Secure Communication in Erasure Networks with State-feedback

THÈSE N^{O} 6303 (2014)

PRÉSENTÉE LE 28 AOÛT 2014

À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS
LABORATOIRE D'ALGORITHMIQUE POUR L'INFORMATION EN RÉSEAUX
PROGRAMME DOCTORAL EN INFORMATIQUE, COMMUNICATIONS ET INFORMATION

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

László CZAP

acceptée sur proposition du jury:

Prof. M. Grossglauser, président du jury
Prof. C. Fragouli, directrice de thèse
Prof. S. Diggavi, rapporteur
Prof. V. M. Prabhakaran, rapporteur
Prof. E. Telatar, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE
Suisse
2014

Riassunto

Sicurezza ed efficienza delle comunicazioni sono due delle principali preoccupazioni nelle reti di oggi e lo saranno in quelle del futuro. La comprensione di come inviare dati efficacemente tramite diversi canali e reti è stata significativamente approfondita nell'ultima decade (vedi per esempio [1–3]). La comprensione di come inviare informazioni in maniera sicura, invece, non ha ancora raggiunto lo stesso livello. Questa tesi contribuisce ad approfondire la comprensione di alcuni aspetti della teoria delle comunicazioni sicure derivando risultati riguardo la capacità di trasmissione segreta per reti con cancellazioni e sviluppando codifiche che garantiscono sicurezza incondizionata (basata sulla teoria dell'informazione) per questo tipo di reti.

La capacità di trasmissione segreta raggiungibile in presenza di un avversario che intercetta la comunicazione viene derivata in vari scenari. In questi scenari la comunicazione avviene tramite canali con cancellazioni e con informazioni di ritorno sullo stato. I risultati presentati forniscono una caratterizzazione per il canale punto-punto, per il canale broadcast con più riceventi, per la rete con canali paralleli, per la rete a V e per la rete triangolare.

Si presenta inoltre uno schema di codifica a due fasi che consiste in una fase di generazione di chiavi e una fase in cui i messaggi criptati vengono inviati. Questo schema sfrutta numerose risorse per garantire la sicurezza: le cancellazioni del canale, l'informazione casuale condivisa e la topologia della rete. Vengono presentati schemi di codifica per gli scenari menzionati nel precedente paragrafo e per reti con cancellazioni con topologia arbitraria. Per tutti i casi per i quali è presentata una caratterizzazione esatta uno schema a due fasi raggiunge la capacità di trasmissione segreta. Tutti gli schemi di codifica proposti usano operazioni lineari e perciò possono servire da base per realizzare dei codici usabili in pratica.

Per le reti viene sviluppata una base di analisi che può essere usata per descrivere schemi di codifica sicuri e per derivare nuove maggioranti esterne. Gli schemi presentati sono descritti, e la loro ottimalità viene provata, usando programmi lineari.

Sono presentate inoltre nuove maggioranti esterne basate sulla teoria dell'informazione. Le dimostrazioni presentate, intuitivamente, trovano una relazione tra velocità di trasmissione dei messaggi e velocità di generazione di della chiave segreta che viene usata per rendere sicuro il messaggio.

I risultati presentati rivelano caratteristiche non triviali della comunicazione sicura in reti con cancellazioni. Si trova che – in contrasto con la comunicazione non sicura – la capacità di

trasmissione sicura di un taglio non può essere ridotta per semplificazione alla somma delle capacità dei canali che formano il taglio. Inoltre la capacità di trasmissione segreta di una rete non può essere ridotta per semplificazione alla capacità di trasmissione segrete minima dei suoi tagli.

Parole chiave: reti con cancellazioni, confidenzialità, segretezza, sicurezza basata sulla teoria dell'informazione, capacità di trasmissione segreta, generazione di chiave segreta

Abstract

The security and efficiency of communication are two of the main concerns for networks of today and the future. Our understanding of how to efficiently send information over various channels and networks has significantly increased in the past decade (see e.g., [1–3]), whereas our understanding of how to *securely* send information has not yet reached the same level. In this thesis, we advance the theory of secure communication by deriving capacity results and by developing coding schemes that provide information-theoretic security for erasure networks.

We characterize the highest achievable secret-message rate in the presence of an eavesdropping adversary in various settings, where communication takes place over erasure channels with state-feedback. Our results provide such a characterization for a point-to-point erasure channel, for a broadcast erasure channel with multiple receivers, for a network with multiple parallel channels, a V-network and for a triangle network.

We introduce several two-phase secure coding schemes that consist of a key generation phase and an encrypted message sending phase. Our schemes leverage several resources for security: channel erasures, feedback, common randomness and the topology of the network. We present coding schemes for all the above mentioned settings as well as for erasure networks with arbitrary topology. In all the cases where we provide exact characterization, a two-phase scheme achieves the secret-message capacity. All our proposed coding schemes use only linear operations and thus can serve as a basis for practical code designs.

For networks, we develop a linear programming framework for describing secure coding schemes and for deriving new outer bounds. We use linear programs to describe our schemes and to prove their optimality.

We derive new information theoretic outer bounds. In our intuitive interpretation, our proofs find the connection between the rate of the message and the rate of a secret key that is required to secure the message.

Our results reveal nontrivial characteristics of secure communication in erasure networks. We find that – in contrast to non-secure communication – the secret message capacity of a cut does not simplify to the sum of the capacities of the channels that form the cut, moreover, the secret message capacity of a network does not simplify to the minimum secret message capacity of its cuts.

Key words: erasure networks, confidentiality, secrecy, information theoretic security, secret message capacity, secret key generation

To my wife…

Acknowledgements

Writing this thesis is a major milestone in my life and in my career. I would have not been able to reach this milestone alone, hence it is most appropriate to devote the first few paragraphs to those who were beside me along this path. I am grateful for all the support – of various form – that I received from them throughout the years.

Foremost, I am indebted to my supervisor, Christina Fragouli. Working together with her has been a source of lot of inspiration both professionally and personally. Her guidance helped me a lot to become a more mature researcher in the way I think about problems, I present results and collaborate with others. I continuously enjoyed her support, sometimes she had more faith in what I was doing than myself, which was a great help in getting over tough periods. I am grateful to her for facilitating working together with other great researchers and for ensuring the possibility of attending workshops and conferences. Her responsible and supportive personality enabled me to focus on research and made my years as a Ph.D. student a fruitful period.

The guidance I received from Christina was supplemented by my close collaborators, Suhas Diggavi and Vinod Prabhakaran. Although physically we were far apart most of the time, I benefited a lot from working together with them. I am especially grateful to Vinod for hosting me as a summer intern in Mumbai. I also had the possibility to spend some time in Los Angeles working closely together with Suhas. These two periods were among the most productive periods of my research.

As a teaching assistant I had the chance to work together with exceptional professors at EPFL: Emre Telatar, Bixio Rimoldi, Peter Wittwer, Olivier Lévêque and Anastasia Ailamaki. My benefit of these collaborations were at least as much as my help in running the various courses.

I thank Francoise Behn for taking care of all administrative issues, organizing travels and for being very helpful every time I had questions or problems. I also thank Damir Laurenzi and Giovanni Cangiani for supporting my access to the IT infrastructure I needed.

The group ARNI provided me a friendly and calm environment. I had the chance to discuss, share ideas and get inspired by others – not only on research related issues. I owe a special thanks to Iris Safaka, with whom I shared the office for three years. Beside her, I received especially lot of help, support and enjoyable lunchtime discussions from Emre Atsan, Lorenzo Keller and Marios Gatzianas. I also thank Lorenzo for translating my abstract into Italian. I

would like to thank all former and current members of ARNI whom I am lucky enough to know: Siddhartha Brahma, Ayan Sengupta, Javad Ebrahimi, Mahdi Jafari Siavoshani, I-Hsiang Wang, Melissa Duarte, Shirin Saeedi Bidhokhti and Ayfer Özgür. I also need to thank Athanasios Papadopoulos from UCLA for being one of the most careful readers of my papers.

My research at EPFL received financial support from the ERC Starting Grant Project NOWIRE ERC-2009-StG-240317, which I appreciate a lot.

I started my years in research at the Budapest University of Technology and Economics (BUTE) in the CrySys laboratory under the supervision of István Vajda and in close collaboration with Levente Buttyán. I thank them for guiding and supporting my first steps as a researcher. I learned the first few lessons about conducting research and academic writing from them. Apart from them, I would like to mention my mentor in my undergraduate years at BUTE, László Zömbik. His rigorous and scientific thinking influenced my approach to problems and my decision to start research. I would like to mention my former colleagues in the CrySys laboratory with whom I spent a memorable period: Boldizsár Bencsáth, Gergely Ács, László Dóra, Péter Schaffer, Tamás Holczer and Ta Vinh Thong.

I have to express my greatest gratitude to my family, especially to my wife, Villő and to my son, Vince. The continuous love and support from Villő were essential for me throughout the years. I am grateful to her for taking a lot of burden from me which helped me to focus and to be more efficient in work. While I was enjoying conferences and internships abroad, it was she who had to handle loneliness and to take care of our < 3 years old child. I appreciate and thank her for all the sacrifices that she has offered. My family, and especially Vince brought the most joy and fun to my life during this period. They make my life and work meaningful.

I receive continuous support from my broader family, my parents (Sarolta and László Czap, but for me Anya and Apa) and my sister, Sarolta. Whatever I become or achieve finds its roots at them.

Last, but not least I am grateful to all friends and family members who offered support. I especially thank my wife's family and our closest friends who were always happy to meet (and sometimes also host) us whenever we visited and managed to maintain our relationship despite of the physical distance.

Lausanne, 12 June 2014

Contents

Abstract (Italian/English) v

Acknowledgements ix

List of figures xiv

List of tables xv

Introduction 1

1 Model and background 5
- 1.1 Communication model . . . 5
  - 1.1.1 Erasure channel with state-feedback . . . 5
- 1.2 Secret-message sending . . . 7
- 1.3 Secret-key generation . . . 9
- 1.4 Adversary model and security notions . . . 10
  - 1.4.1 Adversary . . . 10
  - 1.4.2 Information theoretic security notions . . . 11
  - 1.4.3 Equivalence of security notions . . . 12
- 1.5 Key generation in a point-to-point setting . . . 13

2 Secret-message capacity of a point-to-point channel 17
- 2.1 Related work . . . 18
- 2.2 Model . . . 19
- 2.3 Main result . . . 19
  - 2.3.1 Discussion . . . 19
- 2.4 Coding scheme . . . 21
  - 2.4.1 Principles, example . . . 21
  - 2.4.2 Detailed description . . . 23
  - 2.4.3 Analysis . . . 24
- 2.5 Outer bound . . . 26
  - 2.5.1 Proofs of Lemmas 2.1-2.2 . . . 27
- 2.6 Linear programming formulation . . . 29
- 2.7 Next steps . . . 30

3 Secret-message capacity of a broadcast channel **31**
3.1 Related work . . . 32
3.2 Model . . . 32
3.2.1 Honest-but-curious adversary . . . 32
3.2.2 Dishonest adversary . . . 33
3.3 Non-secure 1-to-M broadcast . . . 36
3.4 Main results . . . 37
3.4.1 Honest-but-curious adversary . . . 37
3.4.2 Dishonest adversary . . . 38
3.5 Coding scheme . . . 40
3.5.1 Honest-but-curious receivers . . . 40
3.5.2 Dishonest adversary . . . 44
3.5.3 Distribution independent scheme . . . 49
3.6 Outer Bound . . . 50
3.6.1 Proof of Theorem 3.2 . . . 51
3.6.2 Interpretation of the converse proof . . . 51
3.6.3 Proofs of Lemmas 3.5-3.8 . . . 52
3.7 Next steps . . . 55

4 Secret-message capacity in networks **57**
4.1 Related Work . . . 58
4.2 Parallel channels . . . 58
4.2.1 Model . . . 58
4.2.2 Main result . . . 59
4.2.3 Coding scheme . . . 60
4.2.4 Outer bound . . . 62
4.3 V-network . . . 67
4.3.1 Model . . . 68
4.3.2 Main result . . . 68
4.3.3 Coding scheme . . . 69
4.3.4 Outer bound . . . 73
4.4 Triangle network . . . 77
4.4.1 Model . . . 77
4.4.2 Main result . . . 78
4.4.3 Coding scheme . . . 80
4.4.4 Analysis . . . 83
4.4.5 Outer bound . . . 85
4.5 Next steps . . . 85

5 Secret-message sending in arbitrary networks **87**
5.1 Related work . . . 88
5.2 Model . . . 88
5.3 Secure network coding over error-free networks . . . 89

5.4 Main result . . . 89
5.4.1 Discussion, numerical examples . . . 91
5.5 Two-phase secure network coding scheme . . . 92
5.5.1 Example . . . 94
5.5.2 Scheme description . . . 94
5.5.3 Discussion . . . 95
5.6 Coding scheme . . . 96
5.6.1 Example . . . 96
5.6.2 Algorithm . . . 97
5.6.3 Multicast . . . 98
5.6.4 Analysis . . . 99
5.7 Outer bounds . . . 100
5.8 Discussion . . . 100
5.8.1 Extension for $z > h$ . . . 100
5.8.2 Intermediate randomness helps . . . 101
5.8.3 Unicast rate $\overset{?}{\geq}$ multicast rate? . . . 101

6 Discussion and open problems **103**

A Summary of often used tools **105**
A.1 One-time pad encryption . . . 105
A.2 Chernoff-Hoeffding bound . . . 106
A.3 MDS matrices . . . 106

B Proofs for Chapter 3 **109**
B.1 Proof of Lemma 3.3 . . . 109
B.2 Proof of Lemma 3.4 . . . 110
B.3 Rate calculation . . . 110
B.3.1 Honest-but-curious adversary . . . 110
B.3.2 Dishonest adversary . . . 111
B.4 Proof of distribution independent security . . . 112
B.5 Proof of Lemma 3.1 . . . 112

C Proofs and calculations for Chapter 4 **117**
C.1 Calculating N_i . . . 117
C.2 Proofs of Lemmas 4.5-4.6 . . . 117
C.2.1 Proof of Lemma 4.5 . . . 117
C.2.2 Proof of Lemma 4.6 . . . 118
C.3 V-network outer bound proof . . . 121
C.3.1 Side calculation . . . 133
C.4 Triangle network outer bound proof . . . 136
C.4.1 Rate constraints . . . 136
C.4.2 Security constraints . . . 136

C.4.3 Time-sharing constraints . . . 137
C.4.4 Distinguishing keys . . . 138
C.4.5 Connecting cuts . . . 139
C.4.6 Trivial constraints . . . 141
C.4.7 Equivalence of LPs . . . 142
C.4.8 Side-calculations . . . 172

D Proofs for Chapter 5 **183**
D.1 Security of two-phase secure network coding . . . 183
D.2 Proof of Theorem 5.1 . . . 184
D.3 Proof of Theorem 5.2 . . . 186
D.3.1 Proof of Lemma D.2 . . . 187
D.4 Proof of Theorem 5.3 . . . 189
D.4.1 Proof of Lemma D.3 . . . 190
D.4.2 Proof of Lemma D.4 . . . 191

Bibliography **193**

Curriculum Vitae **199**

List of Figures

1 Communication settings with a complete characterization. 3

1.1 Example key generation 14

2.1 Point-to-point channel: secret-message and secret-key capacity, $\delta = 0.4$ 20
2.2 Point-to-point channel: secret-message and secret-key capacity, $\delta = \delta_E$ 21

3.1 Non-secure message sending and secret-message sending capacity regions . . 38
3.2 Distribution independent secret-message rate region 39

4.1 Our networks. 57
4.2 Parallel channels: secret-message capacity 60
4.3 V-network: role of common randomness 69
4.4 V-network: capacity and time-sharing rate 70
4.5 Triangle network 77
4.6 Comparison of secret-message rates with/without exploiting erasures and with-/without feedback. 79
4.7 Comparison of secret-message capacities with/without private randomness at U. 80

5.1 Advantage of using feedback as a function of the number of eavesdropped edges 92
5.2 Upper bound/achieved rate (based on Theorem 5.2) 93
5.3 Upper bound/achieved rate (based on Theorems 5.2-5.3) 93
5.4 Secure network coding example 94
5.5 Coding with shared key 94
5.6 Two-phase secure network coding scheme example 94
5.7 Example network – triangle topology 96
5.8 Two-hop line network 100
5.9 Multicast problem example 102

D.1 Transformed network $\mathscr{G}'$ 187
D.2 Transformed network $\mathscr{G}''$ 189

List of Tables

2.1 Example scheme with Eve's channel state known 22

3.1 An example of the protocol run. 41

4.1 Coding across packets from the common randomness 70

Introduction

The security and efficiency of communication are two of the main concerns for networks of today and the future. Our understanding of how to efficiently send information over various channels and networks has significantly increased in the past decade (see e.g., [1–3]), whereas our understanding of how to *securely* send information has not yet reached the same level. In this thesis, we advance the theory of secure communication by deriving capacity results and by developing coding schemes that provide information-theoretic security for erasure networks.

Information-theoretic security can augment the current technology that relies on computational security. The common model of an adversary that protocol designers consider assumes that the adversary has full access to all communications, but that its computational capability is bounded. In many scenarios, however, the adversary has only a limited access to the communication between honest parties. Every communication medium – especially a wireless channel – is noisy, which interferes with the adversary's observation; moreover, in a large network, the adversary is very unlikely to be able to control all the communication channels. Information theory offers a framework for exploring how and to what extent we can benefit from the limited communication presence of the adversary.

In this thesis, we exploit the limited network presence of the adversary to design schemes that can lead to practical protocols. Early results on information-theoretic secrecy were far from practical. For perfect secrecy, not only a secret key has to be pre-shared, but the size of the key also has to be at least the size of the message to be encrypted (Shannon [4][1]). Although Shannon's operation of encryption was extremely simple, no solution for key sharing was available. In erasure networks, however, we can design efficient algorithms that exploit erasures to create a key [5]. We also show that a key size that equals the number of message packets that Eve observes (as opposed to the size of the whole message) is sufficient for security. These contributions enable us to make steps to circumvent both of the main practical issues: we develop algorithms for setting up a shared key and we design encryption that efficiently uses the secret key.

We make use of several network resources simultaneously to design our secure coding schemes. The topology of the network, the erasures that the adversary experiences, the feedback from

[1] For more details on one-time pad encryption, we refer the reader to Appendix A.1 and to [4].

honest parties and the common randomness that network nodes share are all valuable resources that can support secure communication. We investigate the role of these resources and design coding schemes that optimally use them simultaneously.

We highlight that, unlike previous works, we assume only state-feedback. It is well known that feedback does not increase the achievable rate over a single-receiver discrete memoryless channel [6], however it does increase the achievable secret-key and secret-message rate[2], even if the feedback is public (as it was first shown by Maurer [7]). In particular, having feedback enables us to achieve a positive secret-message rate even if an eavesdropper has an observation less noisy than that of the intended receiver. We consider the case when the feedback is limited to the channel state in contrast with e.g., [5,7–9], where an unlimited public discussion channel is assumed. This enables us to make a clear distinction between the secret-key generation and the secret-message sending problem (for more, see Section 1.3), while it makes our model more realistic. Indeed, acknowledgments are part of many communication standards, hence such state-feedback is a resource already available.

We believe that the erasure channel (with state-feedback) is a good starting point for investigating security in networks. This model – although simpler than a general communication model – is suitable for capturing the intricacies and possibilities of operating in a wireless network environment. The applicability of the erasure model is justified through experiments that artificially create erasure channels by injecting interference [10, 11] and through results showing the relevance of an erasure model for a state-dependent Gaussian channel [12]. Thus, results on secure communication in the erasure model are, in themselves, relevant in practice, and they also serve as a first step towards solving the problem with a more general channel model.

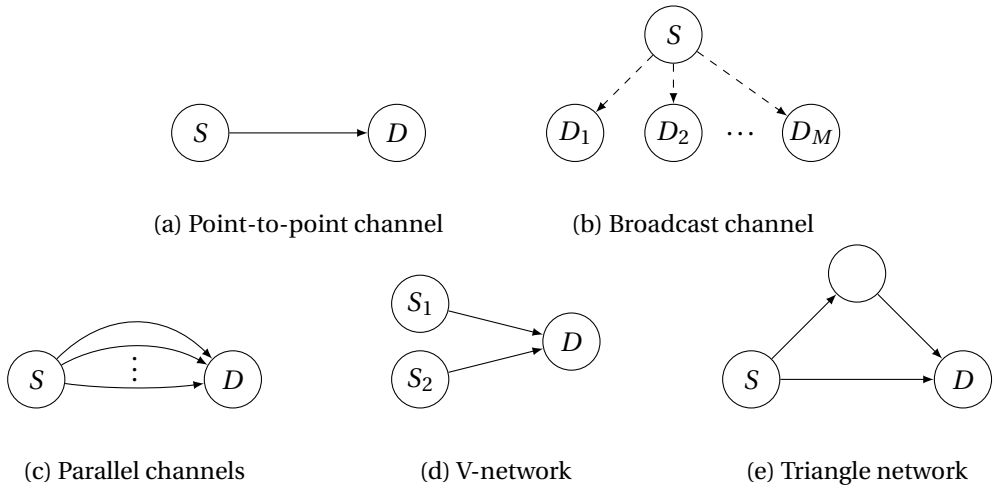
Secure communication over an erasure network significantly differs in nature, both from non-secure communication and from secure communication over an error-free network. Our contributions include the development of new capacity characterizations (coding schemes and outer bounds) that reveal nontrivial properties, of which we find at least three surprising: (1) In all the cases where we derive capacity, a two-phase scheme – which consists of a key generation phase and an encrypted message sending phase – is shown to be optimal. (2) In a network, finding the secret-message capacity of a cut does not simplify to summing up the individual capacities of the channels that form the cut as might be expected. (3) Similarly, finding the secret-message capacity of a network does not simplify to finding the minimum value cut of the network.

Contributions

The main contributions that we present in this thesis are as follows:

- We provide a complete characterization of the secret-message capacity of various communication settings (see Figure 1) in the presence of an eavesdropping adversary: the

[2] At this point we have not yet defined secret-message rate/capacity formally, one can think of it as the highest achievable rate under the constraint of maintaining secrecy. For formal definitions we refer to Chapter 1.



(a) Point-to-point channel (b) Broadcast channel

(c) Parallel channels (d) V-network (e) Triangle network

Figure 1: Communication settings with a complete characterization of secret-message capacity. Causal channel state feedback are sent over a separate noiseless public channel (not shown).

point-to-point channel, the broadcast channel[3], the network with multiple parallel channels, the V-network and the triangle network.

- We design new coding schemes for the above mentioned communication settings as well as for erasure networks with arbitrary topologies. Our coding schemes provide information-theoretic security and use only linear operations. In all the cases where we derive capacity, a two-phase scheme achieves the secret-message capacity.
- We derive new information theoretic outer bounds. Our proofs reflect the two-phase nature of our scheme: in an intuitive interpretation, a lower bound is given for the length of a key generation phase that is required to secure a message. Our outer bounds for networks confirm the properties mentioned above: Finding the secret-message capacity of a cut does not simplify to summing up the individual capacities of the channels that form the cut; furthermore, finding the secret-message capacity of a network does not simplify to finding the minimum value cut of the network.
- We introduce a linear programming approach for describing secure coding schemes for networks. We also use a novel linear programming proof technique for proving optimality.

Significance

Although our results are not dependent on communication technology, we believe that they are the most relevant for wireless communication. Providing security while maintaining large throughput is especially challenging in a wireless network. The wireless spectrum is a scarce resource and the demand of using this resource increases daily. In the last decade, data communication over the wireless medium has become omnipresent. Either through a local WiFi network or a mobile data network, people often use a wireless connection to access

[3]For $M > 3$ a complete characterization is available in some special cases. We refer to Chapter 3 for details.

media content and their data in the cloud. The Cisco Visual Networking Index [13] predicts further rapid growth in the next four years with 10 times the volume of today's mobile data traffic by 2018. As opposed to a wired counterpart, where throughput can be easily extended by deploying new communication cables, the capacity of a wireless channel is limited by nature. As the possibilities of physically protecting the medium are limited, it is easy to launch an eavesdropping attack. Moreover, such an attack is hard to prevent or detect: Therefore, advanced wireless technology and – as part of it – coding techniques are required for an efficient and secure use of the medium.

Our results could serve as the theoretical foundation of practical protocols. Some experimental practical protocols already use wireless channel variations as a source of randomness for secret-key generation [14–17], but our approach is different: We do not extract randomness from channel variability, instead, we take advantage of the adversary's limited presence and the errors that affect the adversary. Our approach promises higher achievable rates; furthermore, we provide a solution not only for key generation, but also for message encryption.

Outline

The thesis consists of 6 chapters. Each chapter – in Chapter 4 each section – has a Main result section, which summarizes the results presented in the given part. We delegate some details and lengthy derivations to the appendices.

Chapter 1 describes our communication model and provides formal definitions. We discuss different security notions and also summarize the most relevant known results that we use through the subsequent chapters.

Chapter 2 presents our results on the point-to-point erasure channel. We both provide a scheme design and derive an outer bound which establishes secret-message capacity of the setting.

Chapter 3 considers a multiterminal broadcast erasure channel. A secure coding scheme is presented that is optimal in all cases where the corresponding non-secure capacity is known. In this chapter we investigate different security definitions and generalize equivalence results for a multiterminal setting.

Chapter 4 is devoted to the secret-message capacity of networks. Using a linear programming approach we derive secret-message capacity of a network that consists of independent parallel channels, the V-network and the triangle network.

Chapter 5 investigates networks with arbitrary topology. Chapters 4 and 5 pursue different goals. In Chapter 4, we aim to design optimal coding schemes and provide complete characterization of the secret-message capacity. In Chapter 5 we design schemes for arbitrary topologies giving up optimality for simpler code design. We propose a two-phase achievability scheme for secure communication.

Chapter 6 concludes the thesis and discusses open questions for future research.

1 Model and background

In this chapter we set up the mathematical framework in which we derive our results. Using information-theoretic tools we give a formal definition of a communication setting, an adversary and the notion of security that we use throughout the thesis. In this setting, a coding scheme describes how a message is encoded and decoded by the communicating parties. Two key properties of a coding scheme are reliability and security. Reliability means that the intended receiver can decode the message without errors with high probability. Security means that secrecy of the message is ensured against the adversary. A *secure coding scheme* – as we define in this section – provides both reliability and security. It is meaningful to define *secret-message capacity*, which is the highest rate of communication that any secure coding scheme could possibly achieve in a given communication setting.

We also overview relevant previous results that serve as background and starting point of our investigation. While doing so, for consistency, we reformulate results using the notions and notation that we will next introduce.

1.1 Communication model

In our communication scenarios we have one sender, one or more receivers and an adversary. To communicate with each other a probabilistic channel is at their disposal. We often refer to the sender as Alice, to the receiver(s) as Bob (and Calvin) and to the adversary as Eve. In all our settings we consider one particular channel model: a discrete memoryless erasure channel with public state-feedback. For technical reasons[1] we assume that erasure probabilities are not 0 or 1. (We can obtain results for these corner cases by taking the limits.).

1.1.1 Erasure channel with state-feedback

The input alphabet of the channel consists of all possible length L vectors over a finite field $\mathbb{F}_q$. Let $\mathscr{X}$ denote the input alphabet of the channel, then $\mathscr{X} = \mathbb{F}_q^L$. We often call one such finite field vector a *packet*. Beside all input symbols, the channel output alphabet $\mathscr{Y}$ contains

[1] Such that for an erasure probability δ terms like $\frac{1}{\delta}$ or $\frac{1}{1-\delta}$ are meaningful.

also an erasure symbol $\perp$: $\mathcal{Y} = \mathcal{X} \cup \{\perp\}$. We denote $X_i \in \mathcal{X}$ the input of the channel in the ith time slot. We use the notation $X^n = (X_1, \ldots, X_n)$. We apply the same shorthand also for other vectors.

Point-to-point setting

We refer to the setting with one sender, one receiver and an adversary as the point-to-point scenario. In this case, the channel outputs two symbols (Y_i, Z_i) in the ith time slot. $Y_i \in \mathcal{Y}$ denotes the output symbol that the receiver (Bob) gets, while $Z_i \in \mathcal{Y}$ denotes the output that Eve observes. Each output symbol is either the input symbol or the erasure symbol. Erasures occur independently for Bob and Eve and also independently, with the same probability in all time slots. We use δ and δ_E to denote the erasure probability toward Bob and Eve respectively. Formally,

$$\Pr\{Y_i, Z_i | X^i Y^{i-1} Z^{i-1}\} = \Pr\{Y_i | X_i\} \Pr\{Z_i | X_i\}, \tag{1.1}$$

$$\Pr\{Y_i | X_i\} = \begin{cases} 1-\delta, & Y_i = X_i \\ \delta, & Y_i = \perp, \end{cases} \tag{1.2}$$

$$\Pr\{Z_i | X_i\} = \begin{cases} 1-\delta_E, & Z_i = X_i \\ \delta_E, & Z_i = \perp. \end{cases} \tag{1.3}$$

With more than one receivers

In a *broadcast* setting there are $M \geq 2$ receivers. In this case we do not distinguish an external adversary, instead we will assume that the adversary is a subset of the receivers. When there are M receivers, the channel output is the tuple $\left(Y_{1,i}, \ldots, Y_{M,i}\right) \in \mathcal{Y}^M$, where $Y_{j,i}$ is the observation of receiver j. We use Y_i as a shorthand for $\left(Y_{1,i}, \ldots, Y_{M,i}\right) \in \mathcal{Y}^M$. Similarly as in the point-to-point case we assume that all erasure events are i.i.d., with erasure probabilities $\delta_1, \ldots, \delta_M$:

$$\Pr\{Y_i | X^i Y^{i-1}\} = \prod_{j=1}^{M} \Pr\left\{Y_{j,i} | X_i\right\} \tag{1.4}$$

$$\Pr\{Y_{j,i} | X_i\} = \begin{cases} 1-\delta_j, & Y_{j,i} = X_i \\ \delta_j, & Y_{j,i} = \perp \end{cases}, \quad \forall j \in \{1, \ldots, M\}. \tag{1.5}$$

With more than one channels

In a *network* setting there are more than one (ℓ) channels and maybe some intermediate nodes addition to the sender and the receivers. Every channel operates as defined in the point-to-point setting, independently of each other. In this case we use the indices of the erasure probabilities and the first indices of variables to denote the index of the channel, e.g., $Z_{3,i}$ denotes the output symbol for Eve on the 3rd channel in the ith time slot, which is an erasure with probability δ_3. We again use Y_i as a shorthand for $\left(Y_{1,i}, \ldots, Y_{\ell,i}\right)$ and Z_i as a shorthand for

$(Z_{1,i},\ldots,Z_{\ell,i})$. We do not consider broadcast and network settings simultaneously, hence in all cases it will be clear whether $Y_{j,i}$ denotes the output of receiver j or that of channel j.

Public feedback

We assume that the receivers send a public acknowledgment after each transmission, i.e., the state of the channel toward the receivers is available strictly causally to all parties. F_i denotes the random variable that describes whether or not an erasure occurred in time slot i toward each receiver. Let $\mathbb{1}_{\{\cdot\}}$ denote the indicator function, then $F_i = \mathbb{1}_{\{Y_i=\perp\}}$ for the point-to-point scenario and $F_i = \left(\mathbb{1}_{\{Y_{1,i}=\perp\}},\ldots,\mathbb{1}_{\{Y_{M,i}=\perp\}}\right)$ for a broadcast or for a network setting. By public we mean that F_i is available to Alice, all the receivers and intermediate nodes in a network and for Eve also before time slot $i+1$. Without giving up rigor, when it is convenient we use indices or letters to denote correct receptions instead of complete channel state. E.g., $F_i = B$ means "only Bob received" the ith transmission.

Note that F_i does not contain information about Z_i or $Z_{j,i}$, but F_i is also available for the adversary. It is reasonable to assume cooperation from the receivers, but not from Eve. Indeed, an acknowledging mechanism is part of common wireless standards, e.g., 802.11 (WiFi), 802.15 (Bluetooth), 802.16 (WiMax). In practice, feedback takes place over the same medium, however, the size of the feedback is only 1 bit per packet, which is assumed to be negligible compared to the packet size. With appropriate batching and source coding, the size of feedback can be further reduced to $h_2(\delta)$ bits [2] in average on a channel with erasure probability δ. For this reason we assume that feedback traffic is negligible, and in our model feedback takes place on a separate, error-free public channel, which is not taken into account in rate calculations.

Further assumptions

Alice, all receivers, and Eve are able to generate private randomness at a practically unlimited rate. We denote the private randomnesses by $\Theta_A,\Theta_B,\ldots,\Theta_E$, or by receiver indices: e.g., Θ_2 is the randomness of receiver 2. These random sources are independent of each other and of any other randomness.

1.2 Secret-message sending

Below we formally define the reliability and security criteria of the secret-message sending problem. In all scenarios the goal of the communication is to reliably and securely send a private message W_j to each receiver j. Definitions are given for M receivers and any number of channels. For the special cases of $M=1$ or a single channel we omit the indices. We provide general definitions here and adapt it for various settings in the subsequent chapters. An $(n,\epsilon,N_1,N_2,\ldots,N_M)$ *coding scheme* sends message W_j which consist of N_j packets to receiver j using n transmissions from Alice. Beside reliable message transmission, a *secure* coding scheme ensures secrecy of the messages.

[2] $h_2(p)$ denotes the binary entropy function: $h_2(p) = -p\log_2(p) - (1-p)\log_2(1-p)$

Definition 1.1. *An* $(n, \epsilon, N_1, N_2, \ldots, N_M)$ coding scheme *consists of the following components: (a) message alphabets* $\mathcal{W}_j = \mathbb{F}_q^{LN_j}$, $j = 1, 2, \ldots, M$, *(b) encoding maps* $f_{k,i}(\cdot)$, $i = 1, 2, \ldots, n$ *for each channel* k, *and (c) decoding maps* $\phi_j(\cdot)$, $j = 1, 2, \ldots, M$ *for each receiver, such that the inputs to the channel are*

$$X_{k,i} = f_{k,i}\left(A_{k,i-1}\right), \quad i = 1, 2, \ldots, n, \tag{1.6}$$

where $A_{k,i-1}$ *denotes all random variables that the sender of channel* k *has access to before the* i*th transmission. A coding scheme provides reliability for each receiver* j*:*

$$\Pr\left\{\phi_j\left(B_{j,n}\right) \neq W_j\right\} < \epsilon, \quad \forall\, j \in \{1, \ldots, M\}, \tag{1.7}$$

where messages $W_j \in \mathcal{W}_j$ *are arbitrary messages in their respective alphabets and* $B_{j,n}$ *denotes all random variables that receiver* j *has access to after the* n*th transmission.*

For all different communication settings we specify $A_{k,i-1}$ and $B_{j,n}$.

Definition 1.2. *In a communication setting a* rate *tuple* $(R_1, \ldots, R_M) \in \mathbb{R}_+^M$ *is achievable if for every* $\epsilon > 0$ *there exists an* $(n, \epsilon, N_1, N_2, \ldots, N_M)$ *coding scheme for which*

$$R_j - \epsilon < \frac{1}{n} N_j, \quad \forall\, j \in \{1, \ldots, M\}. \tag{1.8}$$

Definition 1.3. *In a communication setting the* capacity *region* $\mathcal{R}^M \subset \mathbb{R}_+^M$ *is the set of achievable rate tuples* $(R_1, \ldots, R_M)$.

Of course, for a single receiver, the capacity region is one dimensional and hence it is meaningful to call its maximum simply the capacity. This definition adheres to the usual definition of channel capacity.

The notion of a coding scheme requires reliability, but does not consider yet security. Capacity is the highest message rate that is achievable reliably. When we want to stress that no security criterion is considered we refer to the capacity as *non-secure capacity*. The following definition extends Definition 1.1 with a security requirement.

Definition 1.4. *A* $(n, \epsilon, N_1, N_2, \ldots, N_M)$ *coding scheme as defined by Definition 1.1 is a* secure coding scheme*, if in addition to* (1.6)-(1.7) *it also satisfies*

$$I\left(W_j; E_n\right) < \epsilon, \tag{1.9}$$

for every receiver j*, where* E_n *denotes all random variables that Eve has access to after the* n*th transmission.*

For all different communication settings we specify E_n.

Definition 1.5. *In a communication setting a* secret-message rate *tuple* $(R_1,\dots,R_M)\in\mathbb{R}_+^M$ *is achievable, if for every* $\epsilon>0$ *there exists an* $(n,\epsilon,N_1,N_2,\dots,N_M)$ *secure coding scheme for which*

$$R_j-\epsilon<\frac{1}{n}N_j,\quad \forall\, j\in\{1,\dots,M\}. \tag{1.10}$$

Definition 1.6. *In a communication setting the* secret-message capacity *region* $\mathscr{R}_S^M\subset\mathbb{R}_+^M$ *is the set of achievable secret-message rate tuples* $(R_1,\dots,R_M)$ *(as defined by Definition 1.5).*

Similarly as in the case of capacity, secret-message capacity refers to the maximum of a one dimensional secret-message capacity region.

We note that the definitions above define rate in terms of packets, i.e., rate 1 means conveying one $\mathbb{F}_q^L$ vector per time slot. We use this convention also for entropy and mutual information throughout the thesis which enables us to omit the constant factor $L\log(q)$ from rate and other information expressions.

1.3 Secret-key generation

In Section 1.2 the necessary definitions for the secret-message sending problem were given, which is the main focus of this thesis. The secret-key generation problem is a different, but closely related problem. The goal of a secret-key generation protocol is to set up a secret common randomness – a key – between parties. As opposed to the secret-message sending problem, the sender does not necessarily select the secret randomness a priori, the actual key is also the outcome of the protocol. Clearly, for the secret-key problem, we are not constrained by what common randomness the parties securely agree upon, hence a securely communicated random message can serve as key. It is immediate that the secret-key capacity of a channel is an upper bound for the secret-message capacity.

In the sequel we will see that secret-key generation is a very natural building block for designing a secret-message sending scheme. For this reason, below we give formal definitions regarding secret-key generation. In Section 1.5 we summarize relevant results on secret-key generation for the erasure channel from [5].

For clarification we note that the terms *secrecy capacity and secrecy rate* are commonly used for both secret-key generation and secret-message sending. This is because in many communication settings there is no need for distinction, the two problems are often equivalent. E.g. if the communication is one-way only (no feedback) or when unlimited public discussion is possible, solving the secret-key generation problem solves the secret-message sending problem also. We follow the convention that terms secrecy capacity or secrecy rate are used only in a context, where the two problems are equivalent.

In this section we consider a single receiver (Bob) in the same communication model as defined in Section 1.1.

Definition 1.7. *An* (n,ϵ') key generation scheme *over n transmissions has the following com-*

ponents: (a) encoding maps $f_{k,i}(\cdot)$, $i = 1,2,\ldots,n$ for each channel k, and (b) key computation functions $\mathcal{K}_A(\cdot), \mathcal{K}_B(\cdot)$ for Alice and Bob respectively, such that the inputs to the channel are

$$X_{k,i} = f_{k,i}\left(A_{k,i-1}\right), \quad i = 1,2,\ldots,n, \tag{1.11}$$

where $A_{k,i-1}$ denotes all random variables that the sender of channel k has access to before the ith transmission. Further, a secret key K_{AB} is computed using the key computation functions such that

$$\Pr\{\mathcal{K}_A(A_n) \neq K_{AB}\} < \epsilon', \tag{1.12}$$

$$\Pr\{\mathcal{K}_B(B_n) \neq K_{AB}\} < \epsilon', \tag{1.13}$$

$$I(K_{AB}; E_n) < \epsilon', \tag{1.14}$$

are satisfied. A_n, B_n and E_n denote all random variables that Alice Bob and Eve have access to after the nth transmission.

For all different communication settings we specify $A_{k,i-1}$.

Definition 1.8. *In a communication setting a* key rate $R_K \in \mathbb{R}_+$ *is achievable, if for every $\epsilon' > 0$ there exists an (n,ϵ') key generation scheme over n transmissions for which the secret key K_{AB} satisfies:*

$$R_K - \epsilon < \frac{1}{n} H(K_{AB}). \tag{1.15}$$

Definition 1.9. *In a communication setting the* secret-key capacity $C_K \in \mathbb{R}$ *is the maximum of the set of achievable key rates.*

1.4 Adversary model and security notions

From the many different aspects of security, in this thesis we focus on confidentiality, which means secrecy of a message from a passive adversary.

1.4.1 Adversary

We consider an eavesdropping adversary, Eve, who aims to learn the message that Alice sends to another receiver. In most settings we assume that Eve is passive, she does not transmit any signal. This assumption is valid in a wireless environment, where eavesdropping the channel is easy, and an eavesdropping node does not want to reveal her presence with communication.

We do not make any assumptions on the computational power of Eve. Instead, we only assume that the channel through which she observes the communication is not perfect; erasures occur on her channel also.

In a broadcast setting we will assume that the adversary controls some of the receivers. This gives the chance to Eve to influence how the protocol runs through the feedback the adver-

sarial receiver sends. An adversarial receiver might lie about its channel state in order to learn information about an other receiver's private message. In Chapter 3 we introduce the appropriate security definitions against such an adversary.

In a network setting, the adversary can select a subset of channels to eavesdrop on. We assume that the maximum number of eavesdropped channels is known, but the actual subset of eavesdropped channels (which we sometimes call the *location of Eve*) is not.

1.4.2 Information theoretic security notions

Information theoretic secrecy is defined in terms of mutual information between the message and the observations of the adversary. It is common to distinguish *perfect, strong* and *weak* secrecy [18]. Let W denote a message to be secured and E_n all observations that Eve has access to after n transmissions of the protocol. Perfect secrecy means that

$$I(W, E_n) = 0. \tag{1.16}$$

This definition implies exact statistical independence between the message and Eve's observations.

In most communication settings perfect secrecy is too stringent and not possible to satisfy. Strong secrecy replaces exact statistical independence with asymptotic independence:

$$\lim_{n\to\infty} I(W, E_n) = 0. \tag{1.17}$$

Clearly, strong secrecy is equivalent to our definition of secrecy according to Definition 1.4. The difference between perfect and strong secrecy has a similar flavor as the difference between zero-error communication [6] and the usual definition of reliable communication (e.g., Definition 7.14. in [2]), which asks arbitrarily small, but not exactly zero error probability of decoding.

Strong secrecy can be interpreted as requiring that the information leak toward Eve is negligible. Instead, weak secrecy requires only a negligible *rate* information leak:

$$\lim_{n\to\infty} \frac{1}{n} I(W, E_n) = 0. \tag{1.18}$$

A sub-linear growth of leaked information still satisfies the weak secrecy criterion. This could even mean that Eve observes an asymptotically unbounded number of message bits in cleartext.

Although weak and strong secrecy are not equivalent, in many cases (e.g., [8, 9, 19, 20]) the securely achievable rates are the same under both criteria. Our secret-message capacity results are of this kind also. We provide secure coding schemes that satisfy strong secrecy and we prove outer bounds for weak secrecy. By this we also show that relaxing the security requirement from strong to weak secrecy does not increase the derived secret-message capacity region.

For completeness, we mention that the term weak security is sometimes used in another meaning in the network coding literature (e.g., [21–23]). In that context, the security criterion applies for any part of the message, but not for the message as a whole. If W^k denotes any k length part of a message of size N, then the weak security criterion is

$$I\left(W^k; E_n\right) = 0. \tag{1.19}$$

1.4.3 Equivalence of security notions

We formulate security in information-theoretic terms. In the realm of computational cryptography it is more common to prove security of an encryption scheme by showing distinguishing security or semantic security. To facilitate the interpretation of our results and to make a fair comparison with other schemes possible, we cite a recent result from [24], which shows equivalence between the two approaches.

The results hold for a single receiver Bob and hence for a single message W, that has a distribution P_W. It is common to define the *advantage* of the adversary to express the gain that the adversary obtains by observing a protocol. Considering information-theoretic security, the adversarial advantage expressed in terms of mutual information (mis = mutual information security) is defined as:

$$\mathbf{Adv}^{\text{mis}} = \max_{P_W} I\left(W; E_n\right). \tag{1.20}$$

The notion of semantic security captures the intuition that the probability that an adversary can compute a function g of the message should not increase significantly after observing the protocol compared to the a priori probability of a correct guess. The semantic security advantage is defined as

$$\mathbf{Adv}^{\text{ss}} = \max_{g, P_W} \left\{ \max_{\mathscr{A}} \Pr\left\{\mathscr{A}\left(E_n\right) = g\left(W\right)\right\} - \max_{\mathscr{S}} \Pr\left\{\mathscr{S}\left(P_W, g\right) = g\left(W\right)\right\} \right\}, \tag{1.21}$$

where g is any function of W, $\mathscr{A}$ is any function the adversary may compute after observing the protocol and $\mathscr{S}$ is a simulator trying to compute g without accessing the protocol output. The term simulator to denote guessing functions comes from the intuition that ideally there exists an algorithm (simulator) that simulates the run of a protocol without having access to the message and whose output is indistinguishable from the output of a real protocol. Theorems 1, 5 and 8 from [24] prove the following inequalities:

$$\mathbf{Adv}^{\text{ss}} \leq \sqrt{2 \cdot \mathbf{Adv}^{\text{mis}}} \tag{1.22}$$

$$\mathbf{Adv}^{\text{mis}} \leq 4 \cdot \mathbf{Adv}^{\text{ss}} \log\left(\frac{2^n}{\mathbf{Adv}^{\text{ss}}}\right). \tag{1.23}$$

This result shows that for a single receiver, the requirement (1.9) implies semantic security, because a small ϵ in (1.9) causes that $\mathbf{Adv}^{\text{ss}}$ is also small. In many cases, $\mathbf{Adv}^{\text{ss}} \log\left(\frac{2^n}{\mathbf{Adv}^{\text{ss}}}\right)$ decays

to 0 when $n \to \infty$ and the converse also holds. Although security definitions might look quite different at first sight, there is no fundamental difference between these notions of security.

The above definitions are not directly applicable for more than one receivers, hence in Chapter 3 we extend the notion of semantic security so that it handles joint message distributions. We show a similar equivalence result with distribution independent security, which we introduce in Chapter 3.

We can conclude that the key difference between information-theoretic and computational security is not the level of security they provide, but the model of adversary they consider. Computational cryptography considers an adversary who is computationally bounded, but has a complete observation of a protocol run, while in the information-theoretic model the adversary is unbounded computationally, but cannot completely observe the communication between honest parties. Real adversaries are not unlimited either in their computational power or in their communication capabilities, hence a combination of the two approaches is possible.

1.5 Key generation in a point-to-point setting

We summarize and reformulate the results from [5] where the secret-key generation problem is considered. In [5] the group key generation problem is investigated in a broadcast setting, which means that all receivers in a group are required to compute the same secret key. Furthermore, an unlimited capacity public channel is at their disposal for public discussion. Despite of these differences, the results are directly applicable for our point-to-point setting, because: (a) with one receiver the group key becomes a pairwise key, which matches our definition of key generation; (b) we observe that for a single receiver the key generation scheme in [5] uses the public channel for state-feedback only, which is available in our setting also.

In the point-to-point setting $A_{i-1} = \left(\Theta_A, F^{i-1}\right)$, i.e., the randomness that Alice generates as well as the state of Bob's channel in the first $i-1$ transmissions. Further, $A_n = (\Theta_A, F^n)$, $E_n = (Z^n, \Theta_E, F^n)$, $B_n = (Y^n, \Theta_B, F^n)$. We note that Y^n determines F^n. Theorems 1 and 3 from [5] state the following (using our notation):

Theorem 1.1. *The secret-key capacity in the point-to-point setting over the erasure channel with state-feedback is*

$$C_K = \delta_E (1-\delta), \tag{1.24}$$

furthermore, secret-key capacity is achieved by a linear scheme such that the resulting secret key K_{AB} is uniformly distributed over its alphabet.

Recall that δ_E is the erasure probability on Eve's channel, while δ is the erasure probability for Bob.

We give the proof of the direct part of Theorem 1.1 by describing the key generation scheme

that achieves secret-key capacity. The outer bound directly comes as a special case of a more general result [9].

Key generation scheme

Intuition and an example Before a formal description, we provide intuition and overview the principles behind the key generation scheme. Consider the following example. Alice sends out 5 random packets $X_1, \ldots, X_5$ over the channel. Due to erasures neither Bob nor Eve receives all of these. Assume Bob receives X_1 and X_4, while Eve receives X_2, X_4 and X_5. Figure 1.1 illustrates the situation. If an oracle could tell Alice and Bob which packets Eve has

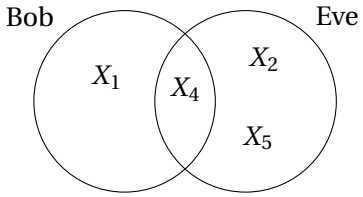


Figure 1.1: Example: packets received by Bob and Eve

received, they could simply use the packet X_1 as their shared key. Recall that by assumption Alice knows the indices of packets received by Bob. Such oracle does not exist, still if Alice and Bob knows that there is one packet that Bob has received but Eve has not, they can both compute $K_{AB} = X_1 \oplus X_4$. This way, either X_1 or X_4 is the unknown packet for Eve, K_{AB} remains secret.

In general, Alice and Bob can securely form as many linear combinations as the number of packets only Bob receives. With an overwhelming probability, the number of such packets is close to its expected value, $n\delta_E(1-\delta)$ after n transmissions. This gives rise to an intuitive interpretation of the secret-key capacity: the achieved key rate corresponds to the probability that a packet gets received by only Bob, but not Eve.

For constructing the linear combinations we need the following property. Assume Bob has received n_B packets and Eve has received n_E of them. A coefficient matrix of size $n_B \times (n_B - n_E)$ is used to produce the linear combinations. Since Eve has n_E packets and we cannot know which ones, we have to assume that she can produce any linear combination of packets with no more than n_E nonzero coefficients. Thus, Eve can compute and subtract any linear combinations of no more than n_E packets from the linear combinations that we produce. For the key to be secret we need that the remaining part is still a linearly independent set of packets. In other words, any $n_B - n_E$ rows of our coefficient matrix has to be a full-rank matrix. By Theorem A.4 in Appendix A.3 such a matrix is the parity check matrix of an MDS code.

Below we present a version of the scheme in [5] adapted for the point-to-point setting. We also reformulate the analysis of the scheme.

Scheme description From n we compute s, s' such that:

$$n = \frac{s}{1-\delta} + \frac{s^{\frac{3}{4}}}{1-\delta} \tag{1.25}$$

$$s = \frac{s'}{\delta_E} + \frac{s'^{\frac{3}{4}}}{\delta_E}. \tag{1.26}$$

1. Alice sends n packets selected i.i.d. uniformly at random from $\mathbb{F}_q^L$.
2. Let X_B denote the row vector of the first s packets that Bob receives. If Bob does not receive at least s packets after n transmissions, an error is declared. Alice and Bob both compute

$$K_{AB} = X_B H_{K_{AB}}, \tag{1.27}$$

where $H_{K_{AB}}$ is a $s \times s'$ matrix and is a parity check matrix of an MDS code. $H_{K_{AB}}$ can be publicly known and used arbitrarily many times.

Analysis From the description it is clear that the scheme satisfies Definition 1.7. The uniform distribution of K_{AB} follows from the uniform distribution of X_B and the MDS property of $H_{K_{AB}}$. Hence, $H(K_{AB}) = s'$ and from (1.25)-(1.26)

$$\lim_{n\to\infty} \frac{H(K_{AB})}{n} = \lim_{n\to\infty} \frac{s'}{n} = \delta_E(1-\delta), \tag{1.28}$$

which shows the rate assertion of Theorem 1.1.

We next show that the protocol succeeds with arbitrarily small error probability. If no error is declared, then both Alice and Bob computes the same key. Let κ denote the number of packets that Bob receives. Then,

$$\mathbb{E}\{\kappa\} = n(1-\delta). \tag{1.29}$$

An error occurs, if Bob receives less than s packets, i.e., $\kappa < s$.

$$\Pr\{\kappa < s\} = \Pr\left\{\kappa < n(1-\delta) - s^{\frac{3}{4}}\right\} = \Pr\left\{\mathbb{E}\{\kappa\} - \kappa > s^{\frac{3}{4}}\right\} \le \Pr\left\{|\mathbb{E}\{\kappa\} - \kappa| > s^{\frac{3}{4}}\right\} \tag{1.30}$$

$$\le 2\exp\left\{-\frac{2s^{\frac{3}{2}}}{n}\right\} = 2\exp\left\{-\frac{2s^{\frac{3}{2}}}{\frac{s}{1-\delta} + \frac{s^{\frac{3}{4}}}{1-\delta}}\right\} < 2\exp\left\{-\frac{2s^{\frac{3}{2}}}{\frac{2s}{1-\delta}}\right\} = 2e^{-(1-\delta)\sqrt{s}}, \tag{1.31}$$

where we used the Chernoff-Hoeffding bound (see in Appendix A.2). With $n \to \infty$, s also grows to infinity, thus the error probability decays to 0.

We use a similar technique to prove the security of the key. The observation of the eavesdropper is $E_n = (Z^n, F^n, \Theta_E)$. We carry out the analysis assuming no error occurred. The private randomness of the eavesdropper is independent from (Z^n, F^n, K_{AB}) and K_{AB} is uniformly

distributed, hence

$$I\left(K_{AB};Z^nF^n\Theta_E\right)=I\left(K_{AB};Z^nF^n\right)=H\left(K_{AB}\right)-H\left(K_{AB}\left|Z^nF^n\right.\right)=s'-H\left(K_{AB}\left|Z^nF^n\right.\right). \quad (1.32)$$

Let X_{BE} denote the set of packets that both Bob and Eve received, while $X_{B\emptyset}$ the set of packets that only Bob has. Further, $H_{K_{AB}}^{BE}$ and $H_{K_{AB}}^{B\emptyset}$ denote the rows of $H_{K_{AB}}$ corresponding to X_{BE} and $X_{B\emptyset}$. Then,

$$\begin{aligned}
H\left(K_{AB}\left|Z^nF^n\right.\right)&=H\left(X_BH_{K_{AB}}\left|Z^nF^n\right.\right)=H\left(\left[\begin{array}{cc}X_{BE} & X_{B\emptyset}\end{array}\right]\left[\begin{array}{c}H_{K_{AB}}^{BE}\\ H_{K_{AB}}^{B\emptyset}\end{array}\right]\left|X_{BE}F^n\right.\right) && (1.33)\\
&=H\left(X_{B\emptyset}H_{K_{AB}}^{B\emptyset}\left|X_{BE}F^n\right.\right)\overset{(a)}{=}H\left(X_{B\emptyset}H_{K_{AB}}^{B\emptyset}\left|F^n\right.\right) && (1.34)\\
&=\sum_{i=0}^{s}H\left(X_{B\emptyset}H_{K_{AB}}^{B\emptyset}\left|\left|X_{B\emptyset}\right|=i,F^n\right.\right)\Pr\left\{\left|X_{B\emptyset}\right|=i\right\} && (1.35)\\
&\overset{(b)}{=}\sum_{i=0}^{s}\min\left\{i,s'\right\}\Pr\left\{\left|X_{B\emptyset}\right|=i\right\}\geq s'\sum_{i=s'}^{s}\Pr\left\{\left|X_{B\emptyset}\right|=i\right\} && (1.36)\\
&=s'\Pr\left\{\left|X_{B\emptyset}\right|\geq s'\right\}=s'\Pr\left\{\left|X_{B\emptyset}\right|\geq\mathbb{E}\left\{\left|X_{B\emptyset}\right|\right\}-s'^{\frac{3}{4}}\right\} && (1.37)\\
&=s'\left(1-\Pr\left\{\mathbb{E}\left\{\left|X_{B\emptyset}\right|\right\}-\left|X_{B\emptyset}\right|>s'^{\frac{3}{4}}\right\}\right) && (1.38)\\
&\geq s'\left(1-\Pr\left\{\left|\mathbb{E}\left\{\left|X_{B\emptyset}\right|\right\}-\left|X_{B\emptyset}\right|\right|>s'^{\frac{3}{4}}\right\}\right) && (1.39)\\
&\geq s'\left(1-2\exp\left\{-\frac{2s'^{\frac{3}{2}}}{s}\right\}\right)>s'\left(1-2\mathrm{e}^{-\delta_E\sqrt{s'}}\right). && (1.40)
\end{aligned}$$

In step (a) we used the independence property of the packets, while in step (b) we used the MDS property of $H_{K_{AB}}$ (see Corollary A.1 in Appendix A.3). In the last step we used again the Chernoff-Hoeffding bound. Substituting back to (1.32) we have:

$$I\left(K_{AB};Z^nF^n\Theta_E\right)\leq s'2\mathrm{e}^{-\delta_E\sqrt{s'}}. \quad (1.41)$$

Since s' grows to infinity when $n\to\infty$, the above mutual information term also decays to 0, hence security of the key is satisfied for any $\epsilon'>0$ if n is sufficiently large. With this, we have shown that for a large enough n the above scheme satisfies Definition 1.8, thus the key rate $\delta_E\left(1-\delta\right)$ is achievable.

2 Secret-message capacity of a point-to-point channel

In this chapter we consider a communication setting with three parties: Alice, the sender, Bob the receiver and Eve, the eavesdropping adversary. Alice aims to send a message to Bob over an eavesdropped erasure channel – as introduced formally in the previous chapter. We refer to this setting as the point-to-point setting or the point-to-point channel, since there are two legitimate communicating parties.

We provide a complete characterization of the secret-message capacity for this setting. We design a secure coding scheme and derive a matching outer bound on the secret-message capacity. Our capacity achieving coding scheme uses linear operations only, both coding and decoding are feasible in polynomial time complexity.

The tools and principles that we introduce in this chapter are used – and further developed – in the subsequent chapters. Most importantly, we introduce a two-phase principle in our coding scheme design. In the first phase Alice and Bob agree on a secret key, while in the second phase they use the key judiciously for encryption in the second phase, when the message is transferred in an encrypted form. Both phases constructively use channel erasures. We often call the phases *key generation phase* and *(encrypted) message sending phase* respectively. In our context, in a message sending phase the message is always encrypted, but for simplicity we sometimes omit the term *encrypted*.

Both phases are optimal by themselves, and our converse proof shows that they are optimal also together. The key generation phase is an optimal scheme for secret-key generation, and the message sending phase conveys the encrypted message reliably using a capacity achieving scheme. The fact that combining the two phases results an optimal secure scheme shows that our message sending phase is using the least possible amount of key for securing and sending the message reliably. We use the principle of encryption that we introduce here also in other scenarios.

The new outer bound that we derive in this chapter provides some further insight to the problem. Although the proof holds for any scheme, it reflects the two phases of our coding scheme. Also, we identify information terms that can be interpreted naturally as generating and consuming a secret key, while an inequality is drawing a balance between the two.

2.1 Related work

The first information-theoretic results on the problem of secure communication over an insecure (eavesdropped) channel date back to Shannon [4]. It was shown that securing a message over an error-free eavesdropped channel requires a pre-shared secret randomness (a key). Furthermore, the entropy of the key has to be at least as large as the entropy of the message. Given such a key, one-time pad encryption implements a perfectly secure encryption. We briefly summarize one-time pad in Appendix A.1.

Wyner's seminal paper [25] considered a noisy channel, where the wiretapper observes a noisy version of the legitimate receiver's observation, or in other words, Eve's channel is a degraded version of Bob's. The secrecy capacity of this wiretap channel is derived. A more general setting was investigated in [26]. Both private and common messages are considered and Eve's channel is arbitrary, not necessarily degraded. A single letter characterization of this setup is provided. The results were generalized also for a Gaussian channel [27] and for a fading channel [28]. However, none of these work consider feedback from the receiver. Applied to erasure channels, these results state that the secrecy capacity is non-zero only if the honest party has a better channel than the eavesdropper, i.e., the erasure probability toward Eve is higher than toward Bob.

Use of feedback and public discussion can improve the secrecy capacity, as was first shown by Maurer [7], followed by more general results for multiple terminals in [8, 9, 29–31]. All these results focus on secret-key generation, however, a cost-free public channel with infinite capacity is also available by assumption. As a result, the secret-message capacity is trivially the same, because the message can be encrypted with the generated secret key using a one-time pad and sent securely on the public channel. In contrast, our setup assumes that only state-feedback is available publicly, and there is no other high-capacity public channel. A similar setup, but with non-causal state-information available only to the transmitter was studied in [32, 33] for the Gaussian problem, where some achievability schemes based on dirty-paper coding were examined.

Feedback can be especially helpful if the feedback channel is not public. In [34] the feedback signal is implicitly used as a random encryption key without the source explicitly knowing the signal. In [35, 36] the wiretap channel with secure feedback is investigated. In [35] the feedback is perfect output feedback, while in [36] a secure rate limited feedback is used as a shared random source. In contrast, in our model feedback is always public and limited to the channel state.

As discussed in Section 1.5, for the broadcast erasure channel with public discussion, a capacity achieving scheme for key generation was proposed in [5] for the group secret-key exchange problem, where the public channel is also considered free and unlimited. In the case of two parties the scheme specializes to requiring only the channel state to be communicated over the public channel. As part of our secure coding scheme we use the key generation algorithm from [5].

2.2 Model

We adapt our definitions to the given scenario. Before the ith transmission Alice knows the message W, her private randomness Θ_A and all previous channel states, hence in Definition 1.1 $A_{i-1} = \left(W, \Theta_A, F^{i-1}\right)$ and the channel inputs are (eq. (1.6) becomes):

$$X_i = f_i\left(W, \Theta_A, F^{i-1}\right), \quad i = 1, 2, \ldots, n. \tag{2.1}$$

Bob receives $Y^n F^n$ and knows Θ_B, thus decodability condition of Definition 1.1 (eq. (1.7)) becomes

$$\Pr\left\{\phi_j\left(Y^n, F^n, \Theta_B\right) \neq W\right\} < \epsilon, \tag{2.2}$$

while Eve learns Z^n, F^n and knows Θ_E, hence the security criterion in Definition 1.4 (eq. (1.9)) is

$$I\left(W; Z^n F^n \Theta_E\right) < \epsilon. \tag{2.3}$$

2.3 Main result

The following theorem is the main result of this chapter.

Theorem 2.1. *The secret-message capacity of the point-to-point erasure channel with state-feedback is*

$$C_{SM} = (1-\delta)\delta_E \frac{1-\delta\delta_E}{1-\delta\delta_E^2}. \tag{2.4}$$

We prove Theorem 2.1 in two steps: we first propose a coding scheme for secret-message sending in Section 2.4, and then give the converse proof in Section 2.5.

2.3.1 Discussion

Comparing (1.24) and (2.4) clearly shows that the secret-key and the secret-message capacities differ, the latter is no larger than the former, which in line with our previous discussion about the relation between the two problems (see Section 1.5).

The role of state-feedback also becomes clear from our result. First, without feedback, there is no difference between the secret-key and the secret-message capacity. Alice does not learn anything during a protocol run, thus if she can compute a key after the protocol run, she could have equally computed it before running the protocol. In contrast, with using feedback, this equivalence does not hold anymore. Second, the achievable rate is increased with the help of feedback and – contrary to the case without feedback – the secret-message capacity is nonzero in all cases.

Given the scheme for secret-key generation, one natural strategy is to generate a secret key,

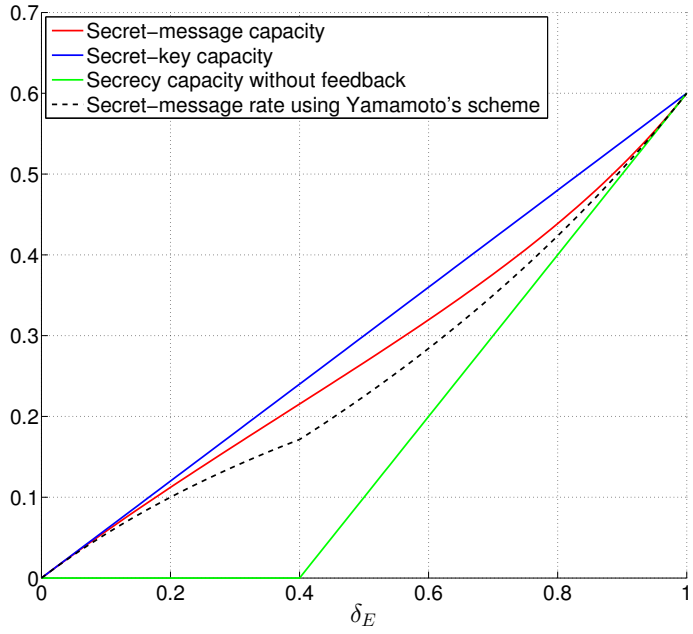


Figure 2.1: Point-to-point channel example: secret-message capacity and secret-key capacity with and without feedback, $\delta = 0.4$

use it as a one-time pad to encrypt the message, and send the encrypted message reliably using a forward error correcting (FEC) code. An improvement over this is possible when Eve has a higher erasure probability than Bob by leveraging secrecy from both the secret key generated and the channel advantage of Bob over Eve using the scheme of Yamamoto [37]. However, our two-phase capacity-achieving scheme demonstrates that one can do even better by exploiting feedback (see Figures 2.1-2.2). The benefits of our scheme come from using ARQ in the message sending phase (as opposed to FEC or a wiretap code) to deliver encrypted message packets to Bob. ARQ focuses on reliable transmission to Bob and hence could repeat (identical) transmissions, with the result that Eve receives fewer *distinct* encrypted message packets. Even when Eve has a lower erasure probability than Bob, the ARQ scheme ensures that Bob has a relative advantage over Eve. Therefore feedback has been used for the purpose of reducing the required key size by tilting the channel advantage towards Bob.

For comparison we plot the secret-message and the secret-key capacity together with the secrecy capacity when no feedback is available for some special parameter values. We plot secret-message rates achieved using FEC and Yamamoto's scheme as well. Figures 2.1-2.2 show the gap between the secret-message and the secret-key capacities as well as the benefit that state-feedback provides. (In Figure 2.2 the secret-message capacity without feedback is 0.)

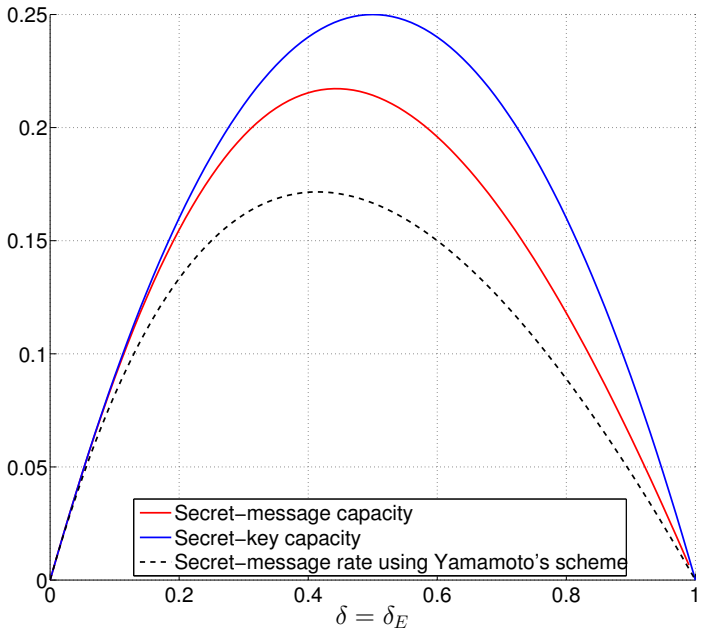


Figure 2.2: Point-to-point channel example: secret-message capacity and secret-key capacity

2.4 Coding scheme

2.4.1 Principles, example

We design a two-phase scheme as follows:

1. *Key generation*: A common secret key is set up between Alice and Bob. For this purpose the optimal key generation scheme is used as described in Section 1.5. We define the exact parameters later.
2. *Encrypted message sending*: From the secret key, linear combinations are produced to obtain an encryption key. The encryption key is larger than the secret key and has the same size as the message. To encrypt the message the encryption key is used as one-time pad. The encrypted packets are then sent to Bob using ARQ, i.e., each packet is repeated until Bob receives it.

In the second phase Eve receives only a subset of the encrypted packets. As a result, although the encryption key has dependent components and thus it is not uniformly distributed, Eve receives a set of packets that are encrypted with a uniformly random key. In one possible interpretation, the message is partially secured by the secret key, and partially by the channel properties. This way a key that is smaller than the size of the message is sufficient for security.

Example In the following example we will assume that Alice and Bob know also Eve's channel state. This assumption is helpful in building intuition, but we stress that the actual scheme does not require state information from the eavesdropper.

An example run of our protocol is shown in Table 2.1. In our example, 8 transmissions are used to convey a message W that consists of three packets: W_1, W_2, W_3. Transmissions 1-3 belong to the key generation phase, while transmissions 4-8 constitute the message sending phase. In the key generation phase, a random packet becomes a key if only Bob receives it. In the second phase, if Bob does not receive a transmission, Alice simply repeats it regardless of whether or not Eve has received (transmissions 5 and 8). In transmission 6, the same key is used again as in transmissions 4-5. Since the previous two transmissions were not received by Eve, this does not risk security. Eve receives transmission 6, thus key K_1 is considered used and another key is used for the following transmission.

	Alice sends	Bob's channel	Eve's channel	Key available	Message decoded
1.	X_1 random	✓	×	$K_1 = X_1$	
2.	X_2 random	×	✓	K_1	
3.	X_3 random	✓	×	$K_1, K_2 = X_3$	
4.	$K_1 \oplus W_1$	×	×	K_1, K_2	
5.	$K_1 \oplus W_1$	✓	×	K_1, K_2	W_1
6.	$K_1 \oplus W_2$	✓	✓	K_2	W_1, W_2
7.	$K_2 \oplus W_3$	×	✓		W_1, W_2
8.	$K_2 \oplus W_3$	✓	✓		W_1, W_2, W_3

Table 2.1: Example scheme with Eve's channel state known. Erasure is indicated by ×, while ✓ denotes correct reception.

The example reveals that for security Alice and Bob needs only as many key packets as the number of distinct packets Eve receives from the ARQ transmissions. Knowing Eve's channel state it is easy to see which key packets can be securely reused. If Alice and Bob do not know exactly which packets Eve receives, but they know *how many*, then they can use coding to make sure Eve receives packets encrypted with independent keys. In our example, if they know that Eve receives two out of the three encrypted packets, they can use encryption keys K_1, K_2, $K_1 \oplus K_2$ to encrypt W_1, W_2, W_3 respectively. No matter which two packets Eve receives, the message remains secure.

To construct the encryption keys we need the following property. Assume Eve receives n_E packets out of the N encrypted packets. Then, we have a secret key of size n_E and a $n_E \times N$ size coefficient matrix is used to produce the encryption keys. For security, we need that any n_E size subset of the encryption keys form an independent set of packets. In other words, any n_E columns of the coefficient matrix has to be a full-rank matrix. By Theorems A.4-A.5 (see Appendix A.3) such a matrix is the generator of an MDS code.

From the channel parameters Alice and Bob can estimate well how many packets Eve receives. If the number of message packets is large, the number of received packets concentrates around

its expected value. The probability that Eve receives an encrypted packet sent using ARQ is:

$$(1-\delta_E)+\delta\delta_E(1-\delta_E)+(\delta\delta_E)^2(1-\delta_E)+\cdots=\frac{1-\delta_E}{1-\delta\delta_E}. \tag{2.5}$$

This suggests that to secure a rate R message, a secret key of rate $R\frac{1-\delta_E}{1-\delta\delta_E}$ is sufficient. This intuition is confirmed by formal analysis in Section 2.4.3.

2.4.2 Detailed description

The (n,ϵ,N) scheme uses n transmissions to convey a message of N packets. $W=(W_1,\ldots,W_N)$ denotes the row vector of the N message packets. We first calculate the following parameters:

$$n=n_1+n_2 \tag{2.6}$$

$$s=N\frac{1-\delta_E}{1-\delta\delta_E}+N^{\frac{3}{4}}\frac{1-\delta_E}{1-\delta\delta_E} \tag{2.7}$$

$$n_2=\frac{N}{1-\delta}+\frac{N^{\frac{3}{4}}}{1-\delta}, \tag{2.8}$$

where n_1 is the number of transmissions needed to create a key of size s with error parameter $\epsilon'<\epsilon$ (for the parameters of the key generation phase we refer the reader to (1.25)-(1.26)).

1. *Key generation phase*: Alice and Bob perform the key generation scheme described in Section 1.5 using n_1 transmissions. By the scheme $X_1,\ldots,X_{n_1}$ are random packets selected uniformly at random. If the key generation fails, declare an error, otherwise let K_{AB} denote the shared secret key. K_{AB} is a row vector of s packets.
2. *Encrypted message sending phase*: Alice computes the encryption key K'_{AB} as follows:

$$K'_{AB}=SG_{K'_{AB}}, \tag{2.9}$$

where $G_{K'_{AB}}$ is a $s\times N$ MDS generator matrix over $\mathbb{F}_q$. The encrypted message W' is computed as a one-time pad with encryption key K'_{AB}:

$$W'=W\oplus K'_{AB}. \tag{2.10}$$

Alice transmits the encrypted packets $W'=(W'_1,\ldots,W'_N)$ using ARQ. Let $X_{n_1+1}=W'_1$, then

$$\forall\, i\in\{2,\ldots,n_2\}: X_{n_1+i}=\begin{cases} X_{n_1+i-1}=W'_j & \text{if } F_{n_1+i-1}=\times \\ W'_{j+1} & \text{if } F_{n_1+i-1}=\checkmark \text{ and } j<N, \\ \text{None} & \text{otherwise,}\end{cases} \tag{2.11}$$

where j denotes the index of the last encrypted packet sent, $\times$ and $\checkmark$ denote erasure and correct reception as Bob's channel state. "None" means that Bob has already received all encrypted packets, but there are still some transmissions left and the channel remains idle. If Bob does not receive all encrypted packets, then an error is declared.

2.4.3 Analysis

In this subsection we prove the direct part of Theorem 2.1. We prove that for a large enough n our (n,ϵ,N) coding scheme satisfies Definition 1.5 showing the achievability of secret-message rate C_{SM}.

From the scheme description it is obvious that (2.1) is satisfied, we need to show security (2.3) and decodability (2.2).

Security

We assume that the key generation phase declared no error, otherwise the scheme is trivially secure. We observe that Θ_E is independent of any other variable and thus $I(W;Z^nF^n\Theta_E) = I(W;Z^nF^n)$. Hence, we can omit Θ_E from the analysis. We have already seen in Section 1.5 that the key generation phase is secure, i.e., for any $\epsilon' > 0$ and a large enough n_1

$$I\left(S;Z^{n_1}F^{n_1}\right) < \epsilon', \tag{2.12}$$

and K_{AB} is uniformly distributed. Let I_E denote the index set of the subset of encrypted packets that Eve receives, and W'^{I_E} the vector of encrypted packets restricted to columns defined by index set I_E. We use the same notation for W, $G_{K'_{AB}}$ also.

$$I\left(W;Z^nF^n\right) = I\left(W;Z^{n_1}W'^{I_E}F^n\right) = I\left(W;Z^{n_1}F^n\right) + I\left(W;W'^{I_E}|Z^{n_1}F^n\right) \tag{2.13}$$

$$\stackrel{(a)}{=} I\left(W;W'^{I_E}|Z^{n_1}F^n\right) = \sum_{i=0}^{N} I\left(W;W'^{I_E}|Z^{n_1}F^n,|I_E|=i\right)\Pr\{|I_E|=i\}, \tag{2.14}$$

where the step (a) follows from the fact that the channel state and the transmissions in the first phase are independent of W. We have that

$$H\left(W'^{I_E}|Z^{n_1}F^n,|I_E|=i\right) \le i \tag{2.15}$$

and

$$H\left(W'^{I_E}|WZ^{n_1}F^n,|I_E|=i\right) = H\left(W^{I_E}\oplus SG^{I_E}_{K'_{AB}}|WZ^{n_1}F^n,|I_E|=i\right) \tag{2.16}$$

$$= H\left(SG^{I_E}_{K'_{AB}}|WZ^{n_1}F^n,|I_E|=i\right) \stackrel{(a)}{=} H\left(SG^{I_E}_{K'_{AB}}|Z^{n_1},|I_E|=i\right) \tag{2.17}$$

$$= H\left(SG^{I_E}_{K'_{AB}}||I_E|=i\right) - I\left(SG^{I_E}_{K'_{AB}};Z^{n_1}||I_E|=i\right) \tag{2.18}$$

$$\ge H\left(SG^{I_E}_{K'_{AB}}||I_E|=i\right) - I\left(S;Z^{n_1}||I_E|=i\right) \tag{2.19}$$

$$\stackrel{(b)}{=} H\left(SG^{I_E}_{K'_{AB}}||I_E|=i\right) - I\left(S;Z^{n_1}\right) \tag{2.20}$$

$$\stackrel{(c)}{\ge} H\left(SG^{I_E}_{K'_{AB}}||I_E|=i\right) - \epsilon' \tag{2.21}$$

$$\stackrel{(d)}{=} \min\{s,i\} - \epsilon', \tag{2.22}$$

where in steps (a) and (b) we used again that transmissions in the first phase are independent of the message and the channel states in the second phase, step (c) is from (2.12), (d) follows from the MDS property of $G_{K'_{AB}}$ (see Corollary A.2 in Appendix A.3). Substituting (2.15) and (2.22) back to (2.14) we get:

$$I\left(W;Z^nF^n\right) \leq \epsilon' + \sum_{i=0}^{N} (i - \min\{s,i\}) \Pr\{|I_E| = i\} = \epsilon' + \sum_{i=0}^{N} \max\{0, i-s\} \Pr\{|I_E| = i\} \tag{2.23}$$

$$= \epsilon' + \sum_{i=s+1}^{N} i \Pr\{|I_E| = i\} \leq \epsilon' + N \Pr\{|I_E| > s\}. \tag{2.24}$$

Eve receives an encrypted packet with probability $\frac{1-\delta_E}{1-\delta\delta_E}$. The receptions of different encrypted packets are independent events, hence $|I_E|$ can be seen as a sum of N independent Bernoulli variables. Hence,

$$\Pr\{|I_E| > s\} = \Pr\left\{|I_E| > N\frac{1-\delta_E}{1-\delta\delta_E} + N^{\frac{3}{4}}\frac{1-\delta_E}{1-\delta\delta_E}\right\} = \Pr\left\{|I_E| - \mathbb{E}\{|I_E|\} > N^{\frac{3}{4}}\frac{1-\delta_E}{1-\delta\delta_E}\right\} \tag{2.25}$$

$$\leq \Pr\left\{||I_E| - \mathbb{E}\{|I_E|\}| > N^{\frac{3}{4}}\frac{1-\delta_E}{1-\delta\delta_E}\right\} \leq 2\mathrm{e}^{-a_{2.26}\sqrt{N}}, \tag{2.26}$$

where $a_{2.26} = 2\left(\frac{1-\delta_E}{1-\delta\delta_E}\right)^2$ is constant from the Chernoff-Hoeffding bound (see Appendix A.2). Substituting back to (2.24):

$$I\left(W;Z^nF^n\right) \leq \epsilon' + 2N\mathrm{e}^{-a_{2.26}\sqrt{N}}. \tag{2.27}$$

From the properties of the key generation phase it follows that ϵ' can be arbitrary small if n_1 is large enough. Also, $N \to \infty$ when $n \to \infty$, hence a large enough n ensures that the above term is smaller than any $\epsilon > 0$ if n is large enough. By this we have shown that the scheme provides security.

Reliability

Clearly, if no error is declared Bob is able to decode W. We have already seen in Section 1.5 that the key generation phase succeeds with arbitrarily small error probability. The second phase fails if Bob does not receive N packets in n_2 transmissions. This error event is similar in nature to the error of the key generation phase and we use the same technique to upper bound its probability. Let $|I_B|$ denote the number of transmissions for which no erasure happens for in the second phase. Error occurs of $|I_B| < N$. Similarly to $|I_E|$, $|I_B|$ is the sum of n_2 Bernoulli variables. Thus,

$$\Pr\{N > |I_B|\} = \Pr\left\{\mathbb{E}\{|I_B|\} - |I_B| > N^{\frac{3}{4}}\right\} \leq \Pr\left\{|\mathbb{E}\{|I_B|\} - |I_B|| > N^{\frac{3}{4}}\right\} \leq 2\exp\left\{-\frac{2N^{\frac{3}{2}}}{n_2}\right\} \tag{2.28}$$

$$\leq 2\mathrm{e}^{-2(1-\delta)\sqrt{N}}. \tag{2.29}$$

The probability of error can be made arbitrarily small by selecting n (and thus N) to be large enough.

Rate calculation

The rate of the above scheme is $\frac{N}{n}$, thus it proves the achievability of secret-message rate $R = \lim_{n\to\infty} \frac{N}{n}$. From the parameter definitions and from the properties of the key generation phase we get for the required key rate:

$$\lim_{n\to\infty} \frac{s}{n} = R\frac{1-\delta_E}{1-\delta\delta_E}. \tag{2.30}$$

Further,

$$\lim_{n\to\infty} \frac{s}{n_1} = \delta_E(1-\delta) \tag{2.31}$$

$$\lim_{n\to\infty} \frac{n_1}{n} = \frac{R(1-\delta_E)}{(1-\delta\delta_E)\delta_E(1-\delta)} \tag{2.32}$$

$$\lim_{n\to\infty} \frac{n_2}{n} = \frac{R}{1-\delta} \tag{2.33}$$

$$1 = \lim_{n\to\infty} \frac{n_1+n_2}{n} = \frac{R(1-\delta_E)}{(1-\delta\delta_E)\delta_E(1-\delta)} + \frac{R}{1-\delta}. \tag{2.34}$$

By rearranging terms in (2.34) we get the claimed expression for the achievable rate by the scheme. □

We also note that the computational complexity of our scheme is polynomial in n. Computationally the most costly operation is the matrix multiplication that is feasible in $O(n^3)$.

2.5 Outer bound

In this section we prove the converse part of Theorem 2.1 by deriving a new outer bound on the achievable secret-message rate. Throughout the proof we assume that the feedback F_i contains Eve's channel state also. Clearly, this information can only help Alice and Bob to achieve a higher rate, hence the derived outer bound is valid for our setting. We use the notation $F_i \in \{B, E, BE, \emptyset\}$ to indicate that the ith transmission was received by "only Bob", "only Eve", "both of them" and "none of them" respectively. Since the bound should hold for any message distribution, we might assume that W is uniformly distributed.

Proof. We have the following inequality:

$$n \geq \sum_{i=1}^{n} H(X_i) \geq H\left(X_i|Y^{i-1}F^{i-1}\right) = \sum_{i=1}^{n} H\left(X_i|Y^{i-1}F^{i-1}W\right) + I\left(X_i; W|Y^{i-1}F^{i-1}\right) \tag{2.35}$$

$$\geq \sum_{i=1}^{n} H\left(X_i|Y^{i-1}Z^{i-1}F^{i-1}W\right) + I\left(X_i; W|Y^{i-1}F^{i-1}\right). \tag{2.36}$$

The following two lemmas give bounds on the last two terms.

Lemma 2.1. *In the defined point-to-point communication model, for any achievable secret-message rate R*

$$\sum_{i=1}^{n} H\left(X_i|Y^{i-1}Z^{i-1}F^{i-1}W\right) \geq \frac{nR(1-\delta_E)}{(1-\delta\delta_E)\delta_E(1-\delta)} - n\mathscr{E}_{2.1}, \tag{2.37}$$

where $\mathscr{E}_{2.1} = \frac{\epsilon}{\delta_E(1-\delta)} + (h_2(\epsilon)+R\epsilon)\frac{1-\delta_E}{\delta_E(1-\delta)(1-\delta\delta_E)}$.

Lemma 2.2. *In the defined point-to-point communication model, for any achievable secret-message rate R*

$$\sum_{i=1}^{n} I\left(X_i; W|Y^{i-1}F^{i-1}\right) \geq \frac{nR}{1-\delta} - n\mathscr{E}_{2.2}, \tag{2.38}$$

where $\mathscr{E}_{2.2} = \frac{h_2(\epsilon)+R\epsilon}{1-\delta}$.

We provide proofs in the next subsection. By substituting back the results of the above two lemmas to (2.36) and dividing by n the proof is complete. We leave the final steps as an exercise for the reader. Note that $\mathscr{E}_{2.1}, \mathscr{E}_{2.2}$ are vanishing error terms, since ϵ is arbitrarily small. □

Although the outer bound holds for any secure coding scheme, (2.36) reflects the structure of (2.34) and the phases of the achievability scheme. Indeed, $\frac{nR}{1-\delta}$ is the (expected) number of transmissions required to complete the message sending phase, while as we have seen, a rate $R\frac{1-\delta_E}{1-\delta\delta_E}$ key is used by the scheme which requires (in expectation) $\frac{nR(1-\delta_E)}{(1-\delta\delta_E)\delta_E(1-\delta)}$ transmissions. In this interpretation, Lemma 2.1 lower bounds the amount of required secret key, i.e., the minimum length of the key generation phase. This is done through drawing a balance between the generated and the used keys, which is a possible interpretation of inequality (2.45) (see in the proof). Once again, this interpretation of information terms is attached to our scheme, but the converse proof is general for any scheme.

2.5.1 Proofs of Lemmas 2.1-2.2

We start with proving Lemma 2.2.

Proof of Lemma 2.2

We observe that the private randomness of the receiver does not help in decoding, because the protocol run does not depend on it. One can formally argue this by observing that

$$I(W; Y^n, F^n, \Theta_B) = I(W; Y^n, F^n). \tag{2.39}$$

In other words, Bob must be able to decode also without using Θ_B. If R is the rate of a coding scheme, then

$$nR = H(W) = H(W) - H\left(W|Y^nF^n\right) + H\left(W|Y^nF^n\right) \overset{(a)}{\leq} I\left(Y^nF^n;W\right) + h_2(\epsilon) + nR\epsilon \tag{2.40}$$

$$= \sum_{i=1}^{n} I\left(Y_iF_i;W|Y^{i-1}F^{i-1}\right) + h_2(\epsilon) + nR\epsilon \overset{(b)}{=} \sum_{i=1}^{n} I\left(Y_i;W|Y^{i-1}F^{i-1}F_i\right) + h_2(\epsilon) + nR\epsilon \tag{2.41}$$

$$\overset{(c)}{=} \sum_{i=1}^{n} I\left(Y_i;W|Y^{i-1}F^{i-1},F_i\in\{B,BE\}\right)\Pr\{F_i\in\{B,BE\}\} + h_2(\epsilon) + nR\epsilon \tag{2.42}$$

$$= (1-\delta)\sum_{i=1}^{n} I\left(X_i;W|Y^{i-1}Z^{i-1}F^{i-1}\right) + h_2(\epsilon) + nR\epsilon \tag{2.43}$$

$$\leq (1-\delta)\sum_{i=1}^{n} I\left(X_i;W|Y^{i-1}Z^{i-1}F^{i-1}\right) + n(1-\delta)\mathscr{E}_{2.2}. \tag{2.44}$$

Step (a) follows from the decodability condition 2.2 using Fano's inequality (see e.g., Theorem 2.47 in [2]) and (2.39). We have (b) from the fact that the channel state F_i is independent of $\left(W,Y^{i-1},F^{i-1}\right)$. Step (c) holds, because the mutual information term gives zero, if $Y_i=\perp$. □

Proof of Lemma 2.1

To complete the proof, we show the following two inequalities:

$$\sum_{i=1}^{n}\delta_E(1-\delta)H\left(X_i|Y^{i-1}Z^{i-1}F^{i-1}W\right) \geq \sum_{i=1}^{n}(1-\delta_E)I\left(X_i;Y^{i-1}|Z^{i-1}F^{i-1}W\right) \tag{2.45}$$

$$\sum_{i=1}^{n} I\left(X_i;Y^{i-1}|Z^{i-1}F^{i-1}W\right) \geq \frac{nR}{1-\delta\delta_E} - \frac{\delta_E(1-\delta)}{1-\delta_E}n\mathscr{E}_{2.1}. \tag{2.46}$$

For (2.45):

$$0 \leq H\left(Y^n|Z^nF^nW\right) = H\left(Y^{n-1}|Z^nF^nW\right) + H\left(Y_n|Y^{n-1}Z^nF^nW\right) \tag{2.47}$$

$$= H\left(Y^{n-1}|Z^{n-1}F^{n-1}W\right) - I\left(Z_nF_n;Y^{n-1}|Z^{n-1}F^{n-1}W\right) + H\left(Y_n|Y^{n-1}Z^nF^nW\right) \tag{2.48}$$

$$\overset{(a)}{=} H\left(Y^{n-1}|Z^{n-1}F^{n-1}W\right) - I\left(Z_n;Y^{n-1}|Z^{n-1}F^nW\right) + H\left(Y_n|Y^{n-1}Z^nF^nW\right) \tag{2.49}$$

$$\overset{(b)}{=} H\left(Y^{n-1}|Z^{n-1}F^{n-1}W\right) - I\left(Z_n;Y^{n-1}|Z^{n-1}F^{n-1}W,F_n\in\{E,EB\}\right)\Pr\{F_n\in\{E,EB\}\}$$
$$+ H\left(Y_n|Y^{n-1}Z^nF^{n-1}W,F_n=B\right)\Pr\{F_n=B\} + H\left(Y_n|Y^{n-1}Z^nF^{n-1}W,F_n=EB\right)\Pr\{F_n=EB\} \tag{2.50}$$

$$= H\left(Y^{n-1}|Z^{n-1}F^{n-1}W\right) - (1-\delta_E)I\left(X_n;Y^{n-1}|Z^{n-1}F^{n-1}W\right)$$
$$+ (1-\delta)\delta_E H\left(X_n|Y^{n-1}Z^{n-1}F^{n-1}W\right) + (1-\delta)(1-\delta_E)H\left(X_n|Y^{n-1}Z^{n-1}X_nF^{n-1}W\right) \tag{2.51}$$

$$= H\left(Y^{n-1}|Z^{n-1}F^{n-1}W\right) - (1-\delta_E)I\left(X_n;Y^{n-1}|Z^{n-1}F^{n-1}W\right)$$
$$+ (1-\delta)\delta_E H\left(X_n|Y^{n-1}Z^{n-1}F^{n-1}W\right). \tag{2.52}$$

In (a) the independence property of the channel states is used, while (b) holds because in case of erasure the terms in question are zero. Doing the same steps recursively gives (2.45).

To get (2.46), we first consider the security condition. In the outer bound proof, we use the weak form of security (see Section 1.4.2 for discussion about different security notions):

$$n\epsilon > I\left(Z^n F^n; W\right) = \sum_{i=1}^{n} I\left(Z_i F_i; W | Z^{i-1} F^{i-1}\right) = \sum_{i=1}^{n} I\left(Z_i; W | Z^{i-1} F^{i-1} F_i\right) \tag{2.53}$$

$$= \sum_{i=1}^{n} I\left(Z_i; W | Z^{i-1} F^{i-1} F_i \in \{E, EB\}\right) \Pr\{F_i \in \{E, EB\}\} = \sum_{i=1}^{n} (1-\delta_E) I\left(X_i; W | Z^{i-1} F^{i-1}\right). \tag{2.54}$$

From Fano's inequality we have that

$$nR \le I\left(Y^n Z^n F^n; W\right) + h_2(\epsilon) + nR\epsilon \tag{2.55}$$

$$\overset{(a)}{=} h_2(\epsilon) + nR\epsilon + \sum_{i=1}^{n} (1-\delta\delta_E) I\left(X_i; W | Y^{i-1} Z^{i-1} F^{i-1}\right) \tag{2.56}$$

$$\le h_2(\epsilon) + nR\epsilon + \sum_{i=1}^{n} (1-\delta\delta_E) I\left(X_i; W | Z^{i-1} F^{i-1}\right) + (1-\delta\delta_E) I\left(X_i; Y^{i-1} | Z^{i-1} F^{i-1} W\right) \tag{2.57}$$

$$\overset{(b)}{<} h_2(\epsilon) + nR\epsilon + \frac{1-\delta\delta_E}{1-\delta_E} n\epsilon + \sum_{i=1}^{n} (1-\delta\delta_E) I\left(X_i; Y^{i-1} | Z^{i-1} F^{i-1} W\right) \tag{2.58}$$

$$\le \frac{(1-\delta\delta_E)\delta_E(1-\delta)}{1-\delta_E} n\mathscr{E}_{2.1} + \sum_{i=1}^{n} (1-\delta\delta_E) I\left(X_i; Y^{i-1} | Z^{i-1} F^{i-1} W\right). \tag{2.59}$$

In (a) we used a derivation similar to Lemma 2.2. To avoid repetition we omitted the details. In step (b) we used (2.54). □

2.6 Linear programming formulation

We provide an alternative formulation of Theorem 2.1 that gives the secret-message capacity as a solution of a linear program. First, we state a theorem for the message sending phase. The theorem is a direct consequence of our achievability proof in Section 2.4.3, noticing that the second phase of our scheme uses the secret key as input, and it does not depend on how the key is generated.

Theorem 2.2. *Consider an erasure channel with state-feedback as defined in our communication model. Assume the sender and the receiver have access to a uniform random key K_{AB} of rate R_K such that for an arbitrarily small ϵ'*

$$I(K_{AB}; E|W) < \epsilon', \tag{2.60}$$

where W denotes the message to be sent and E denotes all the random variables the eavesdropper observes before starting transmissions. Then, a secret-message rate $\min\left\{1-\delta, R_K \frac{1-\delta\delta_E}{1-\delta_E}\right\}$ is achievable. Further, the claimed rate is achievable by the message sending phase of the scheme described in Section 2.4.

The linear programming formulation of our result is the following:

Theorem 2.3. *The secret-message capacity of the point-to-point erasure channel with state feedback is the solution of the following linear program (LP), where parameters $m, k \geq 0$.*

$$\max R, \quad \textit{such that:}$$

$$R \leq (1-\delta)m \tag{2.61}$$

$$m(1-\delta)\frac{1-\delta_E}{1-\delta\delta_E} \leq k\delta_E(1-\delta) \tag{2.62}$$

$$1 \geq m + k. \tag{2.63}$$

We identify constraint (2.61) as a *rate constraint*, constraint (2.62) as a *security constraint* and constraint (2.63) as a *time-sharing constraint*. In the sequel (in Chapter 4) we will see LPs that have a similar structure.

The above LP has a direct closed form solution, which is the same as the secret-message capacity claimed by Theorem 2.1, so no proof is needed beyond what is already shown. Still, we make a reverse argument to show that the solution of the LP gives directly the parameters of a scheme.

Assume the above LP is feasible with some parameter values m, k. Then, nk is the length of a key generation phase, while nm is the length of a message sending phase. Constraint (2.63) ensures that n transmissions suffice. The key generation phase enables to set up a key of rate $k\delta_E(1-\delta)$, while in the message sending phase we have the possibility to transmit a message of rate at most $(1-\delta)m$ (see (2.61)). By Theorem 2.2, constraint (2.62) ensures that the available secret-key rate is sufficient to secure a message of this rate. Thus, a feasible LP gives rise to the parameters of a secure coding scheme that provides achievability of a secret-message rate R.

At this point the LP formulation might seem to be a complicated way of describing formula (2.4), but at the same time it explicitly reflects the components that our scheme is built of. This property will enable us to use key generation and Theorem 2.2 as building blocks to design schemes described through linear programs for more general settings. In Chapter 4 we use this approach extensively.

2.7 Next steps

We investigated the point-to-point erasure channel with state-feedback and provided a complete characterization of its secret-message capacity. We introduced a two-phase coding scheme that achieves capacity in this setting. Our result clearly shows the role of feedback as well as the difference between the secret-key generation and the secret-message sending problem.

In the following chapters we generalize this result in various directions. We investigate the secret-message sending problem for a broadcast channel (Chapter 3) and also for multihop networks (Chapter 4-5).

3 Secret-message capacity of a broadcast channel

This chapter is devoted to the 1-to-M broadcast channel. The broadcast channel models a situation that is common in a wireless environment, namely when the sender, Alice can simultaneously transmit to M receivers. As an example, one can think of an access point with multiple terminals connected to it. In this setting – unless the communication is synchronized – nodes cannot know in advance if they are the intended recipient of a packet. For example, according to the WiFi standard (IEEE 802.11), nodes check the destination MAC address after reception and drop packets not destined to them. This mode of operation makes clear that – apart from some storage space – storing packets not intended for the given node requires no overhead from the devices. Results on network coding have shown that storing such packets is not useless, they can be used as side information to decode a subsequent encoded packet (e.g., [38–41]). In this chapter we make use of side information to maximize efficiency, but at the same time we also ensure message security.

In this chapter message secrecy is considered against the receivers instead of an outsider eavesdropper. Alice aims to send a private message to all receivers, such that they do not learn each other's message. We define two adversary models. An *honest-but-curious* adversary follows the protocol, which in our setting means that she sends honest acknowledgments about her channel state. In contrast, a *dishonest* adversary might lie about her receptions in order to gain information about the message of an other receiver. The crucial difference between the two models is that a dishonest adversary can influence – to some extent – the run of the protocol.

We propose a secure coding scheme for the broadcast channel that is optimal in the honest-but-curious adversary model in all the cases where a non-secure capacity achieving scheme is available. In order to show optimality, we derive a new outer bound for the secret-message capacity region. As a side result, our proof offers a new proof for the known outer bound of the non-secure capacity region derived in [39, 42].

We consider the dishonest adversary model in the special case of two receivers. We show that securing private messages against a dishonest adversary does not come at a compromise in rate, the secret-message capacity region does not change when the stronger adversary is

considered. Thus, the optimality of our scheme is immediate from the optimality result for the honest-but-curious case.

The security notions are adapted to the multiuser setting and also to the dishonest adversary model. We make the observation that in the dishonest adversary model security relies on the independence and on the uniform distribution of the private messages. In some communication scenarios this assumption might be restrictive, hence we introduce the notion of distribution independent security that requires secrecy without any assumption on the joint distribution of the messages. We also provide a scheme that satisfies this stronger security requirement.

3.1 Related work

Besides the results on secure communication summarized in the previous chapter, the most relevant results are on non-secure communication over a broadcast channel, and more specifically over a broadcast erasure channel with state-feedback. The characterization of a broadcast channel is significantly harder than dealing with the point-to-point counterpart. The capacity region of a discrete memoryless broadcast channel (DMBC) is still an open problem (for partial results see e.g., [43–45]). In the case of a point-to-point discrete memoryless channel feedback does not increase capacity [6], however, in some cases feedback increases the capacity of a broadcast channel [46–48]. The broadcast erasure channel with state-feedback is a special case of a DMBC, where capacity results are available [39, 40, 42, 49]. Our focus is on this type of broadcast channel.

Network coding has been used in several capacity achieveing coding schemes [39, 40, 42, 49]. An optimal scheme that uses linear network coding is available for the ≤ 3 receiver broadcast erasure channel with state-feedback under arbitrary channel parameters. Network coding has proven to be a powerful technique to achieve capacity in various network settings (e.g., [50–52]). As investigated in [53, 54], feedback offers benefits in certain scenarios where network coding is used, it is the case in our communication setting also.

In our secure coding scheme we are going to use the non-secure capacity achieving coding scheme proposed in [42]. We briefly summarize the scheme in Section 3.3.

3.2 Model

3.2.1 Honest-but-curious adversary

For the honest-but-curious adversary model we can naturally adapt our definitions in Section 1.2. Before each transmission, Alice has access to her private randomness, the messages $W_1, \ldots, W_M$ for each receiver and the previous channel states. We use the shorthand $W = (W_1, \ldots, W_M)$, thus in Definition 1.1 $A_{i-1} = \left(W, \Theta_A, F^{i-1}\right)$ and hence (1.6) becomes:

$$X_i = f_i\left(W, \Theta_A, F^{i-1}\right), \quad i = 1, 2, \ldots, n. \tag{3.1}$$

We have a decodability condition for each receiver ((1.7) in Definition 1.1):

$$\forall\ 1 \le j \le M : \Pr\left\{\phi_j\left(Y_j^n F^n \Theta_j\right) \neq W_j\right\} < \epsilon, \tag{3.2}$$

We require secrecy of any private message even if the adversary has access to the observations of all other receivers. Hence, our security criterion in Definition 1.4 becomes:

$$\forall\, j \in \{1 \dots M\} :\ I\left(W_j; Y_{-j}^n F^n \Theta_{-j}\right) < \epsilon, \tag{3.3}$$

where Y_{-j}^n stands for $Y_1^n, \dots, Y_{j-1}^n, Y_{j+1}^n, \dots, Y_M^n$ and similarly Θ_{-j} is $\Theta_1, \dots, \Theta_{j-1}, \Theta_{j+1}, \dots, \Theta_M$ denoting the private randomnesses of the receivers other than j.

To make a clear distinction, we denote the secret-message capacity region in the honest-but-curious adversary model as $\mathscr{R}_H^M$.

Following [42], we distinguish two special cases.

Definition 3.1. *We call the channel* symmetric *if the erasure probabilities are all the same:* $\delta_i = \delta_j, \forall\ 1 \le i, j \le M$.

Definition 3.2. *We call a rate vector one-sidedly fair if* $\delta_i \ge \delta_j$ *implies* $R_i \delta_i \ge R_j \delta_j$. *The set of one-sidedly fair rate tuples for a given channel is denoted by* $\Lambda_{\text{osf}}^M \subset \mathbb{R}_+^M$.

3.2.2 Dishonest adversary

For a dishonest adversary, we need a slight modification in Definitions 1.1, 1.4-1.6. This is because we do not aim to provide any guarantee – either reliability or security – for a dishonest party. However, an honest receiver should suffer no harm. Also, as we will see, the distribution of the messages plays an important role in security. First, we assume that the message of the dishonest receiver is uniformly distributed and independent of the other receiver's message. To relax this constraint, we introduce the notion of *distribution independent security*, which requires security independently of the joint message distribution.

We provide definitions for two receivers, Bob and Calvin. This time Alice has no access to the true channel states, only the potentially dishonest acknowledgments. We denote F_i^* the acknowledgments in time slot i. A dishonest receiver can do the following: he can (a) select the marginal distribution of the other user's message arbitrarily; his own message is assumed to be independent of the other user's message and uniformly distributed over its alphabet and the dishonest user does not have a priori access to his own message, and (b) produce dishonest acknowledgments as a (potentially randomized) function of all the information he has access to when producing each acknowledgment (this includes all the packets and the pattern of erasures he received up to and including the current packet he is acknowledging and the acknowledgments sent by the other user over the public channel up to the previous packet). In the sequel σ denotes the acknowledging strategy of the adversary. The definitions below are specific to this chapter.

Thus, $A_{i-1} = \left(W, \Theta_A, F^{*\,i-1}\right)$ and:

Definition 3.3. *An* (n,ϵ,N_1,N_2) *coding scheme* against a dishonest adversary *consists of the following components: (a) message alphabets* $\mathcal{W}_j = \mathbb{F}_q^{LN_j}$, $j=1,2$, *(b) encoding maps* $f_i(\cdot)$, $i=1,2,\ldots,n$, *and (c) decoding maps* $\phi_j(\cdot)$, $j=1,2$ *for each receiver, such that the inputs to the channel are*

$$X_i = f_{k,i}\left(W,\Theta_A,F^{*\,i-1}\right), \quad i=1,2,\ldots,n. \tag{3.4}$$

It provides reliability for each honestly acknowledging *receiver. That is, in case Bob acknowledges honestly*

$$\Pr\left\{\phi_1\left(Y_1^n F^{*\,n}\Theta_B\right)\neq W_1\right\}<\epsilon \tag{3.5}$$

is satisfied and if Calvin acknowledges honestly then

$$\Pr\left\{\phi_j\left(Y_2^n F^{*\,n}\Theta_C\right)\neq W_2\right\}<\epsilon \tag{3.6}$$

is satisfied.

Definition 3.4. *A* (n,ϵ,N_1,N_2) *coding scheme as defined by Definition 3.3 is* secure against a dishonest adversary*, if in addition to reliability it also provides secrecy for each* honestly acknowledging *receiver. That is, in case Bob acknowledges honestly*

$$\max_{P_{W_1},\sigma} I\left(W_1;Y_2^n F^n\Theta_C\right)<\epsilon \tag{3.7}$$

is satisfied and if Calvin acknowledges honestly then

$$\max_{P_{W_2},\sigma} I\left(W_2;Y_1^n F^n\Theta_B\right)<\epsilon \tag{3.8}$$

is satisfied under the assumption that the message of a dishonest receiver is uniformly distributed and is independent of the other message. The maxima are taken over all possible adversarial acknowledging strategies.

Definition 3.5. *A secret-message rate tuple* $(R_1,\ldots,R_M)\in\mathbb{R}_+^M$ *is achievable against a dishonest adversary, if for every* $\epsilon>0$ *there exists an* (n,ϵ,N_1,N_2) *coding scheme that is secure against a dishonest adversary for which*

$$R_j-\epsilon<\frac{1}{n}N_j, \quad \forall\, j\in\{1,2\}. \tag{3.9}$$

Definition 3.6. *The secret-message capacity region* in the dishonest adversary model $\mathscr{R}_{DH}^2\subset\mathbb{R}_+^2$ *is the set of secret-message rate tuples* (R_1,R_2) *that are achievable against a dishonest adversary.*

When defining security against a dishonest receiver, we assumed that the dishonest receiver cannot control his own message distribution. Relaxing this assumption leads to a stronger notion of security.

Definition 3.7. *A (n,ϵ,N_1,N_2) coding scheme as defined by Definition 3.3 is secure against a dishonest adversary* independently of the message distribution*, if in addition to reliability it also provides secrecy for each* honestly acknowledging *receiver irrespective of the joint distribution of the messages W_1,W_2. That is, in case Bob acknowledges honestly*

$$\max_{P_{W_1,W_2},\sigma} I\left(W_1;Y_2^nF^n\Theta_C|W_2\right)<\epsilon \tag{3.10}$$

is satisfied and if Calvin acknowledges honestly then

$$\max_{P_{W_1,W_2},\sigma} I\left(W_2;Y_1^nF^n\Theta_B|W_1\right)<\epsilon \tag{3.11}$$

is satisfied. The maxima are taken over all possible joint message distributions and all possible adversarial acknowledging strategies.

We refer to security definitions (3.10)-(3.11) as *distribution independent security.*

Definition 3.8. *A secret-message rate tuple $(R_1,\ldots,R_M)\in\mathbb{R}_+^M$ is achievable against a dishonest adversary independently of the message distribution, if for every $\epsilon>0$ there exists an (n,ϵ,N_1,N_2) coding scheme that satisfies Definition 3.7 and for which*

$$R_j-\epsilon<\frac{1}{n}N_j,\quad \forall\, j\in\{1,2\}. \tag{3.12}$$

Definition 3.9. *The* distribution independent *secret-message capacity region $\mathscr{R}_{DIS}^2\subset\mathbb{R}_+^2$ is the set of achievable secret-message rate tuples against a dishonest adversary independently of the message distribution.*

We also adapt the definition of semantic security. The current definition of semantic security (1.21) considers only the distribution of one message. Applying the definition directly to the multiuser setting thus implicitly assumes that a dishonest adversary has no control over its own message distribution. This definition of security matches Definition 3.4, however the implicit assumption might be too restrictive. We thus extend the definition of semantic security for two receivers. We give definitions assuming Calvin is the dishonest receiver, symmetric versions hold for security against Bob. We define the adversarial advantage in this case as:

$$\begin{aligned}\mathbf{Adv}_{\text{dis}}^{\text{ss}}=\max_{g,P_{W_1,W_2},\sigma}\Big\{&\max_{\mathscr{A}}\Pr\left\{\mathscr{A}(Y_2^n,F^n,\sigma,W_2)=g(W_1,W_2)\right\}\\&-\max_{\mathscr{S}}\Pr\left\{\mathscr{S}(P_{W_1,W_2},g,W_2)=g(W_1,W_2)\right\}\Big\}.\end{aligned} \tag{3.13}$$

Note that here we let the simulator to have access to the message W_2 which an honestly acknowledging Calvin will learn. The corresponding definition of adversarial advantage for

distribution independent security in information-theoretic terms matches Definition 3.7:

$$\mathbf{Adv}_{\mathrm{dis}}^{\mathrm{mis}} = \max_{P_{W_1,W_2},\sigma} I\left(W_1; Y_2^n F^n \Theta_2 | W_2\right). \tag{3.14}$$

3.3 Non-secure 1-to-M broadcast

Before summarizing our results we review the non-secure message sending problem and the coding schemes proposed in [39, 42]. First, we restate the result that characterizes the non-secure capacity region. Let π denote a permutation of $\{1,2,\ldots,M\}$ and π_i the ith element of the permutation.

Theorem 3.1. *For $M \leq 3$ or for a symmetric channel with $M > 3$, the capacity region $\mathscr{R}^M$ of the 1-to-M broadcast erasure channel with state-feedback is characterized by the following inequality:*

$$\max_{\pi} \sum_{i=1}^{M} \frac{R_{\pi_i}}{1 - \prod_{k=1}^{i} \delta_{\pi_k}} \leq 1, \tag{3.15}$$

where the maximization is taken over all permutations π of $\{1,\ldots,M\}$. Furthermore, if a rate tuple $(R_1,\ldots,R_M) \in \Lambda_{\mathrm{osf}}^M$, then $(R_1,\ldots,R_M) \in \mathscr{R}^M$ if and only if (3.15) *is satisfied.*

Further, it is known [42, 49] that (3.15) is an outer-bound for $\mathscr{R}^M$ in all cases.

Theorem 3.2. *Any rate tuple $(R_1,\ldots,R_M) \in \mathscr{R}^M$ satisfies* (3.15).

We illustrate the proposed scheme for two receivers, Bob and Calvin. Alice sends messages of rate R_1 and R_2 to Bob and Calvin respectively. We refer to the corresponding messages W_1 and W_2 as messages intended for Bob/Calvin or simply Bob's/Calvin's message. According to Theorem 3.2 R_1 and R_2 are such that

$$\frac{R_1}{1-\delta_1} + \frac{R_2}{1-\delta_1\delta_2} \leq 1, \tag{3.16}$$

$$\frac{R_1}{1-\delta_1\delta_2} + \frac{R_2}{1-\delta_2} \leq 1. \tag{3.17}$$

The scheme has two steps:

Step a) Alice repeats every message packet intended for Bob until either Bob or Calvin correctly receives it. She then repeats every message packet intended for Calvin until either Bob or Calvin correctly receives.

Step b) Alice sends the XOR of one of Bob's message packet that only Calvin received and one of Calvin's message packets that only Bob received. The XOR-ed packet simultaneously carries a message packet for both Bob and Calvin. Upon reception, either is able to receive a new message packet. Alice keeps sending such XOR-ed packets until one of the receivers receives all his message. After that, she simply repeats the not yet received message packets until the other receiver is also satisfied.

The use of network coding in Step b) enables us to make use of the side information collected in Step a). The coded transmissions are equally useful for both receivers and make the scheme more efficient than simple time-sharing. The first step takes (in expectation) $\frac{n(R_1+R_2)}{1-\delta_1\delta_2}$ transmissions, while finishing the second step takes $\max\left\{\frac{nR_1}{1-\delta_1}-\frac{nR_1}{1-\delta_1\delta_2},\frac{nR_2}{1-\delta_2}-\frac{nR_2}{1-\delta_1\delta_2}\right\}$. The achieved rate region thus matches (3.16)-(3.17).

The scheme can be generalized for more than two receivers [42]. The proposed schemes have the following structure:

Step a) Alice repeats each message packet until at least one of the three receivers correctly receives it.

Step b) Alice sends linear combinations of the packets that are not received by their intended receiver in Step a).

A key contribution of [42] is in specifying how to construct the linear combinations in Step b). We refer the reader to [42] for the exact constructions, and highlight here the two important properties that we rely on:

- A message packet successfully delivered to its intended receiver in Step a) is never used in Step b).
- For $M=3$, or for a symmetric channel, or for a one-sidedly fair rate tuple the proposed scheme achieves any rate point within the region in (3.15).

3.4 Main results

3.4.1 Honest-but-curious adversary

Our main result for honest-but-curious receivers is the characterization of the secret-message capacity region for sending private messages to M receivers over a broadcast erasure channel, for all the cases where the non-secure capacity region has been characterized, namely, the 2-receiver, 3-receiver, symmetric M-receiver and one-sidedly fair M-receiver cases. For all the mentioned cases, when the capacity region $\mathscr{R}^M$ is known, we prove the following theorem which describes the corresponding secret-message capacity region $\mathscr{R}_H^M$.

Theorem 3.3. *For $M\leq 3$ or for a symmetric channel with $M>3$, the secret-message capacity region $\mathscr{R}_H^M$ is characterized by the following inequality:*

$$\max_{j\in\{1,\dots,M\}}\frac{R_j\left(1-\frac{\prod_{k=1}^M\delta_k}{\delta_j}\right)}{(1-\delta_j)\frac{\prod_{k=1}^M\delta_k}{\delta_j}\left(1-\prod_{k=1}^M\delta_k\right)}+\max_{\pi}\sum_{i=1}^M\frac{R_{\pi_i}}{1-\prod_{k=1}^i\delta_{\pi_k}}\leq 1, \tag{3.18}$$

where the second maximization is taken over all permutations π of $\{1,\dots,M\}$. Furthermore, if a rate tuple $(R_1,\dots,R_M)\in\Lambda_{\text{osf}}^M$, then $(R_1,\dots,R_M)\in\mathscr{R}_H^M$ if and only if (3.18) *is satisfied.*

We prove the achievability part of Theorem 3.3 constructively by describing a coding scheme that achieves any rate tuple in $\mathscr{R}_H^M$ in the mentioned cases. The scheme together with the proof of its properties are given in Section 3.5.1.

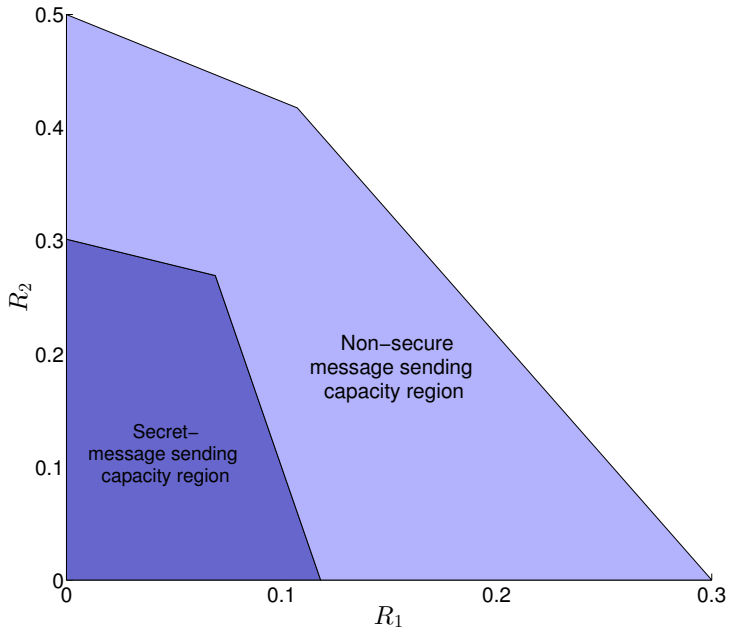


Figure 3.1: Non-secure message sending and secret-message sending capacity regions for $M = 2$, $\delta_1 = 0.7$, $\delta_2 = 0.5$.

We also develop a new converse proof to show that the scheme is optimal. We provide the converse proof in Section 3.6, which completes the proof of Theorem 3.3. Our converse proof inherently provides a new proof of Theorem 3.2. This new proof might be of interest on its own, hence we provide it separately in Section 3.6.1.

Comparing regions $\mathscr{R}^M$ and $\mathscr{R}^M_H$, the first term in (3.18) can be interpreted as the overhead for security. Indeed, in the scheme we present, there is a key generation phase whose duration is proportional to this term. In Figure 3.1 we visualize this overhead for some parameter values.

3.4.2 Dishonest adversary

For the case of a dishonest receiver, we characterize the rate region $\mathscr{R}^2_{DH}$. In particular, we show that $\mathscr{R}^2_{DH} = \mathscr{R}^2_H$, i.e., the same rates are achievable against dishonest receivers as against honest-but-curious receivers. We provide a formal description and proof for $M = 2$. The same ideas can be easily extended for the cases where Theorem 3.3 holds.

Theorem 3.4. *The rate region $\mathscr{R}^2_{DH}$ as defined in Definition 3.6 is the set of all rate pairs $(R_1, R_2) \in \mathbb{R}^2_+$ which satisfy the following two inequalities:*

$$\frac{R_1(1-\delta_2)}{\delta_2(1-\delta_1)(1-\delta_1\delta_2)} + \frac{R_1}{1-\delta_1} + \frac{R_2}{1-\delta_1\delta_2} \leq 1, \tag{3.19}$$

$$\frac{R_2(1-\delta_1)}{\delta_1(1-\delta_2)(1-\delta_1\delta_2)} + \frac{R_1}{1-\delta_1\delta_2} + \frac{R_2}{1-\delta_2} \leq 1. \tag{3.20}$$

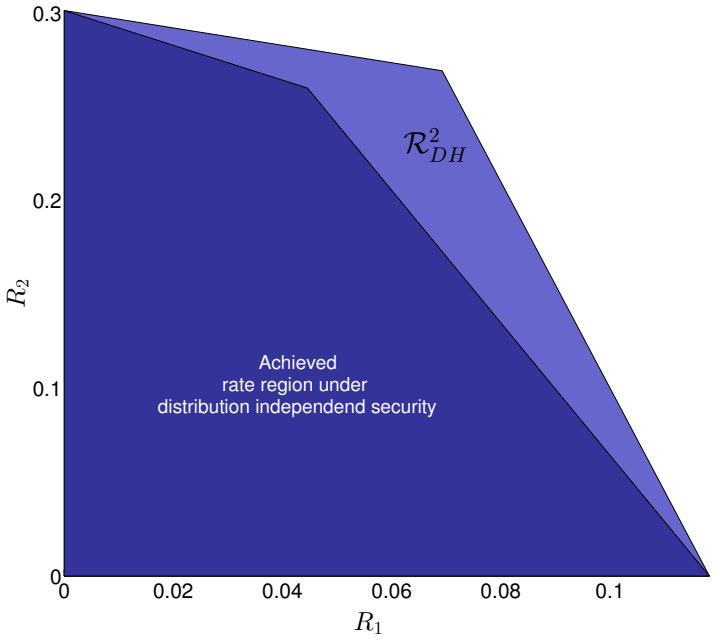


Figure 3.2: Achieved rate region under distribution independent security criterion compared to secret-message capacity region $M = 2$, $\delta_1 = 0.7$, $\delta_2 = 0.5$.

It is clear that $\mathscr{R}^2_{DH} \subseteq \mathscr{R}^2_H$, since the converse developed for the honest-but-curious case provides a valid outer bound. To prove that the region given by (3.19)-(3.20) is achievable, we construct a linear scheme that is secure against dishonest receivers and achieves any pair in the region. The scheme is described in Section 3.5.2.

Theorem 3.4 gives a complete characterization of the problem considering security against a dishonest receiver. Regarding distribution independent security we do not have such a characterization. We construct a scheme that satisfies this stronger security definition, however its optimality is not clear. The next theorem gives the rate region achieved by our scheme.

Theorem 3.5. *If a rate pair* (R_1, R_2) *satisfies*

$$\frac{R_1(1-\delta_2)}{\delta_2(1-\delta_1)(1-\delta_1\delta_2)} + \frac{R_2(1-\delta_1)}{\delta_1(1-\delta_2)(1-\delta_1\delta_2)} + \frac{R_1}{1-\delta_1} + \frac{R_2}{1-\delta_1\delta_2} \leq 1, \tag{3.21}$$

$$\frac{R_1(1-\delta_2)}{\delta_2(1-\delta_1)(1-\delta_1\delta_2)} + \frac{R_2(1-\delta_1)}{\delta_1(1-\delta_2)(1-\delta_1\delta_2)} + \frac{R_1}{1-\delta_1\delta_2} + \frac{R_2}{1-\delta_2} \leq 1, \tag{3.22}$$

then $(R_1, R_2) \in \mathscr{R}^2_{DIS}$.

From the definitions it is clear that $\mathscr{R}^2_{DIS} \subseteq \mathscr{R}^2_{DH}$. We conjecture that there is a fundamental gap between $\mathscr{R}^2_{DIS}$ and $\mathscr{R}^2_{DH}$, but we leave the proof an open question. We illustrate the gap between the rate regions of Theorems 3.4-3.5 in Figure 3.2. The scheme that constructively proves Theorem 3.5 is given in Section 3.5.3.

Corollary: security against an eavesdropper

Consider the special case when $R_2 = 0$. There is only one receiver with nonzero rate and we aim to secure his message against the other, dishonest party. In this setting the other receiver is equivalent to a passive eavesdropper who overhears the communication. Note that the sender does not trust the feedback from the second receiver, so this feedback is simply ignored, or in other words, in this particular setting there is no difference between giving potentially dishonest feedback and not giving any feedback at all. In the end, we have a broadcast channel with one receiver and an eavesdropper against whom we aim to secure a message. Indeed, Theorems 3.4-3.5 give the result of Chapter 2 as a special case when $R_2 = 0$.

Equivalence between security notions

We show the following lemma which implies that the security requirement (3.10) is equivalent to the extended notion of semantic security that we introduced in Section 3.2.2. This result confirms the intuition of Section 1.4.3 that semantic security and information-theoretic security are equivalent [24].

Lemma 3.1.

$$\mathbf{Adv}^{\mathrm{ss}}_{\mathrm{dis}} \leq \sqrt{2 \cdot \mathbf{Adv}^{\mathrm{mis}}_{\mathrm{dis}}} \tag{3.23}$$

$$\mathbf{Adv}^{\mathrm{mis}}_{\mathrm{dis}} \leq 4 \cdot \mathbf{Adv}^{\mathrm{ss}}_{\mathrm{dis}} \log\left(\frac{2^n}{\mathbf{Adv}^{\mathrm{ss}}_{\mathrm{dis}}}\right). \tag{3.24}$$

We provide the proof of the above lemma in Appendix B.5.

3.5 Coding scheme

We constructively prove the direct parts of Theorems 3.3-3.5 in this section.

3.5.1 Honest-but-curious receivers

Following the two-phase concept introduced in Chapter 2, our scheme for the broadcast channel has two main steps:

1. *Key generation.* We create M pairwise keys, each key is shared between Alice and one of the receivers, and it is perfectly secure from all the other receivers even if they collude.
2. *Encrypted broadcast.* Using the keys set up in the first phase, we employ an encrypted version of the non-secure 1-to-M broadcast scheme as described above.

Our scheme uses the property of the encryption we have seen in Chapter 2 that fewer key packets than message packets suffice. In addition, we make use of two key observations. First, there is no need to run the key generation protocol one by one with each receiver, the key generation phase can be done simultaneously with all receivers. Second, in the second step of the non-secure coding scheme, when linear combinations of packets are sent, the receivers

cannot learn any new packet from the other receiver's message, thus if packets in the first step are properly encrypted no further encryption is needed in the second step. It follows that the use of network coding does not compromise security.

Before giving a detailed description we illustrate these properties through an example.

Example

In our example Alice wants to securely send $N_1 = 1$ message packet $W_1 = [W_{1,1}]$ to Bob and $N_2 = 2$ message packets $W_2 = [W_{2,1}, W_{2,2}]$ to Calvin. The example protocol run is found in Table 3.1 below.

Alice sends	Bob's ACK	Calvin's ACK	Bob's key	Calvin's key	Bob decoded	Calvin decoded
X_1 random	✓	×	$K_{B,1} = X_1$			
X_2 random	✓	✓	$K_{B,1}$			
X_3 random	×	✓	$K_{B,1}$	$K_{C,1} = X_3$		
$X_4 = W_{1,1} \oplus K_{B,1}$	×	✓		$K_{C,1}$		
$X_5 = W_{2,1} \oplus K_{C,1}$	×	✓		$K_{C,1}$		$W_{2,1}$
$X_6 = W_{2,2} \oplus K_{C,1}$	×	×		$K_{C,1}$		$W_{2,1}$
$X_7 = X_6$	✓	×				$W_{2,1}$
$X_8 = X_4 \oplus X_7$	✓	✓			$W_{1,1}$	$W_{2,1}, W_{2,2}$

Table 3.1: An example of the protocol run.

Key generation:

a) Alice transmits random (independent and uniformly distributed) packets X_1, X_2, X_3. At the end of this phase, Alice and Bob share a secret key packet $K_{B,1} = X_1$ that Bob received and Calvin did not. Similarly, Alice and Calvin share the secret key packet $K_{C,1} = X_3$. The packet X_2 which was received by both Bob and Calvin is discarded.

Encrypted message transmissions:

b) Alice secures Bob's first message packets with a one-time-pad (using the secret key generated above) and repeatedly transmits an encrypted packet until either Bob or Calvin receive. In our example this happens immediately (X_4). The packet received only by Calvin is a side information which enables us to efficiently use the channel at a later point.

c) In the next few transmissions (X_5-X_7) we do the same with Calvin's packets. As we see, if only Calvin receives (X_5), a part of the message is successfully delivered, however the key used for encryption can be used again securely to encrypt the next message packet (X_6). If neither Bob nor Calvin receive (X_6), the packet is simply repeated (X_7).

d) Once Bob also has a side information (X_7), we send the sum of the two side information packets thereby sending information that is useful simultaneously for both receivers. This happens at transmission X_8, where both Bob and Calvin can decode a novel message packet. Note that at this step we do not need any new keys to secure the transmission.

Detailed description

We need to define a few parameters. We note that in the case of honest-but-curious receivers a variable-length coding scheme that adapts to the actual erasures could be used, because in this case we do not need to rely on the statistical behavior of the channel to ensure the security of the keys. Using such a scheme, in this setting, we could achieve secret-message capacity even without knowing the erasure probabilities of the channel. For consistency with the rest of thesis, we define a fixed-length coding scheme also here.

The length of the secret keys we aim to set up for receiver j (expressed in terms of packets) is s_j, and the length of the key generation phase in terms of transmissions is n_1. We define

$$s_j = N_j \frac{1-\frac{\prod_{k=1}^{M}\delta_k}{\delta_j}}{1-\prod_{k=1}^{M}\delta_k} + \left(N_j \frac{1-\frac{\prod_{k=1}^{M}\delta_k}{\delta_j}}{1-\prod_{k=1}^{M}\delta_k} \right)^{3/4}, \text{ and } \quad n_1 = \max_j \frac{s_j + s_j^{3/4}}{(1-\delta_j)\frac{\prod_{k=1}^{M}\delta_k}{\delta_j}}. \tag{3.25}$$

1. *Key generation*: K_j denotes the key between Alice and receiver j. Alice transmits n_1 random packets $X_1, \ldots, X_{n_1}$ generated uniformly at random over $\mathbb{F}_q^L$. K_j is the vector of the first s_j packets X_i for which $F_i = j$. If there are less than s_j such packets, we stop and declare an error for receiver j.

   That is, Alice transmits random packets, and we treat a packet received by only one receiver as a shared secret between Alice and that receiver.

2. *Encrypted broadcast*: We now follow the two transmission steps in the non-secure protocol, with the following modifications: in Step a), we encrypt the message packets using key packets as we specify in the following; in Step b), we simply reuse the already encrypted packets from Step a) to create the required linear combinations – we do not use additional key packets.

   Step 2.a) Before transmitting each message packet to receiver i, Alice encrypts it by XOR-ing it with a key packet that has either not been used for encryption in the past, or if used, none of the other users received the corresponding transmitted packet. In other words, a key is reused until an encrypted packet is received by any of the other receivers.

   Consider the transmissions to receiver j. Initially, Alice encrypts the first packet for j as $W_{j,1} \oplus K_{j,1}$ and transmits it until it is received by at least one of the receivers. If only receiver j receives this encrypted packet, she reuses the same key packet $K_{j,1}$ to encrypt the next message packet. Subsequently, if for some i and $\ell < N_j$, $k < s_j$: $X_i = W'_{j,\ell} = W_{j,\ell} \oplus K_{j,k}$, then

$$X_{i+1} = \begin{cases} X_i, & \text{if } F_i = \emptyset \\ W'_{1,\ell+1} = W_{j,\ell+1} \oplus K_{j,k}, & \text{if } F_i = j \\ W'_{1,\ell+1} = W_{j,\ell+1} \oplus K_{j,k+1}, & \text{otherwise.} \end{cases} \tag{3.26}$$

An error is declared if the s_j key packets are not sufficient to encrypt all the N_j message packets of W_j. Alice performs similar transmissions for all receivers.

Step 2.b) At the end of Step 2.a), the receivers have received as side information encrypted packets that are not intended for them; we use the same encoding as in Step b) of the non-secure protocol to deliver these packets to their intended receivers.

Analysis of the secure protocol

Condition (3.1) is clearly satisfied by our scheme. We show the other required properties for receiver j, the same arguments apply to any j.

We first argue that our scheme satisfies (3.3). From construction, we create at the end of the first phase a key K_j with

$$I\left(K_j; Y_1^{n_1}, \ldots, Y_{j-1}^{n_1}, Y_{j+1}^{n_1}, \ldots, Y_M^{n_1} F^{n_1}\right) = 0. \tag{3.27}$$

In Step 2.a), every packet $W'_{j,\ell}$ that any of the other receivers *receive* has been encrypted using a different key packet $K_{j,i}$. These key packets, from (3.27), are secret from all other receivers. Thus the packets received by the $M-1$ other receivers together are one-time pad encrypted and hence perfectly secret from them, even if they collude. In Step 2.b), Alice transmits linear combinations of packets $W'_{j,\ell}$ that have not been received by receiver j, but have already been received by at least one of the other $M-1$ receivers – thus, assuming these receivers collude, they do not receive any innovative $W'_{j,\ell}$. This concludes our argument and shows that

$$I\left(W_j; Y_1^n, \ldots, Y_{j-1}^n, Y_{j+1}^n, \ldots, Y_M^n F^n\right) = 0. \tag{3.28}$$

We next prove (3.2) showing that receivers can decode. Trivially, if no error is declared, receiver j can retrieve W_j from W'_j using his key K_j. We show that the probability of declaring an error can be made arbitrarily small. It is enough to consider the following two error events since the other error events are similar: *(i)* we do not obtain s_j key packets for receiver j during the first phase, and *(ii)* s_j key packets are not sufficient in Step 2.a).

(i) Denote by κ the number of packets in the first phase that are received only by receiver j. Then, κ is the sum of n_1 i.i.d. Bernoulli variables with parameter $p = (1-\delta_j)\frac{\prod_{k=1}^M \delta_k}{\delta_j}$. Thus,

$$\mathbb{E}\{\kappa\} = n_1 p = n_1(1-\delta_j)\frac{\prod_{k=1}^M \delta_k}{\delta_j} \geq s_j + s_j^{3/4}.$$

The probability of error event *(i)* equals

$$\Pr\{\kappa < s_j\} \leq \Pr\left\{\mathbb{E}\{\kappa\} - \kappa > s_j^{3/4}\right\} \leq \Pr\left\{|\mathbb{E}\{\kappa\} - \kappa| > s_j^{3/4}\right\} \leq \mathrm{e}^{-a_{3.29}\sqrt{s_j}}, \tag{3.29}$$

for some constant $a_{3.29} > 0$. The last inequality follows from the Chernoff-Hoeffding bound. Selecting n sufficiently large, this error probability can be made arbitrarily small.

(*ii*) This error event is similar, it occurs if the number of packets that only Bob receives is significantly less than its expected value, and the same technique applies. We omit details to avoid repetitive arguments.

With this we have shown that the scheme is secure against an honest-but-curious adversary. A straightforward calculation with the given parameters together with the capacity achieving property of non-secure 1-to-M protocol shows that our proposed schemes achieves any rate tuple within the region given by (3.18), which concludes the proof of achievability of Theorem 3.3. For completeness we provide the rate calculation in Appendix B.3.

3.5.2 Dishonest adversary

We describe a coding scheme for $M = 2$ that provides security for an honestly acknowledging receiver even if the other receiver is dishonest. To achieve this goal our starting points are the results of Chapter 2 for providing secrecy against an eavesdropper and the results of Section 3.5.1 for serving multiple receivers at the same time. Compared to the previous coding schemes our scheme against a dishonest adversary has the following distinguishing features:

- In the key generation phase the set of packets we use to compute the keys for Bob and for Calvin are not disjoint. Despite of this, we show that the produced keys are secure.
- Although the adversary can influence the run of the protocol, we ensure that independently of his acknowledging strategy, he cannot control how many times a given encrypted packet from the other receiver's message appears on the channel. From this property it follows that we can estimate accurately the number of packets the adversary overhears which makes it possible to use the encryption scheme as seen in Chapter 2.
- In the second phase we need coding to make transmissions maximally useful for both users as seen in the previous section. Alice can send an XOR-ed packet only if both receivers have a side information packet. However, a dishonest user might deny having a side information packet and hinder these coded transmissions. In our scheme, we limit the number of transmissions that each step might take, ensuring that the honest receiver does not experience a loss in rate even if no encoded transmissions take place.

The design principle of the scheme is not different from the previous schemes: we have a key generation phase and an encrypted message transmission phase. We apply coding similarly as in Chapter 2 to make security for one receiver independent from the feedback of the other receiver.

The operation of the protocol utilizes a set of parameters which we can directly calculate before the protocol starts, and whose use will be described in the following. Recall that Bob's message W_1 consists of N_1 packets, while Calvin's message W_2 consists of N_2 packets. Similarly as before $F_i, F_i^* \in \{B, C, BC, \emptyset\}$ denotes that "Bob received", "Calvin received", "both received", "none received".

$$s'_B = N_1 \frac{1-\delta_2}{1-\delta_1\delta_2} + \left(N_1 \frac{1-\delta_2}{1-\delta_1\delta_2}\right)^{\frac{3}{4}} \tag{3.30}$$

$$s'_C = N_2 \frac{1-\delta_1}{1-\delta_1\delta_2} + \left(N_2 \frac{1-\delta_1}{1-\delta_1\delta_2}\right)^{\frac{3}{4}} \tag{3.31}$$

$$s_B = \frac{s'_B}{\delta_2} + \frac{s'^{3/4}_B}{\delta_2} \tag{3.32}$$

$$s_C = \frac{s'_C}{\delta_1} + \frac{s'^{3/4}_C}{\delta_1} \tag{3.33}$$

$$n_1 = \max\left\{ \frac{s_B}{1-\delta_1} + \left(\frac{s_B}{1-\delta_1}\right)^{\frac{3}{4}}, \frac{s_C}{1-\delta_2} + \left(\frac{s_C}{1-\delta_2}\right)^{\frac{3}{4}} \right\} \tag{3.34}$$

$$n_{2,1} = \frac{N_1}{1-\delta_1\delta_2} + \left(\frac{N_1}{1-\delta_1\delta_2}\right)^{3/4} \tag{3.35}$$

$$n_{2,2} = \frac{N_2}{1-\delta_1\delta_2} + \left(\frac{N_2}{1-\delta_1\delta_2}\right)^{3/4} \tag{3.36}$$

$$n'_{2,3} = \frac{N_1}{1-\delta_1} + \left(\frac{N_1}{1-\delta_1}\right)^{3/4} - n_{2,1} \tag{3.37}$$

$$n''_{2,3} = \frac{N_2}{1-\delta_2} + \left(\frac{N_2}{1-\delta_2}\right)^{3/4} - n_{2,2} \tag{3.38}$$

$$n = n_1 + n_{2,1} + n_{2,2} + \max\left\{n'_{2,3}, n''_{2,3}\right\}. \tag{3.39}$$

Key Generation

1. Alice transmits n_1 packets $X_1, \ldots, X_{n_1}$. She generates these packets uniformly at random from $\mathbb{F}_q^L$ using her private randomness, and independently of W_1, W_2.
2. Bob and Calvin acknowledge which packets they have received. If Bob receives less than s_B packets we declare a protocol error for him. Similarly for Calvin if he receives less than s_C packets. When an error is declared for both users, the protocol terminates. If not, we continue with the user not in error, as if the user in error did not exist.
3. Let X_1^B be the row vector of the first s_B packets that Bob acknowledged. Alice and Bob create s'_B secret key packets as $K_B = X_1^B H_{K_B}$, where H_{K_B} is a $s_B \times s'_B$ matrix and is a parity check matrix of a $[s_B, s_B - s'_B]$ MDS code. Similarly, using the first s_C packets that Calvin acknowledges, Alice and Calvin create s'_C secret key packets using a matrix H_{K_C}. Matrices H_{K_B}, H_{K_C} are publicly known and fixed in advance.

Encrypted message broadcast
Encryption

4. Alice and Bob produce N_1 linear combinations of their s'_B secret key packets as $K'_B = K_B G_{K'_B}$, where $G_{K'_B}$ is a $s'_B \times N_1$ matrix and is a generator matrix of an $[N_1, s'_B]$ MDS code which is also publicly known. Similarly, Alice and Calvin create N_2 linear combinations of their s'_C key packets.
5. Alice creates N_1 encrypted messages to send to Bob

$$U_{B,i} = W_{1,i} \oplus K'_{B,i}, \quad i = 1 \ldots N_1. \tag{3.40}$$

Let U_B denote the set $\{U_{B,i} : i = 1,\ldots,N_1\}$. She similarly produces a set U_C of N_2 encrypted messages to send to Calvin.

Encrypted transmissions

6. Alice sequentially takes the first encrypted packet from $U_{B,i}$, $i = 1\ldots N_1$, that is not yet acknowledged by either Bob or Calvin and repeatedly transmits it, until it is acknowledged by either receiver. That is, if at time i Alice transmits $X_i = U_{B,j}$ for some $j < N_1$, then

$$X_{i+1} = \begin{cases} X_i, & \text{if } F_i^* = \emptyset \\ U_{B,j+1}, & \text{otherwise.} \end{cases} \tag{3.41}$$

   Alice continues these transmissions until all packets from U_B are acknowledged or $n_{2,1}$ transmissions are already made in this step. In the former case, she continues with the next step. In the latter case, if Bob does not acknowledge $\frac{N_1(1-\delta_1)}{1-\delta_1\delta_2}$ packets, then he is considered to be dishonest and Alice continues with sending only Calvin's packets using ARQ. Similarly, if Calvin does not acknowledge $\frac{N_1(1-\delta_2)}{1-\delta_1\delta_2}$ packets, then he is considered to be dishonest and Alice continues with sending only Bob's packets. In case neither receiver is considered to be dishonest, still U_B is not completely delivered, Alice stops and an error is declared for both receivers.

7. Similarly, Alice sends not-yet-acknowledged encrypted packets from $U_{C,i}$, $i = 1\ldots N_2$, until either Bob or Calvin acknowledges. If at time i Alice transmits $X_i = U_{C,j}$ for some $j < N_2$, then

$$X_{i+1} = \begin{cases} X_i, & \text{if } F_i^* = \emptyset \\ U_{C,j+1}, & \text{otherwise.} \end{cases} \tag{3.42}$$

   Alice continues these transmissions until all packets from U_C are acknowledged or $n_{2,2}$ transmissions are already made in this step. In the former case, she continues with the next step. In the latter case, if Bob does not acknowledge $\frac{N_2(1-\delta_1)}{1-\delta_1\delta_2}$ packets, then he is considered to be dishonest and Alice continues with sending only Calvin's packets using ARQ. Similarly, if Calvin does not acknowledge $\frac{N_2(1-\delta_2)}{1-\delta_1\delta_2}$ packets, then he is considered to be dishonest and Alice continues with sending only Bob's packets. In case neither receiver is considered to be dishonest, still U_C is not completely delivered, Alice stops and an error is declared for both receivers.

8. Let Q_B denote the set of packets that only Calvin acknowledged in Step 6. Similarly, Q_C denotes those packets that only Bob acknowledged in Step 7. Alice sequentially takes packets from Q_B and Q_C. For each transmission, she takes the first packet from Q_B that Bob has not acknowledged together with the first packet from Q_C that Calvin has not yet acknowledged and she transmits the XOR of the two packets. If at time i Alice transmits

$X_i = Q_{B,j} \oplus Q_{C,\ell}$ for some $j < |Q_B|, \ell < |Q_C|$, then

$$X_{i+1} = \begin{cases} X_i, & \text{if } F_i^* = \emptyset, \\ Q_{B,j+1} \oplus Q_{C,\ell}, & \text{if } F_i^* = B, \\ Q_{B,j} \oplus Q_{C,\ell+1}, & \text{if } F_i^* = C, \\ Q_{B,j+1} \oplus Q_{C,\ell+1}, & \text{if } F_i^* = BC. \end{cases} \tag{3.43}$$

Alice continues with the XOR transmissions as long as either receiver acknowledges all his packets. If Bob has already acknowledged all packets from Q_B, Alice repeats packets that are not yet acknowledged by Calvin from Q_C. Similarly, if Calvin has already acknowledged all packets from Q_C, then Alice continues with repeating the remaining packets for Bob from Q_B.

If at any point, the overall number of transmissions would exceed n as defined in (3.39) we stop and declare an error for the party (or parties) who has not acknowledged all his encrypted message packets.

Protocol analysis

We prove that the presented scheme is secure against a dishonest adversary as defined in Definition 3.4 and runs without error with high probability. We use lemmas whose proofs are provided in Appendix B. A simple calculation with the given parameters (delegated to Appendix B.3) shows that it achieves any rate pair in the the region defined by (3.19)-(3.20).

Security In our argument we focus on the secrecy of W_1 against a dishonest Calvin, but the same reasoning works for W_2 against a dishonest Bob as well.

To analyze the secrecy of W_1, we may, without loss of generality, assume that no error was declared for Bob during the key generation phase. Recall that an error is declared for Bob only if Bob fails to acknowledge at least s_B packets. If an error was in fact declared for Bob, no information about Bob's message W_1 is ever transmitted by Alice. However, note that we do account for this error event when we analyze the probability of error for Bob.

We observe that the generation of keys K_B and K_C is no different than the key generation against a passive eavesdropper. Also, K_B depends only on Bob's feedback and Calvin has no control over the protocol run in the key generation phase. Hence, we can rely on the proof of Theorem 1.1 to show that the key generation phase is secure. We have the following lemma.

Lemma 3.2. *When Bob is honest and no error is declared for Bob in the key generation phase,*

$$I\left(K_B; Y_2^{n_1} F^{n_1} \Theta_C\right) \leq s'_B \mathrm{e}^{-a_{3.44}\sqrt{s'_B}}, \tag{3.44}$$

where $a_{3.44} > 0$ is a constant. Further, K_B is uniformly distributed over its alphabet.

The proof of this lemma is the same as the proof of Theorem 1.1 with appropriate substitution

of the parameters. For this reason we omit the detailed proof.

We still need to show that the secrecy condition is satisfied by the scheme even if Calvin chooses any message distribution P_{W_1} and applies any acknowledging strategy, i.e., (3.7) holds. In the proof we omit taking the maximum, but the argument holds for any message distribution and any adversarial strategy, so the statement follows. We have

$$I\left(W_1; Y_2^n F^n \Theta_C\right) \leq I\left(W_1; Y_2^n | Y_2^{n_1} F^n \Theta_C U_C\right), \tag{3.45}$$

where the inequality used the fact that $\Theta_A, \Theta_C, W_2, F^n$ are independent of W_1 and we may express $Y_2^{n_1}, U_C$ as deterministic functions of $\Theta_A, \Theta_C, W_2, F^n$. Let M_B^C be the random variable which denotes the number of distinct packets of U_B that Calvin observes either in its pure form or in a form where the $U_{B,i}$ packet is added with some $U_{C,j}$ packet. We have the following two lemmas:

Lemma 3.3. $H\left(Y_2^n | Y_2^{n_1} F^n \Theta_C U_C H\right) \leq \mathbb{E}\left\{M_B^C\right\}$.

Lemma 3.4. $H\left(Y_2^n | W_1 Y_2^{n_1} F^n \Theta_C U_C\right) \geq \mathbb{E}\left\{\min\left(s'_B, M_B^C\right)\right\} - I\left(K_B; Y_2^{n_1} F^{n_1} \Theta_C\right)$.

The proofs of these lemmas are found in Appendix B.1-B.2. Using these in (3.45), we have

$$I\left(W_1; Y_2^n F^n \Theta_C\right) \leq \mathbb{E}\left\{\max\left(0, M_B^C - s'_B\right)\right\} + I\left(K_B; Y_2^{n_1} F^{n_1} \Theta_C\right). \tag{3.46}$$

Lemma 3.2 gives a bound for the second term. Notice that the probability that Calvin overhears a packet $U_{B,i}$ (where we count overhearing in both pure form or as part of a linear combination), is $\frac{1-\delta_2}{1-\delta_1\delta_2}$ independently of Calvin's acknowledging strategy, because Calvin has no control over how many times a given packet is transmitted (it is repeated until Bob acknowledges). Thus, M_B^C is a sum of N_1 independent random variables, and hence $\mathbb{E}\left\{M_B^C\right\} = N_1 \frac{1-\delta_2}{1-\delta_1\delta_2}$. Since $s'_B = N_1 \frac{1-\delta_2}{1-\delta_1\delta_2} + \left(N_1 \frac{1-\delta_2}{1-\delta_1\delta_2}\right)^{\frac{3}{4}}$, by applying Chernoff-Hoeffding bound we have

$$\mathbb{E}\left\{\max\left(0, M_B^C - s'_B\right)\right\} \leq N_1 \Pr\left\{M_B^C > s'_B\right\} \leq N_1 \mathrm{e}^{-a_{3.47}\sqrt{N_1}}, \tag{3.47}$$

for a constant $a_{3.47} > 0$. Substituting this together with Lemma 3.2 in (3.46) we get

$$I\left(W_1; Y_2^n F^n \Theta_C\right) \leq N_1 \mathrm{e}^{-a_{3.47}\sqrt{N_1}} + s'_B \mathrm{e}^{-a_{3.44}\sqrt{s'_B}}, \tag{3.48}$$

for constants $a_{3.44}, a_{3.47} > 0$. By choosing a large enough value of N_1 (which implies a large enough n), we satisfy (3.7).

Error probability An error happens if (a) Bob receives less than s_B packets in the first phase, or (b) he does not receive N_1 encrypted message packets in steps 6 and 8 before the protocol terminates. Both these error events have the same nature. An error happens if Bob collects significantly fewer packets than he is expected to receive in a particular step. We apply the Chernoff-Hoeffding bound as we did earlier proving that the probability of these events can be made arbitrarily small. We omit details to avoid repetition.

3.5.3 Distribution independent scheme

In the following, we describe a scheme which satisfies the stronger security notion as defined in Definition 3.7. The protocol of Section 3.5.2 cannot satisfy distribution independent security, because if Calvin knows his message *a priori*, then U_C carries information about the packets used in the key generation phase, hence potentially giving him extra information about Bob's key. We can overcome this issue if we modify the key generation phase and make sure that no packet used in generating Calvin's key contributes to Bob's key, thus U_C is conditionally independent of Bob's key given W_2 and Calvin's observation of the protocol. This results in two separate key generation phases, one for Bob and one for Calvin.

Instead of sending n_1 key generation packets as defined in (3.34), we have a key generation of length $n_1^* + n_2^*$, where

$$n_1^* = \frac{s_B}{1-\delta_1} + \left(\frac{s_B}{1-\delta_1}\right)^{\frac{3}{4}} \tag{3.49}$$

$$n_2^* = \frac{s_C}{1-\delta_2} + \left(\frac{s_C}{1-\delta_2}\right)^{\frac{3}{4}}. \tag{3.50}$$

Bob's key is then computed from the first n_1^* packets, while Calvin's key is computed from the next n_2^* packets. All other parameters remain the same as in Section 3.5.2 and the second phase remains unchanged too.

This scheme provides distribution independent security, which property is proved in Appendix B.4. A straightforward rate calculation completes the proof of Theorem 3.5.

3.6 Outer Bound

We show the converse part of Theorem 3.3, by which we conclude the proof of Theorems 3.3 and 3.4. This result proves optimality of the schemes presented in Section 3.5.1 and in Section 3.5.2. Our derivation assumes honest feedback. This provides a valid outer bound for a dishonest adversary also, since honest acknowledging is a valid adversarial strategy. We provide an intuitive interpretation of our proof in Section 3.6.2.

Proof. We present our proof for $M = 3$, the generalization of the same argument for any M is straightforward. We are going to show that for any j and any π

$$\frac{R_j\left(1-\frac{\delta_1\delta_2\delta_3}{\delta_j}\right)}{(1-\delta_j)\frac{\delta_1\delta_2\delta_3}{\delta_j}(1-\delta_1\delta_2\delta_3)}+\frac{R_{\pi_1}}{1-\delta_{\pi_1}}+\frac{R_{\pi_2}}{1-\delta_{\pi_1}\delta_{\pi_2}}+\frac{R_{\pi_3}}{1-\delta_{\pi_1}\delta_{\pi_2}\delta_{\pi_3}}\leq 1 \tag{3.51}$$

holds, which implies the statement of the theorem. Also, to avoid cumbersome notation we show (3.51) for $j = 1$ and $\pi = (1,2,3)$. With simple relabeling, the same argument holds for any j and π.

$$n \geq \sum_{i=1}^{n} H(X_i) \geq \sum_{i=1}^{n} H\left(X_i|Y_1^{i-1}F^{i-1}\right) = \sum_{i=1}^{n} H\left(X_i|Y_1^{i-1}Y_2^{i-1}F^{i-1}\right) + I\left(X_i;Y_2^{i-1}|Y_1^{i-1}F^{i-1}\right) \tag{3.52}$$

$$= \sum_{i=1}^{n} H\left(X_i|Y_1^{i-1}Y_2^{i-1}Y_3^{i-1}F^{i-1}\right) + I\left(X_i;Y_2^{i-1}|Y_1^{i-1}F^{i-1}\right) + I\left(X_i;Y_3^{i-1}|Y_1^{i-1}Y_2^{i-1}F^{i-1}\right) \tag{3.53}$$

$$= \sum_{i=1}^{n} H\left(X_i|W_1W_2W_3Y_1^{i-1}Y_2^{i-1}Y_3^{i-1}F^{i-1}\right) \tag{3.54}$$

$$+ I\left(X_i;Y_2^{i-1}|Y_1^{i-1}F^{i-1}\right) \tag{3.55}$$

$$+ I\left(X_i;Y_3^{i-1}|Y_1^{i-1}Y_2^{i-1}F^{i-1}\right) \tag{3.56}$$

$$+ I\left(X_i;W_1W_2W_3|Y_1^{i-1}Y_2^{i-1}Y_3^{i-1}F^{i-1}\right). \tag{3.57}$$

In the following Lemmas 3.5-3.8 we give bounds on each of the terms (3.54)-(3.57). Combining these results together and taking the asymptotic of both sides gives (3.51) and in turn the statement of the theorem. The proofs of Lemmas 3.6-3.8 are delegated to Section 3.6.3. □

Lemma 3.5. *From conditions (3.1)-(3.3) it follows that*

$$\sum_{i=1}^{n} H\left(X_i|Y_1^{i-1}Y_2^{i-1}Y_3^{i-1}W_1W_2W_3F^{i-1}\right) \geq \frac{nR_1(1-\delta_2\delta_3)}{(1-\delta_1)\delta_2\delta_3(1-\delta_1\delta_2\delta_3)} - n\mathscr{E}_{3.5}, \tag{3.58}$$

where $\mathscr{E}_{3.5} = \mathscr{E}_{3.10}\frac{1-\delta_2\delta_3}{(1-\delta_1)\delta_2\delta_3}$, *and* $\mathscr{E}_{3.10}$ *is a vanishing error constant specified in the proof.*

Lemma 3.6. *From conditions (3.1)-(3.3) it follows that*

$$\sum_{i=1}^{n} I\left(X_i; Y_2^{i-1} | Y_1^{i-1} F^{i-1}\right) \geq \frac{nR_1}{1-\delta_1} - \frac{nR_1}{1-\delta_1\delta_2} - n\mathscr{E}_{3.6}, \tag{3.59}$$

where $\mathscr{E}_{3.6} = \frac{h_2(\epsilon)+\epsilon R_1}{1-\delta_1}$.

Lemma 3.7. *From conditions (3.1)-(3.2) it follows*

$$\sum_{i=1}^{n} I\left(X_i; Y_3^{i-1} | Y_1^{i-1} Y_2^{i-1} F^{i-1}\right) \geq \frac{n(R_1+R_2)}{1-\delta_1\delta_2} - \frac{n(R_1+R_2)}{1-\delta_1\delta_2\delta_3} - n\mathscr{E}_{3.7}, \tag{3.60}$$

where $\mathscr{E}_{3.7} = \frac{h_2(2\epsilon)+2\epsilon(R_1+R_2)}{1-\delta_1\delta_2}$.

Lemma 3.8. *From conditions (3.1)-(3.2) it follows that*

$$\frac{n(R_1+R_2+R_3)}{1-\delta_1\delta_2\delta_3} - n\mathscr{E}_{3.8} \leq \sum_{i=1}^{n} I\left(X_i; W_1W_2W_3 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} F^{i-1}\right) \leq \frac{n(R_1+R_2+R_3)}{1-\delta_1\delta_2\delta_3}, \tag{3.61}$$

where $\mathscr{E}_{3.8} = \frac{h_2(3\epsilon)+3\epsilon(R_1+R_2+R_3)}{1-\delta_1\delta_2\delta_3}$.

3.6.1 Proof of Theorem 3.2

Proof. It is sufficient to prove the inequality for $\pi = (1,2,3)$. By relabeling, the same argument holds for any π. We repeat the first steps of the previous proof and bound term (3.54) by 0:

$$\begin{aligned} n \geq \sum_{i=1}^{n} & I\left(X_i; Y_2^{i-1} | Y_1^{i-1} F^{i-1}\right) && (3.62)\\ & + I\left(X_i; Y_3^{i-1} | Y_1^{i-1} Y_2^{i-1} F^{i-1}\right) && (3.63)\\ & + I\left(X_i; W_1W_2W_3 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} F^{i-1}\right). && (3.64) \end{aligned}$$

Lemmas 3.6-3.8 give bounds on terms (3.62)-(3.64) respectively. Combining these gives the stated inequality. □

3.6.2 Interpretation of the converse proof

To facilitate understanding, beside our formal proof through Lemmas 3.5-3.8 here we provide some intuitive interpretation of terms (3.54)-(3.57) and of the inequalities we derive. Similarly as in the point-to-point setting, we can match terms to steps of our scheme, but we stress that the proof holds for any possible scheme.

In Lemma 3.5 we see the following (here we omit small terms for simplicity):

$$(1-\delta_1)\delta_2\delta_3 \sum_{i=1}^{n} H\left(X_i | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} W_1 W_2 W_3 F^{i-1}\right) \geq \frac{nR_1(1-\delta_2\delta_3)}{1-\delta_1\delta_2\delta_3}. \tag{3.65}$$

The entropy term on the LHS of this inequality accounts for fresh randomness sent by the source. In our scheme we call this the key generation phase. The constant factor $(1-\delta_1)\delta_2\delta_3$ suggests that a random packet becomes a key for receiver 1 if if he is the only one that receives the transmission. The RHS of the inequality corresponds to the expected number of (encrypted) W_1 packets that not only receiver 1 gets, but some other receivers also overhear. These are the packets that need to be secured, thus to be able to secure them, receiver 1 needs at least the same amount of secret key packets. This lower bound on term (3.54) suggests that any scheme has to introduce some source randomness. We find it natural to call it key generation.

Terms (3.55)-(3.57) correspond to the second phase of our protocol. Term (3.57) corresponds to the first step of the message transmission phase (see Step (a)), when the sender ensures that the receivers *together* could decode all the messages. Terms (3.55)-(3.56) account for the encoded transmissions. E.g. (3.55) intuitively corresponds to "a packet that is of interest for receiver 1 known to receiver 2". Indeed, Lemma 3.6 lower bounds this term with the expected number of transmissions that are needed to convey to receiver 1 the side information overheard by receiver 2.

3.6.3 Proofs of Lemmas 3.5-3.8

For technical reasons the order of the proofs does not follow the order of appearance of the Lemmas. We use that the private randomness of the receivers does not help them decoding, as we have seen in the proof of Lemma 2.2.

Proof of Lemma 3.8

$$n(R_1+R_2+R_3)-n\mathscr{E}_{3.8}(1-\delta_1\delta_2\delta_3) \le I\left(Y_1^nY_2^nY_3^nF^n;W_1W_2W_3\right) \tag{3.66}$$

$$=\sum_{i=1}^{n} I\left(Y_{1,i}Y_{2,i}Y_{3,i}F_i;W_1W_2W_3|Y_1^{i-1}Y_2^{i-1}Y_3^{i-1}F^{i-1}\right) \tag{3.67}$$

$$=\sum_{i=1}^{n} I\left(Y_{1,i}Y_{2,i}Y_{3,i};W_1W_2W_3|Y_1^{i-1}Y_2^{i-1}Y_3^{i-1}F^{i-1}F_i\right) \tag{3.68}$$

$$=\sum_{i=1}^{n} \Pr\{F_i\neq\emptyset\}\, I\left(Y_{1,i}Y_{2,i}Y_{3,i};W_1W_2W_3|Y_1^{i-1}Y_2^{i-1}Y_3^{i-1}F^{i-1},F_i\neq\emptyset\right) \tag{3.69}$$

$$=\sum_{i=1}^{n} I\left(X_i;W_1W_2W_3|Y_1^{i-1}Y_2^{i-1}Y_3^{i-1}F^{i-1}\right)(1-\delta_1\delta_2\delta_3). \tag{3.70}$$

Here, the first inequality is Fano's inequality, and we exploited the independence property of F_i. This completes the proof of the first inequality of the lemma. Further, we also see that

$$I\left(Y_1^nY_2^nY_3^nF^n;W_1W_2W_3\right) \le n(R_1+R_2+R_3). \tag{3.71}$$

From (3.66)-(3.70)

$$I\left(Y_1^n Y_2^n Y_3^n F^n; W_1 W_2 W_3\right) = \sum_{i=1}^{n} I\left(X_i; W_1 W_2 W_3 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} F^{i-1}\right)(1-\delta_1\delta_2\delta_3), \tag{3.72}$$

which gives the second inequality of the lemma. □

Proof of Lemma 3.6

From the same type of derivation as we apply in Lemma 3.8, we have that

$$\sum_{i=1}^{n} I\left(X_i; W_1 | Y_1^{i-1} F^{i-1}\right) \geq \frac{nR_1}{1-\delta_1} - n\mathscr{E}_{3.6}, \tag{3.73}$$

$$\sum_{i=1}^{n} I\left(X_i; W_1 | Y_1^{i-1} Y_2^{i-1} F^{i-1}\right) \leq \frac{nR_1}{1-\delta_1\delta_2}. \tag{3.74}$$

Thus,

$$\begin{aligned}
\frac{nR_1}{1-\delta_1} - n\mathscr{E}_{3.6} &\leq \sum_{i=1}^{n} I\left(X_i; W_1 | Y_1^{i-1} F^{i-1}\right) && (3.75)\\
&= \sum_{i=1}^{n} I\left(X_i; W_1 | Y_1^{i-1} Y_2^{i-1} F^{i-1}\right) + I\left(X_i; Y_2^{i-1} | Y_1^{i-1} F^{i-1}\right) - I\left(X_i; Y_2^{i-1} | Y_1^{i-1} F^{i-1} W_1\right)\\
&\leq \sum_{i=1}^{n} I\left(X_i; W_1 | Y_1^{i-1} Y_2^{i-1} F^{i-1}\right) + I\left(X_i; Y_2^{i-1} | Y_1^{i-1} F^{i-1}\right) && (3.76)\\
&\leq \frac{nR_1}{1-\delta_1\delta_2} + \sum_{i=1}^{n} I\left(X_i; Y_2^{i-1} | Y_1^{i-1} F^{i-1}\right) && (3.77)
\end{aligned}$$

□

Proof of Lemma 3.7

The proof follows the same steps as the proof of Lemma 3.6. We omit details. □

Proof of Lemma 3.5

To show the statement, we prove the next two helper lemmas. Combining the results of Lemma 3.9 and Lemma 3.10 completes the proof of Lemma 3.5.

Lemma 3.9. *From the definition of the channel it follows that*

$$\sum_{i=1}^{n} H\left(X_i | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} W_1 W_2 W_3 F^{i-1}\right) \geq \frac{1-\delta_2\delta_3}{(1-\delta_1)\delta_2\delta_3} \sum_{i=1}^{n} I\left(X_i; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} W_1 W_2 W_3 F^{i-1}\right). \tag{3.78}$$

Proof. The proof follows the same steps as we have seen when deriving inequality (2.45). Here

we omit some intermediate steps. Recall that W is a shorthand for (W_1, W_2, W_3).

$$0 \leq H\left(Y_1^n | Y_2^n Y_3^n F^n W\right) = H\left(Y_1^{n-1} | Y_2^n Y_3^n F^n W\right) + H\left(Y_{1,n} | Y_1^{n-1} Y_2^n Y_3^n F^n W\right) \tag{3.79}$$

$$= H\left(Y_1^{n-1} | Y_2^{n-1} Y_3^{n-1} F^{n-1} W\right) - I\left(Y_1^{n-1}; Y_{2,n} Y_{3,n} F_n | Y_2^{n-1} Y_3^{n-1} F^{n-1} W\right) + H\left(Y_{1,n} | Y_1^{n-1} Y_2^n Y_3^n F^n W\right) \tag{3.80}$$

$$= H\left(Y_1^{n-1} | Y_2^{n-1} Y_3^{n-1} F^{n-1} W\right) - I\left(Y_1^{n-1}; X_n | Y_2^{n-1} Y_3^{n-1} F^{n-1} W\right)(1-\delta_2\delta_3) + H\left(X_n | Y_1^{n-1} Y_2^{n-1} Y_3^{n-1} F^{n-1} W\right)(1-\delta_1)\delta_2\delta_3. \tag{3.81}$$

We do the same steps recursively to obtain the statement of the lemma. □

Lemma 3.10. *From conditions (3.1)-(3.3) it follows that*

$$\sum_{i=1}^{n} I\left(X_i; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} F^{i-1} W_1 W_2 W_3\right) \geq \frac{nR_1}{1-\delta_1\delta_2\delta_3} - n\mathscr{E}_{3.10}, \tag{3.82}$$

where $\mathscr{E}_{3.10} = 2\mathscr{E}_{3.10a} + \mathscr{E}_{3.10b} + \mathscr{E}_{3.10c} + \mathscr{E}_{3.10d}$. $\mathscr{E}_{3.10a}, \mathscr{E}_{3.10b}, \mathscr{E}_{3.10c}, \mathscr{E}_{3.10d}$ *are vanishing error terms, defined throughout the proof.*

Proof.

$$\sum_{i=1}^{n} I\left(X_i; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} F^{i-1} W_1 W_2 W_3\right) = \sum_{i=1}^{n} I\left(X_i; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} F^{i-1} W_1\right) \tag{3.83}$$

$$- I\left(X_i; W_2 W_3 | Y_2^{i-1} Y_3^{i-1} F^{i-1} W_1\right) + I\left(X_i; W_2 W_3 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} F^{i-1} W_1\right). \tag{3.84}$$

We bound the terms in (3.83)-(3.84) one by one. First, consider the two terms in (3.84). From the decodability condition and Fano's inequality we have

$$I\left(Y_2^n Y_3^n F^n; W_2 W_3 | W_1\right) \leq I\left(Y_2^n Y_3^n F^n; W_2 W_3\right) + n(h_2(\epsilon) + \epsilon(R_2 + R_3)). \tag{3.85}$$

Following the same kind of derivation as in the proof of Lemma (3.8), we can write

$$\sum_{i=1}^{n} I\left(X_i; W_2 W_3 | Y_2^{i-1} Y_3^{i-1} F^{i-1} W_1\right) \leq \sum_{i=1}^{n} I\left(X_i; W_2 W_3 | Y_2^{i-1} Y_3^{i-1} F^{i-1}\right) + n\mathscr{E}_{3.10a}, \tag{3.86}$$

where $\mathscr{E}_{3.10a} = \frac{h_2(\epsilon)+\epsilon(R_2+R_3)}{1-\delta_2\delta_3}$.

For the other term, we use the independence property of the messages:

$$I\left(Y_1^n Y_2^n Y_3^n F^n; W_2 W_3 | W_1\right) = I\left(Y_1^n Y_2^n Y_3^n F^n; W_2 W_3\right) - I\left(W_1; W_2 W_3\right) + I\left(W_1; W_2 W_3 | Y_1^n Y_2^n Y_3^n F^n\right) \geq I\left(Y_1^n Y_2^n Y_3^n F^n; W_2 W_3\right), \tag{3.87}$$

and thus

$$\sum_{i=1}^{n} I\left(X_i; W_2 W_3 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} F^{i-1} W_1\right) \geq \sum_{i=1}^{n} I\left(X_i; W_2 W_3 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} F^{i-1}\right). \tag{3.88}$$

Now both terms have the form which is similar to those seen in Lemma 3.8. This enables us to bound these terms using the same ideas. Doing so gives

$$-\sum_{i=1}^{n} I\left(X_i; W_2 W_3 | Y_2^{i-1} Y_3^{i-1} F^{i-1} W_1\right) \geq -\frac{n(R_2+R_3)}{1-\delta_2\delta_3} - n\mathscr{E}_{3.10a} \tag{3.89}$$

$$\sum_{i=1}^{n} I\left(X_i; W_2 W_3 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} F^{i-1} W_1\right) \geq \frac{n(R_2+R_3)}{1-\delta_1\delta_2\delta_3} - n\mathscr{E}_{3.10a}. \tag{3.90}$$

It remains to give a bound on the term in (3.83). From the security condition[1] and after a few basic steps (as seen in the derivation of (2.54)) we can arrive to

$$\begin{aligned} n\mathscr{E}_{3.10b} > \sum_{i=1}^{n} I\left(X_i; W_1 | Y_2^{i-1} Y_3^{i-1} F^{i-1}\right) = \sum_{i=1}^{n} -I\left(X_i; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} F^{i-1} W_1\right), \\ + I\left(X_i; W_1 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} F^{i-1}\right) + I\left(X_1; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} F^{i-1}\right), \end{aligned} \tag{3.91}$$

where $\mathscr{E}_{3.10b} = \frac{\epsilon}{1-\delta_2\delta_3}$. From a similar result as we have seen in Lemma 3.8:

$$I\left(X_i; W_1 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} F^{i-1}\right) \geq \frac{nR_1}{1-\delta_1\delta_2\delta_3} - n\mathscr{E}_{3.10c}, \tag{3.92}$$

where $\mathscr{E}_{3.10c} = \frac{h_2(\epsilon)+\epsilon R_1}{1-\delta_1\delta_2\delta_3}$. Further, a symmetric result to Lemma 3.7 shows:

$$I\left(X_1; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} F^{i-1}\right) \geq \frac{n(R_2+R_3)}{1-\delta_2\delta_3} - \frac{n(R_2+R_3)}{1-\delta_1\delta_2\delta_3} - n\mathscr{E}_{3.10d}, \tag{3.93}$$

where $\mathscr{E}_{3.10d} = \frac{h_2(2\epsilon)+2\epsilon(R_2+R_3)}{1-\delta_2\delta_3}$. Applying these bounds in (3.91) results

$$\sum_{i=1}^{n} -I\left(X_i; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} F^{i-1} W_1\right) \geq \frac{n(R_2+R_3)}{1-\delta_2\delta_3} - \frac{n(R_1+R_2+R_3)}{1-\delta_1\delta_2\delta_3} - n(\mathscr{E}_{3.10b} + \mathscr{E}_{3.10c} + \mathscr{E}_{3.10d}). \tag{3.94}$$

Substituting back (3.89)-(3.90) and (3.94) to (3.83) results the claim of the lemma. □

3.7 Next steps

In this chapter we have seen that the techniques developed for a point-to-point channel are applicable (with appropriate modifications) for a broadcast channel. Our two-phase protocol design achieves capacity in this setting also. Indeed, we have seen that the point-to-point channel can be seen as a special case of the broadcast channel.

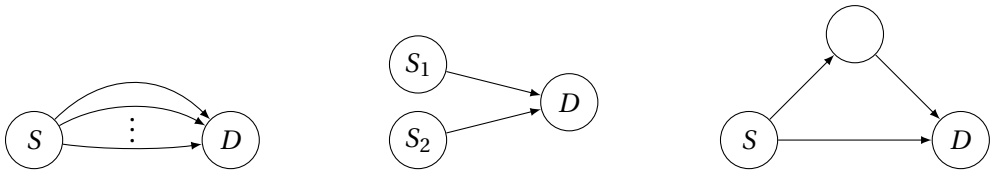
In the following chapters we turn our attention toward networks, where communication takes place over multiple channels.

[1] Recall that in outer bound proofs we use the weak form of security.

4 Secret-message capacity in networks

The current chapter and the next chapter investigate the secret-message sending problem in networks. We restrict our attention to networks where all channels operate independently. In particular, we do not consider broadcasting, instead, our networks are built from the same kind of point-to-point channels that Chapter 2 considers, i.e., all channels are potentially eavesdropped point-to-point erasure channels with state-feedback. This model is well suited for wired networks and also matches today's common wireless practice which aims to operate all wireless link independently using different frequency bands for each connection.

In this chapter we provide exact characterization for a number of simple network topologies. First we look at a network that consists of multiple parallel channels between the source and the destination node. We then consider the V-network, which is an intermediate step toward multihop settings. The V-network has two sources that have a common destination and that have access to a limited rate common randomness. The last topology we consider is the triangle network, where the source is connected to the destination via two paths: one goes through an intermediate node whereas the other is a direct channel. These topologies are depicted in Figure 4.1.



(a) Parallel channels (b) V-network (c) Triangle network

Figure 4.1: Our networks. Causal channel state-feedback are sent over a separate noiseless public channel (not shown).

In all cases we derive the secret-message capacity against an eavesdropper who wiretaps any one channel of her choice. We present optimal coding schemes for each setting. We build on ideas from the previous chapters, but beyond that, each setting requires several new ideas. In particular, to achieve security we can exploit not only the channel erasures, but also

the structure of the network. Intermediate network nodes have the possibility to inject new randomness to the network, while common randomness between nodes becomes a precious resource. These new features make the behavior of even these simple networks nontrivial.

We extensively use the linear programming approach that we have introduced in Section 2.6 for the point-to-point channel. We give both our coding schemes and our outer bounds in form of linear programs, which is a new approach to address network secrecy problems.

4.1 Related Work

Secure network coding considers secure communication over a network that consists of unit capacity, error-free channels [55, 56]. A strong connection between the min-cut, the number of eavesdropped edges and the secrecy capacity of the network was shown. In particular, if the min-cut is h and Eve eavesdrops on z channels, the secrecy capacity is $h-z$. In that setting, the secret-message capacity is the same whether or not the set of wiretapped edges is known. In our networks this is not the case. Note that in the case of error-free channels, state-feedback is superfluous. Hence, the setting of [55, 56] can be seen as a special case of our network when all erasure probabilities are zero. Indeed, our proposed schemes specialize to the secure network coding scheme when the erasure probabilities get close to 0.

Wyner's wiretap channel [25] was generalized for networks [57, 58] and more specifically for wireless erasure networks [59], however none of these works consider feedback.

4.2 Parallel channels

4.2.1 Model

As shown in Figure 4.1a, a source, Alice is connected to the destination, Bob through independent parallel channels. The number of channels is $\ell \geq 1$. The operation of the channels are as we described in Chapter 1. According to Definition 1.1 in this setting (1.6) becomes

$$X_{k,i} = f_{k,i}\left(W, \Theta_A, F^{i-1}\right), \quad i = 1,2,\ldots,n; \quad k = 1,\ldots,\ell. \tag{4.1}$$

Bob simultaneously receives through all channels, hence (1.7) in Definition 1.1 has the form

$$\Pr\left\{\phi\left(Y_1^n,\ldots,Y_\ell^n,F^n,\Theta_B\right) \neq W\right\} < \epsilon. \tag{4.2}$$

The eavesdropper, Eve, might select any one channel to wiretap without Alice or Bob being aware of her choice. This is equivalent to having an eavesdropper on every channel, but these eavesdroppers do not collude. The security criterion (1.9) in Definition 1.4 is thus

$$I\left(W; Z_j^n F^n \Theta_E\right) < \epsilon, \quad \forall\, j \in \{1,\ldots,\ell\}. \tag{4.3}$$

4.2.2 Main result

The following theorem gives the secret-message capacity for the communication setting with ℓ parallel channels.

Theorem 4.1. *The secret-message capacity of the network with ℓ parallel channels is the optimal value of the following linear program, where all parameters $m_i, c_i \geq 0$:*

$$\max R, \quad \textit{such that:}$$

$$R \leq \sum_{i=1}^{\ell} (1-\delta_i)\, m_i \tag{4.4}$$

$$\forall\, i \in \{1,\ldots,\ell\}:$$

$$m_i\,(1-\delta_i)\,\frac{1-\delta_{iE}}{1-\delta_i\delta_{iE}} \leq c_i\delta_{iE}\,(1-\delta_i) + \sum_{j=1, j\neq i}^{\ell} c_j\left(1-\delta_j\right) \tag{4.5}$$

$$1 \geq m_i + c_i. \tag{4.6}$$

We show that through its parameter values, the above linear program describes a coding scheme that achieves secret-message rate R in all cases when the linear program is feasible.

Discussion

The linear program in Theorem 4.1 follows the structure of the LP (2.61)-(2.63). Constraint (4.4) is a rate constraint, constraints (4.5) are security constraints for each channel and constraints (4.6) are time-sharing constraints.

For the special case of $\ell = 1$, we get the same LP as we have seen in Section 2.6 for the point-to-point channel. Consider $\ell = 2$. Then, the linear program is the following.

$$\max R, \quad \text{such that:}$$

$$R \leq (1-\delta_1)\, m_1 + (1-\delta_2)\, m_2 \tag{4.7}$$

$$m_1\,(1-\delta_1)\,\frac{1-\delta_{1E}}{1-\delta_1\delta_{1E}} \leq c_1\delta_{1E}\,(1-\delta_1) + c_2\,(1-\delta_2) \tag{4.8}$$

$$m_2\,(1-\delta_2)\,\frac{1-\delta_{2E}}{1-\delta_2\delta_{2E}} \leq c_2\delta_{2E}\,(1-\delta_2) + c_1\,(1-\delta_1) \tag{4.9}$$

$$1 \geq m_1 + c_1 \tag{4.10}$$

$$1 \geq m_2 + c_2. \tag{4.11}$$

It should be noted that – as opposed to the LP seen in Section 2.6 – the solution of this linear program is not trivial any more. Given Theorem 2.1, it is clear that if we knew that the eavesdropper eavesdrops on the first channel, the secret-message capacity would be

$$(1-\delta_2) + \delta_{1E}\,(1-\delta_1)\,\frac{1-\delta_1\delta_{1E}}{1-\delta_1\delta_{1E}^2}, \tag{4.12}$$

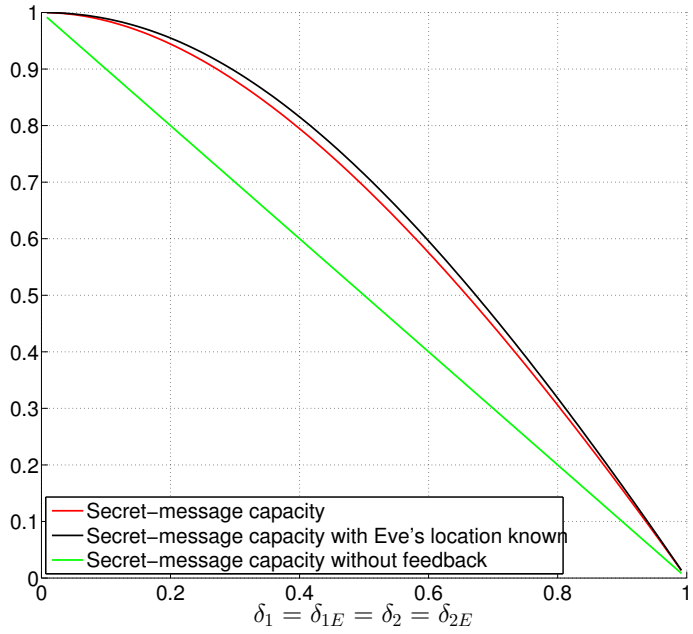


Figure 4.2: Parallel channels: example secret-message capacity with/without knowing Eve's location, with/without feedback

whereas if we knew that she selects the second channel, it would be

$$(1-\delta_1)+\delta_{2E}\,(1-\delta_2)\,\frac{1-\delta_2\delta_{2E}}{1-\delta_2\delta_{2E}^2}. \tag{4.13}$$

One might expect that if her selection is not known we can possibly achieve

$$\min\left\{(1-\delta_2)+\delta_{1E}\,(1-\delta_1)\,\frac{1-\delta_1\delta_{1E}}{1-\delta_1\delta_{1E}^2},\ (1-\delta_1)+\delta_{2E}\,(1-\delta_2)\,\frac{1-\delta_2\delta_{2E}}{1-\delta_2\delta_{2E}^2}\right\}. \tag{4.14}$$

The above formula gives a trivial upper bound, however – as we show here – it is not achievable in general. In some cases the secret-message capacity is strictly smaller than (4.14). To illustrate the gap, we plot one such case in Figure 4.2. In the same figure one can also observe the role of state-feedback. In our example case, if feedback is not available, one cannot do better than using secure network coding after applying an error correction code on each channel.

4.2.3 Coding scheme

Principle

In our scheme we use the ideas that we developed in Chapter 2 for the point-to-point channel. In particular, we have two phases on each channel, a key generation phase and a message

sending phase. This time, the sender sets up two different secret keys, one against each eavesdropper (or in other words, one for each). Then, the message is split into ℓ parts, one part to be sent on each channel. In the second phase channels operate independently. The sender encrypts the message part assigned to each channel using the secret key of the given channel and transmits the encrypted packets using ARQ in the same way as we have seen for the point-to-point channel.

Compared to the point-to-point channel, the novelty of the scheme lies in the key generation phase. We exploit the fact that a key generation packet sent over channel j and correctly received by the destination can serve as a secret key against every eavesdropper, except the one who wiretaps channel j. Clearly, such a packet is a shared randomness between the sender and the destination, and the eavesdroppers on the other channels cannot have any information about it. On the RHS of (4.5) new terms $c_j\left(1-\delta_j\right)$ appear, which corresponds to this observation.

Detailed description

We assume that a feasible set of parameters for the linear program in Theorem 4.1 is given. For a given n, we define the following parameters:

$$\forall i \in \{1,\dots,\ell\}:$$

$$nc_i = \frac{s_i}{1-\delta_i} + \frac{s_i^{\frac{3}{4}}}{1-\delta_i} \tag{4.15}$$

$$s_i = \frac{s_i'}{\delta_{iE}} + \frac{s_i'^{\frac{3}{4}}}{\delta_{iE}}. \tag{4.16}$$

Let N_i be the number of message packets that can be reliably and securely transmitted on channel i using nm_i transmissions and a key of size $s_i' + \sum_{j=1, j\neq i}^{\ell} s_j$. By Theorem 2.2 and from (4.5) such N_i exists such that $\lim_{n\to\infty} \frac{N_i}{n} = m_i\left(1-\delta_i\right)$. For a detailed calculation of N_i see Appendix C.1. Then $N = \sum_{i=1}^{\ell} N_i$, and an (n, ϵ, N) coding scheme is as follows:

Key generation: On channel j Alice sends nc_j packets selected independently and uniformly at random. If the destination does not receive s_i packets on some channel i an error is declared. For channel j the secret key S_j consists of two parts. The first part is the s_j' secure key packets generated from the first s_j packets received through channel j using the key generation scheme in Section 1.5. The second part is the first $\sum_{i=1, i\neq j}^{\ell} s_i$ random packets that the destination receives on channels other than j.

Encrypted message sending: For each channel j, N_j message packets are assigned and Alice performs the second phase of the scheme for a point-to-point channel (as described in Section 2.4), i.e., she encrypts the N_j message packets using the secret key S_j and forwards them with ARQ. If the destination does not receive all encrypted message packets, an error is declared.

Analysis

The security of the above scheme directly follows form Theorem 2.2 and (4.5). The rate of the secret key S_j that we set up for each channel equals $c_i\delta_{iE}(1-\delta_i)+\sum_{j=1,j\neq i}^{\ell} c_j\left(1-\delta_j\right)$, which by Theorem 2.2 and (4.5) is sufficient to secure a message of rate $m_i(1-\delta_i)$. The nm_i transmissions in the second phase support rate $m_i(1-\delta_i)$, thus overall we can send a message of rate $\sum_{i=1}^{\ell} m_i(1-\delta_i)$ such that (4.2) and (4.3) are satisfied for a large enough n.

On each channel two phases run and we have already shown that each of these succeeds with high probability. The error probability of the scheme is upper bounded by the sum of the error probabilities on each channel. The analysis in Section 2.4.3 has already shown that for every channel the probability decays to 0, thus by the same argument the overall error probability is also arbitrarily small for a large enough n. This proves the direct part of Theorem 4.1.

4.2.4 Outer bound

We provide the outer bound also in the form of a linear program. We derive several general information inequalities and transform them into linear constraints by treating entropy and mutual information terms as arbitrary non-negative parameters. As we are going to see, there exists a one-to-one mapping between the parameters we define this way and the parameters of the coding scheme linear program. That is, the outer bound linear program has the same form as the linear program that describes the scheme, which shows immediately that the outer bound matches the achieved rate.

As a first step, we show that assuming independence of simultaneous transmissions on different channels does not reduce the achievable secret-message rate. The theorem below formalizes the following: if a scheme does not satisfy independence of simultaneous transmissions, we can construct another scheme that achieves the same rate and satisfies the assumption as follows. We take ℓ independent copies of the scheme (using independent messages and new independent randomness). In every ℓ time slots we proceed one transmission of each copy such that on all the ℓ edges a different copy of the scheme runs in each time slot. Clearly, the rate does not change and also packets sent in the same time slot are independent.

Let $\mathcal{G}$ be the class of functions that give back a subset of their inputs, such that the selection of the subset does not depend on the input, i.e.

$$\mathcal{G}=\left\{g|g(X)=X^{I_g}, \text{for some } I_g\in 2^{\{1,\ldots,|X|\}}\right\}, \tag{4.17}$$

where X^{I_g} denotes the vector X restricted to indices I_g.

Theorem 4.2. *If there exists a scheme $\mathcal{P}$ that achieves secret-message rate R in our setting, there also exists a scheme that achieves the same rate and for which*

$$I\left(X_{k,i};X_{j,i}|g\left(Y_1^{i-1}Z^{i-1}F^{i-1}W\right)\right)=0 \tag{4.18}$$

for every i, for any $k,j,k\neq j$ and for any function $g\in\mathcal{G}$.

Corollary 4.1. *If for every i, for any $k, j, k \neq j$ and for any function $g \in \mathcal{G}$* (4.18) *holds, then*

$$I\left(Y_{k,i}; Y_{j,i} | g\left(Y^{i-1} Z^{i-1} F^{i-1} W\right)\right) = 0 \tag{4.19}$$

also holds.

Proof. To help readability, we show the theorem for $\ell = 2$, the generalization for any ℓ is trivial. We construct a new scheme that achieves the same rate as follows. Let $X^{\mathscr{P}}$ and $Y^{\mathscr{P}}$ denote the channel inputs and outputs of the given scheme $\mathscr{P}$. We take two independent copies of the scheme $\mathscr{P}$. Let $\mathscr{P}'$ and $\mathscr{P}''$ denote the two copies. The channel inputs defined by $\mathscr{P}'$ and $\mathscr{P}''$ are

$$X_{k,i}^{\mathscr{P}'} = f_{k,i}\left(\Theta'_A, Y^{\mathscr{P}' i-1}, W'\right), \tag{4.20}$$

$$X_{k,i}^{\mathscr{P}''} = f_{k,i}\left(\Theta''_A, Y^{\mathscr{P}'' i-1}, W''\right) \tag{4.21}$$

for some function $f_{k,i}$ defined by the scheme $\mathscr{P}$. Here $\left(\Theta'_A, W'\right)$ and $\left(\Theta''_A, W''\right)$ are independent random variables. For the new scheme

$$\Theta_A = \left(\Theta'_A, \Theta''_A\right), \tag{4.22}$$

$$W = \left(W', W''\right). \tag{4.23}$$

The new scheme uses $2n$ transmissions defined as follows:

$$X_{1,i} = \begin{cases} X_{1,\frac{i+1}{2}}^{\mathscr{P}'}, & \text{if } i \text{ is odd} \\ X_{1,\frac{i}{2}}^{\mathscr{P}''}, & \text{if } i \text{ is even} \end{cases} \tag{4.24}$$

$$X_{2,i} = \begin{cases} X_{2,\frac{i+1}{2}}^{\mathscr{P}''}, & \text{if } i \text{ is odd} \\ X_{2,\frac{i}{2}}^{\mathscr{P}'}, & \text{if } i \text{ is even} \end{cases}. \tag{4.25}$$

The new scheme achieves the same rate as $\mathscr{P}$, since it runs two copies of it in $2n$ time slots. Let F'^n and F''^n denote the channel states relevant for each copies. Also, for i odd and some $g', g'' \in \mathcal{G}$:

$$I\left(X_{1i}; X_{2i} | g\left(Y^{i-1} Z^{i-1} F^n W\right)\right) = H\left(X_{1i} | g'\left(Y^{\mathscr{P}' i-1} Z^{\mathscr{P}' i-1} F'^n W'\right), g''\left(Y^{\mathscr{P}'' i-1} Z^{\mathscr{P}'' i-1} F''^n W''\right)\right) \tag{4.26}$$

$$- H\left(X_{1i} | X_{2i}, g'\left(Y^{\mathscr{P}' i-1} Z^{\mathscr{P}' i-1} F'^n W'\right), g''\left(Y^{\mathscr{P}'' i-1} Z^{\mathscr{P}'' i-1} F''^n W''\right)\right) \tag{4.27}$$

$$= H\left(X_{1i} | g'\left(Y^{\mathscr{P}' i-1} Z^{\mathscr{P}' i-1} F'^n W'\right)\right) - H\left(X_{1i} | g'\left(Y^{\mathscr{P}' i-1} Z^{\mathscr{P}' i-1} F'^n W'\right)\right) = 0, \tag{4.28}$$

where the last step follows, from the independence of variables used by $\mathscr{P}'$ and $\mathscr{P}''$. By

symmetry, we get the same result for i even. Corollary 4.1 is a direct consequence of (4.18) and the fact that channel erasures are independent. □

In the rest of the proof we will assume (4.18)-(4.19). Similarly to our previous outer bound proofs we assume that F_i contains also Eve's channel state information and use the fact that Θ_B does not help decoding. We can now start deriving the inequalities that will serve as constraints in our linear program.

Rate constraint

Lemma 4.1. *For any achievable message rate R:*

$$nR - n\mathscr{E}_{4.1} \leq \sum_{j=1}^{\ell} (1-\delta_j) \sum_{i=1}^{n} I\left(X_{j,i}; W | Y^{i-1} F^{i-1}\right), \tag{4.29}$$

where $\mathscr{E}_{4.1} = h_2(\epsilon) + R\epsilon$.

Proof. We use Fano's inequality and the independence of channel erasures:

$$\begin{aligned} nR - n\mathscr{E}_{4.1} &\leq I\left(Y^n F^n; W\right) = \sum_{i=1}^{n} I\left(Y_i; W | Y^{i-1} F^{i-1}\right) = \sum_{i=1}^{n} \sum_{j=1}^{\ell} I\left(Y_{j,i}; W | Y^{i-1} Y_{1,i} \dots Y_{j-1,i} F^{i-1}\right) \\ &\overset{(a)}{=} \sum_{i=1}^{n} \sum_{j=1}^{\ell} I\left(Y_{j,i}; W | Y^{i-1} F^{i-1}\right) = \sum_{j=1}^{\ell} \sum_{i=1}^{n} (1-\delta_j) I\left(X_{j,i}; W | Y^{i-1} F^{i-1}\right). \end{aligned} \tag{4.30}$$

In step (a) we used Corollary 4.1. □

Security constraint

Lemma 4.2. *For any $k \in \{1, \dots, \ell\}$:*

$$\begin{aligned} &\frac{(1-\delta_k)(1-\delta_{kE})}{1-\delta_k \delta_{kE}} \sum_{i=1}^{n} I\left(X_{k,i}; W | Y^{i-1} F^{i-1}\right) - n\epsilon \leq \\ &\qquad \sum_{i=1}^{n} \delta_{kE} (1-\delta_k) H\left(X_{k,i} | Y^{i-1} F^{i-1} W\right) + \sum_{j=1, j\neq k}^{\ell} (1-\delta_j) H\left(X_{j,i} | Y^{i-1} F^{i-1} W\right). \end{aligned} \tag{4.31}$$

Proof. We are going to show that for any k

$$\begin{aligned} &\sum_{i=1}^{n} (1-\delta_k \delta_{kE}) I\left(X_{k,i}; W | Y^{i-1} Z_k^{i-1} F^{i-1}\right) + \sum_{j=1, j\neq k}^{\ell} (1-\delta_j) I\left(X_{j,i}; W | Y^{i-1} Z_k^{i-1} F^{i-1}\right) \\ &\qquad \geq \sum_{i=1}^{n} \sum_{j=1}^{\ell} (1-\delta_j) I\left(X_{j,i}; W | Y^{i-1} F^{i-1}\right), \end{aligned} \tag{4.32}$$

and

$$\sum_{i=1}^{n} \delta_{kE}(1-\delta_k) H\left(X_{k,i}|Y^{i-1}F^{i-1}W\right) + \sum_{j=1,j\neq k}^{\ell} (1-\delta_j) H\left(X_{j,i}|Y^{i-1}Z_k^{i-1}F^{i-1}W\right) \geq \sum_{i=1}^{n} (1-\delta_{kE}) I\left(X_{k,i};W|Y^{i-1}Z_k^{i-1}F^{i-1}\right) - n\epsilon. \quad (4.33)$$

We eliminate the common term $\sum_{i=1}^{n} I\left(X_{k,i};W|Y^{i-1}Z_k^{i-1}F^{i-1}\right)$ by combining the above two inequalities:

$$\begin{aligned}
&\frac{(1-\delta_k)(1-\delta_{kE})}{1-\delta_k\delta_{kE}} \sum_{i=1}^{n} I\left(X_{k,i};W|Y^{i-1}F^{i-1}\right) - n\epsilon \leq \sum_{i=1}^{n} \delta_{kE}(1-\delta_k) H\left(X_{k,i}|Y^{i-1}F^{i-1}W\right) \\
&\quad + \sum_{j=1,j\neq k}^{\ell} (1-\delta_j)\Big(H\left(X_{j,i}|Y^{i-1}Z_k^{i-1}F^{i-1}W\right) \\
&\quad + \frac{(1-\delta_k)(1-\delta_{kE})}{1-\delta_k\delta_{kE}}\left(I\left(X_{j,i};W|Y^{i-1}Z_k^{i-1}F^{i-1}\right) - I\left(X_{j,i};W|Y^{i-1}F^{i-1}\right)\right)\Big) \quad (4.34)\\
&\overset{(a)}{\leq} \sum_{i=1}^{n} \delta_{kE}(1-\delta_k) H\left(X_{k,i}|Y^{i-1}F^{i-1}W\right) \\
&\quad + \sum_{j=1,j\neq k}^{\ell} (1-\delta_j)\left(H\left(X_{j,i}|Y^{i-1}Z_k^{i-1}F^{i-1}W\right) + \frac{(1-\delta_k)(1-\delta_{kE})}{1-\delta_k\delta_{kE}} I\left(X_{j,i};Z_k^{i-1}|Y^{i-1}F^{i-1}W\right)\right) \quad (4.35)\\
&\leq \sum_{i=1}^{n} \delta_{kE}(1-\delta_k) H\left(X_{k,i}|Y^{i-1}F^{i-1}W\right) \\
&\quad + \sum_{j=1,j\neq k}^{\ell} (1-\delta_j)\Big(H\left(X_{j,i}|Y^{i-1}F^{i-1}W\right) - I\left(X_{j,i};Z_k^{i-1}|Y^{i-1}F^{i-1}W\right) \\
&\quad + \frac{(1-\delta_k)(1-\delta_{kE})}{1-\delta_k\delta_{kE}} I\left(X_{j,i};Z_k^{i-1}|Y^{i-1}F^{i-1}W\right)\Big) \quad (4.36)\\
&\overset{(b)}{\leq} \sum_{i=1}^{n} \delta_{kE}(1-\delta_k) H\left(X_{k,i}|Y^{i-1}F^{i-1}W\right) + \sum_{j=1,j\neq k}^{\ell} (1-\delta_j) H\left(X_{j,i}|Y^{i-1}F^{i-1}W\right), \quad (4.37)
\end{aligned}$$

where in (a) we have used that

$$\begin{aligned}
&I\left(X_{j,i};W|Y^{i-1}Z_k^{i-1}F^{i-1}\right) - I\left(X_{j,i};W|Y^{i-1}F^{i-1}\right) \quad (4.38)\\
&= I\left(X_{j,i};Z_k^{i-1}|Y^{i-1}F^{i-1}W\right) - I\left(X_{j,i};Z_k^{i-1}|Y^{i-1}F^{i-1}\right) \leq I\left(X_{j,i};Z_k^{i-1}|Y^{i-1}F^{i-1}W\right), \quad (4.39)
\end{aligned}$$

and in (b) we have used that $\frac{(1-\delta_k)(1-\delta_{kE})}{1-\delta_k\delta_{kE}} \leq 1$.

It remains to show (4.32) and (4.33). First we show (4.32).

From (4.30),

$$\sum_{j=1}^{\ell}(1-\delta_j)\sum_{i=1}^{n} I\left(X_{j,i};W|Y^{i-1}F^{i-1}\right) = I\left(Y^nF^n;W\right) \le I\left(Y^nZ_k^nF^n;W\right) \tag{4.40}$$

$$\overset{(a)}{=} \sum_{i=1}^{n}(1-\delta_k\delta_{kE})\, I\left(X_{k,i};W|Y^{i-1}Z_k^{i-1}F^{i-1}\right) + \sum_{j=1,j\neq k}^{\ell}(1-\delta_j)\, I\left(X_{j,i};W|Y^{i-1}Z_k^{i-1}F^{i-1}\right). \tag{4.41}$$

The derivation of (a) follows the same steps as (4.30).

As for (4.33), we use the independence property of the channels and Theorem 4.2 to get (we omit some intermediate steps):

$$\begin{aligned}
0 &\le H\left(Y^n|Z^nF^nW\right)\\
&= \sum_{i=1}^{n} -(1-\delta_{kE})\, I\left(X_{k,i};Y^{i-1}|Z_k^{i-1}F^{i-1}W\right) + \delta_{kE}(1-\delta_k)\, H\left(X_{k,i}|Y^{i-1}Z_k^{i-1}F^{i-1}W\right)\\
&\quad + \sum_{j=1,j\neq k}^{\ell}(1-\delta_j)\, H\left(X_{j,i}|Y^{i-1}Z_k^{i-1}F^{i-1}W\right) \qquad (4.42)\\
&\le \sum_{i=1}^{n} -(1-\delta_{kE})\, I\left(X_{k,i};Y^{i-1}|Z_k^{i-1}F^{i-1}W\right) + \delta_{kE}(1-\delta_k)\, H\left(X_{k,i}|Y^{i-1}F^{i-1}W\right)\\
&\quad + \sum_{j=1,j\neq k}^{\ell}(1-\delta_j)\, H\left(X_{j,i}|Y^{i-1}Z_k^{i-1}F^{i-1}W\right). \qquad (4.43)
\end{aligned}$$

Also, from the security criterion, we have that

$$n\epsilon > I\left(Z_k^nF^n;W\right) = \sum_{i=1}^{n}(1-\delta_{kE})\, I\left(X_{k,i};W|Z_k^{i-1}F^{i-1}\right), \tag{4.44}$$

thus,

$$\begin{aligned}
(1-\delta_{kE})\sum_{i=1}^{n} I\left(X_{k,i};Y^{i-1}|Z_k^{i-1}F^{i-1}W\right) &\ge \sum_{i=1}^{n} -I\,(1-\delta_{kE})\left(X_{k,i};W|Z_k^{i-1}F^{i-1}\right)\\
&\quad + (1-\delta_{kE})\, I\left(X_{k,i};W|Y^{i-1}Z_k^{i-1}F^{i-1}\right) \qquad (4.45)\\
&\ge -n\epsilon + \sum_{i=1}^{n}(1-\delta_{kE})\, I\left(X_{k,i};W|Y^{i-1}Z_k^{i-1}F^{i-1}\right). \qquad (4.46)
\end{aligned}$$

Combining (4.43) and (4.46) results (4.33). □

Time-sharing constraint

Lemma 4.3. *For any $j \in \{1,\dots,\ell\}$:*

$$n \ge \sum_{i=1}^{n} H\left(X_{j,i}|Y^{i-1}F^{i-1}W\right) + I\left(X_{j,i};W|Y^{i-1}F^{i-1}\right). \tag{4.47}$$

Proof.

$$n \geq \sum_{i=1}^{n} H\left(X_{j,i}\right) \geq \sum_{i=1}^{n} H\left(X_{j,i}|Y^{i-1}F^{i-1}\right) = \sum_{i=1}^{n} H\left(X_{j,i}|Y^{i-1}F^{i-1}W\right) + I\left(X_{j,i};W|Y^{i-1}F^{i-1}\right). \tag{4.48}$$

□

Outer bound linear program

We use Lemmas 4.1-4.3. We divide the inequalities in the lemmas by n and take the limit with $n \to \infty$, $\epsilon \to 0$ (by this we eliminate the small error terms). We introduce the following relabeling by which we define the variables of our outer bound LP:

$$\hat{m}_i \sim \sum_{i=1}^{n} \frac{1}{n} I\left(X_{j,i};W|Y^{i-1}F^{i-1}\right) \tag{4.49}$$

$$\hat{c}_i \sim \sum_{i=1}^{n} \frac{1}{n} H\left(X_{j,i}|Y^{i-1}F^{i-1}W\right). \tag{4.50}$$

We rewrite the inequalities in Lemmas 4.1-4.3 and get:

$$R \leq \sum_{i=1}^{\ell} (1-\delta_i)\,\hat{m}_i$$

$$\forall\, i \in \{1,\ldots,\ell\}:$$

$$\hat{m}_i\,(1-\delta_i)\,\frac{1-\delta_{iE}}{1-\delta_i\delta_{iE}} \leq \hat{c}_i\delta_{iE}\,(1-\delta_i) + \sum_{j=1,j\neq i}^{\ell} \hat{c}_j\left(1-\delta_j\right)$$

$$1 \geq \hat{m}_i + \hat{c}_i.$$

We treat variables $\hat{m}_i, \hat{c}_i$ as arbitrary non-negative variables and take the maximum value in R, which results an outer bound. By this we arrive to the same linear program as the scheme linear program, proving optimality of our scheme. □

4.3 V-network

In this section we investigate the V-network depicted in Figure 4.1b, where two sources S_1 and S_2 are connected to a common destination through independent erasure channels. The two sources share a rate limited common random source and they both have access to the same message W to be cooperatively sent to D.

This model is motivated by the observation that in a network, nodes cannot share arbitrary amount of randomness. The rate of available common randomness limits the amount of keys that can be generated on one of the channels and used for encryption on the other channel. In one extreme, without any common randomness we get the sum capacity of the two point-to-point channels. In the other extreme, with infinite rate common randomness,

we get the secret-message capacity of two parallel channels. In this section we explore the intermediate cases.

4.3.1 Model

We denote the common randomness Ψ. For simplicity we assume that the common randomness is available in the form of i.i.d. uniform random packets. The rate of common randomness is C_r, that is $H(\Psi) = nC_r$. We assume that S_1 and S_2 share nC_r uniformly random packets. The sources can generate private randomness Θ_1 and Θ_2 at infinite rate. The channel inputs are defined by

$$X_{1,i} = f_{1,i}\left(W, \Psi, \Theta_1, F^{i-1}\right), \tag{4.51}$$

$$X_{2,i} = f_{2,i}\left(W, \Psi, \Theta_2, F^{i-1}\right). \tag{4.52}$$

As before, we assume that Eve observes any one of the two channels. We think of such an eavesdropper as two noncolluding eavesdroppers E_1 and E_2 wiretapping channel 1 and channel 2 respectively. The decodability and security conditions are the same as for the parallel channels' case, (4.2)-(4.3) applies with $\ell = 2$.

4.3.2 Main result

We characterize the secret-message capacity of the V-network.

Theorem 4.3. *The secret-message capacity of the V-network is the optimal value of the following linear program, where all parameters $m_i, c_i, c, k_i, r_i \geq 0$:*

$$\max R, \quad \textit{such that:}$$

$$R \leq (1-\delta_1)\, m_1 + (1-\delta_2)\, m_2 \tag{4.53}$$

$$m_1 (1-\delta_1) \frac{1-\delta_{1E}}{1-\delta_1\delta_{1E}} \leq (c_1 + k_1)\, \delta_{1E}(1-\delta_1) + r_1 \frac{\delta_{1E}(1-\delta_1)}{1-\delta_1\delta_{1E}} + r_2 + c_2(1-\delta_2) \tag{4.54}$$

$$m_2 (1-\delta_2) \frac{1-\delta_{2E}}{1-\delta_2\delta_{2E}} \leq (c_2 + k_2)\, \delta_{2E}(1-\delta_2) + r_2 \frac{\delta_{2E}(1-\delta_2)}{1-\delta_2\delta_{2E}} + r_1 + c_1(1-\delta_1) \tag{4.55}$$

$$1 \geq m_1 + k_1 + c_1 + \frac{r_1}{1-\delta_1} \tag{4.56}$$

$$1 \geq m_2 + k_2 + c_2 + \frac{r_2}{1-\delta_2} \tag{4.57}$$

$$C_r \geq c + r_1 + r_2 \tag{4.58}$$

$$c \geq (1-\delta_1\delta_{1E})\, c_1 + (1-\delta_2)\, c_2 \tag{4.59}$$

$$c \geq (1-\delta_2\delta_{2E})\, c_2 + (1-\delta_1)\, c_1. \tag{4.60}$$

The interpretation of the new variables will be clear from the description of our scheme.

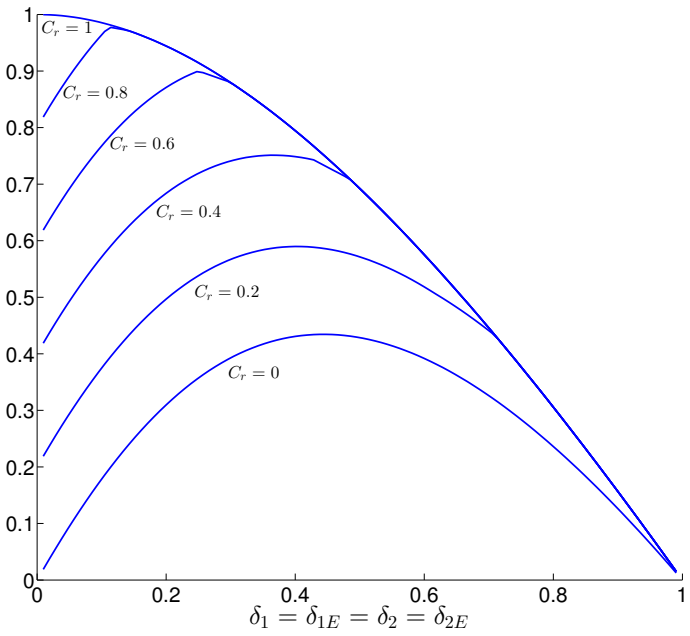


Figure 4.3: Secret-message capacity of the V-network as the function of the rate of common randomness

Discussion

In addition to a rate constraint (4.53), security constraints (4.54)-(4.55) and times sharing constraints (4.56)-(4.57) a new set of constraints (4.58)-(4.60) appears. These new constraints describe the use of the common randomness, hence we refer to them as *common randomness constraints*.

In Figure 4.3 we plot the secret-message capacity of the V-network for some example parameter values with several common randomness rates.

As one can observe, the solution of the LP in Theorem 4.3 does not coincide with the rate that a simple time-sharing between the two extreme cases (rate 0 and rate 1 common randomness) would give. Note that beyond a threshold value, the increase of common randomness rate cannot increase the secret-message capacity. For comparison, we plot the secret-message capacity against the time-sharing rate for some example parameter values on Figure 4.4.

4.3.3 Coding scheme

We prove the direct part of Theorem 4.3. Given any feasible point of the LP in Theorem 4.3 our scheme achieves the value of the LP.

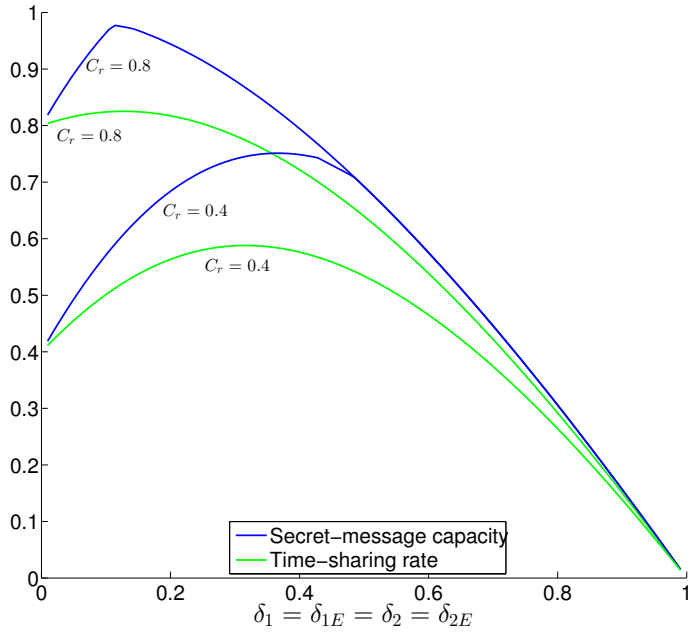


Figure 4.4: Secret-message capacity of the V-network compared to the rate achieved with time-sharing

Principle

We build on our scheme for the parallel channels. The novelty of our scheme for the V-network is the efficient use of common randomness for key generation. We need to distinguish key generation from the private randomness of each source and key generation from the common random source.

In order to use common randomness the most efficiently we introduce two new techniques. We have seen in the previous section that a key generation packet that D receives contributes to the key that is used on the other channel. In the V-network this holds only for packets generated from the common randomness.

Our first observation is that transmissions from the common randomness need not to be independent. Consider the example on Table 4.1. For simplicity we assume that Eve's channel

	Packet	D	E_1	E_2	Key for S_1	Key for S_2
S_1 sends	X_1	×	✓		×	×
S_2 sends	X_2	×		✓	×	×
S_2 sends	$X_1 \oplus X_2$	✓		×	✓	✓

Table 4.1: Coding across packets from the common randomness

state is known. Packets X_1 and X_2 are common random packets sent by S_1 and S_2 respectively.

The first two transmissions are not successful for D, but X_1 gets received by E_1 and X_2 gets received by E_2. Although both X_1 and X_2 are known by one of the eavesdroppers, they can still contribute to both keys. In our example, as a third transmission, S_2 sends $X_1 \oplus X_2$. E_2 does not receive the transmission, hence $X_1 \oplus X_2$ becomes a key against both eavesdroppers. Observe that the important property of a key generation packet is that it is innovative for D and E_1 (taken together) and also for D and E_2 (taken together). In our scheme we use coding to generalize this idea for the case when Eve's channel state is not known. This key generation strategy achieves the same key rates as if a new random packet was sent in each transmission, but it uses the available randomness more efficiently.

Our second observation is that ARQ can be used also as a key generation technique. On one hand, retransmission of a random packet increases the chance of the eavesdropper to overhear it, which results a lower key rate on the given channel. On the other hand, ARQ uses less common randomness than sending always an innovative random packet and a packet *always* provides a key for the other channel. Thus, in some cases ARQ is a reasonable strategy for key generation, unlike when the common randomness is unlimited.

We split the common random packets and apply a time-sharing between the above described new key generation methods.

Detailed description

In our description we focus on the key generation step. Given a secret key for both channels, the encryption and the message sending phase is the same as we have seen in previous section. Building on Theorem 2.2 and the results of the previous section it becomes clear that the availability of a key of sufficient rate (as required by (4.54)-(4.55)) ensures that the claimed secret-message rate (4.53) is achievable. To ease readability we omit detailed parameter definitions for the second phase.

Below we describe the transmissions each source does. Every step results a key of a certain rate, the keys of S_1 and S_2 are the concatenation of these keys. Keys are generated as linear combinations of packets received by D, similarly as we have seen in Section 1.5. Here we only claim the key rate achieved by each step, we delegate parameter definitions and proofs to the analysis of the scheme.

Initialization:

The common randomness Ψ (or potentially a part of it) is split into three disjoint sets of random packets:

$$H(\Psi) \geq n(c + r_1 + r_2). \tag{4.61}$$

We assign nr_1 packets to S_1 and nr_2 packets to S_2. The third part, nc packets will be used commonly.

Key generation:

Step 1) S_1 and S_2 send nk_1 and nk_2 uniform random packets generated from their private randomness. Key rates $k_1\delta_{1E}(1-\delta_1)$ and $k_1\delta_{2E}(1-\delta_2)$ are achieved for S_1 and S_2.

Step 2) S_1 and S_2 send the nr_1 and nr_2 packets from the common randomness using ARQ. Key rates $r_2+\frac{r_1\delta_{1E}(1-\delta_1)}{1-\delta_1\delta_{1E}}$ and $r_1+\frac{r_2\delta_{2E}(1-\delta_2)}{1-\delta_2\delta_{2E}}$ are achieved.

Step 3) The nc packets from the common randomness are arranged into vector C. Out of these packets nc_1+nc_2 linear combination packets are produced to be sent by S_1 and S_2 respectively. The linear combinations are produced as follows:

$$CG=\begin{bmatrix}C_1 & C_2\end{bmatrix}, \tag{4.62}$$

where G is a $nc\times n(c_1+c_2)$ matrix and is a generator of an MDS code. C_1, C_2 are vectors of nc_1 and nc_2 packets. These packets are sent (once each) by S_1 and S_2 respectively. This step creates a rate $\delta_{1E}(1-\delta_1)c_1+c_2(1-\delta_2)$ key for S_1 and a rate $\delta_{2E}(1-\delta_2)c_2+c_1(1-\delta_1)$ for S_2.

Message sending:

Using their keys from the key generation phase, S_1 and S_2 transmit a rate $m_1(1-\delta_1)$ and a rate $m_2(1-\delta_2)$ part of the message. Channels operate independently, they both apply the second phase of the coding scheme for a point-to-point channel. We omit details to avoid repetition.

Analysis

The previously proved properties of the message sending phase together with constraints (4.54)-(4.57) shows that secret-message rate R is achievable, if the claimed key rates are achieved. In the following lemmas we show the key rate of each step. We highlight the ideas here and delegate details to Appendix C.2.

Lemma 4.4. *Step 1 of the key generation phase achieves secret-key rates $k_1\delta_{1E}(1-\delta_1)$ and $k_2\delta_{2E}(1-\delta_2)$ for S_1 and S_2.*

Proof. This step of the key generation is the same as the key generation scheme in Section 1.5. The lemma is a rephrasing of Theorem 1.1. □

Lemma 4.5. *Step 2 of the key generation phase achieves secret-key rates $r_2+\frac{r_1\delta_{1E}(1-\delta_1)}{1-\delta_1\delta_{1E}}$ and $r_1+\frac{r_2\delta_{2E}(1-\delta_2)}{1-\delta_2\delta_{2E}}$ for S_1 and S_2.*

Proof. For the keys of S_1, the security of packets sent on channel 2 is obvious, this part of the key has rate r_2. To create keys from the packets sent using ARQ, we use the same key generation method as we have seen in Section 1.5, with the only difference that instead of $1-\delta_{1E}$, this time the probability that Eve receives a packet is increased to $\frac{1-\delta_{1E}}{1-\delta_1\delta_{1E}}$. Security and small error probability follows from the properties shown in Theorem 1.1. For details we refer the reader to Appendix C.2.1. □

Lemma 4.6. *Step 3 of the key generation phase achieves secret-key rates* $c_2(1-\delta_2)+c_1\delta_{1E}(1-\delta_1)$ *and* $c_1(1-\delta_1)+c_2\delta_{2E}(1-\delta_2)$ *for* S_1 *and* S_2.

Proof. Our proof shows that the linear combination packets C_1 and C_2 can be used as if they were packets generated independently. Constraints (4.59)-(4.60) ensure the property that the packets we send are innovative for the eavesdropper and D taken together. Similarly as before, the proof builds on the MDS property of the generator matrix and on concentration results. For the detailed proof see Appendix C.2.2. □

4.3.4 Outer bound

To derive our outer bound we use the technique that we have seen in Section 4.2.4. However, in this case the linear program that we get after transforming the various information inequalities into linear constraints does not have the same form as the LP in Theorem 4.3. Hence, we apply a series of transformation in order to show that the outer bound linear program has the same optimal value as the linear program that describes our scheme. Throughout the transformation we make sure that the optimal value of the outer bound program does not decrease, however the feasibility region might shrink (the outer bound program has more variables initially). Our possible reduction steps are:

1. Renaming of variables. By any renaming we apply, we make sure that the introduced new variables are non-negative.

2. Eliminating variables. We apply the well known Fourier-Motzkin elimination [60] to reduce the number of variables.

3. Introducing constraints. We introduce constraints that do not follow from the inequalities in the linear program. By this we reduce the feasibility region but do not reduce the optimal value of the program. Our arguments show that if in an optimal point the given constraint is not satisfied, we can apply a transformation on the variables without violating any existing constraints to arrive to another optimal point, where the constraint is satisfied. We conclude that introducing the constraint in question does not lower the value of the program.

4. Dropping constraints. Obviously, by dropping some constraints we cannot decrease the value of the program.

5. Deriving constraints. We add constraints that follow from existing constraints.

Since the transformation is a lengthy process we delegate it to Appendix C.3. Here we only derive the set of constraints from which we form the outer bound linear program. In our derivations we use that parallel transmissions on different channels can be assumed to be independent (see Theorem 4.2, which trivially generalizes for the V-network).

Rate constraints

We use that $H(W) \le nR$ and the same kind of derivation as in the proof of Lemma 4.1 to get the following inequalities. We omit details.

$$nR - \mathscr{E}_{4.1} \le I\left(W; Y^n F^n\right) = \sum_{i=1}^{n} (1-\delta_1)\, I\left(W; X_{1i} | Y^{i-1} F^{n-1}\right) + (1-\delta_2)\, I\left(W; X_{2i} | Y^{i-1} F^{n-1}\right) \le nR \tag{4.63}$$

$$\begin{aligned} nR - \mathscr{E}_{4.1} &\le I\left(W; Y^n Z_1^n F^n\right) \\ &= \sum_{i=1}^{n} (1-\delta_1\delta_{1E})\, I\left(W; X_{1i} | Y^{i-1} Z_1^{i-1} F^{n-1}\right) + (1-\delta_2)\, I\left(W; X_{2i} | Y^{i-1} Z_1^{i-1} F^{n-1}\right) \le nR \end{aligned} \tag{4.64}$$

$$\begin{aligned} nR - \mathscr{E}_{4.1} &\le I\left(W; Y^n Z_2^n F^n\right) \\ &= \sum_{i=1}^{n} (1-\delta_1)\, I\left(W; X_{1i} | Y^{i-1} Z_2^{i-1} F^{n-1}\right) + (1-\delta_2\delta_{2E})\, I\left(W; X_{2i} | Y^{i-1} Z_2^{i-1} F^{n-1}\right) \le nR \end{aligned} \tag{4.65}$$

$$\begin{aligned} nR - \mathscr{E}_{4.1} &\le I\left(W; Y^n Z^n F^n\right) \\ &= \sum_{i=1}^{n} (1-\delta_1\delta_{1E})\, I\left(W; X_{1i} | Y^{i-1} Z^{i-1} F^{n-1}\right) + (1-\delta_2\delta_{2E})\, I\left(W; X_{2i} | Y^{i-1} Z^{i-1} F^{n-1}\right) \\ &\le nR \end{aligned} \tag{4.66}$$

Common randomness constraints

$$\begin{aligned} nCr = H(\Psi) &\ge I\left(\Psi; Y^n Z^n F^n W\right) \\ &= \sum_{i=1}^{n} (1-\delta_1\delta_{1E})\, I\left(X_{1i}; \Psi | Y^{i-1} Z^{i-1} F^{i-1} W\right) + (1-\delta_2\delta_{2E})\, I\left(X_{2i}; \Psi | Y^{i-1} Z^{i-1} F^{i-1} W\right) \end{aligned} \tag{4.67}$$

$$\begin{aligned} 0 \le I\left(\Psi; Z_1^n | Y^n Z_2^n F^{i-1} W\right) &= \\ = \sum_{i=1}^{n} &-(1-\delta_2\delta_{2E})\, I\left(X_{2i}; Z_1^{i-1} | Y^{i-1} Z_2^{i-1} F^{i-1} W\right) - (1-\delta_1)\, I\left(X_{1i}; Z_1^{i-1} | Y^{i-1} Z_2^{i-1} F^{i-1} W\right) \\ &+ \delta_1 (1-\delta_{1E})\, I\left(X_{1i}; \Psi | Y^{i-1} Z^{i-1} F^{i-1} W\right) + (1-\delta_1)\, I\left(X_{1i}; Z_1^{i-1} | Y^{i-1} Z_2^{i-1} W \Psi\right) \end{aligned} \tag{4.68}$$

$$\begin{aligned} 0 \le I\left(\Psi; Z_2^n | Y^n Z_1^n F^{i-1} W\right) &= \\ = \sum_{i=1}^{n} &-(1-\delta_1\delta_{1E})\, I\left(X_{1i}; Z_2^{i-1} | Y^{i-1} Z_1^{i-1} F^{i-1} W\right) - (1-\delta_2)\, I\left(X_{2i}; Z_2^{i-1} | Y^{i-1} Z_1^{i-1} F^{i-1} W\right) \\ &+ \delta_2 (1-\delta_{2E})\, I\left(X_{2i}; \Psi | Y^{i-1} Z^{i-1} F^{i-1} W\right) + (1-\delta_2)\, I\left(X_{2i}; Z_2^{i-1} | Y^{i-1} Z_1^{i-1} W \Psi\right) \end{aligned} \tag{4.69}$$

$$
\begin{aligned}
0 &\le I\left(Y_2^n;\Psi|Y_1^nZ_1^nF^{i-1}W\right)\\
&=\sum_{i=1}^{n}-(1-\delta_1\delta_{1E})\,I\left(X_{1i};Y_2^{i-1}|Y_1^{i-1}Z_1^{i-1}F^{i-1}W\right)+(1-\delta_2)\,I\left(X_{2i};\Psi|Y^{i-1}Z^{i-1}F^{i-1}W\right)\\
&\quad+(1-\delta_2)\,I\left(X_{2i};Z_2^{i-1}|Y^{i-1}Z_1^{i-1}F^{i-1}W\right)-(1-\delta_2)\,I\left(X_{2i};Z_2^{i-1}|Y^{i-1}Z_1^{i-1}W\Psi\right) \qquad (4.70)\\
0 &\le I\left(Y_1^n;\Psi|Y_2^nZ_2^nF^{i-1}W\right)\\
&=\sum_{i=1}^{n}-(1-\delta_2\delta_{2E})\,I\left(X_{2i};Y_1^{i-1}|Y_2^{i-1}Z_2^{i-1}F^{i-1}W\right)+(1-\delta_1)\,I\left(X_{1i};\Psi|Y^{i-1}Z^{i-1}F^{i-1}W\right)\\
&\quad+(1-\delta_1)\,I\left(X_{1i};Z_1^{i-1}|Y^{i-1}Z_2^{i-1}F^{i-1}W\right)-(1-\delta_1)\,I\left(X_{1i};Z_1^{i-1}|Y^{i-1}Z_2^{i-1}W\Psi\right) \qquad (4.71)
\end{aligned}
$$

Distinguishing keys

We interpret the following constraints as distinguishing keys generated by S_1 and S_2. The derivation is similar to that of (2.45).

$$
\begin{aligned}
0 &\le H\left(Y_1^n|Z_1^nF^nW\right)=\sum_{i=1}^{n}-(1-\delta_{1E})\,I\left(X_{1i};Y_1^{i-1}|Z_1^{i-1}F^{i-1}W\right)\\
&\quad+\delta_{1E}(1-\delta_1)\,H\left(X_{1i}|Y^{i-1}Z_1^{i-1}F^{i-1}W\right)+\delta_{1E}(1-\delta_1)\,I\left(X_{1i};Y_2^{i-1}|Y_1^{i-1}Z_1^{i-1}F^{i-1}W\right) \qquad (4.72)\\
&\overset{(a)}{=}\sum_{i=1}^{n}-(1-\delta_{1E})\,I\left(X_{1i};Y^{i-1}|Z_1^{i-1}F^{i-1}W\right)\\
&\quad+\delta_{1E}(1-\delta_1)\,H\left(X_{1i}|Y^{i-1}Z_1^{i-1}F^{i-1}W\right)+(1-\delta_1\delta_{1E})\,I\left(X_{1i};Y_2^{i-1}|Y_1^{i-1}Z_1^{i-1}F^{i-1}W\right) \qquad (4.73)\\
&\overset{(b)}{=}-n\epsilon+\sum_{i=1}^{n}-(1-\delta_{1E})\,I\left(X_{1i};W|Z_1^{i-1}F^{i-1}W\right)\\
&\quad+\delta_{1E}(1-\delta_1)\,H\left(X_{1i}|Y^{i-1}Z_1^{i-1}F^{i-1}W\right)+(1-\delta_1\delta_{1E})\,I\left(X_{1i};Y_2^{i-1}|Y_1^{i-1}Z_1^{i-1}F^{i-1}W\right) \qquad (4.74)
\end{aligned}
$$

In (a) we used that

$$
\sum_{i=1}^{n}I\left(X_{1i};Y^{i-1}|Z_1^{i-1}F^{i-1}W\right)=\sum_{i=1}^{n}I\left(X_{1i};Y_1^{i-1}|Z_1^{i-1}F^{i-1}W\right)+I\left(X_{1i};Y_2^{i-1}|Y_1^{i-1}Z_1^{i-1}F^{i-1}W\right), \qquad (4.75)
$$

while in (b) we used (4.46). From symmetry,

$$
\begin{aligned}
0 &\le -n\epsilon+\sum_{i=1}^{n}-(1-\delta_{2E})\,I\left(X_{2i};W|Z_2^{i-1}F^{i-1}W\right)\\
&\quad+\delta_{2E}(1-\delta_2)\,H\left(X_{2i}|Y^{i-1}Z_2^{i-1}F^{i-1}W\right)+(1-\delta_2\delta_{2E})\,I\left(X_{2i};Y_1^{i-1}|Y_2^{i-1}Z_2^{i-1}F^{i-1}W\right). \qquad (4.76)
\end{aligned}
$$

Time-sharing constraints

The following constraints make sure that no more than n transmissions take place on each channel.

$$\begin{aligned} n &\geq \sum_{i=1}^{n} H(X_{1i}) \geq \sum_{i=1}^{n} H\left(X_{1i}|Y^{i-1}F^{i-1}\right) = \sum_{i=1}^{n} H\left(X_{1i}|Y^{i-1}Z_2^{i-1}F^{i-1}\right) + I\left(X_{1i};Z_2^{i-1}|Y^{i-1}F^{i-1}\right) \\ &= \sum_{i=1}^{n} H\left(X_{1i}|Y^{i-1}Z_2^{i-1}F^{i-1}W\right) + I\left(W;X_{1i}|Y^{i-1}Z_2^{i-1}F^{i-1}\right) + I\left(X_{1i};Z_2^{i-1}|Y^{i-1}F^{i-1}\right) \\ &\geq \sum_{i=1}^{n} H\left(X_{1i}|Y^{i-1}Z_2^{i-1}F^{i-1}W\right) + I\left(W;X_{1i}|Y^{i-1}Z_2^{i-1}F^{i-1}\right) \end{aligned} \tag{4.77}$$

From symmetry:

$$n \geq \sum_{i=1}^{n} H\left(X_{2i}|Y^{i-1}Z_1^{i-1}W\right) + I\left(W;X_{2i}|Y^{i-1}Z_1^{i-1}F^{i-1}\right) \tag{4.78}$$

Also,

$$\begin{aligned} n &\geq \sum_{i=1}^{n} H(X_{1i}) \geq \sum_{i=1}^{n} H\left(X_{1i}|Y^{i-1}F^{i-1}\right) = \sum_{i=1}^{n} H\left(X_{1i}|Y^{i-1}F^{i-1}W\right) + I\left(X_{1i};W|Y^{i-1}F^{i-1}\right) \\ &\geq \sum_{i=1}^{n} H\left(X_{1i}|Y^{i-1}Z_1^{i-1}F^{i-1}W\right) + I\left(W;X_{1i}|Y^{i-1}F^{i-1}\right) \end{aligned} \tag{4.79}$$

holds, and again from symmetry:

$$n \geq \sum_{i=1}^{n} H\left(X_{1i}|Y^{i-1}Z_2^{i-1}F^{i-1}W\right) + I\left(W;X_{1i}|Y^{i-1}F^{i-1}\right) \tag{4.80}$$

$$n \geq \sum_{i=1}^{n} H\left(X_{2i}|Y^{i-1}Z_1^{i-1}F^{i-1}W\right) + I\left(W;X_{2i}|Y^{i-1}F^{i-1}\right) \tag{4.81}$$

$$n \geq \sum_{i=1}^{n} H\left(X_{2i}|Y^{i-1}Z_2^{i-1}F^{i-1}W\right) + I\left(W;X_{2i}|Y^{i-1}F^{i-1}\right). \tag{4.82}$$

Further constraints

We need the following trivial constraints to complete our program.

$$\sum_{i=1}^{n} H\left(X_{1i}|Y^{i-1}Z^{i-1}F^{i-1}W\right) = \sum_{i=1}^{n} H\left(X_{1i}|Y^{i-1}Z_1^{i-1}F^{i-1}W\right) - I\left(X_{1i};Z_2^{i-1}|Y^{i-1}Z_1^{i-1}F^{i-1}W\right) \tag{4.83}$$

$$\sum_{i=1}^{n} H\left(X_{1i}|Y^{i-1}Z^{i-1}F^{i-1}W\right) = \sum_{i=1}^{n} H\left(X_{1i}|Y^{i-1}Z_2^{i-1}F^{i-1}W\right) - I\left(X_{1i};Z_1^{i-1}|Y^{i-1}Z_2^{i-1}F^{i-1}W\right) \tag{4.84}$$

$$\sum_{i=1}^{n} H\left(X_{2i}|Y^{i-1}Z^{i-1}F^{i-1}W\right) = \sum_{i=1}^{n} H\left(X_{2i}|Y^{i-1}Z_1^{i-1}F^{i-1}W\right) - I\left(X_{2i};Z_2^{i-1}|Y^{i-1}Z_1^{i-1}F^{i-1}W\right) \tag{4.85}$$

$$\sum_{i=1}^{n} H\left(X_{2i}|Y^{i-1}Z^{i-1}F^{i-1}W\right) = \sum_{i=1}^{n} H\left(X_{2i}|Y^{i-1}Z_2^{i-1}F^{i-1}W\right) - I\left(X_{2i};Z_1^{i-1}|Y^{i-1}Z_2^{i-1}F^{i-1}W\right) \tag{4.86}$$

$$\sum_{i=1}^{n} H\left(X_{1i}|Y^{i-1}Z^{i-1}F^{i-1}W\right) \geq \sum_{i=1}^{n} I\left(X_{1i};\Psi|Y^{i-1}Z^{i-1}F^{i-1}W\right) \tag{4.87}$$

$$\sum_{i=1}^{n} H\left(X_{2i}|Y^{i-1}Z^{i-1}F^{i-1}W\right) \geq \sum_{i=1}^{n} I\left(X_{2i};\Psi|Y^{i-1}Z^{i-1}F^{i-1}W\right) \tag{4.88}$$

$$\sum_{i=1}^{n} I\left(X_{1i};W|Y^{i-1}Z^{i-1}F^{i-1}\right) \leq \sum_{i=1}^{n} I\left(X_{1i};W|Y^{i-1}Z_1^{i-1}\right) + I\left(X_{1i};Z_2^{i-1}|Y^{i-1}Z_1^{i-1}F^{i-1}W\right) \tag{4.89}$$

$$\sum_{i=1}^{n} I\left(X_{2i};W|Y^{i-1}Z^{i-1}F^{i-1}\right) \leq \sum_{i=1}^{n} I\left(X_{2i};W|Y^{i-1}Z_2^{i-1}\right) + I\left(X_{2i};Z_1^{i-1}|Y^{i-1}Z_2^{i-1}F^{i-1}W\right) \tag{4.90}$$

The transformation shown in Appendix C.3 completes the proof of Theorem 4.3.

4.4 Triangle network

This section considers the triangle network, which consists of three nodes and three channels as it is shown in Figure 4.5. This is the first multi-hop network we investigate and we provide a complete characterization.

4.4.1 Model

We label the channels as Figure 4.5 indicates: channel 1 is the $S-D$ channel, channel 2 is the $S-U$ channel, while channel 3 is the $U-D$ channel. Channels operate as described in

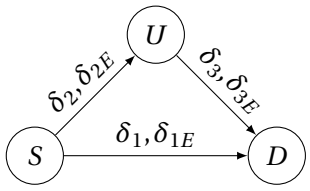


Figure 4.5: Triangle network

Chapter 1. The channel inputs are defined by:

$$X_{k,i} = f_{k,i}\left(W,\Theta_S,F^{i-1}\right), \quad k\in\{1,2\} \tag{4.91}$$

$$X_{3,i} = f_{3,i}\left(Y_2^{i-1},F^{i-1},\Theta_U\right). \tag{4.92}$$

D has to be able to decode from its receptions on channels 1 and 3:

$$\Pr\left\{\phi\left(Y_1^n, Y_3^n, F^n, \Theta_D\right) \neq W\right\} < \epsilon. \tag{4.93}$$

In our model we assume that node U can generate private randomness Θ_U at unlimited rate. In the next section we give an example on how this affects the achievable secret-message rate.

The eavesdropping adversary might select any one channel to wiretap, but network nodes are not aware of her choice. As previously in this chapter, we can think of three noncolluding eavesdroppers. Message W remains secret from each eavesdropper, thus the security condition becomes:

$$I\left(W; Z_k^n \Theta_E F^n\right) < \epsilon, \quad k \in \{1,2,3\}. \tag{4.94}$$

4.4.2 Main result

We characterize the secret-message capacity of the triangle network.

Theorem 4.4. *The secret-message capacity of the triangle is the optimal value of the following linear program, where all parameters $m_i, c_i, c, k_i, r_i \geq 0$:*

$$\max R, \quad \textit{such that:}$$

$$R \leq (1-\delta_1)\, m_1 + (1-\delta_3)\, m_3 \tag{4.95}$$

$$m_1 (1-\delta_1) \frac{1-\delta_{1E}}{1-\delta_1\delta_{1E}} \leq (k_1 + c_1)\, \delta_{1E} (1-\delta_1) + r_3 + c_3 (1-\delta_3) \tag{4.96}$$

$$m_2 (1-\delta_2) \frac{1-\delta_{2E}}{1-\delta_2\delta_{2E}} \leq k_2 \delta_{2E} (1-\delta_2) + k_1 (1-\delta_1) \tag{4.97}$$

$$m_3 (1-\delta_3) \frac{1-\delta_{3E}}{1-\delta_3\delta_{3E}} \leq (k_3 + c_3)\, \delta_{3E} (1-\delta_3) + (k_1 + c_1)(1-\delta_1) + r_3 \delta_{3E} \frac{1-\delta_3}{1-\delta_3\delta_{3E}} \tag{4.98}$$

$$1 \geq m_1 + k_1 + c_1 \tag{4.99}$$

$$1 \geq m_2 + k_2 \tag{4.100}$$

$$1 \geq m_3 + k_3 + c_3 + \frac{r_3}{1-\delta_3} \tag{4.101}$$

$$k_2 (1-\delta_2) \geq c + r_3 \tag{4.102}$$

$$c \geq (1-\delta_1\delta_{1E})\, c_1 + (1-\delta_3)\, c_3 \tag{4.103}$$

$$c \geq (1-\delta_3\delta_{3E})\, c_3 + (1-\delta_1)\, c_1 \tag{4.104}$$

$$(1-\delta_3)\, m_3 = (1-\delta_2)\, m_2 + c_1 (1-\delta_1)\,. \tag{4.105}$$

As usual, we prove Theorem 4.4 in two steps. In the next section we design a coding scheme that achieves the secret-message rate claimed by the theorem. The roles of constraints (4.95)-(4.105) become clear from the description of the scheme. The matching outer bound is provided in Appendix C.4.

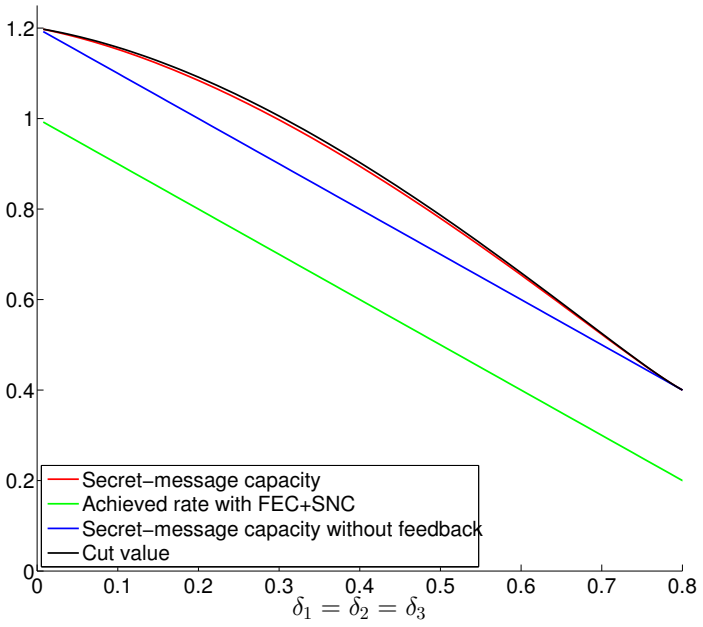


Figure 4.6: Comparison of secret-message rates with/without exploiting erasures and with-/without feedback. In all cases $\delta_{iE} = \delta_i + 0.2$.

Discussion

In the linear program we can observe the same structure that the previous LPs have. There is a rate constraint (4.95), three security constraints (4.96)-(4.98), three time-sharing constraints (4.99)-(4.101) and three common randomness constraints (4.102)-(4.104) with $k2(1-\delta_2)$ playing the role of the common randomness. The last constraint (4.105) describes the operation of the intermediate node as we will see shortly.

The triangle network has two cuts, the $S-UD$ and the $SU-D$ cut. The minimum value of these cuts is an obvious outer bound on the secret-message capacity. Note that the cuts consist of two parallel channels, thus we can find their values using the LP of Theorem 4.1. It might be expected that the solution of the LP in Theorem 4.4 reduces to the min-cut value – as in the case of non-secure message sending. In contrast, in the case of secure message sending, the min-cut value is not achievable in general.

Solving the LP in Theorem 4.4 enables us to evaluate (1) the benefit of exploiting erasures (2) the benefit of exploiting feedback (3) how much private randomness at the relay U can help. Figure 4.6 compares three schemes: secret-message capacity refers to our scheme in Theorem 4.4; we plot secret-message capacity without feedback to show the benefits of exploiting erasures for secrecy yet without using feedback [25, 58]; and finally FEC+SNC refers to applying a link-by-link error correction coding (FEC) and then using the secure network coding scheme [55, 56]. On the same plot, we show the value of the min-cut as well.

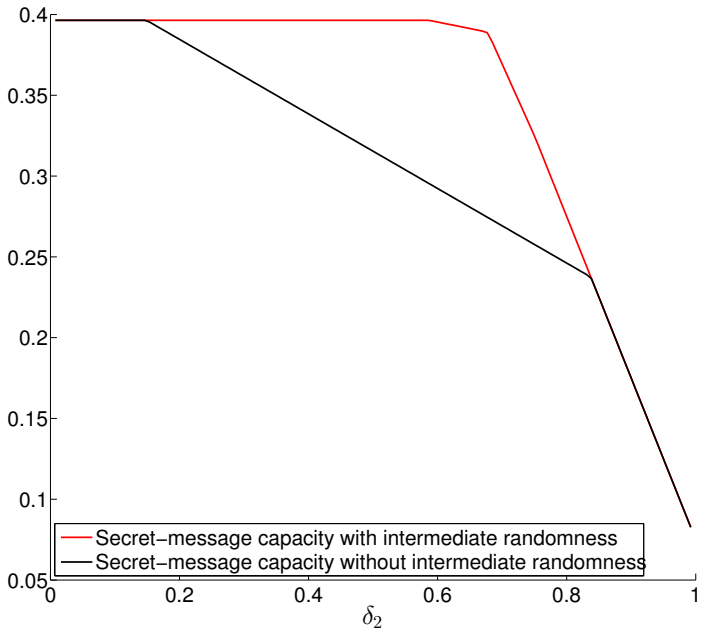


Figure 4.7: Comparison of secret-message capacities with/without private randomness at U. $\delta_1 = \delta_{2E} = 0.8$, $\delta_{1E} = 0.5$, $\delta_3 = \delta_{3E} = 0.3$.

Our other example (Figure 4.7) shows a case when the presence of private randomness at U increases secret-message capacity by more than 40% compared to the case where U does not use private randomness[1]. We note that in many cases the difference is less significant.

4.4.3 Coding scheme

This section proves the direct part of Theorem 4.4. We build our coding scheme for the triangle network using the tools that we have developed through the previous sections. Note that the relay U and the source S will share limited rate common randomness, from random packets that S sends to U through the $S-U$ channel, which is very similar to the V-network setup. Yet there is also a significant difference from the V-network: the relay U does not have the message, it can only receive it by consuming resources of the $S-U$ channel. Beside the encryption and key generation methods we use in the V-network we apply two new ideas explained below.

Principles

Our first observation is that keys generated on the $S-D$ channel can be used for encryption on both the $S-U$ and the $U-D$ channel, even though U does not have access to such keys. The encrypted message packets that travel on the $S-U-D$ path can potentially be encrypted with three types of key: key that only $S-U$ share (say K_{SU}); key that only $U-D$ share (say

[1] We do not prove the capacity result without private randomness at U here, but it can be easily verified using our proof that we obtain capacity by forcing $k_3 = 0$ in the LP.

K_{UD}), and key that S and D share (say K_{SD}). The source can send to U messages encrypted with K_{SU} and K_{SD}. U needs to remove K_{SU} from these encrypted messages (since D does not have it), yet it does not need to remove K_{SD}. That is, U does not need to completely decrypt the message, but instead it can recombine packets and secure part of the message sending phase on the $U-D$ channel with the K_{SD} key. Conceptually, U can use K_{SD} as if it had access to it.

The second new idea we use reduces the number of message packets U needs to receive. Assume that S creates a random packet P from the common randomness that is shares with U (as we did in the V-network). Instead of sending P, S now combines P with a message packet W_k and sends $P' = P \oplus W_k$ to D. Note that D cannot yet decode W_k. Packet P is available for U (since it is from the shared randomness), and it is transmitted on the $U-D$ channel using ARQ. This transmission is utilized in two ways. First, note that $W_k = P' \oplus P$, thus it allows D to decode a message packet. As far as U and D are concerned, transmission of P from U can be considered as part of the message sending phase, but there is no need to further encrypt such transmissions, because P is independent of the message. Second, as far as Eve on the $U-D$ channel is concerned, these are random key generation packets forwarded using ARQ, so they also contribute to the key K_{UD} on the $U-D$ channel.

The latter observation is counter intuitive for our usual flow based interpretation of network traffic. For D it is not always possible to tell through which path a certain message packet has arrived, because it depends on the interpretation of the packet. This gives us some flexibility and it overrules the intuition that it should not be possible to send more message packets on the $U-D$ channel than what was received by U in the message sending phase on the $S-U$ channel.

Detailed description

We use the parameters defined by the LP in Theorem 4.4. We rely on the properties of the key generation techniques that we have proved in the previous section and also on Theorem 2.2 to argue that the available key rate is sufficient to secure the corresponding message sending phase. This enables us to focus on the novelties of the scheme and use key generation and encryption techniques as building blocks while omitting the formal specification of details, which can be deduced from previously shown results. We believe that a detailed formal description would be repetitive and would compromise readability. Our phrasing is thus not completely formal. We often use "number of packets" instead of time fractions or rates. E.g., we say "S sends k random packets", which in the actual scheme means sending k' packets for some k' such that $\lim_{n\to\infty} \frac{1}{n} k' = k$, where n is the overall number of transmissions. Given our previous formal arguments, this does not compromise the mathematical rigor of our proof.

Key generation

1. $S-U$ *channel:* S sends k_2 i.i.d. uniform random packets. The packets that U receives form a $k_2(1-\delta_2)$ rate common randomness shared between S and U.

2. $S-D$ *channel:* S sends k_1 i.i.d. uniform random packets from its private randomness. It then utilizes the $k_2(1-\delta_2)$ packets it has in common with U with a scheme akin to the V-network, but with a modification: it divides the packets to two disjoint sets of c and r_3 packets. From the c packets, it creates two streams of rates c_1 and c_3; c_1 to be sent by S, c_3 by U. These packets are created by expanding the c packets through multiplication with the generator matrix of an MDS code (as seen in (4.62)). S sends the c_1 packets using the idea described above, i.e., XOR-ed with message packets. All c_1 such transmissions use a different packet created from the common randomness, while the same message packet is repeatedly used to form the XOR-ed packets until D acknowledges its reception. Thus all (encrypted) message packets are received, and all received transmissions are independent.
3. $U-D$ *channel:* U sends k_3 i.i.d. uniform random packets from its private randomness. Then, U sends the c_3 packets (each once), and finally sends the r_3 packets using ARQ.

Key rate on each channel

1. $S-U$ *channel:* The k_2 packets enable a key rate $k_2\delta_{2E}(1-\delta_2)$. Moreover, the k_1 packets sent through the $S-D$ channel also contribute to the encryption, resulting in an overall key rate

$$k_2\delta_{2E}(1-\delta_2)+k_1(1-\delta_1). \tag{4.106}$$

2. $S-D$ *channel:* From the $S-D$ channel's perspective there is no difference between the k_1 packets from the private randomness and the c_1 packets that are XOR-ed with message packets. Indeed, all these packets are i.i.d. random packets and they are independent of the message packets that are to be sent in the message sending phase of this channel. Additionally, there are $r_3+c_3(1-\delta_3)$ packets that D receives from U and S can also generate, which adds to a rate

$$(k_1+c_1)\delta_{1E}(1-\delta_1)+r_3+c_3(1-\delta_3). \tag{4.107}$$

3. $U-D$ *channel:* The k_3 private random packets, the c_3 common randomness packets, as well as the r_3 packets sent using ARQ, together result in a secret-key rate $(k_3+c_3)\delta_{3E}(1-\delta_3)+r_3\frac{\delta_{3E}(1-\delta_3)}{1-\delta_3\delta_{3E}}$. Moreover, U will send $c_1(1-\delta_1)$ packets from the $S-U$ common randomness using ARQ. For an eavesdropper on the $U-D$ channel these are random packets independent of the message, thus U can additionally use $c_1(1-\delta_1)\frac{\delta_{3E}(1-\delta_3)}{1-\delta_3\delta_{3E}}$ key packets (the rate follows from the key rate achieved by ARQ). Overall we have a key of rate

$$(k_3+c_3)\delta_{3E}(1-\delta_3)+r_3\frac{\delta_{3E}(1-\delta_3)}{1-\delta_3\delta_{3E}}+c_1(1-\delta_1)\frac{\delta_{3E}(1-\delta_3)}{1-\delta_3\delta_{3E}}. \tag{4.108}$$

Encryption and message sending phase

The message packets are split into: $c_1(1-\delta)$ packets delivered with the c_1 packets; $m_1(1-\delta)$ packets W_1 to be sent through the $S-D$ channel; and $m_2(1-\delta_2)$ packets W_2 to be sent through the $S-U-D$ path.

1. *$S-U$ channel:* Let $K_2^{(1)}$ and $K_2^{(2)}$ denote the matrices formed of the $k_1(1-\delta_1)$ and the $k_2\delta_{2E}(1-\delta_2)$ key packets, respectively. The encrypted packets W_2' are

$$W_2' = W_2 \oplus \left[K_2^{(1)} \quad K_2^{(2)}\right] \underbrace{\begin{bmatrix} G_2^{(1)} \\ G_2^{(2)} \end{bmatrix}}_{G_2} \tag{4.109}$$

   where G_2 is a $(k_1(1-\delta_1)+k_2\delta_{2E}(1-\delta_2))\times m_2(1-\delta_2)$ generator of an MDS code. Packets W_2' are sent using ARQ.

2. *$S-D$ channel:* Similarly, let K_1 denote the key created for the $S-D$ channel.

$$W_1' = W_1 \oplus K_1 G_1, \tag{4.110}$$

   where G_1 is a $(k_1\delta_{1E}(1-\delta_1)+c_3(1-\delta_3)+r3)\times m_1(1-\delta_1)$ MDS code generator. Packets W_1' are sent using ARQ.

3. *$U-D$ channel:* The message sending phase on the $U-D$ channel takes three steps. U first sends the $c_1(1-\delta_1)$ packets from the $S-U$ common randomness using ARQ; these enable D to decode the $c_1(1-\delta_1)$ message packets. U then calculates

$$W_2'' = W_2' \oplus K_2^{(2)} G_2^{(2)} = W_2 \oplus K_1^{(1)} G_2^{(1)}, \tag{4.111}$$

   to remove the $K_2^{(2)}G_2^{(2)}$ that D does not know. U computes

$$\left[W_{3a}' \quad W_{3b}'\right] = W_2'' G_3 = W_2'' \left[G_{3a} \quad G_{3b}\right], \tag{4.112}$$

   where G_3 is an $m_2(1-\delta_2)\times m_2(1-\delta_2)$ invertible matrix such that G_{3a} is of size $m_2(1-\delta_2)\times \min\left\{k_1(1-\delta_1)\frac{1-\delta_3\delta_{3E}}{1-\delta_{3E}}, m_2(1-\delta_2)\right\}$ and $G_2^{(1)}G_{3a}$ is the generator of an MDS code. W_{3a}' are sent using ARQ.

   Finally, let K_3 denote the key that U creates with the different key generation methods, as explained above. It uses K_3 to encrypt the remaining part of the message W_{3b}':

$$W_{3b}'' = W_{3b}' \oplus K_3 G_3', \tag{4.113}$$

   where G_3' is a $|K_3|\times\left(m_2(1-\delta_2)-k_1(1-\delta_1)\frac{1-\delta_3\delta_{3E}}{1-\delta_{3E}}\right)^+$ generator of an MDS code. Packets of W_{3b}'' are sent using ARQ.

4.4.4 Analysis

Rate D receives a rate $m_1(1-\delta_1)$ message from the message sending phase on the $S-D$ channel. In addition, the message sending phase on the $U-D$ channel delivers a message of rate $(1-\delta_3)m_3$ as long as it is not larger than $c_1(1-\delta_1)+m_2(1-\delta_2)$, which is the rate of message packets (or packets interpreted as message packets) that U has access to. The condition $(1-\delta_3)m_3 \le c_1(1-\delta_1)+m_2(1-\delta_2)$ is ensured by (4.105), hence rate $(1-\delta_1)m_1+(1-\delta_3)m_3$ is achieved.

Constraints (4.99)-(4.101) ensure that the scheme described above is feasible, i.e., no more than n transmissions are used on each channel.

Security We need to see if a sufficient key rate is available against all eavesdroppers whenever we send encrypted packets. The security of the scheme then follows from Theorem 2.2.

1. $S-U$ *channel:* It is clear from (4.97) that the key rate (4.106) available on this channel is sufficient to secure a message of rate $m_2(1-\delta_2)$.

2. $S-D$ *channel:* In the same way (4.96) ensures that the key rate (4.107) is sufficient to secure a message of length $m_1(1-\delta_1)$.

3. $U-D$ *channel:* The first set of packets (c_1 packets from the common randomness) are random packets that are independent of the message, thus no encryption is required and they cannot reveal any information to Eve about the message.

   Packets W'_{3a} are of the form

   $$W'_{3a} = W_2 G_{3a} \oplus K_1^{(1)} G_2^{(1)} G_{3a}. \tag{4.114}$$

   We see the same form of encryption as in Theorem 2.2, applied on the linear combination $W_2 G_{3a}$ as message packets and matrix $G_2^{(1)} G_{3a}$ for combining the keys $K_1^{(1)}$. The key rate of $K_1^{(1)}$ is $k_1(1-\delta_1)$, while the rate of W'_{3a} is

   $$k_1(1-\delta_1)\frac{1-\delta_3\delta_{3E}}{1-\delta_{3E}}, \tag{4.115}$$

   hence the rate of $K_1^{(1)}$ is sufficient to secure this message rate by Theorem 2.2.

   Consider the message packets W'_{3b}. The rate of W'_{3b} is $\left(m_2(1-\delta_2) - k_1(1-\delta_1)\frac{1-\delta_3\delta_{3E}}{1-\delta_{3E}}\right)^+$, while the key has rate $(k_3+c_3)\delta_{3E}(1-\delta_3) + (r_3+c_1(1-\delta_1))\frac{\delta_{3E}(1-\delta_3)}{1-\delta_3\delta_{3E}}$ by (4.108). Hence, for security we need that

   $$\begin{aligned} m_2(1-\delta_2) - k_1(1-\delta_1)\frac{1-\delta_3\delta_{3E}}{1-\delta_{3E}} \leq \\ (k_3+c_3)\delta_{3E}(1-\delta_3)\frac{1-\delta_3\delta_{3E}}{1-\delta_{3E}} + (r_3+c_1(1-\delta_1))\frac{\delta_{3E}(1-\delta_3)}{1-\delta_{3E}}. \end{aligned} \tag{4.116}$$

   Using (4.105) we get:

   $$\begin{aligned} m_3(1-\delta_3) - c_1(1-\delta_1) - k_1(1-\delta_1)\frac{1-\delta_3\delta_{3E}}{1-\delta_{3E}} \leq \\ (k_3+c_3)\delta_{3E}(1-\delta_3)\frac{1-\delta_3\delta_{3E}}{1-\delta_{3E}} + (r_3+c_1(1-\delta_1))\frac{\delta_{3E}(1-\delta_3)}{1-\delta_{3E}}. \end{aligned} \tag{4.117}$$

   After rearranging terms this condition becomes constraint (4.98), hence the security of message packets W'_{3b} is ensured by the feasibility of the LP.

This concludes the proof of the direct part of Theorem 4.4. We have seen that the scheme is feasible, it achieves the claimed rate and it ensures security against each eavesdropper. □

4.4.5 Outer bound

We complete the proof of Theorem 4.4 by providing a matching outer bound. We apply the same proof technique as we have used for the V-network. Showing equivalence of the outer bound linear program and the LP in Theorem 4.4 is a lengthy derivation. We delegate the whole proof to Appendix C.4.

4.5 Next steps

In this chapter we further developed our two-phase approach to the secret-message sending problem and derived secret-message capacity of small network settings. We used a linear programming framework to describe our schemes as well as to derive new outer bounds.

Even these small networks show nontrivial behavior. The tools that we used to find an optimal scheme – and especially to prove optimality – do not scale well for a network of arbitrary size. With the size of the network the complexity of the problem increases exponentially. The number of paths, the number of subsets of intermediate nodes that can generate and share randomness grow exponentially. In fact, finding the secret-message capacity is an NP-hard problem even if intermediate nodes do not generate randomness [61]. In the next chapter we consider arbitrary networks and target a trade-off between complexity and achieved secret-message rate.

5 Secret-message sending in arbitrary networks

In the preceding chapter we took the first steps towards understanding the secret-message sending problem in a multihop erasure network. We have seen that an optimal coding scheme needs to combine various key generation techniques and exploit some nontrivial coding solutions even in a small setting with no more than three network nodes. It is also known that finding the secret-message capacity of a general network is as hard as determining the capacity region of mulitple unicast network coding, which is a long-standing open problem [62, 63]. All these indicate that the complexity of pursuing optimality in a network increases with the size of the network.

In this chapter we target a trade-off: our goal is to design a simple scheme that is applicable for a network of arbitrary size and achieves a secret-message rate that is close to the optimum. In our design we follow our two-phase approach. In the previous chapter we have seen that the optimal scheme might require that the length of the phases are different on each channel. In this chapter we use a *global* two-phase scheme, by which we mean that the length of each phase is the same on each channel of the network. Here we also assume that intermediate nodes do not generate private randomness. This assumption fits well with the current networking philosophy of having the intelligence at the edge of the network and keeping intermediate node operations simple. Moreover, it allows a fair comparison with secure network coding [56], where the same assumption was imposed.

Our scheme simultaneously exploit erasures and the topology of the network for secrecy. First, we establish a connection between the two-phase approach and secure network coding. We show that these two are not fundamentally different, a secure network coding scheme can be converted to a two-phase scheme. This observation enables us to design a general scheme that reduces to the optimal two-phase scheme for a network with a single channel (point-to-point setting, see Chapter 2) and at the same time reduces to the optimal secure network coding scheme for an error-free network.

We also derive new outer bounds for the network setting and using these, we evaluate the achieved secret-message rate of our scheme. We believe that our proposed scheme represents a reasonable trade-off between complexity and achieved rate.

5.1 Related work

The most relevant related work for the current chapter is secure network coding by Cai and Yeung [56], where secret-message capacity of an error-free network was derived. In Section 5.3 we provide a short summary of the secure network coding scheme. This work was followed by a number of alternative constructions and extensions [64–66]. In [57] a generalization for erasure networks was presented, however feedback was not considered. The problem remains open in many nontrivial settings, e.g., for unequal link capacities, for nonuniform wiretap sets or with private randomness at intermediate nodes only partial results are available [61, 67–69].

5.2 Model

We introduce notation specific to this chapter and adapt the definitions in Section 1.2.

Communication takes place over a network which is represented by a directed acyclic multi-graph $\mathcal{G}(V,E)$, where V is the set of network nodes and E is the (multi-)set of edges.

Every edge $e=(u,v)\in E$ is an erasure channel with parameters δ,δ_E. The channels operate as described in Section 1.1.

In this chapter we consider not only unicast, but also multicast traffic. A source node $s\in V$ aims to send securely message to a set of destination nodes $D\subset V$. The source S can generate arbitrary amount of private randomness, however other network nodes do not use private randomness.

The multicast capacity of $\mathcal{G}$ with source S and destination nodes D is $h(1-\delta)$, where h denotes the number of edges in the smallest value min-cut between S and any $d\in D$. We introduce parameters $h=t+\ell$, where t is the number of multihop paths between S and D while ℓ is the number of direct S-D links in the smallest value min-cut. In case the smallest value min-cut is not unique, we choose the one that gives the smallest value for ℓ.

The eavesdropper Eve can select arbitrarily up to z edges of the network to wiretap. $A\subseteq E$ denotes the subset of wiretapped edges, where $|A|\leq z$. We assume that z is known as a design parameter, but A is known only by the eavesdropper. We assume that $z\leq h$, but we discuss the possibility of relaxing this restriction in Section 5.8.

The set of incoming and outgoing edges of $v\in V$ are denoted by I_v and O_v. If $\mathcal{E}\subseteq E$ then $Y_{i,\mathcal{E}}$ denotes the set of received packets by the network nodes in the ith time slot on the set of edges $\mathcal{E}$. Similarly for $\mathcal{V}\subseteq V$ the notation $Y_{i,\mathcal{V}}$ denotes the set of packets that the set of nodes $\mathcal{V}$ receives in the ith time slot. In case there are parallel edges the notation (u,v) means the set of edges starting from U and ending at v.

According to Definition 1.1, the channel inputs are defined as

$$X_{i,(s,v)}=f_{i,(s,v)}\left(W,\Theta_s,F^{i-1}\right) \tag{5.1}$$

$$X_{i,(u,v)} = f_{i,(u,v)}\left(Y_{I_u}^{i-1}, F^{i-1}\right), \quad \forall u \neq s. \tag{5.2}$$

For each receiver $d \in D$ the decodability condition (Definition 1.1, (1.7)) becomes:

$$\Pr\left\{\phi_d\left(Y_{I_d}^n\right) \neq W\right\} < \epsilon, \quad \forall d \in D. \tag{5.3}$$

The security requirement against the eavesdropper (Definition 1.4, (1.9)) is

$$I\left(W; Z_A^n, F^n\right) < \epsilon, \quad \forall A \subseteq E, |A| \leq z. \tag{5.4}$$

5.3 Secure network coding over error-free networks

In this section we shortly summarize the work of Cai and Yeung [56]. In the special case of error-free channels $\delta = \delta_E = 0$ our model becomes the same as seen in [56]. A linear coding scheme known as the secure network coding scheme is proposed that has secret-message rate $(h-z)^+$. The scheme uses source randomness of size z and ensures that all destination nodes receive both the message and the additional randomness. It is shown that the secure network coding scheme is optimal in terms of the achieved secret-message rate and it uses the minimum amount of additional randomness any optimal scheme might use.

Let us assume that Eve simply discards any packet that she receives more than once. We denote $z' \leq z$ the number of innovative packets Eve observes. Then, we can write Eve's observation in the following form:

$$Z_A = [\Theta \quad W] \begin{bmatrix} Q_A^{\Theta\mathsf{T}} \\ Q_A^{W\mathsf{T}} \end{bmatrix}. \tag{5.5}$$

Here $Q_A^{\Theta\mathsf{T}}$ is a $z \times z'$ matrix and $Q_A^{W\mathsf{T}}$ is a $(h-z) \times z'$ matrix. The secure network coding scheme has the property that the matrix $Q_A^{\Theta\mathsf{T}}$ has rank z'. Thus, $\Theta Q_A^{\Theta\mathsf{T}}$ is a set of $z' \leq z$ independent uniform random packets, while $WQ_A^{W\mathsf{T}}$ is a set of z' linear combinations of the message packets, hence from Eve's perspective what she observes is some data $WQ_A^{W\mathsf{T}}$ encrypted using one-time pad with key $\Theta Q_A^{\Theta\mathsf{T}}$.

One possible intuitive interpretation of these results is the following: to give perfect security against Eve who has access to at most z innovative packets, we need to send z packets of additional randomness and hence the secret-message capacity of the network is reduced by z compared to its multicast capacity. We use this intuition when we design our scheme for erasure networks.

5.4 Main result

The main result of this chapter is the design of secure coding schemes for arbitrary network topologies.

Theorem 5.1. *In a network, a secret-message rate*

$$R = \frac{h}{z\frac{1-\delta_E}{\kappa(1-\delta\delta_E)} + \frac{1}{1-\delta}}, \tag{5.6}$$

is achievable, where κ *corresponds to the key rate that our key generation phase achieves and it equals*

$$\kappa = h(1-\delta) - (z-t)^+(1-\delta)(1-\delta_E) - \min\{z,t\}(1-\delta)\frac{1-\delta_E}{1-\delta\delta_E}, \tag{5.7}$$

for a unicast problem, whereas it is

$$\kappa = h(1-\delta) - z(1-\delta)\frac{1-\delta_E}{1-\delta\delta_E}. \tag{5.8}$$

for a multicast problem.

Our coding scheme described in Section 5.6 provides a constructive proof of Theorem 5.1 (together with details delegated to Appendix D.2).

Clearly, if $\delta = \delta_E \to 0$ then $\kappa \to h - z$ and $R \to h - z$, hence in this special case the scheme achieves the same rate as the secure network coding scheme. Also, for $h = z = \ell = 1$ we have $\kappa = \delta_E(1-\delta)$ and get back $R = \delta_E(1-\delta)\frac{1-\delta\delta_E}{1-\delta\delta_E^2}$, the optimal rate of a single channel network.

We derive two outer bounds which give a basis for comparison and provides proof of optimality in a few further cases. The first bound is valid for any network and depends on h and z, while the other is valid for networks where $O_s = I_d = h$ and beside h and z, parameter t also plays a role. The ratio between our achieved rate and the outer bound is not larger than $\frac{1}{1-\delta}$. In most cases the gap is even smaller.

Theorem 5.2. *Assuming* $z \le h$*, for the achievable secret-message rate over* $\mathcal{G}$ *it holds that*

$$R \le (1-\delta)(h-z) + z\delta_E(1-\delta)\frac{1-\delta\delta_E}{1-\delta\delta_E^2}. \tag{5.9}$$

We note here that when $z > h$ we can substitute $z = h$ to get a valid upper bound for all cases, since the secret-message capacity cannot decrease by decreasing z.

Theorem 5.3. *Assuming* $O_s = I_d = h$*, for the achievable secret-message rate over* $\mathcal{G}$ *it holds that*

$$R \le (1-\delta)h - \min\{t,z\}\frac{(1-\delta_E)(1-\delta)}{1-\delta\delta_E}. \tag{5.10}$$

We provide the proofs of Theorems 5.2-5.3 in Appendices D.3-D.4. As a corollary of Theorems 5.2-5.3 we have the following optimality result.

Corollary 5.1. *Our scheme achieves secret-message capacity in the following cases:*

1. $h = \ell = z$,
2. $O_s = I_d = h$ *and* $t \geq z$,
3. $\delta_E = 0$ *or* $\delta_E = 1$.

5.4.1 Discussion, numerical examples

Applying a link-by-link error correction first and then using a secure network code directly results a secret-message rate $(h-z)(1-\delta)$. In the case where $\delta \geq \delta_E$, taking into account Eve's erasures, but not using feedback does not allow any better rates [58]. The advantage of exploiting feedback is twofold. First, it allows a higher key generation rate $\kappa \geq (h-z)(1-\delta)$. Second, it allows to reduce the size of the key we need in the second phase from $n_2 z(1-\delta)$ to $n_2 z(1-\delta)\frac{1-\delta_E}{1-\delta\delta_E}$. In this section we illustrate qualitatively how large this advantage is.

One can immediately see that the larger δ_E and z are the larger the advantage of exploiting erasures and feedback is. In particular, if $z = h$ our scheme still achieves a nonzero rate, which is not possible without feedback (assuming $\delta_E \leq \delta$).

In our example we consider the case when $\delta = \delta_E$. In this case, the highest achievable secret-message rate without feedback is $(h-z)(1-\delta)$. We consider a network with parameters $h = t = 10$ and $\delta = 0.3$. We plot in Figure 5.1 the advantage of our scheme as the ratio between R and $(h-z)(1-\delta)$. We see that in this case our scheme achieves a rate up to 3 times higher than the scheme without feedback. With the increase of the network size or $\delta = \delta_E$ the advantage becomes even larger. Note that we have selected the parameter values for the example such that there is no difference between the unicast and the multicast rate of our scheme.

We have optimality result only in some special cases. In order to see how the gap between our outer bound and achieved rate behaves we give a few numerical examples for the cases when there is a gap. Theorem 5.2 holds for any network, while Theorem 5.3 offers a potentially better bound when $O_s = I_d = h$. For cases when our achieved rate for unicast and for multicast differ we evaluate for the unicast rate.

First we consider the bound given by Theorem 5.2. We express the gap between the outer bound and the achieved rate as the ratio between the two (i.e., 1 means no gap). We evaluate for values $z \leq h$. We can observe that the gap takes its largest value if $h = t$, i.e., all paths between S and D are multihop paths. Further, we get the largest gap for $z = h$ and δ_E close to 0. In this case with $\delta_E \to 0$ the gap tends to $\frac{1}{1-\delta}$, which is the largest possible gap we might get. For other cases the gap is more moderate, see Figure 5.2 for a few examples.

For cases when $O_s = I_d = h$ holds, we can take the minimum of our two outer bounds. Note that in this case, when $z \leq t$ we do not have any gap, our scheme is optimal. Beside giving this optimality result, Theorem 5.3 offers an improvement over Theorem 5.2 for some cases also when $t > z$. As an example, on Figure 5.3 we compare the gap that the bound of Theorem 5.2

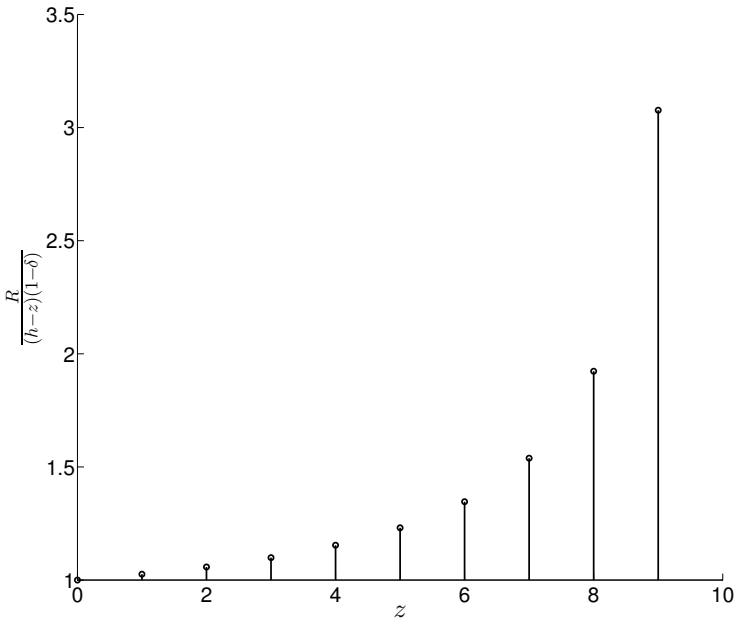


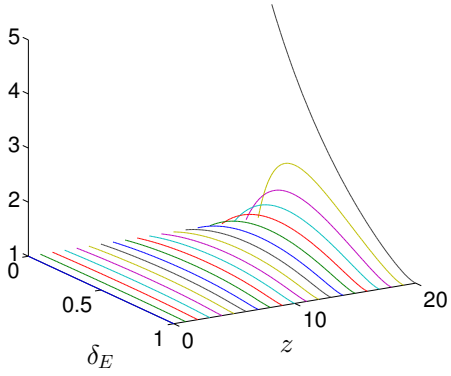
Figure 5.1: Advantage of using feedback as a function of the number of eavesdropped edges z, when $h = t = 10, \delta = \delta_E = 0.3$. In this case multicast and unicast rates are the same.

gives and the gap we get when taking the minimum of the two bounds. We plot the gap for a few specific values and for $z \geq t$ (for $z < t$ there is no gap in the second case).

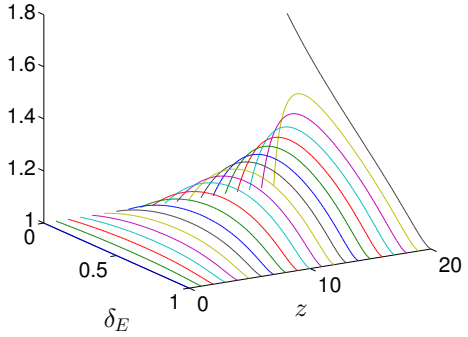
Both our scheme and our upper bounds are general in the sense that beside the min-cut value and the number of direct S-D channels they do not depend on the topology of the network. A more sophisticated network specific analysis could result both in higher achieved rates and in improved upper bounds. However, we see that for most parameter values the rate of the general scheme is already reasonably close to the upper bound.

5.5 Two-phase secure network coding scheme

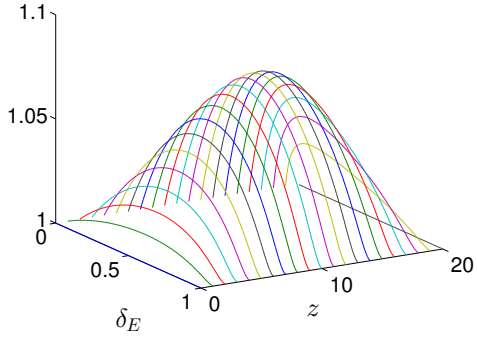
As a first step, in this section we examine only error-free networks. We show that the secure network coding scheme [56] – with an appropriate modification – can be cast as a two-phase scheme. We show that the modification does not affect the achieved rate, hence the two phase secure network coding scheme is also optimal. That is, we provide a new, alternative achievability scheme achieving the secret-message capacity of a error-free network. Separating the two phases makes it possible to consider key generation and message sending separately. This leads to our unified achievability scheme in Section 5.6 that accepts as special cases the (optimal) achievability scheme we provide next for error-free networks and the (optimal) achievability scheme for the point-to-point erasure channel that we described in Chapter 2.



(a) $h = t = 20$, $\delta = 0.8$

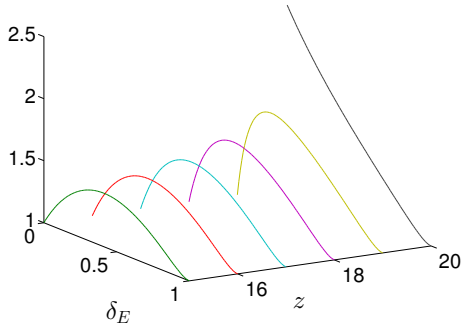


(b) $h = 20$, $t = 10$, $\delta = 0.8$

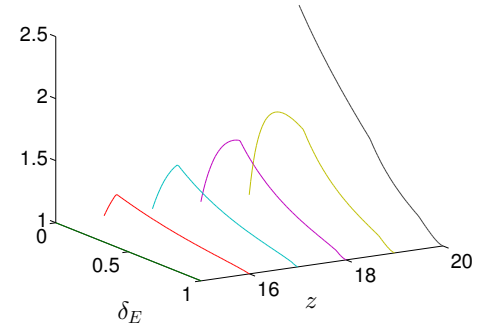


(c) $h = 20$, $t = 0$, $\delta = 0.8$

Figure 5.2: Upper bound/achieved rate (based on Theorem 5.2) for various parameter values



(a) $h = 20$, $t = 15$, $\delta = 0.8$



(b) $O_s = I_d = h = 20$, $t = 15$, $\delta = 0.8$

Figure 5.3: Upper bound/achieved rate (based on Theorems 5.2-5.3) for various parameter values

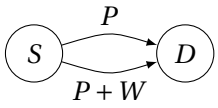


Figure 5.4: Secure network coding example

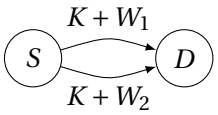
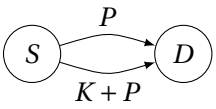
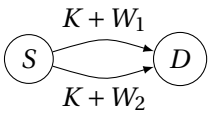


Figure 5.5: Coding with shared key



(a) First time slot



(b) Second time slot

Figure 5.6: Two-phase secure network coding scheme example

5.5.1 Example

For simplicity in this example we assume unicast traffic. Consider the following simple network (Figure 5.4). Source S and destination D are connected through two parallel unit capacity error-free links out of which any one is being wiretapped by Eve ($h = 2, z = 1$). The secrecy capacity of the network is 1, hence S can send securely a unit size message W. For the secure network coding scheme, source S generates a unit size randomness P. As shown in the figure, on one of the links S sends P while on the other link it sends $P + W$. The eavesdropper either sees P or $P + W$, in either case no information about W is leaked.

Assume now, that S and D already share a unit size random key K, which is not known by Eve. Then, as shown in Figure 5.5, S can securely send two unit size messages W_1 and W_2 using K for encryption on both links. Hence, in this case, a unit size shared key enables us to exploit the min-cut capacity of the network. One might ask, how S and D can set up a shared key. The secure network coding scheme offers a way to send any message W securely to D, this message can equally well be a key K. Consider the example in Figure 5.6, where in two time slots two messages are sent securely to D. In the first slot a key is set up, while in the second slot this key is used for encryption. Note that the achieved rate is 1, the same as what the secure network coding scheme achieves. Also the amount of additional randomness remains the same.

5.5.2 Scheme description

The properties that we have seen through our example can be generalized as follows. We call the scheme described below the two-phase secure network coding scheme. As opposed to the secure network coding scheme our scheme uses every link $n = n_1 + n_2$ times, where n_1 and n_2 are the number of time slots used for the two phases respectively. We use a secure network code as a building block, we select one such code at the outset and then in each of the n time slots we use the same code on different inputs. Hence, we have that $Q_{i,e} = Q_{j,e} = Q_e, \forall i, j$.

In our scheme the size of our message is $N = n_2 h$. To securely send a message of this size, we need a shared key K of size $n_2 z$ between S and the destination nodes.

Key generation:

The sender generates a uniformly random K of size $n_2 z$. It also generates additional randomness Θ of size $n_2 \frac{z^2}{h-z}$. The key generation phase consists of $n_1 = n_2 \frac{z}{h-z}$ time slots, in each slot S securely sends $h-z$ packets from K. On edge e in the ith slot we thus send

$$\begin{bmatrix} \Theta(i) & K(i) \end{bmatrix} Q_e, \tag{5.11}$$

where $K(i)$ is the ith $h-z$ length fraction of K: $K(i) = K_{(i-1)(h-z)+1 \ldots i(h-z)}$. Similarly, $\Theta(i)$ is the ith z length fraction of Θ: $\Theta(i) = \Theta_{(i-1)z+1 \ldots iz}$.

Encrypted message sending:

In the second phase we use K for encryption and in each slot h message packets are sent securely. We use again the same secure network code n_2 times. We denote $W(i)$ the first $h-z$ elements of the ith h length fraction of W and $W'(i)$ the last z elements of the same fraction. $K'(i)$ is the ith z length fraction of K. On edge e in the ith slot of the second phase we then send

$$\begin{bmatrix} W'(i) + K'(i) & W(i) \end{bmatrix} Q_e. \tag{5.12}$$

It directly follows from the properties of the secure network code that we use that all destination nodes know K and hence can decode W.

Analysis

Building on the security of the secure network code we show that the scheme is secure. We delegate the proof of security to Appendix D.1.

Our scheme conveys a message of size $n_2 h$ using $n_1 + n_2$ transmissions, thus our rate is

$$\frac{n_2 h}{n_1 + n_2} = \frac{n_2 h}{n_2 \frac{z}{h-z} + n_2} = h - z, \tag{5.13}$$

which is the same as the rate of the secure network coding scheme. We further note that the amount of randomness we use is $|K| + |\Theta| = n_2 \frac{hz}{h-z}$, which is also the same as the amount of randomness that the secure network coding scheme uses to securely send a message of size $n_2 h$. By selecting $n_2 = h - z$ the rate $h - z$ is achieved in a finite block length.

5.5.3 Discussion

In the case of error-free networks separating the two phases does not make any difference in the achievable secret-message rate, since the rate of key generation is the same as the achievable secret-message rate. However, in some cases this might not hold and a higher key generation rate is possible. In those cases designing the two phases separately results in an improved secret-message rate. We have seen that the point-to-point erasure channel is an example, where a gap between the secret-key capacity and the secret-message capacity exists.

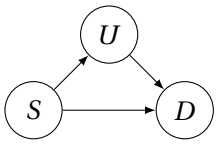


Figure 5.7: Example network – triangle topology

5.6 Coding scheme

We provide the proof of Theorem 5.1 by designing a coding scheme for an arbitrary network. We note that although our model allows that channel state is available for every network node, in our scheme, on every channel, transmissions depend only on the state of the given channel.

We first describe our scheme for a single receiver node $D = \{d\}$, then in Section 5.6.3 we generalize our description for multicast. We also assume here that $z \leq h$. This assumption was necessary in the case of an error free network to achieve any nonzero rate securely, however in an erasure network we can achieve some rate even if $z > h$. We discuss this case in a subsequent section.

5.6.1 Example

Before a detailed description, we explain ideas through the example network in Figure 5.7. This network is a special case of the more general triangle network discussed in the previous chapter. Let $z = 1$. For simplicity, in the example when we calculate the number of received packets we work with expected values instead of random variables.

Key generation:

Source S sends independent random packets over all its outgoing links. Both the destination node D and the intermediate node U receive $n_1(1-\delta)$ packets. On the link between U and D the packets that U received are then sent to D using ARQ. To complete this task U needs n_1 transmissions.

The achievable key rate corresponds to the number of packets that D receives but Eve does not. Eve has three possible choices to select a wiretapped link, and when generating the key we need to consider her worst-case selection.

Case 1: Eve selects the S-D link. In this case the number of packets that both D and Eve receive is $n_1(1-\delta)(1-\delta_E)$.

Case 2: Eve selects the U-D link. Since D eventually receives every packet that U has and every packet is repeated potentially several times, the probability that Eve overhears a certain packet of D is increased to $\frac{1-\delta_E}{1-\delta\delta_E}$. Node U sends $n_1(1-\delta)$ different packets, hence Eve has $n_1(1-\delta)\frac{1-\delta_E}{1-\delta\delta_E}$ packets in common with D.

Case 3: Eve selects the S-U link. We know that all the packets that U receives D also receives. Eve and U have $n_1(1-\delta)(1-\delta_E)$ packets that they both receive, we get the same result as in the first case.

We conclude that Eve's best choice (from her perspective) is the U-D link. Destination D has

$$2n_1(1-\delta) - n_1(1-\delta)\frac{1-\delta_E}{1-\delta\delta_E} \tag{5.14}$$

packets not received by Eve, hence a key rate $1-\delta+\frac{(1-\delta)^2\delta_E}{1-\delta\delta_E}$ is achievable.

Encrypted message transmission:

There are two edge disjoint paths between S and D. Let n_2 be the number of transmissions in the second phase. The message is encrypted in the form that we have already seen: $W_E = W + KG$, where K is the key and G is an MDS generator matrix. W_E is split into two parts and each half of the message is assigned to one of the paths. The message packets are then forwarded towards D using ARQ on each link.

The size of the key K we use has to equal the number of packets Eve receives in the second phase. In this case, the MDS property of G ensures that Eve receives every packet with an independent linear combination of K, ensuring security of the scheme.

Since the same forwarding strategy is applied on each link, regardless of which link Eve selects, she receives a certain packet with probability $\frac{1-\delta_E}{1-\delta\delta_E}$, thus she receives overall $n_2(1-\delta)\frac{1-\delta_E}{1-\delta\delta_E}$ different packets. Hence n_1 and n_2 are chosen such that $|K| = n_2(1-\delta)\frac{1-\delta_E}{1-\delta\delta_E}$.

5.6.2 Algorithm

As a first step we select h edge disjoint paths between S and D. We ignore all other edges of $\mathcal{G}$. The example in the previous section suggests that the achievable rate depends not only on h, z, δ and δ_E, but also on the number of direct S-D links, i.e on ℓ and t. Recall that $h = \ell + t$, where ℓ denotes the number of direct S-D links and t denotes the number of multihop paths.

Key generation:

We define

$$n_1' = n_1 - n_1^{\frac{3}{4}} \tag{5.15}$$

$$\zeta_1 = n_1'(z-t)^+(1-\delta)(1-\delta_E) + n_1'\min\{z,t\}(1-\delta)\frac{1-\delta_E}{1-\delta\delta_E} \tag{5.16}$$

$$|K| = hn_1'(1-\delta) - \zeta_1 - n_1'^{\frac{3}{4}}. \tag{5.17}$$

Parameter ζ_1 corresponds to the number of packets Eve might receive in the first phase[1]. Source S sends at most n_1 random packets on all its h outgoing edges. It stops transmission on each link as soon as $n_1'(1-\delta)$ packets are acknowledged on the given link. Intermediate nodes on each path forward the $n_1'(1-\delta)$ packets that they receive to the next node on the path towards D using ARQ.

[1] More precisely, with a probability arbitrarily close to 1, Eve does not receive more packets than ζ_1.

If D does not receive $hn_1'(1-\delta)$ packets, then an error is declared. Otherwise, let M denote the vector of all the packets that D receives. Both S and D compute

$$K = MH, \tag{5.18}$$

where H is a $\left(hn_1'(1-\delta) \times |K|\right)$ matrix and it is a parity check matrix of an MDS code.

Encryption and message sending:

We find N, n_2 and n_2' such that

$$\zeta_2 = n_2' z (1-\delta) \frac{1-\delta_E}{1-\delta\delta_E} \tag{5.19}$$

$$|K| = \zeta_2 + n_2'^{\frac{3}{4}} \tag{5.20}$$

$$n_2' = n_2 - n_2^{\frac{3}{4}} \tag{5.21}$$

$$N = hn_2'(1-\delta). \tag{5.22}$$

Similarly to ζ_1, parameter ζ_2 corresponds to the number of packets Eve might receive in the second phase. The encrypted message W_E is computed as

$$W_E = W + KG, \tag{5.23}$$

where K is the key from the first phase and G is a $(|K| \times N)$ matrix and it is a generator of an MDS code.

We assign $n_2'(1-\delta)$ packets to each of our paths. These packets are then forwarded on their assigned path to D using ARQ over each link. If D does not receive all the packets of W_E after n_2 transmissions, then an error is declared.

5.6.3 Multicast

In this section we present our scheme for the multicast problem, where there are more than one destination nodes and all of them have to receive the same message securely. Compared to the unicast scheme only a few modifications are needed. To avoid repetition, below we highlight only the differences.

Instead of h edge disjoint paths, first we need to find a network code over $\mathbb{F}_q$ for multicasting at rate $(1-\delta)h$. Again, we can ignore all edges that are not used by the network code.

In the key generation phase we need the following modification. Instead of sending new random packets on the outgoing edges, S selects in advance $n_1' h(1-\delta)$ random packets that are sent reliably to all destination nodes using ARQ on each link and applying the network code that we have chosen. The same network code is used in each time slot. This ensures that all $d \in D$ receive the same set of packets and hence they all can compute the same key.

According to this we modify parameter ζ_1:

$$\zeta_1 = n_1' z (1-\delta) \frac{1-\delta_E}{1-\delta\delta_E}. \tag{5.24}$$

Note that this change implies a change of parameters $|K|$, n_2, n_2', ζ_2 and N, however all formulas remain the same as defined for unicast.

In the second phase the only difference is that instead of forwarding through h edge-disjoint paths we use the network code (together with ARQ) to reliably send the encrypted packets to all destinations.

Another modification is needed in the selection of matrices H and G. Note that in the unicast case intermediate network nodes do not perform any coding, hence Eve might only receive packets that S produces. This property enables to code only at the source using any H and G matrices that have the MDS property. In the case of multicast, intermediate nodes might produce new linear combinations, hence Eve might receive combined packets as well.

As for matrix H, consider the $hn_1'(1-\delta)$ packets that S sends in the key generation phase and all their different linear combinations that the prescribed network code produces. Let $Q_A^{n_1}$ denote a coefficient matrix of size $hn_1'(1-\delta) \times hn_1'(1-\delta) - |K|$ that describes a $hn_1'(1-\delta) - |K|$ size subset of these packets. This subset corresponds to a set of packets that Eve might receive. We will see during the analysis that the probability that Eve receives a larger subset of packets is negligible. We select H such that $\begin{bmatrix} H & Q_A^{n_1} \end{bmatrix}$ is a full rank (in fact invertible) matrix for all possible $Q_A^{n_1}$. This property ensures the security of the generated keys. In [70] it was shown that if H is a parity check matrix of a rank metric code over $\mathbb{F}_{q^L}$, this property is satisfied independently of $Q_A^{n_1}$.

As for matrix G, we consider a $|K|$ size subset of the different encoded packets that Eve might receive during the second phase. Let $Q_A^{n_2 K}$ denote the $|K| \times |K|$ coefficient matrix of Eve's possible receptions that contain the coefficients of packets from K. We select G such that all possible such $Q_A^{n_2 K}$ matrix is invertible. As shown in our analysis this property ensures security of the message. Using again a parity check matrix H' of a rank metric code over $\mathbb{F}_{q^L}$, we can find such a G as the last $|K|$ rows of $\begin{bmatrix} H' & Z \end{bmatrix}^{-1}$, where H' has size $n_2 h(1-\delta) \times n_2 h(1-\delta) - |K|$ and Z is a full rank matrix over $\mathbb{F}_q$ [70].

5.6.4 Analysis

The proof of security and low error probability of our scheme uses the same techniques that we have seen in the previous chapters. We prove that the size of the key that we set up in the first phase is sufficient to encrypt the message in the second phase. We further rely on the observation that due to the ARQ that we apply, Eve receives the most number of different packets if she selects edges from different paths. We delegate details to Appendix D.2.

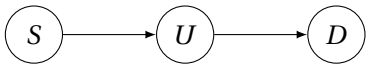


Figure 5.8: Two-hop line network

5.7 Outer bounds

Theorems 5.2-5.3 provide outer bounds on the achievable secret-message rate. We highlight some concepts and provide the complete proofs in Appendix D.3-D.4. When deriving our upper bounds we make the following two assumptions which can only increase the achievable rates: (a) The set of eavesdropped edges are known, hence we restrict Eve to one particular selection of edges. (b) The state of the eavesdropper's channel is also known to every node in the network. Both theorems rely on the assumption that intermediate nodes do not use private randomness.

Theorem 5.2 holds for any network. After a transformation of the network the proof generalizes the converse part of Theorem 2.1.

Theorem 5.3 considers a special class of network. By intuition, the restriction $O_s = I_d = h$ ensures that the deletion of edges that are not on a path towards a destination node does not reduce the achievable secret-message rate.

5.8 Discussion

5.8.1 Extension for $z > h$

Assume $z = 2$ and consider the two-hop line network shown in Figure 5.8. Against this stronger Eve, we can run our scheme as presented in Section 5.6, but with different parameters. We need to calculate how many packets Eve might receive in each phase. We give the calculation in expectation.

In the message sending phase Eve has two independent chances to overhear a certain packet, on each link she receives a given packet with probability $\frac{1-\delta_E}{1-\delta\delta_E}$, hence the number of different packets she receives (in expectation) is:

$$n_2 \frac{1-\delta_E}{1-\delta\delta_E} + n_2 \frac{1-\delta_E}{1-\delta\delta_E} \left(1 - \frac{1-\delta_E}{1-\delta\delta_E}\right). \tag{5.25}$$

In the key generation phase she gets $n_1 (1-\delta)(1-\delta_E)$ packets in common with U on the first link, while she receives a packet with probability $\frac{1-\delta_E}{1-\delta\delta_E}$ on the second link, hence she is expected to get

$$n_1 (1-\delta)(1-\delta_E) + n_1 (1-(1-\delta)(1-\delta_E)) \frac{1-\delta_E}{1-\delta\delta_E} \tag{5.26}$$

packets in common with D, which results a key rate

$$\kappa = \delta_E (1-\delta) - (1-(1-\delta)(1-\delta_E)) \frac{1-\delta_E}{1-\delta\delta_E}. \tag{5.27}$$

To calculate the achievable rate $\frac{(1-\delta)n_2}{n_1+n_2}$ we need to consider n_1 and n_2 such that

$$n_1\kappa = n_2 \frac{1-\delta_E}{1-\delta\delta_E} + n_2 \frac{1-\delta_E}{1-\delta\delta_E}\left(1 - \frac{1-\delta_E}{1-\delta\delta_E}\right). \tag{5.28}$$

Note that for any given network and any given set of wiretapped edges a similar analysis is feasible. After investigating all the $\binom{|E|}{z}$ possible sets of wiretapped edges, we can design our code such that it provides secrecy against all possible eavesdropped sets. However, the worst-case selection of eavesdropped edges and thus the actual rates achieved highly depends on the topology of our network.

5.8.2 Intermediate randomness helps

In our model we assume that no intermediate node in the network can generate additional randomness. Relaxing this assumption makes the problem significantly more difficult which is open even for an error-free network. Consider as an example the two hop line network (Figure 5.8) with $z = 1$. In case U can also randomize, then we can run our optimal protocol for a single channel twice, first to send the message securely from S to U and second to send it from U to D. This strategy achieves a rate $\delta_E(1-\delta)\frac{1-\delta\delta_E}{1-\delta\delta_E^2}$. Without randomization we have seen that the secure capacity is $\delta_E\frac{(1-\delta)^2}{1-\delta\delta_E}$. Clearly, intermediate randomness increases the secure capacity in certain cases.

5.8.3 Unicast rate $\stackrel{?}{\geq}$ multicast rate?

Note that this scheme for multicast (see Section 5.6.3) results in a lower secure multicast rate than the minimum of the secure (unicast) communication rates to the individual destinations achieved by the scheme in the preceding section. There is such a gap in case $z > t$, where t is the number of multihop paths between S and the destination with the lowest secure unicast capacity.

Considering the example on Figure 5.9 with $z = 2$ helps understanding why this gap shows up. If we had a unicast problem, with D_1 as the only destination, then we could achieve a key rate

$$\delta_E(1-\delta) + \frac{\delta_E(1-\delta)^2}{1-\delta\delta_E} \tag{5.29}$$

between S and D_1. However, the unicast scheme requires S to send a new random packet on the S-D_1 channel in each time slot of the key generation phase meaning that S has no control over which packets reach D_1 and contribute to the key. If S applies the same strategy over the S-D_2 channel then D_1 and D_2 share two different keys with S. But in the second phase S

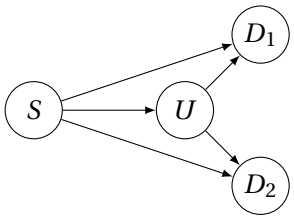


Figure 5.9: Multicast problem example

cannot use two different keys at the same time on the S-U channel, because the destinations would not be able to decode. Our proposed scheme overcomes this issue by ensuring that D_1 and D_2 receive the same set of packets and can generate the same key. Compared to unicast, this solution comes with a sacrifice in the key generation rate. Instead of the unicast key rate in (5.29) the multicast key rate is

$$2\frac{\delta_E(1-\delta)^2}{1-\delta\delta_E}. \tag{5.30}$$

Whether there is a fundamental gap between achievable unicast and multicast rates for $z > t$ remains open.

6 Discussion and open problems

In this thesis we investigated the problem of secret-message sending in various erasure settings. We introduced a two-phase scheme design and a linear programming approach which enabled us to propose optimal coding schemes in several communication scenarios. In particular, we have derived the secret-message capacity of a point-to-point channel, of a broadcast erasure channel[1], a network with independent parallel channels, the V-network and the triangle network. We have shown optimality of our scheme designs using novel techniques for deriving outer bounds. We also considered the secret-message sending problem in an erasure network with arbitrary topology, where we proposed a general coding scheme.

Our research also leads to a few related open problems. In Section 5.8 we have already mentioned some of these. Below we provide some questions of interest for future research.

- We have considered erasure channels only. A very natural direction is to consider channel models other than the erasure channel and further explore the role of feedback in secure communication.
- We only considered networks composed of point-to-point channels that operate independently. It would be relevant for a multihop wireless network to consider broadcast transmissions through multiple hops, which would necessarily bring together ideas from Chapters 3 and 4.
- Our outer bound proofs assume in all cases that the channel state of the eavesdropper is available. Clearly, the outer bound is valid with this assumption and it is tight in many cases – as we have seen. However, we conjecture that in some cases the resulting bound is not achievable without knowing Eve's channel state. We have already seen a similar phenomena for networks, where we have shown that knowing or not knowing the location of Eve (i.e., the set of eavesdropped channels) makes a difference in the achievable secret-message rate – unlike in the case of an error-free network.

  Consider a broadcast erasure channel with two receivers and an eavesdropper. State-feedback is available from the two receivers, who receive a private message each. Assume

[1] For special cases only: see Chapter 3.

we aim to ensure secrecy of both messages against the passive eavesdropper.[2] In this setting we conjecture that a different secret-message rate region is achievable with and without knowing the channel state of the eavesdropper. If our conjecture is true, a new technique is required to derive a better outer bound for the problem.

- As we discussed in Section 3.4.2, it remains open whether or not the distribution independent secret-message capacity region differs from the secret-message capacity region, i.e., whether $\mathscr{R}^2_{DIS} \subset \mathscr{R}^2_{DH}$ or $\mathscr{R}^2_{DIS} = \mathscr{R}^2_{DH}$ holds. We conjecture the former.
- We considered networks of arbitrary topology assuming that all channel parameters are the same. It remains open to generalize ideas for networks with arbitrarily varying channel parameters. We have to note that the problem becomes significantly harder, because a special case of the problem is the same as the secure network coding problem with unequal link capacities, which is known to be NP-hard [67].
- We have derived theoretical results and proposed polynomial time coding schemes for secure communication. There are still several practical challenges to address in order to translate these results into practical protocols.

[2] Note the difference between this setting and the model in Chapter 3. There each message is secured against all other receivers, whereas here we consider all messages to be secured against one eavesdropper.

A Summary of often used tools

For reference here we cite technical results we often use in our derivations. We state theorems without proofs, but we also provide references for an interested reader.

A.1 One-time pad encryption

One-time pad encryption, also known as Vernam cipher was already used for hiding information in the early 20th century, while historical documents prove that the idea appeared as early as 1882 [71]. The name suggests that no encryption key should be used more than once, which is indeed an important security criterion. A rigorous mathematical treatment was presented by Shannon [4], which we take as basis here.

Let W denote random finite field element (with any distribution), which represents a message and K another element of the same field selected uniformly at random. In this thesis we perform operations over packets, hence the finite field is $\mathbb{F}_q^L$ in our case. Then, the encrypted message is

$$W' = W \oplus K, \tag{A.1}$$

where $\oplus$ denotes the addition operation of the field. We use the $\oplus$ operator, because if the field is a binary extension field, the addition is bitwise XOR.

Theorem A.1. *For W' and K as defined above*

$$I(W;W') = 0, \tag{A.2}$$

i.e., one-time pad encryption provides perfect secrecy against an adversary who sees only W', and has no side-information about W or K.

For further details and proofs we refer to [4].

A.2 Chernoff-Hoeffding bound

We use the following form of the Chernoff-Hoeffding bound. For the proof and other related inequalities we refer to [72].

Theorem A.2. *Let $X_1, \ldots, X_n$ denote independent random variables taking values in $\{0,1\}$, and let $X = \sum_{i=1}^{n} X_i$. Further, $\mu = \mathbb{E}\{X\}$. Then,*

$$\Pr\{X \leq \mu - \tau\} \leq \mathrm{e}^{-\frac{2\tau^2}{n}} \tag{A.3}$$

$$\Pr\{X \geq \mu + \tau\} \leq \mathrm{e}^{-\frac{2\tau^2}{n}}. \tag{A.4}$$

As a direct consequence,

$$\Pr\{|X - \mu| \geq \tau\} \leq 2\mathrm{e}^{-\frac{2\tau^2}{n}}. \tag{A.5}$$

A.3 MDS matrices

A linear $[n,k]$ code over $\mathbb{F}_q$ is described by a $k \times n$ generator matrix G. An information (row-) vector x of k symbols over $\mathbb{F}_q$ is assigned the codeword xG. A linear code $\mathscr{C}$ is the set of all codewords, defined by

$$\mathscr{C} = \left\{ c \in \mathbb{F}_q^n | \exists x \in \mathbb{F}_q^k : c = xG \right\}. \tag{A.6}$$

The parity check matrix H of an $[n,k]$ linear code is the $n \times (n-k)$ matrix for which

$$c \in \mathscr{C} \Longleftrightarrow cH = 0. \tag{A.7}$$

An MDS code is a code that meets the Singleton-bound with equality. The Singleton-bound claims the following:

Theorem A.3. *For a code $\mathscr{C}$ over $\mathbb{F}_q$ with block length n and minimum distance $d \leq n$, the number of codewords is upper bounded by*

$$|\mathscr{C}| \leq q^{n-d+1}. \tag{A.8}$$

If $\mathscr{C}$ is a linear $[n,k]$ code, then $d \leq n-k+1$.

MDS codes exist for any parameters $[n,k]$ over $\mathbb{F}_q$, provided q is large enough. E.g. Reed-Solomon codes are linear MDS codes over a filed of size $q > n$.

Theorem A.4. *A linear $[n,k]$ code over $\mathbb{F}_q$ with parity check matrix H is an MDS code if and only if any $(n-k)$ rows of H are linearly independent.*

Theorem A.5. *The dual of an MDS code is again an MDS code. Thus, if H is the parity check matrix of an $[n,k]$ MDS code, then $H^\intercal$ is the generator of an $[n,n-k]$ MDS code.*

Corollary A.1. *Let X denote a row vector of ℓ random packets, i.e., X is an $L \times \ell$ matrix over $\mathbb{F}_q$ such that each column of X is selected independently and uniformly at random from $\mathbb{F}_q^L$. Further, let H^ℓ denote an arbitrary selected subset of $\ell \leq n$ rows of a parity check matrix of a linear $[n,k]$ MDS code over $\mathbb{F}_q$. Then, the entropy of XH^ℓ expressed in terms of packets is:*

$$H\left(XH^\ell\right) = \min\{\ell, n-k\}. \tag{A.9}$$

Corollary A.2. *Let X denote a row vector of k random packets, i.e., X is an $L \times k$ matrix over $\mathbb{F}_q$ such that each column of X is selected independently and uniformly at random from $\mathbb{F}_q^L$. Further, let G^ℓ denote an arbitrary selected subset of $\ell \leq n$ columns of a generator matrix of a linear $[n,k]$ MDS code over $\mathbb{F}_q$. Then, the entropy of XG^ℓ expressed in terms of packets is:*

$$H\left(XG^\ell\right) = \min\{\ell, k\}. \tag{A.10}$$

For proofs of the above theorems and for more insight about error correcting codes we refer to [73, 74].

B Proofs for Chapter 3

B.1 Proof of Lemma 3.3

Let U_B^C be a vector of length N_1 such that the i-th element $U_{B,i}^C$ is $U_{B,i}$ if Calvin observes this $U_{B,i}$ either in the pure form or added with some element of U_C, and $U_{B,i}^C = \perp$ otherwise. Let $1_{B,i}^C$ be the indicator random variable for the event $U_{B,i}^C \neq \perp$, so $M_B^C = \sum_{i=1}^{N_1} 1_{B,i}^C$. The following are information equivalent, i.e., we can express each side as a deterministic function of the other.

$$\left(Y_2^n, F^n, \Theta_C, U_C\right) \equiv \left(U_B^C, Y_2^{n_1}, F^n, \Theta_C, U_C\right). \tag{B.1}$$

Therefore,

$$H\left(Y_2^n F^n \Theta_C U_C\right) = H\left(U_B^C Y_2^{n_1} F^n \Theta_C U_C\right). \tag{B.2}$$

$$H\left(Y_2^n | Y_2^{n_1} F^n \Theta_C U_C\right) = H\left(U_B^C | Y_2^{n_1} F^n \Theta_C U_C\right) \tag{B.3}$$

$$= \sum_{i=1}^{N_1} H\left(U_{B,i}^C | U_B^{C\,i-1} Y_2^{n_1} F^n \Theta_C U_C\right) \tag{B.4}$$

$$= \sum_{i=1}^{N_1} H\left(U_{B,i}^C | 1_{B,i}^C U_B^{C\,i-1} Y_2^{n_1} F^n \Theta_C U_C\right) \tag{B.5}$$

$$\leq \sum_{i=1}^{N_1} H\left(U_{B,i}^C | 1_{B,i}^C\right) \tag{B.6}$$

$$\leq \sum_{i=1}^{N_1} \Pr\left\{1_{B,i}^C = 1\right\} = \mathbb{E}\left\{\sum_{i=1}^{N_1} 1_{B,i}^C\right\}. \tag{B.7}$$

where the third equality follows from the fact that the indicator random variable $1_{B,i}^C$ is a deterministic function of the conditioning random variables. □

B.2 Proof of Lemma 3.4

We adopt the notation for U_B^C and $1_{B,i}^C$ introduced in the proof of Lemma 3.3. In addition, let ${K'}_B^C$ be defined in a similar manner as U_B^C such that ${K'}_{B,i}^C = \perp$ if $U_{B,i}^C = \perp$ and ${K'}_{B,i}^C = {K'}_{B,i}$ otherwise. Also, let 1_B^C be the vector of indicator random variables $1_{B,i}^C$, $j = 1, \ldots, N_1$.

Proceeding as in the proof of Lemma 3.3, we have

$$H\left(Y_2^n | W_1 Y_2^{n_1} F^n \Theta_C U_C\right) = H\left(U_B^C | W_1 Y_2^{n_1} F^n \Theta_C U_C\right) \tag{B.8}$$

$$= H\left({K'}_B^C | W_1 Y_2^{n_1} F^n \Theta_C U_C\right) \geq H\left({K'}_B^C | 1_B^C W_1 Y_2^{n_1} F^n \Theta_C U_C\right) \tag{B.9}$$

$$= H\left({K'}_B^C | 1_B^C\right) - I\left({K'}_B^C; W_1 Y_2^{n_1} F^n \Theta_C U_C | 1_B^C\right). \tag{B.10}$$

But, from the MDS property of $G_{K'_B}$, and the fact that K_B is uniformly distributed over its alphabet, we have

$$H\left({K'}_B^C | 1_B^C\right) = \sum_{i=1}^{N_1} \min\left(i, s'_B\right) \Pr\left\{\sum_{j=1} 1_{B,j}^C = i\right\} = \mathbb{E}\left\{\min\left(s'_B, \sum_{i=1}^{N_1} 1_{B,i}^C\right)\right\}. \tag{B.11}$$

Also,

$$I\left({K'}_B^C; W_1 Y_2^{n_1} F^n \Theta_C U_C | 1_B^C\right) \overset{\text{(a)}}{=} I\left({K'}_B^C; Y_2^{n_1} F^{n_1} | 1_B^C\right) \leq I\left({K'}_B^C 1_B^C; Y_2^{n_1} F^{n_1}\right) \leq I\left(K_B; Y_2^{n_1} F^{n_1} \Theta_C\right). \tag{B.12}$$

where (a) follows from the fact that the distribution of W_2 (uniform and independent of F^n, Θ_A, Θ_C) implies that U_C is independent of Θ_A, F^n and using this we can argue that the following is Markov chain

$${K'}_B^C - \left(1_B^C, Y_2^{n_1}, F^{n_1}\right) - \left(W_1, \Theta_C, U_C\right). \tag{B.13}$$

Substituting back we have the lemma. □

B.3 Rate calculation

B.3.1 Honest-but-curious adversary

The achievable rate for user j is $R_j = \lim_{n\to\infty} \frac{N_j}{n}$. Compared to the non-secure 1-to-M protocol we have an overhead of n_1 transmissions. We have

$$\lim_{n\to\infty} \frac{s_j}{n} = R_j \frac{1 - \frac{\prod_{k=1}^M \delta_k}{\delta_j}}{1 - \prod_{k=1}^M \delta_k}, \tag{B.14}$$

and thus

$$\lim_{n\to\infty}\frac{s_j+s_j^{3/4}}{n}=R_j\frac{1-\frac{\prod_{k=1}^{M}\delta_k}{\delta_j}}{1-\prod_{k=1}^{M}\delta_k}, \tag{B.15}$$

$$\lim_{n\to\infty}\frac{n_1}{n}=\max_{j\in\{1,\dots,M\}}\frac{R_j\left(1-\frac{\prod_{k=1}^{M}\delta_k}{\delta_j}\right)}{\left(1-\delta_j\right)\frac{\prod_{k=1}^{M}\delta_k}{\delta_j}\left(1-\prod_{k=1}^{M}\delta_k\right)}. \tag{B.16}$$

Using Theorem 3.1 the rate assertion of Theorem 3.3 follows.

B.3.2 Dishonest adversary

Similarly as in the honest-but-curious case, we need to compute $\lim_{n\to\infty}\frac{n_1}{n}$ and $\lim_{n\to\infty}\frac{\max\{n_2',n_2''\}}{n}$. It is immediate that

$$\lim_{n\to\infty}\frac{n_2'}{n}=\frac{R_1}{1-\delta_1}+\frac{R_2}{1-\delta_1\delta_2} \tag{B.17}$$

$$\lim_{n\to\infty}\frac{n_2''}{n}=\frac{R_1}{1-\delta_1\delta_2}+\frac{R_2}{1-\delta_2}. \tag{B.18}$$

Further,

$$\lim_{n\to\infty}\frac{s_B'}{n}=R_1\frac{1-\delta_2}{1-\delta_1\delta_2} \tag{B.19}$$

$$\lim_{n\to\infty}\frac{s_C'}{n}=R_2\frac{1-\delta_1}{1-\delta_1\delta_2}, \tag{B.20}$$

from which

$$\lim_{n\to\infty}\frac{s_B}{n}=R_1\frac{1-\delta_2}{\delta_2\left(1-\delta_1\delta_2\right)} \tag{B.21}$$

$$\lim_{n\to\infty}\frac{s_C}{n}=R_2\frac{1-\delta_1}{\delta_1\left(1-\delta_1\delta_2\right)}, \tag{B.22}$$

and

$$\lim_{n\to\infty}\frac{n_1}{n}=\max\left(R_1\frac{1-\delta_2}{\delta_2\left(1-\delta_1\right)\left(1-\delta_1\delta_2\right)},R_2\frac{1-\delta_1}{\delta_1\left(1-\delta_2\right)\left(1-\delta_1\delta_2\right)}\right). \tag{B.23}$$

We also observe that

$$R_1\frac{1-\delta_2}{\delta_2\left(1-\delta_1\right)\left(1-\delta_1\delta_2\right)}>R_2\frac{1-\delta_1}{\delta_1\left(1-\delta_2\right)\left(1-\delta_1\delta_2\right)}\iff\frac{R_1}{1-\delta_1}+\frac{R_2}{1-\delta_1\delta_2}>\frac{R_1}{1-\delta_1\delta_2}+\frac{R_2}{1-\delta_2}. \tag{B.24}$$

From these the rate assertion of Theorem 3.4 follows.

B.4 Proof of distribution independent security

We need to show that if Bob is honest, then (3.10) holds. In the proof we omit taking the maximum, but our argument is true for all joint distributions of (W_1, W_2), hence the property follows.

We can almost identically follow the proof of Section 3.5.2. Similarly to (3.45) we have

$$I\left(W_1; Y_2^n F^n \Theta_C | W_2\right) \leq I\left(W_1; Y_2^n F^n \Theta_C U_C | W_2\right) = I\left(W_1; Y_2^n | Y_2^{n_1+n_2} F^n \Theta_C U_C W_2\right). \tag{B.25}$$

The last step follows because given W_2, variables Θ_A, Θ_C, F^n are independent of W_1, further $Y_2^{n_1+n_2}, U^C$ are deterministic functions of $\Theta_A, \Theta_C, W_2, F^n$. The proofs of Lemmas 3.2 and 3.3 directly give us

$$I\left(K_B; Y_2^{n_1+n_2} F^{n_1+n_2} \Theta_C\right) \leq s'_B \mathrm{e}^{-a_{B.26}\sqrt{s'_B}}, \tag{B.26}$$

$$H\left(Y_2^n | Y_2^{n_1+n_2} F^n \Theta_C U_C W_2\right) \leq \mathbb{E}\left\{M_B^C\right\}, \tag{B.27}$$

under the same conditions as defined in Lemmas 3.2 and 3.3, where $a_{B.26} > 0$ is some constant. We still need to show that

$$H\left(Y_2^n | W_1 W_2 Y_2^{n_1+n_2} F^n \Theta_C U_C\right) \geq \mathbb{E}\left\{\min\left(s'_B, M_B^C\right)\right\} - I\left(K_B; Y_2^{n_1+n_2} F^{n_1+n_2}\right) \tag{B.28}$$

holds. We can again follow the proof of Lemma 3.4, but we have to argue step (a) in (B.12), i.e.,

$$I\left(K'^C_B; W_1 W_2 Y_2^{n_1+n_2} F^n \Theta_C U_C | 1_B^C\right) = I\left(K'^C_B; Y_2^{n_1+n_2} F^{n_1+n_2} | 1_B^C\right), \tag{B.29}$$

where the independent and uniformly distributed property of W_2 was exploited when proving the lemma. To see that equation (B.29) is true under the modified protocol, consider that K'_C is generated from a different set of random packets than K'^C_B, so $K'^C_B - Y_2^{n_1+n_2} - U_C$ is a Markov-chain, and since (Θ_A, F^n) is generated independently of (W_1, W_2, Θ_C), $K'^C_B - \left(Y_2^{n_1+n_2}, 1_B^C, F^{n_1+n_2}\right) - (W_1, W_2, \Theta_C, U_C)$ has the Markov property too.

Having established the three key lemmas for the modified protocol, we can conclude the proof the same way as we have seen in Section 3.5.2. We omit the details to avoid repetitive arguments. □

B.5 Proof of Lemma 3.1

As a first step we define

$$\mathbf{Adv}_{\mathrm{dis}}^{*\mathrm{ss}} = \max_{f, P_{W_1}, w_2, \sigma} \left\{ \max_{\mathscr{A}} \Pr\left\{\mathscr{A}\left(Y_2^n, F^n, \Theta_2, \sigma, w_2\right) = f\left(W_1, w_2\right)\right\} - \max_{\mathscr{S}} \Pr\left\{\mathscr{S}\left(P_{W_1}, f, w_2\right) = f\left(W_1, w_2\right)\right\} \right\}. \tag{B.30}$$

As opposed to $\mathbf{Adv}^{\mathrm{ss}}$ here w_2 is not a random variable but a constant value from $\mathcal{W}_2$. Clearly,

$$\mathbf{Adv}^{*\mathrm{ss}}_{\mathrm{dis}} \leq \mathbf{Adv}^{\mathrm{ss}}_{\mathrm{dis}}, \tag{B.31}$$

because W_2 taking the value w_2 with probability 1 is a particular joint distribution W_1, W_2 can take, so the scope of the maximization is restricted. We show that $\mathbf{Adv}^{*\mathrm{ss}}_{\mathrm{dis}} = \mathbf{Adv}^{\mathrm{ss}}_{\mathrm{dis}}$.

$$\begin{aligned}
\mathbf{Adv}^{\mathrm{ss}}_{\mathrm{dis}} = \max_{f, P_{W_1,W_2}, \sigma} \Big\{ & \max_{\mathscr{A}} \Pr\left\{\mathscr{A}\left(Y_2^n, F^n, \Theta_2, \sigma, W_2\right) = f\left(W_1, W_2\right)\right\} \\
& - \max_{\mathscr{S}} \Pr\left\{\mathscr{S}\left(P_{W_1}, f, W_2\right) = f\left(W_1, W_2\right)\right\} \Big\} && \text{(B.32)} \\
= \max_{f, P_{W_1,W_2}, \sigma} \sum_{w_2} p_{W_2}(w_2) \Big\{ & \max_{\mathscr{A}} \Pr\left\{\mathscr{A}\left(Y_2^n, F^n, \Theta_2, \sigma, W_2\right) = f\left(W_1, W_2\right) | W_2 = w_2\right\} \\
& - \max_{\mathscr{S}} \Pr\left\{\mathscr{S}\left(P_{W_1}, f, \sigma, W_2\right) = f\left(W_1, W_2\right) | W_2 = w_2\right\} \Big\} && \text{(B.33)} \\
= \max_{w_2^*, f, P_{W_1|W_2=w_2^*}, \sigma} \Big\{ & \max_{\mathscr{A}} \Pr\{\mathscr{A}\left(Y_2^n, F^n, \Theta_2, \sigma, w_2^*\right) = f\left(W_1, w_2^*\right)\} \\
& - \max_{\mathscr{S}} \Pr\left\{\mathscr{S}\left(P_{W_1}, f, w_2^*\right) = f\left(W_1, w_2^*\right)\right\} \Big\} && \text{(B.34)} \\
= \max_{f, P_{W_1,w_2^*}, \sigma} \Big\{ & \max_{\mathscr{A}} \Pr\left\{\mathscr{A}\left(Y_2^n, F^n, \Theta_2, \sigma, w_2^*\right) = f\left(W_1, w_2^*\right)\right\} \\
& - \max_{\mathscr{S}} \Pr\left\{\mathscr{S}\left(P_{W_1}, f, w_2^*\right) = f\left(W_1, w_2^*\right)\right\} \Big\} && \text{(B.35)} \\
= \mathbf{Adv}^{*\mathrm{ss}}_{\mathrm{dis}}, & && \text{(B.36)}
\end{aligned}$$

where the second step follows because there is a certain value w_2^* of W_2 that maximizes the expression inside $\{\ldots\}$, and moreover this expression depends on $P_{W_1|W_2}$ only through $P_{W_1|W_2=w_2}$, hence a maximizing joint distribution of W_1, W_2 is when W_2 takes this particular value with probability 1.

We continue the proof in two steps, first we define a notion of distinguishing security applicable for jointly distributed messages by extending a similar definition in [24] and show its equivalence with the above definition of semantic security. Then we show equivalence between this notion of distinguishing security and distribution independent security as defined by $\mathbf{Adv}^{\mathrm{mis}}_{\mathrm{dis}}$.

We define a notion corresponding to distinguishing security by defining the adversarial advantage:

$$\mathbf{Adv}^{\mathrm{ds}}_{\mathrm{dis}} = \max_{\mathscr{A}, w_1^0, w_1^1, w_2, \sigma} 2 \Pr\left\{\mathscr{A}\left(w_1^0, w_1^1, w_2, {}^b Y_2^n, F^n, \Theta_2, \sigma\right) = b\right\} - 1, \tag{B.37}$$

where $w_1^0, w_1^1 \in \mathcal{W}_1$ are possible messages, similarly $w_2 \in \mathcal{W}_2$, b is a variable uniformly distributed over $\{0,1\}$ and is independent of all other variables, while ${}^b Y_2^n$ is Calvin's observation given $W_1 = w_1^b$.

Distinguishing security defined by $\mathbf{Adv}^{\mathrm{ds}}_{\mathrm{dis}}$ is equivalent to semantic security as defined by $\mathbf{Adv}^{*\mathrm{ss}}_{\mathrm{dis}}$ and hence equivalently as defined by $\mathbf{Adv}^{\mathrm{ss}}_{\mathrm{dis}}$. To show that distinguishing security implies semantic security, we can almost identically follow the proof of Theorem 5 from [24], with a slight difference that a conditioning on W_2 appears. Given an adversary $\mathscr{A}_{ss}$ attacking semantic security, we construct an adversary $\mathscr{A}_{ds}$ attacking distinguishing security as follows: $\mathscr{A}_{ds}$ outputs 1, if the adversary attacking semantic security $\mathscr{A}_{ss}$ gives as output $f\left(w_1^1, w_2\right)$, otherwise it returns 0. Then, if W_1^0 and W_1^1 are i.i.d. both having the same distribution as W_1, then

$$\begin{aligned}\Pr\left\{\mathscr{A}_{ds}\left(W_1^0, W_1^1, W_2, {}^1Y_2^n, F^n, \Theta_2, \sigma\right) = 1 | W_2 = w_2\right\} \\ = \Pr\left\{A_{ss}\left(Y_2^n, F^n, \Theta_2, \sigma, W_2\right) = f\left(W_1, W_2\right) | W_2 = w_2\right\}\end{aligned} \tag{B.38}$$

$$\begin{aligned}\Pr\left\{\mathscr{A}_{ds}\left(W_1^0, W_1^1, W_2, {}^0Y_2^n, F^n, \Theta_2, \sigma\right) = 1 | W_2 = w_2\right\} \\ \leq \max_{\mathscr{S}} \Pr\left\{\mathscr{S}\left(P_{W_1}, f, W_2\right) = f\left(W_1, W_2\right) | W_2 = w_2\right\}.\end{aligned} \tag{B.39}$$

Finishing the derivation as in [24] we get

$$\begin{aligned}\Pr\left\{A_{ss}\left(Y_2^n, F^n, \Theta_2, \sigma, W_2\right) = f\left(W_1, W_2\right) | W_2 = w_2\right\} - \max_{\mathscr{S}} \Pr\left\{\mathscr{S}\left(P_{W_1}, f, W_2\right) = f\left(W_1, W_2\right) | W_2 = w_2\right\} \\ \leq \max_{w_1^0, w_1^1, w_2, \mathscr{A}_{ds}, \sigma} 2\Pr\left\{\mathscr{A}_{ds}\left(w_1^0, w_1^1, w_2, {}^bY_2^n, F^n, \Theta_2, \sigma\right) = b\right\} - 1\end{aligned} \tag{B.40}$$

for all $P_{W_1}, f, \mathscr{A}_{ss}, \sigma$, hence taking the maximum over these variables on the LHS and over w_2 on both sides keeps the inequality. This establishes that

$$\mathbf{Adv}^{\mathrm{ss}}_{\mathrm{dis}} = \mathbf{Adv}^{*\mathrm{ss}}_{\mathrm{dis}} \leq \mathbf{Adv}^{\mathrm{ds}}_{\mathrm{dis}} \leq 2\mathbf{Adv}^{\mathrm{ss}}_{\mathrm{dis}}. \tag{B.41}$$

The other direction of implication is a straightforward consequence of the definitions, the scope of maximization in $\mathbf{Adv}^{\mathrm{ds}}_{\mathrm{dis}}$ is a subset of that of $\mathbf{Adv}^{\mathrm{ss}}_{\mathrm{dis}}$, in case of $\mathbf{Adv}^{\mathrm{ds}}_{\mathrm{dis}}$ f is a function that computes b, while P_{W_1,W_2} is such that W_1 uniformly takes the two values w_1^0 and w_1^1 and independently W_2 takes w_2 with probability 1.

What remains to show is that distinguishing security defined by $\mathbf{Adv}^{\mathrm{ds}}_{\mathrm{dis}}$ is equivalent to distribution independent security as defined by $\mathbf{Adv}^{\mathrm{mis}}_{\mathrm{dis}}$. Clearly, for any particular value of w_2,

$$\mathbf{Adv}^{\mathrm{mis}}_{\mathrm{dis}} \geq \max_{P_{W_1}, \sigma} I\left(W_1; Y_2^n F^n \Theta_2 | W_2 = w_2\right). \tag{B.42}$$

If we fix w_2 for the scheme, we can directly invoke Theorem 5 from [24] which proves that

$$\max_{\mathscr{A}, w_1^0, w_1^1, \sigma} 2\Pr\left\{\mathscr{A}\left(w_1^0, w_1^1, w_2, {}^bY_2^n, F^n, \Theta_2, \sigma\right) = b\right\} - 1 \leq \sqrt{2 \max_{P_{W_1}, \sigma} I\left(W_1; Y_2^n F^n \Theta_2 | W_2 = w_2\right)} \tag{B.43}$$

$$\leq \sqrt{2 \cdot \mathbf{Adv}^{\mathrm{mis}}_{\mathrm{dis}}}, \tag{B.44}$$

which holds for every w_2, so we can take the maximum in w_2 on the LHS, which gives in turn

$$\mathbf{Adv}^{\mathrm{ds}}_{\mathrm{dis}} \leq \sqrt{2 \cdot \mathbf{Adv}^{\mathrm{mis}}_{\mathrm{dis}}} \tag{B.45}$$

showing that the distribution independent security implies distinguishing security. The other direction is also true. We can apply the same type of argument as when showing $\mathbf{Adv}^{\mathrm{ss}}_{\mathrm{dis}} = \mathbf{Adv}^{*\mathrm{ss}}_{\mathrm{dis}}$ to get:

$$\mathbf{Adv}^{\mathrm{mis}}_{\mathrm{dis}} = \max_{P_{W_1,W_2},\sigma} I\left(W_1; Y_2^n F^n \Theta_2 | W_2\right) = \max_{w_2, P_{W_1},\sigma} I\left(W_1; Y_2^n F^n \Theta_2 | W_2 = w_2\right). \tag{B.46}$$

Let us denote

$$\mathbf{Adv}^{\mathrm{ds}}(w_2) = \max_{\mathscr{A}, w_1^0, w_1^1, \sigma} 2\Pr\left\{\mathscr{A}\left(w_1^0, w_1^1, w_2, {}^b Y_2^n, F^n, \Theta_2, \sigma\right) = b\right\} - 1. \tag{B.47}$$

We can apply Theorem 4.9 from [24] with a conditioning on $W_2 = w_2$, which implies that for any w_2:

$$\max_{P_{W_1},\sigma} I\left(W_1; Y_2^n F^n \Theta_2 | W_2 = w_2\right) \leq 2\mathbf{Adv}^{\mathrm{ds}}(w_2) \log\left(\frac{2^n}{\mathbf{Adv}^{\mathrm{ds}}(w_2)}\right). \tag{B.48}$$

Since the above is true for any w_2, we can take the maximum in w_2 on both sides resulting

$$\mathbf{Adv}^{\mathrm{mis}}_{\mathrm{dis}} \leq 2\mathbf{Adv}^{\mathrm{ds}} \log\left(\frac{2^n}{\mathbf{Adv}^{\mathrm{ds}}}\right). \tag{B.49}$$

This completes the proof that distribution independent security is equivalent to semantic security defined by $\mathbf{Adv}^{\mathrm{ss}}_{\mathrm{dis}}$. □

C Proofs and calculations for Chapter 4

C.1 Calculating N_i

Let N_i', N_i'' and N_i be defined by the following equalities:

$$s_i' + \sum_{j=1, j\neq i}^{\ell} s_j = N_i' \frac{1-\delta_{iE}}{1-\delta_i\delta_{iE}} + N_i'^{\frac{3}{4}} \frac{1-\delta_{iE}}{1-\delta_i\delta_{iE}} \tag{C.1}$$

$$nm_i = \frac{N_i''}{1-\delta_i} + \frac{N_i''^{\frac{3}{4}}}{1-\delta} \tag{C.2}$$

$$N_i = \min\left\{N_i', N_i''\right\}. \tag{C.3}$$

From the paramter defintions it is clear that

$$\lim_{n\to\infty} \frac{N_i'}{n} = \frac{1-\delta_i\delta_{iE}}{1-\delta_{iE}} \left(c_i\delta_{iE}(1-\delta_i) + \sum_{j=1, j\neq i}^{\ell} c_j\left(1-\delta_j\right) \right) \tag{C.4}$$

$$\lim_{n\to\infty} \frac{N_i''}{n} = m_i(1-\delta_i). \tag{C.5}$$

From (4.5) we know that $\lim_{n\to\infty} \frac{N_i''}{n} \leq \lim_{n\to\infty} \frac{N_i'}{n}$, hence $\lim_{n\to\infty} \frac{N_i}{n} = m_i(1-\delta_i)$.

C.2 Proofs of Lemmas 4.5-4.6

C.2.1 Proof of Lemma 4.5

Define

$$\frac{nr_1}{1-\delta_1} = r_1' + r_1'^{\frac{3}{4}} \tag{C.6}$$

$$\frac{nr_2}{1-\delta_2} = r_2' + r_2'^{\frac{3}{4}} \tag{C.7}$$

$$\delta_{1E}' = \frac{1-\delta_{1E}}{1-\delta_1\delta_{1E}} \tag{C.8}$$

$$\delta'_{2E} = \frac{1-\delta_{1E}}{1-\delta_1\delta_{1E}} \tag{C.9}$$

$$r'_1 = \frac{r''_1}{\delta'_{1E}} + \frac{r'^{\frac{3}{4}}_1}{\delta'_{1E}} \tag{C.10}$$

$$r'_2 = \frac{r'_2}{\delta'_{2E}} + \frac{r'^{\frac{3}{4}}_2}{\delta'_{2E}}. \tag{C.11}$$

S_1 and S_2 continue ARQ for at most $\frac{nr_1}{1-\delta_1}$ and $\frac{nr_2}{1-\delta_2}$ transmissions. Let $X_{1,D}$ denote the first r'_1 packets that D receives on channel 1 (if no such, an error is declared). Then, for S_1 a key is computed as follows:

$$K_{1,2a} = X_{1,D}H_{1,2a}, \tag{C.12}$$

where $H_{1,2a}$ is a $r'_1 \times r''_1$ parity check matrix of an MDS code. Another key $K_{1,2b}$ is formed for S_1 from the first r'_2 packets that D receives on channel 2. The same method is applied to create keys $K_{2,2a}$, $K_{2,2b}$ for S_2.

The security of keys $K_{1,2b}$ and $K_{2,2b}$ is obvious. The generation of $K_{1,2a}$, $K_{2,2a}$ is the same as the key generation scheme in Section 1.5, with the only difference that δ'_{1E} and δ'_{2E} play the role of δ_E. The same analysis of security and error probability applies. The claimed key rates follow directly from the parameter definitions after taking the limits with $n \to \infty$. □

C.2.2 Proof of Lemma 4.6

We provide the proof for the key of S_1, the same derivation applies for S_2. We define the following parameters:

$$nc_1 = nc'_1 + \left(nc'_1\right)^{3/4} \tag{C.13}$$

$$nc_2 = nc'_2 + \left(nc'_2\right)^{3/4} \tag{C.14}$$

$$nc'_1 = \frac{nc''_1}{1-\delta_1\delta_{1E}} + \left(\frac{nc''_1}{1-\delta_1\delta_{1E}}\right)^{3/4} \tag{C.15}$$

$$nc'_2 = \frac{nc''_2}{1-\delta_2} + \left(\frac{nc''_2}{1-\delta_2}\right)^{3/4} \tag{C.16}$$

$$\delta'_{1E} = \frac{\delta_{1E}(1-\delta_1)}{1-\delta_1\delta_{1E}} \tag{C.17}$$

$$nc''_1 = \frac{nc'''_1}{\delta'_{1E}} + \frac{1}{\delta'_{1E}}\left(\frac{2nc'''_1}{\delta'_{1E}}\right)^{3/4}. \tag{C.18}$$

X^D_{C1} denotes the $nc''_1\frac{1-\delta_1}{1-\delta_1\delta_{1E}}$ vector that consists of the first $nc''_1\frac{1-\delta_1}{1-\delta_1\delta_{1E}}$ packets that D receives in this step on the $S_1 - D$ channel, while X^D_{C2} the first nc''_2 packets received from S_2 in the corresponding step. Again, we assume that sufficient number of packets are correctly received to form these matrices, otherwise an error occurs. We later show that the error probability is negligible.

Like before, we use the notation X_{C1}^{A} and $X_{C1}^{A\emptyset}$ for the packets received by $A \in \{D, E, DE\}$ and for those received by S only, where D stands for the destination and E for the eavesdropper on channel 1. With a slight abuse, we denote corresponding index sets by $I_A, I_{A\emptyset}$. Note that the eavesdropper on channel 1 does not receive any packets from C_2.

The key $K_{1,3}$ is computed as follows:

$$K_{1,3} = \left[X_{C2}^{D} \;\; X_{C1}^{D} G_{1,3} \right], \tag{C.19}$$

where $G_{1,3}$ is a matrix of dimension $nc_1'' \frac{1-\delta_1}{1-\delta_1\delta_{1E}} \times nc_1'''$ and it is a parity check matrix of an MDS code.

We show that

$$I\left(K_{1,3}; X_{C1}^{E}\right) \tag{C.20}$$

can be made arbitrarily small by choosing a large enough n.

From the properties of C_1 and C_2 (i.e., the MDS property of G) it follows that X_{C2}^{D} and X_{C1}^{D} are independent as long as

$$c_2'' + c_1'' \frac{1-\delta_1}{1-\delta_1\delta_{1E}} < c, \tag{C.21}$$

which directly follows from (4.59). Thus, we have that

$$H\left(K_{1,3}\right) = H\left(\left[X_{C2}^{D} X_{C1}^{D} G_{1,3}\right]\right) = H\left(X_{C2}^{D}\right) + H\left(X_{C1}^{D} G_{1,3}\right) = n\left(c_2'' + c_1'''\right). \tag{C.22}$$

This already shows that $K_{1,3}$ has uniform distribution and that it achieves the claimed rate. We need to show that $K_{1,3}$ is secret from the eavesdropper on the $S_1 - D$ channel.

$$H\left(K_{1,3}|X_{C1}^{E}S^{n}\right) = H\left(\left[X_{C2}^{D} \;\; X_{C1}^{D}G_{1,3}\right]|X_{C1}^{E}S^{n}\right) \tag{C.23}$$
$$= H\left(X_{C2}^{D}|X_{C1}^{E}S^{n}\right) + H\left(X_{C1}^{D}G_{1,3}|X_{C1}^{E}X_{C2}^{D}S^{n}\right) \tag{C.24}$$
$$= nc_2' + H\left(X_{C1}^{D}G_{1,3}|X_{C1}^{E}X_{C2}^{D}S^{n}\right) \tag{C.25}$$
$$= nc_2' + H\left(X_{C1}^{D\emptyset}G_{1,3}^{I_{D\emptyset}}|X_{C1}^{E}X_{C2}^{D}S^{n}\right) \tag{C.26}$$

Further,

$$H\left(X_{C1}^{D\emptyset}G_{1,3}^{I_{D\emptyset}}|X_{C1}^{E}X_{C2}^{D}S^{n}\right)$$
$$= \sum_{k}\sum_{j} \Pr\left\{|X_{C1}^{E}| + |X_{C1}^{D\emptyset}| = k, |X_{C1}^{D\emptyset}| = j\right\} H\left(X_{C1}^{D\emptyset}G_{1,3}^{I_{D\emptyset}}|X_{C1}^{E}X_{C2}^{D}S^{n}, |X_{C1}^{E}| + |X_{C1}^{D\emptyset}| = k, |X_{C1}^{D\emptyset}| = j\right) \tag{C.27}$$

$$\geq \sum_{k=0}^{n(c-c_2'')} \sum_{j=nc_1'''}^{k} \Pr\left\{|X_{C1}^{E}| + |X_{C1}^{D\emptyset}| = k, |X_{C1}^{D\emptyset}| = j\right\} H\left(X_{C1}^{D\emptyset}G_{1,3}^{I_{D\emptyset}}|X_{C1}^{E}X_{C2}^{D}S^{n}, |X_{C1}^{E}| + |X_{C1}^{D\emptyset}| = k, |X_{C1}^{D\emptyset}| = j\right) \tag{C.28}$$

$$= nc_1''' \sum_{k=0}^{n(c-c_2'')} \sum_{j=nc_1'''}^{k} \Pr\left\{|X_{C1}^E| + |X_{C1}^{D\phi}| = k, |X_{C1}^{D\phi}| = j\right\} \tag{C.29}$$

$$= nc_1''' \sum_{k=0}^{n(c-c_2'')} \Pr\left\{|X_{C1}^E| + |X_{C1}^{D\phi}| = k, |X_{C1}^{D\phi}| \geq nc_1'''\right\}. \tag{C.30}$$

We exploited that $|X_{C1}^E| + |X_{C1}^{D\phi}| \leq n\left(c - c_2''\right)$ implies that $|X_{C1}^E| + |X_{C1}^{D\phi}| + |X_{C2}^D| < c$, thus the columns of $X_{C1}^{D\phi} G_{1,3}^{I_{D\phi}}$ are independent of $\left(X_{C1}^E, X_{C2}^D\right)$.

$$\sum_{k=0}^{n(c-c_2'')} \Pr\left\{|X_{C1}^E| + |X_{C1}^{D\phi}| = k, |X_{C1}^{D\phi}| \geq nc_1'''\right\} \tag{C.31}$$

$$= \sum_{k=0}^{n(c-c_2'')} \Pr\left\{|X_{C1}^E| + |X_{C1}^{D\phi}| = k\right\} \Pr\left\{|X_{C1}^{D\phi}| \geq nc_1''' || X_{C1}^E| + |X_{C1}^{D\phi}| = k\right\} \tag{C.32}$$

$$\geq \sum_{k=nc_1''}^{n(c-c_2'')} \Pr\left\{|X_{C1}^E| + |X_{C1}^{D\phi}| = k\right\} \Pr\left\{|X_{C1}^{D\phi}| \geq nc_1''' || X_{C1}^E| + |X_{C1}^{D\phi}| = k\right\} \tag{C.33}$$

$$\geq \sum_{k=nc_1''}^{n(c-c_2'')} \Pr\left\{|X_{C1}^E| + |X_{C1}^{D\phi}| = k\right\} \Pr\left\{|X_{C1}^{D\phi}| \geq nc_1''' || X_{C1}^E| + |X_{C1}^{D\phi}| = nc_1''\right\} \tag{C.34}$$

$$= \Pr\left\{nc_1'' \leq |X_{C1}^E| + |X_{C1}^{D\phi}| \leq n\left(c - c_2''\right)\right\} \Pr\left\{|X_{C1}^{D\phi}| \geq nc_1''' || X_{C1}^E| + |X_{C1}^{D\phi}| = nc_1''\right\} \tag{C.35}$$

We show that both these latter two probability terms are close to 1. We use the Chernoff-Hoeffding bound and that $\mathbb{E}\left\{|X_{C1}^E| + |X_{C1}^{D\phi}|\right\} = nc_1'\left(1 - \delta_1\delta_{1E}\right)$ to get:

$$\Pr\left\{|X_{C1}^E| + |X_{C1}^{D\phi}| \leq n\left(c - c_2''\right)\right\} \tag{C.36}$$

$$= 1 - \Pr\left\{|X_{C1}^E| + |X_{C1}^{D\phi}| > n\left(c - c_2''\right)\right\} \tag{C.37}$$

$$\geq 1 - \Pr\left\{|X_{C1}^E| + |X_{C1}^{D\phi}| > n\left(c - c_2\left(1 - \delta_2\right)\right)\right\} \tag{C.38}$$

$$\geq 1 - \Pr\left\{||X_{C1}^E| + |X_{C1}^{D\phi}| - \mathbb{E}\left\{|X_{C1}^E| + |X_{C1}^{D\phi}|\right\}| > n\left(c - c_2\left(1 - \delta_2\right) - c_1'\left(1 - \delta_1\delta_{1E}\right)\right)\right\} \tag{C.39}$$

$$\overset{(a)}{\geq} 1 - \Pr\left\{||X_{C1}^E| + |X_{C1}^{D\phi}| - \mathbb{E}\left\{|X_{C1}^E| + |X_{C1}^{D\phi}|\right\}| > \left(1 - \delta_1\delta_{1E}\right)\left(nc_1'\right)^{3/4}\right\} \tag{C.40}$$

$$\geq 1 - \mathrm{e}^{-a_{C.41}\sqrt{nc_1'}}, \tag{C.41}$$

for some constant $a_{C.41} > 0$. In step (a) we used (4.59). Also,

$$\Pr\left\{|X_{C1}^E| + |X_{C1}^{D\phi}| < nc_1''\right\} \tag{C.42}$$

$$\leq \Pr\left\{||X_{C1}^E| + |X_{C1}^{D\phi}| - \mathbb{E}\left\{|X_{C1}^E| + |X_{C1}^{D\phi}|\right\}| > \left(1 - \delta_1\delta_{1E}\right)\left(\frac{nc_1''}{1 - \delta_1\delta_{1E}}\right)^{3/4}\right\} \tag{C.43}$$

$$\leq \mathrm{e}^{-a_{C.44}\sqrt{nc_1''}}, \tag{C.44}$$

for some constant $a_{C.44} > 0$. Thus,

$$\Pr\left\{nc_1'' \le |X_{C1}^E| + |X_{C1}^{D\phi}| \le n\left(c - c_2''\right)\right\} \ge 1 - \mathrm{e}^{-a_{C.41}\sqrt{nc_1'}} - \mathrm{e}^{-a_{C.44}\sqrt{nc_1''}}. \tag{C.45}$$

For the second term:

$$\Pr\left\{|X_{C1}^{D\phi}| \ge nc_1'''||X_{C1}^E| + |X_{C1}^{D\phi}| = nc_1''\right\} = \tag{C.46}$$

$$\Pr\left\{|X_{C1}^E| \le n\left(c_1'' - c_1'''\right)||X_{C1}^E| + |X_{C1}^{D\phi}| = nc_1''\right\} \tag{C.47}$$

$$= 1 - \Pr\left\{|X_{C1}^E| > n\left(c_1'' - c_1'''\right)||X_{C1}^E| + |X_{C1}^{D\phi}| = nc_1''\right\}. \tag{C.48}$$

The probability that a given packet in $\left(X_{C1}^E, X_{C1}^{D\phi}\right)$ is in X_{C1}^E, or in other words the probability that a packet that we know that either D or the eavesdropper received is known to the eavesdropper, equals $\frac{1-\delta_{1E}}{1-\delta_1\delta_{1E}}$, hence

$$\mathbb{E}\left\{|X_{C1}^E|||X_{C1}^E| + |X_{C1}^{D\phi}| = nc_1''\right\} = nc_1''\frac{1-\delta_{1E}}{1-\delta_1\delta_{1E}} = nc_1''\left(1-\delta_{1E}'\right). \tag{C.49}$$

Using this, we can again apply the Chernoff-Hoeffding bound to get:

$$\Pr\left\{|X_{C1}^E| > n\left(c_1'' - c_1'''\right)||X_{C1}^E| + |X_{C1}^{D\phi}| = nc_1''\right\} \le \mathrm{e}^{-a_{C.50}\sqrt{nc_1''}}, \tag{C.50}$$

for some constant $a_{C.50} > 0$. With this we have shown that

$$I\left(K_{1,1}; X_{K1}^E S^n\right) \le \epsilon \tag{C.51}$$

is satisfied if n is sufficiently large.

The error probability of this step can be shown to be arbitrarily small using the techniques we have seen before. We omit details to avoid repetitions. □

C.3 V-network outer bound proof

We summarize our outer bound linear program below. With a slight abuse of formality, we do not introduce a name on all the different entropy and mutual information terms, but from this point we do not use any properties of the information terms except that they are non-negative. We replace some terms with a named variable to help readability or when there is a corresponding variable in the scheme program. As before, we work with the asymptotic form of the inequalities.

Since we use terms as names we can ease notation by omitting $\frac{1}{n}\sum_{i=1}^n$, F^{i-1} (which appears in all terms), and index i. This should also help the reader not to think of these terms as meaningful information terms in the sequel.

$$R = (1-\delta_1)\, I\left(X_1; W|Y\right) + (1-\delta_2)\, I\left(X_2; W|Y\right)$$

$$R = (1-\delta_1\delta_{1E})\, I(X_1; W|YZ_1) + (1-\delta_2)\, I(X_2; W|YZ_1)$$
$$R = (1-\delta_2\delta_{2E})\, I(X_2; W|YZ_2) + (1-\delta_1)\, I(X_1; W|YZ_2)$$
$$R = (1-\delta_1\delta_{1E})\, I(X_1; W|YZ) + (1-\delta_2\delta_{2E})\, I(X_2; W|YZ)$$
$$C_r \geq (1-\delta_1\delta_{1E})\, I(X_1; \Psi|YZW) + (1-\delta_2\delta_{2E})\, I(X_2; \Psi|YZW)$$
$$0 \leq -(1-\delta_2\delta_{2E})\, I(X_2; Z_1|YZ_2W) - (1-\delta_1)\, I(X_1; Z_1|YZ_2W)$$
$$+\delta_1(1-\delta_{1E})\, I(X_1; \Psi|YZW) + (1-\delta_1)\, I(X_1; Z_1|YZ_2W\Psi)$$
$$0 \leq -(1-\delta_1\delta_{1E})\, I(X_1; Z_2|YZ_1W) - (1-\delta_2)\, I(X_2; Z_2|YZ_1W)$$
$$+\delta_2(1-\delta_{2E})\, I(X_2; \Psi|YZW) + (1-\delta_2)\, I(X_2; Z_2|YZ_1W\Psi)$$
$$0 \leq -(1-\delta_{1E})\, I(X_1; W|Z_1W) + \delta_{1E}(1-\delta_1)\, H(X_1|YZ_1W) + (1-\delta_1\delta_{1E})\, I(X_1; Y_2|Y_1Z_1W)$$
$$0 \leq -(1-\delta_{2E})\, I(X_2; W|Z_2W) + \delta_{2E}(1-\delta_2)\, H(X_2|YZ_2W) + (1-\delta_2\delta_{2E})\, I(X_2; Y_1|Y_2Z_2W)$$
$$(1-\delta_1\delta_{1E})\, I(X_1; Y_2|Y_1Z_1W) \leq (1-\delta_2)\, I(X_2; \Psi|YZW)$$
$$+(1-\delta_2)\, I(X_2; Z_2|YZ_1W) - (1-\delta_2)\, I(X_2; Z_2|YZ_1W\Psi)$$
$$(1-\delta_2\delta_{2E})\, I(X_2; Y_1|Y_2Z_2W) \leq (1-\delta_1)\, I(X_1; \Psi|YZW)$$
$$+(1-\delta_1)\, I(X_1; Z_1|YZ_2W) - (1-\delta_1)\, I(X_1; Z_1|YZ_2W\Psi)$$
$$I(X_1; W|YZ) \leq I(X_1; W|YZ_1) + I(X_1; Z_2|YZ_1W)$$
$$I(X_2; W|YZ) \leq I(X_2; W|YZ_2) + I(X_2; Z_1|YZ_2W)$$
$$1 \geq H(X_1|YZ_2W) + I(X_1; W|YZ_2)$$
$$1 \geq H(X_2|YZ_1W) + I(X_2; W|YZ_1)$$
$$1 \geq H(X_1|YZ_1W) + I(X_1; W|Y)$$
$$1 \geq H(X_2|YZ_2W) + I(X_2; W|Y)$$
$$1 \geq H(X_1|YZ_2W) + I(X_1; W|Y)$$
$$1 \geq H(X_2|YZ_1W) + I(X_2; W|Y)$$
$$H(X_1|YZW) = H(X_1|YZ_1W) - I(X_1; Z_2|YZ_1W) \tag{C.52}$$
$$H(X_1|YZW) = H(X_1|YZ_2W) - I(X_1; Z_1|YZ_2W) \tag{C.53}$$
$$H(X_2|YZW) = H(X_2|YZ_1W) - I(X_2; Z_2|YZ_1W) \tag{C.54}$$
$$H(X_2|YZW) = H(X_2|YZ_2W) - I(X_2; Z_1|YZ_2W) \tag{C.55}$$
$$H(X_1|YZW) \geq I(X_1; \Psi|YZW) \tag{C.56}$$
$$H(X_2|YZW) \geq I(X_2; \Psi|YZW) \tag{C.57}$$
$$I(X_1; W|YZ) \leq I(X_1; W|YZ_1) + I(X_1; Z_2|YZ_1W)$$
$$I(X_2; W|YZ) \leq I(X_2; W|YZ_2) + I(X_2; Z_1|YZ_2W)$$

We introduce the following naming. Some variables already match variables in the scheme program, in which case we use the name of the corresponding variable. We distinguish the variable names used only in the outer bound program with overscore. Recall that in the following the unnamed terms are also treated as non-negative variables.

$$I(X_1; W|Y) \sim m_1 \tag{C.58}$$

$$I(X_2;W|Y) \sim m_2 \tag{C.59}$$
$$H(X_1|YZW) \sim I(X_1;\Psi|YZW) + k_1 \tag{C.60}$$
$$H(X_2|YZW) \sim I(X_2;\Psi|YZW) + k_2 \tag{C.61}$$
$$I(X_1;\Psi|YZW) \sim \bar{s}_1 \tag{C.62}$$
$$I(X_1;\Psi|YZW) \sim \bar{s}_2 \tag{C.63}$$
$$I(X_1;Z_2|YZ_1W) \sim \bar{\ell}_1 \tag{C.64}$$
$$I(X_2;Z_1|YZ_2W) \sim \bar{\ell}_2 \tag{C.65}$$
$$I(X_1;Z_1|YZ_2W) \sim \bar{x}_1 \tag{C.66}$$
$$I(X_2;Z_2|YZ_1W) \sim \bar{x}_2 \tag{C.67}$$

Besides these, we also apply equalities (C.52)-(C.55) to eliminate the following terms:

$$H(X_1|YZ_1W) \qquad H(X_1|YZ_2W)$$
$$H(X_2|YZ_1W) \qquad H(X_2|YZ_2W)$$

After these steps we can drop (C.56)-(C.57), and we have the following:

$$R = (1-\delta_1)\, m_1 + (1-\delta_2)\, m_2 \tag{C.68}$$
$$R = (1-\delta_1\delta_{1E})\, I(X_1;W|YZ_1) + (1-\delta_2)\, I(X_2;W|YZ_1) \tag{C.69}$$
$$R = (1-\delta_2\delta_{2E})\, I(X_2;W|YZ_2) + (1-\delta_1)\, I(X_1;W|YZ_2) \tag{C.70}$$
$$R = (1-\delta_1\delta_{1E})\, I(X_1;W|YZ) + (1-\delta_2\delta_{2E})\, I(X_2;W|YZ) \tag{C.71}$$
$$C_r \geq (1-\delta_1\delta_{1E})\, \bar{s}_1 + (1-\delta_1\delta_{1E})\, \bar{s}_2 \tag{C.72}$$
$$0 \leq -(1-\delta_2\delta_{2E})\, \bar{\ell}_2 - (1-\delta_1)\, \bar{x}_1 + \delta_1(1-\delta_{1E})\, \bar{s}_1 + (1-\delta_1)\, I(X_1;Z_1|YZ_2W\Psi) \tag{C.73}$$
$$0 \leq -(1-\delta_1\delta_{1E})\, \bar{\ell}_1 - (1-\delta_2)\, \bar{x}_2 + \delta_2(1-\delta_{2E})\, \bar{s}_2 + (1-\delta_2)\, I(X_2;Z_2|YZ_1W\Psi) \tag{C.74}$$
$$0 \leq \delta_{1E}(1-\delta_1)\left(k_1 + \bar{s}_1 + \bar{\ell}_1\right) - (1-\delta_{1E})\, I(X_1;W|Z_1W) + (1-\delta_1\delta_{1E})\, I(X_1;Y_2|Y_1Z_1W) \tag{C.75}$$
$$0 \leq \delta_{2E}(1-\delta_2)\left(k_2 + \bar{s}_2 + \bar{\ell}_2\right) - (1-\delta_{2E})\, I(X_2;W|Z_2W) + (1-\delta_2\delta_{2E})\, I(X_2;Y_1|Y_2Z_2W) \tag{C.76}$$
$$(1-\delta_1\delta_{1E})\, I(X_1;Y_2|Y_1Z_1W) \leq (1-\delta_2)\, \bar{s}_2 + (1-\delta_2)\, \bar{x}_2 - (1-\delta_2)\, I(X_2;Z_2|YZ_1W\Psi) \tag{C.77}$$
$$(1-\delta_2\delta_{2E})\, I(X_2;Y_1|Y_2Z_2W) \leq (1-\delta_1)\, \bar{s}_1 + (1-\delta_1)\, \bar{x}_1 - (1-\delta_1)\, I(X_1;Z_1|YZ_2W\Psi) \tag{C.78}$$
$$I(X_1;W|YZ) \leq I(X_1;W|YZ_1) + \bar{\ell}_1 \tag{C.79}$$
$$I(X_2;W|YZ) \leq I(X_2;W|YZ_2) + \bar{\ell}_2 \tag{C.80}$$
$$1 \geq k_1 + \bar{s}_1 + \bar{x}_1 + I(X_1;W|YZ_2) \tag{C.81}$$
$$1 \geq k_2 + \bar{s}_2 + \bar{x}_2 + I(X_2;W|YZ_1) \tag{C.82}$$
$$1 \geq k_1 + \bar{s}_1 + \bar{\ell}_1 + m_1 \tag{C.83}$$
$$1 \geq k_2 + \bar{s}_2 + \bar{\ell}_2 + m_2 \tag{C.84}$$
$$1 \geq k_1 + \bar{s}_1 + m_1 + \bar{x}_1 \tag{C.85}$$
$$1 \geq k_2 + \bar{s}_2 + m_2 + \bar{x}_2 \tag{C.86}$$

We eliminate $I(X_2;Z_2|YZ_1W\Psi)$ and $I(X_1;Z_1|YZ_2W\Psi)$ using Fourier-Motzkin elimination.

$$R = (1-\delta_1)\, m_1 + (1-\delta_2)\, m_2$$
$$R = (1-\delta_1\delta_{1E})\, I(X_1;W|YZ_1) + (1-\delta_2)\, I(X_2;W|YZ_1)$$
$$R = (1-\delta_2\delta_{2E})\, I(X_2;W|YZ_2) + (1-\delta_1)\, I(X_1;W|YZ_2)$$
$$R = (1-\delta_1\delta_{1E})\, I(X_1;W|YZ) + (1-\delta_2\delta_{2E})\, I(X_2;W|YZ)$$
$$0 \le \delta_{1E}(1-\delta_1)\left(k_1+\bar{s}_1+\bar{\ell}_1\right) - (1-\delta_{1E})\, I(X_1;W|YZ_1) + (1-\delta_1\delta_{1E})\, I(X_1;Y_2|Y_1Z_1W) \quad \text{(C.87)}$$
$$0 \le \delta_{2E}(1-\delta_2)\left(k_2+\bar{s}_2+\bar{\ell}_2\right) - (1-\delta_{2E})\, I(X_2;W|YZ_2) + (1-\delta_2\delta_{2E})\, I(X_2;Y_1|Y_2Z_2W) \quad \text{(C.88)}$$
$$(1-\delta_1\delta_{1E})\, I(X_1;Y_2|Y_1Z_1W) \le (1-\delta_2\delta_{2E})\,\bar{s}_2 - (1-\delta_1\delta_{1E})\,\bar{\ell}_1 \quad \text{(C.89)}$$
$$(1-\delta_1\delta_{1E})\, I(X_1;Y_2|Y_1Z_1W) \le (1-\delta_2)\,\bar{s}_2 + (1-\delta_2)\,\bar{x}_2 \quad \text{(C.90)}$$
$$(1-\delta_2\delta_{2E})\, I(X_2;Y_1|Y_2Z_2W) \le (1-\delta_1\delta_{1E})\,\bar{s}_1 - (1-\delta_2\delta_{2E})\,\bar{\ell}_2 \quad \text{(C.91)}$$
$$(1-\delta_2\delta_{2E})\, I(X_2;Y_1|Y_2Z_2W) \le (1-\delta_1)\,\bar{s}_1 + (1-\delta_1)\,\bar{x}_1 \quad \text{(C.92)}$$
$$I(X_1;W|YZ) \le I(X_1;W|YZ_1) + \bar{\ell}_1$$
$$I(X_2;W|YZ) \le I(X_2;W|YZ_2) + \bar{\ell}_2$$
$$1 \ge k_1 + \bar{s}_1 + \bar{x}_1 + I(X_1;W|YZ_2)$$
$$1 \ge k_2 + \bar{s}_2 + \bar{x}_2 + I(X_2;W|YZ_1)$$
$$1 \ge k_1 + \bar{s}_1 + \bar{\ell}_1 + m_1$$
$$1 \ge k_2 + \bar{s}_2 + \bar{\ell}_2 + m_2$$
$$1 \ge k_1 + \bar{s}_1 + m_1 + \bar{x}_1$$
$$1 \ge k_2 + \bar{s}_2 + m_2 + \bar{x}_2$$

We apply (C.89)-(C.92) in (C.87) and (C.88). After this, we drop (C.89)-(C.92).

$$R = (1-\delta_1)\, m_1 + (1-\delta_2)\, m_2$$
$$R = (1-\delta_1\delta_{1E})\, I(X_1;W|YZ_1) + (1-\delta_2)\, I(X_2;W|YZ_1) \quad \text{(C.93)}$$
$$R = (1-\delta_2\delta_{2E})\, I(X_2;W|YZ_2) + (1-\delta_1)\, I(X_1;W|YZ_2) \quad \text{(C.94)}$$
$$R = (1-\delta_1\delta_{1E})\, I(X_1;W|YZ) + (1-\delta_2\delta_{2E})\, I(X_2;W|YZ)$$
$$C_r \ge (1-\delta_1\delta_{1E})\,\bar{s}_1 + (1-\delta_1\delta_{1E})\,\bar{s}_2$$
$$(1-\delta_{1E})\, I(X_1;W|YZ_1) \le \delta_{1E}(1-\delta_1)\left(k_1+\bar{s}_1+\bar{\ell}_1\right) + (1-\delta_2)\,\bar{s}_2 + (1-\delta_2)\,\bar{x}_2 \quad \text{(C.95)}$$
$$(1-\delta_{1E})\, I(X_1;W|YZ_1) \le \delta_{1E}(1-\delta_1)\left(k_1+\bar{s}_1+\bar{\ell}_1\right) + (1-\delta_2\delta_{2E})\,\bar{s}_2 - (1-\delta_1\delta_{1E})\,\bar{\ell}_1$$
$$(1-\delta_{2E})\, I(X_2;W|YZ_2) \le \delta_{2E}(1-\delta_2)\left(k_2+\bar{s}_2+\bar{\ell}_2\right) + (1-\delta_1)\,\bar{s}_1 + (1-\delta_1)\,\bar{x}_1 \quad \text{(C.96)}$$
$$(1-\delta_{2E})\, I(X_2;W|YZ_2) \le \delta_{2E}(1-\delta_2)\left(k_2+\bar{s}_2+\bar{\ell}_2\right) + (1-\delta_1\delta_{1E})\,\bar{s}_1 - (1-\delta_2\delta_{2E})\,\bar{\ell}_2$$
$$I(X_1;W|YZ) \le I(X_1;W|YZ_1) + \bar{\ell}_1$$
$$I(X_2;W|YZ) \le I(X_2;W|YZ_2) + \bar{\ell}_2$$
$$1 \ge k_1 + \bar{s}_1 + \bar{x}_1 + I(X_1;W|YZ_2)$$
$$1 \ge k_2 + \bar{s}_2 + \bar{x}_2 + I(X_2;W|YZ_1)$$
$$1 \ge k_1 + \bar{s}_1 + \bar{\ell}_1 + m_1$$

$$1 \geq k_2 + \bar{s}_2 + \bar{\ell}_2 + m_2$$
$$1 \geq k_1 + \bar{s}_1 + m_1 + \bar{x}_1$$
$$1 \geq k_2 + \bar{s}_2 + m_2 + \bar{x}_2$$

In (C.93) we write

$$(1-\delta_1\delta_{1E})\, I(X_1; W|YZ_1) = (1-\delta_{1E})\, I(X_1; W|YZ_1) + \delta_{1E}(1-\delta_1)\, I(X_1; W|YZ_1) \tag{C.97}$$

and apply (C.95) on the first term. We replace (C.93) with the resulting inequality. We do a similar replacement with (C.94) using (C.96).

$$R = (1-\delta_1)\, m_1 + (1-\delta_2)\, m_2$$
$$R \leq \delta_{1E}(1-\delta_1)\left(k_1 + \bar{s}_1 + \bar{\ell}_1\right) + (1-\delta_2)\,\bar{s}_2 + \delta_{1E}(1-\delta_1)\, I(X_1; W|YZ_1)$$
$$+ (1-\delta_2)\, I(X_2; W|YZ_1) + (1-\delta_2)\,\bar{x}_2 \tag{C.98}$$
$$R \leq \delta_{2E}(1-\delta_2)\left(k_2 + \bar{s}_2 + \bar{\ell}_2\right) + (1-\delta_1)\,\bar{s}_1 + \delta_{2E}(1-\delta_2)\, I(X_2; W|YZ_2)$$
$$+ (1-\delta_1)\, I(X_1; W|YZ_2) + (1-\delta_1)\,\bar{x}_1 \tag{C.99}$$
$$R = (1-\delta_1\delta_{1E})\, I(X_1; W|YZ) + (1-\delta_2\delta_{2E})\, I(X_2; W|YZ)$$
$$C_r \geq (1-\delta_1\delta_{1E})\,\bar{s}_1 + (1-\delta_1\delta_{1E})\,\bar{s}_2$$
$$(1-\delta_{1E})\, I(X_1; W|YZ_1) \leq \delta_{1E}(1-\delta_1)\left(k_1 + \bar{s}_1 + \bar{\ell}_1\right) + (1-\delta_2)\,\bar{s}_2 + (1-\delta_2)\,\bar{x}_2 \tag{C.100}$$
$$(1-\delta_{1E})\, I(X_1; W|YZ_1) \leq \delta_{1E}(1-\delta_1)\left(k_1 + \bar{s}_1 + \bar{\ell}_1\right) + (1-\delta_2\delta_{2E})\,\bar{s}_2 - (1-\delta_1\delta_{1E})\,\bar{\ell}_1 \tag{C.101}$$
$$(1-\delta_{2E})\, I(X_2; W|YZ_2) \leq \delta_{2E}(1-\delta_2)\left(k_2 + \bar{s}_2 + \bar{\ell}_2\right) + (1-\delta_1)\,\bar{s}_1 + (1-\delta_1)\,\bar{x}_1 \tag{C.102}$$
$$(1-\delta_{2E})\, I(X_2; W|YZ_2) \leq \delta_{2E}(1-\delta_2)\left(k_2 + \bar{s}_2 + \bar{\ell}_2\right) + (1-\delta_1\delta_{1E})\,\bar{s}_1 - (1-\delta_2\delta_{2E})\,\bar{\ell}_2 \tag{C.103}$$
$$I(X_1; W|YZ) \leq I(X_1; W|YZ_1) + \bar{\ell}_1$$
$$I(X_2; W|YZ) \leq I(X_2; W|YZ_2) + \bar{\ell}_2$$
$$1 \geq k_1 + \bar{s}_1 + \bar{x}_1 + I(X_1; W|YZ_2) \tag{C.104}$$
$$1 \geq k_2 + \bar{s}_2 + \bar{x}_2 + I(X_2; W|YZ_1)$$
$$1 \geq k_1 + \bar{s}_1 + \bar{\ell}_1 + m_1$$
$$1 \geq k_2 + \bar{s}_2 + \bar{\ell}_2 + m_2$$
$$1 \geq k_1 + \bar{s}_1 + m_1 + \bar{x}_1$$
$$1 \geq k_2 + \bar{s}_2 + m_2 + \bar{x}_2$$

We show that there is an optimal point where (C.100) is equality. First, there is an optimal point where either (C.100) or (C.101) is equality. If it is not the case in an optimal point, then we can increase the value of $I(X_1; W|YZ_1)$ without violating any constraints until either becomes tight. Assume that in an optimal point the RHS of (C.101) is smaller than the RHS of (C.100). Then we do the following transform, for some $\Delta > 0$:

$$\bar{\ell}_1 \downarrow \Delta \tag{C.105}$$
$$I(X_1; W|YZ_1) \uparrow \Delta. \tag{C.106}$$

One can observe that the (C.101) remains equality, and the RHS of (C.98) does not change, thus no constraints are violated. As a result either (C.100) becomes equality or $\bar{\ell}_1 = 0$. If the latter occurs, do the following transform:

$$\bar{x_2} \downarrow \Delta \tag{C.107}$$

$$I(X_2; W|YZ_1) \uparrow \Delta. \tag{C.108}$$

Again, the RHS of (C.98) and (C.104) do not change, thus no constraints are violated. Since $\bar{\ell}_1 = 0$, (C.100) becomes tight at least when $\bar{x_2}$ reaches 0.

With a similar argument it can be assumed that (C.102) is also equality. Thus, adding the constraints

$$(1-\delta_1\delta_{1E})\,\bar{\ell}_1 + (1-\delta_2)\,\bar{x_2} \le \delta_2\,(1-\delta_{2E})\,\bar{s_2} \tag{C.109}$$

$$(1-\delta_2\delta_{2E})\,\bar{\ell}_2 + (1-\delta_1)\,\bar{x_1} \le \delta_1\,(1-\delta_{1E})\,\bar{s_1} \tag{C.110}$$

does not restrict the value of the program. We add these two constraints as well as we drop (C.101), (C.103) and change (C.100) and (C.102) to equalities. We also substitute these equalities back to (C.98)-(C.99).

$$R = (1-\delta_1)\,m_1 + (1-\delta_2)\,m_2 \tag{C.111}$$

$$R \le (1-\delta_1\delta_{1E})\,I(X_1; W|YZ_1) + (1-\delta_2)\,I(X_2; W|YZ_1) \tag{C.112}$$

$$R \le (1-\delta_2\delta_{2E})\,I(X_2; W|YZ_2) + (1-\delta_1)\,I(X_1; W|YZ_2) \tag{C.113}$$

$$R = (1-\delta_1\delta_{1E})\,I(X_1; W|YZ) + (1-\delta_2\delta_{2E})\,I(X_2; W|YZ)$$

$$C_r \ge (1-\delta_1\delta_{1E})\,\bar{s_1} + (1-\delta_1\delta_{1E})\,\bar{s_2}$$

$$(1-\delta_{1E})\,I(X_1; W|YZ_1) = \delta_{1E}\,(1-\delta_1)\left(k_1 + \bar{s_1} + \bar{\ell}_1\right) + (1-\delta_2)\,\bar{s_2} + (1-\delta_2)\,\bar{x_2}$$

$$(1-\delta_{2E})\,I(X_2; W|YZ_2) = \delta_{2E}\,(1-\delta_2)\left(k_2 + \bar{s_2} + \bar{\ell}_2\right) + (1-\delta_1)\,\bar{s_1} + (1-\delta_1)\,\bar{x_1}$$

$$(1-\delta_1\delta_{1E})\,\bar{\ell}_1 + (1-\delta_2)\,\bar{x_2} \le \delta_2\,(1-\delta_{2E})\,\bar{s_2}$$

$$(1-\delta_2\delta_{2E})\,\bar{\ell}_2 + (1-\delta_1)\,\bar{x_1} \le \delta_1\,(1-\delta_{1E})\,\bar{s_1}$$

$$I(X_1; W|YZ) \le I(X_1; W|YZ_1) + \bar{\ell}_1$$

$$I(X_2; W|YZ) \le I(X_2; W|YZ_2) + \bar{\ell}_2$$

$$1 \ge k_1 + \bar{s_1} + \bar{x_1} + I(X_1; W|YZ_2) \tag{C.114}$$

$$1 \ge k_2 + \bar{s_2} + \bar{x_2} + I(X_2; W|YZ_1)$$

$$1 \ge k_1 + \bar{s_1} + \bar{\ell}_1 + m_1$$

$$1 \ge k_2 + \bar{s_2} + \bar{\ell}_2 + m_2$$

$$1 \ge k_1 + \bar{s_1} + m_1 + \bar{x_1} \tag{C.115}$$

$$1 \ge k_2 + \bar{s_2} + m_2 + \bar{x_2} \tag{C.116}$$

We show that

$$m_1 = I(X_1; W|YZ_2) \tag{C.117}$$
$$m_2 = I(X_2; W|YZ_1) \tag{C.118}$$

can be assumed without restricting the value of the program. First, we show that

$$m_1 \leq I(X_1; W|YZ_2) \tag{C.119}$$

can be assumed. If in an optimal point this does not hold, then we can increase the value of $I(X_1; W|YZ_2)$ until it holds without violating any constraints. Note that (C.114) is respected due to (C.115). With a similar argument

$$m_2 \leq I(X_2; W|YZ_1) \tag{C.120}$$

can be assumed. We continue the proof in two steps. First we show that at least one of (C.117)-(C.118) can be assumed. Assume the contrary:

$$m_1 < I(X_1; W|YZ_2)$$
$$m_2 < I(X_2; W|YZ_1).$$

If in an optimal point (C.112) is not equality, then one can decrease $I(X_2; W|YZ_1)$ until either (C.117) holds or (C.112) becomes equality. Hence, if (C.117) does not hold, then one can assume that (C.112) is tight. Similarly, we can assume that (C.113) is equality. If $\bar{\ell}_2 > 0$, consider the following transform for some $\Delta > 0$:

$$\bar{\ell}_2 \downarrow \frac{\Delta}{1-\delta_2\delta_{2E}} \tag{C.121}$$
$$\bar{x_1} \uparrow \frac{\Delta}{1-\delta_1} \tag{C.122}$$
$$I(X_1; W|YZ_2) \downarrow \frac{\Delta}{1-\delta_1} \tag{C.123}$$
$$I(X_2; W|YZ_2) \uparrow \frac{\Delta}{1-\delta_2\delta_{2E}} \tag{C.124}$$

As a result of this transform either $\bar{\ell}_2 = 0$ or $m_1 = I(X_1; W|YZ_2)$ occurs. Note that other constraints remain unchanged, except for (C.115). However, if (C.115) was equality, $m_1 \geq I(X_1; W|YZ_2)$ would follow. We can also do a symmetric transform:

$$\bar{\ell}_1 \downarrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.125}$$
$$\bar{x_2} \uparrow \frac{\Delta}{1-\delta_2} \tag{C.126}$$
$$I(X_2; W|YZ_1) \downarrow \frac{\Delta}{1-\delta_2} \tag{C.127}$$

$$I(X_1;W|YZ_1) \uparrow \frac{\Delta}{1-\delta_1\delta_{1E}}. \tag{C.128}$$

Thus, if still neither of (C.117) and (C.118) holds, we can assume $\bar{\ell}_1 = \bar{\ell}_2 = 0$. At this point, we can increase $(1-\delta_2)\, m_2$ and decrease $(1-\delta_1)\, m_1$ by the same amount, until (C.118) holds without violating any constraints (if (C.116) becomes equality, then (C.118) already holds). We note that it cannot happen that m_1 reaches 0 before (C.118) holds, because if $m_1 = 0$, then the RHS of (C.111) is strictly smaller than the RHS of (C.112), thus (C.112) could not be tight. We conclude that there is an optimal point, where at least one of (C.117),(C.118) holds. W.l.o.g. we assume

$$m_2 = I(X_2;W|YZ_1). \tag{C.129}$$

We then have:

$$R = (1-\delta_1)\, m_1 + (1-\delta_2)\, m_2 \tag{C.130}$$
$$R \leq (1-\delta_1\delta_{1E})\, I(X_1;W|YZ_1) + (1-\delta_2)\, m_2 \tag{C.131}$$
$$R \leq (1-\delta_2\delta_{2E})\, I(X_2;W|YZ_2) + (1-\delta_1)\, I(X_1;W|YZ_2) \tag{C.132}$$
$$R = (1-\delta_1\delta_{1E})\, I(X_1;W|YZ) + (1-\delta_2\delta_{2E})\, I(X_2;W|YZ) \tag{C.133}$$
$$C_r \geq (1-\delta_1\delta_{1E})\, \bar{s}_1 + (1-\delta_1\delta_{1E})\, \bar{s}_2 \tag{C.134}$$
$$(1-\delta_{1E})\, I(X_1;W|YZ_1) = \delta_{1E}(1-\delta_1)\left(k_1 + \bar{s}_1 + \bar{\ell}_1\right) + (1-\delta_2)\, \bar{s}_2 + (1-\delta_2)\, \bar{x}_2 \tag{C.135}$$
$$(1-\delta_{2E})\, I(X_2;W|YZ_2) = \delta_{2E}(1-\delta_2)\left(k_2 + \bar{s}_2 + \bar{\ell}_2\right) + (1-\delta_1)\, \bar{s}_1 + (1-\delta_1)\, \bar{x}_1 \tag{C.136}$$
$$(1-\delta_1\delta_{1E})\, \bar{\ell}_1 + (1-\delta_2)\, \bar{x}_2 \leq \delta_2(1-\delta_{2E})\, \bar{s}_2 \tag{C.137}$$
$$(1-\delta_2\delta_{2E})\, \bar{\ell}_2 + (1-\delta_1)\, \bar{x}_1 \leq \delta_1(1-\delta_{1E})\, \bar{s}_1 \tag{C.138}$$
$$I(X_1;W|YZ) \leq I(X_1;W|YZ_1) + \bar{\ell}_1 \tag{C.139}$$
$$I(X_2;W|YZ) \leq I(X_2;W|YZ_2) + \bar{\ell}_2 \tag{C.140}$$
$$1 \geq k_1 + \bar{s}_1 + \bar{x}_1 + I(X_1;W|YZ_2) \tag{C.141}$$
$$1 \geq k_1 + \bar{s}_1 + \bar{\ell}_1 + m_1 \tag{C.142}$$
$$1 \geq k_2 + \bar{s}_2 + \bar{\ell}_2 + \tag{C.143}$$
$$1 \geq k_1 + \bar{s}_1 + m_1 + \bar{x}_1 \tag{C.144}$$
$$1 \geq k_2 + \bar{s}_2 + m_2 + \bar{x}_2 \tag{C.145}$$
$$m_1 \leq I(X_1;W|YZ_2) \tag{C.146}$$

We show that the last constraint can also be made tight, i.e., $m_1 = I(X_1;W|YZ_2)$ can be assumed. If in an optimal point (C.132) is not equality, then one can decrease $I(X_1;W|YZ_2)$ until either (C.146) is tight or (C.132) is equality. Thus, we assume (C.132) is tight. Similarly as in the previous step, if $\bar{\ell}_2 > 0$ we do the following transform:

$$\bar{\ell}_2 \downarrow \frac{\Delta}{1-\delta_2\delta_{2E}} \tag{C.147}$$

$$\bar{x_1} \uparrow \frac{\Delta}{1-\delta_1} \tag{C.148}$$

$$I(X_1;W|YZ_2) \downarrow \frac{\Delta}{1-\delta_1} \tag{C.149}$$

$$I(X_2;W|YZ_2) \uparrow \frac{\Delta}{1-\delta_2\delta_{2E}}. \tag{C.150}$$

Again, either $m_1 = I(X_1;W|YZ_2)$ or $\bar{\ell}_2 = 0$. Hence we assume $\bar{\ell}_2 = 0$. We show that $\bar{\ell}_1 = 0$ can also be assumed. If $\bar{\ell}_1 > 0$, we do the following transform:

$$\bar{\ell}_1 \downarrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.151}$$

$$\bar{s_2} \downarrow \frac{\Delta}{\delta_2(1-\delta_{2E})} \tag{C.152}$$

$$\bar{s_1} \uparrow \frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta \tag{C.153}$$

$$\bar{x_1} \uparrow \frac{\delta_1(1-\delta_{1E})(1-\delta_2\delta_{2E})}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})(1-\delta_1)}\Delta \tag{C.154}$$

$$I(X_1;W|YZ_1) \downarrow \frac{1-\delta_2}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta \tag{C.155}$$

$$I(X_1;W|YZ) \downarrow \frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta \tag{C.156}$$

$$I(X_2;W|YZ_2) \uparrow \frac{\Delta}{\delta_2(1-\delta_{2E})} \tag{C.157}$$

$$I(X_2;W|YZ) \uparrow \frac{\Delta}{\delta_2(1-\delta_{2E})} \tag{C.158}$$

$$I(X_1;W|YZ_2) \downarrow \frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})(1-\delta_1)}\Delta \tag{C.159}$$

$$m_1 \downarrow \frac{1-\delta_2}{\delta_2(1-\delta_{2E})(1-\delta_1)}\Delta \tag{C.160}$$

$$m_2 \uparrow \frac{\Delta}{\delta_2(1-\delta_{2E})} \tag{C.161}$$

A side calculation in Section C.3.1 shows that no constraints are violated by this transform. If $\bar{s_2} = 0$ or $I(X_1;W|YZ_1) = 0$ then $\bar{\ell}_1 = 0$ follows, thus if after this transform $m_1 < I(X_1;W|YZ_2)$ still holds and $\bar{\ell}_1 \neq 0$, then there are two cases:

Case 1. $m_1 = 0$. Do the following transform:

$$I(X_1;W|YZ) \downarrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.162}$$

$$I(X_2;W|YZ) \uparrow \frac{\Delta}{1-\delta_2\delta_{12E}} \tag{C.163}$$

As a result, either $I(X_1;W|YZ) = 0$ (i.e., *Case* 2) or $I(X_2;W|YZ) = I(X_2;W|YZ_2)$ occurs. Assume the latter. If (C.142) was not equality, we could do the following trans-

form:

$$\bar{\ell_1} \downarrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.164}$$

$$\bar{x_2} \uparrow \frac{\Delta}{1-\delta_2} \tag{C.165}$$

$$m_2 \downarrow \frac{\Delta}{1-\delta_2} \tag{C.166}$$

$$m_1 \uparrow \frac{\Delta}{1-\delta_1} \tag{C.167}$$

$$I(X_1;W|YZ_1) \uparrow \frac{\Delta}{1-\delta_1\delta_{1E}}. \tag{C.168}$$

Note that $m_2 = 0$ cannot occur, otherwise (C.130) could not be equality. Thus, we assume (C.142) is equality. From this,

$$\bar{\ell_1} \geq \bar{x_1} + I(X_1;W|YZ_2) \tag{C.169}$$

which implies that (C.139) cannot be tight, because otherwise, if (C.139) is equality, then

$$\begin{aligned}(1-\delta_1\delta_{1E})\, I(X_1;W|YZ) \\ \geq (1-\delta_1\delta_{1E})\, I(X_1;W|YZ_1) + (1-\delta_1\delta_{1E})\, I(X_1;W|YZ_2)\end{aligned}$$

would follow. But from the fact that (C.132) is equality and $I(X_2;W|YZ) = I(X_2;W|YZ_2)$

$$(1-\delta_1\delta_{1E})\, I(X_1;W|YZ) = (1-\delta_1)\, I(X_1;W|YZ_2) \tag{C.170}$$

holds, thus $I(X_1;W|YZ_1) = 0$ and $\bar{\ell_1} = 0$ would follow. Thus, one can decrease $\bar{\ell_1}$ and $I(X_1;W|YZ_1)$ without violating any constraints until $\bar{\ell_1} = 0$ holds. Note that since $m_1 = 0$, (C.131) is not violated.

Case 2. $I(X_1;W|YZ) = 0$. Since (C.132) is equality, from (C.140) and (C.133) it follows that $I(X_1;W|YZ_2) = 0$, thus (C.146) must be equality.

We conclude, that if $m_1 < I(X_1;W|YZ_2)$ holds, then $\bar{\ell_1} = 0$ can be assumed. If this is the case, then one can decrease $(1-\delta_2)\, m_2$ and increase $(1-\delta_1)\, m_1$ by the same amount, until either $m_1 = I(X_1;W|YZ_2)$ or (C.131) is equality. In the latter case we know that

$$(1-\delta_2)\, m_2 > (1-\delta_2\delta_{2E})\, I(X_2;W|YZ_2) \geq (1-\delta_2\delta_{2E})\, I(X_2;W|YZ) \tag{C.171}$$

$$(1-\delta_1)\, m_1 = (1-\delta_1\delta_{1E})\, I(X_1;W|YZ_1) \geq (1-\delta_1\delta_{1E})\, I(X_1;W|YZ)\,. \tag{C.172}$$

These two imply that the RHS of (C.133) is strictly smaller than the RHS of (C.130), which is not possible. Thus we can conclude that $m_1 = I(X_1;W|YZ_2)$ can be assumed without restricting the value of the program.

As a result we have:

$$R = (1-\delta_1)\, m_1 + (1-\delta_2)\, m_2 \tag{C.173}$$
$$R \leq (1-\delta_1\delta_{1E})\, I(X_1; W|YZ_1) + (1-\delta_2)\, m_2 \tag{C.174}$$
$$R \leq (1-\delta_2\delta_{2E})\, I(X_2; W|YZ_2) + (1-\delta_1)\, m_1$$
$$R = (1-\delta_1\delta_{1E})\, I(X_1; W|YZ) + (1-\delta_2\delta_{2E})\, I(X_2; W|YZ) \tag{C.175}$$
$$C_r \geq \bar{s}_1 + \bar{s}_2$$
$$(1-\delta_{1E})\, I(X_1; W|YZ_1) = \delta_{1E}(1-\delta_1)\left(k_1 + \bar{s}_1 + \bar{\ell}_1\right) + (1-\delta_2)\,\bar{s}_2 + (1-\delta_2)\,\bar{x}_2 \tag{C.176}$$
$$(1-\delta_{2E})\, I(X_2; W|YZ_2) = \delta_{2E}(1-\delta_2)\left(k_2 + \bar{s}_2 + \bar{\ell}_2\right) + (1-\delta_1)\,\bar{s}_1 + (1-\delta_1)\,\bar{x}_1 \tag{C.177}$$
$$(1-\delta_1\delta_{1E})\,\bar{\ell}_1 + (1-\delta_2)\,\bar{x}_2 \leq \delta_2(1-\delta_{2E})\,\bar{s}_2$$
$$(1-\delta_2\delta_{2E})\,\bar{\ell}_2 + (1-\delta_1)\,\bar{x}_1 \leq \delta_1(1-\delta_{1E})\,\bar{s}_1$$
$$I(X_1; W|YZ) \leq I(X_1; W|YZ_1) + \bar{\ell}_1 \tag{C.178}$$
$$I(X_2; W|YZ) \leq I(X_2; W|YZ_2) + \bar{\ell}_2 \tag{C.179}$$
$$1 \geq k_1 + \bar{s}_1 + \bar{\ell}_1 + m_1$$
$$1 \geq k_2 + \bar{s}_2 + \bar{\ell}_2 + m_2$$
$$1 \geq k_1 + \bar{s}_1 + m_1 + \bar{x}_1$$
$$1 \geq k_2 + \bar{s}_2 + m_2 + \bar{x}_2$$

From (C.173), (C.174) and (C.176)

$$\frac{(1-\delta_1)(1-\delta_{1E})}{1-\delta_1\delta_{1E}}\, m_1 \leq \delta_{1E}(1-\delta_1)\left(k_1 + \bar{s}_1 + \bar{\ell}_1\right) + (1-\delta_2)\,\bar{s}_2 + (1-\delta_2)\,\bar{x}_2 \tag{C.180}$$

follows. Similarly

$$\frac{(1-\delta_2)(1-\delta_{2E})}{1-\delta_2\delta_{2E}}\, m_2 \leq \delta_{2E}(1-\delta_2)\left(k_2 + \bar{s}_2 + \bar{\ell}_2\right) + (1-\delta_1)\,\bar{s}_1 + (1-\delta_1)\,\bar{x}_1 \tag{C.181}$$

holds. We add these two constraints, while we drop (C.175), (C.176), (C.177), (C.178) and (C.179).

$$R = (1-\delta_1)\, m_1 + (1-\delta_2)\, m_2$$
$$C_r \geq \bar{s}_1 + \bar{s}_2$$
$$\frac{(1-\delta_1)(1-\delta_{1E})}{1-\delta_1\delta_{1E}}\, m_1 \leq \delta_{1E}(1-\delta_1)\left(k_1 + \bar{s}_1 + \bar{\ell}_1\right) + (1-\delta_2)\,\bar{s}_2 + (1-\delta_2)\,\bar{x}_2$$
$$\frac{(1-\delta_2)(1-\delta_{2E})}{1-\delta_2\delta_{2E}}\, m_2 \leq \delta_{2E}(1-\delta_2)\left(k_2 + \bar{s}_2 + \bar{\ell}_2\right) + (1-\delta_1)\,\bar{s}_1 + (1-\delta_1)\,\bar{x}_1$$
$$(1-\delta_1\delta_{1E})\,\bar{\ell}_1 + (1-\delta_2)\,\bar{x}_2 \leq \delta_2(1-\delta_{2E})\,\bar{s}_2 \tag{C.182}$$
$$(1-\delta_2\delta_{2E})\,\bar{\ell}_2 + (1-\delta_1)\,\bar{x}_1 \leq \delta_1(1-\delta_{1E})\,\bar{s}_1 \tag{C.183}$$
$$1 \geq k_1 + \bar{s}_1 + \bar{\ell}_1 + m_1 \tag{C.184}$$
$$1 \geq k_2 + \bar{s}_2 + \bar{\ell}_2 + m_2$$

$$1 \geq k_1 + \bar{s_1} + m_1 + \bar{x_1} \tag{C.185}$$
$$1 \geq k_2 + \bar{s_2} + m_2 + \bar{x_2}$$

We show that there is an optimal point where $\bar{\ell}_1 \leq \bar{x_1}$. If in an optimal point it is not true, then do the following transform:

$$\bar{\ell}_1 \downarrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.186}$$
$$\bar{s_2} \downarrow \frac{\Delta}{\delta_2\left(1-\delta_{2E}\right)} \tag{C.187}$$
$$\bar{s_1} \uparrow \frac{1-\delta_2\delta_{2E}}{\delta_2\left(1-\delta_{2E}\right)\left(1-\delta_1\delta_{1E}\right)}\Delta \tag{C.188}$$
$$\bar{x_1} \uparrow \frac{\delta_1\left(1-\delta_{1E}\right)\left(1-\delta_2\delta_{2E}\right)}{\delta_2\left(1-\delta_{2E}\right)\left(1-\delta_1\delta_{1E}\right)\left(1-\delta_1\right)}\Delta \tag{C.189}$$
$$m_1 \downarrow \frac{1-\delta_2}{\delta_2\left(1-\delta_{2E}\right)\left(1-\delta_1\right)}\Delta \tag{C.190}$$
$$m_2 \uparrow \frac{\Delta}{\delta_2\left(1-\delta_{2E}\right)} \tag{C.191}$$

By this transform, one of the following cases occurs:

Case 1 $\bar{\ell}_1 = 0$. This immediately implies $\bar{\ell}_1 \leq \bar{x_1}$.

Case 2 $m_1 = 0$. In this case, we can decrease $\bar{\ell}_1$ without violating any constraints until $\bar{\ell}_1 = 0$.

Case 3 $\bar{s_2} = 0$. Which also implies $\bar{\ell}_1 = 0$.

Case 4 (C.185) is equality, which together with (C.184) implies $\bar{\ell}_1 \leq \bar{x_1}$.

Similarly we can show that $\bar{\ell}_2 \leq \bar{x_2}$. Let

$$r_1 = \left(\bar{x_1} - \bar{\ell}_1\right)\frac{\left(1-\delta_1\delta_{1E}\right)\left(1-\delta_1\right)}{\delta_1\left(1-\delta_{1E}\right)} \tag{C.192}$$
$$r_2 = \left(\bar{x_2} - \bar{\ell}_2\right)\frac{\left(1-\delta_2\delta_{2E}\right)\left(1-\delta_2\right)}{\delta_2\left(1-\delta_{2E}\right)} \tag{C.193}$$
$$c = \left(1-\delta_1\delta_{1E}\right)\bar{s_1} + \left(1-\delta_2\delta_{2E}\right)\bar{s_2} - r_1 - r_2 \tag{C.194}$$
$$c_1 = \bar{\ell}_1 + \bar{s_1} - \frac{r_1}{1-\delta_1\delta_{1E}} \tag{C.195}$$
$$c_2 = \bar{\ell}_2 + \bar{s_2} - \frac{r_2}{1-\delta_2\delta_{2E}} \tag{C.196}$$

Note that $c, c_1, c_2 \geq 0$ directly follows from (C.182)-(C.183). We get:

$$R = \left(1-\delta_1\right)m_1 + \left(1-\delta_2\right)m_2 \tag{C.197}$$
$$C_r \geq c + r_1 + r_2 \tag{C.198}$$
$$1 \geq k_1 + m_1 + c_1 + \frac{r_1}{1-\delta_1} \tag{C.199}$$
$$1 \geq k_2 + m_2 + c_2 + \frac{r_2}{1-\delta_2} \tag{C.200}$$

$$m_1 \frac{(1-\delta_{1E})(1-\delta_1)}{1-\delta_1\delta_{1E}} \le r_2 + r_1 \frac{\delta_{1E}(1-\delta_1)}{1-\delta_1\delta_{1E}} + c_2(1-\delta_2) + (c_1+k_1)\delta_{1E}(1-\delta_1) \tag{C.201}$$

$$m_2 \frac{(1-\delta_{2E})(1-\delta_2)}{1-\delta_2\delta_{2E}} \le r_1 + r_2 \frac{\delta_{2E}(1-\delta_2)}{1-\delta_2\delta_{2E}} + c_1(1-\delta_1) + (c_2+k_2)\delta_{2E}(1-\delta_2) \tag{C.202}$$

$$(1-\delta_1\delta_{1E})c_1 + (1-\delta_2)c_2 \le c \tag{C.203}$$

$$(1-\delta_2\delta_{2E})c_2 + (1-\delta_1)c_1 \le c \tag{C.204}$$

which is the same as the linear program for the scheme. □

C.3.1 Side calculation

Change of RHS of (C.130):

$$\underbrace{-\frac{1-\delta_2}{\delta_2(1-\delta_{2E})}\Delta}_{\text{from } (1-\delta_1)m_1} + \underbrace{\frac{1-\delta_2}{\delta_2(1-\delta_{2E})}\Delta}_{\text{from } (1-\delta_2)m_2} = 0. \tag{C.205}$$

Change of RHS of (C.131):

$$\underbrace{-\frac{1-\delta_2}{\delta_2(1-\delta_{2E})}\Delta}_{\text{from } (1-\delta_1\delta_{1E})I(X_1;W|YZ_1)} + \underbrace{\frac{1-\delta_2}{\delta_2(1-\delta_{2E})}\Delta}_{\text{from } (1-\delta_2)m_2} = 0. \tag{C.206}$$

Change of RHS of (C.132):

$$\underbrace{\frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})}\Delta}_{\text{from } (1-\delta_2\delta_{2E})I(X_2;W|YZ_2)} \underbrace{-\frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})}\Delta}_{\text{from } (1-\delta_1)I(X_1;W|YZ_2)} = 0. \tag{C.207}$$

Change of RHS of (C.133):

$$\underbrace{-\frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})}\Delta}_{\text{from } (1-\delta_1\delta_{1E})I(X_1;W|YZ)} + \underbrace{\frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})}\Delta}_{\text{from } (1-\delta_2\delta_{2E})I(X_2;W|YZ)} = 0. \tag{C.208}$$

Change of RHS of (C.134):

$$\underbrace{\frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})}\Delta}_{\text{from } (1-\delta_1\delta_{1E})\bar{s}_1} \underbrace{-\frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})}\Delta}_{\text{from } (1-\delta_2\delta_{2E})\bar{s}_2} = 0. \tag{C.209}$$

Change of LHS of (C.135):

$$-\frac{(1-\delta_2)(1-\delta_{1E})}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta \tag{C.210}$$

Change of RHS of (C.135):

$$\underbrace{\frac{\delta_{1E}(1-\delta_1)(1-\delta_2\delta_{2E})}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta}_{\text{from }\delta_{1E}(1-\delta_1)\,\bar{s}_1}-\underbrace{\frac{\delta_{1E}(1-\delta_1)}{1-\delta_1\delta_{1E}}\Delta}_{\text{from }\delta_{1E}(1-\delta_1)\,\bar{\ell}_1}-\underbrace{\frac{1-\delta_2}{\delta_2(1-\delta_{2E})}\Delta}_{\text{from }(1-\delta_2)\,\bar{\sigma}_2} \tag{C.211}$$

$$=\frac{\delta_{1E}(1-\delta_1)(1-\delta_2\delta_{2E})}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta-\frac{\delta_2(1-\delta_{2E})\delta_{1E}(1-\delta_1)}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta-\frac{(1-\delta_2)(1-\delta_1\delta_{1E})}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta \tag{C.212}$$

$$=\frac{\delta_{1E}(1-\delta_1)(1-\delta_2)}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta-\frac{(1-\delta_2)(1-\delta_1\delta_{1E})}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta=-\frac{(1-\delta_2)(1-\delta_{1E})}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta \tag{C.213}$$

Change of LHS of (C.136):

$$\frac{1-\delta_{2E}}{\delta_2(1-\delta_{2E})}\Delta \tag{C.214}$$

Change of RHS of (C.136):

$$\underbrace{-\frac{\delta_{2E}(1-\delta_2)}{\delta_2(1-\delta_{2E})}\Delta}_{\text{from }\delta_{2E}(1-\delta_2)\,\bar{s}_2}+\underbrace{\frac{(1-\delta_2\delta_{2E})(1-\delta_1)}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta}_{\text{from }(1-\delta_1)\,\bar{s}_1}+\underbrace{\frac{\delta_1(1-\delta_{1E})(1-\delta_2\delta_{2E})}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta}_{\text{from }(1-\delta_1)\,I(X_1;Z_1|YZ_2W)} \tag{C.215}$$

$$-\frac{\delta_{2E}(1-\delta_2)}{\delta_2(1-\delta_{2E})}\Delta+\frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})}\Delta=\frac{1-\delta_{2E}}{\delta_2(1-\delta_{2E})}\Delta. \tag{C.216}$$

Change of LHS of (C.137): $-\Delta$, change of RHS: $-\Delta$.
Change of LHS of (C.138):

$$\frac{\delta_1(1-\delta_{1E})(1-\delta_2\delta_{2E})}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta. \tag{C.217}$$

Change of RHS of (C.138):

$$\frac{\delta_1(1-\delta_{1E})(1-\delta_2\delta_{2E})}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta. \tag{C.218}$$

Change of LHS of (C.139):

$$-\frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta. \tag{C.219}$$

Change of RHS of (C.139):

$$\underbrace{-\frac{1-\delta_2}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta}_{\text{from }I(X_1;W|YZ_1)}-\underbrace{\frac{\Delta}{1-\delta_1\delta_{1E}}}_{\text{from }\bar{\ell}_1}=-\frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta. \tag{C.220}$$

Change of LHS of (C.140):

$$\frac{\Delta}{\delta_2(1-\delta_{2E})} \tag{C.221}$$

Change of RHS of (C.140):

$$\frac{\Delta}{\delta_2(1-\delta_{2E})} \tag{C.222}$$

Change of RHS of (C.141):

$$\underbrace{\frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta}_{\text{from } \bar{s}_1} + \underbrace{\frac{\delta_1(1-\delta_{1E})(1-\delta_2\delta_{2E})}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})(1-\delta_1)}\Delta}_{\text{from } I(X_1;Z_1|YZ_2W)} - \underbrace{\frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})(1-\delta_1)}\Delta}_{\text{from } I(X_1;W|YZ_2)} \tag{C.223}$$

$$= \frac{(1-\delta_2\delta_{2E})(1-\delta_1)}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})(1-\delta_1)}\Delta + \frac{\delta_1(1-\delta_{1E})(1-\delta_2\delta_{2E})}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})(1-\delta_1)}\Delta - \frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})(1-\delta_1)}\Delta \tag{C.224}$$

$$= \frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})(1-\delta_1)}\Delta - \frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})(1-\delta_1)}\Delta = 0 \tag{C.225}$$

Change of RHS of (C.142):

$$\underbrace{\frac{1-\delta_2\delta_{2E}}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta}_{\text{from } \bar{s}_1} - \underbrace{\frac{\Delta}{1-\delta_1\delta_{1E}}}_{\text{from } \bar{\ell}_1} - \underbrace{\frac{1-\delta_2}{\delta_2(1-\delta_{2E})(1-\delta_1)}\Delta}_{\text{from } m_1} \tag{C.226}$$

$$= \frac{1-\delta_2}{\delta_2(1-\delta_{2E})(1-\delta_1\delta_{1E})}\Delta - \frac{1-\delta_2}{\delta_2(1-\delta_{2E})(1-\delta_1)}\Delta \leq 0 \tag{C.227}$$

Since the increase of m_2 equals the decrease of $\bar{s}_2$, (C.143) and (C.145) are also respected.

C.4 Triangle network outer bound proof

When deriving our bounds we will use the assumption that the inputs X_{1i}, X_{2i}, X_{3i} of the different channels in the same time slot are generated from different independent random sources. Theorem 4.2 shows that this assumption does not affect capacity. The proof trivially generalizes for the triangle network. We also assume that F_i contains the channel state of the eavesdroppers.

C.4.1 Rate constraints

We use the same kind of derivation as in Section 4.3.4 to get the following inequalities. We omit details.

$$\begin{aligned} nR - \mathscr{E}_{4.1} &\leq I\left(W; Y_1^n Y_2^n F^n\right) \\ &= \sum_{i=1}^{n} (1-\delta_1)\, I\left(X_{1i}; W | Y_1^{i-1} Y_2^{i-1} F^{i-1}\right) + (1-\delta_2)\, I\left(X_{2i}; W | Y_1^{i-1} Y_2^{i-1} F^{i-1}\right) \end{aligned} \tag{C.228}$$

$$\begin{aligned} nR - \mathscr{E}_{4.1} &\leq I\left(W; Y_1^n Y_2^n Z_1^n F^n\right) \\ &= \sum_{i=1}^{n} (1-\delta_1\delta_{1E})\, I\left(X_{1i}; W | Y_1^{i-1} Y_2^{i-1} Z_1^{i-1} F^{i-1}\right) + (1-\delta_2)\, I\left(X_{2i}; W | Y_1^{i-1} Y_2^{i-1} Z_1^{i-1} F^{i-1}\right) \end{aligned} \tag{C.229}$$

$$\begin{aligned} nR - \mathscr{E}_{4.1} &\leq I\left(W; Y_1^n Y_2^n Z_2^n F^n\right) \\ &= \sum_{i=1}^{n} (1-\delta_2\delta_{2E})\, I\left(X_{2i}; W | Y_1^{i-1} Y_2^{i-1} Z_2^{i-1} F^{i-1}\right) + (1-\delta_1)\, I\left(X_{1i}; W | Y_1^{i-1} Y_2^{i-1} Z_2^{i-1} F^{i-1}\right) \end{aligned} \tag{C.230}$$

$$\begin{aligned} nR - \mathscr{E}_{4.1} &\leq I\left(W; Y_1^n Y_3^n F^n\right) \\ &= \sum_{i=1}^{n} (1-\delta_1)\, I\left(X_{1i}; W | Y_1^{i-1} Y_3^{i-1} F^{i-1}\right) + (1-\delta_3)\, I\left(X_{3i}; W | Y_1^{i-1} Y_3^{i-1} F^{i-1}\right) \end{aligned} \tag{C.231}$$

$$\begin{aligned} nR - \mathscr{E}_{4.1} &\leq I\left(W; Y_1^n Y_3^n Z_1^n F^n\right) \\ &= \sum_{i=1}^{n} (1-\delta_1\delta_{1E})\, I\left(X_{1i}; W | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} F^{i-1}\right) + (1-\delta_3)\, I\left(X_{3i}; W | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} F^{i-1}\right) \end{aligned} \tag{C.232}$$

$$\begin{aligned} nR - \mathscr{E}_{4.1} &\leq I\left(W; Y_1^n Y_3^n Z_3^n F^n\right) \\ &= \sum_{i=1}^{n} (1-\delta_3\delta_{3E})\, I\left(X_{3i}; W | Y_1^{i-1} Y_3^{i-1} Z_3^{i-1} F^{i-1}\right) + (1-\delta_1)\, I\left(X_{1i}; W | Y_1^{i-1} Y_3^{i-1} Z_3^{i-1} F^{i-1}\right) \end{aligned} \tag{C.233}$$

C.4.2 Security constraints

We use a derivation similar to the derivation of (2.45). We omit details.

$$0 \leq H\left(Y_1^n Y_2^n | Z_1^n F^n W\right)$$

$$= \sum_{i=1}^{n} -(1-\delta_{1E})\, I\left(Y_1^{i-1}Y_2^{i-1}; X_{1i}|Z_1^{i-1}F^{i-1}W\right) + \delta_{1E}(1-\delta_1)\, H\left(X_{1i}|Y_1^{i-1}Y_2^{i-1}Z_1^{i-1}F^{i-1}W\right)$$
$$+ (1-\delta_2)\, H\left(X_{2i}|Y_1^{i-1}Y_2^{i-1}Z_1^{i-1}F^{i-1}W\right) \tag{C.234}$$

We have a similar inequality for the second eavesdropper.

$$0 \leq \sum_{i=1}^{n} -(1-\delta_{2E})\, I\left(Y_1^{i-1}Y_2^{i-1}; X_{2i}|Z_2^{i-1}F^{i-1}W\right) + \delta_{2E}(1-\delta_2)\, H\left(X_{2i}|Y_1^{i-1}Y_2^{i-1}Z_2^{i-1}F^{i-1}W\right)$$
$$+ (1-\delta_1)\, H\left(X_{1i}|Y_1^{i-1}Y_2^{i-1}Z_2^{i-1}F^{i-1}W\right) \tag{C.235}$$

We have two more inequalities of this kind.

$$0 \leq \sum_{i=1}^{n} -(1-\delta_{3E})\, I\left(Y_1^{i-1}Y_3^{i-1}; X_{3i}|Z_3^{i-1}F^{i-1}W\right) + \delta_{3E}(1-\delta_3)\, H\left(X_{3i}|Y_1^{i-1}Y_3^{i-1}Z_3^{i-1}F^{i-1}W\right)$$
$$+ (1-\delta_1)\, H\left(X_{1i}|Y_1^{i-1}Y_3^{i-1}Z_3^{i-1}F^{i-1}W\right) \tag{C.236}$$
$$0 \leq \sum_{i=1}^{n} -(1-\delta_{1E})\, I\left(Y_1^{i-1}Y_3^{i-1}; X_{1i}|Z_1^{i-1}F^{i-1}W\right) + \delta_{1E}(1-\delta_1)\, H\left(X_{1i}|Y_1^{i-1}Y_3^{i-1}Z_1^{i-1}F^{i-1}W\right)$$
$$+ (1-\delta_3)\, H\left(X_{3i}|Y_1^{i-1}Y_3^{i-1}Z_1^{i-1}F^{i-1}W\right) \tag{C.237}$$

We derive the following inequalities directly from the security criterion:

$$\sum_{i=1}^{n} I\left(X_{1i}; Y_1^{i-1}Y_2^{i-1}|Z_1^{i-1}F^{i-1}W\right)$$
$$= \sum_{i=1}^{n} I\left(X_{1i}; W|Y_1^{i-1}Y_2^{i-1}Z_1^{i-1}F^{i-1}\right) + I\left(X_{1i}; Y_1^{i-1}Y_2^{i-1}|Z_1^{i-1}F^{i-1}\right) - I\left(X_{1i}; W|Z_1^{i-1}F^{i-1}\right) \tag{C.238}$$
$$\geq \sum_{i=1}^{n} I\left(X_{1i}; W|Y_1^{i-1}Y_2^{i-1}Z_1^{i-1}F^{i-1}\right) - n\epsilon \tag{C.239}$$
$$\sum_{i=1}^{n} I\left(X_{2i}; Y_1^{i-1}Y_2^{i-1}|Z_2^{i-1}F^{i-1}W\right) \geq \sum_{i=1}^{n} I\left(X_{2i}; W|Y_1^{i-1}Y_2^{i-1}Z_2^{i-1}F^{i-1}\right) - n\epsilon \tag{C.240}$$
$$\sum_{i=1}^{n} I\left(X_{1i}; Y_1^{i-1}Y_3^{i-1}|Z_1^{i-1}F^{i-1}W\right) \geq \sum_{i=1}^{n} I\left(X_{1i}; W|Y_1^{i-1}Y_3^{i-1}Z_1^{i-1}F^{i-1}\right) - n\epsilon \tag{C.241}$$
$$\sum_{i=1}^{n} I\left(X_{3i}; Y_1^{i-1}Y_3^{i-1}|Z_3^{i-1}F^{i-1}W\right) \geq \sum_{i=1}^{n} I\left(X_{3i}; W|Y_1^{i-1}Y_3^{i-1}Z_3^{i-1}F^{i-1}\right) - n\epsilon. \tag{C.242}$$

C.4.3 Time-sharing constraints

The following few constraints ensure that no more than n transmissions are needed. We omit details.

$$n \geq \sum_{i=1}^{n} H\left(X_{1i}|Y_1^{i-1}Y_2^{i-1}Z_2^{i-1}F^{i-1}W\right) + I\left(X_{1i}; W|Y_1^{i-1}Y_2^{i-1}Z_2^{i-1}F^{i-1}\right) \tag{C.243}$$

$$n \geq \sum_{i=1}^{n} H\left(X_{2i}|Y_1^{i-1}Y_2^{i-1}Z_1^{i-1}F^{i-1}W\right) + I\left(X_{2i};W|Y_1^{i-1}Y_2^{i-1}Z_1^{i-1}F^{i-1}\right) \tag{C.244}$$

$$n \geq \sum_{i=1}^{n} H\left(X_{1i}|Y_1^{i-1}Y_2^{i-1}Z_1^{i-1}F^{i-1}W\right) + I\left(X_{1i};W|Y_1^{i-1}Y_2^{i-1}F^{i-1}\right) + I\left(X_{1i};Z_1^{i-1}|Y_1^{i-1}Y_2^{i-1}F^{i-1}W\right) \tag{C.245}$$

$$n \geq \sum_{i=1}^{n} H\left(X_{2i}|Y_1^{i-1}Y_2^{i-1}Z_1^{i-1}F^{i-1}W\right) + I\left(X_{2i};W|Y_1^{i-1}Y_2^{i-1}F^{i-1}\right) + I\left(X_{2i};Z_1^{i-1}|Y_1^{i-1}Y_2^{i-1}F^{i-1}W\right) \tag{C.246}$$

$$n \geq \sum_{i=1}^{n} H\left(X_{1i}|Y_1^{i-1}Y_2^{i-1}Z_2^{i-1}F^{i-1}W\right) + I\left(X_{1i};W|Y_1^{i-1}Y_2^{i-1}F^{i-1}\right) \tag{C.247}$$

$$n \geq \sum_{i=1}^{n} H\left(X_{2i}|Y_1^{i-1}Y_2^{i-1}Z_2^{i-1}F^{i-1}W\right) + I\left(X_{2i};W|Y_1^{i-1}Y_2^{i-1}F^{i-1}\right) \tag{C.248}$$

$$n \geq \sum_{i=1}^{n} H\left(X_{1i}|Y_1^{i-1}Y_3^{i-1}Z_3^{i-1}F^{i-1}W\right) + I\left(X_{1i};W|Y_1^{i-1}Y_3^{i-1}Z_3^{i-1}F^{i-1}\right) \tag{C.249}$$

$$n \geq \sum_{i=1}^{n} H\left(X_{3i}|Y_1^{i-1}Y_3^{i-1}Z_1^{i-1}F^{i-1}W\right) + I\left(X_{3i};W|Y_1^{i-1}Y_3^{i-1}Z_1^{i-1}F^{i-1}\right) \tag{C.250}$$

$$n \geq \sum_{i=1}^{n} H\left(X_{1i}|Y_1^{i-1}Y_3^{i-1}Z_1^{i-1}F^{i-1}W\right) + I\left(X_{1i};W|Y_1^{i-1}Y_3^{i-1}F^{i-1}\right) \tag{C.251}$$

$$n \geq \sum_{i=1}^{n} H\left(X_{3i}|Y_1^{i-1}Y_3^{i-1}Z_1^{i-1}F^{i-1}W\right) + I\left(X_{3i};W|Y_1^{i-1}Y_3^{i-1}F^{i-1}\right) \tag{C.252}$$

$$n \geq \sum_{i=1}^{n} H\left(X_{1i}|Y_1^{i-1}Y_3^{i-1}Z_3^{i-1}F^{i-1}W\right) + I\left(X_{1i};W|Y_1^{i-1}Y_3^{i-1}F^{i-1}\right) \tag{C.253}$$

$$n \geq \sum_{i=1}^{n} H\left(X_{3i}|Y_1^{i-1}Y_3^{i-1}Z_3^{i-1}F^{i-1}W\right) + I\left(X_{3i};W|Y_1^{i-1}Y_3^{i-1}F^{i-1}\right) \tag{C.254}$$

C.4.4 Distinguishing keys

We distinguish keys that we use on channel 1 and on channel 3 based on where they were generated.

$$\begin{aligned} 0 &\leq H\left(Y_1^n|Z_1^nF^nW\right) \\ &= \sum_{i=1}^{n} -(1-\delta_{1E})\, I\left(X_{1i};Y_1^{i-1}|Z_1^{i-1}F^{i-1}W\right) + \delta_{1E}(1-\delta_1)\, H\left(X_{1i}|Y_1^{i-1}Z_1^{i-1}F^{i-1}W\right) \qquad \text{(C.255)} \\ &= \sum_{i=1}^{n} -(1-\delta_{1E})\, I\left(X_{1i};Y_1^{i-1}|Z_1^{i-1}F^{i-1}W\right) + \delta_{1E}(1-\delta_1)\, H\left(X_{1i}|Y_1^{i-1}Y_3^{i-1}Z_1^{i-1}F^{i-1}W\right) \\ &\quad + \delta_{1E}(1-\delta_1)\, I\left(X_{1i};Y_3^{i-1}|Y_1^{i-1}Z_1^{i-1}F^{i-1}W\right) \qquad \text{(C.256)} \end{aligned}$$

A symmetric inequality holds for the other channel as well.

$$0 \leq \sum_{i=1}^{n} -(1-\delta_{3E})\, I\left(X_{3i}; Y_3^{i-1} | Z_3^{i-1} F^{i-1} W\right) + \delta_{3E}(1-\delta_3)\, H\left(X_{3i} | Y_1^{i-1} Y_3^{i-1} Z_3^{i-1} F^{i-1} W\right) + \delta_{3E}(1-\delta_3)\, I\left(X_{3i}; Y_1^{i-1} | Y_3^{i-1} Z_3^{i-1} F^{i-1} W\right) \tag{C.257}$$

Also,

$$\sum_{i=1}^{n} I\left(X_{1i}; Y_1^{i-1} Y_3^{i-1} | Z_1^{i-1} F^{i-1} W\right) = I\left(X_{1i}; Y_1^{i-1} | Z_1^{i-1} F^{i-1} W\right) + I\left(X_{1i}; Y_3^{i-1} | Y_1^{i-1} Z_1^{i-1} F^{i-1} W\right) \tag{C.258}$$

$$\sum_{i=1}^{n} I\left(X_{3i}; Y_1^{i-1} Y_3^{i-1} | Z_3^{i-1} F^{i-1} W\right) = I\left(X_{3i}; Y_3^{i-1} | Z_3^{i-1} F^{i-1} W\right) + I\left(X_{3i}; Y_1^{i-1} | Y_3^{i-1} Z_3^{i-1} F^{i-1} W\right) \tag{C.259}$$

C.4.5 Connecting cuts

So far all our inequalities hold for either the first cut or the other cut. However, since we want to show that the cut values are not achievable we need to connect the two cuts through some more constraints.

$$0 \leq H\left(Y_2^n | Y_1^n Y_3^n Z_1^n Z_3^n F^n W\right) = H\left(Y_2^{n-1} | Y_1^n Y_3^n Z_1^n Z_3^n F^n W\right) + H\left(Y_{2n} | Y_1^n Y_2^{n-1} Y_3^n Z_1^n Z_3^n F^n W\right) \tag{C.260}$$

$$\begin{aligned} &= H\left(Y_2^{n-1} | Y_1^{n-1} Y_3^{n-1} Z_1^{n-1} Z_3^{n-1} F^{n-1} W\right) \\ &\quad - I\left(Y_{1n} Z_{1n} Y_{3n} Z_{3n}; Y_2^{n-1} | Y_1^{n-1} Y_3^{n-1} Z_1^{n-1} Z_3^{n-1} F^{n-1} W\right) \\ &\quad + H\left(Y_{2n} | Y_1^{n-1} Y_2^{n-1} Y_3^{n-1} Z_1^{n-1} Z_3^{n-1} F^{n-1} W\right) \end{aligned} \tag{C.261}$$

$$\begin{aligned} &= \sum_{i=1}^{n} -(1-\delta_1\delta_{1E})\, I\left(X_{1i}; Y_2^{i-1} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} Z_3^{i-1} F^{i-1} W\right) \\ &\quad - (1-\delta_3\delta_{3E})\, I\left(X_{3i}; Y_2^{i-1} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} Z_3^{i-1} F^{i-1} W\right) + (1-\delta_2)\, H\left(X_{2i} | Y_1^{i-1} Y_2^{i-1} Z_1^{i-1} F^{i-1} W\right) \end{aligned} \tag{C.262}$$

In the last step we used that

$$\sum_{i=1}^{n} I\left(X_{2i}; Y_3^{i-1} Z_3^{i-1} | Y_1^{i-1} Y_2^{i-1} Z_1^{i-1} F^{i-1} W\right) = 0 \tag{C.263}$$

follows from the definition of channel inputs, because $\left(Y_2^{i-1}, W, \Theta_U, F^{i-1}\right)$ determine X_3^{i-1} and thus $\left(Y_3^{i-1}, Z_3^{i-1}\right)$, while X_{2i} is independent of Θ_U. Further,

$$0 \leq I\left(Z_1^n Y_1^n; Y_2^n | Y_3^n Z_3^n F^n W\right) \tag{C.264}$$

$$= I\left(Z_1^{n-1} Y_1^{n-1}; Y_2^n | Y_3^n Z_3^n F^n W\right) + I\left(Z_{1n} Y_{1n}; Y_2^n | Y_1^{n-1} Y_3^n Z_1^{n-1} Z_3^n F^n W\right) \tag{C.265}$$

$$= I\left(Z_1^{n-1} Y_1^{n-1}; Y_2^{n-1} | Y_3^n Z_3^n F^n W\right) + I\left(Y_{2n}; Z_1^{n-1} Y_1^{n-1} | Y_2^{n-1} Y_3^n Z_3^n F^n W\right)$$

$$+ I\left(Z_{1n}Y_{1n}; Y_2^{n-1}|Y_1^{n-1}Y_3^{n-1}Z_1^{n-1}Z_3^{n-1}F^{n-1}W\right) \tag{C.266}$$

$$= I\left(Z_1^{n-1}Y_1^{n-1}; Y_2^{n-1}|Y_3^{n-1}Z_3^{n-1}F^{n-1}W\right) - I\left(Y_{3n}Z_{3n}; Z_1^{n-1}Y_1^{n-1}|Y_3^{n-1}Z_3^{n-1}F^{n-1}W\right)$$
$$+ I\left(Y_{2n}; Z_1^{n-1}Y_1^{n-1}|Y_2^{n-1}Y_3^{n-1}Z_3^{n-1}F^{n-1}W\right) + I\left(Z_{1n}Y_{1n}; Y_2^{n-1}|Y_1^{n-1}Y_3^{n-1}Z_1^{n-1}Z_3^{n-1}F^{n-1}W\right) \tag{C.267}$$

$$= \sum_{i=1}^{n} -(1-\delta_3\delta_{3E})\, I\left(X_{3i}; Y_1^{i-1}|Y_3^{i-1}Z_3^{i-1}F^{i-1}W\right)$$
$$-(1-\delta_3\delta_{3E})\, I\left(X_{3i}; Z_1^{i-1}|Y_1^{i-1}Y_3^{i-1}Z_3^{i-1}F^{i-1}W\right)$$
$$+(1-\delta_2)\, I\left(X_{2i}; Y_1^{i-1}|Y_2^{i-1}F^{i-1}W\right) + (1-\delta_2)\, I\left(X_{2i}; Z_1^{i-1}|Y_1^{i-1}Y_2^{i-1}F^{i-1}W\right)$$
$$+(1-\delta_1\delta_{1E})\, I\left(X_{1i}; Y_2^{i-1}|Y_1^{i-1}Y_3^{i-1}Z_1^{i-1}Z_3^{i-1}F^{i-1}W\right) \tag{C.268}$$

In a similar way,

$$0 \le I\left(Y_3^n; Y_2^n|Y_1^nZ_1^nF^nW\right) = \sum_{i=1}^{n} -(1-\delta_1\delta_{1E})\, I\left(X_{1i}; Y_3^{i-1}|Y_1^{i-1}Z_1^{i-1}F^{i-1}W\right)$$
$$+(1-\delta_2)\, I\left(X_{2i}; Y_3^{i-1}|Y_1^{i-1}Y_2^{i-1}Z_1^{i-1}F^{i-1}W\right) + (1-\delta_3)\, I\left(X_{3i}; Y_2^{i-1}|Y_3^{i-1}Y_1^{i-1}Z_1^{i-1}F^{i-1}W\right) \tag{C.269}$$

$$= \sum_{i=1}^{n} -(1-\delta_1\delta_{1E})\, I\left(X_{1i}; Y_3^{i-1}|Y_1^{i-1}Z_1^{i-1}F^{i-1}W\right)$$
$$+(1-\delta_3)\, I\left(X_{3i}; Y_2^{i-1}|Y_3^{i-1}Y_1^{i-1}Z_1^{i-1}Z_3^{i-1}F^{i-1}W\right)$$
$$+(1-\delta_3)\, I\left(X_{3i}; Z_3^{i-1}|Y_3^{i-1}Y_1^{i-1}Z_1^{i-1}F^{i-1}W\right)$$
$$-(1-\delta_3)\, I\left(X_{3i}; Z_3^{i-1}|Y_2^{i-1}Y_3^{i-1}Y_1^{i-1}Z_1^{i-1}F^{i-1}W\right) \tag{C.270}$$

We used (C.263) again. We further have

$$H\left(Z_3^n|Y_3^nY_1^nZ_1^nF^nW\right) = \sum_{i=1}^{n} \delta_3(1-\delta_{3E})\, H\left(X_{3i}|Y_3^{i-1}Y_1^{i-1}Z_3^{i-1}Z_1^{i-1}F^{i-1}W\right)$$
$$-(1-\delta_1\delta_{1E})\, I\left(X_{1i}; Z_3^{i-1}|Y_3^{i-1}Y_1^{i-1}Z_1^{i-1}F^{i-1}W\right) - (1-\delta_3)\, I\left(X_{3i}; Z_3^{i-1}|Y_3^{i-1}Y_1^{i-1}Z_1^{i-1}F^{i-1}W\right) \tag{C.271}$$

$$H\left(Z_3^n|Y_3^nY_2^nY_1^nZ_1^nW\right) = \sum_{i=1}^{n} \delta_3(1-\delta_{3E})\, H\left(X_{3i}|Y_1^{i-1}Y_2^{i-1}Y_3^{i-1}Z_3^{i-1}Z_1^{i-1}F^{i-1}W\right)$$
$$-(1-\delta_3)\, I\left(X_{3i}; Z_3^{i-1}|Y_1^{i-1}Y_2^{i-1}Y_3^{i-1}Z_1^{i-1}F^{i-1}W\right) \tag{C.272}$$

For the latter equality we used (C.263) and also that for a similar reason

$$\sum_{i=1}^{n} I\left(X_{1i}; Z_3^{i-1}|Y_1^{i-1}Y_2^{i-1}Z_1^{i-1}F^{i-1}W\right) = 0. \tag{C.273}$$

Combining the two equalities gives:

$$0 \le I\left(Z_3^n; Y_2^n | Y_1^n Y_3^n Z_1^n F^n W\right) = \sum_{i=1}^{n} \delta_3 (1-\delta_{3E})\, I\left(X_{3i}; Y_2^{i-1} | Y_1^{i-1} Y_3^{i-1} Z_3^{i-1} Z_1^{i-1} F^{i-1} W\right)$$
$$- (1-\delta_1\delta_{1E})\, I\left(X_{1i}; Z_3^{i-1} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} F^{i-1} W\right) - (1-\delta_3)\, I\left(X_{3i}; Z_3^{i-1} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} F^{i-1} W\right)$$
$$+ (1-\delta_3)\, I\left(X_{3i}; Z_3^{i-1} | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} Z_1^{i-1} F^{i-1} W\right) \tag{C.274}$$

Also,

$$0 \le H\left(Y_1^n | Y_2^n F^n W\right) = \sum_{i=1}^{n} (1-\delta_1)\, H\left(X_{1i} | Y_1^{i-1} Y_2^{i-1} F^{i-1} W\right) - (1-\delta_2)\, I\left(X_{2i}; Y_1^{i-1} | Y_2^{i-1} F^{i-1} W\right)$$
$$= \sum_{i=1}^{n} (1-\delta_1)\, H\left(X_{1i} | Y_1^{i-1} Y_2^{i-1} Z_1^{i-1} F^{i-1} W\right) + (1-\delta_1)\, I\left(X_{1i}; Z_1^{i-1} | Y_1^{i-1} Y_2^{i-1} F^{i-1} W\right)$$
$$- (1-\delta_2)\, I\left(X_{2i}; Y_1^{i-1} | Y_2^{i-1} F^{i-1} W\right) \tag{C.275}$$

Finally,

$$\sum_{i=1}^{n} H\left(X_{1i} | Y_1^{i-1} Y_2^{i-1} Z_1^{i-1} F^{i-1} W\right) = \sum_{i=1}^{n} H\left(X_{1i} | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} Z_1^{i-1} Z_3^{i-1} F^{i-1} W\right)$$
$$= \sum_{i=1}^{n} H\left(X_{1i} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} Z_3^{i-1} F^{i-1} W\right) - I\left(X_{1i}; Y_2^{i-1} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} Z_3^{i-1} F^{i-1} W\right) \tag{C.276}$$

C.4.6 Trivial constraints

$$\sum_{i=1}^{n} H\left(X_{1i} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} Z_3^{i-1} F^{i-1} W\right) = \sum_{i=1}^{n} H\left(X_{1i} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} F^{i-1} W\right)$$
$$- I\left(X_{1i}; Z_3^{i-1} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} F^{i-1} W\right) \tag{C.277}$$
$$\sum_{i=1}^{n} H\left(X_{3i} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} Z_3^{i-1} F^{i-1} W\right) = \sum_{i=1}^{n} H\left(X_{3i} | Y_1^{i-1} Y_3^{i-1} Z_3^{i-1} F^{i-1} W\right)$$
$$- I\left(X_{3i}; Z_1^{i-1} | Y_1^{i-1} Y_3^{i-1} Z_3^{i-1} F^{i-1} W\right) \tag{C.278}$$
$$\sum_{i=1}^{n} H\left(X_{1i} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} Z_3^{i-1} F^{i-1} W\right) = \sum_{i=1}^{n} H\left(X_{1i} | Y_1^{i-1} Y_3^{i-1} Z_3^{i-1} F^{i-1} W\right)$$
$$- I\left(X_{1i}; Z_1^{i-1} | Y_1^{i-1} Y_3^{i-1} Z_3^{i-1} F^{i-1} W\right) \tag{C.279}$$
$$\sum_{i=1}^{n} H\left(X_{3i} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} Z_3^{i-1} F^{i-1} W\right) = \sum_{i=1}^{n} H\left(X_{3i} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} F^{i-1} W\right)$$
$$- I\left(X_{3i}; Z_3^{i-1} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} F^{i-1} W\right) \tag{C.280}$$
$$\sum_{i=1}^{n} H\left(X_{3i} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} Z_3^{i-1} F^{i-1} W\right) \ge \sum_{i=1}^{n} I\left(X_{3i}; Y_2^{i-1} | Y_1^{i-1} Y_3^{i-1} Z_1^{i-1} Z_3^{i-1} F^{i-1} W\right) \tag{C.281}$$

C.4.7 Equivalence of LPs

We apply a series of transformations on the outer bound LP until we get back the LP in Theorem 4.4. The possible steps are as summarized in Section 4.3.4. Here again we work with the asymptotic form of the inequalities and use terms as names of non-negative variables. We ease notation by omitting $\frac{1}{n}\sum_{i=1}^{n}$, F^{i-1} (which appears in all terms), and index i. Summary of the outer bound program:

$$R \le (1-\delta_1)\, I(X_1; W|Y_1Y_2) + (1-\delta_2)\, I(X_2; W|Y_1Y_2)$$
$$R \le (1-\delta_1\delta_{1E})\, I(X_1; W|Y_1Y_2Z_1) + (1-\delta_2)\, I(X_2; W|Y_1Y_2Z_1)$$
$$R \le (1-\delta_2\delta_{2E})\, I(X_2; W|Y_1Y_2Z_2) + (1-\delta_1)\, I(X_1; W|Y_1Y_2Z_2)$$
$$R \le (1-\delta_1)\, I(X_1; W|Y_1Y_3) + (1-\delta_3)\, I(X_3; W|Y_1Y_3)$$
$$R \le (1-\delta_1\delta_{1E})\, I(X_1; W|Y_1Y_3Z_1) + (1-\delta_3)\, I(X_3; W|Y_1Y_3Z_1)$$
$$R \le (1-\delta_3\delta_{3E})\, I(X_3; W|Y_1Y_3Z_3) + (1-\delta_1)\, I(X_1; W|Y_1Y_3Z_3)$$
$$0 \le -(1-\delta_{1E})\, I(Y_1Y_2; X_1|Z_1W) + \delta_{1E}(1-\delta_1)\, H(X_1|Y_1Y_2Z_1W) + (1-\delta_2)\, H(X_2|Y_1Y_2Z_1W) \tag{C.282}$$
$$0 \le -(1-\delta_{2E})\, I(Y_1Y_2; X_2|Z_2W) + \delta_{2E}(1-\delta_2)\, H(X_2|Y_1Y_2Z_2W) + (1-\delta_1)\, H(X_1|Y_1Y_2Z_2W) \tag{C.283}$$
$$0 \le -(1-\delta_{3E})\, I(Y_1Y_3; X_3|Z_3W) + \delta_{3E}(1-\delta_3)\, H(X_3|Y_1Y_3Z_3W) + (1-\delta_1)\, H(X_1|Y_1Y_3Z_3W) \tag{C.284}$$
$$0 \le -(1-\delta_{1E})\, I(Y_1Y_3; X_1|Z_1W) + \delta_{1E}(1-\delta_1)\, H(X_1|Y_1Y_3Z_1W) + (1-\delta_3)\, H(X_3|Y_1Y_3Z_1W)$$
$$I(X_1; Y_1Y_2|Z_1W) \ge I(X_1; W|Y_1Y_2Z_1) \tag{C.285}$$
$$I(X_2; Y_1Y_2|Z_2W) \ge I(X_2; W|Y_1Y_2Z_2) \tag{C.286}$$
$$I(X_1; Y_1Y_3|Z_1W) \ge I(X_1; W|Y_1Y_3Z_1) \tag{C.287}$$
$$I(X_3; Y_1Y_3|Z_3W) \ge I(X_3; W|Y_1Y_3Z_3) \tag{C.288}$$
$$(1-\delta_{1E})\, I(X_1; Y_1|Z_1W) \le \delta_{1E}(1-\delta_1)\, H(X_1|Y_1Y_3Z_1W) + \delta_{1E}(1-\delta_1)\, I(X_1; Y_3|Y_1Z_1W) \tag{C.289}$$
$$(1-\delta_{3E})\, I(X_3; Y_3|Z_3W) \le \delta_{3E}(1-\delta_3)\, H(X_3|Y_1Y_3Z_3W) + \delta_{3E}(1-\delta_3)\, I(X_3; Y_1|Y_3Z_3W)$$
$$I(X_1; Y_1Y_3|Z_1W) = I(X_1; Y_1|Z_1W) + I(X_1; Y_3|Y_1Z_1W) \tag{C.290}$$
$$I(X_3; Y_1Y_3|Z_3W) = I(X_3; Y_3|Z_3W) + I(X_3; Y_1|Y_3Z_3W)$$
$$0 \le -(1-\delta_1\delta_{1E})\, I(X_1; Y_2|Y_1Y_3Z_1Z_3W) - (1-\delta_3\delta_{3E})\, I(X_3; Y_2|Y_1Y_3Z_1Z_3W) + (1-\delta_2)\, H(X_2|Y_1Y_2Z_1W)$$
$$0 \le -(1-\delta_3\delta_{3E})\, I(X_3; Y_1|Y_3Z_3W) - (1-\delta_3\delta_{3E})\, I(X_3; Z_1|Y_1Y_3Z_3W) + (1-\delta_2)\, I(X_2; Y_1|Y_2W) + (1-\delta_2)\, I(X_2; Z_1|Y_1Y_2W) + (1-\delta_1\delta_{1E})\, I(X_1; Y_2|Y_1Y_3Z_1Z_3W)$$
$$0 \le -(1-\delta_1\delta_{1E})\, I(X_1; Y_3|Y_1Z_1W) + (1-\delta_3)\, I(X_3; Y_2|Y_3Y_1Z_1Z_3W) + (1-\delta_3)\, I(X_3; Z_3|Y_3Y_1Z_1W) - (1-\delta_3)\, I(X_3; Z_3|Y_2Y_3Y_1Z_1W)$$
$$0 \le \delta_3(1-\delta_{3E})\, I(X_3; Y_2|Y_1Y_3Z_3Z_1W) - (1-\delta_1\delta_{1E})\, I(X_1; Z_3|Y_1Y_3Z_1W)$$

$$
\begin{aligned}
&-(1-\delta_3)\, I(X_3;Z_3|Y_1Y_3Z_1W)+(1-\delta_3)\, I(X_3;Z_3|Y_1Y_2Y_3Z_1W)\\
&0\le(1-\delta_1)\, H(X_1|Y_1Y_2Z_1W)+(1-\delta_1)\, I(X_1;Z_1|Y_1Y_2W)-(1-\delta_2)\, I(X_2;Y_1|Y_2W)\\
&H(X_1|Y_1Y_2Z_1W)=H(X_1|Y_1Y_3Z_1Z_3W)-I(X_1;Y_2|Y_1Y_3Z_1Z_3W)\\
&H(X_1|Y_1Y_3Z_1Z_3W)=H(X_1|Y_1Y_3Z_1W)-I(X_1;Z_3|Y_1Y_3Z_1W) && \text{(C.291)}\\
&H(X_3|Y_1Y_3Z_1Z_3W)=H(X_3|Y_1Y_3Z_3W)-I(X_3;Z_1|Y_1Y_3Z_3W) && \text{(C.292)}\\
&H(X_1|Y_1Y_3Z_1Z_3W)=H(X_1|Y_1Y_3Z_3W)-I(X_1;Z_1|Y_1Y_3Z_3W) && \text{(C.293)}\\
&H(X_3|Y_1Y_3Z_1Z_3W)=H(X_3|Y_1Y_3Z_1W)-I(X_3;Z_3|Y_1Y_3Z_1W) && \text{(C.294)}\\
&H(X_3|Y_1Y_3Z_1Z_3W)\ge I(X_3;Y_2|Y_1Y_3Z_1Z_3W)\\
&1\ge H(X_1|Y_1Y_2Z_2W)+I(X_1;W|Y_1Y_2Z_2)\\
&1\ge H(X_2|Y_1Y_2Z_1W)+I(X_2;W|Y_1Y_2Z_1)\\
&1\ge H(X_1|Y_1Y_2Z_1W)+I(X_1;W|Y_1Y_2)+I(X_1;Z_1|Y_1Y_2W)\\
&1\ge H(X_2|Y_1Y_2Z_1W)+I(X_2;W|Y_1Y_2)+I(X_2;Z_1|Y_1Y_2W)\\
&1\ge H(X_1|Y_1Y_2Z_2W)+I(X_1;W|Y_1Y_2)\\
&1\ge H(X_2|Y_1Y_2Z_2W)+I(X_2;W|Y_1Y_2)\\
&1\ge H(X_1|Y_1Y_3Z_3W)+I(X_1;W|Y_1Y_3Z_3)\\
&1\ge H(X_3|Y_1Y_3Z_1W)+I(X_3;W|Y_1Y_3Z_1)\\
&1\ge H(X_1|Y_1Y_3Z_1W)+I(X_1;W|Y_1Y_3)\\
&1\ge H(X_3|Y_1Y_3Z_1W)+I(X_3;W|Y_1Y_3)\\
&1\ge H(X_1|Y_1Y_3Z_3W)+I(X_1;W|Y_1Y_3)\\
&1\ge H(X_3|Y_1Y_3Z_3W)+I(X_3;W|Y_1Y_3)
\end{aligned}
$$

We apply (C.285)-(C.286) in (C.282)-(C.283) and keep the resulting inequalities while dropping (C.285)-(C.286) and (C.282)-(C.283).

We show that (C.287)-(C.288) both can be made equalities without reducing the value of the program. If (C.287) is not equality, we apply the following transform (for some $\Delta>0$):

$$
\begin{aligned}
I(X_1;Y_1Y_3|Z_1W)&\downarrow\Delta && \text{(C.295)}\\
I(X_1;Y_1|Z_1W)&\downarrow\Delta. && \text{(C.296)}
\end{aligned}
$$

This transform either makes (C.287) tight or $I(X_1;Y_1|Z_1W)$ becomes 0. No constraints are violated. If the latter happens, we can do the following transform:

$$
\begin{aligned}
I(X_1;Y_1Y_3|Z_1W)&\downarrow\Delta && \text{(C.297)}\\
I(X_1;Y_3|Y_1Z_1W)&\downarrow\Delta. && \text{(C.298)}
\end{aligned}
$$

Note that (C.289) cannot be violated, since the RHS is already 0. Eventually, this transform makes (C.287) tight. A similar argument shows that also (C.288) can be made tight, which allows to eliminate the variables $I(X_1;Y_1Y_3|Z_1W)$ and $I(X_3;Y_1Y_3|Z_3W)$.

Using equalities (C.291)-(C.294) we eliminate $H(X_1|Y_1Y_3Z_1W)$, $H(X_3|Y_1Y_3Z_3W)$, $H(X_1|Y_1Y_3Z_3W)$ and $H(X_3|Y_1Y_3Z_1W)$.

Further, we introduce the following naming:

$$H(X_1|Y_1Y_2Z_1W) \sim k_1 \tag{C.299}$$
$$H(X_2|Y_1Y_2Z_2W) \sim k_2 \tag{C.300}$$
$$I(X_1;W|Y_1Y_3) \sim m_1 \tag{C.301}$$
$$I(X_2;W|Y_1Y_2) \sim m_2 \tag{C.302}$$
$$I(X_3;W|Y_1Y_3) \sim m_3 \tag{C.303}$$
$$I(X_1;Y_2|Y_1Y_3Z_1Z_3W) \sim \bar{s}_1 \tag{C.304}$$
$$I(X_3;Y_2|Y_1Y_3Z_1Z_3W) \sim \bar{s}_3 \tag{C.305}$$
$$I(X_1;Z_3|Y_1Y_3Z_1W) \sim \bar{\ell}_1 \tag{C.306}$$
$$I(X_3;Z_1|Y_1Y_3Z_3W) \sim \bar{\ell}_3 \tag{C.307}$$
$$H(X_3|Y_1Y_3Z_1Z_3W) \sim k_3 + \bar{s}_3 \tag{C.308}$$
$$I(X_1;W|Y_1Y_2Z_1) \sim \bar{x}_{121} \tag{C.309}$$
$$I(X_1;W|Y_1Y_3Z_1) \sim \bar{x}_{131} \tag{C.310}$$
$$I(X_3;W|Y_1Y_3Z_3) \sim \bar{x}_{313} \tag{C.311}$$
$$I(X_2;W|Y_1Y_2Z_2) \sim \bar{x}_{22} \tag{C.312}$$
$$I(X_3;Z_3|Y_1Y_3Z_1W) \sim \bar{z}_3 \tag{C.313}$$

We distinguish the variable names used only in the outer bound program with overscore. We get the following:

$$\begin{aligned}
&R \le (1-\delta_2)\, m_2 + (1-\delta_1)\, I(X_1;W|Y_1Y_2)\\
&R \le (1-\delta_1\delta_{1E})\, \bar{x}_{121} + (1-\delta_2)\, I(X_2;W|Y_1Y_2Z_1)\\
&R \le (1-\delta_2\delta_{2E})\, \bar{x}_{22} + (1-\delta_1)\, I(X_1;W|Y_1Y_2Z_2)\\
&R \le (1-\delta_1)\, m_1 + (1-\delta_3)\, m_3\\
&R \le (1-\delta_1\delta_{1E})\, \bar{x}_{131} + (1-\delta_3)\, I(X_3;W|Y_1Y_3Z_1)\\
&R \le (1-\delta_3\delta_{3E})\, \bar{x}_{313} + (1-\delta_1)\, I(X_1;W|Y_1Y_3Z_3)\\
&0 \le \delta_{1E}(1-\delta_1)\, k_1 - (1-\delta_{1E})\, \bar{x}_{121} + (1-\delta_2)\, H(X_2|Y_1Y_2Z_1W)\\
&0 \le \delta_{2E}(1-\delta_2)\, k_2 - (1-\delta_{2E})\, \bar{x}_{22} + (1-\delta_1)\, H(X_1|Y_1Y_2Z_2W)\\
&0 \le \delta_{3E}(1-\delta_3)\, \bar{\ell}_3 + \delta_{3E}(1-\delta_3)\, k_3 + \delta_{3E}(1-\delta_3)\, \bar{s}_3 - (1-\delta_{3E})\, \bar{x}_{313}\\
&\quad + (1-\delta_1)\, H(X_1|Y_1Y_3Z_1Z_3W) + (1-\delta_1)\, I(X_1;Z_1|Y_1Y_3Z_3W)\\
&0 \le \delta_{1E}(1-\delta_1)\, \bar{\ell}_1 + (1-\delta_3)\, k_3 + (1-\delta_3)\, \bar{s}_3 - (1-\delta_{1E})\, \bar{x}_{131}\\
&\quad + (1-\delta_3)\, \bar{z}_3 + \delta_{1E}(1-\delta_1)\, H(X_1|Y_1Y_3Z_1Z_3W)\\
&(1-\delta_{1E})\, I(X_1;Y_1|Z_1W) \le \delta_{1E}(1-\delta_1)\, \bar{\ell}_1 + \delta_{1E}(1-\delta_1)\, H(X_1|Y_1Y_3Z_1Z_3W)\\
&\quad + \delta_{1E}(1-\delta_1)\, I(X_1;Y_3|Y_1Z_1W)
\end{aligned}$$

$$
\begin{aligned}
&(1-\delta_{3E})\, I(X_3; Y_3|Z_3W) \le \delta_{3E}(1-\delta_3)\,\bar{\ell}_3 + \delta_{3E}(1-\delta_3)\, k_3 + \delta_{3E}(1-\delta_3)\,\bar{s}_3 \\
&\quad + \delta_{3E}(1-\delta_3)\, I(X_3; Y_1|Y_3Z_3W) \\
&\bar{x}_{131} = I(X_1; Y_1|Z_1W) + I(X_1; Y_3|Y_1Z_1W) \qquad \text{(C.314)} \\
&\bar{x}_{313} = I(X_3; Y_3|Z_3W) + I(X_3; Y_1|Y_3Z_3W) \qquad \text{(C.315)} \\
&(1-\delta_1\delta_{1E})\,\bar{s}_1 + (1-\delta_3\delta_{3E})\,\bar{s}_3 \le (1-\delta_2)\, H(X_2|Y_1Y_2Z_1W) \\
&(1-\delta_3\delta_{3E})\,\bar{\ell}_3 \le (1-\delta_1\delta_{1E})\,\bar{s}_1 - (1-\delta_3\delta_{3E})\, I(X_3; Y_1|Y_3Z_3W) \\
&\quad + (1-\delta_2)\, I(X_2; Y_1|Y_2W) + (1-\delta_2)\, I(X_2; Z_1|Y_1Y_2W) \\
&0 \le (1-\delta_3)\,\bar{s}_3 - (1-\delta_1\delta_{1E})\, I(X_1; Y_3|Y_1Z_1W) \\
&\quad + (1-\delta_3)\, I(X_3; Z_3|Y_3Y_1Z_1W) - (1-\delta_3)\, I(X_3; Z_3|Y_2Y_3Y_1Z_1W) \\
&(1-\delta_1\delta_{1E})\,\bar{\ell}_1 \le \delta_3(1-\delta_{3E})\,\bar{s}_3 - (1-\delta_3)\,\bar{z}_3 \\
&\quad + (1-\delta_3)\, I(X_3; Z_3|Y_1Y_2Y_3Z_1W) \\
&0 \le (1-\delta_1)\, k_1 + (1-\delta_1)\, I(X_1; Z_1|Y_1Y_2W) - (1-\delta_2)\, I(X_2; Y_1|Y_2W) \\
&\bar{s}_1 + k_1 = H(X_1|Y_1Y_3Z_1Z_3W) \qquad \text{(C.316)} \\
&1 \ge H(X_1|Y_1Y_2Z_2W) + I(X_1; W|Y_1Y_2Z_2) \\
&1 \ge H(X_2|Y_1Y_2Z_1W) + I(X_2; W|Y_1Y_2Z_1) \\
&1 \ge k_1 + I(X_1; W|Y_1Y_2) + I(X_1; Z_1|Y_1Y_2W) \\
&1 \ge m_2 + H(X_2|Y_1Y_2Z_1W) + I(X_2; Z_1|Y_1Y_2W) \\
&1 \ge H(X_1|Y_1Y_2Z_2W) + I(X_1; W|Y_1Y_2) \\
&1 \ge m_2 + k_2 \\
&1 \ge H(X_1|Y_1Y_3Z_1Z_3W) + I(X_1; Z_1|Y_1Y_3Z_3W) + I(X_1; W|Y_1Y_3Z_3) \\
&1 \ge k_3 + \bar{s}_3 + \bar{z}_3 + I(X_3; W|Y_1Y_3Z_1) \\
&1 \ge m_1 + \bar{\ell}_1 + H(X_1|Y_1Y_3Z_1Z_3W) \\
&1 \ge m_3 + k_3 + \bar{s}_3 + \bar{z}_3 \\
&1 \ge m_1 + H(X_1|Y_1Y_3Z_1Z_3W) + I(X_1; Z_1|Y_1Y_3Z_3W) \\
&1 \ge m_3 + \bar{\ell}_3 + k_3 + \bar{s}_3
\end{aligned}
$$

We use eqalities (C.314)-(C.315) and (C.316) to eliminate variables $I(X_1; Y_1|Z_1W)$, $I(X_3; Y_3|Z_3W)$ and $H(X_1|Y_1Y_3Z_1Z_3W)$.

$$
\begin{aligned}
&R \le (1-\delta_2)\, m_2 + (1-\delta_1)\, I(X_1; W|Y_1Y_2) \\
&R \le (1-\delta_1\delta_{1E})\,\bar{x}_{121} + (1-\delta_2)\, I(X_2; W|Y_1Y_2Z_1) \\
&R \le (1-\delta_2\delta_{2E})\,\bar{x}_{22} + (1-\delta_1)\, I(X_1; W|Y_1Y_2Z_2) \\
&R \le (1-\delta_1)\, m_1 + (1-\delta_3)\, m_3 \\
&R \le (1-\delta_1\delta_{1E})\,\bar{x}_{131} + (1-\delta_3)\, I(X_3; W|Y_1Y_3Z_1) \\
&R \le (1-\delta_3\delta_{3E})\,\bar{x}_{313} + (1-\delta_1)\, I(X_1; W|Y_1Y_3Z_3) \\
&(1-\delta_{1E})\,\bar{x}_{121} \le \delta_{1E}(1-\delta_1)\, k_1 + (1-\delta_2)\, H(X_2|Y_1Y_2Z_1W)
\end{aligned}
$$

$$
\begin{aligned}
&(1-\delta_{2E})\,\bar{x}_{22} \le \delta_{2E}\,(1-\delta_2)\,k_2 + (1-\delta_1)\,H(X_1|Y_1Y_2Z_2W)\\
&(1-\delta_{3E})\,\bar{x}_{313} \le \delta_{3E}\,(1-\delta_3)\,\bar{\ell}_3 + \delta_{3E}\,(1-\delta_3)\,k_3 + \delta_{3E}\,(1-\delta_3)\,\bar{s}_3 + (1-\delta_1)\,\bar{s}_1 + (1-\delta_1)\,k_1\\
&\quad + (1-\delta_1)\,I(X_1;Z_1|Y_1Y_3Z_3W)\\
&(1-\delta_{1E})\,\bar{x}_{131} \le \delta_{1E}\,(1-\delta_1)\,\bar{\ell}_1 + \delta_{1E}\,(1-\delta_1)\,\bar{s}_1 + \delta_{1E}\,(1-\delta_1)\,k_1 + (1-\delta_3)\,k_3 + (1-\delta_3)\,\bar{s}_3 + (1-\delta_3)\,\bar{z}_3\\
&(1-\delta_{1E})\,\bar{x}_{131} \le \delta_{1E}\,(1-\delta_1)\,\bar{\ell}_1 + \delta_{1E}\,(1-\delta_1)\,\bar{s}_1 + \delta_{1E}\,(1-\delta_1)\,k_1 + (1-\delta_1\delta_{1E})\,I(X_1;Y_3|Y_1Z_1W)\\
&(1-\delta_{3E})\,\bar{x}_{313} \le \delta_{3E}\,(1-\delta_3)\,\bar{\ell}_3 + \delta_{3E}\,(1-\delta_3)\,k_3 + \delta_{3E}\,(1-\delta_3)\,\bar{s}_3 + (1-\delta_3\delta_{3E})\,I(X_3;Y_1|Y_3Z_3W)
\end{aligned}
\tag{C.317}
$$

$$
\begin{aligned}
&(1-\delta_1\delta_{1E})\,\bar{s}_1 + (1-\delta_3\delta_{3E})\,\bar{s}_3 \le (1-\delta_2)\,H(X_2|Y_1Y_2Z_1W)\\
&(1-\delta_3\delta_{3E})\,\bar{\ell}_3 \le (1-\delta_1\delta_{1E})\,\bar{s}_1 - (1-\delta_3\delta_{3E})\,I(X_3;Y_1|Y_3Z_3W)\\
&\quad + (1-\delta_2)\,I(X_2;Y_1|Y_2W) + (1-\delta_2)\,I(X_2;Z_1|Y_1Y_2W)
\end{aligned}
\tag{C.318}
$$

$$
\begin{aligned}
&0 \le (1-\delta_3)\,\bar{s}_3 - (1-\delta_1\delta_{1E})\,I(X_1;Y_3|Y_1Z_1W)\\
&\quad + (1-\delta_3)\,I(X_3;Z_3|Y_3Y_1Z_1W) - (1-\delta_3)\,I(X_3;Z_3|Y_2Y_3Y_1Z_1W)\\
&(1-\delta_1\delta_{1E})\,\bar{\ell}_1 \le \delta_3\,(1-\delta_{3E})\,\bar{s}_3 - (1-\delta_3)\,\bar{z}_3\\
&\quad + (1-\delta_3)\,I(X_3;Z_3|Y_1Y_2Y_3Z_1W)\\
&0 \le (1-\delta_1)\,k_1 + (1-\delta_1)\,I(X_1;Z_1|Y_1Y_2W) - (1-\delta_2)\,I(X_2;Y_1|Y_2W)\\
&1 \ge H(X_1|Y_1Y_2Z_2W) + I(X_1;W|Y_1Y_2Z_2)\\
&1 \ge k_1 + I(X_1;W|Y_1Y_2) + I(X_1;Z_1|Y_1Y_2W)\\
&1 \ge H(X_1|Y_1Y_2Z_2W) + I(X_1;W|Y_1Y_2)\\
&1 \ge \bar{s}_1 + k_1 + I(X_1;Z_1|Y_1Y_3Z_3W) + I(X_1;W|Y_1Y_3Z_3)\\
&1 \ge m_1 + \bar{\ell}_1 + \bar{s}_1 + k_1\\
&1 \ge m_1 + \bar{s}_1 + k_1 + I(X_1;Z_1|Y_1Y_3Z_3W)\\
&1 \ge m_2 + H(X_2|Y_1Y_2Z_1W) + I(X_2;Z_1|Y_1Y_2W)\\
&1 \ge m_2 + k_2\\
&1 \ge H(X_2|Y_1Y_2Z_1W) + I(X_2;W|Y_1Y_2Z_1)\\
&1 \ge m_3 + k_3 + \bar{s}_3 + \bar{z}_3\\
&1 \ge k_3 + \bar{s}_3 + \bar{z}_3 + I(X_3;W|Y_1Y_3Z_1)\\
&1 \ge m_3 + \bar{\ell}_3 + k_3 + \bar{s}_3
\end{aligned}
$$

We apply (C.318) in (C.317) and keep only the derived inequality.

$$
\begin{aligned}
&R \le (1-\delta_2)\,m_2 + (1-\delta_1)\,I(X_1;W|Y_1Y_2)\\
&R \le (1-\delta_1\delta_{1E})\,\bar{x}_{121} + (1-\delta_2)\,I(X_2;W|Y_1Y_2Z_1)\\
&R \le (1-\delta_2\delta_{2E})\,\bar{x}_{22} + (1-\delta_1)\,I(X_1;W|Y_1Y_2Z_2)\\
&R \le (1-\delta_1)\,m_1 + (1-\delta_3)\,m_3\\
&R \le (1-\delta_1\delta_{1E})\,\bar{x}_{131} + (1-\delta_3)\,I(X_3;W|Y_1Y_3Z_1)\\
&R \le (1-\delta_3\delta_{3E})\,\bar{x}_{313} + (1-\delta_1)\,I(X_1;W|Y_1Y_3Z_3)
\end{aligned}
$$

$$(1-\delta_{1E})\,\bar{x}_{121} \le \delta_{1E}(1-\delta_1)\,k_1 + (1-\delta_2)\,H(X_2|Y_1Y_2Z_1W)$$
$$(1-\delta_{2E})\,\bar{x}_{22} \le \delta_{2E}(1-\delta_2)\,k_2 + (1-\delta_1)\,H(X_1|Y_1Y_2Z_2W)$$
$$(1-\delta_{3E})\,\bar{x}_{313} \le \delta_{3E}(1-\delta_3)\,\bar{\ell}_3 + \delta_{3E}(1-\delta_3)\,k_3 + \delta_{3E}(1-\delta_3)\,\bar{s}_3 + (1-\delta_1)\,\bar{s}_1 + (1-\delta_1)\,k_1 + (1-\delta_1)\,I(X_1;Z_1|Y_1Y_3Z_3W)$$
$$(1-\delta_{1E})\,\bar{x}_{131} \le \delta_{1E}(1-\delta_1)\,\bar{\ell}_1 + \delta_{1E}(1-\delta_1)\,\bar{s}_1 + \delta_{1E}(1-\delta_1)\,k_1 + (1-\delta_3)\,k_3 + (1-\delta_3)\,\bar{s}_3 + (1-\delta_3)\,\bar{z}_3$$
$$(1-\delta_{1E})\,\bar{x}_{131} \le \delta_{1E}(1-\delta_1)\,\bar{\ell}_1 + \delta_{1E}(1-\delta_1)\,\bar{s}_1 + \delta_{1E}(1-\delta_1)\,k_1 + (1-\delta_1\delta_{1E})\,I(X_1;Y_3|Y_1Z_1W)$$
$$(1-\delta_{3E})\,\bar{x}_{313} \le (1-\delta_1\delta_{1E})\,\bar{s}_1 - (1-\delta_{3E})\,\bar{\ell}_3 + \delta_{3E}(1-\delta_3)\,k_3 + \delta_{3E}(1-\delta_3)\,\bar{s}_3 + (1-\delta_2)\,I(X_2;Y_1|Y_2W) + (1-\delta_2)\,I(X_2;Z_1|Y_1Y_2W) \tag{C.319}$$
$$(1-\delta_1\delta_{1E})\,\bar{s}_1 + (1-\delta_3\delta_{3E})\,\bar{s}_3 \le (1-\delta_2)\,H(X_2|Y_1Y_2Z_1W)$$
$$0 \le (1-\delta_3)\,\bar{s}_3 - (1-\delta_1\delta_{1E})\,I(X_1;Y_3|Y_1Z_1W) + (1-\delta_3)\,I(X_3;Z_3|Y_3Y_1Z_1W) - (1-\delta_3)\,I(X_3;Z_3|Y_2Y_3Y_1Z_1W) \tag{C.320}$$
$$(1-\delta_1\delta_{1E})\,\bar{\ell}_1 \le \delta_3(1-\delta_{3E})\,\bar{s}_3 - (1-\delta_3)\,\bar{z}_3 + (1-\delta_3)\,I(X_3;Z_3|Y_1Y_2Y_3Z_1W) \tag{C.321}$$
$$0 \le (1-\delta_1)\,k_1 + (1-\delta_1)\,I(X_1;Z_1|Y_1Y_2W) - (1-\delta_2)\,I(X_2;Y_1|Y_2W) \tag{C.322}$$
$$1 \ge H(X_1|Y_1Y_2Z_2W) + I(X_1;W|Y_1Y_2Z_2)$$
$$1 \ge k_1 + I(X_1;W|Y_1Y_2) + I(X_1;Z_1|Y_1Y_2W)$$
$$1 \ge H(X_1|Y_1Y_2Z_2W) + I(X_1;W|Y_1Y_2)$$
$$1 \ge \bar{s}_1 + k_1 + I(X_1;Z_1|Y_1Y_3Z_3W) + I(X_1;W|Y_1Y_3Z_3)$$
$$1 \ge m_1 + \bar{\ell}_1 + \bar{s}_1 + k_1$$
$$1 \ge m_1 + \bar{s}_1 + k_1 + I(X_1;Z_1|Y_1Y_3Z_3W)$$
$$1 \ge m_2 + H(X_2|Y_1Y_2Z_1W) + I(X_2;Z_1|Y_1Y_2W)$$
$$1 \ge m_2 + k_2$$
$$1 \ge H(X_2|Y_1Y_2Z_1W) + I(X_2;W|Y_1Y_2Z_1)$$
$$1 \ge m_3 + k_3 + \bar{s}_3 + \bar{z}_3$$
$$1 \ge k_3 + \bar{s}_3 + \bar{z}_3 + I(X_3;W|Y_1Y_3Z_1)$$
$$1 \ge m_3 + \bar{\ell}_3 + k_3 + \bar{s}_3$$

We observe that (C.322) can be made tight by increasing $I(X_2;Y_1|Y_2W)$. We apply the resulting equality in (C.319). At the same time we eliminate the variable $I(X_3;Z_3|Y_2Y_3Y_1Z_1W)$ from (C.320)-(C.321).

$$R \le (1-\delta_2)\,m_2 + (1-\delta_1)\,I(X_1;W|Y_1Y_2)$$
$$R \le (1-\delta_1\delta_{1E})\,\bar{x}_{121} + (1-\delta_2)\,I(X_2;W|Y_1Y_2Z_1)$$
$$R \le (1-\delta_2\delta_{2E})\,\bar{x}_{22} + (1-\delta_1)\,I(X_1;W|Y_1Y_2Z_2)$$
$$R \le (1-\delta_1)\,m_1 + (1-\delta_3)\,m_3$$
$$R \le (1-\delta_1\delta_{1E})\,\bar{x}_{131} + (1-\delta_3)\,I(X_3;W|Y_1Y_3Z_1)$$

$$
\begin{aligned}
&R \le (1-\delta_3\delta_{3E})\,\bar{x}_{313} + (1-\delta_1)\,I(X_1;W|Y_1Y_3Z_3)\\
&(1-\delta_{1E})\,\bar{x}_{121} \le \delta_{1E}(1-\delta_1)\,k_1 + (1-\delta_2)\,H(X_2|Y_1Y_2Z_1W)\\
&(1-\delta_{2E})\,\bar{x}_{22} \le \delta_{2E}(1-\delta_2)\,k_2 + (1-\delta_1)\,H(X_1|Y_1Y_2Z_2W)\\
&(1-\delta_{3E})\,\bar{x}_{313} \le \delta_{3E}(1-\delta_3)\,\bar{\ell}_3 + \delta_{3E}(1-\delta_3)\,k_3 + \delta_{3E}(1-\delta_3)\,\bar{s}_3 + (1-\delta_1)\,\bar{s}_1 + (1-\delta_1)\,k_1\\
&\quad + (1-\delta_1)\,I(X_1;Z_1|Y_1Y_3Z_3W)\\
&(1-\delta_{1E})\,\bar{x}_{131} \le \delta_{1E}(1-\delta_1)\,\bar{\ell}_1 + \delta_{1E}(1-\delta_1)\,\bar{s}_1 + \delta_{1E}(1-\delta_1)\,k_1 + (1-\delta_3)\,k_3 + (1-\delta_3)\,\bar{s}_3 + (1-\delta_3)\,\bar{z}_3
\end{aligned}
\tag{C.323}
$$

$$
\begin{aligned}
&(1-\delta_{1E})\,\bar{x}_{131} \le \delta_{1E}(1-\delta_1)\,\bar{\ell}_1 + \delta_{1E}(1-\delta_1)\,\bar{s}_1 + \delta_{1E}(1-\delta_1)\,k_1 + (1-\delta_1\delta_{1E})\,I(X_1;Y_3|Y_1Z_1W)\\
&(1-\delta_{3E})\,\bar{x}_{313} \le (1-\delta_1\delta_{1E})\,\bar{s}_1 - (1-\delta_{3E})\,\bar{\ell}_3 + (1-\delta_1)\,k_1 + \delta_{3E}(1-\delta_3)\,k_3 + \delta_{3E}(1-\delta_3)\,\bar{s}_3\\
&\quad + (1-\delta_1)\,I(X_1;Z_1|Y_1Y_2W) + (1-\delta_2)\,I(X_2;Z_1|Y_1Y_2W)\\
&(1-\delta_1\delta_{1E})\,\bar{s}_1 + (1-\delta_3\delta_{3E})\,\bar{s}_3 \le (1-\delta_2)\,H(X_2|Y_1Y_2Z_1W)\\
&(1-\delta_1\delta_{1E})\,I(X_1;Y_3|Y_1Z_1W) \le (1-\delta_3)\,\bar{s}_3 + (1-\delta_3)\,I(X_3;Z_3|Y_3Y_1Z_1W)\\
&(1-\delta_1\delta_{1E})\,I(X_1;Y_3|Y_1Z_1W) \le (1-\delta_3\delta_{3E})\,\bar{s}_3 - (1-\delta_1\delta_{1E})\,\bar{\ell}_1\\
&1 \ge H(X_1|Y_1Y_2Z_2W) + I(X_1;W|Y_1Y_2Z_2)\\
&1 \ge k_1 + I(X_1;W|Y_1Y_2) + I(X_1;Z_1|Y_1Y_2W)\\
&1 \ge H(X_1|Y_1Y_2Z_2W) + I(X_1;W|Y_1Y_2)\\
&1 \ge \bar{s}_1 + k_1 + I(X_1;Z_1|Y_1Y_3Z_3W) + I(X_1;W|Y_1Y_3Z_3)\\
&1 \ge m_1 + \bar{\ell}_1 + \bar{s}_1 + k_1\\
&1 \ge m_1 + \bar{s}_1 + k_1 + I(X_1;Z_1|Y_1Y_3Z_3W)\\
&1 \ge m_2 + H(X_2|Y_1Y_2Z_1W) + I(X_2;Z_1|Y_1Y_2W)\\
&1 \ge m_2 + k_2\\
&1 \ge H(X_2|Y_1Y_2Z_1W) + I(X_2;W|Y_1Y_2Z_1)\\
&1 \ge m_3 + k_3 + \bar{s}_3 + \bar{z}_3\\
&1 \ge k_3 + \bar{s}_3 + \bar{z}_3 + I(X_3;W|Y_1Y_3Z_1)\\
&1 \ge m_3 + \bar{\ell}_3 + k_3 + \bar{s}_3
\end{aligned}
$$

We eliminate $I(X_1;Y_3|Y_1Z_1W)$. (C.323) becomes redundant.

$$
\begin{aligned}
&R \le (1-\delta_2)\,m_2 + (1-\delta_1)\,I(X_1;W|Y_1Y_2)\\
&R \le (1-\delta_1\delta_{1E})\,\bar{x}_{121} + (1-\delta_2)\,I(X_2;W|Y_1Y_2Z_1)\\
&R \le (1-\delta_2\delta_{2E})\,\bar{x}_{22} + (1-\delta_1)\,I(X_1;W|Y_1Y_2Z_2) && \text{(C.324)}\\
&R \le (1-\delta_1)\,m_1 + (1-\delta_3)\,m_3\\
&R \le (1-\delta_1\delta_{1E})\,\bar{x}_{131} + (1-\delta_3)\,I(X_3;W|Y_1Y_3Z_1)\\
&R \le (1-\delta_3\delta_{3E})\,\bar{x}_{313} + (1-\delta_1)\,I(X_1;W|Y_1Y_3Z_3) && \text{(C.325)}\\
&(1-\delta_{1E})\,\bar{x}_{121} \le \delta_{1E}(1-\delta_1)\,k_1 + (1-\delta_2)\,H(X_2|Y_1Y_2Z_1W) && \text{(C.326)}\\
&(1-\delta_{1E})\,\bar{x}_{131} \le -(1-\delta_{1E})\,\bar{\ell}_1 + \delta_{1E}(1-\delta_1)\,\bar{s}_1 + \delta_{1E}(1-\delta_1)\,k_1 + (1-\delta_3\delta_{3E})\,\bar{s}_3 && \text{(C.327)}
\end{aligned}
$$

$$(1-\delta_{1E})\,\bar{x}_{131} \le \delta_{1E}(1-\delta_1)\,\bar{\ell}_1 + \delta_{1E}(1-\delta_1)\,\bar{s}_1 + \delta_{1E}(1-\delta_1)\,k_1 + (1-\delta_3)\,\bar{s}_3 + (1-\delta_3)\,I(X_3;Z_3|Y_3Y_1Z_1W) \tag{C.328}$$

$$(1-\delta_{2E})\,\bar{x}_{22} \le \delta_{2E}(1-\delta_2)\,k_2 + (1-\delta_1)\,H(X_1|Y_1Y_2Z_2W)$$

$$(1-\delta_{3E})\,\bar{x}_{313} \le \delta_{3E}(1-\delta_3)\,\bar{\ell}_3 + (1-\delta_1)\,\bar{s}_1 + (1-\delta_1)\,k_1 + \delta_{3E}(1-\delta_3)\,k_3 + \delta_{3E}(1-\delta_3)\,\bar{s}_3 + (1-\delta_1)\,I(X_1;Z_1|Y_1Y_3Z_3W) \tag{C.329}$$

$$(1-\delta_{3E})\,\bar{x}_{313} \le (1-\delta_1\delta_{1E})\,\bar{s}_1 - (1-\delta_{3E})\,\bar{\ell}_3 + (1-\delta_1)\,k_1 + \delta_{3E}(1-\delta_3)\,k_3 + \delta_{3E}(1-\delta_3)\,\bar{s}_3 + (1-\delta_1)\,I(X_1;Z_1|Y_1Y_2W) + (1-\delta_2)\,I(X_2;Z_1|Y_1Y_2W) \tag{C.330}$$

$$(1-\delta_1\delta_{1E})\,\bar{s}_1 + (1-\delta_3\delta_{3E})\,\bar{s}_3 \le (1-\delta_2)\,H(X_2|Y_1Y_2Z_1W)$$

$$1 \ge H(X_1|Y_1Y_2Z_2W) + I(X_1;W|Y_1Y_2Z_2)$$

$$1 \ge k_1 + I(X_1;W|Y_1Y_2) + I(X_1;Z_1|Y_1Y_2W)$$

$$1 \ge H(X_1|Y_1Y_2Z_2W) + I\left(X_1;W|Y_1^{i-1}Y_2\right) \tag{C.331}$$

$$1 \ge \bar{s}_1 + k_1 + I(X_1;Z_1|Y_1Y_3Z_3W) + I(X_1;W|Y_1Y_3Z_3)$$

$$1 \ge m_1 + \bar{\ell}_1 + \bar{s}_1 + k_1$$

$$1 \ge m_1 + \bar{s}_1 + k_1 + I(X_1;Z_1|Y_1Y_3Z_3W)$$

$$1 \ge m_2 + H(X_2|Y_1Y_2Z_1W) + I(X_2;Z_1|Y_1Y_2W)$$

$$1 \ge m_2 + k_2$$

$$1 \ge H(X_2|Y_1Y_2Z_1W) + I(X_2;W|Y_1Y_2Z_1)$$

$$1 \ge m_3 + k_3 + \bar{s}_3 + \bar{z}_3$$

$$1 \ge k_3 + \bar{s}_3 + \bar{z}_3 + I(X_3;W|Y_1Y_3Z_1)$$

$$1 \ge m_3 + \bar{\ell}_3 + k_3 + \bar{s}_3$$

We show that the we can assume that the RHS of (C.328) can be made smaller than the RHS of (C.327). If this assumption is not true at an optimal point and $\bar{\ell}_1 > 0$ we can decrease $\bar{\ell}_1$ until the assumption becomes true or $\bar{\ell}_1 = 0$ without violating any constraints. If $\bar{\ell}_1 = 0$, then we can decrease the value of $I(X_3;Z_3|Y_3Y_1Z_1W)$ until the assumption holds. Note that we reach equality between the RHS of (C.328) and the RHS of (C.327) at least when $I(X_3;Z_3|Y_3Y_1Z_1W)$ reaches 0. Thus,

$$(1-\delta_1\delta_{1E})\,\bar{\ell}_1 + (1-\delta_3)\,I(X_3;Z_3|Y_3Y_1Z_1W) \le \delta_3(1-\delta_{3E})\,\bar{s}_3 \tag{C.332}$$

can be assumed. Given this, (C.327) can be dropped. We can apply the same argument on (C.329)-(C.330) with reducing first $\bar{\ell}_3$ and then if needed $I(X_1;Z_1|Y_1Y_3Z_3W)$. We thus replace (C.330) with the following inequality:

$$(1-\delta_3\delta_{3E})\,\bar{\ell}_3 + (1-\delta_1)\,I(X_1;Z_1|Y_1Y_3Z_3W) \le \tag{C.333}$$

$$\delta_1(1-\delta_{1E})\,\bar{s}_1 + (1-\delta_1)\,I(X_1;Z_1|Y_1Y_2W) + (1-\delta_2)\,I(X_2;Z_1|Y_1Y_2W)\,. \tag{C.334}$$

Independently of this transform, we observe that $I(X_1;W|Y_1Y_2)$ and $I(X_1;W|Y_1Y_2Z_2)$ can be assumed to be equal. If at an optimal point $I(X_1;W|Y_1Y_2) > I(X_1;W|Y_1Y_2Z_2)$, then

$I(X_1;W|Y_1Y_2Z_2)$ can be increased until equality holds without violating any constraints. In case $I(X_1;W|Y_1Y_2) < I(X_1;W|Y_1Y_2Z_2)$ then we can do the following transform for some $\Delta > 0$, until equality holds:

$$I(X_1;W|Y_1Y_2Z_2) \downarrow \frac{\Delta}{1-\delta_1} \tag{C.335}$$

$$H(X_1|Y_1Y_2Z_2W) \uparrow \frac{\Delta}{1-\delta_1} \tag{C.336}$$

$$\bar{x}_{22} \uparrow \frac{\Delta}{1-\delta_{2E}}. \tag{C.337}$$

No constraints are violated by this transform. Note that if (C.331) becomes equality then $I(X_1;W|Y_1Y_2) \geq I(X_1;W|Y_1Y_2Z_2)$ follows, while the RHS of (C.324) is increasing by the transform.

$$R \leq (1-\delta_2)\, m_2 + (1-\delta_1)\, I(X_1;W|Y_1Y_2) \tag{C.338}$$

$$R \leq (1-\delta_1\delta_{1E})\, \bar{x}_{121} + (1-\delta_2)\, I(X_2;W|Y_1Y_2Z_1) \tag{C.339}$$

$$R \leq (1-\delta_2\delta_{2E})\, \bar{x}_{22} + (1-\delta_1)\, I(X_1;W|Y_1Y_2) \tag{C.340}$$

$$R \leq (1-\delta_1)\, m_1 + (1-\delta_3)\, m_3 \tag{C.341}$$

$$R \leq (1-\delta_1\delta_{1E})\, \bar{x}_{131} + (1-\delta_3)\, I(X_3;W|Y_1Y_3Z_1) \tag{C.342}$$

$$R \leq (1-\delta_3\delta_{3E})\, \bar{x}_{313} + (1-\delta_1)\, I(X_1;W|Y_1Y_3Z_3) \tag{C.343}$$

$$(1-\delta_{1E})\, \bar{x}_{121} \leq \delta_{1E}(1-\delta_1)\, k_1 + (1-\delta_2)\, H(X_2|Y_1Y_2Z_1W) \tag{C.344}$$

$$(1-\delta_1\delta_{1E})\, \bar{\ell}_1 + (1-\delta_3)\, I(X_3;Z_3|Y_3Y_1Z_1W) \leq \delta_3(1-\delta_{3E})\, \bar{s}_3 \tag{C.345}$$

$$\begin{aligned}(1-\delta_{1E})\, \bar{x}_{131} \leq{}& \delta_{1E}(1-\delta_1)\, \bar{\ell}_1 + \delta_{1E}(1-\delta_1)\, \bar{s}_1 + \delta_{1E}(1-\delta_1)\, k_1 + (1-\delta_3)\, \bar{s}_3 \\ &+ (1-\delta_3)\, I(X_3;Z_3|Y_3Y_1Z_1W)\end{aligned} \tag{C.346}$$

$$(1-\delta_{2E})\, \bar{x}_{22} \leq \delta_{2E}(1-\delta_2)\, k_2 + (1-\delta_1)\, H(X_1|Y_1Y_2Z_2W) \tag{C.347}$$

$$\begin{aligned}(1-\delta_{3E})\, \bar{x}_{313} \leq{}& \delta_{3E}(1-\delta_3)\, \bar{\ell}_3 + (1-\delta_1)\, \bar{s}_1 + (1-\delta_1)\, k_1 + \delta_{3E}(1-\delta_3)\, k_3 + \delta_{3E}(1-\delta_3)\, \bar{s}_3 \\ &+ (1-\delta_1)\, I(X_1;Z_1|Y_1Y_3Z_3W)\end{aligned} \tag{C.348}$$

$$\begin{aligned}&(1-\delta_3\delta_{3E})\, \bar{\ell}_3 + (1-\delta_1)\, I(X_1;Z_1|Y_1Y_3Z_3W) \leq \delta_1(1-\delta_{1E})\, \bar{s}_1 + (1-\delta_1)\, I(X_1;Z_1|Y_1Y_2W) \\ &\quad + (1-\delta_2)\, I(X_2;Z_1|Y_1Y_2W)\end{aligned} \tag{C.349}$$

$$(1-\delta_1\delta_{1E})\, \bar{s}_1 + (1-\delta_3\delta_{3E})\, \bar{s}_3 \leq (1-\delta_2)\, H(X_2|Y_1Y_2Z_1W) \tag{C.350}$$

$$1 \geq k_1 + I(X_1;W|Y_1Y_2) + I(X_1;Z_1|Y_1Y_2W) \tag{C.351}$$

$$1 \geq H(X_1|Y_1Y_2Z_2W) + I(X_1;W|Y_1Y_2) \tag{C.352}$$

$$1 \geq \bar{s}_1 + k_1 + I(X_1;Z_1|Y_1Y_3Z_3W) + I(X_1;W|Y_1Y_3Z_3) \tag{C.353}$$

$$1 \geq m_1 + \bar{\ell}_1 + \bar{s}_1 + k_1 \tag{C.354}$$

$$1 \geq m_1 + \bar{s}_1 + k_1 + I(X_1;Z_1|Y_1Y_3Z_3W) \tag{C.355}$$

$$1 \geq m_2 + H(X_2|Y_1Y_2Z_1W) + I(X_2;Z_1|Y_1Y_2W) \tag{C.356}$$

$$1 \geq m_2 + k_2 \tag{C.357}$$

$$1 \geq H(X_2|Y_1Y_2Z_1W) + I(X_2;W|Y_1Y_2Z_1) \tag{C.358}$$

$$1 \geq m_3 + k_3 + \bar{s}_3 + \bar{z}_3 \tag{C.359}$$
$$1 \geq k_3 + \bar{s}_3 + \bar{z}_3 + I(X_3; W|Y_1 Y_3 Z_1) \tag{C.360}$$
$$1 \geq m_3 + \bar{\ell}_3 + k_3 + \bar{s}_3 \tag{C.361}$$

We observe that

$$m_1 \leq I(X_1; W|Y_1 Y_3 Z_3) \tag{C.362}$$
$$m_3 \leq I(X_3; W|Y_1 Y_3 Z_1) \tag{C.363}$$

can be assumed. If these are not true at an optimal point then we could increase $I(X_1; W|Y_1 Y_3 Z_3)$ and/or $I(X_3; W|Y_1 Y_3 Z_1)$ until they both hold. No constraints can be violated by this increase.

We next show that (C.362) can be made equality. We do the following transform (T1) ($\Delta > 0$):

$$\bar{\ell}_1 \downarrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.364}$$
$$\bar{s}_3 \downarrow \frac{\Delta}{\delta_3(1-\delta_{3E})} \tag{C.365}$$
$$\bar{s}_1 \uparrow \frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})} \tag{C.366}$$
$$I(X_1; Z_1|Y_1 Y_3 Z_3 W) \uparrow \frac{\Delta(1-\delta_3\delta_{3E})\delta_1(1-\delta_{1E})}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})(1-\delta_1)} \tag{C.367}$$
$$\bar{x}_{131} \downarrow \frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})} \tag{C.368}$$
$$\bar{x}_{313} \uparrow \frac{\Delta}{\delta_3(1-\delta_{3E})} \tag{C.369}$$
$$I(X_1; W|Y_1 Y_3 Z_3) \downarrow \frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})(1-\delta_1)} \tag{C.370}$$
$$m_1 \downarrow \frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})(1-\delta_1)} \tag{C.371}$$
$$m_3 \uparrow \frac{\Delta}{\delta_3(1-\delta_{3E})} \tag{C.372}$$
$$I(X_3; W|Y_1 Y_3 Z_1) \uparrow \frac{\Delta}{\delta_3(1-\delta_{3E})} \tag{C.373}$$

We give a side-calculation in Appendix C.4.8 to help verifying that the transform does not violate any constraints. If (C.362) is not yet equality, we can perform this transform unless any of the variables the transform decreases already equals 0 or (C.355) is equality. In the latter case $m_1 \geq I(X_1; W|Y_1 Y_3 Z_3)$ follows from (C.353) which implies that (C.362) is equality. Otherwise, we have the following cases:

1. $\bar{\ell}_1 = 0$. In this case m_1 can be increased until (C.362) is equality without violating any constraints.
2. $\bar{s}_3 = 0$. In this case (C.345) implies $\bar{\ell}_1 = 0$ and the first case applies.
3. $\bar{x}_{131} = 0$. In this case $\bar{\ell}_1$ can be decreased until it equals 0 without violating any con-

straints. Then, the first case applies.

4. $I(X_1;W|Y_1Y_3Z_3)=0$. In this case (C.362) implies that $m_1=0$ and hence (C.362) is equality.
5. $m_1=0$. In this case from (C.363) it follows that the RHS of (C.341) is strictly smaller than the RHS of (C.342) unless $\bar{x}_{131}=0$. Hence, $\bar{x}_{131}$ can be decreased to 0 without violating any constraints and thus case 3 applies.

We have shown that assuming $m_1=I(X_1;W|Y_1Y_3Z_3)$ does not restrict the value of the program. We now have the following program:

$$
\begin{aligned}
&R\leq(1-\delta_2)\,m_2+(1-\delta_1)\,I(X_1;W|Y_1Y_2)\\
&R\leq(1-\delta_1\delta_{1E})\,\bar{x}_{121}+(1-\delta_2)\,I(X_2;W|Y_1Y_2Z_1)\\
&R\leq(1-\delta_2\delta_{2E})\,\bar{x}_{22}+(1-\delta_1)\,I(X_1;W|Y_1Y_2)\\
&R\leq(1-\delta_1)\,m_1+(1-\delta_3)\,m_3 && \text{(C.374)}\\
&R\leq(1-\delta_1\delta_{1E})\,\bar{x}_{131}+(1-\delta_3)\,I(X_3;W|Y_1Y_3Z_1) && \text{(C.375)}\\
&R\leq(1-\delta_1)\,m_1+(1-\delta_3\delta_{3E})\,\bar{x}_{313} && \text{(C.376)}\\
&(1-\delta_{1E})\,\bar{x}_{121}\leq\delta_{1E}(1-\delta_1)\,k_1+(1-\delta_2)\,H(X_2|Y_1Y_2Z_1W)\\
&(1-\delta_1\delta_{1E})\,\bar{\ell}_1+(1-\delta_3)\,I(X_3;Z_3|Y_3Y_1Z_1W)\leq\delta_3(1-\delta_{3E})\,\bar{s}_3 && \text{(C.377)}\\
&(1-\delta_{1E})\,\bar{x}_{131}\leq\delta_{1E}(1-\delta_1)\,\bar{\ell}_1+\delta_{1E}(1-\delta_1)\,\bar{s}_1+\delta_{1E}(1-\delta_1)\,k_1+(1-\delta_3)\,\bar{s}_3\\
&\quad+(1-\delta_3)\,I(X_3;Z_3|Y_3Y_1Z_1W)\\
&(1-\delta_{2E})\,\bar{x}_{22}\leq\delta_{2E}(1-\delta_2)\,k_2+(1-\delta_1)\,H(X_1|Y_1Y_2Z_2W)\\
&(1-\delta_{3E})\,\bar{x}_{313}\leq\delta_{3E}(1-\delta_3)\,\bar{\ell}_3+(1-\delta_1)\,\bar{s}_1+(1-\delta_1)\,k_1+\delta_{3E}(1-\delta_3)\,k_3+\delta_{3E}(1-\delta_3)\,\bar{s}_3\\
&\quad+(1-\delta_1)\,I(X_1;Z_1|Y_1Y_3Z_3W)\\
&(1-\delta_3\delta_{3E})\,\bar{\ell}_3+(1-\delta_1)\,I(X_1;Z_1|Y_1Y_3Z_3W)\leq\delta_1(1-\delta_{1E})\,\bar{s}_1+(1-\delta_1)\,I(X_1;Z_1|Y_1Y_2W)\\
&+(1-\delta_2)\,I(X_2;Z_1|Y_1Y_2W)\\
&(1-\delta_1\delta_{1E})\,\bar{s}_1+(1-\delta_3\delta_{3E})\,\bar{s}_3\leq(1-\delta_2)\,H(X_2|Y_1Y_2Z_1W)\\
&m_3\leq I(X_3;W|Y_1Y_3Z_1) && \text{(C.378)}\\
&1\geq k_1+I(X_1;W|Y_1Y_2)+I(X_1;Z_1|Y_1Y_2W)\\
&1\geq H(X_1|Y_1Y_2Z_2W)+I(X_1;W|Y_1Y_2)\\
&1\geq m_1+\bar{s}_1+k_1+I(X_1;Z_1|Y_1Y_3Z_3W) && \text{(C.379)}\\
&1\geq m_1+\bar{\ell}_1+\bar{s}_1+k_1 && \text{(C.380)}\\
&1\geq m_2+H(X_2|Y_1Y_2Z_1W)+I(X_2;Z_1|Y_1Y_2W)\\
&1\geq m_2+k_2\\
&1\geq H(X_2|Y_1Y_2Z_1W)+I(X_2;W|Y_1Y_2Z_1)\\
&1\geq m_3+k_3+\bar{s}_3+\bar{z}_3\\
&1\geq k_3+\bar{s}_3+\bar{z}_3+I(X_3;W|Y_1Y_3Z_1)\\
&1\geq m_3+\bar{\ell}_3+k_3+\bar{s}_3
\end{aligned}
$$

We show that $\bar{\ell}_1 \le I(X_1; Z_1|Y_1 Y_3 Z_3 W)$ can be assumed. Do the following transform:

$$\bar{\ell}_1 \downarrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.381}$$

$$\bar{s_3} \downarrow \frac{\Delta}{\delta_3(1-\delta_{3E})} \tag{C.382}$$

$$\bar{s_1} \uparrow \frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})} \tag{C.383}$$

$$I(X_1; Z_1|Y_1 Y_3 Z_3 W) \uparrow \frac{\Delta(1-\delta_3\delta_{3E})\delta_1(1-\delta_{1E})}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})(1-\delta_1)} \tag{C.384}$$

$$\bar{x}_{131} \downarrow \frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})} \tag{C.385}$$

$$\bar{x}_{313} \uparrow \frac{\Delta}{\delta_3(1-\delta_{3E})} \tag{C.386}$$

$$m_1 \downarrow \frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})(1-\delta_1)} \tag{C.387}$$

$$m_3 \uparrow \frac{\Delta}{\delta_3(1-\delta_{3E})} \tag{C.388}$$

$$I(X_3; W|Y_1 Y_3 Z_1) \uparrow \frac{\Delta}{\delta_3(1-\delta_{3E})} \tag{C.389}$$

Observe, that this transform is the same as T1, we need to recheck only the changed constraint (C.376), which is straightforward to verify. We cannot do this transform in the following cases:

1. (C.379) is equality. Then, (C.380) implies that $\bar{\ell}_1 \le I(X_1; Z_1|Y_1 Y_3 Z_3 W)$.
2. $\bar{\ell}_1 = 0$. $\bar{\ell}_1 \le I(X_1; Z_1|Y_1 Y_3 Z_3 W)$ is immediate.
3. $\bar{s_3} = 0$. From (C.377) $\bar{\ell}_1 = 0$ follows, hence $\bar{\ell}_1 \le I(X_1; Z_1|Y_1 Y_3 Z_3 W)$.
4. $\bar{x}_{131} = 0$. Then, $\bar{\ell}_1$ can be reduced to 0 without violating any constraints.
5. $m_1 = 0$. $\bar{x}_{131} = 0$ can be reduced to 0 without violating any constraints, because (C.378) implies that the RHS of (C.375) is no larger than that of (C.374). Then, the previous case applies.

We can add the constraint

$$\bar{\ell}_1 \le I(X_1; Z_1|Y_1 Y_3 Z_3 W) \tag{C.390}$$

and then (C.380) becomes redundant.

$$R \le (1-\delta_2)\, m_2 + (1-\delta_1)\, I(X_1; W|Y_1 Y_2)$$
$$R \le (1-\delta_1\delta_{1E})\, \bar{x}_{121} + (1-\delta_2)\, I(X_2; W|Y_1 Y_2 Z_1)$$
$$R \le (1-\delta_2\delta_{2E})\, \bar{x}_{22} + (1-\delta_1)\, I(X_1; W|Y_1 Y_2)$$
$$R \le (1-\delta_1)\, m_1 + (1-\delta_3)\, m_3$$
$$R \le (1-\delta_1\delta_{1E})\, \bar{x}_{131} + (1-\delta_3)\, I(X_3; W|Y_1 Y_3 Z_1)$$
$$R \le (1-\delta_1)\, m_1 + (1-\delta_3\delta_{3E})\, \bar{x}_{313}$$

$$
\begin{aligned}
&(1-\delta_{1E})\,\bar{x}_{121} \le \delta_{1E}(1-\delta_1)\,k_1 + (1-\delta_2)\,H(X_2|Y_1Y_2Z_1W)\\
&(1-\delta_1\delta_{1E})\,\bar{\ell}_1 + (1-\delta_3)\,I(X_3;Z_3|Y_3Y_1Z_1W) \le \delta_3(1-\delta_{3E})\,\bar{s}_3\\
&(1-\delta_{1E})\,\bar{x}_{131} \le \delta_{1E}(1-\delta_1)\,\bar{\ell}_1 + \delta_{1E}(1-\delta_1)\,\bar{s}_1 + \delta_{1E}(1-\delta_1)\,k_1 + (1-\delta_3)\,\bar{s}_3\\
&\quad + (1-\delta_3)\,I(X_3;Z_3|Y_3Y_1Z_1W)\\
&(1-\delta_{2E})\,\bar{x}_{22} \le \delta_{2E}(1-\delta_2)\,k_2 + (1-\delta_1)\,H(X_1|Y_1Y_2Z_2W)\\
&(1-\delta_{3E})\,\bar{x}_{313} \le \delta_{3E}(1-\delta_3)\,\bar{\ell}_3 + (1-\delta_1)\,\bar{s}_1 + (1-\delta_1)\,k_1 + \delta_{3E}(1-\delta_3)\,k_3 + \delta_{3E}(1-\delta_3)\,\bar{s}_3\\
&\quad + (1-\delta_1)\,I(X_1;Z_1|Y_1Y_3Z_3W)
\end{aligned}
\tag{C.391}
$$

$$
\begin{aligned}
&(1-\delta_3\delta_{3E})\,\bar{\ell}_3 + (1-\delta_1)\,I(X_1;Z_1|Y_1Y_3Z_3W) \le \delta_1(1-\delta_{1E})\,\bar{s}_1 + (1-\delta_1)\,I(X_1;Z_1|Y_1Y_2W)\\
&+ (1-\delta_2)\,I(X_2;Z_1|Y_1Y_2W)\\
&(1-\delta_1\delta_{1E})\,\bar{s}_1 + (1-\delta_3\delta_{3E})\,\bar{s}_3 \le (1-\delta_2)\,H(X_2|Y_1Y_2Z_1W)\\
&m_3 \le I(X_3;W|Y_1Y_3Z_1)\\
&\bar{\ell}_1 \le I(X_1;Z_1|Y_1Y_3Z_3W)
\end{aligned}
\tag{C.392}
$$

$$
\begin{aligned}
&1 \ge k_1 + I(X_1;W|Y_1Y_2) + I(X_1;Z_1|Y_1Y_2W)\\
&1 \ge H(X_1|Y_1Y_2Z_2W) + I(X_1;W|Y_1Y_2)\\
&1 \ge m_1 + \bar{s}_1 + k_1 + I(X_1;Z_1|Y_1Y_3Z_3W)\\
&1 \ge m_2 + H(X_2|Y_1Y_2Z_1W) + I(X_2;Z_1|Y_1Y_2W)\\
&1 \ge m_2 + k_2\\
&1 \ge H(X_2|Y_1Y_2Z_1W) + I(X_2;W|Y_1Y_2Z_1)\\
&1 \ge m_3 + k_3 + \bar{s}_3 + \bar{z}_3\\
&1 \ge k_3 + \bar{s}_3 + \bar{z}_3 + I(X_3;W|Y_1Y_3Z_1)\\
&1 \ge m_3 + \bar{\ell}_3 + k_3 + \bar{s}_3
\end{aligned}
$$

In this system (C.392) can be made equality. In case it is not equality, do the following transform:

$$I(X_1;Z_1|Y_1Y_3Z_3W) \downarrow \frac{\Delta}{1-\delta_1} \tag{C.393}$$

$$m_1 \uparrow \frac{\Delta}{1-\delta_1} \tag{C.394}$$

$$m_3 \downarrow \frac{\Delta}{1-\delta_3} \tag{C.395}$$

$$\bar{\ell}_3 \uparrow \frac{\Delta}{1-\delta_3\delta_{3E}} \tag{C.396}$$

$$\bar{x}_{313} \downarrow \frac{\Delta}{1-\delta_3\delta_{3E}} \tag{C.397}$$

We verify that (C.391) is not violated. Other constraints are straightforward to check. Change of LHS of (C.391):

$$\frac{-\Delta(1-\delta_{3E})}{1-\delta_3\delta_{3E}} \tag{C.398}$$

Change of RHS:

$$\underbrace{\frac{\Delta\delta_{3E}(1-\delta_3)}{1-\delta_3\delta_{3E}}}_{\text{from } \delta_{3E}(1-\delta_3)\bar{\ell}_3} \underbrace{-\Delta}_{\text{from } (1-\delta_1) I(X_1;Z_1|Y_1Y_3Z_3W)} = \frac{-\Delta(1-\delta_{3E})}{1-\delta_3\delta_{3E}} \tag{C.399}$$

We have two cases when we cannot do this transform:

1. $\bar{x}_{313} = 0$. In this case decreasing $I(X_1;Z_1|Y_1Y_3Z_3W)$ until equality holds does not violate any constraints.
2. $m_3 = 0$. In this case decreasing $\bar{x}_{313}$ to 0 does not violate any constraints and the first case applies.

$$R \le (1-\delta_2)\, m_2 + (1-\delta_1)\, I(X_1;W|Y_1Y_2) \tag{C.400}$$
$$R \le (1-\delta_1\delta_{1E})\, \bar{x}_{121} + (1-\delta_2)\, I(X_2;W|Y_1Y_2Z_1)$$
$$R \le (1-\delta_2\delta_{2E})\, \bar{x}_{22} + (1-\delta_1)\, I(X_1;W|Y_1Y_2) \tag{C.401}$$
$$R \le (1-\delta_1)\, m_1 + (1-\delta_3)\, m_3$$
$$R \le (1-\delta_1\delta_{1E})\, \bar{x}_{131} + (1-\delta_3)\, I(X_3;W|Y_1Y_3Z_1)$$
$$R \le (1-\delta_1)\, m_1 + (1-\delta_3\delta_{3E})\, \bar{x}_{313} \tag{C.402}$$
$$(1-\delta_{1E})\, \bar{x}_{121} \le \delta_{1E}(1-\delta_1)\, k_1 + (1-\delta_2)\, H(X_2|Y_1Y_2Z_1W)$$
$$(1-\delta_{1E})\, \bar{x}_{131} \le \delta_{1E}(1-\delta_1)\, \bar{\ell}_1 + \delta_{1E}(1-\delta_1)\, \bar{s}_1 + \delta_{1E}(1-\delta_1)\, k_1 + (1-\delta_3)\, \bar{s}_3$$
$$+ (1-\delta_3)\, I(X_3;Z_3|Y_3Y_1Z_1W)$$
$$(1-\delta_{2E})\, \bar{x}_{22} \le \delta_{2E}(1-\delta_2)\, k_2 + (1-\delta_1)\, H(X_1|Y_1Y_2Z_2W)$$
$$(1-\delta_{3E})\, \bar{x}_{313} \le \delta_{3E}(1-\delta_3)\left(\bar{\ell}_3 + k_3 + \bar{s}_3\right) + (1-\delta_1)\left(\bar{s}_1 + k_1 + \bar{\ell}_1\right) \tag{C.403}$$
$$(1-\delta_1\delta_{1E})\, \bar{\ell}_1 + (1-\delta_3)\, I(X_3;Z_3|Y_3Y_1Z_1W) \le \delta_3(1-\delta_{3E})\, \bar{s}_3$$
$$(1-\delta_3\delta_{3E})\, \bar{\ell}_3 + (1-\delta_1)\, \bar{\ell}_1 \le \delta_1(1-\delta_{1E})\, \bar{s}_1 + (1-\delta_1)\, I(X_1;Z_1|Y_1Y_2W) + (1-\delta_2)\, I(X_2;Z_1|Y_1Y_2W)$$
$$(1-\delta_1\delta_{1E})\, \bar{s}_1 + (1-\delta_3\delta_{3E})\, \bar{s}_3 \le (1-\delta_2)\, H(X_2|Y_1Y_2Z_1W)$$
$$m_3 \le I(X_3;W|Y_1Y_3Z_1)$$
$$1 \ge k_1 + I(X_1;W|Y_1Y_2) + I(X_1;Z_1|Y_1Y_2W) \tag{C.404}$$
$$1 \ge H(X_1|Y_1Y_2Z_2W) + I(X_1;W|Y_1Y_2) \tag{C.405}$$
$$1 \ge m_1 + \bar{s}_1 + k_1 + \bar{\ell}_1 \tag{C.406}$$
$$1 \ge m_2 + H(X_2|Y_1Y_2Z_1W) + I(X_2;Z_1|Y_1Y_2W)$$
$$1 \ge m_2 + k_2$$
$$1 \ge H(X_2|Y_1Y_2Z_1W) + I(X_2;W|Y_1Y_2Z_1)$$
$$1 \ge m_3 + k_3 + \bar{s}_3 + \bar{z}_3$$
$$1 \ge k_3 + \bar{s}_3 + \bar{z}_3 + I(X_3;W|Y_1Y_3Z_1)$$
$$1 \ge m_3 + \bar{\ell}_3 + k_3 + \bar{s}_3$$

In this system one can increase the value of $I(X_1;Z_1|Y_1Y_2W)$, $H(X_1|Y_1Y_2Z_2W)$ and m_1 until

(C.404)-(C.406) all become equality. Then, it follows that we can assume

$$m_1 + \bar{s_1} + \bar{\ell}_1 \geq I(X_1; W|Y_1Y_2) + I(X_1; Z_1|Y_1Y_2W). \tag{C.407}$$

We can also increase $\bar{x}_{313}$ until (C.403) is equality. From this, we can assume

$$(1-\delta_3\delta_{3E})\,\bar{x}_{313} \geq (1-\delta_1)\left(\bar{s_1} + \bar{\ell}_1\right). \tag{C.408}$$

We show that $I(X_1; Z_1|Y_1Y_2W) = 0$ can be assumed. We do the following transform:

$$I(X_1; Z_1|Y_1Y_2W) \downarrow \frac{\Delta}{1-\delta_1} \tag{C.409}$$

$$I(X_2; Z_1|Y_1Y_2W) \uparrow \frac{\Delta}{1-\delta_2} \tag{C.410}$$

$$m_2 \downarrow \frac{\Delta}{1-\delta_2} \tag{C.411}$$

$$k_2 \uparrow \frac{\Delta}{1-\delta_2} \tag{C.412}$$

$$I(X_1; W|Y_1Y_2) \uparrow \frac{\Delta}{1-\delta_1} \tag{C.413}$$

$$H(X_1|Y_1Y_2Z_2W) \downarrow \frac{\Delta}{1-\delta_1} \tag{C.414}$$

$$\bar{x}_{22} \downarrow \frac{\Delta}{1-\delta_2\delta_{2E}}. \tag{C.415}$$

It is straightforward to verify that this transform does not violate any constraints. Note that $H(X_1|Y_1Y_2Z_2W) = 0$ implies $I(X_1; Z_1|Y_1Y_2W) = 0$, so there are two cases when we cannot do this transform:

1. $m_2 = 0$. In this case we can decrease $\bar{x}_{22}$ to 0 without violating any constraints (due to (C.400)).
2. $\bar{x}_{22} = 0$. Then, we can decrease m_2 to 0 without violating any constraints (due to (C.401)).

Hence, if $I(X_1; Z_1|Y_1Y_2W) \neq 0$ we can assume that $0 = m_2 = \bar{x}_{22}$. We observe that in this case (C.408) and (C.407) together imply that (C.402) cannot be equality. Thus, a transform that reduces the RHS of (C.402) while maintaining (C.407) and the equality of (C.403) does not violate (C.402). Apply the following transform of this kind:

$$\bar{\ell}_3 \downarrow \frac{\Delta}{1-\delta_3\delta_{3E}} \tag{C.416}$$

$$I(X_1; Z_1|Y_1Y_2W) \downarrow \frac{\Delta}{1-\delta_1} \tag{C.417}$$

$$\bar{x}_{313} \downarrow \frac{\Delta\delta_{3E}(1-\delta_3)}{(1-\delta_{3E})(1-\delta_3\delta_{3E})} \tag{C.418}$$

$$I(X_1; W|Y_1Y_2) \uparrow \frac{\Delta}{1-\delta_1} \tag{C.419}$$

$$H(X_1|Y_1Y_2Z_2W) \downarrow \frac{\Delta}{1-\delta_1}. \tag{C.420}$$

It is straightforward to verify that all constraints are respected. After this either $I(X_1;Z_1|Y_1Y_2W)=0$, or $\bar{\ell}_3=0$. Note that in case $\bar{x}_{313}=0$, since equality of (C.403) is maintained, $\bar{\ell}_3=0$ follows. We do yet another transform:

$$\bar{\ell}_1 \downarrow \frac{\Delta}{1-\delta_1} \tag{C.421}$$

$$I(X_1;Z_1|Y_1Y_2W) \downarrow \frac{\Delta}{1-\delta_1} \tag{C.422}$$

$$k_1 \uparrow \frac{\Delta}{1-\delta_1}. \tag{C.423}$$

We can do this transform unless $\bar{\ell}_1=0$. In this case, since $\bar{\ell}_3=0$, reducing $I(X_1;Z_1|Y_1Y_2W)$ to 0 does not violate any constraints. We have shown that $I(X_1;Z_1|Y_1Y_2W)=0$ can be assumed.

$$R \le (1-\delta_2)\, m_2 + (1-\delta_1)\, I(X_1;W|Y_1Y_2)$$
$$R \le (1-\delta_1\delta_{1E})\, \bar{x}_{121} + (1-\delta_2)\, I(X_2;W|Y_1Y_2Z_1)$$
$$R \le (1-\delta_2\delta_{2E})\, \bar{x}_{22} + (1-\delta_1)\, I(X_1;W|Y_1Y_2)$$
$$R \le (1-\delta_1)\, m_1 + (1-\delta_3)\, m_3$$
$$R \le (1-\delta_1\delta_{1E})\, \bar{x}_{131} + (1-\delta_3)\, I(X_3;W|Y_1Y_3Z_1)$$
$$R \le (1-\delta_1)\, m_1 + (1-\delta_3\delta_{3E})\, \bar{x}_{313}$$
$$(1-\delta_{1E})\, \bar{x}_{121} \le \delta_{1E}(1-\delta_1)\, k_1 + (1-\delta_2)\, H(X_2|Y_1Y_2Z_1W)$$
$$(1-\delta_{1E})\, \bar{x}_{131} \le \delta_{1E}(1-\delta_1)\, \bar{\ell}_1 + \delta_{1E}(1-\delta_1)\, \bar{s}_1 + \delta_{1E}(1-\delta_1)\, k_1 + (1-\delta_3)\, \bar{s}_3$$
$$\quad + (1-\delta_3)\, I(X_3;Z_3|Y_3Y_1Z_1W)$$
$$(1-\delta_{2E})\, \bar{x}_{22} \le \delta_{2E}(1-\delta_2)\, k_2 + (1-\delta_1)\, H(X_1|Y_1Y_2Z_2W)$$
$$(1-\delta_{3E})\, \bar{x}_{313} \le \delta_{3E}(1-\delta_3)\left(\bar{\ell}_3 + k_3 + \bar{s}_3\right) + (1-\delta_1)\left(\bar{s}_1 + k_1 + \bar{\ell}_1\right)$$
$$(1-\delta_1\delta_{1E})\, \bar{\ell}_1 + (1-\delta_3)\, I(X_3;Z_3|Y_3Y_1Z_1W) \le \delta_3(1-\delta_{3E})\, \bar{s}_3$$
$$(1-\delta_3\delta_{3E})\, \bar{\ell}_3 + (1-\delta_1)\, \bar{\ell}_1 \le \delta_1(1-\delta_{1E})\, \bar{s}_1 + (1-\delta_2)\, I(X_2;Z_1|Y_1Y_2W)$$
$$(1-\delta_1\delta_{1E})\, \bar{s}_1 + (1-\delta_3\delta_{3E})\, \bar{s}_3 \le (1-\delta_2)\, H(X_2|Y_1Y_2Z_1W)$$
$$m_3 \le I(X_3;W|Y_1Y_3Z_1)$$
$$1 \ge k_1 + I(X_1;W|Y_1Y_2) \tag{C.424}$$
$$1 \ge H(X_1|Y_1Y_2Z_2W) + I(X_1;W|Y_1Y_2) \tag{C.425}$$
$$1 \ge m_1 + \bar{s}_1 + k_1 + \bar{\ell}_1 \tag{C.426}$$
$$1 \ge m_2 + H(X_2|Y_1Y_2Z_1W) + I(X_2;Z_1|Y_1Y_2W)$$
$$1 \ge m_2 + k_2$$
$$1 \ge H(X_2|Y_1Y_2Z_1W) + I(X_2;W|Y_1Y_2Z_1)$$
$$1 \ge m_3 + k_3 + \bar{s}_3 + \bar{z}_3$$
$$1 \ge k_3 + \bar{s}_3 + \bar{z}_3 + I(X_3;W|Y_1Y_3Z_1)$$
$$1 \ge m_3 + \bar{\ell}_3 + k_3 + \bar{s}_3$$

We show that $I(X_1;W|Y_1Y_2) = m_1 + \bar{s}_1 + \bar{\ell}_1$ can be assumed. It is enough to show that (C.424)-(C.426) are all equalities. By increasing $H(X_1|Y_1Y_2Z_2W)$ and m_1 we can easily make (C.425) and (C.426) equality. If (C.424) is not equality, we do the following transform:

$$I(X_1;W|Y_1Y_2) \uparrow \frac{\Delta}{1-\delta_1} \tag{C.427}$$

$$H(X_1|Y_1Y_2Z_2W) \downarrow \frac{\Delta}{1-\delta_1} \tag{C.428}$$

$$k_2 \uparrow \frac{\Delta}{1-\delta_2} \tag{C.429}$$

$$m_2 \downarrow \frac{\Delta}{1-\delta_2} \tag{C.430}$$

$$\bar{x}_{22} \downarrow \frac{\Delta}{1-\delta_2\delta_{2E}} \tag{C.431}$$

There are three cases when we cannot do this transform:

1. $H(X_1|Y_1Y_2Z_2W) = 0$. This implies $I(X_1;W|Y_1Y_2) = 1$, i.e., (C.424) is equality.
2. $\bar{x}_{22} = 0$. The following transform does not violate any constraints:

$$I(X_1;W|Y_1Y_2) \uparrow \frac{\Delta}{1-\delta_1} \tag{C.432}$$

$$H(X_1|Y_1Y_2Z_2W) \downarrow \frac{\Delta}{1-\delta_1} \tag{C.433}$$

   If $H(X_1|Y_1Y_2Z_2W)$ reaches 0, the first case applies.
3. $m_2 = 0$. In this case reducing $\bar{x}_{22}$ to 0 does not violate any constraints. Then, the second case applies.

From $I(X_1;W|Y_1Y_2) = m_1 + \bar{s}_1 + \bar{\ell}_1$ and the equalities (C.424)-(C.426) it also follows that $H(X_1|Y_1Y_2Z_2W) = k_1$ can be assumed.

$$R \le (1-\delta_2)\, m_2 + (1-\delta_1)\left(m_1 + \bar{s}_1 + \bar{\ell}_1\right) \tag{C.434}$$

$$R \le (1-\delta_1\delta_{1E})\, \bar{x}_{121} + (1-\delta_2)\, I(X_2;W|Y_1Y_2Z_1) \tag{C.435}$$

$$R \le (1-\delta_1)\left(m_1 + \bar{s}_1 + \bar{\ell}_1\right) + (1-\delta_2\delta_{2E})\, \bar{x}_{22} \tag{C.436}$$

$$R \le (1-\delta_1)\, m_1 + (1-\delta_3)\, m_3 \tag{C.437}$$

$$R \le (1-\delta_1\delta_{1E})\, \bar{x}_{131} + (1-\delta_3)\, I(X_3;W|Y_1Y_3Z_1) \tag{C.438}$$

$$R \le (1-\delta_1)\, m_1 + (1-\delta_3\delta_{3E})\, \bar{x}_{313} \tag{C.439}$$

$$(1-\delta_{1E})\, \bar{x}_{121} \le \delta_{1E}(1-\delta_1)\, k_1 + (1-\delta_2)\, H(X_2|Y_1Y_2Z_1W) \tag{C.440}$$

$$\begin{aligned}(1-\delta_{1E})\, \bar{x}_{131} \le{}& \delta_{1E}(1-\delta_1)\, \bar{\ell}_1 + \delta_{1E}(1-\delta_1)\, \bar{s}_1 + \delta_{1E}(1-\delta_1)\, k_1 + (1-\delta_3)\, \bar{s}_3 \\ &+ (1-\delta_3)\, I(X_3;Z_3|Y_3Y_1Z_1W)\end{aligned} \tag{C.441}$$

$$(1-\delta_{2E})\, \bar{x}_{22} \le \delta_{2E}(1-\delta_2)\, k_2 + (1-\delta_1)\, k_1 \tag{C.442}$$

$$(1-\delta_{3E})\, \bar{x}_{313} \le \delta_{3E}(1-\delta_3)\left(\bar{\ell}_3 + k_3 + \bar{s}_3\right) + (1-\delta_1)\left(\bar{s}_1 + k_1 + \bar{\ell}_1\right) \tag{C.443}$$

$$(1-\delta_1\delta_{1E})\, \bar{\ell}_1 + (1-\delta_3)\, I(X_3;Z_3|Y_3Y_1Z_1W) \le \delta_3(1-\delta_{3E})\, \bar{s}_3 \tag{C.444}$$

$$(1-\delta_3\delta_{3E})\,\bar{\ell}_3+(1-\delta_1)\,\bar{\ell}_1\le\delta_1(1-\delta_{1E})\,\bar{s}_1+(1-\delta_2)\,I(X_2;Z_1|Y_1Y_2W) \tag{C.445}$$

$$(1-\delta_1\delta_{1E})\,\bar{s}_1+(1-\delta_3\delta_{3E})\,\bar{s}_3\le(1-\delta_2)\,H(X_2|Y_1Y_2Z_1W) \tag{C.446}$$

$$m_3\le I(X_3;W|Y_1Y_3Z_1) \tag{C.447}$$

$$1\ge m_1+\bar{s}_1+\bar{\ell}_1+k_1 \tag{C.448}$$

$$1\ge m_2+H(X_2|Y_1Y_2Z_1W)+I(X_2;Z_1|Y_1Y_2W) \tag{C.449}$$

$$1\ge m_2+k_2 \tag{C.450}$$

$$1\ge H(X_2|Y_1Y_2Z_1W)+I(X_2;W|Y_1Y_2Z_1) \tag{C.451}$$

$$1\ge m_3+k_3+\bar{s}_3+\bar{z}_3 \tag{C.452}$$

$$1\ge k_3+\bar{s}_3+\bar{z}_3+I(X_3;W|Y_1Y_3Z_1) \tag{C.453}$$

$$1\ge m_3+\bar{\ell}_3+k_3+\bar{s}_3 \tag{C.454}$$

In this system we can see that one can increase the value of $I(X_2;Z_1|Y_1Y_2W)$, k_2 and $I(X_2;W|Y_1Y_2Z_1)$ until (C.449)-(C.451) all become equalities. We next show that (C.447) can also be made equality. We do the following transform (T2):

$$\bar{\ell}_3\downarrow\frac{\Delta}{1-\delta_3\delta_{3E}} \tag{C.455}$$

$$\bar{s}_1\downarrow\frac{\Delta}{\delta_1(1-\delta_{1E})} \tag{C.456}$$

$$\bar{s}_3\uparrow\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})} \tag{C.457}$$

$$\bar{z}_3\uparrow\frac{\Delta(1-\delta_1\delta_{1E})\delta_3(1-\delta_{3E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})(1-\delta_3)} \tag{C.458}$$

$$\bar{x}_{131}\uparrow\frac{\Delta}{\delta_1(1-\delta_{1E})} \tag{C.459}$$

$$I(X_3;W|Y_1Y_3Z_1)\downarrow\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3)} \tag{C.460}$$

$$m_3\downarrow\frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})(1-\delta_3)} \tag{C.461}$$

$$m_1\uparrow\frac{\Delta}{\delta_1(1-\delta_{1E})} \tag{C.462}$$

$$\bar{x}_{313}\downarrow\frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})} \tag{C.463}$$

$$I(X_1;W|Y_1Y_3Z_3)\uparrow\frac{\Delta}{\delta_1(1-\delta_{1E})} \tag{C.464}$$

The side-calculations that verify that transform T2 respects all constraints are found in Appendix C.4.8. If (C.447) is not yet equality we can do this transform unless any variable the transform decreases is 0 or (C.452) becomes equality. The latter already implies that (C.447) is equality, so we have the following cases:

1. $\bar{\ell}_3=0$. In this case we can increase m_3 without violating any constraints until (C.447) is equality.

2. $I(X_3;W|Y_1Y_3Z_1)=0$. This implies $m_3=0$, hence equality follows.
3. $\bar{x}_{313}=0$. In this case we can decrease $\bar{\ell}_3$ to 0 without violating any constraints and then the first case applies.
4. $m_3=0$. In this case decreasing $\bar{x}_{313}$ to 0 does not violate any constraints, hence the previous case applies.
5. $\bar{s}_1=0$. In this case we do the following transform:

$$I(X_2;Z_1|Y_1Y_2W)\downarrow \frac{\Delta}{1-\delta_2} \tag{C.465}$$
$$H(X_2|Y_1Y_2Z_1W)\uparrow \frac{\Delta}{1-\delta_2} \tag{C.466}$$
$$\bar{\ell}_1\downarrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.467}$$
$$\bar{s}_1\uparrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.468}$$
$$I(X_2;W|Y_1Y_2Z_1)\downarrow \frac{\Delta}{1-\delta_2} \tag{C.469}$$
$$\bar{x}_{121}\uparrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.470}$$

It is straightforward to verify that we can do this transform unless any of the decreased variables equal 0. There are three cases:

(a) $I(X_2;Z_1|Y_1Y_2W)=0$. Since $\bar{s}_1=0$, (C.445) implies that $\bar{\ell}_3=0$ and then case 1) applies.

(b) $I(X_2;W|Y_1Y_2Z_1)=0$. Since the transform maintains the equalities (C.449)-(C.451), $m_2=I(X_2;Z_1|Y_1Y_2W)=0$ also follows, hence the previous case applies.

(c) $\bar{\ell}_1=0$. In this case, we do yet another transform:

$$I(X_2;Z_1|Y_1Y_2W)\downarrow \frac{\Delta}{1-\delta_2} \tag{C.471}$$
$$H(X_2|Y_1Y_2Z_1W)\uparrow \frac{\Delta}{1-\delta_2} \tag{C.472}$$
$$\bar{\ell}_3\downarrow \frac{\Delta}{1-\delta_3\delta_{3E}} \tag{C.473}$$
$$\bar{s}_3\uparrow \frac{\Delta}{1-\delta_3\delta_{3E}} \tag{C.474}$$
$$I(X_2;W|Y_1Y_2Z_1)\downarrow \frac{\Delta}{1-\delta_2} \tag{C.475}$$
$$\bar{x}_{121}\uparrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.476}$$
$$I(X_3;W|Y_1Y_3Z_1)\downarrow \frac{\Delta}{1-\delta_3\delta_{3E}} \tag{C.477}$$
$$\bar{x}_{131}\uparrow \frac{\Delta(1-\delta_3)}{(1-\delta_3\delta_{3E})(1-\delta_1\delta_{1E})} \tag{C.478}$$

Again, it is straightforward to verify the correctness of this transform. There are four cases if (C.447) is not equality:

i. $I(X_2;Z_1|Y_1Y_2W)=0$. Case 5a) applies.
ii. $I(X_2;W|Y_1Y_2Z_1)=0$. Since the transform maintains the equalities (C.449)-(C.451), $m_2=I(X_2;Z_1|Y_1Y_2W)=0$ also follows and the previous case holds.
iii. $\bar{\ell}_3=0$. Case 1) holds.
iv. $I(X_3;W|Y_3Y_1Z_1W)=0$. Case 2) holds.

$$
\begin{aligned}
&R\le(1-\delta_2)\,m_2+(1-\delta_1)\left(m_1+\bar{s}_1+\bar{\ell}_1\right)\\
&R\le(1-\delta_1\delta_{1E})\,\bar{x}_{121}+(1-\delta_2)\,I(X_2;W|Y_1Y_2Z_1)\\
&R\le(1-\delta_1)\left(m_1+\bar{s}_1+\bar{\ell}_1\right)+(1-\delta_2\delta_{2E})\,\bar{x}_{22}\\
&R\le(1-\delta_1)\,m_1+(1-\delta_3)\,m_3\\
&R\le(1-\delta_3)\,m_3+(1-\delta_1\delta_{1E})\,\bar{x}_{131}\\
&R\le(1-\delta_1)\,m_1+(1-\delta_3\delta_{3E})\,\bar{x}_{313}\\
&(1-\delta_{1E})\,\bar{x}_{121}\le\delta_{1E}(1-\delta_1)\,k_1+(1-\delta_2)\,H(X_2|Y_1Y_2Z_1W)\\
&(1-\delta_{1E})\,\bar{x}_{131}\le\delta_{1E}(1-\delta_1)\,\bar{\ell}_1+\delta_{1E}(1-\delta_1)\,\bar{s}_1+\delta_{1E}(1-\delta_1)\,k_1+(1-\delta_3)\,\bar{s}_3\\
&\quad+(1-\delta_3)\,I(X_3;Z_3|Y_3Y_1Z_1W)\\
&(1-\delta_{2E})\,\bar{x}_{22}\le\delta_{2E}(1-\delta_2)\,k_2+(1-\delta_1)\,k_1\\
&(1-\delta_{3E})\,\bar{x}_{313}\le\delta_{3E}(1-\delta_3)\left(\bar{\ell}_3+k_3+\bar{s}_3\right)+(1-\delta_1)\left(\bar{s}_1+k_1+\bar{\ell}_1\right)\\
&(1-\delta_1\delta_{1E})\,\bar{\ell}_1+(1-\delta_3)\,I(X_3;Z_3|Y_3Y_1Z_1W)\le\delta_3(1-\delta_{3E})\,\bar{s}_3\\
&(1-\delta_3\delta_{3E})\,\bar{\ell}_3+(1-\delta_1)\,\bar{\ell}_1\le\delta_1(1-\delta_{1E})\,\bar{s}_1+(1-\delta_2)\,I(X_2;Z_1|Y_1Y_2W)\\
&(1-\delta_1\delta_{1E})\,\bar{s}_1+(1-\delta_3\delta_{3E})\,\bar{s}_3\le(1-\delta_2)\,H(X_2|Y_1Y_2Z_1W)\\
&1\ge m_1+\bar{s}_1+\bar{\ell}_1+k_1\\
&1\ge m_2+H(X_2|Y_1Y_2Z_1W)+I(X_2;Z_1|Y_1Y_2W) \qquad \text{(C.479)}\\
&1\ge m_2+k_2 \qquad \text{(C.480)}\\
&1\ge H(X_2|Y_1Y_2Z_1W)+I(X_2;W|Y_1Y_2Z_1) \qquad \text{(C.481)}\\
&1\ge m_3+k_3+\bar{s}_3+\bar{z}_3 \qquad \text{(C.482)}\\
&1\ge m_3+\bar{\ell}_3+k_3+\bar{s}_3
\end{aligned}
$$

We next show that $I(X_2;Z_1|Y_1Y_2W)=0$ can be assumed in this system. Again, we can increase the value of $I(X_2;Z_1|Y_1Y_2W)$, k_2 and $I(X_2;W|Y_1Y_2Z_1)$ until (C.479)-(C.481) all become equalities. We do the following transform:

$$I(X_2;Z_1|Y_1Y_2W)\downarrow\frac{\Delta}{1-\delta_2} \qquad \text{(C.483)}$$

$$H(X_2|Y_1Y_2Z_1W)\uparrow\frac{\Delta}{1-\delta_2} \qquad \text{(C.484)}$$

$$\bar{\ell}_1\downarrow\frac{\Delta}{1-\delta_1\delta_{1E}} \qquad \text{(C.485)}$$

$$\bar{s_1} \uparrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.486}$$

$$I(X_2;W|Y_1Y_2Z_1) \downarrow \frac{\Delta}{1-\delta_2} \tag{C.487}$$

$$\bar{x}_{121} \uparrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.488}$$

The transform maintains all the inequalities, thus we cannot do this transform in the following two cases:

1. $I(X_2;W|Y_1Y_2Z_1) = 0$. Since the transform maintains the equalities (C.479)-(C.481), $m_2 = I(X_2;Z_1|Y_1Y_2W) = 0$ also follows.
2. $\bar{\ell_1} = 0$. In this case we do the following transform:

$$I(X_2;Z_1|Y_1Y_2W) \downarrow \frac{\Delta}{1-\delta_2} \tag{C.489}$$

$$H(X_2|Y_1Y_2Z_1W) \uparrow \frac{\Delta}{1-\delta_2} \tag{C.490}$$

$$\bar{\ell_3} \downarrow \frac{\Delta}{1-\delta_3\delta_{3E}} \tag{C.491}$$

$$\bar{s_3} \uparrow \frac{\Delta}{1-\delta_3\delta_{3E}} \tag{C.492}$$

$$I(X_2;W|Y_1Y_2Z_1) \downarrow \frac{\Delta}{1-\delta_2} \tag{C.493}$$

$$\bar{x}_{121} \uparrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.494}$$

   This transform maintains all inequalities except (C.482), which case is considered as case 2c below. Hence, there are three cases if $I(X_2;Z_1|Y_1Y_2W)$ is not 0:

   (a) $I(X_2;W|Y_1Y_2Z_1) = 0$. Since the transform maintains the equalities (C.479)-(C.481), $m_2 = I(X_2;Z_1|Y_1Y_2W) = 0$ also follows.

   (b) $\bar{\ell_3} = 0$. Since $\bar{\ell_1} = 0$ we can reduce $I(X_2;Z_1|Y_1Y_2W)$ to 0 without violating any constraints.

   (c) (C.482) is equality. In this case we know that $\bar{z_3} \geq \bar{\ell_3}$. We do yet another transform:

$$I(X_2;Z_1|Y_1Y_2W) \downarrow \frac{\Delta}{1-\delta_2} \tag{C.495}$$

$$H(X_2|Y_1Y_2Z_1W) \uparrow \frac{\Delta}{1-\delta_2} \tag{C.496}$$

$$\bar{\ell_3} \downarrow \frac{\Delta}{1-\delta_3\delta_{3E}} \tag{C.497}$$

$$\bar{s_3} \uparrow \frac{\Delta}{1-\delta_3\delta_{3E}} \tag{C.498}$$

$$I(X_2;W|Y_1Y_2Z_1) \downarrow \frac{\Delta}{1-\delta_2} \tag{C.499}$$

$$\bar{x}_{121} \uparrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.500}$$

$$\bar{z}_3 \downarrow \frac{\Delta}{1-\delta_3\delta_{3E}} \tag{C.501}$$

After this transform either case 2a or case 2b occurs.

Since all transforms maintain equalities (C.479)-(C.481), $m_2 = I(X_2; W|Y_1Y_2Z_1)$ and $k_2 = H(X_2|Y_1Y_2Z_1W)$ can also be assumed.

$$R \le (1-\delta_2)\, m_2 + (1-\delta_1)\left(m_1 + \bar{s}_1 + \bar{\ell}_1\right) \tag{C.502}$$
$$R \le (1-\delta_2)\, m_2 + (1-\delta_1\delta_{1E})\, \bar{x}_{121} \tag{C.503}$$
$$R \le (1-\delta_1)\left(m_1 + \bar{s}_1 + \bar{\ell}_1\right) + (1-\delta_2\delta_{2E})\, \bar{x}_{22} \tag{C.504}$$
$$R \le (1-\delta_1)\, m_1 + (1-\delta_3)\, m_3 \tag{C.505}$$
$$R \le (1-\delta_3)\, m_3 + (1-\delta_1\delta_{1E})\, \bar{x}_{131} \tag{C.506}$$
$$R \le (1-\delta_1)\, m_1 + (1-\delta_3\delta_{3E})\, \bar{x}_{313} \tag{C.507}$$
$$(1-\delta_{1E})\, \bar{x}_{121} \le \delta_{1E}(1-\delta_1)\, k_1 + (1-\delta_2)\, k_2 \tag{C.508}$$
$$\begin{aligned}(1-\delta_{1E})\, \bar{x}_{131} \le{}& \delta_{1E}(1-\delta_1)\, \bar{\ell}_1 + \delta_{1E}(1-\delta_1)\, \bar{s}_1 + \delta_{1E}(1-\delta_1)\, k_1 + (1-\delta_3)\, \bar{s}_3 \\ &+ (1-\delta_3)\, I(X_3; Z_3|Y_3Y_1Z_1W)\end{aligned} \tag{C.509}$$
$$(1-\delta_{2E})\, \bar{x}_{22} \le \delta_{2E}(1-\delta_2)\, k_2 + (1-\delta_1)\, k_1 \tag{C.510}$$
$$(1-\delta_{3E})\, \bar{x}_{313} \le \delta_{3E}(1-\delta_3)\left(\bar{\ell}_3 + k_3 + \bar{s}_3\right) + (1-\delta_1)\left(\bar{s}_1 + k_1 + \bar{\ell}_1\right) \tag{C.511}$$
$$(1-\delta_1\delta_{1E})\, \bar{\ell}_1 + (1-\delta_3)\, I(X_3; Z_3|Y_3Y_1Z_1W) \le \delta_3(1-\delta_{3E})\, \bar{s}_3 \tag{C.512}$$
$$(1-\delta_3\delta_{3E})\, \bar{\ell}_3 + (1-\delta_1)\, \bar{\ell}_1 \le \delta_1(1-\delta_{1E})\, \bar{s}_1 \tag{C.513}$$
$$(1-\delta_1\delta_{1E})\, \bar{s}_1 + (1-\delta_3\delta_{3E})\, \bar{s}_3 \le (1-\delta_2)\, k_2 \tag{C.514}$$
$$1 \ge m_1 + \bar{s}_1 + \bar{\ell}_1 + k_1 \tag{C.515}$$
$$1 \ge m_2 + k_2 \tag{C.516}$$
$$1 \ge m_3 + k_3 + \bar{s}_3 + \bar{z}_3 \tag{C.517}$$
$$1 \ge m_3 + \bar{\ell}_3 + k_3 + \bar{s}_3 \tag{C.518}$$

We show that $\bar{\ell}_3 \le \bar{z}_3$ can be assumed. If in an optimal point this inequality does not hold, do the following transform (T3):

$$\bar{\ell}_3 \downarrow \frac{\Delta}{1-\delta_3\delta_{3E}} \tag{C.519}$$
$$\bar{s}_1 \downarrow \frac{\Delta}{\delta_1(1-\delta_{1E})} \tag{C.520}$$
$$\bar{s}_3 \uparrow \frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})} \tag{C.521}$$
$$\bar{z}_3 \uparrow \frac{\Delta\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})(1-\delta_3)} \tag{C.522}$$
$$m_1 \uparrow \frac{\Delta}{\delta_1(1-\delta_{1E})} \tag{C.523}$$
$$\bar{x}_{131} \uparrow \frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})(1-\delta_1\delta_{1E})} \tag{C.524}$$

$$m_3 \downarrow \frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})(1-\delta_3)} \tag{C.525}$$

$$\bar{x}_{313} \downarrow \frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})} \tag{C.526}$$

We verify this transform in Appendix C.4.8. We can do this transform unless one of the following cases occurs:

1. (C.517) is equality. In this case $\bar{\ell}_3 \le \bar{z}_3$ already follows from (C.518).
2. $\bar{\ell}_3 = 0$. The inequality in question is immediate.
3. $\bar{s}_1 = 0$. In this case (C.513) implies $\bar{\ell}_3 = 0$ and the inequality follows.
4. $\bar{x}_{313} = 0$. Then $\bar{\ell}_3$ can be reduced to 0 without violating any constraints.
5. $m_3 = 0$. In this case $\bar{x}_{313}$ can be reduced to 0 without violating any constraints and the previous case applies.

We introduce the following new variables replacing some of the existing ones:

$$r_3 \sim \left(\bar{z}_3 - \bar{\ell}_3\right) \frac{(1-\delta_3\delta_{3E})(1-\delta_3)}{\delta_3(1-\delta_{3E})} \tag{C.527}$$

$$c_1 \sim \bar{s}_1 + \bar{\ell}_1 \tag{C.528}$$

$$c_3 \sim \bar{\ell}_3 + \bar{s}_3 - \frac{r_3}{1-\delta_3\delta_{3E}} \tag{C.529}$$

$$c \sim (1-\delta_1\delta_{1E})\,\bar{s}_1 + (1-\delta_3\delta_{3E})\,\bar{s}_3 - r_3 \tag{C.530}$$

The non-negativity of the introduced variables is the consequence of $\bar{\ell}_3 \le \bar{z}_3$ and (C.512). The new system is

$$\begin{aligned}
&R \le (1-\delta_2)\,m_2 + (1-\delta_1)\,(m_1+c_1) \qquad \text{(C.531)}\\
&R \le (1-\delta_2)\,m_2 + (1-\delta_1\delta_{1E})\,\bar{x}_{121}\\
&R \le (1-\delta_1)\,(m_1+c_1) + (1-\delta_2\delta_{2E})\,\bar{x}_{22}\\
&R \le (1-\delta_1)\,m_1 + (1-\delta_3)\,m_3\\
&R \le (1-\delta_3)\,m_3 + (1-\delta_1\delta_{1E})\,\bar{x}_{131}\\
&R \le (1-\delta_1)\,m_1 + (1-\delta_3\delta_{3E})\,\bar{x}_{313}\\
&(1-\delta_{1E})\,\bar{x}_{121} \le \delta_{1E}(1-\delta_1)\,k_1 + (1-\delta_2)\,k_2\\
&(1-\delta_{1E})\,\bar{x}_{131} \le \delta_{1E}(1-\delta_1)\,(c_1+k_1) + (1-\delta_3)\,c_3 + r_3\\
&(1-\delta_{2E})\,\bar{x}_{22} \le \delta_{2E}(1-\delta_2)\,k_2 + (1-\delta_1)\,k_1\\
&(1-\delta_{3E})\,\bar{x}_{313} \le \delta_{3E}(1-\delta_3)\,(c_3+k_3) + r_3\frac{\delta_{3E}(1-\delta_3)}{1-\delta_3\delta_{3E}} + (1-\delta_1)\,(c_1+k_1)\\
&(1-\delta_1\delta_{1E})\,c_1 + (1-\delta_3)\,c_3 \le c\\
&(1-\delta_3\delta_{3E})\,c_3 + (1-\delta_1)\,c_1 \le c\\
&c + r_3 \le (1-\delta_2)\,k_2\\
&1 \ge m_1 + c_1 + k_1
\end{aligned}$$

$$1 \geq m_2 + k_2$$
$$1 \geq m_3 + k_3 + c_3 + \frac{r_3}{1-\delta_3}$$

We show that $(1-\delta_2)\, m_2 = (1-\delta_2\delta_{2E})\, \bar{x}_{22}$ can be assumed. If equality does not hold, we have two cases:

1. $(1-\delta_2)\, m_2 < (1-\delta_2\delta_{2E})\, \bar{x}_{22}$. In this case the value of $\bar{x}_{22}$ can be decreased until equality holds without violating any constraints.
2. $(1-\delta_2)\, m_2 > (1-\delta_2\delta_{2E})\, \bar{x}_{22}$. Do the following transform:

$$m_2 \downarrow \frac{\Delta}{1-\delta_2} \tag{C.532}$$
$$k_2 \uparrow \frac{\Delta}{1-\delta_2} \tag{C.533}$$
$$\bar{x}_{121} \uparrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.534}$$

Since (C.531) cannot be equality by assumption, we can always do this transform until equality holds.

$$R \leq (1-\delta_2)\, m_2 + (1-\delta_1)\,(m_1 + c_1)$$
$$R \leq (1-\delta_2)\, m_2 + (1-\delta_1\delta_{1E})\, \bar{x}_{121}$$
$$R \leq (1-\delta_1)\, m_1 + (1-\delta_3)\, m_3$$
$$R \leq (1-\delta_3)\, m_3 + (1-\delta_1\delta_{1E})\, \bar{x}_{131} \tag{C.535}$$
$$R \leq (1-\delta_1)\, m_1 + (1-\delta_3\delta_{3E})\, \bar{x}_{313}$$
$$(1-\delta_{1E})\, \bar{x}_{121} \leq \delta_{1E}\,(1-\delta_1)\, k_1 + (1-\delta_2)\, k_2 \tag{C.536}$$
$$(1-\delta_{1E})\, \bar{x}_{131} \leq \delta_{1E}\,(1-\delta_1)\,(c_1 + k_1) + (1-\delta_3)\, c_3 + r_3$$
$$\frac{(1-\delta_2)\,(1-\delta_{2E})}{1-\delta_2\delta_{2E}}\, m_2 \leq \delta_{2E}\,(1-\delta_2)\, k_2 + (1-\delta_1)\, k_1$$
$$(1-\delta_{3E})\, \bar{x}_{313} \leq \delta_{3E}\,(1-\delta_3)\,(c_3 + k_3) + r_3 \frac{\delta_{3E}\,(1-\delta_3)}{1-\delta_3\delta_{3E}} + (1-\delta_1)\,(c_1 + k_1) \tag{C.537}$$
$$(1-\delta_1\delta_{1E})\, c_1 + (1-\delta_3)\, c_3 \leq c$$
$$(1-\delta_3\delta_{3E})\, c_3 + (1-\delta_1)\, c_1 \leq c$$
$$c + r_3 \leq (1-\delta_2)\, k_2$$
$$1 \geq m_1 + c_1 + k_1$$
$$1 \geq m_2 + k_2$$
$$1 \geq m_3 + k_3 + c_3 + \frac{r_3}{1-\delta_3}$$

We show that $(1-\delta_1)\,(m_1 + c_1) = (1-\delta_1\delta_{1E})\, \bar{x}_{121}$ can be assumed. If equality does not hold, there are two cases:

1. $(1-\delta_1)\,(m_1 + c_1) < (1-\delta_1\delta_{1E})\, \bar{x}_{121}$. In this case the value of $\bar{x}_{121}$ can be decreased until

equality holds without violating any constraints.

2. $(1-\delta_1)(m_1+c_1) > (1-\delta_1\delta_{1E})\,\bar{x}_{121}$. Increase $\bar{x}_{121}$ until (C.536) is equality. From this we know that

$$(1-\delta_{1E})\,\bar{x}_{121} - (1-\delta_{1E})\,\bar{x}_{131} \geq \tag{C.538}$$

$$\geq (1-\delta_2)\,k_2 - \delta_{1E}(1-\delta_1)\,c_1 - (1-\delta_3)\,c_3 - r_3 \geq c - \delta_{1E}(1-\delta_1)\,c_1 - (1-\delta_3)\,c_3 \geq (1-\delta_{1E})\,c_1 \tag{C.539}$$

and hence

$$(1-\delta_1\delta_{1E})\,\bar{x}_{121} - (1-\delta_1\delta_{1E})\,\bar{x}_{131} \geq (1-\delta_1)\,c_1 \tag{C.540}$$

Do the following transform:

$$m_2 \downarrow \frac{\Delta}{1-\delta_2} \tag{C.541}$$

$$k_2 \uparrow \frac{\Delta}{1-\delta_2} \tag{C.542}$$

$$\bar{x}_{121} \uparrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.543}$$

We can do this transform, unless $m_2 = 0$. In this case first decrease $\bar{x}_{131}$ until it is 0 or (C.535) is equality. We then know that

$$(1-\delta_1)\,m_1 \geq (1-\delta_1\delta_{1E})\,\bar{x}_{131} \tag{C.544}$$

Then, increase $\bar{x}_{313}$ until (C.537) is equality. We then know that

$$(1-\delta_3\delta_{3E})\,\bar{x}_{313} \geq (1-\delta_1)\,c_1. \tag{C.545}$$

From these inequalities it follows that decreasing m_1 does not violate any constraints unless (C.544) is equality. In this latter case however, (C.540) implies that

$$(1-\delta_1\delta_{1E})\,\bar{x}_{121} \geq (1-\delta_1)\,(m_1+c_1)\,. \tag{C.546}$$

Thus, we can always decrease m_1 until $(1-\delta_1)(m_1+c_1) = (1-\delta_1\delta_{1E})\,\bar{x}_{121}$.

$$R \leq (1-\delta_1)(m_1+c_1) + (1-\delta_2)\,m_2 \tag{C.547}$$

$$R \leq (1-\delta_1)\,m_1 + (1-\delta_3)\,m_3 \tag{C.548}$$

$$R \leq (1-\delta_3)\,m_3 + (1-\delta_1\delta_{1E})\,\bar{x}_{131} \tag{C.549}$$

$$R \leq (1-\delta_1)\,m_1 + (1-\delta_3\delta_{3E})\,\bar{x}_{313} \tag{C.550}$$

$$\frac{(1-\delta_{1E})(1-\delta_1)}{1-\delta_1\delta_{1E}}(m_1+c_1) \leq \delta_{1E}(1-\delta_1)\,k_1 + (1-\delta_2)\,k_2 \tag{C.551}$$

$$(1-\delta_{1E})\,\bar{x}_{131} \leq \delta_{1E}(1-\delta_1)(c_1+k_1) + (1-\delta_3)\,c_3 + r_3 \tag{C.552}$$

$$\frac{(1-\delta_2)(1-\delta_{2E})}{1-\delta_2\delta_{2E}} m_2 \le \delta_{2E}(1-\delta_2) k_2 + (1-\delta_1) k_1 \tag{C.553}$$

$$(1-\delta_{3E})\bar{x}_{313} \le \delta_{3E}(1-\delta_3)(c_3+k_3) + r_3\frac{\delta_{3E}(1-\delta_3)}{1-\delta_3\delta_{3E}} + (1-\delta_1)(c_1+k_1) \tag{C.554}$$

$$(1-\delta_1\delta_{1E})c_1 + (1-\delta_3)c_3 \le c \tag{C.555}$$

$$(1-\delta_3\delta_{3E})c_3 + (1-\delta_1)c_1 \le c \tag{C.556}$$

$$c + r_3 \le (1-\delta_2)k_2 \tag{C.557}$$

$$1 \ge m_1 + c_1 + k_1 \tag{C.558}$$

$$1 \ge m_2 + k_2 \tag{C.559}$$

$$1 \ge m_3 + k_3 + c_3 + \frac{r_3}{1-\delta_3} \tag{C.560}$$

We show that $(1-\delta_3)m_3 \le (1-\delta_1)c_1 + (1-\delta_2)m_2$ can be assumed. Assume the contrary. Then, we know that (C.548) is not equality. Do the following transform (T4):

$$c_1 \downarrow \frac{\Delta}{\delta_1(1-\delta_{1E})} \tag{C.561}$$

$$m_1 \uparrow \frac{\Delta}{\delta_1(1-\delta_{1E})} \tag{C.562}$$

$$\bar{x}_{131} \uparrow \frac{\Delta}{\delta_1(1-\delta_{1E})} \tag{C.563}$$

$$\bar{x}_{313} \downarrow \frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})} \tag{C.564}$$

$$m_3 \downarrow \frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3)} \tag{C.565}$$

$$r_3 \uparrow \frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})} + \frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})} \tag{C.566}$$

$$c_3 \downarrow \frac{\Delta}{\delta_3(1-\delta_{3E})} \tag{C.567}$$

$$c \downarrow \frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})} + \frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})} \tag{C.568}$$

$$\tag{C.569}$$

The side-calculation in C.4.8 shows that the transform respects all inequalities. We have the following cases:

1. $m_3 = 0$. $(1-\delta_3)m_3 \le (1-\delta_1)c_1 + (1-\delta_2)m_2$ already holds.

2. $\bar{x}_{313} = 0$. If this variable cannot be increased, then (C.554) is equality and thus $c_1 = 0$ follows, and case 3 applies.

3. $c_1 = 0$. In this case we know that (C.555) is not equality unless $c_3 = 0$. We either can do the following transform or $c_3 = 0$ (case 3c below):

$$c_1 \uparrow \frac{\Delta}{1-\delta_1} \tag{C.570}$$

$$m_1 \downarrow \frac{\Delta}{1-\delta_1} \tag{C.571}$$

$$k_2 \uparrow \frac{\Delta}{1-\delta_2} \tag{C.572}$$

$$c \uparrow \Delta \tag{C.573}$$

$$\bar{x}_{313} \uparrow \frac{\Delta}{1-\delta_3\delta_{3E}} \tag{C.574}$$

$$m_2 \downarrow \frac{\Delta}{1-\delta_2} \tag{C.575}$$

It is straightforward to verify that the transform does not violate any constraints. We have thus three cases:

(a) $m_1 = 0$. In this case we know that (C.549) cannot be equality, thus m_3 can be decreased until $(1-\delta_3)\, m_3 \leq (1-\delta_1)\, c_1 + (1-\delta_2)\, m_2$ holds.

(b) $m_2 = 0$. In this case we can decrease $\bar{x}_{313}$ to 0 without violating any constraints. Do the following transform:

$$m_1 \downarrow \frac{\Delta}{1-\delta_1} \tag{C.576}$$

$$c_1 \uparrow \frac{\Delta}{1-\delta_1} \tag{C.577}$$

$$c_3 \downarrow \frac{\Delta}{1-\delta_3} \tag{C.578}$$

$$m_3 \uparrow \frac{\Delta}{1-\delta_3} \tag{C.579}$$

$$\bar{x}_{131} \downarrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.580}$$

$$\bar{x}_{313} \uparrow \frac{\Delta}{1-\delta_3\delta_{3E}} \tag{C.581}$$

We have the following cases:

i. $m_1 = 0$. Then (C.549) cannot be equality ($R = 0$), hence we can decrease m_3 to 0 and then $(1-\delta_3)\, m_3 \leq (1-\delta_1)\, c_1 + (1-\delta_2)\, m_2$ holds.

ii. $c_3 = 0$. Case 3c.

iii. $\bar{x}_{131} = 0$. If this variable cannot be increased then (C.552) is equality, hence $c_3 = 0$ and case 3c applies.

iv. (C.554) is equality. Then $c_3 = 0$ follows and case 3c applies.

(c) $c_3 = 0$. Do the following transform:

$$c_3 \uparrow \frac{\Delta}{1-\delta_3} \tag{C.582}$$

$$m_3 \downarrow \frac{\Delta}{1-\delta_3} \tag{C.583}$$

$$\bar{x}_{131} \uparrow \frac{\Delta}{1-\delta_1\delta_{1E}} \tag{C.584}$$

We can do this transform unless $c = 0$. (Note that if $m_3 = 0$, then $(1-\delta_3)\, m_3 \leq (1-\delta_1)\, c_1 + (1-\delta_2)\, m_2$ already holds.) If $c = 0$, we can either increase c and do the transform or (C.557) is equality. Thus we can assume $c = 0$ and (C.557) is equality. In this case increase $\bar{x}_{131}$ until (C.552) is equality. Then from (C.551) and (C.552):

$$(1-\delta_1)\, m_1 \leq \frac{1-\delta_1\delta_{1E}}{1-\delta_{1E}} \left(\delta_{1E}\,(1-\delta_1)\, k_1 + (1-\delta_2)\, k_2\right) = \frac{1-\delta_1\delta_{1E}}{1-\delta_{1E}} \left(\delta_{1E}\,(1-\delta_1)\, k_1 + r_3\right) \tag{C.585}$$

$$= (1-\delta_1\delta_{1E})\, \bar{x}_{131}. \tag{C.586}$$

This means that (C.549) cannot be equality, and hence m_3 can be decreased until $(1-\delta_3)\, m_3 \leq (1-\delta_1)\, c_1 + (1-\delta_2)\, m_2$ holds.

4. $c_3 = 0$. In this case we know that (C.556) is not equality otherwise we have case 3. Do the following transform:

$$c_1 \downarrow \frac{\Delta}{\delta_1\,(1-\delta_{1E})} \tag{C.587}$$

$$m_1 \uparrow \frac{\Delta}{\delta_1\,(1-\delta_{1E})} \tag{C.588}$$

$$\bar{x}_{131} \uparrow \frac{\Delta}{\delta_1\,(1-\delta_{1E})} \tag{C.589}$$

$$\bar{x}_{313} \downarrow \frac{\Delta\,(1-\delta_1)}{\delta_1\,(1-\delta_{1E})\,(1-\delta_3\delta_{3E})} \tag{C.590}$$

$$m_3 \downarrow \frac{\Delta\,(1-\delta_1\delta_{1E})}{\delta_1\,(1-\delta_{1E})\,(1-\delta_3)} \tag{C.591}$$

$$r_3 \uparrow \frac{\Delta\,(1-\delta_1\delta_{1E})}{\delta_1\,(1-\delta_{1E})} \tag{C.592}$$

$$c \downarrow \frac{\Delta\,(1-\delta_1\delta_{1E})}{\delta_1\,(1-\delta_{1E})} \tag{C.593}$$

It is straightforward to verify that the transform does not violate any constraints. After this transform either of the previous three cases occurs. We add the constraint $(1-\delta_3)\, m_3 \leq (1-\delta_1)\, c_1 + (1-\delta_2)\, m_2$ and drop the constraint (C.551).

$$\begin{aligned}
&(1-\delta_3)\, m_3 \leq (1-\delta_1)\, c_1 + (1-\delta_2)\, m_2 \\
&R \leq (1-\delta_1)\, m_1 + (1-\delta_3)\, m_3 \\
&R \leq (1-\delta_3)\, m_3 + (1-\delta_1\delta_{1E})\, \bar{x}_{131} \\
&R \leq (1-\delta_1)\, m_1 + (1-\delta_3\delta_{3E})\, \bar{x}_{313} \\
&(1-\delta_{1E})\, \bar{x}_{131} \leq \delta_{1E}\,(1-\delta_1)\,(c_1+k_1) + (1-\delta_3)\, c_3 + r_3 \\
&\frac{(1-\delta_2)\,(1-\delta_{2E})}{1-\delta_2\delta_{2E}}\, m_2 \leq \delta_{2E}\,(1-\delta_2)\, k_2 + (1-\delta_1)\, k_1 \\
&(1-\delta_{3E})\, \bar{x}_{313} \leq \delta_{3E}\,(1-\delta_3)\,(c_3+k_3) + r_3 \frac{\delta_{3E}\,(1-\delta_3)}{1-\delta_3\delta_{3E}} + (1-\delta_1)\,(c_1+k_1)
\end{aligned} \tag{C.594}$$

$$
\begin{aligned}
&(1-\delta_1\delta_{1E})\,c_1+(1-\delta_3)\,c_3\le c\\
&(1-\delta_3\delta_{3E})\,c_3+(1-\delta_1)\,c_1\le c\\
&c+r_3\le(1-\delta_2)\,k_2\\
&1\ge m_1+c_1+k_1\\
&1\ge m_2+k_2\\
&1\ge m_3+k_3+c_3+\frac{r_3}{1-\delta_3}
\end{aligned}
$$

We show that $(1-\delta_1)\,m_1=(1-\delta_1\delta_{1E})\,\bar{x}_{131}$ can be assumed. If $(1-\delta_1)\,m_1<(1-\delta_1\delta_{1E})\,\bar{x}_{131}$, then we can decrease $\bar{x}_{131}$ until equality holds without violating any constraint. Assume that $(1-\delta_1)\,m_1>(1-\delta_1\delta_{1E})\,\bar{x}_{131}$. Do the following transform:

$$m_1\downarrow\Delta \tag{C.595}$$

$$k_1\uparrow\Delta \tag{C.596}$$

$$\bar{x}_{313}\uparrow\frac{\Delta\,(1-\delta_1)}{1-\delta_3\delta_{3E}} \tag{C.597}$$

This transform does not violate any constraints ((C.594) cannot be equality), thus m_1 decreases until $(1-\delta_1)\,m_1=(1-\delta_1\delta_{1E})\,\bar{x}_{131}$.

$$(1-\delta_3)\,m_3\le(1-\delta_1)\,c_1+(1-\delta_2)\,m_2 \tag{C.598}$$

$$R\le(1-\delta_1)\,m_1+(1-\delta_3)\,m_3 \tag{C.599}$$

$$R\le(1-\delta_1)\,m_1+(1-\delta_3\delta_{3E})\,\bar{x}_{313} \tag{C.600}$$

$$\frac{(1-\delta_1)\,(1-\delta_{1E})}{1-\delta_1\delta_{1E}}\,m_1\le\delta_{1E}\,(1-\delta_1)\,(c_1+k_1)+(1-\delta_3)\,c_3+r_3 \tag{C.601}$$

$$\frac{(1-\delta_2)\,(1-\delta_{2E})}{1-\delta_2\delta_{2E}}\,m_2\le\delta_{2E}\,(1-\delta_2)\,k_2+(1-\delta_1)\,k_1 \tag{C.602}$$

$$(1-\delta_{3E})\,\bar{x}_{313}\le\delta_{3E}\,(1-\delta_3)\,(c_3+k_3)+r_3\frac{\delta_{3E}\,(1-\delta_3)}{1-\delta_3\delta_{3E}}+(1-\delta_1)\,(c_1+k_1) \tag{C.603}$$

$$(1-\delta_1\delta_{1E})\,c_1+(1-\delta_3)\,c_3\le c \tag{C.604}$$

$$(1-\delta_3\delta_{3E})\,c_3+(1-\delta_1)\,c_1\le c \tag{C.605}$$

$$c+r_3\le(1-\delta_2)\,k_2 \tag{C.606}$$

$$1\ge m_1+c_1+k_1 \tag{C.607}$$

$$1\ge m_2+k_2 \tag{C.608}$$

$$1\ge m_3+k_3+c_3+\frac{r_3}{1-\delta_3} \tag{C.609}$$

We show that $(1-\delta_3)\,m_3=(1-\delta_3\delta_{3E})\,\bar{x}_{313}$ can be assumed. If equality does not hold, then we can always decrease the variable on the larger side of the inequality without violating any constraints until equality holds.

$$(1-\delta_3)\,m_3\le(1-\delta_1)\,c_1+(1-\delta_2)\,m_2 \tag{C.610}$$

$$
\begin{aligned}
&R \le (1-\delta_1)\, m_1 + (1-\delta_3)\, m_3 \\
&\frac{(1-\delta_1)\,(1-\delta_{1E})}{1-\delta_1\delta_{1E}}\, m_1 \le \delta_{1E}\,(1-\delta_1)\,(c_1+k_1) + (1-\delta_3)\, c_3 + r_3 \\
&\frac{(1-\delta_2)\,(1-\delta_{2E})}{1-\delta_2\delta_{2E}}\, m_2 \le \delta_{2E}\,(1-\delta_2)\, k_2 + (1-\delta_1)\, k_1 \\
&\frac{(1-\delta_3)\,(1-\delta_{3E})}{1-\delta_3\delta_{3E}}\, m_3 \le \delta_{3E}\,(1-\delta_3)\,(c_3+k_3) + r_3\frac{\delta_{3E}\,(1-\delta_3)}{1-\delta_3\delta_{3E}} + (1-\delta_1)\,(c_1+k_1) \\
&(1-\delta_1\delta_{1E})\, c_1 + (1-\delta_3)\, c_3 \le c \\
&(1-\delta_3\delta_{3E})\, c_3 + (1-\delta_1)\, c_1 \le c \\
&c + r_3 \le (1-\delta_2)\, k_2 \\
&1 \ge m_1 + c_1 + k_1 \\
&1 \ge m_2 + k_2 \\
&1 \ge m_3 + k_3 + c_3 + \frac{r_3}{1-\delta_3}
\end{aligned}
$$

Finally, we show that (C.610) can be assumed to be equality. If $(1-\delta_3)\, m_3 < (1-\delta_1)\, c_1 + (1-\delta_2)\, m_2$, then m_2 can be decreased until equality holds or $m_2 = 0$. In that case, do the following transform:

$$c_1 \downarrow \Delta \tag{C.611}$$

$$k_1 \uparrow \Delta. \tag{C.612}$$

Eventually (C.610) becomes equality by this transform. No constraints are violated.

$$
\begin{aligned}
&(1-\delta_3)\, m_3 = (1-\delta_1)\, c_1 + (1-\delta_2)\, m_2 \qquad \text{(C.613)} \\
&R \le (1-\delta_1)\, m_1 + (1-\delta_3)\, m_3 \\
&\frac{(1-\delta_1)\,(1-\delta_{1E})}{1-\delta_1\delta_{1E}}\, m_1 \le \delta_{1E}\,(1-\delta_1)\,(c_1+k_1) + (1-\delta_3)\, c_3 + r_3 \\
&\frac{(1-\delta_2)\,(1-\delta_{2E})}{1-\delta_2\delta_{2E}}\, m_2 \le \delta_{2E}\,(1-\delta_2)\, k_2 + (1-\delta_1)\, k_1 \\
&\frac{(1-\delta_3)\,(1-\delta_{3E})}{1-\delta_3\delta_{3E}}\, m_3 \le \delta_{3E}\,(1-\delta_3)\,(c_3+k_3) + r_3\frac{\delta_{3E}\,(1-\delta_3)}{1-\delta_3\delta_{3E}} + (1-\delta_1)\,(c_1+k_1) \\
&(1-\delta_1\delta_{1E})\, c_1 + (1-\delta_3)\, c_3 \le c \\
&(1-\delta_3\delta_{3E})\, c_3 + (1-\delta_1)\, c_1 \le c \\
&c + r_3 \le (1-\delta_2)\, k_2 \\
&1 \ge m_1 + c_1 + k_1 \\
&1 \ge m_2 + k_2 \\
&1 \ge m_3 + k_3 + c_3 + \frac{r_3}{1-\delta_3}
\end{aligned}
$$

The resulting LP is the same as the LP in Theorem 4.4, which concludes the proof. □

C.4.8 Side-calculations

Transform T1

Constraints (C.338)-(C.340) are not affected. Change of RHS of (C.341):

$$\underbrace{-\frac{1-\delta_3}{\delta_3(1-\delta_{3E})}\Delta}_{\text{from } (1-\delta_1)\,m_1}+\underbrace{\frac{1-\delta_3}{\delta_3(1-\delta_{3E})}\Delta}_{\text{from } (1-\delta_3)\,m_3}=0. \tag{C.614}$$

Change of RHS of (C.342):

$$\underbrace{-\frac{1-\delta_3}{\delta_3(1-\delta_{3E})}\Delta}_{\text{from } (1-\delta_1\delta_{1E})\,\bar{x}_{131}}+\underbrace{\frac{1-\delta_3}{\delta_3(1-\delta_{3E})}\Delta}_{\text{from } (1-\delta_3)\,I(X_3;W|Y_1Y_3Z_1)}=0. \tag{C.615}$$

Change of RHS of (C.343):

$$\underbrace{-\frac{1-\delta_3\delta_{3E}}{\delta_3(1-\delta_{3E})}\Delta}_{\text{from } (1-\delta_3\delta_{3E})\,\bar{x}_{313}}+\underbrace{\frac{1-\delta_3\delta_{3E}}{\delta_3(1-\delta_{3E})}\Delta}_{\text{from } (1-\delta_1)\,I(X_1;W|Y_1Y_3Z_3)}=0. \tag{C.616}$$

Constraint (C.344) is not affected. Change of LHS of (C.345):

$$\underbrace{-\Delta}_{\text{from } (1-\delta_1\delta_{1E})\,\bar{\ell}_1} \tag{C.617}$$

Change of RHS of (C.345):

$$\underbrace{-\Delta}_{\text{from } \bar{s}_3} \tag{C.618}$$

Change of LHS of (C.346):

$$\underbrace{-\frac{\Delta(1-\delta_3)(1-\delta_{1E})}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})}}_{\text{from } (1-\delta_{1E})\,\bar{x}_{131}} \tag{C.619}$$

Change of RHS of (C.346):

$$\underbrace{-\frac{\Delta\delta_{1E}(1-\delta_{1E})}{1-\delta_1\delta_{1E}}}_{\text{from } \delta_{1E}(1-\delta_1)\,\bar{\ell}_1}\underbrace{+\frac{\Delta\delta_{1E}(1-\delta_1)(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})}}_{\text{from } \delta_{1E}(1-\delta_1)\,\bar{s}_1}\underbrace{-\frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})}}_{\text{from } (1-\delta_3)\,\bar{s}_3} \tag{C.620}$$

$$=\frac{\Delta\delta_{1E}(1-\delta_1)(1-\delta_3)}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})}-\frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})}=-\frac{\Delta(1-\delta_3)(1-\delta_{1E})}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})} \tag{C.621}$$

Constraint (C.347) is not affected. Change of LHS of (C.348):

$$\underbrace{\frac{\Delta}{\delta_3}}_{\text{from } (1-\delta_{3E})\,\bar{x}_{313}} \tag{C.622}$$

Change of RHS of (C.348):

$$\underbrace{\frac{\Delta(1-\delta_1)(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})}}_{\text{from } (1-\delta_1)\,\bar{s}_1}\underbrace{-\frac{\Delta\delta_{3E}(1-\delta_3)}{\delta_3(1-\delta_{3E})}}_{\text{from } \delta_{3E}(1-\delta_3)\,\bar{s}_3}\underbrace{+\frac{\Delta(1-\delta_3\delta_{3E})\delta_1(1-\delta_{1E})}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})}}_{\text{from } (1-\delta_1)\,I(X_1;Z_1|Y_1Y_3Z_3W)} \tag{C.623}$$

$$=-\frac{\Delta\delta_{3E}(1-\delta_3)}{\delta_3(1-\delta_{3E})}+\frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})}=\frac{\Delta}{\delta_3} \tag{C.624}$$

Change of LHS of (C.349):

$$\underbrace{\frac{\Delta(1-\delta_3\delta_{3E})\delta_1(1-\delta_{1E})}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})}}_{\text{from } (1-\delta_1)\,I(X_1;Z_1|Y_1Y_3Z_3W)} \tag{C.625}$$

Change of RHS of (C.349):

$$\underbrace{\frac{\Delta(1-\delta_3\delta_{3E})\delta_1(1-\delta_{1E})}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})}}_{\text{from } \delta_1(1-\delta_{1E})\,\bar{s}_1} \tag{C.626}$$

Change of LHS of (C.350):

$$\underbrace{\frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})}}_{\text{from } (1-\delta_1\delta_{1E})\,\bar{s}_1}-\underbrace{\frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})}}_{\text{from } (1-\delta_3\delta_{3E})\,\bar{s}_3}=0 \tag{C.627}$$

Constraints (C.351)-(C.352) are not affected. Change of RHS of (C.353):

$$\underbrace{\frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})}}_{\text{from } \bar{s}_1}\underbrace{+\frac{\Delta(1-\delta_3\delta_{3E})\delta_1(1-\delta_{1E})}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})(1-\delta_1)}}_{\text{from } I(X_1;Z_1|Y_1Y_3Z_3W)}\underbrace{-\frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})(1-\delta_1)}}_{\text{from } I(X_1;W|Y_1Y_3Z_3)} \tag{C.628}$$

$$=\frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})(1-\delta_1)}-\frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})(1-\delta_1)}=0 \tag{C.629}$$

Change of RHS of (C.354):

$$\underbrace{-\frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})(1-\delta_1)}}_{\text{from } m_1}\underbrace{-\frac{\Delta}{1-\delta_1\delta_{1E}}}_{\text{from } \bar{\ell}_1}\underbrace{+\frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})}}_{\text{from } \bar{s}_1} \tag{C.630}$$

$$=-\frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})(1-\delta_1)}+\frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})}\leq 0 \tag{C.631}$$

Constraint (C.355) cannot be violated by assumption of the transform. Constraints (C.356)-(C.358) are not affected. Change of RHS of (C.359):

$$\underbrace{\frac{\Delta}{\delta_3\,(1-\delta_{3E})}}_{\text{from } m_3}\underbrace{-\frac{\Delta}{\delta_3\,(1-\delta_{3E})}}_{\text{from } \bar{s}_3}=0 \tag{C.632}$$

Change of RHS of (C.360):

$$\underbrace{\frac{\Delta}{\delta_3\,(1-\delta_{3E})}}_{\text{from } I(X_3;W|Y_1Y_3Z_1)}\quad\underbrace{-\frac{\Delta}{\delta_3\,(1-\delta_{3E})}}_{\text{from } \bar{s}_3}=0 \tag{C.633}$$

Change of RHS of (C.361):

$$\underbrace{\frac{\Delta}{\delta_3\,(1-\delta_{3E})}}_{\text{from } I(X_3;W|Y_1Y_3Z_1)}\quad\underbrace{-\frac{\Delta}{\delta_3\,(1-\delta_{3E})}}_{\text{from } \bar{s}_3}=0 \tag{C.634}$$

Change of LHS of (C.363):

$$\underbrace{\frac{\Delta}{\delta_3\,(1-\delta_{3E})}}_{\text{from } m_3} \tag{C.635}$$

Change of RHS of (C.363):

$$\underbrace{\frac{\Delta}{\delta_3\,(1-\delta_{3E})}}_{\text{from } I(X_3;W|Y_1Y_3Z_1)} \tag{C.636}$$

Transform T2

Change of RHS of (C.434):

$$\underbrace{\frac{\Delta\,(1-\delta_1)}{\delta_1\,(1-\delta_{1E})}}_{\text{from } (1-\delta_1)\,m_1}\underbrace{-\frac{\Delta\,(1-\delta_1)}{\delta_1\,(1-\delta_{1E})}}_{\text{from } (1-\delta_1)\,\bar{s}_1}=0 \tag{C.637}$$

Constraint (C.435) is not affected. Change of RHS of (C.436) is 0, for the same reason as (C.434). Change of RHS of (C.437):

$$\underbrace{\frac{\Delta\,(1-\delta_1)}{\delta_1\,(1-\delta_{1E})}}_{\text{from } (1-\delta_1)\,m_1}\underbrace{-\frac{\Delta\,(1-\delta_1)}{\delta_1\,(1-\delta_{1E})}}_{\text{from } (1-\delta_3)\,m_3}=0 \tag{C.638}$$

Change of RHS of (C.438):

$$\underbrace{\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_1\delta_{1E})\,\bar{x}_{131}} \underbrace{-\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_3)\,I(X_3;W|Y_1Y_3Z_1)} = 0 \tag{C.639}$$

Change of RHS of (C.439):

$$\underbrace{\frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_1)\,m_1} \underbrace{-\frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_3\delta_{3E})\,\bar{x}_{313}} = 0 \tag{C.640}$$

Constraint (C.440) is not affected. Change of LHS of (C.441):

$$\underbrace{\frac{\Delta}{\delta_1}}_{\text{from } (1-\delta_{1E})\,\bar{x}_{131}} \tag{C.641}$$

Change of RHS of (C.441):

$$\underbrace{-\frac{\Delta\delta_{1E}(1-\delta_1)}{\delta_1(1-\delta_{1E})}}_{\text{from } \delta_{1E}(1-\delta_1)\,\bar{s}_1} \underbrace{+\frac{\Delta(1-\delta_3)(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})}}_{\text{from } (1-\delta_3)\,\bar{s}_3} \underbrace{+\frac{\Delta(1-\delta_1\delta_{1E})\delta_3(1-\delta_{3E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})}}_{\text{from } (1-\delta_3)\,\bar{z}_3} \tag{C.642}$$

$$= -\frac{\Delta\delta_{1E}(1-\delta_1)}{\delta_1(1-\delta_{1E})} + \frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})} = \frac{\Delta}{\delta_1} \tag{C.643}$$

Constraint (C.442) is not affected. Change of LHS of (C.443):

$$\underbrace{-\frac{\Delta(1-\delta_1)(1-\delta_{3E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})}}_{\text{from } (1-\delta_{3E})\,\bar{x}_{313}} \tag{C.644}$$

Change of RHS of (C.443):

$$\underbrace{-\frac{\Delta\delta_{3E}(1-\delta_3)}{1-\delta_3\delta_{3E}}}_{\text{from } \delta_{3E}(1-\delta_3)\,\bar{\ell}_3} \underbrace{+\frac{\Delta\delta_{3E}(1-\delta_3)(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})}}_{\text{from } \delta_{3E}(1-\delta_3)\,\bar{s}_3} \underbrace{-\frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_1)\,\bar{s}_2} \tag{C.645}$$

$$= \frac{\Delta\delta_{3E}(1-\delta_3)(1-\delta_1)}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})} - \frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})} = -\frac{\Delta(1-\delta_1)(1-\delta_{3E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})} \tag{C.646}$$

Change of LHS of (C.444):

$$\underbrace{\frac{\Delta(1-\delta_1\delta_{1E})\delta_3(1-\delta_{3E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})}}_{\text{from } (1-\delta_3)\,\bar{z}_3} \tag{C.647}$$

Change of RHS of (C.444):

$$\underbrace{\frac{\Delta\,(1-\delta_1\delta_{1E})\,\delta_3\,(1-\delta_{3E})}{\delta_1\,(1-\delta_{1E})\,(1-\delta_3\delta_{3E})}}_{\text{from } \delta_3\,(1-\delta_{3E})\,\bar{s}_3} \tag{C.648}$$

Change of LHS of (C.445):

$$\underbrace{-\Delta}_{\text{from } (1-\delta_3\delta_{3E})\,\bar{l}_3} \tag{C.649}$$

Change of RHS of (C.445):

$$\underbrace{-\Delta}_{\text{from } \delta_1\,(1-\delta_{1E})\,\bar{s}_1} \tag{C.650}$$

Change of LHS of (C.446):

$$\underbrace{-\frac{\Delta\,(1-\delta_1\delta_{1E})}{\delta_1\,(1-\delta_{1E})}}_{\text{from } (1-\delta_1\delta_{1E})\,\bar{s}_1}\underbrace{+\frac{\Delta\,(1-\delta_1\delta_{1E})}{\delta_1\,(1-\delta_{1E})}}_{\text{from } (1-\delta_3\delta_{3E})\,\bar{s}_3}=0 \tag{C.651}$$

Constraint (C.447) is not violated by assumption. Change of RHS of (C.448):

$$\underbrace{\frac{\Delta}{\delta_1\,(1-\delta_{1E})}}_{\text{from } m_1}\underbrace{-\frac{\Delta}{\delta_1\,(1-\delta_{1E})}}_{\text{from } \bar{s}_1}=0 \tag{C.652}$$

Constraints (C.449)-(C.451) are not affected, (C.452) is not violated by assumption. Change of RHS of (C.453):

$$\underbrace{\frac{\Delta\,(1-\delta_1\delta_{1E})}{\delta_1\,(1-\delta_{1E})\,(1-\delta_3\delta_{3E})}}_{\text{from } \bar{s}_3}+\underbrace{\frac{\Delta\,(1-\delta_1\delta_{1E})\,\delta_3\,(1-\delta_{3E})}{\delta_1\,(1-\delta_{1E})\,(1-\delta_3\delta_{3E})\,(1-\delta_3)}}_{\text{from } \bar{z}_3}-\underbrace{\frac{\Delta\,(1-\delta_1\delta_{1E})}{\delta_1\,(1-\delta_{1E})\,(1-\delta_3)}}_{\text{from } I(X_3;W|Y_1Y_3Z_1)} \tag{C.653}$$

$$\frac{\Delta\,(1-\delta_1\delta_{1E})}{\delta_1\,(1-\delta_{1E})\,(1-\delta_3)}-\frac{\Delta\,(1-\delta_1\delta_{1E})}{\delta_1\,(1-\delta_{1E})\,(1-\delta_3)}=0 \tag{C.654}$$

Change of RHS of (C.454):

$$\underbrace{-\frac{\Delta\,(1-\delta_1)}{\delta_1\,(1-\delta_{1E})\,(1-\delta_3)}}_{\text{from } m_3}\underbrace{-\frac{\Delta}{1-\delta_3\delta_{3E}}}_{\text{from } \bar{\ell}_3}\underbrace{+\frac{\Delta\,(1-\delta_1\delta_{1E})}{\delta_1\,(1-\delta_{1E})\,(1-\delta_3\delta_{3E})}}_{\text{from } \bar{s}_3} \tag{C.655}$$

$$=\underbrace{-\frac{\Delta\,(1-\delta_1)}{\delta_1\,(1-\delta_{1E})\,(1-\delta_3)}}+\frac{\Delta\,(1-\delta_1)}{\delta_1\,(1-\delta_{1E})\,(1-\delta_3\delta_{3E})}\leq 0 \tag{C.656}$$

Transform T3

Change of RHS of (C.502):

$$\underbrace{-\frac{(1-\delta_1)\,\Delta}{\delta_1\,(1-\delta_{1E})}}_{\text{from } (1-\delta_1)\,\bar{s}_1}\underbrace{+\frac{(1-\delta_1)\,\Delta}{\delta_1\,(1-\delta_{1E})}}_{\text{from } (1-\delta_1)\,m_1}=0 \tag{C.657}$$

Constraint (C.503) is not affected, while the RHS of (C.504) changes the same way as (C.502). Change of RHS of (C.505):

$$\underbrace{\frac{(1-\delta_1)\,\Delta}{\delta_1\,(1-\delta_{1E})}}_{\text{from } (1-\delta_1)\,m_1}\underbrace{-\frac{\Delta\,(1-\delta_1)}{\delta_1\,(1-\delta_{1E})}}_{\text{from } (1-\delta_3)\,m_3}=0 \tag{C.658}$$

Change of RHS of (C.506):

$$\underbrace{-\frac{\Delta\,(1-\delta_1)}{\delta_1\,(1-\delta_{1E})}}_{\text{from } (1-\delta_3)\,m_3}\underbrace{+\frac{(1-\delta_1)\,\Delta}{\delta_1\,(1-\delta_{1E})}}_{\text{from } (1-\delta_1\delta_{1E})\,\bar{x}_{131}}=0 \tag{C.659}$$

Change of RHS of (C.507):

$$\underbrace{\frac{(1-\delta_1)\,\Delta}{\delta_1\,(1-\delta_{1E})}}_{\text{from } (1-\delta_1)\,m_1}\underbrace{-\frac{\Delta\,(1-\delta_1)}{\delta_1\,(1-\delta_{1E})}}_{\text{from } (1-\delta_3\delta_{3E})\,\bar{x}_{313}}=0 \tag{C.660}$$

Constraint (C.508) is not affected. Change of LHS of (C.509):

$$\underbrace{\frac{\Delta\,(1-\delta_1)}{\delta_1\,(1-\delta_1\delta_{1E})}}_{\text{from } (1-\delta_{1E})\,\bar{x}_{131}} \tag{C.661}$$

Change of RHS of (C.509):

$$\underbrace{-\frac{\Delta\delta_{1E}\,(1-\delta_1)}{\delta_1\,(1-\delta_{1E})}}_{\text{from } \delta_{1E}\,(1-\delta_1)\,\bar{s}_1}\underbrace{+\frac{\Delta\,(1-\delta_3)\,(1-\delta_1\delta_{1E})}{\delta_1\,(1-\delta_{1E})\,(1-\delta_3\delta_{3E})}}_{\text{from } (1-\delta_3)\,\bar{s}_3}\underbrace{+\frac{\Delta\delta_3\,(1-\delta_{3E})\,(1-\delta_1\delta_{1E})}{\delta_1\,(1-\delta_{1E})\,(1-\delta_3\delta_{3E})}}_{\text{from } (1-\delta_3)\,I\,(X_3;Z_3|Y_3Y_1Z_1W)} \tag{C.662}$$

$$=-\frac{\Delta\delta_{1E}\,(1-\delta_1)}{\delta_1\,(1-\delta_{1E})}+\frac{\Delta\,(1-\delta_1\delta_{1E})}{\delta_1\,(1-\delta_{1E})}=\frac{\Delta}{\delta_1}\geq\frac{\Delta\,(1-\delta_1)}{\delta_1\,(1-\delta_1\delta_{1E})} \tag{C.663}$$

Constraint (C.510) is not affected. Change of LHS of (C.511):

$$\underbrace{-\frac{\Delta\,(1-\delta_{3E})\,(1-\delta_1)}{\delta_1\,(1-\delta_{1E})\,(1-\delta_3\delta_{3E})}}_{\text{from } (1-\delta_{3E})\,\bar{x}_{313}} \tag{C.664}$$

Change of LHS of (C.511):

$$\underbrace{-\frac{\Delta\delta_{3E}(1-\delta_3)}{1-\delta_3\delta_{3E}}}_{\text{from } \delta_{3E}(1-\delta_3)\,\bar{l}_3}\underbrace{+\frac{\Delta\delta_{3E}(1-\delta_3)(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})}}_{\text{from } \delta_{3E}(1-\delta_3)\,\bar{s}_3}\underbrace{-\frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_1)\,\bar{s}_1} \tag{C.665}$$

$$=\frac{\Delta\delta_{3E}(1-\delta_3)(1-\delta_1)}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})}-\frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})}=-\frac{\Delta(1-\delta_{3E})(1-\delta_1)}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})} \tag{C.666}$$

Change of LHS of (C.512):

$$\underbrace{\frac{\Delta\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})}}_{\text{from } (1-\delta_3)\,I(X_3;Z_3|Y_3Y_1Z_1W)} \tag{C.667}$$

Change of RHS of (C.512):

$$\underbrace{\frac{\Delta\delta_3(1-\delta_{3E})(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})}}_{\text{from } \delta_3(1-\delta_{3E})\,\bar{s}_3} \tag{C.668}$$

Change of LHS of (C.513):

$$\underbrace{-\Delta}_{\text{from } (1-\delta_3\delta_{3E})\,\bar{\ell}_3} \tag{C.669}$$

Change of RHS of (C.513):

$$\underbrace{-\Delta}_{\text{from } \delta_1(1-\delta_{1E})\,\bar{s}_1} \tag{C.670}$$

Change of LHS of (C.514):

$$\underbrace{-\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_1\delta_{1E})\,\bar{s}_1}\underbrace{+\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_3\delta_{3E})\,\bar{s}_3}=0 \tag{C.671}$$

Change of RHS of (C.515):

$$\underbrace{\frac{\Delta}{\delta_1(1-\delta_{1E})}}_{\text{from } m_1}\underbrace{-\frac{\Delta}{\delta_1(1-\delta_{1E})}}_{\text{from } s_1}=0 \tag{C.672}$$

Inequality (C.516) is not affected and (C.517) is not violated by assumption. Change of RHS of (C.518):

$$\underbrace{-\frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_3)(1-\delta_{1E})}}_{\text{from } m_3}\underbrace{-\frac{\Delta}{(1-\delta_3\delta_{3E})}}_{\text{from } \bar{\ell}_3}\underbrace{+\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})}}_{\text{from } \bar{s}_3} \tag{C.673}$$

$$= -\frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_3)(1-\delta_{1E})} + \frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})} \leq 0 \tag{C.674}$$

Transform T4

Change of RHS of (C.547):

$$\underbrace{\frac{(1-\delta_1)\Delta}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_1)\, m_1} \underbrace{-\frac{(1-\delta_1)\Delta}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_1)\, c_1} = 0 \tag{C.675}$$

Inequality (C.548) is not violated by assumption. Change of RHS of (C.549):

$$\underbrace{-\frac{(1-\delta_1\delta_{1E})\Delta}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_3)\, m_3} \underbrace{+\frac{(1-\delta_1\delta_{1E})\Delta}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_1\delta_{1E})\, \tilde{x}_{131}} = 0 \tag{C.676}$$

Change of RHS of (C.550):

$$\underbrace{\frac{(1-\delta_1)\Delta}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_1)\, m_1} \underbrace{-\frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_3\delta_{3E})\, \tilde{x}_{313}} = 0 \tag{C.677}$$

Change of LHS of (C.551):

$$\frac{(1-\delta_{1E})(1-\delta_1)}{1-\delta_1\delta_{1E}} \left(\underbrace{\frac{\Delta}{\delta_1(1-\delta_{1E})}}_{\text{from } m_1} \underbrace{-\frac{\Delta}{\delta_1(1-\delta_{1E})}}_{\text{from } c_1} \right) = 0 \tag{C.678}$$

Change of LHS of (C.552):

$$\underbrace{\frac{\Delta}{\delta_1}}_{\text{from } (1-\delta_{1E})\, \tilde{x}_{131}} \tag{C.679}$$

Change of RHS of (C.552):

$$\underbrace{-\frac{\Delta\delta_{1E}(1-\delta_1)}{\delta_1(1-\delta_{1E})}}_{\text{from } \delta_{1E}(1-\delta_1)\, c_1} \underbrace{-\frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})}}_{\text{from } (1-\delta_3)\, c_3} \underbrace{+\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})} + \frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})}}_{\text{from } r_3} = \frac{\Delta}{\delta_1} \tag{C.680}$$

Constraint (C.553) is not affected. Change of LHS of (C.554):

$$\underbrace{-\frac{\Delta(1-\delta_1)(1-\delta_{3E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})}}_{\text{from } (1-\delta_{3E})\, \tilde{x}_{313}} \tag{C.681}$$

Change of RHS of (C.554):

$$\underbrace{-\frac{\Delta\delta_{3E}(1-\delta_3)}{\delta_3(1-\delta_{3E})}}_{\text{from } \delta_{3E}(1-\delta_3)\,c_3} + \underbrace{\frac{\Delta\delta_{3E}(1-\delta_3)(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})} + \frac{\Delta\delta_{3E}(1-\delta_3)^2}{\delta_3(1-\delta_E)(1-\delta_3\delta_{3E})}}_{\text{from } \frac{\delta_{3E}(1-\delta_3)}{1-\delta_3\delta_{3E}}\,r_3} \underbrace{-\frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_1)\,c_1} \tag{C.682}$$

$$= -\frac{\Delta\delta_{3E}(1-\delta_3)\delta_3(1-\delta_{3E})}{\delta_3(1-\delta_{3E})(1-\delta_3\delta_{3E})} + \frac{\Delta\delta_{3E}(1-\delta_3)(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})} - \frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})} \tag{C.683}$$

$$= -\frac{\Delta\delta_{3E}(1-\delta_3)}{1-\delta_3\delta_{3E}} + \frac{\Delta\delta_{3E}(1-\delta_3)(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})} - \frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})} \tag{C.684}$$

$$= \frac{\Delta\delta_{3E}(1-\delta_3)(1-\delta_1)}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})} - \frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})} = -\frac{\Delta(1-\delta_1)(1-\delta_{3E})}{\delta_1(1-\delta_{1E})(1-\delta_3\delta_{3E})} \tag{C.685}$$

Change of LHS of (C.555):

$$\underbrace{-\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_1\delta_{1E})\,c_1} \underbrace{-\frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})}}_{\text{from } (1-\delta_3)\,c_3} \tag{C.686}$$

Change of RHS of (C.555):

$$\underbrace{-\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})} - \frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})}}_{\text{from } c} \tag{C.687}$$

Change of LHS of (C.556):

$$\underbrace{-\frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})}}_{\text{from } (1-\delta_1)\,c_1} \underbrace{-\frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})}}_{\text{from } (1-\delta_3\delta_{3E})\,c_3} \tag{C.688}$$

Change of RHS of (C.556):

$$\underbrace{-\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})} - \frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})}}_{\text{from } c} \tag{C.689}$$

$$= -\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})} + \frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})} - \frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})} + \frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})} - \frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})} - \frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})} \tag{C.690}$$

$$= -1 + 1 - \frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})} - \frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})} = -\frac{\Delta(1-\delta_1)}{\delta_1(1-\delta_{1E})} - \frac{\Delta(1-\delta_3\delta_{3E})}{\delta_3(1-\delta_{3E})} \tag{C.691}$$

Change of LHS of (C.557):

$$\underbrace{-\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})} - \frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})}}_{\text{from } c} \underbrace{+\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})} - \frac{\Delta(1-\delta_3)}{\delta_3(1-\delta_{3E})}}_{\text{from } r_3} = 0 \tag{C.692}$$

Change of LHS of (C.558):

$$\underbrace{\frac{\Delta}{\delta_1(1-\delta_{1E})}}_{\text{from } m_1}\underbrace{-\frac{\Delta}{\delta_1(1-\delta_{1E})}}_{\text{from } c_1}=0 \tag{C.693}$$

Constraint (C.559) is not affected. Change of LHS of (C.560):

$$\underbrace{-\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3)}}_{\text{from } m_3}\underbrace{-\frac{\Delta}{\delta_3(1-\delta_{3E})}}_{\text{from } c_3}\underbrace{+\frac{\Delta(1-\delta_1\delta_{1E})}{\delta_1(1-\delta_{1E})(1-\delta_3)}-\frac{\Delta}{\delta_3(1-\delta_{3E})}}_{\text{from } \frac{r_3}{1-\delta_3}}=0 \tag{C.694}$$

D Proofs for Chapter 5

D.1 Security of two-phase secure network coding

In the first phase we use new independent randomness in each slot, hence we know from the secure network code that

$$I\left(Z_A^{n_1};K,W\right)=0. \tag{D.1}$$

In the second phase we use only randomness of which Eve has no information, hence her observation of the first phase does not help her learning about W. With a slight abuse of our notation we denote Eve's observation in the second phase by $Z_A^{n_2}$. Formally,

$$\begin{aligned}
I\left(Z_A^{n};W\right)&=I\left(Z_A^{n_1}Z_A^{n_2};W\right)=I\left(Z_A^{n_2};W\right)+I\left(Z_A^{n_1};W|Z_A^{n_2}\right) && \text{(D.2)}\\
&\le I\left(Z_A^{n_2};W\right)+I\left(Z_A^{n_1};W\right)+I\left(Z_A^{n_1};Z_A^{n_2}|W\right) && \text{(D.3)}\\
&\overset{(a)}{=} I\left(Z_A^{n_2};W\right)+I\left(Z_A^{n_1};Z_A^{n_2}|W\right) && \text{(D.4)}\\
&\overset{(b)}{\le} I\left(Z_A^{n_2};W\right)+I\left(Z_A^{n_1};K,W|W\right) && \text{(D.5)}\\
&= I\left(Z_A^{n_2};W\right)+I\left(Z_A^{n_1};K|W\right)=I\left(Z_A^{n_2};W\right), && \text{(D.6)}
\end{aligned}$$

where in (a) we used that transmissions in the first phase are independent of the message, and to get (b) we used that $Z_A^{n_2}$ is a function of (K,W). In the last step we used (D.1). Further, in each slot we use independent randomness, thus

$$I\left(Z_A^{n_2};W\right)=\sum_{i=n_1+1}^{n_2}I\left(Z_{i,A};W|Z_A^{i-1}\right)\le\sum_{i=n_1+1}^{n_2}I\left(Z_{i,A};W\right)+I\left(Z_{i,A};Z_A^{i-1}|W\right)=\sum_{i=n_1+1}^{n_2}I\left(Z_{i,A};W\right). \tag{D.7}$$

Hence, we can focus on one time slot of the second phase. Again we can assume that Eve discards any packet received more than once. Let $z'\le z$ be the number of distinct packets she

receives. We can write Eve's observation in the ith time slot of the second phase as

$$Z_{n_1+i,A} = \begin{bmatrix} W'(i) + K'(i) & W(i) \end{bmatrix} Q_A = \begin{bmatrix} W'(i) + K'(i) & W(i) \end{bmatrix} \begin{bmatrix} Q_A^{\Theta\intercal} \\ Q_A^{W\intercal} \end{bmatrix} = K'(i) Q_A^{\Theta\intercal} + W\Phi_{i,A}, \tag{D.8}$$

for some matrix $\Phi_{i,A}$. From the properties of the secure network code and K we see that $K(i)' Q_A^{\Theta\intercal}$ is a set of uniform random packets and $W\Phi_{i,A}$ is at most z' linear combinations of packets from W. Hence, as we noted in the case of secure network coding, from Eve's perspective this is a one time pad encrypted data. From this observation

$$I\left(Z_A^n; W\right) \leq I\left(Z_A^{n_2}; W\right) \leq \sum_{i=n_1+1}^{n_2} I\left(Z_{i,A}; W\right) = 0 \tag{D.9}$$

follows, hence our scheme is secure. □

D.2 Proof of Theorem 5.1

We first show that a key rate κ is achieved. We denote the size of the generated keys k_1,

$$k_1 = |K| = hn_1'(1-\delta) - \zeta_1 - n_1'^{\frac{3}{4}}. \tag{D.10}$$

We denote M the set of packets D receives, M^Z is the subset of these that Eve receives and M^d is the subset that only D receives. The corresponding rows of matrix H are H^Z and H^d. Hence K can be written as

$$K = MH = \begin{bmatrix} M^d & M^Z \end{bmatrix} \begin{bmatrix} H^d \\ H^Z \end{bmatrix}. \tag{D.11}$$

Note that in case the key generation is successful, then $|M|$ is $hn_1'(1-\delta)$, but $|M^d|, |M^Z|$ are not deterministic, they depend on the channel realizations F^{n_1}.

K does not depend on $|M^Z|$, hence

$$I\left(Z_A^{n_1} F^{n_1}; K\right) = I\left(Z_A^{n_1} F^{n_1} |M^Z|; K\right) = I\left(|M^Z|; K\right) + I\left(Z_A^{n_1} F^{n_1}; K||M^Z|\right) = I\left(Z_A^{n_1} F^{n_1}; K||M^Z|\right) \tag{D.12}$$

$$= H\left(K||M^Z|\right) - H\left(K|Z_A^{n_1} F^{n_1}|M^Z|\right) = k_1 - \sum_{i=1}^{n_1} H\left(K|Z_A^{n_1} F^{n_1}|M^Z| = i\right) \Pr\left\{|M^Z| = i\right\} \tag{D.13}$$

We have

$$H\left(K|Z_A^{n_1} F^{n_1}|M^Z| = i\right) = H\left(M^d H^d + M^Z H^Z | M^Z F^{n_1} | M^Z| = i\right) = H\left(M^d H^d | M^Z F^{n_1} | M^Z| = i\right) \tag{D.14}$$

$$= H\left(M^d H^d || M^Z| = i\right) = \min\{|M| - i, k_1\}, \tag{D.15}$$

since H^d is full-rank and $|M^d| = |M| - |M^Z|$. Given this,

$$\begin{aligned} I\left(Z_A^{n_1} F^{n_1}; K\right) = I\left(Z_A^{n_1} F^{n_1} | M^Z|; K\right) &= k_1 - \sum_{i=1}^{|M|-k_1} k_1 \Pr\left\{|M^Z| = i\right\} - \sum_{i=|M|-k_1+1}^{n_1} (|M| - i) \Pr\left\{|M^Z| = i\right\} \\ &\leq k_1 - k_1 \Pr\left\{|M^Z| \leq |M| - k_1\right\} + n_1 \Pr\left\{|M^Z| > |M| - k_1\right\} \\ &= (n_1 + k_1) \Pr\left\{|M^Z| > |M| - k_1\right\}. \end{aligned} \tag{D.16}$$

The probability that Eve receives more than $|M| - k_1$ packets can be bounded as follows:

$$\begin{aligned} \Pr\left\{|M^Z| > |M| - k_1\right\} = \Pr\left\{|M^Z| > \zeta_1 + n_1'^{\frac{3}{4}}\right\} &\leq \Pr\left\{|M^Z| - \mathbb{E}\left\{|M^Z|\right\} > n_1'^{\frac{3}{4}}\right\} \\ &\leq \Pr\left\{\left||M^Z| - \mathbb{E}\left\{|M^Z|\right\}\right| > n_1'^{\frac{3}{4}}\right\} \leq e^{-c_1\sqrt{n_1'}}, \end{aligned} \tag{D.17}$$

for some constant $c_1 > 0$. We used that $\zeta_1 \geq \mathbb{E}\left\{|M^Z|\right\}$ irrespective of Eve's selection. The last inequality follows from the Chernoff-Hoeffding bound. We see from (D.17) and (D.16) that $I\left(Z_A^{n_1} F^{n_1}; K\right)$ can be made arbitrarily small by choosing a large enough n_1. This proves the security of the key.

The key generation fails if D does not receive $hn_1'(1-\delta)$ packets. We calculate the probability of the event that a node who has received $n_1'(1-\delta)$ packet fails to forward all of these to the next node towards D. This event happens if out of n_1 transmissions more than $n_1 - n_1'(1-\delta)$ erasures occur. Let η denote the number of erasures of n_1 transmissions. Then, the probability of the event equals

$$\begin{aligned} \Pr\left\{\eta > n_1 - n_1'(1-\delta)\right\} = \Pr\left\{\eta - \delta n_1 > \left(n_1 - n_1'\right)(1-\delta)\right\} &= \Pr\left\{\eta - \mathbb{E}\left\{\eta\right\} > n_1^{\frac{3}{4}}(1-\delta)\right\} \\ &\leq \Pr\left\{|\eta - \mathbb{E}\left\{\eta\right\}| > n_1^{\frac{3}{4}}(1-\delta)\right\} = e^{-c_2\sqrt{n_1}}, \end{aligned} \tag{D.18}$$

where $c_2 > 0$ is some constant and we used the Chernoff-Hoeffding bound. This shows that the probability of successful forwarding of $n_1'(1-\delta)$ packets can be made arbitrarily close to 1 on each link, hence the probability that D receives $hn_1'(1-\delta)$ packets is also arbitrarily close to 1 by selecting a large enough n_1.

The rate of the key $\lim_{n_1\to\infty} \frac{|K|}{n_1} = \kappa$ directly follows from the parameter values and from the fact that $\lim_{n_1\to\infty} \frac{n_1'}{n_1} = 1$.

Destination D can decode the message if he receives all packets in the second phase. The probability of error has the same nature as in the first phase and thus can be made arbitrarily small by selecting a large enough n_2.

Similarly as in the two phase secure network coding scheme, observing the first phase does

not help Eve to learn about W. Formally, if $I\left(Z_A^{n_1}F^{n_1};K\right)=\epsilon'$, then

$$I\left(Z_A^nF^n;W\right)\leq I\left(Z_A^{n_2}F^{n_2};W\right)+\epsilon', \tag{D.19}$$

where $Z_A^{n_2}$ denotes the packets that Eve receives in the second phase.

$Z_A^{n_2}$ can be written as W_E^Z, which denotes the subset of encrypted packets Eve receives. Let G^Z denote the corresponding rows of G, then

$$W_E^Z=W^Z+KG^Z. \tag{D.20}$$

In case $|W_E^Z|\leq|K|$, then KG^Z is a set of uniformly distributed independent packets, thus W_E^Z is a set of one-time-pad encrypted message packets, hence

$$I\left(Z_A^{n_2}F^{n_2};W\right)=I\left(W_E^ZF^{n_2};W\right)=I\left(W_E^Z;W|F^{n_2}|W_E^Z|\right) \tag{D.21}$$

$$\begin{aligned}&=I\left(W_E^Z;W|F^{n_2}|W_E^Z|\leq|K|\right)\Pr\left\{|W_E^Z|\leq|K|\right\}\\&\quad+I\left(W_E^Z;W|F^{n_2}|W_E^Z|>|K|\right)\Pr\left\{|W_E^Z|>|K|\right\}\end{aligned} \tag{D.22}$$

$$=I\left(W_E^Z;W|F^{n_2}|W_E^Z|>|K|\right)\Pr\left\{|W_E^Z|>|K|\right\}\leq n_2\Pr\left\{|W_E^Z|>|K|\right\}\leq n_2e^{-c_3\sqrt{n_2'}}, \tag{D.23}$$

where $c_3>0$ is a constant. We omit the details of the last step, where we bound the probability of the event that Eve receives significantly more packets than she is expected to. We use the same technique as we have seen earlier. This together with (D.19) shows that for a sufficiently large $n=n_1+n_2$ the scheme satisfies (5.4). The rate assertion follows directly from the parameter definitions.

D.3 Proof of Theorem 5.2

Given a network $\mathscr{G}$ consider the partitioning of the vertices (V_1,V_2) such that $s\in V_1, d\in V_2$ and it has the minimum cut value h. We create a new network $\mathscr{G}'\left(V',E'\right)$ by merging all the nodes in V_1 with S and all the nodes in V_2 with D. We further remove all (d,s) edges. The resulting graph is depicted in Figure D.1. We assume that Eve eavesdrops on the remaining set of edges. The secret-message capacity over $\mathscr{G}'$ cannot be smaller than over $\mathscr{G}$. Clearly, merging nodes and restricting Eve to the remaining edges can only increase capacity. The removal of (d,s) edges does not affect the achievable rates, since no nodes in V_2 can generate randomness, and thus whatever a scheme could send through the (d,s) edges, S can also generate from its randomness and from the public acknowledgments. (Note that the channel states are known to Eve, thus the channel itself cannot be used to generate secure randomness.) We give an upper bound on the secure capacity over $\mathscr{G}'$ which is also a valid upper bound for $\mathscr{G}$.

Using Theorem 4.2 we assume that packets sent in the same time slot on different edges are always all independent.

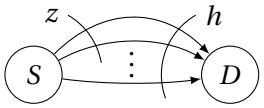


Figure D.1: Transformed network $\mathscr{G}'$ after merging network nodes with S or D

We start from the following inequality:

$$hn \geq \sum_{i=1}^{n} H\left(X_{i,s}\right) \geq \sum_{i=1}^{n} H\left(X_{i,s}|Y_d^{i-1}F^{i-1}\right) \tag{D.24}$$

$$= \sum_{i=1}^{n} H\left(X_{i,s}|Y_d^{i-1}F^{i-1}W\right) + I\left(X_{i,s};W|Y_d^{i-1}F^{i-1}\right) \tag{D.25}$$

$$\geq \sum_{i=1}^{n} H\left(X_{i,s}|Y_d^{i-1}Z_A^{i-1}F^{i-1}W\right) + I\left(X_{i,s};W|Y_d^{i-1}F^{i-1}\right) \tag{D.26}$$

The following two lemmas that provide bounds for the latter two terms in (D.26). Applying these results in (D.26) and rearranging terms provide the claimed upper bound on R.

Lemma D.1.

$$\sum_{i=1}^{n} I\left(X_{i,s};W|Y_d^{i-1}F^{i-1}\right) \geq \frac{nR}{1-\delta} - n\mathscr{E}_{D.1}, \tag{D.27}$$

where $\mathscr{E}_{D.1} = \frac{R\epsilon+h_2(\epsilon)}{1-\delta}$.

Lemma D.2.

$$nR - n\left(h-z\right)\left(1-\delta\right) \leq \sum_{i=1}^{n} \frac{\delta_E\left(1-\delta\right)\left(1-\delta\delta_E\right)}{1-\delta_E} H\left(X_{i,s}|Y_d^{i-1}Z_A^{i-1}F^{i-1}W\right) + n\mathscr{E}_{D.2}, \tag{D.28}$$

where $\mathscr{E}_{D.2} = R\epsilon + h_2\left(\epsilon\right) + \frac{\epsilon(1-\delta\delta_E)}{1-\delta_E}$.

The above two lemmas are generalizations of Lemmas 2.1-2.2 with the use of Theorem 4.2. In the case of Lemma D.1 the generalization is straightforward, hence we omit the proof to avoid repetition. We provide the proof of Lemma D.2 below. □

D.3.1 Proof of Lemma D.2

We show the following two inequalities:

$$\sum_{i=1}^{n} I(X_{i,A};W|Y_d^{i-1}Z_A^{i-1}F^{i-1})$$
$$\geq \frac{nR-n(1-\delta)(h-z)}{1-\delta\delta_E} - n\mathscr{E}_{D.2,a} + \sum_{i=1}^{n} \frac{1-\delta}{1-\delta\delta_E} H(X_{i,E'\setminus A}|Y_d^{i-1}Z_A^{i-1}F^{i-1}W), \tag{D.29}$$

where $\mathscr{E}_{D.2,a} = \frac{\epsilon R+h_2(\epsilon)}{1-\delta\delta_E}$, and

$$\sum_{i=1}^{n} I(X_{i,A};W|Y_d^{i-1}Z_A^{i-1}F^{i-1}) - \frac{n\epsilon}{1-\delta_E} \leq \sum_{i=1}^{n} \frac{1-\delta}{1-\delta_E} H(X_{i,E'\setminus A}|Y_d^{i-1}Z_A^{i-1}F^{i-1}W) + \frac{\delta_E(1-\delta)}{1-\delta_E} H(X_{i,A}|Y_d^{i-1}Z_A^{i-1}F^{i-1}W). \quad \text{(D.30)}$$

We combine these two and get that

$$\frac{nR-n(1-\delta)(h-z)}{1-\delta\delta_E} - n\mathscr{E}_{D.2,a} - \frac{n\epsilon}{1-\delta_E} + \sum_{i=1}^{n} \frac{1-\delta}{1-\delta\delta_E} H(X_{i,E'\setminus A}|Y_d^{i-1}Z_A^{i-1}F^{i-1}W) \leq \sum_{i=1}^{n} \frac{1-\delta}{1-\delta_E} H(X_{i,E'\setminus A}|Y_d^{i-1}Z_A^{i-1}F^{i-1}W) + \frac{\delta_E(1-\delta)}{1-\delta_E} H(X_{i,A}|Y_d^{i-1}Z_A^{i-1}F^{i-1}W). \quad \text{(D.31)}$$

We observe that $\frac{1-\delta}{1-\delta_E} - \frac{1-\delta}{1-\delta\delta_E} \leq \frac{\delta_E(1-\delta)}{1-\delta_E}$ and thus we can merge the entropy terms corresponding to A and $E'\setminus A$ without violating the inequality (we use again the independence property of parallel transmissions). We conclude that

$$\frac{nR-n(1-\delta)(h-z)}{1-\delta\delta_E} - n\mathscr{E}_{D.2,a} - \frac{n\epsilon}{1-\delta_E} \leq \sum_{i=1}^{n} \frac{\delta_E(1-\delta)}{1-\delta_E} H(X_{i,s}|Y_d^{i-1}Z_A^{i-1}F^{i-1}W). \quad \text{(D.32)}$$

What remains is to show (D.29) and (D.30). Consider first (D.29). We use again Fano's inequality:

$$nR - nR\epsilon - h_2(\epsilon) \leq I(Y_d^n Z_A^n F^n;W) \quad \text{(D.33)}$$

$$= \sum_{i=1}^{n} (1-\delta) I(X_{i,E'\setminus A};W|Y_d^{i-1}Z_A^{i-1}F^{i-1}) + (1-\delta\delta_E) I(X_{i,A};W|Y_d^{i-1}Z_A^{i-1}F^{i-1}) \quad \text{(D.34)}$$

$$\leq \sum_{i=1}^{n} (1-\delta)(h-z) - (1-\delta) H(X_{i,E'\setminus A}|Y_d^{i-1}Z_A^{i-1}F^{i-1}W) + (1-\delta\delta_E) I(X_{i,A};W|Y_d^{i-1}Z_A^{i-1}F^{i-1}), \quad \text{(D.35)}$$

where we used the independence property of the channel erasures as well as the independence of packets sent in the same time slot over different channels.

We derive (D.30) as follows.

$$0 \leq H(Y_d^n|Z_A^n F^n W) = H(Y_d^{n-1}|Z_A^n F^n W) + H(Y_{n,d}|Y_d^{n-1}Z_A^n F^n W) \quad \text{(D.36)}$$

$$= H(Y_d^{n-1}|Z_A^{n-1}F^{n-1}W) - I(Z_{n,A}F_n;Y_d^{n-1}|Z_A^{n-1}F^{n-1}W) + H(Y_{n,d}|Y_d^n Z_A^n F^n W) \quad \text{(D.37)}$$

$$= H(Y_d^{n-1}|Z_A^{n-1}F^{n-1}W) - I(Z_{n,A}F_n;Y_d^{n-1}|Z_A^{n-1}F^{n-1}W) + H(Y_{n,E'\setminus A}|Y_d^n Z_A^n F^n W) + H(Y_{n,A}|Y_d^n Z_A^n F^n W) \quad \text{(D.38)}$$

$$= H(Y_d^{n-1}|Z_A^{n-1}F^{n-1}W) - (1-\delta_E) I(X_{n,A};Y_d^{n-1}|Z_A^{n-1}F^{n-1}W) + (1-\delta) H(X_{n,E'\setminus A}|Y_d^{n-1}Z_A^{n-1}F^{n-1}W) + \delta_E(1-\delta) H(X_{n,A}|Y_d^{n-1}Z_A^{n-1}F^{n-1}W) \quad \text{(D.39)}$$

$$= \sum_{i=1}^{n} -(1-\delta_E) I(X_{i,A};Y_d^{i-1}|Z_A^{i-1}F^{i-1}W) + (1-\delta) H(X_{i,E'\setminus A}|Y_d^{i-1}Z_A^{i-1}F^{i-1}W) + \delta_E(1-\delta) H(X_{i,A}|Y_d^{i-1}Z_A^{i-1}F^{i-1}W) \quad \text{(D.40)}$$

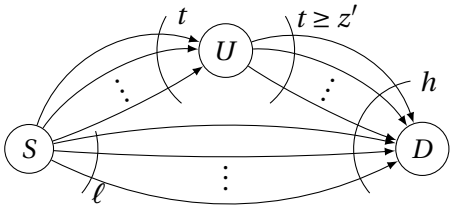


Figure D.2: Transformed network $\mathcal{G}''$ after merging all intermediate nodes and deleting some edges

$$\leq \sum_{i=1}^{n} -(1-\delta_E) I(X_{i,A}; W | Y_d^{i-1} Z_A^{i-1} F^{i-1}) + (1-\delta) H(X_{i,E'\setminus A} | Y_d^{i-1} Z_A^{i-1} F^{i-1} W)$$
$$+ \delta_E (1-\delta) H(X_{i,A} | Y_d^{i-1} Z_A^{i-1} F^{i-1} W) + n\epsilon. \tag{D.41}$$

In the last step we used that

$$\sum_{i=1}^{n} I(X_{i,A}; Y_d^{i-1} | Z_A^{i-1} F^{i-1} W) \geq \sum_{i=1}^{n} I(X_{i,A}; W | Y_d^{i-1} Z_A^{i-1} F^{i-1}) - I(X_{i,A}; W | Z_A^{i-1} F^{i-1}) \tag{D.42}$$

and that from (5.4)

$$n\epsilon > I(Z_A^n F^n; W) = \sum_{i=1}^{n} I(Z_{i,A} F_i; W | Z_A^{i-1} F^{i-1}) = \sum_{i=1}^{n} (1-\delta_E) I(X_{i,A}; W | Z_A^{i-1} F^{i-1}). \tag{D.43}$$

□

D.4 Proof of Theorem 5.3

We proceed similarly as we did in the proof of Theorem 5.2. From a network $\mathcal{G}$ we construct a new graph $\mathcal{G}''$ such that the secure capacity over $\mathcal{G}''$ cannot be smaller than over $\mathcal{G}$. First, we delete all nodes U for which $(d,u) \in E$. Note that this step cannot decrease the secure capacity of network, because if $(d,u) \in E$, then there is no path between U and D, otherwise $\mathcal{G}$ would have a cycle. After this step D has only incoming edges. Next, we merge all intermediate nodes $u \notin \{s,d\}$ into one node. As a result, $\mathcal{G}''$ is a network with three nodes: S, D and U which represents all other nodes. By this we could only increase the achievable rates, hence the upper bound we derive is valid for $\mathcal{G}$. Note that $\mathcal{G}''$ might be cyclic, there might be some edges (u,s). We know that $(d,u) \notin E''$, since D does not have any outgoing edges. As a next step we delete all edges (u,s) from E''. This step cannot reduce the secure capacity of the network, because S knows exactly every packet that U has, hence it can produce any packet that U might send on the (u,s) link. We derive our bound for $z' = \min\{t,z\}$. In case $z > t$ using z' instead of z restricts Eve, hence cannot decrease the secure capacity. Our resulting graph $\mathcal{G}''$ looks as depicted in Figure D.2.

For the same reasons as seen in Theorem 5.2 we might assume that transmissions over different channels in the same time slot are independent. We consider an eavesdropper who wiretaps

on a known z' size subset of the U-D channels.

We have

$$hn \geq \sum_{i=1}^{n} H\left(X_{i,s}\right) \geq \sum_{i=1}^{n} H\left(X_{i,s}|Y_{(s,d)}^{i-1}Y_u^{i-1}F^{i-1}\right) \tag{D.44}$$

$$= \sum_{i=1}^{n} H\left(X_{i,s}|Y_{(s,d)}^{i-1}Y_u^{i-1}F^{i-1}W\right) + I\left(X_{i,s};W|Y_{(s,d)}^{i-1}Y_u^{i-1}F^{i-1}\right) \tag{D.45}$$

$$\geq \sum_{i=1}^{n} H\left(X_{i,(s,u)}|Y_{(s,d)}^{i-1}Y_u^{i-1}F^{i-1}W\right) + I\left(X_{i,s};W|Y_{(s,d)}^{i-1}Y_u^{i-1}F^{i-1}\right) \tag{D.46}$$

$$\geq \sum_{i=1}^{n} H\left(X_{i,(s,u)}|Y_{(s,d)}^{i-1}Y_u^{i-1}Z_A^{i-1}F^{i-1}W\right) + I\left(X_{i,s};W|Y_{(s,d)}^{i-1}Y_u^{i-1}F^{i-1}\right). \tag{D.47}$$

We give bounds on the last two terms seen in (D.47).

Lemma D.3.

$$\sum_{i=1}^{n} I\left(X_{i,s};W|Y_{(s,d)}^{i-1}Y_u^{i-1}F^{i-1}\right) \geq \frac{nR}{1-\delta} - n\mathscr{E}_{D.3}, \tag{D.48}$$

where $\mathscr{E}_{D.3} = \frac{R\epsilon+h_2(\epsilon)}{1-\delta}$.

Lemma D.4.

$$\sum_{i=1}^{n} (1-\delta)\, H\left(X_{i,(s,u)}|Y_{(s,d)}^{i-1}Y_u^{i-1}Z_A^{i-1}F^{i-1}W\right) \geq \sum_{i=1}^{n} (1-\delta\delta_E)\, H\left(X_{i,A}|Y_d^{i-1}Z_A^{i-1}F^{i-1}W\right). \tag{D.49}$$

Lemma D.5.

$$nR - n\left(h-z'\right)(1-\delta) \leq \sum_{i=1}^{n} \frac{\delta_E\,(1-\delta)\,(1-\delta\delta_E)}{1-\delta_E} H\left(X_{i,A}|Y_d^{i-1}Z_A^{i-1}F^{i-1}W\right) + n\mathscr{E}_{D.5}, \tag{D.50}$$

where $\mathscr{E}_{D.5} = R\epsilon + h_2\left(\epsilon\right) + \frac{\epsilon(1-\delta\delta_E)}{1-\delta_E}$.

We give the proof of Lemmas D.3-D.4 in the following subsections. We omit the proof of Lemma D.5, which follows the same line as the proof of Lemma D.2.

We apply the results of Lemmas D.3-D.5 in (D.47) and get the claim of the theorem after rearranging terms. □

D.4.1 Proof of Lemma D.3

We observe that Y_A^n is a function of $\left(Y_u^n, F^n\right)$, and hence

$$I\left(W;Y_d^nF^n\right) \leq I\left(W;Y_{(s,d)}^nY_u^nF^n\right) = \sum_{i=1}^{n} (1-\delta)\, I\left(X_{i,s};W|Y_{(s,d)}^{i-1}Y_u^{i-1}F^{i-1}\right) \tag{D.51}$$

We then use Lemma D.1, from which we know that

$$nR - nh_2(\epsilon) - nR\epsilon \leq I\left(Y_d^n F^n; W\right). \tag{D.52}$$

From these two inequalities the claim of the lemma follows. □

D.4.2 Proof of Lemma D.4

We introduce the notation $P = (s,d) \cup (u,d) \setminus A$, i.e., P denotes the set of not eavesdropped incoming edges of D. We use the fact that Y_A^n is a function of $\left(Y_u^n, F^n\right)$. From this we have

$$H\left(Y_u^n | Z_A^n Y_P^n F^n W\right) \geq H\left(Y_A^n | Z_A^n Y_P^n F^n W\right). \tag{D.53}$$

We expand these terms as follows:

$$\begin{aligned}
H\left(Y_u^n | Z_A^n Y_P^n F^n W\right) &= H\left(Y_u^{n-1} | Z_A^n Y_P^n F^n W\right) + H\left(Y_{n,u} | Z_A^n Y_P^n Y_u^{i-1} F^n W\right) && \text{(D.54)}\\
&= H\left(Y_u^{n-1} | Z_A^{n-1} Y_P^{n-1} F^{n-1} W\right) + H\left(Y_{n,u} | Z_A^{n-1} Y_P^{n-1} Y_u^{i-1} F^{n-1} W\right)\\
&\quad - I\left(Z_{n,A}; Y_u^{n-1} | Z_A^{n-1} Y_P^{n-1} F^{n-1} W\right) - I\left(Y_{n,P}; Y_u^{n-1} | Z_A^{n-1} Y_P^{n-1} F^{n-1} W\right) && \text{(D.55)}\\
&\overset{(a)}{=} H\left(Y_u^{n-1} | Z_A^{n-1} Y_P^{n-1} F^{n-1} W\right) + H\left(Y_{n,u} | Z_A^{n-1} Y_P^{n-1} Y_u^{i-1} F^{n-1} W\right)\\
&\quad - H\left(Z_{n,A} | Z_A^{n-1} Y_P^{n-1} F^{n-1} W\right) - I\left(Y_{n,P}; Y_u^{n-1} | Z_A^{n-1} Y_P^{n-1} F^{n-1} W\right) && \text{(D.56)}\\
&= H\left(Y_u^{n-1} | Z_A^{n-1} Y_P^{n-1} F^{n-1} W\right) + (1-\delta) H\left(X_{n,(s,u)} | Z_A^{n-1} Y_P^{n-1} Y_u^{i-1} F^{n-1} W\right)\\
&\quad - (1-\delta_E) H\left(X_{n,A} | Z_A^{n-1} Y_P^{n-1} F^{n-1} W\right)\\
&\quad - (1-\delta) I\left(X_{n,P}; Y_u^{n-1} | Z_A^{n-1} Y_P^{n-1} F^{n-1} W\right) && \text{(D.57)}\\
&\overset{(b)}{=} \sum_{i=1}^{n} (1-\delta) H\left(X_{i,(s,u)} | Z_A^{i-1} Y_{(s,d)}^{i-1} Y_u^{i-1} F^{i-1} W\right)\\
&\quad - (1-\delta_E) H\left(X_{i,A} | Z_A^{i-1} Y_P^{i-1} F^{i-1} W\right)\\
&\quad - (1-\delta) I\left(X_{i,P}; Y_u^{i-1} | Z_A^{i-1} Y_P^{i-1} F^{i-1} W\right). && \text{(D.58)}
\end{aligned}$$

In (a) we used that $Z_{n,A}$ is a function of $\left(Y_u^{n-1}, F_n\right)$ and F_n is independent of every other variable. In (b) we used recursion and that $Y_{(u,d)}^{i-1}$ is a function of $\left(Y_u^{i-1}, F^{i-1}\right)$. With a similar derivation we get

$$\begin{aligned}
H\left(Y_A^n | Z_A^n Y_P^n F^n W\right) &= \sum_{i=1}^{n} \delta_E (1-\delta) H\left(X_{i,A} | Z_A^{i-1} Y_d^{i-1} F^{i-1} W\right)\\
&\quad - (1-\delta_E) I\left(X_{i,A}; Y_A^{i-1} | Z_A^{i-1} Y_P^{i-1} F^{i-1} W\right)\\
&\quad - (1-\delta) I\left(X_{i,P}; Y_A^{i-1} | Z_A^{i-1} Y_P^{i-1} F^{i-1} W\right) && \text{(D.59)}
\end{aligned}$$

$$= \sum_{i=1}^{n} (1-\delta\delta_E)\, H\left(X_{i,A}|Z_A^{i-1}Y_d^{i-1}F^{i-1}W\right) - (1-\delta_E)\, H\left(X_{i,A}|Z_A^{i-1}Y_P^{i-1}F^{i-1}W\right)$$
$$- (1-\delta)\, I\left(X_{i,P}; Y_A^{i-1}|Z_A^{i-1}Y_P^{i-1}F^{i-1}W\right). \tag{D.60}$$

To get the statement of our lemma we combine (D.53), (D.58) and (D.60) as well as use the fact that Y_A^{i-1} is a function of $\left(Y_u^{i-1}, F^{i-1}\right)$ and thus

$$\sum_{i=1}^{n} I\left(X_{i,P}; Y_A^{i-1}|Z_A^{i-1}Y_P^{i-1}F^{i-1}W\right) \leq \sum_{i=1}^{n} I\left(X_{i,P}; Y_u^{i-1}|Z_A^{i-1}Y_P^{i-1}F^{i-1}W\right). \tag{D.61}$$

□

Bibliography

[1] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[2] R. W. Yeung, *Information Theory and Network Coding*, 1st ed. Springer Publishing Company, Incorporated, 2008.

[3] A. Avestimehr, S. Diggavi, and D. Tse, "Wireless Network Information Flow: a Deterministic Approach," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1872–1905, April 2011.

[4] C. Shannon, "Communication Theory of Secrecy Systems," *The Bell System Technical Journal*, vol. 28, pp. 656–715, Oktober 1949.

[5] M. Jafari Siavoshani, S. Diggavi, C. Fragouli, U. K. Pulleti, and K. Argyraki, "Group Secret Key Generation over Broadcast Erasure Channels," in *Asilomar Conference on Signals, Systems, and Computers*, 2010, pp. 719–723.

[6] C. Shannon, "The Zero Error Capacity of a Noisy Channel," *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 8–19, September 1956.

[7] U. Maurer, "Secret Key Agreement by Public Discussion from Common Information," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

[8] I. Csiszár and P. Narayan, "Secrecy Capacities for Multiple Terminals," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3047–3061, 2004.

[9] ——, "Secrecy Capacities for Multiterminal Channels," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 2437–2452, 2008.

[10] I. Safaka, C. Fragouli, K. J. Argyraki, and S. N. Diggavi, "Creating Shared Secrets Out of Thin Air." in *ACM Workshop on Hot Topics in Networks (HotNets)*, S. Kandula, J. Padhye, E. G. Sirer, and R. Govindan, Eds., 2012, pp. 73–78.

[11] K. Argyraki, S. Diggavi, M. Duarte, C. Fragouli, M. Gatzianas, and P. Kostopoulos, "Creating Secrets Out of Erasures," in *ACM International Conference on Mobile Computing and Networking (MOBICOM)*, New York, NY, USA, 2013, pp. 429–440.

[12] M. Jafari Siavoshani, S. Mishra, S. Diggavi, and C. Fragouli, "Group Secret Key Agreement over State-dependent Wireless Broadcast Channels," in *IEEE International Symposium on Information Theory (ISIT)*, 2011.

[13] "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update." [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf

[14] B. Azimi-Sadjadi, A. Kiayias, A. Mercado, and B. Yener, "Robust Key Generation from Signal Envelopes in Wireless Networks," in *ACM Conference on Computer and Communications Security (CCS)*, New York, NY, USA, 2007, pp. 401–410.

[15] S. Jana, S. N. Premnath, M. Clark, S. K. Kasera, N. Patwari, and S. V. Krishnamurthy, "On the Effectiveness of Secret Key Extraction from Wireless Signal Strength in Real Environments." in *ACM International Conference on Mobile Computing and Networking (MOBICOM)*, K. G. Shin, Y. Zhang, R. Bagrodia, and R. Govindan, Eds. ACM, 2009, pp. 321–332.

[16] C. Ye, S. Mathur, A. Reznik, Y. Shah, W. Trappe, and N. B. Mandayam, "Information-Theoretically Secret Key Generation for Fading Wireless Channels," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 2, pp. 240–254, June 2010.

[17] C. Chen and M. Jensen, "Secret Key Establishment Using Temporally and Spatially Correlated Wireless Channel Coefficients," *IEEE Transactions on Mobile Computing*, vol. 10, no. 2, pp. 205–215, Feb 2011.

[18] M. Bloch and J. Barros, *Physical-Layer Security: from Information Theory to Security Engineering*. Cambridge University Press, 2011.

[19] I. Csiszár, "Almost Independence and Secrecy Capacity," *Problemy Peredachi Informatsii*, vol. 32, no. 1, pp. 48–57, 1996.

[20] U. Maurer and S. Wolf, "Information-theoretic Key Agreement: from Weak to Strong Secrecy for Free," in *International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT)*. Springer, 2000, pp. 351–368.

[21] D. Silva and F. R. Kschischang, "Universal Weakly Secure Network Coding," in *IEEE Information Theory Workshop (ITW)*, 2009, pp. 281–285.

[22] K. Bhattad and K. R. Narayanan, "Weakly Secure Network Coding," in *IEEE Workshop on Network Coding, Theory, and Applications (NetCod)*, 2005.

[23] F. du Pin Calmon, M. Medard, L. M. Zeger, J. Barros, M. M. Christiansen, and K. R. Duffy, "Lists That are Smaller Than Their Parts: a Coding Approach to Tunable Secrecy," in *Allerton Conference on Communication, Control, and Computing*. IEEE, 2012, pp. 1387–1394.

[24] M. Bellare, S. Tessaro, and A. Vardy, "Semantic Security for the Wiretap Channel." in *International Cryptology Conference (CRYPTO)*. Springer, 2012, pp. 294–311.

[25] A. D. Wyner, "The Wire-tap Channel," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[26] I. Csiszár and J. Körner, "Broadcast Channels with Confidential Messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, 1978.

[27] S. Leung-Yan-Cheong and M. Hellman, "The Gaussian Wire-tap Channel," *IEEE Transactions on Information Theory*, vol. 24, no. 4, pp. 451–456, Jul 1978.

[28] J. Barros and M. R. Rodrigues, "Secrecy Capacity of Wireless Channels," in *IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2006, pp. 356–360.

[29] A. Gohari and V. Anantharam, "Information-Theoretic Key Agreement of Multiple Terminals – Part I," *IEEE Transactions on Information Theory*, vol. 56, no. 8, pp. 3973–3996, Aug 2010.

[30] ——, "Information-Theoretic Key Agreement of Multiple Terminals – Part II: Channel Model," *Information Theory, IEEE Transactions on*, vol. 56, no. 8, pp. 3997–4010, Aug 2010.

[31] L. Lai, Y. Liang, and H. Poor, "A Unified Framework for Key Agreement over Wireless Fading Channels," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 480–490, April 2012.

[32] C. Mitrpant, A. Vinck, and Y. Luo, "An Achievable Region for the Gaussian Wiretap Channel with Side Information," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2181–2190, May 2006.

[33] Y. Chen and A. H. Vinck, "Wiretap Channel with Side Information," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 395–402, Jan. 2008.

[34] L. Lai, H. E. Gamal, and H. Poor, "The Wiretap Channel with Feedback: Encryption over the Channel," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5059–5067, 2008.

[35] R. Ahlswede and N. Cai, *Transmission, Identification and Common Randomness Capacities for Wire-Tape Channels with Secure Feedback from the Decoder.*, ser. LNCS. Springer, 2006, vol. 4123.

[36] E. Ardestanizadeh, M. Franceschetti, T. Javidi, and Y. Kim, "Wiretap Channel with Secure Rate-limited Feedback," *IEEE Transactions on Information Theory*, vol. 55, no. 12, pp. 5353–5361, 2009.

[37] H. Yamamoto, "Rate-distortion Theory for the Shannon Cipher System," *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 827–835, 1997.

[38] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft, "XORs in the Air: Practical Wireless Network Coding," *IEEE/ACM Transactions on Networking*, vol. 16, no. 3, pp. 497–510, June 2008.

[39] L. Georgiadis and L. Tassiulas, "Broadcast Erasure Channel with Feedback – Capacity and Algorithms," in *IEEE Workshop on Network Coding, Theory, and Applications, (NetCod)*, 2009, pp. 54–61.

[40] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Multiuser Broadcast Erasure Channel with Feedback – Capacity and Algorithms," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5779–5804, Sept 2013.

[41] Z. Bar-Yossef, Y. Birk, T. S. Jayram, and T. Kol, "Index Coding with Side Information," *Information Theory, IEEE Transactions on*, vol. 57, no. 3, pp. 1479–1494, March 2011.

[42] C. Wang, "On the Capacity of 1-to-K Broadcast Packet Erasure Channels with Channel Output Feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 931–956, 2012.

[43] K. Marton, "A Coding Theorem for the Discrete Memoryless Broadcast Channel," *IEEE Transactions on Information Theory*, vol. 25, no. 3, pp. 306–311, 1979.

[44] A. E. Gamal and E. C. van der Meulen, "A Proof of Marton's Coding Theorem for the Discrete Memoryless Broadcast Channel." *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 120–122, 1981.

[45] G. Kramer, "Capacity Results for the Discrete Memoryless Network," *IEEE Transactions on Information Theory*, vol. 49, no. 1, pp. 4–21, 2003.

[46] G. Dueck, "Partial Feedback for Two-Way and Broadcast Channels," *Information and Control*, vol. 46, no. 1, pp. 1–15, July 1980.

[47] A. E. Gamal, "The Feedback Capacity of Degraded Broadcast Channels," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 379–381, 1978.

[48] L. H. Ozarow and S. K. Leung-Yan-Cheong, "An Achievable Region and Outer Bound for the Gaussian Broadcast Channel with Feedback," *IEEE Transactions on Information Theory*, vol. 30, no. 4, pp. 667–671, 1984.

[49] S. Athanasiadou, M. Gatzianas, L. Georgiadis, and L. Tassiulas, "XOR-based Coding for the 3-user Broadcast Erasure Channel with Feedback," in *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, May 2012, pp. 417–424.

[50] D. S. Lun, M. Médard, R. Koetter, and M. Effros, "On Coding for Reliable Communication over Packet Networks," *Physical Communication*, vol. 1, no. 1, pp. 3–20, 2008.

[51] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear Network Coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[52] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network Information Flow," *IEEE Transactions on Information Theory*, vol. 46, pp. 1204–1216, 2000.

[53] C. Fragouli, D. S. Lun, M. Médard, and P. Pakzad, "On Feedback for Network Coding," in *Conference on Information Sciences and Systems (CISS)*. IEEE, 2007, pp. 248–252.

[54] M. Durvy, C. Fragouli, and P. Thiran, "Towards Reliable Broadcasting Using ACKs," in *IEEE International Symposium on Information Theory (ISIT)*, June 2007, pp. 1156–1160.

[55] N. Cai and R. Yeung, "Secure Network Coding," in *International Symposium on Information Theory (ISIT)*. IEEE, 2005, p. 323.

[56] ——, "Secure Network Coding on a Wiretap Network," *IEEE Transactions on Information Theory,*, vol. 57, no. 1, pp. 424–435, 2011.

[57] T. Cui, T. Ho, and J. Kliewer, "Achievable Strategies for General Secure Network Coding," in *Information Theory and Applications Workshop (ITA)*, Jan 2010, pp. 1–6.

[58] T. Cui, "Coding for Wireless Broadcast and Network Secrecy," Ph.D. dissertation, California Institute of Technology, 2010.

[59] A. Mills, B. Smith, T. Clancy, E. Soljanin, and S. Vishwanath, "On Secure Communication over Wireless Erasure Networks," in *IEEE International Symposium on Information Theory (ISIT)*, 2008, pp. 161–165.

[60] H. P. Williams, "Fourier's Method of Linear Programming and Its Dual," *The American mathematical monthly*, vol. 93, no. 9, pp. 681–695, 1986.

[61] T. Cui, T. Ho, and J. Kliewer, "On Secure Network Coding with Nonuniform Or Restricted Wiretap Sets," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 166–176, Jan 2013.

[62] M. Langberg and M. Médard, "On the Multiple Unicast Network Coding, Conjecture," in *47th Annual Allerton Conference on Communication, Control, and Computing*. IEEE, 2009, pp. 222–227.

[63] W. Huang, T. Ho, M. Langberg, and J. Kliewer, "On Secure Network Coding with Uniform Wiretap Sets," in *IEEE International Symposium on Network Coding (NetCod)*. IEEE, 2013.

[64] K. Jain, "Security Based on Network Topology Against the Wiretapping Attack," *Wireless Communications, IEEE*, vol. 11, no. 1, pp. 68–71, 2004.

[65] J. Feldman, T. Malkin, C. Stein, and R. Servedio, "On the Capacity of Secure Network Coding," in *Allerton Conference on Communication, Control, and Computing*, 2004.

[66] S. Rouayheb and E. Soljanin, "On Wiretap Networks II," in *International Symposium on Information Theory (ISIT)*, 2007.

[67] T. Cui, T. Ho, and J. Kliewer, "On Secure Network Coding with Unequal Link Capacities and Restricted Wiretapping Sets," in *IEEE Information Theory Workshop (ITW)*, 2010.

[68] ——, "On Secure Network Coding with Unequal Link Capacities and Restricted Wiretapping Sets," in *Information Theory Workshop (ITW)*, Aug 2010, pp. 1–5.

[69] F. Cheng and R. Yeung, "Performance Bounds on a Wiretap Network with Arbitrary Wiretap Sets," *IEEE Transactions on Information Theory*, vol. 60, no. 6, pp. 3345–3358, June 2014.

[70] D. Silva and F. R. Kschischang, "Security for Wiretap Networks Via Rank-metric Codes," in *IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2008, pp. 176–180.

[71] S. M. Bellovin, "Frank Miller: Inventor of the One-Time Pad," *Cryptologia*, vol. 35, no. 3, pp. 203–222, Jul. 2011.

[72] W. Hoeffding, "Probability Inequalities for Sums of Bounded Random Variables," *Journal of the American statistical association*, vol. 58, no. 301, pp. 13–30, 1963.

[73] F. MacWilliams and N. Sloane, *The Theory of Error-Correcting Codes*, 2nd ed. Northholland Publishing Company, 1978.

[74] W. C. Huffman and V. Pless, *Fundamentals of Error-correcting Codes*. Cambridge university press, 2003.

Curriculum Vitae

László Czap

Address: EPFL, BC 048, Station 14, 1015 Lausanne, Switzerland
Tel: +41 78 859 1254
E-mail: laszlo.czap@epfl.ch
Personal: Hungarian (Swiss B permit), 29 years old, male

Education

2010-2014 **Ph.D. in Computer and Communication Sciences, EPFL**
Dissertation: *Secure Communication in Erasure Networks with State-feedback*
Supervisor: Prof. Christina Fragouli

2008-2010 **Ph.D. Studies, Budapest University of Technology and Economics (BUTE)**
Research: *Defense against pollution attack in network coding*
Advisor: Prof. István Vajda

2003-2008 **M.Sc. in Computer Science, BUTE**
Specialized on computer security
Thesis: *Security Analysis of Anonymizer Networks*
Advisor: László Zömbik

Experience, projects

2010-2014 **EPFL**
NOWIRE project: Derivation of new capacity results applicable in wireless networks, design of polynomial time optimal coding schemes that provide information theoretic security.

2008-2010 **BUTE**
WSAN4CIP project: Design, analysis and implementation of new defense mechanism for coding based sensor storage.
BIONETS project: Design and analysis of new data dissemination techniques for delay tolerant networks achieving up to 30% improvement in throughput. Design of a new signature scheme for securing data dissemination.
Master's Thesis: Implementation of a new attack against the Tor network with detailed analysis. Design and implementation of defense against the attack.

Languages

Hungarian	Native language
English	Fluent (C1)
French	Basic (A2)
German	Basic (A2)

Publications

In submission:

[1] L. Czap, V. Prabhakaran, S. Diggavi, and C. Fragouli, "Secure Network Coding with Erasures and Feedback," *submitted to IEEE Transactions on Information Theory.*

[2] ——, "Secret communication over broadcast erasure channels with state-feedback," *submitted to IEEE Transactions on Information Theory.*

Journal papers:

[3] L. Buttyán, L. Czap, and I. Vajda, "Detection and Recovery From Pollution Attacks in Coding Based Distributed Storage Schemes," *IEEE Transactions on Dependable and Secure Computing*, vol. 7, no. 4, pp. 824 – 838, 2011.

[4] L. Czap and I. Vajda, "Secure Network Coding in DTNs," *IEEE Communications Letters*, vol. 15, no. 1, pp. 28 – 30, 2011.

Conference papers:

[5] L. Czap, V. M. Prabhakaran, S. Diggavi, and C. Fragouli, "Triangle Network Secrecy," in *IEEE International Symposium on Information Theory (ISIT)*, Honolulu, 2014.

[6] L. Czap, V. Prabhakaran, S. Diggavi, and C. Fragouli, "Secure Network Coding with Erasures and Feedback," in *Annual Allerton Conference on Communication, Control and Computing*. IEEE, 2013.

[7] L. Czap, V. Prabhakaran, and S. Diggavi, "Exploiting Common Randomness: a Resource for Network Secrecy," in *IEEE Information Theory Workshop (ITW)*, 2013.

[8] L. Czap, V. Prabhakaran, S. Diggavi, and C. Fragouli, "Securing Broadcast Against Dishonest Receivers," in *IEEE International Symposium on Network Coding (NETCOD)*, 2013.

[9] L. Czap, V. M. Prabhakaran, S. Diggavi, and C. Fragouli, "Broadcasting Private Messages Securely," in *IEEE International Symposium on Information Theory (ISIT)*, Boston, 2012, pp. 428–432, **Best student paper award candidate**.

[10] L. Czap, V. Prabhakaran, S. Diggavi, and C. Fragouli, "On Interactive Message Secrecy Over Erasure Networks," in *International Symposium on Communications, Control, and Signal Processing (ISCCSP)*, 2012, (Invited paper).

[11] L. Czap, V. M. Prabhakaran, C. Fragouli, and S. Diggavi, "Secret Message Capacity of Erasure Broadcast Channels with Feedback," in *IEEE Information Theory Workshop (ITW)*, 2011.

[12] L. Czap and C. Fragouli, "Secure Key Exchange in Wireless Networks," in *IEEE International Symposium on Network Coding (NETCOD)*, 2011.

[13] L. Buttyán, L. Czap, and I. Vajda, "Pollution Attack Defense for Coding Based Sensor Storage," in *IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC)*, 2010.

[14] ——, "Securing Coding Based Distributed Storage in Wireless Sensor Networks," in *IEEE International Workshop on Wireless and Sensor Networks Security (WSNS)*, 2008, pp. 821 – 827.

Other:

[15] L. Czap and I. Vajda, "Signatures for multi-source network coding," Cryptology ePrint Archive, Report 2010/328, 2010.