case, the code is such that the DF decoder cancels the interference of CF code. *Case I:* A DPC scheme is applied to X_B for cancelling the interference X_A, while for the relay branch of the channel this is similar to [1]. Hence, the auxiliary RVs (U_1, U_2) are set to

$$U_1 = X_A = \tilde{X}_A + \sqrt{\frac{\overline{\beta}\alpha P}{P_1}} X_1, \tag{6.13}$$

$$U_2 = X_B + \gamma X_A, \tag{6.14}$$

where β is the correlation coefficient between the relay and source, and $\tilde{X}_A$ and X_1 are independent. Notice that in this case, instead of only Y_2, we have also $\hat{Z}_2$ present in the rate, which is chosen to as $\hat{Z}_2 = Z_2 + \hat{\mathcal{N}}_2$. Thus DPC should be also able to cancel the interference in both, received and compressed signals which have different noise levels. Calculation should be done again with $(Y_2, \hat{Z}_2)$ which are the main message X_B and the interference X_A. We can show that the optimum γ has a similar form to the classical DPC with the noise term replaced by an equivalent noise which is like the harmonic mean of the noise in $(Y_2, \hat{Z}_2)$. The optimum γ^* is given by

$$\begin{aligned} \gamma^* &= \frac{\overline{\alpha} P}{\overline{\alpha} P + N_{t1}}, \\ N_{t1} &= \left[(d_{z_2}^{\delta}(\tilde{N}_2 + \widehat{N}_2))^{-1} + (d_{y_2}^{\delta}(N_2))^{-1} \right]^{-1}. \end{aligned} \tag{6.15}$$

As we can see the equivalent noise is twice of the harmonic mean of the other noise terms.

From Corollary 13, we can see that the optimal γ^* and the current definitions yield the rates

$$\begin{aligned} R_1^* &= \min \left\{ I(U_1; Z_1 | X_1), I(U_1, X_1; Y_1) \right\} \\ &= \max_{0 \leq \beta \leq 1} \min \left\{ C\left(\frac{\alpha \beta P}{\overline{\alpha} P + d_{z_1}^{\delta} \tilde{N}_1} \right), C\left(\frac{\alpha \dfrac{P}{d_{y_1}^{\delta}} + \dfrac{P_1}{d_{z_1 y_1}^{\delta}} + 2\sqrt{\dfrac{\overline{\beta} \alpha P P_1}{d_{y_1}^{\delta} d_{z_1 y_1}^{\delta}}}}{\dfrac{\overline{\alpha} P}{d_{y_1}^{\delta}} + N_1} \right) \right\}, \end{aligned} \tag{6.16}$$

$$\begin{aligned} R_2^* &= I(U_2; Y_2, \hat{Z}_2 | X_2) - I(U_1, X_1; U_2) \\ &= C\left(\frac{\overline{\alpha} P}{d_{y_2}^{\delta} N_2} + \frac{\overline{\alpha} P}{d_{z_2}^{\delta}(\widehat{N}_2 + \tilde{N}_2)} \right), \end{aligned} \tag{6.17}$$

where $C(x) = \frac{1}{2}\log(1+x)$. Note that since (X_A, X_B) are chosen independent, destination 1 sees X_B as an additional channel noise. The compression noise is chosen as follows

$$\widehat{N}_2 = \left(P\left(\frac{1}{d_{y_2}^{\delta} N_2} + \frac{1}{d_{z_2}^{\delta} \tilde{N}_2} \right) + 1 \right) / \frac{P_2}{d_{y_2}^{\delta} N_2}. \tag{6.18}$$

Case 2: We use a DPC scheme for Y_2 to cancel the interference X_1, and next we use a DPC scheme for Y_1 to cancel X_B. For this case, the auxiliary RVs (U_1, U_2) are chosen as

$$\begin{aligned} &U_1 = X_A + \lambda\, X_B \text{ with } \; X_A = \tilde{X}_A + \sqrt{\frac{\overline{\beta} \alpha P}{P_1}} X_1, \\ &U_2 = X_B + \gamma X_1. \end{aligned} \tag{6.19}$$

From Corollary 13, the rates with the current definitions are

$$R_1 = \min\big\{I(U_1;Z_1|X_1), I(U_1,X_1;Y_1)\big\} - I(U_1;U_2|X_1), \tag{6.20}$$

$$R_2 = I(U_2;Y_2,\hat{Z}_2|X_2) - I(X_1;U_2). \tag{6.21}$$

The argument for destination 2 is similar than before but it differs in the DPC. Here only X_1 can be canceled and then X_A remains as additional noise. The optimum γ^* similar to [29] is given by

$$\gamma^* = \sqrt{\frac{\overline{\beta}\alpha P}{P_1}}\frac{\overline{\alpha}P}{\overline{\alpha}P + N_{t2}}, \tag{6.22}$$

$$N_{t2} = \big((d^{\delta}_{z_2}(\tilde{N}_2 + \widehat{N}_2) + \beta\alpha P)^{-1} + (d^{\delta}_{y_2}(N_2) + \beta\alpha P)^{-1}\big)^{-1}, \tag{6.23}$$

and

$$R_2^* = C\left(\frac{\overline{\alpha}P}{d^{\delta}_{y_2}N_2 + \beta\alpha P} + \frac{\overline{\alpha}P}{d^{\delta}_{z_2}(\widehat{N}_2 + \tilde{N}_2) + \beta\alpha P}\right). \tag{6.24}$$

For destination 1, the achievable rate is the minimum of two terms, where the first one is given by

$$\begin{aligned} R_{11}^{(\beta,\lambda)} &= I(U_1;Z_1|X_1) - I(U_1;U_2|X_1) \\ &= \frac{1}{2}\log\left(\frac{\alpha\beta P(\alpha\beta P + \overline{\alpha}P + d^{\delta}_{z_1}\tilde{N}_1)}{d^{\delta}_{z_1}\tilde{N}_1(\alpha\beta P + \lambda^2\overline{\alpha}P) + (1-\lambda)^2\overline{\alpha}P\alpha\beta P}\right). \end{aligned} \tag{6.25}$$

The second term is $R_{12} = I(U_1X_1;Y_1) - I(U_1;U_2|X_1)$, where the first mutual information can be decomposed into two terms $I(X_1;Y_1)$ and $I(U_1;Y_1|X_1)$. Notice that regardless of the former, the rest of the terms in the expression of rate R_{12} are similar to R_{11}. The main codeword is $\tilde{X}_A$, while X_B and $\mathcal{N}_1$ represent the random state and the noise, respectively. After adding the term $I(X_1;Y_1)$, we obtain

$$R_{12}^{(\beta,\lambda)} = \frac{1}{2}\log\left[\frac{\alpha\beta P d^{\delta}_{y_1}\left(\frac{P}{d^{\delta}_{y_1}} + \frac{P_1}{d^{\delta}_{z_1y_1}} + 2\sqrt{\frac{\overline{\beta}\alpha P P_1}{d^{\delta}_{y_1}d^{\delta}_{z_1y_1}}} + N_1\right)}{N_1 d^{\delta}_{y_1}(\alpha\beta P + \lambda^2\overline{\alpha}P) + (1-\lambda)^2\overline{\alpha}P\alpha\beta P}\right]. \tag{6.26}$$

Based on expressions (6.26) and (6.25), the maximum achievable rate follows as

$$R_1^* = \max_{0\leq\beta,\lambda\leq 1}\min\big\{R_{11}^{(\beta,\lambda)}, R_{12}^{(\beta,\lambda)}\big\}. \tag{6.27}$$

It should be noted that the constraints for $\widehat{N}_2$ is still the same as (6.18).

6.4.3 CF-CF region for Gaussian BRC

We now investigate the Gaussian BRC for the CF-CF region, where the relays are collocated with the destinations. In this setting, we set

$$\begin{aligned}\hat{Z}_1 &= Z_1 + \hat{\mathcal{N}}_1,\\ \hat{Z}_2 &= Z_2 + \hat{\mathcal{N}}_2,\end{aligned} \tag{6.28}$$

where $\hat{\mathcal{N}}_1, \hat{\mathcal{N}}_2$ are zero-mean Gaussian noises of variances $\hat{N}_1, \hat{N}_2$. As for the classical broadcast channel, by using superposition coding, we decompose $X = X_A + X_B$ as a sum of two independent RVs such that $\mathbb{E}\left\{X_A^2\right\} = \alpha P$ and $\mathbb{E}\left\{X_B^2\right\} = \overline{\alpha} P$, where $\overline{\alpha} = 1 - \alpha$. The codewords (X_A, X_B) contain the information intended to receivers Y_1 and Y_2. A DPC scheme is applied to X_B for canceling the interference X_A, while for the relay branch of the channel is similar to [1]. Hence, the auxiliary RVs (U_1, U_2) are set to

$$U_1 = X_A, \quad U_2 = X_B + \gamma X_A. \tag{6.29}$$

Notice that in this case, instead of only Y_2, we have also $\hat{Z}_2$ present in the rate. Thus DPC should be also able to cancel the interference in both, received and compressed signals which have different noise levels. Calculation should be done again with $(Y_2, \hat{Z}_2)$ which are the main message X_B and the interference X_A. We can show that the optimum γ has a similar form to the classical DPC with the noise term replaced by an equivalent noise which is like the harmonic mean of the noise in $(Y_2, \hat{Z}_2)$. The optimum

$$\begin{aligned}\gamma^* &= \frac{\overline{\alpha} P}{\overline{\alpha} P + N_{t1}},\\ N_{t1} &= \left[1/(d_{z_2}^{\delta}(\tilde{N}_2 + \widehat{N}_2)) + 1/(d_{y_2}^{\delta} N_2)\right]^{-1}.\end{aligned} \tag{6.30}$$

As we can see, the equivalent noise is twice of the harmonic mean of the other noise terms. For calculating the rates, we use the Theorem 26 with $U_0 = \phi$, which yields the rates

$$\begin{aligned}R_1^* &= I(U_1; Y_1, \hat{Z}_1 | X_1)\\ &= C\left(\frac{\alpha P}{d_{y_1}^{\delta} N_1 + \overline{\alpha} P} + \frac{\alpha P}{d_{z_1}^{\delta}(\widehat{N}_1 + \tilde{N}_1) + \overline{\alpha} P}\right),\end{aligned} \tag{6.31}$$

$$\begin{aligned}R_2^* &= I(U_2; Y_2, \hat{Z}_2 | X_2) - I(U_1 X_1; U_2)\\ &= C\left(\frac{\overline{\alpha} P}{d_{y_2}^{\delta} N_2} + \frac{\overline{\alpha} P}{d_{z_2}^{\delta}(\widehat{N}_2 + \tilde{N}_2)}\right).\end{aligned} \tag{6.32}$$

Note that since (X_A, X_B) are chosen independent, destination 1 sees X_B as an additional channel noise. The compression noise is chosen as follows

$$\hat{N}_1 = \tilde{N}_1 \left[P \left(\frac{1}{d_{y_1}^\delta N_1} + \frac{1}{d_{z_1}^\delta \tilde{N}_1} \right) + 1 \right] / \frac{P_1}{d_{z_1 y_1}^\delta N_1},$$
$$\hat{N}_2 = \tilde{N}_2 \left[P \left(\frac{1}{d_{y_2}^\delta N_2} + \frac{1}{d_{z_2}^\delta \tilde{N}_2} \right) + 1 \right] / \frac{P_2}{d_{z_2 y_2}^\delta N_2}. \tag{6.33}$$

Common-rate: Define $X = U_0$ and evaluate the Theorem 26 for $U_1 = U_2 = \phi$. The goal is to send common-information at rate R_0. It is easy to verify the following results based on the theorem 26:

$$R_0 \leq \min \left\{ C \left(\frac{P}{d_{y_1}^\delta N_1} + \frac{P}{d_{z_1}^\delta (\widehat{N}_1 + \tilde{N}_1)} \right), C \left(\frac{P}{d_{y_2}^\delta N_2} + \frac{P}{d_{z_2}^\delta (\widehat{N}_2 + \tilde{N}_2)} \right) \right\}. \tag{6.34}$$

The constraint for the compression noise remains unchanged, exactly like the previous section.

6.4.4 Source is oblivious to the cooperative strategy adopted by the relay

6.4.4.1 Compound RC

Consider first lower and upper bounds on the common-rate for the DF-CF region. The definition of the channels remain the same. We set $X = U + \sqrt{\frac{\overline{\beta} P}{P_1}} X_1$ and evaluate Corollary 12. The goal is to send common-information at rate R_0. It is easy to verify that the two DF rates result in

$$R_{DF} \leq \min \left\{ C \left(\frac{\beta P}{d_{z_1}^\delta \tilde{N}_1} \right), C \left(\frac{\frac{P}{d_{y_1}^\delta} + \frac{P_1}{d_{z_1 y_1}^\delta} + 2\sqrt{\frac{\overline{\beta} P P_1}{d_{y_1}^\delta d_{z_1 y_1}^\delta}}}{N_1} \right) \right\}, \tag{6.35}$$

R_{DF} is the achievable rate for the destination Y_1. For the destination Y_2, the CF rate $I(X; Y_2, \hat{Z}_2 | X_2)$ follows as

$$R_{CF} \leq C \left(\frac{P}{d_{y_2}^\delta N_2} + \frac{P}{d_{z_2}^\delta (\widehat{N}_2 + \tilde{N}_2)} \right). \tag{6.36}$$

The upper bound from Theorem 28 turns into the next rate

$$C = \max_{0\leq\beta_1,\beta_2\leq 1} \min\Bigg\{ C\left(\beta_1 P\left[\frac{1}{d_{z_1}^{\delta}\tilde{N}_1}+\frac{1}{d_{y_1}^{\delta}N_1}\right]\right), C\left(\frac{\frac{P}{d_{y_1}^{\delta}}+\frac{P_1}{d_{z_1y_1}^{\delta}}+2\sqrt{\frac{\overline{\beta}_1 PP_1}{d_{y_1}^{\delta}d_{z_1y_1}^{\delta}}}}{N_1}\right),$$
$$C\left(\beta_2 P\left[\frac{1}{d_{z_2}^{\delta}\tilde{N}_2}+\frac{1}{d_{y_2}^{\delta}N_2}\right]\right), C\left(\frac{\frac{P}{d_{y_2}^{\delta}}+\frac{P_2}{d_{z_2y_2}^{\delta}}+2\sqrt{\frac{\overline{\beta}_2 PP_2}{d_{y_2}^{\delta}d_{z_2y_2}^{\delta}}}}{N_2}\right)\Bigg\}. \tag{6.37}$$

Observe that the rate (6.36) is exactly the same as the Gaussian CF [14]. This means that DF regular encoding can also be decoded with the CF strategy, as well for the case with collocated relay and receiver (similar to [34]). By using the proposed coding it is possible to send common information at the minimum rate between CF and DF schemes $R_0 = \min\{R_{DF}, R_{CF}\}$ (i.e. expressions (6.35) to (6.36)). For the case of private information, we have shown that any pair of rates $(R_{DF} \leq R_1^*, R_{CF} \leq R_2^*)$ given by (6.24) and (6.27) are admissible and thus (R_{DF}, R_{CF}) can be simultaneously sent.

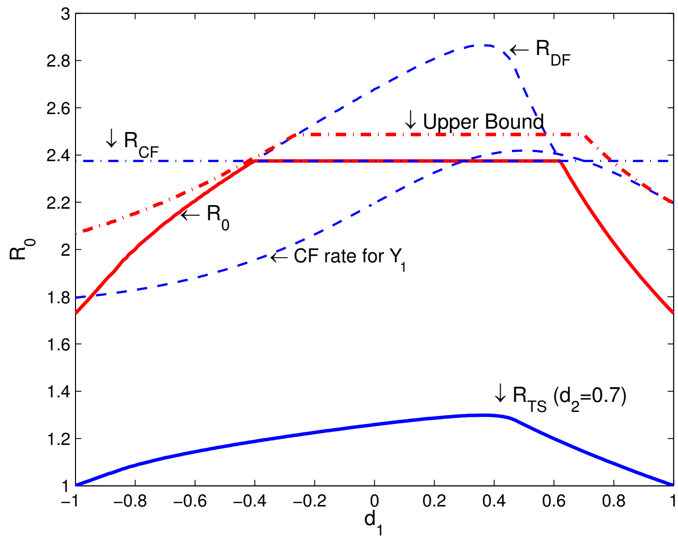


Figure 6.7: Common-rate of the Gaussian BRC with DF-CF strategies

Fig. 6.7 shows numerical evaluation of R_0 for the common-rate case. All channel noises are set to the unit variance and $P = P_1 = P_2 = 10$. The distance between X and (Y_1, Y_2)

is 1, while $d_{z_1} = d_1$, $d_{z_1 y_1} = 1 - d_1$, $d_{z_2} = d_2$, $d_{z_2 y_2} = 1 - d_2$. The relay 1 moves with $d_1 \in [-1, 1]$ and Fig. 6.7 presents rates as a function of d_1. But the position of the relay 2 is assumed to be fixed to $d_2 = 0.7$ so R_{CF} which does not depend on d_1, is a constant function of d_1. On the other hand R_{DF} is dependent on d_1. CF rate for Y_1 is also plotted which corresponds to the case where the first relay uses CF. This setting serves to compare the performances of our coding schemes regarding the position of the relay. It can be seen that one can achieve the minimum between the two possible CF and DF rates. These rates are also compared with a naive time-sharing strategy which consists in using DF scheme $\tau\%$ of the time and CF scheme $(1-\tau)\%$ of the time [2]. Time-sharing yields the achievable rate

$$R_{TS} = \max_{0 \le \tau \le 1} \min\{\tau R_{DF}, (1-\tau) R_{CF}\}.$$

Notice that with the proposed coding scheme significant gains can be achieved when the relay is close to the source (i.e. DF scheme is more suitable), compared to the worst case.

6.4.4.2 Composite RC

Consider now a composite model where the relay is collocated with the source with probability p (refer to it as the first channel) and with the destination with probability $1-p$ (refer to it as the second channel). Therefore DF scheme is the suitable strategy for the first channel while CF scheme performs better on the second one. For any triple of rates (R_0, R_1, R_2) we define the expected rate as

$$R_{av} = R_0 + pR_1 + (1-p)R_2.$$

Expected rate based on the proposed coding strategy is compared to conventional strategies. Alternative coding schemes for this scenario are possible where the encoder can simply invest on one coding scheme DF or CF, which is useful when one probability is high. There are different ways to proceed:

- Send information via DF scheme at the best possible rate between both channels. Then the worst channel cannot decode and thus the expected rate becomes $p_{DF}^{\max} R_{DF}^{\max}$, where $R_{DF}^{\max}$ is the DF rate achieved on the best channel and $p_{DF}^{\max}$ is its probability.

2. One should not confuse time-sharing in compound settings with conventional time-sharing which yields convex combination of rates.

– Send information via the DF scheme at the rate of the worst (second) channel and hence both users can decode the information at rate $R_{DF}^{\min}$. Finally the next expected rate is achievable by investing on only one coding scheme

$$R_{av}^{DF} = \max \left\{ p_{DF}^{\max} R_{DF}^{\max}, R_{DF}^{\min} \right\}.$$

– By investing on CF scheme with the same arguments as before the expected rate writes as

$$R_{av}^{CF} = \max \left\{ p_{CF}^{\max} R_{CF}^{\max}, R_{CF}^{\min} \right\},$$

with definitions of $(R_{CF}^{\min}, R_{CF}^{\max}, p_{CF}^{\max})$ similar to before.

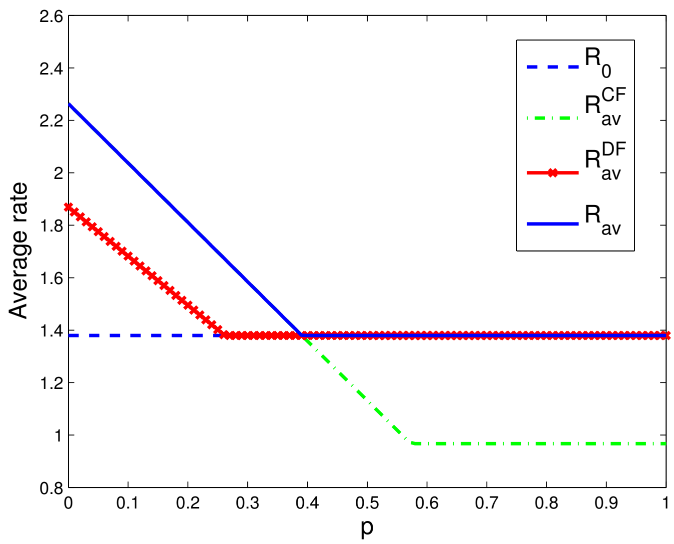


Figure 6.8: Expected rate of the composite Gaussian relay channel

Fig. 6.8 shows numerical evaluation of the average rate. All channel noises are set to the unit variance and $P = P_1 = P_2 = 10$. The distance between X and (Y_1, Y_2) is $(3, 1)$, while $d_{z_1} = 1$, $d_{z_1 y_1} = 2$, $d_{z_2} = 0.9$, $d_{z_2 y_2} = 0.1$. As one can see, the common rate strategy provides a fixed rate all time which is always better than the worst case. However in one corner the full investments on one rate performs better since the high probability of one channel reduces the effect of the other one. Based on the proposed coding scheme, i.e. using the private coding and common coding at the same time, one can cover the corner points and always doing better than both full investments strategies. It is worth to note that in this corner area, only private information of one channel is needed.

6.4.5 Source is oblivious to the presence of relay

We now focus on a scenario where the source user is unaware of the relay's presence. This scenario arises, for example, when the informed relay decides by itself to help the destination whenever cooperative relaying is efficient, e.g. the channel conditions are good enough. In this case, the BRC would have a single relay node. It is assumed here that there is no common information, then we set $X_2 = \{\emptyset\}$ and $Z_2 = Y_2$. The Gaussian BRC is defined here by

$$\begin{aligned} Y_1 &= X + X_1 + \mathcal{N}_1, \\ Y_2 &= X + \mathcal{N}_2, \\ Z_1 &= X + \widehat{\mathcal{N}}_1. \end{aligned} \tag{6.38}$$

The definitions are exactly same as before. As for the classical broadcast channel, by using superposition coding, we decompose X as a sum of two independent RVs such that $\mathbb{E}\left\{X_A^2\right\} = \alpha P$ and $\mathbb{E}\left\{X_B^2\right\} = \overline{\alpha} P$, where $\overline{\alpha} = 1 - \alpha$. The codewords (X_A, X_B) contain the information for user Y_1 and user Y_2, respectively. We use a DPC scheme applied to X_B for canceling the interference X_A, while the relay branch of the channel is similar to [1]. Hence, the auxiliary RVs (U_1, U_2) are set to

$$\begin{aligned} U_1 &= X_A = \tilde{X}_A + \sqrt{\frac{\overline{\beta}\alpha P}{P_1}} X_1, \\ U_2 &= X_B + \gamma X_A, \end{aligned} \tag{6.39}$$

where β is the correlation coefficient between the relay and source, and $\tilde{X}_A$ and X_1 are independent.

The distance between the relay and the source is denoted by d_1, between the relay and destination 1 by $1-d_1$ and between destination 2 and the source by d_2. The new Gaussian BRC writes as: $Z_1 = X/d_1 + \widehat{\mathcal{N}}_1$, $Y_1 = X + X_1/(1-d_1) + \mathcal{N}_1$ and $Y_2 = X/d_2 + \mathcal{N}_2$. From the previous section, the achievable rates are

$$\begin{aligned} R_1^* &= \max_{\beta \in [0,1]} \min \left\{ C\left(\frac{\alpha\beta P}{\overline{\alpha}P + d_1^2 \widehat{N}_1}\right), C\left(\frac{\alpha P + \frac{P_1}{(1-d_1)^2} + \frac{2\sqrt{\overline{\beta}\alpha P P_1}}{|1-d_1|}}{\overline{\alpha}P + N_1}\right) \right\}, \\ R_2^* &= C\left(\frac{\overline{\alpha}P}{d_2^2 N_2}\right). \end{aligned} \tag{6.40}$$

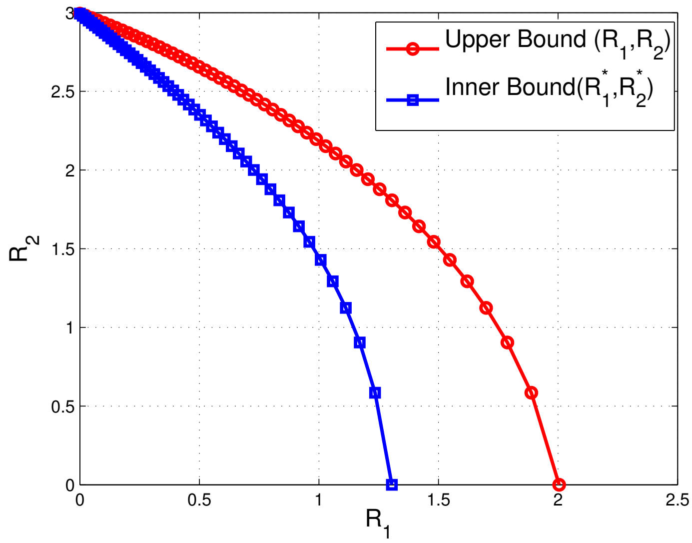


Figure 6.9: Inner bound on the capacity of the Gaussian BRC.

Note that since (X_A, X_B) are chosen independent the destination 1 sees X_B as channel noise. The following outer bound is also presented for this channel

$$R_1 \leq \max_{\beta\in[0,1]} \min\left\{C\left(\frac{\alpha\beta P}{\overline{\alpha}P + d_1^2\widehat{N}_1} + \frac{\alpha\beta P}{\overline{\alpha}P + N_1}\right), C\left(\frac{\alpha P + \frac{P_1}{(1-d_1)^2} + \frac{2\sqrt{\beta\alpha P P_1}}{|1-d_1|}}{\overline{\alpha}P + N_1}\right)\right\},$$

$$R_2 \leq C\left(\frac{\overline{\alpha}P}{d_2^2 N_2}\right). \tag{6.41}$$

Note that if the relay channel is degraded the bound in (6.41) reduces to the region of (6.40) and thus we have the capacity of this channel according to the theorem 34. Fig. 6.9 shows a numerical evaluation of these rates. All channel noises are set to the unit variance and $P = P_1 = 10$. We assume that destination 2, which does not possess a relay, is the closest to the source $d_2 = 0.4$, while the distance between the relay and the source is set to $d_1 = 1.4$. The broadcast strategy provides significant gains compare to the simple time-sharing scheme, which consists in sharing over time the information for both destinations.

Chapter 7

Selective Coding Strategy for Composite Unicast Networks

7.1 Introduction

Multiterminal networks are the essential part of modern telecommunication systems. Wireless Mobile systems, Computer networks, Sensor and Ad hoc networks are some examples of multiterminal networks. They are usually a combination of common basic networks as broadcast channels, interference channels, multiple access channels and relay channels. The vast development of practical networks during recent years revitalized the interest in network information theory. Particularly multicast networks were studied from various aspects. The cutset bound for the general multicast networks was established in [15,16]. Network coding theorem for graphical multicast network was studied in [121] where the max-flow min-cut theorem for network information flow was presented for the point-to-point communication network. Capacity of networks that have deterministic channels with no interference at the receivers was discussed in [48]. Capacity of wireless erasure multicast networks was determined in [122]. A deterministic approximation of general networks was proposed by Avestimehr *et al.* [19]. Lower bound for general deterministic multicast networks was presented and it was shown that their scheme achieves the cut-set upper bound to within a constant gap. Recently Lim *et al.* discussed Noisy Network Coding (NNC) scheme for the general multicast networks which includes all the previous bounds [18]. Kramer *et al.* developed an inner bound for a point-to-point general network using Decode-and-Forward which achieves the capacity of the degraded networks [14].

The relay channel is the essential part of multiterminal networks. Cover and El Gamal [1] provided the main contribution by developing Decode-and-Forward (DF) and Compress-and-Forward (CF) schemes, as the central cooperative strategies for relay networks. Particularly CF strategy was developed using Wyner-Ziv coding at the relay and sequential decoding at the destination. Lately El Gamal-Mohseni-Zahedi [40] developed an alternative scheme for CF, whose achievable rate turns out to be equivalent to the original CF rate. The idea was fully exploited to develop the Noisy Network Coding (NNC) scheme [18] for general relay networks. NNC is based on sending the same message in all blocks –repetitive encoding– and non-unique compression index decoding where compression does not rely on binning. Authors in [42] showed that the same region can be achieved using backward decoding and joint decoding of both compression indexes and messages. In their scheme the communication takes place over $B + L$ blocks, the first B-blocks are used to transmit the message while the last L-blocks serve to decode the compression index. The region is achieved by letting both (B, L) tend to infinity.

The common assumption made for the above scenarios is that the probability distribution (PD) of the network remains fixed during the communication and it is available to all nodes beforehand. Nevertheless, the time-varying nature of wireless channels, e.g. due to fading and user mobility, does not allow terminals to have full knowledge of all channel parameters involved in the communication. In particular, without feedback channel state information (CSI) cannot be available to the encoder ends. During years, an ensemble of works has been carried out addressing both theoretical and practical aspects of communication problems in presence of channel uncertainty. Perhaps, from an information-theoretic view point the compound channel first introduced by Wolfowitz [21] is one of the most important model to deal with channel uncertainty, which continues to attract much of attention from researchers (see [35] and references therein). Composite models are more adapted to deal with wireless scenarios since unlike compound they address channel uncertainty by introducing a PD $\mathbb{P}_\theta$ on the channels. These models consist of a set of conditional PDs from which the current channel index θ –vector of parameters– is drawn according to $\mathbb{P}_\theta$ and remains fixed during the communication. Capacity for this class of channels has been widely studied beforehand (see [27] and references therein), for wireless scenarios via the well-known notion of outage capacity (see [36] and references therein) and oblivious cooperation over fading Gaussian channels in [28, 37, 38].

In this work, we investigate the composite relay channel where the channel index $\theta \in \Theta$ is randomly drawn according to $\mathbb{P}_\theta$. The channel draw $\theta = (\theta_r, \theta_d)$ remains fixed during the communication, however it is assumed to be unknown at the source, fully known at the destination and partly known θ_r at the relay end. Although a compound approach can guarantee asymptotically vanishing error probability regardless of θ, it would be not an adequate choice for most of wireless scenarios where the worst possible channel index yields non-positive rates. A different approach to this problem which is mostly preferred when dealing with wireless models consists in selecting the coding rate r regardless of the current channel index. Hence the encoder cannot necessarily guarantee –no matter the value of r– arbitrary small error probability. Actually, the asymptotic error probability becomes the measure of interest characterizing the reliability function [39]. Moreover, it turns out that depending on the draw of channel parameters, there may not be a unique relay function –between best known cooperative strategies– that minimizes the error probability. However, since CSI is not available at all nodes the relay function should be made independently of its channel measurement and this will become the bottleneck in the code performance. We present a novel coding strategy from which the relay can select, based on its channel measurement θ_r, the adequate coding strategy.

To this purpose, achievable rates are first derived for the two-relay network with mixed coding strategy. This region improves the achievable region for two relay networks with mixed strategies in [14]. As a matter of fact, it is shown that the same code for this two-relay network works as well for the composite relay channel where the relay is allowed to select either DF or CF scheme. Here the source sends the information regardless of the relay function. More specifically, we show that the recent CF scheme [40] can simultaneously work with DF scheme. Furthermore, only CSI from the source-to-relay channel is needed to decide –at the relay end– about the adequate relay function to be implemented. So the relay does not need to know full CSI to decide about the strategy. This idea can be extended to the general composite networks with multiple relays. To this purpose, a similar coding should be developed such that it can be selected whether a relay in the network uses DF or CF. The achievable region is presented which generalizes NNC to the case of mixed coding strategy, i.e. both DF and CF. It is also shown that DF relays can exploit the help of CF relays using offset coding.

This chapter is organized as follows. Section 7.2 presents definitions while Section 7.3 presents main results for the case of single relay network. Section 7.4 presents the result for the case of multiple relay networks and the sketch of the proofs are relegated to the appendix. Section 7.5 provides the application examples and numerical evaluation for the slow-fading Gaussian relay channels and concluding remarks.

7.2 Problem Definition

The composite relay channel consists of a set of relay channels $\left\{P_{Y_{1\theta}^n Z_{1\theta_r}^n | X^n X_{1\theta_r}^n}\right\}_{n=1}^{\infty}$ indexed by vectors of parameters $\theta = (\theta_d, \theta_r)$ with $(\theta_d, \theta_r) \in \Theta$, where $\theta_r \in \Theta_r$ denotes all parameters affecting the relay output and $\theta_d \in \Theta_d$ are the remainder parameters involved in the communication. Let $\mathbb{P}_\theta$ be a joint probability measure on $\Theta = \Theta_d \times \Theta_r$ and define each channel by a conditional PD $\left\{P_{Y_{1\theta} Z_{1\theta_r} | X X_1} : \mathcal{X} \times \mathcal{X}_1 \longmapsto \mathcal{Y}_1 \times \mathcal{Z}_1\right\}$. We assume a memoryless relay channel that implies the next decomposition

$$P_{Y_{1\theta}^n Z_{1\theta_r}^n | X^n X_{1\theta_r}^n}(\underline{y}_1, \underline{z}_1 | \underline{x}, \underline{x}_1) = \prod_{i=1}^{n} P_{Y_{1\theta} Z_{1\theta_r} | X X_{1\theta_r}}(y_1, z_1 | x, x_1),$$

where channel inputs are denoted by $\underline{x} = (x_1, \dots, x_n) \in \mathcal{X}^n$, relay inputs by $\underline{x_1} = (x_{1,1}, \dots, x_{1,n}) \in \mathcal{X}_1^n$, relay observations by $\underline{z}_1 = (z_{1,1}, \dots, z_{1,n}) \in \mathcal{Z}_1^n$ and channel outputs by $\underline{y}_1 = (y_{1,1}, \dots, y_{1,n}) \in \mathcal{Y}_1^n$. The channel parameters affecting relay and destination outputs $\theta = (\theta_r, \theta_d)$ are drawn according to the joint PD $\mathbb{P}_\theta$ and remain fix during the communication. However, the specific draw of θ is assumed to be unknown at the source, fully known at the destination and partly known θ_r at the relay end. Notice that θ_r is enough to know $P_{Z_{1\theta_r}^n | X^n X_{1\theta_r}^n}$ and hence the relay knows [1] all parameters related to its own channel.

Definition 14 (code and achievability) *A code-$\mathcal{C}(n, M_n, r)$ for the composite relay channel consists of:*

- *An encoder mapping $\{\varphi : \mathcal{M}_n \longmapsto \mathcal{X}^n\}$,*
- *A decoder mapping $\{\phi_\theta : \mathcal{Y}_1^n \longmapsto \mathcal{M}_n\}$,*
- *A set of relay functions $\left\{f_{i,\theta_r} : \mathcal{Z}_1^{i-1} \longmapsto \mathcal{X}_1\right\}_{i=1}^n$, for some set of uniformly distributed messages $W \in \mathcal{M}_n = \left\{1, \dots, M_n\right\}$. Note that only partial CSI at the relay is assumed (denoted by θ_r) which is mainly related to the source-relay link.*

An error probability $0 \leq \epsilon < 1$ is said to be r-achievable, if there exists a code-$\mathcal{C}(n, M_n, r)$ with rate satisfying

$$\liminf_{n \to \infty} \frac{1}{n} \log M_n \geq r \tag{7.1}$$

and average error probability

$$\limsup_{n \to \infty} \mathbb{E}_\theta \left[\Pr\left\{\phi_\theta(Y_{1\theta}^n) \neq W \middle| \theta\right\}\right] \leq \epsilon. \tag{7.2}$$

1. We emphasize that there is no loss of generality by assuming this because the index θ is fix during the communication and thus every destination can perfectly know its own channel parameters.

The infimum of all r-achievable error probabilities $\bar{\epsilon}(r)$ is defined as

$$\bar{\epsilon}(r) = \inf \{0 \leq \epsilon < 1 \, : \, \epsilon \text{ is } r\text{-achievable}\} \,. \tag{7.3}$$

We emphasize that for channels satisfying the strong converse property the quantity (7.3) coincides with the usual definition of the outage probability, which turns to be the (asymptotic) average error probability.

7.3 Bounds on the Average Error Probability

In the present setting, we assume that the source is not aware of the specific draw $\theta \sim \mathbb{P}_\theta$ and hence, the coding rate r and the coding strategy –e.g. DF or CF scheme– must be chosen independently of the channel draw. Furthermore, both remain fixed during the communication regardless of the channel measurement at the relay end. We want to characterize the smallest possible average error probability as defined by (7.2), as a function of the coding rate r. In the next chapter [38], it is shown that the average error probability $\bar{\epsilon}(r)$ can be bounded as follows

$$\mathbb{P}_\theta(\underline{r} \in \mathcal{S}_\theta) \leq \bar{\epsilon}(r) \leq \inf_{\phi} \mathbb{P}_\theta(\underline{r} \notin \mathcal{R}_\theta(\phi)), \tag{7.4}$$

where S_θ is the max-flow min-cut bound

$$S_\theta = \min \{ I(X; Z_{1\theta_r} Y_{1\theta} | X_{1\theta_r}), I(XX_{1\theta_r}; Y_{1\theta} \}$$

and $\mathcal{R}_\theta$ is an achievable rate of the relay channel for a given θ and ϕ as the set of all encoding functions φ.

7.3.1 Decode-and-Forward (DF) and Compress-and-Forward (CF) Schemes

Let us first assume that DF scheme is selected and thus the outage probability becomes an upper bound on the average error probability, which is given by

$$P_{\text{out}}^{\text{DF}}(r) = \min_{p(x,x_1)} \mathbb{P}_\theta \big[r > \min\{ I(X; Z_{1\theta_r} | X_1), I(XX_1; Y_{1\theta}) \} \big]. \tag{7.5}$$

Notice that since the source is unaware of $\theta = (\theta_r, \theta_d)$ and $p(x, x_1)$ must be known at both the relay and source end, then $p(x_1)$ cannot be independently optimized on θ_r to minimize the outage probability.

Consider now the case of CF, for which the source does not need to know $p(x_1)$. So the relay can choose $p(x_1)$ to minimize the outage probability conditioned on each θ_r. This process leads to two steps optimization and by using noisy network coding scheme [18], the outage probability of CF scheme is written as

$$P_{\text{out}}^{\text{CF}}(r) = \min_{p(x,q)} \mathbb{E}_{\theta_r}\Big[\min_{p(x_1|q)p(\hat{z}_1|x_1,z_1,q)} \mathbb{P}_{\theta|\theta_r}\big[r > \min\{I(X;\hat{Z}_1Y_{1\theta}|X_{1\theta_r}Q), \\ I(XX_{1\theta_r};Y_{1\theta}|Q) - I(Z_{1\theta_r};\hat{Z}_1|XX_{1\theta_r}Y_{1\theta}Q)\}\big|\theta_r\big]\Big]. \tag{7.6}$$

Then it is easy to see that the strategy minimizing the outage probability will provide the tightest upper bound on the average error probability. From (7.4), it holds for all rate r that

$$\bar{\epsilon}(r) \leq \min\left\{P_{\text{out}}^{\text{DF}}(r), P_{\text{out}}^{\text{CF}}(r)\right\}. \tag{7.7}$$

The central question that arises here is whether the error probability (7.7) can be improved by some kind of smart coding strategy. Specially, the relay would select the better strategy instantaneously according to its channel measurement θ_r. To this purpose, the source code should be capable of being used simultaneously with DF and CF schemes. In the next section, we first prove an achievable rate for the two-relay network where one relay node employs DF coding while the other one uses CF coding, and then by relying on this coding we introduce the Selective Coding Strategy (SCS).

7.3.2 Selective Coding Strategy (SCS)

Consider the two-relay channel as shown in Fig. 7.1. This channel is characterized by conditional PD $\left\{P_{Y_1Z_1Z_2|XX_1X_2} : \mathcal{X}\times\mathcal{X}_1\times\mathcal{X}_2 \longmapsto \mathcal{Y}_1\times\mathcal{Z}_1\times\mathcal{Z}_2\right\}$ for any source input $x\in\mathcal{X}$, relay inputs $x_1\in\mathcal{X}_1$ and $x_2\in\mathcal{X}_2$, channel output $y_1\in\mathcal{Y}_1$, relay observations $z_1\in\mathcal{Z}_1$ and $z_2\in\mathcal{Z}_2$. The achievable rate is defined as usual for this channel. The next theorem provides a lower bound on the capacity of this channel where one relay function uses DF scheme while the other one implements CF scheme.

Theorem 35 (two-relay network) *An inner bound on the capacity of the two-relay network is given by all rates satisfying*

$$R \leq \max_{P\in\mathcal{P}} \min\Big\{I(X;Z_1|X_1Q), \max\Big\{I(XX_1;Y_1|Q), \\ \min\Big[I(XX_1;\hat{Z}_2Y_1|X_2Q), I(XX_1X_2;Y_1|Q) - I(Z_2;\hat{Z}_2|Y_1XX_1X_2Q)\Big]\Big\}\Big\} \tag{7.8}$$

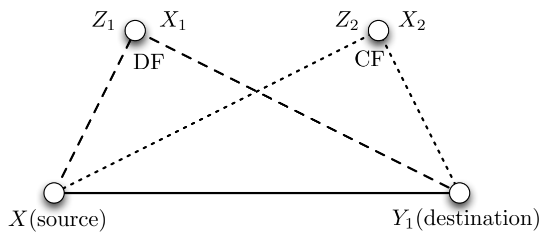


Figure 7.1: Two-relay network.

and the set of all admissible PDs $\mathcal{P}$ is defined as

$$\mathcal{P} = \left\{P_{QX_2X_1XY_1Z_1Z_2\hat{Z}_2} = P_Q P_{X_2|Q} P_{XX_1|Q} P_{Y_1Z_1Z_2|XX_1X_2Q} P_{\hat{Z}_2|X_2Z_2Q}\right\}.$$

The proof of Theorem 35, presented in appendix B.1, is based on superposition coding, DF and CF coding schemes, backward and joint decoding at the destination.

The second maximum in (7.8) determines whether the second relay, which uses CF, is really increasing the rate or it is better to ignore it. It can be seen from the proof that the relay 2 increases the rate, if the following condition holds:

$$I(X_2;Y_1|XX_1Q) \geq I(Z_2;\hat{Z}_2|Y_1XX_1X_2Q) \tag{7.9}$$

Note that the region in Theorem 35 improves the region in [14] by not using the CF relay when its link is too noisy, and by using noisy network coding which improves the CF constraint to (7.9).

Remark 19 *The source code X is superimposed on DF relay code X_1. Observe the last two terms in* (7.8)*. These expressions are the condition of successful decoding at the destination while the first term is the condition of successful decoding of X_1 for the first relay. If one compares the last two terms with usual CF rate, it can be seen that they present similar form with the minor difference that the source code X_1 has been replaced by the pair (X, X_1).*

Consider next a scenario for which the first relay X_1 is not present but the source is still using superposition coding. In this case, we deal with a single relay channel, CF scheme and superposition coding at the source. Along almost the same lines as the proof for Theorem 35, achievable rates can be developed for this case, but there is no need for the condition of successful decoding of X_1 at the relay 1 because such relay is not present. Then we would get only the last two terms in (7.8) as the final rate. This observation means that superposition coding via (X, X_1) achieves the same rate as usual CF with the pair (X, X_1) instead of X. After applying the proper Markov chains, it turns out that the final region remains intact. In other words, DF code can simultaneously be used with CF scheme without any loss of performance. Based on this remark, the following result for the composite relay channel can be derived (see definition 14). We shall refer to this as to the *Selective Coding Strategy (SCS)* and we will see that it can further improve the upper bound on the average error probability.

Theorem 36 (SCS with partial CSI at relay) *The average error probability of the composite relay channel with partial CSI θ_r at the relay can be upper bounded by*

$$\bar{\epsilon}(r) \leq \min_{p(x,x_1,q)} \inf_{\mathcal{D}_{DF}\subseteq\Theta_r} \mathbb{E}_{\theta_r}\Big\{\mathbb{P}_{\theta|\theta_r}\big[r > I_{DF}, \theta_r \in \mathcal{D}_{DF}\big|\theta_r\big] + \min_{p(x_2|q)p(\hat{z}_2|x_2,z_1,q)} \mathbb{P}_{\theta|\theta_r}\big[r > I_{CF}, \theta_r \notin \mathcal{D}_{DF}\big|\theta_r\big]\Big\} \quad (7.10)$$

where (X_1, X_2) denote the corresponding relay inputs selected as follows

$$X_{1\theta_r} = \begin{cases} X_1 & \text{if } \theta_r \in \mathcal{D}_{DF} \\ X_2 & \text{if } \theta_r \notin \mathcal{D}_{DF} \end{cases}$$

and the quantities I_{DF}, I_{CF} are defined by

$$I_{DF} = \min\big\{I(X; Z_{1\theta_r}|X_1 Q), I(XX_1; Y_{1\theta}|Q)\big\}, \quad (7.11)$$

$$I_{CF} = \max\Big\{\min\big[I(X; \hat{Z}_2 Y_{1\theta}|X_2 Q), I(XX_2; Y_{1\theta}) - I(Z_{1\theta_r}; \hat{Z}_2|Y_{1\theta}XX_2Q)\big], I(X; Y_{1\theta})\Big\}. \quad (7.12)$$

Furthermore, the optimal decision region in (7.10) *is given by the set*

$$\mathcal{D}^{\star}_{DF} = \big\{\theta_r \in \Theta_r : I(X; Z_{1\theta_r}|X_1 Q) > r\big\}. \quad (7.13)$$

The proof of Theorem 36 is given in appendix B.1.2, but it is a direct consequence of the achievability proof in Theorem 35 and some subtleties. First of all, we emphasize that the same code proposed in Theorem 35 can be used for the composite relay setting. Basically, the relay has two set of codebooks X_1 and X_2, and it sends $X_{1\theta_r} = X_1$ when condition $\theta_r \in \mathcal{D}_{\mathrm{DF}}$ holds and otherwise it sends $X_{1\theta_r} = X_2$. Obviously, X_1 corresponds to DF while X_2 corresponds CF scheme. Therefore, since the error probability can be made arbitrary small for each relay function, as shown in Fig. 7.1, then the source does not need to know the specific relay function implemented. With this technique, the relay can select the coding strategy according to its channel measurement θ_r, which must improve the overall error probability.

Secondly, observe that for the case of CF there may be the additional condition (7.9) for decoding. However, since the destination is assumed to know θ and consequently the coding strategy, it can be aware if condition (7.9) is or not satisfied. In the case where it fails, destination will treat the relay input as noise –without perform its decoding– and then the condition for unsuccessful decoding simple becomes $\{r > I(X; Y_{1\theta})\}$.

We remark that SCS is at least as good as only DF or CF schemes. On the other hand, it can be shown that in general the best choice for $\mathcal{D}_{\mathrm{DF}}$ is the region for which the relay can decode the source message. Because $I(XX_2; Y_{1\theta})$ is the max-flow min-cut bound and it is bigger than I_{CF}, so if r is bigger than $I(XX_2; Y_{1\theta})$ then it will be bigger than I_{CF} too. So when the decoding at the relay is successful, CF cannot do better than DF and, then the optimal choice is DF scheme. As a matter of fact, full CSI at the relay is not necessary to decide on the best cooperative scheme and CSI on the source-to-relay link is enough to this purpose. Nevertheless full CSI further improves the source-coding description that the relay sends to the destination. This yields the following result which is an extension of Theorem 36.

Corollary 14 (SCS with full CSI at relay) *The average error probability of the composite relay channel with full CSI $\theta = (\theta_r, \theta_d)$ at the relay can be upper bounded by*

$$\bar{\epsilon}(r) \leq \min_{p(x,x_1,q)} \inf_{\mathcal{D}_{DF} \subseteq \Theta_r} \left\{ \mathbb{P}_\theta\left[r > I_{DF}, \theta_r \in \mathcal{D}_{DF}\right] + \mathbb{P}_\theta\left[r > I_{CF}, \theta_r \notin \mathcal{D}_{DF}\right] \right\} \tag{7.14}$$

where (X_1, X_2) denote the corresponding relay inputs selected as follows

$$X_{1\theta_r} = \begin{cases} X_1 & \text{if } \theta_r \in \mathcal{D}_{DF} \\ X_2 & \text{if } \theta_r \notin \mathcal{D}_{DF} \end{cases}$$

and the quantities I_{DF}, I_{CF} are defined by

$$I_{DF} = \min\left\{ I(X; Z_{1\theta_r}|X_1Q), I(XX_1; Y_{1\theta}|Q) \right\}, \tag{7.15}$$

$$I_{CF} = \max_{p(x_2|q)p(\hat{z}_2|x_2,z_1,q)} \min\left\{ I(X; \hat{Z}_2 Y_{1\theta}|X_2Q), I(XX_2; Y_{1\theta}|Q) - I(Z_{1\theta_r}; \hat{Z}_2|Y_{1\theta}XX_2Q) \right\}. \tag{7.16}$$

Furthermore, the optimal decision region in (7.14) *is given by the set*

$$\mathcal{D}^{\star}_{DF} = \big\{\theta_r \in \Theta_r :\ I(X; Z_{1\theta_r}|X_1 Q) > r\big\}. \tag{7.17}$$

The proof of Corollary 14 easily follows from the same lines than that of Theorem 36. Nevertheless, it would be worth to mention here that since full CSI is available at the relay, the relay input can be optimized over $\theta = (\theta_r, \theta_d)$ and then I_{CF} can never be less than the capacity of the link source-to-destination.

7.4 Composite Multiple Relay Networks

In this section, first an achievable region is presented for single source single destination multi-relay networks where part of relays are using DF and the rest are using CF. The result is used to prove the selective coding strategy for the composite multi-relay channels.

The composite multiple relay channel consists of a set of multiple relay channels denoted as follows:

$$\big\{\mathbb{W}^n_\theta = P_{Y^n_{1,\theta} Z^n_{1,\theta_r} Z^n_{2,\theta_r} \dots Z^n_{N,\theta_r}|X^n X^n_{1,\theta_r} X^n_{2,\theta_r} X^n_{N,\theta_r}}\big\}_{n=1}^{\infty}.$$

Similar to the case of single relay channels, they are indexed by vectors of parameters $\theta = (\theta_d, \theta_r)$ with $\theta_d, \theta_r \in \Theta$, where θ_r denotes all parameters affecting the relays' output and θ_d are the remaining parameters involved in the communication. Let $\mathbb{P}_\theta$ be a joint probability measure on Θ and define each channel by a conditional PD

$$\big\{P_{Y_{1,\theta} Z_{1,\theta_r} \dots Z_{N,\theta_r}|X X_{1,\theta_r} \dots X_{N,\theta_r}} : \mathcal{X} \times \mathcal{X}_1 \times \dots \times \mathcal{X}_N \longmapsto \mathcal{Y}_1 \times \mathcal{Z}_1 \times \dots \times \mathcal{Z}_N\big\}.$$

We assume again a memoryless multiple relay channel. The channel parameters affecting relay and destination outputs $\theta = (\theta_r, \theta_d)$ are drawn according to the joint PD $\mathbb{P}_\theta$ and remain fixed during the communication. However, the specific draw of θ is assumed to be unknown at the source, fully known at the destination and partly known (θ_r) at the relays end.

Definition 15 (code and achievability) *A code-$\mathcal{C}(n, M_n, r)$ for the composite multiple relay channels consists of:*

- *An encoder mapping* $\{\varphi : \mathcal{M}_n \longmapsto \mathcal{X}^n\}$,
- *A decoder mapping* $\{\phi_\theta : \mathcal{Y}^n_1 \longmapsto \mathcal{M}_n\}$,

– *A set of relay functions* $\left\{ f_{i,\theta_r}^{(k)} : \mathcal{Z}_k^{i-1} \longmapsto \mathcal{X}_k \right\}_{i=1}^{n}$ *for* $k \in \mathcal{N}$. *Only partial CSI at the relay is assumed (denoted by* θ_r*) which is mainly related to the source-relay link.*

An error probability $0 \leq \epsilon < 1$ *is said to be* r*-achievable, if there exists a code-*$\mathcal{C}(n, M_n, r)$ *with rate satisfying*

$$\liminf_{n\to\infty} \frac{1}{n} \log M_n \geq r \tag{7.18}$$

and average error probability

$$\limsup_{n\to\infty} \mathbb{E}_{\theta} \left[\Pr \left\{ \phi_{\theta}(Y_{1\theta}^n) \neq W \middle| \theta \right\} \right] \leq \epsilon. \tag{7.19}$$

The infimum of all r*-achievable error probabilities* $\bar{\epsilon}(r)$ *is defined as*

$$\bar{\epsilon}(r) = \inf \left\{ 0 \leq \epsilon < 1 \, : \, \epsilon \textit{ is } r\textit{-achievable} \right\}. \tag{7.20}$$

We emphasize that for channels satisfying the strong converse property, (7.20) *coincides with the definition of the outage probability, which is the (asymptotic) average error probability.*

Suppose that $\mathcal{N} = \{1, ..., N\}$ and for any $\mathcal{S} \subseteq \mathcal{N}$, $X_{\mathcal{S}} = \{X_i, i \in \mathcal{S}\}$. In a similar manner, $\bar{\epsilon}(r)$ can be bounded as follows:

$$\mathbb{P}_{\theta}(\underline{r} \in \mathcal{S}_{\theta}) \leq \bar{\epsilon}(r) \leq \inf_{\phi} \mathbb{P}_{\theta}(\underline{r} \notin \mathcal{R}_{\theta}(\phi)), \tag{7.21}$$

where $\mathcal{R}_{\theta}$ is an achievable rate of the unicast network for a given θ, and $\mathcal{S}_{\theta}$ is the infimum of all rates such that every code with such rate yields error probability tending to one, and ϕ as the set of all encoding functions φ. It can be shown that S_{θ} can be replaced with max-flow min-cut bound.

7.4.1 The Multiple Relay Networks

Before bounding the expected error for the composite model, we look to the multiple relay channel characterized by $\left\{ P_{Y_1^n Z_1^n Z_2^n \ldots Z_N^n | X^n X_1^n X_2^n \ldots X_N^n} \right\}_{n=1}^{\infty}$. This channel is not composite and so known to every user. The following theorem presents an inner bound on the capacity of multiple relay channels which generalizes the previous NNC bound on the multiple relay networks.

Theorem 37 (Mixed Noisy Network Coding) *For the multiple relay channel, the following rate is achievable:*

$$R \leq \max_{P\in\mathcal{P}} \max_{\mathcal{V}\subseteq\mathcal{N}, \mathcal{T}\in\Upsilon(\mathcal{V})} \min \left\{ \min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S}), \min_{i\in\mathcal{V}^c} I(X; Z_i | X_{\mathcal{V}^c} Q) \right\} \tag{7.22}$$

where

$$R_{\mathcal{T}}(\mathcal{S}) = I(XX_{\mathcal{V}^c}X_{\mathcal{S}}; \hat{Z}_{\mathcal{S}^c}Y_1|X_{\mathcal{S}^c}Q) - I(Z_{\mathcal{S}}; \hat{Z}_{\mathcal{S}}|XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{S}^c}Y_1Q),$$

for $\mathcal{T} \subseteq \mathcal{V} \subseteq \mathcal{N}$ *and* $\mathcal{V}^c = \mathcal{N} - \mathcal{V}$, $\mathcal{S}^c = \mathcal{T} - \mathcal{S}$. *As convention take* $\min_{\emptyset} = +\infty$. *Moreover* $\Upsilon(\mathcal{V})$ *is defined as follows:*

$$\Upsilon(\mathcal{V}) = \{\mathcal{T} \subseteq \mathcal{V} : \text{ for all } \mathcal{S} \subseteq \mathcal{T}, Q_{\mathcal{T}}(\mathcal{S}) \geq 0\}, \tag{7.23}$$

where $Q_{\mathcal{T}}(\mathcal{S})$ *is defined as follows:*

$$Q_{\mathcal{T}}(\mathcal{S}) = I(X_{\mathcal{S}}; \hat{Z}_{\mathcal{S}^c}Y_1|XX_{\mathcal{S}^c\cup\mathcal{V}^c}Q) - I(Z_{\mathcal{S}}; \hat{Z}_{\mathcal{S}}|XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{S}^c}Y_1Q).$$

The set of all admissible distributions $\mathcal{P}$ *is defined as follows:*

$$\mathcal{P} = \Big\{P_{QXX_{\mathcal{N}}Z_{\mathcal{N}}\hat{Z}_{\mathcal{N}}Y_1} = P_Q P_{XX_{\mathcal{V}^c}|Q} P_{Y_1Z_{\mathcal{N}}|XX_{\mathcal{N}}Q} \prod_{j\in\mathcal{V}} P_{X_j|Q} P_{\hat{Z}_j|X_jZ_jQ}\Big\}.$$

Remark 20 *It can be shown using the same technique as [41] that the optimization in* (7.22) *can be done over* $\mathcal{T} \subseteq \mathcal{V}$ *instead of* $\mathcal{T} \in \Upsilon(\mathcal{V})$. *So* (7.22) *can be rewritten as follows:*

$$R \leq \max_{P\in\mathcal{P}} \max_{\mathcal{T}\subseteq\mathcal{V}\subseteq\mathcal{N}} \min\Big\{\min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S}), \min_{i\in\mathcal{V}^c} I(X; Z_i|X_{\mathcal{V}^c}Q)\Big\}. \tag{7.24}$$

To prove this, it is enough to prove that (7.24) *is included in* (7.22). *In other words it is enough to show that each* $\mathcal{T} \subseteq \mathcal{V}$ *in* $\Upsilon(\mathcal{V})^c$ *does not affect the maximum in* (7.24). *First the following equality can be verified, using the same idea as [41], for* $\mathcal{A} \subseteq \mathcal{S} \subseteq \mathcal{T}$:

$$R_{\mathcal{T}}(\mathcal{S}) = R_{\mathcal{T}\cap\mathcal{A}^c}(\mathcal{S}\cap\mathcal{A}^c) + Q_{\mathcal{T}}(\mathcal{A}). \tag{7.25}$$

Now for each $\mathcal{T} \subseteq \mathcal{V}$ *and* $\mathcal{T} \in \Upsilon(\mathcal{V})^c$, *according to the definition, there is* $\mathcal{A} \subseteq \mathcal{T}$ *such that* $Q_{\mathcal{T}}(\mathcal{A}) < 0$. *From* (7.25), *it can be seen that for each* $\mathcal{S} \subseteq \mathcal{T}\mathcal{A}^c$, $R_{\mathcal{T}}(\mathcal{S}\cup\mathcal{A}) < R_{\mathcal{T}\cap\mathcal{A}^c}(\mathcal{S})$, *which means that replacing* $\mathcal{T}$ *with* $\mathcal{T}\cap\mathcal{A}^c$ *increases the final rate. In other words for each* $\mathcal{T} \subseteq \mathcal{V}$ *and* $\mathcal{T} \in \Upsilon(\mathcal{V})^c$, *there is* $\mathcal{T}' \subset \mathcal{T} \subseteq \mathcal{V}$, *not necessarily in* $\Upsilon(\mathcal{V})^c$ *such that the region with respect to* $\mathcal{T}'$ *is increased which finishes the proof.*

The consequence of this observation is that for each $\mathcal{T}$ *and* $\mathcal{A} \subseteq \mathcal{T}$, *such that* $Q_{\mathcal{T}}(\mathcal{A}) < 0$ *it is enough to ignore these relays in* $\mathcal{A}$ *and not to use their compression. The region* (7.24) *is easier to be dealt with particularly in composite setting.*

In the previous theorem by choosing $\mathcal{V} = \mathcal{N}$ the theorem reduces to SNNC region as in [41, 42] which is equivalent to NNC region [18]. So the theorem 37 generalizes and includes the previous NNC scheme and it provides a potentially larger region. For instance for the single degraded relay channel it achieves the capacity which is not the case for NNC. In fact the relays are divided into two groups. First groups in $\mathcal{V}^c$ are using DF and those in $\mathcal{V}$ are using CF. However a set $\mathcal{T}$ of relays in $\mathcal{V}^c$ can be helpful and increase the rate only if they jointly satisfy (7.23). Otherwise it is better to consider them as noise.

The proof is in general inspired by the proof of the theorem 35 in the sense that instead of X_1, X_2, we have $X_{\mathcal{V}^c}, X_{\mathcal{V}}$. The proof is presented in the appendix B.2.

In the previous theorem, there is no cooperation between the relays of DF and CF. More particularly the relays that are using DF, those in $\mathcal{V}^c$ decode the source message alone and without any help of other relays as it can be seen in the region. However it is possible that the relays in $\mathcal{V}^c$ use the help of those in $\mathcal{V}$ by decoding the transmitted compression indices. This means that each relay in $\mathcal{V}^c$ acts as a potential destination and uses a similar NNC scheme to decode the source message. The next theorem proves the result for this network.

Theorem 38 (Cooperative Mixed Noisy Network Coding) *For the multiple relay channel, the following rate is achievable:*

$$R \leq \max_{P\in\mathcal{P}} \max_{\mathcal{V}\subseteq\mathcal{N}} \min \left(\max_{\mathcal{T}\in\Upsilon(\mathcal{V})} \min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S}), \min_{k\in\mathcal{V}^c} \max_{\mathcal{T}_k\in\Upsilon_k(\mathcal{V})} \min_{\mathcal{S}\subseteq\mathcal{T}_k} R^{(k)}_{\mathcal{T}_k}(\mathcal{S}) \right). \tag{7.26}$$

where

$$R_{\mathcal{T}}(\mathcal{S}) = I(XX_{\mathcal{V}^c}X_{\mathcal{S}};\hat{Z}_{\mathcal{S}^c}Y_1|X_{\mathcal{S}^c}Q) - I(Z_{\mathcal{S}};\hat{Z}_{\mathcal{S}}|XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{S}^c}Y_1Q),$$

with ($\mathcal{S}^c = \mathcal{T} - \mathcal{S}$), *and*

$$R^{(k)}_{\mathcal{T}_k}(\mathcal{S}) = I(X;\hat{Z}_{\mathcal{T}_k}Z_k|X_{\mathcal{V}^c}X_{\mathcal{T}_k}Q) + I(X_{\mathcal{S}};Z_k|X_{\mathcal{V}^c\cup\mathcal{S}^c}Q) - I(\hat{Z}_{\mathcal{S}};Z_{\mathcal{S}}|X_{\mathcal{V}^c\cup\mathcal{T}_k}\hat{Z}_{\mathcal{S}^c}Z_kQ)$$

with ($\mathcal{S}^c = \mathcal{T}_k - \mathcal{S}$), *for* $\mathcal{T}, \mathcal{T}_k \subseteq \mathcal{V} \subseteq \mathcal{N}$ *and* $\mathcal{V}^c = \mathcal{N} - \mathcal{V}$. *Moreover* $\Upsilon(\mathcal{V})$ *and* $\Upsilon_k(\mathcal{V})$ *are defined as follows:*

$$\begin{aligned} \Upsilon(\mathcal{V}) &= \{\mathcal{T} \subseteq \mathcal{V} : \text{ for all } \mathcal{S} \subseteq \mathcal{T}, Q_{\mathcal{T}}(\mathcal{S}) \geq 0\}, \\ \Upsilon_k(\mathcal{V}) &= \{\mathcal{T} \subseteq \mathcal{V} : \text{ for all } \mathcal{S} \subseteq \mathcal{T}, Q^{(k)}_{\mathcal{T}}(\mathcal{S}) \geq 0\} \end{aligned} \tag{7.27}$$

where $Q_{\mathcal{T}}(\mathcal{S})$ *and* $Q^{(k)}_{\mathcal{T}}(\mathcal{S})$ *are defined as follows:*

$$\begin{aligned} Q_{\mathcal{T}}(\mathcal{S}) &= I(X_{\mathcal{S}};\hat{Z}_{\mathcal{S}^c}Y_1|XX_{\mathcal{S}^c\cup\mathcal{V}^c}Q) - I(Z_{\mathcal{S}};\hat{Z}_{\mathcal{S}}|XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{S}^c}Y_1Q), \\ Q^{(k)}_{\mathcal{T}}(\mathcal{S}) &= I(X_{\mathcal{S}};Z_k|X_{\mathcal{V}^c\cup\mathcal{S}^c}Q) - I(\hat{Z}_{\mathcal{S}};Z_{\mathcal{S}}|XX_{\mathcal{V}^c\cup\mathcal{T}}\hat{Z}_{\mathcal{S}^c}Z_kQ). \end{aligned}$$

The set of all admissible distributions $\mathcal{P}$ is defined as before.

The proof is presented in the appendix B.3. The only difference with regard to the theorem 37 is that the DF relays are using the noisy network coding scheme to decode the source message. To this purpose the CF relays do not use binning and the DF relays use joint message and compression index decoding using forward decoding. For the rest, we make some observation about the previous theorem.

Note that $R^{(k)}_{\mathcal{T}_k}(\mathcal{S})$ is strictly less than when NNC is used in the relay using backward decoding. The reason as discussed in [42] is that the gain in NNC is achieved by delaying the decoding. In the relays that are using DF, delaying the decoding is not possible and that's the reason behind the loss in the rate.

Moreover following the similar argument in remark 20, the optimization in (7.26) can be done over the subsets of $\mathcal{V}$ instead of $\Upsilon(\mathcal{V})$. This will give the following region:

$$R \leq \max_{P\in\mathcal{P}} \max_{\mathcal{V}\subseteq\mathcal{N}} \min \left(\max_{\mathcal{T}\subseteq\mathcal{V}} \min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S}), \min_{k\in\mathcal{V}^c} \max_{\mathcal{T}_k\in\Upsilon_k(\mathcal{V})} \min_{\mathcal{S}\subseteq\mathcal{T}_k} R^{(k)}_{\mathcal{T}_k}(\mathcal{S}) \right). \tag{7.28}$$

Finally take a network where all the relays are using CF. However the destination uses forward decoding instead of backward decoding which is the same decoding method as the DF relays' decoding method in the theorem 38. As a consequence of the theorem 38, we can obtain an achievable region with forward decoding for this network which is the following corollary.

Corollary 15 *The following region presents an inner bound on the capacity of multiple relay channels:*

$$R \leq \max_{P\in\mathcal{P}} \max_{\mathcal{T}\subseteq\Upsilon(\mathcal{N})} \min_{\mathcal{S}\subseteq\mathcal{T}} R^{FD}_{\mathcal{T}}(\mathcal{S}), \tag{7.29}$$

where

$$R^{FD}_{\mathcal{T}}(\mathcal{S}) = I(X;\hat{Z}_{\mathcal{T}}Y_1|X_{\mathcal{T}}Q) + I(X_{\mathcal{S}};Y_1|X_{\mathcal{S}^c}Q) - I(\hat{Z}_{\mathcal{S}};Z_{\mathcal{S}}|X_{\mathcal{T}}\hat{Z}_{\mathcal{S}^c}Y_1Q) \quad (\mathcal{S}^c = \mathcal{T}-\mathcal{S}).$$

Moreover $\Upsilon(\mathcal{N})$ is defined as follows:

$$\Upsilon(\mathcal{N}) = \{\mathcal{T}\subseteq\mathcal{N}: \textit{ for all } \mathcal{S}\subseteq\mathcal{T}, Q_{\mathcal{T}}(\mathcal{S})\geq 0\}, \tag{7.30}$$

where $Q_{\mathcal{T}}(\mathcal{S})$ is defined as follows:

$$Q_{\mathcal{T}}(\mathcal{S}) = I(X_{\mathcal{S}};Y_1|X_{\mathcal{S}^c}Q) - I(\hat{Z}_{\mathcal{S}};Z_{\mathcal{S}}|XX_{\mathcal{T}}\hat{Z}_{\mathcal{S}^c}Y_1Q).$$

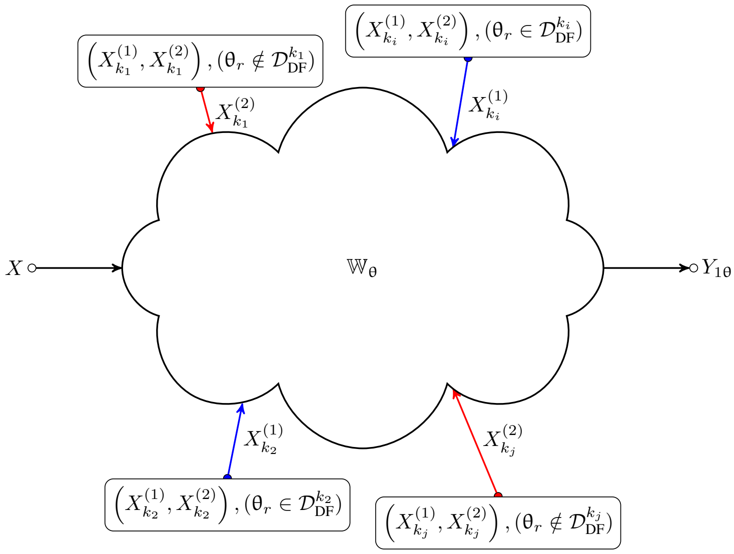


Figure 7.2: Composite Multiple Relay Network

The previous region is not as good as Noisy Network Coding, due to forward decoding but however it is potentially better than using binning and other forward decoding techniques [42]. Particularly the condition which determines the optimization region in [42], i.e. $Q^*_{\mathcal{T}}(\mathcal{S}) \geq 0$ is defined as follows:

$$Q^*_{\mathcal{T}}(\mathcal{S}) = I(X_{\mathcal{S}}; Y_1 | X_{\mathcal{S}^c} Q) - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | X_{\mathcal{T}} \hat{Z}_{\mathcal{S}^c} Y_1 Q).$$

This will create a smaller optimization region that $\Upsilon(\mathcal{N})$ because:

$$I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | X_{\mathcal{T}} \hat{Z}_{\mathcal{S}^c} Y_1 Q) \geq I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | X X_{\mathcal{T}} \hat{Z}_{\mathcal{S}^c} Y_1 Q).$$

So the joint-forward decoding without binning performs potentially better that the joint-forward decoding with binning.

7.4.2 The Composite Multiple Relay Networks

After finding the achievable rates for the multiple relay networks, we move to the composite networks. Suppose that the channel parameters affecting relays and destination outputs are $\theta = (\theta_r, \theta_d)$. The relays are indexed by $\mathcal{N} = \{1, ..., N\}$.

7.4.3 Non-Selective Coding Strategy

The common option is that each relay fixes its coding strategy (Decode-and-Forward or Compress-and-Forward) regardless of θ. In other words $\mathcal{V}$ is chosen beforehand. This will lead to the following bound on the expected error probability for the composite multiple relay channels with partial CSI θ_r at the relays:

$$\bar{\epsilon}(r) \leq \inf_{\mathcal{V}\subseteq\mathcal{N}} \min_{p(x,x_{\mathcal{V}^c},q)} \mathbb{E}_{\theta_r}\left\{ \min_{\prod_{j\in\mathcal{V}} p(x_j|q)p(\hat{z}_j|x_j z_j q)} \mathbb{P}_{\theta|\theta_r}\left[r > I_{\text{CMNNC}}(\mathcal{V}) \middle| \theta_r \right] \right\} \tag{7.31}$$

And $I_{\text{MNNC}}(\mathcal{V})$ is defined as follows ($\theta = (\theta_r, \theta_d)$):

$$I_{\text{CMNNC}}(\mathcal{V}) = \min\left(\max_{\mathcal{T}\subseteq\mathcal{V}} \min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S},\theta), \min_{k\in\mathcal{V}^c} \max_{\mathcal{T}_k\in\Upsilon_k(\mathcal{V})} \min_{\mathcal{S}\subseteq\mathcal{T}_k} R_{\mathcal{T}_k}^{(k)}(\mathcal{S},\theta) \right). \tag{7.32}$$

with

$$\begin{aligned} R_{\mathcal{T}}(\mathcal{S},\theta) &= I(X X_{\mathcal{V}^c} X_{\mathcal{S}}; \hat{Z}_{\mathcal{S}^c} Y_{1\theta} | X_{\mathcal{S}^c} Q) - I(Z_{\mathcal{S}\theta_r}; \hat{Z}_{\mathcal{S}} | X X_{\mathcal{T}\cup\mathcal{V}^c} \hat{Z}_{\mathcal{S}^c} Y_{1\theta} Q) \quad (\mathcal{S}^c = \mathcal{T} - \mathcal{S}), \\ R_{\mathcal{T}_k}^{(k)}(\mathcal{S},\theta) &= I(X; \hat{Z}_{\mathcal{T}_k} Z_{k\theta_r} | X_{\mathcal{V}^c} X_{\mathcal{T}_k} Q) + I(X_{\mathcal{S}}; Z_{k\theta_r} | X_{\mathcal{V}^c\cup\mathcal{S}^c} Q) \\ &\quad - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}\theta_r} | X_{\mathcal{V}^c\cup\mathcal{T}_k} \hat{Z}_{\mathcal{S}^c} Z_{k\theta_r} Q) \qquad (\mathcal{S}^c = \mathcal{T}_k - \mathcal{S}). \end{aligned}$$

for $\mathcal{T}, \mathcal{T}_k \subseteq \mathcal{V} \subseteq \mathcal{N}$ and $\mathcal{V}^c = \mathcal{N} - \mathcal{V}$. Similarly $\Upsilon_k(\mathcal{V})$ is defined as follows:

$$\Upsilon_k(\mathcal{V}) = \{\mathcal{T} \subseteq \mathcal{V} : \text{ for all } \mathcal{S} \subseteq \mathcal{T}, Q_{\mathcal{T}}^{(k)}(\mathcal{S},\theta_r) \geq 0\} \tag{7.33}$$

where $Q_{\mathcal{T}}^{(k)}(\mathcal{S},\theta_r)$ is defined as follows:

$$Q_{\mathcal{T}}^{(k)}(\mathcal{S},\theta_r) = I(X_{\mathcal{S}}; Z_{k\theta_r} | X_{\mathcal{V}^c\cup\mathcal{S}^c} Q) - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}\theta_r} | X X_{\mathcal{V}^c\cup\mathcal{T}} \hat{Z}_{\mathcal{S}^c} Z_{k\theta_r} Q).$$

It can be seen that only CF relays can adapt the probability distribution of the compressed version of their output to the channel and the other relays have to fix both their coding and its distribution regardless of CSI available to them. As we saw in the case of two relay channels a coding can be developed which makes possible for the relays to choose their coding based on the available CSI.

7.4.4 Selective Coding Strategy

Similar to the case of two relay channels, each relay has many codebooks, one for the case that it uses Decode-and-Forward and the others for the case of Compress-and-Forward. Each relay according to its CSI (θ_r) switch between its code for DF and CF. We assume that θ_r is known to all the relays. Each relay k has a decision region $\mathcal{D}_{\text{DF}}^{(k)}$

such that for all $\theta_r \in \mathcal{D}_{\mathrm{DF}}^{(k)}$, the relay k uses Decode-and-Forward and otherwise it uses Compress-and-Forward. Now for each $\mathcal{V} \subseteq \mathcal{N}$ define $\mathcal{D}_{\mathcal{V}}$ as follows:

$$\mathcal{D}_{\mathcal{V}} = \left(\bigcap_{k \in \mathcal{V}^c} \mathcal{D}_{\mathrm{DF}}^{(k)} \right) \bigcap \left(\bigcap_{k \in \mathcal{V}} \mathcal{D}_{\mathrm{DF}}^{(k)\,c} \right)$$

If $\theta_r \in \mathcal{D}_{\mathcal{V}}$, then $\theta_r \notin \mathcal{D}_{\mathrm{DF}}^{(k)}$ for all $k \in \mathcal{V}$ and $\theta_r \in \mathcal{D}_{\mathrm{DF}}^{(k)}$ for all $k \notin \mathcal{V}$. So the relay k, for each $k \in \mathcal{V}$ uses CF and the relay k' for $k' \in \mathcal{V}^c$ uses DF. The ensemble of decision regions of relays will thus provide the regions $\mathcal{D}_{\mathcal{V}}$ which are mutually disjoint and together form a partitioning over the set Θ_r. Now if $\theta_r \in \mathcal{D}_{\mathcal{V}}$, we have a multiple relay network where the relays in $\mathcal{V}$ are using CF. The achievable rate corresponding to this case is known from the theorem 37.

Look at the Figure 7.2. Each relay has two main codebooks, $X_{(k)}^{(1)}$ and $X_{(k)}^{(2)}$. The first code ($X_{(k)}^{(1)}$) is transmitted when $\theta_r \in \mathcal{D}_{\mathrm{DF}}^{(k)}$. This code is based on DF strategy so the relay k decodes the source message and transmits it to the destination. However the source, not knowing whether the relay k is sending $X_{(k)}^{(1)}$ or not, uses superposition coding and superimpose its code over $X_{(k)}^{(1)}$. If the relay k sends $X_{(k)}^{(1)}$ then this will become DF relaying. However in this case the source has to choose the joint distribution between the relay k and itself a priori without knowing θ_r. So the optimum correlation between the source and the relay cannot be adaptively found based on θ_r.

On the other hand if $\theta_r \notin \mathcal{D}_{\mathrm{DF}}^{(k)}$, CF strategy is used. Note that unlike DF, the code which is used for CF, $X_{(k)}^{(2)}$, is independent of the source code and so its probability distribution can be chosen adaptively based on θ_r. The optimum choice for $\mathcal{D}_{\mathcal{V}}$ will potentially give a better outage probability than the case that each relay is using a fixed coding for all θ_r. This idea will lead to the following theorem.

Theorem 39 (SCS with partial CSI) *The average error probability of the composite multiple relay channels with partial CSI θ_r at the relays can be upper bounded by*

$$\bar{\epsilon}(r) \leq \min_{p(x,x_{\mathcal{N}}^{(1)},q)} \inf_{\{\mathcal{D}_{\mathcal{V}},\mathcal{V}\subseteq\mathcal{N}\}\in\Pi(\Theta_r,N)} \sum_{\mathcal{V}\subseteq\mathcal{N}} \mathbb{E}_{\theta_r}\left\{ \min_{\prod_{j\in\mathcal{V}} p(x_j^{(2)}|q)p(\hat{z}_j|x_j^{(2)}z_j q)} \mathbb{P}_{\theta|\theta_r}\left[r > I_{MNNC}(\mathcal{V}), \theta_r \in \mathcal{D}_{\mathcal{V}} \middle| \theta_r \right] \right\} \tag{7.34}$$

$\Pi\left(\Theta_r, N\right)$ is the set of all partitioning over Θ_r into at most 2^N disjoint sets. $I_{MNNC}(\mathcal{V})$

is defined as follows $\theta = (\theta_r, \theta_d)$*:*

$$I_{MNNC}(\mathcal{V}) = \max_{\mathcal{T}\subseteq\mathcal{V}} \min\left\{ \min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S},\theta), \min_{i\in\mathcal{V}^c} I(X; Z_{i\theta_r}|X_{\mathcal{N}}^{(1)}Q)\right\} \tag{7.35}$$

with

$$R_{\mathcal{T}}(\mathcal{S},\theta) = I(XX_{\mathcal{V}^c}^{(1)}X_{\mathcal{S}}^{(2)}; \hat{Z}_{\mathcal{S}^c}Y_{1\theta}|X_{\mathcal{S}^c}^{(2)}Q) - I(Z_{\mathcal{S}\theta_r}; \hat{Z}_{\mathcal{S}}|XX_{\mathcal{T}}^{(2)}X_{\mathcal{V}^c}^{(1)}\hat{Z}_{\mathcal{S}^c}Y_{1\theta}Q),$$

where $(X_k^{(1)}, X_k^{(2)})$ *denote the corresponding relay inputs selected as follows*

$$X_{k\theta_r} = \begin{cases} X_k^{(1)} & \text{if } \theta_r \in \mathcal{D}_{DF}^k \\ X_k^{(2)} & \text{if } \theta_r \notin \mathcal{D}_{DF}^k \end{cases} \quad (\mathcal{D}_{DF}^k = \bigcup_{\mathcal{V}\subset\mathcal{N}, k\notin\mathcal{V}} \mathcal{D}_{\mathcal{V}})$$

In other words for $\theta_r \in \mathcal{D}_{\mathcal{V}}$*, the following Markov chain holds:*

$$X_{\mathcal{V}}^{(1)}X_{\mathcal{V}^c}^{(2)} \ominus XX_{\mathcal{V}^c}^{(1)}X_{\mathcal{V}}^{(2)} \ominus Y_{1\theta}Z_{\mathcal{N}\theta_r}.$$

It can be seen that using superposition code does not change $R_{\mathcal{T}}(\mathcal{S},\theta)$ but unlike the case of single relay channel, the condition of correct decoding at DF relays is changed from $I(X; Z_{i\theta_r}|X_{\mathcal{V}^c}^{(1)}Q)$ to $I(X; Z_{i\theta_r}|X_{\mathcal{N}}^{(1)}Q)$. The proof for this theorem is presented in the appendix B.4.

The next corollary presents the full CSI case.

Corollary 16 (SCS with Full CSI) *The average error probability of the composite multiple relay channels with full CSI at the relays can be upper bounded by*

$$\bar{\epsilon}(r) \leq \min_{p(x,x_{\mathcal{N}}^{(1)},q)} \inf_{\{\mathcal{D}_{\mathcal{V}},\mathcal{V}\subseteq\mathcal{N}\}\in\Pi(\Theta_r,N)} \sum_{\mathcal{V}\subseteq\mathcal{N}} \mathbb{P}_\theta\left[r > I_{MNNC}(\mathcal{V}), \theta_r \in \mathcal{D}_{\mathcal{V}}\right] \tag{7.36}$$

$\Pi(\Theta_r, N)$ *is the set of all partitioning over* Θ_r *into at most* 2^N *disjoint sets.* $I_{MNNC}(\mathcal{V})$ *is defined as follows* $\theta = (\theta_r, \theta_d)$*:*

$$I_{MNNC}(\mathcal{V}) = \max_{\prod_{j\in\mathcal{V}} p(x_j^{(2)}|q)p(\hat{z}_j|x_j^{(2)}z_jq)} \max_{\mathcal{T}\subseteq\mathcal{V}} \min\left\{ \min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S},\theta), \min_{i\in\mathcal{V}^c} I(X; Z_{i\theta_r}|X_{\mathcal{N}}^{(1)}Q)\right\} \tag{7.37}$$

with

$$R_{\mathcal{T}}(\mathcal{S},\theta) = I(XX_{\mathcal{V}^c}^{(1)}X_{\mathcal{S}}^{(2)}; \hat{Z}_{\mathcal{S}^c}Y_{1\theta}|X_{\mathcal{S}^c}^{(2)}Q) - I(Z_{\mathcal{S}\theta_r}; \hat{Z}_{\mathcal{S}}|XX_{\mathcal{T}}^{(2)}X_{\mathcal{N}}^{(1)}\hat{Z}_{\mathcal{S}^c}Y_{1\theta}Q),$$

where $(X_k^{(1)}, X_k^{(2)})$ *denote the corresponding relay inputs selected as follows*

$$X_{k\theta_r} = \begin{cases} X_k^{(1)} & \text{if } \theta_r \in \mathcal{D}_{DF}^k \\ X_k^{(2)} & \text{if } \theta_r \notin \mathcal{D}_{DF}^k \end{cases} \quad (\mathcal{D}_{DF}^k = \bigcup_{\mathcal{V}\subset\mathcal{N}, k\notin\mathcal{V}} \mathcal{D}_{\mathcal{V}})$$

The proof of the corollary remains the same as the theorem 39 with the difference that the relays that are using CF can choose the optimum probability distribution for the compression knowing both θ_r and θ_d.

The theorem 39 and the corollary 16 can be changed to the case that instead of achievability scheme of the theorem 37, the cooperative mixed noisy network coding of the theorem 38 is used. We state the following theorem without proof which will be used in the next section.

Theorem 40 (SCS with partial CSI-Cooperative relays) *The average error probability of the composite multiple relay channels with partial CSI θ_r at the relays can be upper bounded by*

$$\bar{\epsilon}(r) \leq \min_{p(x,x_{\mathcal{N}}^{(1)},q)} \inf_{\{\mathcal{D}_{\mathcal{V}},\mathcal{V}\subseteq\mathcal{N}\}\in\Pi(\Theta_r,N)} \tag{7.38}$$

$$\sum_{\mathcal{V}\subseteq\mathcal{N}} \mathbb{E}_{\theta_r}\Big\{ \min_{\prod_{j\in\mathcal{V}} p(x_j^{(2)}|q)p(\hat{z}_j|x_j^{(2)}z_j q)} \mathbb{P}_{\theta|\theta_r}\big[r > I_{CMNNC}(\mathcal{V}), \theta_r \in \mathcal{D}_{\mathcal{V}} \big| \theta_r\big] \Big\} \tag{7.39}$$

$\Pi\left(\Theta_r, N\right)$ is the set of all partitioning over Θ_r into at most 2^N disjoint sets. $I_{MNNC}(\mathcal{V})$ is defined as follows $\theta = (\theta_r, \theta_d)$:

$$I_{CMNNC}(\mathcal{V}) = \min\left(\max_{\mathcal{T}\subseteq\mathcal{V}} \min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S},\theta), \min_{k\in\mathcal{V}^c} \max_{\mathcal{T}_k\in\Upsilon_k(\mathcal{V})} \min_{\mathcal{S}\subseteq\mathcal{T}_k} R_{\mathcal{T}_k}^{(k)}(\mathcal{S},\theta) \right). \tag{7.40}$$

with

$$\begin{aligned}
R_{\mathcal{T}}(\mathcal{S},\theta) &= I(XX_{\mathcal{V}^c}X_{\mathcal{S}};\hat{Z}_{\mathcal{S}^c}Y_{1\theta}|X_{\mathcal{S}^c}Q) - I(Z_{\mathcal{S}\theta_r};\hat{Z}_{\mathcal{S}}|XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{S}^c}Y_{1\theta}Q) \quad (\mathcal{S}^c = \mathcal{T}-\mathcal{S}),\\
R_{\mathcal{T}_k}^{(k)}(\mathcal{S},\theta) &= I(X;\hat{Z}_{\mathcal{T}_k}Z_{k\theta_r}|X_{\mathcal{V}^c}X_{\mathcal{T}_k}Q) + I(X_{\mathcal{S}};Z_{k\theta_r}|X_{\mathcal{V}^c\cup\mathcal{S}^c}Q)\\
&\quad - I(\hat{Z}_{\mathcal{S}};Z_{\mathcal{S}\theta_r}|X_{\mathcal{V}^c\cup\mathcal{T}_k}\hat{Z}_{\mathcal{S}^c}Z_{k\theta_r}Q) \qquad (\mathcal{S}^c = \mathcal{T}_k-\mathcal{S}).
\end{aligned}$$

for $\mathcal{T},\mathcal{T}_k\subseteq\mathcal{V}\subseteq\mathcal{N}$ and $\mathcal{V}^c = \mathcal{N}-\mathcal{V}$. Moreover $\Upsilon_k(\mathcal{V})$ is defined as follows:

$$\Upsilon_k(\mathcal{V}) = \{\mathcal{T}\subseteq\mathcal{V}: \text{ for all } \mathcal{S}\subseteq\mathcal{T}, Q_{\mathcal{T}}^{(k)}(\mathcal{S},\theta_r)\geq 0\} \tag{7.41}$$

where $Q_{\mathcal{T}}^{(k)}(\mathcal{S},\theta_r)$ is defined as follows:

$$Q_{\mathcal{T}}^{(k)}(\mathcal{S},\theta_r) = I(X_{\mathcal{S}};Z_{k\theta_r}|X_{\mathcal{V}^c\cup\mathcal{S}^c}Q) - I(\hat{Z}_{\mathcal{S}};Z_{\mathcal{S}\theta_r}|XX_{\mathcal{V}^c\cup\mathcal{T}}\hat{Z}_{\mathcal{S}^c}Z_{k\theta_r}Q).$$

$(X_k^{(1)}, X_k^{(2)})$ denote the corresponding relay inputs selected similar to the theorem 39.

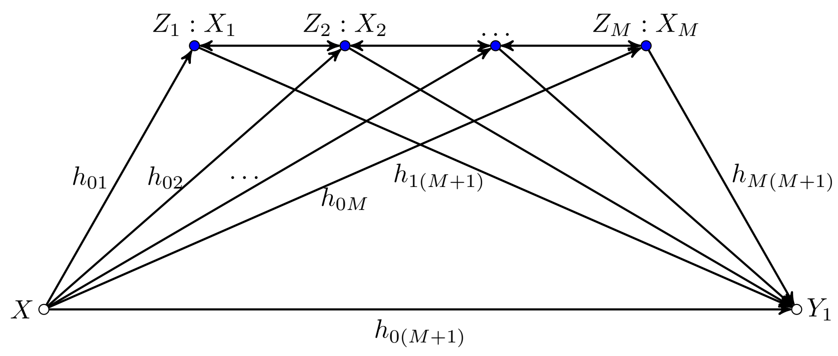


Figure 7.3: Gaussian Multiple Relay Channels

7.5 Application Example and Discussions

7.5.1 Gaussian Fading Multiple Relay Channels

In this part, we evaluate the achievable regions for fading Gaussian multiple relay channels for a given draw of fading coefficients. These bounds will be the main part of the procedure of bounding the expected error. Consider a fading Gaussian network with M relays, single source and destination making $M + 2$ nodes in general, Fig. 7.3. The relays are indexed as usual, however we associate the source with the index 0, i.e. $X_0 = X$ and the destination with the index $M + 1$. We denote by $\mathcal{M} = \{1, ..., M\}$ relays index set, $\mathcal{T} = \{0, 1, ..., M\}$ transmitters index set, and $\mathcal{R} = \{1, ..., M, M + 1\}$ receivers index set. By h_{ij} we denote the fading coefficients from the node i to the node j. Suppose that for a given fading coefficients, the relays with index in $\mathcal{V}^c$ use DF, and those in $\mathcal{V}$ use CF. with following input and output relation:

$$\mathbf{Y}(\mathcal{M}) = H(\mathcal{S}, \mathcal{M})\mathbf{X}(\mathcal{S}) + \mathcal{N}(\mathcal{M}) \tag{7.42}$$

$$\hat{\mathbf{Z}}(\mathcal{M}) = \mathbf{Z}(\mathcal{M}) + \hat{\mathcal{N}}(\mathcal{M}) \tag{7.43}$$

where

$$\mathbf{Y}(\mathcal{S}) = \begin{bmatrix} Z_{i_1} \\ \vdots \\ Z_{i_k} \\ Y_1 \end{bmatrix}, \mathbf{Z}(\mathcal{S}) = \begin{bmatrix} Z_{i_1} \\ \vdots \\ Z_{i_k} \end{bmatrix}, \mathbf{X}(\mathcal{S}) = \begin{bmatrix} X_{i_1} \\ \vdots \\ X_{i_k} \end{bmatrix}, \mathcal{N}(\mathcal{S}) = \begin{bmatrix} \mathcal{N}_{i_1} \\ \vdots \\ \mathcal{N}_{i_k} \end{bmatrix}, i_j \in \mathcal{S}.$$

$\hat{\mathcal{N}}(\mathcal{S})$ and $\hat{\mathbf{Z}}(\mathcal{S})$ are defined in a similar manner. $\hat{N}_k$ is equal to zero for those not in $\mathcal{V}$. On the other hand $H(\mathcal{S}_1, \mathcal{S}_2)$ is defined as $[h_{ij} i \in \mathcal{S}_1, j \in \mathcal{S}_2]$ where evidently $h_{ii} = 0$. It is assumed $\mathcal{N}_k$ is the noise at the receiver k with zero mean and variance N_k. The source transmits with the power P and the relay k with the power P_k. By $\mathbf{N}(\mathcal{S})$ and $\hat{\mathbf{N}}(\mathcal{S})$ we denote the noise variance matrix. The covariance Matrix between channel inputs is $\mathbf{K}(\mathcal{S}_1, \mathcal{S}_2) = \left[\sqrt{P_i P_j}\rho_{ij}\right]$ for $i \in \mathcal{S}_1, j \in \mathcal{S}_2$. Moreover we define $\mathbf{K}(\mathcal{S}) = \mathbf{K}(\mathcal{S}, \mathcal{S})$. Again the relays in $\mathcal{V}$ are generated independently which makes their covariance matrix diagonal.

The region in theorem 37 and 38 can be extended to Gaussian cases. Starting with the region in the theorem 37, we start to evaluate $R_{\mathcal{T}}(\mathcal{S})$ for $\mathcal{S} \subseteq \mathcal{T} \subseteq \mathcal{V}$. Note that all the relays in $\mathcal{T}^c = \mathcal{V} - \mathcal{T}$ are considered as noise and only those CF relays in $\mathcal{T}$ contribute to the final rate. Using this notation the following region can be presented for the case of non-cooperative multiple relay channels.

Corollary 17 *For the fading Gaussian multiple relay channel, the following region is inner bound on the capacity:*

$$R \leq \max_{P \in \mathcal{P}} \max_{\mathcal{T} \subseteq \mathcal{V} \subseteq \mathcal{N}} \min \left\{ \min_{\mathcal{S} \subseteq \mathcal{T}} R_{\mathcal{T}}(\mathcal{S}), \min_{k \in \mathcal{V}^c} I_{DF}(k) \right\}$$

where

$$R_{\mathcal{T}}(\mathcal{S}) = \frac{1}{2} \log \left(\frac{\left| H(\mathcal{S}', \mathcal{D}) \mathbf{K}(\mathcal{S}') H(\mathcal{S}', \mathcal{D})^T + \mathbf{N}(\mathcal{D}) + \hat{\mathbf{N}}(\mathcal{D}) \right|}{\left| H(\mathcal{T}^c, \mathcal{D} \cup \mathcal{S}) \mathbf{K}(\mathcal{T}^c) H(\mathcal{T}^c, \mathcal{D} \cup \mathcal{S})^T + \mathbf{N}(\mathcal{D} \cup \mathcal{S}) + \hat{\mathbf{N}}(\mathcal{D} \cup \mathcal{S}) \right|} \right) - \frac{1}{2} \log \left(\frac{1}{\left| \hat{\mathbf{N}}(\mathcal{S}) \right|} \right). \tag{7.44}$$

By $\mathcal{D}$ we denote $\mathcal{S}^c \cup \{M+1\}$ which includes the index of Y_1. By $\mathcal{S}'$ we denote $\mathcal{V}^c \cup \mathcal{T}^c \cup \mathcal{S} \cup \{0\}$. $I_{DF}(k)$ is as follows:

$$I_{DF}(k) = \frac{1}{2} \log \left(1 + \frac{|h_{0k}|^2 (1-\beta) P}{H(\mathcal{V}, \{k\}) \mathbf{K}(\mathcal{V}) H(\mathcal{V}, \{k\})^T + N_k} \right). \tag{7.45}$$

With β defined as follows:

$$\beta = \frac{\left(\sum_{i=0, k_i \in \mathcal{V}^c}^{i=|\mathcal{V}^c|-1} (-1)^i \sqrt{P_{k_i}} \rho_{0k_i} \left| \begin{array}{cc} \frac{1}{\sqrt{P}} \mathbf{G}^T & \mathbf{K}_i(\mathcal{V}^c) \end{array} \right| \right)}{|\mathbf{K}(\mathcal{V}^c)|} (\mathbf{G} = K(\{0\}, \mathcal{V}^c) = \left[\sqrt{P P_i} \rho_{0i} \right], i \in \mathcal{V}^c)$$

$R_{\mathcal{T}}(\mathcal{S})$ is indeed the rate where those relays in $\mathcal{S}^c$ helps to Y_1 to decode together the source message. The previous rate however is for non-cooperative multiple relay channels. To evaluate the region in (7.28), we have to focus on $R^{(k)}_{\mathcal{T}_k}(\mathcal{S})$. The following corollary presents the region of the cooperative multiple relay channels.

Corollary 18 *For the fading Gaussian multiple relay channel, the following region is inner bound on the capacity:*

$$R \leq \max_{P\in\mathcal{P}} \max_{\mathcal{V}\subseteq\mathcal{N}} \min\left(\max_{\mathcal{T}\subseteq\mathcal{V}} \min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S}), \min_{k\in\mathcal{V}^c} \max_{\mathcal{T}_k\in\Upsilon_k(\mathcal{V})} \min_{\mathcal{S}\subseteq\mathcal{T}_k} R^{(k)}_{\mathcal{T}_k}(\mathcal{S})\right).$$

where

$$R_{\mathcal{T}}(\mathcal{S}) = \frac{1}{2}\log\left(\frac{\left|H(\mathcal{S}',\mathcal{D})\mathbf{K}(\mathcal{S}')H(\mathcal{S}',\mathcal{D})^T + \mathbf{N}(\mathcal{D}) + \hat{\mathbf{N}}(\mathcal{D})\right|}{\left|H(\mathcal{T}^c,\mathcal{D}\cup\mathcal{S})\mathbf{K}(\mathcal{T}^c)H(\mathcal{T}^c,\mathcal{D}\cup\mathcal{S})^T + \mathbf{N}(\mathcal{D}\cup\mathcal{S}) + \hat{\mathbf{N}}(\mathcal{D}\cup\mathcal{S})\right|}\right) - \frac{1}{2}\log\left(\frac{1}{\left|\hat{\mathbf{N}}(\mathcal{S})\right|}\right). \tag{7.46}$$

$$R^{(k)}_{\mathcal{T}_k}(\mathcal{S}) = \frac{1}{2}\log\left(\frac{\begin{vmatrix} H(\mathcal{S}'_k,\mathcal{D}_k)\mathbf{K}(\mathcal{S}'_k)H(\mathcal{S}'_k,\mathcal{D}_k)^T + \mathbf{N}(\mathcal{D}_k) + \hat{\mathbf{N}}(\mathcal{D}_k) & H(\mathcal{S}'_k,\mathcal{D}_k)\mathbf{K}(\mathcal{S}'_k,\mathcal{V}^c) \\ H(\mathcal{S}'_k,\mathcal{D}_k)\mathbf{K}(\mathcal{S}'_k,\mathcal{V}^c)^T & \mathbf{K}(\mathcal{V}^c)\end{vmatrix}}{\left|H(\mathcal{T}_k^c,\mathcal{D}_k\cup\mathcal{S})\mathbf{K}(\mathcal{T}_k^c)H(\mathcal{T}_k^c,\mathcal{D}_k\cup\mathcal{S})^T + \mathbf{N}(\mathcal{D}_k\cup\mathcal{S}) + \hat{\mathbf{N}}(\mathcal{D}_k\cup\mathcal{S})\right| |\mathbf{K}(\mathcal{V}^c)|}\right)$$
$$+ \frac{1}{2}\log\left(1 + \frac{H(\mathcal{S},\{k\})\mathbf{K}(\mathcal{S})H(\mathcal{S},\{k\})^T}{\left(|h_{0k}|^2(1-\beta)P + H(\mathcal{T}_k^c,\{k\})\mathbf{K}(\mathcal{T}_k^c)H(\mathcal{T}_k^c,\{k\})^T + N_k\right)}\right) - \frac{1}{2}\log\left(\frac{1}{\left|\hat{\mathbf{N}}(\mathcal{S})\right|}\right). \tag{7.47}$$

With $\mathcal{D}_k = \mathcal{S}^c\cup\{k\}$ *and* $\mathcal{S}'_k = \mathcal{T}_k\cup\{0\}$.

The general proof for both corollaries is presented in the appendix B.5. These bounds will provide the basis for numerical results of the next sections.

7.5.2 Gaussian Fading Single Relay Channels

We now consider an application example to the fading Gaussian relay channel defined by the following destination and relay outputs, respectively

$$Y_1 = H_1X + H_3X_1 + \mathcal{N}, \tag{7.48}$$

$$Z_1 = \frac{H_2}{D^\alpha}X + \tilde{\mathcal{N}}, \tag{7.49}$$

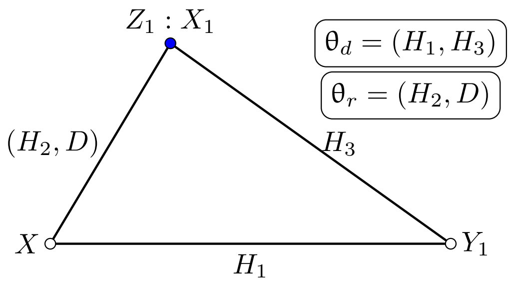


Figure 7.4: Fading Gaussian relay channel

where $\mathcal{N}$ and $\tilde{\mathcal{N}}$ are the additive noises of unit variance, i.i.d. circularly symmetric complex Gaussian RVs with zero-mean and unit variance. In addition to this, (H_1, H_2, H_3) are independent zero mean unit variance circularly symmetric complex Gaussian RVs and D is the random source-to-relay distance, assumed to be uniformly distributed over the interval $[0, 1]$. The average power of source X and relay X_1 must not exceed powers P and P_1, respectively. It is assumed that the source is not aware of the channel measurements (H_1, H_2, H_3), the relay only knows (H_2, D) and the destination is fully aware of all fading coefficients. By choosing $\theta_r = (H_2, D)$ and $\theta_d = (H_1, H_3)$, this model is special case of the composite relay channel. We aim to evaluate the asymptotic error probability based on the bounds derived from Theorem 36 and Corollary 14, and compare them to the usual upper bounds from DF and CF schemes (7.7) and the absolute (CB) lower bound (7.4).

The DF rate for this channel is given by

$$I_{\text{DF}} = \min\left\{\mathcal{C}\left(\beta|H_2|^2 QP\right), \mathcal{C}\left(|H_1|^2 P + |H_3|^2 P_1 + 2\sqrt{\overline{\beta} P P_1}\mathbb{R}e\{H_1 H_3^\star\}\right)\right\}$$

where $0 \leq \beta \leq 1$, $Q = D^{-2\alpha}$ and $\mathcal{C}(x) = \log_2(1 + x)$. The CF rate is given by

$$I_{\text{CF}} = \max\left\{I'_{\text{CF}}, \mathcal{C}\left(\frac{|H_1|^2 P}{|H_3|^2 P_1 + 1}\right)\right\},$$

$$I'_{\text{CF}} = \min\left\{\mathcal{C}\left(|H_1|^2 P + \frac{|H_2|^2 P}{\widehat{N}_2 + 1}\right), \mathcal{C}\left(|H_1|^2 P + |H_3|^2 P_1\right) - \mathcal{C}\left(\frac{1}{\widehat{N}_2}\right)\right\}$$

where the compression noise is $\widehat{N}_2$ for which we have $\hat{Z}_1 = Z_1 + \widehat{N}_2$. Finally, the asymptotically error probability based on DF and CF schemes can be easily calculated by using the above expressions, (7.5) and (7.6). If full CSI is available at the relay, then $\widehat{N}_2$ can be optimally chosen as

$$\widehat{N}_2 = \frac{\left(P\left(|H_1|^2 + Q|H_2|^2\right) + 1\right)}{|H_3|^2 P_1}. \tag{7.50}$$

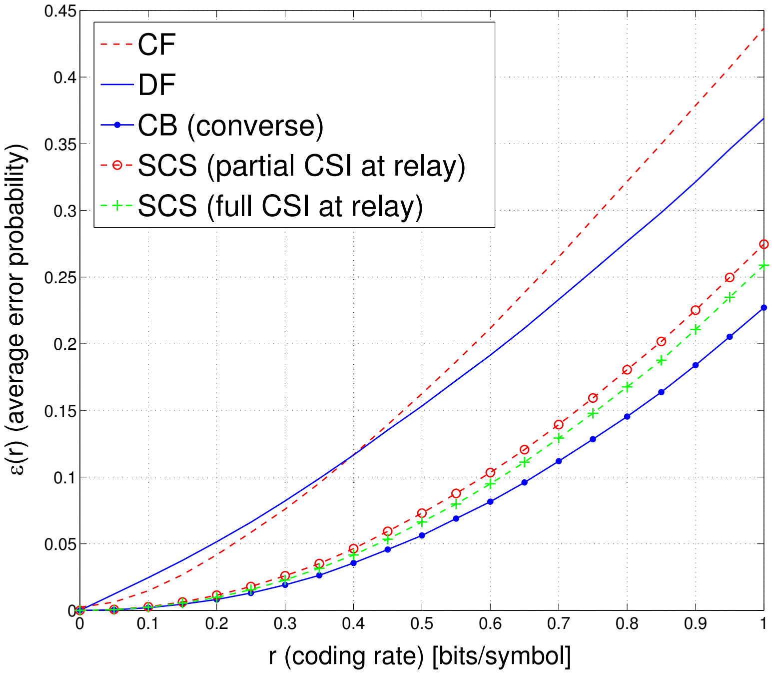


Figure 7.5: The asymptotic error probability $\bar{\epsilon}(r)$ vs. the coding rate r. for single relay channel

But in general only $|H_2|^2$ is available at the relay and so it chooses some constant $\widehat{N}_2$ and optimize the error probability. We saw that the relay can select the proper coding strategy based on $|H_2|$, i.e., selective coding strategy SCS). In this case, if the source-to-relay channel is not good enough for allow decoding and hence DF scheme, so switches to use CF scheme as the relay function. Otherwise, DF scheme is utilized at the relay. It turns out that the optimum decision region $\mathcal{D}_{\text{DF}}$ is given by the set $\mathcal{D}^{\star}_{\text{DF}} = \big\{H_2 : r \leq \mathcal{C}\left(\beta|H_2|^2QP\right)\big\}$.

Fig. 7.5 presents numerical plots of the asymptotic error probability for $P_1 = 2$, $P = 1$. For the case of partially CSI at the relay, we observe that SCS clearly outperforms simple DF and CF schemes. Moreover, notice that full CSI (H_1, H_2, H_3, D) at the relay improves the error probability only through the choice of the best possible compression noise $\widehat{N}_2$. This guarantee that CF scheme cannot perform less favorable than direct transmission. Finally, it can be seen that the upper bound (achievable error probability) resulting from SCS is very close to the max-flow min-cut bound (converse bound for error probability).

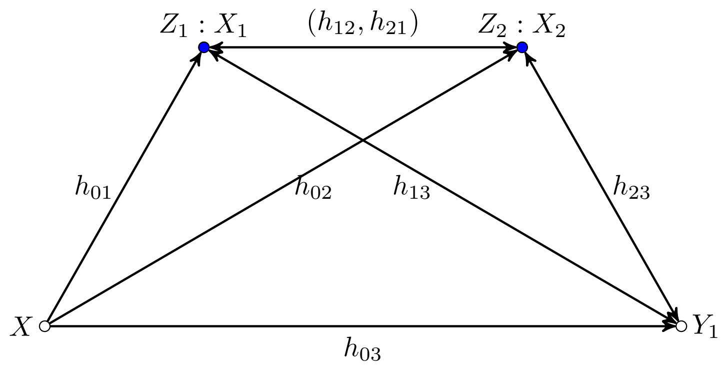


Figure 7.6: Fading Gaussian two relay channel

7.5.3 Gaussian Fading Two Relay Channels

We move to another example which is Gaussian fading two relay channels, Fig 7.6 defined by the following relations:

$$Z_1 = h_{01}X + h_{21}X_2 + \mathcal{N}_1, \tag{7.51}$$

$$Z_2 = h_{02}X + h_{12}X_1 + \mathcal{N}_2, \tag{7.52}$$

$$Y_1 = h_{03}X + h_{13}X_1 + h_{23}X_2 + \mathcal{N}_3, \tag{7.53}$$

Similarly $\mathcal{N}_i$'s are the additive noises, i.i.d. circularly symmetric complex Gaussian RVs with zero-mean. h_{ij}'s are independent zero mean circularly symmetric complex Gaussian RVs which make the fading matrix $\mathbf{H}$. The average power of source X, X_1 and X_2 must not exceed powers P, P_1 and P_2 respectively. It is assumed that the source is not aware of fading coefficients and the relays all know all fading coefficients except h_{i3}'s and the destination is fully aware of all fading coefficients. There are following possibilities to choose the proper cooperative strategy

1. DF Coding (DF): both relays use Decode-and-Forward to transmit the information.
2. CF Coding (CF): both relays use Compress-and-Forward to transmit the information. However here the destination can consider one or both relays as noise to prevent the performance degradation.
3. Mixed Coding: One relay uses DF and the other one uses CF. However there are two possibilities here:

(a) Non-Cooperative Mixed Coding (MC): DF relay decodes the source message independently and without the help of other relays.

(b) Cooperative Mixed Coding (CMC): DF relay exploits the compressed version of CF relay's output to decode the source message.

We first present the achievable rates for each of these cases for a given draw of fading coefficients using the general formula presented in the previous section. For the first case we have the following rate as the inner bound of the capacity ($P_0 = P$):

$$I_{\mathrm{DF}}(\mathbf{H}) = \min\left\{\frac{1}{2}\log\left(1+\frac{|h_{03}|^2P+|h_{13}|^2P_1+|h_{23}|^2P_2+\sum_{0\le i<j\le 2}2\rho_{ij}\sqrt{P_iP_j}\mathbb{R}e\{h_{i3}h_{j3}^\star\}}{N_3}\right),\right.$$
$$\left.\frac{1}{2}\log\left(1+\frac{|h_{01}|^2(1-\beta)P}{N_1}\right), \frac{1}{2}\log\left(1+\frac{|h_{02}|^2(1-\beta)P}{N_2}\right)\right\}, \tag{7.54}$$

where β is defined as follows:

$$\beta = \frac{\rho_{01}^2+\rho_{02}^2-2\rho_{01}\rho_{02}\rho_{12}}{1-\rho_{12}^2}.$$

As for the second case, where all relays use CF coding, the following rate can be achieved:

$$\begin{aligned}
I_{\mathrm{CF}}(\mathbf{H}) = \min&\left\{\frac{1}{2}\log\left(1+P\left(\frac{|h_{01}|^2}{N_1+\hat{N}_1}+\frac{|h_{02}|^2}{N_2+\hat{N}_2}+\frac{|h_{03}|^2}{N_3}\right)\right),\right.\\
&\frac{1}{2}\log\left(1+\frac{|h_{03}|^2P+|h_{13}|^2P_1+|h_{23}|^2P_2}{N_3}\right)-\frac{1}{2}\log\left(1+\frac{N_1}{\hat{N}_1}\right)-\frac{1}{2}\log\left(1+\frac{N_2}{\hat{N}_2}\right),\\
&\frac{1}{2}\log\left(1+\frac{|h_{03}|^2P+|h_{13}|^2P_1}{N_3}+\frac{|h_{02}|^2P+|h_{12}|^2P_1}{N_2+\hat{N}_2}+\frac{PP_1|h_{02}h_{13}-h_{03}h_{12}|^2}{N_3(N_2+\hat{N}_2)}\right)\\
&\quad-\frac{1}{2}\log\left(1+\frac{N_1}{\hat{N}_1}\right),\\
&\frac{1}{2}\log\left(1+\frac{|h_{03}|^2P+|h_{23}|^2P_2}{N_3}+\frac{|h_{01}|^2P+|h_{21}|^2P_2}{N_1+\hat{N}_1}+\frac{PP_2|h_{01}h_{23}-h_{03}h_{21}|^2}{N_3(N_1+\hat{N}_1)}\right)\\
&\quad\left.-\frac{1}{2}\log\left(1+\frac{N_2}{\hat{N}_2}\right)\right\}.
\end{aligned} \tag{7.55}$$

Note that in the previous rate, the destination decodes all the compression indices. Let's denote by $I_{\mathrm{CF}}^{(i)}(\mathbf{H})$ the rate when only the compressed version of the relay i is used. In this case the other relay input appears as noise for the destination and also for the other relay however this so called total noise at the relay and the destination are correlated due to the presence of the other relay's input. For this case the rate is as follows:

$$I_{\mathrm{CF}}^{(1)}(\mathbf{H}) = \min\left\{\frac{1}{2}\log\left(1+\frac{\left(|h_{03}|^2(N_1+\hat{N}_1)+|h_{01}|^2N_3\right)P+PP_2|h_{01}h_{23}-h_{03}h_{21}|^2}{N_3(N_1+\hat{N}_1)+P_2\left(N_3|h_{21}|^2+(N_1+\hat{N}_1)|h_{23}|^2\right)}\right),\right.$$
$$\left.\frac{1}{2}\log\left(1+\frac{|h_{03}|^2P+|h_{13}|^2P_1}{|h_{23}|^2P_2+N_3}\right)-\frac{1}{2}\log\left(1+\frac{N_1}{\hat{N}_1}+\frac{N_3|h_{21}|^2P_2}{\hat{N}_1(|h_{23}|^2P_2+N_3)}\right)\right\}. \tag{7.56}$$

As for the next step, we move to the case of non-cooperative mixed coding. The following rate is achievable when X_1 uses DF:

$$I_{\mathrm{MC}}(\mathbf{H}) = \min\left\{\frac{1}{2}\log\left(1+\frac{|h_{01}|^2(1-\rho_{01}^2)P}{N_1+P_2|h_{21}|^2}\right),\right.$$
$$\frac{1}{2}\log\left(1+\frac{|h_{03}|^2P+|h_{13}|^2P_1+|h_{23}|^2P_2+2\rho_{01}\sqrt{PP_1}\mathbb{R}e\{h_{03}h_{13}^\star\}}{N_3}\right)-\frac{1}{2}\log\left(1+\frac{N_2}{\hat{N}_2}\right),$$
$$\frac{1}{2}\log\left(1+\frac{|h_{03}|^2P+|h_{13}|^2P_1+2\rho_{01}\sqrt{PP_1}\mathbb{R}e\{h_{03}h_{13}^\star\}}{N_3}\right.$$
$$+\frac{|h_{02}|^2P+|h_{12}|^2P_1+2\rho_{01}\sqrt{PP_1}\mathbb{R}e\{h_{02}h_{12}^\star\}}{N_2+\hat{N}_2}$$
$$\left.\left.+\frac{PP_1(1-\rho_{01}^2)|h_{02}h_{13}-h_{03}h_{12}|^2}{N_3(N_2+\hat{N}_2)}+\frac{2\rho_{01}\sqrt{PP_1}\alpha}{N_3(N_2+\hat{N}_2)}\right)\right\}, \tag{7.57}$$

where α is:

$$\alpha = (1-\mathbb{R}e\{h_{13}h_{03}^\star\})(|h_{02}|^2P+|h_{12}|^2P_1)+(1-\mathbb{R}e\{h_{12}h_{02}^\star\})(|h_{03}|^2P+|h_{13}|^2P_1).$$

Finally for the cooperative mixed coding only the first term changes (by $(a)^-$ we mean the negative part of a):

$$I_{\mathrm{CMC}}(\mathbf{H}) = \min\left\{\frac{1}{2}\log\left(1+P(1-\rho_{01}^2)\left(\frac{|h_{01}|^2}{N_1}+\frac{|h_{02}|^2}{N_2+\hat{N}_2}\right)\right)+\right.$$
$$\left(\frac{1}{2}\log\left(1+\frac{|h_{21}|^2P_2}{N_1+|h_{01}|^2(1-\rho_{01}^2)P}\right)-\frac{1}{2}\log\left(1+\frac{N_2}{\hat{N}_2}+\frac{N_1|h_{02}|^2(1-\rho_{01}^2)P}{\hat{N}_2(|h_{01}|^2(1-\rho_{01}^2)P+N_1)}\right)\right)^-,$$
$$\frac{1}{2}\log\left(1+\frac{|h_{03}|^2P+|h_{13}|^2P_1+|h_{23}|^2P_2+2\rho_{01}\sqrt{PP_1}\mathbb{R}e\{h_{03}h_{13}^\star\}}{N_3}\right)-\frac{1}{2}\log\left(1+\frac{N_2}{\hat{N}_2}\right),$$
$$\frac{1}{2}\log\left(1+\frac{|h_{03}|^2P+|h_{13}|^2P_1+2\rho_{01}\sqrt{PP_1}\mathbb{R}e\{h_{03}h_{13}^\star\}}{N_3}\right.$$
$$+\frac{|h_{02}|^2P+|h_{12}|^2P_1+2\rho_{01}\sqrt{PP_1}\mathbb{R}e\{h_{02}h_{12}^\star\}}{N_2+\hat{N}_2}$$
$$\left.\left.+\frac{PP_1(1-\rho_{01}^2)|h_{02}h_{13}-h_{03}h_{12}|^2}{N_3(N_2+\hat{N}_2)}+\frac{2\rho_{01}\sqrt{PP_1}\alpha}{N_3(N_2+\hat{N}_2)}\right)\right\}. \tag{7.58}$$

(7.54) to (7.58) can be used to calculate the bounds on the expected error. As an example consider the Gaussian fading two-relay network, depicted in Fig. 7.6, which is defined by the following relations:

$$Z_1 = \frac{h_{01}}{d^{\alpha}} X + h_{21} X_2 + \mathcal{N}_1, \tag{7.59}$$

$$Z_2 = h_{02} X + h_{12} X_1 + \mathcal{N}_2, \tag{7.60}$$

$$Y_1 = h_{03} X + h_{13} X_1 + h_{23} X_2 + \mathcal{N}_3. \tag{7.61}$$

Define $\mathcal{N}_i$'s to be additive noises, i.i.d. circularly symmetric complex Gaussian RVs with zero-mean and unit variance; let h_{ij}'s be independent zero-mean circularly symmetric complex Gaussian RVs again with unit variance. Set the fading matrix $\mathbf{H}$, and d is the random path-loss. The average power of the source and relay inputs X, X_1 and X_2 must not exceed powers P, P_1 and P_2, respectively. Compression is obtained by adding an additive noise $\hat{Z}_1 = Z_1 + \hat{\mathcal{N}}_1$, $\hat{Z}_2 = Z_2 + \hat{\mathcal{N}}_2$. It is assumed that the source is not aware of fading coefficients, the relays know all fading coefficients except h_{i3}'s and the destination is fully aware of everything. The source and relay powers are respectively 1 and 10.

The possibilities to choose the proper cooperative strategy are as follows. One is when both relays use DF to transmit the information, namely full DF case. Next case is when both relays use CF to transmit the information, call it full CF. Here the destination can consider one or both relays as noise to prevent the performance degradation. Then another case is when one relay uses DF and the other one uses CF, namely Mixed Coding. Finally the relays can select their coding strategy based on the channel parameters, namely Selective Coding.

The figure 7.7 presents the numerical analysis of these strategies. d is chosen with uniform distribution between 0 and 0.1 which means that the first relay is always around the source. Given this assumption, we suppose that the first relay uses DF in case of mixed coding and it is the other relay which is using CF. It can be seen that none of the non-selective strategies like full DF, full CF and Mixed Coding is not in general the best regardless of fading coefficients. However if one lets the relay select their strategy given the fading coefficients, this selective coding will lead to significant improvement compared to the cutset bound and the other strategies.

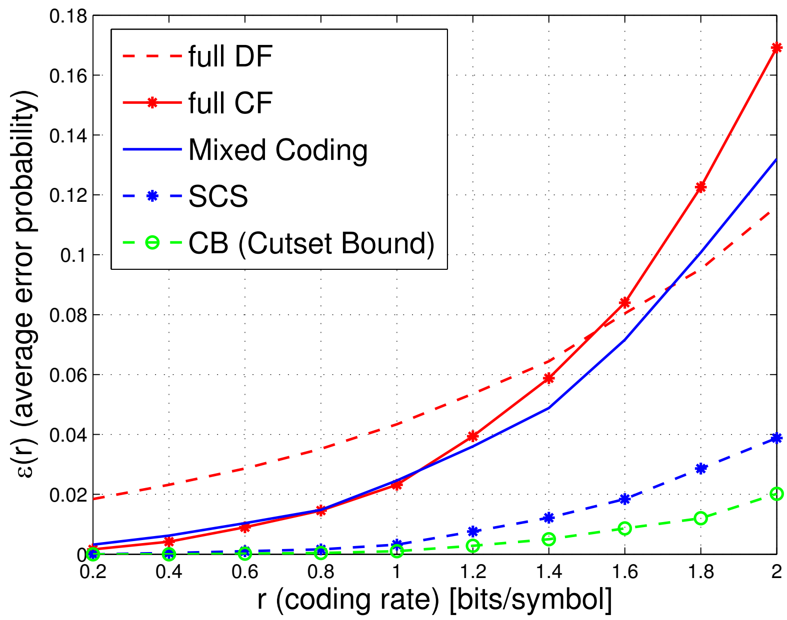


Figure 7.7: Asymptotic error probability $\bar{\epsilon}(r)$ vs. coding rate r. for two relay channels

Chapter 8

The Asymptotic Spectrum of the EP for Composite Networks

binary symmetric averaged channel. On the other hand, due to the unavailability of capacity and ϵ-capacity region in general, the asymptotic spectrum of EP can be bounded using available achievable rate regions and new region called *full error region*. Every code with a rate belonging to this region yields EP equal to one. In this sense for the networks satisfying strong converse condition, asymptotic spectrum of EP coincides with the outage probability. To this purpose it is shown that the cutset bounds provide an upper bound for the full error region. Hence those networks with the cutset bound as the capacity region satisfy strong converse and the notion of outage probability represents the asymptotic spectrum of EP.

8.1 Introduction

Multiterminal networks are the essential part of modern telecommunication systems. Wireless Mobile systems, Computer networks, Sensor and Ad hoc networks are some examples of multiterminal networks. They are usually a combination of common basic networks as broadcast channels, interference channels, multiple access channels and relay channels. The vast development of practical networks during recent years revitalized the interest in network information theory. Particularly multicast networks were studied from various aspects. The cutset bound for the general multicast networks was established in [15,16]. Network coding theorem for graphical multicast network was studied in [121] where the max-flow min-cut theorem for network information flow was presented for the point-to-point communication network. Capacity of networks that have deterministic channels with no interference at the receivers was discussed in [48]. Capacity of wireless erasure multicast networks was determined in [122]. A deterministic approximation of general networks was proposed by Avestimehr *et al.* [19]. Lower bound for general deterministic multicast networks was presented and it was shown that their scheme achieves the cut-set upper bound to within a constant gap. Recently Lim *et al.* discussed a Noisy Network Coding (NNC) scheme for the general multicast networks which includes all the previous bounds [18]. Their approach is based on using Compress-and-Forward as the cooperative strategy [1, 40]. NNC is based on sending the same message in all blocks –repetitive encoding– and non-unique compression index decoding where compression does not rely on binning. Authors in [42] showed that the same region can be achieved using backward decoding and joint decoding of both compression indexes and messages. Kramer *et al.* developed an inner bound for a point-to-point general network using Decode-and-Forward which achieves the capacity of the degraded multicast networks [14].

However the focus has been on the way we can improve the achievability schemes and increase the achievable rate regions and getting bounds on capacity region. Moreover it has been assumed that the probability distribution governing the communication is known to all the users. However the statistics of communication channels are not always available to all terminals. The time-varying nature of wireless channels, e.g. due to fading and user mobility, does not allow all terminals with full knowledge of all channel parameters involved in the communication. Particularly for wireless scenarios the encoders do not know the channels due to fading and user mobility. During years, an ensemble of works has been done on channel models for uncertainty. The compound channel has been introduced by Wolfowitz [21] and continued to attract much of attention from researchers [44, 106]. Nevertheless these models deal with the uncertainty problem in a non-probabilistic way, yielding in general zero capacity for fading channels. Another related model is averaged (or mixed) channels introduced by Ahlswede [45] and further studied in [43] where capacity coincides with that of compound channels.

To deal with the uncertainty in the wireless channels, the notion of composite channels has been introduced. Composite channels consist of a set of channels where the current channel is draw from the set with a probability distribution (PD). This channel was recently investigated in [27, 123, 124]. Channel uncertainty has been studied beforehand in case of fading single user channels [25] and relay channel with oblivious cooperation [28,38] where broadcasting strategy was used to adapt the rate to the channel in function. The composite unlike compound channel introduces a probability distribution to model the uncertainty. Most of the time, in practice arbitrary small error probability (EP) cannot be guaranteed for all channels in the set. Hence an arbitrary small EP constraint may yield a null achievable rate for some composite channels. In these cases the conventional formulation of capacity is not adequate to assess the performance of the channel.

Here instead of finding achievable rates for arbitrary small EP, we fix the rate r and look at the characteristics of EP for the fixed rate r in the composite networks. The rate is fixed and then we look at EP for each channel draw. In this case EP is considered as a random variable which is function of the channel parameter. The notion of outage probability, meaning the probability that a code of rate r cannot be reliably decoded, has been extensively used as a measure of performance for fading scenarios [34]. For instance, it is commonplace that encoders without state information send their messages by using fixed-rate codes and then the outage probability is used to measure the performance. The relation of capacity and outage probability was discussed by Effros *et al.* for the general channels [27]. However it can be seen that the notion of outage probability is not precise enough to characterize the EP for the channels not satisfying strong converse. Different notions are introduced to study EP. The notion of *asymptotic spectrum of EP* for (r, ϵ) is defined as the asymptotic probability that EP falls over ϵ. It is shown that this notion implies other available notions used to measure the performance of composite networks.

However ϵ-capacity is not known in general for the general networks. It is shown that the asymptotic spectrum of EP can be bounded using available achievable rate regions. Particularly the concept of *full error region* is defined for the general networks where every code with transmission rates in this region yields EP equal to one. For the composite networks this region is also a random region. The asymptotic spectrum of EP at rate r is bounded by the probability that the full error region includes the transmission rates. For the general point to point networks the full error region is reduced to a value called *full error capacity*. Furthermore, it turns out that for channels satisfying the strong-converse property [43] the EP coincides with the outage probability. In this sense the performance of composite networks can be studied using available achievable rate regions and full error regions. Specifically the channels satisfying strong converse are of huge interest because the asymptotic spectrum of EP coincides with the outage probability for them.

Various achievable rates have been developed for the general networks however strong converse and full error regions are yet to be studied extensively. Strong converse for the discrete memoryless single user channels was proved by various authors [43, 125–128]. Ash provided an example for channels that do not satisfy strong converse [129]. ϵ-capacity of binary symmetric averaged channels was derived by Kieffer [46]. The necessary and sufficient condition for a channel to satisfy strong converse was provided by Han-Verdu in [43, 44] using information spectrum methods. The strong converse for multiple access channels and broadcast channels has been also proved in [10, 43, 130, 131].

In this chapter we prove that the closure of cutset bounds for the discrete memoryless networks falls into the full error region. In other words for each code with transmission rates not satisfying the cutset bound the probability of error tends to one. This will provide a bound on the asymptotic spectrum of EP for the composite discrete memoryless networks. As a result the multiterminal networks that have the cutset bound as capacity satisfy strong converse and thus the asymptotic spectrum of EP coincides with the outage probability. To prove this result the information spectrum method is used.

This chapter is organized as follows. In the next section the main definitions are provided along with the composite relay channels. Then the notion of EP is discussed and various measure of performance is studied using EP. The result of these measures are compared in case of binary symmetric averaged channel. In the next section we study the bounds on the asymptotic spectrum of EP. Particularly the bounds are discussed for the case of the composite discrete memoryless networks and it is shown that the cut-set bound provides an upper bound on the full error capacity. The sketch of proof is relegated to the appendix. It is shown that for the composite deterministic networks outage probability is a good measure.

8.2 Main Definitions and Background

8.2.1 Notation and Background

For the rest, we denote the random variables (RV) either by upper case letters $Y, X, ...$, bold letters $\mathbf{M}_n$ or by bold Greek letters $\boldsymbol{\epsilon}, \boldsymbol{\theta}, ...$ The particular variables are denoted by lower case letters and ordinary Greek letters. By $\underline{\mathrm{X}}$ we denote $X^n = (X_1, \dots, X_n)$ where subscripts are time indices. Similarly by $\underline{\mathrm{X}}_k$ we denote the channel input of the node k, $(X_{1,k}, \dots, X_{n,k})$ and $\underline{\mathrm{Y}}_k$ denotes the vector of length n indicating the channel observations at node k . By $\underline{\mathrm{X}}_{\mathcal{S}}$ we denote $(\underline{\mathrm{X}}_k)_{k\in\mathcal{S}}$.

The *information density*[1] is defined by [44]

$$i(\mathbf{M}_n;\underline{\mathrm{Y}}) \triangleq \log \frac{\mathbb{P}_{Y^n|\mathbf{M}_n}(\underline{\mathrm{Y}}|\mathbf{M}_n)}{\mathbb{P}_{Y^n}(\underline{\mathrm{Y}})},$$

for an arbitrary sequence of n-dimensional outputs $\underline{\mathrm{Y}} = (Y_1,\dots,Y_n) \in \mathscr{Y}^n$ and $\mathbf{M}_n$ an uniform RV over the index set $\mathcal{M}_n = \{1,\dots,M_n\}$. We will use lim sup*in probability* of the random sequence Z_n, which is defined as

$$\text{p-}\limsup_{n\to\infty} Z_n \triangleq \inf\left\{\beta : \lim_{n\to\infty} \Pr\{Z_n > \beta\} = 0\right\},$$

and equally the lim inf *in probability* of the random sequence Z_n, which is defined as

$$\text{p-}\liminf_{n\to\infty} Z_n \triangleq \sup\left\{\alpha : \lim_{n\to\infty} \Pr\{Z_n < \alpha\} = 0\right\}.$$

8.2.2 Definition of Composite Multiterminal Network (CMN)

We begin with the description of the Composite Multiterminal Network (CMN) with m-nodes, which is characterized by a set of conditional probability distributions (PDs):

$$\mathcal{W}_\theta = \left\{\mathbb{P}_{Y^n_{1\theta}\cdots Y^n_{m\theta}|X^n_{1\theta}\cdots X^n_{m\theta}} : \mathscr{X}^n_1 \times \cdots \times \mathscr{X}^n_m \longmapsto \mathscr{Y}^n_1 \times \cdots \times \mathscr{Y}^n_m\right\}_{n=1}^{\infty}$$

indexed with any vector of parameters $\theta \in \Theta$, and where each node $i = \{1,\dots,m\}$ is equipped with a transmitter $\underline{\mathrm{X}}_{i\theta} \in \mathscr{X}^n_i$ and a receiver $\underline{\mathrm{Y}}_{i\theta} \in \mathscr{Y}^n_i$, as described in Fig. 8.1. The entries of the matrices $\mathbb{P}_{Y^n_{1\theta}\cdots Y^n_{m\theta}|X^n_{1\theta}\cdots X^n_{m\theta}}$ will often be written as $\mathbb{W}^n_\theta$. In addition, the network is said to be non-stationary memoryless if the joint PD of the multiterminal channel decomposes as

$$\mathbb{P}_{Y^n_{1\theta}\cdots Y^n_{m\theta}|X^n_{1\theta}\cdots X^n_{m\theta}}(\underline{\mathrm{y}}_1,\cdots,\underline{\mathrm{y}}_m|\underline{\mathrm{x}}_1,\cdots,\underline{\mathrm{x}}_m) = \prod_{t=1}^{n} \mathbb{W}_{\theta,t}(y_{1t},\cdots,y_{mt}|x_{1t},\cdots,x_{mt})$$

with source inputs $\underline{\mathrm{x}}_k = (x_{k1},\dots,x_{kn}) \in \mathscr{X}^n_k$ and channel outputs $\underline{\mathrm{y}}_k = (y_{k1},\dots,y_{kn}) \in \mathscr{Y}^n_k$, for all $k = \{1,\dots,m\}$, where $\left\{\mathbb{W}_{\theta,t}\right\}_{t=1}^{\infty}$ is a sequence of single-letter multiterminal channels. Similarly, the channel is said to be stationary and memoryless if $\mathbb{W}_{\theta,t} = \mathbb{W}_\theta$ for all $t = \{1,2,\dots,n\}$.

1. Let $\mathbb{P}_{Y^n|\mathbf{M}_n} = \dfrac{\mathbb{P}_{Y^n|\mathbf{M}_n}(d\underline{\mathrm{Y}}|\mathbf{M}_n)}{\mathbb{P}_{Y^n}(d\underline{\mathrm{Y}})}$ denote the *Radon-Nikodym derivative* between two probability measures on $\mathscr{Y}^n$ with values on a singular set assumed conventionally to be $+\infty$. Then, $\dfrac{\mathbb{P}_{Y^n|\mathbf{M}_n}(d\underline{\mathrm{Y}}|\mathbf{M}_n)}{\mathbb{P}_{Y^n}(d\underline{\mathrm{Y}})}$ is defined to be $\mathbb{P}_{Y^n|\mathbf{M}_n}$, which is obviously a random variable (RV).

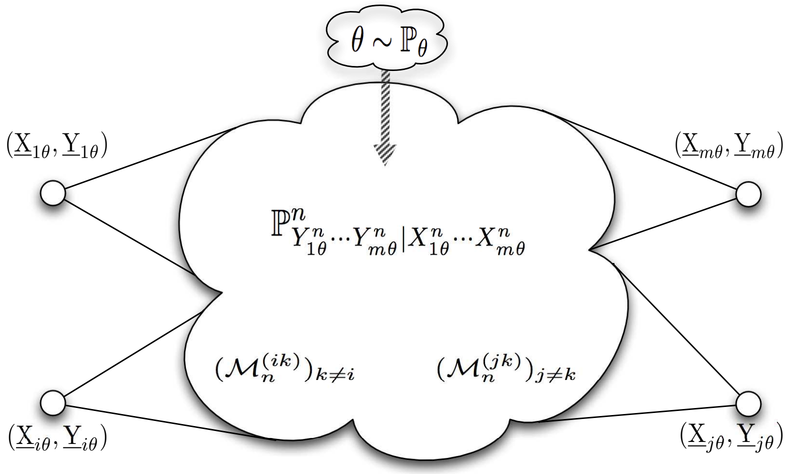


Figure 8.1: Composite Multiterminal Network (CMN)

Let $\mathbb{P}_\theta$ denote any arbitrary PD on the set of network parameters (or channel indices) Θ. Before the communication starts a channel index $\theta \in \Theta$ is assumed to be drawn from $\mathbb{P}_\theta$ and remains fixed during the entire transmission.

The set $\mathcal{M}_n^{(ki)} \triangleq \{1, \ldots, M_n^{(ki)}\}$ represents the set of possible messages to be sent (in n channel uses) from source k to the i-th destination with $i \in \{1, \ldots, m\} \setminus \{k\}$. If there are no messages intended to node i from node k we will set $\mathcal{M}_n^{(ki)} = \emptyset$.

Definition 16 (code and error probability) *An $\left(n, M_n^{(kj)}, (\epsilon_{n,\theta})_{\theta\in\Theta}\right)$-code for the CMN consists of:*

– *A sequence of encoding mappings for $t = \{1, \ldots, n\}$ at each node $k \in \{1, \ldots, m\}$,*

$$\varphi_{t,\theta}^{(k)} : \bigotimes_{i=\{1,\ldots,m\}\setminus\{k\}} \mathcal{M}_n^{(ki)} \otimes \mathscr{Y}_k^{t-1} \longmapsto \mathscr{X}_k$$

where $\mathcal{M}_n^{(ki)}$ is the message set from source node k intended to destination node i, for every $i = \{1, \ldots, m\} \setminus \{k\}$. Transmitted symbols $x_{kt} = \varphi_{t,\theta}^{(k)}(\underline{w}, y_k^{t-1})$ are function of past received symbols y_k^{t-1} and all messages to be sent from node k

$$\underline{w} \in \bigotimes_{i=\{1,\ldots,m\}\setminus\{k\}} \mathcal{M}_n^{(ki)}.$$

– *A decoder mapping at each node $k \in \{1, \ldots, m\}$,*

$$\phi_{n,\theta}^{(jk)} : \mathscr{Y}_k^n \otimes \bigotimes_{i\in\{1,\ldots,m\}\setminus\{k\}} \mathcal{M}_n^{(ki)} \longmapsto \mathcal{M}_n^{(jk)}$$

for all source node $j \neq k \in \{1,\dots,m\}$. *This decoding mapping is related to the message destined for destination node* k *from source node* j. *Decoding sets corresponding to each decoding mapping are defined by* $\mathcal{D}_{l,\theta}^{(jk)} \triangleq \phi_{n,\theta}^{(jk)^{-1}}(l)$ *for all messages* $l \in \mathcal{M}_n^{(jk)}$, *which corresponds to the decoding sets for messages* l *intended to node* k *from node* j.

– *The error event* $\mathcal{E}_\theta^{(jk)}(l) \triangleq \left\{Y_{k\theta}^n \notin \mathcal{D}_{l,\theta}^{(jk)}\right\}$ *for all pairs* $j \neq k \in \{1,...,m\}$ *and every* $l \in \mathcal{M}_n^{(jk)}$ *is defined as the event that the message* l *from node* j *cannot be correctly decoded at destination* k. *Hence the corresponding probability of error, based on each decoding set, is defined by*

$$\mathbf{e}_{n,\theta}^{(jk)}(l) \triangleq \Pr\left(Y_{k\theta}^n \notin \mathcal{D}_{l,\theta}^{(jk)} | \mathbf{M}_n^{(jk)} = l\right), \tag{8.1}$$

where $\mathbf{M}_n^{(jk)}$ *denotes the RV corresponding to the transmitted message. Assuming a uniform PD over the message sets, the average error probability (EP) is defined as*

$$\epsilon_{n,\theta}^{(jk)} \triangleq \frac{1}{M_n^{(jk)}} \sum_{l=1}^{M_n^{(jk)}} \mathbf{e}_{n,\theta}^{(jk)}(l) \tag{8.2}$$

and the maximum EP as

$$\epsilon_{\max,n,\theta}^{(jk)} \triangleq \max_{l \in \mathcal{M}_n^{(jk)}} \mathbf{e}_{n,\theta}^{(jk)}(l) \geq \epsilon_{n,\theta}^{(jk)}. \tag{8.3}$$

Therefore the error event for the network is the union of all error events $\mathcal{E}_\theta^{(jk)}(l)$ *over all sources* j *and destinations* k *with corresponding messages* l. *The EP of the network writes as*

$$\epsilon_{n,\theta} \triangleq \Pr\left(\bigcup_{j\neq k \in \{1,\dots,m\}} \bigcup_{l \in \mathcal{M}_n^{(jk)}} \left\{Y_{k\theta}^n \notin \mathcal{D}_{l,\theta}^{(jk)}, \mathbf{M}_n^{(jk)} = l\right\}\right) \tag{8.4}$$

where it is easy to check that

$$\begin{aligned} \epsilon_{n,\theta} &\leq \sum_{j\neq k \in \{1,\dots,m\}} \epsilon_{n,\theta}^{(jk)} \\ &\leq \sum_{j\neq k \in \{1,\dots,m\}} \epsilon_{\max,n,\theta}^{(jk)}. \end{aligned} \tag{8.5}$$

Through this chapter average EP will be used. Notice that in the CMN setting the error probabilities $\epsilon_{\max,n,\theta}^{(jk)}$, $\epsilon_{n,\theta}^{(jk)}$ and $\epsilon_{n,\theta}$ are RVs, indeed they are functions of the random channel index $\uptheta$. For instance, for any fix $\uptheta = \theta$ the CMN is reduced to a conventional multiterminal network. For the rest, we also use the notion of C to designate a code.

One thing is important to note here. In general it is reasonable to assume if $\mathcal{M}_n^{(ki)} \neq \emptyset$ for all k, i, in other words if each node is sending something and each node is receiving something then each node can be cognizant of $\uptheta$ because the encoder and the decoder functions can now be chosen based on θ, i.e. CSIT and CSIR are both available. In this case, if everybody knows which channel is in operation, then the problem is reduced to usual multiterminal network and there is no need for further study of the subject beyond usual methods.

However in the real networks, not each node transmits and receives in the same time and moreover not each node is source and destination. If the node k is only transmitter which means that $\mathcal{Y}_k = \emptyset$, then there is no way for it to have information of the channel in operation and so it is indeed oblivious. On the other hand if the node k is only receiver which means that $\mathcal{X}_k = \emptyset$, then there is no way for it to send the information regarding its observation of the channel to other users, and it means that the other users are necessarily oblivious to CSI of this user. On the other hand there are users which serve as relays which means that $\mathcal{M}_n^{(ik)} = \emptyset$ and $\mathcal{M}_n^{(ki)} = \emptyset$ for all $i \neq k$. These users can only partly know CSI and they are not naturally cognizant of CSI of the users without channel inputs.

The model which seems to be more adapted to the practical scenario consists of three types of nodes as in Fig. 8.2. Those nodes with index belonging to the set $\mathcal{T}$ are only transmitters and sources. Node i where $i \in \mathcal{T}$ transmits $\underline{X}_i$ independent of $\uptheta$ because it cannot have access to it. Those nodes with index in $\mathcal{D}$ are only destinations. These nodes cannot transmit their own observation to other nodes however they can have access to CSI of all the transmitters. Finally those nodes with index in $\mathcal{R}$ simultaneously transmit and receive information. Relays are essentially part of these nodes. They cannot have access to full CSI because of the presence of nodes in $\mathcal{D}$ and can only be partly aware of CSI. We can suppose that $\uptheta$ is indeed composed of two parts $\uptheta_d$ and $\uptheta_r$ where the nodes in $\mathcal{R}$ can only be cognizant of $\uptheta_r$. For the rest it is always presupposed that we are dealing with the kind of network where full CSI is not available at each node. We next define the notion of achievability and capacity region for the CMN.

Definition 17 (achievability and capacity) *An error probability* $0 \leq \epsilon < 1$ *and a vector of rates* $\underline{r} = (r_{jk})_{j \neq k \in \{1,\dots,m\}}$ *are said achievable for the multiterminal network, if*

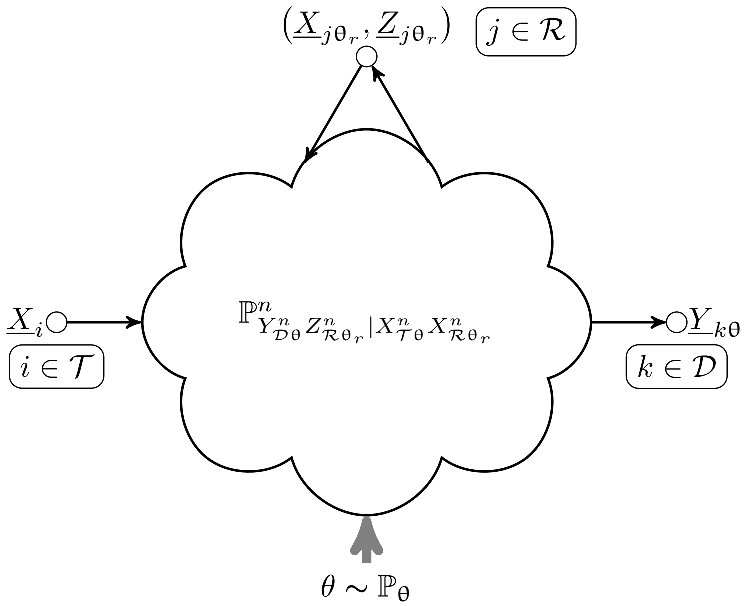


Figure 8.2: Composite Multiterminal Network (CMN) with $\theta = (\theta_r, \theta_d)$

there exists an $\big(n, M_n^{(jk)}, (\epsilon_{n,\theta})_{\theta\in\Theta}\big)$*-code such that for all pairs* $j \neq k \in \{1, \dots, m\}$

$$\liminf_{n\to\infty} \frac{1}{n} \log M_n^{(jk)} \geq r_{jk} \tag{8.6}$$

and

$$\limsup_{n\to\infty} \sup_{\theta\in\Theta} \epsilon_{n,\theta} \leq \epsilon. \tag{8.7}$$

Then the ϵ*-capacity region* $\mathcal{C}_\epsilon$ *is defined by the region composed of all achievable rates satisfying* (8.7)*. Similarly,* $(\epsilon, \epsilon_{n,\theta})$ *in definition* (8.7) *can be replaced by* $(\epsilon^{(jk)}, \epsilon_{n,\theta}^{(jk)})$ *which corresponds to the tolerated error from the source node* j *to the destination node* k*. Hence the notion of* $\underline{\epsilon}$*-capacity region* $\mathcal{C}_{\underline{\epsilon}}$ *can be defined in a similar way, by setting the vector* $\underline{\epsilon} = (\epsilon^{(jk)})_{j\neq k\in\{1,\dots,m\}}$*. Finally, the notion of capacity region is defined as*

$$\mathcal{C} \triangleq \lim_{\underline{\epsilon}\to 0} \mathcal{C}_{\underline{\epsilon}}.$$

Notice that the reliability function (8.7) has been chosen in the strongest sense. In general, the preceding definitions may lead to null achievable rates because every node have to fix the rate such that $\epsilon_{n,\theta}$ is less than ϵ and the worst possible index $\theta \in \Theta$ can have zero ϵ-capacity, for any $0 \leq \epsilon < 1$. Furthermore, in wireless networks it is rare to have non-zero rate for the worst channel draws but yet it is desirable to send the information and measure performance somehow. As we will see in next section, by relying on the PD $\mathbb{P}_\theta$, several different notions for reliability can be suggested.

8.2.3 Reliability Function for Composite Networks

An alternative approach is the study of the behavior of error probability $\epsilon^{(jk)}_{n,\theta}$, $\epsilon_{n,\theta}$ as n goes to infinity for fixed rates. For the rest, we will focus on the study of the error probability. We assume that $\epsilon_{n,\theta}$ converges to ϵ_θ almost everywhere to facilitate of working with limits. However the results remain valid if we replace this with the much weaker assumption that $\epsilon_{n,\theta}$ converges in distribution to ϵ_θ. Because $\epsilon_{n,\theta}$ is uniformly integrable, the working with limits stays intact. [2]

Definition 18 (reliability functions) *The value $0 \leq \epsilon < 1$ is said to be achievable for a tuple of rates $\underline{r} = (r_{jk})_{j\neq k \in \{1,\dots,m\}}$ based on following reliability functions, if there exists an $\left(n, M_n^{(jk)}, (\epsilon_{n,\theta})_{\theta\in\Theta}\right)$-code such that for all pairs $j \neq k \in \{1,\dots,m\}$ the rates satisfy*

$$\liminf_{n\to\infty} \frac{1}{n} \log M_n^{(jk)} \geq r_{jk},$$

and $\epsilon_{n,\theta}$ satisfies certain reliability condition, listed in below.

- *If we look at the limit of $\epsilon_{n,\theta}$ as $n \to \infty$ using the notion of Convergence almost everywhere (a.e), then ϵ is said to be achievable if the limit is less than or equal to ϵ almost everywhere. This means that:*

$$\mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} \leq \epsilon) = 1. \tag{8.8}$$

It guarantees that for all subsets in Θ having non-zero measure the asymptotic EP will be not larger than ϵ. Hence based on a.e. convergence, the drawback caused by zero measure events can be removed and the reliability function relaxed. However it can be seen that:

$$\begin{aligned}
\mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > \epsilon) &= \mathbb{E}(\mathbf{1}[\lim_{n\to\infty} \epsilon_{n,\theta} > \epsilon]) \\
&= \mathbb{E}(\lim_{n\to\infty} \mathbf{1}[\epsilon_{n,\theta} > \epsilon]) \\
&\overset{(a)}{=} \lim_{n\to\infty} \mathbb{E}(\mathbf{1}[\epsilon_{n,\theta} > \epsilon]) \\
&= \lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \epsilon)
\end{aligned}$$

2. It is always possible to use lim sup to assure the convergence however the equalities are not all valid and turn into inequalities for some cases.

where (a) comes from Lebesgue dominated convergence theorem. This means that the notion convergence in probability is equivalent to usually looser notion of Convergence in probability for this case. This also means that ϵ is said to be achievable if:

$$\lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \epsilon) = 0$$

The achievable EP can also be characterized by

$$\epsilon_{\text{-}p}(\underline{r}, \mathsf{C}) = p\text{-}\limsup_{n\to\infty} \epsilon_{n,\theta} \tag{8.9}$$

where we have

$$p\text{-}\limsup_{n\to\infty} \epsilon_{n,\theta} = \inf\left\{\alpha : \lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \alpha) = 0\right\}. \tag{8.10}$$

This means that ϵ is achievable if there is a code C such that ϵ is bigger than or equal to $\epsilon_{\text{-}p}(\underline{r}, \mathsf{C})$ which means that with probability 1, $\epsilon_{n,\theta}$ cannot exceed ϵ.
However, the problem with these notions is that for each $\epsilon < 1$ there may be no code satisfying $\mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > \epsilon) = 0$ or $\lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \epsilon) = 0$. In other words, for each $0 \leq \epsilon < 1$ and each code, there is non-zero probability that the error falls over it (e.g. wireless networks). However the condition in (8.10) can be relaxed to

$$\epsilon_{\text{-}\delta}(\underline{r}, \mathsf{C}) = \delta\text{-}\limsup_{n\to\infty} \epsilon_{n,\theta} \tag{8.11}$$

where

$$\delta\text{-}\limsup_{n\to\infty} \epsilon_{n,\theta} = \inf\left\{\alpha : \lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \alpha) \leq \delta\right\}, \tag{8.12}$$

for any constant $0 \leq \delta < 1$. We call this notion $\delta-$achievable EP. This means that ϵ is $\delta-$achievable if there is a code C such that $\epsilon_{\text{-}\delta}(\underline{r}, \mathsf{C})$ is less than ϵ which means that there is a code such that $\epsilon_{n,\theta}$ is less than ϵ with at least the probability $1 - \delta$.

- *In the same lines, the average error probability can be characterized for a code C as follows*

$$\bar{\epsilon}(\underline{r}, \mathsf{C}) = \lim_{n\to\infty} \mathbb{E}_\theta[\epsilon_{n,\theta}]. \tag{8.13}$$

This definition may lead to the definition according to which ϵ is said to be achievable if there is a code C such that ϵ is bigger than the average error which implies the existence of codes with EP less than ϵ in $\mathcal{L}^1$ but not everywhere, meaning that for some $\theta \in \Theta$ the asymptotic EP may fall over $\bar{\epsilon}$. This shows that the average error is not precise enough to characterize the error probability, as it will be shown later. It should be mentioned here that the expected EP is equivalent to the definition of EP for the averaged channel in [45].

– *The throughput error probability is defined for a code* C *by*

$$\epsilon_T(\underline{r}, \mathsf{C}) = \sup_{0\leq\alpha<1} \lim_{n\to\infty} \alpha\, \mathbb{P}_\theta(\epsilon_{n,\theta} > \alpha). \tag{8.14}$$

The throughput EP takes into account the desired error probability ϵ *and the probability the error falls over it. So if the error probability falls over a big* ϵ *with small probability, the throughput EP takes it into the account.* ϵ *is said achievable with respect to this measure if there is code* C *such that* ϵ *is bigger than* $\epsilon_T(\underline{r}, \mathsf{C})$*.*

It is particularly interesting to define *the smallest achievable EP*, characterized by

$$\epsilon_{\text{-p}}(\underline{r}) = \inf_{\mathsf{C}} \epsilon_{\text{-p}}(\underline{r}, \mathsf{C}) = \inf_{\mathsf{C}} \text{p-}\limsup_{n\to\infty} \epsilon_{n,\theta} \tag{8.15}$$

where the infimum is taken over all codes. This means that for ϵ smaller than $\epsilon_{\text{-p}}(\underline{r})$, for all codes, we have:

$$\lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \epsilon) > 0.$$

The smallest achievable EP is a good indicator of the composite channels. Notice that the meaning of the smallest achievable EP ϵ is that information can be sent at rate $\underline{r}$ with diminishing probability that EP falls over ϵ.

However, the same problem with this notion is that for some cases this value is not non-trivial. Then the same idea can be used here to relax the previous condition and to define $\delta-$*smallest achievable EP*:

$$\epsilon_{\text{-}\delta}(\underline{r}) = \inf_{\mathsf{C}} \epsilon_{\text{-}\delta}(\underline{r}, \mathsf{C}) = \inf_{\mathsf{C}} \delta\text{-}\limsup_{n\to\infty} \epsilon_{n,\theta} \tag{8.16}$$

where infimum is taken again over all codes. This means that for ϵ bigger than $\epsilon_{\text{-}\delta}(\underline{r})$, there is a code such that $\epsilon_{n,\theta}$ is less than ϵ with at least the probability $1-\delta$.

Therefore, on one hand the expected EP (8.13) may not always be the adequate reliability function for the CMN, but on the other hand (8.9) may yield very pessimistic rates. Furthermore, depending on the target application there may be different reliability functions of interest for composite models. Hence the question whether there exists an universal measure of reliability from which the others can be derived arises. The next section introduces such fundamental quantity, referred to as the *asymptotic spectrum of error probability* (ASEP).

8.2.4 Asymptotic Spectrum of Error Probability

In the previous section, we discussed different notions of achievability for an error. The smallest achievable error was defined for a fixed $\underline{r}$ using different criteria. Now we investigate the asymptotic cumulative PD of EP for the fixed vector of transmission rates $\underline{r} = (r_{jk})_{j \neq k \in \{1,\ldots,m\}}$, which is given by

$$\lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} \leq \epsilon),$$

for every $0 \leq \epsilon \leq 1$.

Definition 19 (asymptotic spectrum of EP) *For every $0 \leq \epsilon \leq 1$ and transmission rates $\underline{r} = (r_{jk})_{j \neq k \in \{1,\ldots,m\}}$, the asymptotic spectrum of EP for a given code C, $\mathcal{E}(\underline{r}, \epsilon, \mathsf{C})$ is defined as*

$$\mathcal{E}(\underline{r}, \epsilon, \mathsf{C}) = \lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \epsilon). \tag{8.17}$$

The asymptotic spectrum of EP for CMN is defined as:

$$\mathcal{E}(\underline{r}, \epsilon) = \inf_{\mathsf{C}} \lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \epsilon) \tag{8.18}$$

where the infimum is taken over all $\left(n, M_n^{(jk)}, (\epsilon_{n,\theta})_{\theta\in\Theta}\right)$-codes with rates satisfying

$$\liminf_{n\to\infty} \frac{1}{n} \log M_n^{(jk)} \geq r_{jk},$$

for all pairs $j \neq k \in \{1, \ldots, m\}$.

The notion of $\mathcal{E}(\underline{r}, \epsilon)$ indicates intuitively what is the smallest probability that the error falls over ϵ. It will be shown that this notion is the most general measure for the performance of composite networks and implies all other notions.

It is also interesting to see in particular that for a given transmission rate $\underline{r}$, what the possible probability of errors are; in other words, to find if the asymptotic value of $\epsilon_{n,\theta}$ is less than a desired value or not. This idea underlies the idea of achievable error for the composite multiterminal networks.

The next theorem provides a relation between the asymptotic spectrum of EP and the other notions introduced before.

Theorem 41 *For the composite networks with rate $\underline{r}$, the asymptotic spectrum of EP implies other reliability functions introduced before. The smallest achievable EP and*

$\delta-$smallest achievable EP can be obtained as follows:

$$\epsilon_{\text{-}p}(\underline{r}) = \inf\{0 \leq \epsilon < 1 : \mathcal{E}(\underline{r}, \epsilon) = 0\}$$
$$\epsilon_{\text{-}\delta}(\underline{r}) = \inf\{0 \leq \epsilon < 1 : \mathcal{E}(\underline{r}, \epsilon) \leq \delta\}.$$

The throughput EP for a code C*is obtained as follows*

$$\epsilon_T(\underline{r}, \mathsf{C}) = \sup_{0 \leq \epsilon < 1} \epsilon \mathcal{E}(\underline{r}, \epsilon, \mathsf{C}).$$

Finally the expected EP for a code C *is obtained as follows:*

$$\bar{\epsilon}(\underline{r}, \mathsf{C}) = \int_0^1 \mathcal{E}(\underline{r}, \epsilon, \mathsf{C}) d\epsilon.$$

Proof The proof of first three equalities follows directly from the definition. For the last inequality, using the fact that $\epsilon_{n,\theta}$ is positive and bounded we have:

$$\begin{aligned}\bar{\epsilon}(\underline{r}, \mathsf{C}) &= \lim_{n\to\infty} \mathbb{E}[\epsilon_{n,\theta}] \\ &\overset{(a)}{=} \mathbb{E}[\lim_{n\to\infty} \epsilon_{n,\theta}] \\ &= \int_0^{+\infty} \mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > t) dt \\ &= \int_0^1 \mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > t) dt \\ &= \int_0^1 \lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > t) dt\end{aligned}$$

where (a) comes from Lebesgue dominated convergence theorem. And this will conclude the proof.

The previous theorem states that the expected EP is not necessarily achievable in the strict sense for a given rate r. Therefore, it is possible in general that EP falls over the expected EP. This observation shows that the expected error, though indicative, would not always be a proper measure for EP.

We are interested in behavior of the error probability $\epsilon_{n,\theta}$ which is itself a random variable. In particular the characterization of the asymptotic spectrum of EP is of great interest for the random networks. This will give us a better criteria over the error probability that can be achieved in a channel compared to the outage probability. Specifically, this will be useful for the cases where the transmitter fixes a rate r regardless of the channel in operation. This is usually the case in practice where the rate is determined by the media in use. However it is interesting to see the relation between outage probability and the asymptotic spectrum of EP. In the next section we consider how to characterize the asymptotic spectrum of EP and the notion is interrelated with the outage probability.

8.3 Main Results

Consider a general composite channel. It can be observed that the probability distribution of $\epsilon_{n,\theta}$ as $n \to \infty$ is directly related to the probability that the rate vector $\underline{r}$ falls into ϵ-capacity region $\mathcal{C}_{\epsilon,\theta}$, where $\mathcal{C}_{\epsilon,\theta}$ is a random set with θ as random parameter. Suppose that the transmission is made at the rate of $\underline{r}$ over a non-composite channel. Then if the code achieves the error probability ϵ then its rate should necessarily belong to ϵ-capacity region. On the other hand if the rate belongs to ϵ-capacity region, then there is a code such that it achieves the error probability ϵ.

However in the case of composite networks, the transmitter, unaware of θ, has a single code for all θ. Then if the rate does not belongs to $\mathcal{C}_{\epsilon,\theta}$ for a θ, then $\epsilon_{n,\theta}$ will exceed ϵ for sure. But if the rate belongs to $\mathcal{C}_{\epsilon,\theta}$ for a θ, then it is not guaranteed that $\epsilon_{n,\theta}$ will not exceed ϵ. Because although there is code such that the EP is less than ϵ but this may not be the code used by the transmitter. This will lead to the following theorem.

Theorem 42 *For composite multiterminal network with the random parameter θ we have:*

$$\mathbb{P}_\theta(\limsup_{n\to\infty} \epsilon_{n,\theta} > \epsilon) \geq \mathbb{P}_\theta(\underline{r} \notin \mathcal{C}_{\epsilon,\theta}), \tag{8.19}$$

where $\mathcal{C}_{\epsilon,\theta}$ is the ϵ-capacity of the network W_θ for a given θ, and $0 \leq \epsilon < 0$.

Proof According to the definition, for each θ, $\underline{r}$ is inside $\mathcal{C}_{\epsilon,\theta}$ if $\limsup_{n\to\infty} \epsilon_{n,\theta} \leq \epsilon$. This gives us the proof for the theorem.

In the theorem 42 ϵ can be replaced with $\epsilon^{(ij)}$ and respectively $\epsilon_{n,\theta}$ to $\epsilon_{n,\theta}^{(ij)}$. This change of the definition also changes the definition of ϵ-capacity to $\underline{\epsilon}$-capacity and thereby the theorem remains valid under the change.

Suppose that the transmitters that are unaware of the channel, fix their encoding function based on $\varphi_t^{(k)}$ and define ϕ as the ensemble of these functions. For each θ and ϕ, define $\mathcal{R}_{\epsilon,\theta}(\phi)$ as ϵ-achievable region such as if the rate belongs to it, then the EP is less or equal than ϵ for the choice of ϕ. Now we have:

$$\mathcal{E}(\underline{r}, \epsilon, \mathsf{C}) = \mathbb{P}_\theta(\underline{r} \notin \mathcal{R}_{\epsilon,\theta}(\phi)).$$

This presents an upper bound on the asymptotic spectrum of EP. Moreover, by taking the limit outside of $\mathbb{P}(\lim_{n\to\infty} \epsilon_{n,\theta} < \epsilon)$, we get the following corollary.

Corollary 19 *For the error probability $\epsilon_{n,\theta}$ and ϵ-capacity defined as before, the asymptotic spectrum of EP is as follows:*

$$\inf_{\phi} \mathbb{P}_\theta(\underline{r} \notin \mathcal{R}_{\epsilon,\theta}(\phi)) \geq \mathcal{E}(\underline{r}, \epsilon) = \lim_{n\to\infty} \mathbb{P}_\theta(\epsilon_{n,\theta} > \epsilon) \geq \mathbb{P}_\theta(\underline{r} \notin \mathcal{C}_{\epsilon,\theta}). \tag{8.20}$$

There are composite channels like composite binary symmetric channels (CBSC) where a unique codeword, uniformly distributed code for CBSC, yields the best code for each channel. In this case, we have the following equality:

$$\mathcal{E}(\underline{r}, \epsilon) = \mathbb{P}_\theta(\underline{r} \notin \mathcal{C}_{\epsilon,\theta}). \tag{8.21}$$

Indeed the next example is one instance of these composite channels with unique best code. We take a closer look at these notions and their relation with $\epsilon-$capacity.

Example (Composite Binary Symmetric Averaged Channel) [46]: A binary symmetric averaged channel with three parameters is defined as the set of three binary symmetric channels $(\mathbb{B}_1, \mathbb{B}_2, \mathbb{B}_3)$ with the following parameters:

$$p_1 < p_2 < p_3 \leq \frac{1}{2}$$

The coefficients of the averaged are $\alpha_1, \alpha_2, \alpha_3$ such that:

$$\alpha_1 + \alpha_2 + \alpha_3 = 1.$$

The averaged channel is then defined as $\mathbb{B} = \alpha_1\mathbb{B}_1 + \alpha_2\mathbb{B}_2 + \alpha_3\mathbb{B}_3$. The capacity of binary symmetric channel with the parameter p is known as:

$$\mathcal{C}(p) = 1 - H(p).$$

Kieffer calculated the capacity of averaged binary symmetric channel and showed that the channel does not satisfy strong converse. Moreover $\epsilon-$capacity of this channel is characterized as follows:

$$\mathcal{C}_\epsilon = \begin{cases} \mathcal{C}(p_3) & 0 < \epsilon < \alpha_3 \\ \mathcal{C}(\lambda(p_2, p_3)) & \epsilon = \alpha_3 \\ \mathcal{C}(p_2) & \alpha_3 < \epsilon < \alpha_3 + \alpha_2 \\ \mathcal{C}(\lambda(p_1, p_2)) & \epsilon = \alpha_3 + \alpha_2 \\ \mathcal{C}(p_1) & \alpha_3 + \alpha_2 < \epsilon < 1 \end{cases} \tag{8.22}$$

where $\lambda(p_1, p_2)$ is defined as:

$$\lambda(p, q) = \frac{\log\left(\frac{1-p}{1-q}\right)}{\log\left(\frac{1-p}{1-q}\right) + \log\left(\frac{q}{p}\right)}.$$

Now suppose that there is randomness associated with this channel. For instance suppose that $\mathbf{p}_3$ takes its value randomly between p_2 and $\frac{1}{2}$ with the measure $\mathbb{P}_{\mathbf{p}_3}$. In other words, the channel parameter θ is $\mathbf{p}_3$. Then the asymptotic spectrum of error probability is as follows:

$$\mathcal{E}(\underline{r}, \epsilon) = \begin{cases} \mathbb{P}_{\mathbf{p}_3}\big(r > \mathcal{C}(\mathbf{p}_3)\big) & 0 < \epsilon < \alpha_3 \\ \mathbb{P}_{\mathbf{p}_3}\big(r > \mathcal{C}(\lambda(p_2, \mathbf{p}_3))\big) & \epsilon = \alpha_3 \\ \mathbf{1}[r > \mathcal{C}(p_2)] & \alpha_3 < \epsilon < \alpha_3 + \alpha_2 \\ \mathbf{1}[r > \mathcal{C}(\lambda(p_1, p_2))] & \epsilon = \alpha_3 + \alpha_2 \\ \mathbf{1}[r > \mathcal{C}(p_1)] & \alpha_3 + \alpha_2 < \epsilon < 1 \end{cases} \tag{8.23}$$

To obtain the smallest achievable EP, we have to take a look at the smallest value of ϵ such that the asymptotic error probability will not exceed it. In this example, the only randomness is due to $\mathbf{p}_3$ and the source is aware of the fact that if it transmits a code with the rate $r > \mathcal{C}(p_2)$, then the code will not be decoded correctly. So for the rest suppose that the source transmits a code with $r \leq \mathcal{C}(p_2)$. Now the last three terms in the asymptotic spectrum of EP are automatically zero and we have:

$$\mathcal{E}(\underline{r}, \epsilon) = \begin{cases} \mathbb{P}_{\mathbf{p}_3}\big(r > \mathcal{C}(\mathbf{p}_3)\big) & 0 < \epsilon < \alpha_3 \\ \mathbb{P}_{\mathbf{p}_3}\big(r > \mathcal{C}(\lambda(p_2, \mathbf{p}_3))\big) & \epsilon = \alpha_3 \\ 0 & \alpha_3 < \epsilon < 1 \end{cases} \tag{8.24}$$

For this case it can be seen that the smallest achievable EP is as follows:

$$\epsilon_{\text{p-}}(\underline{r}) = \inf\left\{0 \leq \epsilon < 1 : \mathcal{E}(\underline{r}, \epsilon) = 0\right\} \leq \alpha_3.$$

In other word for this channel, the probability of error will be under α_3 with probability 1 for $r < \mathcal{C}(p_2)$. On the other hand the expected error can be calculated as follows:

$$\bar{\epsilon}(\underline{r}) = \int_0^1 \mathcal{E}(\underline{r}, \epsilon) d\epsilon = \int_0^{\alpha_3} \mathbb{P}_{\mathbf{p}_3}\big(r > \mathcal{C}(\mathbf{p}_3)\big) d\epsilon = \alpha_3 \times \mathbb{P}_{\mathbf{p}_3}\big(r > \mathcal{C}(\mathbf{p}_3)\big).$$

It can be directly seen that the expected error dismisses the information about the error for the error probability at the point $\epsilon = \alpha_3$. This shows once again that the expected error is not enough general to provide full information. Finally the throughput EP is calculated as follows:

$$\bar{\epsilon}_T(\underline{r}) = \sup_{0 \leq \epsilon < 1} \epsilon \mathcal{E}(\underline{r}, \epsilon) = \alpha_3 \times \mathbb{P}_{\mathbf{p}_3}\big(r > \mathcal{C}(\lambda(p_2, \mathbf{p}_3))\big).$$

Here the information about the error for ϵ less than α_3 is lost in the notion. This example clearly shows the relation between all these notions and how the asymptotic spectrum of EP is the notion that implies all other notions and includes all the information related to the probability of error in the composite channels.

However the main problem is that the capacity is not known in general for most of the multiterminal networks and consequently neither is ϵ-capacity. So we have to look for ways to characterize the asymptotic spectrum of EP in other ways.

One option is to analyze the relation between the notion of outage probability and the asymptotic spectrum of EP. The outage probability P_{out} is defined as the probability that a code with the rate $\underline{r}$, cannot be correctly decoded which means that it has non-zero error. The outage probability is then equal to:

$$P_{out} = \mathbb{P}_\theta(\underline{r} \notin \mathcal{C}_\theta).$$

Now suppose that, each channel for each θ satisfies the strong converse condition. It means that for each channel, every code with the rate vector outside the capacity region yields asymptotically the error probability 1. This also means that

$$\mathcal{C}_\theta = \mathcal{C}_{\epsilon\theta} \;\; 0 \leq \epsilon < 1.$$

So for each θ, the asymptotic error probability, i.e. $\limsup_{n\to\infty} \epsilon_{n,\theta}$ takes as value either zero or one. Moreover if there is unique best code for the composite channel then from (8.21), it follows that the asymptotic error probability can be considered as a Bernoulli trial with parameter P_{out} where P_{out} is the outage probability (figure 8.3). So those channels satisfying strong converse condition are of particular interest because the notion of outage probability in these cases coincide with the notion of asymptotic spectrum of EP.

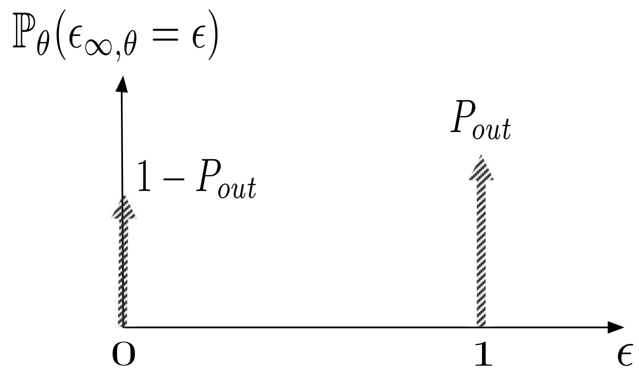


Figure 8.3: Multiterminal Network with Strong Capacity

Another option is to bound the asymptotic spectrum of EP. There are various inner bounds and upper bounds, achievable rates and converse proofs for multiterminal networks. Consider a composite multiterminal network with the parameter θ where an achievable region is known for each θ and ϕ as before, $\mathcal{R}_\theta(\phi)$. Now if the rate $\underline{r}$ is inside the region then the error probability tends to zero and will be less than ϵ for $0 < \epsilon < 1$. For the rate $\underline{r}$, number of those channels with the error probability bigger than ϵ is less or equal to the number of channels with non-zero error probability which implies that the asymptotic spectrum of EP is essentially less or equal than the probability that the rate $\underline{r}$ is not inside the achievable region.

Similarly for the rate $\underline{r}$, number of those channels with the error probability bigger than ϵ is less or equal to the number of channels with error probability equal to one. For a given channel, it is interesting to see for which values of $\underline{r}$, the error probability will tend to one. Apparently for the channels satisfying strong converse, the rates bigger than capacity yield the error probability 1. This leads to the following definition which will be useful for the characterization of the asymptotic EP.

Definition 20 *Consider a multiterminal channel $\mathbb{W}^n$ with m sources and destinations. The full error region is the region $\mathcal{S} \subset \mathbb{R}_+^{m(m-1)}$ such that for all codes $\big(n, M_n^{(ij)}, \epsilon_n\big)$, if the rate vector $\underline{r} = \big(\liminf_{n\to\infty} \frac{1}{n} \log M_n^{(ij)}\big)$ is inside the region $\mathcal{S}$ then the error probability tends to one:*

$$\lim_{n\to\infty} \epsilon_n = 1.$$

The previous definition is simply the definition that we get the error probability 1 for all the nodes if the rate of the codes belong to this region. In this definition, the notion of full error region was defined for the whole network. It can be also defined particularly for a point-to-point communication. In this case the region is determined by a single value called *the full error capacity* $\mathcal{S}$ which is defined as the infimum of all rates for which every code with such rate will yield asymptotically error probability 1.

Using this definition, the following theorem provides the limits over the probability distribution of the error.

Theorem 43 *For composite multiterminal network with the random parameter θ we have:*

$$\mathbb{P}_\theta(\underline{r} \in \mathcal{S}_\theta) \leq \mathcal{E}(\underline{r}, \epsilon) \leq \inf_{\phi} \mathbb{P}_\theta(\underline{r} \notin \mathcal{R}_\theta(\phi)), \tag{8.25}$$

where $\mathcal{R}_\theta$ is the achievable region of the network $\mathbb{W}_\theta$ for a given θ, and $\mathcal{S}_\theta$ is the full error region of this channel for a given θ.

Proof To prove the theorem we start from the definition of the asymptotic spectrum of EP and using the convergence of EP:

$$\begin{aligned}
\mathcal{E}(\underline{r}, \epsilon) &= \mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > \epsilon) \\
&= \mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > \epsilon \text{ and } \underline{r} \in \mathcal{S}_\theta) + \mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > \epsilon \text{ and } \underline{r} \notin \mathcal{S}_\theta) \\
&= \mathbb{P}_\theta(1 > \epsilon \text{ and } \underline{r} \in \mathcal{S}_\theta) + \mathbb{P}_\theta(\lim_{n\to\infty} \epsilon_{n,\theta} > \epsilon \text{ and } \underline{r} \notin \mathcal{S}_\theta) \\
&\geq \mathbb{P}_\theta(\underline{r} \in \mathcal{S}_\theta)
\end{aligned}$$

The proof of the next part is also concluded from the corollary 19 by using the fact that $\mathcal{C}_\theta$ is included in $\mathcal{C}_{\epsilon,\theta}$.

Interestingly $\mathbb{P}_\theta(\underline{r} \notin \mathcal{R}_\theta(\phi))$ is the outage probability. Once again one can see if the channel satisfies the strong converse condition, i.e. $\mathcal{S}_\theta = \mathcal{C}_\theta$ and it has a unique best encoding code then the asymptotic error probability will be equal to outage probability, which supports the operational meaning of this notion.

There are various achievable regions available for multiterminal networks [18] but there is not much about the full error region. We try to provide some results in this direction for the case of discrete memoryless multiterminal channels with $W_\theta = \mathbb{P}_{Y_\theta^{(1)},\ldots,Y_\theta^{(m)}|X_\theta^{(1)},\ldots,X_\theta^{(m)}}$. Various achievable rates can be found for these channels which provide the inner bound over the probability distribution according to the corollary 19.

On the other hand the well known upper bound known for multiterminal network is the cut-set bound [15, 47]. This states that any rate outside the region formed by cutset bound will have non-zero EP. In the next theorem, we prove that the error is necessarily one for any rate inside this region. This result provides an upper bound on the full error region.

Now we focus without loss of generality on the case that a group of source nodes $S \subset \{1, 2, ..., m\}$ send information to destination nodes S^c with the rate vector $\underline{r} = (r^{ij})_{i\in S, j\in S^c}$. The definition of achievability will be limited to the case where $i \in S$ and $j \in S^c$.

Theorem 44 *Consider a discrete memoryless multiterminal channel with* m *nodes. For all codes* $\left(n, M_n^{(ij)}, \epsilon_n\right)$*, Suppose that the rates of the code,* $\underline{r} = \left(\liminf_{n\to\infty} \frac{1}{n} \log M_n^{(ij)}\right)$*, falls outside the following closure for all* $S \subset \{1, 2, ..., m\}$

$$\mathscr{S}_{CB} = co \bigcup_{P\in\mathcal{P}} \Big\{(R(S) \geq 0) : R(S) < I(X_S; Y_{S^c}|X_{S^c})\Big\}$$

where

$$R(S) = \sum_{i\in S, j\in S^c} R_{ij}.$$

In other words suppose that $\underline{r} \notin \mathscr{S}_{CB}$*. Then* $\lim_{n\to\infty} \epsilon_n = 1$*.*

Proof Proof is presented at the appendix C.1.

This theorem implies that the cut-set bound also provides a bound on the full error region. Indeed all those bounds in network information theory that are obtained using the cutset bound, in the same time provide the bound for the full error region.

Using the cutset bound and the available achievable regions like those developed in [18], one can obtain bounds on the asymptotic spectrum of error probability. However we do not assume that all users are in the same time receivers and transmitters. The channel is assumed to be composed of sources, relays and destinations as in Figure 8.2. Suppose that each source $i \in \mathcal{T}$ sends the message to the destinations in the set $\mathcal{D}$ and all other users in $\mathcal{R}$ are relays. The source i sends the message with the rate R_i to all the destinations. The cutset bound for this channel is characterized by:

$$\mathscr{S}^*_{\text{CB}} = co \bigcup_{P\in\mathcal{P}} \Big\{(R(S) \geq 0) : R(S) < \min_{S\subseteq\mathcal{T}} \min_{S'\subseteq\mathcal{R}} \min_{d\in\mathcal{D}} I(X_S X_{S'}; Z_{S'^c} Y_d | X_{S^c} X_{S'^c})\Big\}$$

where $S^c = \mathcal{T} - S$, $S'^c = \mathcal{R} - S'$.

On the other hand an inner bound was developed for this channel using Compress-and-Forward as cooperative strategy in [18]. The following theorem is re-statement of Noisy Network Coding theorem for this channel.

Theorem 45 (Lim-Kim-El Gamal-Chung [18]) *An inner bound on the capacity region of DM network with sources in $\mathcal{T}$, the relays in $\mathcal{R}$ and the destinations in D is given by*

$$\mathscr{R}_{IB} = co \bigcup_{P \in \mathcal{P}} \mathscr{R}_{NNC} \tag{8.26}$$

where

$$\mathscr{R}_{NNC} = \Big\{(R(S) \geq 0) : R(S) < \min_{S \subseteq \mathcal{T}} \min_{S' \subseteq \mathcal{R}} \min_{d \in \mathcal{D}} I(X_S X_{S'}; \hat{Z}_{S'^c} Y_d | X_{S^c} X_{S'^c} Q) \\ - I(Z(S'); \hat{Z}(S') | X(\mathcal{R}) X(\mathcal{T}) \hat{Z}(S'^c) Y_d Q)\Big\}$$

where $S^c = \mathcal{T} - S$, $S'^c = \mathcal{R} - S'$ and $R(S) = \sum_{k \in S} R_k$.

Now take the composite multiterminal network with the parameter θ. And suppose that the sources use the previous Noisy Network Coding scheme for the communication. However, unlike non-composite cases, the sources cannot choose the probability distribution of the channel P from $\mathcal{P}$ to maximize the region because they are not aware of θ. So the probability distribution should be picked up beforehand not to minimize the outage probability.

Now the regions $\mathscr{R}_{\mathrm{NNC}}, \mathscr{S}^*_{\mathrm{CB}}$ can be parametrized using θ as $\mathscr{R}_{\mathrm{NNC},\theta}$ and $\mathscr{S}^*_{\mathrm{CB},\theta}$. These regions can be exploited to provide the following bound over the asymptotic spectrum of EP using the theorem 43.

Corollary 20 *Asymptotic spectrum of EP for the rate $\underline{r}$ and each ϵ satisfies the following bounds:*

$$\mathbb{P}_\theta(\underline{r} \in \mathscr{S}^*_{CB,\theta}) \leq \mathcal{E}(\underline{r}, \epsilon) \leq \min_{P \in \mathcal{P}} \mathbb{P}_\theta(\underline{r} \notin \mathscr{R}_{NNC,\theta}). \tag{8.27}$$

Note that the probability distribution is chosen such that it minimizes the outage probability.

The noisy network coding theorem is a tight bound for group of channels. For the case of deterministic network without interference [48] or the case of finite field linear deterministic networks $Y_k = \sum_{i=1}^m g_{ik} X_i$ [19], if we choose $\hat{Z}_k = Z_k$ for $k \in \{1, ..., m\}$ then it can be seen that the bounds of noisy network coding is tight and coincides with the cutset bound. However it is only for the finite field linear deterministic network that the optimum value is obtained by independent and uniform distribution of probabilities.

Now consider a composite finite field linear deterministic network where the channel in operation is chosen from the set of the finite field linear deterministic networks, indexed by $\theta \sim \mathbb{P}_\theta$. Each channel satisfies the strong converse and moreover there is unique best encoding function, i.e. a unique optimum probability distribution for all channels. Then the outage probability is the asymptotic spectrum of EP in this network and the following corollary can be obtained.

Corollary 21 *For the composite finite field linear deterministic network, the asymptotic spectrum of EP for the rate* $\underline{r}$ *and each* ϵ *is as follows:*

$$\mathcal{E}(\underline{r}, \epsilon) = \mathbb{P}_\theta(\underline{r} \notin \mathscr{C}_{DN,\theta}). \tag{8.28}$$

where $\mathscr{C}_{DN,\theta}$ *for a given* θ *is defined:*

$$\mathscr{C}_{DN,\theta} = \Big\{(R(S) \geq 0) : R(S) < \min_{S \subseteq \mathcal{R}} \min_{d \in \mathcal{D}} H(Z_{S^c\theta} Y_{d\theta} | X_{\mathcal{T}} X_{S^c\theta})\Big\},$$

where the input distribution is chosen at each source as independent and uniformly distributed.

It is interesting to see that the right hand side of (21) is independent of ϵ which means that the outage probability is a sufficient measure for the performance of this network.

Chapter 9

Conclusion and Future Work

In this chapter we conclude the thesis and highlight some future directions.

9.1 Summary and Conclusion

In this thesis we developed cooperative strategies for multiterminal networks with channel uncertainty. By uncertainty here, it was meant that the channel in operation is chosen from a set of channels and the source is not aware of the choice.

First we studied the single relay channel with the channel uncertainty, which consists of a set of single relay channels out of which the channel in operation is chosen. This model was called *simultaneous relay channel*. The idea was to use the broadcasting approach which means the transmission of a non-zero rate, not necessarily equal, for each channel in the set. The simultaneous relay channel was studied with a set composed of two relay channels. The broadcasting approach turns the problem to the analysis of the broadcast relay channel with two relays. Hence we investigated cooperative strategies for the broadcast relay channel. Several novel schemes have been considered, for which inner and outer bounds on the capacity region were derived.

Depending on the nature of the channels involved, it is well-known that the best way to cover the information from relays to destinations will not be the same. Based on the best known cooperative strategies, namely, *Decode-and-Forward* (DF) and *Compress-and-Forward* (CF), achievable regions for three scenarios of interest have been analyzed. These may be summarized as follows: (i) both relay nodes use DF schemes, (ii) one relay node uses CF scheme while the other one uses CF scheme and (iii) both relay nodes use CF scheme. In particular, for the region (ii) it is shown that *block-Markov coding* works with

CF scheme without incurring performance losses. These inner bounds are shown to be tight for some cases, yielding capacity results for the semi-degraded BRC with common relay (BRC-CR) and two Gaussian degraded BRC-CRs. Whereas our bounds seem to be not tight for the case of degraded BRC-CR. An outer bound on the capacity region of the general BRC was also derived. One should emphasize that when the relays are not present, this bound reduces to the best known outer bound for general broadcast channels (referred to as UVW-outer bound). Similarly, when only one relay channel is present at once, this bound reduces to the cut-set bound for the general relay channel. Finally, application examples for Gaussian channels have been studied and the corresponding achievable rates were computed for all inner bounds.

It should be worth to mention that the inner and outer bounds obtained for broadcast relay channels with two relays are rather complicated. For instance, the DF-DF achievable region involves 16 bounds, and yet the model corresponds to the simultaneous relay channel with only two possibilities. This reveals the complexity of using broadcasting approaches for cooperative networks with numerous channels in the set. In these cases, there is another approach which consists in fixing the rate and studying the behavior of the network with respect to a performance criteria. Although a compound approach can guarantee asymptotically zero-error probability for all channel indices, the worst possible index may yield in general non-positive rates for most of wireless scenarios. In this direction, the composite relay channel was discussed in the next chapter where the channel index $\theta \in \Theta$ is randomly drawn according to $\mathbb{P}_\theta$. The channel draw $\theta = (\theta_r, \theta_d)$ remains fix during the communication, however it is assumed to be unknown at the source, fully known at the destination and partly known θ_r at the relay end. The coding rate r is fixed regardless of the current channel index. The asymptotic error probability is chosen as the measure of characterizing the performance. It was shown that instead of choosing a unique relay function for all possible channels, a novel coding strategy can be adopted where the relay can select –based on its channel measurement θ_r– the adequate coding strategy.

To this purpose, achievable rates were first derived for the two-relay network with mixed coding strategy. This region improves the achievable region for two relay networks with mixed strategies in [14]. As a matter of fact, it is shown that the same code for this two-relay network works as well for the composite relay channel where the relay is allowed to select either DF or CF scheme. Whereas the source sends the information regardless of the relay function. More specifically, we showed that the recent CF scheme [40] can simultaneously work with DF scheme. Furthermore, only CSI from the source-to-relay channel is needed to decide –at the relay end– about the adequate relay function to be implemented. So the relay does not need full CSI to decide about the strategy. This idea was further extended to general composite networks with multiple relays and single source and destination. A similar coding was developed to allow the selection to the relays in the network whether to use DF or CF schemes. The achievable region presented generalizes NNC to the case of mixed coding strategy. It was also shown that the relay using DF scheme can exploit via offset coding the help of those using CF scheme. An application example to the case of fading Gaussian relay channel was also considered, where SCS clearly outperforms the well-known DF and CF schemes.

Finally, the asymptotic behavior of error probability is studied independently for composite multiterminal networks. We showed that the notion of outage probability in general is not enough precise to characterize the error probability. Instead the notion of asymptotic spectrum of error probability is introduced as a novel performance measure for composite networks. The notion of *asymptotic spectrum of EP* for (r, ϵ) is defined as the asymptotic probability that error probability falls over ϵ for a fixed rate r. It is shown that this notion implies other available notions used to measure the performance of composite networks. As a matter of fact, the behavior of EP is directly related to the ϵ-capacity of the network.

We showed that the asymptotic spectrum of EP can be bounded using available achievable rate regions and a new region called *full error region*. For networks satisfying strong converse condition, the asymptotic spectrum of EP coincides with the conventional notion of the outage probability. Finally, it was shown that the cutset bound provides an outer bound on the full error region of multiterminal networks. In other words, each code with transmission rates not satisfying the cutset bound the probability of error tends to one.

9.2 Future Work

We first discuss the broadcast relay channel where we observed that the relay must help both destinations and the tricky part is how to share this help between common and private information. Particularly, in the case of the physically degraded broadcast relay channel with common relay, as shown in Fig. 9.1, the relay has to help both destinations. This implies decoding of both messages and forward them to the destinations. Theorem 32 gives one way to share the relay help between common and private information. Essentially, the relay uses V to help common information and X_1 to help private information. In this case, the choice of V distinct from X_1 appears to be necessary, because $V = \emptyset$ would

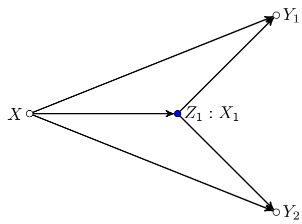


Figure 9.1: Broadcast relay channel with common relay

remove the help of relay for the common information and then the region will become clearly suboptimal. Note that for the case of $Y_1 = Y_2$, it would imply that the help of the relay is not exploited. Similarly, setting $V = X_1$ will lead to a similar problem when $Y_2 = \ emptyset$.

This problem can be more clearly explained from another perspective. We remark that for a single relay and destination channel, the source uses superposition coding to superimpose the source codeword over the relay codeword. Now in the case of broadcast relay channels, the source has to provide two separate codes for each destination. Here there are two source codes destined for each destination, and the relay help for each of those messages. None of source codes can be superimposed on the whole relay code since it limits the relay help for the other user. For instance, suppose that the source code U_1 for the first destination is superimposed on V_1 and V_2, the relay helps for Y_1 and Y_2. Then it can be seen that the rate of V_2 becomes limited by the condition of correctly decoding U_1, which is clearly not good enough. Another way would be to superimpose U_1 only on the code V_1. However this causes another problem. Now that U_1 is not superimposed over V_2 these variables do not have full dependence anymore. In these cases, it seems not possible to show the converse. Finally, Marton coding can remove this problem of correlation but at the price of appearing the negative terms in the inner bounds, which again renders difficult the task of proving converse. One perspective for future work is to explore a proper code for the problem of superimposing one DF code code on the top of another DF code.

Regarding composite multiterminal networks, only unicast settings were considered in this thesis. However, it would be worth to investigate similar coding for multicast composite networks. The task is not a straightforward generalization of the current results. It should be emphasized that the use of conventional noisy network coding in multi-sources problems does not significantly differ from the unicast setting. Whereas in presence of selective coding, we are interested in using part of relays with DF scheme, which poses the problem of dynamical selection at relays of the sources they tend to help. Notice that the source which has the best channel quality with respect to a relay, can dynamically change in a composite setting. Is it possible to develop a selective coding for these cases such that the relay can dynamically choose the source with which it wants to cooperate? On the other hand, similar problem rises for the case of multicast networks with multiple destinations. It is possible that a relay is better to use CF for one destination and at the same time to help via DF scheme for another one. It is interesting to see whether there is a possibility of developing a code capable of using DF for some destinations while using CF for other ones.

Appendix A

Appendix chapter 2

A.1 Proof of Theorem 24

Before starting the proof, we remind the notion of typical sequences that are used for the proofs.

Definition 21 (Typical Sequences) *The set of A_ϵ of ϵ-typical n-sequences $(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ..., \mathbf{x}^{(k)})$, called also ϵ-strong typical, is defined by*

$$A_\epsilon(X^{(1)}, X^{(2)}, ..., X^{(k)}) = \Big\{(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ..., \mathbf{x}^{(k)}) : \\ \left|\frac{1}{n}N(x^{(1)}, x^{(2)}, ..., x^{(k)}; \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ..., \mathbf{x}^{(k)}) - p(x^{(1)}, x^{(2)}, \cdots, x^{(k)})\right| \\ < \epsilon \left\|\mathscr{X}^{(1)} \times \mathscr{X}^{(2)} \times \cdots \times \mathscr{X}^{(k)}\right\|, \textit{for } (x^{(1)}, x^{(2)}, ..., x^{(k)}) \in \mathscr{X}^{(1)} \times \mathscr{X}^{(2)} \times, \cdots, \times \mathscr{X}^{(k)}\Big\},$$

where $N(s; \mathbf{s})$ is the number of indices in $\mathbf{s}$, $i = \{1, 2, \ldots, n\}$ such that $s_i = s$.

The following lemma is the fundamental AEP results for typical sequences [47].

Lemma 1 *For any $\epsilon > 0$, there exists an integer n such that $A_\epsilon(\mathbf{S})$ satisfies*

(i) $\mathbf{P}\Big\{A_\epsilon(\mathbf{S})\Big\} \geq 1 - \epsilon$, *for all* $\mathbf{S} \subseteq \Big\{\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, ..., \mathbf{X}^{(k)}\Big\}$,

(ii) $\mathbf{s} \in A_\epsilon(\mathbf{S}) \Rightarrow \left|-\frac{1}{n}\log p(\mathbf{s}) - H(\mathbf{S})\right| < \epsilon$,

(iii) $(1 - \epsilon)2^{n(H(\mathbf{S})-\epsilon)} \leq \|A_\epsilon(\mathbf{S})\| \leq 2^{n(H(\mathbf{S})+\epsilon)}$.

To prove the theorem, first split the private information W_b into non-negative indices (S_{0b}, S_b, S_{b+2}) with $b = \{1, 2\}$. Then, merge the common information W_0 with a part of private information (S_{01}, S_{02}) into a single message. Hence we obtain that $R_b = S_{b+2} +$

$S_b + S_{0b}$ where this operation can be seen in Fig. 6.4. For the sake of notation, it is defined that $\underline{u} = u_1^n$ for each r.v. u. Let consider the main steps for codebook generation, encoding and decoding procedures.

Code Generation:

(i) Generate 2^{nT_0} i.i.d. sequences $\underline{v}_0$ each with PD

$$P_{V_0}(\underline{v}_0) = \prod_{j=1}^{n} p_{V_0}(v_{0j}),$$

and index them as $\underline{v}_0(r_0)$ with $r_0 = \left[1 : 2^{nT_0}\right]$.

(ii) For each $\underline{v}_0(r_0)$, generate 2^{nT_0} i.i.d. sequences $\underline{u}_0$ each with PD

$$P_{U_0|V_0}(\underline{u}_0|\underline{v}_0(r_0)) = \prod_{j=1}^{n} p_{U_0|V_0}(u_{0j}|v_{0j}(r_0)),$$

and index them as $\underline{u}_0(r_0, t_0)$ with $t_0 = \left[1 : 2^{nT_0}\right]$.

(iii) For $b \in \{1, 2\}$ and each $\underline{v}_0(r_0)$, generate 2^{nT_b} i.i.d. sequences $\underline{x}_b$ each with PD

$$P_{X_b|V_0}(\underline{x}_b|\underline{v}_0(r_0)) = \prod_{j=1}^{n} p_{X_b|V_0}(x_{bj}|v_{0j}(r_0)),$$

and index them as $\underline{x}_b(r_0, r_b)$ with $r_b = \left[1 : 2^{nT_b}\right]$.

(iv) Partition the set $\left\{1, \ldots, 2^{nT_0}\right\}$ into $2^{n(R_0+S_{01}+S_{02})}$ cells (similarly to [12]) and label them as $S_{w_0,s_{01},s_{02}}$. In each cell there are $2^{n(T_0-R_0-S_{01}-S_{02})}$ elements.

(v) For each $b = \{1, 2\}$ and every pair $\big(\underline{u}_0(r_0, t_0), \underline{x}_b(r_0, r_b)\big)$ chosen in the bin (w_0, s_{01}, s_{02}), generate 2^{nT_b} i.i.d. sequences $\underline{u}_b$ each with PD

$$\begin{aligned} P_{U_b|U_0X_bV_0}\big(\underline{u}_b|\underline{u}_0(r_0, t_0),&\underline{x}_b(r_0, r_b), \underline{v}_0(r_0)\big) \\ &= \prod_{j=1}^{n} p_{U_b|U_0X_bV_0}(u_{bj}|u_{0j}(r_0, t_0), x_{bj}(r_0, r_b), v_{0j}(r_0)), \end{aligned}$$

and index them as $\underline{u}_b(r_0, t_0, r_b, t_b)$ with $t_b = \left[1 : 2^{nT_b}\right]$.

(vi) For $b = \{1, 2\}$, partition the set $\left\{1, \ldots, 2^{nT_b}\right\}$ into 2^{nS_b} cells and label them as S_{s_b}. In each cell there are $2^{n(T_b-S_b)}$ elements.

(vii) For each $b = \{1, 2\}$ and every pair of sequences $\big(\underline{u}_1(r_0, t_0, r_1, t_1),\ \underline{u}_2(r_0, t_0, r_2, t_2)\big)$ chosen in the bin (s_1, s_2), generate $2^{nT_{b+2}}$ i.i.d. sequences $\underline{u}_{b+2}$ each with PD

$$P_{U_{b+2}|U_b}(\underline{u}_{b+2}|\underline{u}_b(r_0, t_0, r_b, t_b)) = \prod_{j=1}^{n} p_{U_{b+2}|U_b}(u_{(b+2)j}|u_{bj}(r_0, t_0, r_b, t_b)).$$

Index them as $\underline{u}_{b+2}(r_0, t_0, r_b, t_b, t_{b+2})$ with $t_{b+2} \in \left[1, 2^{nT_{b+2}}\right]$.

(viii) For $b = \{1,2\}$, partition the set $\{1,\dots,2^{nT_{b+2}}\}$ into $2^{nS_{b+2}}$ cells and label them as $S_{s_{b+2}}$. In each cell there are $2^{n(T_{b+2}-S_{b+2})}$ elements.

(ix) Finally, use a deterministic function for generating $\underline{x}$ as $f(\underline{u}_3, \underline{u}_4)$ indexed by $\underline{x}(r_0, t_0, r_1, r_2, t_1, t_2, t_3, t_4)$.

Encoding Part: The transmission is done in $B+1$ block. The encoding in block i is as follows:

(i) First, reorganize the current message (w_{0i}, w_{1i}, w_{2i}) into $(w_{0i}, s_{01i}, s_{02i}, s_{1i}, s_{2i}, s_{3i}, s_{4i})$.

(ii) Then for each $b = \{1,2\}$, relay b already knows about the index $(t_{0(i-1)}, t_{b(i-1)})$, so it sends $\underline{x}_b\big(t_{0(i-1)}, t_{b(i-1)}\big)$.

(iii) For each $\underline{v}_0(t_{0(i-1)})$, the encoder searches for an index t_{0i} at the cell $S_{w_{0i},s_{01i},s_{02i}}$ such that $\underline{u}_0\big(t_{0(i-1)}, t_{0i}\big)$ is jointly typical with $\big(\underline{x}_1(t_{0(i-1)}, t_{1(i-1)}), \underline{x}_2(t_{0(i-1)}, t_{2(i-1)}), \underline{v}_0(t_{0(i-1)})\big)$. The success of this step requires that [12]

$$T_0 - R_0 - S_{01} - S_{02} \geq I(U_0; X_1, X_2|V_0). \tag{A.1}$$

(iv) For each $b = \{1,2\}$ and every cell $S_{s_{bi}}$, define the set $\mathscr{L}_b$ to be the set of all sequences $\underline{u}_b\big(t_{0(i-1)}, t_{0i}, t_{b(i-1)}, t_{bi}\big)$ for $t_{bi} \in S_{s_{bi}}$ which are jointly typical with

$$\big(\underline{x}_{\overline{b}}(t_{0(i-1)}, t_{\overline{b}(i-1)}), \underline{v}_0(t_{0(i-1)}), \underline{u}_0(t_{0(i-1)}, t_{0i}), \underline{x}_b(t_{0(i-1)}, t_{b(i-1)})\big)$$

where $\overline{b} = \{1,2\} \setminus \{b\}$. In order to create $\mathscr{L}_b$, we look for the $\underline{u}_b$-index inside the cell $S_{s_{bi}}$ and find $\underline{u}_b$ such that it belongs to the set of ϵ-typical n-sequences $A_\epsilon^n(V_0U_0X_1X_2U_b)$.

(v) Then search for a pair $(\underline{u}_1 \in \mathscr{L}_1, \underline{u}_2 \in \mathscr{L}_2)$ such that $\big(\underline{u}_1(t_{0(i-1)}, t_{0i}, t_{1(i-1)}, t_{1i})$ $, \underline{u}_2(t_{0(i-1)}, t_{0i}, t_{2(i-1)}, t_{2i})\big)$ are jointly typical given the RVs $\big(\underline{v}_0(t_{0(i-1)}), \underline{x}_2(t_{0(i-1)}, t_{2(i-1)}),$ $\underline{x}_1(t_{0(i-1)}, t_{1(i-1)})\ , \underline{u}_0(t_{0(i-1)}, t_{0i})\big)$. The success of coding steps (iv) and (v) requires

$$\begin{aligned} T_b - S_b &\geq I(U_b; X_{\overline{b}}|X_b, U_0, V_0), \\ T_1 + T_2 - S_1 - S_2 &\geq I(U_1; X_2|X_1, U_0, V_0) + I(U_2; X_1|X_2, U_0, V_0) \\ &\quad + I(U_2; U_1|X_1, X_2, U_0, V_0). \end{aligned} \tag{A.2}$$

Notice that the first inequality in the above expression, for $b = \{1,2\}$, guarantees the existence of non-empty sets $(\mathscr{L}_1, \mathscr{L}_2)$, and the last one is for the step (iv).

(vi) The encoder searches for index $t_{3i} \in S_{s_{3i}}$ and $t_{4i} \in S_{s_{4i}}$, such that $\underline{u}_3\big(t_{0(i-1)}, t_{0i}, t_{1(i-1)}, t_{1i}, t_{3i}\big)$ and $\underline{u}_4\big(t_{0(i-1)}, t_{0i}, t_{2(i-1)}, t_{2i}, t_{4i}\big)$ are jointly typical given each chosen typical pair of

$\underline{u}_1(t_{0(i-1)}, t_{0i}, t_{1(i-1)}, t_{1i})$ and $\underline{u}_2(t_{0(i-1)}, t_{0i}, t_{2(i-1)}, t_{2i})$. The success of this encoding step requires

$$T_3 + T_4 - S_3 - S_4 \geq I(U_3; U_4|U_1, U_2, X_1, X_2, U_0, V_0). \tag{A.3}$$

(vii) Once the encoder found $(t_{0i}, t_{1i}, t_{2i}, t_{3i}, t_{4i})$ (based on the code generation) corresponding to $(w_{0i},$
$s_{01i}, s_{02i}, s_{1i}, s_{2i}, s_{3i}, s_{4i})$, it transmits $\underline{x}(r_{0(i-1)}, t_{0i}, r_{1(i-1)}, r_{2(i-1)}, t_{1i}, t_{2i}, t_{3i}, t_{4i})$. t_{0i} carries the common message after bit recombination and Marton coding. t_{1i}, t_{3i} and t_{2i}, t_{4i} are respectively private information for Y_1 and Y_2. t_{3i} and t_{4i} correspond to partial coding and are transmitted directly to the destination.

Decoding Part: To decode the messages at block i, the relays assume that all the messages up to block $i-1$ have been correctly decoded and decode the current messages in the same block. The destinations use backward decoding assuming correctly decoded messages until block $i+1$.

(i) First for $b = \{1, 2\}$, the relay b after receiving z_{bi} tries to decode (t_{0i}, t_{bi}). The relay is aware of (V_0, X_b) because it is supposed to know about $(t_{0(i-1)}, t_{b(i-1)})$. The relay b declares that the pair (t_{0i}, t_{bi}) is sent if the following conditions are simultaneously satisfied:

(a) $\underline{u}_0(t_{0(i-1)}, t_{0i})$ is jointly typical with $\big(z_{bi}, \underline{v}_0(t_{0(i-1)}), \underline{x}_b(t_{0(i-1)}, t_{b(i-1)})\big)$.

(b) $\underline{u}_b(t_{0(i-1)}, t_{0i}, t_{b(i-1)}, t_{bi})$ is jointly typical with $\big(z_{bi}, \underline{v}_0(t_{0(i-1)}), \underline{x}_b(t_{0(i-1)}, t_{b(i-1)})\big)$.

Notice that $\underline{u}_0$ has been generated independent of $\underline{x}_b$ and hence $\underline{x}_b$ does not appear in the given part of mutual information. This is an important issue that may increase the region. Constraints for reliable decoding are:

$$T_b < I(U_b; Z_b|U_0, V_0, X_b), \tag{A.4}$$

$$T_b + T_0 < I(U_b; Z_b|U_0, V_0, X_b) + I(U_0; Z_b, X_b|V_0). \tag{A.5}$$

Remark 21 *The intuition behind expressions* (A.4) *and* (A.5) *is as follows. Since the relay knows $\underline{x}_{b(i-1)}$ we are indeed decreasing the cardinality of the set of possible $\underline{u}_0$, which without additional knowledge is 2^{nT_0}. The new set of possible ($\underline{u}_0$, $\mathscr{L}_{X_b}$) can be defined as all $\underline{u}_0$ jointly typical with $\underline{x}_{b(i-1)}$. It can be shown [13] that $\mathbb{E}[\|\mathscr{L}_{X_b}\|] = 2^{n[T_0 - I(U_0;X_b|V_0)]}$, which proves our claim on the reduction of cardinality. One can see that after simplification* (A.5) *using* (A.1)*, $I(U_0; Z_b, X_b|V_0)$ is removed and the final bound reduces to $I(U_0, U_b; Z_b|V_0, X_b)$.*

(ii) For each $b \in \{1,2\}$ destination b, after receiving $y_{b(i+1)}$, tries to decode the relay-forwarded information (t_{0i}, t_{bi}), knowing $(t_{0(i+1)}, t_{b(i+1)})$. It also tries to decode the direct information $t_{(b+2)(i+1)}$. Backward decoding is used to decode index (t_{0i}, t_{bi}). The decoder declares that $(t_{0i}, t_{bi}, t_{(b+2)(i+1)})$ is sent if the following constraints are simultaneously satisfied:

   (a) $\big(\underline{v}_0(t_{0i}), \underline{u}_0(t_{0i}, t_{0(i+1)}), y_{b(i+1)}\big)$ are jointly typical,

   (b) $\big(\underline{x}_b(t_{0(i)}, t_{b(i)}), \underline{v}_0(t_{0i}), \underline{u}_0(t_{0i}, t_{0(i+1)})\big)$ and $y_{b(i+1)}$ are jointly typical,

   (c) $\big(\underline{u}_b(t_{0i}, t_{0(i+1)}, t_{bi}, t_{b(i+1)}), \underline{u}_{b+2}(t_{0i}, t_{0(i+1)}, t_{bi}, t_{b(i+1)}, t_{b(i+1)})\big)$ and $\big(y_{b(i+1)}, \underline{v}_0(t_{0i})$ $, \underline{u}_0(t_{0i}, t_{0(i+1)}), \underline{x}_b\big(t_{0(i)}, t_{b(i)}\big)\big)$ are jointly typical.

Notice that in the decoding step (iib) the destination knows about $t_{0(i+1)}$, which has been chosen such that $(\underline{u}_0, \underline{x}_b)$ are jointly typical and this information contributes to decrease the cardinality of all possible $\underline{x}_b$ (similarly to what happened in decoding at the relay). Hence U_0 in step (iib) does not appear in the given part of mutual information. From this we have that the main constraints for successful decoding are as follows:

$$T_{b+2} < I(U_{b+2}; Y_b | U_0, V_0, X_b, U_b), \tag{A.6}$$

$$T_{b+2} + T_b < I(U_{b+2}, U_b, X_b; Y_b | U_0, V_0), \tag{A.7}$$

$$T_{b+2} + T_b + T_0 < I(V_0, U_0; Y_b) + I(X_b; Y_b, U_0 | V_0) + I(U_{b+2}, U_b; Y_b | U_0, V_0, X_b). \tag{A.8}$$

Observe that U_0 increases the bound in (A.7). Similarly using (A.1), and after removing the common term $I(U_0; X_b | V_0)$, one can simplify the bound in (A.8) to $I(U_{b+2}, U_b, X_b, V_0, U_0; Y_b)$.

(iii) Theorem 24 follows by applying Fourier-Motzkin elimination to (A.1)-(A.8) and using the non-negativity of the rates. This concludes the proof.

A.2 Proof of Theorem 25

Reorganize first private messages w_i, $i = \{1,2\}$ into (s'_i, s_i) with non-negative rates (S'_i, S_i) where $R_i = S'_i + S_i$. Merge (s'_1, s'_2, w_0) to one message s_0 with rate $S_0 = R_0 + S'_1 + S'_2$. For the sake of notation, it is assumed that $\underline{u} = u_1^n$. Let consider the main steps for codebook generation, encoding and decoding procedures.

Code Generation:

(i) Generate 2^{nS_0} i.i.d. sequences $\underline{v}_0$ with PD

$$P_{V_0}(\underline{v}_0) = \prod_{j=1}^{n} p_{V_0}(v_{0j})$$

and index them as $\underline{v}_0(r_0)$ with $r_0 = \left[1 : 2^{nS_0}\right]$.

(ii) For each $\underline{v}_0(r_0)$, generate 2^{nS_0} i.i.d. sequences $\underline{u}_0$ with PD

$$P_{U_0|V_0}(\underline{u}_0|\underline{v}_0(r_0)) = \prod_{j=1}^{n} p_{U_0|V_0}(u_{0j}|v_{0j}(r_0)),$$

and index them as $\underline{u}_0(r_0, s_0)$ with $s_0 = \left[1 : 2^{nS_0}\right]$.

(iii) For each $\underline{v}_0(r_0)$, generate 2^{nT_1} i.i.d. sequences $\underline{x}_1$ with PD

$$P_{X_1|V_0}(\underline{x}_1|\underline{v}_0(r_0)) = \prod_{j=1}^{n} p_{X_1|V_0}(x_{1j}|v_{0j}(r_0)),$$

and index them as $\underline{x}_1(r_0, r_1)$ with $r_1 = \left[1 : 2^{nT_1}\right]$.

(iv) Generate $2^{nR_{x_2}}$ i.i.d. sequences $\underline{x}_2$ with PD

$$P_{X_2}(\underline{x}_2) = \prod_{j=1}^{n} p_{X_2}(x_{2j})$$

as $\underline{x}_2(r_2)$, where $r_2 = \left[1 : 2^{nR_{x_2}}\right]$.

(v) For each $\underline{x}_2(r_2)$ generate $2^{n\hat{R}_2}$ i.i.d. sequences $\underline{\hat{z}}_2$ with PD

$$P_{\hat{Z}_2|X_2}(\underline{\hat{z}}_2|\underline{x}_2(r_2)) = \prod_{j=1}^{n} p_{\hat{Z}_2|X_2}(\hat{z}_{2j}|x_{2j}(r_2)),$$

and index them as $\underline{\hat{z}}_2(r_2, \hat{s})$, where $\hat{s} = \left[1 : 2^{n\hat{R}_2}\right]$.

(vi) Partition the set $\left\{1, \ldots, 2^{n\hat{R}_2}\right\}$ into 2^{nR_2} cells and label them as S_{r_2}. In each cell there are $2^{n(\hat{R}_2 - R_2)}$ elements.

(vii) For each pair $\left(\underline{u}_0(r_0, s_0), \underline{x}_1(r_0, r_1)\right)$, generate 2^{nT_1} i.i.d. sequences $\underline{u}_1$ with PD

$$\begin{aligned} P_{U_1|U_0X_1V_0}(\underline{u}_1|\underline{u}_0(r_0, s_0),&\underline{x}_1(r_0, r_1), \underline{v}_0(r_0)) \\ &= \prod_{j=1}^{n} p_{U_1|U_0V_0X_1}(u_{1j}|u_{0j}(r_0, s_0), x_{1j}(r_0, r_1), v_{0j}(r_0)), \end{aligned}$$

and index them as $\underline{u}_1(r_0, s_0, r_1, t_1)$, where $t_1 = \left[1 : 2^{nT_1}\right]$.

(viii) For each $\underline{u}_0(r_0, s_0)$, generate 2^{nT_2} i.i.d. sequences $\underline{u}_2$ with PD

$$P_{U_2|U_0V_0}(\underline{u}_2|\underline{u}_0(r_0,s_0),\underline{v}_0(r_0)) = \prod_{j=1}^{n} p_{U_2|U_0V_0}(u_{2j}|u_{0j}(r_0,s_0),v_{0j}(r_0)),$$

and index them as $\underline{u}_2(r_0, s_0, t_2)$, where $t_2 = \left[1 : 2^{nT_2}\right]$.

(ix) For $b = \{1, 2\}$, partition the set $\left\{1 : \dots, 2^{nT_b}\right\}$ into 2^{nS_b} subsets and label them as S_{s_b}. In each subset, there are $2^{n(T_b - S_b)}$ elements.

(x) Finally, use a deterministic function for generating $\underline{x}$ as $f\left(\underline{u}_1, \underline{u}_2\right)$ indexed by $\underline{x}(r_0, s_0, r_1, t_1, t_2)$.

Encoding Part: In block i, the source wants to send (w_{0i}, w_{1i}, w_{2i}) by reorganizing them into (s_{0i}, s_{1i}, s_{2i}). Encoding steps are as follows:

(i) DF relay knows $(s_{0(i-1)}, t_{1(i-1)})$ so it sends $\underline{x}_1\big(s_{0(i-1)}, t_{1(i-1)}\big)$.

(ii) CF relay knows from the previous block that $\hat{s}_{i-1} \in S_{r_{2i}}$ and it sends $\underline{x}_2(r_{2i})$.

(iii) Then for each subset $S_{s_{2i}}$, create the set $\mathscr{L}$ consisting of those index t_{2i} such that $t_{2i} \in S_{s_{2i}}$, and $\underline{u}_2\big(s_{0(i-1)}, s_{0i}, t_{2i}\big)$ is jointly typical with $\underline{x}_1\big(s_{0(i-1)}, t_{1(i-1)}\big), \underline{v}_0\big(s_{0(i-1)}\big), \underline{u}_0\big(s_{0(i-1)}, s_{0i}\big)$.

(iv) Then look for $t_{1i} \in S_{s_{1i}}$ and $t_{2i} \in \mathscr{L}$ such that $\big(\underline{u}_1(s_{0(i-1)}, s_{0i}, t_{1(i-1)}, t_{1i}), \underline{u}_2(s_{0(i-1)}, s_{0i}, t_{2i})\big)$ are jointly typical given the RVs $\underline{v}_0(s_{0(i-1)}), \underline{x}_1(s_{0(i-1)}, t_{1(i-1)})$, and with $\underline{u}_0(s_{0(i-1)}, s_{0i})$. The constraints for the successful coding steps (iii) and (iv) are:

$$T_2 - S_2 \geq I(U_2; X_1|U_0, V_0), \tag{A.9}$$

$$T_1 + T_2 - S_1 - S_2 \geq I(U_2; U_1, X_1|U_0, V_0). \tag{A.10}$$

The first inequality guarantees the existence of non-empty sets $\mathscr{L}$.

(v) From (s_{0i}, s_{1i}, s_{2i}), the source finds (t_{1i}, t_{2i}) and sends $\underline{x}(s_{0(i-1)}, s_{0i}, t_{1(i-1)}, t_{1i}, t_{2i})$.

Decoding Part: After the transmission of the block $i + 1$, the DF relay starts to decode the messages of block $i+1$ with the assumption that all messages up to block i have been correctly decoded. Destination 1 waits until the last block and uses backward decoding (similarly to [14]). The second destination first decodes $\hat{Z}_2$ and then uses it with Y_2 to decode the messages while the second relay tries to find $\hat{Z}_2$ of the current block.

(i) DF relay tries to decode $(s_{0(i+1)}, t_{1(i+1)})$. The conditions for reliable decoding are:

$$T_1 + S_0 < I(U_0, U_1; Z_1|X_1V_0), \tag{A.11}$$

$$T_1 < I(U_1; Z_1|U_0, V_0, X_1). \tag{A.12}$$

(ii) Destination 1 tries to decode (s_{0i}, t_{1i}) subject to

$$T_1 + S_0 < I(X_1, V_0, U_0, U_1; Y_1), \tag{A.13}$$

$$T_1 < I(U_1, X_1; Y_1|U_0, V_0). \tag{A.14}$$

(iii) CF relay searches for $\hat{s}_i$ after receiving $\underline{z}_2(i)$ such that $\big(\underline{x}_2(r_{2i}), \underline{z}_2(i), \underline{\hat{z}}_2(\hat{s}_i, r_{2i})\big)$ are jointly typical subject to

$$\hat{R}_2 \geq I(Z_2; \hat{Z}_2|X_2). \tag{A.15}$$

(iv) Destination 2 searches for $r_{2(i+1)}$ such that $\big(\underline{y}_2(i+1), \underline{x}_2(r_{2(i+1)})\big)$ is jointly typical. Then in finds $\hat{s}_i$ such that $\hat{s}_i \in S_{r_{2(i+1)}}$ and $\big(\underline{\hat{z}}_2(\hat{s}_i, r_{2i}), \underline{y}_2(i), \underline{x}_2(r_{2i})\big)$ is jointly typical. Conditions for reliable decoding are:

$$R_{x_2} \leq I(X_2; Y_2), \tag{A.16}$$

$$\hat{R}_2 \leq R_{x_2} + I(\hat{Z}_2; Y_2|X_2). \tag{A.17}$$

(v) Decoding of CF user in block i is done with the assumption of correct decoding of (s_{0l}, t_{2l}) for $l \leq i-1$. The pair (s_{0i}, t_{2i}) are decoded as the message such that $(\underline{v_0}(s_{0(i-1)}), \underline{u_0}(s_{0(i-1)}, s_{0i}), \underline{u_2}(s_{0(i-1)}$
$, s_{0i}, t_{2i}), \underline{y}_2(i), \underline{\hat{z}}_2(\hat{s}_i, , r_{2i}), \underline{x}_2(r_{2i}))$ and $(\underline{v}_0(s_{0i}), \underline{y}_2(i+1), \underline{\hat{z}}_2(\hat{s}_{i+1}, r_{2(i+1)}), \underline{x}_2(r_{2(i+1)}))$ are all jointly typical. This leads to the next constraints

$$S_0 + T_2 \leq I(V_0 U_0 U_2; Y_2 \hat{Z}_2|X_2), \tag{A.18}$$

$$T_2 \leq I(U_2; Y_2 \hat{Z}_2|V_0 U_0 X_2). \tag{A.19}$$

It is interesting to remark that regular coding allows us to use the same code for DF and CF scenarios, while keeping the same final CF rate.

After decoding of (s_{0i}, s_{1i}, s_{2i}) at destinations, the original messages (w_{0i}, w_{1i}, w_{2i}) can be extracted. One can see that the rate region of Theorem 25 follows form equations (A.9)-(A.19), the equalities between the original rates and reorganized rates, the fact that all the rates are positive and by using Fourier-Motzkin elimination. Similarly to [1], the necessary condition $I(X_2; Y_2) \geq I(Z_2; \hat{Z}_2|X_2, Y_2)$ follows from (A.15) and (A.17).

A.3 Proof of Theorem 26

Reorganize first private messages w_i, $i = \{1,2\}$ into (s'_i, s_i) with non-negative rates (S'_i, S_i) where $R_i = S'_i + S_i$. Merge (s'_1, s'_2, w_0) to one message s_0 with rate $S_0 = R_0 + S'_1 + S'_2$.

For the sake of notation, it is assumed that $\underline{u} = u_1^n$.

Code Generation:

(i) Generate 2^{nS_0} i.i.d. sequences $\underline{u}_0$ with PD

$$P_{U_0}(\underline{u}_0) = \prod_{j=1}^{n} p_{U_0}(u_{0j}),$$

and index them as $\underline{u}_0(s_0)$ with $s_0 = \left[1 : 2^{nS_0}\right]$.

(ii) Generate $2^{nR_{x_b}}$ i.i.d. sequences $\underline{x}_b$ with PD

$$P_{X_b}(\underline{x}_b) = \prod_{j=1}^{n} p_{X_b}(x_{bj})$$

as $\underline{x}_b(r_b)$, where $r_b = \left[1 : 2^{nR_{x_b}}\right]$ for $b = \{1, 2\}$.

(iii) For each $\underline{x}_b(r_b)$ generate $2^{n\hat{R}_b}$ i.i.d. sequences $\underline{\hat{z}}_b$ each with PD

$$P_{\hat{Z}_b|X_b}(\underline{\hat{z}}_b|\underline{x}_b(r_b)) = \prod_{j=1}^{n} p_{\hat{Z}_b|X_b}(\hat{z}_{bj}|x_{bj}(r_b)),$$

and index them as $\underline{\hat{z}}_b(r_b, \hat{s}_b)$, where $\hat{s}_b = \left[1 : 2^{n\hat{R}_b}\right]$ for $b = \{1, 2\}$.

(iv) Partition the set $\left\{1, \ldots, 2^{n\hat{R}_b}\right\}$ into $2^{nR_{x_b}}$ cells and label them as S_{r_2}. In each cell there are $2^{n(\hat{R}_b - R_{x_b})}$ elements.

(v) For each pair $\underline{u}_0(s_0)$, generate 2^{nT_b} i.i.d. sequences $\underline{u}_b$ with PD

$$P_{U_b|U_0}(\underline{u}_b|\underline{u}_0(s_0)) = \prod_{j=1}^{n} p_{U_b|U_0}(u_{bj}|u_{0j}(s_0)),$$

and index them as $\underline{u}_b(s_0, t_b)$, where $t_b = \left[1 : 2^{nT_b}\right]$ for $b = \{1, 2\}$.

(vi) For $b = \{1, 2\}$, partition the set $\left\{1, \ldots, 2^{nT_b}\right\}$ into 2^{nS_b} subsets and label them as S_{s_b}. In each subset, there are $2^{n(T_b - S_b)}$ elements for $b = \{1, 2\}$.

(vii) Finally, use a deterministic function for generating $\underline{x}$ as $f\left(\underline{u}_1, \underline{u}_2\right)$ indexed by $\underline{x}(s_0, t_1, t_2)$.

Encoding Part: In block i, the source wants to send (w_{0i}, w_{1i}, w_{2i}) by reorganizing them into (s_{0i}, s_{1i}, s_{2i}). Encoding steps are as follows:

(i) Relay b knows from the previous block that $\hat{s}_{b(i-1)} \in S_{r_{bi}}$ and it sends $\underline{x}_b(r_{bi})$ for $b = \{1, 2\}$.

(ii) Look for $t_{1i} \in S_{s_{1i}}$ and $t_{2i} \in S_{s_{2i}}$ such that $\big(\underline{u}_1(s_{0i}, t_{1i}), \underline{u}_2(s_{0i}, t_{2i})\big)$ are jointly typical given the RV $\underline{u}_0(s_{0i})$. The constraints for guaranteeing the success of this step is given by

$$T_1 + T_2 - S_1 - S_2 \geq I(U_2; U_1|U_0). \tag{A.20}$$

At the end, choose one pair $(t_{1(i-1)}, t_{2(i-1)})$ satisfying these conditions.

(iii) From (s_{0i}, s_{1i}, s_{2i}), the source finds (t_{1i}, t_{2i}) and sends $\underline{x}(s_{0i}, t_{1i}, t_{2i})$.

Decoding Part: In each block the relays start to find $\hat{s}_{bi}$ for that block. After the transmission of the block $i+1$, the destinations decode $\hat{s}_{bi}$ and then use it to find $\hat{Z}_b$ which along with Y_b is used to decode the messages.

(i) Relay b searches for $\hat{s}_{bi}$ after receiving $\underline{z}_b(i)$ such that $\big(\underline{x}_b(r_{bi}), \underline{z}_b(i), \underline{\hat{z}}_b(\hat{s}_{bi}, r_{bi})\big)$ are jointly typical subject to

$$\hat{R}_b \geq I(Z_b; \hat{Z}_b|X_b). \tag{A.21}$$

(ii) Destination b searches for $r_{b(i+1)}$ such that $\big(\underline{y}_b(i+1), \underline{x}_b(r_{b(i+1)})\big)$ is jointly typical. Then in finds $\hat{s}_{bi}$ such that $\hat{s}_{bi} \in S_{r_{b(i+1)}}$ and $\big(\underline{\hat{z}}_b(\hat{s}_{bi}, r_{bi}), \underline{y}_b(i), \underline{x}_b(r_{bi})\big)$ are jointly typical. Conditions for reliable decoding are:

$$R_{x_b} \leq I(X_b; Y_b), \;\; \hat{R}_b \leq R_{x_b} + I(\hat{Z}_b; Y_b|X_b). \tag{A.22}$$

(iii) Decoding in block i is done such that $(\underline{u_0}(s_{0i}), \underline{u_b}(s_{0i}, t_{bi}), \underline{y}_b(i), \underline{\hat{z}}_b(\hat{s}_{bi}, r_{bi}), \underline{x}_b(r_{bi}))$ are all jointly typical. This leads to the next constraints

$$S_0 + T_b \leq I(U_0, U_b; Y_b\hat{Z}_b|X_b), \tag{A.23}$$

$$T_b \leq I(U_b; Y_b, \hat{Z}_b|U_0, X_b). \tag{A.24}$$

After decoding of (s_{0i}, s_{1i}, s_{2i}) at destinations, the original messages (w_{0i}, w_{1i}, w_{2i}) can be extracted. One can see that the rate region of Theorem 26 follows form equations (A.20)-(A.24), the equalities between the original rates and reorganized rates, the fact that all the rates are positive and by using Fourier-Motzkin elimination technique. Similarly to [1], the necessary condition $I(X_b; Y_b) \geq I(Z_b; \hat{Z}_b|X_b, Y_b)$ follows from (A.21) and (A.22) for $b = \{1, 2\}$.

A.4 Proof of Theorem 27

Before proceeding the proof we state the following lemmas which is the generalization of a similar equality in [10] and it can be proved in a similar way.

Lemma 2 *For the random variable W, and the ensemble of n random variables $\mathbf{S}_j = (S_{j1}, S_{j2}$
$, ..., S_{jn})$ for $j \in \{1, 2, ..., M\}$ and $\mathbf{T}_k = (T_{k1}, T_{k2}, ..., T_{kn})$ for $k \in \{1, 2, ..., N\}$, the following equality holds:*

$$\sum_{i=1}^{n} I(T^n_{1(i+1)}, T^n_{2(i+1)}, ..., T^n_{N(i+1)}; S_{1i}, S_{2i}, ..., S_{Mi}|W, S_1^{i-1}, S_2^{i-1}, ..., S_M^{i-1}) = \sum_{i=1}^{n} I(S_1^{i-1}, S_2^{i-1}, ..., S_M^{i-1}; T_{1i}, T_{2i}, ..., T_{Ni}|W, T^n_{1(i+1)}, T^n_{2(i+1)}, ..., T^n_{N(i+1)}). \tag{A.25}$$

The proof can be done using the same procedure as [10]. Also the following equation will be used during the proofs.

$$I(A; B|D) - I(A; C|D) = I(A; B|C, D) - I(A; C|B, D). \tag{A.26}$$

For any code $(n, \mathcal{W}_0, \mathcal{W}_1, \mathcal{W}_2, P_e^{(n)})$ (i.e. with rates (R_0, R_1, R_2)), Fano's inequality will lead to:

$$\begin{aligned} H(W_0|\mathbf{Y}_2) &\leq P_e^{(n)} nR_0 + 1 \stackrel{\Delta}{=} n\epsilon_0, \\ H(W_1|\mathbf{Y}_1) &\leq H(W_0, W_1|\mathbf{Y}_1) \\ &\leq P_e^{(n)} n(R_0 + R_1) + 1 \stackrel{\Delta}{=} n\epsilon_1, \end{aligned}$$

$$\begin{aligned} H(W_2|\mathbf{Y}_2) &\leq H(W_0, W_2|\mathbf{Y}_2) \\ &\leq P_e^{(n)} n(R_0 + R_2) + 1 \stackrel{\Delta}{=} n\epsilon_2, \end{aligned}$$

We start with the following inequality:

$$\begin{aligned} n(R_0 + R_1 + R_2) - n(\epsilon_0 + \epsilon_1 + \epsilon_2) &\leq I(W_0; \mathbf{Y}_1) + I(W_1; \mathbf{Y}_1) + I(W_2; \mathbf{Y}_2) \\ &\leq I(W_0; \mathbf{Y}_1) + I(W_1; \mathbf{Y}_1, W_0, W_2) + I(W_2; \mathbf{Y}_2, W_0) \\ &\leq I(W_0, W_1, W_2; \mathbf{Y}_1) - I(W_2; \mathbf{Y}_1|W_0) + I(W_2; \mathbf{Y}_2|W_0). \end{aligned} \tag{A.27}$$

We can bound the first term of (A.27) on the right hand side as follows:

$$
\begin{aligned}
I(W_0, W_1, W_2; \mathbf{Y}_1) &= \sum_{i=1}^{n} I(W_0, W_1, W_2; Y_{1i}|Y_1^{i-1}) \\
&= \sum_{i=1}^{n} \left[H(Y_{1i}|Y_1^{i-1}) - H(Y_{1i}|Y_1^{i-1}, W_0, W_1, W_2)\right] \\
&\overset{(a)}{\leq} \sum_{i=1}^{n} \left[H(Y_{1i}) - H(Y_{1i}|Y_1^{i-1}, W_0, W_1, W_2, Y_{2(i+1)}^{n})\right] \\
&\overset{(b)}{=} \sum_{i=1}^{n} \left[H(Y_{1i}) - H(Y_{1i}|V_i, U_{1i}, U_{2i})\right] \\
&= \sum_{i=1}^{n} I(V_i, U_{1i}, U_{2i}; Y_{1i})
\end{aligned}
$$

where (a) is due to the fact that conditioning decreases the entropy and (b) is based on the definitions of $V_i = (W_0, Y_1^{i-1}, Y_{2(i+1)}^{n})$,$U_{1i} = (W_1, Y_1^{i-1}, Y_{2(i+1)}^{n})$ and $U_{2i} = (W_2, Y_1^{i-1}, Y_{2(i+1)}^{n})$. Now we continue with the proof as follows

$$
\begin{aligned}
I(W_2; \mathbf{Y}_2|W_0) - I(W_2; \mathbf{Y}_1|W_0) &= \sum_{i=1}^{n} \left[I(W_2; Y_{2i}|W_0, Y_{2(i+1)}^{n}) - I(W_2; Y_{1i}|W_0, Y_1^{i-1})\right] \\
&= \sum_{i=1}^{n} \Big[I(W_2, Y_1^{i-1}; Y_{2i}|W_0, Y_{2(i+1)}^{n}) - I(Y_1^{i-1}; Y_{2i}|W_2, W_0, Y_{2(i+1)}^{n}) \\
&\quad - I(W_2, Y_{2(i+1)}^{n}; Y_{1i}|W_0, Y_1^{i-1}) + I(Y_{2(i+1)}^{n}; Y_{1i}|W_2, W_0, Y_1^{i-1})\Big] \\
&\overset{(c)}{=} \sum_{i=1}^{n} \left[I(W_2, Y_1^{i-1}; Y_{2i}|W_0, Y_{2(i+1)}^{n}) - I(W_2, Y_{2(i+1)}^{n}; Y_{1i}|W_0, Y_1^{i-1})\right] \\
&= \sum_{i=1}^{n} \Big[I(W_2; Y_{2i}|W_0, Y_1^{i-1}, Y_{2(i+1)}^{n}) + I(Y_1^{i-1}; Y_{2i}|W_0, Y_{2(i+1)}^{n}) \\
&\quad - I(W_2; Y_{1i}|W_0, Y_1^{i-1}, Y_{2(i+1)}^{n}) - I(Y_{2(i+1)}^{n}; Y_{1i}|W_0, Y_1^{i-1})\Big] \\
&\overset{(d)}{=} \sum_{i=1}^{n} \left[I(W_2; Y_{2i}|W_0, Y_1^{i-1}, Y_{2(i+1)}^{n}) - I(W_2; Y_{1i}|W_0, Y_1^{i-1}, Y_{2(i+1)}^{n})\right],
\end{aligned}
$$

where (c) and (d) are due to Lemma 2 by choosing $M = N = 1$ and $\mathbf{T}_1 = \mathbf{Y}_1, \mathbf{S}_1 = \mathbf{Y}_2$, and respectively $W = (W_0, W_2)$ and $W = W_0$. Now the right hand side of (A.27) can be simplified as

$$n(R_0+R_1+R_2)-n(\epsilon_0+\epsilon_1+\epsilon_2) \leq \sum_{i=1}^{n}\left[I(V_i,U_{1i},U_{2i};Y_{1i})+I(U_{2i};Y_{2i}|V_i)-I(U_{2i};Y_{1i}|V_i)\right]$$
$$=\sum_{i=1}^{n}\left[I(V_i;Y_{1i})+I(U_{2i};Y_{2i}|V_i)+I(U_{1i},U_{2i};Y_{1i}|V_i)-I(U_{2i};Y_{1i}|V_i)\right]$$
$$=\sum_{i=1}^{n}\left[I(V_i;Y_{1i})+I(U_{2i};Y_{2i}|V_i)+I(U_{1i};Y_{1i}|U_{2i},V_i)\right], \tag{A.28}$$

yielding the first inequality. Now we move to the next inequality

$$n(R_0+R_1+R_2)-n(\epsilon_0+\epsilon_1+\epsilon_2) \leq I(W_0,W_1,W_2;\mathbf{Y}_1)-I(W_2;\mathbf{Y}_1|W_0)+I(W_2;\mathbf{Y}_2|W_0)$$
$$\leq I(W_0,W_1,W_2;\mathbf{Y}_1,\mathbf{Z}_1)-I(W_2;\mathbf{Y}_1,\mathbf{Z}_1|W_0)+I(W_2;\mathbf{Y}_2,\mathbf{Z}_2|W_0). \tag{A.29}$$

By using a similar method we obtain

$$I(W_0,W_1,W_2;\mathbf{Y}_1,\mathbf{Z}_1)=\sum_{i=1}^{n} I(W_0,W_1,W_2;Y_{1i},Z_{1i}|Y_1^{i-1},Z_1^{i-1})$$
$$=\sum_{i=1}^{n}\left[H(Y_{1i},Z_{1i}|Y_1^{i-1},Z_1^{i-1})-H(Y_{1i},Z_{1i}|Y_1^{i-1},Z_1^{i-1},W_0,W_1,W_2)\right]$$
$$\overset{(e)}{=}\sum_{i=1}^{n}\left[H(Y_{1i},Z_{1i}|Y_1^{i-1},Z_1^{i-1},X_{1i})-H(Y_{1i},Z_{1i}|Y_1^{i-1},Z_1^{i-1},X_{1i},W_0,W_1,W_2)\right]$$
$$\overset{(f)}{\leq}\sum_{i=1}^{n}\left[H(Y_{1i},Z_{1i}|X_{1i})-H(Y_{1i},Z_{1i}|Y_1^{i-1},Z_1^{i-1},W_0,W_1,W_2,X_{1i},Y_{2(i+1)}^{n},Z_{2(i+1)}^{n})\right]$$
$$=\sum_{i=1}^{n} I(V_i,V_{1i},U_{1i},U_{2i};Y_{1i},Z_{1i}|X_{1i}),$$

where (e) follows because X_{1i} is a function of the past relay output, (f) is the result of decreasing entropy by its conditioning and V_{1i} is denoted by $(Z_1^{i-1},Z_{2(i+1)}^{n})$. In a similar way to above we can obtain

$$I(W_2;\mathbf{Y}_2,\mathbf{Z}_2|W_0)-I(W_2;\mathbf{Y}_1,\mathbf{Z}_1|W_0)$$
$$=\sum_{i=1}^{n}\left[I(W_2;Y_{2i},Z_{2i}|W_0,Y_{2(i+1)}^{n},Z_{2(i+1)}^{n})-I(W_2;Y_{1i},Z_{1i}|W_0,Y_1^{i-1},Z_1^{i-1})\right]$$
$$\overset{(g)}{\leq}\sum_{i=1}^{n}\Big[I(W_2;Y_{2i},Z_{2i}|W_0,X_{1i},Y_1^{i-1},Z_1^{i-1},Y_{2(i+1)}^{n},Z_{2(i+1)}^{n})$$
$$-I(W_2;Y_{1i}|W_0,X_{1i},Y_1^{i-1},Z_1^{i-1},Y_{2(i+1)}^{n},Z_{2(i+1)}^{n})\Big],$$

where the step (g) can be proven by using the same procedure as the steps (c) and (d). Then

$$\begin{aligned}
&n(R_0+R_1+R_2)-n(\epsilon_0+\epsilon_1+\epsilon_2)\\
&\qquad\le \sum_{i=1}^{n}\big[I(V_i,V_{1i},U_{1i},U_{2i};Y_{1i},Z_{1i}|X_{1i})+I(U_{2i};Y_{2i},Z_{2i}|V_i,V_{1i},X_{1i})\\
&\qquad\qquad -I(U_{2i};Y_{1i},Z_{1i}|V_i,V_{1i},X_{1i})\big]\\
&\qquad=\sum_{i=1}^{n}\big[I(V_i,V_{1i};Y_{1i},Z_{1i}|X_{1i})+I(U_{2i};Y_{2i},Z_{2i}|V_i,V_{1i},X_{1i})\\
&\qquad\qquad +I(U_{1i};Y_{1i},Z_{1i}|X_{1i},U_{2i},V_i,V_{1i})\big].
\end{aligned}\tag{A.30}$$

Now take the following inequality

$$\begin{aligned}
n(R_0+R_1+R_2)-n(\epsilon_0+\epsilon_1+\epsilon_2)&\le I(W_0;\mathbf{Y}_2)+I(W_1;\mathbf{Y}_1)+I(W_2;\mathbf{Y}_2)\\
&\le I(W_0,W_1,W_2;\mathbf{Y}_2)-I(W_1;\mathbf{Y}_2|W_0)+I(W_1;\mathbf{Y}_1|W_0).
\end{aligned}\tag{A.31}$$

We again bound the first term on the right hand side as follows similar to previous one

$$\begin{aligned}
I(W_0,W_1,W_2;\mathbf{Y}_2)&=\sum_{i=1}^{n}I(W_0,W_1,W_2;Y_{2i}|Y_{2(i+1)}^n)\\
&=\sum_{i=1}^{n}\big[H(Y_{2i}|Y_{2(i+1)}^n)-H(Y_{2i}|Y_{2(i+1)}^n,W_0,W_1,W_2)\big]\\
&\le\sum_{i=1}^{n}\big[H(Y_{2i})-H(Y_{2i}|Y_{2(i+1)}^n,W_0,W_1,W_2,Y_1^{i-1})\big]\\
&=\sum_{i=1}^{n}\big[H(Y_{2i})-H(Y_{2i}|Y_{2(i+1)}^n,W_0,W_1,W_2,Y_1^{i-1})\big]\\
&=\sum_{i=1}^{n}I(V_i,U_{1i},U_{2i};Y_{2i}).
\end{aligned}$$

Now for the next terms we obtain

$$
\begin{aligned}
I(W_1;\mathbf{Y}_1|W_0) - I(W_1;\mathbf{Y}_2|W_0) &= \sum_{i=1}^{n}\big[I(W_1;Y_{1i}|W_0,Y_1^{i-1}) - I(W_1;Y_{2i}|W_0,Y_{2(i+1)}^n)\big]\\
&= \sum_{i=1}^{n}\big[I(W_1,Y_{2(i+1)}^n;Y_{1i}|W_0,Y_1^{i-1}) - I(Y_{2(i+1)}^n;Y_{1i}|W_1,W_0,Y_1^{i-1})\\
&- I(W_1,Y_1^{i-1};Y_{2i}|W_0,Y_{2(i+1)}^n) + I(Y_1^{i-1};Y_{2i}|W_1,W_0,Y_{2(i+1)}^n)\big]\\
&\overset{(h)}{=} \sum_{i=1}^{n}\big[I(W_1,Y_{2(i+1)}^n;Y_{1i}|W_0,Y_1^{i-1}) - I(W_1,Y_1^{i-1};Y_{2i}|W_0,Y_{2(i+1)}^n\big])\\
&= \sum_{i=1}^{n}\big[I(W_1;Y_{1i}|W_0,Y_1^{i-1},Y_{2(i+1)}^n) + I(Y_{2(i+1)}^n;Y_{1i}|W_0,Y_1^{i-1})\\
&- I(W_1;Y_{2i}|W_0,Y_1^{i-1},Y_{2(i+1)}^n) - I(Y_1^{i-1};Y_{2i}|W_0,Y_{2(i+1)}^n)\big]\\
&\overset{(i)}{=} \sum_{i=1}^{n}\big[I(W_1;Y_{1i}|W_0,Y_1^{i-1},Y_{2(i+1)}^n) - I(W_1;Y_{2i}|W_0,Y_1^{i-1},Y_{2(i+1)}^n)\big],
\end{aligned}
$$

where (h) and (i) are due to Lemma 2 by choosing $M = N = 1$ and $\mathbf{T}_1 = \mathbf{Y}_1, \mathbf{S}_1 = \mathbf{Y}_2$, and respectively $W = (W_0, W_1)$ and $W = W_0$. Now we simplify the right hand side of (A.31) to

$$
\begin{aligned}
n(R_0+R_1+R_2) - n(\epsilon_0+\epsilon_1+\epsilon_2) &\le \sum_{i=1}^{n}\big[I(V_i,U_{1i},U_{2i};Y_{2i}) + I(U_{1i};Y_{1i}|V_i) - I(U_{1i};Y_{2i}|V_i)\big]\\
&= \sum_{i=1}^{n}\big[I(V_i;Y_{2i}) + I(U_{1i};Y_{1i}|V_i) + I(U_{2i};Y_{2i}|U_{1i},V_i)\big]. \qquad \text{(A.32)}
\end{aligned}
$$

We can see the symmetry between (A.28) and (A.32). Another inequality, symmetric to (A.30) and (A.29) can be proved in a same way

$$
\begin{aligned}
n(R_0+R_1+R_2) - n(\epsilon_0+\epsilon_1+\epsilon_2) &\le I(W_0,W_1,W_2;\mathbf{Y}_2) - I(W_1;\mathbf{Y}_2|W_0) + I(W_1;\mathbf{Y}_1|W_0)\\
&\le I(W_0,W_1,W_2;\mathbf{Y}_2,\mathbf{Z}_2) + I(W_1;\mathbf{Y}_1,\mathbf{Z}_1|W_0) - I(W_1;\mathbf{Y}_2,\mathbf{Z}_2|W_0). \qquad \text{(A.33)}
\end{aligned}
$$

Now by following similar steps we can also show

$$
\begin{aligned}
I(W_0,W_1,W_2;\mathbf{Y}_2,\mathbf{Z}_2) &= \sum_{i=1}^{n} I(W_0,W_1,W_2;Y_{2i},Z_{2i}|Y_{2(i+1)}^n,Z_{2(i+1)}^n)\\
&= \sum_{i=1}^{n}\big[H(Y_{2i},Z_{2i}|Y_{2(i+1)}^n,Z_{2(i+1)}^n) - H(Y_{2i},Z_{2i}|Y_{2(i+1)}^n,Z_{2(i+1)}^n,W_0,W_1,W_2)\big]\\
&\overset{(j)}{\leq} \sum_{i=1}^{n}\big[H(Y_{2i},Z_{2i}) - H(Y_{2i},Z_{2i}|Y_{2(i+1)}^n,Z_{2(i+1)}^n,Y_1^{i-1},Z_1^{i-1},W_0,W_1,W_2)\\
&= \sum_{i=1}^{n} I(V_i,V_{1i},U_{1i},U_{2i};Y_{2i},Z_{2i})\big]\\
&\overset{(k)}{=} \sum_{i=1}^{n}\big[I(V_i,V_{1i};Y_{2i},Z_{2i}) + I(U_{1i},U_{2i};Y_{2i},Z_{2i}|V_i,V_{1i},X_{1i})\big],
\end{aligned}
$$

where (k) is because X_{1i} is a function of the past relay output (V_{1i}) and (j) is the result of decreasing entropy by its conditioning. In a similar way to before we can show

$$
\begin{aligned}
&I(W_1;\mathbf{Y}_1,\mathbf{Z}_1|W_0) - I(W_1;\mathbf{Y}_2,\mathbf{Z}_2|W_0)\\
&\quad= \sum_{i=1}^{n}\big[I(W_2;Y_{1i},Z_{1i}|W_0,Y_1^{i-1},Z_1^{i-1})\\
&\qquad\qquad - I(W_1;Y_{2i},Z_{2i}|W_0,Y_{2(i+1)}^n,Z_{2(i+1)}^n)\big]\\
&\quad\overset{(l)}{\leq} \sum_{i=1}^{n}\big[I(W_1;Y_{1i}|W_0,X_{1i},Y_1^{i-1},Z_1^{i-1},Y_{2(i+1)}^n,Z_{2(i+1)}^n)\\
&\qquad\qquad - I(W_1;Y_{2i},Z_{2i}|W_0,X_{1i},Y_1^{i-1},Z_1^{i-1},Y_{2(i+1)}^n,Z_{2(i+1)}^n)\big],
\end{aligned}
$$

where step (l) can be proved using the same procedure as for steps (e) and (f). Finally, we found

$$
\begin{aligned}
&n(R_0+R_1+R_2) - n(\epsilon_0+\epsilon_1+\epsilon_2)\\
&\quad\leq \sum_{i=1}^{n}\big[I(V_i,V_{1i};Y_{2i},Z_{2i}) + I(U_{1i},U_{2i};Y_{2i},Z_{2i}|V_i,V_{1i},X_{1i})\\
&\qquad\qquad + I(U_{1i};Y_{1i},Z_{1i}|V_i,V_{1i},X_{1i}) - I(U_{1i};Y_{2i},Z_{2i}|V_i,V_{1i},X_{1i})\big]\\
&\quad= \sum_{i=1}^{n}\big[I(V_i,V_{1i};Y_{2i},Z_{2i}) + I(U_{2i};Y_{2i},Z_{2i}|V_i,V_{1i},U_{1i},X_{1i})\\
&\qquad\qquad + I(U_{1i};Y_{1i},Z_{1i}|X_{1i},V_i,V_{1i})\big].
\end{aligned}
\tag{A.34}
$$

The inequalities (A.28), (A.30), (A.32) and (A.34) are related to the sum of R_0, R_1, R_2. For the rest of the proof we focus on the following inequalities:

$$\begin{aligned} nR_0 &\leq I(W_0; \mathbf{Y}_2) + n\epsilon_0, \\ n(R_0 + R_1) &\leq I(W_0; \mathbf{Y}_2) + I(W_1; \mathbf{Y}_1|W_0) + n(\epsilon_0 + \epsilon_1), \\ n(R_0 + R_2) &\leq I(W_0; \mathbf{Y}_1) + I(W_2; \mathbf{Y}_2|W_0) + n(\epsilon_0 + \epsilon_2). \end{aligned}$$

Starting from the last inequality, we have

$$\begin{aligned} &n(R_0 + R_1) - n(\epsilon_0 + \epsilon_1) \leq I(W_0; \mathbf{Y}_2) + I(W_1; \mathbf{Y}_1|W_0) \\ &\quad = \sum_{i=1}^{n} \big[I(W_0; Y_{2i}|Y_{2(i+1)}^n) + I(W_1; Y_{1i}|Y_1^{i-1}, W_0) \big] \\ &\quad = \sum_{i=1}^{n} \big[I(W_0, Y_1^{i-1}; Y_{2i}|Y_{2(i+1)}^n) - I(Y_1^{i-1}; Y_{2i}|W_0, Y_{2(i+1)}^n) + I(W_1; Y_{1i}|Y_1^{i-1}, W_0) \big] \\ &\quad \overset{(a')}{=} \sum_{i=1}^{n} \big[I(W_0, Y_1^{i-1}; Y_{2i}|Y_{2(i+1)}^n) - I(Y_{2(i+1)}^n; Y_{1i}|W_0, Y_1^{i-1}) + I(W_1; Y_{1i}|Y_1^{i-1}, W_0) \big] \\ &\quad \overset{(b')}{=} \sum_{i=1}^{n} \big[I(W_0, Y_1^{i-1}; Y_{2i}|Y_{2(i+1)}^n) + I(W_1; Y_{1i}|Y_{2(i+1)}^n, Y_1^{i-1}, W_0) \\ &\quad - I(Y_{2(i+1)}^n; Y_{1i}|W_1, W_0, Y_1^{i-1}) \big] \\ &\quad \leq \sum_{i=1}^{n} \big[I(W_0, Y_{2(i+1)}^n, Y_1^{i-1}; Y_{2i}) + I(W_1; Y_{1i}|Y_1^{i-1}, Y_{2(i+1)}^n, W_0) \big] \\ &\quad \leq \sum_{i=1}^{n} \big[I(V_i; Y_{2i}) + I(U_{1i}; Y_{1i}|V_i) \big], \end{aligned} \tag{A.35}$$

where (a') comes from the Lemma 2 with choosing $M = N = 1, S_1 = Y_1, T_1 = Y_2, W = W_0$, (b') comes from the (A.26). With a similar procedure it can be proved that

$$\begin{aligned} n(R_0 + R_2) - n(\epsilon_0 + \epsilon_2) &\leq I(W_0; \mathbf{Y}_1) + I(W_2; \mathbf{Y}_2|W_0) \\ &\leq \sum_{i=1}^{n} \big[I(V_i; Y_{1i}) + I(U_{2i}; Y_{2i}|V_i) \big]. \end{aligned} \tag{A.36}$$

Now we move to the next inequality

$$
\begin{aligned}
n(R_0+R_1) - n(\epsilon_0+\epsilon_1) &\leq I(W_0;\mathbf{Y}_2) + I(W_1;\mathbf{Y}_1|W_0) \\
&\leq I(W_0;\mathbf{Y}_2,\mathbf{Z}_2) + I(W_1;\mathbf{Y}_1,\mathbf{Z}_1|W_0) \\
&= \sum_{i=1}^{n} \big[I(W_0;Y_{2i},Z_{2i}|Y_{2(i+1)}^n,Z_{2(i+1)}^n) + I(W_1;Y_{1i},Z_{1i}|Y_1^{i-1},Z_1^{i-1},W_0) \big] \\
&= \sum_{i=1}^{n} \big[I(W_0,Z_1^{i-1},Y_1^{i-1};Z_{2i},Y_{2i}|Y_{2(i+1)}^n,Z_{2(i+1)}^n) \\
&\qquad - I(Y_1^{i-1},Z_1^{i-1};Y_{2i},Z_{2i}|W_0,Y_{2(i+1)}^n,Z_{2(i+1)}^n) \\
&\qquad + I(W_1;Y_{1i},Z_{1i}|Y_1^{i-1},Z_1^{i-1},W_0) \big] \\
&\overset{(c')}{=} \sum_{i=1}^{n} \big[I(W_0,Z_1^{i-1},Y_1^{i-1};Z_{2i},Y_{2i}|Y_{2(i+1)}^n,Z_{2(i+1)}^n) \\
&\qquad - I(Y_{2(i+1)}^n,Z_{2(i+1)}^n;Y_{1i},Z_{1i}|W_0,Y_1^{i-1},Z_1^{i-1}) \\
&\qquad + I(W_1;Y_{1i},Z_{1i}|Y_1^{i-1},Z_1^{i-1},W_0) \big] \\
&\overset{(d')}{=} \sum_{i=1}^{n} \big[I(W_0,Z_1^{i-1},Y_1^{i-1};Z_{2i},Y_{2i}|Y_{2(i+1)}^n,Z_{2(i+1)}^n) \\
&\qquad + I(W_1;Y_{1i},Z_{1i}|Y_1^{i-1},Z_1^{i-1},Y_{2(i+1)}^n,Z_{2(i+1)}^n,W_0) \\
&\qquad - I(Y_{2(i+1)}^n,Z_{2(i+1)}^n;Y_{1i},Z_{1i}|W_1,W_0,Y_1^{i-1},Z_1^{i-1}) \big] \\
&\leq \sum_{i=1}^{n} \big[I(W_0,Z_1^{i-1},Y_1^{i-1};Z_{2i},Y_{2i}|Y_{2(i+1)}^n,Z_{2(i+1)}^n) \\
&\qquad + I(W_1;Y_{1i},Z_{1i}|Y_1^{i-1},Z_1^{i-1},Y_{2(i+1)}^n,Z_{2(i+1)}^n,W_0) \big] \\
&\overset{(e')}{\leq} \sum_{i=1}^{n} \big[I(W_0,Y_1^{i-1},Z_1^{i-1},Z_{2(i+1)}^n,Y_{2(i+1)}^n;Z_{2i},Y_{2i}) \\
&\qquad + I(W_1;Y_{1i},Z_{1i}|Y_1^{i-1},Z_1^{i-1},Y_{2(i+1)}^n,Z_{2(i+1)}^n,W_0,X_{1i}) \big].
\end{aligned}
$$

Now by using the previous definitions, we obtain

$$
n(R_0+R_1) - n(\epsilon_0+\epsilon_1) = \sum_{i=1}^{n} \big[I(V_i,V_{1i};Z_{2i},Y_{2i}) + I(U_{1i};Y_{1i},Z_{1i}|V_i,V_{1i},X_{1i}) \big], \tag{A.37}
$$

where (c') comes from the Lemma 2 by choosing $M = N = 2$, $T_1 = Y_2, S_1 = Y_1, T_2 = Z_2, S_2 = Z_1, W = W_0$, (d') comes from (A.26), (e') is due to the fact that X_{1i} is a function of Z_1^{i-1}. And finally the proof of the final sum rate is as follows

$$\begin{aligned}
n(R_0+R_2)-n(\epsilon_0+\epsilon_2) &\le I(W_0;\mathbf{Y}_1)+I(W_2;\mathbf{Y}_2|W_0)\\
&\le I(W_0;\mathbf{Y}_1,\mathbf{Z}_1)+I(W_2;\mathbf{Y}_2,\mathbf{Z}_2|W_0)\\
&=\sum_{i=1}^{n}\big[I(W_0;Y_{1i},Z_{1i}|Y_1^{i-1},Z_1^{i-1})+I(W_2;Y_{2i},Z_{2i}|Y_{2(i+1)}^n,Z_{2(i+1)}^n,W_0)\big]\\
&=\sum_{i=1}^{n}\big[I(W_0,Y_{2(i+1)}^n,Z_{2(i+1)}^n;Z_{1i},Y_{1i}|Y_1^{i-1},Z_1^{i-1})\\
&\qquad -I(Y_{2(i+1)}^n,Z_{2(i+1)}^n;Y_{1i},Z_{1i}|W_0,Y_1^{i-1},Z_1^{i-1})\\
&\qquad +I(W_2;Y_{2i},Z_{2i}|Y_{2(i+1)}^n,Z_{2(i+1)}^n,W_0)\big]\\
&\overset{(f')}{=}\sum_{i=1}^{n}\big[I(W_0,Y_{2(i+1)}^n,Z_{2(i+1)}^n;Z_{1i},Y_{1i}|Y_1^{i-1},Z_1^{i-1})\\
&\qquad -I(Y_1^{i-1},Z_1^{i-1};Y_{2i},Z_{2i}|W_0,Y_{2(i+1)}^n,Z_{2(i+1)}^n)\\
&\qquad +I(W_2;Y_{2i},Z_{2i}|Y_{2(i+1)}^n,Z_{2(i+1)}^n,W_0)\big]\\
&\overset{(g')}{=}\sum_{i=1}^{n}\big[I(W_0,Y_{2(i+1)}^n,Z_{2(i+1)}^n;Z_{1i},Y_{1i}|Y_1^{i-1},Z_1^{i-1})\\
&\qquad +I(W_2;Y_{2i},Z_{2i}|Y_1^{i-1},Z_1^{i-1},Y_{2(i+1)}^n,Z_{2(i+1)}^n,W_0)\\
&\qquad -I(Y_1^{i-1},Z_1^{i-1};Y_{2i},Z_{2i}|W_2,W_0,Y_{2(i+1)}^n,Z_{2(i+1)}^n)\big]\\
&\le\sum_{i=1}^{n}\big[I(W_0,Y_{2(i+1)}^n,Z_{2(i+1)}^n;Z_{1i},Y_{1i}|Y_1^{i-1},Z_1^{i-1})\\
&\qquad +I(W_2;Y_{2i},Z_{2i}|Y_1^{i-1},Z_1^{i-1},Y_{2(i+1)}^n,Z_{2(i+1)}^n,W_0)\big]\\
&\overset{(h')}{=}\sum_{i=1}^{n}\big[I(W_0,Y_{2(i+1)}^n,Z_{2(i+1)}^n;Z_{1i},Y_{1i}|Y_1^{i-1},Z_1^{i-1},X_{1i})\\
&\qquad +I(W_2;Y_{2i},Z_{2i}|Y_1^{i-1},Z_1^{i-1},Y_{2(i+1)}^n,Z_{2(i+1)}^n,W_0,X_{1i})\big]\\
&\le\sum_{i=1}^{n}\big[I(W_0,Y_1^{i-1},Z_1^{i-1},Y_{2(i+1)}^n,Z_{2(i+1)}^n;Z_{1i},Y_{1i}|X_{1i})\\
&\qquad +I(W_2;Y_{2i},Z_{2i}|Y_1^{i-1},Z_1^{i-1},Y_{2(i+1)}^n,Z_{2(i+1)}^n,W_0,X_{1i})\big].
\end{aligned}$$

Again using previous definitions we obtain

$$n(R_0+R_2)-n(\epsilon_0+\epsilon_2)\le\sum_{i=1}^{n}I(V_i,V_{1i};Z_{1i},Y_{1i}|X_{1i})+I(U_{2i};Y_{2i},Z_{2i}|V_i,V_{1i},X_{1i}),\tag{A.38}$$

where (f') comes from the Lemma 2 with the choice $M = N = 2$, $S_1 = Y_1, T_1 = Y_2, S_2 = Z_1, T_2 = Z_2, W = W_0$, (g') comes from (A.26), (h') is due to the fact that X_{1i} is a function of Z_1^{i-1}.

Finally, we prove the first two inequalities

$$\begin{aligned} n(R_0 + R_1) - n(\epsilon_0 + \epsilon_1) &\leq I(W_0, W_1; \mathbf{Y}_1) \\ &= \sum_{i=1}^{n} I(W_0, W_1; Y_{1i} | Y_1^{i-1}) \\ &\leq \sum_{i=1}^{n} I(Y_1^{i-1}, W_0, W_1; Y_{1i}) \\ &\leq \sum_{i=1}^{n} I(Y_{2(i+1)}^{n}, Y_1^{i-1}, W_0, W_1; Y_{1i}) \\ &= \sum_{i=1}^{n} I(V_i, U_{1i}; Y_{1i}), \end{aligned} \tag{A.39}$$

and similarly we derive

$$n(R_0 + R_2) - n(\epsilon_0 + \epsilon_2) \leq \sum_{i=1}^{n} I(V_i, U_{2i}; Y_{2i}). \tag{A.40}$$

The next step is to prove another bound on the sum rate $R_0 + R_1$

$$\begin{aligned} n(R_0 + R_1) - n(\epsilon_0 + \epsilon_1) &\leq I(W_0, W_1; \mathbf{Y}_1, \mathbf{Z}_1) \\ &= \sum_{i=1}^{n} I(W_0, W_1; Y_{1i}, Z_{1i} | Y_1^{i-1}, Z_1^{i-1}) \\ &= \sum_{i=1}^{n} I(W_0, W_1; Y_{1i}, Z_{1i} | Y_1^{i-1}, Z_1^{i-1}, X_{1i}) \\ &\leq \sum_{i=1}^{n} I(Y_1^{i-1}, Z_1^{i-1}, W_0, W_1; Y_{1i}, Z_{1i} | X_{1i}) \\ &\leq \sum_{i=1}^{n} I(Y_{2(i+1)}^{n}, Z_{2(i+1)}^{n}, Y_1^{i-1}, Z_1^{i-1}, W_0, W_1; Y_{1i}, Z_{1i} | X_{1i}) \\ &= \sum_{i=1}^{n} I(V_i, V_{1i}, U_{1i}; Y_{1i}, Z_{1i} | X_{1i}). \end{aligned} \tag{A.41}$$

Similarly for the sum rate $R_0 + R_2$

$$\begin{aligned}
n(R_0 + R_2) - n(\epsilon_0 + \epsilon_2) &\le I(W_0, W_2; \mathbf{Y}_2, \mathbf{Z}_2) \\
&= \sum_{i=1}^{n} I(W_0, W_2; Y_{2i}, Z_{2i} | Y_{2(i+1)}^n, Z_{2(i+1)}^n) \\
&\le \sum_{i=1}^{n} I(Y_{2(i+1)}^n, Z_{2(i+1)}^n, W_0, W_2; Y_{2i}, Z_{2i}) \\
&\le \sum_{i=1}^{n} I(Y_{2(i+1)}^n, Z_{2(i+1)}^n, Y_1^{i-1}, Z_1^{i-1}, W_0, W_2; Y_{2i}, Z_{2i}) \\
&= \sum_{i=1}^{n} I(V_i, V_{1i}, U_{2i}; Y_{2i}, Z_{2i}) \\
&= \sum_{i=1}^{n} \big[I(V_i, V_{1i}; Y_{2i}, Z_{2i}) + I(U_{2i}; Y_{2i}, Z_{2i} | V_i, V_{1i}) \big] \\
&\stackrel{(i')}{=} \sum_{i=1}^{n} \big[I(V_i, V_{1i}; Y_{2i}, Z_{2i}) + I(U_{2i}; Y_{2i}, Z_{2i} | V_i, V_{1i}, X_{1i}) \big], \quad \text{(A.42)}
\end{aligned}$$

where (i') is due to the fact that X_{1i} is function of Z_1^{i-1} and so function of V_{1i}.

And at last we bound the rate R_0

$$\begin{aligned}
nR_0 - n\epsilon_0 &\le I(W_0; \mathbf{Y}_1) \\
&= \sum_{i=1}^{n} I(W_0; Y_{1i} | Y_1^{i-1}) \\
&\le \sum_{i=1}^{n} I(Y_1^{i-1}, W_0; Y_{1i}) \\
&\le \sum_{i=1}^{n} I(Y_{2(i+1)}^n, Y_1^{i-1}, W_0; Y_{1i}) \\
&= \sum_{i=1}^{n} I(V_i; Y_{1i}). \quad \text{(A.43)}
\end{aligned}$$

Similarly for Y_2

$$\begin{aligned}
nR_0 - n\epsilon_0 &\le I(W_0; \mathbf{Y}_2) \\
&\le \sum_{i=1}^{n} I(V_i; Y_{2i}). \quad \text{(A.44)}
\end{aligned}$$

The rest of the proof is as usual with resort to an independent time-sharing RV Q and applying it to (A.28)-(A.44) which yields the final region.

A.5 Proof of Theorem 30

Note that the upper bound can be proved to be a special case of the outer bound presented in Theorem 29 in semi-degraded BRC. But we prove the converse indpendently here. For proving the upper bound in the Theorem 30, we start with the fact that the user 1 is decoding all the information. For any code $(n, \mathcal{W}_1, \mathcal{W}_2, P_e^{(n)})$ (i.e. (R_1, R_2)), we start from Fano's inequality:

$$H(W_2|\mathbf{Y}_2) \leq P_e^{(n)} nR_2 + 1 \overset{\Delta}{=} n\epsilon_0,$$
$$H(W_1|\mathbf{Y}_1) \leq P_e^{(n)} nR_1 + 1 \overset{\Delta}{=} n\epsilon_1,$$

and

$$\begin{aligned} nR_2 &\leq I(W_2;\mathbf{Y}_2) + n\epsilon_0, \\ n(R_1+R_2) - n\epsilon_0 - n\epsilon_1 &\leq I(W_2;\mathbf{Y}_2) + I(W_1;\mathbf{Y}_1) \\ &\leq I(W_2;\mathbf{Y}_2) + I(W_1;\mathbf{Y}_1, W_2) \\ &\leq I(W_2;\mathbf{Y}_2) + I(W_1;\mathbf{Y}_1|W_2). \end{aligned}$$

Before starting the proof, we state the following lemma.

Lemma 3 *For the BRC-CR with the condition $X \ominus (Y_1, X_1) \ominus Z_1$, the following relation holds*

$$H(Y_{1i}|Y_1^{i-1}, W_2) = H(Y_{1i}|Y_1^{i-1}, Z_1^{i-1}, X_1^i, W_2).$$

Proof

$$\begin{aligned} H(Y_{1i}|Y_1^{i-1}, W_2) &= H(Y_{1i}|Y_{11}, Y_{12}, ..., Y_{1(i-1)}, W_2) \\ &\overset{(a)}{=} H(Y_{1i}|Y_{11}, X_{11}, Y_{12}, ..., Y_{1(i-1)}, W_2) \\ &\overset{(b)}{=} H(Y_{1i}|Y_{11}, X_{11}, Z_{11}, Y_{12}, ..., Y_{1(i-1)}, W_2) \\ &\overset{(c)}{=} H(Y_{1i}|Y_{11}, X_{11}, Z_{11}, X_{12}, Y_{12}, ..., Y_{1(i-1)}, W_2) \\ &\vdots \\ &= H(Y_{1i}|Y_{11}, X_{11}, Z_{11}, Y_{12}, X_{12}, Z_{12} ..., Y_{1(i-1)}, X_{1(i-1)}, Z_{1(i-1)}, X_{1i}, W_2) \\ &= H(Y_{1i}|Y_1^{i-1}, Z_1^{i-1}, X_1^i, W_2), \end{aligned}$$

where (a) follows since $X_{1i} = f_{1,i}(Z_1^{i-1})$, for $i = 1$, X_{11} is chosen as constant because the argument of the function is empty, so it can be added for free, (b) is due to the Markovity assumption of the lemma where given X_{11}, Y_{11}, Z_{11} can be added for free. Now $X_{12} = f_{1,2}(Z_{11})$ and it can be added for free and this justifies (c). With the same argument, we can continue to add first $Z_{1(j-1)}$ given $Y_{1(j-1)}, X_{1(j-1)}$ and then X_{1j} given $Z_{1(j-1)}$ until $j = i$ and this will conclude the proof.

By setting $U_i = (Y_2^{i-1}, Z_1^{i-1}, X_1^{i-1}, W_2)$, it can be shown that

$$
\begin{aligned}
I(W_1; \mathbf{Y}_1 | W_2) &= \sum_{i=1}^{n} I(W_1; Y_{1i} | Y_1^{i-1}, W_2) \\
&= \sum_{i=1}^{n} \left[H(Y_{1i} | Y_1^{i-1}, W_2) - H(Y_{1i} | Y_1^{i-1}, W_2, W_1) \right] \\
&\overset{(a)}{\leq} \sum_{i=1}^{n} \left[H(Y_{1i} | Y_1^{i-1}, Z_1^{i-1}, X_1^{i}, W_2) - H(Y_{1i} | X_i, X_{1i}, Y_1^{i-1}, W_2, W_1) \right] \\
&\overset{(b)}{=} \sum_{i=1}^{n} \left[H(Y_{1i} | Y_1^{i-1}, Y_2^{i-1}, Z_1^{i-1}, X_1^{i}, W_2) - H(Y_{1i} | X_i, X_{1i}, Y_1^{i-1}, W_2, W_1) \right] \\
&\overset{(c)}{=} \sum_{i=1}^{n} \left[H(Y_{1i} | Y_1^{i-1}, Y_2^{i-1}, Z_1^{i-1}, X_1^{i}, W_2) - H(Y_{1i} | X_i, X_{1i}) \right] \\
&\overset{(d)}{\leq} \sum_{i=1}^{n} \left[H(Y_{1i} | Y_2^{i-1}, Z_1^{i-1}, X_1^{i-1}, W_2, X_{1i}) - H(Y_{1i} | X_i, X_{1i}, Y_2^{i-1}, Z_1^{i-1}, X_1^{i-1}, W_2) \right] \\
&= \sum_{i=1}^{n} I(X_i; Y_{1i} | Y_2^{i-1}, Z_1^{i-1}, X_1^{i-1}, W_2, X_{1i}) \\
&= \sum_{i=1}^{n} I(X_i, X_{1i}; Y_{1i} | U_i, X_{1i}),
\end{aligned}
$$

where (a) results from the Lemma 3, (b) results from the Markov chain $Y_{2i} \ominus (Z_{1i}, X_{1i}) \ominus X_i$ while (c) and (d) is because Y_{1i} depends only on (X_i, X_{1i}).

For the next bound we have

$$
\begin{aligned}
I(W_2;\mathbf{Y}_2) &\leq I(W_2;\mathbf{Y}_2,\mathbf{Z}_1)\\
&= \sum_{i=1}^{n} I(W_2;Y_{1i},Z_{1i}|Y_1^{i-1},Z_1^{i-1})\\
&= \sum_{i=1}^{n} \left[H(W_2|Y_1^{i-1},Z_1^{i-1}) - H(W_2|Y_1^{i},Z_1^{i})\right]\\
&\overset{(e)}{\leq} \sum_{i=1}^{n} \left[H(W_2|Z_1^{i-1},X_1^{i}) - H(W_2|X_1^{i},Z_1^{i})\right]\\
&= \sum_{i=1}^{n} \left[H(Z_{1i}|Z_1^{i-1},X_1^{i-1},X_{1i}) - H(Z_{1i}|X_{1i},X_1^{i-1},Z_1^{i-1},W_2)\right]\\
&\overset{(f)}{=} \sum_{i=1}^{n} \left[H(Z_{1i}|Z_1^{i-1},X_1^{i-1},X_{1i}) - H(Z_{1i}|X_{1i},Z_1^{i-1},X_1^{i-1},Y_2^{i-1},W_2)\right]\\
&\leq \sum_{i=1}^{n} \left[H(Z_{1i}|X_{1i}) - H(Z_{1i}|X_{1i},Z_1^{i-1},X_1^{i-1},Y_2^{i-1},W_2)\right]\\
&= \sum_{i=1}^{n} I(Z_1^{i-1},X_1^{i-1},Y_2^{i-1},W_2;Z_{1i}|X_{1i})\\
&= \sum_{i=1}^{n} I(U_i;Z_{1i}|X_{1i}).
\end{aligned}
$$

Based on the definition X_{1i} is available given Z_1^{i-1}. But Z_1^{i-1} also includes Z_1^j for all the $j \leq i-1$, therefore given Z_1^{i-1}, $X_{11}, X_{12}, ..., X_{1(i-1)}$ and thus X_1^i are also available. This justifies (e). Then with Z_1^{i-1}, X_1^{i-1} and using Markovity between (Z_1, X_1) and (Y_2), one can say that Y_2^{i-1} is also available given Z_1^{i-1}. Step (f) results from this fact.

For the last inequality, we have

$$
\begin{aligned}
I(W_2;\mathbf{Y}_2) &= \sum_{i=1}^{n} I(W_2;Y_{2i}|Y_2^{i-1})\\
&\leq \sum_{i=1}^{n} I(Y_2^{i-1},W_0;Y_{2i})\\
&\leq \sum_{i=1}^{n} I(Z_1^{i-1},X_1^{i-1},Y_2^{i-1},W_2;Y_{2i}) = \sum_{i=1}^{n} I(U_i;Y_{2i}).
\end{aligned}
$$

Finally, the bound can be proved using an independent time sharing RV Q.

A.6 Proof of Theorem 31

We now prove the outer bound in Theorem 31. First, notice that the second bound is the capacity of a degraded relay channel, shown in [1]. Regarding the fact that destination 1 is decoding all the information, the bound can be reached by using the same method. Therefore the focus is on the other bounds. For any code $(n, \mathcal{W}_0, \mathcal{W}_1, P_e^{(n)})$ (i.e. (R_0, R_1)), we want to show that if the error probability goes to zero then the rates satisfy the conditions in Theorem 31. From Fano's inequality we have that

$$\begin{aligned} H(W_0|\mathbf{Y}_2) &\leq P_e^{(n)} nR_0 + 1 \stackrel{\Delta}{=} n\epsilon_0, \\ H(W_1|\mathbf{Y}_1) &\leq H(W_0, W_1|\mathbf{Y}_1) \leq P_e^{(n)} n(R_0 + R_1) + 1 \stackrel{\Delta}{=} n\epsilon_1, \end{aligned}$$

and

$$\begin{aligned} nR_0 &\leq I(W_0; \mathbf{Y}_2) + n\epsilon_0, \\ n(R_0 + R_1) - n\epsilon_0 - n\epsilon_1 &\leq I(W_0; \mathbf{Y}_2) + I(W_1; \mathbf{Y}_1) \leq I(W_0; \mathbf{Y}_2) + I(W_1; \mathbf{Y}_1, W_0), \\ &\leq I(W_0; \mathbf{Y}_2) + I(W_1; \mathbf{Y}_1|W_0). \end{aligned}$$

By setting $U_i = (Y_2^{i-1}, W_0)$, it can be shown that

$$\begin{aligned} I(W_1; \mathbf{Y}_1|W_0) &= \sum_{i=1}^{n} \left[I(W_1; Y_{1i}|Y_1^{i-1}, W_0) \right] \\ &= \sum_{i=1}^{n} \left[H(Y_{1i}|Y_1^{i-1}, W_0) - H(Y_{1i}|Y_1^{i-1}, W_0, W_1) \right] \\ &\stackrel{(a)}{\leq} \sum_{i=1}^{n} \left[H(Y_{1i}|Y_2^{i-1}, W_0) - H(Y_{1i}|X_i, X_{1i}, Y_1^{i-1}, W_0, W_1) \right] \\ &\stackrel{(b)}{=} \sum_{i=1}^{n} \left[H(Y_{1i}|Y_2^{i-1}, W_0) - H(Y_{1i}|X_i, X_{1i}) \right] \\ &\stackrel{(c)}{\leq} \sum_{i=1}^{n} \left[I(X_i, X_{1i}; Y_{1i}|Y_2^{i-1}, W_0) \right. \\ &= \sum_{i=1}^{n} \left. I(X_i, X_{1i}; Y_{1i}|U_i) \right], \end{aligned}$$

where (a) results from the degradedness between Y_1 and Y_2, where (b) and (c) require Markov chain Y_{1i} and (X_i, X_{1i}). Similarly, we have that

$$
\begin{aligned}
I(W_1;\mathbf{Y}_1|W_0) &\le I(W_1;\mathbf{Y}_1,\mathbf{Z}_1|W_0)\\
&= \sum_{i=1}^{n}\left[I(W_1;Y_{1i},Z_{1i}|Y_1^{i-1},Z_1^{i-1},W_0)\right]\\
&= \sum_{i=1}^{n}\left[H(W_1|Y_1^{i-1},Z_1^{i-1},W_0)-H(W_1|Y_1^{i},Z_1^{i},W_0)\right]\\
&\overset{(d)}{\le} \sum_{i=1}^{n}\left[H(W_1|Z_1^{i-1},X_{1i},W_0)-H(W_1|X_{1i},Z_1^{i},W_0)\right]\\
&= \sum_{i=1}^{n}\left[H(Z_{1i}|Z_1^{i-1},X_{1i},W_0)-H(Z_{1i}|X_{1i},Z_1^{i-1},W_0,W_1)\right]\\
&\le \sum_{i=1}^{n}\left[H(Z_{1i}|Z_1^{i-1},X_{1i},W_0)-H(Z_{1i}|X_i,X_{1i},Z_1^{i-1},W_0,W_1)\right]\\
&\overset{(e)}{\le} \sum_{i=1}^{n}\left[H(Z_{1i}|Y_2^{i-1},X_{1i},W_0)-H(Z_{1i}|X_i,X_{1i})\right]\\
&\overset{(f)}{=} \sum_{i=1}^{n}\left[H(Z_{1i}|Y_2^{i-1},X_{1i},W_0)-H(Z_{1i}|X_i,X_{1i},Y_2^{i-1},W_0)\right]\\
&= \sum_{i=1}^{n} I(X_i;Z_{1i}|X_{1i},Y_2^{i-1},W_0) = \sum_{i=1}^{n} I(X_i;Z_{1i}|X_{1i},U_i).
\end{aligned}
$$

Based on the definition X_{1i} can be obtained via Z_1^{i-1}, so given Z_1^{i-1} one can have X_1^{i-1}, and then with Z_1^{i-1}, X_1^{i-1} and using Markovity between (Z_1, X_1) and (Y_1, Y_2), one can say that (Y_1^{i-1}, Y_2^{i-1}) is also available given Z_1^{i-1}. Step (d) and (e) result from this fact. Markovity of Z_{1i} and (X_i, X_{1i}) has been used for (e) and (f). For the first inequality, we have

$$
\begin{aligned}
I(W_0;\mathbf{Y}_2) &= \sum_{i=1}^{n} I(W_0;Y_{2i}|Y_2^{i-1})\\
&\le \sum_{i=1}^{n} I(U_i;Y_{2i}).
\end{aligned}
$$

Finally, the bound can be proved using an independent time sharing RV Q.

A.7 Proof of Theorem 33

The proof of Theorem 33 in [33] is established directly for Gaussian models. Inner bound and outer bound provided there are of different forms. Their equivalence is established later using tuning technique which consists in tuning the inner bound to match the outer bound. In our case, these bounds can be obtained using Theorem 32 and 31. The outer bound is the same as [33] and the inner bound includes the inner bound of [33]. The equivalence of these bounds can be established then using the same argument as [33].

The inner bound is obtained using Theorem 33. Choose U and X_1 conditionally independent given V. This means that the source divides its power to θP and $\overline{\theta}P$ for the first and the second user and the relay does the same with dividing its power into $\theta_r P_1$ and $\overline{\theta_r}P_1$. Then γ and ρ represents correlation coefficient respectively between (U,V) and between (X_1,X). The inner bound is then calculated in the same way as [33] and involves less equation as presented there. The region is as follows:

$$\begin{aligned}
R_0 &\leq C\left(\frac{\overline{\theta}P + \overline{\theta_r}P_1 + 2\sqrt{\gamma\overline{\theta}\overline{\theta_r}PP_1}}{N_2 + \theta P + \theta_r P_1 + 2\sqrt{\rho\theta\theta_r PP_1}}\right)\\
R_1 &\leq \min\left\{C\left(\theta\overline{\rho}\frac{P}{\tilde{N}_1}\right), C\left(\frac{\theta P + \theta_r P_1 + 2\sqrt{\rho\theta\theta_r PP_1}}{N_1}\right)\right\}\\
R_0 + R_1 &\leq C\left((\theta\overline{\rho} + \overline{\theta}\overline{\gamma})\frac{P}{\tilde{N}_1}\right)
\end{aligned}$$

The region includes the region in [33].

We now focus on the upper bound which is calculated using Theorem 31. Let $h(\cdot)$ denotes the differential entropy where

$$I(U;Y_2) = h(Y_2) - h(Y_2|U).$$

We start by bounding $\sum_{i=1}^{n} h(Y_{2i})$. This can be bounded by

$$\sum_{i=1}^{n} h(Y_{2i}) \leq \frac{n}{2}\log\left[2\pi e(N_2 + P + P_1 + 2\sqrt{\overline{\beta}PP_1})\right],$$

where

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\mathbb{E}^2(X_i|X_{1i}) = \overline{\beta}P.$$

On the other hand, it can be shown that

$$\begin{aligned}\sum_{i=1}^{n} h(\mathcal{N}_{2i}) &= \sum_{i=1}^{n} h(Y_{2i}|U_i, X_i, X_{1i}) \\ &\leq \sum_{i=1}^{n} h(Y_{2i}|U_i) \\ &\leq \sum_{i=1}^{n} h(Y_{2i}),\end{aligned}$$

and as a result

$$\begin{aligned}\frac{n}{2}\log\left[2\pi e N_2\right] &\leq \sum_{i=1}^{n} h(Y_{2i}|U_i) \\ &\leq \frac{n}{2}\log\left[2\pi e(N_2 + P + P_1 + 2\sqrt{\beta P P_1})\right],\end{aligned}$$

so there exists $\alpha \in [0,1]$ such that

$$\sum_{i=1}^{n} h(Y_{2i}|U_i) = \frac{n}{2}\log\left[2\pi e(N_2 + \alpha(P + P_1 + 2\sqrt{\beta P P_1}))\right].$$

Using the entropy power inequality we have

$$\exp\left[\frac{2}{n}h(\mathbf{Y}_1|U)\right] \leq \exp\left[\frac{2}{n}h(\mathbf{Y}_2|U)\right] - \exp\left[\frac{2}{n}h(\mathcal{N}_2 - \mathcal{N}_1)\right],$$

and hence

$$\sum_{i=1}^{n} h(Y_{1i}|U_i) \leq \frac{n}{2}\log\left[2\pi e(N_1 + \alpha(P + P_1 + 2\sqrt{\beta P P_1}))\right].$$

On the other hand we have

$$\begin{aligned}&I(X, X_1; Y_1|U) = h(Y_1|U) - h(Y_1|X, X_1, U), \\ &h(Y_1|X, X_1, U) = h(\mathcal{N}_1).\end{aligned}$$

Using the constraints introduced before, the bounds are easily obtained by direct calculation. Finally, the calculation of $\sum_{i=1}^{n} I(X_i, Z_{1i}|X_{1i})$ is done like [1] by bounding

$$\sum_{i=1}^{n} h(Z_{1i}|X_{1i}) \leq \frac{n}{2}\log\left[2\pi e(\tilde{N}_2 + \beta P)\right]$$

with the similar definition of β as before. Then we obtain

$$\begin{aligned}I(X; Z_1|U, X_1) &= h(Z_1|U, X_1) - h(Z_1|X, X_1), \\ h(\tilde{\mathcal{N}}_1) &\leq h(Z_1|U, X_1) \leq h(Z_1|X_1), \\ h(Z_1|X, X_1) &= h(\tilde{\mathcal{N}}_1).\end{aligned}$$

Using the bound of $h(Z_1|X_1)$, it can be said that there is γ such that

$$\sum_{i=1}^{n} h(Z_{1i}|U_i, X_{1i}) = \frac{n}{2} \log(2\pi e(\tilde{N}_1 + \beta\gamma P)).$$

Using this we can bound $I(X; Z_1|U, X_1)$ as presented in the theorem. But it can be seen that γ appears only here and hence one can choose $\gamma = 1$ to maximize the region.

This concludes the proof since, as the author has proven in [33], the inner bound meets the upper bound.

A.8 Proof of Theorem 34

The direct part can be easily proved by using (6.40) and removing d_1 and d_2 from the definition of the channel. For the converse we start with the following lemma.

Lemma 4 *Any pair of rates (R_1, R_2) in the capacity region $\mathscr{C}_{BRC\text{-}PC}$ of the degraded BRC-PC satisfy the following inequalities*

$$\begin{aligned} nR_1 &\leq \sum_{i=1}^{n} I(U_i, X_{1i}; Y_{1i}) + n\epsilon_1, \\ nR_1 + nR_2 &\leq \sum_{i=1}^{n} I(U_i; Z_{1i}|X_{1i}) + I(X_i; Y_{2i}|U_i, X_{1i}) + n\epsilon_2. \end{aligned}$$

Proof This lemma can be obtained by taking $U_i = (W_1, Y_1^{i-1}, Z_1^{i-1}, Y_{2(i+1)}^n)$ and similar steps as in Appendix A.4. For this reason, we will not repeat the proof here. Note that only the degradedness between the relay and the first destination is necessary for the proof.

Now for the Gaussian degraded BRC-PC defined as before, we calculate the preceding bounds. The calculation follows the same steps as in Appendix A.6. We start by bounding $h(Z_{1i}|U_i, X_{1i})$ where it can be seen that

$$\begin{aligned} h(\tilde{\mathcal{N}}_{1i}) &= h(Z_{1i}|U_i, X_i, X_{1i}) \\ &\leq h(Z_{1i}|U_i, X_{1i}) \\ &\leq h(Z_{1i}) \\ &= h(X_i + \tilde{\mathcal{N}}_{1i}). \end{aligned}$$

Using this fact it can be said that

$$\begin{aligned}\frac{n}{2}\log\left[2\pi e\tilde{N}_1\right] &= \sum_{i=1}^{n} h(\tilde{\mathcal{N}}_{1i}) \\ &\leq \sum_{i=1}^{n} h(Z_{1i}|U_i, X_{1i}) \\ &\leq \sum_{i=1}^{n} h(X_i + \tilde{\mathcal{N}}_{1i}) \\ &= \frac{n}{2}\log\left[2\pi e(\tilde{N}_1 + P)\right].\end{aligned}$$

The previous condition implies that there is $\alpha \in [0,1]$ such that

$$\sum_{i=1}^{n} h(Z_{1i}|U_i, X_{1i}) = \frac{n}{2}\log\left[2\pi e(\tilde{N}_1 + \overline{\alpha}P)\right].$$

Note that the previous condition means that

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\mathbb{E}^2(X_i|U_i, X_{1i}) = \alpha P.$$

Now take the following inequalities

$$0 \leq \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\mathbb{E}^2(X_i|X_{1i}) \leq \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\mathbb{E}^2(X_i|U_i, X_{1i}) = \alpha P.$$

This is the result of $\mathbb{E}\mathbb{E}^2(X|Y) \leq \mathbb{E}\mathbb{E}^2(X|Y,Z)$ which can be proved using Jensen inequality. Similarly the previous condition implies that there exists $\beta \in [0,1]$ such that

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\mathbb{E}^2(X_i|X_{1i}) = \overline{\beta}\alpha P.$$

From this equality, we get the following inequalities by following the same technique as [1]

$$\sum_{i=1}^{n} h(Z_{1i}|X_{1i}) \leq \frac{n}{2}\log\left[2\pi e(\tilde{N}_1 + \overline{\alpha}P + \alpha\beta P)\right].$$

Also using this fact $h(Y_{1i})$ can be bounded by

$$\sum_{i=1}^{n} h(Y_{1i}) \leq \frac{n}{2}\log\left[2\pi e(N_1 + P + P_1 + 2\sqrt{\alpha\overline{\beta}PP_1})\right].$$

From the degradedness of Y_1 respect to Z_1 and Y_2, and using entropy power inequality we obtain

$$\sum_{i=1}^{n} h(Y_{1i}|U_i, X_{1i}) \geq \frac{n}{2} \log\left[2\pi e(N_1 + \overline{\alpha}P)\right],$$
$$\sum_{i=1}^{n} h(Y_{2i}|U_i, X_{1i}) \leq \frac{n}{2} \log\left[2\pi e(N_2 + \overline{\alpha}P)\right],$$

and these bounds prove the upper bound and conclude the proof.

Appendix B

Appendix chapter 3

B.1 Proof of Theorem 35 and Corollary 36

B.1.1 Proof of Theorem 35

Assume a set $\mathcal{M}_n$ of size 2^{nR} of message indices W to be transmitted. Transmission is done in $B+L$ blocks, each of them of length n, and decoding at the destination is done backwardly. At the last $L-1$ blocks, the last compression index is first decoded and then all compression indices and transmitted messages are jointly decoded. Table B.1 shows the messages and codewords over different blocks.

Code generation:

(i) Randomly and independently generate 2^{nR} sequences $\underline{x}_1$ drawn i.i.d. from

$$P^n_{X_1}(\underline{x}_1) = \prod_{j=1}^{n} P_{X_1}(x_{1j}).$$

Index them as $\underline{x}_1(r_1)$ with index $r_1 \in \left[1, 2^{nR}\right]$.

(ii) Randomly and independently generate $2^{n\hat{R}}$ sequences $\underline{x}_2$ drawn i.i.d. from

$$P^n_{X_2}(\underline{x}_2) = \prod_{j=1}^{n} P_{X_2}(x_{2j}).$$

Index them as $\underline{x}_2(r_2)$, where $r_2 \in \left[1, 2^{n\hat{R}}\right]$ for $\hat{R} = I(Z_2;\hat{Z}_2|X_2) + \epsilon$.

(iii) For each $\underline{x}_2(r_2)$, randomly and conditionally independently generate $2^{n\hat{R}}$ sequences $\underline{\hat{z}}_2$ each with probability

$$P^n_{\hat{Z}_2|X_2}(\underline{\hat{z}}_2|\underline{x}_2(r_2)) = \prod_{j=1}^{n} P_{\hat{Z}_2|X_2}(\hat{z}_{2j}|x_{2j}(r_2)).$$

Table B.1: Coding for the two-relay channel

$b=1$	$b=2$	...	$b=B$	$b=B+1$	$b=B+2$	...	$b=B+L$
$\underline{x}_1(1)$	$\underline{x}_1(w_1)$	...	$\underline{x}_1(w_{(B-1)})$	$\underline{x}_1(w_B)$	$\underline{x}_1(1)$	...	$\underline{x}_1(1)$
$\underline{x}(1,w_1)$	$\underline{x}(w_1,w_2)$	...	$\underline{x}(w_{(B-1)},w_B)$	$\underline{x}(w_B,1)$	$\underline{x}(1,1)$	...	$\underline{x}(1,1)$
$\underline{x}_2(1)$	$\underline{x}_2(l_1)$	...	$\underline{x}_2(l_{(B-1)})$	$\underline{x}_2(l_B)$	$\underline{x}_2(l_{(B+1)})$	...	$\underline{x}_2(l_{(B+1)})$
$\underline{\hat{z}}_2(1,l_1)$	$\underline{\hat{z}}_2(l_1,l_2)$	...	$\underline{\hat{z}}_2(l_{(B-1)},l_B)$	$\underline{\hat{z}}_2(l_B,l_{(B+1)})$	.	.	.
$\underline{z}_1(1)$	$\underline{z}_1(2)$	...	$\underline{z}_1(B)$	$\underline{z}_1(B+1)$	$\underline{z}_1(B+2)$	...	$\underline{z}_1(B+L)$
$\underline{z}_2(1)$	$\underline{z}_2(2)$	...	$\underline{z}_2(B)$	$\underline{z}_2(B+1)$	$\underline{z}_2(B+2)$	...	$\underline{z}_2(B+L)$
$\underline{y}_1(1)$	$\underline{y}_1(2)$	...	$\underline{y}_1(B)$	$\underline{y}_1(B+1)$	$\underline{y}_1(B+2)$	...	$\underline{y}_1(B+L)$

Index them as $\underline{\hat{z}}_2(r_2,\hat{s})$, where $\hat{s} \in \left[1, 2^{n\hat{R}}\right]$.

(iv) For each $\underline{x}_1(r_1)$, randomly and conditionally independently generate 2^{nR} sequences $\underline{x}$ drawn i.i.d. from

$$P^n_{X|X_1}(\underline{x}|\underline{x}_1(r_1)) = \prod_{j=1}^{n} P_{X|X_1}(x_j|x_{1j}).$$

Index them as $\underline{x}(r_1,w)$, where $w \in \left[1, 2^{nR}\right]$.

(v) Provide the corresponding codebooks to the relays, the encoder and the decoder ends.

Encoding part:

(i) In every block $i=[1:B]$, the source sends w_i based on $\underline{x}\big(w_{(i-1)},w_i\big)$. Moreover, for blocks $i=[B+1:B+L]$, the source sends the dummy message $w_i=1$ known to all users.

(ii) For every block $i=[1:B+L]$, the relay 1 knows $w_{(i-1)}$ since by assumption and $w_0=1$, so it sends $\underline{x}_1\big(w_{(i-1)}\big)$.

(iii) For each $i=[1:B+1]$ and after receiving $\underline{z}_2(i)$, relay 2 searches for at least one index l_i with $l_0=1$ such that

$$\big(\underline{x}_2(l_{(i-1)}), \underline{z}_2(i), \underline{\hat{z}}_2(l_{(i-1)},l_i)\big) \in \mathcal{A}^n_\epsilon[X_2Z_2\hat{Z}_2].$$

The probability of finding such l_i goes to one as n goes to infinity.

(iv) For $i=[1:B+1]$, relay 2 knows from the previous block $l_{(i-1)}$ and it sends $\underline{x}_2(l_{(i-1)})$. Moreover, relay 2 repeats l_{B+1} for $i=[B+2:B+L]$.

Decoding part:

(i) After the transmission of the block $i = [1 : B]$, the first relay starts to decode the messages of block i with the assumption that all messages up to block $i-1$ have been correctly decoded. Relay 1 searches for the unique index $\hat{w}_i \in \mathcal{M}_n$ such that:

$$\big(\underline{x}\big(w_{(i-1)}, \hat{w}_i\big), \underline{x}_1\big(w_{(i-1)}\big), \underline{z}_1(i)\big)\big) \in \mathcal{A}_\epsilon^n[XX_1Z_1].$$

By following similar argument than [1], the probability of error goes to zero as n goes to infinity provided that:

$$R \leq I(X; Z_1|X_1). \tag{B.1}$$

(ii) If the condition (7.9) fails, i.e. if

$$I(\hat{Z}_2; Z_2|Y_1XX_1X_2) > I(X_2; Y_1|XX_1)$$

then the destination finds the unique $\hat{w}_{(i-1)}$ with the assumption that w_i has been decoded successfully using backward decoding such that

$$\Big(\underline{x}\big(\hat{w}_{(i-1)}, w_i\big), \underline{x}_1\big(\hat{w}_{(i-1)}\big), \underline{y}_1(i)\big)\Big) \in \mathcal{A}_\epsilon^n[XX_1Y_1].$$

The probability of error for this decoding goes to zero if

$$R \leq I(XX_1; Y_1). \tag{B.2}$$

(iii) If the condition (7.9) holds, then the destination waits until the last block and uses backward decoding [42]. It searches for the unique index $\hat{l}_{B+1}$ such that for all $b \in [B+2 : B+L]$, we have

$$\big(\underline{x}_2(\hat{l}_{B+1}), \underline{x}(1,1), \underline{x}_1(1), \underline{y}_1(b)\big) \in \mathcal{A}_\epsilon^n[XX_2X_1Y_1].$$

Define the following events:

$$\begin{aligned}
\mathcal{E}_1 &= \Big\{ \big(\underline{x}_2(l_{B+1}), \underline{x}(1,1), \underline{x}_1(1), \underline{y}_1(b)\big) \notin \mathcal{A}_\epsilon^n[XX_2X_1Y_1] \Big\}, \\
\mathcal{E}_2 &= \Big\{ \big(\underline{x}_2(\hat{l}), \underline{x}(1,1), \underline{x}_1(1), \underline{y}_1(b)\big) \in \mathcal{A}_\epsilon^n[XX_2X_1Y_1] \\
&\qquad \text{for some } \hat{l}_{B+1} \neq l_{B+1} \text{ and all } b \in [B+2 : B+L] \Big\}.
\end{aligned}$$

The probability of error is bounded as follows

$$\Pr\left(\hat{l}_{B+1} \neq l_{B+1}\right) \leq \Pr(\mathcal{E}_1) + \Pr(\mathcal{E}_2).$$

The first probability on the right side goes to zero as $n \to \infty$. The second probability also goes to zero as n goes to infinity provided that

$$\begin{aligned}\Pr(\mathcal{E}_2) \leq & \sum_{\hat{l} \neq l_{B+1}} \Pr\left[\bigcap_{b=[B+2:B+L]} \left\{\left(\underline{x}_2(\hat{l}), \underline{x}(1,1), \underline{x}_1(1), \underline{y}_1(b)\right) \in \mathcal{A}_\epsilon^n[XX_2X_1Y_1]\right\}\right] \\ \leq & \left(2^{n\hat{R}} - 1\right)\left[2^{-n(I(X_2;XX_1Y_1)-\epsilon_1)}\right]^{L-1},\end{aligned}$$

which leads to the following condition for $\Pr(\mathcal{E}_2)$ going to zero:

$$I(\hat{Z}_2; Z_2|X_2) + \epsilon_2 \leq (L-1)I(X_2; XX_1Y_1). \tag{B.3}$$

(iv) After finding l_{B+1} and since $w_{(B+1)} = 1$, the destination decodes jointly (w_b, l_b) for each $b = [1:B]$ where decoding is performed backwardly with the assumption that (w_{b+1}, l_{b+1}) have been correctly decoded. It finds the unique pair of indices $(\hat{w}_b, \hat{l}_b)$ such that

$$\left(\underline{x}(\hat{w}_b, w_{(b+1)}), \underline{x}_1(\hat{w}_b), \underline{y}_1(b+1), \underline{x}_2(\hat{l}_b), \underline{\hat{z}}_2(\hat{l}_b, l_{(b+1)})\right) \in \mathcal{A}_\epsilon^n[XX_1X_2\hat{Z}_2Y_1].$$

Consider the following error events associated with this step:

$$\begin{aligned}\mathcal{E}_3 =& \left\{\left(\underline{x}(w_b, w_{(b+1)}), \underline{x}_1(w_b), \underline{y}_1(b+1), \underline{x}_2(l_b), \underline{\hat{z}}_2(l_b, l_{(b+1)})\right) \notin \mathcal{A}_\epsilon^n[XX_1X_2\hat{Z}_2Y_1]\right\},\\ \mathcal{E}_4 =& \Big\{\left(\underline{x}(w_b, w_{(b+1)}), \underline{x}_1(w_b), \underline{y}_1(b+1), \underline{x}_2(\hat{l}), \underline{\hat{z}}_2(\hat{l}, l_{(b+1)})\right) \\ & \qquad \in \mathcal{A}_\epsilon^n[XX_1X_2\hat{Z}_2Y_1] \text{ for some } \hat{l} \neq l_b\Big\},\\ \mathcal{E}_5 =& \Big\{\left(\underline{x}(\hat{w}, w_{(b+1)}), \underline{x}_1(\hat{w}), \underline{y}_1(b+1), \underline{x}_2(l_b), \underline{\hat{z}}_2(l_b, l_{(b+1)})\right) \\ & \qquad \in \mathcal{A}_\epsilon^n[XX_1X_2\hat{Z}_2Y_1] \text{ for some } \hat{w} \neq w_b\Big\},\\ \mathcal{E}_6 =& \Big\{\left(\underline{x}(\hat{w}, w_{(b+1)}), \underline{x}_1(\hat{w}), \underline{y}_1(b+1), \underline{x}_2(\hat{l}), \underline{\hat{z}}_2(\hat{l}, l_{(b+1)})\right) \\ & \qquad \in \mathcal{A}_\epsilon^n[XX_1X_2\hat{Z}_2Y_1] \text{ for some } \hat{l} \neq l_b, \hat{w} \neq w_b\Big\}.\end{aligned}$$

The probability of error of this step can be bounded by

$$\Pr\left((\hat{w}_b, \hat{l}_b) \neq (w_b, l_b)\right) \leq \Pr(\mathcal{E}_3) + \Pr(\mathcal{E}_4) + \Pr(\mathcal{E}_5) + \Pr(\mathcal{E}_6).$$

First probability on the right side goes to zero as $n \to \infty$. Second and third probabilities on the right side also go to zero as $n \to \infty$ provided that

$$I(\hat{Z}_2; Z_2|X_2) + \epsilon_3 < I(X_2\hat{Z}_2; XX_1Y_1), \tag{B.4}$$

$$R \leq I(XX_1; Y_1\hat{Z}_2|X_2). \tag{B.5}$$

The last probability is bounded as follows:

$$\begin{aligned}
&\Pr(\mathcal{E}_6) \leq \\
&\sum_{\hat{l}\neq l_b, \hat{w}\neq w_b} \Pr\left[\left(\underline{x}(\hat{w}, w_{(b+1)}), \underline{x}_1(\hat{w}), \underline{y}_1(b+1), \underline{x}_2(\hat{l}), \underline{\hat{z}}_2(\hat{l}, l_{(b+1)})\right) \in \mathcal{A}_\epsilon^n[XX_1X_2\hat{Z}_2Y_1]\right] \\
&\leq \left(2^{n(R+\hat{R})} - 1\right)\left(2^{n(H(XX_1X_2Y_1\hat{Z}_2) - H(XX_1) - H(\hat{Z}_2X_2) - H(Y_1) + \epsilon_4)}\right).
\end{aligned}$$

Now given the following inequality the last probability also tends to zero as $n \to \infty$

$$R + I(\hat{Z}_2; Z_2|X_2) + \epsilon_5 < I(XX_1X_2; Y_1) + I(\hat{Z}_2; XX_1Y_1|X_2). \tag{B.6}$$

Particularly (B.6) can be simplified as

$$\begin{aligned}
R + \epsilon_5 &< I(XX_1X_2; Y_1) + I(XX_1Y_1; \hat{Z}_2|X_2) - I(\hat{Z}_2; Z_2|X_2) \\
&= I(XX_1X_2; Y_1) - I(\hat{Z}_2; Z_2|Y_1XX_1X_2).
\end{aligned} \tag{B.7}$$

Then from the inequality (B.4) we obtain

$$I(\hat{Z}_2; Z_2|Y_1XX_1X_2) \leq I(X_2; Y_1|XX_1), \tag{B.8}$$

where we used the fact that $I(\hat{Z}_2; Z_2|X_2) = I(\hat{Z}_2; Z_2XX_1Y_1|X_2)$. The inequality (B.4) holds because the condition (7.9) holds. By choosing L large enough, not necessarily infinite, the condition (B.3) holds too. By letting (B, n) tend to infinity such that $\frac{L}{B}$ goes to zero too, we get the following conditions if (7.9) holds:

$$R \leq \min\{I(XX_1X_2; Y_1) - I(\hat{Z}_2; Z_2|Y_1XX_1X_2), I(XX_1; \hat{Z}_2Y_1|X_2)\}. \tag{B.9}$$

And the following if (7.9) fails:

$$R \leq I(XX_1; Y_1). \tag{B.10}$$

However we have $I(XX_1; Y_1) \leq I(XX_1; \hat{Z}_2Y_1|X_2)$ and also we have

$$\begin{aligned}
&I(XX_1X_2; Y_1) - I(\hat{Z}_2; Z_2|Y_1XX_1X_2) = \\
&\qquad I(XX_1; Y_1) + \left[I(X_2; Y_1|XX_1) - I(\hat{Z}_2; Z_2|Y_1XX_1X_2)\right].
\end{aligned}$$

This means that if the condition (7.9) holds, (B.9) leads to a bigger region than (B.10), i.e.:

$$I(XX_1;Y_1) \leq \min\{I(XX_1X_2;Y_1) - I(\hat{Z}_2;Z_2|Y_1XX_1X_2), I(XX_1;\hat{Z}_2Y_1|X_2)\},$$

which means that decoding the compression index and using CF leads to a better rate. And if (7.9) fails it can be seen that (B.9) is worse that (B.10), i,e:

$$I(XX_1;Y_1) > \min\{I(XX_1X_2;Y_1) - I(\hat{Z}_2;Z_2|Y_1XX_1X_2), I(XX_1;\hat{Z}_2Y_1|X_2)\},$$

which means that it is better not to use CF at all. In other words the region obtained as such coincides with the following region:

$$R \leq \max\Big\{I(XX_1;Y_1), \min\{I(XX_1X_2;Y_1) - I(\hat{Z}_2;Z_2|Y_1XX_1X_2), I(XX_1;\hat{Z}_2Y_1|X_2)\}\Big\} \tag{B.11}$$

This region with (B.1) will prove the theorem. At the end the time sharing random variable Q can be added using usual manipulations. Note that unlike [42] it is not necessary that $L \to \infty$.

B.1.2 Proof of Corollary 36

Consider now the composite relay channel with parameters $\theta = (\theta_d, \theta_r)$. Transmission is done over $B + L$ blocks as well as for Theorem 35. Assume the same code generation than for Theorem 35 with transmission rate fixed to r.

Encoding part:

(i) In every block $i = [1 : B]$, the source sends $w_i \in [1, 2^{nr}]$ based on $\underline{x}\big(w_{(i-1)}, w_i\big)$. Moreover, for blocks $i = [B + 1 : B + L]$, the source sends the dummy message $w_i = 1$ known to all users.

(ii) The relay knows θ_r. If $\theta_r \in \mathcal{D}_{\mathrm{DF}}$, it sends a codeword $\underline{x}_{1\theta_r} = \underline{x}_1$ from the first codebook and uses it for the rest of communication. In other words, the relay function for this choice is DF scheme. In block i, the relay uses its decoder output $\hat{w}_{(i-1)}$ $(w_0 = 1)$ and it sends $\underline{x}_1\big(\hat{w}_{(i-1)}\big)$.
Otherwise, if $\theta_r \notin \mathcal{D}_{\mathrm{DF}}$, then the relay picks the codebook of codewords $\underline{x}_{1\theta_r} = \underline{x}_2$. The relay function in this case is CF scheme. In this case, after receiving the corresponding output that we denote $\underline{z}_{1,\theta_r}(i)$, the relay searches for at least one index l_i where $l_0 = 1$, such that

$$\big(\underline{x}_2(l_{(i-1)}), \underline{z}_{1,\theta_r}(i), \underline{\hat{z}}_2(l_{(i-1)}, l_i)\big) \in \mathcal{A}_\epsilon^n[X_2 Z_{1,\theta_r} \hat{Z}_2].$$

The probability of finding such l_i goes to one as n goes to infinity. Note that the typical set used for such coding is known to the relay because it knows θ_r. For $i = [1 : B+1]$, the relay knows from the previous block $l_{(i-1)}$ and it sends $\underline{x}_2(l_{(i-1)})$. Moreover, the relay repeats l_{B+1} for $i = [B+2 : B+L]$.

Decoding part:

1. For every block $i = [1 : B+L]$, the relay decodes w_i exactly similar to the Theorem 35. For a fixed r and given θ_r, the condition (B.1) should be satisfied to have correct decoding. The error is declared and relay declares a random output $\hat{w}$ if:

$$r > I(X; Z_{1\theta_r}|X_1). \tag{B.12}$$

2. The decoder knows θ and hence θ_r. If $\theta_r \in \mathcal{D}_{\text{DF}}$ then if the inequality (B.12) is satisfied, the error is declared. Otherwise it uses the DF decoder to decode the message and if the following inequality holds true, then an error is declared

$$r > I(X, X_1; Y_{1\theta}). \tag{B.13}$$

Therefore if $\theta_r \in \mathcal{D}_{\text{DF}}$, an error is declared if $r > I_{\text{DF}}$ with

$$I_{\text{DF}} = \min\left\{I(X; Z_{1\theta_r}|X_1), I(X, X_1; Y_{1\theta})\right\}. \tag{B.14}$$

Consider now the step when $\theta_r \notin \mathcal{D}_{\text{DF}}$. Note that in this case the relay input is $X_{1,\theta_r} = X_2$ and we have the following Markov chain:

$$X_1 \multimap X \multimap (Y_{1\theta}, Z_{1\theta_r}, X_2). \tag{B.15}$$

Moreover the decoder knows whether the following inequality is satisfied subject to the Markov chain (B.15)

$$I(X_2; Y_{1\theta}|XX_1) \geq I(Z_{1,\theta_r}; \hat{Z}_2|Y_{1\theta}XX_1X_2).$$

Now the decoder applies the exact same decoding procedure as in Theorem 35. It can be seen that the decoding conditions at destination do not change even if X_1 is not really transmitted only $I(XX_1; Y_{1\theta}) = I(X; Y_{1\theta})$. The only change is in the Markov chains. It can be seen that the previous inequality corresponds to (7.9) for the composite channels. The destination declares an error for $\theta_r \notin \mathcal{D}_{\text{DF}}$ if $r > I_{\text{CF}}$ where

$$I_{\text{CF}} = \max\left\{ \min\Big(I(XX_1; Y_{1\theta}\hat{Z}_2|X_2), I(XX_1X_2; Y_{1\theta}) - I(\hat{Z}_2; Z_{1\theta_r}|Y_{1\theta}XX_1X_2)\Big), I(X; Y_{1\theta})\right\}. \tag{B.16}$$

Using (B.14) and (B.16), the error event denoted by the indicator function $\mathbf{1}_E$ is as follows

$$\mathbf{1}_E = \mathbf{1}[\theta_r \in \mathcal{D}_{\text{DF}} \text{ and } r > I_{\text{DF}}] + \mathbf{1}[\theta_r \notin \mathcal{D}_{\text{DF}} \text{ and } r > I_{\text{CF}}]. \tag{B.17}$$

|q Taking the expected value from two sides give the expected probability of error. The expected value is taken in two steps. For each θ_r, the expected error is calculated using $\mathbb{P}_{\theta|\theta_r}$. The relay chooses the distribution of X_{2,θ_r} and $\hat{Z}_{2,\theta_r}$ to minimize this error:

$$\mathbb{E}_{\theta|\theta_r}[\mathbf{1}_E] = \begin{cases} \mathbb{P}_{\theta|\theta_r}\{r > I_{\text{DF}}|\theta_r\} & \theta_r \in \mathcal{D}_{\text{DF}} \\ \min\limits_{p(x_2)p(\hat{z}_2|x_2,z_{1\theta_r})} \mathbb{P}_{\theta|\theta_r}\{r > I_{\text{CF}}|\theta_r\} & \theta_r \notin \mathcal{D}_{\text{DF}}. \end{cases}$$

At the next step, the expected value is taken over θ_r and is minimized over $\mathcal{D}_{\text{DF}}$ and $p(x, x_1)$.

$$\begin{aligned} \overline{\epsilon}(r) = \min_{p(x,x_1)} \inf_{\mathcal{D}_{\text{DF}} \subseteq \Theta_r} \mathbb{E}_{\theta_r} \Big\{ & \mathbb{P}_{\theta|\theta_r}\big(r > I_{\text{DF}}, \theta_r \in \mathcal{D}_{\text{DF}}|\theta_r\big) \\ & + \min_{p(x_2)p(\hat{z}_2|x_2,z_{1\theta_r})} \mathbb{P}_{\theta|\theta_r}\big(r > I_{\text{CF}}, \theta_r \notin \mathcal{D}_{\text{DF}}|\theta_r\big) \Big\}. \end{aligned} \tag{B.18}$$

At the end time sharing RV Q can be added to the region, however the optimization should be done outside the expectation.

Optimizing over $\mathcal{D}_{DF}$ region. Suppose that the relay is not able to decode the message. Then, use DF scheme shall lead to an error event while CF scheme can perform much better. So if the condition $r > I(X; Z_{1\theta_r}|X_1)$ is satisfied, the best guess of the relay would be CF scheme. Now the question is what the proper guess would be if the relay can decode the message. In other words, the relevant question is whether CF scheme may turn to perform better than DF scheme when the relay is able to decode. Obviously, this is not the case. As a matter of fact, if the relay decodes and uses DF scheme, an error may occur for $r > I(XX_1; Y_{1\theta})$. Nevertheless, since X_2 is independent of X, but X is in general correlated with X_1, we have that

$$I(XX_1; Y_{1\theta}) \geq I(X, X_2; Y_{1\theta}).$$

As a consequence of this, if an error occurs with DF scheme while the relay can decode the message, then the error will happen anyway with CF scheme. So it is still better to use DF scheme, meaning that for any $r < I(X; Z_{1\theta_r}|X_1)$, it is clearly optimal to select DF scheme. Indeed, to have the knowledge of θ_r at the relay is enough to decide about the coding strategy.

B.2 Proof of Theorem 37

Fix P, $\mathcal{V}, \mathcal{T}$ such that they maximize the right hand side of (7.22). Then similarly assume a set $\mathcal{M}_n$ of size 2^{nR} of message indices W to be transmitted. Transmission is done in $B+L$ blocks, each of them of length n, and decoding at the destination is done backwardly. At the last $L-1$ blocks, the last compression index is first decoded and then all compression indices and transmitted messages are jointly decoded. By $\underline{x}_{\mathcal{S}}$ we denote $(\underline{x}_i)_{i\in\mathcal{S}}$.

Code generation:

(i) Randomly and independently generate 2^{nR} sequences $\underline{x}_{\mathcal{V}^c}$ drawn i.i.d. from

$$P^n_{X_{\mathcal{V}^c}}(\underline{x}_{\mathcal{V}^c}) = \prod_{j=1}^{n} P_{X_{\mathcal{V}^c}}(x_{\mathcal{V}^c j}).$$

Index them as $\underline{x}_{\mathcal{V}^c}(r)$ with index $r \in \left[1, 2^{nR}\right]$. This step will provide $|\mathcal{V}^c|$ different codebooks $\left(\underline{x}_k(r), r \in [1, 2^{nR}]\right)$ for each $k \in \mathcal{V}^c$, each with 2^{nR} codes. However the codes in each codebook corresponding to an index are jointly generated based on $P^n_{X_{\mathcal{V}^c}}$ and are not in general independent.

(ii) For each $\underline{x}_{\mathcal{V}^c}(r)$, randomly and conditionally independently generate 2^{nR} sequences $\underline{x}$ drawn i.i.d. from

$$P^n_{X|X_{\mathcal{V}^c}}(\underline{x}|\underline{x}_{\mathcal{V}^c}(r)) = \prod_{j=1}^{n} P_{X|X_{\mathcal{V}^c}}(x_j|x_{\mathcal{V}^c j}).$$

Index them as $\underline{x}(r,w)$, where $w \in \left[1, 2^{nR}\right]$.

(iii) For each $k \in \mathcal{T}$, randomly and independently generate $2^{n\hat{R}_k}$ sequences $\underline{x}_k$ drawn i.i.d. from

$$P^n_{X_k}(\underline{x}_k) = \prod_{j=1}^{n} P_{X_k}(x_{kj}).$$

Index them as $\underline{x}_k(r_k)$, where $r_k \in \left[1, 2^{n\hat{R}_k}\right]$ for $\hat{R}_k = I(Z_k; \hat{Z}_k|X_k) + \epsilon$.

(iv) For each $k \in \mathcal{T}$ and each $\underline{x}_k(r_k)$, randomly and conditionally independently generate $2^{n\hat{R}_k}$ sequences $\underline{\hat{z}}_k$ each with probability

$$P^n_{\hat{Z}_k|X_k}(\underline{\hat{z}}_k|\underline{x}_k(r_k)) = \prod_{j=1}^{n} P_{\hat{Z}_k|X_k}(\hat{z}_{kj}|x_{kj}(r_k)).$$

Index them as $\underline{\hat{z}}_k(r_k, \hat{s}_k)$, where $\hat{s}_k \in \left[1, 2^{n\hat{R}_k}\right]$.

(v) Provide the corresponding codebooks to the relays, the encoder and the decoder ends.

Encoding part:

(i) In every block $i = [1 : B]$, the source sends w_i based on $\underline{x}\big(w_{(i-1)}, w_i\big)$. Moreover, for blocks $i = [B+1 : B+L]$, the source sends the dummy message $w_i = 1$ known to all users.

(ii) For every block $i = [1 : B+L]$, and each $k \in \mathcal{V}^c$, the relay k knows $w_{(i-1)}$ by assumption and $w_0 = 1$, so it sends $\underline{x}_k\big(w_{(i-1)}\big)$.

(iii) For each $i = [1 : B+1]$, each $k \in \mathcal{T}$, the relay k after receiving $\underline{z}_k(i)$, searches for at least one index l_{ki} with $l_{k0} = 1$ such that

$$\big(\underline{x}_k(l_{k(i-1)}), \underline{z}_k(i), \underline{\hat{z}}_k(l_{k(i-1)}, l_{ki})\big) \in \mathcal{A}_\epsilon^n[X_k Z_k \hat{Z}_k].$$

The probability of finding such l_{ki} goes to one as n goes to infinity due to the choice of $\hat{R}_k$.

(iv) For $i = [1 : B+1]$ and $k \in \mathcal{T}$, relay k knows from the previous block $l_{k(i-1)}$ and it sends $\underline{x}_k(l_{k(i-1)})$. Moreover, relay k repeats $l_{k(B+1)}$ for $i = [B+2 : B+L]$.

Decoding part:

(i) After the transmission of the block $i = [1 : B]$ and for each $k \in \mathcal{V}^c$, the relay k decodes the message of block i with the assumption that all messages up to block $i-1$ have been correctly decoded. Since the relay k knows the message $w_{(i-1)}$ and so $\underline{x}_k\big(w_{(i-1)}\big)$, it also knows all the other $\underline{x}_{k'}\big(w_{(i-1)}\big)$ for $k' \in \mathcal{V}^c$ given the code generation. Relay k searches for the unique index $\hat{w}_i \in \mathcal{M}_n$ such that:

$$\big(\underline{x}\big(w_{(i-1)}, \hat{w}_i\big), \underline{x}_{\mathcal{V}^c}\big(w_{(i-1)}\big), \underline{z}_k(i)\big) \in \mathcal{A}_\epsilon^n[X X_{\mathcal{V}^c} Z_1].$$

By following similar argument in [1], the probability of error goes to zero as n goes to infinity provided that:

$$R < I(X; Z_k | X_{\mathcal{V}^c}). \tag{B.19}$$

(ii) The decoding in the destination is done backwardly. It first decodes the last compression indices sent by the relays in $\mathcal{T}$. The destination waits until the last block and then it jointly searches for the unique indices $\Big(\hat{l}_{k(B+1)}\Big)_{k\in\mathcal{T}}$ such that for all $b \in [B+2 : B+L]$ the following condition holds:

$$\bigg(\Big(\underline{x}_k(\hat{l}_{k(B+1)})\Big)_{k\in\mathcal{T}}, \underline{x}(1,1), \underline{x}_{\mathcal{V}^c}(1), \underline{y}_1(b)\bigg) \in \mathcal{A}_\epsilon^n[X X_{\mathcal{T}} X_{\mathcal{V}^c} Y_1].$$

Define the following events that can cause error in the previous decoding step.

$$\mathcal{E}_0 = \left\{ \left(\left(\underline{x}_k(l_{k(B+1)})\right)_{k\in\mathcal{T}}, \underline{x}(1,1), \underline{x}_{\mathcal{V}^c}(1), \underline{y}_1(b) \right) \notin \mathcal{A}_\epsilon^n[XX_{\mathcal{T}}X_{\mathcal{V}^c}Y_1] \right\},$$

$$\mathcal{E}_{\mathcal{S}} = \Big\{ \left(\left(\underline{x}_k(\hat{l}_{k(B+1)})\right)_{k\in\mathcal{S}}, \left(\underline{x}_k(l_{k(B+1)})\right)_{k\in\mathcal{S}^c}, \underline{x}(1,1), \underline{x}_{\mathcal{V}^c}(1), \underline{y}_1(b) \right) \in \mathcal{A}_\epsilon^n[XX_{\mathcal{T}\cup\mathcal{V}^c}Y_1]$$
$$\text{for some } \hat{l}_{k(B+1)} \neq l_{k(B+1)}, \text{ and all } b \in [B+2 : B+L] \Big\}.$$

The last event is the event that there be a jointly typical sequences with correct indices for the relays in $\mathcal{S}^c = \mathcal{T} - \mathcal{S}$ and wrong indices for the relays in $\mathcal{S}$. The probability of error is bounded as follows

$$\Pr\left(\hat{l}_{B+1} \neq l_{B+1}\right) \leq \Pr(\mathcal{E}_0) + \sum_{\mathcal{S}\subseteq\mathcal{T}} \Pr(\mathcal{E}_{\mathcal{S}}).$$

The first probability on the right side goes to zero as $n \to \infty$. $\Pr(\mathcal{E}_{\mathcal{S}})$ can be bounded as follows:

$$\Pr(\mathcal{E}_{\mathcal{S}}) \leq \sum_{\hat{l}_{k(B+1)} \neq l_{k(B+1)}, k\in\mathcal{S}} \Pr\Big[\bigcap_{b=[B+2:B+L]} \left(\left(\underline{x}_k(\hat{l}_{k(B+1)})\right)_{k\in\mathcal{S}}, \left(\underline{x}_k(l_{k(B+1)})\right)_{k\in\mathcal{S}^c}, \underline{x}(1,1), \underline{x}_{\mathcal{V}^c}(1), \underline{y}_1(b) \right) \in \mathcal{A}_\epsilon^n[XX_{\mathcal{T}}X_{\mathcal{V}^c}Y_1] \Big]$$
$$\leq \prod_{k\in\mathcal{S}} \left(2^{n\hat{R}_k} - 1\right) \left[2^{-n(I(X_{\mathcal{S}};XX_{\mathcal{S}^c\cup\mathcal{V}^c}Y_1)-\epsilon_1)}\right]^{L-1}.$$

Now this probability goes to zero as n goes to infinity provided that for all $\mathcal{S} \subseteq \mathcal{T}$:

$$\sum_{k\in\mathcal{S}} I(\hat{Z}_k; Z_k|X_k) + \epsilon_2 \leq (L-1) I(X_{\mathcal{S}}; XX_{\mathcal{S}^c\cup\mathcal{V}^c}Y_1). \tag{B.20}$$

(iii) After finding correctly $l_{k(B+1)}$ for all $k \in \mathcal{T}$ and since $w_{(B+1)} = 1$, the destination decodes jointly the message and all the compression indices $(w_b, l_{\mathcal{T}b})$ for each $b = [1 : B]$ where $l_{\mathcal{T}b} = (l_{kb})_{k\in\mathcal{T}}$. The decoding is performed backwardly with the assumption that $(w_{b+1}, l_{\mathcal{T}(b+1)})$ have been correctly decoded. The destination finds the unique pair of indices $(\hat{w}_b, \hat{l}_{\mathcal{T}b})$ such that

$$\left(\underline{x}(\hat{w}_b, w_{(b+1)}), \underline{x}_{\mathcal{V}^c}(\hat{w}_b), \underline{y}_1(b+1), \left(\underline{x}_k(\hat{l}_{kb}), \hat{\underline{z}}_k(\hat{l}_{kb}, l_{k(b+1)})\right)_{k\in\mathcal{T}}\right) \in \mathcal{A}_\epsilon^n[XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{T}}Y_1].$$

Consider the following error events associated with this step ($\mathcal{S} \subseteq \mathcal{T}, \mathcal{S}^c = \mathcal{T} - \mathcal{S}$):

$$\begin{aligned}
\mathcal{E}_0 =&\Big\{ \Big(\underline{x}(w_b, w_{(b+1)}), \underline{x}_{\mathcal{V}^c}(w_b), \underline{y}_1(b+1), \big(\underline{x}_k(l_{kb}), \underline{\hat{z}}_k(l_{kb}, l_{k(b+1)})\big)_{k\in\mathcal{T}}\Big) \\
&\qquad \notin \mathcal{A}_\epsilon^n[XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{T}}Y_1]\Big\}, \\
\mathcal{E}_{\mathcal{S}} =&\Big\{\Big(\underline{x}(w_b, w_{(b+1)}), \underline{x}_{\mathcal{V}^c}(w_b), \underline{y}_1(b+1), \Big(\underline{x}_k(\hat{l}_{kb}), \underline{\hat{z}}_k(\hat{l}_{kb}, l_{k(b+1)})\Big)_{k\in\mathcal{S}}, \\
&\qquad \big(\underline{x}_k(l_{kb}), \underline{\hat{z}}_k(l_{kb}, l_{k(b+1)})\big)_{k\in\mathcal{S}^c}\Big) \\
&\qquad \in \mathcal{A}_\epsilon^n[XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{T}}Y_1] \text{ for some } \hat{l}_{kb} \neq l_{kb}, \text{ and } k \in \mathcal{S}\Big\}, \\
\mathcal{E}_{w,\mathcal{S}} =&\Big\{\Big(\underline{x}(\hat{w}_b, w_{(b+1)}), \underline{x}_{\mathcal{V}^c}(\hat{w}_b), \underline{y}_1(b+1), \Big(\underline{x}_k(\hat{l}_{kb}), \underline{\hat{z}}_k(\hat{l}_{kb}, l_{k(b+1)})\Big)_{k\in\mathcal{S}}, \\
&\qquad \big(\underline{x}_k(l_{kb}), \underline{\hat{z}}_k(l_{kb}, l_{k(b+1)})\big)_{k\in\mathcal{S}^c}\Big) \\
&\qquad \in \mathcal{A}_\epsilon^n[XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{T}}Y_1] \text{ for some } \hat{w} \neq w_b, \hat{l}_{kb} \neq l_{kb}, \text{ and } k \in \mathcal{S}\Big\}.
\end{aligned}$$

In fact $\mathcal{E}_{\mathcal{S}}$ is the event that there is jointly typical codes with correct message index but wrong compression indicies for the relays in $\mathcal{S}$. On the other hand $\mathcal{E}_{w,\mathcal{S}}$ is the event that there is jointly typical codes with wrong message index and wrong compression indicies for the relays in $\mathcal{S}$. The probability of error of this step is thus bounded by

$$\Pr\Big((\hat{w}_b, \hat{l}_{\mathcal{T}b}) \neq (w_b, l_{\mathcal{T}b})\Big) \leq \Pr(\mathcal{E}_0) + \sum_{\mathcal{S}\subseteq\mathcal{T}} \left(\Pr(\mathcal{E}_{\mathcal{S}}) + \Pr(\mathcal{E}_{w,\mathcal{S}})\right).$$

First probability on the right side goes to zero as $n \to \infty$. $\Pr(\mathcal{E}_{\mathcal{S}})$ goes to zero as $n \to \infty$ provided that

$$\sum_{k\in\mathcal{S}} I(\hat{Z}_k; Z_k|X_k) + \epsilon_3 < \sum_{k\in\mathcal{S}} H(\hat{Z}_k|X_k) + H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}\hat{Z}_{\mathcal{S}^c}Y_1) - H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}\hat{Z}_{\mathcal{T}}Y_1|X_{\mathcal{S}}),$$

which can be written as:

$$\epsilon_3 < \sum_{k\in\mathcal{S}} H(\hat{Z}_k|Z_kX_k) + H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}\hat{Z}_{\mathcal{S}^c}Y_1) - H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}\hat{Z}_{\mathcal{T}}Y_1|X_{\mathcal{S}}).$$

The preceding inequality can be simplified using the fact that $\hat{Z}_k$ is independent of other random variables given (X_k, Z_k):

$$
\begin{aligned}
&\sum_{k\in\mathcal{S}} H(\hat{Z}_k|Z_kX_k) + H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}\hat{Z}_{\mathcal{S}^c}Y_1) - H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}\hat{Z}_{\mathcal{T}}Y_1|X_{\mathcal{S}})\\
=&\sum_{k\in\mathcal{S}} H(\hat{Z}_k|Z_{\mathcal{S}}X_{\mathcal{S}}) + H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}\hat{Z}_{\mathcal{S}^c}Y_1) - H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}\hat{Z}_{\mathcal{T}}Y_1|X_{\mathcal{S}})\\
=&\sum_{j=1}^{|\mathcal{S}|} H(\hat{Z}_{o(j)}|Z_{\mathcal{S}}X_{\mathcal{S}}) + H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}\hat{Z}_{\mathcal{S}^c}Y_1) - H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}\hat{Z}_{\mathcal{T}}Y_1|X_{\mathcal{S}})\\
=&\sum_{j=1}^{|\mathcal{S}|} H(\hat{Z}_{o(j)}|\hat{Z}_{o(1)}...\hat{Z}_{o(j-1)}Z_{\mathcal{S}}X_{\mathcal{S}}) + H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}\hat{Z}_{\mathcal{S}^c}Y_1) - H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}\hat{Z}_{\mathcal{T}}Y_1|X_{\mathcal{S}})\\
=&H(\hat{Z}_{\mathcal{S}}|Z_{\mathcal{S}}X_{\mathcal{S}}) + H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}\hat{Z}_{\mathcal{S}^c}Y_1) - H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}\hat{Z}_{\mathcal{T}}Y_1|X_{\mathcal{S}}),
\end{aligned}
$$

where $o : [1, |\mathcal{S}|] \to \mathcal{S}$ is an arbitrary ordering over $\mathcal{S}$. This manipulation gives us the following:

$$
\begin{aligned}
\epsilon_3 &< I(XX_{\mathcal{V}^c\cup\mathcal{S}^c}\hat{Z}_{\mathcal{S}^c}Y_1; X_{\mathcal{S}}) - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}}|XX_{\mathcal{V}^c\cup\mathcal{T}}\hat{Z}_{\mathcal{S}^c}Y_1)\\
&= I(\hat{Z}_{\mathcal{S}^c}Y_1; X_{\mathcal{S}}|XX_{\mathcal{V}^c\cup\mathcal{S}^c}) - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}}|XX_{\mathcal{V}^c\cup\mathcal{T}}\hat{Z}_{\mathcal{S}^c}Y_1).
\end{aligned}
\tag{B.21}
$$

Given the fact that $\mathcal{T} \in \Upsilon(\mathcal{V})$, the last inequality holds for each $\mathcal{S} \subseteq \mathcal{T}$.

As the next step, we bound the probability $\Pr(\mathcal{E}_{w,\mathcal{S}})$ as follows:

$$
\begin{aligned}
&\Pr(\mathcal{E}_{w,\mathcal{S}}) \leq\\
&\sum_{\hat{l}_{kb}\neq l_{kb},\hat{w}\neq w_b} \Pr\Big[\Big(\underline{x}(\hat{w}_b, w_{(b+1)}), \underline{x}_{\mathcal{V}^c}(\hat{w}_b), \underline{y}_1(b+1), \Big(\underline{x}_k(\hat{l}_{kb}), \underline{\hat{z}}_k(\hat{l}_{kb}, l_{k(b+1)})\Big)_{k\in\mathcal{S}},\\
&\qquad \big(\underline{x}_k(l_{kb}), \underline{\hat{z}}_k(l_{kb}, l_{k(b+1)})\big)_{k\in\mathcal{S}^c}\Big) \in \mathcal{A}_\epsilon^n[XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{T}}Y_1]\Big]\\
&\leq \left(2^{nR} - 1\right)\prod_{k\in\mathcal{S}}\left(2^{n\hat{R}_k} - 1\right) \times\\
&\left(2^{n(H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}Y_1\hat{Z}_{\mathcal{T}}|X_{\mathcal{S}}) - H(XX_{\mathcal{V}^c}) - H(Y_1\hat{Z}_{\mathcal{S}^c}X_{\mathcal{S}^c}) - \sum_{k\in\mathcal{S}} H(\hat{Z}_k|X_k) + \epsilon_4)}\right).
\end{aligned}
$$

Now given the following inequality the last probability also tends to zero as $n \to \infty$

$$
\begin{aligned}
R + \sum_{k\in\mathcal{S}} I(\hat{Z}_k; Z_k|X_k)) + \epsilon_5 \leq -H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}Y_1\hat{Z}_{\mathcal{T}}|X_{\mathcal{S}}) + H(XX_{\mathcal{V}^c})\\
+ H(Y_1\hat{Z}_{\mathcal{S}^c}X_{\mathcal{S}^c}) + \sum_{k\in\mathcal{S}} H(\hat{Z}_k|X_k).
\end{aligned}
\tag{B.22}
$$

The inequality (B.22) is then simplified as follows

$$
\begin{aligned}
R+\epsilon_5 &\le -H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}Y_1\hat{Z}_{\mathcal{T}}|X_{\mathcal{S}}) + H(XX_{\mathcal{V}^c}) + H(Y_1\hat{Z}_{\mathcal{S}^c}X_{\mathcal{S}^c}) + \sum_{k\in\mathcal{S}} H(\hat{Z}_k|Z_kX_k)\\
&\le -H(XX_{\mathcal{V}^c\cup\mathcal{S}^c}Y_1\hat{Z}_{\mathcal{T}}|X_{\mathcal{S}}) + H(XX_{\mathcal{V}^c}) + H(Y_1\hat{Z}_{\mathcal{S}^c}X_{\mathcal{S}^c}) + H(\hat{Z}_{\mathcal{S}}|Z_{\mathcal{S}}X_{\mathcal{S}})\\
&\le I(XX_{\mathcal{V}^c}X_{\mathcal{S}};Y_1\hat{Z}_{\mathcal{S}^c}|X_{\mathcal{S}^c}) - H(\hat{Z}_{\mathcal{S}}|XX_{\mathcal{V}^c\cup\mathcal{T}}Y_1\hat{Z}_{\mathcal{S}^c}) + H(\hat{Z}_{\mathcal{S}}|Z_{\mathcal{S}}X_{\mathcal{S}})\\
&\le I(XX_{\mathcal{V}^c}X_{\mathcal{S}};Y_1\hat{Z}_{\mathcal{S}^c}|X_{\mathcal{S}^c}) - I(\hat{Z}_{\mathcal{S}};Z_{\mathcal{S}}|XX_{\mathcal{V}^c\cup\mathcal{T}}Y_1\hat{Z}_{\mathcal{S}^c}) \qquad \text{(B.23)}
\end{aligned}
$$

By choosing finite L but large enough, inequalities (B.19) and (B.23) prove Theorem 37, where the rate is achieved by letting (B, n) tend to infinity. At the end time sharing random variable Q can be added.

B.3 Proof of Theorem 38

The coding for the Cooperative Mixed Noisy Network Coding (CMNNC) is done with a difference regarding the previous theorems. To see the difference look at the table B.2. In this table, we present the coding scheme for CMNNC- two relay networks. The relay 1 uses DF to help the source so it has to decode the source messages successively and not backwardly. But on the other hand the relay 1 wants to exploit the help of the relay 2 to decode the source message. So it does not start decoding until it retrieves the compression index. To this purpose the relay 1 uses offset decoding which means that it waits two blocks instead of one to decode the source message and the compression index. In block $b = 2$, the relay 1 decodes l_1 and w_1. Equally the source code at the block $b+2$ is correlated with the relay 1 code from the block b and not the block $b + 1$. The price we pay here is one block delay. The source has to wait until $b = B + 2$ to start Backward decoding. The compression index l_{B+2} is repeated until the block $B + L$.

The proof for the multiple relay networks follows the same idea. Fix P, $\mathcal{V}, \mathcal{T}$ and $\mathcal{T}_k$'s such that they maximize the right hand side of (7.26). Note that $\mathcal{T}, \mathcal{T}_k \subseteq \mathcal{V}$. Then again assume a set $\mathcal{M}_n$ of size 2^{nR} of message indices W to be transmitted, again in $B + L$ blocks, each of them of length n. At the last $L - 2$ blocks, the last compression index is first decoded and then all compression indices and transmitted messages are jointly decoded. The relays in $\mathcal{V}^c$ starts to decode after the block 2.

Code generation:

(i) The code generation for the sources and the relays in $\mathcal{V}^c$ remains the same as the appendix B.2. Generate them as before like $(\underline{x}_{\mathcal{V}^c}(r), \underline{x}(r, w))$ and provide them to all

Table B.2: Coding for CMNNC

$b=1$	$b=2$	$b=3$	...	$b=B+2$	$b=B+3$	...	$b=B+L$
$\underline{x}_1(1)$	$\underline{x}_1(1)$	$\underline{x}_1(w_1)$	...	$\underline{x}_1(w_B)$	$\underline{x}_1(1)$	...	$\underline{x}_1(1)$
$\underline{x}(1,w_1)$	$\underline{x}(1,w_2)$	$\underline{x}(w_1,w_3)$	...	$\underline{x}(w_B,1)$	$\underline{x}(1,1)$	...	$\underline{x}(1,1)$
$\underline{x}_2(1)$	$\underline{x}_2(l_1)$	$\underline{x}_2(l_2)$	...	$\underline{x}_2(l_{(B+1)})$	$\underline{x}_2(l_{(B+2)})$	...	$\underline{x}_2(l_{(B+2)})$
$\underline{\hat{z}}_2(1,l_1)$	$\underline{\hat{z}}_2(l_1,l_2)$	$\underline{\hat{z}}_2(l_2,l_3)$	...	$\underline{\hat{z}}_2(l_{(B+1)},l_{(B+2)})$	.	.	.
$\underline{z}_1(1)$	$\underline{z}_1(2)$	$\underline{z}_1(3)$	...	$\underline{z}_1(B+2)$	$\underline{z}_1(B+3)$	...	$\underline{z}_1(B+L)$
$\underline{z}_2(1)$	$\underline{z}_2(2)$	$\underline{z}_2(3)$	...	$\underline{z}_2(B+2)$	$\underline{z}_2(B+3)$	...	$\underline{z}_2(B+L)$
$\underline{y}_1(1)$	$\underline{y}_1(2)$	$\underline{y}_1(3)$	...	$\underline{y}_1(B+2)$	$\underline{y}_1(B+3)$	...	$\underline{y}_1(B+L)$

users.

(ii) For each $k \in \mathcal{T} \cup (\cup_{k\in\mathcal{V}^c}\mathcal{T}_k)$, randomly and independently generate $2^{n\hat{R}_k}$ sequences $\underline{x}_k$ drawn i.i.d. from

$$P^n_{X_k}(\underline{x}_k) = \prod_{j=1}^{n} P_{X_k}(x_{kj}).$$

Index them as $\underline{x}_k(r_k)$, where $r_k \in \left[1, 2^{n\hat{R}_k}\right]$ for $\hat{R}_k = I(Z_k; \hat{Z}_k|X_k) + \epsilon$.

(iii) For each $k \in \mathcal{T} \cup (\cup_{k\in\mathcal{V}^c}\mathcal{T}_k)$ and each $\underline{x}_k(r_k)$, randomly and conditionally independently generate $2^{n\hat{R}_k}$ sequences $\underline{\hat{z}}_k$ each with probability

$$P^n_{\hat{Z}_k|X_k}(\underline{\hat{z}}_k|\underline{x}_k(r_k)) = \prod_{j=1}^{n} P_{\hat{Z}_k|X_k}(\hat{z}_{kj}|x_{kj}(r_k)).$$

Index them as $\underline{\hat{z}}_k(r_k, \hat{s}_k)$, where $\hat{s}_k \in \left[1, 2^{n\hat{R}_k}\right]$.

(iv) Provide the corresponding codebooks to the relays, the encoder and the decoder ends.

Encoding part:

(i) In every block $i = [1 : B]$, the source sends w_i using $\underline{x}\big(w_{(i-2)}, w_i\big)$ ($w_0 = w_{-1} = 1$). Moreover, for blocks $i = [B+1 : B+L]$, the source sends the dummy message $w_i = 1$ known to all users.

(ii) For every block $i = [1 : B + L]$, and each $k \in \mathcal{V}^c$, the relay k knows $w_{(i-2)}$ by assumption and $w_0 = w_{-1} = 1$, so it sends $\underline{x}_k\big(w_{(i-2)}\big)$.

(iii) For each $i = [1 : B + 2]$, each $k \in \mathcal{T} \cup (\cup_{k\in\mathcal{V}^c}\mathcal{T}_k)$, the relay k after receiving $\underline{z}_k(i)$, searches for at least one index l_{ki} with $l_{k0} = 1$ such that

$$\big(\underline{x}_k(l_{k(i-1)}), \underline{z}_k(i), \underline{\hat{z}}_k(l_{k(i-1)}, l_{ki})\big) \in \mathcal{A}^n_\epsilon[X_k Z_k \hat{Z}_k].$$

The probability of finding such l_{ki} goes to one as n goes to infinity due to the choice of $\hat{R}_k$.

(iv) For $i = [1 : B{+}2]$ and $k \in \mathcal{T} \cup (\cup_{k\in\mathcal{V}^c}\mathcal{T}_k)$, relay k knows from the previous block $l_{k(i-1)}$ and it sends $\underline{x}_k(l_{k(i-1)})$. Moreover, relay k repeats $l_{k(B+2)}$ for $i = [B+3 : B+L]$, which means for $L-2$ blocks.

Decoding part:

(i) After the transmission of the block $i = [1 : B+1]$ and for each $k \in \mathcal{V}^c$, the relay k decodes the message w_i and the compression index $l_{\mathcal{T}_k i}$ with the assumption that all messages and compression indices up to block $i-1$ have been correctly decoded. Since the relay k knows the message $w_{(i-2)}, w_{(i-1)}$ and so $\underline{x}_k\big(w_{(i-2)}\big)$ and $\underline{x}_k\big(w_{(i-1)}\big)$, it also knows all the other $\underline{x}'_k\big(w_{(i-2)}\big)$ and $\underline{x}_{k'}\big(w_{(i-1)}\big)$ for $k' \in \mathcal{V}^c$ given the code generation. Relay k searches for the unique index $(\hat{w}_b, \hat{l}_{\mathcal{T}_k b})$ by looking at two consecutive blocks b and $b+1$ such that:

$$\Big(\underline{x}(w_{(b-2)}, \hat{w}_b), \underline{x}_{\mathcal{V}^c}(w_{(b-2)}), \underline{z}_k(b), \Big(\underline{x}_k(l_{k(b-1)}), \underline{\hat{z}}_k(l_{k(b-1)}, \hat{l}_{kb})\Big)_{k\in\mathcal{T}_k}\Big)$$
$$\in \mathcal{A}_\epsilon^n[XX_{\mathcal{T}_k\cup\mathcal{V}^c}\hat{Z}_{\mathcal{T}_k}Z_k] \text{ and } \Big(\underline{x}_{\mathcal{V}^c}(w_{(b-1)}), \underline{z}_k(b+1), \Big(\underline{x}_k(\hat{l}_{kb})\Big)_{k\in\mathcal{T}_k}\Big) \in \mathcal{A}_\epsilon^n[X_{\mathcal{T}_k\cup\mathcal{V}^c}Z_k].$$

Define the following events ($\mathcal{S} \subseteq \mathcal{T}, \mathcal{S}^c = \mathcal{T}_k - \mathcal{S}$):

$$\mathcal{E}_0 = \Big\{ \Big(\underline{x}(w_{(b-2)}, w_b), \underline{x}_{\mathcal{V}^c}(w_{(b-2)}), \underline{z}_k(b), \big(\underline{x}_k(l_{k(b-1)}), \underline{\hat{z}}_k(l_{k(b-1)}, l_{kb})\big)_{k\in\mathcal{T}_k}\Big) \notin$$
$$\mathcal{A}_\epsilon^n[XX_{\mathcal{T}_k\cup\mathcal{V}^c}\hat{Z}_{\mathcal{T}_k}Z_k] \text{ or } \Big(\underline{x}_{\mathcal{V}^c}(w_{(b-1)}), \underline{z}_k(b+1), (\underline{x}_k(l_{kb}))_{k\in\mathcal{T}_k}\Big) \notin \mathcal{A}_\epsilon^n[X_{\mathcal{T}_k\cup\mathcal{V}^c}Z_k].\Big\},$$

$$\mathcal{E}_{\mathcal{S}} = \Big\{\Big(\underline{x}(w_{(b-2)}, w_b), \underline{x}_{\mathcal{V}^c}(w_{(b-2)}), \underline{z}_k(b), \Big(\underline{x}_k(l_{k(b-1)}), \underline{\hat{z}}_k(l_{k(b-1)}, \hat{l}_{kb})\Big)_{k\in\mathcal{S}},$$
$$\big(\underline{x}_k(l_{k(b-1)}), \underline{\hat{z}}_k(l_{k(b-1)}, l_{kb})\big)_{k\in\mathcal{S}^c}\Big) \in \mathcal{A}_\epsilon^n[XX_{\mathcal{T}_k\cup\mathcal{V}^c}\hat{Z}_{\mathcal{T}_k}Z_k] \text{ and}$$
$$\Big(\underline{x}_{\mathcal{V}^c}(w_{(b-1)}), \underline{z}_k(b+1), \Big(\underline{x}_k(\hat{l}_{kb})\Big)_{k\in\mathcal{S}}, (\underline{x}_k(l_{kb}))_{k\in\mathcal{S}^c}\Big) \in \mathcal{A}_\epsilon^n[X_{\mathcal{T}_k\cup\mathcal{V}^c}Z_k]$$
$$\text{for some } \hat{l}_{kb} \neq l_{kb}, \text{ and } k \in \mathcal{S}\Big\},$$

$$\mathcal{E}_{w,\mathcal{S}} = \Big\{\Big(\underline{x}(w_{(b-2)}, \hat{w}_b), \underline{x}_{\mathcal{V}^c}(w_{(b-2)}), \underline{z}_k(b), \Big(\underline{x}_k(l_{k(b-1)}), \underline{\hat{z}}_k(l_{k(b-1)}, \hat{l}_{kb})\Big)_{k\in\mathcal{S}},$$
$$\big(\underline{x}_k(l_{k(b-1)}), \underline{\hat{z}}_k(l_{k(b-1)}, l_{kb})\big)_{k\in\mathcal{S}^c}\Big) \in \mathcal{A}_\epsilon^n[XX_{\mathcal{T}_k\cup\mathcal{V}^c}\hat{Z}_{\mathcal{T}_k}Z_k] \text{ and}$$
$$\Big(\underline{x}_{\mathcal{V}^c}(w_{(b-1)}), \underline{z}_k(b+1), \Big(\underline{x}_k(\hat{l}_{kb})\Big)_{k\in\mathcal{S}}, (\underline{x}_k(l_{kb}))_{k\in\mathcal{S}^c}\Big) \in \mathcal{A}_\epsilon^n[X_{\mathcal{T}_k\cup\mathcal{V}^c}Z_k]$$
$$\text{for some } \hat{w} \neq w_b, \hat{l}_{kb} \neq l_{kb}, \text{ and } k \in \mathcal{S}\Big\}.$$

Then the probability of error is bounded as follows:

$$\Pr\left((\hat{w}_b, \hat{l}_{\mathcal{T}_k b}) \neq (w_b, l_{\mathcal{T}_k b})\right) \leq \Pr(\mathcal{E}_0) + \sum_{\mathcal{S} \subseteq \mathcal{T}} \left(\Pr(\mathcal{E}_{\mathcal{S}}) + \Pr(\mathcal{E}_{w,\mathcal{S}})\right).$$

$\Pr(\mathcal{E}_0)$ goes to zero as n goes to infinity due to the code generation and the encoding process. $\Pr(\mathcal{E}_{\mathcal{S}})$ goes to zero as $n \to \infty$ provided that

$$\begin{aligned}\sum_{k \in \mathcal{S}} I(\hat{Z}_k; Z_k | X_k) + \epsilon_3 < \sum_{k \in \mathcal{S}} H(\hat{Z}_k | X_k) + \sum_{k \in \mathcal{S}} H(X_k) + H(X X_{\mathcal{V}^c \cup \mathcal{T}_k} \hat{Z}_{\mathcal{S}^c} Z_k) \\ + H(X_{\mathcal{V}^c \cup \mathcal{S}^c} Z_k) - H(X X_{\mathcal{V}^c \cup \mathcal{T}_k} \hat{Z}_{\mathcal{T}_k} Z_k) - H(X_{\mathcal{V}^c \cup \mathcal{T}_k} Z_k).\end{aligned}$$

This can be written as:

$$\epsilon_3 < \sum_{k \in \mathcal{S}} H(\hat{Z}_k | Z_k X_k) + I(X_{\mathcal{S}}; X_{\mathcal{V}^c \cup \mathcal{S}^c} Z_k) - H(\hat{Z}_{\mathcal{S}} | X X_{\mathcal{V}^c \cup \mathcal{T}_k} \hat{Z}_{\mathcal{S}^c} Z_k).$$

The preceding inequality can be simplified using the fact that $\hat{Z}_k$ is independent of other random variables given (X_k, Z_k):

$$\begin{aligned}&\sum_{k \in \mathcal{S}} H(\hat{Z}_k | Z_k X_k) + I(X_{\mathcal{S}}; X_{\mathcal{V}^c \cup \mathcal{S}^c} Z_k) - H(\hat{Z}_{\mathcal{S}} | X X_{\mathcal{V}^c \cup \mathcal{T}_k} \hat{Z}_{\mathcal{S}^c} Z_k) \\ =& H(\hat{Z}_{\mathcal{S}} | Z_{\mathcal{S}} X_{\mathcal{S}}) + I(X_{\mathcal{S}}; X_{\mathcal{V}^c \cup \mathcal{S}^c} Z_k) - H(\hat{Z}_{\mathcal{S}} | X X_{\mathcal{V}^c \cup \mathcal{T}_k} \hat{Z}_{\mathcal{S}^c} Z_k),\end{aligned}$$

which gives us the following:

$$\epsilon_3 < I(Z_k; X_{\mathcal{S}} | X_{\mathcal{V}^c \cup \mathcal{S}^c}) - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | X X_{\mathcal{V}^c \cup \mathcal{T}_k} \hat{Z}_{\mathcal{S}^c} Z_k). \tag{B.24}$$

Given the fact that $\mathcal{T}_k \in \Upsilon_k(\mathcal{V})$, the last inequality holds for each $\mathcal{S} \subseteq \mathcal{T}_k$. Finally $\Pr(\mathcal{E}_{\mathcal{S}})$ goes to zero as $n \to \infty$ provided that

$$\begin{aligned}R + \sum_{k \in \mathcal{S}} I(\hat{Z}_k; Z_k | X_k) + \epsilon_5 < \sum_{k \in \mathcal{S}} H(\hat{Z}_k | X_k) + \sum_{k \in \mathcal{S}} H(X_k) + H(X X_{\mathcal{V}^c}) \\ + H(X_{\mathcal{T}_k} \hat{Z}_{\mathcal{S}^c} Z_k | X_{\mathcal{V}^c}) + H(X_{\mathcal{V}^c \cup \mathcal{S}^c} Z_k) \\ - H(X X_{\mathcal{V}^c \cup \mathcal{T}_k} \hat{Z}_{\mathcal{T}_k} Z_k) - H(X_{\mathcal{V}^c \cup \mathcal{T}_k} Z_k).\end{aligned}$$

This will give the following condition:

$$\begin{aligned}R + \epsilon_5 &< I(X; \hat{Z}_{\mathcal{S}^c} Z_k | X_{\mathcal{V}^c} X_{\mathcal{T}_k}) + I(X_{\mathcal{S}}; Z_k | X_{\mathcal{V}^c} X_{\mathcal{S}^c}) - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | X X_{\mathcal{V}^c \cup \mathcal{T}_k} \hat{Z}_{\mathcal{S}^c} Z_k) \\ &= I(X; \hat{Z}_{\mathcal{T}} Z_k | X_{\mathcal{V}^c} X_{\mathcal{T}_k}) + I(X_{\mathcal{S}}; Z_k | X_{\mathcal{V}^c} X_{\mathcal{S}^c}) - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | X_{\mathcal{V}^c \cup \mathcal{T}_k} \hat{Z}_{\mathcal{S}^c} Z_k).\end{aligned} \tag{B.25}$$

(ii) The decoding in the destination remains in general same than before. It jointly searches for the unique indices $\left(\hat{l}_{k(B+2)}\right)_{k\in\mathcal{T}}$ such that for all $b \in [B+3 : B+L]$ the following condition holds:

$$\left(\left(\underline{x}_k(\hat{l}_{k(B+1)})\right)_{k\in\mathcal{T}}, \underline{x}(1,1), \underline{x}_{\mathcal{V}^c}(1), \underline{y}_1(b)\right) \in \mathcal{A}_\epsilon^n[XX_{\mathcal{T}}X_{\mathcal{V}^c}Y_1].$$

The probability of error goes to zero as n goes to infinity provided that for all $\mathcal{S} \subseteq \mathcal{T}$:

$$\sum_{k\in\mathcal{S}} I(\hat{Z}_k; Z_k|X_k) + \epsilon_2 \leq (L-2)I(X_{\mathcal{S}}; XX_{\mathcal{S}^c\cup\mathcal{V}^c}Y_1). \tag{B.26}$$

(iii) After finding correctly $l_{k(B+2)}$ for all $k \in \mathcal{T}$ and since $w_{(B+1)} = 1$, the destination decodes jointly the message and all the compression indices $(w_b, l_{\mathcal{T}(b+1)})$ for each $b = [1 : B]$ where $l_{\mathcal{T}b} = (l_{kb})_{k\in\mathcal{T}}$. The decoding is performed backwardly with the assumption that $(w_{b+2}, l_{\mathcal{T}(b+2)})$ have been correctly decoded. The destination finds the unique pair of indices $(\hat{w}_b, \hat{l}_{\mathcal{T}(b+1)})$ such that

$$\left(\underline{x}(\hat{w}_b, w_{(b+2)}), \underline{x}_{\mathcal{V}^c}(\hat{w}_b), \underline{y}_1(b+2), \left(\underline{x}_k(\hat{l}_{k(b+1)}), \underline{\hat{z}}_k(\hat{l}_{k(b+1)}, l_{k(b+2)})\right)_{k\in\mathcal{T}}\right)$$
$$\in \mathcal{A}_\epsilon^n[XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{T}}Y_1].$$

Following the same procedure as the proof in the appendix B.2, the probability of error goes to zero as n goes to infinity given the following conditions:

$$\epsilon_3 < I(\hat{Z}_{\mathcal{S}^c}Y_1; X_{\mathcal{S}}|XX_{\mathcal{V}^c\cup\mathcal{S}^c}) - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}}|XX_{\mathcal{V}^c\cup\mathcal{T}}\hat{Z}_{\mathcal{S}^c}Y_1), \tag{B.27}$$

$$R + \epsilon_5 \leq I(XX_{\mathcal{V}^c}X_{\mathcal{S}}; Y_1\hat{Z}_{\mathcal{S}^c}|X_{\mathcal{S}^c}) - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}}|XX_{\mathcal{V}^c\cup\mathcal{T}}\hat{Z}_{\mathcal{S}^c}Y_1). \tag{B.28}$$

By choosing finite L but large enough, inequalities (B.25) and (B.28) prove Theorem 38, where the rate is achieved by letting (B, n) tend to infinity. At the end time sharing random variable Q can be added.

B.4 Proof of the theorem 39

The proof is in general similar to the corollary 36.

Consider the composite multiple relay channel with parameters $\theta = (\theta_d, \theta_r)$. Transmission is done over $B + L$ blocks like the theorem 37. Suppose that every relay knows $\mathcal{D}_{\mathcal{V}}$. i.e. decision of other relays for each θ_r. Suppose that the channel in function is with index $\theta = (\theta_r, \theta_d)$.

Code generation:

(i) The relay knows θ_r and for each θ_r, relay k generates two main codes $\left(\underline{x}_k^{(1)}, \underline{x}_k^{(2)}\right)$.

(a) Randomly and independently generate 2^{nR} sequences $\underline{x}_{\mathcal{N}}^{(1)}$ drawn i.i.d. from

$$P^n_{X_{\mathcal{N}}^{(1)}}(\underline{x}_{\mathcal{N}}^{(1)}) = \prod_{j=1}^{n} P_{X_{\mathcal{N}}^{(1)}}(x_{\mathcal{N}j}^{(1)}).$$

Index them as $\underline{x}_{\mathcal{N}}^{(1)}(r)$ with index $r \in \left[1, 2^{nR}\right]$. This code is also known to the source, so it cannot depend on θ_r.

(b) Randomly and independently generate $2^{n\hat{R}_k}$ sequences $\underline{x}_k^{(2)}$ drawn i.i.d. from

$$P^n_{X_k^{(2)}}(\underline{x}_k^{(2)}) = \prod_{j=1}^{n} P_{X_k^{(2)}}(x_{kj}^{(2)}).$$

Index them as $\underline{x}_k^{(2)}(r_k)$, where $r_k \in \left[1, 2^{n\hat{R}_k}\right]$ for $\hat{R}_k = I(Z_{k\theta_r}; \hat{Z}_k | X_k^{(2)}) + \epsilon$.

(c) For each $\underline{x}_k^{(2)}(r_k)$, randomly and conditionally independently generate $2^{n\hat{R}_k}$ sequences $\underline{\hat{z}}_k$ each with probability

$$P^n_{\hat{Z}_k|X_k^{(2)}}(\underline{\hat{z}}_k|\underline{x}_k^{(2)}(r_k)) = \prod_{j=1}^{n} P_{\hat{Z}_k|X_k^{(2)}}(\hat{z}_{kj}|x_{kj}^{(2)}(r_k)).$$

Index them as $\underline{\hat{z}}_k(r_k, \hat{s}_k)$, where $\hat{s}_k \in \left[1, 2^{n\hat{R}_k}\right]$. This code generation can depend on θ_r

(ii) For each $\underline{x}_{\mathcal{N}}^{(1)}(r)$, randomly and conditionally independently generate 2^{nR} sequences $\underline{x}$ drawn i.i.d. from

$$P^n_{X|X_{\mathcal{N}}^{(1)}}(\underline{x}|\underline{x}_{\mathcal{N}}^{(1)}(r)) = \prod_{j=1}^{n} P_{X|X_{\mathcal{N}}^{(1)}}(x_j|x_{\mathcal{N}j}^{(1)}).$$

Index them as $\underline{x}(r, w)$, where $w \in \left[1, 2^{nR}\right]$.

Encoding part:

(i) In every block $i = [1 : B]$, the source sends w_i based on $\underline{x}\big(w_{(i-1)}, w_i\big)$. Moreover, for blocks $i = [B+1 : B+L]$, the source sends the dummy message $w_i = 1$ known to all users.

(ii) The relay k knows θ_r. If $\theta_r \in \mathcal{D}_{\mathrm{DF}}^{(k)}$, for every block $i = [1 : B+L]$, the relay k knows $w_{(i-1)}$ by assumption and $w_0 = 1$, so it sends $\underline{x}_k^{(1)}\big(w_{(i-1)}\big)$.

If $\theta_r \notin \mathcal{D}_{\text{DF}}^{(k)}$, the relay k after receiving $\underline{z}_{k\theta_r}(i)$, searches for at least one index l_{ki} with $l_{k0} = 1$ such that

$$\big(\underline{x}_k^{(2)}(l_{k(i-1)}), \underline{z}_{k\theta_r}(i), \underline{\hat{z}}_k(l_{k(i-1)}, l_{ki})\big) \in \mathcal{A}_\epsilon^n[X_k Z_{k\theta_r} \hat{Z}_k].$$

The probability of finding such l_{ki} goes to one as n goes to infinity. The relay k knows from the previous block $l_{k(i-1)}$ and it sends $\underline{x}_k^{(2)}(l_{k(i-1)})$. Moreover, relay k repeats $l_{k(B+1)}$ for $i = [B+2 : B+L]$.

Decoding part:

1. After the transmission of the block $i = [1 : B]$, if $\theta_r \in \mathcal{D}_{\text{DF}}^{(k)}$ the relay k decodes the message of block i with the assumption that all messages up to block $i-1$ have been correctly decoded. The relay k searches for the unique index $\hat{w}_i \in \mathcal{M}_n$ such that:

$$\Big(\underline{x}\big(w_{(i-1)}, \hat{w}_i\big), \underline{x}_{\mathcal{N}}^{(1)}\big(w_{(i-1)}\big), \underline{z}_{k\theta_r}(i)\Big) \in \mathcal{A}_\epsilon^n[X X_{\mathcal{N}} Z_{k\theta_r}].$$

The outage event is occurred provided that:

$$r > I(X; Z_{k\theta_r} | X_{\mathcal{N}}^{(1)}). \tag{B.29}$$

We emphasize that not all $X_{\mathcal{N}}^{(1)}$ are in function, but only $X_{\mathcal{V}^c}^{(1)}$ for $\theta_r \in \mathcal{D}_{\mathcal{V}}$.

2. The decoding in the destination is done backwardly. It knows θ, $\mathcal{D}_{\mathcal{V}}$ and therefore $\mathcal{V}$, i.e. it knows each relay is using Decode-and-Forward or Compress-and-Forward. Moreover it chooses $\mathcal{T}$ to maximize (7.35). It first decodes the last compression indices sent by the relays in $\mathcal{T}$. It jointly searches for the unique indices $\Big(\hat{l}_{k(B+1)}\Big)_{k\in\mathcal{T}}$ such that for all $b \in [B+2 : B+L]$ the following condition holds:

$$\Big(\Big(\underline{x}_k(\hat{l}_{k(B+1)})\Big)_{k\in\mathcal{T}}, \underline{x}(1,1), \underline{x}_{\mathcal{V}^c}(1), \underline{y}_{1\theta}(b)\Big) \in \mathcal{A}_\epsilon^n[X X_{\mathcal{T}} X_{\mathcal{V}^c} Y_{1\theta}].$$

.

After finding correctly $l_{k(B+1)}$ for all $k \in \mathcal{T}$ and since $w_{(B+1)} = 1$, the destination decodes jointly the message and all the compression indices $(w_b, l_{\mathcal{T}b})$ for each $b = [1 : B]$ where $l_{\mathcal{T}b} = (l_{kb})_{k\in\mathcal{T}}$. The decoding is performed backwardly with the assumption that $(w_{b+1}, l_{\mathcal{T}(b+1)})$ have been correctly decoded. The destination finds the unique pair of indices $(\hat{w}_b, \hat{l}_{\mathcal{T}b})$ such that

$$\Big(\underline{x}(\hat{w}_b, w_{(b+1)}), \underline{x}_{\mathcal{V}^c}(\hat{w}_b), \underline{y}_{1\theta}(b+1), \Big(\underline{x}_k(\hat{l}_{kb}), \underline{\hat{z}}_k(\hat{l}_{kb}, l_{k(b+1)})\Big)_{k\in\mathcal{T}}\Big)$$
$$\in \mathcal{A}_\epsilon^n[X X_{\mathcal{T}\cup\mathcal{V}^c} \hat{Z}_{\mathcal{T}} Y_{1\theta}].$$

It can be seen from the theorem 37 that the error is occurred if:

$$r > \min_{\mathcal{S} \subseteq \mathcal{T}} R_{\mathcal{T}}(\mathcal{S}, \theta), \tag{B.30}$$

for

$$R_{\mathcal{T}}(\mathcal{S}, \theta) = I(XX^{(1)}_{\mathcal{V}^c} X^{(2)}_{\mathcal{S}}; \hat{Z}_{\mathcal{S}^c} Y_{1\theta} | X^{(2)}_{\mathcal{S}^c}) - I(Z_{\mathcal{S}\theta_r}; \hat{Z}_{\mathcal{S}} | XX^{(2)}_{\mathcal{T}} X^{(1)}_{\mathcal{N}} \hat{Z}_{\mathcal{S}^c} Y_{1\theta}),$$

Note that $\mathcal{T}$ is chosen in such a way that the right hand side is in its maximum value. From the discussion around (7.24) we know that this set belongs to $\Upsilon(\mathcal{V})$ and so $Q_{\mathcal{T}}(\mathcal{A}) \geq 0$ for each $\mathcal{A} \subseteq \mathcal{T}$.

3. Using (B.29) and (B.30), the error event denoted by the indicator function $\mathbf{1}_E$ is as follows

$$\mathbf{1}_E = \sum_{\mathcal{V} \subseteq \mathcal{N}} \mathbf{1} \left[\theta_r \in \mathcal{D}_{\mathcal{V}} \text{ and } r > I_{\text{MNNC}}(\mathcal{V}) \right], \tag{B.31}$$

where

$$I_{\text{MNNC}}(\mathcal{V}) = \max_{\mathcal{T} \subseteq \mathcal{V}} \min \left\{ \min_{\mathcal{S} \subseteq \mathcal{T}} R_{\mathcal{T}}(\mathcal{S}, \theta), \min_{i \in \mathcal{V}^c} I(X; Z_{i\theta_r} | X^{(1)}_{\mathcal{N}}) \right\}.$$

As before, the expected value is taken in two steps. For each θ_r, the expected error is calculated with $\mathbb{P}_{\theta|\theta_r}$. The relays chooses the distribution $\prod_{j\in\mathcal{V}} p(x_j^{(2)}) p(\hat{z}_j | x_j^{(2)} z_j)$ to minimize the conditional expectation for each θ_r and for each $\mathcal{D}_{\mathcal{V}}$. This will lead to the following:

$$\mathbb{E}_{\theta|\theta_r}[\mathbf{1}_E] = \sum_{\mathcal{V} \subseteq \mathcal{N}} \min_{\prod_{j\in\mathcal{V}} p(x_j^{(2)}) p(\hat{z}_j | x_j^{(2)} z_j)} \mathbb{P}_{\theta|\theta_r} \left[r > I_{\text{MNNC}}(\mathcal{V}), \theta_r \in \mathcal{D}_{\mathcal{V}} \middle| \theta_r \right]$$

At the next step, the expected value is taken over θ_r and is minimized over decision regions $\mathcal{D}_{\mathcal{V}}$ and $p(x, x^{(1)}_{\mathcal{N}})$ which will lead to the following:

$$\bar{\epsilon}(r) \leq \min_{p(x, x^{(1)}_{\mathcal{N}})} \inf_{\{\mathcal{D}_{\mathcal{V}}, \mathcal{V} \subseteq \mathcal{N}\} \in \Pi(\Theta_r, N)} \sum_{\mathcal{V} \subseteq \mathcal{N}} \mathbb{E}_{\theta_r} \left\{ \min_{\prod_{j\in\mathcal{V}} p(x_j^{(2)}) p(\hat{z}_j | x_j^{(2)} z_j)} \mathbb{P}_{\theta|\theta_r} \left[r > I_{\text{MNNC}}(\mathcal{V}), \theta_r \in \mathcal{D}_{\mathcal{V}} \middle| \theta_r \right] \right\}$$

At the end time sharing RV Q can be added to the region, however the optimization should be done outside the expectation.

B.5 Proof of the Corollaries 17 and 18

Proof of the corollary 17: We start with $R_{\mathcal{T}}(\mathcal{S})$. We know:

$$\begin{aligned} R_{\mathcal{T}}(\mathcal{S}) =& I(XX_{\mathcal{V}^c}X_{\mathcal{S}};\hat{Z}_{\mathcal{S}^c}Y_1|X_{\mathcal{S}^c}) - I(Z_{\mathcal{S}};\hat{Z}_{\mathcal{S}}|XX_{\mathcal{T}\cup\mathcal{V}^c}\hat{Z}_{\mathcal{S}^c}Y_1) \\ =& h(\hat{Z}_{\mathcal{S}^c}Y_1|X_{\mathcal{S}^c}) + h(\hat{Z}_{\mathcal{S}}|Z_{\mathcal{S}}X_{\mathcal{S}}) - h(\hat{Z}_{\mathcal{T}}Y_1|XX_{\mathcal{T}\cup\mathcal{V}^c}) \end{aligned}$$

Now suppose that $\mathcal{D} = \mathcal{S}^c \cup \{M+1\}$ and $\mathcal{S}' = \mathcal{V}^c \cup \mathcal{T}^c \cup \mathcal{S} \cup \{0\}$. We start with finding differential entropies:

$$\begin{aligned} h(\hat{Z}_{\mathcal{T}}Y_1|XX_{\mathcal{V}^c}X_{\mathcal{T}}) &= h(H(\mathcal{T}^c,\mathcal{S}\cup\mathcal{D})\mathbf{X}(\mathcal{T}^c) + \mathcal{N}(\mathcal{S}\cup\mathcal{D}) + \hat{\mathcal{N}}(\mathcal{S}\cup\mathcal{D})) = \\ &\frac{1}{2}\log\left((2\pi e)^{|\mathcal{S}+\mathcal{S}^c+1|}\left|H(\mathcal{T}^c,\mathcal{S}\cup\mathcal{D})\mathbf{K}(\mathcal{T}^c)H(\mathcal{T}^c,\mathcal{S}\cup\mathcal{D})^T + \mathbf{N}(\mathcal{S}\cup\mathcal{D}) + \hat{\mathbf{N}}(\mathcal{S}\cup\mathcal{D})\right|\right) \end{aligned}$$

$$\begin{aligned} h(\hat{Z}_{\mathcal{S}^c}Y_1|X_{\mathcal{S}^c}) =& h(H(\mathcal{S}',\mathcal{D})\mathbf{X}(\mathcal{S}') + \mathcal{N}(\mathcal{D}) + \hat{\mathcal{N}}(\mathcal{D})) = \\ & \frac{1}{2}\log\left((2\pi e)^{|\mathcal{S}^c|+1}\left|H(\mathcal{S}',\mathcal{D})\mathbf{K}(\mathcal{S}')H(\mathcal{S}',\mathcal{D})^T + \mathbf{N}(\mathcal{D}) + \hat{\mathbf{N}}(\mathcal{D})\right|\right) \end{aligned}$$

$$h(\hat{Z}_{\mathcal{S}}|Z_{\mathcal{S}}X_{\mathcal{S}}) = \frac{1}{2}\log\left((2\pi e)^{|\mathcal{S}|}\left|\hat{\mathbf{N}}(\mathcal{S})\right|\right)$$

The previous differential entropies will lead to the rate for $R_{\mathcal{T}}(\mathcal{S})$.

On the other hand $I_{\mathrm{DF}}(k) = I(X;Z_k|X_{\mathcal{V}^c})$ can be bounded in a similar way. First Z_k can be re-written as follows:

$$Z_k = h_{0k}X + H(\mathcal{M},\{k\})\mathbf{X}(\mathcal{M}) + \mathcal{N}_k.$$

To obtain the rate, first we calculate $h(Z_k|X_{\mathcal{V}^c})$. Then it is obvious that $h(Z_k|XX_{\mathcal{V}^c})$ is the entropy of the noise seen at the DF relay caused by its own Gaussian noise and the interference of all CF relays and can be written as follows:

$$h(Z_k|XX_{\mathcal{V}^c}) = \frac{1}{2}\log\left((2\pi e)\left(H(\mathcal{V},\{k\})\mathbf{K}(\mathcal{V})H(\mathcal{V},\{k\})^T + N_k\right)\right).$$

To calculate $h(Z_k|X_{\mathcal{V}^c})$, it can be seen that

$$h(X_{\mathcal{V}^c}) = \frac{1}{2}\log\left((2\pi e)^{|\mathcal{V}^c|}|\mathbf{K}(\mathcal{V}^c)|\right).$$

So it is enough to calculate $h(Z_k, X_{\mathcal{V}^c})$, which leads to the following ($\mathbf{G} = \left[\sqrt{PP_i}\rho_{0i}\right], i \in \mathcal{V}^c$):

$$h(Z_k, X_{\mathcal{V}^c}) = \frac{1}{2}\log\left((2\pi e)^{|\mathcal{V}^c|+1}\begin{vmatrix} |h_{0k}|^2 P + H(\mathcal{V},\{k\})\mathbf{K}(\mathcal{V})H(\mathcal{V},\{k\})^T + N_k & h_{0k}\mathbf{G} \\ h_{0k}^*\mathbf{G}^T & \mathbf{K}(\mathcal{V}^c) \end{vmatrix}\right).$$

The determinant of the covariance matrix in previous rate can be written as:

$$\begin{vmatrix} |h_{0k}|^2 P + H(\mathcal{V},\{k\})\mathbf{K}(\mathcal{V})H(\mathcal{V},\{k\})^T + N_k & h_{0k}\mathbf{G} \\ h_{0k}^*\mathbf{G} & \mathbf{K}(\mathcal{V}^c) \end{vmatrix} =$$
$$\left(|h_{0k}|^2 P + H(\mathcal{V},\{k\})\mathbf{K}(\mathcal{V})H(\mathcal{V},\{k\})^T + N_k\right) \times |\mathbf{K}(\mathcal{V}^c)| - |h_{0k}|^2 P \times$$
$$\left(\sum_{i=0,k_i\in\mathcal{V}^c}^{i=|\mathcal{V}^c|-1} (-1)^i \sqrt{P_{k_i}}\rho_{0k_i} \begin{vmatrix} \frac{1}{\sqrt{P}}\mathbf{G}^T & \mathbf{K}_i(\mathcal{V}^c) \end{vmatrix}\right).$$

$\mathbf{K}_i(\mathcal{V}^c)$ is the covariance matrix $\mathbf{K}(\mathcal{V}^c)$ with the column i removed. Now we can see that:

$$h(Z_k|X_{\mathcal{V}^c}) = \frac{1}{2}\log\left((2\pi e)\left(|h_{0k}|^2(1-\beta)P + H(\mathcal{V},\{k\})\mathbf{K}(\mathcal{V})H(\mathcal{V},\{k\})^T + N_k\right)\right).$$

And β is defined as follows:

$$\beta = \frac{\left(\sum_{i=0,k_i\in\mathcal{V}^c}^{i=|\mathcal{V}^c|-1} (-1)^i \sqrt{P_{k_i}}\rho_{0k_i} \begin{vmatrix} \frac{1}{\sqrt{P}}\mathbf{G}^T & \mathbf{K}_i(\mathcal{V}^c) \end{vmatrix}\right)}{|\mathbf{K}(\mathcal{V}^c)|}$$

Based on this $I_{\mathrm{DF}}(k)$ can be calculated as follows:

$$I_{\mathrm{DF}}(k) = \frac{1}{2}\log\left(1 + \frac{|h_{0k}|^2(1-\beta)P}{H(\mathcal{V},\{k\})\mathbf{K}(\mathcal{V})H(\mathcal{V},\{k\})^T + N_k}\right). \tag{B.32}$$

$I_{\mathrm{DF}}(k)$ and $R_{\mathcal{T}}(\mathcal{S})$ will provide the region according to the theorem 37.

$$R \le \max_{P\in\mathcal{P}} \max_{\mathcal{T}\subseteq\mathcal{V}\subseteq\mathcal{N}} \min\left\{\min_{\mathcal{S}\subseteq\mathcal{T}} R_{\mathcal{T}}(\mathcal{S}), \min_{k\in\mathcal{V}^c} I_{\mathrm{DF}}(k)\right\}.$$

Proof of the corollary 17: To prove this corollary it is enough to bound $R^{(k)}_{\mathcal{T}_k}(\mathcal{S})$. We use (B.25) to evaluate it.

$$\begin{aligned} R^{(k)}_{\mathcal{T}_k}(\mathcal{S}) &= I(X;\hat{Z}_{\mathcal{S}^c}Z_k|X_{\mathcal{V}^c}X_{\mathcal{T}_k}) + I(X_{\mathcal{S}};Z_k|X_{\mathcal{V}^c}X_{\mathcal{S}^c}) - I(\hat{Z}_{\mathcal{S}};Z_{\mathcal{S}}|XX_{\mathcal{V}^c\cup\mathcal{T}_k}\hat{Z}_{\mathcal{S}^c}Z_k) \\ &= h(\hat{Z}_{\mathcal{S}^c}Z_k|X_{\mathcal{V}^c}X_{\mathcal{T}_k}) + h(\hat{Z}_{\mathcal{S}}|Z_{\mathcal{S}}X_{\mathcal{S}}) - h(\hat{Z}_{\mathcal{T}_k}Z_k|XX_{\mathcal{V}^c\cup\mathcal{T}_k}) + I(X_{\mathcal{S}};Z_k|X_{\mathcal{V}^c}X_{\mathcal{S}^c}) \end{aligned}$$

Now suppose that $\mathcal{D}_k = \mathcal{S}^c \cup \{k\}$ and $\mathcal{S}'_k = \mathcal{V}^c \cup \mathcal{T}_k^c \cup \mathcal{S} \cup \{0\}$. We re-state the following differential entropy:

$$h(\hat{Z}_{\mathcal{T}_k}Z_k|XX_{\mathcal{V}^c}X_{\mathcal{T}_k}) = h(H(\mathcal{T}_k^c,\mathcal{S}\cup\mathcal{D}_k)\mathbf{X}(\mathcal{T}_k^c) + \mathcal{N}(\mathcal{S}\cup\mathcal{D}_k) + \hat{\mathcal{N}}(\mathcal{S}\cup\mathcal{D}_k)) =$$
$$\frac{1}{2}\log\left((2\pi e)^{|\mathcal{S}|+|\mathcal{S}^c|+1}\left|H(\mathcal{T}_k^c,\mathcal{D}_k\cup\mathcal{S})\mathbf{K}(\mathcal{T}_k^c)H(\mathcal{T}_k^c,\mathcal{D}_k\cup\mathcal{S})^T + \mathbf{N}(\mathcal{D}_k\cup\mathcal{S}) + \hat{\mathbf{N}}(\mathcal{D}_k\cup\mathcal{S})\right|\right).$$

To find $h(\hat{Z}_{\mathcal{S}^c} Z_k | X_{\mathcal{V}^c} X_{\mathcal{T}_k})$ we write it as:

$$h(\hat{Z}_{\mathcal{S}^c} Z_k | X_{\mathcal{V}^c} X_{\mathcal{T}_k}) = h(\hat{Z}_{\mathcal{S}^c} Z_k X_{\mathcal{V}^c} | X_{\mathcal{T}_k}) - h(X_{\mathcal{V}^c}).$$

Again we have $h(X_{\mathcal{V}^c}) = \frac{1}{2}\log\left((2\pi e)^{|\mathcal{V}^c|}|\mathbf{K}(\mathcal{V}^c)|\right)$ and then the following manipulation is straightforward:

$$h(\hat{Z}_{\mathcal{S}^c} Z_k X_{\mathcal{V}^c} | X_{\mathcal{T}_k}) =$$
$$\frac{1}{2}\log\left((2\pi e)^{|\mathcal{V}^c|+|\mathcal{S}^c|+1}\begin{vmatrix} H(\mathcal{S}'_k,\mathcal{D}_k)\mathbf{K}(\mathcal{S}'_k)H(\mathcal{S}'_k,\mathcal{D}_k)^T + \mathbf{N}(\mathcal{D}_k) + \hat{\mathbf{N}}(\mathcal{D}_k) & H(\mathcal{S}'_k,\mathcal{D}_k)\mathbf{K}(\mathcal{S}'_k,\mathcal{V}^c) \\ H(\mathcal{S}'_k,\mathcal{D}_k)\mathbf{K}(\mathcal{S}'_k,\mathcal{V}^c)^T & \mathbf{K}(\mathcal{V}^c) \end{vmatrix}\right).$$

Then we move to the case of $I(X_{\mathcal{S}}; Z_k | X_{\mathcal{V}^c} X_{\mathcal{S}^c})$. Similar to the evaluation of $I_{\mathrm{DF}}(k)$, we can write down:

$$h(Z_k X_{\mathcal{V}^c} | X_{\mathcal{T}_k}) =$$
$$\frac{1}{2}\log\left((2\pi e)^{|\mathcal{V}^c|+1}\begin{vmatrix} |h_{0k}|^2 P + H(\mathcal{T}_k^c,\{k\})\mathbf{K}(\mathcal{T}_k^c)H(\mathcal{T}_k^c,\{k\})^T + N_k & h_{0k}\mathbf{G} \\ h_{0k}^*\mathbf{G}^T & \mathbf{K}(\mathcal{V}^c) \end{vmatrix}\right).$$
$$h(Z_k X_{\mathcal{V}^c} | X_{\mathcal{S}_c}) =$$
$$\frac{1}{2}\log\left((2\pi e)^{|\mathcal{V}^c|+1}\begin{vmatrix} |h_{0k}|^2 P + H(\mathcal{T}_k^c\cup\mathcal{S},\{k\})\mathbf{K}(\mathcal{T}_k^c\cup\mathcal{S})H(\mathcal{T}_k^c\cup\mathcal{S},\{k\})^T + N_k & h_{0k}\mathbf{G} \\ h_{0k}^*\mathbf{G}^T & \mathbf{K}(\mathcal{V}^c) \end{vmatrix}\right).$$

The preceding differential entropies are simplified to:

$$h(Z_k | X_{\mathcal{V}^c} X_{\mathcal{T}_k}) = \frac{1}{2}\log\left((2\pi e)\left(|h_{0k}|^2(1-\beta)P + H(\mathcal{T}_k^c,\{k\})\mathbf{K}(\mathcal{T}_k^c)H(\mathcal{T}_k^c,\{k\})^T + N_k\right)\right).$$
$$h(Z_k | X_{\mathcal{V}^c} X_{\mathcal{S}^c}) =$$
$$\frac{1}{2}\log\left((2\pi e)\left(|h_{0k}|^2(1-\beta)P + H(\mathcal{T}_k^c\cup\mathcal{S},\{k\})\mathbf{K}(\mathcal{T}_k^c\cup\mathcal{S})H(\mathcal{T}_k^c\cup\mathcal{S},\{k\})^T + N_k\right)\right).$$

And so:

$$I(X_{\mathcal{S}}; Z_k | X_{\mathcal{V}^c} X_{\mathcal{S}^c}) = \frac{1}{2}\log\left(1 + \frac{H(\mathcal{S},\{k\})\mathbf{K}(\mathcal{S})H(\mathcal{S},\{k\})^T}{\left(|h_{0k}|^2(1-\beta)P + H(\mathcal{T}_k^c,\{k\})\mathbf{K}(\mathcal{T}_k^c)H(\mathcal{T}_k^c,\{k\})^T + N_k\right)}\right).$$

Finally we have:

$$h(\hat{Z}_{\mathcal{S}} | Z_{\mathcal{S}} X_{\mathcal{S}}) = \frac{1}{2}\log\left((2\pi e)^{|\mathcal{S}|}\left|\hat{\mathbf{N}}(\mathcal{S})\right|\right).$$

Using all this, can be written as follows:

$$R_{\mathcal{T}_k}^{(k)}(\mathcal{S}) =$$

$$\frac{1}{2}\log\left(\frac{\begin{vmatrix} H(\mathcal{S}'_k,\mathcal{D}_k)\mathbf{K}(\mathcal{S}'_k)H(\mathcal{S}'_k,\mathcal{D}_k)^T+\mathbf{N}(\mathcal{D}_k)+\hat{\mathbf{N}}(\mathcal{D}_k) & H(\mathcal{S}'_k,\mathcal{D}_k)\mathbf{K}(\mathcal{S}'_k,\mathcal{V}^c) \\ H(\mathcal{S}'_k,\mathcal{D}_k)\mathbf{K}(\mathcal{S}'_k,\mathcal{V}^c)^T & \mathbf{K}(\mathcal{V}^c) \end{vmatrix}}{\left|H(\mathcal{T}_k^c,\mathcal{D}_k\cup\mathcal{S})\mathbf{K}(\mathcal{T}_k^c)H(\mathcal{T}_k^c,\mathcal{D}_k\cup\mathcal{S})^T+\mathbf{N}(\mathcal{D}_k\cup\mathcal{S})+\hat{\mathbf{N}}(\mathcal{D}_k\cup\mathcal{S})\right|\left|\mathbf{K}(\mathcal{V}^c)\right|}\right)$$

$$+\frac{1}{2}\log\left(1+\frac{H(\mathcal{S},\{k\})\mathbf{K}(\mathcal{S})H(\mathcal{S},\{k\})^T}{\left(|h_{0k}|^2(1-\beta)P+H(\mathcal{T}_k^c,\{k\})\mathbf{K}(\mathcal{T}_k^c)H(\mathcal{T}_k^c,\{k\})^T+N_k\right)}\right)-\frac{1}{2}\log\left(\frac{1}{\left|\hat{\mathbf{N}}(\mathcal{S})\right|}\right).$$

Appendix C

Appendix chapter 4

C.1 Proof of Theorem

We start by some definitions. Let the random variable $\mathbf{M}_n^{(ij)}$ denote the sent message chosen with uniform distribution over $\mathcal{M}_n^{(ij)}$. m denotes the number of nodes. For two sets $S_1, S_2 \subset \{1, ..., m\}$ with cardinalities $|S_1|, |S_2|$, let $\mathbf{M}_n^{(S_1 S_2)}$ be an $|S_1||S_2|$-tuple as follows:

$$\mathbf{M}_n^{(S_1 S_2)} = \left(\mathbf{M}_n^{(ij)}\right)_{i \in S_1, j \in S_2}.$$

Hence $\mathbf{M}_n^{(S_1 S_2)}$ denotes a random variable uniformly distributed over the following set:

$$\mathcal{M}_n^{(S_1 S_2)} = \bigotimes_{i \in S_1, j \in S_2} \mathcal{M}_n^{(ij)}$$

where $\bigotimes$ is the Cartesian product of sets. This variable represents the random vector corresponding the messages transmitted from the nodes in S_1 to the nodes in S_2 and its particular value is denoted by the vector $\mathbf{l}^{(S_1 S_2)} = (l^{(ij)}) \in \mathcal{M}_n^{(S_1 S_2)}$. With each $\mathbf{l}^{(S_1 S_2)} = (l^{(ij)})$ is associated the set $\mathcal{D}_{\mathbf{l}^{(S_1 S_2)}}^{(S_1, S_2)c} = \bigcup_{i \in S_1, j \in S_2} \mathcal{D}_{l^{(ij)}}^{(ij)\,c}$ which is equivalent to the error events for some $i \in S, j \in S^c$ if $\mathbf{l}^{(S_1 S_2)}$ is transmitted.

$M_n^{(S_1 S_2)} = \prod_{i \in S_1, j \in S_2} M_n^{(ij)}$ represents the cardinality of $\mathcal{M}_n^{(S_1 S_2)}$, or in other words the sum rate transmitted from the nodes in S_1 to the nodes in S_2. We are particularly interested in the case $S_1 = S$ and $S_2 = S^c$ which is a cut from the network. It is also of interest to consider $S_1 = S_2 = \{1, ..., m\}$ where we are concerned with the whole network. The respective variables in this case are denoted by $\mathbf{M}_n, \mathcal{M}_n, \mathbf{l}, M_n$ and $\mathcal{D}_{\mathbf{l}}^c$. Similarly define $Y_S^n = (Y_j^n)$ for $j \in S$ where Y_S^n takes its values over the set $\mathscr{Y}_S^n = \bigotimes_{j \in S} \mathscr{Y}_j^n$.

The following theorem is Verdu-Han's lemma [44] adapted for the multiterminal networks.

Theorem 46 *For all codes* $\left(n, M_n^{(ij)}, \epsilon_n\right)$, $i \neq j \in \{1, ..., m\}$ *the error probability satisfy:*

$$\epsilon_n \geq \Pr\left[\frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n) \leq \frac{1}{n}\log M_n^{(SS^c)} - \gamma\right] - \exp(-\gamma n). \tag{C.1}$$

For every $\gamma > 0$ *and every* $S \subseteq \{1, ..., m\}$, *where*

$$i(X^n; Y^n) = \log\frac{\mathbb{P}_{X_n Y^n}(Y^n \big| X_n)}{\mathbb{P}_{Y^n}(Y^n)}. \tag{C.2}$$

Consequently the following inequality holds:

$$\epsilon_n \geq \Pr\left[\textit{for some } S; \frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n) \leq \frac{1}{n}\log M_n^{(SS^c)} - \gamma\right] - (2^m - 1)\exp(-\gamma n). \tag{C.3}$$

Proof The inequality above is equal to the following:

$$\epsilon_n \geq \Pr\left[\frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n) \leq \sum_{i\in S, j\in S^c}\frac{1}{n}\log M_n^{(ij)} - \gamma\right] - \exp(-\gamma n). \tag{C.4}$$

The proof follows the same line as Verdu-Han. Consider a $\left(n, M_n^{(ij)}, \epsilon_n\right)$, $i \neq j \in \{1, ..., m\}$ code. By the definition (8.4) and those introduced above, we have:

$$\epsilon_n = \frac{1}{M_n}\sum_{\mathbf{l}\in\mathcal{M}_n}\Pr(\mathcal{D}_l{}^c | \mathbf{M}_n = \mathbf{l})$$

We first prove the following inequality:

$$\beta + \epsilon_n \geq \Pr\left[\mathbb{P}(\mathbf{M}_n^{(SS^c)} | Y_{S^c}^n) \leq \beta\right]. \tag{C.5}$$

Then by choosing $\beta = \exp(-\gamma n)$, the theorem follows.

The starting point is similar to [44] with small additional step. For each $\mathbf{l}^{(SS^c)} = (l^{(ij)}) \in \mathcal{M}_n^{(SS^c)}$ define the sets $B_{\mathbf{l}^{(SS^c)}}$ and $\mathfrak{L}$ as follows:

$$\begin{aligned} B_{\mathbf{l}^{(SS^c)}} =& \{y_{S^c}^n \in \mathscr{Y}^{(S^c)n} : P(\mathbf{l}^{(SS^c)} | y_{S^c}^n) \leq \beta\} \\ \mathfrak{L} =& \bigcup_{\mathbf{l}^{(SS^c)}\in\mathcal{M}_n^{(SS^c)}} B_{\mathbf{l}^{(SS^c)}}. \end{aligned}$$

Then we have another lemma:

Lemma 5 *For a cut in the network dividing nodes into* S, S^c, *define the error probability of the cut* $\epsilon_n^{(SS^c)}$ *as:*

$$\epsilon_n^{(SS^c)} = \frac{1}{M_n^{(SS^c)}} \sum_{\mathbf{l}^{(SS^c)} \in \mathcal{M}_n^{(SS^c)}} \Pr(\mathcal{D}_{\mathbf{l}^{(SS^c)}}^{(SS^c)c} | \mathbf{l}^{(SS^c)})$$

Then $\epsilon_n^{(SS^c)}$ *satisfies the following inequality:*

$$\epsilon_n^{(SS^c)} \leq \epsilon_n$$

Proof By using the assumption of uniform distribution over the input message, it is easy to verify the following steps:

$$\begin{aligned}
\epsilon_n^{(SS^c)} &= \frac{1}{M_n^{(SS^c)}} \sum_{\mathbf{l}^{(SS^c)} \in \mathcal{M}_n^{(SS^c)}} \Pr(\mathcal{D}_{l^{(SS^c)}}^{(SS^c)c} | \mathbf{M}_n^{(SS^c)} = l^{(SS^c)}) \\
&= \sum_{\mathbf{l}^{(SS^c)} \in \mathcal{M}_n^{(SS^c)}} \Pr(\mathcal{D}_{l^{(SS^c)}}^{(SS^c)c}, \mathbf{M}_n^{(SS^c)} = l^{(SS^c)}) + \sum_{\mathbf{l} \in \mathcal{M}_n} P(\mathcal{D}_{l^{(SS^c)}}^{(SS^c)c}, \mathbf{l}) \\
&\leq \sum_{\mathbf{l} \in \mathcal{M}_n} P(\mathcal{D}_{\mathbf{l}}{}^c, \mathbf{l}) = \epsilon_n
\end{aligned}$$

We will use this lemma for the proof. Now note that:

$$\Pr\left[\mathbb{P}(\mathbf{M}_n^{(SS^c)} | Y_{S^c}^n) \leq \beta\right] = \Pr(\mathfrak{L}).$$

Using this, the proof is as follows:

$$\begin{aligned}
\Pr(\mathfrak{L}) &= \sum_{\mathbf{l}^{(SS^c)} \in \mathcal{M}_n^{(SS^c)}} \Pr(\mathbf{l}^{(SS^c)}, B_{\mathbf{l}^{(SS^c)}}) \\
&= \sum_{\mathbf{l}^{(SS^c)} \in \mathcal{M}_n^{(SS^c)}} \Pr(\mathbf{l}^{(SS^c)}, B_{\mathbf{l}^{(SS^c)}} \cap \mathcal{D}_{\mathbf{l}^{(SS^c)}}^{(SS^c)c}) + \sum_{\mathbf{l}^{(SS^c)} \in \mathcal{M}_n^{(SS^c)}} \Pr(\mathbf{l}^{(SS^c)}, B_{\mathbf{l}^{(SS^c)}} \cap \mathcal{D}_{\mathbf{l}^{(SS^c)}}^{(SS^c)}) \\
&= \frac{1}{M_n^{(SS^c)}} \sum_{\mathbf{l}^{(SS^c)} \in \mathcal{M}_n^{(SS^c)}} \Pr(B_{\mathbf{l}^{(SS^c)}} \cap \mathcal{D}_{\mathbf{l}^{(SS^c)}}^{(SS^c)c} | \mathbf{l}^{(SS^c)}) \\
&\qquad + \sum_{\mathbf{l}^{(SS^c)} \in \mathcal{M}_n^{(SS^c)}} \Pr(\mathbf{l}^{(SS^c)} | B_{\mathbf{l}^{(SS^c)}} \cap \mathcal{D}_{\mathbf{l}^{(SS^c)}}^{(SS^c)}) P(B_{\mathbf{l}^{(SS^c)}} \cap \mathcal{D}_{\mathbf{l}^{(SS^c)}}^{(SS^c)}) \\
&\overset{(a)}{\leq} \frac{1}{M_n^{(SS^c)}} \sum_{\mathbf{l}^{(SS^c)} \in \mathcal{M}_n^{(SS^c)}} \Pr(\mathcal{D}_{\mathbf{l}^{(SS^c)}}^{(SS^c)c} | \mathbf{l}^{(SS^c)}) + \sum_{\mathbf{l}^{(SS^c)} \in \mathcal{M}_n^{(SS^c)}} \beta \times \Pr(B_{\mathbf{l}^{(SS^c)}} \cap \mathcal{D}_{\mathbf{l}^{(SS^c)}}^{(SS^c)}) \\
&\overset{(b)}{\leq} \epsilon_n + \beta \sum_{\mathbf{l}^{(SS^c)} \in \mathcal{M}_n^{(SS^c)}} \Pr(B_{\mathbf{l}^{(SS^c)}} \cap \mathcal{D}_{\mathbf{l}^{(SS^c)}}^{(SS^c)})
\end{aligned}$$

(a)comes from the definition of $B_{\mathbf{l}^{(SS^c)}}$ and (b) comes from the lemma. On the other hand we have:

$$\begin{aligned}\sum_{\mathbf{l}^{(SS^c)}\in\mathcal{M}_n^{(SS^c)}} \Pr(B_{\mathbf{l}^{(SS^c)}}\cap\mathcal{D}_{\mathbf{l}^{(SS^c)}}^{(SS^c)}) &\leq \sum_{\mathbf{l}^{(SS^c)}\in\mathcal{M}_n^{(SS^c)}} \Pr(\mathcal{D}_{\mathbf{l}^{(SS^c)}}^{(SS^c)})\\ &= \sum_{\mathbf{l}^{(SS^c)}\in\mathcal{M}_n^{(SS^c)}} \Pr\left(\Big(\bigcup_{i\in S,j\in S^c}\mathcal{D}_{l^{(ij)}}^{(ij)\,c}\Big)^c\right)\\ &= \sum_{\mathbf{l}^{(SS^c)}\in\mathcal{M}_n^{(SS^c)}} \Pr\left(\Big(\bigcap_{i\in S,j\in S^c}\mathcal{D}_{l^{(ij)}}^{(ij)}\Big)\right)\\ &= \Pr\Big(\bigcup_{\mathbf{l}^{(SS^c)}\in\mathcal{M}_n^{(SS^c)}}\Big(\bigcap_{i\in S,j\in S^c}\mathcal{D}_{l^{(ij)}}^{(ij)}\Big)\Big)\leq 1\end{aligned}$$

Hence we have:

$$\Pr(\mathfrak{L}) \leq \epsilon_n + \beta.$$

which results in the first inequality.

The second inequality is followed using the same manipulation, similar to the inequality developed in [43] for multiple access channels.

After stating the proof for this theorem, we are ready to start the proof. The general idea behind the proof is to prove that for all S, the discrete memoryless channel and the rate outside the closure of cut-set bound the following holds for all $\gamma > 0$.

$$\Pr\left[\text{for some } S; \frac{1}{n}i(\mathbf{M}_n^{(SS^c)};Y_{S^c}^n) \leq \sum_{i\in S,j\in S^c}\frac{1}{n}\log M_n^{(ij)} - \gamma\right] \to 1. \tag{C.6}$$

In other words, the information spectrum of $\frac{1}{n}i(\mathbf{M}_n^{(SS^c)};Y_{S^c}^n)$ is placed on the left-hand side of cut-set bound values. Then based on the theorem 46 the probability of error goes ϵ_n to one. The proof follows two general steps. First it is shown that the information density is less than another random variable called $U_n^{(S)}$ in probability as $n\to\infty$. In the next step it is shown that the random variable $U_n^{(S)}$ is less than the cut set bound in probability. So if $\frac{1}{n}\log M_n^{(ij)} - \gamma$ is bigger than the cut set bound as $n\to\infty$ then it is easy to see that the probability in (C.6) will tend to 1 as $n\to\infty$.

C.1.1 First Step

For a general (not necessarily i.i.d.) inputs and outputs define

$$U_n^{(S)} = \frac{1}{n}\sum_{j=1}^{n} \log \frac{\mathbb{W}_j(Y_{S^c j}|X_{S^c j}, X_{Sj})}{\mathbb{P}_{Y_{S^c j}, X_{S^c j}}(Y_{S^c j}|X_{S^c j})}$$

where $\mathbb{W}_j$ is the channel in j^{th} use. By assumption that the channel is discrete memoryless, $\mathbb{W}_j$ is the same, namely $\mathbb{W}$ for all j. However we keep the index until the end.

Note that Y_k^n is not i.i.d. in general. However if X_k^n's are i.i.d. for each $k \in \{1, ..., m\}$ then $U_n^{(S)}$ turns into the sum of i.i.d. random variables and tends to its expected value a.e. by the law of large number which is $I(X_S; Y_{S^c}|X_{S^c})$. However no restriction is put over the input distribution. For simplicity, we drop the subscripts in probability when implicitly clear. For $U_n^{(S)}$ defined as such we prove the following lemma:

Lemma 6 *For $U_n^{(S)}$ defined as above we have:*

$$\lim_{n\to\infty} P\Big(\text{for all } S; U_n^{(S)} > \frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n)\Big) = 1. \tag{C.7}$$

Proof By the independence of $\mathbf{M}^{(SS^c)}$ from $\mathbf{M}^{((SS^c)^c)}$, we have:.

$$i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n) = i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n|\mathbf{M}_n^{((SS^c)^c)}) - i(\mathbf{M}_n^{(SS^c)}; \mathbf{M}_n^{((SS^c)^c)}|Y_{S^c}^n) \tag{C.8}$$

We continue to simplify the first term on the right hand side:

$$\begin{aligned}
i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n|\mathbf{M}_n^{((SS^c)^c)}) &= \log \frac{\mathbb{P}(Y_{S^c}^n|\mathbf{M}_n^{(SS^c)}, \mathbf{M}_n^{((SS^c)^c)})}{\mathbb{P}(Y_{S^c}^n|\mathbf{M}_n^{((SS^c)^c)})} \\
&= \sum_{j=1}^{n} \log \frac{\mathbb{P}(Y_{S^c j}|Y_{S^c}^{j-1}, \mathbf{M}_n^{(SS^c)}, \mathbf{M}_n^{((SS^c)^c)})}{\mathbb{P}(Y_{S^c j}|Y_{S^c}^{j-1}, \mathbf{M}_n^{((SS^c)^c)})} \\
&\overset{(a)}{=} \sum_{j=1}^{n} \log \frac{\mathbb{P}(Y_{S^c j}|Y_{S^c}^{j-1}, X_{S^c j}, \mathbf{M}_n^{(SS^c)}, \mathbf{M}_n^{((SS^c)^c)})}{\mathbb{P}(Y_{S^c j}|Y_{S^c}^{j-1}, X_{S^c j}, \mathbf{M}_n^{((SS^c)^c)})} \\
&= \sum_{j=1}^{n} \log \frac{\mathbb{P}(Y_{S^c j}|X_{S^c j}, X_{Sj})}{\mathbb{P}(Y_{S^c j}|X_{S^c j})} \\
&+ \sum_{j=1}^{n} \log \frac{\mathbb{P}(Y_{S^c j}|Y_{S^c}^{j-1}, X_{S^c j}, \mathbf{M}_n^{(SS^c)}, \mathbf{M}_n^{((SS^c)^c)})}{\mathbb{P}(Y_{S^c j}|X_{Sj}, Y_{S^c}^{j-1}, X_{S^c j}, \mathbf{M}_n^{(SS^c)}, \mathbf{M}_n^{((SS^c)^c)})} \\
&+ \sum_{j=1}^{n} \log \frac{\mathbb{P}(Y_{S^c j}|X_{S^c j})}{\mathbb{P}(Y_{S^c j}|Y_{S^c}^{j-1}, X_{S^c j}, \mathbf{M}_n^{((SS^c)^c)})}
\end{aligned}$$

where (a) is due to the fact that X_{S^cj} is a function of $Y_{S^c}^{j-1}, \mathbf{M}_n^{((SS^c)^c)}$. Now using (C.8), the definition of $U_n^{(S)}$ and the notion of information density we get:

$$nU_n^{(S)} - i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n) = i(\mathbf{M}_n^{(SS^c)}; \mathbf{M}_n^{((SS^c)^c)} | Y_{S^c}^n) + \sum_{j=1}^{n} i(Y_{S^cj}; X_{Sj} | Y_{S^c}^{j-1}, X_{S^cj}, \mathbf{M}_n^{(SS^c)}, \mathbf{M}_n^{((SS^c)^c)}) + \sum_{j=1}^{n} i(Y_{S^cj}; \mathbf{M}_n^{((SS^c)^c)}, Y_{S^c}^{j-1} | X_{S^cj}) \tag{C.9}$$

Now consider the following inequality for two sequences of random variables $\{X_n\}_{n=1}^{\infty}$ and $\{Y_n\}_{n=1}^{\infty}$ [43]:

$$\text{p-}\liminf_{n\to\infty}(X_n + Y_n) \geq \text{p-}\liminf_{n\to\infty} X_n + \text{p-}\liminf_{n\to\infty} Y_n.$$

Now if we take liminf in probability from both sides and dividing by $\frac{1}{n}$, we have the following inequality:

$$\begin{aligned}\text{p-}\liminf_{n\to\infty}\left(U_n^{(S)} - \frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n)\right) &\geq \text{p-}\liminf_{n\to\infty} \frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; \mathbf{M}_n^{((SS^c)^c)} | Y_{S^c}^n) \\ &+ \text{p-}\liminf_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} i(Y_{S^cj}; X_{Sj} | Y_{S^c}^{j-1}, X_{S^cj}, \mathbf{M}_n^{(SS^c)}, \mathbf{M}_n^{((SS^c)^c)}) \\ &+ \text{p-}\liminf_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} i(Y_{S^cj}; \mathbf{M}_n^{((SS^c)^c)}, Y_{S^c}^{j-1} | X_{S^cj})\end{aligned} \tag{C.10}$$

According to the positivity of inf-mutual information rate [43,44], three terms on the right hand side of the inequality are positive and so it can be said that liminf in probability of $U_n^{(S)} - \frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n)$ will be positive. Then by definition of liminf in probability, the probability of $\frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n) < U_n^{(S)}$ goes to one as $n \to \infty$ which means:

$$\lim_{n\to\infty} \Pr\left(U_n^{(S)} - \frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n) > 0\right) = \lim_{n\to\infty} \Pr\left(U_n^{(S)} > \frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n)\right) = 1. \tag{C.11}$$

Now because this holds for all S and the probability goes to one, then the next equation follows immediately:

$$\lim_{n\to\infty} \Pr\left(\text{for all } S; U_n^{(S)} > \frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n)\right) = 1. \tag{C.12}$$

C.1.2 Second Step

The second step is based on the following lemma.

Lemma 7 *For $U_n^{(S)}$ defined as above, there is a probability distribution $\mathbb{P}_{X_S X_{S^c}}$ for each $x_{\mathcal{N}}^n$ such that:*

$$\lim_{n\to\infty} \mathbb{E}\left[\Pr\left(U_n^{(S)} < I(X_S; Y_{S^c}|X_{S^c})\right)|X_{\mathcal{N}}^n\right] = 1,$$

Proof For each $x_S^n \in \mathscr{X}_S^n, x_{S^c}^n \in \mathscr{X}_{S^c}^n$, define:

$$U_n^{(S)}(x_S^n, x_{S^c}^n) = \frac{1}{n}\sum_{j=1}^{n} \log \frac{\mathbb{W}_j(Y_{S^cj}|x_{S^cj}, x_{Sj})}{\mathbb{P}(Y_{S^cj}|x_{S^cj})}$$

Note that given $x_S^n, x_{S^c}^n$, the probability distribution $\mathbb{W}^n(Y_{S^c}^n|x_S^n, x_{S^c}^n)$ and using the fact that the channel is memoryless, it can be seen that $\log \frac{\mathbb{W}_j(Y_{S^cj}|x_{S^cj},x_{Sj})}{\mathbb{P}(Y_{S^cj}|x_{S^cj})}$ for all j are independent.

Here, we have to use the assumption of discreteness of the channel and their finite cardinality which guarantees the finite variance of each $\log \frac{\mathbb{W}_j(Y_{S^cj}|x_{S^cj},x_{Sj})}{\mathbb{P}(Y_{S^cj}|x_{S^cj})}$ in the previous sum. Using Chebyshev's inequality it can be seen that:

$$\Pr\left(U_n^{(S)}(x_S^n, x_{S^c}^n) > \mathbb{E}[U_n^{(S)}(x_S^n, x_{S^c}^n)] + \delta \middle| x_S^n, x_{S^c}^n\right) < \frac{\sigma^2}{n\delta^2}.$$

In other words:

$$\Pr\left(U_n^{(S)}(x_S^n, x_{S^c}^n) \leq \mathbb{E}[U_n^{(S)}(x_S^n, x_{S^c}^n)] + \delta \middle| x_S^n, x_{S^c}^n\right) > 1 - \frac{\sigma^2}{n\delta^2}.$$

And for a network with m nodes it can be seen that:

$$\Pr\left(\text{for all } S; U_n^{(S)}(x_S^n, x_{S^c}^n) \leq \mathbb{E}[U_n^{(S)}(x_S^n, x_{S^c}^n)] + \delta \middle| x_{\mathcal{N}}^n\right) > 1 - \frac{2^m\sigma^2}{n\delta^2},$$

where $\mathcal{N} = S \cup S^c = \{1, ..., m\}$. Now assume that X_S, X_{S^c} follow the PD $\mathbb{P}_{X_S X_{S^c}}$ which is the empirical distribution of $(x_S^n, x_{S^c}^n)$. Then we can say that:

$$\mathbb{E}[U_n^{(S)}(x_S^n, x_{S^c}^n)] = I(X_S; Y_S|X_{S^c}).$$

Note that here we have to assume that the channel $\mathbb{W}_j$ is the same for all j. So we get:

$$\Pr\left(\text{for all } S; U_n^{(S)}(x_S^n, x_{S^c}^n) \leq I(X_S; Y_S|X_{S^c}) + \delta \middle| x_{\mathcal{N}}^n\right) > 1 - \frac{2^m\sigma^2}{n\delta^2},$$

Now for each $x_{\mathcal{N}}^n$, the previous probability goes to 1 as $n \to \infty$ which proves the theorem.

C.1.3 Final Step

We are ready to prove the theorem. First consider the following inequalities

$$\Pr\Big[\text{for some } S; \frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n) \leq \sum_{i\in S, j\in S^c} \frac{1}{n}\log M_n^{(ij)} - \gamma\Big] \geq$$

$$\sum_{x_{\mathcal{N}}^n} \mathbb{P}_{X_{\mathcal{N}}^n}(x_{\mathcal{N}}^n) \Pr\Big(\text{for all } S; U_n^{(S)} > \frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n), U_n^{(S)} < I(X_S; Y_{S^c}|X_{S^c})|x_{\mathcal{N}}^n\Big)$$

$$\times \Pr\Big[\text{for some } S; \frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n) \leq \sum_{i\in S, j\in S^c} \frac{1}{n}\log M_n^{(ij)} - \gamma$$

$$\Big|\text{for all } I(X_S; Y_{S^c}|X_{S^c}) > U_n^{(S)} > \frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n), x_{\mathcal{N}}^n\Big]$$

$$\geq \sum_{x_{\mathcal{N}}^n} \mathbb{P}_{X_{\mathcal{N}}^n}(x_{\mathcal{N}}^n) \Pr\Big(\text{for all } S; U_n^{(S)} > \frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n), U_n^{(S)} < I(X_S; Y_{S^c}|X_{S^c})|x_{\mathcal{N}}^n\Big)$$

$$\times \mathbf{1}\Big[\Big(\frac{1}{n}\log M_n^{(ij)}\Big)_{i,j\in\mathcal{N}, i\neq j} \notin \mathscr{S}_{\text{CB}}\Big].$$

In the last step, $\left(\frac{1}{n}\log M_n^{(ij)}\right)_{i,j\in\mathcal{N}, i\neq j} \notin \mathscr{S}_{\text{CB}}$, means that, if the rate $\left(\frac{1}{n}\log M_n^{(ij)}\right)_{i,j\in\mathcal{N}, i\neq j}$ is outside the closure of cut set bound then for each $\mathbb{P}_{X_S X_{S^c}}$, there are some S such that:

$$\sum_{i\in S, j\in S^c} \frac{1}{n}\log M_n^{(ij)} > I(X_S; Y_{S^c}|X_{S^c}).$$

and so we have:

$$\sum_{i\in S, j\in S^c} \frac{1}{n}\log M_n^{(ij)} > \frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n),$$

which justifies the last step.

Using the last inequality with the lemmas we have:

$$\epsilon_n \geq \Pr\left[\frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n) \leq \sum_{i\in S, j\in S^c} \frac{1}{n}\log M_n^{(ij)} - \gamma\right] - (2^m - 1)\exp(-\gamma n)$$

$$\geq \sum_{x_{\mathcal{N}}^n} \mathbb{P}_{X_{\mathcal{N}}^n}(x_{\mathcal{N}}^n) \Pr\Big(\text{for all } S; U_n^{(S)} > \frac{1}{n} i(\mathbf{M}_n^{(SS^c)}; Y_{S^c}^n), U_n^{(S)} < I(X_S; Y_{S^c}|X_{S^c})|x_{\mathcal{N}}^n\Big)$$

$$\times \mathbf{1}\Big[\Big(\frac{1}{n}\log M_n^{(ij)}\Big)_{i,j\in\mathcal{N}, i\neq j} \notin \mathscr{S}_{\text{CB}}\Big] - (2^m - 1)\exp(-\gamma n). \tag{C.13}$$

From the first and second step we know that the probability on the right-hand side of (C.13) converges to 1 as $n \to \infty$. By taking the limit on both sides of (C.13), we obtain for all γ

$$\lim_{n\to\infty} \epsilon_n \geq \lim_{n\to\infty} \mathbf{1}\Big[\Big(\frac{1}{n}\log M_n^{(ij)}\Big)_{i,j\in\mathcal{N}, i\neq j} \notin \mathscr{S}_{\text{CB}}\Big]. \tag{C.14}$$

Then it is easy to see that the rates not satisfying the cutset bounds for each S, i.e. falling outside the closure of cutset bounds, lead to the error proability 1.

Bibliography

[1] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *Information Theory, IEEE Trans. on*, vol. IT-25, pp. 572–584, 1979.

[2] A. Carleial, "Multiple-access channels with different generalized feedback signals," *Information Theory, IEEE Transactions on*, vol. 28, pp. 841–850, Nov 1982.

[3] F. M. J. Willems, "Informationtheoretical results for the discrete memoryless multiple access channel," *Doctor in de Wetenschappen Proefschrift dissertation, Katholieke Univ. Leuven, Leuven, Belgium*, Oct. 1982.

[4] A. E. Gamal, "The capacity of a class of broadcast channels," *IEEE Trans. Information Theory*, vol. IT-25, pp. 166–169, March 1979.

[5] A. E. Gamal and M. Aref, "The capacity of semideterministic relay channel," *IEEE Trans. Information Theory*, vol. IT-28, p. 536, May 1982.

[6] P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *Information Theory, IEEE Transactions on*, vol. 19, pp. 197–207, Mar 1973.

[7] P. Bergmans, "A simple converse for broadcast channels with additive white gaussian noise (corresp.)," *Information Theory, IEEE Transactions on*, vol. 20, pp. 279–280, mar 1974.

[8] R. G. Gallager, "Capacity and coding for degraded broadcast channels," *Probl. Peredachi Inf.*, vol. 10, pp. 3–14, July-Sept 1974.

[9] R. Ahlswede and J. Korner, "Source coding with side information and a converse for degraded broadcast channels," *Information Theory, IEEE Transactions on*, vol. 21, pp. 629–637, nov 1975.

[10] J. Korner and K. Marton, "General broadcast channels with degraded message sets," *Information Theory, IEEE Transactions on*, vol. 23, pp. 60–64, jan 1977.

[11] J. Korner and K. Marton, “A source network problem involving the comparison of two channels ii,” *Trans. Colloquium Inform. Theory, Keszthely, Hungary,*, Aug. 1975.

[12] K. Marton, “A coding theorem for the discrete memoryless broadcast channel,” *Information Theory, IEEE Transactions on*, vol. 25, pp. 306–311, May 1979.

[13] A. El Gamal and E. Van der Meulen, “A proof of Marton’s coding theorem for the discrete memoryless broadcast channel,” *Information Theory, IEEE Transactions on*, vol. IT-27, pp. 120–122, 1981.

[14] G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” *Information Theory, IEEE Transactions on*, vol. 51, pp. 3037–3063, Sept. 2005.

[15] P. Elias, A. Feinstein, and C. Shannon, “A note on the maximum flow through a network,” *Information Theory, IRE Transactions on*, vol. 2, no. 4, pp. 117 –119, 1956.

[16] D. R. Ford, L. R.; Fulkerson, “Maximal flow through a network,” *Canadian Journal of Mathematics*, vol. 8, p. 399–404, 1956.

[17] G. B. Dantzig and D. R. Fulkerson, “On the max flow min cut theorem of networks,” *Ann. Math. Studies, Princeton, New Jersey*, no. 38, 1956.

[18] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, “Noisy network coding,” *Information Theory, IEEE Transactions on*, vol. 57, pp. 3132 –3152, may 2011.

[19] A. Avestimehr, S. Diggavi, and D. Tse, “Wireless network information flow: A deterministic approach,” *Information Theory, IEEE Transactions on*, vol. 57, pp. 1872 –1905, april 2011.

[20] R. L. Dobrusin, “Optimal information transfer over channel with unknown parameters,” *Radiotech. i Elektron.*, vol. 4, pp. 1951–1956, 1959.

[21] J. Wolfowitz, “Simultaneous channels,” *Arch. Rat. Mech. Anal.*, vol. 4, pp. 371–386, 1960.

[22] L. B. David Blackwell and A. J. Thomasian, “The capacity of a class of channels,” *Ann. Math. Statist.*, vol. 30, no. 4, pp. 1229–1241, 1959.

[23] S. Verdu and S. Shamai, “Variable-rate channel capacity,” *Information Theory, IEEE Transactions on*, vol. 56, pp. 2651–2667, June 2010.

[24] T. Cover, "Broadcast channels," *IEEE Trans. Information Theory*, vol. IT-18, pp. 2–14, 1972.

[25] S. Shamai, "A broadcast strategy for the gaussian slowly fading channel," *Information Theory. 1997. Proceedings., 1997 IEEE International Symposium on*, pp. 150–, Jun-4 Jul 1997.

[26] S. S. L. Ozarow and A. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Information Theory*, vol. 43, p. 359–378, May 1994.

[27] M. Effros, A. Goldsmith, and Y. Liang, "Generalizing capacity: New definitions and capacity theorems for composite channels," *Information Theory, IEEE Transactions on*, vol. 56, no. 7, pp. 3069 –3087, 2010.

[28] M. Katz and S. Shamai, "Cooperative schemes for a source and an occasional nearby relay in wireless networks," *Information Theory, IEEE Transactions on*, vol. 55, pp. 5138–5160, nov. 2009.

[29] A. Behboodi and P. Piantanida, "On the simultaneous relay channel with informed receivers," in *IEEE International Symposium on Information Theory (ISIT 2009)*, pp. 1179–1183, 28 June-July 3 2009.

[30] Y. Liang and G. Kramer, "Rate regions for relay broadcast channels," *Information Theory, IEEE Transactions on*, vol. 53, pp. 3517–3535, Oct. 2007.

[31] C. Nair and A. El Gamal, "An outer bound to the capacity region ofthe broadcast channel," *Information Theory, IEEE Transactions on*, vol. 53, pp. 350–355, jan. 2007.

[32] C. Nair, "A note on outer bounds for broadcast channel." http://arxiv.org/abs/1101.0640v1, jan 2011.

[33] S. Bhaskaran, "Gaussian degraded relay broadcast channel," *Information Theory, IEEE Transactions on*, vol. 54, pp. 3699–3709, Aug. 2008.

[34] M. Katz and S. Shamai, "Transmitting to colocated users in wireless ad hoc and sensor networks," *Information Theory, IEEE Transactions on*, vol. 51, pp. 3540–3563, Oct. 2005.

[35] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Trans. Information Theory*, vol. 44, pp. 2148–2177, October 1998.

[36] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *Information Theory, IEEE Transactions on*, vol. 44, pp. 2619 –2692, oct 1998.

[37] A. Behboodi and P. Piantanida, "Broadcasting over the relay channel with oblivious cooperative strategy," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, pp. 1098 –1103, 29 2010-oct. 1.

[38] A. Behboodi and P. Piantanida, "On the asymptotic error probability of composite relay channels," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, 2011.

[39] M. Hayashi, "Information spectrum approach to second-order coding rate in channel coding," *Information Theory, IEEE Transactions on*, vol. 55, pp. 4947 –4966, nov. 2009.

[40] A. M. M. El Gamal and S. Zahedi, "Bounds on capacity and minimum energy-per-bit for AWGN relay channels," *IEEE Trans. Information Theory*, vol. 52, pp. 1545–1561, Apr. 2006.

[41] G. Kramer and J. Hou, "On message lengths for noisy network coding," in *IEEE Information Theory Workshop*, pp. 430 –431, October 2011.

[42] X. Wu and L.-L. Xie, "On the optimal compressions in the compress-and-forward relay schemes," Feb. 2011.

[43] T. S. Han, *Information-Spectrum Methods in Information Theory.* Springer-Verlag, 2003.

[44] S. Verdu and T. S. Han, "A general formula for channel capacity," *Information Theory, IEEE Transactions on*, vol. 40, pp. 1147 –1157, July 1994.

[45] R. Ahlswede, "The weak capacity of averaged channel," *Wahtscheinlichkeitstheorie und Vrew. Gebiete*, vol. 11, pp. 61 –73, 1968.

[46] J. Kieffer, "ϵ-capacity of binary symmetric averaged channels," *Information Theory, IEEE Transactions on*, vol. 53, pp. 288 –303, jan. 2007.

[47] T. Cover and J. Thomas, *Elements of Information Theory.* Wiley Series in Telecomunications, Wiley & Sons New York, 1991.

[48] N. Ratnakar and G. Kramer, "The multicast capacity of deterministic relay networks with no interference," *Information Theory, IEEE Transactions on*, vol. 52, pp. 2425 – 2432, june 2006.

[49] C. E. Shannon, “A mathematical theory of communication,” *Bell System Tech. J.*, vol. 27, pp. 379 – 423, 623 – 656, 1948.

[50] C. E. Shannon, “Two-way communication channels,” *Proc. Fourth Berkeley Symp. on Math. Statist. and Prob.*, vol. 1, pp. 611 – 644, 1961.

[51] R. Ahlswede, “Multi-way communication channels,” *in B.N. Petrov (Ed.), Second International Symposium on Information Theory: Tsahkadsor, Armenia, USSR, Akadémiai Kiadó, Budapest*, pp. 23–51, Sept. 1971.

[52] H. Liao, “Multiple access channels,” *Ph.D. thesis, Department of Electrical Engineering, University of Hawaii, Honolulu*, 1972.

[53] T. M. Cover, “Open problems in information theory,” *In Proceedings of the IEEE–USSR Joint Workshop in Information Theory, Moscow*, pp. 35–36, 1975.

[54] R. Ahlswede, “Open problems in communication and computation,” *Contributions to “”, T.M. Cover and B. Gopinath (Ed.), Springer-Verlag.*, pp. 35–36, 1987.

[55] E. C. van der Meulen, “Three-terminal communication channels,” *Adv. Appl. Prob.*, vol. 3, pp. 120–154, 1971.

[56] C. Politis, T. Oda, S. Dixit, A. Schieder, H.-Y. Lach, M. Smirnov, S. Uskela, and R. Tafazolli, “Cooperative networks for the future wireless world,” *Communications Magazine, IEEE*, vol. 42, pp. 70–79, sept. 2004.

[57] D. Soldani and S. Dixit, “Wireless relays for broadband access [radio communications series],” *Communications Magazine, IEEE*, vol. 46, pp. 58–66, march 2008.

[58] R. Pabst, B. Walke, D. Schultz, P. Herhold, H. Yanikomeroglu, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. Falconer, and G. Fettweis, “Relay-based deployment concepts for wireless and mobile broadband radio,” *Communications Magazine, IEEE*, vol. 42, pp. 80–89, sept. 2004.

[59] J. Laneman, D. Tse, and G. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *Information Theory, IEEE Transactions on*, vol. 50, pp. 3062–3080, Dec. 2004.

[60] A. Sendonaris, E. Erkip, and B. Aazhang, “User cooperation diversity. part i. system description,” *Communications, IEEE Transactions on*, vol. 51, pp. 1927–1938, nov. 2003.

[61] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. part ii. implementation aspects and performance analysis," *Communications, IEEE Transactions on*, vol. 51, pp. 1939–1948, nov. 2003.

[62] Y. T. J. Fan, "Cmimo configurations for relay channels: Theory and practice," *Wireless Communications, IEEE Transactions on*, vol. 6, pp. 1774–1786, May 2007.

[63] T. Hunter, S. Sanayei, and A. Nosratinia, "Outage analysis of coded cooperation," *Information Theory, IEEE Transactions on*, vol. 52, pp. 375–391, feb. 2006.

[64] R. C. King, "Multiple access channels with generalized feedback," *Ph.D. dissertation, Stanford Univ., Stanford, CA*, Mar. 1978.

[65] C.-M. Zeng, F. Kuhlmann, and A. Buzo, "Achievability proof of some multiuser channel coding theorems using backward decoding," *Information Theory, IEEE Transactions on*, vol. 35, pp. 1160–1165, nov 1989.

[66] J. Laneman and G. Kraner, "Window decoding for the multiaccess channel with generalized feedback," *Information Theory, 2004. ISIT 2004. Proceedings. International Symposium on*, p. 281, june-2 july 2004.

[67] H.-F. Chong, M. Motani, and H. K. Garg, "Generalized backward decoding strategies for the relay channel," *Information Theory, IEEE Transactions on*, vol. 53, pp. 394–401, jan. 2007.

[68] M. R. Aref, "Information flow in relay networks," *Ph.D. dissertation, Stanford Univ., Stanford, CA*, 1980.

[69] Z. Zhang, "Partial converse for a relay channel," *Information Theory, IEEE Transactions on*, vol. 34, pp. 1106–1110, Sep 1988.

[70] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *Information Theory, IEEE Transactions on*, vol. 57, pp. 6463 –6486, oct. 2011.

[71] L.-L. Xie and P. Kumar, "An achievable rate for the multiple-level relay channel," *Information Theory, IEEE Transactions on*, vol. 51, pp. 1348–1358, april 2005.

[72] M. Gastpar and M. Vetterli, "On the capacity of large gaussian relay networks," *Information Theory, IEEE Transactions on*, vol. 51, pp. 765–779, March 2005.

[73] Y. Liang and V. Veeravalli, "Gaussian orthogonal relay channels: Optimal resource allocation and capacity," *Information Theory, IEEE Transactions on*, vol. 51, pp. 3284–3289, Sept. 2005.

[74] B. Schein and R. G. Gallagar, "The gaussian parallel relay network," *Proc IEEE Int. Symp. Info. Theory*, p. 22, Sorrento, Italy, Jun. 2000.

[75] C. W. Sung, P. Hu, and K. Leung, "The multiuser gaussian parallel relay network," *Information Theory and Its Applications, 2008. ISITA 2008. International Symposium on*, pp. 1–6, dec. 2008.

[76] S. Rezaei, S. Gharan, and A. Khandani, "A new achievable rate for the gaussian parallel relay channel," *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*, pp. 194–198, 28 2009-july 3 2009.

[77] Y. Kochman, A. Khina, U. Erez, and R. Zamir, "Rematch and forward for parallel relay networks," *Information Theory, 2008. ISIT 2008. IEEE International Symposium on*, pp. 767–771, july 2008.

[78] I. Maric and R. Yates, "Forwarding strategies for gaussian parallel-relay networks," *Information Theory, 2004. ISIT 2004. Proceedings. International Symposium on*, p. 269, june-2 july 2004.

[79] A. Host-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *Information Theory, IEEE Transactions on*, vol. 51, pp. 2020–2040, june 2005.

[80] Y. Liang, V. Veeravalli, and H. Poor, "Resource allocation for wireless fading relay channels: Max-min solution," *Information Theory, IEEE Transactions on*, vol. 53, pp. 3432–3453, oct. 2007.

[81] C. Ng and A. Goldsmith, "The impact of csi and power allocation on relay channel capacity and cooperation strategies," *Wireless Communications, IEEE Transactions on*, vol. 7, pp. 5380–5389, december 2008.

[82] L. Lai, K. Liu, and H. El Gamal, "The three-node wireless network: achievable rates and cooperation strategies," *Information Theory, IEEE Transactions on*, vol. 52, pp. 805–828, march 2006.

[83] A. E. Gamal and S. Zahedi, "Capacity of a class of relay channels with orthogonal components," *IEEE Trans. Information Theory*, vol. 51, pp. 1815–1818, May 2005.

[84] R. Tannious and A. Nosratinia, "Relay channel with private messages," *Information Theory, IEEE Transactions on*, vol. 53, pp. 3777–3785, oct. 2007.

[85] M. Aleksic, P. Razaghi, and W. Yu, "Capacity of a class of modulo-sum relay channels," *Information Theory, IEEE Transactions on*, vol. 55, no. 3, pp. 921 –930, 2009.

[86] G. Kramer and A. van Wijngaarden, "On the white gaussian multiple-access relay channel," in *Information Theory, 2000. Proceedings. IEEE International Symposium on*, pp. 40–, 2000.

[87] L. Sankar, N. Mandayam, and H. Poor, "On the sum-capacity of degraded gaussian multiple-access relay channels," *Information Theory, IEEE Transactions on*, vol. 55, pp. 5394–5411, Dec. 2009.

[88] L.-L. Xie and P. Kumar, "Multisource, multidestination, multirelay wireless networks," *Information Theory, IEEE Transactions on*, vol. 53, pp. 3586–3595, Oct. 2007.

[89] L. Sankar, G. Kramer, and N. Mandayam, "Offset encoding for multiple-access relay channels," *Information Theory, IEEE Transactions on*, vol. 53, pp. 3814–3821, oct. 2007.

[90] Y. Liang and V. Veeravalli, "The impact of relaying on the capacity of broadcast channels," in *Information Theory, 2004. ISIT 2004. Proceedings. International Symposium on*, pp. 403–, June-2 July 2004.

[91] R. Dabora and S. Servetto, "On the rates for the general broadcast channel with partially cooperating receivers," in *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on*, pp. 2174–2178, Sept. 2005.

[92] A. Reznik, S. Kulkarni, and S. Verdu, "Broadcast-relay channel: capacity region bounds," in *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on*, pp. 820–824, Sept. 2005.

[93] R. Dabora and S. Servetto, "Broadcast channels with cooperating decoders," *Information Theory, IEEE Transactions on*, vol. 52, pp. 5438–5454, Dec. 2006.

[94] Y. Liang and V. V. Veeravalli, "Cooperative relay broadcast channels," *Information Theory, IEEE Transactions on*, vol. 53, pp. 900–928, March 2007.

[95] S. I. Bross, "On the discrete memoryless partially cooperative relay broadcast channel and the broadcast channel with cooperative decoders," *IEEE Trans. Information Theory*, vol. IT-55, pp. 2161–2182, May 2009.

[96] O. Simeone, D. Gündüz, and S. Shamai, "Compound relay channel with informed relay and destination," in *47th Annual Allerton Conference on Comm., Control, and Computing, Monticello, IL*, Sept. 2009.

[97] A. Behboodi and P. Piantanida, "Capacity of a class of broadcast relay channels," in *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, pp. 590–594, 2010.

[98] A. Steiner and S. Shamai, "Single user broadcasting protocols over a two hop relay fading channel," *IEEE Trans. Information Theory*, vol. IT-52, pp. 4821–4838, Nov 2006.

[99] M. Katz and S. Shamai, "Oblivious cooperation in colocated wireless networks," *Information Theory, 2006 IEEE Int. Symp. on*, pp. 2062–2066, July 2006.

[100] S. Shamai and A. Steiner, "A broadcast approach for a single-user slowly fading MIMO channel," *Information Theory, IEEE Transactions on*, vol. 49, pp. 2617–2635, Oct. 2003.

[101] D. B. L. Blackwell and A. J. Thomasian, "The capacity of a class of channels," *Ann. Math. Stat.*, vol. 31, pp. 558–567, 1960.

[102] J. R. K. Nayak, "Graph capacities and zero-error transmission over compound channels," *Information Theory, IEEE Transactions on*, vol. 51, pp. 4374–4378, Dec. 2005.

[103] P. Mitran, N. Devroye, and V. Tarokh, "On compound channels with side information at the transmitter," *Information Theory, IEEE Transactions on*, vol. 52, pp. 1745–1755, April 2006.

[104] A. P. V. Raja and P. Viswanath, "The two-user compound interference channel," *Information Theory, IEEE Transactions on*, vol. 55, pp. 5100–5120, Nov. 2009.

[105] S. Denic, C. Charalambous, and S. Djouadi, "Information theoretic bounds for compound mimo gaussian channels," *Information Theory, IEEE Transactions on*, vol. 55, pp. 1603–1617, April 2009.

[106] A. T. I. Lapidoth, "The compound channel capacity of a class of finite-state channels," *Information Theory, IEEE Transactions on*, vol. 44, pp. 973–983, May 1998.

[107] O. G. D. Simeone, H. G. A. Poor, and S. Shamai, "Compound multiple-access channels with partial cooperation," *Information Theory, IEEE Transactions on*, vol. 55, pp. 2425–2441, June 2009.

[108] B. P. H. Shrader, "Feedback capacity of the compound channel," *Information Theory, IEEE Transactions on*, vol. 55, pp. 3629–3644, Aug. 2009.

[109] P. Piret, "Binary codes for compound channels (corresp.)," *Information Theory, IEEE Transactions on*, vol. 31, pp. 436–440, May 1985.

[110] H. Weingarten, T. Liu, S. Shamai, Y. Steinberg, and P. Viswanath, "The capacity region of the degraded multiple-input multiple-output compound broadcast channel," *Information Theory, IEEE Transactions on*, vol. 55, pp. 5011–5023, Nov. 2009.

[111] A. Gamal, "The capacity of the physically degraded gaussian broadcast channel with feedback (corresp.)," *Information Theory, IEEE Transactions on*, vol. 27, pp. 508–511, jul 1981.

[112] A. Gamal, "The feedback capacity of degraded broadcast channels (corresp.)," *Information Theory, IEEE Transactions on*, vol. 24, pp. 379–381, may 1978.

[113] L. Ozarow and S. Leung-Yan-Cheong, "An achievable region and outer bound for the gaussian broadcast channel with feedback (corresp.)," *Information Theory, IEEE Transactions on*, vol. 30, pp. 667–671, jul 1984.

[114] T. M. Cover, "Comments on broadcast channels," *IEEE Trans. Information Theory*, vol. 44, pp. 2524–2530, Oct. 1998.

[115] S. Gelfand and M. Pinsker, "Capacity of a broadcast channel with one deterministic component," *Probl. Inform. Transm.*, vol. 16, pp. 17–25, Jan.-Mar. 1980.

[116] M. Pinsker, "Capacity of noiseless broadcast channels," *Probl. Inform. Transm.*, vol. 14, pp. 28–334, Apr.-Jun. 1978.

[117] T. Han, "The capacity region for the deterministic broadcast channel with a common message (corresp.)," *Information Theory, IEEE Transactions on*, vol. 27, pp. 122–125, jan 1981.

[118] C. Nair and A. El Gamal, "The capacity region of a class of 3-receiver broadcast channels with degraded message sets," *Information Theory, 2008. ISIT 2008. IEEE International Symposium on*, pp. 1706–1710, july 2008.

[119] C. Nair and A. E. Gamal, "The capacity region of a class of 3-receiver broadcast channels with degraded message sets," *IEEE Trans. Information Theory*, vol. IT-55, pp. 4479–4493, October 2009.

[120] C. Nair and A. El Gamal, "An outer bound to the capacity region of the broadcast channel," *Information Theory, 2006 IEEE International Symposium on*, pp. 2205–2209, july 2006.

[121] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung, "Network information flow," *Information Theory, IEEE Transactions on*, vol. 46, pp. 1204 –1216, jul 2000.

[122] A. Dana, R. Gowaikar, R. Palanki, B. Hassibi, and M. Effros, "Capacity of wireless erasure networks," *Information Theory, IEEE Transactions on*, vol. 52, pp. 789 –804, march 2006.

[123] A. Goldsmith and M. Medard, "Capacity of time-varying channels with causal channel side information," *Information Theory, IEEE Transactions on*, vol. 53, no. 3, pp. 881 –899, 2007.

[124] P. Piantanida, G. Matz, and P. Duhamel, "Outage behavior of discrete memoryless channels under channel estimation errors," *Information Theory, IEEE Transactions on*, vol. 55, no. 9, pp. 4221 –4239, 2009.

[125] A. Feinstein, "A new basic theorem on information theory," *Information Theory, IRE Transactions on*, vol. 4, pp. 2 –22, 1954.

[126] S. Arimoto, "On the converse to the coding theorem for discrete memoryless channels (corresp.)," *Information Theory, IEEE Transactions on*, vol. 19, pp. 357 – 359, may 1973.

[127] R. G. Gallager, *Information Theory and Reliable Communication.* Wiley & Sons New York, 1968.

[128] J. Wolfowitz, *Coding Theorems of Information Theory.* Wiley & Sons New York, 1964.

[129] R. Ash, "A simple example of a channel for which the strong converse fails (corresp.)," *Information Theory, IEEE Transactions on*, vol. 11, pp. 456 – 457, jul 1965.

[130] J. Korner and K. Marton, "Images of a set via two channels and their role in multiuser communication," *Information Theory, IEEE Transactions on*, vol. 23, pp. 751 – 761, nov 1977.

[131] T. S. Han, "An information-spectrum approach to capacity theorems for the general multiple-access channel," *Information Theory, IEEE Transactions on*, vol. 44, pp. 2773 –2795, nov 1998.