INFORMATION-THEORETIC SECURITY UNDER COMPUTATIONAL, BANDWIDTH, AND RANDOMIZATION CONSTRAINTS

A Dissertation
Presented to
The Academic Faculty

By

Rémi A. Chou

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
in
Electrical and Computer Engineering

School of Electrical and Computer Engineering
Georgia Institute of Technology
August 2015

INFORMATION-THEORETIC SECURITY UNDER COMPUTATIONAL, BANDWIDTH, AND RANDOMIZATION CONSTRAINTS

Approved by:

Dr. Faramarz Fekri, Committee Chair
Professor, School of ECE
Georgia Institute of Technology

Dr. Matthieu R. Bloch, Advisor
Associate Professor, School of ECE
Georgia Institute of Technology

Dr. John R. Barry
Professor, School of ECE
Georgia Institute of Technology

Dr. Steven W. McLaughlin
Professor, School of ECE
Georgia Institute of Technology

Dr. Prasad Tetali
Professor, School of Mathematics
Georgia Institute of Technology

Date Approved: July 2015

ACKNOWLEDGMENTS

I have been very fortunate to have Dr. Matthieu Bloch as a teacher and as a thesis advisor. I have greatly benefited from his wise guidance and have learned a lot from his approach to research. I am grateful for his support and for his confidence in me, which very much contributed to making these years of doctoral studies an enjoyable experience.

I would next like to thank Dr. Emmanuel Abbe and Dr. Jörg Kliewer for their very valuable comments and suggestions on my research, as well as for their support.

I also wish to thank Dr. Faramarz Fekri, Dr. John Barry, Dr. Steven McLaughlin, and Dr. Prasad Tetali for being part of my dissertation committee.

My colleagues largely contributed to providing me with a productive and pleasant work environment. Many thanks to the Communication Architectures Research Group, including Alex, Guillaume, Ishaq, and Keerthi, and to all my labmates of the Communication and Information Theory Lab, including Elnaz, Hyoungsuk, Sarwat, and Ubaid.

I spent memorable moments that are indissociable from this period of doctoral studies with many remarkable persons, mostly from GTL, Supélec, Georgia Tech, FSK, and SKA, and in particular with Andréas and Bilal, in Metz, Alex and Guillaume, inside and outside the lab in Metz and Atlanta, Matthieu, Chloé, Spyros, Stef, Marco, and Elnaz, in Atlanta. I wish to extend my greetings to all of them.

I naturally also wish to thank my family for their excellent support.

TABLE OF CONTENTS

LIST OF FIGURES

NOTATION AND ABBREVIATIONS

Notation

$\mathbb{R}$	set of real numbers
$\mathbb{N}$	set of natural numbers
$\mathcal{X}$	generic alphabet or set
$\lvert\mathcal{X}\rvert$	cardinality of $\mathcal{X}$
x	generic element of alphabet $\mathcal{X}$
$\lvert x\rvert$	absolute value of x
$[x]^+$	positive part of x, i.e., $\max(0,x)$
$\lceil\cdot\rceil$	ceiling function
$\lfloor\cdot\rfloor$	floor function
$[\![x,y]\!]$	set of integers between $\lfloor x\rfloor$ and $\lceil y\rceil$
$\mathbb{1}$	indicator function
X	random variable implicitly defined on alphabet $\mathcal{X}$
p_X	probability distribution of random variable X
$p_{X\mid Y}$	conditional probability distribution of X given Y
$X \sim p_X$	random variable X with distribution p_X
$\mathcal{N}(\mu,\sigma^2)$	Gaussian distribution with mean μ and variance σ^2
$\mathcal{B}(p)$	Bernoulli distribution with parameter $p \in [0,1]$

$\mathbb{V}(\cdot,\cdot)$	variational distance between two distributions
$\mathbb{D}(\cdot\|\cdot)$	divergence between two distributions
$\mathrm{Var}(X)$	variance of random variable X
$\mathbb{E}_X$	expected value over random variable X
$H(X)$	Shannon entropy of discrete random variable X
H_b	binary entropy function
H_∞	min-entropy of discrete random variable X
$h(X)$	differential entropy of continuous random variable X
$I(X;Y)$	mutual information between random variables X and Y
$\overline{\lim}_{x\to x_0} f(x)$	limit superior of $f(x)$ as x goes to x_0
$\underline{\lim}_{x\to x_0} f(x)$	limit inferior of $f(x)$ as x goes to x_0

Abbreviations

BMS	Binary memoryless source
BSC	Binary symmetric channel
CMS	Continuous memoryless source
DBMS	Degraded binary memoryless source
DMS	Discrete memoryless source
MS	Memoryless source
SK	Secret-key
WSK	Wiretap secret-key

SUMMARY

The objective of this thesis is to develop and analyze coding schemes for information-theoretic security, which could bridge a gap between theory and practice. We focus on two fundamental models for information-theoretic security: secret-key generation for a source model and secure communication over the wire-tap channel. Many results for these models only prove the existence of codes, and few attempts have been made to design practical schemes. The schemes we would like to propose should account for practical constraints to avoid oversimplifying the problems. From a practical point of view, many constraints should be taken into account; we, however, restrict our study to the following ones: (i) computationally bounded legitimate users, in particular, one should not solely rely on proofs showing the existence of codes with exponential complexity in the block-length; (ii) a rate-limited public communication channel for the secret-key generation model, to account for bandwidth constraints; (iii) a non-uniform and rate-limited source of randomness at the encoder for the wire-tap channel model, since a perfectly uniform and rate-unlimited source of randomness might be an expensive resource. The main contributions of this thesis are coding schemes for secret-key generation and wire-tap channel models that satisfy the aforementioned constraints.

CHAPTER 1
INTRODUCTION

Secure communications and data privacy in large-scale networks have become a major concern with economical and safety issues at stake not only for individuals but also for companies, or governments. For instance, the increasing amount of personal data collected in databases is threatening users privacy, while the nature of the transmission medium of wireless communication, over which a significant amount of sensitive data is carried, is prone to malicious and undetected acts of eavesdropping.

Information-theoretic security aims at enhancing security and privacy-preserving properties of future and emerging information and communication systems. It includes providing solid mathematical foundations, through an analysis of the fundamental security and communication limits under an information-theoretic framework, as well as practical solutions. Taking the example of wireless communication, all upper layers of typical communication protocols already have their own set of cryptographic primitives, whereas the physical layer, at which channel coding is implemented and on which all others layers rely, is currently not intrinsically secured against eavesdropping. Information-theoretic security could be used to secure this layer and thus enhance the security of wireless communication protocols. Unlike complexity-based cryptography, it would also have the advantage of making no assumption on the computational power of adversaries, and thus of being everlastingly usable.

Nevertheless, little progress has been made toward a widespread use of systems implementing physical-layer security since the introduction of information-theoretic security. To date, information-theoretic security still needs to be further explored and better understood to pursue this goal. Specifically, the need exists for models with as few simplifying assumptions as possible and for constructive schemes.

1.1 Background on information-theoretic security

Cryptography has a long history and is intimately related to the development of communication systems and recent wars. Prior to the 19th century, most cryptographic techniques relied on alphabet substitutions or letter permutations and transpositions. However, at the beginning of the 19th century and the beginning of the 20th century, the inventions of telegraphy and wireless telegraphy, respectively, exacerbated the need for better encryption. Moreover, these inventions played a key role during the American Civil-War, World War I, and World War II. Consequently, new cryptographic techniques emerged.

A major breakthrough in cryptography occurred in the aftermath of World War II, when Shannon publicly released its work on secure communication in 1949 [1]. As illustrated in Figure 1, Shannon formalized the problem as follows. Consider two legitimate users, Alice and Bob, who share a secret key K. Alice aims at securely sending a message M to Bob, called plaintext, over a public error-free channel. On Alice's side, an encoder takes as inputs the message M and the secret-key K. The output of the encoder C is publicly transmitted by Alice to Bob and is called the ciphertext. The objective for Bob is to recover M from the ciphertext C and the secret-key K, whereas Eve should obtain no information about M given C. The latter statement is quantified in an information-theoretic sense as $I(M;C) = 0$, where I is Shannon's mutual information. Shannon proved that an encoder exists such that the message M can be securely transmitted if the key K satisfies three conditions. Specifically, the key K must be distributed according to a uniform distribution, this key must be at least as large as the entropy of the message M, and finally, this key must only be used once.

The one-time pad, illustrated in Figure 2, is an instance of secure communication under this model. The encoding consists in performing the modulo-2 addition, between the plaintext M and the secret-key K. The result is a uniform sequence from

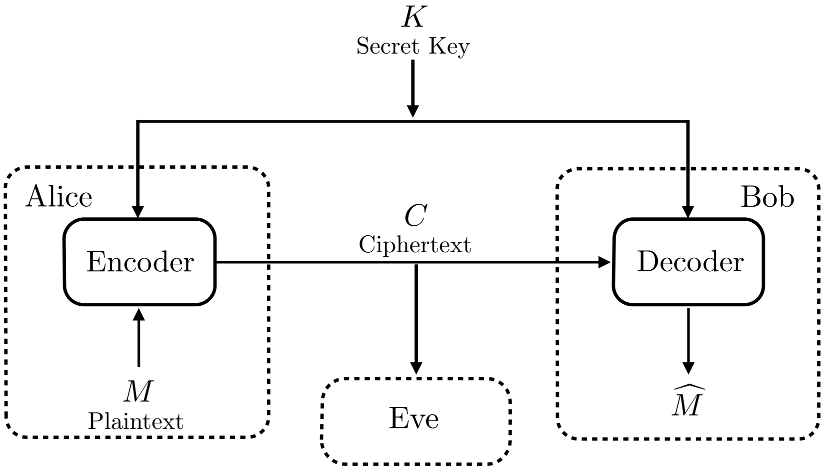


Figure 1. Model for symmetric encryption.

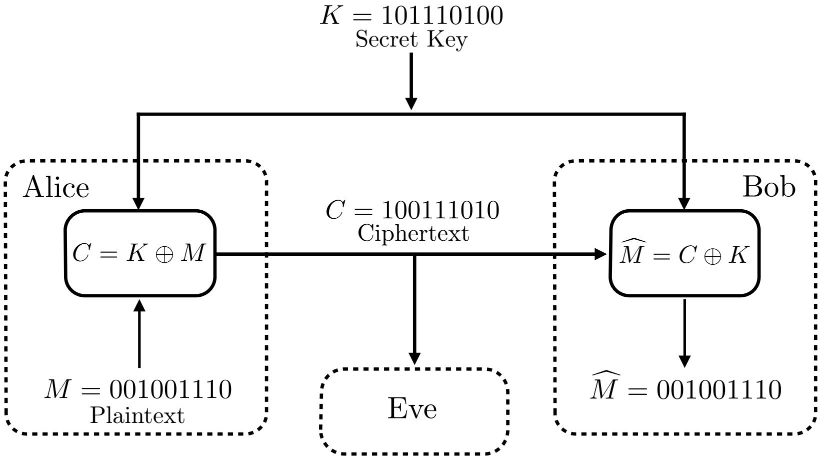


Figure 2. An instantiation of symmetric encryption: the one-time pad.

which Eve gains no information about M. As for Bob, he recovers M by performing the modulo-2 addition of the ciphertext C with the secret-key K.

The problem of secure communication against a computationally unbounded eavesdropper is thus solved by the one-time pad. However, the difficulty is transferred to the problem of secret-key generation, which still remains a challenging task in itself. In fact, no efficient method for information-theoretic secret-key generation is currently known.

1.2 Secret-key generation

Information-theoretic secret-key generation protocols are a fundamental primitive for information-theoretic security, as explained in the previous section. We first formally describe such problems and then provide a literature survey.

1.2.1 Model

Information-theoretic secret-key generation was first formally introduced in [2, 3], and can abstractly be described as follows in a multi-terminal configuration [4]. Let $m \geqslant 2$ be the number of terminals that wish to generate a common secret-key. We set $\mathcal{M} \triangleq [\![1, m]\!]$, and let $\mathcal{Z}$ and $\mathcal{X}_i$, for $i \in \mathcal{M}$, be arbitrary finite alphabets. We define $\mathcal{X}_{\mathcal{M}} \triangleq (\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_m)$ and consider a discrete memoryless source $(\mathcal{X}_{\mathcal{M}}\mathcal{Z}, p_{X_{\mathcal{M}}Z})$, where $X_{\mathcal{M}} \triangleq (X_1, X_2, \dots, X_m)$. For $i \in \mathcal{M}$, Terminal i observes n realizations of the component X_i of $(\mathcal{X}_{\mathcal{M}}\mathcal{Z}, p_{X_{\mathcal{M}}Z})$, whereas an eavesdropper observes n realizations of the component Z. The source is assumed to be outside the control of all parties, but its statistics are known to all parties. Communication is allowed between terminals over an authenticated[1] noiseless public channel with communication rate $R_p \in \mathbb{R}^+ \cup \{+\infty\}$. All the public inter-terminal communications are collectively denoted by $\mathbf{F}$ and are subject to the constraint $H(\mathbf{F}) \leqslant NR_p$. The case of $m = 2$ legitimate users and two-way one-round public communication with $\mathbf{F} \triangleq (A, B)$ is illustrated in Figure 3.

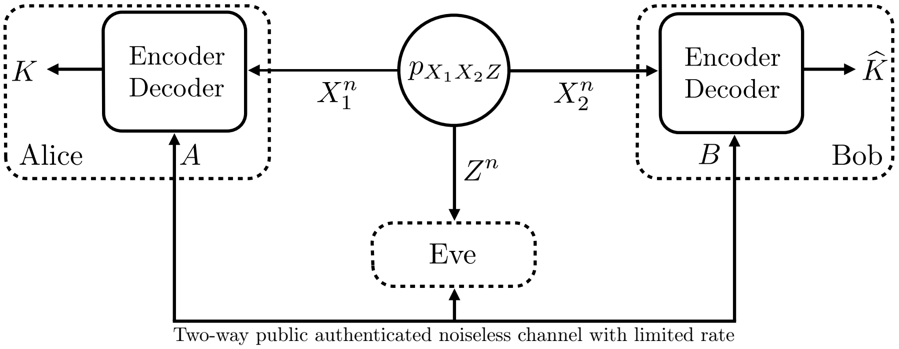


Figure 3. Model for secret-key generation: two-users case and two-way one-round public communication with $\mathbf{F} \triangleq (A, B)$.

[1] In others words, Eve has total access to Alice and Bobs messages, but cannot tamper with the messages over the channel.

The rules by which the legitimate users compute the messages they exchange over the public channel and agree on a key define a secret-key generation strategy. The performance of a secret-key generation strategy that allows the m terminals to agree on the key K is measured in terms of the following metrics.

- The average probability of error between the keys:

$$\lim_{n\to\infty} \mathbb{P}[\exists i \in \mathcal{M}, K \neq K_i] = 0,$$

- The information leakage to the eavesdropper:

$$\lim_{n\to\infty} I(K; Z^n\mathbf{F}) = 0,^{2} \tag{1}$$

- The uniformity of the key:

$$\lim_{n\to\infty} \log\lceil 2^{nR}\rceil - H(K) = 0.$$

Moreover, the maximum number of secret-key bits per observation is called the wiretap secret-key (WSK) capacity. This quantity is simply called the secret-key (SK) capacity for the special case $Z = \emptyset$, i.e., when the eavesdropper has no correlated observation of the source.

We briefly comment on two hypotheses made in this model.

- We assume the existence of a memoryless source with known but uncontrollable statistics. In practice, it can, for instance, be obtained in a wireless communication setting [5–7]. Assume that we denote $c_{A\to B}$ the channel from Alice to Bob, $c_{B\to A}$ the channel from Bob to Alice, $c_{A\to E}$ the channel from Alice to Eve, and $c_{B\to E}$ the channel from Bob to Eve. We can then set X as the channel gain of $c_{A\to B}$, Y as the channel gain of $c_{B\to A}$, and Z as the pair of channel gains for $(c_{A\to E}$,$c_{B\to E})$.

[2] This condition corresponds to strong secrecy, whereas $\lim_{n\to\infty} I(K; Z^n\mathbf{F})/n = 0$ corresponds to weak secrecy.

- We also assume the existence of an authenticated public channel. In practice, a solution to ensure authentication would be to have the legitimate users share a secret sequence of random bits. This solution is acceptable since the size of this secret sequence can be chosen in the order of the logarithm of the length of the message [2, 8], which is negligible compared to the length of the key generated.

We now list additional constraints that will be taken into account in this thesis to avoid oversimplifying the model.

- **Computationally bounded legitimate users**. A secret-key generation scheme should be implementable by computationally bounded users, and not solely rely on a proof showing existence of codes with exponential complexity in the number of observations n.

- **Rate-limited public communication**. The public communication constraint R_p should be considered finite. Indeed, channels with unlimited communication rate do not exist. Moreover, we can expect sharp performance degradation for applications in which strong bandwidth constraints hold, as for instance in a wireless sensor network. We will see that the main difficulty introduced by this constraint is the need for vector quantization of the source observation.

From a practical point of view, many other constraints should be taken into account, such as finite block-length, unknown eavesdropper's statistics, or arbitrary eavesdropper's alphabet. Proposing a scheme that is able to account for all these constraints is a challenging problem. A first step towards this goal will consist in dealing with subsets of these constraints.

We conclude this section by pointing out other models for secret-key generation. The model we have introduced is called the "source model" for secret-key generation. A variant of this model is the "channel model" for secret-key generation, in which the source is partially controlled by one of the legitimate users. Upper and lower bounds of the WSK capacity for the channel model in different settings, often derived from bounds of WSK capacity for the source model or obtained with similar techniques, can be found in [2,3,9,10]. Secret-key generation is also studied in the quantum setting, and relies on arguments of a totally different nature, such as quantum entanglement or quantum superposition. Quantum secret-key generation first appeared in [11,12] and has since then attracted interest, as well [13–15].

1.2.2 Literature Survey

Closed-form expressions and bounds for the WSK capacity with $m = 2$, i.e. two legitimate users, have been established for a large variety of models [2–4,9,10,16–22]. However, usual achievability proofs only prove existence of codes, and do not always provide direct insight into the design of practical key-generation strategies.

The only exception is sequential strategies when rate-unlimited public communication is considered. The main benefit of such strategies is to successively deal with reliability and secrecy by means of a reconciliation protocol and privacy amplification, respectively. Indeed, reconciliation can be efficiently implemented with LDPC codes [23] and privacy amplification can be performed with extractors [24,25]. Note that uncertainty about the eavesdropper's statistics is addressed, since extractors are universal and can thus be chosen such that security holds when the statistics of Z are known to belong to a given set $\mathcal{S}_Z$. Specifically, the length of the output of the extractors is chosen such that security holds for p_Z^*, where $p_Z^* \triangleq \arg\max_{p_Z \in \mathcal{S}_Z} I(X;Z)$. Moreover, in the case of non-memoryless source, [26] addresses the finite-length regime with a sequential strategy. A finite-length analysis of privacy amplification is also provided in [27], and in [28] by means of malleable extractors [28,29].[3] Note also that, for a related model,[4] [30] deals with computationally bounded legitimate users, and

[3] The model considered in [28] is the following. The legitimate users observe the same component of a non-memoryless source, while the eavesdropper observes a correlated component of the source. Moreover, two-way one-round public communication over an unauthenticated channel with unlimited capacity is assumed.

[4] The model considered in [30] is the following. The legitimate users observe the components of

the finite-length regime.

To the best of our knowledge, only non-constructive schemes deal with rate-limited public communication. For discrete memoryless sources, the WSK capacity with one-way rate-limited public communication, and bounds for the WSK capacity with two-way one-round rate-limited public communication are provided in [4]. The WSK capacity for one-way rate-limited public communication is extended to the case of continuous degraded memoryless sources in [20].

For a multi-terminal setting, that is, the number m of legitimate users is such that $m > 2$, upper and lower bounds for the WSK capacity are derived in [4, 9, 21, 31]. The analysis of such a setting is considerably more involved than the case of two legitimate users. Moreover, most results only hold when the eavesdropper observes the inter-terminal public communication, but has no side observation Z of the source observed by the legitimate users. Again, for these settings, the proofs in the literature only provide existence of codes but no explicit code constructions. We can, though, mention the exception of [32,33], that are based on explicit algorithms for tree packing, and [34], that relies on channel coding. The protocol proposed in the latter reference is, however, computationally intractable, because it requires standard arrays that grow exponentially with the number of source observations.

1.3 Communication over a wire-tapped channel

In this section, we discuss a model related to secret-key generation, called the wire-tap channel. We first formally introduce the problem and then review relevant known results.

1.3.1 Model

Communication over the wire-tap channel model can be seen as a secret-key generation problem for the channel model, in which the source of randomness stems from

a non-memoryless source, that are close with respect to certain metrics, while the eavesdropper has no observations of the source. Moreover, one-way public communication over an unauthenticated channel is assumed.

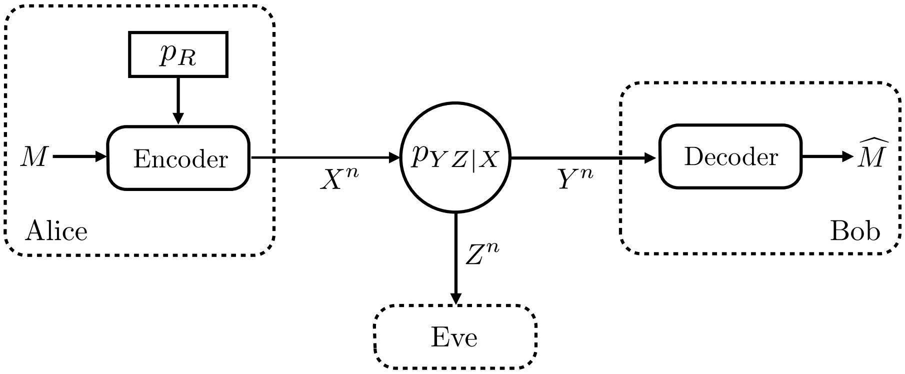


Figure 4. The wire-tap channel model.

the transmission medium, in which there is no error-free public channel for communication, and in which the key is fixed by the transmitter ahead of time. Because of this distinction, coding mechanisms for the wire-tap channel and secret-key generation are very different. The wire-tap channel model was first introduced in [35], and can be described as follows. As illustrated in Figure 4, consider two legitimate users, Alice and Bob, connected by a communication channel, which is abstracted by the conditional probability distribution $p_{Y|X}$. Consider also an eavesdropper, Eve, which observes the communication of Alice and Bob through a channel defined by the conditional probability distribution $p_{Z|X}$. Alice aims at secretly transmitting to Bob a message M, which is encoded in X^n, and received as Y^n and Z^n by Bob and Eve, respectively. It is assumed that Alice has access to a source of uniform randomness $(\mathcal{R}, p_R)$ to randomize the encoding of M. Similar to the secret-key generation model described in Section 2.2.1, the performance of a coding scheme for the wiretap channel is measured in terms of the following metrics.

- The average probability of error between M and $\widehat{M}$:

$$\lim_{n\to\infty} \mathbb{P}[M \neq \widehat{M}] = 0,$$

- Information leakage, measured by the mutual information between the message

M and all the information available to Eve through Z^n:

$$\lim_{n \to \infty} I(M; Z^n) = 0.$$

The highest rate at which Alice can securely transmit messages to Bob is called the secrecy capacity. It is shown in [35], that the secrecy capacity is strictly positive if Bob's channel, $p_{Y|X}$, is less noisy than Eve's channel, $p_{Z|X}$ – see, for instance, [23, Proposition 3.6] for a formal description of "less noisy". In the latter case, unlike in Section 2.1 for symmetric encryption, secure communication is possible without the need of a shared secret-key for the legitimate users, by harnessing the communication noise introduced by the transmission medium.

We now list additional constraints that will be taken into account in the thesis to avoid oversimplifying the model.

- **Computationally bounded legitimate users**. A coding scheme for the wire-tap channel should be implementable by computationally bounded users, and not solely rely on a proof showing existence of code with exponential complexity in the number of observations n.

- **Non-uniform and rate-limited source of randomness**. As depicted in Figure 4, an implicit assumption made in the original wire-tap channel model [35] is the availability at the encoder of a perfectly uniform and rate-unlimited source of randomness. Such a resource may be expensive or not available at all. Consequently, we introduce the constraint that only a possibly non-uniform and rate-limited source of randomness is available at the encoder.

- **Bandwidth efficiency**. The secrecy capacity, which is always less than the capacity, suggests that secrecy can only be achieved at the cost of reducing communication rates. This decrease in achievable communication rates could be a factor that may hinder the adoption of physical-layer security schemes in communication systems.

Similar to secret-key generation schemes, from a practical point of view, other constraints, such as finite block-length, unknown eavesdropper's statistics, or arbitrary eavesdropper's alphabet, should be taken into account. We will, however, restrict our analysis to the three aforementioned constraints.

1.3.2 Literature survey

Although closed-form expressions for the secrecy capacity are known [35, 36], traditional achievability proofs only prove the existence of codes, and, again, do not always provide direct insight into the design of practical coding schemes.

Recent works have tackled the constraint of computationally bounded legitimate users with polar codes [37], when symmetric channels are assumed [38, 39]. We can also mention a constructive scheme with efficiently invertible extractors based on finite field multiplication [40, 41] for symmetric or additive channels, and LDPC-based constructions [42–44] for erasure channels. However, none of these solutions allows the treatment of arbitrary channels.

Rate-limited randomness at the encoder has been studied in [45, 46]. In [45], the authors precisely analyze the trade-off between the rates of secret message, public message, and local uniform randomness in the broadcast channel with confidential messages. Moreover, [46, 47] investigates the case of non-uniform randomness at the encoder.

Bandwidth efficiency can be improved by multiplexing public and confidential messages. This idea implicitly appears in the original work of Csiszár and Körner [36], and is explicitly formalized in [48, 49]. In [50], the authors analyze the possibility of guaranteeing secrecy with dependent messages and non-uniform randomization.

The constraints of unknown eavesdropper's statistics is, for instance, addressed in [51–53]. [52] handles a Gaussian multiple-input multiple-output setting and shows that if the eavesdropper's statistics are unknown, one can obtain a strictly positive secrecy capacity when the legitimate users have more antennas than the eavesdropper. However, no constructive scheme is known in this case. The case of an arbitrary eavesdropper's alphabet and arbitrary wiretap channels can be treated with information-spectrum methods using non-constructive schemes, see for instance [54].

A non-asymptotic treatment of the wire-tap channel model by means of non-constructive schemes is also possible, see for instance [14].

Finally, the assumption regarding a source of uniform randomness available at the encoder is partially relaxed with non-constructive schemes in [46, 47, 55], where non-uniform or rate-limited sources of randomness are considered.

1.4 Outline of the dissertation and related publications

Chapters 2 and 3 are related to secret-key generation, while Chapters 4 and 5 are related to secure communication over a wire-tap channel. Each chapter may be read independently from the other chapters.

Chapter 2 considers secret-key generation with two-way one-round rate-limited communication between two legitimate users. Specifically, we study a sequential key-generation strategy that handles reliability and secrecy successively, and show its optimality (under the assumption of degraded sources for two-way communication). We, however, show that although reliability and secrecy can be treated successively, they might not always be treated independently, thereby exhibiting the limits of sequential strategies to rate-limited public communication. Chapter 2 is based on the results obtained in the following references:

- **R. Chou** and M. Bloch. Separation of Reliability and Secrecy in Rate-Limited Secret-Key Generation, in *IEEE transactions on Information Theory*, Vol. 60, no. 8. 2014.

- **R. Chou** and M. Bloch. One-Way Rate-Limited Sequential Key-Distillation. Proc. of *IEEE International Symposium on Information Theory (ISIT)*. 2012.

Chapter 3 considers polar coding for different models of secret-key generation. Specifically, we propose secret-key capacity-achieving and low-complexity schemes for the following models: (i) the degraded binary memoryless source (DBMS) model with rate-unlimited public communication, (ii) the DBMS model with one-way rate-limited public communication, (iii) the 1-to-m broadcast model, (iv) the Markov tree model with uniform marginals, (v) several models for biometric systems. Chapter 3 is based on the results obtained in the following references:

- **R. Chou**, M. Bloch, and E. Abbe. Polar Coding for Secret-Key Generation. Accepted to *IEEE transactions on Information Theory*. May, 2015.
  Available at http://arxiv.org/abs/1305.4746

- **R. Chou**, M. Bloch, and E. Abbe. Polar Coding for Secret-Key Generation. Proc. of *IEEE Information Theory Workshop (ITW)*. 2013.

Chapter 4 considers a source-channel coding scheme for the wiretap channel. We show that multiplexing unprotected and protected data allows, first, to avoid the necessity of additional randomness at the encoder and, second, to efficiently use the bandwidth available between the legitimate users. Specifically, the overall communication rate of the same channel without secrecy constraints is maintained. The scheme leverage results about lossless source coding with uniform encoder output. Chapter 4 is based on the results obtained in the following references:

- **R. Chou**, M. Bloch, B. Vellambi, and J. Kliewer. Source-Channel Coding Schemes for Achieving Strong Security at Negligible Cost. To be submitted to *IEEE transactions on Information Theory*. 2015.

- **R. Chou** and M. Bloch. Uniform Distributed Source Coding for the Multiple Access Wiretap Channel. Proc. of *IEEE Conf. on Communications and Network Security (CNS)*. 2014.

- **R. Chou** and M. Bloch. Data Compression with Nearly Uniform Output. Proc. of *IEEE International Symposium on Information Theory (ISIT)*. 2013.

Chapter 5 considers polar coding for the wiretap channel model. Specifically, we propose a low-complexity and secrecy capacity achieving scheme. Our scheme extends previous work by using an optimal rate of uniform randomness in the stochastic encoder, and avoiding assumptions regarding the symmetry or degraded nature of the channels. Moreover, we describe a close conceptual connection between our coding scheme and a random binning proof of the secrecy capacity region. An extension to the broadcast channel with confidential messages is also proposed. Chapter 5 is based on the results obtained in the following references:

- **R. Chou** and M. Bloch. Polar Coding for the Broadcast Channel with Confidential Messages and Constrained Randomization. Submitted to *IEEE transactions on Information Theory*. November, 2014.
  Available at http://arxiv.org/abs/1411.0281

- **R. Chou**, M. Bloch. Polar Coding for the Broadcast Channel with Confidential Messages. Proc. of *IEEE Information Theory Workshop (ITW)*. 2015.

CHAPTER 2

SEPARATION OF RELIABILITY AND SECRECY IN SECRET-KEY GENERATION

2.1 Summary

For a discrete or a continuous source model, we study in this chapter the problem of secret-key generation with two-way one-round of rate-limited public communication between two legitimate users. Although we do not provide new bounds on the wiretap secret-key (WSK) capacity for the discrete source model, we use an alternative achievability scheme that may be useful for practical applications. As a side result, we conveniently extend known bounds to the case of a continuous source model. Specifically, we consider a sequential key-generation strategy, that implements a rate-limited reconciliation step to handle reliability, followed by a privacy amplification step performed with extractors to handle secrecy. We prove that such a sequential strategy achieves the best known bounds for the rate-limited WSK capacity (under the assumption of degraded sources in the case of two-way communication). However, we show that, unlike the case of rate-unlimited public communication, achieving the reconciliation capacity in a sequential strategy does not necessarily lead to achieving the best known bounds for the WSK capacity. Consequently, reliability and secrecy can be treated successively but not independently, thereby exhibiting a limitation of sequential strategies for rate-limited public communication. Nevertheless, we provide scenarios for which reliability and secrecy can be treated successively and independently, such as the two-way rate-limited SK capacity, the one-way rate-limited WSK capacity for degraded binary symmetric sources, and the one-way rate-limited WSK capacity for Gaussian degraded sources. This chapter is based on the results obtained in [56, 57].

2.2 Introduction

A sequential key-generation strategy consists of (i) a reconciliation step, during which Alice and Bob communicate over the public channel to agree on a common bit sequence, which might not be totally hidden from Eve, (ii) a privacy amplification step, during which Alice and Bob apply a deterministic function to their shared sequence to generate their common secret key, this time completely unknown from Eve. The main benefit of sequential key-generation strategies is to separate how one deals with reliability and secrecy,[1] and thus to provide a perhaps more practical key-generation design. Indeed, reconciliation can be efficiently implemented with LDPC codes [58,59] and privacy amplification can be performed with extractors [24, 25]. While sequential key-generation is studied in [23, 25] for a public channel of unlimited capacity, we focus on the performance of sequential key-generation strategies in the case of rate-limited public communication.[2]

Although, we do not improve the rate-limited WSK capacity bounds for the discrete source model, we provide an achievability scheme that might be easier to translate into practical designs. Specifically, we show that sequential strategies, that are known to be optimal for rate-unlimited public communication, are also optimal for rate-limited communication. We, however, also qualify the robustness of sequential strategies to rate-limited public communication, as we show in this case that it may not be optimal to achieve the reconciliation capacity in a sequential strategy. That is, reliability and secrecy can be handled successively but not necessarily independently, thereby limiting the coding scheme flexibility. The main results of this chapter are:

[1] We mean that the key-generation can be performed by the succession of two protocols, one, free from any secrecy constraint, dealing with reliability, and the other dealing with secrecy. A stronger result would be that optimizing both protocols independently, in a sense defined in Section 2.3.4, leads to the best possible key-generation strategy. In Section 2.5, we prove that this stronger result holds in some scenarios.

[2] Note that the achievability scheme of [21, Theorem 4.1], which only holds for Gaussians sources and when there is no side information at the eavesdropper, is very close to the sequential approach that we study, even though their model is different in that it deals with a quantized source and unrestricted public communication.

- an alternative achievability scheme that separates reliability and secrecy by means of a reconciliation protocol and a privacy amplification step performed with extractors, which achieves

  (i) the best known bound of the two-way one-round rate-limited WSK capacity for degraded sources in Theorem 2.4.3;

  (ii) the one-way rate-limited WSK capacity in Theorem 2.4.4;

  (iii) the two-way one-round rate-limited SK capacity (no side information at the eavesdropper) in Theorem 2.4.5;

  As a side result, we extend the bounds for a discrete source model in [16], to the case of a continuous source model in Corollary 2.4.2 (the case of the one-way rate-limited WSK capacity is treated in [60], but only for degraded sources) ;

- scenarios for which achieving the reconciliation capacity is optimal in a sequential key-generation strategy, as it is not necessarily the case in general when constraints are imposed on public communication. Such results are important to obtain a flexible coding scheme; Specifically, we treat the case of

  (i) the two-way rate-limited SK capacity in Section 2.5.1;

  (ii) the one-way rate-limited WSK capacity for degraded binary symmetric sources in Section 2.5.2;

  (iii) the one-way rate-limited WSK capacity for degraded Gaussian sources in Section 2.5.3;

  As side results, we obtain a characterization of the rate-limited reconciliation capacity in Proposition 2.5.2, which corresponds to the best trade-off between the length of the sequence shared by Alice and Bob after reconciliation and the quantity of information publicly exchanged; we also obtain a closed-form expression of the one-way WSK capacity for degraded binary symmetric sources with Proposition 2.5.4, as illustrated in Example 2.5.2.

Our proofs techniques mainly rely on the analysis of randomness extraction with extractors, Wyner-Ziv coding, and a fine analysis with robust typicality [61] to extend the discrete case to a continuous setting. The determination of the one-way WSK capacity for degraded binary symmetric sources relies on perhaps less standard techniques, as we use the Krein-Milman Theorem to simplify a convex optimization problem under convex constraints.

The remainder of the chapter is organized as follows. In Section 2.3, we introduce the problem and provide some background on the topic. Specifically, we formally introduce the problem studied in Section 2.3.1, and recall known bounds for the secret-key capacity in Section 2.3.2. In Section 2.3.3, we describe the two steps of a sequential strategy and recall known bounds achieved by such a strategy. In Section 2.3.4, we introduce the notion of independence between the two steps of a sequential strategy, when constraints are imposed on public communication. In Section 2.4, we prove that the sequential application of reconciliation and privacy amplification with extractors is an optimal key-generation strategy. In Section 2.5, we provide scenarios for which these two phases can be treated independently of each other. Specifically, we provide the case of the two-way SK capacity in Section 2.5.1, the one-way WSK capacity for degraded binary symmetric sources in Section 2.5.2, and the one-way WSK capacity for degraded Gaussian sources in Section 2.5.3. All proofs are gathered in the appendices to streamline presentation.

2.3 Problem statement and background

2.3.1 Model

We consider in this chapter a special case of the model introduced in Section 1.2.1. As illustrated in Figure 5, two legitimate users, Alice and Bob, and one eavesdropper, Eve, observe the realizations of a memoryless source (MS) $(\mathcal{X}\mathcal{Y}\mathcal{Z}, p_{XYZ})$, that can

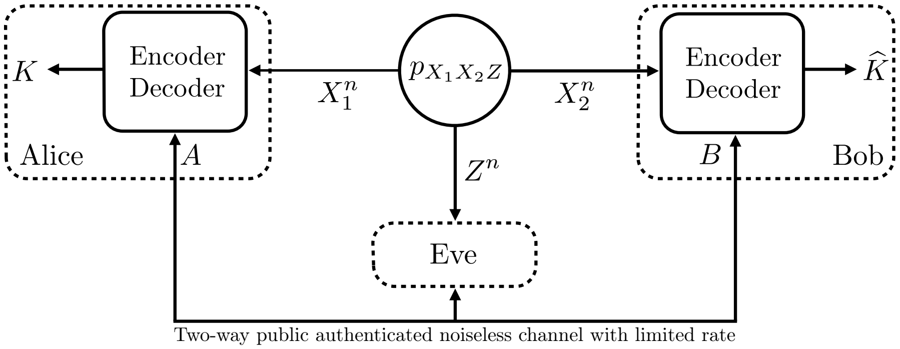


Figure 5. Source model for secret-key generation.

be either discrete (DMS) or continuous (CMS). The three components X, Y and Z, are observed by Alice, Bob, and Eve, respectively. The MS is assumed to be outside the control of all parties, but its statistics are known. Alice and Bob's objective is to process their observations and agree on a key K, about which Eve should have no information. We assume a two-way one-round communication between Alice and Bob, that is, we suppose that Alice first sends a message to Bob, and that in return Bob sends a message to Alice.[3] We also assume that the messages are exchanged over an authenticated noiseless public channel with limited rate; in others words, Eve has total access to Alice and Bob's messages, but cannot tamper with the messages over the channel. We now formally define a key-generation strategy.

Definition 2.3.1. *A $\left(2^{nR}, n, R_1, R_2\right)$ key-generation strategy $\mathcal{S}_n$ for a source model with MS $(\mathcal{X}\mathcal{Y}\mathcal{Z}, p_{XYZ})$ consists of*

- *a key alphabet $\mathcal{K} = [\![1, 2^{nR}]\!]$;*
- *two alphabets $\mathcal{A}$, $\mathcal{B}$ respectively used by Alice and Bob to communicate over the public channel;*
- *two encoding functions $f_0 : \mathcal{X}^n \to \mathcal{A}$, $g_0 : \mathcal{Y}^n \times \mathcal{A} \to \mathcal{B}$;*
- *two functions $\kappa_a : \mathcal{X}^n \times \mathcal{B} \to \mathcal{K}$, $\kappa_b : \mathcal{Y}^n \times \mathcal{A} \to \mathcal{K}$;*

[3] One could also suppose that Bob is the one who sends messages, in which case one only needs to exchange the role of X and Y in the following.

and operates as follows.

- *Alice observes X^n while Bob observes Y^n;*
- *Alice transmits $A = f_0(X^n)$ subject to $H(A) \leqslant nR_1$;*
- *Bob transmits $B = g_0(Y^n, A)$ subject to $H(B) \leqslant nR_2$;*
- *Alice computes $K = \kappa_a(X^n, B)$ while Bob computes $\hat{K} = \kappa_b(Y^n, A)$.*

The performance of a $\left(2^{nR}, n, R_1, R_2\right)$ key-generation strategy $\mathcal{S}_n$ is measured in terms of the average probability of error between the key K generated by Alice and the key $\hat{K}$ generated by Bob

$$\mathbf{P}_e(\mathcal{S}_n) \triangleq \mathbb{P}[K \neq \hat{K} | \mathcal{S}_n],$$

in terms of the information leakage to the eavesdropper

$$\mathbf{L}(\mathcal{S}_n) \triangleq I(K; Z^n AB | \mathcal{S}_n),$$

and in terms of the uniformity of the key

$$\mathbf{U}(\mathcal{S}_n) \triangleq \log \left\lceil 2^{nR} \right\rceil - H(K | \mathcal{S}_n).$$

Definition 2.3.2. *A WSK rate R is achievable for a source model if there exists a sequence of $\left(2^{nR}, n, R_1, R_2\right)$ key-generation strategies $\left\{\mathcal{S}_n\right\}_{n \geqslant 1}$ such that*

$$\lim_{n\to\infty} \mathbf{P}_e(\mathcal{S}_n) = 0 \text{ (reliability)},$$

$$\lim_{n\to\infty} \mathbf{L}(\mathcal{S}_n) = 0 \text{ (strong secrecy)},$$

$$\lim_{n\to\infty} \mathbf{U}(\mathcal{S}_n) = 0 \text{ (strong uniformity)}.$$

Moreover, the WSK capacity of a source model with MS $(\mathcal{XYZ}, p_{XYZ})$ is the supremum of achievable WSK rates, and is denoted by C_{WSK}. In the following, we also consider situations in which the eavesdropper has access to the public messages exchanged by Alice and Bob, but has no side information Z^n. In such cases, the WSK capacity is simply called the secret-key (SK) capacity and is denoted by C_{SK}.

2.3.2 Known bounds for $C_{\mathbf{WSK}}$ and $C_{\mathbf{SK}}$

For convenience, we recall known results regarding the model described in Section 2.3.1. Note that these results only hold for DMS.

Theorem 2.3.1 ([16, Theorems 2.5, 2.6]). *Let $(\mathcal{X}\mathcal{Y}\mathcal{Z}, p_{XYZ})$ be a DMS.*

(a) For $R_1, R_2 \in \mathbb{R}^+$, the two-way one-round WSK capacity satisfies

$$C_{\mathrm{WSK}}(R_1, R_2) \geqslant R_{\mathrm{WSK}}(R_1, R_2),$$

where

$$R_{\mathrm{WSK}}(R_1, R_2) \triangleq \max_{U,V} \left([I(Y;U) - I(Z;U)]^+ + [I(X;V|U) - I(Z;V|U)]^+ \right)$$

subject to

$$\begin{aligned} &R_1 \geqslant I(X;U) - I(Y;U), \\ &R_2 \geqslant I(Y;V|U) - I(X;V|U), \\ &U — X — YZ,\ V — YU — XZ, \\ &|\mathcal{U}| \leqslant |\mathcal{X}| + 2, |\mathcal{V}| \leqslant |\mathcal{Y}|. \end{aligned}$$

(b) For $R_1 \in \mathbb{R}^+$, the one-way WSK capacity is

$$C_{\mathrm{WSK}}(R_1, 0) = \max_{U,V} (I(Y;V|U) - I(Z;V|U))$$

subject to

$$\begin{aligned} &R_1 \geqslant I(X;V) - I(Y;V), \\ &U — V — X — YZ, \\ &|\mathcal{U}|, |\mathcal{V}| \leqslant |\mathcal{X}| + 2. \end{aligned}$$

Corollary 2.3.1 ([16, Theorems 2.2, 2.3, 2.4]). *Let $(\mathcal{X}\mathcal{Y}, p_{XY})$ be a DMS.*

(a) For $R_1, R_2 \in \mathbb{R}^+$, the two-way one-round SK capacity is

$$C_{\mathrm{SK}}(R_1, R_2) = \max_{U,V} (I(Y;U) + I(X;V|U))$$

subject to
$$\begin{aligned}&R_1 \geqslant I(X;U) - I(Y;U),\\&R_2 \geqslant I(Y;V|U) - I(X;V|U),\\&U—X—Y,\ V—YU—X,\\&|\mathcal{U}|\leqslant |\mathcal{X}|+2, |\mathcal{V}|\leqslant |\mathcal{Y}|.\end{aligned}$$

(b) For $R_1 \in \mathbb{R}^+$, the one-way SK capacity is

$$C_{\text{SK}}(R_1, 0) = \max_U I(Y;U)$$

subject to
$$\begin{aligned}&R_1 \geqslant I(X;U) - I(Y;U),\\&U—X—Y,\\&|\mathcal{U}|\leqslant |\mathcal{X}|+1.\end{aligned}$$

2.3.3 Sequential strategy

In the following, we use the term sequential key-generation strategy, for a key-generation strategy consisting of the succession of a reconciliation protocol and a privacy amplification with extractors.

2.3.3.1 Reconciliation

During the reconciliation phase, Alice and Bob send messages to each other over an authenticated public channel with limited rate. Alice and Bob then process their observations to agree on a common bit sequence S. At this stage the sequence is not subject to any secrecy constraint. Formally, a two-way one-round rate-limited reconciliation protocol is defined as follows.

Definition 2.3.3. *Let $R_1, R_2 \in \mathbb{R}^+$. A rate-limited reconciliation protocol $\mathcal{R}_n(R_1, R_2)$, noted $\mathcal{R}_n$ for convenience, for a source model with MS $(\mathcal{XY}, p_{XY})$ consists of*

- *an alphabet $\mathcal{S} = [\![1, M]\!]$;*
- *two alphabets $\mathcal{A}$, $\mathcal{B}$ respectively used by Alice and Bob to communicate over the public channel;*

- *two encoding functions $f : \mathcal{X}^n \to \mathcal{A}$, $g : \mathcal{Y}^n \times \mathcal{A} \to \mathcal{B}$;*

- *two functions $\eta_a : \mathcal{X}^n \times \mathcal{B} \to \mathcal{S}$, $\eta_b : \mathcal{Y}^n \times \mathcal{A} \to \mathcal{S}$;*

and operates as follows

- *Alice observes X^n while Bob observes Y^n;*

- *Alice transmits $A = f(X^n)$ subject to $H(A) \leqslant nR_1$;*

- *Bob transmits $B = g(Y^n, A)$ subject to $H(B) \leqslant nR_2$;*

- *Alice computes $S = \eta_a(X^n, B)$ while bob computes $\hat{S} = \eta_b(Y^n, A)$.*

The reliability performance of a reconciliation protocol is measured in terms of the average probability of error

$$\mathbf{P}_e(\mathcal{R}_n) \triangleq \mathbb{P}[S \neq \hat{S} | \mathcal{R}_n].$$

In addition, since the reconciliation protocol, which generates the common sequence S, is followed by the privacy amplification step to generate a secret-key, it is desirable to leak as little information as possible over the public channel. As in [23] we define the reconciliation rate of a reconciliation protocol as

$$\mathbf{R}(\mathcal{R}_n) \triangleq \frac{1}{n} \left[H(S|\mathcal{R}_n) - H(AB|\mathcal{R}_n) \right].$$

Definition 2.3.4. *For a given (R_1, R_2), a reconciliation rate R is achievable, if there exists a sequence of rate-limited reconciliation protocols $\{\mathcal{R}_n\}_{n \geqslant 1}$ such that*

$$\lim_{n \to \infty} \mathbf{P}_e(\mathcal{R}_n) = 0 \text{ and } \varliminf_{n \to \infty} \mathbf{R}(\mathcal{R}_n) \geqslant R.$$

Moreover, the two-way one-round rate-limited reconciliation capacity $C_{\text{rec}}(R_1, R_2)$ of a MS $(\mathcal{X}\mathcal{Y}, p_{XY})$ is the supremum of achievable reconciliation rates.

Intuitively, the reconciliation capacity characterizes the best trade-off between the length of the sequence shared by Alice and Bob after reconciliation and the quantity of information publicly exchanged.

2.3.3.2 Privacy amplification

During the privacy amplification phase, Alice and Bob generate their secret key by applying a deterministic function, on which they publicly agreed ahead of time, to their common sequence S obtained after reconciliation. This phase is performed with extractors [62], which are functions that take as input a sequence of n arbitrarily distributed bits and output a sequence of k nearly uniformly distributed bits, using another input of d truly uniformly distributed bits. Define the min-entropy of a discrete random variable X as

$$H_\infty(X) \triangleq -\log\left(\max_x p_X(x)\right),$$

and, for two discrete random variables X and Y, the conditional min-entropy of X given Y as

$$H_\infty(X|Y) \triangleq \sum_y p_Y(y) H_\infty(X|Y=y).$$

The following theorem provides a lower bound on the size of the key, on which the legitimate users agree.

Theorem 2.3.2 ([25], [23, Theorem 4.6])**.** *Let $S \in \{0,1\}^n$ be the random variable that represents the common sequence shared by Alice and Bob, and let E be the random variable that represents the total knowledge about S available to Eve. Let e be a particular realization of E. If Alice and Bob know that*

$$H_\infty(S|E=e) \geqslant \gamma n, \text{ for some } \gamma \in]0,1[,$$

then there exists an extractor $g : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^k$ with $d \leqslant n\delta(n)$ and $k \geqslant n(\gamma - \delta(n))$, where $\delta(n)$ satisfies $\lim_{n\to+\infty} \delta(n) = 0$.
Moreover, if U_d is a random variable uniformly distributed on $\{0,1\}^d$ and Alice and Bob choose $K = g(S, U_d)$ as their secret key, then

$$H(K|U_d, E=e) \geqslant k - \delta^*(n),$$

with $\delta^*(n) = 2^{-\sqrt{n}/\log n}(k + \sqrt{n}/\log n)$.

Note that, the size d of the uniformly distributed input sequence is negligible, compared to n, so that the effect on the rate of public communication is negligible. Moreover, extractors that extract almost the entire min-entropy of the input S and require comparatively negligible amount of uniform randomness can be efficiently constructed [62].

2.3.3.3 Known results concerning sequential strategies

For a DMS, in the absence of rate constraint between Alice and Bob, i.e. $R_1, R_2 = +\infty$, [25], [23, Theorem 4.7] state that one can handle reliability and secrecy successively to achieve the WSK capacity $C_{\text{WSK}}(+\infty, +\infty)$, by means of a reconciliation step, and a privacy amplification step. Figure 6 schematically illustrates the role of each step in terms of information shared by each party. At the beginning of the protocol, we assume, without loss of generality [2, 63], that Bob has an advantage over Eve in terms of the amount of information he has about Alice's observations of the source. The reconciliation step aims at correcting the discrepancies between Alice's and Bob's observations. Hence, after this step, Alice and Bob share a common sequence S, while Eve has gained some information about S from the public communication that occurred during reconciliation. Finally, the privacy amplification step allows Alice and Bob to extract from S a shorter sequence K totally independent from Eve's total knowledge.

2.3.4 Independence between reconciliation and privacy amplification

In this section, we define a notion of independence between reconciliation and privacy amplification, when constraints hold on the public communication rate. As explained earlier, we would like to ensure that reliability and secrecy can be handled not only successively but also independently, to obtain a flexible coding scheme. We will show in the following section that the reconciliation capacity is given by the following proposition.

Proposition 2.3.1. *Let $(\mathcal{X}\mathcal{Y}, p_{XY})$ be a DMS. Let $R_1 \in \mathbb{R}_+$. The reconciliation*

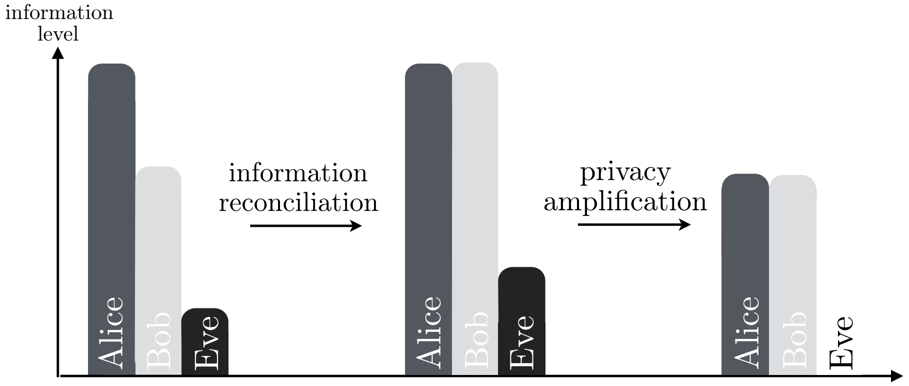


Figure 6. Schematic representation of information shared between the users and the eavesdropper during a sequential strategy for secret-key generation.

capacity $C_{\text{rec}}(R_1, 0)$ *is given by*

$$C_{\text{rec}}(R_1, 0) = C_{\text{SK}}(R_1, 0).$$

As shown in Example 2.3.1, unlike the case of rate-unlimited communication, in the case of rate-limited communication, it is not necessarily optimal to first achieve the reconciliation capacity in Proposition 2.3.1 and then to perform privacy amplification. In other words, if a sequential strategy is known to achieve the secret-key capacity, it does not tell us at which rate we should perform the reconciliation step. In the following, we say that reconciliation and privacy amplification are independent if achieving the reconciliation capacity in a sequential strategy leads to achieving the secret-key capacity.

Example 2.3.1. *Consider the scenario presented in Figure 7, in which* $|\mathcal{X}| = |\mathcal{Y}| = |\mathcal{Z}| = 2$*,* $X — Y — Z$ *forms a Markov chain, and* $X \sim \mathcal{B}(p)$*. We assume a one-way rate-limited public communication, i.e* $R_1 \in \mathbb{R}$ *and* $R_2 = 0$*. We set the parameters as follows.* $R_1 = H(X|Y)/3$*,* $p = 0.23$*,* $\beta_1 = 0.01$*,* $\beta_2 = 0.03$*,* $\gamma_1 = 0.03$ *and* $\gamma_2 = 0.01$*. We note* H_b *the binary entropy function and define for* $p \in [0, 1]$*,* $\bar{p} \triangleq 1 - p$*.*

We will show in the next section that a sequential strategy achieves the WSK *capacity* $C_{\mathrm{WSK}}(R_1,0)$. *Moreover, we can show that*

$$C_{\mathrm{WSK}}(R_1,0) = \max_{\alpha_1,\alpha_2}(f-g)(\alpha_1,\alpha_2), \quad \textit{subject to } (h-f)(\alpha_1,\alpha_2) = R_1, \tag{2}$$

$$C_{\mathrm{rec}}(R_1,0) = \max_{\alpha_1,\alpha_2} f(\alpha_1,\alpha_2), \quad \textit{subject to } (h-f)(\alpha_1,\alpha_2) = R_1, \tag{3}$$

where

$$f(\alpha_1,\alpha_2) \triangleq H_b(p_y) - p_u H_b(a) - \bar{p}_u H_b(b),$$

$$g(\alpha_1,\alpha_2) \triangleq H_b(p_z) - p_u H_b(c) - \bar{p}_u H_b(d),$$

$$h(\alpha_1,\alpha_2) \triangleq H_b(p) - p_u H_b(\alpha_1) - \bar{p}_u H_b(\alpha_2),$$

with $p_u = (\bar{\alpha}_2 - p)/(\bar{\alpha}_2 - \alpha_1)$, $p_y = \bar{p}\bar{\beta}_1 + p\beta_2$, $p_z = p_y\bar{\gamma}_1 + \bar{p}_y\gamma_2$, $a = \alpha_1\beta_2 + \bar{\alpha}_1\bar{\beta}_1$, $b = \alpha_2\bar{\beta}_1 + \bar{\alpha}_2\beta_2$, $c = \bar{\gamma}_1 a + \gamma_2\bar{a}$, $d = \bar{\gamma}_1 b + \gamma_2\bar{b}$.

Numerically,

$$C_{\mathrm{WSK}}(R_1,0) > 0.050 > 0.045 > (f-g)(\alpha_1^*,\alpha_2^*),$$

where (α_1^*,α_2^*) *achieves* $C_{\mathrm{rec}}(R_1,0)$. *Hence, for this example, achieving the reconciliation capacity in a sequential key-generation is not optimal and incurs a rate loss above* 10%.

Remark 2.3.1. *Deriving (2) and (3) is not straightforward. We used Proposition 2.5.3 given in the following sections, which shows that equality holds in the public communication rate constraint (4) and that* $|\mathcal{U}| \leqslant |\mathcal{X}|$.

In Section 2.4, for $R_1, R_2 \in \mathbb{R}_+$, we study the achievability of $R_{\mathrm{WSK}}(R_1,R_2)$, $C_{\mathrm{WSK}}(R_1,0)$, given in Theorem 2.3.1 and $C_{\mathrm{SK}}(R_1,R_2)$ given in Corollary 2.3.1, with a sequential key-generation strategy. Moreover, in Section 2.5, we identify scenarios for which reconciliation and privacy are independent in the sense defined in this section.

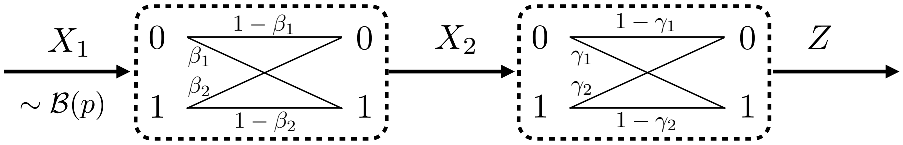


Figure 7. Example of a binary DMS studied in Example 2.3.1.

2.4 Sequential strategies achieve the best know bounds of $C_{\mathbf{WSK}}$ and $C_{\mathbf{SK}}$

In this section, we provide one of our main result. That is, the successive combination of reconciliation and privacy amplification, achieves the best known rates of the secret-key capacity (under the assumption of degraded sources in the case of two-way communication), when constraints are imposed on the public communication. As a side result, we extend known bounds of C_{WSK} and C_{SK} for DMS to the case of CMS.

Before we state our results, we provide a high-level description of our coding schemes. The main difficulty introduced by rate-limited public communication is the need for vector quantization of the source in the reconciliation step to better control the amount of information sent over the public channel. We use Wyner-Ziv coding, i.e. lossy source coding with side information, to handle this part. The privacy amplification step is performed with extractors, that is, functions that take as input a sequence of n arbitrarily distributed bits and a sequence of d truly uniformly distributed bits to output a sequence of k nearly uniformly distributed bits. Specifically, Alice and Bob publicly agree on a sequence U_d of d truly uniformly distributed bits, and use the extractors, using S and U_d as inputs, to form their secret-key K. Observe that the extractors must be chosen such that the size d of U_d must be negligible, compared to n, so that the effect of the transmission of U_d on the rate of public communication is negligible. Moreover, the extractors must extract almost the entire min-entropy of the input, that is, all the randomness of the input, to maximize the length of the secret-key generated. As already mentioned, such extractors are, for instance, explicitly constructed in [62].

The next major difficulty is to combine reconciliation and privacy amplification, since the output of the reconciliation step is not an independently and identically distributed random variable because of the vector quantization. Crucial ingredients to successfully perform this combination are Markov's Lemma (see for instance [61]) and a repetition of the reconciliation protocol. A schematic representation of the scheme is illustrated in Figure 8 for one-way rate-limited public communication. Specifically, a reconciliation protocol, operating over sequences of size n, is repeated m times. Then, an overall reconciliation step is performed, followed by an overall privacy amplification step. After the overall reconciliation step, Alice and Bob have agreed on a common sequence $S \triangleq U^N$, where $N \triangleq n \times m$. Finally, Alice and Bob perform privacy amplification by using an extractor with inputs U^N and a uniform random variable U_d. Note that the eavesdropper's total knowledge is thus Z^N, his observation of the source, $F \triangleq (A, B)$, the public communication of Alice and Bob, and U_d. We now state our results as follows.

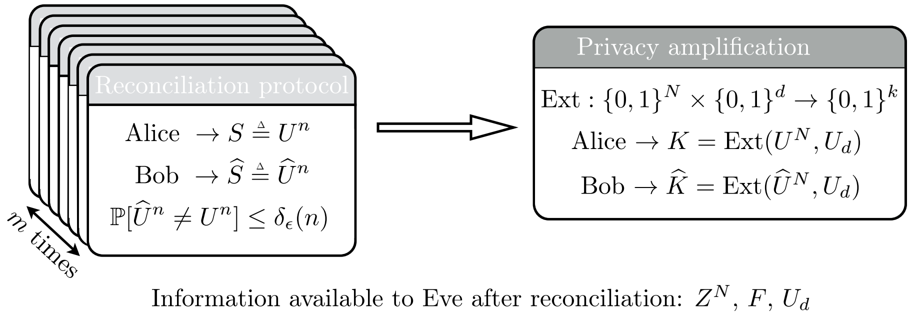

Reconciliation protocol
Alice $\to S \triangleq U^n$
Bob $\to \widehat{S} \triangleq \widehat{U}^n$
$\mathbb{P}[\widehat{U}^n \neq U^n] \leq \delta_\epsilon(n)$
m times
Privacy amplification
Ext : $\{0,1\}^N \times \{0,1\}^d \to \{0,1\}^k$
Alice $\to K = \text{Ext}(U^N, U_d)$
Bob $\to \widehat{K} = \text{Ext}(\widehat{U}^N, U_d)$
Information available to Eve after reconciliation: Z^N, F, U_d

Figure 8. Schematic representation of sequential-key generation for one-way rate-limited public communication. The scheme starts with m repetitions of a reconciliation protocol, during which Alice and Bob agree on U^n, a quantized version of X^n, with an error probability bounded by $\delta_\epsilon(n)$, with $\lim_{n\to\infty} \delta_\epsilon(n) = 0$, and $\epsilon > 0$. Privacy amplification with extractors is then performed to form the shared secret-key.

Theorem 2.4.3. *Let $(\mathcal{XYZ}, p_{XYZ})$ be a MS such that $X — Y — Z$. For $R_1, R_2 \in \mathbb{R}^+$, all WSK rates R that satisfy*

$$R < R_{\text{WSK}}(R_1, R_2)$$

are achievable with sequential key-generation strategies.

Proof. See Appendix 2.A. □

Remark 2.4.2. *Note that we assume $X—Y—Z$. For two-way communication, the necessity of this hypothesis might be an inherent weakness of a scheme that consists of a successive design of reconciliation and privacy amplification, rather than a joint design as in [16] (see the proof in Appendix 2.A for more details). Observe, however, that for a one-way public communication, in Theorem 2.4.4, this assumption is not required.*

Theorem 2.4.4. *Let $(\mathcal{X}\mathcal{Y}\mathcal{Z}, p_{XYZ})$ be a MS. For $R_1 \in \mathbb{R}^+$, all WSK rates R that satisfy*

$$R < C_{\text{WSK}}(R_1, 0)$$

are achievable with sequential key-generation strategies.

Proof. See Appendix 2.B. □

Theorem 2.4.5. *Let $(\mathcal{X}\mathcal{Y}, p_{XY})$ be a MS. For $R_1, R_2 \in \mathbb{R}^+$, all SK rates R that satisfy*

$$R < C_{\text{SK}}(R_1, R_2)$$

are achievable with sequential key-generation strategies.

We omit the proof of Theorem 2.4.5, which is similar to the one of Theorem 2.4.3 without the random variable Z.

Note that putting constraints on the public communication leads to auxiliary random variables in the expression of the secret-key capacity and the reconciliation capacity, as seen in Section 2.3. Hence, as demonstrated in Example 2.3.1, auxiliary random variables that achieve the reconciliation capacity, may not achieve the secret-key capacity. In other words, reliability and secrecy can be handled successively, but cannot necessarily be treated independently, as defined in Section 2.3.4. Nevertheless, in the next section, we identify scenarios for which reconciliation and privacy amplification can be treated independently.

As a side result, we have extended known bounds for the secret-key capacity for DMS to the case of CMS. We summarize this result in the following corollary, which is directly deduced from Theorems 2.4.3, 2.4.4, and 2.4.5.

Corollary 2.4.2. *Let* $(\mathcal{XYZ}, p_{XYZ})$ *be a MS.*

(a) Assume that $X — Y — Z$*. For* $R_1, R_2 \in \mathbb{R}^+$*, the two-way WSK achievable bound* $R_{\text{WSK}}(R_1, R_2)$ *given in Theorem 2.3.1.a, remains valid for a CMS.*

(b) For $R_1 \in \mathbb{R}^+$*, the expression of the one-way WSK capacity* $C_{\text{WSK}}(R_1, 0)$ *given in Theorem 2.3.1.b, remains valid for a CMS.*

(c) For $R_1, R_2 \in \mathbb{R}^+$*, the two-way SK capacity* $C_{\text{SK}}(R_1, R_2)$ *given in Corollary 2.3.1, remains valid for a CMS.*

2.5 Scenarios for which independence holds between reliability and secrecy

As seen in the Example 2.3.1, achieving the reconciliation capacity might not lead to achieving the secret-key capacity. In this section, we identify special cases for which independence holds between reconciliation and privacy amplification. Specifically, we prove that independence holds for the two-way one-round SK capacity, the one-way WSK capacity in the case of binary symmetric degraded sources, and the one-way WSK capacity in the case of Gaussian degraded sources. As a side result, we obtain an expression for the two-way rate-limited reconciliation capacity and a closed-form expression for the secret-key capacity $C_{\text{WSK}}(R_1, 0)$ in the case of degraded binary symmetric sources.

2.5.1 Two-way rate-limited SK capacity

In this section, we consider the two-way rate-limited SK capacity. That is, the eavesdropper has no correlated observation of the source.

We first show that the two-way rate-limited SK capacity is equal to the two-way rate-limited reconciliation capacity in the following proposition.

Proposition 2.5.2. *Let $(\mathcal{X}\mathcal{Y}, p_{XY})$ be a MS. For $R_1, R_2 \in \mathbb{R}^+$, the rate-limited reconciliation capacity $C_{\text{rec}}(R_1, R_2)$ is*

$$C_{\text{rec}}(R_1, R_2) = C_{\text{SK}}(R_1, R_2).$$

Proof. See Appendix 2.C. □

Hence, by Proposition 2.5.2, the auxiliary random variables that achieve the reconciliation capacity, also achieve the secret-key capacity; combined with Theorem 2.4.5, we obtain the following corollary.

Corollary 2.5.3. *Let $(\mathcal{X}\mathcal{Y}, p_{XY})$ be a MS and $R_1, R_2 \in \mathbb{R}^+$. The two-way rate-limited SK capacity $C_{\text{SK}}(R_1, R_2)$ is achievable by a sequential strategy, moreover, reconciliation and privacy amplification steps can be handled independently, as defined in Section 2.3.4.*

2.5.2 One-way rate-limited WSK capacity for degraded binary symmetric sources

In this section, we assume a degraded DMS. We first refine Proposition 2.5.2 and Theorem 2.3.1.b in the following proposition.

Proposition 2.5.3. *Let $(\mathcal{X}\mathcal{Y}\mathcal{Z}, p_{XYZ})$ be a DMS such that $X — Y — Z$. Assume $R_1 \in \mathbb{R}^+$ and $R_2 = 0$. We tighten the rate constraint in (4), (6) and the range constraint in (5), (7) as follows.*

(a) The one-way rate-limited reconciliation capacity is

$$C_{\text{rec}}(R_1, 0) = \max_U I(Y; U)$$

subject to

$$R_1 = I(X; U|Y), \tag{4}$$

$$U — X — Y,$$

$$|\mathcal{U}| \leqslant |\mathcal{X}|. \tag{5}$$

(b) The one-way rate-limited secret-key capacity is

$$C_{\mathrm{WSK}}(R_1, 0) = \max_U \left(I(Y;U) - I(Z;U) \right)$$

subject to

$$R_1 = I(X;U|Y), \tag{6}$$

$$U - X - Y - Z,$$

$$|\mathcal{U}| \leqslant |\mathcal{X}|. \tag{7}$$

Proof. See Appendix 2.D. □

Remark 2.5.3. *The expression of the WSK capacity in Proposition 2.5.3.b is obtained from Theorem 2.3.1.b and is due to Watanabe [60]. We refine this result by proving that equality holds in the rate constraint and by improving the range constraint of $\mathcal{U}$; The argument used to show the equality in the rate constraint of Propositions 2.5.3.a and 2.5.3.b, is one that applies to various convex maximization problems: the maximum principle (see Appendix 2.D). This refinement is critical for the analysis of binary sources, especially to solve the optimization problem for the WSK capacity in Proposition 2.5.4, and thus to determine the WSK capacity for degraded binary symmetric sources in Example 2.5.2.*

Remark 2.5.4. *As soon as R_1 is at least $H(X|Y)$, $C_{\mathrm{rec}}(R_1, 0)$ (resp. $C_{\mathrm{WSK}}(R_1, 0)$) attains the same maximum $I(X;Y)$ (resp. $I(X;Y) - I(X;Z)$) as in the case $R_1 = +\infty$.*

The solution of the maximization problem in Proposition 2.5.3.b can be obtained explicitly, when the source has symmetry properties.

Proposition 2.5.4. *Let $(\mathcal{X}\mathcal{Y}\mathcal{Z}, p_{XYZ})$ be a DMS such that $X - Y - Z$. Assume that $|\mathcal{X}| = 2$ and let $R_1 \in \mathbb{R}_+^*$.*

If the channels $p_{Y|X}$ and $p_{Z|X}$ are symmetric [64], then the auxiliary random variable U achieving $C_{\mathrm{WSK}}(R_1, 0)$ in Proposition 2.5.3.b, is such that the test-channel

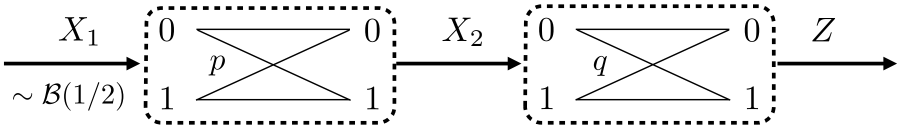


Figure 9. Binary DMS studied in Example 2.5.2.

$p_{U|X}$ is a BSC with parameter β_0, with β_0, any of the two symmetric solutions of

$$R_1 = I(U;X) - I(U;Y).$$

Proof. See Appendix 2.E. □

Although the result stated in Proposition 2.5.4 seems intuitive and non-surprising, the proof is not straightforward, as a crucial step is the improvements proposed in Proposition 2.5.3. Hence, if the channels $p_{Y|X}$ and $p_{Z|X}$ are symmetric, by Proposition 2.5.4, the auxiliary random variable U achieving $C_{\text{rec}}(R_1, 0)$ in Proposition 2.5.3.a also achieves $C_{\text{WSK}}(R_1, 0)$ in Proposition 2.5.3.b; combined with Theorem 2.4.4, we obtain the following corollary.

Corollary 2.5.4. *Let $(\mathcal{X}\mathcal{Y}\mathcal{Z}, p_{XYZ})$ be a DMS such that X—Y—Z and $|\mathcal{X}| = 2$. Let $R_1 \in \mathbb{R}_+^*$. We assume the channels $p_{Y|X}$ and $p_{Z|X}$ to be symmetric. The one-way rate-limited WSK capacity $C_{\text{WSK}}(R_1, 0)$ is achievable by a sequential strategy, moreover, reconciliation and privacy amplification steps can be handled independently, as defined in Section 2.3.4.*

The following example illustrates Proposition 2.5.4 and Corollary 2.5.4.

Example 2.5.2. *As depicted in Figure 9, assume that X and Y (respectively Y and Z) are connected by a binary symmetric channel (BSC) with crossover probability p (respectively q). We also assume $X \sim \mathcal{B}(1/2)$ to obtain simpler expressions; however, the application of Proposition 2.5.4 remains valid for $X \sim \mathcal{B}(\alpha)$, $\alpha \in [0,1]$. By*

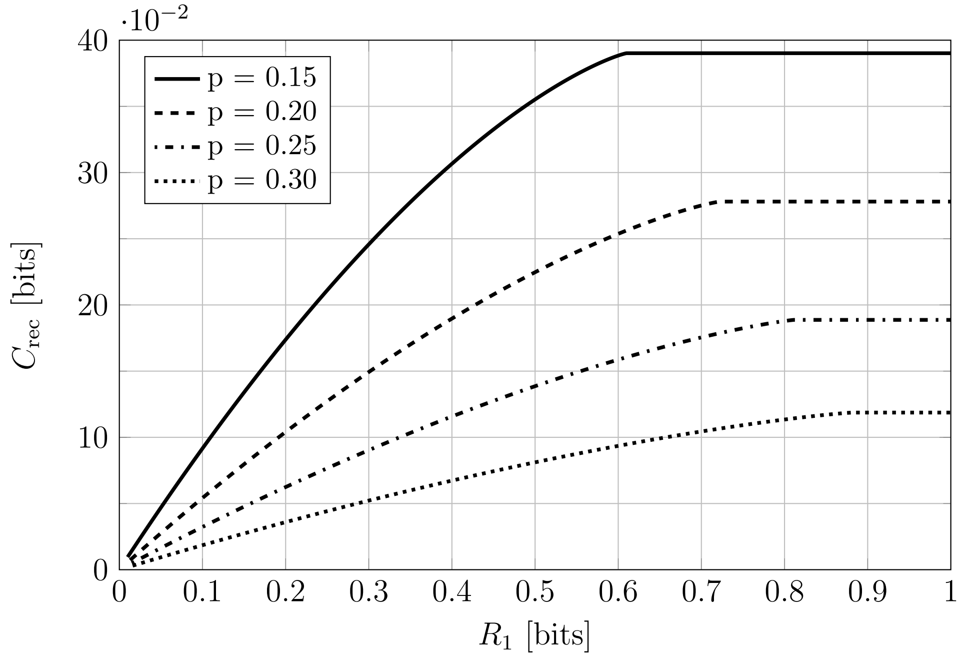


Figure 10. Reconciliation capacity $C_{\mathbf{rec}}(R_1, 0)$.

Proposition 2.5.4, the reconciliation capacity is

$$C_{\text{rec}}(R_1, 0) = \begin{cases} 1 - H_b(p \star \beta_0), & \textit{if } R_1 \leqslant H(X|Y), \\ 1 - H_b(p), & \textit{if } R_1 \geqslant H(X|Y), \end{cases}$$

and the WSK capacity is

$$C_{\text{WSK}}(R_1, 0) = \begin{cases} H_b\left(p \star \beta_0 \star q\right) - H_b(p \star \beta_0), & \textit{if } R_1 \leqslant H(X|Y), \\ H_b(p \star q) - H_b(p), & \textit{if } R_1 \geqslant H(X|Y), \end{cases}$$

with β_0, any of the two symmetric solutions of the equation $H_b(p \star \beta_0) - H_b(\beta_0) = R_1$ and where, for $p, q \in [0,1]$, we have defined the following associative and commutative operation $p \star q \triangleq p(1-q) + (1-p)q$; observe that $[0,1]$ is closed with respect to $\star$.

Figure 11 (resp. Figure 10) illustrates Remark 2.5.4 and the fact that the reconciliation capacity $C_{\text{rec}}(R_1, 0)$ (resp. the secret key-capacity $C_{\text{WSK}}(R_1, 0)$) is monotonically increasing in the communication rate constraint.

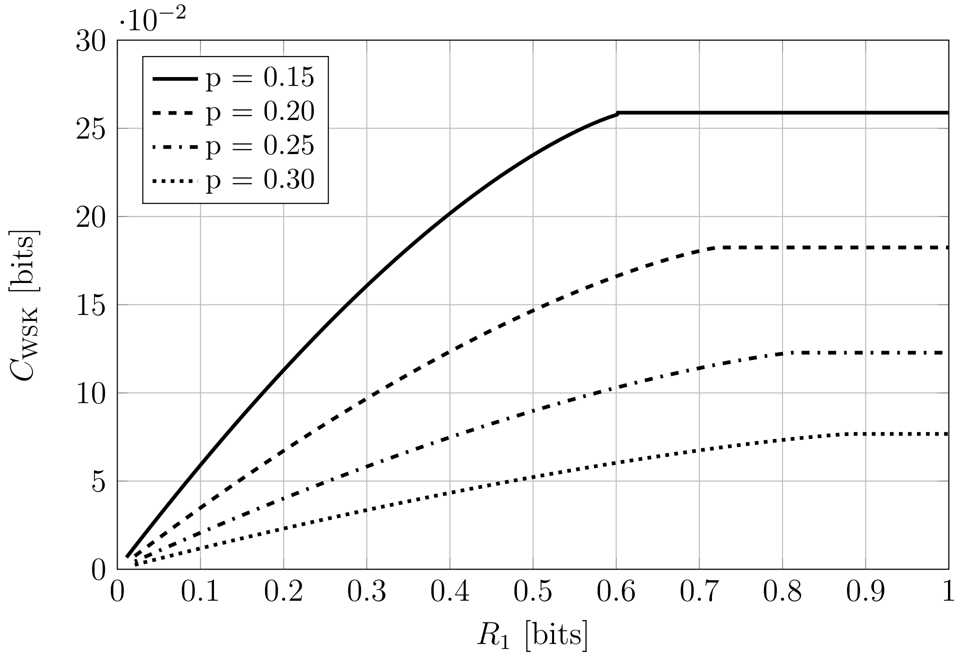


Figure 11. WSK capacity $C_{\mathbf{WSK}}(R_1, 0)$ ($q = 0.2$).

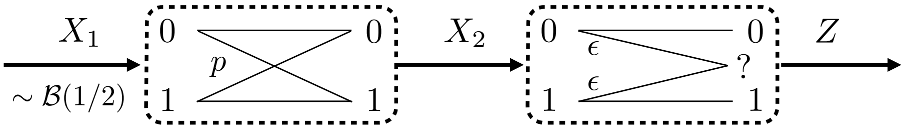


Figure 12. Binary erasure channel studied in Example 2.5.2.

Corollary 2.5.4 states that choosing a test-channel $p_{U|X}$ as a BSC with parameter β_0, achieves $C_{\text{rec}}(R_1, 0)$ and $C_{\text{WSK}}(R_1, 0)$, so that reconciliation and privacy amplification can be designed independently. Consequently, for any other channel $p_{Z|Y}$, as long as $p_{Z|X}$ stays symmetric, the reconciliation capacity and the optimal reconciliation protocol for sequential key-generation remains the same. It is for instance the case if we choose $p_{Z|Y}$ as a binary erasure channel (BEC), as depicted in Figure 12. Moreover, in this case, Proposition 2.5.4 still allows us to determine the WSK capacity:

$$C_{\text{WSK}}^{(erasure)}(R_1, 0) = \begin{cases} \epsilon(1 - H_b(p \star \beta_0)), & \textit{if } R_1 \leqslant H(X|Y), \\ \epsilon(1 - H_b(p)), & \textit{if } R_1 \geqslant H(X|Y), \end{cases}$$

where ϵ is the erasure probability characterizing $p_{Z|Y}$.

Remark 2.5.5. *We can show that the sequential strategy used in this section can also be applied to similar models for biometric secrecy [65].*

2.5.3 One-way rate-limited WSK capacity for degraded Gaussian sources

In this section, we consider a degraded Gaussian MS with one-way rate-limited public communication. We assume that X, Y, and Z are zero-mean correlated Gaussian sources on $\mathbb{R}$, and that Alice, Bob, and Eve know the covariance matrix of (X, Y, Z). We first refine the reconciliation capacity and the secret-key capacity to give the counterpart of Proposition 2.5.3. We then provide the reconciliation capacity and the secret-key capacity, and show that reconciliation and privacy amplification can be treated independently. We also briefly discuss the performance of vector quantization compared to scalar quantization for the reconciliation step, thereby providing a counterpart of Remark 2.5.4.

Proposition 2.5.5. *Let $(\mathcal{XYZ}, p_{XYZ})$ be a zero-mean Gaussian MS such that X—Y—Z. Assume $R_1 \in \mathbb{R}^+$ and $R_2 = 0$.*

(a) The one-way rate-limited reconciliation capacity is

$$C_{\text{rec}}(R_1, 0) = \max_U I(Y; U)$$

subject to

$$\begin{aligned} &R_1 = I(X; U|Y), \\ &U\text{—}X\text{—}Y, . \end{aligned} \tag{8}$$

(b) The one-way rate-limited WSK capacity is

$$C_{\text{WSK}}(R_1, 0) = \max_U (I(Y; U) - I(Z; U))$$

subject to

$$\begin{aligned} &R_1 = I(X; U|Y), \\ &U\text{—}X\text{—}Y\text{—}Z, \end{aligned}$$

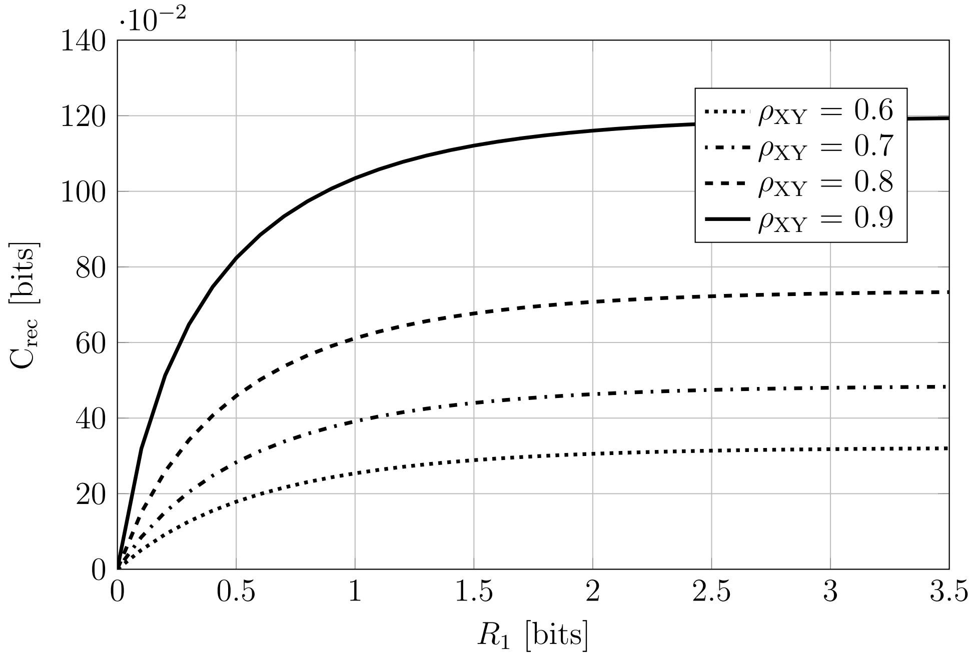


Figure 13. Reconciliation capacity $C_{\mathbf{rec}}(R_1, 0)$ for different correlation coefficients ρ_{XY}.

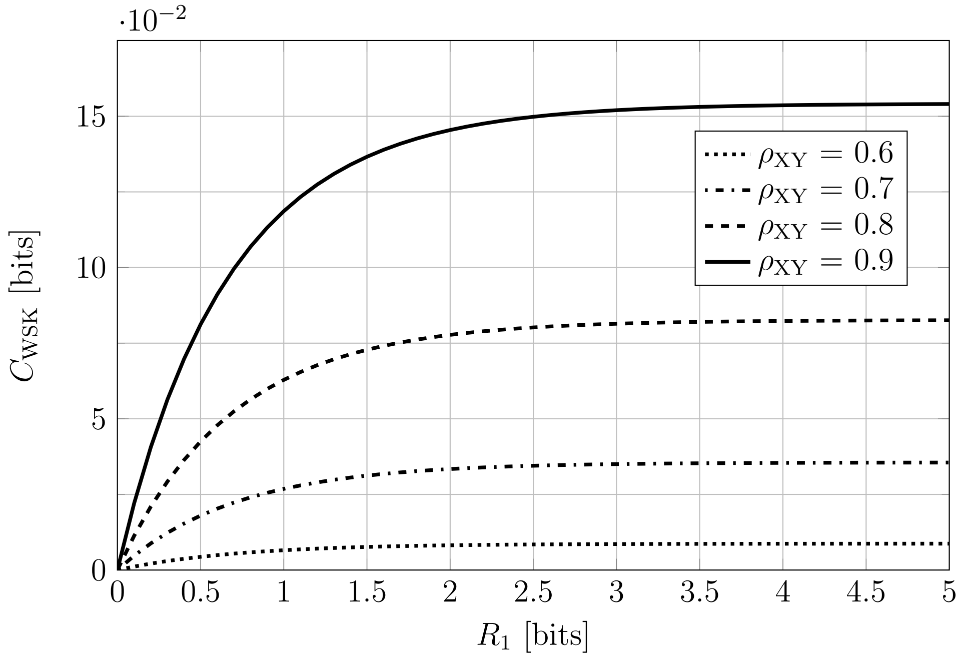


Figure 14. WSK capacity $C_{\mathbf{WSK}}(R_1, 0)$, for different correlation coefficients ρ_{XY} ($\rho_{XZ} = 0.1$, $\rho_{YZ} = 0.4$).

Proposition 2.5.5 follows from Proposition 2.5.6.

Proposition 2.5.6. *Assume that* $(\mathcal{XYZ}, p_{XYZ})$ *is a degraded zero-mean Gaussian source. Let* $R_1 \in \mathbb{R}_+$.

The auxiliary random variable U *achieving* $C_{\text{rec}}(R_1, 0)$ *in Proposition 2.5.5.a is a zero-mean Gaussian with variance*

$$\sigma_0 \triangleq \sigma_x \left(1 + (1 - \rho_{XY})(e^{2R_1} - 1)^{-1}\right)$$

that satisfies the rate-constraint (8), where ρ_{XY} *is the correlation coefficient between* X *and* Y*. Moreover, the same auxiliary random variable* U *achieves* $C_{\text{WSK}}(R_1, 0)$ *in Proposition 2.5.5.b.*

(a) The one-way rate-limited reconciliation capacity is given by

$$C_{\text{rec}}(R_1, 0) = \frac{1}{2} \log_2 \frac{1 - \left(\rho_{XY} e^{-R_1}\right)^2}{1 - \rho_{XY}^2}.$$

(b) The one-way rate-limited WSK capacity is

$$C_{\text{WSK}}(R_1, 0) = \frac{1}{2} \log_2 \frac{(1 - \rho_{YZ}^2)(1 - \rho_{XZ}^2) - \left(\rho_{XY} - \rho_{YZ}\rho_{XZ}\right)^2 e^{-2R_1}}{(1 - \rho_{YZ}^2)(1 - \rho_{XZ}^2) - \left(\rho_{XY} - \rho_{YZ}\rho_{XZ}\right)^2}.$$

Proof. (b) is due to Watanabe [60], and the proof of (a) is similar to the one of (b). □

Proposition 2.5.6 states that both arguments of the maximum for the auxiliary random variable U, in (a) and (b) of Proposition 2.5.5 are identical; combined with Theorem 2.4.4, we deduce the following corollary.

Corollary 2.5.5. *Assume that* $(\mathcal{XYZ}, p_{XYZ})$ *is a degraded zero-mean Gaussian source. Let* $R_1 \in \mathbb{R}_+$*. The one-way rate-limited WSK capacity* $C_{\text{WSK}}(R_1, 0)$ *is achievable by a sequential strategy, moreover, reconciliation and privacy amplification steps can be handled independently, as defined in Section 2.3.4.*

As shown by Proposition 2.5.6.a (resp. Proposition 2.5.6.b), and as illustrated in Figure 13 (resp. Figure 14), the reconciliation capacity (resp. the WSK capacity) does not reach $I(X;Y)$ (resp. $I(X;Y) - I(X;Z)$) when R_1 exceed a certain value. As mentioned in [60] and Remark 2.5.4, unlike the case of discrete random variables, $C_{\text{rec}}(R_1, 0)$ (resp. $C_{\text{WSK}}(R_1, 0)$) can only approach $I(X;Y)$ (resp. $I(X;Y) - I(X;Z)$) asymptotically. Nevertheless, we show in the following proposition a continuous counterpart of Remark 2.5.4.

The achievability of $C_{\text{WSK}}(R_1, 0)$ with our sequential strategy is based on Wyner-Ziv coding. For a practical implementation, additional structure needs to be introduced, for instance with vector quantization. Since scalar quantization is the simplest and often the most computationally efficient type of quantization, it is natural to ask how scalar quantization performs compared to vector quantization. We answer this question in the following proposition.

Proposition 2.5.7. *Let $n \in \mathbb{N}^*$, and $a > 0$. Define U as a uniformly quantized version of X. Specifically,*

$$\forall k \in [\![1, n]\!], p_U(u_k) \triangleq \int_{t_k}^{t_{k+1}} p_X(x)dx, \text{ with } t_k \triangleq a(2\tfrac{k-1}{n-1} - 1).$$

If n is large enough, then

$$|I(X;Y) - I(Y;U)| \leqslant \epsilon(a) + a \cdot K e^{h(X|Y) - R_1},$$

where R_1 is the communication rate constraint, K is a constant, and $\epsilon(a)$ decreases exponentially fast to zero as a goes to infinity.

Proof. See Appendix 2.F. □

Proposition 2.5.7 gives a continuous counterpart of Remark 2.5.4. Indeed, when $R_1 > h(X|Y)$, by Proposition 2.5.7, if X is quantized finely enough, then $I(Y;U)$ approach $I(X;Y)$ exponentially fast as R_1 increases.

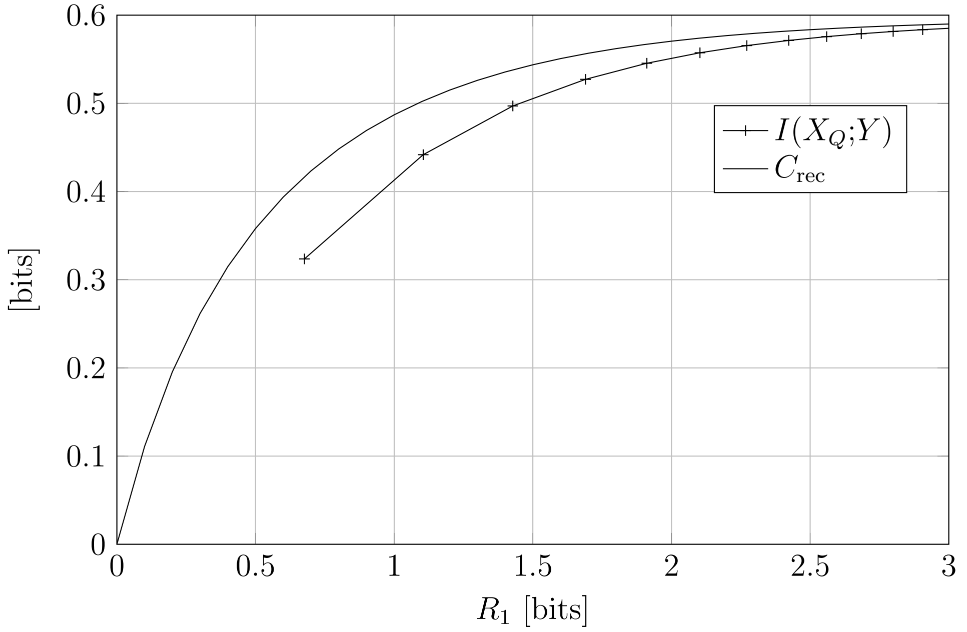


Figure 15. Reconciliation capacity obtained for a scalar quantization of X with $\rho_{XY} = 0.75$, $h(X|Y) \approx 1$.

Hence, the improvement of vector quantization compared to scalar quantization decays rapidly as the communication rate increases beyond $h(X|Y)$. Note that, in practice, we can optimize the scalar quantization, so that the loss could be even smaller than predicted by Proposition 2.5.7. Figure 15 illustrates this point by comparing the reconciliation capacity with numerical values of achievable rates obtained when X is scalar-quantized.[4] Nevertheless, for low communication rates, Figure 15 shows that vector quantization improves the performance; in this case, we could implement, for instance, trellis coded vector quantization (TCVQ) [66].

[4] We have increased the number of interval of quantization of X from 2 to 15 and chosen their bounds by a standard gradient method to maximize $I(X_Q; Y)$.

2.6 Concluding remarks

We have shown that the one-way rate-limited capacity is achievable by a sequential strategy that separates reliability and secrecy thanks to a reconciliation step followed by a privacy amplification step with extractors; in the case of two-way communication, the sequential design seems to suffer a loss of performance compared to the joint design and similar secret key rates have only been established for degraded sources or when there is no side information at the eavesdropper (SK capacity). We have also qualified robustness of sequential strategy to rate-limited communication, by showing that achieving the reconciliation capacity in a sequential strategy is, unlike the case of rate-unlimited communication, not necessarily optimal. We further provide scenarios for which it stays optimal. As a side result, we have extended known bounds of the WSK capacity for a discrete source model to the case of a continuous source model, and derived a closed-form expression of the one-way rate-limited capacity for degraded binary symmetric sources.

A strength of sequential key-generation is to easily translate into practical designs. Even more interestingly, the proposed scheme can be made very flexible with the following modifications.

Rate-compatible reconciliation: We can adapt to the characteristics of the legitimate users by the use of rate-compatible LDPC codes, to perform the reconciliation phase, as demonstrated in [67, 68]. Note, however, that vector quantization might be required, which could complexify the reconciliation phase.

Rate-compatible privacy amplification: Privacy amplification can also be performed with universal families of hash functions, in which case the counterpart of Theorem 2.3.2 is found in [24].[5] Hence, one can design privacy amplification methods easily adjustable to the characteristics of the eavesdropper's observations, if we make k

[5] However, it requires more random bits than extractors, on the order of n random bits, since functions must be chosen at random in universal families. Consequently, our scheme needs to be adapted to account for it.

vary in the following universal family of hash functions $\mathcal{H} = \{\mathrm{GF}(2^n) \to \{0,1\}^k, x \mapsto (k \text{ bits of the product } xy) | y \in \mathrm{GF}(2^n)\}$, where the k bits are fixed but their position can be chosen arbitrarily [69].

APPENDICES

2.A Proof of Theorem 2.4.3

2.A.1 Discrete case

Let $\epsilon > 0$. Let $R_1, R_2 \in \mathbb{R}^+$. Let $m, n \in \mathbb{N}$, and define $N \triangleq nm$. Let $k \in \mathbb{N}$ to be determined later. Consider a sequential key-distillation strategy $\mathcal{S}_N$ that consists of

- m repetitions of a reconciliation protocol $\mathcal{R}_n$ based on Wyner-Ziv coding. The protocol $\mathcal{R}_n$ operates as described in Appendix 2.C.2. Hence, after one repetition of the protocol, Alice obtains $S^n \triangleq U^n\hat{V}^n$, whereas Bob has $\hat{S}^n \triangleq \hat{U}^nV^n$ and $\mathbb{P}[\hat{S}^n \neq S^n|\mathcal{R}_n] \leqslant P_e(\epsilon, n)$.[6] In addition, the information disclosed over the public channel during the m repetitions of the reconciliation protocol is upper bounded by

  $$\log|\mathcal{A}|^m + \log|\mathcal{B}|^m = N(I(U;X) - I(U;Y) + I(V;Y|U) - I(V;X|U) + r_0(\epsilon)),$$

  with $\lim_{\epsilon\to 0} r_0(\epsilon) = 0$.[7] An additional round of reconciliation is then performed to ensure $\mathbb{P}[(\hat{S}^n)^m \neq (S^n)^m|\mathcal{R}_n] \leqslant \delta_e(m)$, where $\lim_{m\to\infty} \delta_e(m) = 0$, for any fixed n. We note $\log|\mathcal{C}|^m$ the information communicated to perform this last step. Hence, the overall information disclosed is upper bounded by $l_{rec} \triangleq \log(|\mathcal{A}|^N|\mathcal{B}|^N|\mathcal{C}|^m)$, that is

  $$l_{rec} = N(I(U;X) - I(U;Y) + I(V;Y|U) - I(V;X|U) + r_1(\epsilon, n)), \tag{9}$$

  $$\text{with } r_1(\epsilon, n) \triangleq \frac{1+\epsilon}{n} H(S^n|\hat{S}^n) + r_0(\epsilon) \tag{10}$$

  arbitrarily small for n large enough by Fano's inequality, so that the communication rates R_1 and R_2 remain asymptotically unchanged.

- privacy amplification, based on extractors with output size k, at the end of which

[6] By Appendix 2.C.2, $P_e(\epsilon, n)$ decreases exponentially to zero as $n\epsilon^2$ goes to infinity.

[7] $r_0(\epsilon) \triangleq 6\epsilon H(U) + 12\epsilon H(V|U)$ by Appendix 2.C.2.

Alice computes her key $K \triangleq g(S^N, U_d)$, while Bob computes $\hat{K} \triangleq g(\hat{S}^N, U_d)$, where U_d is a sequence of d uniformly distributed random bits.

The total information available to Eve after reconciliation consists of her observation Z^N, the public messages A^m and B^m, respectively sent by Alice and Bob, the public message C^m, and U_d. The strategy $\mathcal{S}_N$ is also known to Eve, but we omit the conditioning on $\mathcal{S}_N$ for convenience.

We first show that, for a suitable choice of the output size k, the quantity $k - H(K|U_d Z^N A^m B^m C^m)$ vanishes to zero for N large enough. Then, we show that the corresponding WSK rate achieves the lower bound on the WSK capacity of Theorem 2.3.1. We first state Lemma 2.1.1, a refined version of the results in [23, 25], that is obtained by using the notion of robust typicality developed in the appendix of [61], to later extend our result to the continuous case.

Lemma 2.1.1 ([23, 25], Refined version). *Consider a DMS $(\mathcal{X}\mathcal{Z}, p_{XZ})$ and define the random variable Θ as*

$$\Theta \triangleq \begin{cases} 1 & \text{if } (X^n, Z^n) \in \mathcal{T}_{2\epsilon}^n(XZ) \text{ and } Z^n \in \mathcal{T}_{\epsilon}^n(Z), \\ 0 & \text{otherwise.} \end{cases}$$

Then, $\mathbb{P}[\Theta = 1] \geqslant 1 - \delta_\epsilon^0(n)$, with $\delta_\epsilon^0(n) \triangleq 2|S_X|e^{-\epsilon^2 n\mu_X/3} + 2|S_{XZ}|e^{-\epsilon^2 n\mu_{XZ}/3}$, where $S_X \triangleq \{x \in \mathcal{X} : p(x) > 0\}$ and $\mu_X \triangleq \min_{x\in S_X} p(x)$. Moreover, if $z^n \in \mathcal{T}_\epsilon^n(Z)$,

$$H_\infty(X^n|Z^n = z^n, \Theta = 1) \geqslant n(H(X|Z) - \delta^0(\epsilon)) + \log(1 - \delta_\epsilon^1(n)),$$

where $\delta^0(\epsilon) \triangleq \epsilon H(X|Z)$ and $\delta_\epsilon^1(n) \triangleq 2|S_{X,Z}|e^{-\epsilon^2 n\mu_{X,Z}/6}$.

Let us start by defining the following random variables

$$\Theta \triangleq \begin{cases} 1 & \text{if } (S^N, Z^N) \in \mathcal{T}_{2\epsilon}^m(U^nV^nZ^n) \text{ and } Z^N \in \mathcal{T}_{\epsilon}^m(Z^n), \\ 0 & \text{otherwise.} \end{cases}$$

$$\Upsilon \triangleq \begin{cases} 1 & \text{if } H_\infty(S^N|z^N, a^m, b^m, c^m, \Theta = 1) \geqslant H_\infty(S^N|z^N, \Theta = 1) - l_{rec} - \sqrt{N}, \\ 0 & \text{otherwise.} \end{cases}$$

By Lemma 2.1.1 applied to the DMS $(\mathcal{U}^n\mathcal{V}^n\mathcal{Z}^n, p_{U^nV^nZ^n})$, $\mathbb{P}[\Theta = 1] \geqslant 1 - \delta_\epsilon^0(m)$, and by [25, Lemma 10], $\mathbb{P}[\Upsilon = 1] \geqslant 1 - 2^{-\sqrt{N}}$. Hence, $\mathbb{P}[\Upsilon = 1, \Theta = 1] \geqslant 1 - \delta_\epsilon^0(m) - 2^{-\sqrt{N}}$, and

$$\begin{aligned} H(K|U_dZ^NA^mB^mC^m) \geqslant \left(1 - \delta_\epsilon^0(m) - 2^{-\sqrt{N}}\right) \\ \times H(K|U_dZ^NA^mB^mC^m, \Upsilon = 1, \Theta = 1). \quad (11) \end{aligned}$$

To lower bound $H(K|U_dZ^NA^mB^mC^m, \Upsilon = 1, \Theta = 1)$, we first lower bound $H_\infty(S^N|z^N, a^m, b^m, c^m, \Theta = 1, \Upsilon = 1)$ to be able to use Theorem 2.3.2. By definition of Υ,

$$\begin{aligned} & H_\infty(S^N|z^N, a^m, b^m, c^m, \Theta = 1, \Upsilon = 1) \\ & \geqslant H_\infty(S^N|Z^N = z^N, \Theta = 1) - l_{rec} - \sqrt{N} \\ & \overset{(a)}{\geqslant} m(H(S^n|Z^n) - nr_2(\epsilon, n, m)) - l_{rec}, \quad (12) \end{aligned}$$

where (a) follows from Lemma 2.1.1 with

$$r_2(\epsilon, n, m) \triangleq \epsilon \frac{H(S^n|Z^n)}{n} - N^{-1}\log(1 - \delta_\epsilon^1(m)) + N^{-1/2}.^{8} \quad (13)$$

[8] The m repetitions of the protocol $\mathcal{R}_n$ allow us to link $H_\infty(\cdot)$ to $H(\cdot)$.

We now lower bound $H(S^n|Z^n)$.

$$\begin{aligned}
H(S^n|Z^n) &= H(\hat{S}^n|Z^n) + H(S^n|\hat{S}^n Z^n) - H(\hat{S}^n|S^n Z^n) \\
&\overset{(b)}{\geqslant} H(\hat{S}^n|Z^n) - \delta_\epsilon(n) \\
&= I(Y^n;\hat{S}^n|Z^n) + H(\hat{S}^n|Y^n Z^n) - \delta_\epsilon(n) \\
&= H(Y^n|Z^n) - H(Y^n|Z^n\hat{S}^n) + H(\hat{U}^n|Y^n Z^n) \\
&\qquad\qquad\qquad + H(V^n|Y^n\hat{U}^n Z^n) - \delta_\epsilon(n) \\
&\overset{(c)}{=} nH(Y|Z) - H(Y^n|Z^n\hat{S}^n) + H(\hat{U}^n|Y^n Z^n) - \delta_\epsilon(n), \qquad (14)
\end{aligned}$$

where (b) follows from Fano's inequality where $\lim_{n\to\infty}\delta_\epsilon(n) = 0$ by the exponential decrease of $P_e(\epsilon, n)$ with $\epsilon^2 n$, and (c) holds because V^n is a function of $(Y^n\hat{U}^n)$, and the Y_i's and Z_i's are i.i.d.. We first lower bound $H(\hat{U}^n|Y^n Z^n)$.

$$\begin{aligned}
H(\hat{U}^n|Y^n Z^n) &= H(U^n|Y^n Z^n) + H(\hat{U}^n|U^n Y^n Z^n) - H(U^n|\hat{U}^n Y^n Z^n) \\
&\overset{(d)}{\geqslant} H(U^n|Y^n Z^n) - \delta_\epsilon(n) \\
&\geqslant I(X^n;U^n|Y^n Z^n) - \delta_\epsilon(n) \\
&\overset{(e)}{=} nH(X|YZ) - H(X^n|Y^n Z^n U^n) - \delta_\epsilon(n), \qquad (15)
\end{aligned}$$

where (d) follows from Fano's inequality where $\lim_{n\to\infty}\delta_\epsilon(n) = 0$ by the exponential decrease of $P_e(\epsilon, n)$ with $\epsilon^2 n$, and (e) holds since the X_i's, Y_i's , and Z_i's are i.i.d.. Then, define

$$\Gamma \triangleq \begin{cases} 1 & \text{if } (X^n, U^n, Y^n, Z^n) \in \mathcal{T}^n_{2\epsilon}(XUYZ), \\ 0 & \text{otherwise.} \end{cases}$$

$$\Delta \triangleq \begin{cases} 1 & \text{if } (X^n, U^n) \in \mathcal{T}^n_{\epsilon}(XU), \\ 0 & \text{otherwise.} \end{cases}$$

so that,

$$\begin{aligned}
&H(X^n|Y^nZ^nU^n)\\
&\leqslant H(X^n\Gamma\Delta|Y^nZ^nU^n)\\
&= H(\Gamma\Delta|Y^nZ^nU^n) + H(X^n|Y^nZ^nU^n\Gamma\Delta)\\
&\leqslant 2 + \sum_{\delta,\gamma\in\{0,1\}} \mathbb{P}[\Gamma=\gamma|\Delta=\delta]\mathbb{P}[\Delta=\delta]H(X^n|Y^nZ^nU^n,\Gamma=\gamma,\Delta=\delta)\\
&\overset{(f)}{\leqslant} 2 + H(X^n|Y^nZ^nU^n,\Gamma=1,\Delta=1) + n(2\delta_\epsilon^2(n)+\delta_\epsilon^4(n))\log|\mathcal{X}|, \qquad (16)
\end{aligned}$$

where (f) holds since $\mathbb{P}[\Delta = 0] \triangleq \delta_\epsilon^2(n)$,[9] and $\mathbb{P}[\Gamma = 0|\Delta = 1] \leqslant \delta_\epsilon^4(n)$.[10] Indeed, we can apply Markov Lemma [70] (see the version given in [61]), since we have U^n—X^n—Y^nZ^n and for every (x^n, y^n, z^n), $p(y^nz^n|x^n) = \prod_{i=1}^{n} p_{YZ|X}(y_iz_i|x_i)$. Then,

$$\begin{aligned}
&H(X^n|Y^nZ^nU^n,\Gamma=1,\Delta=1)\\
&= \sum_{y^n,z^n,u^n} p(y^n,z^n,u^n|1,1)H(X^n|y^n,z^n,u^n,\Gamma=1,\Delta=1)\\
&\leqslant \sum_{y^n,z^n,u^n} p(y^n,z^n,u^n|1,1)\log|\mathcal{T}_{2\epsilon}^n(X|y^n,z^n,u^n)|\\
&\leqslant \sum_{y^n,z^n,u^n} p(y^n,z^n,u^n|1,1)(nH(X|YZU)(1+2\epsilon))\\
&\leqslant nH(X|YZU)(1+2\epsilon). \qquad (17)
\end{aligned}$$

Hence, combining (15), (16), and (17), we obtain

$$H(\hat{U}^n|Y^nZ^n) \geqslant n(H(X|YZ) - H(X|YZU) - r_3(\epsilon,n)), \qquad (18)$$

where

$$r_3(\epsilon,n) \triangleq 2H(X|YZU)\epsilon + (2\delta_\epsilon^2(n)+\delta_\epsilon^4(n))\log|\mathcal{X}| + 2/n + \delta_\epsilon(n)/n. \qquad (19)$$

[9] We have $\delta_\epsilon^2(n) \leqslant P_e(\epsilon,n)$ by Appendix 2.C.2.

[10] By Markov Lemma, we have $\delta_\epsilon^4(n) \triangleq 2|S_{UXYZ}|e^{-\epsilon^2 n\mu_{UXYZ}/6}$.

We now lower bound the term $-H(Y^n|Z^n\hat{S}^n)$ in (14). Define

$$\Gamma_1 \triangleq \begin{cases} 1 & \text{if } (Y^n, \hat{U}^n, V^n, Z^n) \in \mathcal{T}_{2\epsilon}^n(YUVZ), \\ 0 & \text{otherwise.} \end{cases}$$

$$\Delta_1 \triangleq \begin{cases} 1 & \text{if } (Y^n, \hat{U}^n, V^n) \in \mathcal{T}_{\epsilon}^n(YUV), \\ 0 & \text{otherwise.} \end{cases}$$

We can write

$$\begin{aligned} & H(Y^n|Z^n\hat{S}^n) \\ & \leqslant H(Y^n\Gamma_1\Delta_1|Z^n\hat{S}^n) \\ & = H(\Gamma_1\Delta_1|Z^n\hat{S}^n) + H(Y^n|Z^n\hat{S}^n\Gamma_1\Delta_1) \\ & \leqslant 2 + \sum_{\delta_1,\gamma_1\in\{0,1\}} \mathbb{P}[\Gamma_1 = \gamma_1|\Delta_1 = \delta_1]\mathbb{P}[\Delta_1 = \delta_1]H(Y^n|Z^n\hat{S}^n, \Gamma_1 = \gamma_1, \Delta_1 = \delta_1) \\ & \overset{(g)}{\leqslant} 2 + H(Y^n|Z^n\hat{S}^n, \Gamma_1 = 1, \Delta_1 = 1) + n(2\delta_\epsilon^3(n) + \delta_\epsilon^5(n))\log|\mathcal{Y}|, \end{aligned} \tag{20}$$

where (g) holds since $\mathbb{P}[\Delta_1 = 0] \triangleq \delta_\epsilon^3(n)$,[11] and $\mathbb{P}[\Gamma_1 = 0|\Delta_1 = 1] \leqslant \delta_\epsilon^5(n)$.[12] Indeed, we can apply Markov Lemma, since we have for every (y^n, z^n), $p(z^n|y^n) = \prod_{i=1}^{n} p_{Z|Y}(z_i|y_i)$, and $(\hat{U}^nV^n)$—Y^n—Z^n, which follows from the assumption X—Y—Z.[13]

$$\begin{aligned} H(Y^n|Z^n\hat{S}^n, \Gamma_1 = 1, \Delta_1 = 1) &= \sum_{z^n,\hat{s}^n} p(z^n, \hat{s}^n|1,1)H(Y^n|z^n, \hat{s}^n, \Gamma_1 = 1, \Delta_1 = 1) \\ &\leqslant \sum_{z^n,\hat{s}^n} p(z^n, \hat{s}^n|1,1)\log|\mathcal{T}_{2\epsilon}^n(Y|z^n, \hat{s}^n)| \\ &\leqslant \sum_{z^n,\hat{s}^n} p(z^n, \hat{s}^n|1,1)(nH(Y|ZUV)(1+2\epsilon)) \\ &\leqslant nH(Y|ZUV)(1+2\epsilon). \end{aligned} \tag{21}$$

[11] We have $\delta_\epsilon^3(n) \leqslant P_e(\epsilon, n)$ by Appendix 2.C.2.

[12] By Markov Lemma, we have $\delta_\epsilon^5(n) \triangleq 2|S_{UVYZ}|e^{-\epsilon^2 n\mu_{UVYZ}/6}$.

[13] Note that the assumption of degraded sources is only necessary here. The use of this hypothesis is the weakness, at least for two-way communication (for one-way communication this assumption is not necessary), of a proof that consists of a successive design of reconciliation and privacy amplification, rather than a joint design as in [16], where the joint design is exploited to get the joint typicality of $(V^n, Y^n, \hat{U}^n, Z^n)$.

Hence by (20), (21),

$$H(Y^n|Z^nU^nV^n) \leqslant n(H(Y|ZUV) + r_4(\epsilon, n)), \tag{22}$$

where

$$r_4(\epsilon, n) \triangleq 2H(Y|ZUV)\epsilon + (2\delta_\epsilon^3(n) + \delta_\epsilon^5(n)) \log|\mathcal{Y}|+2/n. \tag{23}$$

Combining (14), (18), (22),

$$\begin{aligned} H(S^n|Z^n) \geqslant n[&H(Y|Z) + H(X|YZ) - H(X|YZU) \\ &- H(Y|ZUV) - r_3(\epsilon, n) - r_4(\epsilon, n)] - \delta_\epsilon(n). \end{aligned} \tag{24}$$

Then, remark that

$$\begin{aligned} &H(Y|Z) + H(X|YZ) - H(X|YZU) - H(Y|ZUV) \\ &= I(Y; UV|Z) + I(X; U|YZ) \\ &= H(U|Z) + H(V|UZ) - H(V|UYZ) - H(U|XYZ) \\ &\overset{(h)}{\geqslant} H(U|Z) + H(V|UZ) - H(V|UY) - H(U|X) \\ &= I(U; X) - I(U; Z) - I(V; Z|U) + I(V; Y|U), \end{aligned} \tag{25}$$

where (h) holds because conditioning reduces entropy. Hence, by (9), (12), (24), and (95)

$$\begin{aligned} &H_\infty(S^N|z^N, a^m, b^m, c^m, \Theta = 1, \Upsilon = 1) \\ &\qquad \geqslant N[I(U; Y) + I(V; X|U) - I(U; Z) - I(V; Z|U) - r_5(\epsilon, n, m)], \end{aligned} \tag{26}$$

where

$$r_5(\epsilon, n, m) \triangleq r_1(\epsilon, n) + r_2(\epsilon, n, m) + r_3(\epsilon, n) + r_4(\epsilon, n) + \delta_\epsilon(n)/n. \tag{27}$$

Set k to be less than the lower bound in (26) by $\sqrt{N}$:

$$k \triangleq \lfloor N[I(U; Y) + I(V; X|U) - I(U; Z) - I(V; Z|U) - r_5(\epsilon, N)] - \sqrt{N} \rfloor. \tag{28}$$

Now with (26) and (28), we can apply Theorem 2.3.2 to lower bound $H(K|U_d Z^N A^m B^m C^m, \Upsilon = 1, \Theta = 1)$ by $k - \delta^*(N)$, where

$$\delta^*(N) = 2^{-\sqrt{N}/\log N} \left(k + \sqrt{N}/\log N\right).$$

Thus, we can finally lower bound $H(K|U_d Z^N A^m B^m C^m)$ in (11):

$$\begin{aligned} H(K|U_d Z^N A^m B^m C^m) &\geqslant \left(1 - \delta_\epsilon^0(m) - 2^{-\sqrt{N}}\right)(k - \delta^*(N)) \\ &= k - r_6(\epsilon, n, m), \end{aligned}$$

where

$$r_6(\epsilon, n, m) \triangleq \left(1 - \delta_\epsilon^0(m) - 2^{-\sqrt{N}}\right)\delta^*(N) + \left(\delta_\epsilon^0(m) + 2^{-\sqrt{N}}\right)k.$$

Moreover, the leakage is such that

$$I(K; U_d Z^N A^m B^m C^m) = H(K) - H(K|U_d Z^N A^m B^m C^m) \leqslant r_6(\epsilon, n, m), \tag{29}$$

with $r_6(\epsilon, n, m)$ vanishing to zero for a fixed n as m goes to infinity. The keys computed by Alice and Bob are asymptotically the same for a fixed n as m goes to infinity, since

$$\mathbb{P}[K \neq \hat{K}] \leqslant \mathbb{P}[(S^n)^m \neq (\hat{S}^n)^m] \leqslant \delta_e(m). \tag{30}$$

Then, by (10), (13), (19), (23), (27), we have that $r_5(\epsilon, n, m)$ vanishes to zero for n large enough and as m goes to infinity, thus the secret key rate $R \triangleq k/N$ is asymptotically as close as desired to

$$I(U;Y) - I(U;Z) + I(V;X|U) - I(V;Z|U).$$

Note that it is not exactly the bound proposed in Theorem 2.3.1.a for the WSK capacity. We finish the proof as follows. If $I(V;X|U) \leqslant I(V;Z|U)$, in the reconciliation we set $R_2 = 0$ so that the asymptotic secret key rate is now as close as desired to

$$I(U;Y) - I(U;Z) + [I(V;X|U) - I(V;Z|U)]^+.$$

Then, if $I(U;Y) \leqslant I(U;Z)$, in the reconciliation protocol, we choose $S^n = V^n$ (see the beginning of the proof), and we assume that U^N is provided by a genie to Eve. Consequently, we obtain instead of Equation (12),

$$\begin{aligned} &H_\infty(V^N|z^N, u^N, b^m, c^m, \Theta = 1, \Upsilon = 1) \\ &\qquad \geqslant m(H(V^n|Z^nU^n) - nr_2(\epsilon, n, m)) - N(I(V;Y|U) - I(V;X|U) - r_1(\epsilon, n)), \end{aligned}$$

and conclude in the same manner, to obtain an asymptotic secret key rate as close as desired to

$$[I(U;Y) - I(U;Z)]^+ + [I(V;X|U) - I(V;Z|U)]^+.$$

2.A.2 Continuous case

We use the following lemma to extend the result to the continuous case by means of quantization.

Lemma 2.1.2 ([71–73]). *Let X and Y be two real-valued random variables with probability distribution $\mathbb{P}_X$ and $\mathbb{P}_Y$ respectively. Let $\mathcal{E}_{\Delta_1} = \{E_i\}_{i \in \mathcal{I}}$, $\mathcal{F}_{\Delta_2} = \{F_j\}_{j \in \mathcal{J}}$ be two partitions of X and Y such that for any $i \in \mathcal{I}, \mathbb{P}_X(E_i) = \Delta_1$, for any $j \in \mathcal{J}, \mathbb{P}_Y(F_j) = \Delta_2$, where $\Delta_1, \Delta_2 > 0$. Let X_{Δ_1}, Y_{Δ_2} be the quantized version of X, Y with respect to the partitions $\mathcal{E}_{\Delta_1}$, $\mathcal{F}_{\Delta_2}$ respectively. Then, we have*

$$I(X;Y) = \lim_{\Delta_1, \Delta_2 \to 0} I(X_{\Delta_1}, Y_{\Delta_2}).$$

Note that a quantization of the eavesdropper observation Z^n might underestimate its knowledge from the legitimate users point of view and implicitly increase the leakage. However, by Lemma 2.1.2, for any $\delta > 0$, if the quantized version $Z^n_{\Delta^n}$ of Z^n is fine enough, then the leakage is not compromised and

$$|I(K;MZ^n) - I(K;MZ^n_{\Delta^n})| < \delta.$$

This argument is also used in [6, 74, 75].

We perform the quantization as follows. As in Lemma 2.1.2, we jointly quantify X, Y, Z, U and V to form X_{Δ_X}, Y_{Δ_Y}, $Z_{\Delta_Z}, U_{\Delta_U}$, V_{Δ_V} such that $\Delta_X = \Delta_Y = \Delta_Z = \Delta_U = \Delta_V = l^{-b}$ and $|\mathcal{X}_{\Delta_X}|= |\mathcal{Y}_{\Delta_Y}|= |\mathcal{Z}_{\Delta_Z}|= |\mathcal{U}_{\Delta_U}|= |\mathcal{V}_{\Delta_V}|= l^b$ with $b > 0$. We now apply the proof of the discrete case to the random variables X_{Δ_X}, Y_{Δ_Y}, $Z_{\Delta_Z}, U_{\Delta_U}$, V_{Δ_V}. By Lemma 2.1.2, we can fix l large enough such that

$$\begin{aligned} |I(U_{\Delta_U}; Y_{\Delta_Y}) - I(U;Y)|&< \delta/4, \\ |I(V_{\Delta_V}; X_{\Delta_X}|U_{\Delta_U}) - I(V;X|U)|&< \delta/4, \\ |I(U_{\Delta_U}; Z_{\Delta_Z}) - I(U;Z)|&< \delta/4, \\ |I(V_{\Delta_V}; Z_{\Delta_Z}|U_{\Delta_U}) - I(V;Z|U)|&< \delta/4, \end{aligned}$$

and Equation (28) becomes

$$k \geqslant \lfloor N[I(Y;U) - I(V;X|U) - I(U;Z) - I(V;Z|U) - r_5(\epsilon, n, m) - \delta] - \sqrt{N} \rfloor.$$

At this point, we cannot conclude with the last inequality. Indeed, in the term $r_5(\epsilon, n, m)$ are hidden the following terms:

$$\begin{gathered} H(X_{\Delta_X}|ZY_{\Delta_Y}U_{\Delta_U})\epsilon \text{ (see (19))}, \\ H(Y_{\Delta_Y}|Z_{\Delta_Z}U_{\Delta_U}V_{\Delta_V})\epsilon \text{ (see (23))}, \\ H(U_{\Delta_U})\epsilon \text{ and } H(V_{\Delta_V}|U_{\Delta_U})\epsilon \text{ (by definition of } r_0(\epsilon)), \end{gathered}$$

which do not vanish to 0 as l get large. Now, if we choose $\epsilon = n^{-a}$, where $a \in]0, 1/2[$, so that for $i \in \{0, 1, 2, 3, 5\}$, $\delta_\epsilon^i(n)$ vanishes as n get large for l fixed,[14] then the asymptotic secret-key rate, for n large enough and as m goes to infinity becomes as close as desired to

$$I(Y;U) - I(V;X|U) - I(U;Z) - I(V;Z|U).$$

Moreover, the leakage in (29), and the key error probability between Alice an Bob in (30), still vanish to zero for n large enough and as m goes to infinity.

[14] Recall that $P_e(\epsilon, n)$ decreases exponentially to zero as $n\epsilon^2$ goes to infinity.

2.B Proof of Theorem 2.4.4

As in [4], Theorem 2.4.4 is not directly deduced from Theorem 2.4.3. We first consider the case of one-way public communication, in which Alice sends messages to Bob, a first time with rate R_1' and a second time with rate R_2'. For this scenario we note C_{rec}^* the reconciliation capacity.

We can modify the proof of Proposition 2.5.2 to obtain for $R_1', R_2' \in \mathbb{R}^+$,

$$C_{\mathrm{rec}}^*(R_1', R_2') \geqslant \max_{U,V} \left[I(U;Y) + I(V;Y|U) \right]$$

subject to

$$R_1' \geqslant I(X;U|Y) \tag{31}$$

$$R_2' \geqslant I(V;X|YU) \tag{32}$$

$$U \text{—} X \text{—} Y,\ V \text{—} UX \text{—} Y.$$

Then, we can modify the proof of Theorem 2.4.3 to prove that we can achieve the rate

$$R_{\mathrm{WSK}}^*(R_1', R_2') \triangleq \max_{U,V} \left([I(Y;U) - I(Z;U)]^+ + [I(Y;V|U) - I(Z;V|U)]^+ \right),$$

subject to rate constraints (31), (32) and Markov conditions

$$U \text{—} X \text{—} YZ,\ V \text{—} UX \text{—} YZ, \tag{33}$$

by a reconciliation phase followed by a privacy amplification phase performed with extractors, and this time without the assumption $X \to Y \to Z$. Note that Markov condition

$$U \text{—} V \text{—} X \text{—} YZ, \tag{34}$$

implies Markov conditions (33), and that if Markov condition (34) holds, then the rate constraint (32) becomes

$$R_2' \geqslant I(X;V|U) - I(Y;V|U) \geqslant I(X;V) - I(Y;V) - R_1'.$$

Hence, for $R_1', R_2' > 0$ satisfying $R_1' + R_2' = R_1$,

$$R^*_{\text{WSK}}(R_1', R_2') \geqslant \max_{U,V}[I(Y;V|U) - I(Z;V|U)],$$

subject to rate constraint $R_1 \geqslant I(X;V) - I(Y;V)$ and Markov condition (34). We conclude by observing that $C_{\text{WSK}}(R_1, 0) \geqslant R^*_{\text{WSK}}(R_1', R_2')$.

2.C Proof of Proposition 2.5.2

2.C.1 Converse

Let $R_1, R_2 \in \mathbb{R}^+$. We first establish the rate constraints on R_1 and R_2. We have

$$\begin{aligned} nR_1 &\geqslant H(A) \\ &\geqslant I(A;X^n) - I(A;Y^n) \\ &\overset{(a)}{=} n[I(A;X_J|\tilde{U}) - I(A;Y_J|\tilde{U})] \\ &\overset{(b)}{=} n[I(U;X_J) - I(U;Y_J)], \end{aligned} \tag{35}$$

where (a) holds by [76, Lemma 4.1], if we set $\tilde{U} \triangleq X^{J-1}Y^N_{J+1}J$ and J is a random variable uniformly distributed on $[\![1, n]\!]$, independent of all previous random variables, (b) holds if we set $U \triangleq A\tilde{U}$, since X_J and $\tilde{U}$ are independent.

Similarly, we have

$$\begin{aligned} nR_2 \;&\geqslant\; H(B|A) \\ &\overset{(c)}{\geqslant} H(B|X^n) + H(\hat{S}|S) - n\delta(\epsilon) \\ &\overset{(d)}{\geqslant} I(\hat{S};B|X^n) + H(\hat{S}|BX^n) - n\delta(\epsilon) \\ &= H(\hat{S}|X^n) - n\delta(\epsilon) && (36) \\ &= H(\hat{S}|A) - I(\hat{S};X^n|A) - n\delta(\epsilon) \\ &\overset{(e)}{=} I(\hat{S};Y^n|A) - I(\hat{S};X^n|A) - n\delta(\epsilon) \\ &\overset{(f)}{=} n[I(V;Y_J|U) - I(V;X_J|U)] - n\delta(\epsilon), \end{aligned}$$

where (c) holds because A is a function of X^n and by Fano's inequality, since for any $\epsilon > 0$, there exists a reconciliation protocol such that $\mathbb{P}[S \neq \hat{S}] \leqslant \delta(\epsilon)$,[15] (d) holds since $S = \eta_a(X^n, B)$, (e) holds since $\hat{S} = \eta_b(Y^n, A)$, (f) holds by [76, Lemma 4.1] and if we set $V \triangleq \hat{S}$.

We now determine the reconciliation capacity bound.

$$\begin{aligned}
I(\hat{S}; X^n) &= \sum_{i=1}^{n} I(\hat{S}; X_i | X^{i-1}) \\
&\overset{(a)}{=} \sum_{i=1}^{n} I(\hat{S} X^{i-1}; X_i) \\
&\leqslant \sum_{i=1}^{n} I(\hat{S} X^{i-1} Y_{i+1}^{n}; X_i) \\
&= n \sum_{i=1}^{n} \mathbb{P}[J = i] I(\hat{S} X^{J-1} Y_{J+1}^{n}; X_J | J = i) \\
&= n I(\hat{S} \tilde{U}; X_J | J) \\
&\leqslant n I(VU; X_J), \qquad (37)
\end{aligned}$$

where (a) holds because the X_i's are i.i.d.. Then,

$$\begin{aligned}
H(\hat{S}) - H(AB) &= I(\hat{S}; X^n) + H(\hat{S}|X^n) - H(A) - H(B|A) \\
&\overset{(b)}{\leqslant} n I(VU; X_J) - H(A) + n\delta(\epsilon) \\
&\overset{(c)}{\leqslant} n[I(V; X_J|U) + I(U; Y_J) + \delta(\epsilon)],
\end{aligned}$$

where (b) holds by (37) and since $H(\hat{S}|X^n) \leqslant H(B|A) + n\delta(\epsilon)$ by (36), and (c) holds by (35).

For a DMS, standard techniques [76] show that $|\mathcal{U}| \leqslant |\mathcal{X}| + 2$ and $|\mathcal{V}| \leqslant |\mathcal{Y}|$.

2.C.2 Achievability

The proof for a DMS is similar to Wyner-Ziv coding [77], we only describe the protocol. In the following, for $n \in \mathbb{N}$ and $\epsilon > 0$, we note $\mathcal{T}_\epsilon^n(X)$ the set of ϵ-letter-typical sequences [78] (also called "robustly typical sequence" in [61]) with respect to p_X. We

[15] $\delta(\epsilon)$ denotes a function of ϵ such that $\lim_{\epsilon \to 0} \delta(\epsilon) = 0$.

also define conditional typical sets as follows, $\mathcal{T}_{\epsilon}^{n}(Y|x^n) \triangleq \{y^n : (x^n, y^n) \in \mathcal{T}_{\epsilon}^{n}(XY)\}$. We note $\mu_X \triangleq \min_{x \in supp(p_X)} p_X(x)$. Let $\epsilon > 0$, and define $\epsilon_1 \triangleq \frac{1}{2}\epsilon$, $\epsilon_2 \triangleq 2\epsilon$.

Code construction: Fix a joint probability distribution p_{UX} on $\mathcal{U} \times \mathcal{X}$ and p_{UVY} on $\mathcal{U} \times \mathcal{V} \times \mathcal{Y}$.

- Define
$$R_u = I(X;U|Y) + 6\epsilon H(U),$$
$$R'_u = I(Y;U) - 3\epsilon H(U).$$

  Generate $2^{n(R_u + R'_u)}$ codewords, labeled $u^n(\omega, \nu)$ with $(\omega, \nu) \in [\![1, 2^{nR_u}]\!] \times [\![1, 2^{nR'_u}]\!]$, by generating the symbols $u_i(\omega, \nu)$ for $i \in [\![1, n]\!]$ and $(\omega, \nu) \in [\![1, 2^{nR_u}]\!] \times [\![1, 2^{nR'_u}]\!]$ independently according to p_U.

- Define
$$R_v = I(V;Y|XU) + 6\epsilon_2 H(V|U),$$
$$R'_v = I(V;X|U) - 3\epsilon_2 H(V|U).$$

  For each (ω, ν), generate $2^{n(R_v + R'_v)}$ codewords, labeled $v^n(\omega, \nu, k, l)$ with $(k, l) \in [\![1, 2^{nR_v}]\!] \times [\![1, 2^{nR'_v}]\!]$, by generating the symbols $v_i(\omega, \nu, k, l)$ for $i \in [\![1, n]\!]$ and $(k, l) \in [\![1, 2^{nR_v}]\!] \times [\![1, 2^{nR'_v}]\!]$ independently according to $p_{V|U=u_i(\omega,\nu)}$.

Step1. At Alice's side:

- Given x^n, find a pair (ω, ν) s.t $(x^n, u^n(\omega, \nu)) \in \mathcal{T}_{\epsilon}^{n}(XU)$. If we find several pairs, we choose the smallest one (by lexicographic order). If we fail we choose $(\omega, \nu) = (1, 1)$.
- Define $s_1^n \triangleq u^n(\omega, \nu)$.
- Transmit $a \triangleq \omega$.

Step2. At Bob's side:

- Given y^n and a, find $\tilde{\nu}$ s.t $(y^n, u^n(\omega, \tilde{\nu})) \in \mathcal{T}_{\epsilon}^{n}(YU)$ and define $\hat{s}_1^n \triangleq u^n(\omega, \tilde{\nu})$. If there is one or more such $\tilde{\nu}$, choose the lowest, otherwise set $\tilde{\nu} = 1$. Find a pair

(k,l) such that $(\hat{s}_1^n, y^n, v^n(\omega,\tilde{\nu},k,l)) \in \mathcal{T}_{\epsilon_2}^n(UYV)$. If there is one or more such (k,l), choose the lowest, otherwise set $(k,l)=(1,1)$.

- Transmit $b=k$.
- Define $\hat{s}_2^n \triangleq v^n(\omega,\tilde{\nu},k,l)$ and $\hat{s}^n \triangleq (\hat{s}_1^n, \hat{s}_2^n)$.

Step3. At Alice's side:

- Given $s_1^n = u^n(\omega,\nu)$ and b, find $\tilde{l}$ s.t $(x^n, s_1^n, v^n(\omega,\tilde{\nu},k,\tilde{l})) \in \mathcal{T}_{\epsilon_2}^n(XUV)$. If there is one or more such $\tilde{l}$, choose the lowest, otherwise set $\tilde{l}=1$.
- Define $s_2^n \triangleq v^n(\omega,\tilde{\nu},k,\tilde{l})$ and $s^n \triangleq (s_1^n, s_2^n)$.

We can show by standard arguments that there exists a code, such that after one repetition of the protocol, Alice obtains $S^n = U^n\hat{V}^n$, whereas Bob has $\hat{S}^n = \hat{U}^n V^n$ with $\mathbb{P}[\hat{U}^n \neq U^n] \leqslant \delta_\epsilon(n)$,[16] $\mathbb{P}[\hat{V}^n \neq V^n] \leqslant \delta_\epsilon(n)$, $\mathbb{P}[\hat{S}^n \neq S^n|\mathcal{R}_n] \leqslant P_e(\epsilon,n)$[17] and (U^n,X^n), $(\hat{U}^n,Y^n)$, $(\hat{U}^n,Y^n,V^n)$, $(U^n,\hat{V}^n,X^n)$ jointly typical with probability approaching one for n large.

To extend the result to a CMS, we proceed as in the proof of Theorem 2.4.3.

2.D Proof of Proposition 2.5.3

2.D.1 Proof of Part i)

The achievability and converse proof can be found in [56], it remains to prove that equality holds in the rate constraint (4) and that $|\mathcal{U}|\leqslant|\mathcal{X}|$.

2.D.1.1 Equality constraint

We start with the following lemma.

Lemma 2.4.3. *$f(U) \triangleq I(Y;U)$ and $f_1(U) \triangleq I(X;U|Y)$ are convex in $p_{U|X}$.*

Proof. Let $\lambda \in [0,1]$, let U_1, U_2 defined by $p_{U_1|X}$ and $p_{U_2|X}$ respectively, be s.t. U_1—X—Y and U_2—X—Y.

[16] $\delta_\epsilon(n)$ denotes a function of ϵ and n such that $\lim_{n\to\infty}\delta_\epsilon(n)=0$.

[17] We can show that $P_e(\epsilon,n)$ decreases exponentially to zero as $n\epsilon^2$ goes to infinity.

We introduce the random variable $Q \in \{1, 2\}$ independent of all others and set $U = U_Q$.

$$Q \triangleq \begin{cases} 1 & \text{with probability } \lambda, \\ 2 & \text{with probability } 1 - \lambda. \end{cases}$$

$$\begin{aligned} I(Y;U) &\leqslant I(Y;UQ) \\ &= I(Y;Q) + I(Y;U|Q) \\ &\overset{(a)}{=} I(Y;U|Q) \\ &= \lambda I(Y;U_1) + (1-\lambda) I(Y;U_2), \end{aligned}$$

where (a) holds since Y and Q are independent.

$$\begin{aligned} I(X;U|Y) &\leqslant I(X;UQ|Y) \\ &= I(X;Q|Y) + I(X;U|YQ) \\ &\overset{(b)}{=} I(X;U|YQ) \\ &= \lambda(I(X;U_1|Y) + (1-\lambda)(I(X;U_2|Y), \end{aligned}$$

where (b) holds because $H(X|YQ) = H(X|Y)$, since Q and (X,Y) are independent.

□

By Lemma 2.4.3, $f(U)$ and $f_1(U)$ are convex in $p_{U|X}$. Define $\Delta \triangleq \{\mathbf{u} \in \mathbb{R}^{|\mathcal{U}||\mathcal{X}|} : \forall i, j \in [\![1, |\mathcal{U}|]\!] \times [\![1, |\mathcal{X}|]\!], \sum_{k=1}^{|\mathcal{U}|} u_{kj} = 1, u_{ij} \geqslant 0\}$, and $\mathcal{C} \triangleq \{\mathbf{u} \in \Delta : f_1(\mathbf{u}) \leqslant R_1\}$. We first show that $\mathcal{C}$ is convex compact, with extreme points in $\{\mathbf{u} \in \Delta : f_1(\mathbf{u}) = R_1\}$:

- $\mathcal{C}$ is the preimage of $[0, R_1]$ by the continuous function f_1, thus $\mathcal{C}$ is closed. We deduce that $\mathcal{C}$ is compact, since $\mathcal{C} \subset [0,1]^{|\mathcal{U}||\mathcal{X}|}$ and $[0,1]^{|\mathcal{U}||\mathcal{X}|}$ is compact.

- $\mathcal{C}$ is convex by convexity of f_1, since the sublevels of a convex function are convex sets.

- Let $\mathbf{u}_1 \in \mathcal{C}$ s.t. $f_1(\mathbf{u}_1) = R_1 - \delta$, with $\delta > 0$. By continuity of f_1, $\exists \epsilon_0, \forall \mathbf{u} \in \mathcal{B}(\mathbf{u}_1, \epsilon_0), |f_1(\mathbf{u}) - f_1(\mathbf{u}_1)| < \delta$. Let $\mathbf{u}_0 \in \mathcal{B}(\mathbf{u}_1, \epsilon_0) \backslash \{\mathbf{u}_1\}$, $\lambda \in \left\{-\frac{1}{2}, +\frac{1}{2}\right\}$ and $\mathbf{u}_\lambda = \lambda \mathbf{u}_0 + (1-\lambda)\mathbf{u}_1$.

  Then $||\mathbf{u}_\lambda - \mathbf{u}_1|| = ||\lambda(\mathbf{u}_0 - \mathbf{u}_1)|| \leqslant |\lambda|\epsilon_0$, which means $\mathbf{u}_\lambda \in \mathcal{C}$. Hence, $\frac{1}{2}\mathbf{u}_{\lambda=+1/2} + \frac{1}{2}\mathbf{u}_{\lambda=-1/2} = \mathbf{u}_1$, and we conclude that $\mathbf{u}_1$ is not an extreme point. Hence, the set of extreme points of $\mathcal{C}$ is a subset of $\{\mathbf{u} \in \Delta : f_1(\mathbf{u}) = R_1\}$.

Since f is continuous, it reaches a maximum $\mathbf{u}_{max}$ on the compact $\mathcal{C}$. Then, since f is convex and $\mathcal{C}$ is a convex compact, by the Krein-Milman Theorem,[18] $\mathbf{u}_{max}$ is a convex linear combination of extreme points of $\mathcal{C}$ (existence of such extreme points comes directly from the Krein-Milman theorem, since $\mathcal{C} \neq \emptyset$). Hence, $\mathbf{u}_{max} = \sum_{k=1}^{n} \lambda_k \mathbf{u}_k$, with $\sum_{k=1}^{n} \lambda_k = 1$, $\lambda_1, \lambda_2, \ldots, \lambda_n \geqslant 0$ and $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ extreme points of $\mathcal{C}$. By convexity of f,

$$f(\mathbf{u}_{max}) \leqslant \sum_{k=1}^{n} \lambda_k f(\mathbf{u}_k) \leqslant \sum_{k=1}^{n} \lambda_k f(\mathbf{u}_{max}) = f(\mathbf{u}_{max}),$$

$$\text{thus } \sum_{k=1}^{n} \lambda_k (f(\mathbf{u}_{max}) - f(\mathbf{u}_k)) = 0,$$

which means that there exists $i \in [\![1, n]\!]$ s.t $f(\mathbf{u}_{max}) = f(\mathbf{u}_i)$. We conclude that $\mathbf{u}_{max}$ is an extreme point of $\mathcal{C}$. This result is known as the maximum principle [79].

2.D.1.2 Cardinality bound $|\mathcal{U}| \leqslant |\mathcal{X}|$

This result is a special case of a more general one that we prove in Appendix 2.D.2.2.

2.D.2 Proof of Part ii)

The proof is partially found in [60] and all that remains to be proved are the equality in the communication rate constraint and the range constraint $|\mathcal{U}| \leqslant |\mathcal{X}|$.

2.D.2.1 Equality in the constraint

To prove that equality holds in the constraint for the argument of the maximum in Proposition 2.5.3.b, we can reuse the proof of Proposition 2.5.3.a in Appendix 2.D.1.1,

[18] A compact convex subset of a locally convex topological vector space is the closed convex hull of the set of its extreme points. Actually, only a weaker version is used since a finite dimensional space is considered.

so that we only need to show that $f(U) = I(Y;U) - I(Z;U)$ is convex in $p_{U|X}$. To obtain the convexity of f, we replace (X,Y) by (Y,Z) in the function f_1 of Lemma 2.4.3.

2.D.2.2 Range constraint $|\mathcal{U}|\leqslant |\mathcal{X}|$

The proof relies on a technique used in [80].

Define

$$\begin{aligned}\mathcal{R} &\triangleq \{(R,R_1) : R \geqslant I(Y;U) - I(Z;U),\\ &\qquad\qquad R_1 \geqslant I(X;U) - I(Y;U), \text{ with } U—X—Y—Z\},\\ \mathcal{C} &\triangleq \{(R,R_1) : R \geqslant I(Y;U) - I(Z;U),\\ &\qquad\qquad R_1 = I(X;U) - I(Y;U), \text{ with } U—X—Y—Z\}.\end{aligned}$$

Note that the capacity region $\mathcal{C}$ is from Proposition 2.5.3.b and that the equality in the communication rate constraint is crucial to make it a subset of $\mathcal{R}$. By [80, Lemma 3],

$$\mathcal{R} = \left\{(R,R_1) : \forall \lambda_1,\lambda_2 \in \mathbb{R}^+, \lambda_1 R + \lambda_2 R_1 \geqslant G(\lambda_1,\lambda_2)\right\},$$

where $\forall \lambda_1,\lambda_2 \in \mathbb{R}^+$,

$$G(\lambda_1,\lambda_2) \triangleq \inf_{\substack{U \text{ s.t}\\ U—X—Y—Z}} [\lambda_1(I(Y;U) - I(Z;U)) + \lambda_2(I(X;U) - I(Y;U))].$$

Consequently $G(\lambda_1,\lambda_2)$ is sufficient information to describe $\mathcal{R}$. Then, we show that for all $\lambda_1,\lambda_2 \in \mathbb{R}^+$, $G(\lambda_1,\lambda_2)$ can be achieved by considering a discrete random variable U such that $|\mathcal{U}|\leqslant |\mathcal{X}|$.

Let $\lambda_1,\lambda_2 \in \mathbb{R}^+$, let $\mathcal{P}$ in [80, Lemma 2] be the $|\mathcal{X}|$-dimensional probability simplex, and let $\mathcal{X} = \{x_i\}_{i=1}^{|\mathcal{X}|}$. Consider $\mathcal{P}$ as a set of elements of the form

$$\left(\mathbb{P}[X = x_1|U = u], \mathbb{P}[X = x_2|U = u], \ldots,\ \mathbb{P}[X = x_{|\mathcal{X}|}|U = u]\right),$$

with $u \in \mathcal{U}$. Then, each probability distribution on U defines a measure μ on $\mathcal{P}$. Define $H_P(X)$, $H_P(Y)$, and $H_P(Z)$ as the entropies of X, Y, and Z respectively, when the distribution of X is $P \in \mathcal{P}$. Define

$$f_1(P) \triangleq \lambda_1(H_P(Z) - H_P(Y)) + \lambda_2(H_P(Y) - H_P(X))$$
$$f_j(P) \triangleq P(x_j), \text{ for } j \in [\![2, |\mathcal{X}|]\!].$$

Let P_X^* achieve $G(\lambda_1, \lambda_2)$, and let μ^* be such that $\int_{\mathcal{P}} P\mu^*(dP) = P_X^*$. Denote by $H^*(X)$ the entropy of X under probability distribution P_X^*. Then, by [80, Lemma 2], there exists $P_1, P_2, \dots, P_{|\mathcal{X}|}$, and $\alpha_1, \alpha_2, \dots, \alpha_{|\mathcal{X}|}$ such that, $\sum_{i=1}^{|\mathcal{X}|} \alpha_i = 1$,

$$\forall j \in [\![2, |\mathcal{X}|]\!], P_X^*(x_j) = \int_{\mathcal{P}} f_j(P)\mu^*(dP) = \sum_{i=1}^{|\mathcal{X}|} \alpha_i f_j(P_i),$$

and,

$$\begin{aligned}&\lambda_1(H^*(Z|U) - H^*(Y|U)) + \lambda_2(H^*(Y|U) - H^*(X|U))\\ &= \int_{\mathcal{P}} f_1(P)\mu^*(dP) = \sum_{i=1}^{|\mathcal{X}|} \alpha_i f_1(P_i).\end{aligned}$$

From $P_X^*(x_j)$, $j \in [\![2, |\mathcal{X}|]\!]$, we can compute $H^*(X)$, $H^*(Y)$, and $H^*(Z)$, then

$$\begin{aligned}&\lambda_1(H^*(Y) - H^*(Y|U) - H^*(Z) + H^*(Z|U))\\ &\qquad + \lambda_2(H^*(X) - H^*(X|U) - H^*(Y) + H^*(Y|U))\\ &= \lambda_1(I^*(Y;U) - I^*(Z;U)) + \lambda_2(I^*(X;U) - I^*(Y;U))\\ &= G(\lambda_1, \lambda_2).\end{aligned}$$

We have thus shown that we can choose U such that $|\mathcal{U}| \leqslant |\mathcal{X}|$ to achieve $G(\lambda_1, \lambda_2)$. Consequently, it is enough to consider U such that $|\mathcal{U}| \leqslant |\mathcal{X}|$, to form the set $\mathcal{R}$, as well as the set $\mathcal{C}$, since $\mathcal{C} \subset \mathcal{R}$.

2.E Proof of Proposition 2.5.4

If $R_1 \geqslant H(X|Y)$, then by Proposition 2.5.3.b $C_{\text{WSK}}(R_1, 0) = \mathbb{I}(X;Y)$. Assume $R_1 \in]0; H(X|Y)[$ in the following. We note $\mathcal{X} = \{0, 1\}$ and by Proposition 2.5.3.b, we can assume $\mathcal{U} = \{u_1, u_2\}$. We note $\beta_1 = p(X = 1|U = u_1)$ and $\beta_2 = p(X = 0|U = u_2)$. We can write

$$
\begin{aligned}
& I(U; X) - I(U; Y) - (H(X) - H(Y)) \\
& = -\sum_{i=1,2} p(u_i)[H(X|U = u_i) - H(Y|U = u_i)] \\
& = -\sum_{i=1,2} p(u_i)[H_b(\beta_i) - H(Y|U = u_i)] \\
& = -\sum_{i=1,2} p(u_i) \left[H_b(\beta_i) + \textstyle\sum_{y\in\mathcal{Y}} p(y|u_i) \log p(y|u_i) \right], \qquad (38)
\end{aligned}
$$

with $\forall y \in \mathcal{Y}$,

$$p(y|u_1) = (1 - \beta_1)p(y|X = 0) + \beta_1 p(y|X = 1), \qquad (39)$$

$$p(y|u_2) = \beta_2 p(y|X = 0) + (1 - \beta_2)p(y|X = 1). \qquad (40)$$

Moreover, since the channel $p_{Y|X}$ is symmetric, there exists a permutation $\pi \in \mathfrak{S}_{|\mathcal{Y}|}$ such that

$$\forall y \in \mathcal{Y}, \forall x \in \mathcal{X}, p(y|x) = p(\pi(y)|x \oplus 1), \qquad (41)$$

where $\oplus$ denotes the modulo 2 operation. Thus by (38), (39), (40), (41) there exists $g_{Y|X}$[19] such that $H(Y|U = u_1) = g_{Y|X}(\beta_1)$, $H(Y|U = u_2) = g_{Y|X}(\beta_2)$. Then,

$$I(U; X) - I(U; Y) - (H(X) - H(Y)) = -\sum_{i=1,2} p(u_i) \left[H_b(\beta_i) - g_{Y|X}(\beta_i) \right]. \qquad (42)$$

Similarly, by using that the channel $p_{Z|X}$ is symmetric, there exists $g_{Z|X}$ such that $H(Z|U = u_1) = g_{Z|X}(\beta_1)$ and $H(Z|U = u_2) = g_{Z|X}(\beta_2)$. Thus, we also have

$$I(U; Y) - I(U; Z) - (H(Y) - H(Z)) = -\sum_{i=1,2} p(u_i) \left[g_{Y|X}(\beta_i) - g_{Z|X}(\beta_i) \right]. \qquad (43)$$

[19] The exact description of $g_{Y|X}$ is not important here, what matters is that $H(Y|U = u_1)$ and $H(Y|U = u_2)$ can be expressed with the same function.

Consider the region $\mathcal{R}_1 \triangleq \bigcup\limits_{\beta_0 \in [0,1]} \mathcal{R}_{\beta_0}$ and $\mathcal{R}_2 \triangleq \bigcup\limits_{(\beta_1,\beta_2)\in[0,1]^2} \mathcal{R}_{\beta_1,\beta_2}$, with

$$\begin{aligned}\mathcal{R}_{\beta_0} \triangleq \{(R,R_1) : R &\leqslant H(Y) - H(Z) - g_{Y|X}(\beta_0) + g_{Z|X}(\beta_0),\\ &\qquad R_1 \leqslant H(X) - H(Y) - H_b(\beta_0) + g_{Y|X}(\beta_0)\},\\ \mathcal{R}_{\beta_1,\beta_2} &\triangleq \{(R,R_1) : R \leqslant I(Y;U) - I(Z;U),\;\; R_1 \leqslant I(X;U) - I(Y;U)\}.\end{aligned}$$

We can verify that both regions $\mathcal{R}_1$ and $\mathcal{R}_2$ are convex and that $\mathcal{R}_1 \subset \mathcal{R}_2$. We will use a similar technique as in [81], based on Lemma 2.5.4, to show that $\mathcal{R}_1 = \mathcal{R}_2$.[20] Then, thanks to the refinement proposed in Proposition 2.5.3.b (equality in the constraint), we will be able to conclude for any $R_1 \in \mathbb{R}_+$,

$$\begin{aligned}C_{\text{WSK}}(R_1,0) = \max_{\beta_0\in[0,1]} &\big(H(Y) - H(Z) - g_{Y|X}(\beta_0) + g_{Z|X}(\beta_0)\big)\\ &\text{such that } R_1 = H(X) - H(Y) - H_b(\beta_0) + g_{Y|X}(\beta_0).\end{aligned}$$

Lemma 2.5.4 ([81] [79])**.** *Let $\mathcal{C} \subset \mathbb{R}^d$ be convex. Let $\mathcal{C}_1 \subset \mathcal{C}_2$ be two bounded convex subsets of $\mathcal{C}$, closed relative to $\mathcal{C}$. If every supporting hyperplanes of $\mathcal{C}_2$ intersects with $\mathcal{C}_1$, then $\mathcal{C}_1 = \mathcal{C}_2$.*

Let $(R, R_1) \in \mathcal{R}_2$, and let $\alpha \in [0,1]$, then we have by (42), (43)

$$\begin{aligned}\alpha R + (1-\alpha)R_1 &\leqslant \alpha(I(Y;U) - I(Z;U)) + (1-\alpha)(I(X;U) - I(Y;U))\\ &= \sum_{i=1,2} p(u_i)[\alpha(H(Y) - H(Z) - g_{Y|X}(\beta_i) + g_{Z|X}(\beta_i))\\ &\qquad + (1-\alpha)(H(X) - H(Y) - H_b(\beta_i) + g_{Y|X}(\beta_i))]\\ &\leqslant \alpha(H(Y) - H(Z) - g_{Y|X}(\beta^*) + g_{Z|X}(\beta^*)) + (1-\alpha)\\ &\qquad \times (H(X) - H(Y) - H_b(\beta^*) + g_{Y|X}(\beta^*)),\end{aligned} \tag{44}$$

[20] Note that the fact that R_1 and R are both lower bounds in $\mathcal{R}_1$ and $\mathcal{R}_2$ is crucial to show $\mathcal{R}_1 = \mathcal{R}_2$. The same argument cannot apply if R is a lower bound and R_1 an upper bound, whence the importance of the equality in the constraint shown in Proposition 2.5.3.b.

where

$$\beta^* \triangleq \underset{\beta}{\operatorname{argmax}}(\alpha(H(Y) - H(Z) - g_{Y|X}(\beta) + g_{Z|X}(\beta)) \\ + (1-\alpha)(1 - H(Y) - H_b(\beta) + g_{Y|X}(\beta))).$$

With the last inequality, we show that every supporting plane of $\mathcal{R}_2$ intersects $\mathcal{R}_1$. Note that the weight coefficients of (R, R_1) have been taken of the form $(\alpha, 1-\alpha)$ with $\alpha \in [0,1]$, because by positivity and convexity of $\mathcal{R}_2$, we only needed to consider hyperplanes (lines) with negative slope to apply Lemma 2.5.4.

Let (R^0, R_1^0) be a boundary point of $\mathcal{R}_2$. There exists a supporting hyperplane $\mathcal{H}_0$ at (R^0, R_1^0) defined by $(\alpha^0, 1-\alpha^0)$. By Equation (44), there exists $\beta_0^* \in [0,1]$ such that

$$\alpha^0 R^0 + (1-\alpha^0) R_1^0 \leqslant \alpha^0 R^* + (1-\alpha^0) R_1^*,$$

where

$$(R^*, R_1^*) \triangleq (H(Y) - H(Z) - g_{Y|X}(\beta_0^*) + g_{Z|X}(\beta_0^*), H(X) - H(Y) - H_b(\beta_0^*) + g_{Y|X}(\beta_0^*)).$$

Then, since $(R^*, R_1^*) \in \mathcal{R}_1 \subset \mathcal{R}_2$, we also have, by definition of $\mathcal{H}_0$

$$\alpha^0 R^* + (1-\alpha^0) R_1^* \leqslant \alpha^0 R^0 + (1-\alpha^0) R_1^0.$$

Hence, $\alpha^0 R^* + (1-\alpha^0) R_1^* = \alpha^0 R^0 + (1-\alpha^0) R_1^0$, and thus $(R^*, R_1^*) \in \mathcal{H}_0$.

2.F Proof of Proposition 2.5.7

Consider $X \sim \mathcal{N}(0, \sigma_x^2)$, $N \sim \mathcal{N}(0, \sigma_n^2)$, $Y = X + N$. We have $\sigma_y^2 = \sigma_x^2 + \sigma_n^2$ and

$$p_X(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{x^2}{2\sigma_x^2}\right],$$

$$p_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi}} \frac{\sigma_y}{\sigma_x \sigma_n} \exp\left[-\frac{1}{2\sigma_n^2}\frac{\sigma_y^2}{\sigma_x^2}\left(x - \frac{\sigma_x^2}{\sigma_y^2} y\right)^2\right].$$

Let $l \in \mathbb{N}^*$ and $k \in [\![1, l]\!]$. Define $t_k \triangleq a(2\frac{k-1}{l-1} - 1)$ and $\Delta \triangleq \frac{2a}{l-1}$. Let U be a scalar quantized version of X, defined as follows.

$$p_U(u_k) \triangleq \int_{t_k}^{t_{k+1}} p_X(x) dx = p_X(\bar{x}_k)\Delta,$$

$$\forall y \in \mathcal{Y}, p_{U|Y}(u_k|y) \triangleq p_{X|Y}(\bar{x}_k|y)\Delta,$$

where $\bar{x}_k \in [t_k, t_{k+1}]$ by the mean value theorem for integration. Hence,

$$H(U) = S_U - \log\Delta, \text{ with } S_U \triangleq -\Delta \sum_k p_X(\bar{x}_k)\log p_X(\bar{x}_k).$$

Observe that S_U is a Riemann sum that approaches $h(X) = -\int p_X \log p_X$. Thus, if we set $f(x) \triangleq -p_X(x)\log p_X(x)$, we can show that for any $a \in \mathbb{R}^+$,[21]

$$\begin{aligned}|h(X) - S_U| &= \left|\int f - S_U\right| \\ &\leqslant \left|\int_{-\infty}^{-a} f + \int_a^{+\infty} f\right| + \left|S_U - \int_{-a}^{a} f\right| \\ &\leqslant \epsilon_1(a) + K_1(a)\Delta,\end{aligned}$$

with $K_1(a) \triangleq a\max_{[-a,a]}|f'|$, $\epsilon_1(a) \triangleq e^{-\frac{a^2}{2\sigma_x^2}}[\alpha_1 a + \beta_1]$, and α_1, β_1 constants.

Similarly, if we define

$$S_{U|Y} \triangleq -\Delta \sum_k \int_y p_{XY}(\bar{x}_k, y)\log p_{X|Y}(\bar{x}_k|y)dy,$$

and $g(x) \triangleq \int p_{XY}(x,y)\log p_{X|Y}(x|y)dy$, then, as previously, we can show that for any $a \in \mathbb{R}^+$,

$$|h(X|Y) - S_{U|Y}| \leqslant \epsilon_2(a) + K_2(a)\Delta,$$

with $K_2(a) \triangleq a\max_{[-a,a]}|g'|$, $\epsilon_2(a) \triangleq e^{-\frac{a^2}{2\sigma_x^2}}[\alpha_2 a + \beta_2]$, and α_2, β_2 constants. Thus,

$$\log\Delta - (\epsilon_2(a) + K_2(a)\Delta) \leqslant h(X|Y) - H(U|Y) \leqslant \log\Delta + \epsilon_2(a) + K_2(a)\Delta.$$

Hence, for any $a \in \mathbb{R}^+$, if we take Δ small enough, then $|\log\Delta| \gg \epsilon_2(a) + K_2(a)\Delta$, such that $h(X|Y) - H(U|Y) \approx \log\Delta$, and

$$\begin{aligned}|I(X;Y) - I(Y;U)| &= |h(X) - S_U + S_{U|Y} - h(X|Y)| \\ &\leqslant \epsilon(a) + K(a)\Delta \\ &\leqslant \epsilon(a) + K(a)\exp[h(X|Y) - H(U|Y)] \\ &= \epsilon(a) + K(a)\exp[h(X|Y) - R_1],\end{aligned}$$

[21] We used a standard Riemann sum error bound, and $\text{erfc}(x) \leqslant e^{-x^2}$.

where $\epsilon(a) \triangleq \epsilon_1(a) + \epsilon_2(a)$, $K(a) \triangleq K_1(a) + K_2(a)$.

To sum up, Δ chosen small enough ensures that $I(Y;U)$ approaches $I(X;Y)$ exponentially fast as $R_1 > h(X|Y)$ increases.

CHAPTER 3

POLAR CODING SCHEMES FOR SECRET-KEY GENERATION

3.1 Summary

Practical implementations of secret-key generation are often based on sequential strategies, which, as seen in the previous chapter, handle reliability and secrecy in two successive steps, called reconciliation and privacy amplification. In this chapter, we propose an alternative approach based on polar codes that jointly deals with reliability and secrecy. Specifically, we propose secret-key capacity-achieving polar coding schemes for the following models: (i) the degraded binary memoryless source (DBMS) model with rate-unlimited public communication, (ii) the DBMS model with one-way rate-limited public communication, (iii) the 1-to-m broadcast model and (iv) the Markov tree model with uniform marginals. For models (i) and (ii) our coding schemes remain valid for non-degraded sources, although they may not achieve the secret-key capacity. For models (i), (ii) and (iii), our schemes rely on pre-shared secret seed of negligible rate; however, we provide special cases of these models for which no seed is required. Finally, we show an application of our results to secrecy and privacy for biometric systems. We thus provide the first examples of low-complexity secret-key capacity-achieving schemes that are able to handle vector quantization for model (ii), or multiterminal communication for models (iii) and (iv). This chapter is based on the results obtained in [82, 83].

3.2 Introduction

This chapter presents low-complexity secret-key capacity-achieving schemes based on polar codes [37] for some classes of source models. Note that polar codes have already been successfully used for secrecy in the context of the symmetric wire-tap channel model [38, 39, 84–86], and for the Slepian-Wolf coding problem [87–90], which is particularly relevant to secret-key generation. Note also that in [91], the journal version of [89], a first application of polar coding to a basic secret key generation setting was proposed. Unlike sequential methods, which successively handle reliability and secrecy, our schemes jointly deal with reliability and secrecy (see Definition 3.3.2 for more details). Both the sequential reliability-secrecy approach, and the direct approach with polar codes have their advantages. On the one hand, we have seen in Chapter 2 that sequential methods offer flexibility in design by separating reliability and secrecy and, unlike polar coding schemes, remain optimal for two-way rate-limited communication and continuous non-degraded sources. On the other hand, polar coding schemes may be easier to design and operate at lesser complexity in some scenarios. They also appear to be convenient to deal with vector quantization when the public communication is rate-limited.

The main result of this chapter is to develop polar coding schemes that achieve the secret-key capacity for the following models.

- The degraded binary memoryless source (DBMS) model with rate-unlimited public communication;
- The DBMS model with one-way rate-limited public communication;
- The 1-to-m broadcast model;
- The Markov tree model with uniform marginals.

For the first two models, the proposed polar coding schemes may also be used to generate secret keys for non-degraded sources, although they may not achieve the secret-key capacity. For the first three models, we assume that the legitimate users initialize their communication with a shared secret seed,[1] whose length is negligible

[1] If one assumes an authenticated public channel [2, 3] a shared small secret seed in the order of the logarithm of the length of the messages is also required for authentication [8].

compared to the number of source samples used to generate a key. As shown in Sections 3.5-3.7, there also exist special cases of the source statistics for which no seed is required.

Note that [92], obtained independently from the present work, develops an alternative polar coding solution for the BMS model with rate-unlimited public communication. The major difference between their approach and ours is that their construction is sequential, i.e., it *successively* deals with reliability and secrecy by means of reconciliation and privacy amplification, whereas our approach *jointly* deals with reliability and secrecy. The construction in [92, Th. 7] has the advantage of not requiring a seed. On the other hand, our protocol only requires one "polarization layer," whose construction is efficient, whereas the sequential approach of [92] requires an inner and an outer layer, the latter having no known efficient code construction as discussed in [92, Section III.C].

The remainder of the chapter is organized as follows. Section 3.3 formally introduces some notation and recall the general multi-terminal secret-key generation problem, which encompasses all the models specialized in subsequent sections. Section 3.4 describes polar coding primitives used in our proposed schemes. Section 3.5, describes a secret-key capacity-achieving scheme with polar codes for the DBMS model with unlimited communication rate. Section 3.6 provides a secret-key capacity-achieving scheme with polar codes for the DBMS model with one-way rate-limited public communication. Section 3.7 develops a secret-key capacity-achieving scheme with polar codes for the 1-to-m broadcast model. Section 3.8 studies a Markov tree model with uniform marginals and provides a secret-key capacity-achieving scheme with polar codes in special cases. Finally, Section 3.9, shows how to apply the results to the related problem of privacy and secrecy for some biometric systems.

3.3 Definitions and notation

We start by introducing some notation used throughout the chapter. For $n \in \mathbb{N}$ and $N \triangleq 2^n$, we let $G_N \triangleq \left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]^{\otimes n}$ be the source polarization transform defined in [87]. We note the components of a vector, $X^{1:N}$ with superscripts, i.e., $X^{1:N} \triangleq (X^1, X^2, \dots, X^N)$. For any set $\mathcal{A} \triangleq \{i_j\}_{j=1}^{|\mathcal{A}|}$ of indices in $[\![1, N]\!]$, we define $U^{1:N}[\mathcal{A}] \triangleq \left[U^{i_1}, U^{i_2}, \dots, U^{i_{|\mathcal{A}|}}\right]$.

We now describe the general model for multiterminal secret-key generation introduced in Section 1.2.1 in a more formal way. Let $m \geqslant 2$ be the number of terminals that wish to generate a common secret-key. Set $\mathcal{M} \triangleq [\![1, m]\!]$, and let $\mathcal{Z}$ and $\mathcal{X}_i$, for $i \in \mathcal{M}$ be arbitrary finite alphabets. Define $\mathcal{X}_\mathcal{M}$ as the Cartesian product of $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_m$. Consider a discrete memoryless multiple source $(\mathcal{X}_\mathcal{M}\mathcal{Z}, p_{X_\mathcal{M}Z})$, where $X_\mathcal{M} \triangleq (X_1, X_2, \dots, X_m)$ and the Cartesian product $\mathcal{X}_\mathcal{M} \times \mathcal{Z}$ is abbreviated as $\mathcal{X}_\mathcal{M}\mathcal{Z}$. For $i \in \mathcal{M}$, Terminal i observes the component X_i of $(\mathcal{X}_\mathcal{M}\mathcal{Z}, p_{X_\mathcal{M}Z})$, whereas an eavesdropper observes the component Z. The source is assumed to be outside the control of all parties, but its statistics are known to all parties. Communication is allowed between terminals over an authenticated noiseless public channel with communication rate $R_p \in \mathbb{R}^+ \cup \{+\infty\}$. A key-generation strategy is then formally defined as follows.

Definition 3.3.1. *Let $R_p \in \mathbb{R}^+ \cup \{+\infty\}$. Let $\mathcal{K}$ be a key alphabet of size 2^{NR}. The protocol defined by the following steps is called a $(2^{NR}, N, R_p)$ key-generation strategy with public communication, and is denoted by $\mathcal{S}_N$.*

1. *Terminal i, $i \in \mathcal{M}$, observes $X_i^{1:N}$.*

2. *The m terminals communicate, possibly interactively, over the public channel. All the public inter-terminal communications are collectively denoted by $\mathbf{F}$ and satisfy $H(\mathbf{F}) \leqslant NR_p$.*

3. *Terminal i, $i \in \mathcal{M}$, computes $K_i(X_i^{1:N}, \mathbf{F}) \in \mathcal{K}$.*

Let K be a random variable taking values in $\mathcal{K}$. The performance of a key-generation strategy $\mathcal{S}_N$ that allows the terminals in $\mathcal{M}$ to agree on the key K is measured in terms of the average probability of error between the keys

$$\mathbf{P}_e(\mathcal{S}_N) \triangleq \mathbb{P}[\exists i \in \mathcal{M} : K \neq K_i],$$

the information leakage to the eavesdropper

$$\mathbf{L}(\mathcal{S}_N) \triangleq I(K; Z^{1:N}\mathbf{F}),$$

the uniformity of the key

$$\mathbf{U}(\mathcal{S}_N) \triangleq \log\lceil 2^{NR} \rceil - H(K).$$

Definition 3.3.2. *A secret-key rate R is achievable if there exists a sequence of $(2^{NR}, N, R_p)$ key-generation strategies $\{\mathcal{S}_N\}_{N\geqslant 1}$ such that*

$$\begin{aligned} \lim_{N\to\infty} \mathbf{P}_e(\mathcal{S}_N) &= 0, \quad \textit{(reliability)} \\ \lim_{N\to\infty} \mathbf{L}(\mathcal{S}_N) &= 0, \quad \textit{(strong secrecy)} \\ \lim_{N\to\infty} \mathbf{U}(\mathcal{S}_N) &= 0. \quad \textit{(uniformity)} \end{aligned}$$

Moreover, the supremum of achievable rates is called the secret-key capacity and is denoted $C_{WSK}(R_p)$. In the special case where Eve has no access to the component Z of the source, the secret-key capacity is denoted $C_{SK}(R_p)$. One also says that perfect secrecy is achieved if $\mathbf{L}(\mathcal{S}_N) = 0$.

In this chapter, we develop low-complexity secret-key capacity-achieving schemes based on polar codes for special cases of the general model presented in Definition 3.3.1. In the following, the blocklength, N, used by the legitimate users is a power of 2. Moreover, we say that the legitimate users share a secret seed, if they share a secret sequence of $d_N \in \mathbb{N}$ uniformly distributed bits, and we define the seed rate as d_N/N. To avoid modifying the secret-key capacity with the introduction of a seed, we only consider schemes with vanishing seed rate.

3.4 Polar coding primitives for secret-key generation

We describe two polar coding primitives that capture the essence of our secret-key generation schemes. The first one implements source coding with side information, while the second one implements privacy amplification. Unlike the previous chapter, polar coding constructions will allow us to perform these two steps simultaneously instead of successively, as we will see in the next sections.

Later, in Chapter 5, a connection between these polar coding primitives and random binning will also be developed and exploited in our polar coding schemes for the wiretap channel.

3.4.1 Source polarization

Consider a discrete memoryless source defined by the distribution $(\mathcal{X}\mathcal{Y}, p_{XY})$, where N is a power of two and $|\mathcal{X}|= 2$. Polar source coding [87] can be seen as the decomposition of $X^{1:N}$, into N bit sources. Specifically, for $i \in [\![1, N]\!]$, the i-th source is defined by U^i, where $U^{1:N} \triangleq X^{1:N} G_N$. As N goes to infinity, any of these N resulting source has either an entropy essentially equal to one or essentially equal to zero, that is, is either totally random or deterministic. More formally, for $\delta_N \triangleq 2^{-N^{\beta}}$, $\beta < 1/2$, we define the set of "high entropy bits" given $Y^{1:N}$ as

$$\mathcal{H}_{X|Y} \triangleq \{i \in [\![1, N]\!] : H(U^i | U^{1:i-1} Y^{1:N}) > \delta_N\}.$$

The following theorem shows how asymptotically optimal lossless compression of $X^{1:N}$ can be performed.

Theorem 3.4.1 ([87])**.** *$X^{1:N}$ can be reconstructed with error probability in $O(N\delta_N)$ from $U^{1:N}[\mathcal{H}_{X|Y}]$ and $Y^{1:N}$ by successive cancellation decoding, whose complexity is in $O(N \log N)$. Moreover, the encoding rate is optimal because*

$$\lim_{N \to \infty} \frac{|\mathcal{H}_{X|Y}|}{N} = H(X|Y).$$

3.4.2 Privacy amplification

Consider a discrete memoryless source defined by the distribution $(\mathcal{X}\mathcal{Z}, p_{XZ})$, where N is a power of two and $|\mathcal{X}|= 2$. Similar to source polarization, we define $U^{1:N} \triangleq X^{1:N}G_N$. We also define, for $\delta_N \triangleq 2^{-N^\beta}$, $\beta < 1/2$, a set of "very high entropy bits" given $Z^{1:N}$ as

$$\mathcal{V}_{X|Z} \triangleq \{i \in [\![1, N]\!] : H(U^i|U^{1:i-1}Z^{1:N}) > 1 - \delta_N\}.$$

The rate of $|\mathcal{V}_{X|Z}|$ is given in the following lemma.

Lemma 3.4.1. *The set $\mathcal{V}_{X|Z}$ is such that*

$$\lim_{N\to+\infty} \frac{|\mathcal{V}_{X|Z}|}{N} = H(X|Z).$$

Proof. See Appendix 3.A.3. □

Note that $\lim_{N\to+\infty}|\mathcal{H}_{X|Y}|/N = H(X|Y)$ follows from [87], but Lemma 3.4.1 requires a different proof based on Lemma 3.1.16 in the appendix.

We claim that the bits $U^{1:N}[\mathcal{V}_{X|Z}]$ are almost uniformly distributed and independent from $Z^{1:N}$. We make this statement clear in the following proposition.

Proposition 3.4.1. *Define $K \triangleq U^{1:N}[\mathcal{V}_{X|Z}]$. We have*

$$I\left(K; Z^{1:N}\right) \leqslant N\delta_N \text{ (independence with } Z^{1:N}\text{)},$$
$$|K|-H(K) \leqslant N\delta_N \text{ (uniformity).}$$

Moreover, the rate of $K = U^{1:N}[\mathcal{V}_{X|Z}]$ is optimal because

$$\lim_{N\to+\infty} \frac{|\mathcal{V}_{X|Z}|}{N} = H(X|Z).$$

Proof. We write

$$
\begin{aligned}
& I\left(K;Z^{1:N}\right) + |K| - H(K) \\
& = |U^{1:N}[\mathcal{V}_{X|Z}]| - H\left(U^{1:N}[\mathcal{V}_{X|Z}]|Z^{1:N}\right) \\
& = |\mathcal{V}_{X|Z}| - H\left(U^{1:N}[\mathcal{V}_{X|Z}]|Z^{1:N}\right) \\
& \overset{(a)}{\leqslant} |\mathcal{V}_{X|Z}| - \sum_{i \in \mathcal{V}_{X|Z}} H(U^i|U^{1:i-1}Z^{1:N}) \\
& \overset{(b)}{\leqslant} |\mathcal{V}_{X|Z}|\delta_N \\
& \leqslant N\delta_N,
\end{aligned}
$$

where (a) holds because conditioning reduces entropy, (b) holds by definition of $\mathcal{V}_{X|Z}$. □

Proposition 3.4.1 thus provides a polar coding counterpart to privacy amplification performed with extractors in Section 2.3.3.2.

3.5 Model 1: Secret-key generation with rate-unlimited public communication

The precise model and known results are described in Section 3.5.1. Our proposed polar coding scheme is given in Section 3.5.2 and analyzed in Section 3.5.3.

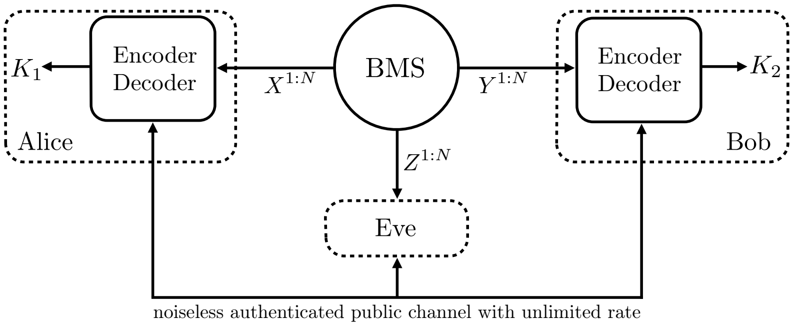


Figure 16. Model 1: Secret-key generation for the BMS model with rate-unlimited public communication.

3.5.1 Secret-key generation model

As illustrated in Figure 16, Model 1 consists of $m = 2$ legitimate terminals. We use $\mathcal{X}$ instead of $\mathcal{X}_1$ and $\mathcal{Y}$ instead of $\mathcal{X}_2$ for convenience. We assume that $\mathcal{X} = \{0, 1\}$ and that the public channel has an unlimited communication rate $R_p = +\infty$. We call this setup the BMS model with rate-unlimited public communication. The following results are known for this model.

Theorem 3.5.2 ([2,3])**.** *Consider a BMS $(\mathcal{XYZ}, p_{XYZ})$. If $X \to Y \to Z$, then the secret-key capacity $C_{\text{WSK}}(+\infty)$ is*

$$C_{\text{WSK}}(+\infty) = I(X;Y) - I(X;Z).$$

Moreover, the secret-key capacity can be achieved by one-way communication.

When the eavesdropper has no access to the source component Z, one obtains the following expression for the secret-key capacity.

Corollary 3.5.1 ([2, 3])**.** *Consider a BMS $(\mathcal{XY}, p_{XY})$. The secret-key capacity $C_{\text{SK}}(+\infty)$ is*

$$C_{\text{SK}}(+\infty) = I(X;Y).$$

Moreover, the secret-key capacity can be achieved using only one-way communication.

Such a model is motivated by the sources of randomness that can be generated from wireless communication channel gains [5–7]. In such settings, the wireless channel gains $c_{A\to B}$ characterizing the channel from Alice to Bob, $c_{B\to A}$ characterizing the channel from Bob to Alice, and the pair ($c_{A\to E}$, $c_{B\to E}$), characterizing the channels to Eve, may be used as the variables X, Y, and Z, respectively, of Model 1.

3.5.2 Polar coding scheme

In the following, we assume that $I(X;Y) - I(X;Z) > 0$ but we do not assume that $X \to Y \to Z$ forms a Markov chain; we discuss at the end of the section how the coding scheme simplifies when $X \to Y \to Z$ holds.

Let $n \in \mathbb{N}$ and $N \triangleq 2^n$. Set $U^{1:N} \triangleq X^{1:N} G_N$. For $\delta_N \triangleq 2^{-N^\beta}$, where $\beta \in]0, 1/2[$, define the following sets

$$\mathcal{V}_{X|Z} \triangleq \left\{ i \in [\![1, N]\!] : H\left(U^i | U^{1:i-1} Z^{1:N}\right) \geqslant 1 - \delta_N \right\},$$
$$\mathcal{H}_{X|Y} \triangleq \left\{ i \in [\![1, N]\!] : H\left(U^i | U^{1:i-1} Y^{1:N}\right) \geqslant \delta_N \right\}.$$

The exact encoding and decoding algorithms are given in Algorithm 1 and Algorithm 2, respectively, and we provide here a high-level discussion of their operation. The set $\mathcal{H}_{X|Y}$ is the set of indices containing "high-entropy bits" such that $U^{1:N}[\mathcal{H}_{X|Y}]$ allows Bob to near losslessly reconstruct $U^{1:N}$ from $Y^{1:N}$ [87]. In our coding scheme, Alice therefore publicly transmits $U^{1:N}[\mathcal{H}_{X|Y}]$ to allow Bob to reconstruct $U^{1:N}$. By construction, the set $\mathcal{V}_{X|Z}$ is the set of indices containing "very-high entropy bits" such that $U^{1:N}[\mathcal{V}_{X|Z}]$ is almost uniform and independent of the eavesdropper's observations $Z^{1:N}$. Consequently, the secret-key should be chosen as a subvector of $U^{1:N}[\mathcal{V}_{X|Z}]$; specifically, since $U^{1:N}[\mathcal{H}_{X|Y}]$ is publicly transmitted, it is natural to use $U^{1:N}[\mathcal{V}_{X|Z} \backslash \mathcal{H}_{X|Y}]$ as the secret key. Unfortunately, $\mathcal{H}_{X|Y} \not\subset \mathcal{V}_{X|Z}$ in general, so that the public communication of $U^{1:N}[\mathcal{H}_{X|Y}]$ leaks some information about $U^{1:N}[\mathcal{V}_{X|Z} \backslash \mathcal{H}_{X|Y}]$. To circumvent this issue, our protocol uses a secret seed to protect the transmission of the bits in positions $\mathcal{H}_{X|Y} \backslash \mathcal{V}_{X|Z}$ with a one-time-pad. In addition, our scheme operates over k blocks of size N to handle non-degraded sources and to make the seed rate negligible. In every Block $i \in [\![1, k]\!]$ Alice generates a secret key K_i together with a seed $\widetilde{K}_i$ used in the next block. Overall, Alice obtains a vector of secret keys $K_{1:k} \triangleq [K_1, K_2, \ldots, K_k]$ while Bob obtains a vector of estimates $\widehat{K}_{1:k} \triangleq [\widehat{K}_1, \widehat{K}_2, \ldots, \widehat{K}_k]$.

Remark 3.5.1. *For convenience, Algorithm 1 does not distinguish the last block from the others; however, there is no need to create a seed in Block k, so that one may actually use $U_k^{1:N}[\mathcal{V}_{X|Z} \backslash \mathcal{H}_{X|Y}]$ as the key K_k and slightly increase the key rate. For a large number of blocks k, this distinction has negligible impact on the achievable rates.*

Algorithm 1: Alice's encoding algorithm for Model 1

```
Require: $\widetilde{K}_0$, a secret key of size $|\mathcal{H}_{X|Y} \backslash \mathcal{V}_{X|Z}|$ shared by Alice and Bob
         beforehand; for every Block $i \in [\![1, k]\!]$, the observations $X_i^{1:N}$ from
         the source; $\mathcal{A}_{XYZ}$ a fixed subset of $\mathcal{V}_{X|Z} \backslash \mathcal{H}_{X|Y}$ with size
         $|\mathcal{H}_{X|Y} \backslash \mathcal{V}_{X|Z}|$.
for Block $i = 1$ to $k$ do
    $U_i^{1:N} \leftarrow X_i^{1:N} G_N$
    $\widetilde{K}_i \leftarrow U_i^{1:N}[\mathcal{A}_{XYZ}]$ {Fraction of the key used as a seed for the next block}
    $K_i \leftarrow U_i^{1:N}[(\mathcal{V}_{X|Z} \backslash \mathcal{H}_{X|Y}) \backslash \mathcal{A}_{XYZ}]$
    $F_i \leftarrow U_i^{1:N}[\mathcal{V}_{X|Z} \cap \mathcal{H}_{X|Y}]$
    $F_i' \leftarrow U_i^{1:N}[\mathcal{H}_{X|Y} \backslash \mathcal{V}_{X|Z}]$
    Transmit $M_i \leftarrow [F_i, F_i' \oplus \widetilde{K}_{i-1}]$ publicly to Bob
end
return : $K_{1:k} \leftarrow [K_1, K_2, \dots, K_k]$
```

Algorithm 2: Bob's decoding algorithm for Model 1

```
Require: $\widetilde{K}_0$, a secret key of size $|\mathcal{H}_{X|Y} \backslash \mathcal{V}_{X|Z}|$ shared by Alice and Bob
         beforehand; the set $\mathcal{A}_{XYZ}$ defined in Algorithm 1; for every Block
         $i \in [\![1, k]\!]$, the observations $Y_i^{1:N}$ from the source and the message $M_i$
         transmitted by Alice.
for Block $i = 1$ to $k$ do
    Form $U_i^{1:N}[\mathcal{H}_{X|Y}]$ from $M_i$ and $\widetilde{K}_{i-1}$
    Create an estimate $\widehat{U}_i^{1:N}$ of $U_i^{1:N}$ with the successive cancellation decoder
    of [87]
    $\widehat{K}_i \leftarrow \widehat{U}_i^{1:N}[(\mathcal{V}_{X|Z} \backslash \mathcal{H}_{X|Y}) \backslash \mathcal{A}_{XYZ}]$
    $\widetilde{K}_i \leftarrow \widehat{U}_i^{1:N}[\mathcal{A}_{XYZ}]$
end
return : $\widehat{K}_{1:k} \leftarrow [\widehat{K}_1, \widehat{K}_2, \dots, \widehat{K}_k]$
```

Remark 3.5.2. *The need for a seed is not an artifact of our proof, but a fundamental requirement of our single polarization approach to generate secret keys and public messages. In fact, a memoryless source cannot be near losslessly compressed at a rate close to the entropy and simultaneously ensure that the encoded messages are nearly uniformly distributed in variational distance [93, Section V]. In the context of secret-key generation with polar codes, this translates into the condition* $\mathcal{H}_{X|Y} \not\subset \mathcal{V}_{X|Z}$ *and in the impossibility of simultaneously ensuring strong secrecy and reliability. Our solution follows ideas from Section 4.4 in Chapter 4, showing that the impossibility may be circumvented if the encoder and the decoder share a small seed beforehand; without seed, only weak secrecy would be ensured.*

As shown in Section 3.5.3, a careful analysis of the algorithms leads to the following result.

Theorem 3.5.3. *Consider a BMS* $(\mathcal{X}\mathcal{Y}\mathcal{Z}, p_{XYZ})$. *Assume that Alice and Bob share a secret seed. The secret-key rate* $I(X;Y) - I(X;Z)$ *is achieved by the polar coding scheme of Algorithm 1 and Algorithm 2, which involves a chaining of* k *blocks of size* N, *and whose computational complexity is* $O(kN \log N)$. *Moreover, the seed rate can be chosen in* $o\left(2^{-N^{\alpha}}\right)$, $\alpha < 1/2$.

Proof. See Section 3.5.3. □

Corollary 3.5.2. *When* $X \to Y \to Z$, *the secret-key capacity of Theorem 3.5.2 is achieved by the polar coding scheme of Algorithm 1 and Algorithm 2. Moreover, one does not need to encode over several blocks, i.e., one can choose* $k = 1$, *and the seed rate is* $o(N)$. *However, encoding over several blocks for this case allows one to reduce the seed rate from* $o(N)$ *to* $o(2^{-N^{\alpha}})$, $\alpha < 1/2$.

Proof. See Appendix 3.A.1. □

Note that, in the special case of a symmetric degraded BMS,[2] Corollary 3.5.2 may

[2] That is, when X, Y, and Z are connected by symmetric channels.

be indirectly obtained from wiretap codes and [39], following the approach of [2], [23, Section 4.2.1]. However, this indirect proof might not translate into practical implementations because it requires much more public channel communication.

Although the seed rate in Theorem 3.5.3 or Corollary 3.5.2 may be made arbitrarily small, it is valuable to identify examples for which no seed is required. We provide two such examples in Proposition 3.5.2, which corresponds to the privacy amplification setting of [24], and in Proposition 3.5.3, which corresponds to a case when the source has uniform marginals and the eavesdropper has no access to correlated observations of the source.

Proposition 3.5.2. *Consider a BMS $(\mathcal{X}\mathcal{Y}\mathcal{Z}, p_{XYZ})$. Assume that Alice and Bob have the same observations, i.e., $X = Y$; then the secret-key capacity $C_{WSK} = H(X|Z)$ is achievable with a polar coding scheme, whose computational complexity is $O(N \log N)$.*

Proof. See Section 3.4.2. □

Proposition 3.5.3. *Consider a BMS $(\mathcal{X}\mathcal{Y}, p_{XY})$ with $X \sim \mathcal{B}(1/2)$. The secret-key capacity $C_{SK}(+\infty)$ given in Corollary 3.5.1 is achievable with perfect secrecy with a polar coding scheme, whose computational complexity is $O(N \log N)$.*

Proof. See Appendix 3.A.2. □

Note that the model studied in Proposition 3.5.3 includes [34, Model 1] as a special case, and does not require the construction of a standard array, whose size grows exponentially with the blocklength.

3.5.3 Analysis of polar coding scheme: proof of Theorem 3.5.3

A functional dependence graph of the block encoding scheme of Section 3.5.2 is depicted in Figure 17 to help the reader identify the dependencies among the variables introduced by the block-coding scheme.

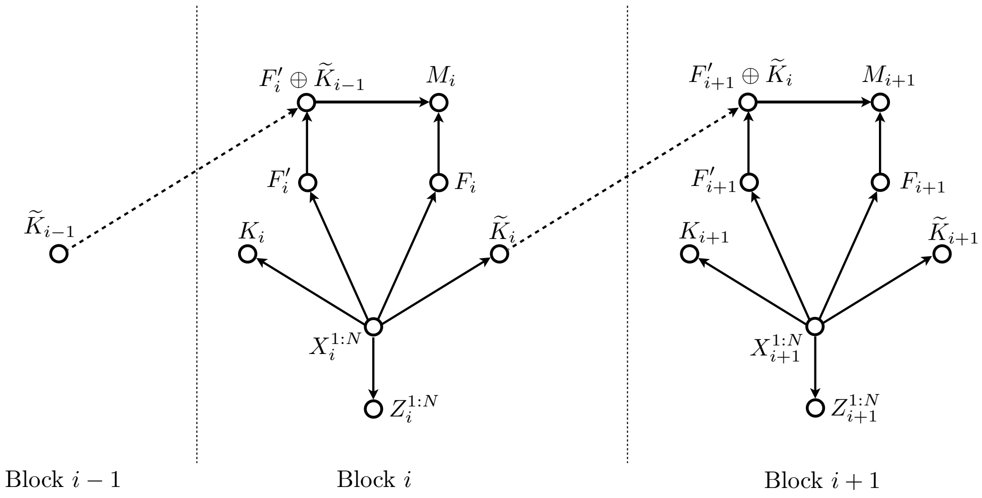


Figure 17. Functional dependence graph of the proposed block encoding scheme

3.5.3.1 Existence of $\mathcal{A}_{XYZ}$

Observe that $|\mathcal{V}_{X|Z}\backslash\mathcal{H}_{X|Y}|-|\mathcal{H}_{X|Y}\backslash\mathcal{V}_{X|Z}|= |\mathcal{V}_{X|Z}|-|\mathcal{H}_{X|Y}|$. Hence, by Lemma 3.4.1 and [87], we have

$$\lim_{N\to\infty}(|\mathcal{V}_{X|Z}\backslash\mathcal{H}_{X|Y}|-|\mathcal{H}_{X|Y}\backslash\mathcal{V}_{X|Z}|)/N = H(X|Z) - H(X|Y).$$

Since $I(X;Y) - I(X;Z) > 0$ by assumption, we conclude that

$$|\mathcal{V}_{X|Z}\backslash\mathcal{H}_{X|Y}|-|\mathcal{H}_{X|Y}\backslash\mathcal{V}_{X|Z}|> 0$$

for N large enough and $\mathcal{A}_{XYZ}$ exists.

3.5.3.2 Asymptotic key rate

The length of the overall key generated is

$$\begin{aligned}|K_{1:k}| &= \sum_{i=1}^{k}|K_i| \\ &= k|(\mathcal{V}_{X|Z}\backslash\mathcal{H}_{X|Y})\backslash\mathcal{A}_{XYZ}| \\ &= k(|\mathcal{V}_{X|Z}\backslash\mathcal{H}_{X|Y}|-|\mathcal{H}_{X|Y}\backslash\mathcal{V}_{X|Z}|) \\ &= k(|\mathcal{V}_{X|Z}|-|\mathcal{H}_{X|Y}|).\end{aligned}$$

Hence, by Lemma 3.4.1 and [87], the asymptotic key rate is

$$\lim_{N\to\infty}\frac{|K_{1:k}|}{kN} \geqslant I(X;Y) - I(X;Z).$$

3.5.3.3 Reliability

Let $i \in [\![2, k]\!]$. Note that F_i' is correctly received only when Bob possesses a correct estimate of the seed $\widetilde{K}_{i-1}$, i.e., when $U_{i-1}^{1:N}$ is correctly reconstructed. We note $\widehat{F}_i'$ the estimate of F_i' formed by Bob from $\widehat{U}_{i-1}^{1:N}$ and define the event $\mathcal{E}_{F_i'} \triangleq \{F_i' \neq \widehat{F}_i'\}$. Then,

$$\begin{aligned}
\mathbb{P}[K_i \neq \widehat{K}_i] &\leqslant \mathbb{P}[U_i^{1:N} \neq \widehat{U}_i^{1:N}] \\
&= \mathbb{P}[U_i^{1:N} \neq \widehat{U}_i^{1:N} | \mathcal{E}_{F_i'}^c]\mathbb{P}[\mathcal{E}_{F_i'}^c] + \mathbb{P}[U_i^{1:N} \neq \widehat{U}_i^{1:N} | \mathcal{E}_{F_i'}]\mathbb{P}[\mathcal{E}_{F_i'}] \\
&\leqslant \mathbb{P}[U_i^{1:N} \neq \widehat{U}_i^{1:N} | \mathcal{E}_{F_i'}^c] + \mathbb{P}[\mathcal{E}_{F_i'}] \\
&\leqslant \mathbb{P}[U_i^{1:N} \neq \widehat{U}_i^{1:N} | \mathcal{E}_{F_i'}^c] + \mathbb{P}[U_{i-1}^{1:N} \neq \widehat{U}_{i-1}^{1:N}] \\
&\overset{(a)}{\leqslant} N\delta_N + \mathbb{P}[U_{i-1}^{1:N} \neq \widehat{U}_{i-1}^{1:N}] \\
&\overset{(b)}{\leqslant} (i-1)N\delta_N + \mathbb{P}[U_1^{1:N} \neq \widehat{U}_1^{1:N}] \\
&\overset{(c)}{\leqslant} iN\delta_N,
\end{aligned}$$

where (a) follows because Bob can reconstruct $U_i^{1:N}$ from $(F_i, F_i') = U_i^{1:N}[\mathcal{H}_{X|Y}]$ and $Y_i^{1:N}$ with error probability less than $N\delta_N$ [87], (b) holds by induction, (c) holds by [87] and because $\widetilde{K}_0$ is known to Bob. Using the union bound,

$$\begin{aligned}
\mathbf{P}_e(\mathcal{S}_N) &= \mathbb{P}[K_{1:k} \neq \widehat{K}_{1:k}] \\
&\leqslant \mathbb{P}[\bigcup_{i=1}^{k} (K_i \neq \widehat{K}_i)] \\
&\leqslant \sum_{i=1}^{k} \mathbb{P}[K_i \neq \widehat{K}_i] \\
&\leqslant \sum_{i=1}^{k} iN\delta_N \\
&= \frac{k(k+1)}{2} N\delta_N. \qquad (45)
\end{aligned}$$

3.5.3.4 Key uniformity

We first prove the uniformity of the key in each block i using the following lemma.

Lemma 3.5.2. *In every block $i \in [\![1,k]\!]$, the vector $[K_i, \widetilde{K}_i]$ is nearly uniform, in the sense that*

$$|K_i|+|\widetilde{K}_i|-H(K_i\widetilde{K}_i) \leqslant N\delta_N.$$

In particular, $|\widetilde{K}_i|-H(\widetilde{K}_i) \leqslant N\delta_N$ and $|K_i|-H(K_i) \leqslant N\delta_N$.

Proof.

$$\begin{aligned}
|K_i|+|\widetilde{K}_i|-H(K_i\widetilde{K}_i) &= |K_i|+|\widetilde{K}_i|-H(U_i^{1:N}[\mathcal{V}_{X|Z}\backslash\mathcal{H}_{X|Y}]) \\
&\overset{(a)}{\leqslant} |K_i|+|\widetilde{K}_i|-\sum_{j\in\mathcal{V}_{X|Z}\backslash\mathcal{H}_{X|Y}} H(U_i^j|U_i^{1:j-1}) \\
&\overset{(b)}{\leqslant} |K_i|+|\widetilde{K}_i|-\sum_{j\in\mathcal{V}_{X|Z}\backslash\mathcal{H}_{X|Y}} (1-\delta_N) \\
&= (|K_i|+|\widetilde{K}_i|)\delta_N \\
&\leqslant N\delta_N,
\end{aligned}$$

where (a) holds because conditioning reduces entropy, (b) holds by definition of $\mathcal{V}_{X|Z}$ and because conditioning reduces entropy. Finally, note that since $|K_i|-H(K_i|\widetilde{K}_i) > 0$, we have

$$\begin{aligned}
|\widetilde{K}_i|-H(\widetilde{K}_i) &\leqslant |\widetilde{K}_i|-H(\widetilde{K}_i) + |K_i|-H(K_i|\widetilde{K}_i) \\
&= |K_i|+|\widetilde{K}_i|-H(K_i\widetilde{K}_i).
\end{aligned}$$

□

It remains to show that the overall key $K_{1:k}$ is uniform, as well. Specifically, we have

$$\begin{aligned}
H(K_{1:k}) &= \sum_{i=1}^{k} H(K_i|K_{1:i-1}) \\
&\overset{(a)}{=} \sum_{i=1}^{k} H(K_i) \\
&\overset{(b)}{\geqslant} \sum_{i=1}^{k} (|K_i| - N\delta_N) \\
&= |K_{1:k}| - kN\delta_N,
\end{aligned}$$

where (a) holds because $X_i^{1:N}$ is independent of of $X_{1:i-1}^{1:N}$ for any $i \in [\![1,k]\!]$, and (b) holds by Lemma 3.5.2. Hence,

$$\mathbf{U}(\mathcal{S}_N) = |K_{1:k}| - H(K_{1:k}) \leqslant kN\delta_N. \tag{46}$$

3.5.3.5 Strong secrecy

We first show that secrecy holds for each block using the following lemma .

Lemma 3.5.3. *For each Block $i \in [\![1,k]\!]$, $[K_i, \widetilde{K}_i]$ is a secret key. Specifically,*

$$I\left(K_i\widetilde{K}_i; M_i Z_i^{1:N}\right) \leqslant 2N\delta_N.$$

Proof. We have

$$
\begin{aligned}
& I(K_i\widetilde{K}_i; F_i Z_i^{1:N}) \\
& = H(K_i\widetilde{K}_i) - H(K_i\widetilde{K}_i|F_i Z_i^{1:N}) \\
& \leqslant |K_i| + |\widetilde{K}_i| - H(K_i\widetilde{K}_i F_i|Z_i^{1:N}) + H(F_i|Z_i^{1:N}) \\
& \leqslant |K_i| + |\widetilde{K}_i| + |F_i| - H(K_i\widetilde{K}_i F_i|Z_i^{1:N}) \\
& \overset{(a)}{=} |\mathcal{V}_{X|Z}\backslash\mathcal{H}_{X|Y}| + |\mathcal{V}_{X|Z}\cap\mathcal{H}_{X|Y}| - H(U_i^{1:N}[(\mathcal{V}_{X|Z}\backslash\mathcal{H}_{X|Y})\cup(\mathcal{V}_{X|Z}\cap\mathcal{H}_{X|Y})]|Z_i^{1:N}) \\
& = |\mathcal{V}_{X|Z}| - H(U_i^{1:N}[\mathcal{V}_{X|Z}]|Z_i^{1:N}) \\
& \overset{(b)}{\leqslant} |\mathcal{V}_{X|Z}| - \sum_{j\in\mathcal{V}_{X|Z}} H(U_i^j|U_i^{1:j-1}Z_i^{1:N}) \\
& \overset{(c)}{\leqslant} |\mathcal{V}_{X|Z}| - \sum_{j\in\mathcal{V}_{X|Z}} (1-\delta_N) \\
& = |\mathcal{V}_{X|Z}|\delta_N \\
& \leqslant N\delta_N,
\end{aligned}
\tag{47}
$$

where (a) holds by definition of K_i, $\widetilde{K}_i$, and F_i, (b) holds because conditioning reduces entropy, (c) holds by definition of $\mathcal{V}_{X|Z}$. Therefore, we obtain

$$
\begin{aligned}
& I(K_i\widetilde{K}_i; M_i Z_i^{1:N}) \\
& \overset{(d)}{=} I(K_i\widetilde{K}_i; F_i(F_i'\oplus\widetilde{K}_{i-1})Z_i^{1:N}) \\
& = I(K_i\widetilde{K}_i; F_i Z_i^{1:N}) + I(K_i\widetilde{K}_i; F_i'\oplus\widetilde{K}_{i-1}|F_i Z_i^{1:N}) \\
& \overset{(e)}{\leqslant} N\delta_N + I(K_i\widetilde{K}_i F_i Z_i^{1:N} F_i'; F_i'\oplus\widetilde{K}_{i-1}) \\
& = N\delta_N + H(F_i'\oplus\widetilde{K}_{i-1}) - H(F_i'\oplus\widetilde{K}_{i-1}|K_i\widetilde{K}_i F_i Z_i^{1:N} F_i') \\
& = N\delta_N + H(F_i'\oplus\widetilde{K}_{i-1}) - H(\widetilde{K}_{i-1}|K_i\widetilde{K}_i F_i Z_i^{1:N} F_i') \\
& = N\delta_N + H(F_i'\oplus\widetilde{K}_{i-1}) - H(\widetilde{K}_{i-1}) \\
& \leqslant N\delta_N + |\widetilde{K}_{i-1}| - H(\widetilde{K}_{i-1}) \\
& \overset{(f)}{\leqslant} 2N\delta_N,
\end{aligned}
$$

where (d) holds by definition of M_i, (e) holds by (47), (f) holds by Lemma 3.5.2. □

We now state two lemmas that will be used to show that secrecy holds for the global scheme.

Lemma 3.5.4. *For $i \in [\![1,k]\!]$, we have for N large enough*

$$I(K_i;\widetilde{K}_i) \leqslant \delta_N^*,$$

where

$$\delta_N^* \triangleq -3\sqrt{2N\log 2} \times N2^{-N^\beta/2}\log_2\left(3\sqrt{2N\log 2}\times 2^{-N^\beta/2}\right). \tag{48}$$

Proof. See Appendix 3.A.4 □

Lemma 3.5.5. *For $i \in [\![2,k]\!]$, define*

$$\widetilde{L}_e^{1:i} \triangleq I\left(K_{1:i}\widetilde{K}_i; M_{1:i}Z_{1:i}^{1:N}\right).$$

We have

$$\widetilde{L}_e^{1:i} - \widetilde{L}_e^{1:i-1} \leqslant I\left(K_i\widetilde{K}_i; M_iZ_i^{1:N}\right) + I\left(K_{i-1};\widetilde{K}_{i-1}\right).$$

Proof. See Appendix 3.A.5. □

We thus obtain

$$\begin{aligned}
\mathbf{L}(\mathcal{S}_N) &= I(K_{1:k}; M_{1:k}Z_{1:k}^{1:N})\\
&\leqslant \widetilde{L}_e^{1:k}\\
&= \sum_{i=2}^{k}(\widetilde{L}_e^{1:i} - \widetilde{L}_e^{1:i-1}) + \widetilde{L}_e^{1}\\
&\overset{(a)}{\leqslant} \sum_{i=2}^{k}\left(I\left(K_i\widetilde{K}_i; M_iZ_i^{1:N}\right) + I\left(K_{i-1};\widetilde{K}_{i-1}\right)\right) + \widetilde{L}_e^{1}\\
&\leqslant \sum_{i=1}^{k} I\left(K_i\widetilde{K}_i; M_iZ_i^{1:N}\right) + \sum_{i=2}^{k} I\left(K_{i-1};\widetilde{K}_{i-1}\right)\\
&\overset{(b)}{\leqslant} 2kN\delta_N + (k-1)\delta_N^*,
\end{aligned} \tag{49}$$

where (a) follows by Lemma 3.5.5, (b) follows by Lemma 3.5.4 and Lemma 3.5.3.

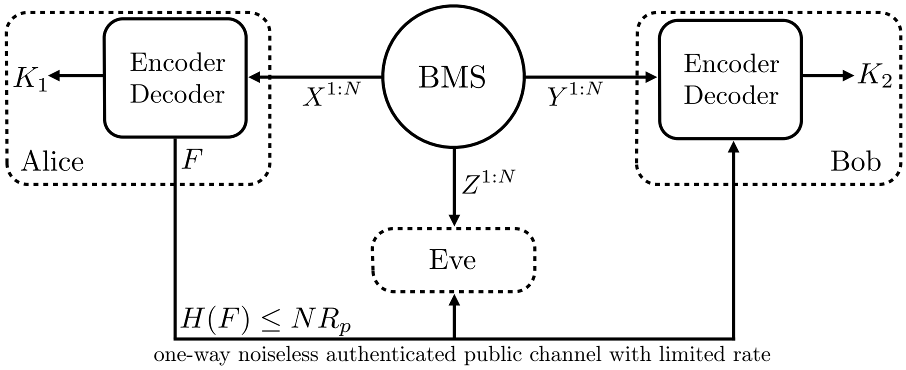


Figure 18. Model 2: Secret-key generation for the BMS model with one-way rate-limited public communication.

3.5.3.6 Seed rate

The seed rate required to initialize the coding scheme is negligible since

$$\lim_{k\to\infty}\lim_{N\to\infty}\frac{|\mathcal{H}_{X|Y}\backslash\mathcal{V}_{X|Z}|}{kN} \leqslant \lim_{k\to\infty}\frac{H(X|Y)}{k} = 0.$$

Note that the seed rate may be chosen to decrease exponentially fast to zero with N since we may choose $k = 2^{N^\alpha}$, $\alpha < \beta$ and still have $\lim_{N\to\infty}\mathbf{P}_e(\mathcal{S}_N) = 0$ by (45), $\lim_{N\to\infty}\mathbf{U}_e(\mathcal{S}_N) = 0$ by (46), and $\lim_{N\to\infty}\mathbf{L}_e(\mathcal{S}_N) = 0$ by (49) and (48).

3.6 Model 2: Secret-key generation with rate-limited public communication

We now move to the second key generation model, which differs from Model 1 by restricting the public communication to be rate-limited and one way from Alice to Bob. The organization follows that of Section 3.6.

3.6.1 Secret-key generation model

As illustrated in Figure 18, we set again $m = 2$ and we use $\mathcal{X}$ instead of $\mathcal{X}_1$, $\mathcal{Y}$ instead of $\mathcal{X}_2$ for convenience. We assume that $\mathcal{X} = \{0, 1\}$ and that Alice and Bob are constrained to only use one-way communication over an authenticated noiseless public channel with limited rate $R_p \in \mathbb{R}$. We call this setup the BMS model with rate-limited public communication. The following results are known for the model.

Theorem 3.6.4. *Let $(\mathcal{X}\mathcal{Y}\mathcal{Z}, p_{XYZ})$ be a BMS and $R_p \in \mathbb{R}_+$ be the public communication rate. If $X \rightarrow Y \rightarrow Z$, then the one-way rate-limited secret-key capacity is*[3]

$$C_{\text{WSK}}(R_p) = \max_{U} \left(I(Y;U) - I(Z;U)\right)$$

subject to

$$\begin{aligned} &R_p = I(U;X) - I(U;Y), \\ &U \rightarrow X \rightarrow Y \rightarrow Z, \\ &|\mathcal{U}| \leqslant |\mathcal{X}|. \end{aligned}$$

Closed form expressions of the secret-key capacity are only known for specific sources. See the following example.

Example 3.6.1. *Assume $\mathcal{X} = \mathcal{Y} = \mathcal{Z} = \{0,1\}$ and $X \sim \mathcal{B}(1/2)$. Set $Y \triangleq X \oplus B_1$ and $Z \triangleq Y \oplus B_2$, with $B_1 \sim \mathcal{B}(p)$, $B_2 \sim \mathcal{B}(q)$, where $\oplus$ denotes the modulo-2 addition. Then, by Example 2.5.2 in Chapter 2, the secret-key capacity is*

$$C_{WSK}(R_p) \triangleq \begin{cases} H_b(p \star \beta_0 \star q) - H_b(p \star \beta_0), & \text{if } R_p \leqslant H(X|Y), \\ H_b(p \star q) - H_b(p), & \text{if } R_p \geqslant H(X|Y), \end{cases}$$

where β_0 must satisfy[4]

$$H_b(p \star \beta_0) - H_b(\beta_0) = R_p, \tag{50}$$

$H_b(\cdot)$ is the binary entropy function, and the associative and commutative operation $\star$ is defined as $p \star \beta_0 = (1-\beta_0)p + \beta_0(1-p)$.

When the eavesdropper has no access to the source component Z, one obtains the following expression for the secret-key capacity.

[3] See Proposition 2.5.3 and Remark 2.5.3 in Chapter 2.

[4] Note that (50) has two symmetric solutions.

Corollary 3.6.3. *Let $(\mathcal{X}\mathcal{Y}\mathcal{Z}, p_{XYZ})$ be a BMS and $R_p \in \mathbb{R}_+$ be the public communication rate. The one-way rate-limited secret-key capacity is*

$$C_{\mathrm{SK}}(R_p) = \max_U I(Y;U)$$

subject to

$$R_p = I(U;X) - I(U;Y),$$

$$U \rightarrow X \rightarrow Y,$$

$$|\mathcal{U}| \leqslant |\mathcal{X}|.$$

The practical justification for Model 2 is similar to that for Model 1; however, Model 2 allows us to account for rate-limited communication constraints, which is relevant in applications with stringent bandwidth constraints, such as wireless sensor networks. We will also see in Section 3.9 that such constraint may account for privacy leakage constraints in biometric systems.

The main challenge in designing a coding scheme for Model 2 is to address the problem of vector quantization with side information at the receiver. Previous polar coding results on lossy source coding with lossless reconstruction of the vector quantized version of the source are reported in [94,95]; our contribution is to extend these results when side information is available at the receiver, and to show how to apply such technique to secret-key generation with rate-limited public communication.

3.6.2 Polar coding scheme

Let $n \in \mathbb{N}$ and $N \triangleq 2^n$. Fix a joint probability distribution p_{XU} such that $I(Y;U) - I(Z;U) > 0$, but we do not assume $X \rightarrow Y \rightarrow Z$. Denote $V^{1:N} \triangleq U^{1:N} G_N$, the polar transform of a vector $U^{1:N}$ with i.i.d components according to the marginal distribution p_U. For $\delta_N \triangleq 2^{-N^\beta}$, where $\beta \in]0,1/2[$, define the following sets.

$$\begin{aligned}
\mathcal{H}_U &\triangleq \left\{i \in [\![1,N]\!] : H\left(V^i|V^{1:i-1}\right) \geqslant \delta_N\right\},\\
\mathcal{V}_{U|Z} &\triangleq \left\{i \in [\![1,N]\!] : H\left(V^i|V^{1:i-1}Z^{1:N}\right) \geqslant 1-\delta_N\right\},\\
\mathcal{V}_{U|Y} &\triangleq \left\{i \in [\![1,N]\!] : H\left(V^i|V^{1:i-1}Y^{1:N}\right) \geqslant 1-\delta_N\right\},\\
\mathcal{H}_{U|Y} &\triangleq \left\{i \in [\![1,N]\!] : H\left(V^i|V^{1:i-1}Y^{1:N}\right) \geqslant \delta_N\right\},\\
\mathcal{H}_{U|X} &\triangleq \left\{i \in [\![1,N]\!] : H\left(V^i|V^{1:i-1}X^{1:N}\right) \geqslant \delta_N\right\}.
\end{aligned}$$

The encoding and decoding algorithms are given in Algorithm 3 and Algorithm 4. The high-level principles are similar to that of Algorithm 1 and Algorithm 2, and we only highlight here the differences. Instead of directly operating on the source symbols, Alice first constructs a vector quantized version $\widetilde{V}^{1:N}$ of $X^{1:N}$, whose distribution is close to that of $V^{1:N}$. This statement is made more precise in Lemma 3.6.6, but a crucial part of the proof is to introduce a stochastic encoder, as in successive cancellation encoding for lossy source coding [94, 95]. The randomness R_1 used in the encoder is publicly transmitted to Bob and reused over several blocks so that its rate vanishes to zero as the number of blocks increases; however, reusing R_1 creates additional dependencies between the variables of the different blocks, which must be carefully taken into account in the secrecy analysis. The choice of public messages and keys is then similar to those in Section 3.5.2, using $\widetilde{V}^{1:N}$ instead of $X^{1:N}$.

Remark 3.6.3. *One may actually use $U_k^{1:N}[\mathcal{V}_{U|Z}\backslash\mathcal{H}_{U|Y}]$ as the key K_k and slightly increase the key rate in Algorithm 3. However, one does not distinguish the last block from the others for convenience – see Remark 3.5.1.*

As shown in Section 3.6.3, the analysis of Algorithm 3 and Algotithm 4 leads to the following result.

Theorem 3.6.5. *Consider a BMS $(\mathcal{X}\mathcal{Y}\mathcal{Z}, p_{XYZ})$. Assume that Alice and Bob share a secret seed and let $R_p \in \mathbb{R}^+$ be the public communication rate. The secret-key rate*

Algorithm 3: Alice's encoding algorithm for Model 2

```
Require: $\widetilde{K}_0$, a secret key of size $|(\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X})\backslash\mathcal{V}_{U|Z}|$ shared by Alice and
         Bob beforehand; for every Block $i \in [\![1,k]\!]$, the observations $X_i^{1:N}$
         from the source; $\mathcal{A}_{UYZ}$ a subset of $\mathcal{V}_{U|Z}\backslash\mathcal{H}_{U|Y}$ with size
         $|(\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X})\backslash\mathcal{V}_{U|Z}|$; a vector $R_1$ of $|\mathcal{V}_{U|X}|$ uniformly distributed bits.
Transmit $R_1$ publicly to Bob
for Block $i = 1$ to $k$ do
    $R_i \leftarrow R_1$
    $\widetilde{V}_i^{1:N}[\mathcal{V}_{U|X}] \leftarrow R_i$
    Given $X_i^{1:N}$, successively draw the remaining bits of $\widetilde{V}_i^{1:N}$ according to
    $\widetilde{p}_{V_i^{1:N}X_i^{1:N}} \triangleq \prod_{j=1}^{N} \widetilde{p}_{V_i^j|V_i^{j-1}X^{1:N}} p_{X^{1:N}}$ with

        $\widetilde{p}_{V_i^j|V_i^{1:j-1}X^{1:N}}(v^j|\widetilde{V}_i^{1:j-1}X_i^{1:N})$
        $\triangleq \begin{cases} p_{V^j|V^{1:j-1}X^{1:N}}(v^j|\widetilde{V}_i^{1:j-1}X_i^{1:N}) & \text{if } j \in \mathcal{H}_U\backslash\mathcal{V}_{U|X} \\ p_{V^j|V^{1:j-1}}(v^j|\widetilde{V}_i^{1:j-1}) & \text{if } j \in \mathcal{H}_U^c \end{cases}.$     (51)

    $\widetilde{K}_i \leftarrow \widetilde{V}_i^{1:N}[\mathcal{A}_{UYZ}]$
    $K_i \leftarrow \widetilde{V}_i^{1:N}[(\mathcal{V}_{U|Z}\backslash\mathcal{H}_{U|Y})\backslash\mathcal{A}_{UYZ}]$
    $F_i \leftarrow \widetilde{V}_i^{1:N}[(\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X}) \cap \mathcal{V}_{U|Z}]$
    $F_i' \triangleq \widetilde{V}_i^{1:N}[(\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X})\backslash\mathcal{V}_{U|Z}]$
    Transmit $M_i \leftarrow [F_i, F_i' \oplus \widetilde{K}_{i-1}]$ publicly to Bob.
end
return : $K_{1:k} \leftarrow [K_1, K_2, \ldots, K_k]$
```

```
Algorithm 4: Bob's decoding algorithm for Model 2
Require: The secret-key $\widetilde{K}_0$ and the set $\mathcal{A}_{UYZ}$ defined in Algorithm 3; for
         every Block $i \in [\![1, k]\!]$, the observations $Y_i^{1:N}$ from the source, the
         message $M_i$. transmitted by Alice; the vector $R_1$ transmitted by
         Alice.
for Block $i = 1$ to $k$ do
    Form $\widetilde{V}_i^{1:N}[\mathcal{H}_{U|Y}]$ from $M_i$ and $\widetilde{K}_{i-1}$
    Create an estimate $\widehat{V}_i^{1:N}$ of $V_i^{1:N}$ with the successive cancellation decoder
    of [87]
    $\widehat{K}_i \leftarrow \widehat{V}_i^{1:N}[(\mathcal{V}_{U|Z} \backslash \mathcal{H}_{U|Y}) \backslash \mathcal{A}_{UYZ}]$
    $\widetilde{K}_i \leftarrow \widehat{V}_i^{1:N}[\mathcal{A}_{UYZ}]$
end
return : $\widehat{K}_{1:k} \leftarrow [\widehat{K}_1, \widehat{K}_2, \ldots, \widehat{K}_k]$
```

defined by

$$\begin{aligned} & \max_U \left(I(Y;U) - I(Z;U)\right) \\ \text{subject to} \quad & R_p = I(U;X) - I(U;Y), \\ & U \to X \to Y, \\ & |\mathcal{U}| \leqslant |\mathcal{X}|. \end{aligned}$$

is achieved by the polar coding scheme of Algorithm 3 and Algorithm 4, which involves a chaining of k blocks of size N, and whose computational complexity is $O(kN \log N)$. Moreover, the seed rate can be chosen in $o\left(2^{-N^\alpha}\right)$, $\alpha < 1/2$.

Proof. See Section 3.6.3. □

The following corollary states sufficient conditions to avoid block encoding.

Corollary 3.6.4. *If $X \to Y \to Z$, $X \sim \mathcal{B}(1/2)$, and the test-channels $p_{Y|X}$ and $p_{Z|X}$ are symmetric,[5] then the secret-key capacity of Theorem 3.6.4 is achieved by the polar coding scheme for Block 1 in Algorithm 3 with $\mathcal{A}_{UYZ} = \emptyset$, R_1 a constant sequence, and a seed rate in $o(N)$.*

[5] As in Example 3.6.1 for instance

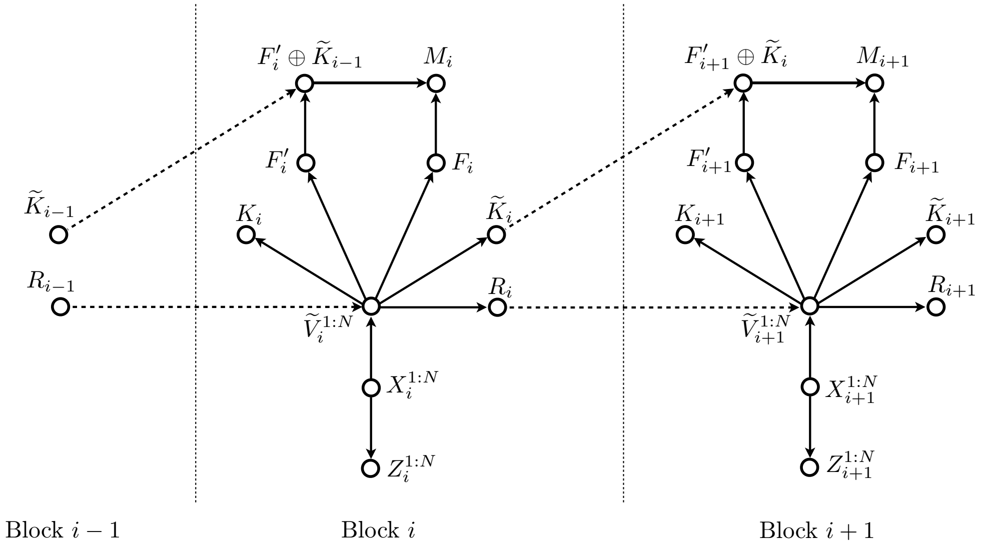


Figure 19. Functional dependence graph of the block encoding scheme

Proof. See Appendix 3.B.1. □

Finally, the following proposition provides sufficient conditions to avoid block encoding and a pre-shared seed. The proof is similar to that of Theorem 3.6.5 and Corollary 3.6.4 and is omitted.

Proposition 3.6.4. *If the eavesdropper has no access to correlated observations of the source, $X \sim \mathcal{B}(1/2)$, and the test-channel $p_{Y|X}$ is symmetric, then the secret-key capacity of Corollary 3.6.3 is achieved by the polar coding scheme for Block 1 in Algorithm 3 with $\mathcal{A}_{UYZ} = \emptyset$, $Z = \emptyset$, $F_1' = \emptyset$, $K_1 \triangleq \widetilde{V}_1^{1:N}[\mathcal{H}_{U|Y}^c]$, $F_1 \triangleq \widetilde{V}_1^{1:N}[\mathcal{H}_{U|Y} \backslash \mathcal{V}_{U|X}]$, and R_1 a constant sequence.*

3.6.3 Analysis of polar coding scheme: Proof of Theorem 3.6.5

A functional dependence graph for the coding scheme of Section 3.6.2 is depicted in Figure 33 for convenience.

3.6.3.1 Preliminary result

Lemma 3.6.6. *For every $i \in [\![1, k]\!]$, the random variable $\widetilde{V}_i^{1:N}$ resulting from Algorithm 3 has a joint distribution $\widetilde{p}_{X_i^{1:N}V_i^{1:N}} \triangleq \widetilde{p}_{V_i^{1:N}|X^{1:N}} p_{X^{1:N}}$ with $X_i^{1:N}$ such that*

$$\mathbb{D}(p_{X^{1:N}V^{1:N}}||\widetilde{p}_{X_i^{1:N}V_i^{1:N}}) \leqslant N\delta_N,$$

Hence, by Pinsker's inequality

$$\mathbb{V}(p_{X^{1:N}V^{1:N}}, \widetilde{p}_{X_i^{1:N}V_i^{1:N}}) \leqslant \sqrt{2\log 2}\sqrt{N\delta_N}.$$

Proof. See Appendix 3.B.2. □

3.6.3.2 Existence of $\mathcal{A}_{UYZ}$

Observe that

$$\begin{aligned}|\mathcal{V}_{U|Z}\backslash\mathcal{H}_{U|Y}|-|(\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X})\backslash\mathcal{V}_{U|Z}| &= |\mathcal{V}_{U|Z}|-|\mathcal{H}_{U|Y}|+|(\mathcal{H}_{U|Y}\cap\mathcal{V}_{U|X})\backslash\mathcal{V}_{U|Z}| \\ &\geqslant |\mathcal{V}_{U|Z}|-|\mathcal{H}_{U|Y}|.\end{aligned}$$

Hence, by Lemma 3.4.1 and [87], we have

$$\lim_{N\to\infty}(|\mathcal{V}_{U|Z}\backslash\mathcal{H}_{U|Y}|-|(\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X})\backslash\mathcal{V}_{U|Z}|)/N \geqslant H(U|Z)-H(U|Y).$$

Since $I(Y;U)-I(Z;U)>0$, $|\mathcal{V}_{U|Z}\backslash\mathcal{H}_{U|Y}|-|(\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X})\backslash\mathcal{V}_{U|Z}|>0$ for N large enough, and we conclude that $\mathcal{A}_{UYZ}$ exists.

3.6.3.3 Communication rate

The total communication is

$$\begin{aligned}|R_1|+\sum_{i=1}^{k}(|F_i|+|F_i'|) &= |R_1|+\sum_{i=1}^{k}|\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X}| \\ &= |\mathcal{V}_{U|X}|+k|\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X}| \\ &= |\mathcal{V}_{U|X}|+k(|\mathcal{H}_{U|Y}|-|\mathcal{V}_{U|X}|)\end{aligned}$$

where the last equality holds because $U\to X\to Y$ and thus $\mathcal{V}_{U|X}\subset\mathcal{V}_{U|Y}\subset\mathcal{H}_{U|Y}$. Hence, the communication rate is by Lemma 3.4.1 and [87],

$$\lim_{N\to\infty}\frac{|\mathcal{V}_{U|X}|+k(|\mathcal{H}_{U|Y}|-|\mathcal{V}_{U|X}|)}{kN} = I(X;U)-I(Y;U)+\frac{H(U|X)}{k}.$$

3.6.3.4 Key rate

The length of the key generated is

$$\begin{aligned}|K_{1:k}| &= \sum_{i=1}^{k} |K_i| \\ &= k|(\mathcal{V}_{U|Z} \backslash \mathcal{H}_{U|Y}) \backslash \mathcal{A}_{UYZ}| \\ &= k(|\mathcal{V}_{U|Z}| - |\mathcal{H}_{U|Y}| + |(\mathcal{H}_{U|Y} \cap \mathcal{V}_{U|X}) \backslash \mathcal{V}_{U|Z}|) \\ &\geqslant k(|\mathcal{V}_{U|Z}| - |\mathcal{H}_{U|Y}|).\end{aligned}$$

Hence, the key rate is by Lemma 3.4.1 and [87],

$$\lim_{N \to \infty} \frac{|K_{1:k}|}{kN} \geqslant I(Y;U) - I(Z;U).$$

3.6.3.5 Reliability

For $i \in [\![1,k]\!]$, Bob forms $\widehat{V}_i^{1:N}$ from $(F_i, F_i', R_i) = \widetilde{V}_i^{1:N}[\mathcal{H}_{U|Y}]$ and $Y_i^{1:N}$ with the successive cancellation encoder of [87]. Consider an optimal coupling [94,96] between $\widetilde{p}_{V_i^{1:N}}$ and $p_{V_i^{1:N}}$ such that $\mathbb{P}[\mathcal{E}] = \mathbb{V}(\widetilde{p}_{V_i^{1:N}}, p_{V_i^{1:N}})$, where $\mathcal{E} \triangleq \{\widetilde{V}_i^{1:N} \neq V_i^{1:N}\}$. For $i \in [\![2,k]\!]$, note that F_i' is correctly received only when Bob has $\widetilde{K}_{i-1}$, i.e., when $\widetilde{V}_{i-1}^{1:N}$ is correctly reconstructed. We note $\widehat{F}_i'$ the estimate of F_i' formed by Bob from $\widetilde{V}_{i-1}^{1:N}$

and define $\mathcal{E}_{F_i'} \triangleq \{F_i' \neq \widehat{F}_i'\}$. We then have

$$
\begin{aligned}
&\mathbb{P}[\widehat{V}_i^{1:N} \neq \widetilde{V}_i^{1:N}] \\
&= \mathbb{P}[\widehat{V}_i^{1:N} \neq \widetilde{V}_i^{1:N} | \mathcal{E} \cup \mathcal{E}_{F_i'}]\mathbb{P}[\mathcal{E} \cup \mathcal{E}_{F_i'}] + \mathbb{P}[\widehat{V}_i^{1:N} \neq \widetilde{V}_i^{1:N} | \mathcal{E}^c \cap \mathcal{E}_{F_i'}^c]\mathbb{P}[\mathcal{E}^c \cap \mathcal{E}_{F_i'}^c] \\
&\leqslant \mathbb{P}[\mathcal{E} \cup \mathcal{E}_{F_i'}] + \mathbb{P}[\widehat{V}_i^{1:N} \neq \widetilde{V}_i^{1:N} | \mathcal{E}^c \cap \mathcal{E}_{F_i'}^c] \\
&\leqslant \mathbb{P}[\mathcal{E}] + \mathbb{P}[\mathcal{E}_{F_i'}] + \mathbb{P}[\widehat{V}_i^{1:N} \neq \widetilde{V}_i^{1:N} | \mathcal{E}^c \cap \mathcal{E}_{F_i'}^c] \\
&= \mathbb{V}(\widetilde{p}_{V_i^{1:N}}, p_{V_i^{1:N}}) + \mathbb{P}[\mathcal{E}_{F_i'}] + \mathbb{P}[\widehat{V}_i^{1:N} \neq \widetilde{V}_i^{1:N} | \mathcal{E}^c \cap \mathcal{E}_{F_i'}^c] \\
&= \mathbb{V}(\widetilde{p}_{V_i^{1:N}}, p_{V_i^{1:N}}) + \mathbb{P}[\mathcal{E}_{F_i'}] + \mathbb{P}[\widehat{V}_i^{1:N} \neq V_i^{1:N} | \mathcal{E}^c \cap \mathcal{E}_{F_i'}^c] \\
&\leqslant \mathbb{V}(\widetilde{p}_{X_i^{1:N}V_i^{1:N}}, p_{X_i^{1:N}V_i^{1:N}}) + \mathbb{P}[\mathcal{E}_{F_i'}] + \mathbb{P}[\widehat{V}_i^{1:N} \neq V_i^{1:N} | \mathcal{E}^c \cap \mathcal{E}_{F_i'}^c] \\
&\overset{(a)}{\leqslant} \sqrt{2\log 2}\sqrt{N\delta_N} + \mathbb{P}[\mathcal{E}_{F_i'}] + \mathbb{P}[\widehat{V}_i^{1:N} \neq V_i^{1:N} | \mathcal{E}^c \cap \mathcal{E}_{F_i'}^c] \\
&\overset{(b)}{\leqslant} \sqrt{2\log 2}\sqrt{N\delta_N} + \mathbb{P}[\mathcal{E}_{F_i'}] + N\delta_N \\
&\leqslant \sqrt{2\log 2}\sqrt{N\delta_N} + N\delta_N + \mathbb{P}[\widehat{V}_{i-1}^{1:N} \neq \widetilde{V}_{i-1}^{1:N}] \\
&\overset{(c)}{\leqslant} (i-1)(\sqrt{2\log 2}\sqrt{N\delta_N} + N\delta_N) + \mathbb{P}[\widehat{V}_1^{1:N} \neq \widetilde{V}_1^{1:N}] \\
&\overset{(d)}{\leqslant} i(\sqrt{2\log 2}\sqrt{N\delta_N} + N\delta_N),
\end{aligned}
$$

where (a) holds by Lemma 3.6.6, (b) holds because $\mathbb{P}[\widehat{V}_i^{1:N} \neq V_i^{1:N} | \mathcal{E}^c \cap \mathcal{E}_{F_i'}^c] \leqslant N\delta_N$ by [87], (c) holds by recurrence, (d) holds by [87] and because $\widetilde{K}_0$ is known to Bob.

Hence, $\mathbb{P}[K_i \neq \widehat{K}_i] \leqslant i(\sqrt{2\log 2}\sqrt{N\delta_N} + N\delta_N)$. Then, similar to Section 3.5.3.3, we obtain with a union bound

$$\mathbf{P}_e(\mathcal{S}_N) \leqslant \frac{k(k+1)}{2}(\sqrt{2\log 2}\sqrt{N\delta_N} + N\delta_N). \tag{52}$$

3.6.3.6 Key uniformity

We first show the key is nearly uniform for every block in the following lemma.

Lemma 3.6.7. *For every block $i \in [\![1, k]\!]$, the vector $[K_i, \widetilde{K}_i, F_i, R_1]$ is nearly uniform, in the sense that*

$$\mathbb{V}(p_{K_i,\widetilde{K}_i,F_i,R_1}, q_{\mathcal{U}_{K,\widetilde{K},F,R}}) \leqslant 2\sqrt{2\log 2}\sqrt{N\delta_N},$$

where $q_{\mathcal{U}_{K,\widetilde{K},F,R}}$ is the uniform distribution over $[\![1, 2^{|(\mathcal{V}_{U|Z}\backslash\mathcal{H}_{U|Y})\cup((\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X})\cap\mathcal{V}_{U|Z})\cup\mathcal{V}_{U|X}|}]\!]$.

Proof. We have

$$
\begin{aligned}
&\mathbb{V}(p_{K_i,\widetilde{K}_i,F_i,R_i}, q_{\mathcal{U}_{K,\widetilde{K},F,R}})\\
&\overset{(a)}{\leqslant} \mathbb{V}(\widetilde{p}_{V_i^{1:N}[\mathcal{V}_U]}, q_{\mathcal{U}_{\mathcal{V}_U}})\\
&\overset{(b)}{\leqslant} \mathbb{V}(\widetilde{p}_{V_i^{1:N}[\mathcal{V}_U]}, p_{V_i^{1:N}[\mathcal{V}_U]}) + \mathbb{V}(p_{V_i^{1:N}[\mathcal{V}_U]}, q_{\mathcal{U}_{\mathcal{V}_U}})\\
&\overset{(c)}{\leqslant} \sqrt{2N\delta_N \log 2} + \mathbb{V}(p_{V_i^{1:N}[\mathcal{V}_U]}, q_{\mathcal{U}_{\mathcal{V}_U}})\\
&\overset{(d)}{\leqslant} \sqrt{2\log 2}\sqrt{N\delta_N} + \sqrt{2\log 2}\sqrt{\mathbb{D}(p_{V_i^{1:N}[\mathcal{V}_U]}||q_{\mathcal{U}_{\mathcal{V}_U}})}\\
&= \sqrt{2\log 2}\sqrt{N\delta_N} + \sqrt{2\log 2}\sqrt{|\mathcal{V}_U| - H(V_i^{1:N}[\mathcal{V}_U])}\\
&\overset{(e)}{\leqslant} 2\sqrt{2\log 2}\sqrt{N\delta_N},
\end{aligned}
$$

where (a) holds because $\mathcal{V}_{U|Z} \subset \mathcal{V}_U$ and $\mathcal{V}_{U|X} \subset \mathcal{V}_U$ with $q_{\mathcal{U}_{\mathcal{V}_U}}$ the uniform distribution over $[\![1, 2^{|\mathcal{V}_U|}]\!]$, (b) holds by the triangle inequality, (c) holds by Lemma 3.6.6, (d) holds by Pinsker's inequality, (e) holds because similar to the proof of Lemma 3.5.2 $|\mathcal{V}_U| - H(V_i^{1:N}[\mathcal{V}_U]) \leqslant N\delta_N$. □

From Lemma 3.6.7, we derive the following lemmas.

Lemma 3.6.8. *For $i \in [\![1, k]\!]$, we have for N large enough*

$$|K_i| + |\widetilde{K}_i| - H(K_i\widetilde{K}_i) \leqslant \delta_N^{(1)},$$

where

$$\delta_N^{(1)} \triangleq 2\sqrt{2\log 2}\sqrt{N\delta_N}(N - \log_2(2\sqrt{2\log 2}\sqrt{N\delta_N})). \tag{53}$$

In particular, we also have $|K_i| - H(K_i) \leqslant \delta_N^{(1)}$ and $|\widetilde{K}_i| - H(\widetilde{K}_i) \leqslant \delta_N^{(1)}$.

Proof. See Appendix 3.B.3. □

Lemma 3.6.9. *For $i \in [\![1, k]\!]$, we have for N large enough*

$$I(K_i; \widetilde{K}_i R_1) \leqslant \delta_N^{(2)} \quad \textit{and} \quad I(\widetilde{K}_i; R_1) \leqslant \delta_N^{(2)},$$

where

$$\delta_N^{(2)} \triangleq 6\sqrt{2\log 2}\sqrt{N\delta_N}(N - \log_2(6\sqrt{2\log 2}\sqrt{N\delta_N})). \tag{54}$$

Proof. See Appendix 3.B.4. □

We now show that the global key $K_{1:k}$ is uniform. Specifically, we have

$$\begin{aligned}
H(K_{1:k}) &= \sum_{i=1}^{k} H(K_i|K_{1:i-1}) \\
&\geqslant \sum_{i=1}^{k} H(K_i|K_{1:i-1}R_1) \\
&\overset{(a)}{=} \sum_{i=1}^{k} H(K_i|R_1) \\
&= \sum_{i=1}^{k} H(K_i) - \sum_{i=1}^{k} I(K_i;R_1) \\
&\overset{(b)}{\geqslant} \sum_{i=1}^{k} H(K_i) - k\delta_N^{(2)} \\
&\overset{(c)}{\geqslant} \sum_{i=1}^{k} (|K_i| - \delta_N^{(1)}) - k\delta_N^{(2)} \\
&= |K_{1:k}| - k(\delta_N^{(1)} + \delta_N^{(2)})
\end{aligned}$$

where (a) holds because $K_i \to R_1 \to K_{1:i-1}$ for any $i \in [\![1,k]\!]$, (b) holds by Lemma 3.6.9, (c) holds by Lemma 3.6.8. Hence,

$$\mathbf{U}(\mathcal{S}_N) = |K_{1:k}| - H(K_{1:k}) \leqslant k(\delta_N^{(1)} + \delta_N^{(2)}). \tag{55}$$

3.6.3.7 Strong secrecy

Because of the successive cancellation encoding, the secrecy analysis is more involved than for Model 1.

Lemma 3.6.10. *For $i \in [\![1,k]\!]$, we have for N large enough*

$$I(\widetilde{V}_i^{1:N}[\mathcal{V}_{U|Z}]; Z_i^{1:N}) \leqslant \delta_N^{(3)},$$

where

$$\delta_N^{(3)} \triangleq 3\sqrt{2\log 2}\sqrt{N\delta_N}(N - \log_2(3\sqrt{2\log 2}\sqrt{N\delta_N})). \tag{56}$$

Proof. See Appendix 3.B.5. □

The following lemma shows that secrecy holds for each block.

Lemma 3.6.11. *For each Block $i \in \llbracket 1, k \rrbracket$, $[K_i, \widetilde{K}_i]$ is a secret key in the sense that*

$$I\left(K_i\widetilde{K}_i; R_1 M_i Z_i^{1:N}\right) \leqslant 2\delta_N^{(1)} + \delta_N^{(2)} + \delta_N^{(3)}.$$

Proof. By the proof of Lemma 3.6.7, we have

$$\mathbb{V}(\widetilde{p}_{V_i^{1:N}[\mathcal{V}_{U|Z}]}, q_{\mathcal{U}_{\mathcal{V}_{U|Z}}}) \leqslant 2\sqrt{2\log 2}\sqrt{N\delta_N},$$

where $q_{\mathcal{U}_{\mathcal{V}_{U|Z}}}$ is the uniform distribution over $\llbracket 1, 2^{|\mathcal{V}_{U|Z}|} \rrbracket$, and by the proof of Lemma 3.6.8, we have

$$|\mathcal{V}_{U|Z}| - H(V_i^{1:N}[\mathcal{V}_{U|Z}]) \leqslant \delta_N^{(1)}. \tag{57}$$

We have

$$\begin{aligned}
&I(K_i\widetilde{K}_i; R_1 F_i Z_i^{1:N})\\
&= H(K_i\widetilde{K}_i) - H(K_i\widetilde{K}_i|R_1 F_i Z_i^{1:N})\\
&\leqslant |K_i| + |\widetilde{K}_i| - H(K_i\widetilde{K}_i R_1 F_i Z_i^{1:N}) + H(R_1 F_i Z_i^{1:N})\\
&= |K_i| + |\widetilde{K}_i| - H(K_i\widetilde{K}_i R_1 F_i|Z_i^{1:N}) + H(F_i R_1|Z_i^{1:N})\\
&\leqslant |K_i| + |\widetilde{K}_i| + |F_i| + |R_1| - H(K_i\widetilde{K}_i R_1 F_i|Z_i^{1:N})\\
&\overset{(a)}{\leqslant} |\mathcal{V}_{U|Z}| - H(\widetilde{V}_i^{1:N}[\mathcal{V}_{U|Z}]|Z_i^{1:N})\\
&= |\mathcal{V}_{U|Z}| - H(\widetilde{V}_i^{1:N}[\mathcal{V}_{U|Z}]) + I(\widetilde{V}_i^{1:N}[\mathcal{V}_{U|Z}]; Z_i^{1:N})\\
&\overset{(b)}{\leqslant} \delta_N^{(1)} + I(\widetilde{V}_i^{1:N}[\mathcal{V}_{U|Z}]; Z_i^{1:N})\\
&\overset{(c)}{\leqslant} \delta_N^{(1)} + \delta_N^{(3)},
\end{aligned} \tag{58}$$

where (a) holds because $(K_i, \widetilde{K}_i, R_1, F_i)$ is a subvector of $\widetilde{V}_i^{1:N}[\mathcal{V}_{U|Z}]$ noting that $\mathcal{V}_{U|X} \subset \mathcal{V}_{U|Z}$ since $U \to X \to Z$, (b) holds (57), (c) holds by Lemma 3.6.10.

Then, we obtain

$$\begin{aligned}
&I(K_i\widetilde{K}_i; R_1 M_i Z_i^{1:N}) - I(K_i\widetilde{K}_i; R_1 F_i Z_i^{1:N})\\
&\overset{(d)}{=} I(K_i\widetilde{K}_i; F_i' \oplus \widetilde{K}_{i-1} | R_1 F_i Z_i^{1:N})\\
&\overset{(e)}{\leqslant} I(R_1 K_i \widetilde{K}_i F_i F_i' Z_i^{1:N}; F_i' \oplus \widetilde{K}_{i-1})\\
&= H(F_i' \oplus \widetilde{K}_{i-1}) - H(F_i' \oplus \widetilde{K}_{i-1} | R_1 K_i \widetilde{K}_i F_i F_i' Z_i^{1:N})\\
&= H(F_i' \oplus \widetilde{K}_{i-1}) - H(\widetilde{K}_{i-1} | R_1 K_i \widetilde{K}_i F_i F_i' Z_i^{1:N})\\
&\overset{(f)}{=} H(F_i' \oplus \widetilde{K}_{i-1}) - H(\widetilde{K}_{i-1} | R_1)\\
&\leqslant |\widetilde{K}_{i-1}| - H(\widetilde{K}_{i-1} | R_1)\\
&= |\widetilde{K}_{i-1}| - H(\widetilde{K}_{i-1}) + I(\widetilde{K}_{i-1}; R_1)\\
&\overset{(g)}{\leqslant} \delta_N^{(1)} + \delta_N^{(2)}, \qquad (59)
\end{aligned}$$

where (d) holds by definition of M_i, (e) holds by the chain rule and positivity of mutual information, (f) holds because $\widetilde{K}_{i-1} \to R_1 \to K_i \widetilde{K}_i F_i F_i' Z_i^{1:N}$, (g) holds by Lemma 3.6.8 and Lemma 3.6.9. Finally, we conclude combining (58) and (59). □

We now state a lemma that will be used to show that secrecy holds for the global scheme.

Lemma 3.6.12. *For $i \in [\![2, k]\!]$, define*

$$\widetilde{L}_e^{1:i} \triangleq I\left(K_{1:i}\widetilde{K}_i; R_1 M_{1:i} Z_{1:i}^{1:N}\right).$$

We have

$$\widetilde{L}_e^{1:i} - \widetilde{L}_e^{1:i-1} \leqslant I\left(K_i\widetilde{K}_i; R_1 M_i Z_i^{1:N}\right) + \sum_{j=1}^{i-1} I\left(K_j; R_1\right) + I\left(K_{i-1}; \widetilde{K}_{i-1} R_1\right).$$

Proof. See Appendix 3.B.6. □

We thus obtain

$$
\begin{aligned}
&\mathbf{L}(\mathcal{S}_N)\\
&= I(K_{1:k}; M_{1:k} Z_{1:k}^{1:N})\\
&\leqslant \widetilde{L}_e^{1:k}\\
&= \sum_{i=2}^{k} (\widetilde{L}_e^{1:i} - \widetilde{L}_e^{1:i-1}) + \widetilde{L}_e^1\\
&\overset{(a)}{\leqslant} \sum_{i=2}^{k} \left(I\left(K_i \widetilde{K}_i; R_1 M_i Z_i^{1:N}\right) + \sum_{j=1}^{i-1} I\left(K_j; R_1\right) + I\left(K_{i-1}; \widetilde{K}_{i-1} R_1\right)\right) + \widetilde{L}_e^1\\
&\overset{(b)}{\leqslant} \sum_{i=2}^{k} \left(I\left(K_i \widetilde{K}_i; R_1 M_i Z_i^{1:N}\right) + i\delta_N^{(2)}\right) + \widetilde{L}_e^1\\
&= \frac{(k-1)(k+2)}{2}\delta_N^{(2)} + \widetilde{L}_e^1 + \sum_{i=2}^{k} I\left(K_i \widetilde{K}_i; R_1 M_i Z_i^{1:N}\right)\\
&\overset{(c)}{\leqslant} \frac{(k-1)(k+2)}{2}\delta_N^{(2)} + k(2\delta_N^{(1)} + \delta_N^{(2)} + \delta_N^{(3)}) \qquad (60)
\end{aligned}
$$

where (a) follows from Lemma 3.6.12, (b) holds by Lemma 3.6.9, (c) holds by Lemma 3.6.11.

3.6.3.8 Seed rate

The seed rate required to initialize the coding scheme is

$$\lim_{k\to\infty} \lim_{N\to\infty} \frac{|(\mathcal{H}_{U|Y} \backslash \mathcal{V}_{U|X}) \backslash \mathcal{V}_{U|Z}|}{kN} \leqslant \lim_{k\to\infty} \frac{H(U|Y)}{k} = 0.$$

Note that the seed rate could be chosen decrease exponentially fast to zero with N, since we may choose $k = 2^{N^\alpha}$, $\alpha < \beta$, and still have $\lim_{N\to\infty} \mathbf{P}_e(\mathcal{S}_N) = 0$ by (52), $\lim_{N\to\infty} \mathbf{U}_e(\mathcal{S}_N) = 0$ by (55), and $\lim_{N\to\infty} \mathbf{L}_e(\mathcal{S}_N) = 0$ by (60) along with (53), (54), (56).

3.7 Model 3: A multiterminal broadcast model

In this section, we develop a polar coding scheme for a multiterminal broadcast model. Sections 3.7.1- 3.7.3 analyze a model with an arbitrary number of terminals but specific source statistics. The extension of the model to general sources is discussed in Section 3.7.4 for the case of three terminals.

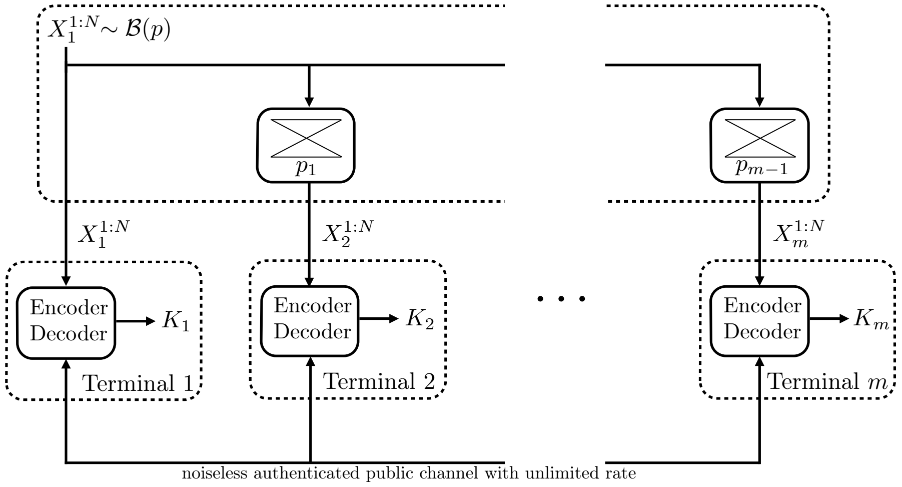


Figure 20. Model 3: Secret-key generation for the 1-to-m broadcast model

3.7.1 Secret-key generation model

As illustrated in Figure 20, we assume that every Terminal $i \in \mathcal{M}\backslash\{1\}$ observes a degraded version of the observation of Terminal 1. For $i \in \mathcal{M}$, we assume that $\mathcal{X}_i = \{0,1\}$ and for $i \in \mathcal{M}\backslash\{1\}$, we set $X_i = X_1 \oplus B_i$, with $X_1 \sim \mathcal{B}(p)$ and $B_i \sim \mathcal{B}(p_{i-1})$, $p_i \in [0,1]$, independent of X_1. Furthermore, we suppose that the eavesdropper does not have access to an observation of the source. We call this setup the 1-to-m broadcast model, and we recall expression of the secret-key capacity in the next proposition.

Proposition 3.7.5 ([16]). *Consider the* 1*-to-*m *broadcast model. The secret-key capacity* $C_{SK}(+\infty)$ *is given by*

$$C_{SK}(+\infty) = \min_{i \in \mathcal{M}\backslash\{1\}} I(X_1; X_i).$$

3.7.2 Polar coding scheme

Define $i_{\min} \triangleq \text{argmin}_{i\in\mathcal{M}\backslash\{1\}} I(X_1;X_i)$ such that $i_{\min} - 1 = \text{argmax}_{i\in\mathcal{M}\backslash\{m\}} p_i$. Let $n \in \mathbb{N}$ and $N \triangleq 2^n$. We set $U^{1:N} \triangleq X_1^{1:N} G_N$. For $\delta_N \triangleq 2^{-N^\beta}$, where $\beta \in]0, 1/2[$, define for $j \in \mathcal{M}\backslash\{1\}$ the sets

$$\mathcal{H}_{X_1|X_j} \triangleq \left\{ i \in [\![1,N]\!] : H\left(U^i | U^{1:i-1} X_j^{1:N}\right) \geqslant \delta_N \right\}.$$

We also define the sets

$$\mathcal{V}_{X_1} \triangleq \left\{ i \in [\![1, N]\!] : H\left(U^i | U^{1:i-1}\right) \geqslant 1 - \delta_N \right\},$$
$$\mathcal{H}_{X_1} \triangleq \left\{ i \in [\![1, N]\!] : H\left(U^i | U^{1:i-1}\right) \geqslant \delta_N \right\}.$$

The encoding and decoding algorithms are given in Algorithm 5 and Algorithm 6, respectively. The high-level principle behind the operation of the algorithm is the following. The set $\mathcal{H}_{X_1|X_i}$ contains the indices such that $U^{1:N}[\mathcal{H}_{X_1|X_i}]$ allows Terminal $i \in \mathcal{M}\backslash\{1\}$ to near losslessly reconstruct $U^{1:N}$ from $X_i^{1:N}$ by [87]. Using a universality argument formalized in Lemma 3.7.13, we will show that it is actually sufficient to transmit $U^{1:N}[\mathcal{H}_{X_1|X_{i_{\min}}}]$ to allow all the terminals to near losslessly reconstruct $U^{1:N}$. The secret key common to all terminals may then be chosen as a subset of $U^{1:N}[\mathcal{V}_{X_1}]$; since $U^{1:N}[\mathcal{H}_{X_1|X_{i_{\min}}}]$ has been publicly transmitted, the secret-key is chosen as $U^{1:N}[\mathcal{V}_{X_1}\backslash\mathcal{H}_{X_1|X_{i_{\min}}}]$. In general, $\mathcal{H}_{X_1|X_{i_{\min}}} \not\subset \mathcal{V}_{X_1}$, and the public communication may leak some information about the key; consequently, as in Model 1 and Model 2, the protocol requires a pre-shared seed to protect the transmission of $U^{1:N}[H_{X_1|X_{i_{\min}}}\backslash\mathcal{V}_{X_1}]$.

```
Algorithm 5: Encoding algorithm for Terminal 1 in Model 3
  Require: $\widetilde{K}$, a secret key of size $|\mathcal{H}_{X_1|X_{i_{\min}}}\backslash\mathcal{V}_{X_1}|$ shared by all terminals
           beforehand; the observations $X_1^{1:N}$ from the source.
  $U^{1:N} \leftarrow X_1^{1:N} G_N$
  $K \leftarrow U^{1:N}[\mathcal{V}_{X_1}\backslash\mathcal{H}_{X_1|X_{i_{\min}}}]$
  $F \triangleq U^{1:N}[\mathcal{V}_{X_1} \cap \mathcal{H}_{X_1|X_{i_{\min}}}]$
  $F' \triangleq U^{1:N}[\mathcal{H}_{X_1|X_{i_{\min}}}\backslash\mathcal{V}_{X_1}]$
  Transmit $M \leftarrow [F, F' \oplus \widetilde{K}]$ publicly to Terminals $\{\mathcal{X}_j\}_{j\in\mathcal{M}\backslash 1}$
  return : $K$
```

As shown in Section 3.7.3, we have the following result.

Theorem 3.7.6. *Consider the* 1*-to-*m *broadcast model of Section 3.7.1. Assume that all terminals share a seed, whose rate can be chosen in* $o(N)$*. The secret-key*

```
Algorithm 6: Decoding algorithm for Terminal $j \in \mathcal{M} \setminus \{1\}$ for Model 3
    Require: $\widetilde{K}$, a secret key of size $|\mathcal{H}_{X_1|X_{i_{\min}}} \backslash \mathcal{V}_{X_1}|$ shared by all terminals
             beforehand; the observations $X_j^{1:N}$ from the source, the message $M$
             transmitted by Terminal 1.
    Form $\widehat{U}^{1:N}$ from $M$ and $\widetilde{K}$ using the successive cancellation decoder of [87].
    $\widehat{K} \leftarrow \widehat{U}^{1:N}[\mathcal{V}_{X_1} \backslash \mathcal{H}_{X_1|X_{i_{\min}}}]$
    return : $\widehat{K}$
```

capacity $C_{SK}(+\infty)$ given in Proposition 3.7.5 is achieved by the polar coding scheme in Algorithm 5 and Algorithm 6, whose computational complexity is $O(N \log N)$.

Proof. See Section 3.7.3. □

The following corollary shows that no seed is required when the source has uniform marginals.

Corollary 3.7.5. *Consider the* 1*-to-*m *broadcast model. Assume that the source has uniform marginal, that is, $X_1 \sim \mathcal{B}(1/2)$. The secret-key capacity $C_{SK}(+\infty)$ given in Proposition 3.7.5 is achievable with perfect secrecy with the polar coding scheme of Algorithm 5 and Algorithm 6 choosing $F' = \emptyset$ and replacing the set $\mathcal{V}_{X_1}$ by $\mathcal{H}_{X_1}$ wherever it appears. .*

We omit the proof of Corollary 3.7.5, which is similar to the ones of Theorem 3.7.6 and Proposition 3.5.3. Note that the model studied in Corollary 3.7.5 is a particular case of [34, Model 3]. However, the construction proposed in [34, Model 3] relies again on a standard array, whose size grows exponentially with the blocklength.

3.7.3 Analysis of polar coding scheme: Proof of Theorem 3.7.6

3.7.3.1 Key rate

Similar to the proof of Theorem 3.6.5, we can show that the key rate is

$$\lim_{N \to +\infty} \frac{|\mathcal{V}_{X_1} \backslash \mathcal{H}_{X_1|X_{i_{\min}}}|}{N} = I(X_1; X_{i_{\min}}).$$

3.7.3.2 Seed rate

Similar to the proof of Theorem 3.6.5, we can show that the seed rate is

$$\lim_{N\to+\infty} \frac{|\mathcal{H}_{X_1|X_{i_{\min}}} \backslash \mathcal{V}_{X_1}|}{N} = 0.$$

3.7.3.3 Reliability

We make use of the following lemma.

Lemma 3.7.13. *For $j \in \mathcal{M}\backslash\{1, i_{\min}\}$, we have $\mathcal{H}_{X_1|X_j} \subset \mathcal{H}_{X_1|X_{i_{\min}}}$.*

Proof. Let $j \in \mathcal{M}\backslash\{1, i_{\min}\}$. We define $\tilde{B}^{(j)}_{i_{\min}} \triangleq B_j + \Delta_j$, with Δ_j independent of B_j and such that $p_{\tilde{B}^{(j)}_{i_{\min}}} = p_{B_{i_{\min}}}$. We set $\tilde{X}^{(j)}_{i_{\min}} \triangleq X_1 + \tilde{B}^{(j)}_{i_{\min}}$. Hence, since $B_{i_{\min}} \sim \mathcal{B}(p_{i_{\min-1}})$, we have for any $x, y \in \{0, 1\}$,

$$\begin{aligned} p_{\tilde{X}^{(j)}_{i_{\min}}|X_1}(x|y) &= (1 - \mathbb{1}\{x = y\})p_{i_{\min-1}} + \mathbb{1}\{x = y\}(1 - p_{i_{\min-1}}) \\ &= p_{X_{i_{\min}}|X_1}(x|y), \end{aligned}$$

that is, $p_{X_1\tilde{X}^{(j)}_{i_{\min}}} = p_{X_1X_{i_{\min}}}$. We now define the sets

$$\mathcal{H}_{X_1|\tilde{X}^{(j)}_{i_{\min}}} \triangleq \left\{ i \in [\![1, N]\!] : H\left(U_i|U^{i-1}\left(\tilde{X}^{(j)}_{i_{\min}}\right)^{1:N}\right) \geqslant \delta_N \right\}.$$

By the data processing equality, we have $\mathcal{H}_{X_1|X_j} \subset \mathcal{H}_{X_1|\tilde{X}^{(j)}_{i_{\min}}}$ but we also have $\mathcal{H}_{X_1|\tilde{X}^{(j)}_{i_{\min}}} = \mathcal{H}_{X_1|X_{i_{\min}}}$ since $p_{X_1\tilde{X}^{(j)}_{i_{\min}}} = p_{X_1X_{i_{\min}}}$, whence $\mathcal{H}_{X_1|X_j} \subset \mathcal{H}_{X_1|X_{i_{\min}}}$. □

By [87, Theorem 3] and by Lemma 3.7.13, for $j \in \mathcal{M}\backslash\{1\}$, Terminal j can reconstruct K from $[F, F'] = U^N[\mathcal{H}_{X_1|X_{i_{\min}}}] \supset U^N[\mathcal{H}_{X_1|X_j}]$ with error probability $\mathbf{P}_e(\mathcal{S}_N) \leqslant N\delta_N$.

3.7.3.4 *Strong secrecy and key uniformity*

Secrecy and uniformity hold since,

$$\begin{aligned}
\mathbf{L}(\mathcal{S}_N)+\mathbf{U}(\mathcal{S}_N) &= I\left(K;F\right)+\log|\mathcal{K}|-H(K)\\
&= |K|-H\left(K|F\right)\\
&= |K|-H\left(KF\right)+H(F)\\
&\leqslant |F|+|K|-H\left(KF\right)\\
&= |\mathcal{V}_{X_1}\cap\mathcal{H}_{X_1|X_{i_{\min}}}|+|\mathcal{V}_{X_1}\backslash\mathcal{H}_{X_1|X_{i_{\min}}}|-H(U^{1:N}[\mathcal{V}_{X_1}])\\
&= |\mathcal{V}_{X_1}|-H(U^{1:N}[\mathcal{V}_{X_1}])\\
&\leqslant N\delta_N,
\end{aligned}$$

where the last inequality can be shown as in the proof of Theorem 3.6.5.

3.7.4 An extension to general sources

The multiterminal model described in Section 3.7.1 only considers binary symmetric channels between the components of the source. A natural question is whether a similar coding scheme may be developed for general sources. We answer this by the affirmative for the case of three terminals; however, the coding scheme is significantly more involved than the one in Section 3.7.2. In the following, we can assume without loss of generality that

$$I(X_1;X_2)=\max_{j\in\{1,2,3\}}\min_{i\in\{1,2,3\}\backslash\{j\}}I(X_j;X_i).$$

Let $n\in\mathbb{N}$ and $N\triangleq 2^n$. We note $U^{1:N}\triangleq {X_2}^{1:N}G_N$, and for $\delta_N\triangleq 2^{-N^\beta}$, where $\beta\in]0,1/2[$, we define the following sets

$$\begin{aligned}
\mathcal{V}_{X_2} &\triangleq \left\{i\in[\![1,N]\!]:H\left(U^i|U^{1:i-1}\right)\geqslant 1-\delta_N\right\},\\
\mathcal{H}_{X_2|X_1} &\triangleq \left\{i\in[\![1,N]\!]:H\left(U^i|U^{1:i-1}X_1^{1:N}\right)\geqslant \delta_N\right\},\\
\mathcal{H}_{X_2|X_3} &\triangleq \left\{i\in[\![1,N]\!]:H\left(U^i|U^{1:i-1}X_3^{1:N}\right)\geqslant \delta_N\right\}.
\end{aligned}$$

We also define

$$\mathcal{K}_{X_\mathcal{M}} \triangleq (\mathcal{V}_{X_2} \backslash \mathcal{H}_{X_2|X_1}) \backslash \mathcal{H}_{X_2|X_3} \quad \text{and} \quad \bar{\mathcal{K}}_{X_\mathcal{M}} \triangleq (\mathcal{V}_{X_2} \backslash \mathcal{H}_{X_2|X_1}) \cap \mathcal{H}_{X_2|X_3},$$

which are such that $\mathcal{V}_{X_2} \backslash \mathcal{H}_{X_2|X_1} = \mathcal{K}_{X_\mathcal{M}} \cup \bar{\mathcal{K}}_{X_\mathcal{M}}$ and $\mathcal{K}_{X_\mathcal{M}} \cap \bar{\mathcal{K}}_{X_\mathcal{M}} = \emptyset$. Finally, we define

$$\begin{aligned}
\mathcal{F}_{X_2|X_1} &\triangleq \mathcal{H}_{X_2|X_1} \cap \mathcal{V}_{X_2}, \\
\bar{\mathcal{F}}_{X_2|X_1} &\triangleq \mathcal{H}_{X_2|X_1} \backslash \mathcal{V}_{X_2}, \\
\mathcal{F}_{X_2|X_3} &\triangleq \mathcal{H}_{X_2|X_3} \cap \mathcal{V}_{X_2}, \\
\bar{\mathcal{F}}_{X_2|X_3} &\triangleq \mathcal{H}_{X_2|X_3} \backslash \mathcal{V}_{X_2},
\end{aligned}$$

which are such that $\mathcal{H}_{X_2|X_1} = \mathcal{F}_{X_2|X_1} \cup \bar{\mathcal{F}}_{X_2|X_1}$, $\mathcal{F}_{X_2|X_1} \cap \bar{\mathcal{F}}_{X_2|X_1} = \emptyset$, $\mathcal{H}_{X_2|X_3} = \mathcal{F}_{X_2|X_3} \cup \bar{\mathcal{F}}_{X_2|X_3}$, and $\mathcal{F}_{X_2|X_3} \cap \bar{\mathcal{F}}_{X_2|X_3} = \emptyset$.

The encoding and decoding algorithms are provided in Algorithm 7, Algorithm 8, and Algorithm 9. The underlying principle is to make Terminals 1 and 3 reconstruct $X_2^{1:N}$ and to choose the secret key as a subset of $U^{1:N}$. For the public communication, we perform universal source coding with side information with an idea similar to [97]. Terminal 2 thus performs encoding over k blocks of size N to transmit the side information necessary to reconstruct $X_2^{1:kN}$ at Terminals 1 and 3. Specifically, Terminal 1 decodes the blocks in order from 1 to k, so that it is able to estimate $U_i^{1:N}[\mathcal{H}_{X_2|X_1}]$ by processing the observations and the public communication in blocks 1 to i. In contrast, Terminal 3 decodes the blocks in reverse order starting from k down to 1, so that it is able to estimate $U_i^{1:N}[\mathcal{H}_{X_2|X_3}]$ by processing the observations and the public communication in blocks k down to i. One of the challenges is to extract a uniform key from $U_{1:k}^{1:N}$ independent of the public communication messages, which we address by protecting some of the public communication corresponding to Block i with part of the secret-key extracted in Block $i-1$. Moreover, similar to Algorithms 1, 3, a small secret seed must be shared by the users to protect the bits in positions $\mathcal{H}_{X_2|X_1}\backslash\mathcal{V}_{X_2} \cup \mathcal{H}_{X_2|X_3}\backslash\mathcal{V}_{X_2}$, which must be revealed to allow reconstruction of the secret-key by Terminals 1, 3, but that may also leak information about the secret-key.

The following remarks clarify why Algorithms 7, 8, 9 achieve the desired behavior.

Remark 3.7.4. *In every block i, Terminal 1 observes $M_i = [F_i^{(1)} \oplus \bar{K}_{i-1}, F_i^{(2)}, F_i' \oplus \widetilde{K}_i]$. Using its estimate of the key $\bar{K}_i$ from the previous block, Terminal 1 estimates $[F_i^{(1)}, F_i^{(2)}, F_i']$, which contains $U_i^{1:N}[\mathcal{H}_{X_2|X_1}]$ by construction. Hence, Terminal 1 has ability to run the successive cancellation decoder and reconstruct $U_i^{1:N}$.*

Remark 3.7.5. *In Block k, Terminal 3 has access to $F_k^{(2)}$, F_k', and $\bar{F}_k$ using M_k and $\widetilde{K}_k$. Since $\mathcal{F}_{X_\mathcal{M}} \subset \mathcal{F}_{X_2|X_1}\backslash\mathcal{F}_{X_2|X_3}$, note that*

$$\begin{aligned}
\mathcal{F}_{X_2|X_1}\backslash\mathcal{F}_{X_\mathcal{M}} &= \mathcal{F}_{X_2|X_1} \cap \mathcal{F}^c_{X_\mathcal{M}} \\
&\supset \mathcal{F}_{X_2|X_1} \cap (\mathcal{F}_{X_2|X_1}\backslash\mathcal{F}_{X_2|X_3})^c \\
&= \mathcal{F}_{X_2|X_1} \cap \mathcal{F}_{X_2|X_3}.
\end{aligned}$$

Hence $U_k^{1:N}[\mathcal{F}_{X_2|X_1} \cap \mathcal{F}_{X_2|X_3}] \subset F_k^{(2)}$, which combined with $\bar{F}_k$ and F_k' allows Terminal 3 to obtain $U_k^{1:N}[\mathcal{H}_{X_2|X_3}]$. Hence, Terminal 3 has the ability to run the successive cancellation decoder and reconstruct $\widehat{U}_k^{1:N}$.

For Block $i \in [\![k-1, 1]\!]$, observe that if $\widehat{U}_{i+1}^{1:N}[\mathcal{F}_{X_\mathcal{M}}] = U_{i+1}^{1:N}[\mathcal{F}_{X_\mathcal{M}}]$, then we have

$$\begin{aligned}
&[F_{i+1}^{(1)} \oplus \bar{K}_i \oplus \widehat{U}_{i+1}^{1:N}[\mathcal{F}_{X_\mathcal{M}}], F_i^{(2)}, F_i'] \\
&= [U_i^{1:N}[\bar{\mathcal{K}}_{X_\mathcal{M}}], F_i^{(2)}, F_i'] \\
&= [U_i^{1:N}[\mathcal{F}_{X_2|X_3}\backslash\mathcal{F}_{X_2|X_1}], F_i^{(2)}, F_i'] \\
&\supset [U_i^{1:N}[\mathcal{F}_{X_2|X_3}\backslash\mathcal{F}_{X_2|X_1}], U_i^{1:N}[\mathcal{F}_{X_2|X_1} \cap \mathcal{F}_{X_2|X_3}], F_i'] \\
&\supset U_i^{1:N}[\mathcal{H}_{X_2|X_3}].
\end{aligned}$$

Consequently, Terminal 3 can form an estimate of $U_i^{1:N}[\mathcal{H}_{X_2|X_3}]$ with

$$[F_{i+1}^{(1)} \oplus \bar{K}_i \oplus \widehat{U}_{i+1}^{1:N}[\mathcal{F}_{X_\mathcal{M}}], F_i^{(2)}, F_i']$$

and apply the successive cancellation decoder to form $\widehat{U}_i^{1:N}$ *an estimate of* $U_i^{1:N}$.

Theorem 3.7.7. *Assume the general setting of Section 3.3 with* $m = 3$, $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{X}_3 = \{0,1\}$, *rate-unlimited public communication, i.e.,* $R_p = +\infty$, *and* $Z = \emptyset$, *i.e., the eavesdropper does not have access to the observation of the source component* Z. *Assume that all terminals share a seed, whose rate can be chosen in* $o(N)$. *The secret-key rate*

$$\max_{j \in \{1,2,3\}} \min_{i \in \{1,2,3\} \backslash \{j\}} I(X_j; X_i)$$

is achieved by the polar coding scheme of Algorithm 7 and Algorithms 8, 9, which involves a chaining of k *blocks of size* N, *and whose complexity is* $O(kN \log N)$.

Proof. See Appendix 3.C. □

As a corollary we obtain the following result for a broadcast model with three terminals.

Corollary 3.7.6. *Assume the broadcast setting of Section 3.7.1 with* $m = 3$, $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{X}_3 = \{0,1\}$, *and an arbitrary distribution* $p_{X_{\mathcal{M}}}$. *Assume that all terminals share a seed, whose rate can be chosen in* $o(N)$. *The secret-key key capacity* $C_s(+\infty) = \min(I(X_1; X_2), I(X_2; X_3))$ *is achieved by the polar coding scheme of Algorithm 7 and Algorithms 8, 9, which involves a chaining of* k *blocks of size* N, *and whose complexity is* $O(kN \log N)$.

3.8 Model 4: Multiterminal markov tree model with uniform marginals

3.8.1 Secret-key generation model

The final model for which we develop a polar coding scheme was first introduced in [34, Model 3]. We assume that all the observation alphabets are $\mathcal{X}_i = \{0,1\}$ for $i \in \mathcal{M}$. As illustrated in Figure 21, consider a tree $\mathcal{T}$ with vertex set $\mathcal{V}(\mathcal{T}) \triangleq \mathcal{M}$

Algorithm 7: Encoding algorithm for Terminal 2 in Model 3

```
Require: $k$ independent secret keys $\{\widetilde{K}_i\}_{i\in\llbracket 1,k\rrbracket}$ of size $|\bar{\mathcal{F}}_{X_2|X_1} \cup \bar{\mathcal{F}}_{X_2|X_3}|$ shared
         by all terminals beforehand; for every block $i \in \llbracket 1,k\rrbracket$, the
         observations $(X_2)_i^{1:N}$ from the source. $\mathcal{F}_{X_{\mathcal{M}}}$, a subset of
         $\mathcal{F}_{X_2|X_1} \backslash \mathcal{F}_{X_2|X_3}$ with size $|\bar{\mathcal{K}}_{X_{\mathcal{M}}}|$.
for Block $i = 1$ to $k$ do
    if $i = 1$ then
        $U_1^{1:N} \leftarrow (X_2)_1^{1:N} G_N$
        $K_1 \leftarrow U_1^{1:N}[\mathcal{K}_{X_{\mathcal{M}}}]$
        $\bar{K}_1 \leftarrow U_1^{1:N}[\bar{\mathcal{K}}_{X_{\mathcal{M}}}]$
        $F_1 \leftarrow U_1^{1:N}[\mathcal{F}_{X_2|X_1}]$
        $F_1' \leftarrow U_1^{1:N}[\bar{\mathcal{F}}_{X_2|X_1} \cup \bar{\mathcal{F}}_{X_2|X_3}]$
        Transmit $M_1 \leftarrow [F_1, F_1' \oplus \widetilde{K}_1]$ publicly to all Terminals
    else if $i = k$ then
        $U_k^{1:N} \leftarrow (X_2)_k^{1:N} G_N$
        $K_k \leftarrow U_k^{1:N}[\mathcal{K}_{X_{\mathcal{M}}} \cup \mathcal{F}_{X_{\mathcal{M}}}]$
        $F_k^{(1)} \leftarrow U_k^{1:N}[\mathcal{F}_{X_{\mathcal{M}}}]$
        $F_k^{(2)} \leftarrow U_k^{1:N}[\mathcal{F}_{X_2|X_1} \backslash \mathcal{F}_{X_{\mathcal{M}}}]$
        $F_k' \leftarrow U_k^{1:N}[\bar{\mathcal{F}}_{X_2|X_1} \cup \bar{\mathcal{F}}_{X_2|X_3}]$
        $\bar{F}_k \leftarrow U_k^{1:N}[\mathcal{F}_{X_2|X_3} \backslash \mathcal{F}_{X_2|X_1}]$
        Transmit $M_k \leftarrow [F_k^{(1)} \oplus \bar{K}_{k-1}, F_k^{(2)}, F_k' \oplus \widetilde{K}_k, \bar{F}_k]$ publicly to all Terminals
    else
        $U_i^{1:N} \leftarrow (X_2)_i^{1:N} G_N$
        $K_i \leftarrow U_i^{1:N}[\mathcal{K}_{X_{\mathcal{M}}} \cup \mathcal{F}_{X_{\mathcal{M}}}]$
        $\bar{K}_i \leftarrow U_i^{1:N}[\bar{\mathcal{K}}_{X_{\mathcal{M}}}]$
        $F_i^{(1)} \leftarrow U_i^{1:N}[\mathcal{F}_{X_{\mathcal{M}}}]$
        $F_i^{(2)} \leftarrow U_i^{1:N}[\mathcal{F}_{X_2|X_1} \backslash \mathcal{F}_{X_{\mathcal{M}}}]$
        $F_i' \leftarrow U_i^{1:N}[\bar{\mathcal{F}}_{X_2|X_1} \cup \bar{\mathcal{F}}_{X_2|X_3}]$
        Transmit $M_i \leftarrow [F_i^{(1)} \oplus \bar{K}_{i-1}, F_i^{(2)}, F_i' \oplus \widetilde{K}_i]$ publicly to all Terminals
    end
end
return : $K_{1:k} \leftarrow [K_1, K_2, \ldots, K_k]$.
```

Algorithm 8: Decoding algorithm for Terminal 1 in Model 3

```
Require: Secret keys $\{\widetilde{K}_i\}_{i \in [\![1,k]\!]}$ of size $|\bar{\mathcal{F}}_{X_2|X_1} \cup \bar{\mathcal{F}}_{X_2|X_3}|$ shared with
         Terminal 2; for every block $i \in [\![1,k]\!]$, the observations $(X_1)_i^{1:N}$ from
         the source; the set $\mathcal{F}_{X_\mathcal{M}}$ defined in Algorithm 7.
for Block $i = 1$ to $k$ do
    if $i = 1$ then
        Form $[F_1, F_1']$ from $M_1$ and $\widetilde{K}_1$ and extract and estimate $U_1^{1:N}[\mathcal{H}_{X_2|X_1}]$
        {See Remark 3.7.4 for a justification}
        Form $\widehat{U}_1^{1:N}$ with the successive cancellation decoder of [87]
        $\widehat{K}_1 \leftarrow \widehat{U}_1^{1:N}[\mathcal{K}_{X_\mathcal{M}}]$
    else
        Estimate $[F_i^{(1)}, F_i^{(2)}, F_i']$ from $M_i$, $\widehat{U}_{i-1}^{1:N}$, and $\widetilde{K}_i$ and extract an estimate
        of $U_i^{1:N}[\mathcal{H}_{X_2|X_1}]$
        Form $\widehat{U}_i^{1:N}$ with the successive cancellation decoder of [87]
        $\widehat{K}_i \leftarrow \widehat{U}_i^{1:N}[\mathcal{K}_{X_\mathcal{M}}]$
    end
end
return : $\widehat{K}_{1:k} \leftarrow [\widehat{K}_1, \widehat{K}_2, \ldots, \widehat{K}_k]$.
```

Algorithm 9: Decoding algorithm for Terminal 3 in Model 3

```
Require: Secret keys $\{\widetilde{K}_i\}_{i \in [\![1,k]\!]}$ of size $|\bar{\mathcal{F}}_{X_2|X_1} \cup \bar{\mathcal{F}}_{X_2|X_3}|$ shared with
         Terminal 2; for every block $i \in [\![1,k]\!]$, the observations $(X_3)_i^{1:N}$ from
         the source; $\mathcal{F}_{X_\mathcal{M}}$ used in encoding.
for Block $i = k$ to 1 do
    if $i = k$ then
        Form $[F_k^{(2)}, F_k', \bar{F}_k]$ from $M_k$ and $\widetilde{K}_k$ and extract an estimate of
        $U_k^{1:N}[\mathcal{H}_{X_2|X_3}]$ {See Remark 3.7.5 for a justification}
        Form $\widehat{U}_k^{1:N}$ with the successive cancellation decoder of [87]
        $\widehat{K}_1 \leftarrow \widehat{U}_1^{1:N}[\mathcal{K}_{X_\mathcal{M}}]$
    else
        Estimate $[\bar{K}_i, F_i^{(2)}, F_i']$ from $M_i$, $\widehat{U}_{i+1}^{1:N}$, and $\widetilde{K}_i$ and extract an estimate
        of $U_i^{1:N}[\mathcal{H}_{X_2|X_3}]$
        Form $\widehat{U}_i^{1:N}$ with the successive cancellation decoder of [87]
        $\widehat{K}_i \leftarrow \widehat{U}_i^{1:N}[\mathcal{K}_{X_\mathcal{M}}]$
    end
end
return : $\widehat{K}_{1:k} \leftarrow [\widehat{K}_1, \widehat{K}_2, \ldots, \widehat{K}_k]$.
```

and edge set $\mathcal{E}(\mathcal{T})$. The joint probability distribution $p_{X_{\mathcal{M}}}$ is characterized as follows. $\forall (i,j) \in \mathcal{E}(\mathcal{T}), \forall x_i, x_j \in \{0,1\}$,

$$p_{X_i X_j}(x_i, x_j) \triangleq \frac{1}{2}(1-p_{i,j})\mathbb{1}\{x_i = x_j\} + \frac{1}{2}p_{i,j}(1 - \mathbb{1}\{x_i = x_j\}),$$

which means that $p_{X_i} = p_{X_j}$ is uniform and the test channel between X_i and X_j is a binary symmetric channel with paramater $p_{i,j}$.

Furthermore, we suppose that the eavesdropper does not have access to the observation of the source component Z. This setup is called the Markov tree model with uniform marginals. The expression of the secret-key capacity is recalled in the following proposition.

Proposition 3.8.6 ([16])**.** *Consider the Markov tree model with uniform marginal. The secret-key capacity $C_{SK}(+\infty)$ is given by*

$$C_{SK}(+\infty) = I(X_{n_0}; X_{n_1}),$$

where $(n_0, n_1) \triangleq \text{argmin}_{(i,j)\in\mathcal{E}(\mathcal{T})} I(X_i; X_j)$.

Note that the construction proposed in [34] is not low-complexity since it relies on the construction of a standard array, whose size grows exponentially with the blocklength.

3.8.2 Polar coding scheme

We first introduce some notations for the coding scheme. For any $i \in \mathcal{M}$, we note $\mathcal{N}^j(i)$ the set of vertices in $\mathcal{V}(\mathcal{T})$ that are at distance j from vertex i. Recall that we note $(n_0, n_1) \triangleq \text{argmin}_{(i,j)\in\mathcal{E}(\mathcal{T})} I(X_i; X_j)$. We also consider for the encoding process the tree $\mathcal{T}$ as a rooted tree with root X_{n_0}. An example is depicted in Figure 21.

Let $n \in \mathbb{N}$ and $N \triangleq 2^n$. For $j \in \mathcal{M}$, we set $U_j^{1:N} \triangleq X_j^{1:N} G_N$. For $j_1 \in \mathcal{M}$, $j_2 \in \mathcal{M}\backslash\{j_1\}$, and $\delta_N \triangleq 2^{-N^\beta}$, $\beta \in]0, 1/2[$, we define the sets

$$\mathcal{H}_{X_{j_1}|X_{j_2}} \triangleq \left\{ i \in [\![1, N]\!] : H\left(U_{j_1}^i | U_{j_1}^{1:i-1} X_{j_2}^{1:N}\right) \geqslant \delta_N \right\}.$$

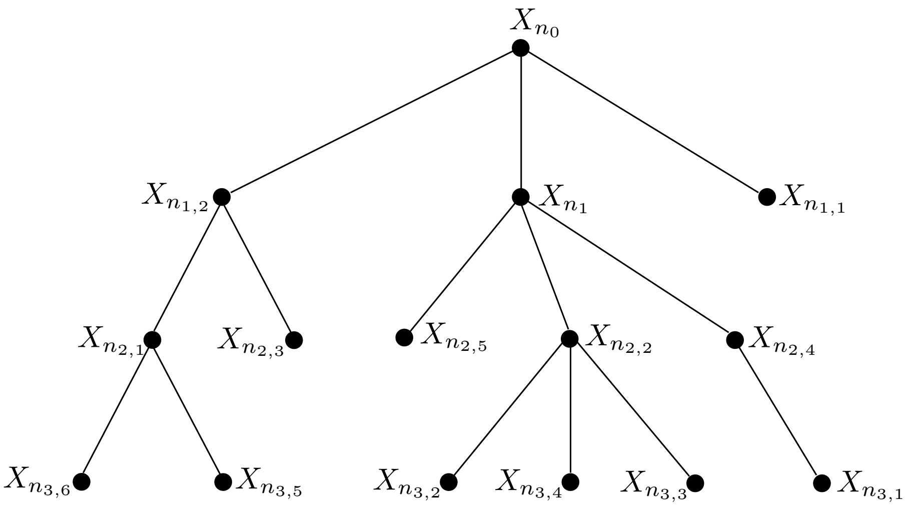


Figure 21. Example of Markov tree model with uniform marginal for $m = 15$. Each vertex represent the random variable observed by a given terminal, and each edge can be seen as a binary symmetric test channel. We have noted $(n_0, n_1) \triangleq$ argmin$_{(i,j)\in\mathcal{E}(\mathcal{T})} I(X_i; X_j)$, $\mathcal{N}^1(n_0) \triangleq \{n_1, n_{1,1}, n_{1,2}\}$, $\mathcal{N}^2(n_0) \triangleq \{n_{2,i}\}_{i\in[\![1,5]\!]}$, $\mathcal{N}^3(n_0) \triangleq \{n_{3,i}\}_{i\in[\![1,6]\!]}$.

The exact encoding and decoding algorithms are given in Algorithm 10 and Algorithm 11. The principle of their operation is to have all terminal reconstruct $U_{n_0}^{1:N}$ and choose the key as a subvector of $U_{n_0}^{1:N}$. The idea behind the inter-terminal communication, which is illustrated in Figure 22, is to take advantage of the tree structure to make all Terminals reconstruct $X_{n_0}^{1:N}$; the source uniformity plays a crucial role to develop a universal result in Lemma 3.8.14, similar to the one obtained for the broadcast model in Lemma 3.7.13. Although the assumption of uniform marginal is required in our proof, a side benefit is that no pre-shared seed is needed to ensure strong secrecy.

We note $\mathcal{F}$ the set of indices (i, j) for which $F_{i,j}$ is defined. We note the collective inter-terminals communication as $\mathbf{F} \triangleq \{F_{i,j}\}_{(i,j)\in\mathcal{F}}$.

The analysis of the scheme in Section 3.8.3 leads to the following result.

Theorem 3.8.8. *Consider the Markov tree model with uniform marginals. The secret-key capacity $C_{SK}(+\infty)$ given in Proposition 3.8.6 is achievable with perfect secrecy with the polar coding scheme of Section 3.8.2, whose computational complexity is $O(N \log N)$.*

Algorithm 10: Encoding algorithm for Model 4

```
$F_{n_0} \leftarrow U_{n_0}^{1:N}\left[\mathcal{H}_{X_{n_0}|X_{n_1}}\right]$.
Terminal $n_0$ transmits $F_{n_0}$ publicly.
Define $d$ as the maximal distance between the vertex $n_0$ and the vertices in $\mathcal{V}(\mathcal{T})$.
for $i = 1$ to $d - 1$ do
    for $j \in \mathcal{N}^i(n_0)$ do
        if $\mathcal{N}^1(j) \cap \mathcal{N}^{i+1}(n_0) \neq \emptyset$ then
            Define $j^* \triangleq \underset{\tilde{j} \in \mathcal{N}^1(j) \cap \mathcal{N}^{i+1}(n_0)}{\operatorname{argmax}} p_{\tilde{j},j}$.
            $F_{i,j} \leftarrow U_j^{1:N}\left[\mathcal{H}_{X_j|X_{j^*}}\right]$ ,
            Terminal $j$ transmits $F_{i,j}$ publicly
        end
    end
end
return : $K \leftarrow U_{n_0}^{1:N}\left[\mathcal{H}^c_{X_{n_0}|X_{n_1}}\right]$
```

Algorithm 11: Decoding algorithm for Model 4

```
Require: Observations from the source, and public messages $\mathbf{F}$.
With $F_{n_0} = U_{n_0}^{1:N}\left[\mathcal{H}_{X_{n_0}|X_{n_1}}\right]$, the terminals in $\mathcal{N}^1(n_0)$ estimate $X_{n_0}^{1:N}$ with the successive cancellation decoder of [87], and then form $\widehat{K}$ an estimate of $K$.
Let $k \in [\![1, d-1]\!]$, $j \in \mathcal{N}^{k+1}(n_0)$ and $i = \mathcal{N}^k(n_0) \cap \mathcal{N}^1(j)$. With $F_{k,i}$ Terminal $j$ estimates $X_i^{1:N}$ with the successive cancellation decoder of [87]. By iterating, Terminal $j$ is successively able to form the estimate of $X_{i_{k-1}}^{1:N}$, $X_{i_{k-2}}^{1:N}$, ..., $X_{i_1}^{1:N}$, $X_{n_0}^{1:N}$, for some $i_1 \in \mathcal{N}^1(n_0)$, $i_2 \in \mathcal{N}^2(n_0)$, ..., $i_k \in \mathcal{N}^{k-1}(n_0)$.
Finally, Terminal $j$ forms $\widehat{K}$ an estimate of $K$ from its estimate of $X_{n_0}^{1:N}$.
return : $\widehat{K}$
```

Proof. See Section 3.8.3. □

Note again that for this model no pre-shared seed is required because the marginal of $p_{X_{\mathcal{M}}}$ are uniform.

3.8.3 Analysis of polar coding scheme: Proof of Theorem 3.8.8

3.8.3.1 Key Rate

From [87], we obtain the key rate

$$\lim_{N\to\infty} \frac{|\mathcal{H}^c_{X_{n_0}|X_{n_1}}|}{N} = 1 - \lim_{N\to\infty} \frac{|\mathcal{H}_{X_{n_0}|X_{n_1}}|}{N} = 1 - H(X_{n_0}|X_{n_1}) = I(X_{n_0}; X_{n_1}).$$

3.8.3.2 Reliability

We first show that for $k \in [\![1, d]\!]$, Terminal $j \in \mathcal{N}^k(n_0)$ can reconstruct X_{j_0} with $j_0 \triangleq \mathcal{N}^1(j) \cap \mathcal{N}^{k-1}(n_0)$ from F_{k-1,j_o}. Specifically, we establish the following.

Lemma 3.8.14. *Let* $k \in [\![1, d]\!]$*,* $j \in \mathcal{N}^k(n_0)$*,* $j_0 \triangleq \mathcal{N}^1(j) \cap \mathcal{N}^{k-1}(n_0)$*. Define* $\mathcal{D}_{k,j_0} \triangleq \mathcal{N}^1(j_0) \cap \mathcal{N}^k(n_0)$*, and* $i^* \triangleq \arg\max_{\tilde{i}\in\mathcal{D}_{k,j_0}} p_{\tilde{i},j_0}$*. We have*

$$\forall i \in \mathcal{D}_{k,j_0},\ \mathcal{H}_{X_{j_0}|X_i} \subset \mathcal{H}_{X_{j_0}|X_{i^*}}.$$

Proof. For $i \in \mathcal{D}$, define $\bar{X}_i \triangleq X_{j_0} + B_i$, with $B_i \sim \mathcal{B}(p_{i,j_0})$. By Lemma 3.7.13, we now that for any $i \in \mathcal{D}$, $\mathcal{H}_{X_{j_0}|\bar{X}_i} \subset \mathcal{H}_{X_{j_0}|\bar{X}_{i^*}}$. Then, observe that for any $i \in \mathcal{D}$, for any $x, y \in \{0, 1\}$,

$$\begin{aligned}
&p_{\bar{X}_i X_{j_0}}(x, y)\\
&= p_{X_{j_0}}(y) p_{\bar{X}_i|X_{j_0}}(x|y)\\
&= \frac{1}{2}(\mathbb{1}\{x = y\}(1 - p_{i,j_0}) + p_{i,j_0}(1 - \mathbb{1}\{x = y\}))\\
&= p_{X_i X_{j_0}}(x, y),
\end{aligned}$$

Hence, $\mathcal{H}_{X_{j_0}|X_i} = \mathcal{H}_{X_{j_0}|\bar{X}_i} \subset \mathcal{H}_{X_{j_0}|\bar{X}_{i^*}} = \mathcal{H}_{X_{j_0}|X_{i^*}}$

□

Lemma 3.8.14 is similar to Lemma 3.7.13; however, unlike Lemma 3.7.13, the proof of Lemma 3.8.14 requires uniform marginals.

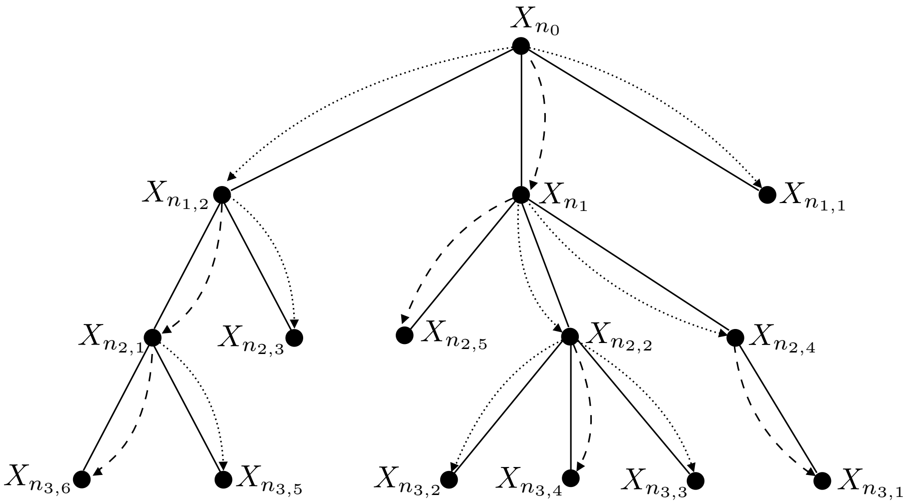


Figure 22. Example for the reconstruction process. A dashed-line from Terminal i to Terminal j represents a public transmission from Terminal i of the information necessary for Terminal j to reconstruct X_i. A dotted-line from Terminal i to Terminal j represents a "virtual communication" and means that Terminal j is able to reconstruct X_i from the information corresponding to the dashed-line leaving Terminal i – this illustrates Lemma 3.8.14. For this example we have assumed $I(X_{n_{1,2}}; X_{n_{2,1}}) \leqslant I(X_{n_{1,2}}; X_{n_{2,3}})$, $I(X_{n_1}; X_{n_{2,5}}) \leqslant \min\{I(X_{n_1}; X_{n_{2,i}})\}_{i \in \{2,4\}}$, $I(X_{n_{2,1}}; X_{n_{3,6}}) \leqslant I(X_{n_{2,1}}; X_{n_{3,5}})$, $I(X_{n_{2,2}}; X_{n_{3,4}}) \leqslant \min\{I(X_{n_{2,2}}; X_{n_{3,i}})\}_{i \in \{2,3\}}$. All in all, all the terminals can reconstruct X_{n_0}.

Now, observe that with $F_{n_0} = U_{n_0}^{1:N}\left[\mathcal{H}_{X_{n_0}|X_{n_1}}\right]$, all terminals in $\mathcal{N}^1(n_0)$ can reconstruct $X_{n_0}^{1:N}$ with error probability $O(N\delta_N)$ by Lemma 3.8.14 and [87]. We then show by induction that all terminals can reconstruct $X_{n_0}^{1:N}$ with error probability $O(N\delta_N)$. Assume that for $k \in [\![1, d-1]\!]$, $X_{n_0}^{1:N}$ can be reconstructed with error probability $O(N\delta_N)$ from any $X_j^{1:N}$, where $j \in \mathcal{N}^k(n_0)$. Let $j \in \mathcal{N}^{k+1}(n_0)$ and $i = \mathcal{N}^k(n_0) \cap \mathcal{N}^1(j)$. With $F_{k,i}$ Terminal j can reconstruct $X_i^{1:N}$ with error probability $O(N\delta_N)$ by Lemma 3.8.14 and [87]. Then, since $X_i^{1:N} \in \mathcal{N}^k(n_0)$, Terminal j can also reconstruct $X_{n_0}^{1:N}$ with error probability $O(N\delta_N)$ by induction hypothesis.

We conclude that all terminals can reconstruct $X_{n_0}^{1:N}$ and thus $K = U_{n_0}^{1:N}\left[\mathcal{H}_{X_{n_0}|X_{n_1}}^c\right]$ with error probability $\mathbf{P}_e(\mathcal{S}_N) = O(N\delta_N)$. The global reconstruction process is illustrated in Figure 22.

3.8.3.3 Key Uniformity

By definition of the model, X_{n_0} is uniform, hence, $U_{n_0}^{1:N}$ and $K \triangleq U_{n_0}^{1:N}\left[\mathcal{H}_{X_{n_0}|n_1}^c\right]$ are also uniform.

3.8.3.4 Perfect Secrecy

We first introduce an equivalent model as follows. We start by defining for $i \in \mathcal{N}^1(n_0)$, $\bar{X}_i \triangleq X_{n_0} + B_i$, with $B_i \sim \mathcal{B}(p_{i,n_0})$. Then, for $k \in [\![2, d]\!]$, for $i \in \mathcal{N}^k(n_0)$, define $i_0 \triangleq \mathcal{N}^{k-1}(n_0) \cap \mathcal{N}^1(i)$, and $\bar{X}_i \triangleq \bar{X}_{i_0} + B_i$, with $B_i \sim \mathcal{B}(p_{i,i_0})$. Consequently, similar to the proof of Lemma 3.8.14, we have

$$p_{\bar{X}_{\mathcal{M}}} = p_{X_{\mathcal{M}}}. \tag{61}$$

Moreover, for $j \in \mathcal{M} \backslash \{n_0\}$. We have

$$\bar{U}_j^{1:N} = U_{n_0}^{1:N} \bigoplus_{i \in \mathcal{P}_{n_0,j}} \widetilde{B}_i^{1:N},$$

where $\mathcal{P}_{n_0,j}$ denotes the set of vertices that form a path between X_{n_0} and X_j including j and excluding n_0, $\widetilde{B}_i^N \triangleq B_i^N G_N$, and $\bar{U}_j^{1:N} \triangleq \bar{X}_j^{1:N} G_N$, $i \in \mathcal{M} \backslash \{n_0\}$. Recall that for $(i, j) \in \mathcal{F}$,

$$F_{i,j} = U_j^{1:N} \left[\mathcal{H}_{X_j|X_{j^*}}\right].$$

We define

$$\bar{F}_{i,j} \triangleq \bar{U}_j^{1:N} \left[\mathcal{H}_{X_j|X_{j^*}}\right] = U_{n_0}^{1:N} \left[\mathcal{H}_{X_j|X_{j^*}}\right] \bigoplus_{i \in \mathcal{P}_{n_0,j}} \widetilde{B}_i^{1:N} \left[\mathcal{H}_{X_j|X_{j^*}}\right], \tag{62}$$

and

$$\bar{\mathbf{F}} \triangleq \{\bar{F}_{i,j}\}_{(i,j)\in\mathcal{F}}. \tag{63}$$

Lemma 3.8.15. *For $j \in \mathcal{M} \backslash \{n_0\}$, $\mathcal{H}_{X_j|X_{j^*}} \subset \mathcal{H}_{X_{n_0}|X_{n_1}}$.*

Proof. Let $j \in \mathcal{M} \backslash \{n_0\}$. Let r_j be such that $p_{n_0,n_1} = p_{j,j^*} \star r_j$ (such r_j exists by definition of (n_0, n_1)), where $\star$ is defined as in Example 3.6.1. We define $\Delta_j^{(1)} \sim \mathcal{B}(p_{j,j^*})$ and $\Delta_j^{(2)} \sim \mathcal{B}(r_j)$ such that $B_{n_1} = \Delta_j^{(1)} + \Delta_j^{(2)}$. We define the dummy random variables $\bar{\bar{X}}_{j^*} \triangleq X_{n_0} + \Delta_j^{(1)}$ and $\bar{\bar{X}}_{n_1} \triangleq X_{n_0} + \Delta_j^{(1)} + \Delta_j^{(1)}$. Then, for any $x, y \in \{0, 1\}$, and by uniformity of the marginals of $p_{X_{\mathcal{M}}}$,

$$\begin{aligned}
p_{\bar{\bar{X}}_{j^*} X_{n_0}}(x,y) &= p_{X_{n_0}}(y) p_{\bar{\bar{X}}_{j^*}|X_{n_0}}(x|y) \\
&= \frac{1}{2} p_{\bar{\bar{X}}_{j^*}|X_{n_0}}(x|y) \\
&= \frac{1}{2} \left[(1 - \mathbb{1}\{x = y\}) p_{j,j^*} + (1 - p_{j,j^*}) \mathbb{1}\{x = y\} \right] \\
&= \frac{1}{2} p_{X_{j^*}|X_j}(x|y) \\
&= p_{X_{j^*} X_j}(x,y),
\end{aligned}$$

so that $\mathcal{H}_{X_j|X_{j^*}} = \mathcal{H}_{X_{n_0}|\bar{\bar{X}}_{j^*}}$. Similarly, we have $p_{X_{n_1} X_{n_0}} = p_{\bar{\bar{X}}_{n_1} X_{n_0}}$ so that $\mathcal{H}_{X_{n_0}|X_{n_1}} = \mathcal{H}_{X_{n_0}|\bar{\bar{X}}_{n_1}}$. Hence, by the data processing inequality, we obtain

$$\mathcal{H}_{X_j|X_{j^*}} = \mathcal{H}_{X_{n_0}|\bar{\bar{X}}_{j^*}} \subset \mathcal{H}_{X_{n_0}|\bar{\bar{X}}_{n_1}} = \mathcal{H}_{X_{n_0}|X_{n_1}}.$$

□

We can now show that perfect secrecy holds as follows.

$$\begin{aligned}
&\mathbf{L}(\mathcal{S}_N) \\
&= I(K; \mathbf{F}) \\
&= I\left(U_{n_0}^{1:N} \left[\mathcal{H}_{X_{n_0}|X_{n_1}}^c \right] ; \mathbf{F} \right) \\
&\overset{(a)}{=} I\left(\bar{U}_{n_0}^{1:N} \left[\mathcal{H}_{X_{n_0}|X_{n_1}}^c \right] ; \bar{\mathbf{F}} \right) \\
&\overset{(b)}{\leqslant} I\left(\bar{U}_{n_0}^{1:N} \left[\mathcal{H}_{X_{n_0}|X_{n_1}}^c \right] ; \bar{U}_{n_0}^{1:N} \left[\mathcal{H}_{X_{n_0}|X_{n_1}} \right], \widetilde{B}_{[\![1,m]\!] \backslash \{n_0\}}^{1:N} \left[\mathcal{H}_{X_{n_0}|X_{n_1}} \right] \right) \\
&= I\left(\bar{U}_{n_0}^{1:N} \left[\mathcal{H}_{X_{n_0}|X_{n_1}}^c \right] ; \bar{U}_{n_0}^{1:N} \left[\mathcal{H}_{X_{n_0}|X_{n_1}} \right] \right) \\
&\qquad + I\left(\bar{U}_{n_0}^{1:N} \left[\mathcal{H}_{X_{n_0}|X_{n_1}}^c \right] ; \widetilde{B}_{[\![1,m]\!] \backslash \{n_0\}}^{1:N} \left[\mathcal{H}_{X_{n_0}|X_{n_1}} \right] | \bar{U}_{n_0}^{1:N} \left[\mathcal{H}_{X_{n_0}|X_{n_1}} \right] \right) \\
&\overset{(c)}{=} I\left(\bar{U}_{n_0}^{1:N} \left[\mathcal{H}_{X_{n_0}|X_{n_1}}^c \right] ; \widetilde{B}_{[\![1,m]\!] \backslash \{n_0\}}^{1:N} \left[\mathcal{H}_{X_{n_0}|X_{n_1}} \right] | \bar{U}_{n_0}^{1:N} \left[\mathcal{H}_{X_{n_0}|X_{n_1}} \right] \right) \\
&\leqslant I\left(\bar{U}_{n_0}^{1:N} ; \widetilde{B}_{[\![1,m]\!] \backslash \{n_0\}}^{1:N} \left[\mathcal{H}_{X_{n_0}|X_{n_1}} \right] \right) \\
&\overset{(d)}{=} 0,
\end{aligned}$$

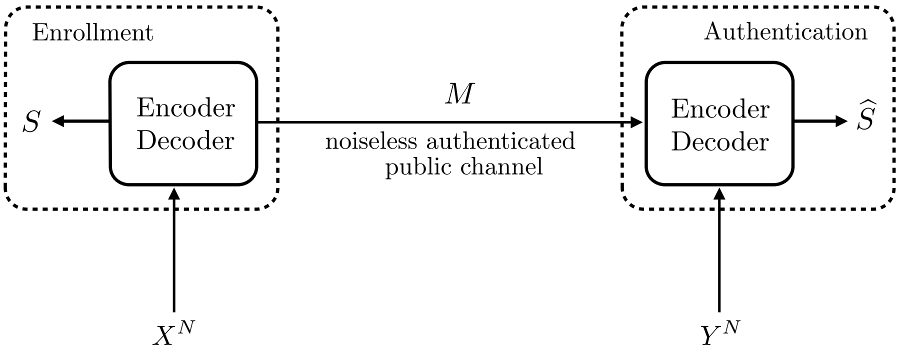


Figure 23. Model for biometric secret generation

where (a) follows by (61), (62), and (63), (b) follows from Lemma 3.8.15 and Equation (62), (c) follows by uniformity of $\bar{U}_{n_0}^{1:N}$, (d) holds by independence of $\bar{U}_{n_0}^{1:N}$ and $\widetilde{B}_{[\![1,m]\!]\backslash\{n_0\}}^{1:N}$. We have thus shown perfect secrecy.

3.9 Application to secrecy and privacy for biometric systems

In this final section, we show how the results obtained for Model 2 may be applied to the related problems of secrecy and privacy for biometric systems [65,98–100]. As noted in [65], the main difficulty in constructing practical codes for such problems is the need for vector quantization; we show here that polar codes offer a low-complexity solution and provably optimal solutionsfor the models studied in [65].

3.9.1 Biometric system models

Consider two biometric sequences $X^{1:N}$ and $Y^{1:N}$ distributed according to the memoryless source $(\mathcal{X}\mathcal{Y}, p_{XY})$. Assume that $X^{1:N}$ is an enrollment sequence and $Y^{1:N}$ an authentication sequence observed by an encoder and a decoder, respectively. In [65], four different models are considered. We only deal with the "generated-secret systems" and the "generated-secret systems with zero leakage," as codes for the latter models can be used for the "chosen-secret systems" and the "chosen-secret systems with zero leakage" using a masking technique [65].

3.9.1.1 Generated-secret systems

A biometric secret generation strategy $\mathcal{S}_N^{\text{bio}}$ is illustrated in Figure 23 and is formally defined as follows.

Definition 3.9.3. *Let $R \in \mathbb{R}^+$. Let $\mathcal{S}$ be an alphabet of size 2^{NR}. The protocol defined by the following steps is called a $(2^{NR}, N, R)$ biometric secret generation strategy.*

- *The encoder observes the enrollment sequence $X^{1:N}$;*
- *The encoder generates a secret $S \in \mathcal{S}$ from $X^{1:N}$;*
- *The encoder transmits publicly to the decoder helper data M;*
- *The decoder observes the authentication sequence $Y^{1:N}$, and computes $\widehat{S} \in \mathcal{S}$.*

The performance of a biometric secret generation strategy is measured in terms of

- the average probability of error between the biometric secrets with $\mathbf{P}_e(\mathcal{S}_N^{\text{bio}}) \triangleq \mathbb{P}[S \neq \widehat{S}]$,
- the information leakage of M on S with $\mathbf{L}(\mathcal{S}_N^{\text{bio}}) \triangleq I(M;S)$,
- the privacy leakage of M on $X^{1:N}$ with $\mathbf{P}_{\text{c}}(\mathcal{S}_N^{\text{bio}}) \triangleq I(M;X^{1:N}|S)$ (conditional case), or $\mathbf{P}_{\text{u}}(\mathcal{S}_N^{\text{bio}}) \triangleq I(M;X^{1:N})$ (unconditional case),
- the uniformity of the biometric secret $\mathbf{U}(\mathcal{S}_N^{\text{bio}}) \triangleq \log\lceil 2^{NR}\rceil - H(S)$.

Definition 3.9.4. *For a fixed privacy leakage threshold L, a biometric secret rate R and information is achievable if there exists a sequence of $(2^{NR}, N, R)$ secret-key* *generation strategies* $\left\{\mathcal{S}_N^{\text{bio}}\right\}_{N\geqslant 1}$ *such that*

$$\begin{aligned}
\lim_{N\to\infty} \mathbf{P}_e(\mathcal{S}_N^{\text{bio}}) &= 0, \ \textit{(reliability)}\\
\lim_{N\to\infty} \mathbf{L}(\mathcal{S}_N^{\text{bio}}) &= 0, \ \textit{(strong secrecy)}\\
\lim_{N\to\infty} \mathbf{P}_{\text{c}}(\mathcal{S}_N^{\text{bio}})/N &\leqslant L, \ \textit{(privacy leakage)}\\
\lim_{N\to\infty} \mathbf{U}(\mathcal{S}_N^{\text{bio}}) &= 0. \ \textit{(uniformity)}
\end{aligned}$$

Moreover, the supremum of achievable rates is called the biometric secret capacity and is denoted $C^{\text{c}}_{Bio}(L)$*. For the unconditional case,* $\mathbf{P}_{\text{c}}(\mathcal{S}_N^{\text{bio}})$ *is replaced with* $\mathbf{P}_{\text{u}}(\mathcal{S}_N^{\text{bio}})$*, and the biometric secret capacity and is denoted by* $C^{\text{u}}_{Bio}(L)$*.*

Note that we require a stronger security metric than in [65]. The biometric secret capacities are known and recalled below.

Theorem 3.9.9 ([65])**.** *Let* $(\mathcal{X}\mathcal{Y}, p_{XY})$ *be a BMS and* $L \in \mathbb{R}_+$ *be a privacy leakage threshold. The conditional and unconditional biometric secret capacities are equal* $C^{\text{c}}_{Bio}(L) = C^{\text{u}}_{Bio}(L)$*, moreover,*

$$C^{\text{c}}_{Bio}(L) = \max_U I(Y;U)$$

subject to

$$\begin{aligned}
&L = I(U;X) - I(U;Y),\\
&U \to X \to Y,\\
&|\mathcal{U}| \leqslant |\mathcal{X}|.
\end{aligned}$$

Remark 3.9.6. *The equality* $L = I(U;X) - I(U;Y)$ *and the range constraint* $|\mathcal{U}| \leqslant |\mathcal{X}|$ *are obtained from Proposition 2.5.3.*

3.9.1.2 Generated-secret systems with zero leakage

A biometric secret generation strategy with zero leakage $\mathcal{S}_N^{\text{bioZ}}$ is describes in Figure 24 and is formally defined as follows.

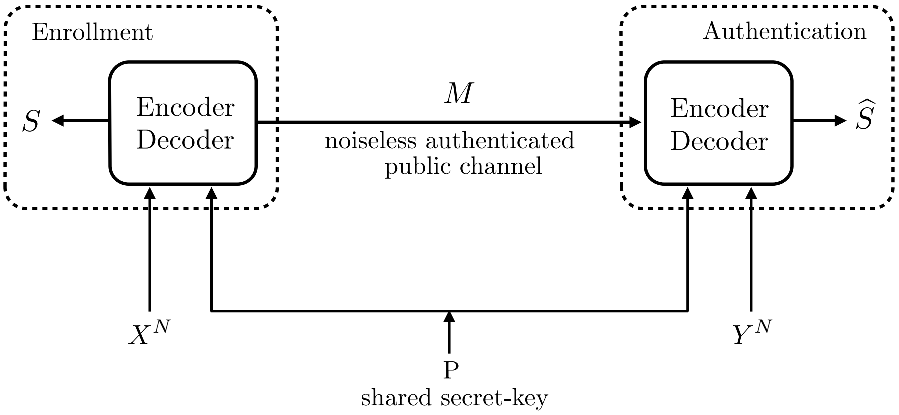


Figure 24. Model for biometric secret generation with zero leakage

Definition 3.9.5. *Let $R \in \mathbb{R}^+$. Let $\mathcal{S}$ be an alphabet of size 2^{NR}. Assume that the encoder and decoder share a uniformly distributed secret-key P beforehand. The protocol defined by the following steps is called a $(2^{NR}, N, R)$ biometric secret generation strategy with zero leakage.*

- *The encoder observes the enrollment sequence $X^{1:N}$;*
- *The encoder generates a secret $S \in \mathcal{S}$ from $X^{1:N}$ and P;*
- *The encoder transmits publicly to the decoder helper data M which is a function of $X^{1:N}$ and P;*
- *The decoder observes the authentication sequence $Y^{1:N}$, and computes $\widehat{S} \in \mathcal{S}$ from $Y^{1:N}$ and P.*

The performance of a biometric secret generation strategy with zero leakage is measured in terms of

- the average probability of error between the biometric secrets with $\mathbf{P}_e(\mathcal{S}_N^{\text{bio}}) \triangleq \mathbb{P}[S \neq \widehat{S}]$,
- the information leakage of M on S and $X^{1:N}$ with $\mathbf{L}_{\text{c}}(\mathcal{S}_N^{\text{bio}}) \triangleq I(SX^{1:N}; M)$ (conditional case), or $\mathbf{L}_{\text{u}}(\mathcal{S}_N^{\text{bio}}) \triangleq I(S; M) + I(X^{1:N}; M)$ (unconditional case),

- the length of the secret-key P with $\mathbf{H}(\mathcal{S}_N^{\text{bioZ}}) \triangleq |P|{-}H(P)$,

- the uniformity of the biometric secret $\mathbf{U}(\mathcal{S}_N^{\text{bio}}) \triangleq \log\lceil 2^{NR}\rceil - H(S)$.

Definition 3.9.6. *For a fixed secret-key length K, a biometric secret rate R is achievable with zero leakage if there exists a sequence of $(2^{NR}, N, R)$ biometric secret generation strategies with zero leakage $\left\{\mathcal{S}_N^{\text{bioZ}}\right\}_{N\geqslant 1}$ such that*

$$\begin{aligned}
\lim_{N\to\infty} \mathbf{P}_e(\mathcal{S}_N^{\text{bioZ}}) =&0, \ \textit{(reliability)}\\
\lim_{N\to\infty} \mathbf{L}_{\text{c}}(\mathcal{S}_N^{\text{bioZ}}) =&0, \ \textit{(strong secrecy)}\\
\lim_{N\to\infty} \mathbf{H}(\mathcal{S}_N^{\text{bioZ}})/N \leqslant&K, \ \textit{(secret-key length)}\\
\lim_{N\to\infty} \mathbf{U}(\mathcal{S}_N^{\text{bioZ}}) =&0. \ \textit{(uniformity)}
\end{aligned}$$

Moreover, the supremum of achievable rates is called the zero-leakage biometric secret capacity and is denoted $C^{\text{c}}_{BioZ}(L)$. For the unconditional case $\mathbf{P}_{\text{c}}(\mathcal{S}_N^{\text{bioZ}})$ is replaced with $\mathbf{P}_{\text{u}}(\mathcal{S}_N^{\text{bioZ}})$, and the zero-leakage biometric secret capacity and is denoted $C^{\text{u}}_{BioZ}(L)$.

Note that we require a stronger security metric than in [65]. The zero-leakage biometric secret capacities are known and recalled below.

Theorem 3.9.10 ([65])**.** *Let $(\mathcal{X}\mathcal{Y}, p_{XY})$ be a BMS and $K \in \mathbb{R}_+$ be a fixed length. The conditional and unconditional zero-leakage biometric secret capacities are equal $C^{\text{c}}_{BioZ}(K) = C^{\text{u}}_{BioZ}(K))$, moreover,*

$$C^{\text{c}}_{BioZ}(L) = \max_U I(Y;U) + K$$

subject to

$$\begin{aligned}
&K = I(U;X) - I(U;Y),\\
&U \to X \to Y,\\
&|\mathcal{U}|\leqslant |\mathcal{X}|.
\end{aligned}$$

Remark 3.9.7. *The equality $K = I(U;X) - I(U;Y)$ and the range constraint $|\mathcal{U}|\leqslant |\mathcal{X}|$ are obtained from Proposition 2.5.3.*

3.9.2 Polar coding scheme for generated-secret systems

Let $n \in \mathbb{N}$ and $N \triangleq 2^n$. Fix a joint probability distribution p_{XU}. We note $V^{1:N} \triangleq U^{1:N} G_N$. For $\delta_N \triangleq 2^{-N^\beta}$, where $\beta \in]0, 1/2[$, define the following sets

$$\begin{aligned}
\mathcal{H}_U &\triangleq \left\{ i \in [\![1, N]\!] : H\left(V^i | V^{1:i-1}\right) \geqslant \delta_N \right\}, \\
\mathcal{V}_U &\triangleq \left\{ i \in [\![1, N]\!] : H\left(V^i | V^{1:i-1}\right) \geqslant 1 - \delta_N \right\}, \\
\mathcal{V}_{U|X} &\triangleq \left\{ i \in [\![1, N]\!] : H\left(V^i | V^{1:i-1} X^{1:N}\right) \geqslant 1 - \delta_N \right\}, \\
\mathcal{H}_{U|Y} &\triangleq \left\{ i \in [\![1, N]\!] : H\left(V^i | V^{1:i-1} Y^{1:N}\right) \geqslant \delta_N \right\}, \\
\mathcal{H}_{U|X} &\triangleq \left\{ i \in [\![1, N]\!] : H\left(V^i | V^{1:i-1} X^{1:N}\right) \geqslant \delta_N \right\}.
\end{aligned}$$

The scheme proposed is a special case (it corresponds to the case $Z = \emptyset$) of the scheme in Section 3.6.2. However, for completeness and clarity, we provide its detailed description in Algorithm 12 and Algorithm 13 with the notation of the biometric secret generation problem. We formally define a biometric key generation strategy $\mathcal{S}_N^{\text{bio}}$ as follows.

Remark 3.9.8. *One may actually use $S_k^{1:N}[\mathcal{V}_U \backslash \mathcal{H}_{U|Y}]$ as the S_k and slightly increase the biometric secret rate in Algorithm 12. However, one does not distinguish the last block from the others for convenience – see Remark 3.5.1.*

Based on the results established for Model 2 in Section 3.6, we obtain the following.

Theorem 3.9.11. *Consider a BMS $(\mathcal{XY}, p_{XY})$. Assume that the encoder and the decoder share a secret seed. For any $L \in \mathbb{R}$, the biometric secret capacities $C_{Bio}^{\text{c}}(L)$, and $C_{Bio}^{\text{u}}(L)$, are achieved by the polar coding scheme of Algorithm 12 and Algorithms 13, which involves a chaining of k blocks of size N, and whose complexity is $O(kN \log N)$. Moreover, the seed rate is in $o\left(2^{-N^\alpha}\right)$, $\alpha < 1/2$.*

Theorem 3.9.11 is a direct consequence of Theorem 3.6.5 for the particular case $Z = \emptyset$, since

$$\max(\mathbf{P}_{\text{c}}(\mathcal{S}_N^{\text{bio}}), \mathbf{P}_{\text{u}}(\mathcal{S}_N^{\text{bio}})) \leqslant H(M).$$

Algorithm 12: Encoding algorithm for generated secret systems

```
Require: $\widetilde{S}_0$, a secret key of size $|(\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X})\backslash\mathcal{V}_U|$; $\mathcal{A}_{UXY}$ be any subset of
         $\mathcal{V}_U\backslash\mathcal{H}_{U|Y}$ with size $|(\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X})\backslash\mathcal{V}_U|$; Observations $X_i^{1:N}$ in every
         block $i \in [\![1,k]\!]$; a vector $R_1$ of uniformly distributed bits with size
         $|\mathcal{V}_{U|X}|$.
Transmit $R_1$ publicly.
for Block $i = 1$ to $k$ do
    $\widetilde{V}_i^{1:N}[\mathcal{V}_{U|X}] \leftarrow R_1$
    Given $X_i^{1:N}$, successively draw the remaining bits of $\widetilde{V}_i^{1:N}$ according to
    $\widetilde{p}_{V_i^{1:N}X_i^{1:N}} \triangleq \prod_{j=1}^{N} \widetilde{p}_{V_i^j|V_i^{j-1}X^{1:N}} p_{X^{1:N}}$ with

        $\widetilde{p}_{V_i^j|V_i^{1:j-1}X^{1:N}}(v^j|\widetilde{V}_i^{1:j-1}X_i^{1:N}) \triangleq \begin{cases} p_{V^j|V^{1:j-1}X^{1:N}}(v^j|\widetilde{V}_i^{1:j-1}X_i^{1:N}) & \text{if } i \in \mathcal{H}_U\backslash\mathcal{V}_{U|X} \\ p_{V^j|V^{1:j-1}}(v^j|\widetilde{V}_i^{1:j-1}) & \text{if } i \in \mathcal{H}_U^c \end{cases}$    (64)

    $\widetilde{S}_i \leftarrow \widetilde{V}_i^{1:N}[\mathcal{A}_{UXY}]$
    $S_i \leftarrow \widetilde{V}_i^{1:N}[(\mathcal{V}_U\backslash\mathcal{H}_{U|Y})\backslash\mathcal{A}_{UXY}]$
    $F_i \leftarrow \widetilde{V}_i^{1:N}[(\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X}) \cap \mathcal{V}_U]$
    $F'_i \leftarrow \widetilde{V}_i^{1:N}[(\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X})\backslash\mathcal{V}_U]$
    Transmit $M_i \leftarrow [F_i, F'_i \oplus \widetilde{S}_{i-1}, R_1]$ publicly
end
return : $S_{1:k} \leftarrow [S_1, S_2, \ldots, S_k]$
```

Algorithm 13: Decoding algorithm for generated secret systems

```
Require: The secret-key $\widetilde{S}_0$, and the set $\mathcal{A}_{UXY}$ defined in Algorithm 12;
         Observations $Y_i^{1:N}$ and message $M_i$ transmitted by other party in
         every block $i \in [\![1,k]\!]$, vector $R_1$.
for Block $i = 1$ to $k$ do
    Form $\widetilde{V}_i^{1:N}[\mathcal{H}_{U|Y}]$ from $(F_i, F'_i) = \widetilde{V}_i^{1:N}[\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X}]$ and
    $R_1 = R_i = \widetilde{V}_i^{1:N}[\mathcal{V}_{U|X}]$.
    Create estimate $\widehat{V}_i^{1:N}$ of $\widetilde{V}_i^{1:N}$ with the successive cancellation decoder
    of [87]
    $\widehat{S}_i \leftarrow \widehat{V}_i^{1:N}[(\mathcal{V}_U\backslash\mathcal{H}_{U|Y})\backslash\mathcal{A}_{UXY}]$
    $\widetilde{S}_i \leftarrow \widehat{V}_i^{1:N}[\mathcal{A}_{UXY}]$
end
return : $\widehat{S}_{1:k} \triangleq [\widehat{S}_1, \widehat{S}_2, \ldots, \widehat{S}_k]$.
```

Note also that for $i \in [\![0, k-1]\!]$, $\widetilde{S}_i = o(N)$.

3.9.3 Polar coding scheme for generated-secret systems with zero leakage

The encoding and decoding algorithms are given in Algorithm 14 and Algorithm 15. The difference with the scheme of Section 3.9.2 is that the public communication is protected with a secret-key shared by the encoder and the decoder.

Algorithm 14: Encoding algorithm for generated secret systems with zero leakage

```
Require: $k$ secret keys $\{P_i\}_{i \in [\![1,k]\!]}$ of size $|\mathcal{H}_{U|Y} \backslash \mathcal{V}_{U|X}|$; observations $X_i^{1:N}$ in
         every block $i \in [\![1,k]\!]$; a vector $R_1$ of uniformly distributed bits with
         size $|\mathcal{V}_{U|X}|$.
Transmit $R_1$ publicly.
for Block $i = 1$ to $k$ do
    $\widetilde{V}_i^{1:N}[\mathcal{V}_{U|X}] \leftarrow R_1$
    Given observations $X_i^{1:N}$, successively draw the remaining bits of $\widetilde{V}_i^{1:N}$
    according to $\widetilde{p}_{VX}$ defined by (64).
    $F_i \leftarrow \widetilde{V}_i^{1:N}[(\mathcal{H}_{U|Y} \backslash \mathcal{V}_{U|X}) \cap \mathcal{V}_U]$
    $F'_i \leftarrow \widetilde{V}_i^{1:N}[(\mathcal{H}_{U|Y} \backslash \mathcal{V}_{U|X}) \backslash \mathcal{V}_U]$
    $S_i \leftarrow [\widetilde{V}_i^{1:N}[\mathcal{V}_U \backslash \mathcal{H}_{U|Y}], F_i]$
    Transmit $M_i \leftarrow [F_i, F'_i] \oplus P_i$ publicly
end
return : $S_{1:k} \leftarrow [S_1, S_2, \ldots, S_k]$
```

Algorithm 15: Decoding algorithm for generated secret systems with zero leakage

```
Require: the secret key $P_i$, $M_i$ transmitted by other party, observations $Y_i^{1:N}$
         in every block $i \in [\![1,k]\!]$, and vector $R_1$.
for Block $i = 1$ to $k$ do
    Form $\widetilde{V}_i^{1:N}[\mathcal{H}_{U|Y}]$ from $(F_i, F'_i) = \widetilde{V}_i^{1:N}[\mathcal{H}_{U|Y} \backslash \mathcal{V}_{U|X}]$ and
    $R_1 = R_i = \widetilde{V}_i^{1:N}[\mathcal{V}_{U|X}]$.
    Create estimate $\widehat{V}_i^{1:N}$ of $\widetilde{V}_i^{1:N}$ with the successive cancellation decoder
    of [87]
    $\widehat{S}_i \leftarrow [\widehat{V}_i^{1:N}[\mathcal{V}_U \backslash \mathcal{H}_{U|Y}], F_i]$
end
return : $\widehat{S}_{1:k} \triangleq [\widehat{S}_1, \widehat{S}_2, \ldots, \widehat{S}_k]$.
```

The performance of the algorithms is ensured by the following result.

Theorem 3.9.12. *Consider a BMS* $(\mathcal{X}\mathcal{Y}, p_{XY})$*. For any* $P \in \mathbb{R}$*, the zero-leakage biometric secret capacities* $C^{\mathrm{c}}_{BioZ}(K)$*, and* $C^{\mathrm{u}}_{BioZ}(K)$*, are achieved by the polar coding scheme of Algorithm 14 and Algorithms 15, which involves a chaining of* k *blocks of size* N*, and whose complexity is* $O(kN \log N)$*.*

Remark that one only needs to prove that $C^{\mathrm{c}}_{\mathrm{BioZ}}(K)$ is achieved in Theorem 3.9.12, since a code that achieves $C^{\mathrm{c}}_{\mathrm{BioZ}}(K)$ also achieves $C^{\mathrm{u}}_{\mathrm{BioZ}}(K)$ by [65]. The proof of Theorem 3.9.12 for $C^{\mathrm{c}}_{\mathrm{BioZ}}(K)$ is similar to the proof of Theorem 3.6.5 and is thus omitted. To show that $S_i = [\widetilde{V}_i^{1:N}[\mathcal{V}_U \backslash \mathcal{H}_{U|Y}], F_i]$, $i \in [\![1, k]\!]$, is uniform one can use Lemma 3.6.7, then, similar to Theorem 3.6.5, one can show that $S_{1:k}$ is also uniform and that strong secrecy holds. Note also that for $i \in [\![0, k-1]\!]$, $F'_i = o(N)$.

3.10 Conclusion

We have proposed low-complexity secret-key capacity-achieving schemes based on polar coding for several classes of sources. Our schemes jointly handle secrecy and reliability, which contrasts with sequential methods that successively perform reconciliation and privacy amplification. Although sequential methods apply to more general classes of sources, our polar coding schemes may be easier to design and may operate with lesser complexity. Nevertheless, the price to be paid for low complexity is that our schemes often require a pre-shared seed, whose rate is negligible compared to the blocklength. When the eavesdropper has no access to correlated observations of the source, and when the source has uniform marginals, we have identified several configurations, including multiterminal models, for which no pre-shared seed is required. Finally, we have applied our polar coding schemes to privacy and secrecy for some biometric systems.

Our polar coding schemes are particularly convenient to handle rate-limited public communication and vector quantization, which are often the major hurdle in designing optimal key-generation schemes.

APPENDICES

3.A Proofs for Model 1 in Section 3.5

3.A.1 Proof of Corollary 3.5.2

We perform the same encoding as in Section 3.5.2 for Block 1. Define the set

$$\mathcal{H}_{X|Z} \triangleq \left\{ i \in \llbracket 1, N \rrbracket : H\left(U^i | U^{1:i-1} Z^{1:N}\right) \geqslant \delta_N \right\}.$$

We have

$$\begin{aligned} |F_1'| &= |\mathcal{H}_{X|Y} \backslash \mathcal{V}_{X|Z}| \\ &\overset{(a)}{\leqslant} |\mathcal{H}_{X|Z} \backslash \mathcal{V}_{X|Z}| \\ &\overset{(b)}{=} |\mathcal{H}_{X|Z}| - |\mathcal{V}_{X|Z}|, \end{aligned}$$

where (a) holds because $\mathcal{H}_{X|Y} \subset \mathcal{H}_{X|Z}$ since we have assumed $X \to Y \to Z$, (b) holds because $\mathcal{V}_{X|Z} \subset \mathcal{H}_{X|Z}$.

We conclude by Lemma 3.4.1 and [87] that $|F_1'| = o(N)$.

3.A.2 Proof of Proposition 3.5.3

3.A.2.1 Polar coding scheme

Let $\delta > 0$, $\beta \in]0, 1/2[$. Let $n \in \mathbb{N}$ and $N \triangleq 2^n$. We set $U^{1:N} \triangleq X^{1:N} G_N$. We define for $\delta_N \triangleq 2^{-N^\beta}$, $\beta < 1/2$, the following sets

$$\mathcal{H}_{X|Y} \triangleq \left\{ i \in \llbracket 1, N \rrbracket : H\left(U^i | U^{1:i-1} Y^{1:N}\right) \geqslant \delta_N \right\},$$
$$\mathcal{H}_{X} \triangleq \left\{ i \in \llbracket 1, N \rrbracket : H\left(U^i | U^{1:i-1}\right) \geqslant \delta_N \right\}.$$

We define a key-generation strategy $\mathcal{S}_N$ as follows. Define the key as $K \triangleq U^{1:N}[\mathcal{H}_X \backslash \mathcal{H}_{X|Y}]$, and the public message as $F \triangleq U^{1:N}[\mathcal{H}_{X|Y}]$.

3.A.2.2 Scheme analysis

Observe that $\mathcal{H}_{X|Y} \subset \mathcal{H}_X$, because conditioning reduces entropy. We thus have by [87], a key rate equal to

$$\lim_{N \to +\infty} \frac{|\mathcal{H}_X \backslash \mathcal{H}_{X|Y}|}{N} = \lim_{N \to +\infty} \frac{|\mathcal{H}_X| - |\mathcal{H}_{X|Y}|}{N} = H(X) - H(X|Y) = I(X;Y).$$

Note that the key K is uniform because $X^{1:N}$ is uniform, that is

$$\mathbf{U}_e(\mathcal{S}_N) = 0.$$

Then, by [87, Theorem 3], Bob can reconstruct K from F with an error probability satisfying

$$\mathbf{P}_e(\mathcal{S}_N) \leqslant N\delta_N.$$

Finally, by the key uniformity and because $(\mathcal{H}_X \backslash \mathcal{H}_{X|Y}) \cap \mathcal{H}_{X|Y} = \emptyset$, we have

$$H(K|F) = H\left(U^{1:N}[\mathcal{H}_X \backslash \mathcal{H}_{X|Y}] | U^{1:N}[\mathcal{H}_{X|Y}]\right) = H\left(U^{1:N}[\mathcal{H}_X \backslash \mathcal{H}_{X|Y}]\right) = H(K),$$

which means that we obtain perfect secrecy, that is

$$\mathbf{L}(\mathcal{S}_N) = I(K;F) = H(K) - H(K|F) = 0.$$

3.A.3 Proof of Lemma 3.4.1

As in [87], for a pair of random variables (X,Y) distributed according to p_{XY} over $\mathcal{X} \times \mathcal{Y}$, we define the Bhattacharyya parameter as

$$Z(X|Y) = 2\sum_y p_Y(y)\sqrt{p_{X|Y}(0|y)p_{X|Y}(1|y)}.$$

We will need the following counterpart of [87, Proposition 1] that is proved using the same technique as [94, Lemma 20].

Lemma 3.1.16. *If (X_1, Y_1) and (X_2, Y_2) are two independent drawings of (X, Y), then*

$$Z\left(X_1 \oplus X_2 | Y_1^2\right) \geqslant \sqrt{2Z(X|Y)^2 - Z(X|Y)^4}.$$

Proof. We have for any v_1, $v_2 \in \mathcal{X}$, y_1, $y_2 \in \mathcal{Y}$,

$$p_{X_1 \oplus X_2, X_2, Y_1, Y_2}(v_1, v_2, y_1, y_2) = p_{XY}(v_1 + v_2, y_1)p_{XY}(v_2, y_2).$$

Hence,

$$Z\left(X_1 \oplus X_2 | Y_1^2\right) \\ = 2 \sum_{y_1, y_2} \left(\sum_{v_2} p_{XY}\left(v_2, y_1\right) p_{XY}\left(v_2, y_2\right) \cdot \sum_{v_2'} p_{XY}\left(1 + v_2', y_1\right) p_{XY}\left(v_2', y_2\right) \right)^{1/2},$$

which can be rewritten as

$$\begin{aligned} &Z\left(X_1 \oplus X_2 | Y_1^2\right) \\ &= \frac{1}{2} Z\left(X_1 | Y_1\right) Z\left(X_2 | Y_2\right) \\ &\quad \times \sum_{y_1, y_2} P_1\left(y_1\right) P_2\left(y_2\right) \sqrt{A\left(y_1\right)^2 + A\left(y_2\right)^2 - 4}, \end{aligned}$$

where, for $i \in [\![1, 2]\!]$,

$$P_i\left(y_i\right) \triangleq \frac{2\sqrt{p_{XY}\left(0, y_i\right) p_{XY}\left(1, y_i\right)}}{Z\left(X_i | Y_i\right)}$$

and

$$A\left(y_i\right) \triangleq \sqrt{\frac{p_{XY}\left(0, y_i\right)}{p_{XY}\left(1, y_i\right)}} + \sqrt{\frac{p_{XY}\left(1, y_i\right)}{p_{XY}\left(0, y_i\right)}}.$$

As observed in [94, Lemma 20], for $i \in [\![1, 2]\!]$, $A\left(y_i\right)^2 \geqslant 4$, by the arithmetic-geometric inequality, and $x \mapsto \sqrt{x^2 + a}$ is convex for $a > 0$. Hence, since for $i \in [\![1, 2]\!]$, P_i defines a probability distribution over $\mathcal{Y}$, by Jensen's inequality applied twice

$$\begin{aligned} &Z\left(X_1 \oplus X_2 | Y_1^2\right) \\ &\geqslant \frac{1}{2} Z\left(X_1 | Y_1\right) Z\left(X_2 | Y_2\right) \\ &\quad \times \sqrt{\left(\mathbb{E}_{P_1}\left[A\left(y_1\right)\right]\right)^2 + \left(\mathbb{E}_{P_2}\left[A\left(y_2\right)\right]\right)^2 - 4}. \end{aligned}$$

We conclude by substituting $\mathbb{E}_{P_i}\left[A\left(y_i\right)\right] = \frac{2}{Z\left(X_i | Y_i\right)}$, for $i \in [\![1, 2]\!]$. □

Let $\alpha \in]\beta, 1/2[$. Define the sets

$$\mathcal{F}_{X|Z} \triangleq \left\{ i \in [\![1, N]\!] : Z\left(U^i | U^{1:i-1} Z^{1:N}\right) \geqslant 1 - 2^{-N^\alpha} \right\},$$

$$\mathcal{H}_{X|Z} \triangleq \left\{ i \in [\![1, N]\!] : H\left(U^i | U^{1:i-1} Z^{1:N}\right) \geqslant \delta_N \right\}.$$

Similar to [94, Theorem 19], which relies on the result in [101], we can show with Lemma 3.1.16

$$\lim_{N\to+\infty}|\mathcal{F}_{X|Z}|/N = H(X|Z).$$

But, by [87, Proposition 2], for N large enough, $|\mathcal{F}_{X|Z}|\leqslant |\mathcal{V}_{X|Z}|$, hence, $\lim_{N\to+\infty}|\mathcal{V}_{X|Z}|/N \geqslant H(X|Z)$. Since we also have $\lim_{N\to+\infty}|\mathcal{H}_{X|Z}|/N = H(X|Z)$, by [87], and $\mathcal{V}_{X|Z} \subset \mathcal{H}_{X|Z}$, we conclude

$$\lim_{N\to+\infty}|\mathcal{V}_{X|Z}|/N = H(X|Z).$$

3.A.4 Proof of Lemma 3.5.4

Let $i \in [\![1,k]\!]$, we note $q_{\mathcal{U}_{K,\widetilde{K}}}$ the uniform distribution over $[\![1, 2^{|K_i|+|\widetilde{K}_i|}]\!]$. We have,

$$\begin{aligned}
\mathbb{V}\left(p_{K_i\widetilde{K}_i}, p_{K_i}p_{\widetilde{K}_i}\right) &\overset{(a)}{\leqslant} \mathbb{V}\left(p_{K_i\widetilde{K}_i}, q_{\mathcal{U}_{K,\widetilde{K}}}\right) + \mathbb{V}\left(q_{\mathcal{U}_{K,\widetilde{K}}}, p_{K_i}p_{\widetilde{K}_i}\right),\\
&\overset{(b)}{\leqslant} \mathbb{V}\left(p_{K_i\widetilde{K}_i}, q_{\mathcal{U}_{K,\widetilde{K}}}\right) + \mathbb{V}\left(q_{\mathcal{U}_{K,\widetilde{K}}}, q_{\mathcal{U}_K}p_{\widetilde{K}_i}\right) + \mathbb{V}\left(q_{\mathcal{U}_K}p_{\widetilde{K}_i}, p_{K_i}p_{\widetilde{K}_i}\right)\\
&= \mathbb{V}\left(p_{K_i\widetilde{K}_i}, q_{\mathcal{U}_{K,\widetilde{K}}}\right) + \mathbb{V}\left(q_{\mathcal{U}_{\widetilde{K}}}, p_{\widetilde{K}_i}\right) + \mathbb{V}\left(q_{\mathcal{U}_K}, p_{K_i}\right)\\
&\leqslant 3\mathbb{V}\left(p_{K_i\widetilde{K}_i}, q_{\mathcal{U}_{K,\widetilde{K}}}\right)\\
&\overset{(c)}{\leqslant} 3\sqrt{2N\log 2} \times 2^{-N^\beta/2},
\end{aligned} \tag{65}$$

where (a), (b) hold by the triangle inequality, (c) holds by Pinsker's inequality and Lemma 3.5.2.

Then, for N large enough ($|\tilde{K}|> 4$), we have

$$\begin{aligned}
I(K_i;\widetilde{K}_i) &\leqslant \mathbb{V}(p_{K_i\widetilde{K}_i}, p_{K_i}p_{\widetilde{K}_i}) \log_2 \frac{|\tilde{K}|}{\mathbb{V}(p_{K_i\widetilde{K}_i}, p_{K_i}p_{\widetilde{K}_i})}\\
&\leqslant \mathbb{V}(p_{K_i\widetilde{K}_i}, p_{K_i}p_{\widetilde{K}_i}) \log_2|\tilde{K}| - \mathbb{V}(p_{K_i\widetilde{K}_i}, p_{K_i}p_{\widetilde{K}_i}) \log_2 \mathbb{V}(p_{K_i\widetilde{K}_i}, p_{K_i}p_{\widetilde{K}_i})\\
&\leqslant \delta_N^*,
\end{aligned}$$

where $\delta_N^* \triangleq -3\sqrt{2N\log 2} \times N2^{-N^\beta/2} \log_2\left(3\sqrt{2N\log 2} \times 2^{-N^\beta/2}\right)$ by (65) and because $x \mapsto x\log x$ is decreasing for $x > 0$ small enough.

3.A.5 Proof of Lemma 3.5.5

Let $i \in [\![2,k]\!]$. By applying the chain rule of mutual information repeatedly, we obtain

$$\widetilde{L}_e^{1:i} = \alpha_i + \beta_i + \gamma_i, \tag{66}$$

where

$$\begin{aligned}
\alpha_i &\triangleq I\left(K_i\widetilde{K}_i; M_i Z_i^{1:N}\right),\\
\beta_i &\triangleq I\left(K_{1:i-1}; Z_i^{1:N} M_i | K_i\widetilde{K}_i\right),\\
\gamma_i &\triangleq I\left(K_{1:i}\widetilde{K}_i; Z_{1:i-1}^{1:N} M_{1:i-1} | Z_i^{1:N} M_i\right).
\end{aligned}$$

Then, note that

$$\begin{aligned}
\gamma_i &\leqslant I\left(K_{1:i}\widetilde{K}_{i-1:i} Z_i^{1:N} M_i; Z_{1:i-1}^{1:N} M_{1:i-1}\right)\\
&= I\left(K_{1:i-1}\widetilde{K}_{i-1}; Z_{1:i-1}^{1:N} M_{1:i-1}\right) + I\left(K_i\widetilde{K}_i Z_i^{1:N} M_i; Z_{1:i-1}^{1:N} M_{1:i-1} | K_{1:i-1}\widetilde{K}_{i-1}\right)\\
&= \widetilde{L}_e^{1:i-1},
\end{aligned} \tag{67}$$

where the last equality follows from $K_i\widetilde{K}_i Z_i^{1:N} M_i \to K_{1:i-1}\widetilde{K}_{i-1} \to Z_{1:i-1}^{1:N} M_{1:i-1}$.

We also have,

$$\begin{aligned}
\beta_k &\leqslant I\left(K_{1:i-1}; Z_i^{1:N} M_i \widetilde{K}_{i-1} | K_i\widetilde{K}_i\right)\\
&= I\left(K_{1:i-1}; \widetilde{K}_{i-1} | K_i\widetilde{K}_i\right) + I\left(K_{1:i-1}; Z_i^{1:N} M_i | K_i\widetilde{K}_{i-1:i}\right)\\
&\overset{(a)}{=} I\left(K_{1:i-1}; \widetilde{K}_{i-1} | K_i\widetilde{K}_i\right)\\
&\overset{(b)}{\leqslant} I\left(K_{1:i-1}; \widetilde{K}_{i-1}\right)\\
&= I\left(K_{i-1}; \widetilde{K}_{i-1}\right) + I\left(K_{1:i-2}; \widetilde{K}_{i-1} | K_{i-1}\right)\\
&\leqslant I\left(K_{i-1}; \widetilde{K}_{i-1}\right) + I\left(K_{1:i-2}; \widetilde{K}_{i-1} K_{i-1}\right)\\
&\leqslant I\left(K_{i-1}; \widetilde{K}_{i-1}\right) + I\left(X_{1:i-2}^{1:N}; X_{i-1}^{1:N}\right)\\
&\overset{(c)}{=} I\left(K_{i-1}; \widetilde{K}_{i-1}\right)
\end{aligned} \tag{68}$$

where (a) holds by $K_{1:i-1} \to K_i\widetilde{K}_{i-1:i} \to Z_i^{1:N}M_i$, (b) holds by $K_{1:i-1} \to \widetilde{K}_{i-1} \to K_i\widetilde{K}_i$, (c) holds by independence between $X_{1:i-2}^{1:N}$ and $X_{i-1}^{1:N}$.

Finally, we conclude combining (66), (67), and (68).

3.B Proofs for Model 2 in Section 3.6

3.B.1 Proof of Corollary 3.6.4

We perform the same encoding as in Section 3.6.2 for Block 1. Note that $C_{\text{WSK}}(R_p)$ is obtained for U uniform by Proposition 2.5.4 since X is uniform and the tests-channel $p_{Y|X}$ and $p_{Z|X}$ are uniform. Hence, the rate R_1 of randomness to perform successive cancellation encoding can be set equal to zero by [94]. We also have

$$\begin{aligned}|F_1'| &= |(\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|X})\backslash\mathcal{V}_{U|Z}|\\ &\overset{(a)}{\leqslant} |\mathcal{H}_{U|Z}\backslash\mathcal{V}_{U|Z}|\\ &\overset{(b)}{=} |\mathcal{H}_{U|Z}| - |\mathcal{V}_{U|Z}|,\end{aligned}$$

where (a) holds because $\mathcal{H}_{U|Y} \subset \mathcal{H}_{U|Z}$ since we have assumed $X \to Y \to Z$, (b) holds because $\mathcal{V}_{U|Z} \subset \mathcal{H}_{U|Z}$.

We conclude by Lemma 3.4.1 and [87] that $|F_1'|= o(N)$.

3.B.2 Proof of Lemma 3.6.6

Using the notation of [73] for conditional relative entropy, we have for $i \in [\![1, k]\!]$

$$
\begin{aligned}
&\mathbb{D}(p_{X^{1:N}U^{1:N}}||\widetilde{p}_{X_i^{1:N}U_i^{1:N}})\\
&\overset{(a)}{=} \mathbb{D}(p_{X^{1:N}V^{1:N}}||\widetilde{p}_{X_i^{1:N}V_i^{1:N}})\\
&\overset{(b)}{=} \mathbb{D}(p_{V^{1:N}|X^{1:N}}||\widetilde{p}_{V_i^{1:N}|X^{1:N}})\\
&\overset{(c)}{=} \sum_{j=1}^{N} \mathbb{D}(p_{V^j|V^{1:j-1}X^{1:N}}||\widetilde{p}_{V_i^j|V_i^{1:j-1}X^{1:N}})\\
&\overset{(d)}{=} \sum_{j\in\mathcal{V}_{U|X}} \sum_{j\in\mathcal{H}_U^c} \mathbb{D}(p_{V^j|V^{1:j-1}X^{1:N}}||\widetilde{p}_{V_i^j|V_i^{1:j-1}X^{1:N}})\\
&\overset{(e)}{=} \sum_{j\in\mathcal{V}_{U|X}} (1 - H(V^j|V^{1:j-1}X^{1:N}))\\
&\qquad + \sum_{j\in\mathcal{H}_U^c} (H(V^j|V^{1:j-1}) - H(V^j|V^{1:j-1}X^{1:N}))\\
&\leqslant |\mathcal{V}_{U|X}|\delta_N + \sum_{j\in\mathcal{H}_U^c} H(V^j|V^{1:j-1})\\
&\leqslant |\mathcal{V}_{U|X}|\delta_N + |\mathcal{H}_U^c|\delta_N\\
&\leqslant N\delta_N,
\end{aligned}
$$

where (a) holds by invertibility of G_n, (b) and (c) hold by the chain rule for divergence, (d) and (e) hold by (51) and by uniformity of the components of $\widetilde{V}_i^{1:N}$ in $\mathcal{V}_{U|X}$.

3.B.3 Proof of Lemma 3.6.8

We have by [102]

$$
\begin{aligned}
|K_i|+|\widetilde{K}_i|-H(K_i\widetilde{K}_i) &\leqslant \mathbb{V}(p_{K_i\widetilde{K}_i}, q_{\mathcal{U}_{K,\widetilde{K}}}) \log_2 \frac{|K_i|+|\widetilde{K}_i|}{\mathbb{V}(p_{K_i\widetilde{K}_i}, q_{\mathcal{U}_{K,\widetilde{K}}})}\\
&\leqslant N\mathbb{V}(p_{K_i\widetilde{K}_i}, q_{\mathcal{U}_{K,\widetilde{K}}}) - \mathbb{V}(p_{K_i\widetilde{K}_i}, q_{\mathcal{U}_{K,\widetilde{K}}}) \log_2 \mathbb{V}(p_{K_i\widetilde{K}_i}, q_{\mathcal{U}_{K,\widetilde{K}}})\\
&\leqslant 2\sqrt{2\log 2}\sqrt{N\delta_N}(N - \log_2(2\sqrt{2\log 2}\sqrt{N\delta_N})),
\end{aligned}
$$

where the last inequality holds for N large enough by Lemma 3.6.7 and because $x \mapsto x\log x$ is decreasing for $x > 0$ small enough.

3.B.4 Proof of Lemma 3.6.9

We only prove the first inequality, the other is obtained similarly. Let $i \in [\![1,k]\!]$. We have,

$$
\begin{aligned}
&\mathbb{V}\left(p_{K_i\widetilde{K}_iR_1}, p_{K_i}p_{\widetilde{K}_iR_1}\right)\\
&\overset{(a)}{\leqslant} \mathbb{V}\left(p_{K_i\widetilde{K}_iR_1}, q_{\mathcal{U}_{K,\widetilde{K},R}}\right) + \mathbb{V}\left(q_{\mathcal{U}_{K,\widetilde{K},R}}, p_{K_i}p_{\widetilde{K}_iR_1}\right),\\
&\overset{(b)}{\leqslant} \mathbb{V}\left(p_{K_i\widetilde{K}_iR_1}, q_{\mathcal{U}_{K,\widetilde{K},R}}\right) + \mathbb{V}\left(q_{\mathcal{U}_{K,\widetilde{K},R}}, q_{\mathcal{U}_K}p_{\widetilde{K}_iR_1}\right) + \mathbb{V}\left(q_{\mathcal{U}_K}p_{\widetilde{K}_iR_1}, p_{K_i}p_{\widetilde{K}_iR_1}\right)\\
&= \mathbb{V}\left(p_{K_i\widetilde{K}_iR_1}, q_{\mathcal{U}_{K,\widetilde{K},R}}\right) + \mathbb{V}\left(q_{\mathcal{U}_{\widetilde{K},R}}, p_{\widetilde{K}_iR_1}\right) + \mathbb{V}\left(q_{\mathcal{U}_K}, p_{K_i}\right)\\
&\leqslant 3\mathbb{V}\left(p_{K_i\widetilde{K}_iR_1}, q_{\mathcal{U}_{K,\widetilde{K},R}}\right)\\
&\overset{(c)}{\leqslant} 6\sqrt{2\log 2}\sqrt{N\delta_N}, \qquad (69)
\end{aligned}
$$

where (a), (b) hold by the triangle inequality, (c) holds by Pinsker's inequality and Lemma 3.6.7.

Then, for N large enough ($|\tilde{K}| > 4$), we have

$$
\begin{aligned}
&I(K_i;\widetilde{K}_iR_1)\\
&\leqslant \mathbb{V}(p_{K_i\widetilde{K}_iR_1}, p_{K_i}p_{\widetilde{K}_iR_1}) \log_2 \frac{|\tilde{K}|}{\mathbb{V}(p_{K_i\widetilde{K}_iR_1}, p_{K_i}p_{\widetilde{K}_iR_1})}\\
&\leqslant N\mathbb{V}(p_{K_i\widetilde{K}_iR_1}, p_{K_i}p_{\widetilde{K}_iR_1}) - \mathbb{V}(p_{K_i\widetilde{K}_iR_1}, p_{K_i}p_{\widetilde{K}_iR_1}) \log_2 \mathbb{V}(p_{K_i\widetilde{K}_iR_1}, p_{K_i}p_{\widetilde{K}_iR_1})\\
&\leqslant \delta_N^{(2)},
\end{aligned}
$$

where $\delta_N^{(2)} \triangleq 6\sqrt{2\log 2}\sqrt{N\delta_N}(N - \log_2(6\sqrt{2\log 2}\sqrt{N\delta_N}))$ by (69) and because $x \mapsto x\log x$ is decreasing for $x > 0$ small enough.

3.B.5 Proof of Lemma 3.6.10

We have

$$\begin{aligned}
&\mathbb{V}(\widetilde{p}_{V_i^{1:N}[\mathcal{V}_{U|Z}]Z_i^{1:N}}, p_{V_i^{1:N}[\mathcal{V}_{U|Z}]Z_i^{1:N}}) \\
&\leqslant \mathbb{V}(\widetilde{p}_{V_i^{1:N}X_i^{1:N}Z_i^{1:N}}, p_{V_i^{1:N}X_i^{1:N}Z_i^{1:N}}) \\
&= \mathbb{V}(\widetilde{p}_{Z_i^{1:N}|V_i^{1:N}X_i^{1:N}}\widetilde{p}_{V_i^{1:N}X_i^{1:N}}, p_{Z_i^{1:N}|V_i^{1:N}X_i^{1:N}}p_{V_i^{1:N}X_i^{1:N}}) \\
&= \mathbb{V}(\widetilde{p}_{Z_i^{1:N}|X_i^{1:N}}\widetilde{p}_{V_i^{1:N}X_i^{1:N}}, p_{Z_i^{1:N}|X_i^{1:N}}p_{V_i^{1:N}X_i^{1:N}}) \\
&= \mathbb{V}(\widetilde{p}_{V_i^{1:N}X_i^{1:N}}, p_{V_i^{1:N}X_i^{1:N}}) \\
&\leqslant \sqrt{2\log 2}\sqrt{N\delta_N}, \qquad (70)
\end{aligned}$$

where the last inequality follows by Lemma 3.6.6, and

$$\begin{aligned}
&\mathbb{V}(p_{V_i^{1:N}[\mathcal{V}_{U|Z}]Z_i^{1:N}}, \widetilde{p}_{V_i^{1:N}[\mathcal{V}_{U|Z}]}p_{Z_i^{1:N}}) \\
&\leqslant \mathbb{V}(p_{V_i^{1:N}[\mathcal{V}_{U|Z}]Z_i^{1:N}}, p_{V_i^{1:N}[\mathcal{V}_{U|Z}]}p_{Z_i^{1:N}}) + \mathbb{V}(p_{V_i^{1:N}[\mathcal{V}_{U|Z}]}p_{Z_i^{1:N}}, \widetilde{p}_{V_i^{1:N}[\mathcal{V}_{U|Z}]}p_{Z_i^{1:N}}) \\
&\overset{(a)}{\leqslant} \mathbb{V}(p_{V_i^{1:N}[\mathcal{V}_{U|Z}]Z_i^{1:N}}, p_{V_i^{1:N}[\mathcal{V}_{U|Z}]}p_{Z_i^{1:N}}) + \sqrt{2\log 2}\sqrt{N\delta_N} \\
&\overset{(b)}{\leqslant} \sqrt{2\log 2}\sqrt{\mathbb{D}(p_{V_i^{1:N}[\mathcal{V}_{U|Z}]Z_i^{1:N}}||p_{V_i^{1:N}[\mathcal{V}_{U|Z}]}p_{Z_i^{1:N}})} + \sqrt{2\log 2}\sqrt{N\delta_N} \\
&= \sqrt{2\log 2}\sqrt{I(V_i^{1:N}[\mathcal{V}_{U|Z}]; Z_i^{1:N})} + \sqrt{2\log 2}\sqrt{N\delta_N} \\
&\leqslant \sqrt{2\log 2}\sqrt{|\mathcal{V}_{U|Z}| - H(V_i^{1:N}[\mathcal{V}_{U|Z}]|Z_i^{1:N})} + \sqrt{2\log 2}\sqrt{N\delta_N} \\
&\overset{(c)}{\leqslant} 2\sqrt{2\log 2}\sqrt{N\delta_N}, \qquad (71)
\end{aligned}$$

where (a) holds by (70), (b) holds by Pinsker's inequality, (c) holds because similar to the proof of Lemma 3.5.3 $|\mathcal{V}_{U|Z}| - H(V_i^{1:N}[\mathcal{V}_{U|Z}]|Z_i^{1:N}) \leqslant N\delta_N$.

Hence, by (70) and (128)

$$\begin{aligned}
&\mathbb{V}(\widetilde{p}_{V_i^{1:N}[\mathcal{V}_{U|Z}]Z_i^{1:N}}, \widetilde{p}_{V_i^{1:N}[\mathcal{V}_{U|Z}]}p_{Z_i^{1:N}}) \\
&\leqslant \mathbb{V}(\widetilde{p}_{V_i^{1:N}[\mathcal{V}_{U|Z}]Z_i^{1:N}}, p_{V_i^{1:N}[\mathcal{V}_{U|Z}]Z_i^{1:N}}) + \mathbb{V}(p_{V_i^{1:N}[\mathcal{V}_{U|Z}]Z_i^{1:N}}, \widetilde{p}_{V_i^{1:N}[\mathcal{V}_{U|Z}]}p_{Z_i^{1:N}}) \\
&\leqslant 3\sqrt{2\log 2}\sqrt{N\delta_N}, \qquad (72)
\end{aligned}$$

and for N large enough

$$\begin{aligned}
&I(\widetilde{V}_i^{1:N}[\mathcal{V}_{U|Z}]; Z_i^{1:N}) \\
&\leqslant \mathbb{V}(\widetilde{p}_{V_i^{1:N}[\mathcal{V}_{U|Z}]Z_i^{1:N}}, \widetilde{p}_{V_i^{1:N}[\mathcal{V}_{U|Z}]}p_{Z_i^{1:N}}) \log_2 \frac{|\mathcal{V}_{U|Z}|}{\mathbb{V}(\widetilde{p}_{V_i^{1:N}[\mathcal{V}_{U|Z}]Z_i^{1:N}}, \widetilde{p}_{V_i^{1:N}[\mathcal{V}_{U|Z}]}p_{Z_i^{1:N}})} \\
&\leqslant 3\sqrt{2\log 2}\sqrt{N\delta_N}(N - \log_2(3\sqrt{2\log 2}\sqrt{N\delta_N})).
\end{aligned}$$

3.B.6 Proof of Lemma 3.6.12

Let $i \in [\![2, k]\!]$. By applying the chain rule of mutual information repeatedly, we obtain

$$\widetilde{L}_e^{1:i} = \alpha_i + \beta_i + \gamma_i, \tag{73}$$

where

$$\begin{aligned}
\alpha_i &\triangleq I\left(K_i\widetilde{K}_i; R_1 M_i Z_i^{1:N}\right), \\
\beta_i &\triangleq I\left(K_{1:i-1}; R_1 M_i Z_i^{1:N} | K_i\widetilde{K}_i\right), \\
\gamma_i &\triangleq I\left(K_{1:i}\widetilde{K}_i; M_{1:i-1}Z_{1:i-1}^{1:N} | R_1 M_i Z_i^{1:N}\right).
\end{aligned}$$

Then, note that

$$\begin{aligned}
\gamma_i &\overset{(a)}{\leqslant} I\left(K_{1:i}\widetilde{K}_{i-1:i}M_iZ_i^{1:N}; M_{1:i-1}Z_{1:i-1}^{1:N}|R_1\right) \\
&= I\left(K_{1:i-1}\widetilde{K}_{i-1}; M_{1:i-1}Z_{1:i-1}^{1:N}|R_1\right) + I\left(K_i\widetilde{K}_iZ_i^{1:N}M_i; M_{1:i-1}Z_{1:i-1}^{1:N}|R_1K_{1:i-1}\widetilde{K}_{i-1}\right) \\
&\overset{(b)}{=} I\left(K_{1:i-1}\widetilde{K}_{i-1}; M_{1:i-1}Z_{1:i-1}^{1:N}|R_1\right) \\
&\leqslant I\left(K_{1:i-1}\widetilde{K}_{i-1}; R_1M_{1:i-1}Z_{1:i-1}^{1:N}\right) \\
&\overset{(c)}{\leqslant} \widetilde{L}_e^{1:i-1},,
\end{aligned} \tag{74}$$

where (a) and (c) hold by the chain rule and positivity of mutual information, (b) holds because $K_i\widetilde{K}_iZ_i^{1:N}M_i \to R_1K_{1:i-1}\widetilde{K}_{i-1} \to M_{1:i-1}Z_{1:i-1}^{1:N}$.

We also have,

$$\begin{aligned}
\beta_i &\overset{(d)}{\leqslant} I\left(K_{1:i-1}; R_1 M_i Z_i^{1:N} \widetilde{K}_{i-1} | K_i \widetilde{K}_i\right) \\
&= I\left(K_{1:i-1}; \widetilde{K}_{i-1} R_1 | K_i \widetilde{K}_i\right) + I\left(K_{1:i-1}; M_i Z_i^{1:N} | K_i \widetilde{K}_{i-1:i} R_1\right) \\
&\overset{(e)}{=} I\left(K_{1:i-1}; \widetilde{K}_{i-1} R_1 | K_i \widetilde{K}_i\right) \\
&\overset{(f)}{\leqslant} I\left(K_{1:i-1}; \widetilde{K}_{i-1} R_1\right) \\
&= I\left(K_{1:i-1}; R_1\right) + I\left(K_{1:i-1}; \widetilde{K}_{i-1} | R_1\right) \\
&= I\left(K_{1:i-1}; R_1\right) + I\left(K_{i-1}; \widetilde{K}_{i-1} | R_1\right) + I\left(K_{1:i-2}; \widetilde{K}_{i-1} | K_{i-1} R_1\right) \\
&\overset{(g)}{=} I\left(K_{1:i-1}; R_1\right) + I\left(K_{i-1}; \widetilde{K}_{i-1} | R_1\right) \\
&\leqslant I\left(K_{1:i-1}; R_1\right) + I\left(K_{i-1}; \widetilde{K}_{i-1} R_1\right) \\
&= I\left(K_{1:i-2}; R_1 | K_{i-1}\right) + I\left(K_{i-1}; R_1\right) + I\left(K_{i-1}; \widetilde{K}_{i-1} R_1\right) \\
&\overset{(h)}{\leqslant} I\left(K_{1:i-2}; R_1\right) + I\left(K_{i-1}; R_1\right) + I\left(K_{i-1}; \widetilde{K}_{i-1} R_1\right) \\
&\overset{(i)}{\leqslant} \sum_{j=1}^{i-1} I\left(K_j; R_1\right) + I\left(K_{i-1}; \widetilde{K}_{i-1} R_1\right) \qquad (75)
\end{aligned}$$

where (d) holds by the chain rule and positivity of mutual information, (e) holds because $K_{1:i-1} \to K_i \widetilde{K}_{i-1:i} R_1 \to M_i Z_i^{1:N}$, (f) holds because $K_{1:i-1} \to \widetilde{K}_{i-1} R_1 \to K_i \widetilde{K}_i$, (g) holds because $K_{1:i-2} \to K_{i-1} R_1 \to \widetilde{K}_{i-1}$, (h) holds because $K_{1:i-2} \to R_1 \to K_{i-1}$, (i) holds by recurrence.

Finally, we conclude combining (73), (74), and (75).

3.C Proof of Theorem 3.7.7

1) Existence of $\mathcal{F}_{X_\mathcal{M}}$: The set $\mathcal{F}_{X_\mathcal{M}}$ exists because we have assumed $I(X_2; X_1) \leqslant I(X_2; X_3)$, i.e., $H(X_2|X_1) \geqslant H(X_2|X_3)$. Indeed,

$$\begin{aligned}
|\mathcal{F}_{X_2|X_1} \backslash \mathcal{F}_{X_2|X_3}| - |\bar{\mathcal{K}}_{X_\mathcal{M}}| &= |\mathcal{F}_{X_2|X_1} \backslash \mathcal{F}_{X_2|X_3}| - |\mathcal{F}_{X_2|X_3} \backslash \mathcal{F}_{X_2|X_1}| \\
&= |\mathcal{F}_{X_2|X_1}| - |\mathcal{F}_{X_2|X_3}|,
\end{aligned}$$

and $\lim_{N\to\infty}(|\mathcal{F}_{X_2|X_1}|-|\mathcal{F}_{X_2|X_3}|)/N = H(X_2|X_1)-H(X_2|X_3)$ by Lemma 3.4.1 and [87].

2) Key Rate: The key rate is

$$\begin{aligned}
&\frac{|\mathcal{K}_{X_\mathcal{M}}|+(k-1)|\mathcal{K}_{X_\mathcal{M}}\cup\mathcal{F}_{X_\mathcal{M}}|}{kN}\\
&\overset{(a)}{=}\frac{|\mathcal{K}_{X_\mathcal{M}}|+(k-1)(|\mathcal{K}_{X_\mathcal{M}}|+|\mathcal{F}_{X_\mathcal{M}}|)}{kN}\\
&=\frac{|\mathcal{K}_{X_\mathcal{M}}|+|\mathcal{F}_{X_\mathcal{M}}|}{N}-\frac{|\mathcal{F}_{X_\mathcal{M}}|}{kN}\\
&=\frac{|\mathcal{K}_{X_\mathcal{M}}|+|\bar{\mathcal{K}}_{X_\mathcal{M}}|}{N}-\frac{|\bar{\mathcal{K}}_{X_\mathcal{M}}|}{kN}\\
&=\frac{|\mathcal{V}_{X_2}\backslash\mathcal{H}_{X_2|X_1}|}{N}-\frac{|\bar{\mathcal{K}}_{X_\mathcal{M}}|}{kN}\\
&\geqslant\frac{|\mathcal{V}_{X_2}\backslash\mathcal{H}_{X_2|X_1}|}{N}-\frac{|\mathcal{V}_{X_2}\backslash\mathcal{H}_{X_2|X_1}|}{kN}\\
&\xrightarrow{N\to\infty} I(X_1;X_2)\left(1-\frac{1}{k}\right)\\
&\xrightarrow{k\to\infty} I(X_1;X_2),
\end{aligned}$$

where (a) holds because $\mathcal{F}_{X_\mathcal{M}}\cap\mathcal{K}_{X_\mathcal{M}}=\emptyset$, and where we have used Lemma 3.4.1 and [87] for the first limit.

3) Reliability: We only detail the reliability analysis for Terminal 3, since reliability for Terminal 1 is similar.

Let $i\in[\![1,k-1]\!]$. Note that Terminal 3 forms an accurate estimate of $U_i^{1:N}[\mathcal{H}_{X_2|X_3}]$ only when $U_{i+1}^{1:N}$ is correctly reconstructed. We note $\widehat{U}_i^{1:N}[\mathcal{H}_{X_2|X_3}]$ the estimate of $U_i^{1:N}[\mathcal{H}_{X_2|X_3}]$ formed by Terminal 3 and define $\mathcal{E}_i\triangleq\{\widehat{U}_i^{1:N}[\mathcal{H}_{X_2|X_3}]\neq U_i^{1:N}[\mathcal{H}_{X_2|X_3}]\}$.

Hence,

$$\begin{aligned}
\mathbb{P}[K_i \neq \widehat{K}_i] &\leqslant \mathbb{P}[U_i^{1:N} \neq \widehat{U}_i^{1:N}] \\
&= \mathbb{P}[U_i^{1:N} \neq \widehat{U}_i^{1:N} | \mathcal{E}_i^c]\mathbb{P}[\mathcal{E}_i^c] + \mathbb{P}[U_i^{1:N} \neq \widehat{U}_i^{1:N} | \mathcal{E}_i]\mathbb{P}[\mathcal{E}_i] \\
&\leqslant \mathbb{P}[U_i^{1:N} \neq \widehat{U}_i^{1:N} | \mathcal{E}_i^c] + \mathbb{P}[\mathcal{E}_i] \\
&\leqslant \mathbb{P}[U_i^{1:N} \neq \widehat{U}_i^{1:N} | \mathcal{E}_i^c] + \mathbb{P}[U_{i+1}^{1:N} \neq \widehat{U}_{i+1}^{1:N}] \\
&\overset{(a)}{\leqslant} N\delta_N + \mathbb{P}[U_{i+1}^{1:N} \neq \widehat{U}_{i+1}^{1:N}] \\
&\overset{(b)}{\leqslant} (k-i)N\delta_N + \mathbb{P}[U_k^{1:N} \neq \widehat{U}_k^{1:N}] \\
&\overset{(c)}{\leqslant} (k-i+1)N\delta_N,
\end{aligned}$$

where (a) holds because by [87], Terminal 3 can reconstruct $U_i^{1:N}$ from $U_i^{1:N}[\mathcal{H}_{X_2|X_3}]$ and $(X_3)_i^{1:N}$ with error probability less than $N\delta_N$, (b) holds by recurrence, (c) holds similarly as previous equations.

Then, by the union bound,

$$\begin{aligned}
\mathbf{P}_e(\mathcal{S}_N) &= \mathbb{P}[K_{1:k} \neq \widehat{K}_{1:k}] \\
&\leqslant \mathbb{P}[\cup_{i=1}^{k}(K_i \neq \widehat{K}_i)] \\
&\leqslant \sum_{i=1}^{k} \mathbb{P}[K_i \neq \widehat{K}_i] \\
&\leqslant \sum_{i=1}^{k} (k-i+1)N\delta_N \\
&= \frac{k(k+1)}{2} N\delta_N.
\end{aligned}$$

4) Key Uniformity: Similar to Lemma 3.5.2 we have the key uniformity for each block.

Lemma 3.3.17. *Uniformity of* $[K_i, \bar{K}_i]$ *holds for each block, where* $i \in [\![1, k-1]\!]$*. Specifically,*

$$|K_i| + |\bar{K}_i| - H(K_i \bar{K}_i) \leqslant N\delta_N.$$

Hence, we also have

$$|\bar{K}_i| - H(\bar{K}_i) \leqslant N\delta_N,$$

$$|K_i| - H(K_i) \leqslant N\delta_N.$$

The global key $K_{1:k}$ is asymptotically uniform as, similar to the proof of Theorem 3.5.3 in Section 3.5.3.4, we have

$$\mathbf{U}(\mathcal{S}_N) = |K_{1:k}| - H(K_{1:k}) \leqslant kN\delta_N.$$

5) Strong Secrecy: Similar to Lemma 3.5.3, we obtain the following result showing that secrecy holds for each block.

Lemma 3.3.18. *Let $i \in [\![1,k]\!]$. For each Block i, secrecy of $[K_i, \bar{K}_i]$ holds. Specifically, we have*

$$I\left(K_i\bar{K}_i; M_i\right) \leqslant 2N\delta_N.$$

Similar to Lemmas 3.5.4 and 3.5.5 we also have the following lemmas.

Lemma 3.3.19. *For $i \in [\![1,k]\!]$, we have for N large enough*

$$I(K_i;\bar{K}_i) \leqslant \delta_N^*,$$

where

$$\delta_N^* \triangleq -3\sqrt{2N\log 2} \times N2^{-N^\beta/2}\log_2\left(3\sqrt{2N\log 2} \times 2^{-N^\beta/2}\right).$$

Lemma 3.3.20. *For $i \in [\![2,k]\!]$, define*

$$\widetilde{L}_e^{1:i} \triangleq I(K_{1:i}\bar{K}_i; M_{1:i}).$$

We have

$$\widetilde{L}_e^{1:i} - \widetilde{L}_e^{1:i-1} \leqslant I\left(K_i\bar{K}_i; M_i\right) + I\left(K_{i-1};\bar{K}_{i-1}\right).$$

Similar to the proof of Theorem 3.5.3, using Lemmas 3.3.18, 3.3.20, 3.3.19, we obtain

$$\mathbf{L}(\mathcal{S}_N) \leqslant 2kN\delta_N + (k-1)\delta_N^*.$$

6) Seed Rate: The seed rate is

$$\begin{aligned}\frac{\sum_{i=1}^{k}|\widetilde{K}_i|}{kN} &= \frac{k|\bar{\mathcal{F}}_{X_2|X_1} \cup \bar{\mathcal{F}}_{X_2|X_3}|}{kN}\\ &\leqslant \frac{|\bar{\mathcal{F}}_{X_2|X_1}|+|\bar{\mathcal{F}}_{X_2|X_3}|}{N}\\ &\xrightarrow{N\to\infty} 0,\end{aligned}$$

where we have used Lemma 3.4.1 and [87].

CHAPTER 4

MULTIPLEXING PUBLIC AND CONFIDENTIAL MESSAGES OVER THE WIRETAP CHANNEL

4.1 Summary

In this chapter, we propose and analyze a source-channel coding architecture over a wiretap channel, in which secrecy is achieved by multiplexing public and confidential messages. Our main contribution is to circumvent the assumption that random numbers with perfectly uniform distributions are available, and to show that strong secrecy may be achieved "at negligible cost", in the sense of maintaining the overall communication rate of the same channel without secrecy constraints. Our source-channel coding architecture relies on a standard wiretap code combined with a modified source code, which we call a "uniform compression code," in which a small shared secret seed is used to enhance the uniformity of the source code output. We carry out an extensive analysis of uniform compression codes and characterize the optimal size of the seed.

4.2 Introduction

The objective of this chapter is to show that the cost of secrecy can be made negligible in the sense that it needs not incur a reduction in overall communication rate and need not require extra randomness resources. The crux of our approach is to analyze the wiretap channel model illustrated in Figure 25. Unlike the wiretap channel model presented in Section 1.3.1, here, the encoder is *deterministic* and is only used to *multiplex* a confidential source with a public source, and the objective is then to maximize the sum-rate of secret and public communication. The idea of multiplexing messages to achieve secrecy already implicitly appears in the original work of Csiszár and Körner [36], and is explicitly formalized in [48, 49]; however, our approach differs in that we relax the common assumption that messages are exactly uniformly

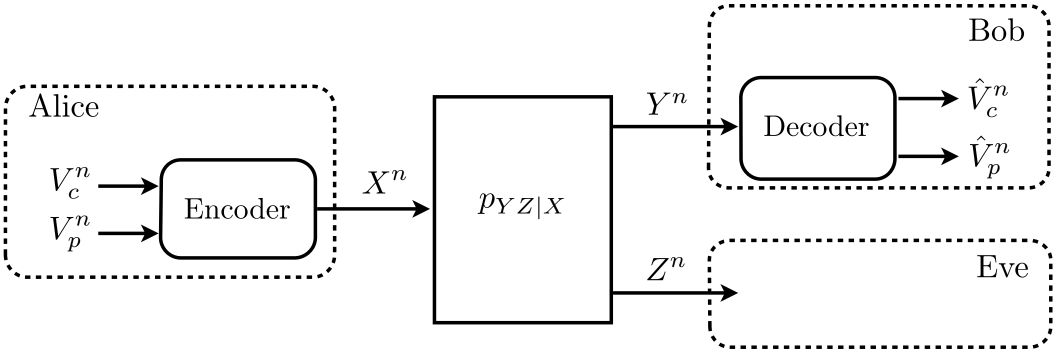


Figure 25. Multiplexing of confidential and public sources in the absence of additional local randomness at the transmitter. The confidential source must be reconstructible by the receiver and kept secret from the eavesdropper. The public source should be reconstructible by the receiver, and information may be leaked to the eavesdropper.

distributed, which is unrealistic even if messages are compressed with optimal source codes [93, 103], and we consider a strong notion of security.

The main contribution of this chapter is a source-channel coding architecture that achieves information-theoretic secrecy over this channel model. Our scheme, illustrated in Figure 26, combines a wiretap code designed to operate with perfectly uniform randomization with a modified source encoder, which compresses data while simultaneously ensuring good uniformity properties. This architecture explicitly requires the encoder and the decoder to share in advance a small secret seed K; however, we will see in Section 4.5 that the seed rate can be made arbitrarily small. Note that a secret key is anyway required for authentication [104, 105].

The remainder of the chapter is organized as follows. In Section 4.3, we formally describe the communication model under consideration. In Section 4.4, we show how to render the output of a source code nearly uniform. In Section 4.5, we prove that the architecture shown in Figure 26 achieves near-optimal performance using the result of Section 4.4. Finally, Section 4.6 concludes the chapter with some perspectives for future work.

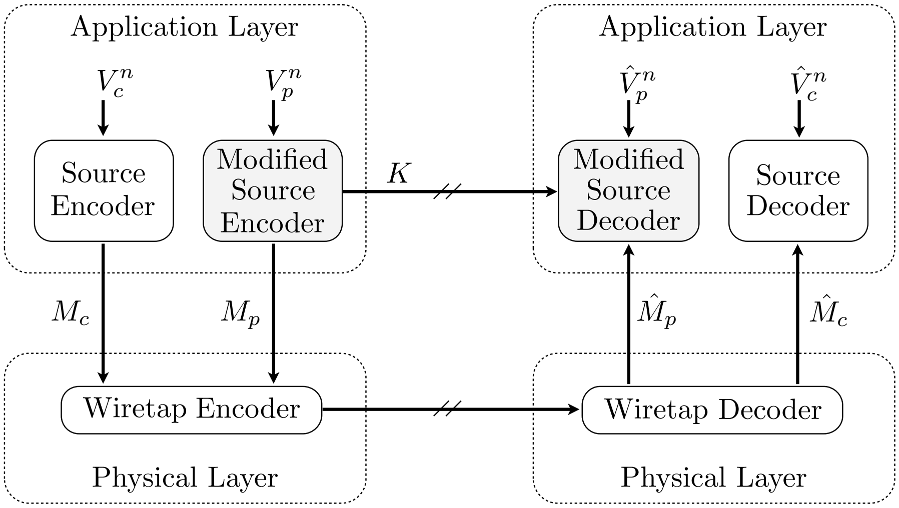


Figure 26. Proposed architecture to multiplex secure and public sources.

4.3 Preliminaries and Problem Statement

4.3.1 Wiretap channel model

Let $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ be finite alphabets. As illustrated in Figure 25, we consider a discrete memoryless wiretap channel $\big(\mathcal{X}, p_{YZ|X}, \mathcal{Y} \times \mathcal{Z}\big)$. The channel $\big(\mathcal{X}, p_{Y|X}, \mathcal{Y}\big)$ is called the main channel while the channel $\big(\mathcal{X}, p_{Z|X}, \mathcal{Z}\big)$ is called the eavesdropper's channel. We assume that the transmitter, Alice, wishes to transmit the realizations of the DMSs (V_c, p_{V_c}) and (V_p, p_{V_p}). Both sources are to be reconstructed without errors by the receiver observing Y^n, Bob, while the source (V_c, p_{V_c}) should be kept secret from the eavesdropper observing Z^n, Eve.

Definition 4.3.1. *A code for $\mathcal{C}_n$ the wiretap channel consists of the following.*

- *A deterministic encoding function $f_n : \mathcal{V}_c^n \times \mathcal{V}_p^n \to \mathcal{X}^n$, which maps n symbols of the confidential source and n symbols of the public source to a codeword of length n;*
- *A decoding function $g_n : \mathcal{Y}^n \to (\mathcal{V}_c^n \times \mathcal{V}_p^n)$, which maps a sequence of n channel output observations to n symbols of the confidential source and n symbols of the public source.*

The performance of $\mathcal{C}_n$ is measured in terms of the average probability of error

$$P_e(\mathcal{C}_n) \triangleq \mathbb{P}\left[(V_c^n, V_p^n) \neq g_n(Y^n)\right],$$

and in terms of the secrecy metric

$$S(\mathcal{C}_n) \triangleq \max_{v_c^n \in \mathcal{V}_c^n} \mathbb{V}\big(p_{Z^n|V_c^n = v_c^n}, p_{Z^n}\big).$$

4.3.2 Source-channel coding theorem

Theorem 4.3.1. *Consider a confidential DMS (V_c, p_{V_c}) and a public DMS (V_p, p_{V_p}) to be transmitted over a wiretap channel $\big(\mathcal{X}, p_{YZ|X}, \mathcal{Y} \times \mathcal{Z}\big)$. For any random variable $U \in \mathcal{U}$ such that $U - X - YZ$, if*

$$\begin{cases} H(V_c) + H(V_p) < I(X;Y) \\ H(V_c) < I(X;Y|U) - I(X;Z|U) \\ H(V_p) > I(X;Z|U) \end{cases},$$

then there exists a sequence of codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ such that

$$\lim_{n\to\infty} P_e(\mathcal{C}_n) = \lim_{n\to\infty} S(\mathcal{C}_n) = 0.$$

Conversely, if there exists a sequence of codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ such that $\lim_{n\to\infty} P_e(\mathcal{C}_n) = \lim_{n\to\infty} S(\mathcal{C}_n) = 0$, then there must exist a random variable $U \in \mathcal{U}$ such that $U - X - YZ$ and

$$\begin{cases} H(V_c) + H(V_p) \leqslant I(X;Y) \\ H(V_c) \leqslant I(X;Y|U) - I(X;Z|U) \\ H(V_p) \geqslant I(X;Z|U) \end{cases}.$$

Although the result might seem intuitive, the achievability proof does not follow from standard arguments and known results because we do not assume the existence of a local source of uniform random numbers; consequently, the encoder must only operate on the sequences emitted by the sources. The main contribution of this chapter is the achievability proof detailed in Section 4.5 using the architecture of Figure 26. The converse follows by combining the proofs in [45, 46].

Remark 4.3.1. *Unlike the capacity region of the broadcast channel with confidential messages, the information constraints in Theorem 4.3.1 do not include an auxiliary random variable V such that $U - V - X - YZ$. This result is not surprising, as this extra random variable accounts for the addition of artificial noise (channel prefixing) in the encoder, which is not allowed by our model, as we require all encoder inputs to be decoded at the receiver. The random variable U is merely a time-sharing random variable [46, 55].*

4.4 Uniform compression codes

Consider a DMS $(\mathcal{X}, p_X)$. Let $n \in \mathbb{N}$, $d_n \in \mathbb{N}$, and let U_{d_n} be a uniform random variable over $\mathcal{U}_{d_n} \triangleq [\![1, 2^{d_n}]\!]$, independent of X^n. In the following we refer to U_{d_n} as the *seed* and d_n as its length. As illustrated in Figure 27, our objective is to design a source code to compress and reconstruct the DMS $(\mathcal{X}, p_X)$ with the assistance of a seed U_{d_n}.

Definition 4.4.2. *A $(2^{nR}, n, 2^{d_n})$ uniform compression code $\mathcal{C}_n$ for a DMS $(\mathcal{X}, p_X)$ consists of*

- *A message set $\mathcal{M}_n \triangleq [\![1, M_n]\!]$, with $M_n \triangleq 2^{nR}$,*
- *A seed set $\mathcal{U}_{d_n} \triangleq [\![1, 2^{d_n}]\!]$,*
- *An encoding function $\phi_n : \mathcal{X}^n \times \mathcal{U}_{d_n} \to \mathcal{M}_n$,*
- *A decoding function $\psi_n : \mathcal{M}_n \times \mathcal{U}_{d_n} \to \mathcal{X}^n$.*

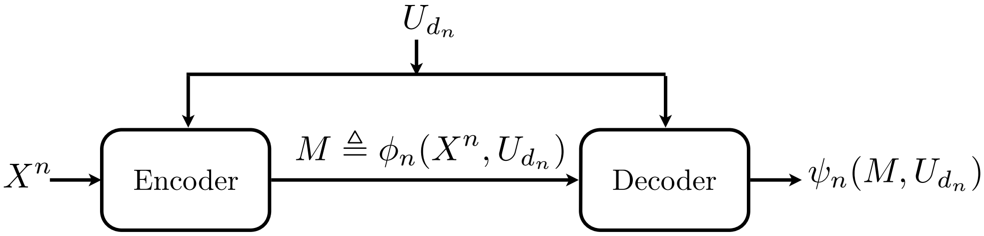


Figure 27. Source encoder and decoder with uniform outputs.

The performance of the code is measured in terms of the average probability of error and the uniformity of its output as

$$\begin{aligned} P_e(\phi_n, \psi_n) &\triangleq \mathbb{P}[X^n \neq \psi_n(\phi_n(X^n, U_{d_n}), U_{d_n})], \\ U_e(\phi_n) &\triangleq \mathbb{V}[p_{\phi_n(X^n, U_{d_n})}, p_{U_{M_n}}], \end{aligned}$$

where U_{M_n} has uniform distribution over $\mathcal{M}_n$.

Remark 4.4.2. *Uniformity could be measured with the stronger metric* $U_e'(\phi_n) \triangleq \mathbb{D}[p_{\phi_n(X^n, U_{d_n})}, p_{U_{M_n}}]$*, where* $\mathbb{D}(\cdot,\cdot)$ *is the Kullback-Leibler divergence; however, by [102, Lemma 2.7],* $U_e(\phi_n)$ *can be replaced by* $U_e'(\phi_n)$*, if* $\lim\limits_{n\to\infty} nU_e(\phi_n) = 0$*, which will be the case.*

Definition 4.4.3. *A rate* R *is achievable, if there exists a sequence of* $(2^{nR}, n, 2^{d_n})$ *uniform compression codes* $\{\mathcal{C}_n\}_{n\geqslant 1}$ *for the DMS* $(\mathcal{X}, p_X)$*, such that*

$$\lim_{n\to\infty} \frac{1}{n} \log M_n \leqslant R, \ \ \lim_{n\to\infty} \frac{d_n}{n} = 0, \lim_{n\to\infty} P_e(\phi_n, \psi_n) = 0, \ \ \textit{and} \ \ \lim_{n\to\infty} U_e(\phi_n) = 0.$$

Our main result in this section is the characterization of the infimum of achievable rates with uniform compression codes as well as the optimal scaling of the seed length d_n. In the following, we use the Landau notation to characterize the limiting behavior of the seed scaling, with the convention that for any real functions f and g, $f = \Omega(g)$ means $f = o(g)$ is false.

Proposition 4.4.1. *Let* $(\mathcal{X}, p_X)$ *be a DMS. Then,*

$$\inf\{R : R \textit{ is achievable with a uniform compression code}\} = H(X).$$

Moreover, the optimal seed length d_n *for* $(2^{nR}, n, 2^{d_n})$ *code verifies*

$$d_n \in \Omega(n^{1/2}) \cap O(n^{1/2+\epsilon}) \quad \textit{for any } \epsilon > 0. \tag{76}$$

Proof. See Appendix 4.A. □

As a first attempt to develop a more practical scheme for uniform compression codes, we propose an achievability scheme for Proposition 4.4.1 based on invertible extractors [106]. We start by recalling known facts about extractors.

Definition 4.4.4 ([106])**.** *Let $\epsilon > 0$. Let $m, d, l \in \mathbb{N}$ and let $t \in \mathbb{R}^+$. A polynomial time probabilistic function* Ext : $\{0,1\}^m \times \{0,1\}^d \mapsto \{0,1\}^l$ *is called a (m, d, l, t, ϵ)-extractor, if for all binary source X satisfying $H_\infty(X) \geqslant t$, we have*

$$\mathbb{V}(p_{\text{Ext}(X,U_d)}, p_{U_l}) \leqslant \epsilon,$$

where U_d is a sequence of d uniformly distributed bits, $\mathcal{U}_l$ is the uniform distribution over $\{0,1\}^l$.
Moreover, a (m, d, l, t, ϵ)-extractor is said invertible if the input can be reconstructed from the output and U_d.

It can be shown [106,107] that there exists *explicit* invertible (m, d, m, t, ϵ)-extractor such that

$$d = m - t + 2\log m + 2\log\frac{1}{\epsilon} + O(1). \tag{77}$$

The following proposition shows that one can establish optimal uniform compression codes using such invertible extractors.

Proposition 4.4.2. *Let $(\mathcal{X}, p_X)$ be a binary memoryless source. For any $R > H(X)$ and for any $\epsilon > 0$, the rate R can be achieved with a sequence of uniform compression codes such that*

- *the seed length scales as $d_n = \Theta(n^{1/2+\epsilon})$;*
- *the encoder $\phi_n : \mathcal{X}^n \times \mathcal{U}_{d_n} \to \mathcal{M}_n$ is composed of a typical-sequence based source code combined with an invertible extractor as described in Figure 28.*

Proof. See Appendix 4.B. □

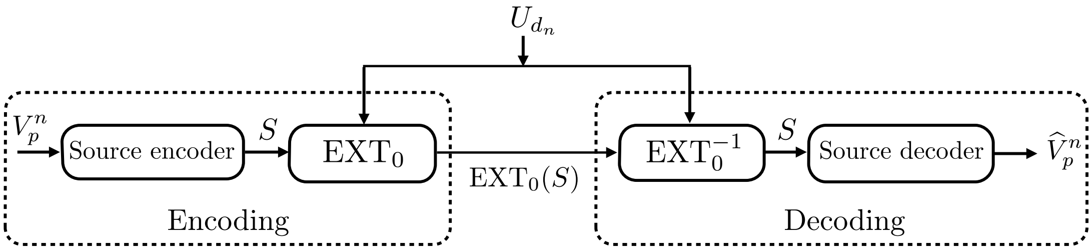


Figure 28. Encoding/Decoding scheme for Proposition 4.4.2. The encoder/decoder is obtained from a typical-sequence based source code, and $\mathbf{EXT}_0$ is an invertible extractor.

Unfortunately, this scheme is not fully explicit because it relies on a typical-sequence based compression. To provide at least one explicit example, we finally develop a uniform compression code based on polar codes for a binary memoryless source $(\mathcal{X}, p_X)$, $\mathcal{X} \triangleq \{0,1\}$. Let $\beta \in]0, 1/2[$, $n \in \mathbb{N}$, $N \triangleq 2^n$, and $\delta_N \triangleq 2^{-N^\beta}$. Let $G_N \triangleq \left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]^{\otimes n}$ be the source polarization transform defined in [87], and set $A^N \triangleq X^N G_N$. For any set $\mathcal{A} \triangleq \{i_j\}_{j=1}^{|\mathcal{A}|}$ of indices in $\llbracket 1, N \rrbracket$, we define $A^N[\mathcal{A}] \triangleq \left[A_{i_1}, A_{i_2}, \dots, A_{i_{|\mathcal{A}|}}\right]$. We also define the sets

$$\begin{aligned} \mathcal{V}_X &\triangleq \left\{ i \in \llbracket 1, N \rrbracket : H\left(A_i | A^{i-1}\right) > 1 - \delta_N \right\}, \\ \mathcal{H}_X &\triangleq \left\{ i \in \llbracket 1, N \rrbracket : H\left(A_i | A^{i-1}\right) > \delta_N \right\}. \end{aligned}$$

A polar-based uniform compression code is obtained by defining the encoding function ϕ_N as follows.

$$\phi_N : (X^N, U_{|\mathcal{H}_X \backslash \mathcal{V}_X|}) \mapsto \left(A^N[\mathcal{V}_X], A^N[\mathcal{H}_X \backslash \mathcal{V}_X] \oplus U_{|\mathcal{H}_X \backslash \mathcal{V}_X|} \right)$$

Proposition 4.4.3. *Let $(\mathcal{X}, p_X)$ be a binary memoryless source. For any $R < H(X)$, the rate R is achievable with a sequence of polar-based uniform compressions codes with length N such that the seed length $|\mathcal{H}_X \backslash \mathcal{V}_X|$ scales as $o(N)$. In addition, the complexity of the encoding is $O(N \log N)$.*

Proof. See Appendix 4.C. □

4.5 Source-channel coding architecture based on uniform compression codes

Recall that our objective is to circumvent the impossibility of generating uniform random numbers with source codes [93, Theorem 4]. The approach to overcome this impossibility is to introduce a small shared uniformly distributed sequence, which we call a "seed", and to use the result of Section 4.4. While, the price paid is that the emitter and the receiver must now share a seed of negligible rate, we will show how to further reduce the seed rate.

4.5.1 Achievability of Theorem 4.3.1 based on uniform compression codes

The uniform compression codes of Section 4.4 may now be combined with known wiretap codes (as depicted in Figure 26), whose properties we recall in the following lemma.

Lemma 4.5.1 (Adapted from [46, Proposition 1])**.** *Consider a Discrete Memoryless Channel (DMC) $(\mathcal{X}, p_{YZ|X}, \mathcal{Y} \times \mathcal{Z})$, in which a message $M \in [\![1, 2^{nR}]\!]$ is encoded by means of a uniform auxiliary message $M_p \in [\![1, 2^{nR_p}]\!]$. If there exists a joint distribution p_{UXYZ} that factorizes as $p_U p_{X|U} p_{YZ|X}$ such that*

$$R + R_p < \mathbb{I}(X;Y) \tag{78}$$

$$R < I(X;Y|U) - I(X;Z|U) \tag{79}$$

$$R_p > I(X;Z|U), \tag{80}$$

then there exists a sequence of wiretap codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ such that

$$\lim_{n\to\infty} \max_m \mathbb{P}\left[\hat{M}_c \neq M_c | M_c = m\right] = 0,$$

$$\lim_{n\to\infty} \max_m \mathbb{P}\left[\hat{M}_p \neq M_p | M_c = m\right] = 0,$$

$$\lim_{n\to\infty} \max_m \mathbb{V}\big(p_{Z^n|M_c=m}, p_{Z^n}\big) = 0.$$

Let $\epsilon > 0$. Going back again to the setting of Section 4.3.1, we encode the confidential DMS using a traditional source code, and the public DMS using a uniform compression code as in Proposition 4.4.1. The corresponding source encoder-decoder pairs are denoted (f_n^c, g_n^c) and (f_n^p, g_n^p), respectively, and we set $M_c \triangleq f_n^c(V_c^n) \in [\![1, 2^{nR_c}]\!]$ and $M_p \triangleq f_n^p(V_p^n) \in [\![1, 2^{nR_p}]\!]$. We assume n large enough so that

$$\mathbb{P}[(V_c^n, V_p^n) \neq (g_n^c(M_c), g_n^p(M_p))] \leqslant \epsilon, \tag{81}$$

$$\mathbb{V}\Big(p_{M_p}, p_{U_{nR_p}}\Big) < \epsilon, \tag{82}$$

where U_{nR_p} is uniformly distributed over $[\![1, 2^{nR_p}]\!]$. Under the conditions (78)-(80) of Lemma 4.5.1, which are met whenever

$$\begin{aligned} H(V_c) + H(V_p) &< I(X;Y), \\ H(V_c) &< I(X;Y|U) - I(X;Z|U), \\ H(V_p) &> I(X;Z|U), \end{aligned}$$

for n sufficiently large there exists a wiretap code $\mathcal{C}_n$ so that for any m_c, and for $\tilde{M}_p$ distributed according to $p_{U_{nR_p}}$, the uniform distribution over $[\![1, 2^{nR_p}]\!]$,

$$\mathbb{P}\left[\hat{\tilde{M}}_p \neq \tilde{M}_p | M_c = m_c\right] < \epsilon, \tag{83}$$

$$\mathbb{P}\left[\hat{\tilde{M}}_c \neq M_c | M_c = m_c\right] < \epsilon, \tag{84}$$

$$\mathbb{V}\big(\tilde{p}_{Z^n|M_c=m_c}, \tilde{p}_{Z^n}\big) \leqslant \epsilon, \tag{85}$$

where $(\hat{\tilde{M}}_p, \hat{\tilde{M}}_c)$ is the estimate of $(\tilde{M}_p, M_c)$ by the decoder of $\mathcal{C}_n$, and for any z^n,

$$\tilde{p}_{Z^n}(z^n) \triangleq \sum_{m_c=1}^{2^{nR_c}} \sum_{m_p=1}^{2^{nR_p}} p_{Z^n|M_c=m_c,M_p=m_p}(z^n) p_{M_c}(m_c) p_{U_{nR_p}}(m_p).$$

Note that (83)-(85) holds by Lemma 4.5.1 because we have assumed $\tilde{M}_p$ uniformly distributed. We now study the consequences of using the wiretap code $\mathcal{C}_n$ with M_p (not exactly uniformly distributed) instead of $\tilde{M}_p$. Specifically, we note $(\hat{M}_p, \hat{M}_c)$ the resulting estimate of (M_p, M_c) by the decoder of $\mathcal{C}_n$, and define for any z^n,

$$p_{Z^n}(z^n) \triangleq \sum_{m_c=1}^{2^{nR_c}} \sum_{m_p=1}^{2^{nR_p}} p_{Z^n|M_c=m_c,M_p=m_p}(z^n) p_{M_c}(m_c) p_{M_p}(m_p).$$

We then have for any m_c,

$$
\begin{aligned}
&\mathbb{V}\big(p_{Z^n|M_c=m_c}, p_{Z^n}\big)\\
&\overset{(a)}{\leqslant} \mathbb{V}\big(p_{Z^n|M_c=m_c}, \tilde{p}_{Z^n|M_c=m_c}\big) + \mathbb{V}\big(\tilde{p}_{Z^n|M_c=m_c}, \tilde{p}_{Z^n}\big) + \mathbb{V}(\tilde{p}_{Z^n}, p_{Z^n})\\
&\overset{(b)}{\leqslant} \epsilon + \mathbb{V}\big(p_{Z^n|M_c=m_c}, \tilde{p}_{Z^n|M_c=m_c}\big) + \mathbb{V}(\tilde{p}_{Z^n}, p_{Z^n})\\
&\overset{(c)}{\leqslant} \epsilon + \sum_{z^n}\sum_{m_p} p_{Z^n|M_c=m_c,M_p=m_p}(z^n)\left|p_{M_p}(m_p) - p_{U_{nR_p}}(m_p)\right| + \mathbb{V}(\tilde{p}_{Z^n}, p_{Z^n})\\
&= \epsilon + \mathbb{V}\Big(p_{M_p}, p_{U_{nR_p}}\Big) + \mathbb{V}(\tilde{p}_{Z^n}, p_{Z^n})\\
&\overset{(d)}{\leqslant} 2\epsilon + \mathbb{V}(\tilde{p}_{Z^n}, p_{Z^n})\\
&\overset{(e)}{\leqslant} 2\epsilon + \sum_{z^n}\sum_{m_c,m_p} p_{M_c}(m_c)p_{Z^n|M_c=m_c,M_p=m_p}(z^n)\left|p_{M_p}(m_p) - p_{U_{nR_p}}(m_p)\right|\\
&= 2\epsilon + \mathbb{V}\Big(p_{M_p}, p_{U_{nR_p}}\Big)\\
&\overset{(f)}{\leqslant} 3\epsilon,
\end{aligned}
\tag{86}
$$

where (a), (c), and (e) follow by the triangle inequality, (b) holds by (85), (d) and (f) hold by (82).

Consider then an optimal coupling [96] between M_p and $\tilde{M}_p$ such that $\mathbb{P}[\mathcal{E}] = \mathbb{V}(p_{M_p}, p_{U_{nR_p}})$, where $\mathcal{E} \triangleq \{M_p \neq \tilde{M}_p\}$. We have for any m_c,

$$
\begin{aligned}
&\mathbb{P}\left[\hat{M}_p \neq M_p | M_c = m_c\right]\\
&= \mathbb{P}\left[\hat{M}_p \neq M_p | M_c = m_c, \mathcal{E}^c\right]\mathbb{P}\left[\mathcal{E}^c\right] + \mathbb{P}\left[\hat{M}_p \neq M_p | M_c = m_c, \mathcal{E}\right]\mathbb{P}\left[\mathcal{E}\right]\\
&\leqslant \mathbb{P}\left[\hat{M}_p \neq M_p | M_c = m_c, \mathcal{E}^c\right] + \mathbb{P}\left[\mathcal{E}\right]\\
&= \mathbb{P}\left[\hat{M}_p \neq M_p | M_c = m_c, \mathcal{E}^c\right] + \mathbb{V}(p_{M_p}, p_{U_{nR_p}})\\
&\leqslant 2\epsilon,
\end{aligned}
$$

where the last inequality follows from (82) and (83). Similarly, using (82) and (84), we have for any m_c,

$$
\mathbb{P}\left[\hat{M}_c \neq M_c | M_c = m_c\right] \leqslant 2\epsilon.
$$

Encoding the sources into codewords as $f_n(f_n^c(V_c^n), f_n^p(V_p^n))$, and forming estimates from the channel output Y^n as $\hat{V}_c^n \triangleq g_n^c(g_n(Y^n))$, and $\hat{V}_p^n \triangleq g_n^p(g_n(Y^n))$, we obtain again

$$\begin{aligned}
&\mathbb{P}[(V_c^n, V_p^n) \neq (\hat{V}_c^n, \hat{V}_p^n)] \\
&\leqslant \mathbb{P}[(V_c^n, V_p^n) \neq (\hat{V}_c^n, \hat{V}_p^n)|(\hat{M}_p, \hat{M}_c) = (M_p, M_c)] + \mathbb{P}[(\hat{M}_p, \hat{M}_c) \neq (M_p, M_c)] \\
&\leqslant 5\epsilon,
\end{aligned}$$

and for any $v_c^n \in \mathcal{V}_c^n$, by noting that $p_{Z^n|V_c^n=v_c^n} = p_{Z^n|V_c^n=v_c^n, M_c=f_n^c(v_c^n)} = p_{Z^n|M_c=f_n^c(v_c^n)}$, where the last equality holds because $Z^n \to M_c \to V_c^n$, we have by (86)

$$\mathbb{V}\big(p_{Z^n|V_c^n=v_c^n}, p_{Z^n}\big) \leqslant 3\epsilon.$$

Since $\epsilon > 0$ can be chosen arbitrarily small, we obtain again the achievability part of Theorem 4.3.1.

4.6 Conclusion

We have proposed and analyzed a source-channel coding architecture for multiplexing confidential and public messages that achieves information-theoretic secrecy over the wiretap channel. Our architecture exploits uniform compression codes that output nearly uniform messages. By showing that secrecy can be achieved without extra randomness resources, and without reducing the overall rate of reliable communication, we show that secrecy can be achieved at negligible cost and provide a step towards integrating physical-layer security into communication systems.

While our architecture introduces a new coding scheme at the application layer, another approach consisting in modifying the physical-layer of the protocol stack, could be possible relying on wiretap codes that do not require uniform randomization [46]. This topic is left for future work and will be addressed in [108].

An important issue that we have not addressed is the design of *universal* wiretap codes that merely require that the public message carries enough randomness, and do not require the knowledge of the statistics. Some results in this direction are already available in [47].

APPENDICES

4.A Proof of Proposition 4.4.1

4.A.1 Achievability

There exists a sequence of $(2^{nR}, n, 2^{d_n})$ uniform compression codes $\{\mathcal{C}_n\}_{n \in \mathbb{N}^*}$ such that C is achievable with a seed length d_n scaling as

$$d_n = \Theta(n^{1/2+\epsilon}),$$

where $\epsilon > 0$ is arbitrary.

Proof. Let $\epsilon_1 > 0$, $\epsilon > 0$, $n \in \mathbb{N}$, $d_n \in \mathbb{N}$, $R > 0$. Define $M_n \triangleq 2^{nR}$ and $\mathcal{M}_n \triangleq [\![1, M_n]\!]$. Consider a random mapping $\Phi : \mathcal{X}^n \times \mathcal{U}_{d_n} \to \mathcal{M}_n$, and its associated decoder $\Psi : \mathcal{M}_n \times \mathcal{U}_{d_n} \to \mathcal{X}^n$. Given $(m, u_{d_n}) \in \mathcal{M}_n \times \mathcal{U}_{d_n}$, the decoder outputs $\hat{x}^n$ if it is the unique sequence such that $\hat{x}^n \in \mathcal{T}_{\epsilon_1}^n(X)$ and $\Phi(\hat{x}^n, u_{d_n}) = m$; otherwise it outputs an error. We let $M \triangleq \Phi(X^n, U_{d_n})$, and define $\mathbf{P}_e \triangleq \mathbb{P}[X^n \neq \Psi(\Phi(X^n, U_{d_n}), U_{d_n})]$, $\mathbf{U}_e \triangleq \mathbb{V}\left(p_M, p_{\mathcal{U}_{M_n}}\right)$.

- We first determine a condition over R to ensure $\mathbb{E}_\Phi\left[\mathbf{U}_e\right] \leqslant \epsilon$. Remark that

$$\forall m \in \mathcal{M}_n, p_M(m) = \sum_{x^n} \sum_{u} p(x^n, u) \mathbb{1}\{\Phi(x^n, u) = m\},$$

hence, on average $\forall m \in \mathcal{M}_n$, $\mathbb{E}_\Phi\left[p_M(m)\right] = 2^{-nR}$, which allows us to write

$$\begin{aligned}\mathbb{E}_\Phi\left[\mathbf{U}_e\right] &= \mathbb{E}_\Phi\left[\sum_m \left|p_M(m) - \mathbb{E}_\Phi\left[p_M(m)\right]\right|\right] \\ &\leqslant \sum_{i=1}^{2} \mathbb{E}_\Phi\left[\sum_m \left|p_M^{(i)}(m) - \mathbb{E}_\Phi\left[p_M^{(i)}(m)\right]\right|\right], \end{aligned} \tag{87}$$

where $\forall m \in \mathcal{M}_n$, $\forall i \in [\![1, 2]\!]$,

$$p_M^{(i)}(m) = \sum_{x^n \in \mathcal{A}_i} \sum_{u} p(x^n, u) \mathbb{1}\{\Phi(x^n, u) = m\},$$

with $\mathcal{A}_1 \triangleq \mathcal{T}_{\epsilon_1}^n(X)$ and $\mathcal{A}_2 \triangleq \mathcal{A}_1^c$. After some manipulations we bound the second term in (87) as follows

$$\mathbb{E}_\Phi\left[\sum_m \left|p_M^{(2)}(m) - \mathbb{E}_\Phi\left[p_M^{(2)}(m)\right]\right|\right] \leqslant 4|\mathcal{X}|e^{-n\epsilon_1^2\mu_X}, \tag{88}$$

with $\mu_X = \min\limits_{x\in\mathrm{supp}(P_X)} P_X(x)$. Then, we bound the first term in (87) by Jensen's inequality

$$\mathbb{E}_\Phi\left[\sum_m \left|p_M^{(1)}(m) - \mathbb{E}_\Phi\left[p_M^{(1)}(m)\right]\right|\right] \leqslant \sum_m \sqrt{\mathrm{Var}_\Phi\left(p_M^{(1)}(m)\right)}. \tag{89}$$

Moreover, after some manipulations, we obtain

$$\begin{aligned}
\mathrm{Var}_\Phi\left(p_M^{(1)}(m)\right) &= \sum_{x^n\in\mathcal{T}_{\epsilon_1}^n(X)} \sum_u p(x^n,u)^2 \mathrm{Var}_\Phi\left(\mathbb{1}\{\Phi(x^n,u)=m\}\right)\\
&\leqslant \sum_{x^n\in\mathcal{T}_{\epsilon_1}^n(X)} \sum_u p(x^n,u)^2 \mathbb{E}_\Phi\left[(\mathbb{1}\{\Phi(x^n,u)=m\})^2\right]\\
&= \sum_{x^n\in\mathcal{T}_{\epsilon_1}^n(X)} \sum_u p(x^n,u)^2 \mathbb{E}_\Phi\left[\mathbb{1}\{\Phi(x^n,u)=m\}\right]\\
&= \sum_{x^n\in\mathcal{T}_{\epsilon_1}^n(X)} \sum_u p(x^n)^2 p(u)^2 2^{-nR}\\
&= \sum_{x^n\in\mathcal{T}_{\epsilon_1}^n(X)} p(x^n)^2 2^{-d} 2^{-nR}\\
&\leqslant \sum_{x^n\in\mathcal{T}_{\epsilon_1}^n(X)} \exp_2\left[-2n(1-\epsilon_1)H(X)\right] 2^{-d}\frac{1}{M_n}\\
&\leqslant \exp_2\left[n(1+\epsilon_1)H(X)\right]\exp_2\left[-2n(1-\epsilon_1)H(X)\right] 2^{-d}2^{-nR}\\
&\leqslant \exp_2\left[-n(1-3\epsilon_1)H(X)\right] 2^{-d_n}2^{-nR}.
\end{aligned} \tag{90}$$

Thus, by combining (89) and (90), we obtain

$$\begin{aligned}
&\mathbb{E}_\Phi\left[\sum_m \left|p_M^{(1)}(m) - \mathbb{E}_\Phi\left[p_M^{(1)}(m)\right]\right|\right]\\
&\leqslant \sum_m \sqrt{\exp_2\left[-n(1-3\epsilon_1)H(X)\right] 2^{-d_n}2^{-nR}}\\
&= \sqrt{M_n}\exp_2\left[-\frac{n}{2}\left((1-3\epsilon_1)H(X) + \frac{d_n}{n}\right)\right]\\
&\leqslant \exp_2\left[\frac{n}{2}\left(R-(1-3\epsilon_1)H(X) - \frac{d_n}{n}\right)\right].
\end{aligned} \tag{91}$$

Hence, if $R < H(X) + \frac{d_n}{n} - 3\epsilon_1 H(X)$, then asymptotically $\mathbb{E}_\Phi\left[\mathbf{U}_e\right] \leqslant \epsilon$ by (88) and (91).

- We now derive a condition over R to ensure $\mathbb{E}_\Phi[\mathbf{P}_e] \leqslant \epsilon$. We define $\mathcal{E}_0 \triangleq \{X^n \notin \mathcal{T}^n_{\epsilon_1}(X)\}$, and $\mathcal{E}_1 \triangleq \{\exists \hat{x}^n \neq X^n, \Phi(\hat{x}^n, U) = \Phi(X^n, U)$ and $\hat{x}^n \in \mathcal{T}^n_{\epsilon_1}(X)\}$ so that by the union bound, $\mathbb{E}_\Phi[\mathbf{P}_e] \leqslant \mathbb{P}[\mathcal{E}_0] + \mathbb{P}[\mathcal{E}_1]$. We have

$$\mathbb{P}[\mathcal{E}_0] \leqslant 2|\mathcal{X}|e^{-n\epsilon_1^2 \mu_X}, \tag{92}$$

and defining $\mathbf{P}(x^n, \hat{x}^n, u) \triangleq \mathbb{P}[\exists \hat{x}^n \neq x^n, \Phi(\hat{x}^n, u) = \Phi(x^n, u)$ and $\hat{x}^n \in \mathcal{T}^n_{\epsilon_1}(X)]$, we have

$$\begin{aligned}
\mathbb{P}[\mathcal{E}_1] &= \sum_{x^n} \sum_u p(x^n, u)\mathbf{P}(x^n, \hat{x}^n, u) \\
&\leqslant \sum_{x^n} \sum_u p(x^n, u) \sum_{\substack{\hat{x}^n \in \mathcal{T}^n_{\epsilon_1}(X) \\ \hat{x}^n \neq x^n}} \mathbb{P}[\Phi(\hat{x}^n, u) = \Phi(x^n, u)] \\
&= \sum_{x^n} \sum_u p(x^n, u) \sum_{\substack{\hat{x}^n \in \mathcal{T}^n_{\epsilon_1}(X) \\ \hat{x}^n \neq x^n}} 2^{-nR} \\
&\leqslant \sum_{x^n} \sum_u p(x^n, u) |\mathcal{T}^n_{\epsilon_1}(X)| 2^{-nR} \\
&\leqslant \sum_{x^n} \sum_u p(x^n, u) \exp_2\left[nH(X)(1+\epsilon_1)\right] 2^{-nR} \\
&\leqslant \exp_2\left[n(H(X)(1+\epsilon_1) - R)\right].
\end{aligned} \tag{93}$$

Hence, if $R > H(X) + \epsilon_1 H(X)$, then asymptotically $\mathbb{E}_\Phi(\mathbf{P}_e) \leqslant \epsilon$ by (92) and (93).

All in all, if R is such that

$$H(X) + \epsilon_1 H(X) < R < H(X) + \frac{d_n}{n} - 3\epsilon_1 H(X),$$

then asymptotically by the selection lemma, $\mathbb{E}_\Phi[\mathbf{U}_e] \leqslant \epsilon$ and $\mathbb{E}_\Phi[\mathbf{P}_e] \leqslant \epsilon$. Thus, we choose d_n such that

$$4n\epsilon_1 H(X) < d_n \leqslant 4n\epsilon_1 H(X) + 1,$$

to obtain

$$H(X) + \epsilon_1 H(X) < H(X) + \frac{d_n}{n} - 3\epsilon_1 H(X).$$

We can also choose $\epsilon_1 = n^{-1/2+\epsilon_b}$, with any $\epsilon_b > 0$,[1] so that for any $\epsilon_a > \epsilon_b$

$$4n^{\epsilon_b - \epsilon_a} H(X) < \frac{d_n}{n^{1/2+\epsilon_a}} \leqslant 4n^{\epsilon_b-\epsilon_a} H(X) + n^{-1/2-\epsilon_a},$$

which means $d_n = o(n^{1/2+\epsilon_a})$. Finally, by means of the selection lemma applied to $\mathbf{P}_e$ and $\mathbf{U}_e$, there exists a realization of Φ such that $\mathbf{U}_e \leqslant \epsilon$ and $\mathbf{P}_e \leqslant \epsilon$. □

4.A.2 Converse

We first show that any achievable rate R must satisfy $R \geqslant H(X)$. Assume that R is an achievable rate. We note $M \triangleq \phi_n(X^n, U_{d_n})$. We have

$$\begin{aligned} nR &\geqslant H(M) \\ &= H(M|U_{d_n}) + I(U_{d_n}; M) \\ &= I(X^n; M|U_{d_n}) + I(U_{d_n}; M) \\ &= H(X^n|U_{d_n}) - H(X^n|MU_{d_n}) + I(U_{d_n}; M) \\ &\overset{(a)}{=} H(X^n|U_{d_n}) + I(U_{d_n}; M) - n\delta(\epsilon) \\ &\overset{(b)}{\geqslant} H(X^n|U_{d_n}) - n\delta(\epsilon) \\ &\overset{(c)}{=} nH(X) - n\delta(\epsilon), \end{aligned}$$

where (a) holds by Fano's inequality and $\lim_{\epsilon \to 0} \delta(\epsilon) = 0$, (b) holds by positivity of the mutual information, and (c) holds by independence of X^n and U_{d_n}.

Hence it remains to show an upper bound for the optimal scaling of d_n. It is done by means of a second order asymptotic study. We consider an arbitrary source $\mathbf{X} \triangleq \{X^n\}_{n=1}^{\infty}$, where X^n is a random variable taking values in $\mathcal{X}^n$ subject to P_{X^n}. Specifically, we generalize some results of [93] to our setup, and show that if $d_n = o(\sqrt{n})$, with n the code length, then the trade-off between error probability and

[1] See Equations (88) and (92).

uniformity of [93] cannot be improved. In the following, we use the notation $\mathbf{P}_e \triangleq P_e(\phi_n, \psi_n)$ and $\mathbf{U}_e \triangleq U_e(\phi_n)$, for a code $\mathcal{C}_n \triangleq (\phi_n, \psi_n, \mathcal{M}_n)$.

For the fixed-length source coding problem, for $\epsilon > 0$, for $\mathbf{d} \triangleq \{d_n\}_n \in \mathbb{R}_+^{\mathbb{N}}$ and for a code $\mathcal{C}_n \triangleq (\phi_n, \psi_n, \mathcal{M}_n)$, we define the following first order asymptotics

$$a_0 \triangleq R(\mathbf{d}, \epsilon|\mathbf{X}) \triangleq \inf_{\{\mathcal{C}_n\}} \left\{ \overline{\lim} \left[\frac{1}{n} \log M_n \right] : \overline{\lim}\ \mathbf{P}_e < \epsilon \right\},$$

$$a_0^+ \triangleq R_+(\mathbf{d}, \epsilon|\mathbf{X}) \triangleq \inf_{\{\mathcal{C}_n\}} \left\{ \underline{\lim} \left[\frac{1}{n} \log M_n \right] : \overline{\lim}\ \mathbf{P}_e < \epsilon \right\},$$

as well as the following second order asymptotics

$$R(\mathbf{d}, \epsilon, a_0\,|\mathbf{X}) \triangleq \inf_{\{\mathcal{C}_n\}} \left\{ \overline{\lim} \left[\frac{1}{\sqrt{n}} \log \frac{M_n}{e^{na_0}} \right] : \overline{\lim}\ \mathbf{P}_e < \epsilon \right\},$$

$$R_+(\mathbf{d}, \epsilon, a_0^+|\mathbf{X}) \triangleq \inf_{\{\mathcal{C}_n\}} \left\{ \underline{\lim} \left[\frac{1}{\sqrt{n}} \log \frac{M_n}{e^{na_0^+}} \right] : \overline{\lim}\ \mathbf{P}_e < \epsilon \right\}.$$

For the intrinsic randomness problem, for $\epsilon > 0$, for $\mathbf{d} \in \mathbb{R}_+^{\mathbb{N}}$, and for a code $\mathcal{C}'_n \triangleq (\phi_n, \mathcal{M}_n)$, we define the following first order asymptotics

$$a \triangleq S(\mathbf{d}, \epsilon|\mathbf{X}) \triangleq \sup_{\{\mathcal{C}'_n\}} \left\{ \underline{\lim} \left[\frac{1}{n} \log M_n \right] : \overline{\lim}\ \mathbf{U}_e < \epsilon \right\},$$

$$a^- \triangleq S_-(\mathbf{d}, \epsilon|\mathbf{X}) \triangleq \sup_{\{\mathcal{C}'_n\}} \left\{ \overline{\lim} \left[\frac{1}{n} \log M_n \right] : \overline{\lim}\ \mathbf{U}_e < \epsilon \right\},$$

as well as the following second order asymptotics

$$S(\mathbf{d}, \epsilon, a|\mathbf{X}) \triangleq \sup_{\{\mathcal{C}'_n\}} \left\{ \underline{\lim} \left[\frac{1}{\sqrt{n}} \log \frac{M_n}{e^{na}} \right] : \overline{\lim}\ \mathbf{U}_e < \epsilon \right\},$$

$$S_-(\mathbf{d}, \epsilon, a^-|\mathbf{X}) \triangleq \sup_{\{\mathcal{C}'_n\}} \left\{ \overline{\lim} \left[\frac{1}{\sqrt{n}} \log \frac{M_n}{e^{na^-}} \right] : \overline{\lim}\ \mathbf{U}_e < \epsilon \right\}.$$

We express the first order and the second order asymptotics, defined above, in the following lemmas.

Lemma 4.1.2. *Let $\epsilon > 0$. Let $\mathbf{d} \in \mathbb{R}_+^{\mathbb{N}}$. The first order asymptotics have the following* *expression*

$$\begin{aligned} R(\mathbf{d},\epsilon|\mathbf{X}) &= \overline{H}(\mathbf{0},1-\epsilon|\mathbf{X}),\\ R_+(\mathbf{d},\epsilon|\mathbf{X}) &= \underline{H}(\mathbf{0},1-\epsilon|\mathbf{X}),\\ S(\mathbf{d},\epsilon|\mathbf{X}) &= \underline{H}(\mathbf{d},\epsilon|\mathbf{X}),\\ S_-(\mathbf{d},\epsilon|\mathbf{X}) &= \overline{H}(\mathbf{d},\epsilon|\mathbf{X}), \end{aligned}$$

where,

$$\underline{H}(\mathbf{d},\epsilon|\mathbf{X}) \triangleq \inf_x \left\{ x \colon \overline{\lim}\ \mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} < x - \frac{d_n}{n}\right] \geqslant \epsilon \right\},$$
$$\overline{H}(\mathbf{d},\epsilon|\mathbf{X}) \triangleq \inf_x \left\{ x \colon \underline{\lim}\ \mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} < x - \frac{d_n}{n}\right] \geqslant \epsilon \right\}.$$

Proof. We proceed as in [93] with the lemmas derived in the proof of Proposition 4.1.3.

□

Lemma 4.1.3. *Let $\epsilon > 0$. Let $\mathbf{d} \in \mathbb{R}_+^{\mathbb{N}}$. The second order asymptotics have the following expression*

$$\begin{aligned} R(\mathbf{d},\epsilon,a_0|\mathbf{X}) &= \overline{H}(\mathbf{0},1-\epsilon,a_0|\mathbf{X}),\\ R_+(\mathbf{d},\epsilon,a_0^+|\mathbf{X}) &= \underline{H}(\mathbf{0},1-\epsilon,a_0^+|\mathbf{X}),\\ S(\mathbf{d},\epsilon,a_1|\mathbf{X}) &= \underline{H}(\mathbf{d},\epsilon,a_1|\mathbf{X}),\\ S_-(\mathbf{d},\epsilon,a_1^-|\mathbf{X}) &= \overline{H}(\mathbf{d},\epsilon,a_1^-|\mathbf{X}), \end{aligned}$$

where,

$$\underline{H}(\mathbf{d},\epsilon,a|\mathbf{X}) \triangleq \inf_x \left\{ x \colon \overline{\lim}\ \mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} < a + \frac{x}{\sqrt{n}} - \frac{d_n}{n}\right] \geqslant \epsilon \right\},$$
$$\overline{H}(\mathbf{d},\epsilon,a|\mathbf{X}) \triangleq \inf_x \left\{ x \colon \underline{\lim}\ \mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} < a + \frac{x}{\sqrt{n}} - \frac{d_n}{n}\right] \geqslant \epsilon \right\}.$$

Proof. Let $\epsilon > 0$. If we decide not to use the additional randomness available, then by [93] we obtain

$$\begin{aligned} R(\mathbf{d},\epsilon,a|\mathbf{X}) \leqslant R(\mathbf{0},\epsilon,a|\mathbf{X}) &= \overline{H}(\mathbf{0},1-\epsilon,a|\mathbf{X}),\\ R_+(\mathbf{d},\epsilon,a|\mathbf{X}) \leqslant R_+(\mathbf{0},\epsilon,a|\mathbf{X}) &= \underline{H}(\mathbf{0},1-\epsilon,a|\mathbf{X}), \end{aligned}$$

We proceed as in [93, Theorem 3], using [109, Lemma 1.3.2], which remains unchanged when additional randomness is available at encoder and decoder, to obtain

$$\begin{aligned} R(\mathbf{d}, \epsilon, a|\mathbf{X}) &\geqslant \overline{H}(\mathbf{0}, 1-\epsilon, a|\mathbf{X}), \\ R_+(\mathbf{d}, \epsilon, a|\mathbf{X}) &\geqslant \underline{H}(\mathbf{0}, 1-\epsilon, a|\mathbf{X}). \end{aligned}$$

From [109, Lemma 2.1.2] we now derive a lemma similar to [93, Lemma 4] for the metrics $\mathbf{U}_e$.

Lemma 4.1.4. *For any $\phi_n : \mathcal{X}^n \times \mathcal{U}_{d_n} \to \mathcal{M}_n$ and for any $\gamma_n \in]0, M_n[$*

$$\mathbf{U}_e \geqslant \mathbb{P}\left[P_{X^n}(x^n) > \frac{2^{d_n}}{\gamma_n}\right] - \frac{\gamma_n}{M_n}.$$

Proof. Let $\phi_n : \mathcal{X}^n \times \mathcal{U}_{d_n} \to \mathcal{M}_n$. We apply [109, Lemma 2.1.2] to ϕ_n so that for any $n \in \mathbb{N}$, for any a, for any $\gamma > 0$

$$\begin{aligned} \mathbf{U}_e &= \mathbb{V}[\phi_n(X^n, U_{d_n}), \mathcal{U}_{M_n}] \\ &\geqslant \mathbb{P}[(X^n, U_{d_n}) \notin S_n'(a)] - \mathbb{P}[U_{M_n} \in T_n(a+\gamma)] - e^{-n\gamma} \\ &= \mathbb{P}[X^n \notin S_n(a - d_n/n)] - \mathbb{P}[U_{M_n} \in T_n(a+\gamma)] - e^{-n\gamma}, \end{aligned}$$

where

$$\begin{aligned} S_n'(a) &\triangleq \left\{(x^n, u_{d_n}) \in \mathcal{X}^n \times \mathcal{U}_{d_n} : \frac{1}{n} \log \frac{1}{P_{X^n U_{d_n}}(x^n, u_{d_n})} \geqslant a\right\} \\ &= \left\{(x^n, u_{d_n}) \in \mathcal{X}^n \times \mathcal{U}_{d_n} : \frac{1}{n} \log \frac{1}{P_{X^n}(x^n)} \geqslant a - \frac{d_n}{n}\right\}, \\ S_n(a) &\triangleq \left\{x^n \in \mathcal{X}^n : \frac{1}{n} \log \frac{1}{P_{X^n}(x^n)} \geqslant a\right\}, \\ T_n(a) &\triangleq \left\{u \in \mathcal{U}_{M_n} : \frac{1}{n} \log \frac{1}{P_{U_{M_n}}(u)} < a\right\}. \end{aligned}$$

for any $\gamma_n \in]0, M_n[$, we choose $\gamma \triangleq \frac{1}{n} \log \frac{M_n}{\gamma_n}$ and $a \triangleq \frac{1}{n} \log \gamma_n$, such that $a + \gamma =$

$\frac{1}{n}\log M_n$ and $\mathbb{P}[U_{M_n} \in T_n(a+\gamma)] = 0$. Hence, we obtain

$$\begin{aligned}\mathbf{U}_e &\geqslant \mathbb{P}[X^n \notin S_n(a - d_n/n)] - e^{-n\gamma} \\ &= \mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(x^n)} < a - d_n/n\right] - e^{-n\gamma} \\ &= \mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(x^n)} < \frac{1}{n}\log(\gamma_n \times 2^{-d_n})\right] - \frac{\gamma_n}{M_n}.\end{aligned}$$

□

Then, with Lemma 4.1.4 we proceed as in [93, Theorem 3] to obtain

$$\begin{aligned}S(\mathbf{d}, \epsilon, a|\mathbf{X}) &\leqslant \underline{H}(\mathbf{d}, 1-\epsilon, a|\mathbf{X}), \\ S_{-}(\mathbf{d}, \epsilon, a|\mathbf{X}) &\leqslant \overline{H}(\mathbf{d}, 1-\epsilon, a|\mathbf{X}).\end{aligned}$$

Finally, from [109, Lemma 2.1.1] we derive a lemma similar to [93, Lemma 3] for the metric $\mathbf{U}_e$.

Lemma 4.1.5. *For any $M_n > 0$ and for any $\gamma_n > M_n$, there exists $\phi_n : \mathcal{X}^n \times \mathcal{U}_{d_n} \to \mathcal{M}_n$ such that*

$$\mathbf{U}_e \leqslant \mathbb{P}\left[P_{X^n}(x^n) > \frac{2^{d_n}}{\gamma_n}\right] - \frac{M_n}{\gamma_n}.$$

Proof. By [109, Lemma 2.1.1], there exists $\phi_n : \mathcal{X}^n \times \mathcal{U}_{d_n} \to \mathcal{M}_n$ such that for any a, for any $\gamma > 0$

$$\begin{aligned}\mathbf{U}_e &= \mathbb{V}(\phi_n(X^n, U_{d_n}), \mathcal{U}_{M_n}) \\ &\leqslant \max\left(\mathbb{P}[(X^n, U_{d_n}) \notin S'_n(a+\gamma)], \mathbb{P}[U_{M_n} \in T_n(a)]\right) + e^{-n\gamma} \\ &\leqslant \max\left(\mathbb{P}[X^n \notin S_n(a+\gamma - d_n/n)], \mathbb{P}[U_{M_n} \in T_n(a)]\right) + e^{-n\gamma},\end{aligned}$$

For any $\gamma_n > M_n$, we choose $\gamma \triangleq \frac{1}{n}\log\frac{\gamma_n}{M_n}$ and $a \triangleq \frac{1}{n}\log M_n$, such that $a + \gamma = \frac{1}{n}\log(\gamma_n)$ and $\mathbb{P}[U_{M_n} \in T_n(a)] = 0$. Hence, we obtain

$$\begin{aligned}\mathbf{U}_e &\leqslant \mathbb{P}[X^n \notin S_n(a+\gamma - d_n/n)] + e^{-n\gamma} \\ &= \mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(x^n)} < a + \gamma - d_n/n\right] + e^{-n\gamma} \\ &= \mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(x^n)} < \frac{1}{n}\log(\gamma_n) - d_n/n\right] + \frac{M_n}{\gamma_n}.\end{aligned}$$

□

We conclude, as in [93, Theorem 3], using Lemma 4.1.5, that

$$\begin{aligned} S(\mathbf{d},\epsilon,a|\mathbf{X}) &\geqslant \underline{H}(\mathbf{d},1-\epsilon,a|\mathbf{X}), \\ S_{-}(\mathbf{d},\epsilon,a|\mathbf{X}) &\geqslant \overline{H}(\mathbf{d},1-\epsilon,a|\mathbf{X}). \end{aligned}$$

□

From the first order and the second order asymptotics derived in Lemma 4.1.2 and Lemma 4.1.3, we study the trade-off between $\mathbf{P}_e$ and $\mathbf{U}_e$, for i.i.d. sources following the same method as in [93]. We consider the intrinsic randomness problem for the code $\mathcal{C}_n' = (\phi_n, \mathcal{M}_n)$ and the fixed-length source coding for the code $\mathcal{C}_n = (\phi_n, \psi_n, \mathcal{M}_n)$. We want to know whether there exists a sequence of triplet $\{(\phi_n, \psi_n, \mathcal{M}_n)\}_{n\in\mathbb{N}}$ such that $\overline{\lim}\ \mathbf{P}_e = \epsilon$ and $\overline{\lim}\ \mathbf{U}_e = \epsilon'$, where $\epsilon, \epsilon' \in]0,1[$ can be chosen arbitrarily small, while ensuring d_n negligible compared to n. We first simplify the first order asymptotics of Lemma 4.1.2, when $d_n = o(n)$.

Lemma 4.1.6. *Let $\mathbf{d} \in \mathbb{R}_+^{\mathbb{N}}$. Assume i.i.d. sources and assume $d_n = o(n)$. Then, $\overline{H}(\mathbf{0},\epsilon|\mathbf{X})$, $\underline{H}(\mathbf{0},\epsilon|\mathbf{X})$, $\underline{H}(\mathbf{d},\epsilon|\mathbf{X})$, $\overline{H}(\mathbf{d},\epsilon|\mathbf{X})$, $\underline{H}(\mathbf{0},\epsilon|\mathbf{X})$, $\overline{H}(\mathbf{0},\epsilon|\mathbf{X})$ are all equal to $H(X)$.*

Proof. By the law of large number we already have

$$\left\{\overline{H}(\mathbf{0},\epsilon|\mathbf{X}), \underline{H}(\mathbf{0},\epsilon|\mathbf{X}), \underline{H}(\mathbf{0},\epsilon|\mathbf{X}), \overline{H}(\mathbf{0},\epsilon|\mathbf{X})\right\} = \{H(X)\}.$$

Then, for any $\epsilon_0 > 0$, since $d_n = o(n)$, we have

$$\overline{\lim}\mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} < x - \frac{d_n}{n}\right] \geqslant \overline{\lim}\mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} < x - \epsilon_0\right],$$

thus,

$$\begin{aligned}\underline{H}(\mathbf{d},\epsilon|\mathbf{X}) &= \inf_x \left\{x : \overline{\lim}\mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} < x - \frac{d_n}{n}\right] \geqslant \epsilon\right\} \\ &\leqslant \inf_x \left\{x : \overline{\lim}\mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} < x - \epsilon_0\right] \geqslant \epsilon\right\} \\ &= \epsilon_0 + \inf_x \left\{x : \overline{\lim}\mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} < x\right] \geqslant \epsilon\right\} \\ &= \epsilon_0 + \underline{H}(\mathbf{0},\epsilon|\mathbf{X}). \qquad (94)\end{aligned}$$

We also have by Proposition 4.1.2

$$\underline{H}(\mathbf{d},\epsilon|\mathbf{X}) = S_-(\mathbf{d},\epsilon|\mathbf{X}) \geqslant S_-(0,\epsilon|\mathbf{X}) = \underline{H}(\mathbf{0},\epsilon|\mathbf{X}), \qquad (95)$$

hence, by (94), (95), since ϵ_0 is arbitrary, we have

$$\underline{H}(\mathbf{d},\epsilon|\mathbf{X}) = \underline{H}(\mathbf{0},\epsilon|\mathbf{X}).$$

Similarly,

$$\overline{H}(\mathbf{d},\epsilon|\mathbf{X}) = \overline{H}(\mathbf{0},\epsilon|\mathbf{X}).$$

□

Proposition 4.1.4 (Converse)**.** *Let* $\mathbf{d} \in \mathbb{R}_+^{\mathbb{N}}$*. Assume i.i.d. sources. If* $d_n = o(\sqrt{n})$*, then*

$$\overline{\lim}\ \mathbf{P}_e + \overline{\lim}\ \mathbf{U}_e \geqslant 1.$$

Proof. We prove the two statements in order. Note that, for i.i.d. sources, by Lemma 4.1.2 and Lemma 4.1.6, all the first asymptotics considered are equal, hence by definition of the second order asymptotics, the following must hold

$$S_-(\mathbf{d},\epsilon',a|\mathbf{X}) \geqslant \overline{\lim}\left[\frac{1}{\sqrt{n}}\log\frac{M_n}{e^{na}}\right] \geqslant R(\mathbf{d},\epsilon,a|\mathbf{X}), \qquad (96)$$

$$S(\mathbf{d},\epsilon',a|\mathbf{X}) \geqslant \underline{\lim}\left[\frac{1}{\sqrt{n}}\log\frac{M_n}{e^{na}}\right] \geqslant R_+(\mathbf{d},\epsilon,a|\mathbf{X}). \qquad (97)$$

Assume $d_n = o(\sqrt{n})$. By Equations (96), (97), we have by Lemma 4.1.3

$$\overline{H}(\mathbf{d}, \epsilon', a|\mathbf{X}) \geqslant \overline{H}(\mathbf{0}, 1-\epsilon, a|\mathbf{X}), \tag{98}$$

$$\underline{H}(\mathbf{d}, \epsilon', a|\mathbf{X}) \geqslant \underline{H}(\mathbf{0}, 1-\epsilon, a|\mathbf{X}). \tag{99}$$

Remark that for any $\epsilon_0 > 0$, since $d_n = o(\sqrt{n})$, we have

$$\overline{\lim}\ \mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} < a + \frac{b - d_n/\sqrt{n}}{\sqrt{n}}\right] \geqslant \overline{\lim}\ \mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} < a + \frac{b-\epsilon_0}{\sqrt{n}}\right],$$

hence,

$$\begin{aligned}
&\underline{H}(\mathbf{d}, \epsilon', a|\mathbf{X}) \\
&= \inf_b \left\{ b : \overline{\lim}\,\mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} < a + \frac{b - d_n/\sqrt{n}}{\sqrt{n}}\right] \geqslant \epsilon \right\} \\
&\leqslant \inf_b \left\{ b : \overline{\lim}\,\mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} < a + \frac{b - \epsilon_0}{\sqrt{n}}\right] \geqslant \epsilon \right\} \\
&= \epsilon_0 + \inf_b \left\{ b : \overline{\lim}\,\mathbb{P}\left[\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} < a + \frac{b}{\sqrt{n}}\right] \geqslant \epsilon \right\} \\
&= \epsilon_0 + \underline{H}(\mathbf{0}, \epsilon', a|\mathbf{X}),
\end{aligned}$$

and similarly

$$\overline{H}(\mathbf{d}, \epsilon', a|\mathbf{X}) \leqslant \epsilon_0 + \overline{H}(\mathbf{0}, \epsilon', a|\mathbf{X}).$$

Thus, by (98), (99), we have

$$\epsilon_0 + \overline{H}(\mathbf{0}, \epsilon', a|\mathbf{X}) \geqslant \overline{H}(\mathbf{d}, \epsilon', a|\mathbf{X}) \geqslant \overline{H}(\mathbf{0}, 1-\epsilon, a|\mathbf{X}),$$

$$\epsilon_0 + \underline{H}(\mathbf{0}, \epsilon', a|\mathbf{X}) \geqslant \underline{H}(\mathbf{d}, \epsilon', a|\mathbf{X}) \geqslant \underline{H}(\mathbf{0}, 1-\epsilon, a|\mathbf{X}),$$

which means

$$\overline{H}(\mathbf{0}, \epsilon', a|\mathbf{X}) \geqslant \overline{H}(\mathbf{0}, 1-\epsilon, a|\mathbf{X}),$$

$$\underline{H}(\mathbf{0}, \epsilon', a|\mathbf{X}) \geqslant \underline{H}(\mathbf{0}, 1-\epsilon, a|\mathbf{X}),$$

since ϵ_0 is arbitrary. Thus, for i.i.d. sources, since $\overline{H}(\mathbf{0}, \epsilon, a|\mathbf{X})$ and $\underline{H}(\mathbf{0}, \epsilon, a|\mathbf{X})$ are continuous and increasing w.r.t. ϵ, we find that

$$\overline{\lim}\ \mathbf{P}_e + \overline{\lim}\ \mathbf{U}_e \geqslant 1.$$

□

4.B Proof of Proposition 4.4.2

Let $\epsilon > 0$, $\delta > 0$ and $n \in \mathbb{N}$. Let t, m, and d_n to be expressed later. We know from [106, 107] that there exists an invertible (m, d, m, t, ϵ)-extractor EXT_0, such that (77) is satisfied. Assume that the emitter and the receiver share a sequence U_{d_n} of d_n uniformly distributed bits. As described in Figure 28, we proceed in two steps to encode X^n. First, we perform a typical sequence based compression of X^n to form S, we note this operation $\phi'_n : \mathcal{X}^n \to \mathcal{M}'_n$, such that $S \triangleq \phi'_n(X^n)$, and we note $\psi'_n : \mathcal{M}'_n \to \mathcal{X}^n$ the inverse operation such that

$$\lim_{n\to\infty} \mathbb{P}[X^n \neq \psi'_n \circ \phi'_n(X^n)] = 0. \tag{100}$$

Note that this compression implies

$$\limsup_{n\to\infty} \frac{1}{n} \log||\phi'_n|| \leqslant H(X) + \delta. \tag{101}$$

Then, we apply the extractor EXT_0 to S and U_{d_n}, to form the encoded message $M = \mathrm{EXT}_0(S, U_{d_n})$. We define the encoding function $\phi_n : \mathcal{X}^n \times \mathcal{U}_{d_n} \to \mathcal{M}_n$ as

$$\phi_n(X^n, U_{d_n}) \triangleq M = \mathrm{EXT}_0(\phi'_n(X^n), U_{d_n}),$$

and the decoding function $\psi_n : \mathcal{M}_n \times \mathcal{U}_{d_n} \to \mathcal{X}^n$ as

$$\psi_n(M, U_{d_n}) \triangleq \psi'_n(\mathrm{EXT}_0^{-1}(M, U_{d_n})) = \psi'_n(S) = \psi'_n \circ \phi'_n(X^n), \tag{102}$$

which is possible since EXT_0 is invertible. Note that by (100), (102), we have

$$\lim_{n\to\infty} \mathbb{P}[X^n \neq \psi_n(\phi_n(X^n, U_{d_n}), U_{d_n})] = \lim_{n\to\infty} \mathbb{P}[X^n \neq \psi'_n \circ \phi'_n(X^n)] = 0,$$

and since the sizes of the input and output of the extractor are the same, by (101), we have

$$\limsup_{n\to\infty} \frac{1}{n} \log||\phi_n|| \leqslant H(X) + \delta.$$

Moreover, [106, 107] also shows that $\mathbf{U}_e \leqslant \epsilon$. It remains to show that for any $\epsilon_b > 0$, we can choose $d_n \triangleq \Theta(n^{1/2+\epsilon_b})$. Let $\epsilon_0 > 0$. We first compute

$$\begin{aligned} p_S(s) &= \mathbb{P}[(X^n = s \in \mathcal{T}^n_{\epsilon_0}(X)) \text{ or } (X^n \notin \mathcal{T}^n_{\epsilon_0}(X) \text{ and s is chosen uniformly in } \mathcal{T}^n_{\epsilon_0}(X))] \\ &\leqslant 2^{-n(1-\epsilon_0)H(X)} + \frac{\delta_{\epsilon_0}(n)}{|\mathcal{T}^n_{\epsilon_0}(X)|} \\ &\leqslant 2^{-n(1-\epsilon_0)H(X)} + \frac{\delta_{\epsilon_0}(n)}{1-\delta_{\epsilon_0}(n)} 2^{-n(1-\epsilon_0)H(X)} \\ &= 2^{-n(1-\epsilon_0)H(X)} \left(1 + \frac{\delta_{\epsilon_0}(n)}{1-\delta_{\epsilon_0}(n)}\right), \end{aligned}$$

where $\mathcal{T}^n_{\epsilon_0}(X)$ is the ϵ_0-letter-typical set with respect to p_X [78], $\delta_{\epsilon_0}(n) \triangleq 2|\mathcal{X}|e^{-n\epsilon_0^2\mu_X}$, with $\mu_X \triangleq \min\limits_{x \in supp(p_X)} p_X(x)$.

Hence,

$$H_\infty(S) = -\log(\max p_S(s)) \geqslant n(1-\epsilon_0)H(X) - \log\left[1 + \frac{\delta_{\epsilon_0}(n)}{1-\delta_{\epsilon_0}(n)}\right].$$

We define

$$t \triangleq n(1-\epsilon_0)H(X) - \log\left[1 + \frac{\delta_{\epsilon_0}(n)}{1-\delta_{\epsilon_0}(n)}\right]. \tag{103}$$

Thus, since the input size m of the extractor verifies $m \leqslant \lceil n(1+\epsilon_0)H(X) \rceil$, by (77) and (103) we obtain

$$\begin{aligned} d_n &\leqslant n(1+\epsilon_0)H(X) - t + 2\log[n(1+\epsilon_0)H(X)] + 2\log\frac{1}{\epsilon} + O(1) \\ &= 2n\epsilon_0 H(X) + \log\left[1 + \frac{\delta_{\epsilon_0}(n)}{1-\delta_{\epsilon_0}(n)}\right] + 2\log[n(1+\epsilon_0)H(X)] + 2\log\frac{1}{\epsilon} + O(1). \end{aligned}$$

Then, we choose $\epsilon_0 = n^{-1/2+\epsilon_b}$ and $\epsilon_b > 0$,[2] such that for any $\epsilon_a > \epsilon_b$

$$\frac{d_n}{n^{1/2+\epsilon_a}} \leqslant 2n^{\epsilon_b - \epsilon_a} H(X) + \frac{2}{n^{1/2+\epsilon_a}} \log\frac{1}{\epsilon} + O\left(\frac{\log n}{n^{1/2+\epsilon_a}}\right),$$

which means $d_n = o(n^{1/2+\epsilon_a})$.

[2] The probability of error of the compression scheme is dominated by a term similar to $\delta_{\epsilon_0}(n)$.

4.C Proof of Proposition 4.4.3

Let $\beta \in]0,1/2[$. Let $n \in \mathbb{N}$ and $N \triangleq 2^n$. We set $A^N \triangleq X^N G_N$. We define the following sets.

$$\mathcal{V}_X \triangleq \left\{ i \in [\![1,N]\!] : H\left(A_i | A^{i-1}\right) > 1 - \delta_N \right\},$$
$$\mathcal{H}_X \triangleq \left\{ i \in [\![1,N]\!] : H\left(A_i | A^{i-1}\right) > \delta_N \right\}.$$

These sets cardinalities satisfy the following properties.

Lemma 4.3.7. *The sets $\mathcal{H}_X$ and $\mathcal{V}_X$ verify*

1. $\lim_{N \to +\infty} |\mathcal{H}_X|/N = H(X)$,
2. $\lim_{N \to +\infty} |\mathcal{V}_X|/N = H(X)$,
3. $\lim_{N \to +\infty} |\mathcal{H}_X \backslash \mathcal{V}_X|/N = 0$.

Proof. 1) follows from [87]. 2) follows from Lemma 3.4.1. 3) holds by 1) and 2) since $\mathcal{V}_X \subset \mathcal{H}_X$. □

Lemma 4.3.8. *The output of the encoder $A^N[\mathcal{V}_X]$ is near uniformly distributed in divergence.*

Proof. We have

$$H\left(A^N[\mathcal{V}_X]\right) = \sum_{i \in \mathcal{V}_X} H\left(A_i | A^{i-1}[\mathcal{V}_X]\right) \geqslant \sum_{i \in \mathcal{V}_X} H\left(A_i | A^{i-1}\right) \geqslant |\mathcal{V}_X|(1 - \delta_N),$$

where the first inequality holds because conditioning reduces entropy and the last inequality follows from the definition of $\mathcal{V}_X$. We thus obtain

$$\log 2^{|\mathcal{V}_X|} - H(A^N[\mathcal{V}_X]) \leqslant |\mathcal{V}_X| \delta_N \leqslant N \delta_N.$$

□

Finally, by [87], the receiver can reconstruct X^N from $A^N[\mathcal{V}_X]$ and $I_0 \triangleq A^N[\mathcal{H}_X \backslash V_X]$, where I_0 is encrypted via a one-time pad with the uniform seed shared by Alice and Bob. Hence, by Lemmas 4.3.7, 4.3.8, we obtain a polar code construction for a uniform compression code, whose seed length scales as $o(N)$.

CHAPTER 5

POLAR CODING SCHEMES FOR THE BROADCAST CHANNEL WITH CONFIDENTIAL MESSAGES

5.1 Summary

We develop a low-complexity polar coding scheme for the discrete memoryless broadcast channel with confidential messages under strong secrecy and randomization constraints. This model encompasses the wiretap channel model presented in Section 1.3. Our scheme extends previous work by using an optimal rate of uniform randomness in the stochastic encoder, and avoiding assumptions regarding the symmetry or degraded nature of the channels. The price paid for these extensions is that the encoder and decoders are required to share a secret seed of negligible size and to increase the block length through chaining. We also highlight a close conceptual connection between the proposed polar coding scheme and a random binning proof of the secrecy capacity region. This chapter is based on the results obtained in [110, 111]

5.2 Introduction

In this chapter, we develop a low-complexity polar coding scheme for the broadcast channel with confidential messages [36]. Rather than view randomness as a free resource, which could be used to simulate random numbers at arbitrary rate with no cost, we adopt the point of view put forward in [46,55], in which any randomness used for stochastic encoding must be explicitly accounted for. In particular, our proposed polar coding scheme exploits the optimal rate of randomness identified in [55] and relies on polar codes for channel prefixing.

When specialized to Wyner's wiretap model (see Section 1.3), our scheme is also related to [92], but with a number of notable distinctions. Specifically, while no pre-shared secret seed is required in [92], the coding scheme therein relies on a two-layer construction for which no efficient code construction is presently known [92, Section

3.3]. In contrast, our coding scheme requires a pre-shared secret seed, but at the benefit of only using a single layer of polarization.

The remaining of the chapter is organized as follows. Section 5.3 formally introduces the notation and the model under investigation. Section 5.4 develops a random binning proof of the results in [55], which serves as a guideline for the design of the polar coding scheme. Section 5.5 describes the proposed polar coding scheme in details, while Section 5.6 provides its detailed analysis. Section 5.7 offers some concluding remarks.

5.3 Problem statement

5.3.1 Notation

For $n \in \mathbb{N}$ and $N \triangleq 2^n$, we let $G_n \triangleq \left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]^{\otimes n}$ be the source polarization transform defined in [87]. We note the components of a vector, $X^{1:N}$, of size N, with superscripts, i.e., $X^{1:N} \triangleq (X^1, X^2, \dots, X^N)$. When the context makes clear that we are dealing with vectors, we write X^N in place of $X^{1:N}$.

5.3.2 Channel model and capacity region

We consider the problem of secure communication over a discrete memoryless broadcast channel $(\mathcal{X}, p_{YZ|X}, \mathcal{Y}, \mathcal{Z})$ illustrated in Figure 29. This model generalizes the wiretap channel model presented in Section 1.3. The marginal probabilities $p_{Y|X}$ and $p_{Z|X}$ define two DMCs $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$ and $(\mathcal{X}, p_{Z|X}, \mathcal{Z})$, which we refer to as Bob's channel and Eve's channel, respectively.

Definition 5.3.1. *A $(2^{NR_O}, 2^{NR_M}, 2^{NR_S}, 2^{NR_R}, N)$ code $\mathcal{C}_N$ for the broadcast channel consists of*

- *a common message set $\mathcal{O} \triangleq [\![1, 2^{NR_O}]\!]$*
- *a private message set $\mathcal{M} \triangleq [\![1, 2^{NR_M}]\!]$*
- *a confidential message set $\mathcal{S} \triangleq [\![1, 2^{NR_S}]\!]$*

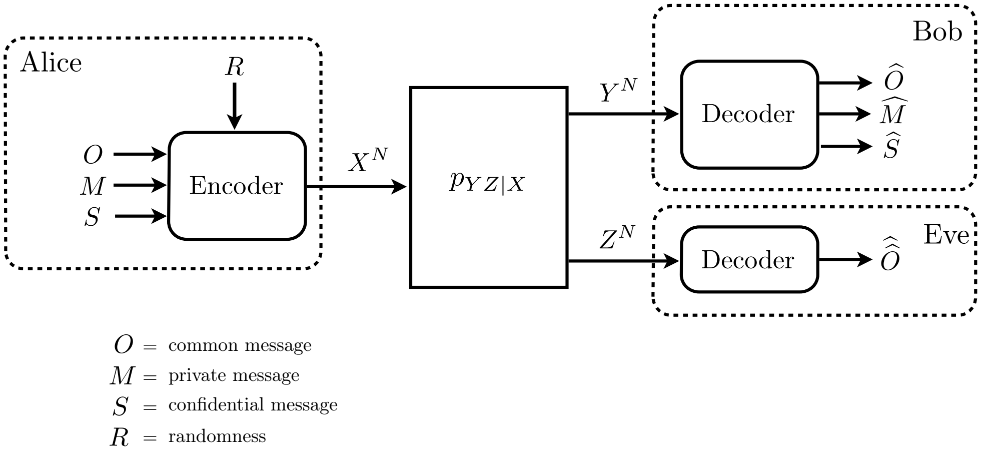


Figure 29. Communication over a broadcast channel with confidential messages. O is a common message that must be reconstructed by both Bob and Eve. S is a confidential message that must be reconstructed by Bob and kept secret from Eve. M is a private message that must be reconstructed by Bob, but may neither be secret nor reconstructed by Eve. R represents an additional randomization sequence used at the encoder.

- *a randomization sequence set* $\mathcal{R} \triangleq [\![1, 2^{NR_R}]\!]$
- *an encoding function* $f : \mathcal{O} \times \mathcal{M} \times \mathcal{S} \times \mathcal{R} \to \mathcal{X}^N$*, which maps the messages* (o, m, s) *and the randomness* r *to a codeword* x^N
- *a decoding function* $g : \mathcal{Y}^N \to \mathcal{O} \times \mathcal{M} \times \mathcal{S}$*, which maps each observation of Bob's channel* y^N *to the messages* $(\hat{o}, \hat{m}, \hat{s})$
- *a decoding function* $h : \mathcal{Z}^N \to \mathcal{O}$*, which maps each observation of Eve's channel* z^N *to the message* $\hat{\hat{o}}$

For uniformly distributed O, M, S, and R, the performance of a $(2^{NR_O}, 2^{NR_M}, 2^{NR_S}, 2^{NR_R}, N)$ code $\mathcal{C}_N$ for the broadcast channel is measured in terms of its probability of error

$$\mathbf{P}_e(\mathcal{C}_N) \triangleq \mathbb{P}\left[(\widehat{O}, \widehat{M}, \widehat{S}) \neq (O, M, S) \text{ or } \widehat{\widehat{O}} \neq O\right],$$

and its leakage of information about the confidential message to Eve

$$\mathbf{L}_e(\mathcal{C}_N) \triangleq I(S; Z^N).$$

Definition 5.3.2. *A rate tuple (R_O, R_M, R_S, R_R) is achievable for the broadcast channel if there exists a sequence of $(2^{NR_O}, 2^{NR_M}, 2^{NR_S}, 2^{NR_R}, N)$ codes $\{\mathcal{C}_N\}_{N\geqslant 1}$ such that*

$$\lim_{N\to\infty} \mathbf{P}_e(\mathcal{C}_N) = 0, (\textit{reliability condition})$$

$$\lim_{N\to\infty} \mathbf{L}_e(\mathcal{C}_N) = 0. (\textit{strong secrecy})$$

The achievable region $\mathcal{R}_{\mathrm{BCC}}$ is defined as the closure of the set of all achievable rate quadruples.

The exact characterization of $\mathcal{R}_{\mathrm{BCC}}$ was obtained in [55].

Theorem 5.3.1 ([55])**.** *$\mathcal{R}_{\mathrm{BCC}}$ is the closed convex set consisting of the quadruples (R_O, R_M, R_S, R_R) for which there exist auxiliary random variables (U, V) such that $U - V - X - (Y, Z)$ and*

$$R_O \leqslant \min[I(U;Y), I(U;Z)],$$

$$R_O + R_M + R_S \leqslant I(V;Y|U) + \min[I(U;Y), I(U;Z)],$$

$$R_S \leqslant I(V;Y|U) - I(V;Z|U),$$

$$R_M + R_R \geqslant I(X;Z|U),$$

$$R_R \geqslant I(X;Z|V).$$

The main contribution of the present work is to develop a polar coding scheme achieving the rates in $\mathcal{R}_{\mathrm{BCC}}$.

5.4 A binning approach to code design: from random binning to polar binning

In this section, we argue that our construction of polar codes for the broadcast channel with confidential messages is essentially the constructive counterpart of a *random binning* proof of the region $\mathcal{R}_{\mathrm{BCC}}$. While *random coding* is often the natural tool to address channel coding problems, random binning is already found in [112] to establish the strong secrecy of the wiretap channel, and is the tool of choice in quantum information theory [113]; there has also been a renewed interest for random binning proofs in multi-user information theory, motivated in part by [114]. In Section 5.4.1, we sketch a random binning proof of the characterization of $\mathcal{R}_{\text{BCC}}$ established in [55], which may be viewed as a refinement of the analysis in [114] to obtain a more precise characterization of the stochastic encoder. While the results we derive are not new, we use this alternative proof in Section 5.4.2 to obtain high-level insight into the construction of polar codes. The main benefit is to clearly highlight the crucial steps of the construction in Section 5.5 and of its analysis in Section 5.6. In particular, the rate conditions developed in the random binning proof of Section 5.4.1 directly translate into the definition of the polarization sets in Section 5.4.2.

5.4.1 Information-theoretic random binning

Information-theoretic random binning proofs rely on the following well-known lemmas. We use the notation $\delta(N)$ to denote an unspecified positive function of N that vanishes as N goes to infinity.

Lemma 5.4.1 (Source-coding with side information)**.** *Consider a DMS $(\mathcal{X}\times\mathcal{Y}, p_{XY})$. For each $x^N \in \mathcal{X}^N$, assign an index $\Phi(x^N) \in [\![1, 2^{NR}]\!]$ uniformly at random. If $R > \mathbb{H}(X|Y)$, then $\exists N_0$ such that $\forall N \geqslant N_0$, there exists a deterministic function $g_N : [\![1, 2^{NR}]\!] \times \mathcal{Y}^N \to \mathcal{X}^N : (\Phi(x^N), y^N) \mapsto \hat{x}^N$ such that*

$$\mathbb{E}_\Phi\big(\mathbb{V}\big(p_{X^N X^N}, p_{X^N g_N(Y^N)}\big)\big) \leqslant \delta(N).$$

Lemma 5.4.2 (Privacy amplification, channel intrinsic randomness, output statistics of random binning)**.** *Consider a DMS $(\mathcal{X}\times\mathcal{Z}, p_{XZ})$ and let $\epsilon > 0$. For each $x^N \in \mathcal{X}^N$, assign an index $\Psi(x^N) \in [\![1, 2^{NR}]\!]$ uniformly at random. Denote by q_M the uniform distribution on $[\![1, 2^{NR}]\!]$. If $R < \mathbb{H}(X|Z)$, then $\exists N_0$ such that $\forall N \geqslant N_0$*

$$\mathbb{E}_\Psi\big(\mathbb{V}\big(p_{\Psi(X^N)Z^N}, q_M p_{Z^N}\big)\big) \leqslant \delta(N).$$

One may obtain more explicit results regarding the convergence to zero in Lemma 5.4.1 and Lemma 5.4.2, but we ignore this for brevity.

The principle of a random binning proof of Theorem 5.3.1 is to consider a DMS $(\mathcal{U}\times\mathcal{V}\times\mathcal{X}\times\mathcal{Y}\times\mathcal{Z},p_{UVXYZ})$ such that $U-V-X-YZ$, and to assign two types of indices to source sequences by random binning. The first type identifies subset of sequences that play the roles of codebooks, while the second type labels sequences with indices that can be thought of as messages. As explained in the next paragraphs, the crux of the proof is to show that the binning can be "inverted," so that the sources may be generated from independent choices of uniform codebooks and messages.

Common message encoding. We introduces two indices $\psi^U \in \llbracket 1,2^{N\rho_U}\rrbracket$ and $o\in\llbracket 1,2^{NR_O}\rrbracket$ by random binning on u^N such that:

- $\rho_U > \max\left(\mathbb{H}(U|Y),\mathbb{H}(U|Z)\right)$, so that Lemma 5.4.1 ensures that the knowledge of ψ^U allows Bob and Eve to reconstruct the sequence u^N with high probability knowing y^N or z^N, respectively;
- $\rho_U + R_O < \mathbb{H}(U)$, so that Lemma 5.4.2 ensures that the indices ψ^U and o are almost uniformly distributed and independent of each other.

The binning scheme induces a joint distribution $p_{U^N\Psi^U O}$. To convert the binning scheme into a channel coding scheme, Alice operates as follows. Upon sampling indices $\tilde{\psi}^U\in\llbracket 1,2^{N\rho_U}\rrbracket$ and $\tilde{o}\in\llbracket 1,2^{NR_O}\rrbracket$ from independent uniform distributions, Alice *stochastically* encodes them into a sequence $\tilde{u}^N$ drawn according to $p_{U^N|\Psi^U O}(\tilde{u}^N|\tilde{\psi}^U,\tilde{o})$. The choice of rates above guarantees that the joint distribution $p_{\tilde{U}^N\tilde{\Psi}^U\tilde{O}}$ approximates the distribution $p_{U^N\Psi^U O}$ in variational distance, so that disclosing $\tilde{\psi}^U$ allows Bob and Eve to decode the sequence $\tilde{u}^N$.

Secret and private message encoding. Following the same approach, we introduce three indices $\psi^{V|U}\in\llbracket 1,2^{N\rho_{V|U}}\rrbracket$, $s\in\llbracket 1,2^{NR_S}\rrbracket$, and $m\in\llbracket 1,2^{NR_M}\rrbracket$ by random binning on v^N such that

- $\rho_{V|U} > \mathbb{H}(V|UY)$, to ensure that knowing $\psi^{V|U}$, u^N, and y^N, Bob may reconstruct the sequence v^N;

- $\rho_{V|U} + R_S < \mathbb{H}(V|UZ)$, to ensure that the indices are almost uniformly distributed and independent of each other, as well as of the source sequences U^N and Z^N.

The binning scheme induces a joint distribution $p_{V^N U^N \Psi^{V|U} SM}$. To obtain a channel coding scheme, Alice encodes the realizations of independent and uniformly distributed indices $\tilde{\psi}^{V|U} \in [\![1, 2^{N\rho_{V|U}}]\!]$, $\tilde{s} \in [\![1, 2^{NR_S}]\!]$, $\tilde{m} \in [\![1, 2^{NR_M}]\!]$, and the sequence $\tilde{u}^N$, into a sequence $\tilde{v}^N$ drawn according to the distribution

$$p_{V^N|U^N \Psi^{V|U} SM}(\tilde{v}^N|\tilde{u}^N, \tilde{\psi}^{V|U}, \tilde{s}, \tilde{m}).$$

The resulting joint distribution is again a close approximation of $p_{V^N U^N \Psi^{V|U} SM}$, so that the scheme inherits the reliability and secrecy properties of the random binning scheme upon disclosing $\tilde{\psi}^{V|U}$.

Channel prefixing. Finally, we introduce the indices $\psi^{X|V} \in [\![1, 2^{N\rho_{V|X}}]\!]$ and $r \in [\![1, 2^{NR_R}]\!]$ by random binning on x^N such that

- $\rho_{X|V} < \mathbb{H}(X|VZ)$ to ensure that $\psi^{X|V}$ is independent of the source sequences V^N and Z^N;

- $\rho_{X|V} + R_R < \mathbb{H}(X|V)$ to ensure that the indices are almost uniformly distributed and independent of each other, as well as of the source sequences V^N.

The binning scheme induces a joint distribution $p_{X^N V^N U^N \Psi^{X|V} R}$. To obtain a channel prefixing scheme, Alice encodes the realizations of uniformly distributed indices $\tilde{\psi}^{X|V}$ and $\tilde{r}$, and the previously obtained $\tilde{v}^N$ into a sequence $\tilde{x}^N$ drawn according to $p_{X^N|V^N \Psi^{X|V} R}(\tilde{x}^N|\tilde{v}^N \tilde{\psi}^{X|V} \tilde{r})$. The resulting joint distribution induced is once again a close approximation of $p_{X^N V^N U^N \Psi^{X|V} R}$.

Chaining to de-randomize the codebooks. The downside of the schemes described earlier is that they require sharing the indices $\tilde{\psi}^U$, $\tilde{\psi}^{V|U}$, and $\tilde{\psi}^{X|V}$, identifying the codebooks between Alice, Bob, and Eve; however, the rate cost may be amortized by reusing the *same* indices over sequences of k blocks. Specifically, the union bound shows that the average error probability over k blocks is at most k times that of an individual block, and a hybrid argument shows that the information leakage over k blocks is at most k times that of an individual block. Consequently, for k and N large enough, the impact on the transmission rates is negligible.

Total amount of randomness. The total amount of randomness required for encoding includes not only the explicit random numbers used for channel prefixing but also all the randomness required in the stochastic encoding to approximate the source distribution. One can show that the rate randomness specifically used in the stochastic encoding is negligible; we omit the proof of this result for random binning, but this is analyzed precisely for polar codes in Section 5.6.

By combining all the rate constraints above and perform Fourier-Motzkin elimination, one recovers the rates in Theorem 5.3.1.

5.4.2 Binning with polar codes

The main observation to translate the analysis of Section 5.4.1 into a polar coding scheme is that Lemma 5.4.1 and Lemma 5.4.2 have the following counterparts in terms of source polarization.

Lemma 5.4.3 (adapted from [87]). *Consider a DMS $(\mathcal{X} \times \mathcal{Y}, p_{XY})$. For each $x^{1:N} \in \mathbb{F}_2^N$ polarized as $u^{1:N} \triangleq G_n x^{1:N}$, let $u^{1:N}[\mathcal{H}_{X|Y}]$ denote the high entropy bits of $u^{1:N}$ in positions $\mathcal{H}_{X|Y} \triangleq \{i \in [\![1, N]\!] : \mathbb{H}\big(U^i|U^{1:i-1}Y^N\big) > \delta_N\}$ and $\delta_N \triangleq 2^{-N^\beta}$ with $\beta \in]0, \frac{1}{2}[$. For every $i \in [\![1, N]\!]$, sample $\tilde{u}^{1:N}$ from the distribution*

$$\tilde{p}_{U^i|U^{1:i-1}}(\tilde{u}^i|\tilde{u}^{1:i-1}) \triangleq \begin{cases} \mathbb{1}\left\{\tilde{u}^i = u^i\right\} & \text{if } i \in \mathcal{H}_{Y|X} \\ p_{U^i|U^{1:i-1}Y^N}(\tilde{u}^i|\tilde{u}^{1:i-1}y^N) & \text{if } i \in \mathcal{H}_{Y|X}^c. \end{cases}$$

and create $\tilde{x}^{1:N} = \tilde{u}^{1:N} G_n$. Then,

$$\mathbb{V}(p_{X^{1:N}X^{1:N}}, p_{X^{1:N}\tilde{X}^N}) = O(N\delta_N),$$

and $\lim_{N\to\infty} \frac{1}{N}\left|\mathcal{H}_{X|Y}\right| = \mathbb{H}(X|Y)$*.*

In other words, the high entropy bits in positions $\mathcal{H}_{X|Y}$ play the same role as the random binning index in Lemma 5.4.1. However, note that the construction of $\tilde{x}^{1:N}$ in Lemma 5.4.3 is explicitly stochastic.

Lemma 5.4.4 (adapted from Section 3.4.2)**.** *Consider a DMS $(\mathcal{X}\times\mathcal{Z}, p_{XZ})$. For each $x^{1:N} \in \mathbb{F}_2^N$ polarized as $u^{1:N} \triangleq G_n x^{1:N}$, let $u^{1:N}[\mathcal{V}_{X|Z}]$ denote the very high entropy bits of $u^{1:N}$ in positions $\mathcal{V}_{X|Z} \triangleq \{i \in [\![1,N]\!] : \mathbb{H}\big(U^i|U^{1:i-1}Z^{1:N}\big) > 1-\delta_N\}$ and $\delta_N \triangleq 2^{-N^\beta}$ with $\beta \in]0, \frac{1}{2}[$. Denote by $q_{U^{1:N}[\mathcal{V}_{X|Z}]}$ the uniform distribution of bits in positions $\mathcal{V}_{X|Z}$. Then,*

$$\mathbb{V}\Big(p_{U^{1:N}[\mathcal{V}_{X|Z}]Z^{1:N}}, q_{U^{1:N}[\mathcal{V}_{X|Z}]}p_{Z^{1:N}}\Big) = O(\sqrt{N\delta_N}),$$

and $\lim_{N\to\infty} \frac{1}{N}\left|\mathcal{V}_{X|Z}\right| = \mathbb{H}(X|Z)$ *by Lemma 3.4.1.*

The very high entropy bits in positions $\mathcal{V}_{X|Z}$ therefore play the same role as the random binning index in Lemma 5.4.2.

This suggests that any result obtained from random binning could also be derived using source polarization as a linear and low-complexity alternative; intuitively, information theoretic constraints resulting from Lemma 5.4.1 translate into the use of "high entropy" sets $\mathcal{H}$, while those resulting from Lemma 5.4.2 translate into the use of "very high entropy" sets $\mathcal{V}$. However, unlike the indices resulting from random binning, the high entropy and very high entropy sets may not necessarily be aligned, and the precise design of a polar coding scheme requires more care.

In the remainder of the chapter, we consider a DMS $(\mathcal{U}\times\mathcal{V}\times\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}, p_{UVXYZ})$ such that $U-V-X-YZ$, $I(V;Y|U) - I(V;Z|U) > 0$, and $|\mathcal{X}|= |\mathcal{U}|= |\mathcal{V}|= 2$. The extension to larger alphabets is obtained following ideas in [115]. We also assume without loss of generality $I(U;Y) \leqslant I(U;Z)$, the case $I(U;Y) > I(U;Z)$ is treated similarly.

Common message encoding. Define the polar transform of $U^{1:N}$, as $A^{1:N} \triangleq$

$U^{1:N}G_n$ and the associated sets

$$\mathcal{H}_U \triangleq \left\{i \in [\![1,N]\!] : H(A^i|A^{1:i-1}) > \delta_N\right\}, \tag{104}$$

$$\mathcal{V}_U \triangleq \left\{i \in [\![1,N]\!] : H(A^i|A^{1:i-1}) > 1-\delta_N\right\}, \tag{105}$$

$$\mathcal{H}_{U|Y} \triangleq \left\{i \in [\![1,N]\!] : H(A^i|A^{1:i-1}Y^{1:N}) > \delta_N\right\}, \tag{106}$$

$$\mathcal{H}_{U|Z} \triangleq \left\{i \in [\![1,N]\!] : H(A^i|A^{1:i-1}Z^{1:N}) > \delta_N\right\}. \tag{107}$$

If we could guarantee that $\mathcal{H}_{U|Z} \subseteq \mathcal{H}_{U|Y} \subseteq \mathcal{V}_U$, then we could directly mimic the information-theoretic random binning proof. We would use random bits in positions $\mathcal{H}_{U|Z}$ to identify the code, random bits in positions $\mathcal{V}_U \setminus \mathcal{H}_{U|Z}$ for the message, successive cancellation encoding to compute the bits in positions $\mathcal{V}_U^c$ and approximate the source distribution, and chaining to amortize the rate cost of the bits in positions $\mathcal{H}_{U|Z}$. Unfortunately, the inclusion $\mathcal{H}_{U|Y} \subseteq \mathcal{H}_{U|Z}$ is not true in general, and one must also use chaining as to "realign" the sets of indices. Furthermore, only the inclusions $\mathcal{H}_{U|Z} \subseteq \mathcal{H}_U$ and $\mathcal{H}_{U|Y} \subseteq \mathcal{H}_U$ are true in general, so that the bits in positions $\mathcal{H}_{U|Z} \cap \mathcal{V}_U^c$ and $\mathcal{H}_{U|Y} \cap \mathcal{V}_U^c$ must be transmitted separately. The precise coding scheme is detailed in Section 5.5.1.

Secret and private messages encoding. Define the polar transform of $V^{1:N}$ as $B^{1:N} \triangleq V^{1:N}G_n$ and the associated sets

$$\mathcal{V}_{V|U} \triangleq \left\{i \in [\![1,N]\!] : H(B^i|B^{1:i-1}U^{1:N}) > 1-\delta_N\right\}, \tag{108}$$

$$\mathcal{V}_{V|UZ} \triangleq \left\{i \in [\![1,N]\!] : H(B^i|B^{1:i-1}U^{1:N}Z^{1:N}) > 1-\delta_N\right\}, \tag{109}$$

$$\mathcal{H}_{V|UY} \triangleq \left\{i \in [\![1,N]\!] : H(B^i|B^{1:i-1}U^{1:N}Y^{1:N}) > \delta_N\right\}, \tag{110}$$

$$\mathcal{V}_{V|UY} \triangleq \left\{i \in [\![1,N]\!] : H(B^i|B^{1:i-1}U^{1:N}Y^{1:N}) > 1-\delta_N\right\}, \tag{111}$$

$$\mathcal{M}_{UVZ} \triangleq \mathcal{V}_{V|U} \backslash \mathcal{V}_{V|UZ}. \tag{112}$$

If the inclusion $\mathcal{H}_{V|UY} \subseteq \mathcal{V}_{V|UZ}$ were true, then we would place random bits identifying the codebook in positions $\mathcal{H}_{V|UY}$, random bits describing the secret message in positions $\mathcal{V}_{V|UZ} \setminus \mathcal{H}_{V|UY}$, random bits describing the private message in positions

$\mathcal{V}_{V|U} \setminus \mathcal{V}_{V|UZ}$, use successive cancellation encoding to compute the bits in positions $\mathcal{V}_{V|U}^c$ and approximate the source distribution, and use chaining to amortize the rate cost of the bits in positions $\mathcal{H}_{V|UY}$. This is unfortunately again not directly possible in general, and one needs to exploit chaining to realign the indices, and transmit the bits in positions $\mathcal{H}_{V|UY} \cap \mathcal{V}_{V|U}^c$ separately and secretly to Bob. The precise coding scheme is detailed in Section 5.5.2.

Channel prefixing. Finally, define the polar transform of $X^{1:N}$ as $T^{1:N} \triangleq X^{1:N} G_n$ and the associated sets

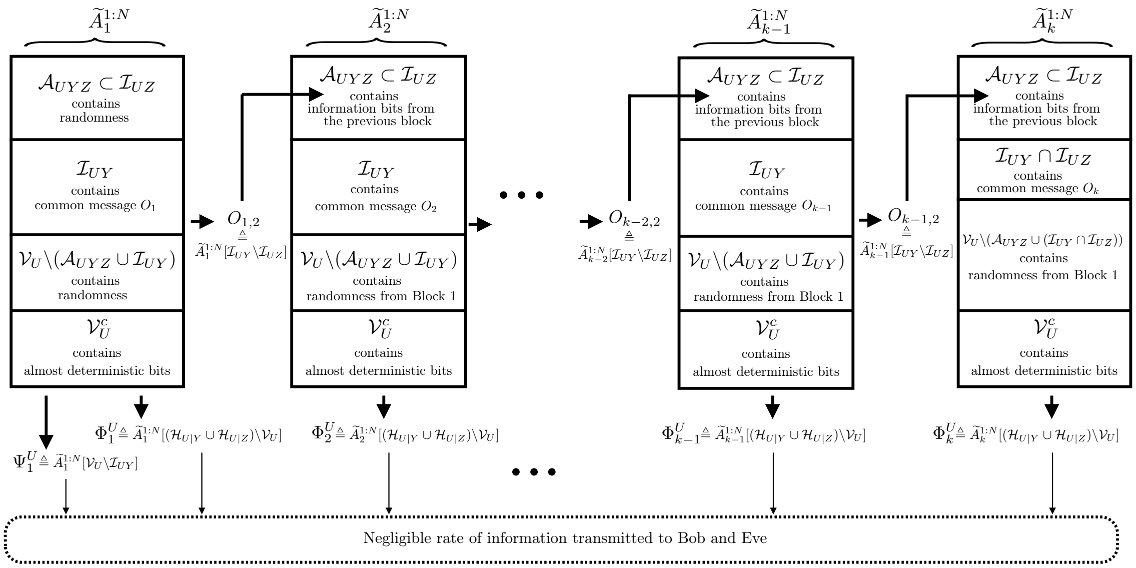
$$\mathcal{V}_{X|V} \triangleq \left\{ i \in [\![1, N]\!] : H(T^i | T^{1:i-1} V^{1:N}) > 1 - \delta_N \right\}, \tag{113}$$

$$\mathcal{V}_{X|VZ} \triangleq \left\{ i \in [\![1, N]\!] : H(T^i | T^{1:i-1} V^{1:N} Z^{1:N}) > 1 - \delta_N \right\}. \tag{114}$$

One performs channel prefixing by placing random bits identifying the code in positions $\mathcal{V}_{X|VZ}$, random bits describing the randomization sequence in positions $\mathcal{V}_{X|V} \setminus \mathcal{V}_{X|VZ}$, and using successive cancellation encoding to compute the bits in positions $\mathcal{V}_{X|V}^c$ and approximate the source distribution. Chaining is finally used to amortize the cost of randomness for describing the code. The precise coding scheme is detailed in Section 5.5.3.

5.5 Polar coding scheme

In this section, we describe the details of the polar coding scheme resulting from the discussion of the previous section. Recall that the joint probability distribution p_{UVXYZ} of the original source is fixed and defined as in Section 5.4.2. As alluded to earlier, we perform the encoding over k blocks of size N. We use the subscript $i \in [\![1, k]\!]$ to denote random variables associated to encoding Block i. The chaining constructions corresponding to the encoding of the common, secret, and private messages, and randomization sequence, are described in Section 5.5.1, Section 5.5.2, and Section 5.5.3, respectively. Although each chaining is described independently, all messages should be encoded in every block before moving to the next. Specifically, in



$\widetilde{A}_1^{1:N}$
$\widetilde{A}_2^{1:N}$
$\widetilde{A}_{k-1}^{1:N}$
$\widetilde{A}_k^{1:N}$
$\mathcal{A}_{UYZ} \subset \mathcal{I}_{UZ}$ contains randomness
$\mathcal{I}_{UY}$ contains common message O_1
$\mathcal{V}_U \backslash (\mathcal{A}_{UYZ} \cup \mathcal{I}_{UY})$ contains randomness
$\mathcal{V}_U^c$ contains almost deterministic bits
$O_{1,2} \triangleq \widetilde{A}_1^{1:N}[\mathcal{I}_{UY} \backslash \mathcal{I}_{UZ}]$
$\mathcal{A}_{UYZ} \subset \mathcal{I}_{UZ}$ contains information bits from the previous block
$\mathcal{I}_{UY}$ contains common message O_2
$\mathcal{V}_U \backslash (\mathcal{A}_{UYZ} \cup \mathcal{I}_{UY})$ contains randomness from Block 1
$\mathcal{V}_U^c$ contains almost deterministic bits
$O_{k-2,2} \triangleq \widetilde{A}_{k-2}^{1:N}[\mathcal{I}_{UY} \backslash \mathcal{I}_{UZ}]$
$\mathcal{A}_{UYZ} \subset \mathcal{I}_{UZ}$ contains information bits from the previous block
$\mathcal{I}_{UY}$ contains common message O_{k-1}
$\mathcal{V}_U \backslash (\mathcal{A}_{UYZ} \cup \mathcal{I}_{UY})$ contains randomness from Block 1
$\mathcal{V}_U^c$ contains almost deterministic bits
$O_{k-1,2} \triangleq \widetilde{A}_{k-1}^{1:N}[\mathcal{I}_{UY} \backslash \mathcal{I}_{UZ}]$
$\mathcal{A}_{UYZ} \subset \mathcal{I}_{UZ}$ contains information bits from the previous block
$\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}$ contains common message O_k
$\mathcal{V}_U \backslash (\mathcal{A}_{UYZ} \cup (\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}))$ contains randomness from Block 1
$\mathcal{V}_U^c$ contains almost deterministic bits
$\Psi_1^U \triangleq \widetilde{A}_1^{1:N}[\mathcal{V}_U \backslash \mathcal{I}_{UY}]$
$\Phi_1^U \triangleq \widetilde{A}_1^{1:N}[(\mathcal{H}_{U|Y} \cup \mathcal{H}_{U|Z}) \backslash \mathcal{V}_U]$
$\Phi_2^U \triangleq \widetilde{A}_2^{1:N}[(\mathcal{H}_{U|Y} \cup \mathcal{H}_{U|Z}) \backslash \mathcal{V}_U]$
$\Phi_{k-1}^U \triangleq \widetilde{A}_{k-1}^{1:N}[(\mathcal{H}_{U|Y} \cup \mathcal{H}_{U|Z}) \backslash \mathcal{V}_U]$
$\Phi_k^U \triangleq \widetilde{A}_k^{1:N}[(\mathcal{H}_{U|Y} \cup \mathcal{H}_{U|Z}) \backslash \mathcal{V}_U]$
Negligible rate of information transmitted to Bob and Eve

Figure 30. Chaining for the encoding of the $\widetilde{A}_i^{1:N}$'s, which corresponds to the encoding of the common messages.

every block $i \in [\![1, k-1]\!]$, Alice successively encodes the common message, the secret and private messages, and performs channel prefixing, before she moves to the next block $i+1$.

5.5.1 Common message encoding

In addition to the polarization sets defined in (104)-(107) we also define

$$\begin{aligned}
\mathcal{I}_{UY} &\triangleq \mathcal{V}_U \backslash \mathcal{H}_{U|Y}, \\
\mathcal{I}_{UZ} &\triangleq \mathcal{V}_U \backslash \mathcal{H}_{U|Z}, \\
\mathcal{A}_{UYZ} &\triangleq \text{any subset of } \mathcal{I}_{UZ} \backslash \mathcal{I}_{UY} \text{ with size } |\mathcal{I}_{UY} \backslash \mathcal{I}_{UZ}|.
\end{aligned}$$

Note that $\mathcal{A}_{UYZ}$ exists since we have assumed $I(U;Y) \leqslant I(U;Z)$. In fact,

$$|\mathcal{I}_{UZ} \backslash \mathcal{I}_{UY}| - |\mathcal{I}_{UY} \backslash \mathcal{I}_{UZ}| = |\mathcal{I}_{UZ}| - |\mathcal{I}_{UY}| \geqslant 0.$$

The encoding procedure with chaining is summarized in Figure 30.

In Block 1, the encoder forms $\widetilde{U}_1^{1:N}$ as follows. Let O_1 be a vector of $|\mathcal{I}_{UY}|$ uniformly distributed information bits that represents the common message to be reconstructed by Bob and Eve. Upon observing a realization o_1, the encoder samples

$\widetilde{a}_1^{1:N}$ from the distribution $\widetilde{p}_{A_1^{1:N}}$ defined as

$$\widetilde{p}_{A_1^j|A_1^{1:j-1}}(a_1^j|a_1^{1:j-1}) \triangleq \begin{cases} \mathbb{1}\left\{a_1^j = o_1^j\right\} & \text{if } j \in \mathcal{I}_{UY} \\ 1/2 & \text{if } j \in \mathcal{V}_U \backslash \mathcal{I}_{UY} \\ p_{A^j|A^{1:j-1}}(a_1^j|a_1^{1:j-1}) & \text{if } j \in \mathcal{V}_U^c \end{cases}, \tag{115}$$

where the components of o_1 have been indexed by the set of indices $\mathcal{I}_{UY}$ for convenience, so that $O_1 \triangleq \widetilde{A}_1^{1:N}[\mathcal{I}_{UY}]$. The random bits that identify the codebook and that are required to reconstruct $\widetilde{A}_1^{1:N}$ are $\widetilde{A}_1^{1:N}[\mathcal{H}_{U|Z}]$ for Eve and $\widetilde{A}_1^{1:N}[\mathcal{H}_{U|Y}]$ for Bob. Moreover, we note

$$\begin{aligned} \Psi_1^U &\triangleq \widetilde{A}_1^{1:N}[\mathcal{V}_U \backslash \mathcal{I}_{UY}] = \widetilde{A}_1^{1:N}[\mathcal{V}_U \cap \mathcal{H}_{U|Y}], \\ \Phi_1^U &\triangleq \widetilde{A}_1^{1:N}[(\mathcal{H}_{U|Y} \cup \mathcal{H}_{U|Z}) \cap \mathcal{V}_U^c]. \end{aligned}$$

Both Ψ_1^U and Φ_1^U are publicly transmitted to both Bob and Eve. Note that, unlike in the random binning proof, the use of polarization forces us to distinguish the part Ψ_1^U that is nearly uniform from the part Φ_1^U that is not. We show later that the rate cost of this additional transmission is negligible. We also write $O_1 \triangleq [O_{1,1}, O_{1,2}]$, where $O_{1,1} \triangleq \widetilde{A}_1^{1:N}[\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}]$ and $O_{1,2} \triangleq \widetilde{A}_1^{1:N}[\mathcal{I}_{UY} \backslash \mathcal{I}_{UZ}]$. We will retransmit $O_{1,2}$ in the next block. Finally, we compute $\widetilde{U}_1^{1:N} \triangleq \widetilde{A}_1^{1:N} G_n$.

In Block $i \in [\![2, k-1]\!]$, the encoder forms $\widetilde{A}_1^{1:N}$ as follows. Let O_i be a vector of $|\mathcal{I}_{UY}|$ uniformly distributed information bits representing the common message in that block. Upon observing the realizations o_i and o_{i-1}, the encoder draws $\widetilde{a}_i^{1:N}$ from the distribution $\widetilde{p}_{A_i^{1:N}}$ defined as follows.

$$\widetilde{p}_{A_i^j|A_i^{1:j-1}}(a_i^j|a_i^{1:j-1}) \triangleq \begin{cases} \mathbb{1}\left\{a_i^j = o_i^j\right\} & \text{if } j \in \mathcal{I}_{UY} \\ \mathbb{1}\left\{a_i^j = o_{i-1,2}^j\right\} & \text{if } j \in \mathcal{A}_{UYZ} \\ \mathbb{1}\left\{a_i^j = (\psi_1^U)^j\right\} & \text{if } j \in \mathcal{V}_U \backslash (\mathcal{I}_{UY} \cup \mathcal{A}_{UYZ}) \\ p_{A^j|A^{1:j-1}}(a_i^j|a_i^{1:j-1}) & \text{if } j \in \mathcal{V}_U^c \end{cases}, \tag{116}$$

where the components of o_i, $o_{i-1,2}$, and ψ_1^U, have been indexed by the set of indices $\mathcal{I}_{UY}$, $\mathcal{A}_{UYZ}$, and $\mathcal{V}_U \backslash (\mathcal{I}_{UY} \cup \mathcal{A}_{UYZ})$, respectively. Consequently, note that

$$O_i = \widetilde{A}_i^{1:N}[\mathcal{I}_{UY}] \text{ and } O_{i-1,2} = \widetilde{A}_i^{1:N}[\mathcal{A}_{UYZ}].$$

The random bits that identify the codebook and that are required to reconstruct $\widetilde{A}_i^{1:N}$ are $\widetilde{A}_i^{1:N}[\mathcal{H}_{U|Y}]$ for Bob and $\widetilde{A}_i^{1:N}[\mathcal{H}_{U|Z}]$ for Eve. Parts of these bits depend on messages in previous blocks. For the others, we define

$$\Psi_i^U \triangleq \widetilde{A}_i^{1:N}[\mathcal{V}_U \backslash (\mathcal{I}_{UY} \cup \mathcal{A}_{UYZ})],$$
$$\Phi_i^U \triangleq \widetilde{A}_i^{1:N}[(\mathcal{H}_{U|Y} \cup \mathcal{H}_{U|Z}) \backslash \mathcal{V}_U].$$

Note that the bits in Ψ_i^U are reusing the bits in Ψ_1^U; however, it is necessary to make the bits Φ_i^U available to both Bob and Eve, to enable the reconstruction of O_i. We show later that this entails a negligible rate cost. Finally, we write $O_i \triangleq [O_{i,1}, O_{i,2}]$, where $O_{i,1} \triangleq \widetilde{A}_i^{1:N}[\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}]$ and $O_{i,2} \triangleq \widetilde{A}_i^{1:N}[\mathcal{I}_{UY} \backslash \mathcal{I}_{UZ}]$, and we retransmit $O_{i,2}$ in the next block, We finally compute $\widetilde{U}_i^{1:N} \triangleq \widetilde{A}_i^{1:N} G_n$.

Finally, the encoder forms $\widetilde{A}_k^{1:N}$ in Block k, as follows. Let O_k be a vector of $|\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}|$ uniformly distributed bits representing the common message in that block. Given realizations o_k and o_{k-1}, the encoder samples $\widetilde{a}_k^{1:N}$ from the distribution $\widetilde{p}_{A_k^{1:N}}$ defined as follows.

$$\widetilde{p}_{A_k^j|A_k^{1:j-1}}(a_k^j|a_k^{1:j-1}) \triangleq \begin{cases} \mathbb{1}\left\{a_k^j = o_k^j\right\} & \text{if } j \in \mathcal{I}_{UY} \cap \mathcal{I}_{UZ} \\ \mathbb{1}\left\{a_k^j = o_{k-1,2}^j\right\} & \text{if } j \in \mathcal{A}_{UYZ} \\ \mathbb{1}\left\{a_k^j = (\psi_1^U)^j\right\} & \text{if } j \in \mathcal{V}_U \backslash (\mathcal{A}_{UYZ} \cup (\mathcal{I}_{UY} \cap \mathcal{I}_{UZ})) \\ p_{A^j|A^{1:j-1}}(a_k^j|a_k^{1:j-1}) & \text{if } j \in \mathcal{V}_U^c \end{cases}, \tag{117}$$

where the components of o_k, $o_{k-1,2}$, and ψ_1^U have been indexed by the set of indices $\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}$, $\mathcal{A}_{UYZ}$, and $\mathcal{V}_U \backslash (\mathcal{A}_{UYZ} \cup (\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}))$, respectively. Consequently,

$$O_k = \widetilde{A}_k^{1:N}[\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}],\ O_{k-1,2} = \widetilde{A}_k^{1:N}[\mathcal{A}_{UYZ}].$$

The random bits that identify the codebook and that are required to reconstruct $\widetilde{A}_k^{1:N}$ are $\widetilde{A}_k^{1:N}[\mathcal{H}_{U|Y}]$ for Bob and $\widetilde{A}_k^{1:N}[\mathcal{H}_{U|Z}]$ for Eve. Parts of these bits depend on messages in previous blocks. For the others, we define

$$\Psi_k^U \triangleq \widetilde{A}_k^{1:N}[\mathcal{V}_U \backslash (\mathcal{A}_{UYZ} \cup (\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}))],$$
$$\Phi_k^U \triangleq \widetilde{A}_k^{1:N}[(\mathcal{H}_{U|Y} \cup \mathcal{H}_{U|Z}) \backslash \mathcal{V}_U],$$

and note that Ψ_k^U merely reuses the bits of Ψ_1^U. Φ_k^U is made available to both Bob and Eve to help them reconstruct O_k, but this incurs a negligible rate cost.

The public transmission of $(\Psi_1^U, \Phi_{1:k}^U)$ to perform the reconstruction of the common message is taken into account in the secrecy analysis in Section 5.6.

5.5.2 Secret and private message encoding

In addition to the polarization set defined in (108)-(112), we also define

$$\mathcal{B}_{V|UY} \triangleq \text{a fixed subset of } \mathcal{V}_{V|UZ} \text{ with size } |\mathcal{V}_{V|UY} \cup ((\mathcal{H}_{V|UY} \backslash \mathcal{V}_{V|UY}) \cap \mathcal{V}_{V|U}))|$$
$$\mathcal{M}_{UVZ} \triangleq \mathcal{V}_{V|U} \backslash \mathcal{V}_{V|UZ}.$$

The encoding procedure with chaining is summarized in Figure 31.

In Block 1, the encoder forms $\widetilde{V}_1^{1:N}$ as follows. Let S_1 be a vector of $|\mathcal{V}_{V|UZ}|$ uniformly distributed bits representing the secret message and let M_1 be a vector of $|\mathcal{M}_{UVZ}|$ uniformly distributed bits representing the private message to be reconstructed by Bob. Given a confidential message s_1, a private message m_1, and $\widetilde{u}_1^{1:N}$ resulting from the encoding of the common message, the encoder samples $\widetilde{b}_1^{1:N}$ from the distribution $\widetilde{p}_{B_1^{1:N}}$ defined as follows.

$$\widetilde{p}_{B_1^j|B_1^{1:j-1}U_1^{1:N}}(b_1^j|b_1^{1:j-1}\widetilde{u}_1^{1:N}) \triangleq \begin{cases} \mathbb{1}\left\{b_1^j = s_1^j\right\} & \text{if } j \in \mathcal{V}_{V|UZ} \\ \mathbb{1}\left\{b_1^j = m_1^j\right\} & \text{if } j \in \mathcal{M}_{UVZ} \\ p_{B^j|B^{1:j-1}U^{1:N}}(b_1^j|b_1^{1:j-1}\widetilde{u}_1^{1:N}) & \text{if } j \in \mathcal{V}_{V|U}^c \end{cases}, \quad (118)$$

where the components of s_1 and m_1 have been indexed by the set of indices $\mathcal{V}_{V|UZ}$ and $\mathcal{M}_{UVZ}$, respectively. Consequently, note that $S_1 = \widetilde{B}_1^{1:N}[\mathcal{V}_{V|UZ}]$ and $M_1 =$

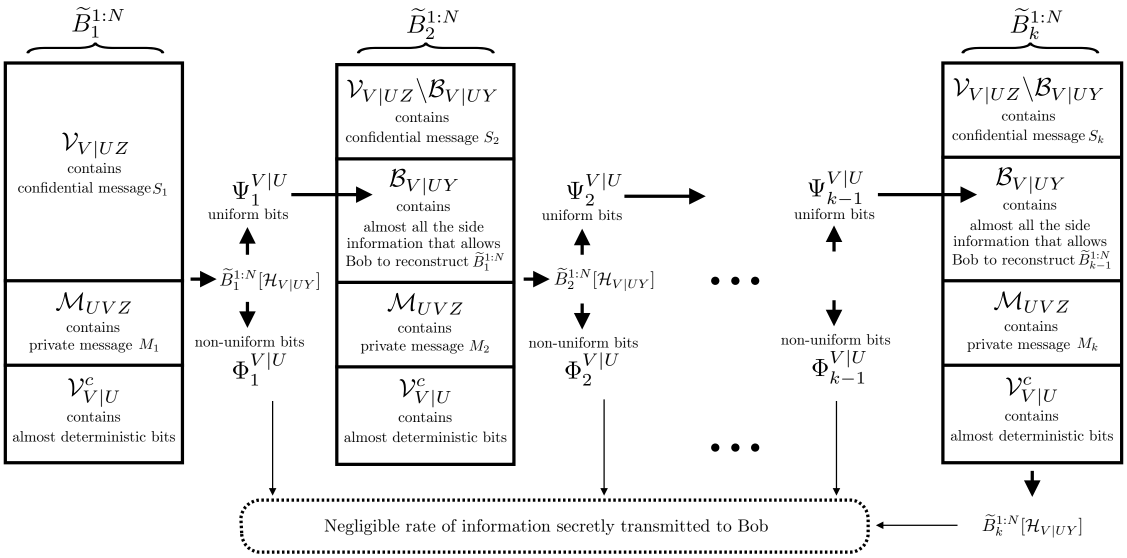


Figure 31. Chaining for the encoding of the $\widetilde{B}_i^{1:N}$'s, which corresponds to the encoding of the private and confidential messages.

$\widetilde{B}_1^{1:N}[\mathcal{M}_{UVZ}]$. The random bits that identify the codebook required for reconstruction are those in positions $\mathcal{H}_{V|UY}$, which we split as

$$\Psi_1^{V|U} \triangleq \widetilde{B}_1^{1:N}[\mathcal{V}_{V|UY} \cup ((\mathcal{H}_{V|UY}\backslash\mathcal{V}_{V|UY}) \cap \mathcal{V}_{V|U}))],$$
$$\Phi_1^{V|U} \triangleq \widetilde{B}_1^{1:N}[(\mathcal{H}_{V|UY}\backslash\mathcal{V}_{V|UY}) \cap \mathcal{V}_{V|U}^c].$$

Note that $\Psi_1^{V|U}$ is uniformly distributed but $\Phi_1^{V|U}$ is not. Consequently, we may reuse $\Psi_1^{V|U}$ in the next block but we cannot reuse $\Phi_1^{V|U}$. We instead share $\Phi_1^{V|U}$ secretly between Alice and Bob and we show later that this may be accomplished with negligible rate cost. Finally, define $\widetilde{V}_1^{1:N} \triangleq \widetilde{B}_1^{1:N} G_n$.

In Block $i \in [\![2, k]\!]$, the encoder forms $\widetilde{V}_i^{1:N}$ as follows. Let S_i be a vector of $|\mathcal{V}_{V|UZ}\backslash\mathcal{B}_{V|UY}|$ uniformly distributed bits and M_i be a vector of $|\mathcal{M}_{UVZ}|$ uniformly distributed bits that represent the secret and private message in block i, respectively. Given a private message m_i, a confidential message s_i, $\psi_{i-1}^{V|U}$, and $\widetilde{u}_i^{1:N}$ resulting from the encoding of the common message, the encoder draws $\widetilde{b}_i^{1:N}$ from the distribution

$\widetilde{p}_{B_i^{1:N}}$ defined as follows.

$$\widetilde{p}_{B_i^j|B_i^{1:j-1}U_i^{1:N}}(b_i^j|b_i^{1:j-1}\widetilde{u}_i^{1:N}) \triangleq \begin{cases} \mathbb{1}\left\{b_i^j = s_i^j\right\} & \text{if } j \in \mathcal{V}_{V|UZ}\backslash\mathcal{B}_{V|UY} \\ \mathbb{1}\left\{b_i^j = \left(\psi_{i-1}^{V|U}\right)^j\right\} & \text{if } j \in \mathcal{B}_{V|UY} \\ \mathbb{1}\left\{b_i^j = m_i^j\right\} & \text{if } j \in \mathcal{M}_{UVZ} \\ p_{B^j|B^{1:j-1}U^{1:N}}(b_1^j|b_1^{1:j-1}\widetilde{u}_i^{1:N}) & \text{if } j \in \mathcal{V}_{V|U}^c \end{cases}, \tag{119}$$

where the components of s_i, $\psi_{i-1}^{V|U}$, and m_i have been indexed by the set of indices $\mathcal{V}_{V|UZ}\backslash\mathcal{B}_{V|UY}$, $\mathcal{B}_{V|UY}$, and $\mathcal{M}_{UVZ}$ respectively, so that $S_i = \widetilde{B}_i^{1:N}[\mathcal{V}_{V|UZ}\backslash\mathcal{B}_{V|UY}]$, $\Psi_{i-1}^{V|U} = \widetilde{B}_i^{1:N}[\mathcal{B}_{V|UY}]$, and $M_i = \widetilde{B}_i^{1:N}[\mathcal{M}_{UVZ}]$. The random bits that identify the codebook required for reconstruction are those in positions $\mathcal{H}_{V|UY}$, which we split as

$$\begin{aligned} \Psi_i^{V|U} &\triangleq \widetilde{B}_i^{1:N}[\mathcal{V}_{V|UY} \cup ((\mathcal{H}_{V|UY}\backslash\mathcal{V}_{V|UY}) \cap \mathcal{V}_{V|U}))], \\ \Phi_i^{V|U} &\triangleq \widetilde{B}_i^{1:N}[(\mathcal{H}_{V|UY}\backslash\mathcal{V}_{V|UY}) \cap \mathcal{V}_{V|U}^c]. \end{aligned}$$

Again, $\Psi_i^{V|U}$ is uniformly distributed but $\Phi_i^{V|U}$ is not, so that we reuse $\Psi_i^{V|U}$ in the next block but we share $\Phi_i^{V|U}$ securely between Alice and Bob. We show later that the cost of sharing $\Phi_i^{V|U}$ is negligible. In Block k, Alice securely shares $(\Psi_k^{V|U}, \Phi_{1:k}^{V|U})$ with Bob. Finally, define $\widetilde{V}_i^{1:N} \triangleq \widetilde{B}_i^{1:N} G_n$.

5.5.3 Channel prefixing

The channel prefixing procedure with chaining is illustrated in Figure 32.

In Block 1, the encoder forms $\widetilde{X}_1^{1:N}$ as follows. Let R_1 be a vector of $|\mathcal{V}_{X|V}\backslash\mathcal{V}_{X|VZ}|$ uniformly distributed bits representing the randomness required for channel prefixing. Given a randomization sequence r_1 and $\widetilde{v}_1^{1:N}$ resulting from the encoding of secret and

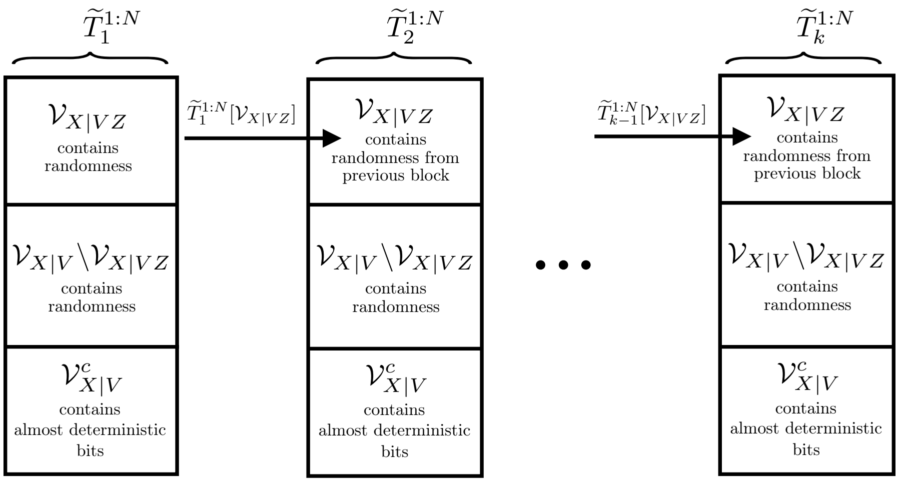

$\widetilde{T}_1^{1:N}$
$\widetilde{T}_2^{1:N}$
$\widetilde{T}_k^{1:N}$
$\mathcal{V}_{X|VZ}$
contains randomness
$\widetilde{T}_1^{1:N}[\mathcal{V}_{X|VZ}]$
$\mathcal{V}_{X|VZ}$
contains randomness from previous block
$\widetilde{T}_{k-1}^{1:N}[\mathcal{V}_{X|VZ}]$
$\mathcal{V}_{X|VZ}$
contains randomness from previous block
$\mathcal{V}_{X|V} \backslash \mathcal{V}_{X|VZ}$
contains randomness
$\mathcal{V}_{X|V} \backslash \mathcal{V}_{X|VZ}$
contains randomness
$\mathcal{V}_{X|V} \backslash \mathcal{V}_{X|VZ}$
contains randomness
$\mathcal{V}_{X|V}^c$
contains almost deterministic bits
$\mathcal{V}_{X|V}^c$
contains almost deterministic bits
$\mathcal{V}_{X|V}^c$
contains almost deterministic bits

Figure 32. Chaining for the encoding of the $\widetilde{T}_i^{1:N}$'s, which corresponds to channel prefixing.

private messages, the encoder draws $\widetilde{t}_1^{1:N}$ from the distribution $\widetilde{p}_{T_1^{1:N}}$ defined as follows.

$$\widetilde{p}_{T_1^j|T_1^{1:j-1}V_1^{1:N}}(t_1^j|t_1^{1:j-1}\widetilde{v}_1^{1:N}) \triangleq \begin{cases} 1/2 & \text{if } j \in \mathcal{V}_{X|VZ} \\ \mathbb{1}\left\{t_1^j = r_1^j\right\} & \text{if } j \in \mathcal{V}_{X|V} \backslash \mathcal{V}_{X|VZ} \\ p_{T^j|T^{1:j-1}V^{1:N}}(t_1^j|t_1^{1:j-1}\widetilde{v}_1^{1:N}) & \text{if } j \in \mathcal{V}_{X|V}^c \end{cases}, \tag{120}$$

where the components of r_1 have been indexed by the set of indices $\mathcal{V}_{X|V} \backslash \mathcal{V}_{X|VZ}$, so that $R_1 = \widetilde{T}_i^{1:N}[\mathcal{V}_{X|V} \backslash \mathcal{V}_{X|VZ}]$. The random bits that identify the codebook are those in position $\mathcal{V}_{X|VZ}$, which we denote

$$\Psi_1^{X|V} \triangleq \widetilde{T}_1^{1:N}[\mathcal{V}_{X|VZ}].$$

Finally, compute $\widetilde{X}_1^{1:N} \triangleq \widetilde{T}_1^{1:N} G_n$, which is transmitted over the channel $W_{YZ|X}$. We note $Y_1^{1:N}$, $Z_1^{1:N}$ the corresponding channel outputs.

In Block $i \in [\![2, k]\!]$, the encoder forms $\widetilde{X}_i^{1:N}$ as follows. Let R_i be a vector of $|\mathcal{V}_{X|V} \backslash \mathcal{V}_{X|VZ}|$ uniformly distributed bits representing the randomness required for channel prefixing in block i. Given a randomization sequence r_i and $\widetilde{v}_i^{1:N}$ resulting from the encoding of secret and private messages, the encoder draws $\widetilde{t}_i^{1:N}$ from the distribution $\widetilde{p}_{T_i^{1:N}}$ defined as follows.

$$\widetilde{p}_{T_i^j|T_i^{1:j-1}V_i^{1:N}}(t_i^j|t_i^{1:j-1}\widetilde{v}_i^{1:N}) \triangleq \begin{cases} \mathbb{1}\left\{t_i^j = \widetilde{t}_{i-1}^j\right\} & \text{if } j \in \mathcal{V}_{X|VZ} \\ \mathbb{1}\left\{t_i^j = r_i^j\right\} & \text{if } j \in \mathcal{V}_{X|V}\backslash\mathcal{V}_{X|VZ} \\ p_{T^j|T^{1:j-1}V^{1:N}}(t_i^j|t_i^{1:j-1}\widetilde{v}_i^{1:N}) & \text{if } j \in \mathcal{V}_{X|V}^c \end{cases}, \tag{121}$$

where the components of r_i have been indexed by the set of indices $\mathcal{V}_{X|V}\backslash\mathcal{V}_{X|VZ}$, so that $R_i = \widetilde{T}_i^{1:N}[\mathcal{V}_{X|V}\backslash\mathcal{V}_{X|VZ}]$. Note that the random bits describing the codebook are $\Psi_i^{X|V} \triangleq \widetilde{T}_i^{1:N}[\mathcal{V}_{X|VZ}]$, and are reused from the previous block. Finally, define $\widetilde{X}_i^{1:N} \triangleq \widetilde{T}_i^{1:N}G_n$ and transmit it over the channel $W_{YZ|X}$. We note $Y_i^{1:N}$, $Z_i^{1:N}$ the corresponding channel outputs.

5.5.4 Decoding

The decoding procedure is as follows.

Reconstruction of the common message by Bob. Bob forms the estimate $\widehat{A}_{1:k}^{1:N}$ of $\widetilde{A}_{1:k}^{1:N}$ as follows. In Block 1, Bob knows (Ψ_1^U, Φ_1^U), which contains all the bits $\widetilde{A}_1^{1:N}[\mathcal{H}_{U|Y}]$ by construction. Bob runs the successive cancellation decoder for source coding with side information of [87] using $Y_1^{1:N}$ and $\widetilde{A}_1^{1:N}[\mathcal{H}_{U|Y}]$. In Block $i \in [\![2,k]\!]$, Bob estimates $\widetilde{A}_i^{1:N}[\mathcal{H}_{U|Y}]$ with $(\Psi_1^U, \widehat{A}_{i-1}^{1:N}[\mathcal{I}_{UY}\backslash\mathcal{I}_{UZ}], \Phi_i^U)$, and uses this estimate along with $Y_i^{1:N}$ to run the successive cancellation decoder for source coding with side information.

Reconstruction of the common message by Eve. Eve forms the estimate $\widehat{\widehat{A}}_{1:k}^{1:N}$ of $\widetilde{A}_{1:k}^{1:N}$ starting from Block k and going backwards as follows. In Block k, Eve knows (Ψ_k^U, Φ_k^U), which contains all the bits in $\widetilde{A}_k^{1:N}[\mathcal{H}_{U|Z}]$ by construction. Eve runs the successive cancellation decoder for source coding with side information using $Z_k^{1:N}$ and $\widetilde{A}_k^{1:N}[\mathcal{H}_{U|Z}]$. For $i \in [\![1, k-1]\!]$, Eve estimates $\widetilde{A}_{k-i}^{1:N}[\mathcal{H}_{U|Z}]$ with $(\Psi_1^U, \widehat{\widehat{A}}_{k-i+1}^{1:N}[\mathcal{A}_{UYZ}], \Phi_{k-i}^U)$, and uses this estimate along with $Z_{k-i}^{1:N}$ to run the successive cancellation decoder for source coding with side information.

Reconstruction of the private and confidential messages by Bob. Bob forms the estimate $\widehat{B}_{1:k}^{1:N}$ of $\widetilde{B}_{1:k}^{1:N}$ as follows starting with Block k. In Block k, given $(\Psi_k^{V|U}, \Phi_k^{V|U}, Y_k^{1:N}, \widehat{U}_k^{1:N})$, Bob estimates $\widetilde{B}_k^{1:N}$ with the successive cancellation decoder for source coding with side information. From $\widetilde{B}_k^{1:N}$, an estimate $\widehat{\Psi}_{k-1}^{V|U} \triangleq \widehat{B}_k^{1:N}[\mathcal{V}_{V|UY}]$ of $\Psi_{k-1}^{V|U}$ is formed. For $i \in [\![1, k-1]\!]$, given $(\widehat{\Psi}_{k-i}^{V|U}, \Phi_{k-i}^{V|U}, Y_{k-i}^{1:N}, \widehat{U}_{k-i}^{1:N})$, Bob estimates $\widetilde{B}_{k-i}^{1:N}$ with the successive cancellation decoder for source coding with side information. From $\widetilde{B}_{k-i}^{1:N}$, an estimate of $\Psi_{k-i-1}^{V|U}$ is formed. Once all the estimates $\widehat{B}_{1:k}^{1:N}$ have been formed, Bob extracts the estimates $\widehat{S}_{1:k}$ and $\widehat{M}_{1:k}$ of $S_{1:k}$ and $M_{1:k}$, respectively.

5.6 Analysis of polar coding scheme

We now analyze in details the characteristics and performances of the polar coding scheme described in Section 5.5. Specifically, we show the following.

Theorem 5.6.2. *Consider a discrete memoryless broadcast channel* $(\mathcal{X}, p_{YZ|X}, \mathcal{Y}, \mathcal{Z})$. *The coding scheme of Section 5.4, whose complexity is* $O(N \log N)$ *achieves the region* $\mathcal{R}_{\text{BCC}}$.

The result of Theorem 5.6.2, follows in four steps. First, we show that the polar coding scheme of Section 5.5 approximates the statistics of the original DMS $(\mathcal{U} \times \mathcal{V} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}, p_{UVXYZ})$ from which the polarization sets were defined. Second, we show that the various messages rates are indeed those in $\mathcal{R}_{\text{BCC}}$. Third, we show that the probability of decoding error vanishes with the block length. Finally, we show that the information leakage vanishes with the block length.

5.6.1 Approximation of original DMS statistics

Recall that the vectors $\widetilde{A}_i^{1:N}$, $\widetilde{B}_i^{1:N}$, $\widetilde{V}_i^{1:N}$, and $\widetilde{X}_i^{1:N}$, generated in block $i \in [\![1, k]\!]$ do not have the exact joint distribution of the vectors $A^{1:N}$, $B^{1:N}$, $V^{1:N}$, and $X^{1:N}$, induced by the source polarization of the original DMS $(\mathcal{U} \times \mathcal{V} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}, p_{UVXYZ})$. However, the following lemmas show that the joint distributions are close to one another, which is crucial for the subsequent reliability and secrecy analysis.

Lemma 5.6.5. *For $i \in [\![1,k]\!]$, we have*

$$\mathbb{D}(p_{U^{1:N}}, \widetilde{p}_{U_i^{1:N}}) = \mathbb{D}(p_{A^{1:N}}, \widetilde{p}_{A_i^{1:N}}) \leqslant N\delta_N.$$

Hence, by Pinsker's inequality

$$\mathbb{V}(p_{A^{1:N}}, \widetilde{p}_{A_i^{1:N}}) \leqslant \delta_N^{(U)},$$

where $\delta_N^{(U)} \triangleq \sqrt{2\log 2}\sqrt{N\delta_N}$.

Proof. See Appendix 5.A. □

Lemma 5.6.6. *For $i \in [\![1,k]\!]$, we have*

$$\mathbb{D}(p_{V^{1:N}U^{1:N}}||\widetilde{p}_{V_i^{1:N}U_i^{1:N}}) = \mathbb{D}(p_{B^{1:N}U^{1:N}}||\widetilde{p}_{B_i^{1:N}U_i^{1:N}}) \leqslant 2N\delta_N.$$

Hence, by Pinsker's inequality

$$\mathbb{V}(p_{B^{1:N}U^{1:N}}, \widetilde{p}_{B_i^{1:N}U_i^{1:N}}) \leqslant \delta_N^{(UV)},$$

where $\delta_N^{(UV)} \triangleq 2\sqrt{\log 2}\sqrt{N\delta_N}$.

Proof. See Appendix 5.B. □

Lemma 5.6.7. *For $i \in [\![1,k]\!]$, we have*

$$\mathbb{D}(p_{X^{1:N}V^{1:N}}||\widetilde{p}_{X_i^{1:N}V_i^{1:N}}) = \mathbb{D}(p_{T^{1:N}V^{1:N}}||\widetilde{p}_{T_i^{1:N}V_i^{1:N}}) \leqslant 3N\delta_N.$$

Hence, by Pinsker's inequality

$$\mathbb{V}(p_{X^{1:N}V^{1:N}}, \widetilde{p}_{X_i^{1:N}V_i^{1:N}}) \leqslant \delta_N^{(XV)},$$

where $\delta_N^{(XV)} \triangleq \sqrt{2\log 2}\sqrt{3N\delta_N}$.

Proof. See Appendix 5.C. □

Combining the three previous lemmas, we obtain the following.

Lemma 5.6.8. *For $i \in [\![1,k]\!]$, we have*

$$\mathbb{V}(p_{U^{1:N}V^{1:N}X^{1:N}Y^{1:N}Z^{1:N}}, \widetilde{p}_{U_i^{1:N}V_i^{1:N}X_i^{1:N}Y_i^{1:N}Z_i^{1:N}}) \leqslant \delta_N^{(P)}.$$

where $\delta_N^{(P)} \triangleq \sqrt{2\log 2}\sqrt{N\delta_N}(2\sqrt{2}+\sqrt{3})$.

Proof. See Appendix 5.D. □

As noted in [116], upper-bounding the divergence with a chain rule is easier than directly upper-bounding the variational distance as in [94, 95].

5.6.2 Transmission rates

We now analyze the rate of common message, confidential message, private message, and randomization sequence, used at the encoder, as well as the different sum rates and the rate of additional information sent to Bob and Eve.

Common message rate. The overall rate R_O of common information bits transmitted satisfies

$$\begin{aligned} R_O &= \frac{(k-1)|\mathcal{I}_{UY}|+|\mathcal{I}_{UY}\cap\mathcal{I}_{UZ}|}{kN} \\ &= \frac{|\mathcal{I}_{UY}|}{N} - \frac{|\mathcal{I}_{UY}\backslash\mathcal{I}_{UZ}|}{kN} \\ &\geqslant \frac{|\mathcal{I}_{UY}|}{N} - \frac{|\mathcal{I}_{UY}|}{kN} \\ &\xrightarrow{N\to\infty} I(Y;U) - \frac{I(Y;U)}{k} \\ &\xrightarrow{k\to\infty} I(Y;U), \end{aligned}$$

where we have used [87]. Since we also have $R_O \leqslant \frac{|\mathcal{I}_{UY}|}{N} \xrightarrow{N\to\infty} I(Y;U)$, we conclude

$$R_O \xrightarrow{N\to\infty, k\to\infty} I(Y;U).$$

Confidential message rate. First, observe that

$$\begin{aligned} |\Psi_1^{V|U}| &= |\mathcal{V}_{V|UY} \cup ((\mathcal{H}_{V|UY}\backslash\mathcal{V}_{V|UY})\cap\mathcal{V}_{V|U}))| \\ &\leqslant |\mathcal{V}_{V|UY}|+|\mathcal{H}_{V|UY}\backslash\mathcal{V}_{V|UY}| \\ &= |\mathcal{V}_{V|UY}|+|\mathcal{H}_{V|UY}|-|\mathcal{V}_{V|UY}| \\ &\leqslant |\mathcal{H}_{V|UY}|, \end{aligned}$$

and $|\Psi_1^{V|U}|\geqslant|\mathcal{V}_{V|UY}|$. Hence, since $\lim_{N\to\infty}|\mathcal{V}_{V|UY}|/N = H(V|UY)$ by Lemma 3.4.1 and

$\lim_{N\to\infty}|\mathcal{H}_{V|UY}|/N = H(V|UY)$ by [87], we have

$$\lim_{N\to\infty}\frac{|\Psi_1^{V|U}|}{N} = H(V|UY).$$

Then, the overall rate R_S of secret information bits transmitted is

$$\begin{aligned}
R_S =& \frac{|\mathcal{V}_{V|UZ}|+(k-1)|\mathcal{V}_{V|UZ}\backslash\mathcal{B}_{V|UY}|}{kN}\\
=& \frac{|\mathcal{V}_{V|UZ}|+(k-1)(|\mathcal{V}_{V|UZ}|-|\mathcal{B}_{V|UY}|)}{kN}\\
=& \frac{|\mathcal{V}_{V|UZ}|-|\mathcal{B}_{V|UY}|}{N}+\frac{|\mathcal{B}_{V|UY}|}{kN}\\
=& \frac{|\mathcal{V}_{V|UZ}|-|\Psi_1^{V|U}|}{N}+\frac{|\Psi_1^{V|U}|}{kN}\\
\xrightarrow{N\to\infty}& I(V;Y|U)-I(V;Z|U)+\frac{H(V|UY)}{k}\\
\xrightarrow{k\to\infty}& I(V;Y|U)-I(V;Z|U).
\end{aligned}$$

Private message rate. The overall rate R_M of private information bits transmitted is

$$\begin{aligned}
R_M =& \frac{k|\mathcal{M}_{UVZ}|}{kN}\\
=& \frac{|\mathcal{V}_{V|U}\backslash\mathcal{V}_{V|UZ}|}{N}\\
=& \frac{|\mathcal{V}_{V|U}|-|\mathcal{V}_{V|UZ}|}{N}\\
\xrightarrow{N\to\infty}& I(V;Z|U),
\end{aligned}$$

where we have used Lemma 3.4.1.

Randomization rate. The uniform random bits used in the stochastic encoder includes those of the randomization sequence for channel prefixing, as well as those required to identify the codebooks and run the successive cancellation encoding. Using Lemma 3.4.1, we find that the rate required to identify the codebook for the common message is

$$\frac{|\mathcal{V}_U \backslash \mathcal{I}_{UY}|}{kN} \leqslant \frac{|\mathcal{V}_U|}{kN} \xrightarrow{N\to\infty} \frac{H(U|Y)}{k} \xrightarrow{k\to\infty} 0.$$

Similarly, the rate required to identify the codebook for the secret and private messages corresponds to the rate of $(\Psi_k^{V|U}, \Phi_k^{V|U})$, which is transmitted to Bob to allow him to reconstruct $\widetilde{B}_{1:k}^{1:N}$,

$$\begin{aligned} &\frac{|(\Psi_k^{V|U}, \Phi_k^{V|U})|}{kN} \\ &= \frac{|\widetilde{B}_k^{1:N}[\mathcal{H}_{V|UY}]|}{kN} \\ &\xrightarrow{N\to\infty} \frac{H(V|UY)}{k} \\ &\xrightarrow{k\to\infty} 0, \end{aligned}$$

where we have used [87].

The randomization sequence rate used in channel prefixing is

$$\begin{aligned} &\frac{|\mathcal{V}_{X|V}|+(k-1)|\mathcal{V}_{X|V}\backslash\mathcal{V}_{X|VZ}|}{kN} \\ &= \frac{|\mathcal{V}_{X|V}\backslash\mathcal{V}_{X|VZ}|}{N} + \frac{|\mathcal{V}_{X|VZ}|}{kN} \\ &= \frac{|\mathcal{V}_{X|V}|-|\mathcal{V}_{X|VZ}|}{N} + \frac{|\mathcal{V}_{X|VZ}|}{kN} \\ &\xrightarrow{N\to\infty} I(X;Z|V) + \frac{H(X|VZ)}{k}, \\ &\xrightarrow{k\to\infty} I(X;Z|V), \end{aligned}$$

where we have used Lemma 3.4.1. We now show that the rate of uniform bits required for successive cancellation encoding in (115), (116), (117), (118), (119), (120), (121) is negligible trough a series of lemmas.

Lemma 5.6.9. *For $i \in [\![1,k]\!]$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j\in\mathcal{V}_U^c} H(\widetilde{A}_i^j|\widetilde{A}_i^{1:j-1}) = 0.$$

Proof. See Appendix 5.E. □

Lemma 5.6.10. *For $i \in [\![1, k]\!]$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j \in \mathcal{V}^c_{V|U}} H(\widetilde{B}_i^j | \widetilde{B}_i^{1:j-1} \widetilde{U}_i^{1:N}) = 0.$$

Proof. See Appendix 5.F. □

Lemma 5.6.11. *For $i \in [\![1, k]\!]$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j \in \mathcal{V}^c_{X|V}} H(\widetilde{T}_i^j | \widetilde{T}_i^{1:j-1} \widetilde{V}_i^{1:N}) = 0.$$

The proof of Lemma 5.6.11 is similar to that of Lemma 5.6.10 using Lemma 5.6.7 in place of Lemma 5.6.6.

Hence, the overall randomness rate R_R used at the encoder is asymptotically

$$R_R \xrightarrow{N\to\infty, k\to\infty} I(X; Z|V).$$

Sum rates. The sum of the private message rate R_M and the randomness rate R_R is asymptotically

$$\begin{aligned} R_M + R_R \xrightarrow{N\to\infty, k\to\infty} & I(V; Z|U) + I(X; Z|V) \\ & \overset{(a)}{=} H(Z|U) - H(Z|UV) + H(Z|V) - H(Z|XV) \\ & = H(Z|U) - H(Z|XV) \\ & \overset{(b)}{=} H(Z|U) - H(Z|XU) \\ & = I(X; Z|U), \end{aligned}$$

where (a) and (b) hold by $U - V - X - Z$.

Moreover, the sum of the common message rate R_O, the private message rate R_M, and the confidential message rate R_S is asymptotically

$$\begin{aligned} R_O + R_M + R_S \xrightarrow{N\to\infty, k\to\infty} & I(Y; U) + I(V; Z|U) + I(V; Y|U) - I(V; Z|U) \\ & = I(Y; U) + I(V; Y|U). \end{aligned}$$

Seed Rate. The rate of the secret sequence that must be shared between the legitimate users to initialize the coding scheme is

$$\begin{aligned}
&\frac{|\Psi_k^{V|U}|+k|\Phi_1^{V|U}|}{kN}\\
&= \frac{|\Psi_k^{V|U}|}{kN} + \frac{|\Phi_1^{V|U}|}{N}\\
&\leqslant \frac{|\mathcal{H}_{V|UY}|}{kN} + \frac{|\mathcal{H}_{V|UY}\backslash\mathcal{V}_{V|UY}|}{N}\\
&\leqslant \frac{|\mathcal{H}_{V|UY}|}{kN} + \frac{|\mathcal{H}_{V|UY}|-|\mathcal{V}_{V|UY}|}{N}\\
&\xrightarrow{N\to\infty} \frac{H(V|Y)}{k}\\
&\xrightarrow{k\to\infty} 0,
\end{aligned}$$

where we have used Lemma 3.4.1 and [87].

Moreover the rate of public communication from Alice to both Bob and Eve is

$$\begin{aligned}
&\frac{|\Psi_1^{U}|+|\Phi_{1:k}^{U}|}{kN}\\
&\leqslant \frac{|\Psi_1^{U}|+k|\mathcal{H}_U\backslash\mathcal{V}_U|}{kN}\\
&= \frac{|\mathcal{V}_U\backslash\mathcal{I}_{UY}|+k(|\mathcal{H}_U|-|\mathcal{V}_U|)}{kN}\\
&\leqslant \frac{|\mathcal{H}_{U|Y}|+k(|\mathcal{H}_U|-|\mathcal{V}_U|)}{kN}\\
&= \frac{|\mathcal{H}_{U|Y}|}{kN} + \frac{|\mathcal{H}_U|-|\mathcal{V}_U|}{N}\\
&\xrightarrow{N\to\infty} \frac{H(U|Y)}{k}\\
&\xrightarrow{k\to\infty} 0.
\end{aligned}$$

5.6.3 Average probability of error

We first show that Eve and Bob can reconstruct the common messages $O_{1:k}^{1:N}$ with small probability. For $i \in [\![1,k]\!]$, consider an optimal coupling [94,96] between $\widetilde{p}_{U_i^{1:N}Y_i^{1:N}}$ and $p_{U^{1:N}Y^{1:N}}$ such that $\mathbb{P}[\mathcal{E}_{U_i,Y_i}] = \mathbb{V}(\widetilde{p}_{U_i^{1:N}Y_i^{1:N}}, p_{U^{1:N}Y^{1:N}})$, where $\mathcal{E}_{U_i,Y_i} \triangleq \{(\widetilde{U}_i^{1:N}, \widetilde{Y}_i^{1:N}) \neq (U^{1:N}, Y^{1:N})\}$. Define also for $i \in [\![2,k]\!]$, $\mathcal{E}_i \triangleq \{\widehat{A}_{i-1}^{1:N}[\mathcal{I}_{UY}\backslash\mathcal{I}_{UZ}] \neq \widetilde{A}_{i-1}^{1:N}[\mathcal{I}_{UY}\backslash\mathcal{I}_{UZ}]\}$.

We have

$$\begin{aligned}
&\mathbb{P}[O_i \neq \widehat{O}_i] \\
&= \mathbb{P}[\widehat{U}_i^{1:N} \neq \widetilde{U}_i^{1:N}] \\
&= \mathbb{P}[\widehat{U}_i^{1:N} \neq \widetilde{U}_i^{1:N} | \mathcal{E}_{U_i,Y_i}^c \cap \mathcal{E}_i^c]\mathbb{P}[\mathcal{E}_{U_i,Y_i}^c \cap \mathcal{E}_i^c] \\
&\qquad\quad + \mathbb{P}[\widehat{U}_i^{1:N} \neq \widetilde{U}_i^{1:N} | \mathcal{E}_{U_i,Y_i} \cup \mathcal{E}_i]\mathbb{P}[\mathcal{E}_{U_i,Y_i} \cup \mathcal{E}_i], \\
&\leqslant \mathbb{P}[\widehat{U}_i^{1:N} \neq \widetilde{U}_i^{1:N} | \mathcal{E}_{U_i,Y_i}^c \cap \mathcal{E}_i^c] + \mathbb{P}[\mathcal{E}_{U_i,Y_i} \cup \mathcal{E}_i] \\
&\overset{(a)}{\leqslant} N\delta_N + \mathbb{P}[\mathcal{E}_{U_i,Y_i}] + \mathbb{P}[\mathcal{E}_i] \\
&\overset{(b)}{\leqslant} N\delta_N + \delta_N^{(P)} + \mathbb{P}[\mathcal{E}_i] \\
&\leqslant N\delta_N + \delta_N^{(P)} + \mathbb{P}[\widehat{U}_{i-1}^{1:N} \neq \widetilde{U}_{i-1}^{1:N}] \\
&\overset{(c)}{\leqslant} (i-1)(N\delta_N + \delta_N^{(P)}) + \mathbb{P}[\widehat{U}_1^{1:N} \neq \widetilde{U}_1^{1:N}] \\
&\overset{(d)}{\leqslant} i(N\delta_N + \delta_N^{(P)}), \qquad (122)
\end{aligned}$$

where (a) follows from the error probability of source coding with side information [87] and the union bound, (b) holds by the optimal coupling and Lemma 5.6.8, (c) holds by induction, (d) holds similar to the previous inequalities. We thus have by the union bound and (122)

$$\begin{aligned}
\mathbb{P}[O_{1:k}^{1:N} \neq \widehat{O}_{1:k}^{1:N}] &\leqslant \sum_{i=1}^{k} \mathbb{P}[O_i \neq \widehat{O}_i] \\
&\leqslant \frac{k(k+1)}{2}(N\delta_N + \delta_N^{(P)}).
\end{aligned}$$

We similarly obtain for Eve

$$\mathbb{P}[O_{1:k}^{1:N} \neq \widehat{\widehat{O}}_{1:k}^{1:N}] \leqslant \frac{k(k+1)}{2}(N\delta_N + \delta_N^{(P)}).$$

Next we show how Bob can recover the secret and private messages. Informally, the decoding process of the confidential and private messages $(M_{1:k}, S_{1:k})$ for Bob is as follows. Reconstruction starts with Block k. Given $(\Psi_k^{V|U}, \Phi_k^{V|U}, Y_k^{1:N}, \widehat{U}_k^{1:N})$, Bob can reconstruct $\widetilde{V}_k^{1:N}$, from which $\Psi_{k-1}^{V|U}$ is deduced. Then, for $i \in [\![1, k-1]\!]$, given

$(\Psi_{k-i}^{V|U}, \Phi_{k-i}^{V|U}, Y_{k-i}^{1:N}, \widehat{U}_{k-i}^{1:N})$, Bob can reconstruct $\widetilde{V}_{k-i}^{1:N}$, from which $\Psi_{k-i-1}^{V|U}$ is deduced. Finally, $S_{1:k}$ can be recovered from $\widetilde{V}_{1:k}^{1:N}$.

Formally, the analysis is as follows. For $i \in [\![1, k]\!]$, consider an optimal coupling [96] between $\widetilde{p}_{U_i^{1:N} V_i^{1:N} Y_i^{1:N}}$ and $p_{U^{1:N} V^{1:N} Y^{1:N}}$ such that

$$\mathbb{P}[\mathcal{E}_{U_i,V_i,Y_i}] = \mathbb{V}(\widetilde{p}_{U_i^{1:N} V_i^{1:N} Y_i^{1:N}}, p_{U^{1:N} V^{1:N} Y^{1:N}}),$$

where $\mathcal{E}_{U_i,V_i,Y_i} \triangleq \{(\widetilde{U}_i^{1:N}, \widetilde{V}_i^{1:N}, Y_i^{1:N}) \neq (U^{1:N}, V^{1:N}, Y^{1:N})\}$. Define also for $i \in [\![1, k-1]\!]$, $\mathcal{E}_{\Psi_i^{V|U}} \triangleq \{\widehat{\Psi}_i^{V|U} \neq \Psi_i^{V|U}\}$, $\mathcal{E}_{\widetilde{U}_i} \triangleq \{\widehat{U}_i^{1:N} \neq \widetilde{U}_i^{1:N}\}$, and $\mathcal{E}_{\Psi_i^{V|U},\widetilde{U}_i} \triangleq \mathcal{E}_{\Psi_i^{V|U}} \cup \mathcal{E}_{\widetilde{U}_i}$.

For $i \in [\![1, k-1]\!]$, we have

$$\begin{aligned}
&\mathbb{P}[(M_i, S_i) \neq (\widehat{M_i}, \widehat{S}_i)] \\
&\overset{(a)}{=} \mathbb{P}[\widetilde{V}_i \neq \widehat{V}_i] \\
&= \mathbb{P}[\widetilde{V}_i \neq \widehat{V}_i | \mathcal{E}^c_{U_i,V_i,Y_i} \cap \mathcal{E}^c_{\Psi_i^{V|U},\widetilde{U}_i}]\mathbb{P}[\mathcal{E}^c_{U_i,V_i,Y_i} \cap \mathcal{E}^c_{\Psi_i^{V|U},\widetilde{U}_i}] \\
&\qquad\qquad + \mathbb{P}[\widetilde{V}_i \neq \widehat{V}_i | \mathcal{E}_{U_i,V_i,Y_i} \cup \mathcal{E}_{\Psi_i^{V|U},\widetilde{U}_i}]\mathbb{P}[\mathcal{E}_{U_i,V_i,Y_i} \cup \mathcal{E}_{\Psi_i^{V|U},\widetilde{U}_i}] \\
&\leqslant \mathbb{P}[\widetilde{V}_i \neq \widehat{V}_i | \mathcal{E}^c_{U_i,V_i,Y_i} \cap \mathcal{E}^c_{\Psi_i^{V|U},\widetilde{U}_i}] + \mathbb{P}[\mathcal{E}_{U_i,V_i,Y_i} \cup \mathcal{E}_{\Psi_i^{V|U},\widetilde{U}_i}] \\
&\leqslant \mathbb{P}[\widetilde{V}_i \neq \widehat{V}_i | \mathcal{E}^c_{U_i,V_i,Y_i} \cap \mathcal{E}^c_{\Psi_i^{V|U},\widetilde{U}_i}] + \mathbb{P}[\mathcal{E}_{U_i,V_i,Y_i}] + \mathbb{P}[\mathcal{E}_{\Psi_i^{V|U},\widetilde{U}_i}] \\
&\leqslant \mathbb{P}[\widetilde{V}_i \neq \widehat{V}_i | \mathcal{E}^c_{U_i,V_i,Y_i} \cap \mathcal{E}^c_{\Psi_i^{V|U},\widetilde{U}_i}] + \mathbb{P}[\mathcal{E}_{U_i,V_i,Y_i}] + \mathbb{P}[\mathcal{E}_{\Psi_i^{V|U}}] + \mathbb{P}[\mathcal{E}_{\widetilde{U}_i}] \\
&\overset{(b)}{\leqslant} \mathbb{P}[\widetilde{V}_i \neq \widehat{V}_i | \mathcal{E}^c_{U_i,V_i,Y_i} \cap \mathcal{E}^c_{\Psi_i^{V|U},\widetilde{U}_i}] + \mathbb{P}[\mathcal{E}_{U_i,V_i,Y_i}] + \mathbb{P}[\widetilde{V}_{i+1} \neq \widehat{V}_{i+1}] + \mathbb{P}[\widehat{U}_i^{1:N} \neq \widetilde{U}_i^{1:N}] \\
&\overset{(c)}{\leqslant} N\delta_N + \mathbb{P}[\mathcal{E}_{U_i,V_i,Y_i}] + \mathbb{P}[\widetilde{V}_{i+1} \neq \widehat{V}_{i+1}] + \mathbb{P}[\widehat{U}_i^{1:N} \neq \widetilde{U}_i^{1:N}] \\
&\overset{(d)}{\leqslant} N\delta_N + \delta_N^{(P)} + \mathbb{P}[\widetilde{V}_{i+1} \neq \widehat{V}_{i+1}] + \mathbb{P}[\widehat{U}_i^{1:N} \neq \widetilde{U}_i^{1:N}] \\
&\overset{(e)}{\leqslant} (i+1)\left(N\delta_N + \delta_N^{(P)}\right) + \mathbb{P}[\widetilde{V}_{i+1} \neq \widehat{V}_{i+1}] \\
&\overset{(f)}{\leqslant} (i+1)(k-i)\left(N\delta_N + \delta_N^{(P)}\right) + \mathbb{P}[\widetilde{V}_k \neq \widehat{V}_k] \\
&\overset{(g)}{\leqslant} (i+1)(k-i+1)\left(N\delta_N + \delta_N^{(P)}\right),
\end{aligned}$$

where (a) holds because $\widetilde{V}_i$ contains $(M_i, S_i, \Psi_{i-1}^{V|U})$ by construction, (b) holds because $\widetilde{V}_{i+1}$ contains $\Psi_i^{V|U}$ by construction, (c) follows from the error probability of lossless source coding with side information [87], (d) holds by the optimal coupling and

Lemma 5.6.8, (e) holds by (122), (f) holds by induction, (g) is obtained similar to the previous inequalities.

Hence,

$$\begin{aligned}&\mathbb{P}[(M_{1:k}, S_{1:k}) \neq (\widehat{M}_{1:k}, \widehat{S}_{1:k})]\\&\leqslant \sum_{i=1}^{k} \mathbb{P}[(M_i, S_i) \neq (\widehat{M}_i, \widehat{S}_i)]\\&\leqslant \sum_{i=1}^{k} (i+1)(k-i+1) \left(N\delta_N + \delta_N^{(P)}\right)\\&= \left(\frac{k(k+1)(k+2)}{6} + k\right) \left(N\delta_N + \delta_N^{(P)}\right). \qquad (123)\end{aligned}$$

5.6.4 Information leakage

The functional dependence graph for the coding scheme of Section 5.4 is given in Figure 33. For the secrecy analysis the following term must be upper bounded

$$I(S_{1:k}; \Psi_1^U \Phi_{1:k}^U Z_{1:k}^N).$$

Note that we have introduced $(\Psi_1^U, \Phi_{1:k}^U)$, since these random variables have been made available to Eve. Recall that $\Phi_{1:k}^U$ is additional information transmitted to Bob and Eve to reconstruct the common messages $O_{1:k}$. Recall also that $\Psi_1^U \supset \Psi_i^U$, $i \in [\![2, k]\!]$, as it is the randomness reused among all the blocks that allows the transmission of the common messages $O_{1:k}$. We start by proving that secrecy holds for a given block $i \in [\![2, k]\!]$ in the following lemma.

Lemma 5.6.12. *For $i \in [\![2, k]\!]$ and N large enough,*

$$I(S_i \Psi_{i-1}^{V|U}; Z_i^{1:N} \Phi_i^U \Psi_1^U) \leqslant \delta_N^{(*)},$$

where

$$\delta_N^{(*)} \triangleq \sqrt{2\log 2}\sqrt{N\delta_N}(1 + 6\sqrt{2} + 3\sqrt{3})(N - \log_2(\sqrt{2\log 2}\sqrt{N\delta_N}(1 + 6\sqrt{2} + 3\sqrt{3}))).$$

Proof. See Appendix 5.G. □

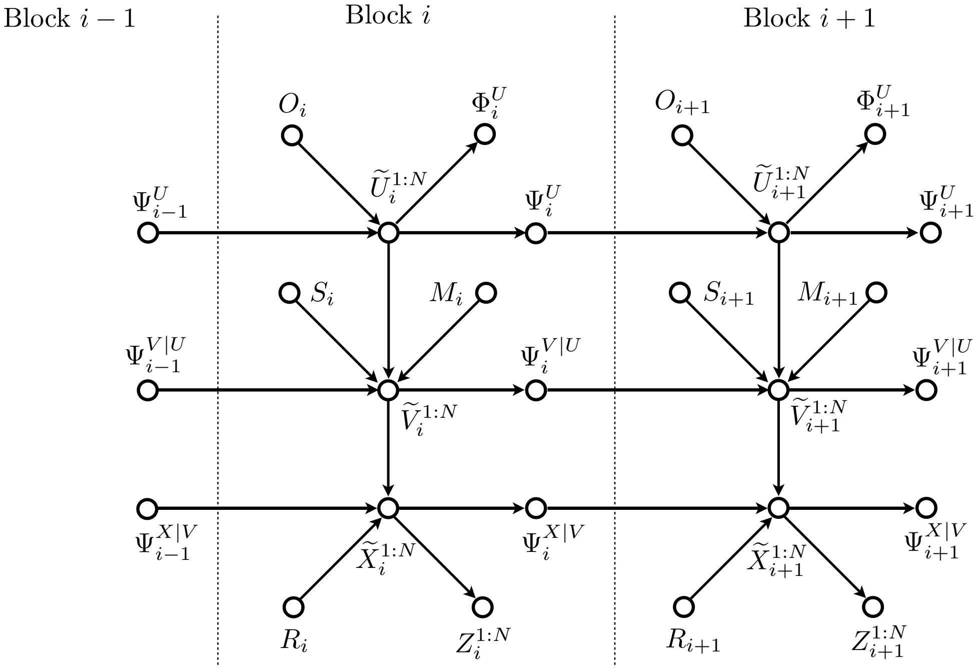


Figure 33. Functional dependence graph of the block encoding scheme. For Block i, O_i is the common message, M_i is the private message, S_i is the confidential message. $\Psi_i^{V|U}$ is the side information retransmitted in the next block to allow Bob to reconstruct M_i and S_i given $\Phi_i^{V|U}$ and its observations $Y_{1:k}^{1:N}$. Ψ_i^U is the randomness used to form $\widetilde{U}_i^{1:N}$, $\Psi_i^U \subset \Psi_1^U$ is reused from the previous block. R_i and $\Psi_i^{X|V}$ represent the randomness necessary at the encoder to form $\widetilde{X}_i^{1:N}$ where $\Psi_i^{X|V} = \Psi_1^{X|V}$ is reused from the previous block. Finally, Φ_i^U is information, whose rate is negligible, sent to Bob and Eve to allow them to reconstruct the common messages.

Recall that for channel prefixing in the encoding process we reuse some randomness $\Psi_1^{X|V}$ among all the blocks so that $\Psi_1^{X|V} = \Psi_i^{X|V}$, $i \in [\![2,k]\!]$. We show in the following lemma that $\Psi_1^{X|V}$ is almost independent from $(Z_i^{1:N}, \Psi_{i-1}^{V|U}, S_i, \Phi_i^U, \Psi_i^U)$. This fact will be useful in the secrecy analysis of the overall scheme.

Lemma 5.6.13. *For $i \in [\![2,k]\!]$ and N large enough,*

$$I(\Psi_1^{X|V}; Z_i^{1:N} \Psi_{i-1}^{V|U} S_i \Phi_i^U \Psi_i^U) \leqslant \delta_N^{(*)},$$

where $\delta_N^{(*)}$ is defined as in Lemma 5.6.12.

Proof. See Appendix 5.H. □

Using Lemmas 5.6.12 and 5.6.13, we show in the following lemma a recurrence relation that will make the secrecy analysis over all blocks easier.

Lemma 5.6.14. *Let $i \in [\![1,k-1]\!]$. Define $\widetilde{L}_i \triangleq I(S_{1:k}; \Psi_1^U \Phi_{1:i}^U Z_{1:i}^{1:N})$. We have*

$$\widetilde{L}_{i+1} - \widetilde{L}_i \leqslant 3\delta_N^{(*)}.$$

Proof. See Appendix 5.I. □

We then have

$$\begin{aligned}
\widetilde{L}_1 &= I(S_{1:k}; \Psi_1^U \Phi_1^U Z_1^{1:N}) \\
&= I(S_1; \Psi_1^U \Phi_1^U Z_1^{1:N}) + I(S_{2:k}; \Psi_1^U \Phi_1^U Z_1^{1:N} | S_1) \\
&\overset{(a)}{\leqslant} \delta_N^{(*)} + I(S_{2:k}; \Psi_1^U \Phi_1^U Z_1^{1:N} | S_1) \\
&\leqslant \delta_N^{(*)} + I(S_{2:k}; \Psi_1^U \Phi_1^U Z_1^{1:N} S_1) \\
&\overset{(b)}{=} \delta_N^{(*)},
\end{aligned}$$

where (a) follows from Lemma 5.6.12, (b) follows from independence of $S_{2:k}$ and the random variables of Block 1.

Hence, strong secrecy follows from Lemma 5.6.14 by remarking that

$$\begin{aligned}
I(S_{1:k}; \Psi_1^U \Phi_{1:k}^U Z_{1:k}^N) &= \widetilde{L}_1 + \sum_{i=1}^{k-1} (\widetilde{L}_{i+1} - \widetilde{L}_i) \\
&\leqslant \delta_N^{(*)} + (k-1)(3\delta_N^{(*)}) \\
&= (3k-2)\delta_N^{(*)}.
\end{aligned}$$

5.7 Conclusion

Our proposed polar coding scheme for the broadcast channel with confidential messages and constrained randomization provides an explicit low-complexity scheme achieving the capacity region of [55]. Although the presence of auxiliary random variables and the need to re-align polarization sets through chaining introduces rather involved notation, the coding scheme is conceptually close to a binning proof of the capacity region, in which polarization is used in place of random binning. We believe that a systematic use of this connection will effectively allow one to translate any results proved with output statistics of random binning [114] into a polar coding scheme.

APPENDICES

5.A Proof of Lemma 5.6.5

Let $i \in [\![2, k-1]\!]$. We have

$$\begin{aligned}
&\mathbb{D}(p_{U^{1:N}} || \widetilde{p}_{U_i^{1:N}}) \\
&\overset{(a)}{=} \mathbb{D}(p_{A^{1:N}} || \widetilde{p}_{A_i^{1:N}}) \\
&\overset{(b)}{=} \sum_{j=1}^{N} \mathbb{D}(p_{A^j|A^{1:j-1}} || \widetilde{p}_{A_i^j|A_i^{1:j-1}}) \\
&\overset{(c)}{=} \sum_{j \in \mathcal{V}_U} \mathbb{D}(p_{A^j|A^{1:j-1}} || \widetilde{p}_{A_i^j|A_i^{1:j-1}}) \\
&\overset{(d)}{=} \sum_{j \in \mathcal{V}_U} (1 - H(A^j|A^{1:j-1})) \\
&\overset{(e)}{\leqslant} |\mathcal{V}_U| \delta_N \\
&\leqslant N\delta_N, \qquad (124)
\end{aligned}$$

where (a) holds by invertibility of G_n, (b) holds by the chain rule, (c) holds by (116), (d) holds by (116) and uniformity of O_i and $O_{i-1,2}$, (e) holds by definition of $\mathcal{V}_U$.

Similarly for $i \in \{1, k\}$, using (115) and (117) we also have

$$\mathbb{D}(p_{U^{1:N}} || \widetilde{p}_{U_i^{1:N}}) \leqslant N\delta_N. \qquad (125)$$

5.B Proof of Lemma 5.6.6

Let $i \in [\![2,k]\!]$. We have

$$\begin{aligned}
&\mathbb{D}(p_{B^{1:N}|U^{1:N}}||\widetilde{p}_{B_i^{1:N}|U_i^{1:N}}) \\
&\overset{(a)}{=} \sum_{j=1}^{N} \mathbb{D}(p_{B^j|B^{1:j-1}U^{1:N}}||\widetilde{p}_{B_i^j|B_i^{1:j-1}U_i^{1:N}}) \\
&\overset{(b)}{=} \sum_{j\in\mathcal{V}_{V|U}} \mathbb{D}(p_{B^j|B^{1:j-1}U^{1:N}}||\widetilde{p}_{B_i^j|B_i^{1:j-1}U_i^{1:N}}) \\
&\overset{(c)}{=} \sum_{j\in\mathcal{V}_{V|U}} (1-H(B^j|B^{1:j-1}U^{1:N})) \\
&\overset{(d)}{\leqslant} |\mathcal{V}_{V|U}|\delta_N \\
&\leqslant N\delta_N, \qquad (126)
\end{aligned}$$

where (a) holds by the chain rule, (b) holds by (119), (c) holds by (119) and uniformity of $\Psi_{i-1}^{V|U}$, S_i, and M_i, (d) holds by definition of $\mathcal{V}_{V|U}$.

Then,

$$\begin{aligned}
&\mathbb{D}(p_{V^{1:N}U^{1:N}}||\widetilde{p}_{V_i^{1:N}U_i^{1:N}}) \\
&\overset{(a)}{=} \mathbb{D}(p_{B^{1:N}U^{1:N}}||\widetilde{p}_{B_i^{1:N}U_i^{1:N}}) \\
&\overset{(b)}{=} \mathbb{D}(p_{B^{1:N}|U^{1:N}}||\widetilde{p}_{B_i^{1:N}|U_i^{1:N}}) + \mathbb{D}(p_{U^{1:N}}||\widetilde{p}_{U_i^{1:N}}) \\
&\overset{(c)}{\leqslant} 2N\delta_N,
\end{aligned}$$

where (a) holds by invertibility of G_n, (b) holds by the chain rule, (c) holds by (126) and Lemma 5.6.5.

Similarly, using (118) and Lemma 5.6.5, we have

$$\mathbb{D}(p_{V^{1:N}U^{1:N}}||\widetilde{p}_{V_1^{1:N}U_1^{1:N}})\leqslant 2N\delta_N.$$

5.C Proof of Lemma 5.6.7

Let $i \in [\![2, k]\!]$. We have

$$\begin{aligned}
&\mathbb{D}(p_{T^{1:N}|V^{1:N}} || \widetilde{p}_{T_i^{1:N}|V_i^{1:N}}) \\
&\overset{(a)}{=} \sum_{j=1}^{N} \mathbb{D}(p_{T^j|T^{1:j-1}V^{1:N}} || \widetilde{p}_{T_i^j|T_i^{1:j-1}V_i^{1:N}}) \\
&\overset{(b)}{=} \sum_{j\in\mathcal{V}_{X|V}} \mathbb{D}(p_{T^j|T^{1:j-1}V^{1:N}} || \widetilde{p}_{T_i^j|T_i^{1:j-1}V_i^{1:N}}) \\
&\overset{(c)}{=} \sum_{j\in\mathcal{V}_{X|V}} (1 - H(T^j|T^{1:j-1}V^{1:N})) \\
&\overset{(d)}{\leqslant} |\mathcal{V}_{X|V}|\delta_N \\
&\leqslant N\delta_N,
\end{aligned} \tag{127}$$

where (a) holds by the chain rule, (b) holds by (121), (c) holds by (121) and uniformity of the bits in $\widetilde{T}_i^{1:N}[\mathcal{V}_{X|V}]$, (d) holds by definition of $\mathcal{V}_{X|V}$.

Then,

$$\begin{aligned}
&\mathbb{D}(p_{X^{1:N}V^{1:N}} || \widetilde{p}_{X_i^{1:N}V_i^{1:N}}) \\
&\overset{(a)}{=} \mathbb{D}(p_{T^{1:N}V^{1:N}} || \widetilde{p}_{T_i^{1:N}V_i^{1:N}}) \\
&\overset{(b)}{=} \mathbb{D}(p_{T^{1:N}|V^{1:N}} || \widetilde{p}_{T_i^{1:N}|V_i^{1:N}}) + \mathbb{D}(p_{V^{1:N}} || \widetilde{p}_{V_i^{1:N}}) \\
&\overset{(c)}{\leqslant} 3N\delta_N,
\end{aligned}$$

where (a) holds by invertibility of G_n, (b) holds by the chain rule, (c) holds by (127) and Lemma 5.6.6.

Similarly, using (120) and Lemma 5.6.6, we have

$$\mathbb{D}(p_{X^{1:N}V^{1:N}} || \widetilde{p}_{X_1^{1:N}V_1^{1:N}}) \leqslant 3N\delta_N.$$

5.D Proof of Lemma 5.6.8

We have

$$
\begin{aligned}
&\mathbb{V}(p_{U^{1:N}V^{1:N}X^{1:N}Y^{1:N}Z^{1:N}}, \widetilde{p}_{U_i^{1:N}V_i^{1:N}X_i^{1:N}Y_i^{1:N}Z_i^{1:N}}) \\
&= \mathbb{V}(p_{Y^{1:N}Z^{1:N}|U^{1:N}V^{1:N}X^{1:N}}p_{U^{1:N}V^{1:N}X^{1:N}}, \widetilde{p}_{Y_i^{1:N}Z_i^{1:N}|U_i^{1:N}V_i^{1:N}X_i^{1:N}}\widetilde{p}_{U_i^{1:N}V_i^{1:N}X_i^{1:N}}) \\
&\overset{(a)}{=} \mathbb{V}(p_{Y^{1:N}Z^{1:N}|X^{1:N}}p_{U^{1:N}V^{1:N}X^{1:N}}, \widetilde{p}_{Y_i^{1:N}Z_i^{1:N}|X_i^{1:N}}\widetilde{p}_{U_i^{1:N}V_i^{1:N}X_i^{1:N}}) \\
&\overset{(b)}{=} \mathbb{V}(p_{U^{1:N}V^{1:N}X^{1:N}}, \widetilde{p}_{U_i^{1:N}V_i^{1:N}X_i^{1:N}}) \\
&= \mathbb{V}(p_{X^{1:N}|U^{1:N}V^{1:N}}p_{U^{1:N}V^{1:N}}, \widetilde{p}_{X_i^{1:N}|U_i^{1:N}V_i^{1:N}}\widetilde{p}_{U_i^{1:N}V_i^{1:N}}) \\
&\overset{(c)}{=} \mathbb{V}(p_{X^{1:N}|V^{1:N}}p_{U^{1:N}V^{1:N}}, \widetilde{p}_{X_i^{1:N}|V_i^{1:N}}\widetilde{p}_{U_i^{1:N}V_i^{1:N}}) \\
&\overset{(d)}{\leqslant} \mathbb{V}(p_{X^{1:N}|V^{1:N}}p_{U^{1:N}V^{1:N}}, \widetilde{p}_{X_i^{1:N}|V_i^{1:N}}p_{U^{1:N}V^{1:N}}) \\
&\qquad\qquad\qquad + \mathbb{V}(\widetilde{p}_{X_i^{1:N}|V_i^{1:N}}p_{U^{1:N}V^{1:N}}, \widetilde{p}_{X_i^{1:N}|V_i^{1:N}}\widetilde{p}_{U_i^{1:N}V_i^{1:N}}) \\
&= \mathbb{V}(p_{X^{1:N}|V^{1:N}}p_{U^{1:N}V^{1:N}}, \widetilde{p}_{X_i^{1:N}|V_i^{1:N}}p_{U^{1:N}V^{1:N}}) + \mathbb{V}(p_{U^{1:N}V^{1:N}}, \widetilde{p}_{U_i^{1:N}V_i^{1:N}}) \\
&\overset{(e)}{\leqslant} \mathbb{V}(p_{X^{1:N}|V^{1:N}}p_{U^{1:N}V^{1:N}}, \widetilde{p}_{X_i^{1:N}|V_i^{1:N}}p_{U^{1:N}V^{1:N}}) + \delta_N^{(UV)} \\
&= \mathbb{V}(p_{X^{1:N}|V^{1:N}}p_{V^{1:N}}, \widetilde{p}_{X_i^{1:N}|V_i^{1:N}}p_{V^{1:N}}) + \delta_N^{(UV)} \\
&\overset{(f)}{\leqslant} \mathbb{V}(p_{X^{1:N}|V^{1:N}}p_{V^{1:N}}, \widetilde{p}_{X_i^{1:N}V_i^{1:N}}) + \mathbb{V}(\widetilde{p}_{X_i^{1:N}V_i^{1:N}}, \widetilde{p}_{X_i^{1:N}|V_i^{1:N}}p_{V^{1:N}}) + \delta_N^{(UV)} \\
&= \mathbb{V}(p_{X^{1:N}V^{1:N}}, \widetilde{p}_{X_i^{1:N}V_i^{1:N}}) + \mathbb{V}(\widetilde{p}_{V_i^{1:N}}, p_{V^{1:N}}) + \delta_N^{(UV)} \\
&\leqslant \mathbb{V}(p_{X^{1:N}V^{1:N}}, \widetilde{p}_{X_i^{1:N}V_i^{1:N}}) + \mathbb{V}(p_{U^{1:N}V^{1:N}}, \widetilde{p}_{U_i^{1:N}V_i^{1:N}}) + \delta_N^{(UV)} \\
&\overset{(g)}{\leqslant} 2\delta_N^{(UV)} + \delta_N^{(XV)},
\end{aligned}
$$

where (a) and (c) follow from the Markov condition $U \to V \to X \to (YZ)$ and $\widetilde{U}_i^{1:N} \to \widetilde{V}_i^{1:N} \to \widetilde{X}_i^{1:N} \to (Y_i^{1:N}Z_i^{1:N})$, (b) follows from $p_{Y^{1:N}Z^{1:N}|X^{1:N}} = \widetilde{p}_{Y_i^{1:N}Z_i^{1:N}|X_i^{1:N}}$ and [117, Lemma 17], (d) holds by the triangle inequality, (e) holds by Lemma 5.6.6, (f) hold by the triangle inequality, (g) holds by Lemmas 5.6.6 and 5.6.7.

5.E Proof of Lemma 5.6.9

We have for $i \in [\![1, k]\!]$, for $j \in \mathcal{V}_U^c$,

$$\begin{aligned}
&|H(\widetilde{A}_i^j|\widetilde{A}_i^{1:j-1}) - H(A^j|A^{1:j-1})| \\
&\leqslant |H(\widetilde{A}_i^{1:j}) - H(A^{1:j})| + |H(\widetilde{A}_i^{1:j-1}) - H(A^{1:j-1})| \\
&\overset{(a)}{\leqslant} \mathbb{V}(p_{A^{1:j}}, \widetilde{p}_{A_i^{1:j}}) \log \frac{2^j}{\mathbb{V}(p_{A^{1:j}}, \widetilde{p}_{A_i^{1:j}})} + |H(\widetilde{A}_i^{1:j-1}) - H(A^{1:j-1})| \\
&\overset{(b)}{\leqslant} \delta_N^{(U)} \left(N - \log_2 \delta_N^{(U)}\right) + |H(\widetilde{A}_i^{1:j-1}) - H(A^{1:j-1})| \\
&\leqslant 2\delta_N^{(U)} \left(N - \log_2 \delta_N^{(U)}\right) \\
&\triangleq \delta_N^{(A)},
\end{aligned}$$

where (a) holds by [102], (b) holds by Lemma 5.6.5 and because $x \mapsto x \log x$ is decreasing for $x > 0$ small enough.

Hence, we obtain

$$\begin{aligned}
&\sum_{j \in \mathcal{V}_U^c} H(\widetilde{A}_i^j|\widetilde{A}_i^{1:j-1}) \\
&= \sum_{j \in \mathcal{H}_U^c} \sum_{j \in \mathcal{H}_U \backslash \mathcal{V}_U} H(\widetilde{A}_i^j|\widetilde{A}_i^{1:j-1}) \\
&\leqslant |\mathcal{H}_U \backslash \mathcal{V}_U| + \sum_{j \in \mathcal{H}_U^c} H(\widetilde{A}_i^j|\widetilde{A}_i^{1:j-1}) \\
&= |\mathcal{H}_U| - |\mathcal{V}_U| + \sum_{j \in \mathcal{H}_U^c} H(\widetilde{A}_i^j|\widetilde{A}_i^{1:j-1}) \\
&\leqslant |\mathcal{H}_U| - |\mathcal{V}_U| + \sum_{j \in \mathcal{H}_U^c} (H(A^j|A^{1:j-1}) + \delta_N^{(A)}) \\
&\leqslant |\mathcal{H}_U| - |\mathcal{V}_U| + |\mathcal{H}_U^c|(\delta_N + \delta_N^{(A)}) \\
&\leqslant |\mathcal{H}_U| - |\mathcal{V}_U| + N(\delta_N + \delta_N^{(A)}),
\end{aligned}$$

and we obtain the result by Lemma 3.4.1 and [87].

5.F Proof of Lemma 5.6.10

We have for $i \in [\![1,k]\!]$, for $j \in \mathcal{V}_{V|U}^c$,

$$\begin{aligned}
&|H(\widetilde{B}_i^j|\widetilde{B}_i^{1:j-1}\widetilde{U}_i^{1:N}) - H(B^j|B^{1:j-1}U^{1:N})| \\
&\leqslant |H(\widetilde{B}_i^{1:j}\widetilde{U}_i^{1:N}) - H(B^{1:j}U^{1:N})| + |H(\widetilde{B}_i^{1:j-1}\widetilde{U}_i^{1:N}) - H(B^{1:j-1}U^{1:N})| \\
&\overset{(a)}{\leqslant} \mathbb{V}(p_{B^{1:j}U^{1:N}}, \widetilde{p}_{B_i^{1:j}U_i^{1:N}}) \log \frac{2^{j+N}}{\mathbb{V}(p_{B^{1:j}U^{1:N}}, \widetilde{p}_{B_i^{1:j}U_i^{1:N}})} \\
&\qquad\qquad\qquad\qquad\qquad + |H(\widetilde{B}_i^{1:j-1}\widetilde{U}_i^{1:N}) - H(B^{1:j-1}U^{1:N})| \\
&\overset{(b)}{\leqslant} \delta_N^{(UV)} \left(2N - \log_2 \delta_N^{(UV)}\right) + |H(\widetilde{B}_i^{1:j-1}\widetilde{U}_i^{1:N}) - H(B^{1:j-1}U^{1:N})| \\
&\leqslant 2\delta_N^{(UV)} \left(2N - \log_2 \delta_N^{(UV)}\right) \\
&\triangleq \delta_N^{(B)},
\end{aligned}$$

where (a) holds by [102], (b) holds by Lemma 5.6.6 and because $x \mapsto x \log x$ is decreasing for $x > 0$ small enough.

Then,

$$\begin{aligned}
&\sum_{j \in \mathcal{V}_{V|U}^c} H(\widetilde{B}_i^j|\widetilde{B}_i^{1:j-1}\widetilde{U}_i^{1:N}) \\
&= \sum_{j \in \mathcal{H}_{V|U}^c} \sum_{j \in \mathcal{H}_{V|U} \backslash \mathcal{V}_{V|U}} H(\widetilde{B}_i^j|\widetilde{B}_i^{1:j-1}\widetilde{U}_i^{1:N}) \\
&\leqslant |\mathcal{H}_{V|U} \backslash \mathcal{V}_{V|U}| + \sum_{j \in \mathcal{H}_{V|U}^c} H(\widetilde{B}_i^j|\widetilde{B}_i^{1:j-1}\widetilde{U}_i^{1:N}) \\
&= |\mathcal{H}_{V|U}| - |\mathcal{V}_{V|U}| + \sum_{j \in \mathcal{H}_{V|U}^c} H(\widetilde{B}_i^j|\widetilde{B}_i^{1:j-1}\widetilde{U}_i^{1:N}) \\
&\leqslant |\mathcal{H}_{V|U}| - |\mathcal{V}_{V|U}| + \sum_{j \in \mathcal{H}_{V|U}^c} (H(B^j|B^{1:j-1}U^{1:N}) + \delta_N^{(B)}) \\
&\leqslant |\mathcal{H}_{V|U}| - |\mathcal{V}_{V|U}| + |\mathcal{H}_{V|U}^c|(\delta_N + \delta_N^{(B)}) \\
&\leqslant |\mathcal{H}_{V|U}| - |\mathcal{V}_{V|U}| + N(\delta_N + \delta_N^{(B)}),
\end{aligned}$$

and we obtain the result by Lemma 3.4.1 and [87].

5.G Proof of Lemma 5.6.12

We have

$$\begin{aligned}
&\mathbb{V}(p_{B^{1:N}[\mathcal{V}_{V|UZ}]U^{1:N}Z^{1:N}}, \widetilde{p}_{B_i^{1:N}[\mathcal{V}_{V|UZ}]}\widetilde{p}_{U_i^{1:N}Z_i^{1:N}})\\
&\leqslant \mathbb{V}(p_{B^{1:N}[\mathcal{V}_{V|UZ}]U^{1:N}Z^{1:N}}, p_{B^{1:N}[\mathcal{V}_{V|UZ}]}p_{U^{1:N}Z^{1:N}})\\
&\qquad\qquad + \mathbb{V}(p_{B^{1:N}[\mathcal{V}_{V|UZ}]}p_{U^{1:N}Z^{1:N}}, \widetilde{p}_{B^{1:N}[\mathcal{V}_{V|UZ}]}\widetilde{p}_{U^{1:N}Z^{1:N}})\\
&\overset{(a)}{\leqslant} \mathbb{V}(p_{B^{1:N}[\mathcal{V}_{V|UZ}]U^{1:N}Z^{1:N}}, p_{B^{1:N}[\mathcal{V}_{V|UZ}]}p_{U^{1:N}Z^{1:N}})\\
&\qquad\qquad + \mathbb{V}(p_{B^{1:N}[\mathcal{V}_{V|UZ}]}, \widetilde{p}_{B^{1:N}[\mathcal{V}_{V|UZ}]}) + \mathbb{V}(p_{U^{1:N}Z^{1:N}}, \widetilde{p}_{U^{1:N}Z^{1:N}})\\
&\overset{(b)}{\leqslant} \mathbb{V}(p_{B^{1:N}[\mathcal{V}_{V|UZ}]U^{1:N}Z^{1:N}}, p_{B^{1:N}[\mathcal{V}_{V|UZ}]}p_{U^{1:N}Z^{1:N}}) + 2\delta_N^{(P)}\\
&\overset{(d)}{\leqslant} \sqrt{2\log 2}\sqrt{\mathbb{D}(p_{B^{1:N}[\mathcal{V}_{V|UZ}]U^{1:N}Z^{1:N}}||p_{B^{1:N}[\mathcal{V}_{V|UZ}]}p_{U^{1:N}Z^{1:N}})} + 2\delta_N^{(P)}\\
&= \sqrt{2\log 2}\sqrt{I(B^{1:N}[\mathcal{V}_{V|UZ}];U^{1:N}Z^{1:N})} + 2\delta_N^{(P)}\\
&\overset{(c)}{\leqslant} \sqrt{2\log 2}\sqrt{N\delta_N} + 2\delta_N^{(P)}, \qquad (128)
\end{aligned}$$

where (a) follows from the triangle inequality, (b) holds by Lemma 5.6.8, (c) holds by Pinsker's inequality, (d) holds because using the fact that conditioning reduces entropy we have

$$\begin{aligned}
&I(B^{1:N}[\mathcal{V}_{V|UZ}];U^{1:N}Z^{1:N})\\
&= H(B^{1:N}[\mathcal{V}_{V|UZ}]) - H(B^{1:N}[\mathcal{V}_{V|UZ}]|U^{1:N}Z^{1:N})\\
&\leqslant |\mathcal{V}_{V|UZ}| - \sum_{j\in\mathcal{V}_{V|UZ}} H(B^j|B^{1:j-1}U^{1:N}Z^{1:N})\\
&\leqslant |\mathcal{V}_{V|UZ}| + |\mathcal{V}_{V|UZ}|(\delta_N - 1)\\
&\leqslant N\delta_N.
\end{aligned}$$

We then obtain

$$
\begin{aligned}
&\mathbb{V}(\widetilde{p}_{B_i^{1:N}[\mathcal{V}_{V|UZ}]U_i^{1:N}Z_i^{1:N}}, \widetilde{p}_{B_i^{1:N}[\mathcal{V}_{V|UZ}]}\widetilde{p}_{U_i^{1:N}Z_i^{1:N}})\\
&\overset{(a)}{\leqslant} \mathbb{V}(\widetilde{p}_{B_i^{1:N}[\mathcal{V}_{V|UZ}]U_i^{1:N}Z_i^{1:N}}, p_{B^{1:N}[\mathcal{V}_{V|UZ}]U^{1:N}Z^{1:N}})\\
&\qquad\qquad\qquad + \mathbb{V}(p_{B^{1:N}[\mathcal{V}_{V|UZ}]U^{1:N}Z^{1:N}}, \widetilde{p}_{B_i^{1:N}[\mathcal{V}_{V|UZ}]}\widetilde{p}_{U_i^{1:N}Z_i^{1:N}})\\
&\overset{(b)}{\leqslant} \sqrt{2\log 2}\sqrt{N\delta_N} + 3\delta_N^{(P)}, \qquad (129)
\end{aligned}
$$

where (a) holds by the triangle inequality, (b) holds by Lemma 5.6.8, and (128).

Then, for N large enough by [102],

$$
\begin{aligned}
&I(S_i\Psi_{i-1}^{V|U}; Z_i^{1:N}\Phi_i^U\Psi_i^U)\\
&\leqslant I(\widetilde{B}_i^{1:N}[\mathcal{V}_{V|UZ}]; Z_i^{1:N}\widetilde{U}_i^{1:N})\\
&\leqslant \mathbb{V}(\widetilde{p}_{B_i^{1:N}[\mathcal{V}_{V|UZ}]U_i^{1:N}Z_i^{1:N}}, \widetilde{p}_{B_i^{1:N}[\mathcal{V}_{V|UZ}]}\widetilde{p}_{U_i^{1:N}Z_i^{1:N}})\\
&\qquad\quad \times \log_2 \frac{|\mathcal{V}_{V|UZ}|}{\mathbb{V}(\widetilde{p}_{B_i^{1:N}[\mathcal{V}_{V|UZ}]U_i^{1:N}Z_i^{1:N}}, \widetilde{p}_{B_i^{1:N}[\mathcal{V}_{V|UZ}]}\widetilde{p}_{U_i^{1:N}Z_i^{1:N}})}\\
&\leqslant \sqrt{2\log 2}\sqrt{N\delta_N}(1+6\sqrt{2}+3\sqrt{3})(N - \log_2(\sqrt{2\log 2}\sqrt{N\delta_N}(1+6\sqrt{2}+3\sqrt{3}))),
\end{aligned}
$$

where we have used (129) and that $x \mapsto x\log x$ is decreasing for $x > 0$ small enough.

5.H Proof of Lemma 5.6.13

By the triangle inequality we can write

$$
\begin{aligned}
&\mathbb{V}(p_{T^{1:N}[\mathcal{V}_{X|VZ}]U^{1:N}V^{1:N}Z^{1:N}}, \widetilde{p}_{T_i^{1:N}[\mathcal{V}_{X|VZ}]}\widetilde{p}_{U_i^{1:N}V_i^{1:N}Z_i^{1:N}})\\
&\leqslant \mathbb{V}(p_{T^{1:N}[\mathcal{V}_{X|VZ}]U^{1:N}V^{1:N}Z^{1:N}}, p_{T^{1:N}[\mathcal{V}_{X|VZ}]}p_{U^{1:N}V^{1:N}Z^{1:N}})\\
&\qquad\qquad\qquad + \mathbb{V}(p_{T^{1:N}[\mathcal{V}_{X|VZ}]}p_{U^{1:N}V^{1:N}Z^{1:N}}, \widetilde{p}_{T_i^{1:N}[\mathcal{V}_{X|VZ}]}\widetilde{p}_{U_i^{1:N}V_i^{1:N}Z_i^{1:N}})\\
&\overset{(a)}{\leqslant} \mathbb{V}(p_{T^{1:N}[\mathcal{V}_{X|VZ}]U^{1:N}V^{1:N}Z^{1:N}}, p_{T^{1:N}[\mathcal{V}_{X|VZ}]}p_{U^{1:N}V^{1:N}Z^{1:N}}) + 2\delta_N^{(P)}\\
&\overset{(b)}{\leqslant} \sqrt{2\log 2}\sqrt{\mathbb{D}(p_{T^{1:N}[\mathcal{V}_{X|VZ}]U^{1:N}V^{1:N}Z^{1:N}}, p_{T^{1:N}[\mathcal{V}_{X|VZ}]}p_{U^{1:N}V^{1:N}Z^{1:N}})} + 2\delta_N^{(P)}\\
&= \sqrt{2\log 2}\sqrt{I(T^{1:N}[\mathcal{V}_{X|VZ}]; Z^{1:N}U^{1:N}V^{1:N})} + 2\delta_N^{(P)}\\
&\overset{(c)}{\leqslant} \sqrt{2\log 2}\sqrt{N\delta_N} + 2\delta_N^{(P)}, \qquad (130)
\end{aligned}
$$

where (a) holds by the triangle inequality and Lemma 5.6.8, (b) holds by Pinsker's inequality, (c) holds because using the fact that conditioning reduces entropy and $U - V - X$ we have

$$\begin{aligned}
&I(T^{1:N}[\mathcal{V}_{X|VZ}]; Z^{1:N}U^{1:N}V^{1:N})\\
&\leqslant |\mathcal{V}_{X|VZ}| - \sum_{j \in \mathcal{V}_{X|VZ}} H(T^j | T^{1:j-1}Z^{1:N}U^{1:N}V^{1:N})\\
&= |\mathcal{V}_{X|VZ}| - \sum_{j \in \mathcal{V}_{X|VZ}} H(T^j | T^{1:j-1}Z^{1:N}V^{1:N})\\
&\leqslant |\mathcal{V}_{X|VZ}| + |\mathcal{V}_{X|VZ}|(\delta_N - 1)\\
&\leqslant N\delta_N.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\mathbb{V}(\widetilde{p}_{T_i^{1:N}[\mathcal{V}_{X|VZ}]U_i^{1:N}V_i^{1:N}Z_i^{1:N}}, \widetilde{p}_{T_i^{1:N}[\mathcal{V}_{X|VZ}]}\widetilde{p}_{U_i^{1:N}V_i^{1:N}Z_i^{1:N}})\\
&\leqslant \mathbb{V}(\widetilde{p}_{T_i^{1:N}[\mathcal{V}_{X|VZ}]U_i^{1:N}V_i^{1:N}Z_i^{1:N}}, p_{T^{1:N}[\mathcal{V}_{X|VZ}]U^{1:N}V^{1:N}Z^{1:N}})\\
&\qquad + \mathbb{V}(p_{T^{1:N}[\mathcal{V}_{X|VZ}]U^{1:N}V^{1:N}Z^{1:N}}, \widetilde{p}_{T_i^{1:N}[\mathcal{V}_{X|VZ}]}\widetilde{p}_{U_i^{1:N}V_i^{1:N}Z_i^{1:N}})\\
&\leqslant \sqrt{2\log 2}\sqrt{N\delta_N} + 3\delta_N^{(P)},
\end{aligned} \tag{131}$$

where (a) holds by the triangle inequality, (b) holds by Lemma 5.6.8, and (130).

Then, for N large enough by [102],

$$\begin{aligned}
&I(\Psi_i^{X|V}; Z_i^{1:N}\Psi_{i-1}^{V|U}S_i\Phi_i^U\Psi_i^U)\\
&= I(\widetilde{T}_i^{1:N}[\mathcal{V}_{X|VZ}]; Z_i^{1:N}\widetilde{B}_i^{1:N}[\mathcal{H}_{V|UZ}]\Phi_i^U\Psi_i^U)\\
&\leqslant I(\widetilde{T}_i^{1:N}[\mathcal{V}_{X|VZ}]; Z_i^{1:N}\widetilde{B}_i^{1:N}\widetilde{U}_i^{1:N})\\
&\overset{(a)}{=} I(\widetilde{T}_i^{1:N}[\mathcal{V}_{X|VZ}]; Z_i^{1:N}\widetilde{V}_i^{1:N}\widetilde{U}_i^{1:N})\\
&\leqslant \mathbb{V}(\widetilde{p}_{T_i^{1:N}[\mathcal{V}_{X|VZ}]U_i^{1:N}V_i^{1:N}Z_i^{1:N}}, \widetilde{p}_{T_i^{1:N}[\mathcal{V}_{X|VZ}]}\widetilde{p}_{U_i^{1:N}V_i^{1:N}Z_i^{1:N}})\\
&\qquad\qquad \times \log_2 \frac{|\mathcal{V}_{X|VZ}|}{\mathbb{V}(\widetilde{p}_{T_i^{1:N}[\mathcal{V}_{X|VZ}]U_i^{1:N}V_i^{1:N}Z_i^{1:N}}, \widetilde{p}_{T_i^{1:N}[\mathcal{V}_{X|VZ}]}\widetilde{p}_{U_i^{1:N}V_i^{1:N}Z_i^{1:N}})}\\
&\overset{(b)}{\leqslant} \sqrt{2\log 2}\sqrt{N\delta_N}(1 + 6\sqrt{2} + 3\sqrt{3})(N - \log_2(\sqrt{2\log 2}\sqrt{N\delta_N}(1 + 6\sqrt{2} + 3\sqrt{3}))),
\end{aligned}$$

where (a) holds by invertibility of G_n, (b) holds by (131) and because $x \mapsto x \log x$ is decreasing for $x > 0$ small enough.

5.I Proof of Lemma 5.6.14

Let $i \in [\![1, k-1]\!]$. We have

$$
\begin{aligned}
&\widetilde{L}_{i+1} - \widetilde{L}_i \\
&= I(S_{1:k}; \Psi_1^U \Phi_{1:i+1}^U Z_{1:i+1}^{1:N}) - I(S_{1:k}; \Psi_1^U \Phi_{1:i}^U Z_{1:i}^{1:N}) \\
&= I(S_{1:k}; \Phi_{i+1}^U Z_{i+1}^{1:N} | \Psi_1^U \Phi_{1:i}^U Z_{1:i}^{1:N}) \\
&= I(S_{1:i+1}; \Phi_{i+1}^U Z_{i+1}^{1:N} | \Psi_1^U \Phi_{1:i}^U Z_{1:i}^{1:N}) + I(S_{i+2:k}; \Phi_{i+1}^U Z_{i+1}^{1:N} | \Psi_1^U \Phi_{1:i}^U Z_{1:i}^{1:N} S_{1:i+1}) \\
&\overset{(a)}{\leqslant} I(S_{1:i+1} \Phi_{1:i}^U Z_{1:i}^{1:N}; \Phi_{i+1}^U Z_{i+1}^{1:N} | \Psi_1^U) + I(S_{i+2:k}; \Phi_{1:i+1}^U Z_{1:i+1}^{1:N} S_{1:i+1} \Psi_1^U) \\
&\overset{(b)}{=} I(S_{1:i+1} \Phi_{1:i}^U Z_{1:i}^{1:N}; \Phi_{i+1}^U Z_{i+1}^{1:N} | \Psi_1^U) \\
&= I(S_{i+1}; \Phi_{i+1}^U Z_{i+1}^{1:N} | \Psi_1^U) + I(S_{1:i} \Phi_{1:i}^U Z_{1:i}^{1:N}; \Phi_{i+1}^U Z_{i+1}^{1:N} | \Psi_1^U S_{i+1}) \\
&\leqslant I(S_{i+1}; \Phi_{i+1}^U Z_{i+1}^{1:N} \Psi_1^U) + I(S_{1:i} \Phi_{1:i}^U Z_{1:i}^{1:N}; \Phi_{i+1}^U Z_{i+1}^{1:N} | \Psi_1^U S_{i+1}) \\
&\overset{(c)}{\leqslant} \delta_N^{(*)} + I(S_{1:i} \Phi_{1:i}^U Z_{1:i}^{1:N}; \Phi_{i+1}^U Z_{i+1}^{1:N} | \Psi_1^U S_{i+1}) \\
&\leqslant \delta_N^{(*)} + I(S_{1:i} \Phi_{1:i}^U Z_{1:i}^{1:N}; \Phi_{i+1}^U Z_{i+1}^{1:N} S_{i+1} | \Psi_1^U) \\
&\overset{(d)}{\leqslant} \delta_N^{(*)} + I(S_{1:i} \Phi_{1:i}^U Z_{1:i}^{1:N} \Psi_i^{V|U} \Psi_i^{X|V}; \Phi_{i+1}^U Z_{i+1}^{1:N} S_{i+1} | \Psi_1^U) \\
&= \delta_N^{(*)} + I(\Psi_i^{V|U} \Psi_i^{X|V}; \Phi_{i+1}^U Z_{i+1}^{1:N} S_{i+1} | \Psi_1^U) \\
&\qquad\qquad\qquad + I(S_{1:i} \Phi_{1:i}^U Z_{1:i}^{1:N}; \Phi_{i+1}^U Z_{i+1}^{1:N} S_{i+1} | \Psi_i^{V|U} \Psi_i^{X|V} \Psi_1^U) \\
&\overset{(e)}{=} \delta_N^{(*)} + I(\Psi_i^{V|U} \Psi_i^{X|V}; \Phi_{i+1}^U Z_{i+1}^{1:N} S_{i+1} | \Psi_1^U) \\
&\leqslant \delta_N^{(*)} + I(\Psi_i^{V|U} \Psi_i^{X|V} \Psi_1^U; S_{i+1}) + I(\Psi_i^{V|U} \Psi_i^{X|V}; \Phi_{i+1}^U Z_{i+1}^{1:N} | \Psi_1^U S_{i+1}) \\
&\overset{(f)}{=} \delta_N^{(*)} + I(\Psi_i^{V|U} \Psi_i^{X|V}; \Phi_{i+1}^U Z_{i+1}^{1:N} | \Psi_1^U S_{i+1}) \\
&= \delta_N^{(*)} + I(\Psi_i^{V|U}; \Phi_{i+1}^U Z_{i+1}^{1:N} | \Psi_1^U S_{i+1}) + I(\Psi_i^{X|V}; \Phi_{i+1}^U Z_{i+1}^{1:N} | \Psi_i^{V|U} \Psi_1^U S_{i+1}) \\
&\leqslant \delta_N^{(*)} + I(\Psi_i^{V|U} S_{i+1}; \Phi_{i+1}^U Z_{i+1}^{1:N} \Psi_1^U) + I(\Psi_i^{X|V}; \Phi_{i+1}^U Z_{i+1}^{1:N} \Psi_i^{V|U} \Psi_1^U S_{i+1}) \\
&\overset{(g)}{\leqslant} 3\delta_N^{(*)},
\end{aligned}
$$

where (a) holds by the chain rule and positivity of mutual information, (b) holds by independence of $S_{i+2:k}$ with all the random variables of the previous blocks, (c) holds by Lemma 5.6.12, in (d) we introduce the random variable $\Psi_i^{V|U}$ and $\Psi_i^{X|V}$ to be able to break the dependencies between the random variables of block $(i+1)$ and the random variables of the previous blocks, (e) holds because $S_{1:i}\Phi_{1:i}^{U}Z_{1:i}^{1:N} \rightarrow \Psi_i^{V|U}\Psi_i^{X|V}\Psi_1^{U} \rightarrow \Phi_{i+1}^{U}Z_{i+1}^{1:N}S_{i+1}$, (f) holds because $(\Psi_i^{V|U}, \Psi_i^{X|V}, \Psi_i^{U})$ is independent of S_{i+1}, (g) holds by Lemmas 5.6.12, 5.6.13 and because $\Psi_i^{X|V}$ is constant equal to $\Psi_1^{X|V}$.

CHAPTER 6
CONCLUSION

Secure communication will remain a major concern with the amount of sensitive data such as medical records, financial transactions, or control information, transmitted over wireless networks. Information-theoretic security has the potential of enhancing security of future wireless networks by adding a level of security at the physical layer.

Privacy is also a growing concern with the increasing amount of private information collected in databases or the introduction of smart meters by utility providers, e.g., electricity, gas, water, to monitor individual consumptions. Information-theoretic security could also provide strong mathematical foundations and practical solutions to such problems.

However, the limit of many information-theoretic models for secure communication networks is an *over-simplification* of the problems studied. Moreover, many achievability theorems rely on random coding techniques that are impractical for *computationally bounded users.* Consequently, the need exists for models with as few simplifying assumptions as possible and for constructive schemes that could bridge the gap between information theory and coding theory. A partial answer to theses issues starts with the study of fundamental primitives for information-theoretic security, such as secret-key generation between two parties and communication over a channel tapped by an eavesdropper. Although these problems appeared in the literature several decades ago, only few practical coding schemes have been proposed until now.

6.1 Contributions

In Chapters 2, 3, we have addressed the problem of secret-key generation, for which we have accounted for *bandwidth constraint* and *computationally bounded legitimate users.*

In Chapter 2, we have studied sequential key-generation strategies, whose strength is to translate into practical designs – with the caveat of potentially high-complexity vector quantization required for the reconciliation step. Specifically, we have shown that the best known bound for rate-limited secret-key capacity are often achievable by a sequential strategy that separates reliability and secrecy thanks to a reconciliation step followed by a privacy amplification step with extractors. However, we have also qualified robustness and flexibility of sequential strategy to rate-limited communication, by showing that achieving the reconciliation capacity in a sequential strategy is, unlike the case of rate-unlimited communication, not necessarily optimal. We have further provided scenarios for which it stays optimal.

In Chapter 3, we have proposed low-complexity secret-key capacity-achieving schemes based on polar coding for several classes of sources. Unlike the sequential strategies proposed in Chapter 2, our polar coding schemes jointly handle secrecy and reliability. The price to be paid for low complexity is that our schemes often require a pre-shared seed, whose rate is negligible compared to the blocklength. Nevertheless, our polar coding schemes are the first provably optimal and low complexity scheme to handle rate-limited public communication and multi-terminal scenarios, which are often the major hurdle in designing optimal key-generation schemes.

In Chapters 4, 5, we have addressed the problem of secret communication over a wire-tapped channel, for which we have accounted for *computationally bounded legitimate users*, *bandwidth efficiency*, and *imperfect and rate-limited randomness* available at the encoder.

In Chapter 4, we have showed that multiplexing unprotected and protected data allows, first, to avoid the necessity of additional randomness at the encoder and, second, to efficiently use the bandwidth available between the legitimate users. Specifically, the overall communication rate of the same channel without secrecy constraints is maintained. The scheme leverage the results about the fundamental limits of lossless source coding with uniform encoder output. We have proposed an extension to multiple access channel in [118].

In Chapter 5, we have proposed a low-complexity and capacity-achieving scheme based on polar codes for the wiretap channel extended to a broadcast setting, in which a common message sent over the channel must be reconstructed by two users and a confidential message must be reconstructed by one user but must remain concealed from the other user. The main contribution, compared to previous works, is a scheme that deals with potentially asymmetric and non-degraded channels, and that also takes into account the cost of channel simulation. The resulting scheme is also optimal in terms of the amount of randomness used at the stochastic encoder.

Finally, the present work on secret-key generation and the wiretap channel model has generated tools whose interest and application go beyond the area of information-theoretic security.

Random binning with polar codes. In Chapter 5, we have drawn a parallel between random binning and polar coding scheme for the wiretap channel. It directly sheds light on the underlying fundamental mechanisms of the coding scheme, which could, at first glance, appear very ad hoc. Moreover, this parallel has the potential to allow a direct translation of any random binning achievability proof to a low-complexity polar coding scheme.

Data compression with uniform encoder output. It has been widely believed that the encoder output of compression codes were random number generators until T. S. Han formally proved it wrong. However, as shown in Chapter 4, this impossibility can be overcome when encoder and decoder share a seed, i.e., a small sequence of random numbers. Moreover, we have characterized the optimal length of the latter. We have also provided a practical coding scheme based on polar codes. Subsequent work in collaboration with the co-authors of [119] has shown that a seed is even unnecessary under lossy reconstruction.

Coding for channel resolvability. Channel resolvability characterizes the amount of randomness required to simulate a process at the output of a channel and plays a key role in the analysis of secure communication over wiretap channels. However, channel resolvability is a primitive that is also useful for the analysis of other problems such as the common information between random variables, and agents coordination in network. We have proposed low-complexity channel resolvability codes based on efficiently invertible extractors in [120], and based on polar codes in [121] using polar coding techniques developed in Chapter 2 and Chapter 5. We have shown that the latter construction yields optimal polar coding schemes for the problem of empirical coordination and strong coordination in two-node network [122].

6.2 Perspectives

As discussed in Chapter 1, many other constraints should be taken into account in information-theoretic models from a practical point of view. We list, below, some of them. Being able to successfully tackle all these constraints is highly challenging.

Finite-length regime. Asymptotic settings are of interest as a first approximation, as they provide insight for practical designs into the optimality of a given strategy. However, in practical applications, the finite-length regime should be considered to account for computational and data storage limitations. While several work have studied this problem, e.g. [14, 26, 27, 123], finding practical coding schemes nearly optimal in the finite-length regime remains elusive.

Unknown eavesdroppers statistics. The physical position of the eavesdropper, and thus the statistics of its observations, may not be known by the legitimate users. Compound models [124] represents an interesting way to model uncertainties about the eavesdropper statistics. MIMO settings could also be of interest as suggested by [52, 75], in which the requirement of the eavesdropper's statistics knowledge to ensure secrecy is replaced by a condition on the number of antennas available at the eavesdropper. Specifically, strictly positive secrecy rate are achievable provided that the legitimate users have more antennas than the eavesdropper.

Multi-user settings. The study and understanding of multi-users setting is primordial for large-scale application of information-theoretic security in tomorrow's communication networks. As alluded to earlier, the transition from point-to-point to network communication is a challenging task. Results on closed-form expressions for the secret-key capacity in multi-user settings, e.g. [4, 9, 31], or for the secrecy capacity of wiretap channels in multiple access or broadcast settings, e.g. [125–127], are only known for very specific scenarios and still remain incomplete in general.

Practical validations. Last but not least, practical implementations or proof of concept need to support theory. Although an increasing number of works have recently studied practical implementations of information-theoretic security [5, 7, 128–130], the gap between theory and practice remains far from being bridged.

REFERENCES

[1] C. Shannon, "Communication Theory of Secrecy Systems," *Bell System Technical Journal*, vol. 28, no. 4, pp. 656–715, 1949.

[2] U. Maurer, "Secret Key Agreement by Public Discussion from Common Information," *IEEE Trans. Inf. Theory*, vol. 39, pp. 733–742, 1993.

[3] R. Ahlswede and I. Csiszár, "Common Randomness in Information Theory and Cryptography Part I: Secret Sharing," *IEEE Trans. Inf. Theory*, vol. 39, pp. 1121–1132, 1993.

[4] I. Csiszár and P. Narayan, "Secrecy Capacities for Multiple Terminals.," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3047–3061, 2004.

[5] C. Ye, S. Mathur, A. Reznik, Y. Shah, W. Trappe, and N. B. Mandayam, "Information-Theoretically Secret Key Generation for Fading Wireless Channels," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 2, pp. 240–254, 2010.

[6] A. Pierrot, R. Chou, and M. Bloch, "Experimental Aspects of Secret Key Generation in Indoor Wireless Environments," in *IEEE Int. Workshop on Signal Processing Advances in Wireless Communications*, pp. 557–561, 2013.

[7] A. Pierrot, R. Chou, and M. Bloch, "The Effect of Eavesdropper's Statistics in Experimental Wireless Secret-Key Generation," *arXiv preprint arXiv:1312.3304*, 2013.

[8] M. Wegman and J. Carter, "New Hash Functions and their Use in Authentication and Set Equality," *Journal of Computer and System Sciences*, vol. 22, no. 3, pp. 265 – 279, 1981.

[9] I. Csiszár and P. Narayan, "Secrecy Capacities for Multiterminal Channel Models," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2437–2452, 2008.

[10] A. Gohari and V. Anantharam, "Information-Theoretic Key Agreement of Multiple TerminalsPart II: Channel Model," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3997–4010, 2010.

[11] C. Bennett and G. Brassard, "Quantum Cryptography: Public Key Distribution and Coin Tossing," in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, vol. 175, New York, 1984.

[12] A. Ekert, "Quantum Cryptography based on Bell's Theorem," *Physical review letters*, vol. 67, no. 6, pp. 661–663, 1991.

[13] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, "Quantum Cryptography," *Reviews of modern physics*, vol. 74, no. 1, pp. 145–195, 2002.

[14] M. Hayashi, "General Nonasymptotic and Asymptotic Formulas in Channel Resolvability and Identification Capacity and their Application to the Wiretap Channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1562–1575, 2006.

[15] G. Van Assche, *Quantum Cryptography and Secret-Key Distillation.* Cambridge University Press, 2006.

[16] I. Csiszár and P. Narayan, "Common Randomness and Secret Key Generation with a Helper," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 344–366, 2000.

[17] C. Ye and P. Narayan, "The Secret Key Private Key Capacity Region for Three Terminals," in *IEEE Int. Symp. Inf. Theory*, pp. 2142–2146, 2005.

[18] I. Csiszár and P. Narayan, "Capacity of a Shared Secret Key," in *IEEE Int. Symp. Inf. Theory*, pp. 2593–2596, 2010.

[19] A. Gohari and V. Anantharam, “Information-Theoretic Key Agreement of Multiple TerminalsPart I,” *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3973–3996, 2010.

[20] S. Watanabe and Y. Oohama, “Secret Key Agreement from Vector Gaussian Sources by Rate Limited Public Communication,” in *IEEE Int. Symp. Inf. Theory*, pp. 2597–2601, 2010.

[21] S. Nitinawarat and P. Narayan, “Secret Key Generation for Correlated Gaussian Sources,” *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3373–3391, 2012.

[22] I. Csiszár and P. Narayan, “Secrecy Generation for Multiaccess Channel Models,” *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 17–31, 2013.

[23] M. Bloch and J. Barros, *Physical-Layer Security: from Information Theory to Security Engineering.* Cambridge University Press, 2011.

[24] C. Bennett, G. Brassard, and U. Maurer, “Generalized Privacy Amplification,” *IEEE Trans. Inf. Theory*, vol. 41, pp. 1915–1923, 1995.

[25] U. Maurer and S. Wolf, “Information-Theoretic Key Agreement: From Weak to Strong Secrecy for Free,” in *Lecture Notes in Computer Science*, pp. 351–368, Springer-Verlag, 2000.

[26] R. Renner and S. Wolf, “Simple and Tight Bounds for Information Reconciliation and Privacy Amplification,” in *Advances in Cryptology-ASIACRYPT*, pp. 199–216, Springer, 2005.

[27] S. Watanabe and M. Hayashi, “Non-Asymptotic Analysis of Privacy Amplification via Rényi Entropy and Inf-Spectral Entropy,” in *IEEE Int. Symp. Inf. Theory*, pp. 2715–2719, 2013.

[28] Y. Dodis and D. Wichs, "Non-Malleable Extractors and Symmetric Key Cryptography from Weak Secrets," in *ACM Symposium on Theory of Computing*, pp. 601–610, 2009.

[29] G. Cohen, R. Raz, and G. Segev, "Non-Malleable Extractors with Short Seeds and Applications to Privacy Amplification," in *IEEE Conference on Computational Complexity*, pp. 298–308, 2012.

[30] Y. Dodis, J. Katz, L. Reyzin, and A. Smith, "Robust Fuzzy Extractors and Authenticated Key Agreement from Close Secrets," in *Advances in Cryptology*, pp. 232–250, Springer, 2006.

[31] C. Chan and L. Zheng, "Multiterminal Secret Key Agreement," *submitted to IEEE Trans. Inf. Theory*, 2010.

[32] S. Nitinawarat, C. Ye, A. Barg, P. Narayan, and A. Reznik, "Secret Key Generation for a Pairwise Independent Network Model," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6482–6489, 2010.

[33] S. Nitinawarat and P. Narayan, "Perfect Omniscience, Perfect Secrecy, and Steiner Tree Packing," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6490–6500, 2010.

[34] C. Ye and P. Narayan, "Secret Key and Private Key Constructions for Simple Multiterminal Source Models," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 639–651, 2012.

[35] A. Wyner, "The Wire-tap Channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[36] I. Csiszár and J. Korner, "Broadcast Channels with Confidential Messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, 1978.

[37] E. Arikan, "Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, 2009.

[38] H. Mahdavifar and A. Vardy, "Achieving the Secrecy Capacity of Wiretap Channels using Polar Codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6428–6443, 2011.

[39] E. Şaşoğlu and A. Vardy, "A New Polar Coding Scheme for Strong Security on Wiretap Channels," in *IEEE Int. Symp. Inf. Theory*, pp. 1117–1121, 2013.

[40] M. Hayashi, "Exponential Decreasing Rate of Leaked Information in Universal Random Privacy Amplification," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3989–4001, 2011.

[41] M. Bellare, S. Tessaro, and A. Vardy, "Semantic Security for the Wiretap channel," in *Advances in Cryptology*, pp. 294–311, Springer, 2012.

[42] A. Thangaraj, S. Dihidar, A. Calderbank, S. McLaughlin, and J.-M. Merolla, "Applications of LDPC Codes to the Wiretap Channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 8, pp. 2933–2945, 2007.

[43] A. Subramanian, A. Thangaraj, M. Bloch, and S. McLaughlin, "Strong Secrecy on the Binary Erasure Wiretap Channel Using Large-Girth LDPC codes," *IEEE Trans. Inf. Forensics and Security*, vol. 6, no. 3, pp. 585–594, 2011.

[44] V. Rathi, R. Urbanke, M. Andersson, and M. Skoglund, "Rate-Equivocation Optimal Spatially Coupled LDPC Codes for the BEC Wiretap Channel," in *IEEE Int. Symp. Inf. Theory*, pp. 2393–2397, 2011.

[45] S. Watanabe and Y. Oohama, "Broadcast Channels with Confidential Messages by Randomness Constrained Stochastic Encoder," in *Proc. of IEEE Int. Symp. Inf. Theory*, pp. 61–65, 2012.

[46] M. Bloch and J. Kliewer, "On Secure Communication with Constrained Randomization," in *IEEE Int. Symp. Inf. Theory*, pp. 1172–1176, IEEE, 2012.

[47] M. Hayashi and R. Matsumoto, "Secure Multiplex Coding with Dependent and Non-Uniform Multiple Messages," *arXiv preprint arXiv:1202.1332*, 2012.

[48] D. Kobayashi, H. Yamamoto, and T. Ogawa, "How to Attain the Ordinary Channel Capacity Securely in Wiretap Channels," in *Proc. of IEEE Inf. Theory Workshop*, pp. 13–18, 2005.

[49] J. Xu and B. Chen, "Broadcast Confidential and Public Messages," in *Proc. of 42nd Annual Conference on Information Sciences and Systems*, (Princeton, NJ), pp. 630–635, March 2008.

[50] M. Hayashi and R. Matsumoto, "Secure Multiplex Coding with Dependent and Non-Uniform Multiple Messages," in *Proc. of 50th Annual Allerton Conference on Communication, Control, and Computing*, pp. 954–959, 2012.

[51] I. Bjelakovic, H. Boche, and J. Sommerfeld, "Capacity Results for Arbitrarily Varying Wiretap Channels," in *Information Theory, Combinatorics, and Search Theory* (H. Aydinian, F. Cicalese, and C. Deppe, eds.), vol. 7777 of *Lecture Notes in Computer Science*, pp. 123–144, Springer Berlin Heidelberg, 2013.

[52] X. He and A. Yener, "MIMO Wiretap Channels with Arbitrarily Varying Eavesdropper Channel States," *arXiv preprint arXiv:1007.4801*, 2010.

[53] E. MolavianJazi, M. Bloch, and J. N. Laneman, "Arbitrary Jamming Can Preclude Secure Communications," in *Proc. 47th Annual Allerton Conference*

on Communication, Control, and Computing, (Monticello, IL), pp. 1069–1075, September 2009.

[54] M. Bloch and J. Laneman, “On the Secrecy Capacity of Arbitrary Wiretap channels,” in *Proceedings of 46th Allerton Conference on Communication, Control, and Computing*, pp. 818–825, 2008.

[55] S. Watanabe and Y. Oohama, “Broadcast Channels with Confidential Messages by Randomness Constrained Stochastic Encoder,” in *IEEE Int. Symp. Inf. Theory*, pp. 61–65, 2012.

[56] R. Chou and M. Bloch, “One-Way Rate-Limited Sequential Key Distillation,” in *IEEE Int. Symp. Inf. Theory*, 2012.

[57] R. Chou and M. Bloch, “Separation of Reliability and Secrecy in Rate-Limited Secret-Key Generation,” *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4941–4957, 2014.

[58] M. Bloch, A. Thangaraj, S. McLaughlin, and J.-M. Merolla, “LDPC-Based Gaussian Key Reconciliation,” in *IEEE Inf. Theory Workshop*, pp. 116–120, 2006.

[59] D. Elkouss, A. Leverrier, R. Alleaume, and J. Boutros, “Efficient Reconciliation Protocol for Discrete-Variable Quantum Key Distribution,” *IEEE Int. Symp. Inf. Theory*, pp. 1879–1883, 2009.

[60] S. Watanabe and Y. Oohama, “Secret Key Agreement from Correlated Gaussian Sources by Rate Limited Public Communication,” *IEICE Trans. Fundamentals*, vol. E93A, 2010.

[61] A. Orlitsky and J. Roche, “Coding for Computing,” *IEEE Trans. Inf. Theory*, vol. 47, no. 3, 2001.

[62] S. Vadhan, “Extracting All the Randomness from a Weakly Random Source,” tech. rep., Electronic Colloquium on Computational Complexity, 1998.

[63] M. Gander and U. Maurer, “On the Secret-Key Rate of Binary Random Variables,” in *IEEE Int. Symp. Inf. Theory*, p. 351, 1994.

[64] R. Gallager, *Information Theory and Reliable Communication.* John Wiley and Sons, New York, 1968.

[65] T. Ignatenko and F. Willems, “Biometric Systems: Privacy and Secrecy Aspects,” *IEEE Trans. Inf. Forens. Security*, vol. 4, no. 4, pp. 956–973, 2009.

[66] Y. Yang, S. Cheng, Z. Xiong, and W. Zhao, “Wyner-Ziv Coding Based on TCQ and LDPC Codes,” *IEEE Trans. Commun.*, vol. 57, no. 2, 2009.

[67] D. Elkouss, J. Martinez, D. Lancho, and V. Martin, “Rate Compatible Protocol for Information Reconciliation: An Application to QKD,” *IEEE Inf. Theory Workshop*, pp. 145–149, 2010.

[68] K. Kasai, R. Matsumoto, and K. Sakaniwa, “Information Reconciliation for QKD with Rate-Compatible Non-Binary LDPC Codes,” in *ISITA'10*, pp. 922–927, 2010.

[69] L. Carter and M. Wegman, “Universal Classes of Hash Functions,” *Journal of Computer and System Sciences*, vol. 18, no. 2, pp. 143–154, 1979.

[70] T. Berger, *Multiterminal Source Coding.* The Information Theory Approach to Communications, G.Longo, Ed. New York: Springer-Verlag, 1978.

[71] R. M. Fano, *Transmission of Information: A Statistical Theory of Communications.* M.I.T. Press, 1961.

[72] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes.* Holden-Day, 1964.

[73] T. Cover and J. Thomas, *Elements of Information Theory.* Wiley, 1991.

[74] J. Barros and M. Bloch, "Strong Secrecy for Wireless Channels," in *Information Theoretic Security*, pp. 40–53, Springer, 2008.

[75] R. Chou and M. Bloch, "Secret-Key Generation with Arbitrarily Varying Eavesdroppers Channel," in *IEEE GlobalSIP*, pp. 277–280, 2013.

[76] R. Ahlswede and I. Csiszár, "Common Randomness in Information Theory and Cryptography Part II: CR Capacity," *IEEE Trans. Inf. Theory*, vol. 44, pp. 225–240, 1998.

[77] A. Wyner and J. Ziv, "The Rate Distortion Function for Source Coding with Side Information at the Decoder," *IEEE Trans. Inf. Theory*, vol. 22(1), pp. 1–10, 1973.

[78] G. Kramer, "Topics in Multi-User Information Theory," *Foundations and Trends in Communications and Information Theory*, vol. 4, pp. 265–444, 2007.

[79] R. Rockafellar, *Convex Analysis.* Princeton University Press, Princeton, NJ, 2011.

[80] M. Salehi, "Cardinality Bounds on Auxiliary Variables in Multiple-User Theory via the Method of Ahlswede and Körner," tech. rep., Electronic Colloquium on Computational Complexity, 1998.

[81] A. E. Gamal and Y.-H. Kim, *Network Information Theory.* Cambridge University Press, 2011.

[82] R. Chou, M. Bloch, and E. Abbe, "Polar Coding for Secret-Key Generation," in *IEEE Inf. Theory Workshop*, 2013.

[83] R. Chou, M. Bloch, and E. Abbe, "Polar Coding for Secret-Key Generation," *Accepted to IEEE Trans. Inf. Theory. Available at arXiv:1305.4746v3*, 2015.

[84] E. Hof and S. Shamai, "Secrecy-Achieving Polar-Coding," in *IEEE Inf. Theory Workshop*, pp. 1–5, 2010.

[85] O. Koyluoglu and H. El Gamal, "Polar Coding for Secure Transmission and Key Agreement," in *IEEE Int. Symp. on Personal Indoor and Mobile Radio Communications*, pp. 2698–2703, 2010.

[86] M. Andersson, V. Rathi, R. Thobaben, J. Kliewer, and M. Skoglund, "Nested Polar Codes for Wiretap and Relay Channels," *IEEE Communications Letters*, vol. 14, no. 8, pp. 752–754, 2010.

[87] E. Arikan, "Source Polarization," in *IEEE Int. Symp. Inf. Theory*, pp. 899–903, 2010.

[88] S. Korada and R. Urbanke, "Polar Codes for Slepian-Wolf, Wyner-Ziv, and Gelfand-Pinsker," in *IEEE Inf. Theory Workshop*, pp. 1–5, 2010.

[89] E. Abbe, "Randomness and Dependencies Extraction via Polarization," in *Information Theory and Applications Workshop*, pp. 1–7, 2011.

[90] E. Şaşoğlu, "Polar Coding Theorems for Discrete Systems," *EPFL Thesis*, no. 5219, 2011.

[91] E. Abbe, "Randomness and Dependencies Extraction via Polarization, with Applications to Slepian-wolf Coding and Secrecy," *to appear in IEEE Trans. Inf. Theory*, 2015.

[92] D. Sutter, J. Renes, and R. Renner, "Efficient One-Way Secret-Key Agreement and Private Channel Coding via Polarization," *arXiv preprint arXiv:1304.3658*, 2013.

[93] M. Hayashi, “Second-Order Asymptotics in Fixed-Length Source Coding and Intrinsic Randomness,” *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4619–4637, 2008.

[94] S. Korada and R. Urbanke, “Polar Codes are Optimal for Lossy Source Coding,” *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1751–1768, 2010.

[95] J. Honda and H. Yamamoto, “Polar Coding Without Alphabet Extension for Asymmetric Models,” *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7829–7838, 2013.

[96] D. Aldous, “Random Walks on Finite Groups and Rapidly Mixing Markov Chains,” in *Séminaire de Probabilités XVII 1981/82*, pp. 243–297, Springer, 1983.

[97] M. Ye and A. Barg, “Universal Source Polarization and an Application to a Multi-User Problem,” in *Proc. of the Annual Allerton Conf. on Communication Control and Computing*, 2014.

[98] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, “Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data,” *SIAM Journal on Computing*, vol. 38, no. 1, pp. 97–139, 2008.

[99] L. Lai, S.-W. Ho, and H. V. Poor, “Privacy-Security Tradeoffs in Biometric Security Systems,” in *Annual Allerton Conf. on Communication Control and Computing*, pp. 268–273, 2008.

[100] S. Rane, Y. Wang, S. Draper, and P. Ishwar, “Secure Biometrics: Concepts, Authentication Architectures, and Challenges,” *Signal Processing Magazine, IEEE*, vol. 30, no. 5, pp. 51–64, 2013.

[101] E. Arikan and I. E. Telatar, “On the Rate of Channel Polarization,” in *Proc. of IEEE Int. Symp. Inf. Theory*, pp. 1493–1495, 2009.

[102] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* Cambridge Univ Pr, 1981.

[103] T. S. Han, “Folklore in Source Coding: Information-Spectrum Approach,” *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 747–753, 2005.

[104] P. Gemmell and M. Naor, “Codes for Interactive Authentication,” in *Advances in CryptologyCRYPTO93*, pp. 355–367, Springer, 1994.

[105] E. Gilbert, F. MacWilliams, and N. Sloane, “Codes which detect deception,” *Bell Syst. Tech. J.*, vol. 53, no. 3, pp. 405–424, 1974.

[106] Y. Dodis, “On Extractors, Error-Correction and Hiding All Partial Information,” in *IEEE Inf. Theory Workshop*, 2005.

[107] Y. Dodis and A. Smith, “Entropic Security and the Encryption of High-Entropy Messages,” in *Theory of Cryptography*, 2005.

[108] R. Chou, M. Bloch, B. Vellambi, and J. Kliewer, “Source-Channel Coding Schemes for Achieving Strong Secrecy at Negligible Cost,” *to be submitted to IEEE Trans. Inf. Theory.*

[109] T. S. Han, *Information-Spectrum Methods in Information Theory*, vol. 50. Springer, 2002.

[110] R. Chou and M. R. Bloch, “Polar Coding for the Broadcast Channel with Confidential Messages and Constrained Randomization,” *arXiv preprint arXiv:1411.0281*, 2014.

[111] R. Chou and M. Bloch, “Polar Coding for the Broadcast Channel with Confidential Messages,” in *IEEE Inf. Theory Workshop*, 2015.

[112] I. Csiszár, “Almost Independence and Secrecy Capacity,” *Problems of Information Transmission*, vol. 32, pp. 40–47, January-March 1996.

[113] J. Renes and R. Renner, “Noisy Channel Coding via Privacy Amplification and information reconciliation,” *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7377–7385, 2011.

[114] M. H. Yassaee, M. R. Aref, and A. Gohari, “Achievability Proof via Output Statistics of Random Binning,” in *Proc. of IEEE Int. Symp. Inf. Theory*, (Boston, MA), pp. 1044–1048, July 2012.

[115] E. Şaşoğlu, “Polar Codes for Discrete Alphabets,” in *IEEE Int. Symp. Inf. Theory*, pp. 2137–2141, 2012.

[116] N. Goela, E. Abbe, and M. Gastpar, “Polar Codes for Broadcast Channels,” *arXiv preprint arXiv:1301.6150*, 2013.

[117] P. Cuff, *Communication in Networks for Coordinating Behavior.* PhD thesis, Stanford Univ., CA., 2009.

[118] R. Chou and M. Bloch, “Uniform Distributed Source Coding for the Multiple Access Wiretap Channel,” in *IEEE Conf. on Communications and Network Security (CNS): Wokshop on Physical-layer Methods for Wireless Security*, 2014.

[119] B. Vellambi, M. Bloch, R. Chou, , and J. Kliewer, “Lossless and Lossy Source Compression with Near-Uniform Outputs: Is Common Randomness Always Required?,” *submitted to IEEE Int. Symp. Inf. Theory*, 2015.

[120] R. Chou, M. Bloch, and J. Kliewer, “Low-Complexity Channel Resolvability Codes for the Symmetric Multiple-Access Channel,” *IEEE Inf. Theory Workshop*, 2014.

[121] R. Chou, M. Bloch, and J. Kliewer, “Polar Coding for Empirical and Strong Coordination through Distribution Approximation,” *to be submitted to IEEE Trans. Inf. Theory.*

[122] R. Chou, M. Bloch, and J. Kliewer, “Polar Coding for Empirical and Strong Coordination through Distribution Approximation,” *IEEE Int. Symp. Inf. Theory*, 2015.

[123] H. Tyagi and S. Watanabe, “Converses for Secret Key Agreement and Secure Computing,” *arXiv preprint arXiv:1404.5715*, 2014.

[124] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai, “Compound wiretap channels,” *EURASIP J. Wirel. Commun. Netw.*, 2009.

[125] E. Tekin and A. Yener, “The General Gaussian Multiple-Access and Two-Way Wiretap Channels: Achievable Rates and Cooperative Jamming,” *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2735–2751, 2008.

[126] Y. Liang and H. V. Poor, “Multiple-Access Channels with Confidential Messages,” *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 976–1002, 2008.

[127] E. Ersen and U. Sennur, “Secrecy Capacity of a Class of Broadcast Channels with an Eavesdropper,” *EURASIP Journal on Wireless Communications and Networking*, 2009.

[128] H. Imai, K. Kobara, and K. Morozov, “On the Possibility of Key Agreement Using Variable Directional Antenna,” in *Proc. of 1st Joint Workshop on Information Security*, 2006.

[129] R. Wilson, D. Tse, and R. Scholtz, “Channel Identification: Secret Sharing Using Reciprocity in Ultrawideband Channels,” *IEEE Trans. Inf. Forensics and Security*, vol. 2, no. 3, pp. 364–375, 2007.

[130] J. Wallace and R. Sharma, "Automatic Secret Keys from Reciprocal MIMO Wireless Channels: Measurement and Analysis," *IEEE Trans. Inf. Forensics and Security*, vol. 5, no. 3, pp. 381–392, 2010.

VITA

Rémi Chou received the Engineering degree from Supélec, France, and the M.Sc. degree in Electrical and Computer Engineering from the Georgia Institute of Technology, Atlanta, in 2011, where he is currently pursuing the Ph.D. degree. His research interests include information theory, coding theory, and signal processing.