CODING TECHNIQUES FOR MULTI-USER PHYSICAL-LAYER SECURITY

A Dissertation
Presented to
The Academic Faculty

By

Alexandre J. Pierrot

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

CODING TECHNIQUES FOR MULTI-USER PHYSICAL-LAYER SECURITY

Approved by:

Dr. Matthieu R. Bloch, Advisor
Associate Professor,
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Mary Ann Weitnauer
Professor,
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. John R. Barry
Professor,
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Éric Feron
Dutton/Ducoffe Professor,
School of Aerospace Software Engineering
Georgia Institute of Technology

Dr. Steven W. McLaughlin
Professor and Steve W. Chaddick School Chair,
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: July 2, 2015

This thesis is dedicated to the memory of my grandfather

Louis Bœuf (1936–2006)

ACKNOWLEDGMENT

I truly owe this thesis to my beloved family, who provided with its love and support for the last 26 years. My mom, Sandrine, who taught me her sense of esthetics and details, my dad, Guillaume, who triggered my engineer curiosity, my grandmother, Paulette, who surrounded me with love and pride, my grandparents, Martine and Jean-Marie, who made me a citizen of the world, and finally my late grandfather, Louis, who asked me to design a rocket when I was 5, are the true authors of this work. I am confident that my dear youngest siblings, Maëva and Enzo, will benefit from the same love and will achieve even more.

I also found the force to go through these last five years thanks to my forever friends Alexandra, Isabelle, and Maxime, who have been constantly here for me. I will always remember that they have traveled for 14 hours to attend my proposal without hesitation. Together we have accumulated 40 years of study in Law, Medicine, and Engineering, after going to the same local high school in Marseille, supporting each other in this long adventure.

In France, Pierre-Alexandre, Manas, and Guillaume have been the best Ph.D. accomplices I could have had to clear the gray sky of Metz. Almost two years ago, Atlanta welcomed me like an old friend thanks to Romain and Andrew, Aude and Manu, Flo, Tim, and all the other AE folks. Atlanta also brought Kevin into my life, who has been an amazingly supportive partner to run the last miles, and hopefully the upcoming ones.

Last but not least, there is my advisor Matthieu, or should I say my friend Matthieu, who supported me far beyond being a simple professor: as a student, I was lucky to have found a devoted and passionate professor; as a person, I was lucky to have found an invaluable support. I am proud to be his first student and wish him the best for the next lucky ones working with him.

REMERCIEMENTS

Je dois, en réalité cette thèse à ma famille chérie, qui m'a donné tout son amour et son soutient lors de ces 26 dernières années. Les véritables auteurs de cette thèse sont ma mère, Sandrine, qui m'a enseigné son sens de l'esthétique et du détail, mon père, Guillaume, qui a su stimuler ma curiosité d'ingénieur, ma grand-mère, Paulette, qui m'a entouré de son amour et de sa fierté, mes grands parents, Martine et Jean-Marie, qui ont fait de moi un véritable citoyen du monde, et enfin, feu mon grand-père, Louis, qui m'a demandé de lui concevoir une fusée quand j'avais cinq ans. Je sais que mes plus jeunes frère et sœur adorés, Maëva et Enzo, bénéficieront du même amour et en feront encore bien plus.

J'ai aussi trouvé la force de traverser toutes ces années grâce à mes amis d'enfance, Alexandra, Isabelle et Maxime qui ont été là pour moi à chaque instant. Je me rappellerai toujours qu'ils ont voyagé 14 heures pour venir à mon proposal sans se poser la moindre question. Ensemble, nous avons cumulé plus de 40 ans d'études en Droit, Médecine et Ingénierie, se soutenant les uns les autres dans cette longue aventure.

En France, Pierre-Alexandre, Manas et Guillaume ont été les meilleurs comparses de doctorat que j'aurais pu avoir pour éclaircir le ciel gris de Metz. Il y a quasiment deux ans, Atlanta m'a accueilli comme un vieil ami grâce à Romain et Andrew, Aude et Manu, Flo, Tim et tous les autres de l'école de Génie Aérospatial. Atlanta a aussi apporté dans ma vie Kevin, qui a été un partenaire merveilleux durant ces dernières années, et je l'espère pour celles à venir.

Enfin, et pas des moindres, il y a mon directeur de thèse, Matthieu, ou devrais-je dire mon ami, Matthieu, qui m'a soutenu bien au delà de son rôle de professeur. En tant qu'étudiant, j'ai eu l'opportunité de travailler avec un professeur dévoué et passionné. En tant que personne, j'ai eu la chance de trouver un soutient indéfectible. Je suis fier d'être son premier étudiant et je lui souhaite le meilleur pour ceux à venir.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

Notation	Definition	Page
$\mathbb{F}_q$	Galois field with q elements	–
$\mathbb{N}$	set of natural numbers ($\mathbb{N}^*$ excludes 0)	–
$\mathbb{R}$	field of real numbers	–
$\mathbb{C}$	field of complex numbers	–
$\jmath$	imaginary unit $\jmath^2 = -1$	–
$\mathfrak{M}_{m,n}(\mathbb{K})$	set of matrices of dimension $m \times n$ with elements in the field $\mathbb{K}$	–
$\mathrm{GL}_n(\mathbb{K})$	general linear group of degree n over the field $\mathbb{K}$	–
$\mathcal{X}$	alphabet or set	–
$\lvert\mathcal{X}\rvert$	cardinality of $\mathcal{X}$	–
$\mathrm{cl}(\mathcal{X})$	closure of set $\mathcal{X}$	–
$\mathbb{1}\{\varpi\}$	indicator function: 1 if predicate ϖ is true, 0 otherwise	–
$\{x_i\}_n$	set with n elements $\{x_1, \ldots, x_n\}$	–
x	generic element of alphabet $\mathcal{X}$	–
$\lvert x\rvert$	absolute value of x	–
$\lceil x \rceil$	unique integer n such that $x \leqslant n < x + 1$	–
$\lfloor x \rfloor$	unique integer n such that $x - 1 < n \leqslant x$	–
$[\![x, y]\!]$	sequence of integers between $\lfloor x \rfloor$ and $\lceil y \rceil$	–
x^+	positive part of x, that is $x^+ = \max(x, 0)$	–
$\mathrm{sign}(x)$	$+1$ if $x \geqslant 0$, -1 otherwise	–
x^n	sequence $x_1, \ldots, x_n$	–
ϵ	usually, a "small" positive real number	–
$\delta(\epsilon)$	a function of ϵ such that $\lim_{\epsilon \to 0} \delta(\epsilon) = 0$	–
$\delta_\epsilon(n)$	a function of ϵ and n such that $\lim_{n \to \infty} \delta_\epsilon(n) = 0$	–

Notation	Definition	Page
$\delta(n)$	a function of n such that $\lim_{n\to\infty} \delta(n) = 0$	–
$\mathbf{x}$	column vector containing the n elements $x_1, x_2, \dots x_n$	–
$\mathbf{x}^\intercal$	transpose of $\mathbf{x}$	–
$\mathrm{wt}(\mathbf{x})$	Hamming weight of codeword $\mathbf{x}$	–
$\mathbf{H}$	matrix	–
$(h_{ij})_{m,n}$	$m \times n$ matrix whose elements are h_{ij} with $i \in [\![1, m]\!]$ and $j \in [\![1, n]\!]$	–
X	random variable implicitly defined on alphabet $\mathcal{X}$	–
p_X	probability distribution of random variable X	–
$X \sim p_X$	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	–
$p_{X\|Y}$	conditional probability distribution of X given Y	–
$\mathbb{E}_X$	expected value over the random variable X	–
$\mathrm{Var}(X)$	variance of the random variable X	–
$\mathbb{P}_X$	probability of an event over X	–
$\mathbb{H}(X)$	Shannon entropy of the discrete random variable X	14
$\mathbb{H}_\mathrm{b}$	binary entropy function	14
$\mathbb{H}_\alpha(X)$	Rényi entropy of order α of the discrete random variable X	18
$\mathbb{H}_\infty(X)$	min-entropy of the discrete random variable X	18
$\mathbb{H}_\infty^\epsilon(X)$	ϵ-smooth min-entropy of the discrete random variable X	19
$\mathbb{I}(X;Y)$	mutual information between random variables X and Y	15
$\mathbb{V}(p_X, p_{X'})$	(total) variational distance between p_X and $p_{X'}$	19
$\mathbb{D}(p_X\|\|p_{X'})$	Kullback-Leiber divergence between p_X and $p_{X'}$	20
$\mathbb{D}_\alpha(p_X\|\|p_{X'})$	Rényi divergence of order α between p_X and $p_{X'}$	21
$\mathcal{T}_\epsilon^n(X)$	strong typical set with respect to p_X	22
$\mathcal{T}_\epsilon^n(XY)$	strong joint-typical set with respect to p_{XY}	23

Notation	Definition	Page
$\mathcal{T}_\epsilon^n(XY\|x^n)$	conditional strong typical set with respect to p_{XY} and x^n	23
$\mathcal{A}_\epsilon^n(X)$	weak typical set with respect to p_X	24
$\mathcal{A}_\epsilon^n(XY)$	joint weak typical set with respect to p_{XY}	25
$\mathbf{P}_\mathrm{e}(\mathscr{C})$	probability of error of a code $\mathscr{C}$	–
$\mathbf{L}(\mathscr{C})$	information leakage of a code $\mathscr{C}$	–
$\mathbf{U}(\mathscr{S})$	uniformity of keys guaranteed by the strategy $\mathscr{S}$	–
$\operatorname*{p\text{-liminf}}_{n\to\infty}$	limit inferior in probability as n goes to ∞	45
$\operatorname*{p\text{-liminf}}_{n\to\infty}$	limit superior in probability as n goes to ∞	45

LIST OF ACRONYMS

Acronym	Definition	Ref.	Page
AES	Advanced Encryption Standard	[68]	2
AWGN	Additive White Gaussian Noise	[81]	70
BEC	Binary Erasure Channel	[23]	29
BP	Belief Propagation	[69]	104
BPSK	Binary Phase Shift Keying	[81]	100
BSC	Binary Symmetric Channel	[23]	29
CSI	Channel-State Information	–	97
DE	Density Evolution	[69]	108
DES	Data Encryption Standard	[68]	2
DMC	Discrete Memoryless Channel	[23]	29
DMS	Discrete Memoryless Source	[23]	32
LDPC	Low-Density Parity-Check	[69]	99
MAC	Multiple-Access Channel	[23]	57
MIMO	Multiple-Input Multiple-Output	–	81
NIST	National Institute of Standard and Technology	[68]	10
RSA	Rivest-Shamir-Adleman	[68]	2
SC-LDPC	Spatially-Coupled Low-Density Parity-Check	[57]	104
SDR	Software-Defined Radio	–	123
TWWTC	Two-Way WireTap Channel	[16]	52
USRP	Universal Software Radio Peripheral	[31]	123
WARP	Wireless Open-Access Research Platform	[70]	116

CHAPTER 1
INTRODUCTION

DSMOX LZWBW YFDGU LPJKK PCLCF AXT

This apparently meaningless sequence of letters is the encryption[1] of "Georgia Institute of Technology" by the Enigma machine, one of the most famous encryption devices in history. The Enigma machine, which was invented in 1919 by Hugo Koch and looked like a simple typing machine, played a pivotal role in the 20th century. This encryption device was used extensively by the German army during World War II and was considered unbreakable. However, the Allies and the team of cryptanalysts at Bletchley Park led by Alan Turing and Gordon Welchman proved them wrong. Even if the most complicated version of this machine used by the *Kriegsmarine* had nearly 159 quintillion possible settings, the efforts of the cryptanalysts allowed the Allies to break the code and disclose essential pieces of information regarding the German U-boats. Turing's team used several flaws in the encryption mechanism that yielded residual correlation between the cyphertext and the plaintext. This correlation originated from a major design flaw of the reflector ring, which was introduced to perform both encryption and decryption with the same device and settings (reciprocity). Using such a mechanism prevents a letter from being encrypted as itself. From this observation, Turing developed a method to locate *cribs*, which are sequences one could expect to be sent (e.g. "Weather Report"), within the cyphertext. Using other flaws in the encryption mechanism and an electromechanical device called the *Bombe*, the Allies were finally able to collect valuable intelligence and gain a significant advantage over the Germans. Breaking the Enigma code, or at least being able to decipher some messages, helped the Allies weaken the U-boats' dominance in the North Atlantic Ocean and conduct the Normandy landings. Without Enigma, the operation could have been delayed

[1]The setting for encryption/decryption for this introduction is PHD for the key setting, AJP for the ring setting, rotors I, II and III, and plugboard setting PO ML IU KJ NH YT GB VF RE DC.

by at least a year, with millions of lives at stake and a potential stalemate between the Allies and the Axis.

Half a century later, cryptography is no longer a science limited to war times and has settled into everyone's life, whether one is using a credit card, filing taxes online, connecting to a Wi-Fi network or placing a phone call. Even if the Enigma machine looks quite primitive compared to the modern cryptographic systems we use every day (e.g., RSA, DES, AES), the most recent algorithms are not impervious to attacks. The endless growth of networks has made security one of the main challenges of the 21st century to guarantee end-user privacy, protection of industrial secrets, and military confidentiality.

Thus far, the development of cryptography has followed a succession of new encryption schemes and new attacks, leading to sophisticated solutions (e.g., elliptic curves, quantum cryptography). Although there is no guarantee that widespread systems such as AES or RSA will remain secure forever,[2] the main problem may be hidden elsewhere. More than any potential weakness of a given cryptographic scheme, the way the end-user uses this scheme can cause major security breaches and compromise the overall security of an information system. For instance, using a sophisticated encryption algorithm for an email account is ineffective when the password is as common as `123456` or `password`. Similarly, using an AES encryption algorithm for a Wi-Fi hotspot provides little security if the password remains the same for years.

The paradigm behind information-theoretic security is different in many extents. If cryptographic security relies on the computational complexity of cracking the system, information-theoretic security guarantees statistical independence between the secret messages and eavesdropped communications. The security of such systems does not rely on the expected secrecy of a key, but rather on the fact that the eavesdropper observation does not provide any information regarding the legitimate messages. The goal is to provide mechanisms immune to cryptanalysis and end-user shortcoming.

[2] If cryptanalysis can be used against some of these algorithms, using long enough keys still suffices to consider them as secure for public use.

Physical-layer security provides secrecy at virtually no cost and opens promising perspectives by leveraging the imperfections of the physical medium. Physical-layer security mechanisms are proposed as another layer of security to improve the security of existing communications systems, requiring neither extensive changes nor additional end-user inputs. Physical-layer security exploits not only the random behavior of a communication channel, but also the feedback and interference arising from bidirectional communications.

This dissertation focuses on some of the main information-theoretic security techniques for multi-user communications and encompasses both theoretical and experimental results. Theory shows the fundamental limits of those techniques and gives insight into the design of actual physical-layer security systems. Experimentation then shows how well those systems would work in a real setting, highlighting additional issues that must be taken into account before practical implementation.

This dissertation is organized as follows. This chapter provides further details on the basic mechanisms behind the notion of physical-layer security. Chapter 2 recalls the main information theory notions used to prove the results presented in this dissertation. Chapter 3 introduces the simplest theoretical model to analyze multi-user physical-layer security, namely the two-way wiretap channel. Chapter 4 focuses on the challenges of implementing of secret-key generation mechanisms based on channel variations. Chapter 5 presents practical codes designed to exploit codeword interference between users for secrecy purposes.

1.1 From Mathematical Cryptography to Physical-Layer Security

The rapid development of information networks and the increasing diversity of exchanged data bring both promising advances and new potential issues. For commercial purposes, the historical challenge was to transfer data as fast as possible, whereas the problem of secrecy was only a concern for military, diplomatic, and industrial data. This difference was and remains due to the required balance between data sensitivity and security complexity. However, new information technology paradigms—such as cloud storage and computing, near field communications, biometric passports, and online banking—make the problem of security part of everyone's

life. The historical approach, called *cryptography*, is to use allegedly complicated mathematical problems to encrypt data. Cryptographic techniques range from the simple "Cæsar's cypher," which consists in permuting letters according to a predefined order, to the elaborate elliptic curves cryptography. This evolution stems from the fact that, whenever an attack is designed to break a cryptographic scheme, it has to be patched or superseded. The exponential growth of the available computational power makes more and more schemes obsolete. New problems have recently appeared with the diversification of users. If military or diplomatic agencies generally know the rules to guarantee that a cryptographic scheme is secure—especially regarding its secret-key characteristics—it might not be the case for an average user. For instance, few people use distinct complex passwords across different services and people usually keep the same Wi-Fi password forever. The main security concern does not lie in the strength of the cryptographic primitives used to protect data but in the way people use them. One solution to overcome this problem is to avoid input from the end user, for instance, by providing a key. *Physical-layer security*, which directly uses the inherent random behavior of communication channels to provide secrecy, is one of the solutions to avoid a user external input. For instance, it may be possible to refresh a Wi-Fi password automatically by using some random variations of the wireless link without the user help. The idea is not to replace the well-tested cryptographic schemes, but to augment them in a cost-effective manner. The analysis and the development of physical-layer security primitives involve the notion of information-theoretic security, which is presented in the next section.

1.2 Information-Theoretic Security: the Wiretap Channel

The first information theoretic model that introduces physical layer security was provided by Wyner in 1975 [107]. The analysis of this model called *the degraded wiretap channel* introduces many mathematical tools to consider security constraints in a communication system. As illustrated in Figure 1.1, this model is an extension of the point-to-point channel introduced by Shannon in which a sender called Alice tries to communicate with a receiver called Bob. Wyner's model further considers a third user called Eve that eavesdrops a *degraded* version of

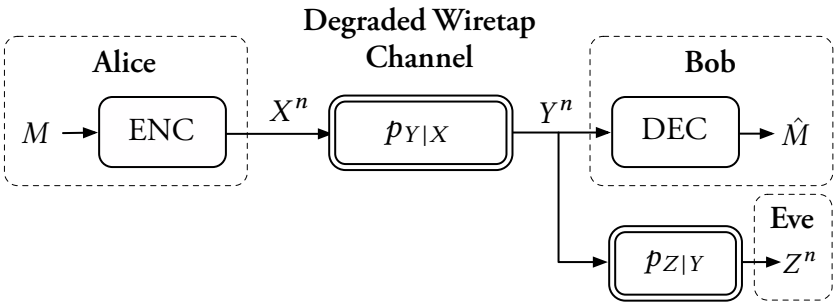

Figure 1.1 – *Communications over the degraded wiretap channel.*

the main channel output. The objective is to ensure both reliable communications between Alice and Bob and secrecy with respect to Eve, which should be prevented from getting information about transmitted messages.

Mathematically, the channel is a one-input two-output channel with transition probability $p_{Y^nZ^n|X^n}$, which is, in the case of a *memoryless* degraded wiretap channel,

$$p_{Y^nZ^n|X^n}(y^n, z^n|x^n) = \prod_{i=1}^{n} p_{Y|X}\left(y^{(i)}|x^{(i)}\right) p_{Z|Y}\left(z^{(i)}|y^{(i)}\right).$$

The reliability metric for a wiretap code $\mathscr{C}$ is, as usual, the average probability of error $\mathbf{P}_e(\mathscr{C}) \triangleq \mathbb{P}(\hat{M} \neq M|\mathscr{C})$ between the sent message M and the estimated received message $\hat{M}$.

The secrecy metric measures statistical independence with the mutual information $\mathbb{I}(M; Z^n)$. For a given code $\mathscr{C}$, the quantity $\mathbf{L}(\mathscr{C}) \triangleq \mathbb{I}(M; Z^n|\mathscr{C})$ is called the information *leakage* and corresponds to a *strong secrecy* metric. For cases difficult to analyze with a strong secrecy criterion, there exists a *weak secrecy* metric called the *leakage rate* $\underline{\mathbf{L}}(\mathscr{C}) \triangleq \frac{1}{n}\mathbb{I}(M; Z^n|\mathscr{C})$. Leakage rate is a weaker metric because of the normalization by the blocklength n. Many other metrics can be considered with different levels of guaranteed security [17].

If sent messages are chosen uniformly at random in a set of 2^{nR} elements, R corresponds to the rate of the code. A rate is strongly (resp. weakly) achievable if there exists a code $\mathscr{C}$ such that both $\mathbf{P}_e(\mathscr{C})$ and $\mathbf{L}(\mathscr{C})$ (resp. $\underline{\mathbf{L}}(\mathscr{C})$) go to zero as n goes to infinity; the supremum of all achievable rates is called the strong (resp. weak) *secrecy capacity* of the channel.

In the case of weakly symmetric channels, [60] shows that the secrecy capacity is the difference between the capacity of the main channel and the capacity of the eavesdropper channel.

In particular it means that the secrecy capacity is zero as soon as the eavesdropper's channel is less noisy than the main channel—that is, $\mathbb{I}(X;Y) \leqslant \mathbb{I}(X;Z)$.

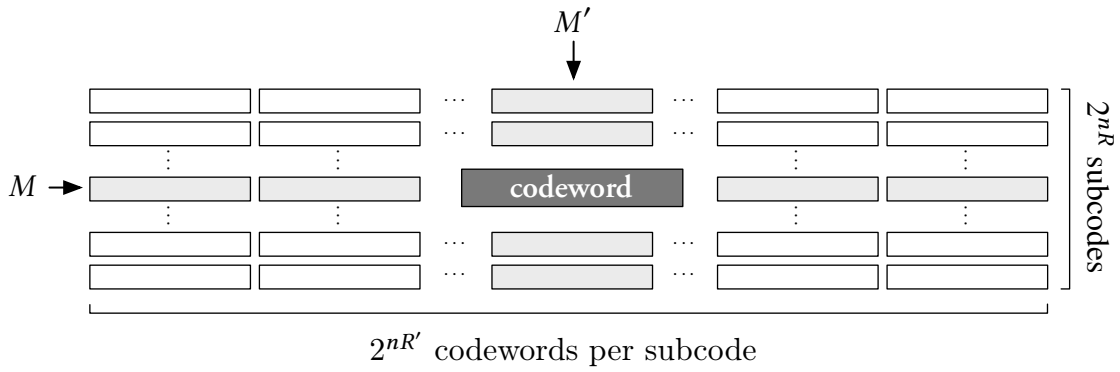

Figure 1.2 – *Nested structure of a wiretap code.*

Random encoding The crucial part to provide secrecy is to introduce randomness in the encoder. Intuitively, the randomness is introduced to mislead and confuse the eavesdropper, but it can be formally proven with information theoretic tools. The randomness is introduced with a random auxiliary message M' drawn uniformly at random in $[\![1, 2^{nR}]\!]$. As represented in Figure 1.2, a *wiretap code* has a *nested* structure: message M selects one of 2^{nR} subcodes and M' picks a codeword within the $2^{nR'}$ possible codewords of the subcode.

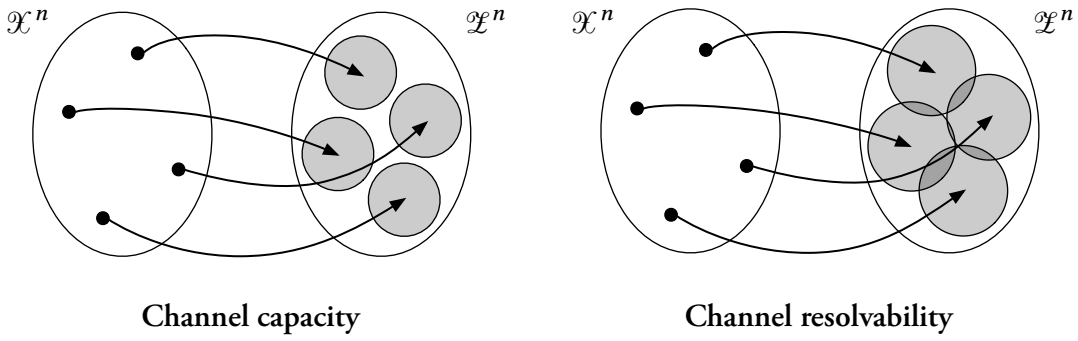

Figure 1.3 – *Capacity versus resolvability.*

Capacity vs. resolvability To find the capacity of a wiretap channel, one has to identify the optimal auxiliary message R' rate that introduces enough randomness to mislead the eavesdropper without wasting the communication rate. There are two conceptual approaches to

deal with this problem, which are illustrated in Figure 1.3. The first called *coding for channel capacity* relies on the fact that $\mathbb{I}(M;Z^n) = \mathbb{I}(X^n;Z^n) - \mathbb{H}(M') + \mathbb{H}(M'|MX^n)$. It may seem logical that $\mathbb{H}(M')$ cancels $\mathbb{I}(X^n;Z^n)$, if the encoder is random enough; however, dealing with $\mathbb{H}(M'|MX^n)$ is far less intuitive. Indeed, to cancel $\mathbb{H}(M'|MX^n)$ the code must allow a virtual user obtaining Z^n and M to retrieve M'. This can be done if subcodes are *capacity-achieving codes* for Eve's channel with auxiliary rates R' lower than the eavesdropper's channel capacity. The second approach, called *coding for channel resolvability*, attempts to design subcodes such that Eve's output distribution is independent of the sent messages, making them indiscernible. This time, the solution is to work at auxiliary message rates R' higher than the eavesdropper's channel. Even if both approaches give the same secrecy capacity under a weak secrecy criterion, it is only possible to provide strong secrecy with resolvability-based codes [17, 64].

1.3 Multi-User Information-Theoretic Security

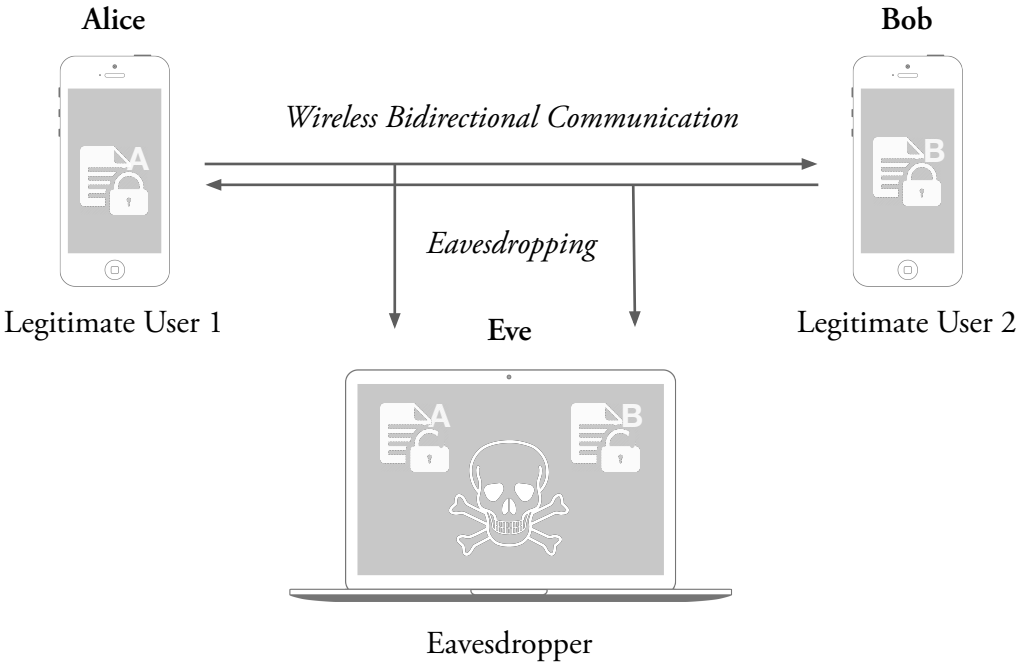

Figure 1.4 – *Simple multi-user scheme with eavesdropper.*

The degraded wiretap channel introduced by Wyner does not take into account several aspects of muti-user communications. For the sake of illustration, consider the setting represented in Figure 1.4, which consists of two legitimate users trying to communicate *reliably* and *secretly* with respect to a third-party user that eavesdrops on the conversation. Even if the goals (reliability and secrecy) are the same as for the wiretap channel, considering bidirectional communications brings new ways of providing secrecy. For instance, the legitimate parties may use the *interference* between transmitted signals to "hide" secret information, and the *feedback* to exchange side information to "increase" secrecy. A more formal analysis is required to clarify the notions of "hiding information" and "increasing secrecy".

Considering a multi-user scheme with two users communicating in two directions as represented in Figure 1.4 takes into account both feedback and interference. A scheme involving more users would allow to consider other securing techniques such as relaying [58] or multi-user secret-key generation [25]. Another limitation related to this model is the assumption that the eavesdropper doesn't communicate over the channel. For instance, an active eavesdropper would be able to jam the communication and interfere with the legitimate users.

1.3.1 Cooperative Jamming and Coded cooperative jamming

A natural attempt to increase secure communication rates consists in jamming Eve with noise to decrease her signal-to-noise ratio. This strategy, called *cooperative jamming (with noise)* [62], forces one user to stop transmitting information to jam the eavesdropper. To overcome this limitation, Alice and Bob can use codewords whose interference also has a detrimental effect on Eve without sacrificing as much information rate. This scheme is called *coded cooperative jamming* and was introduced by Tekin and Yener [92, 93].

1.3.2 Key Exchange

Key exchange works as follows: one user sacrifices part of its secret rate to transmit a secret-key to the other, which then uses the secret-key to encrypt a message, thus augmenting its secret rate. This mechanism illustrated in Figure 1.5 only transfers secure rate from one user to the other, but does not generate new secrecy.

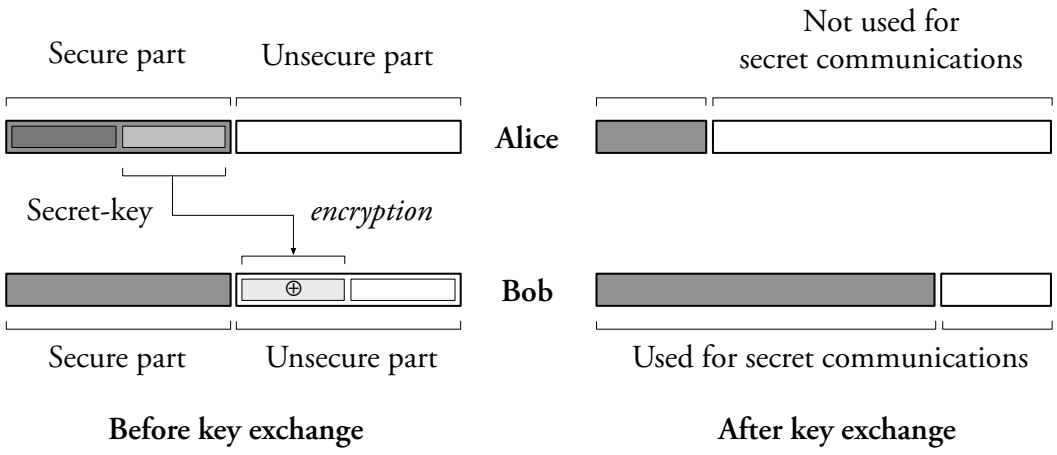

FIGURE 1.5 – *Secret-key exchange.*

1.3.3 Secret-Key Generation Strategy

It is also possible to provide secrecy by using the physical layer as a source of randomness. Suppose Alice, Bob and Eve observe the outputs of a correlated continuous source, respectively denoted $(X, Y, Z) \in \mathscr{X} \times \mathscr{Y} \times \mathscr{Z}$. To agree on a common key, the legitimate users exchange messages over an authenticated, noiseless, public channel with unlimited capacity. Let $\mathscr{K} \triangleq [\![1, 2^{nR}]\!]$ denote the alphabet for the key, where R is called the *secret-key rate*.

A secret-key generation strategy with unlimited public communication $\mathscr{S}_n$ consists of the following operations: Alice observes n realizations of the source X^n while Bob observes Z^n; Alice transmits a public message F; Alice computes $K(X^n) \in \mathscr{K}$ while Bob computes $\hat{K}(Y^n, F) \in \mathscr{K}$.

The performance of such a secret-key generation strategy $\mathscr{S}_n$ is assessed in terms of:

1. *reliability*, measured by the average probability of error $\mathbf{P}_{\mathrm{d}}(\mathscr{S}_n) \triangleq \mathbb{P}(K \neq \hat{K}|\mathscr{S}_n)$;
2. *(strong) secrecy*, measured by the leakage $\mathbf{L}(\mathscr{S}_n) \triangleq \mathbb{I}(K; Z^n F|\mathscr{S}_n)$;
3. *(strong) uniformity*, measured by the quantity $\mathbf{U}(\mathscr{S}_n) \triangleq \log\lceil 2^{nR}\rceil - \mathbb{H}(K|\mathscr{S}_n)$.

The conditioning on $\mathscr{S}_n$ reflects the fact that the eavesdropper knows the strategy.

A secret-key rate R is achievable if all those three metrics tend to zero as n goes to infinity, and the supremum of all these achievable rates is called the *secret-key capacity*, which is denoted by C_s.

1.4 Experimental Aspects

The information theory results rely on the existence of random phenomena. For several physical reasons (noise, environment modifications, mobility, etc.), a wireless channel exhibits some variability that can be exploited as a source of randomness in a security system. In particular, one can exploit the *reciprocity* of wireless channels to generate strongly correlated observations between two users, and the *diversity* between channels so that an external eavesdropper obtains little information about the correlated observations.

Several works, such as [49, 63, 72, 98, 109–111] and references therein, study the problem of secret-key generation with reconciliation and privacy amplification [9, 67]. They used different solutions to induce the source of randomness from the channel (gain, phase, etc.). However, the security analysis is often performed with metrics (probability of error for the eavesdropper, NIST tests, decorrelation, etc.) that do not guarantee information-theoretic security. From an information-theoretic perspective, their analyses do not suffice to ensure secrecy, which is one of the objectives addressed in the next chapters.

1.5 Contributions

This dissertation extends multiple aspects of multi-user communications from the theoretical model to practical implementations.

Section 2.3, parts of which have been published in [74], investigates the separation of channel intrinsic randomness and channel resolvability. The proposed joint exponents are compared to the tandem exponents obtained with a separate approach. This proves at once, achievability results for channel intrinsic randomness, random number generation, and channel resolvability.

Chapter 3, parts of which have been published in [76], considers the problem of secure communications over the two-way wiretap channel under a strong secrecy criterion through a resolvability-based approach. This also improves on previous works by developing an achievable region based on strategies that exploit both the interference at the eavesdropper's terminal and cooperation between legitimate users. This chapter shows how the artificial noise created by

cooperative jamming induces a source of common randomness that can be used for secret-key agreement. The proposed coding technique shows significant improvements for different configurations of the Gaussian two-way wiretap channel.

Chapter 4, parts of which have been published in [75], analyzes the practical limitations of a secret-key generation system from channel gain variations in a narrowband wireless environment. In particular, different assumptions usually made for theoretical purposes are verified with an actual system based on software-defined radios. It is important not only to characterize the source of common randomness induced by channel gain variations, but also to estimate achievable secret-key rates in the finite key-length regime. The secret-key generation system based on channel gain variations is extremely sensitive to external modifications of the environment, and such system should adapt accordingly to guarantee a given level of information-theoretic secrecy.

Chapter 5, parts of which have been published in [77], presents a practical coded cooperative jamming scheme for the problem of secure communications over the two-way wiretap channel. This scheme uses low-density parity-check (LDPC) based codes whose codewords interfere at the eavesdropper's terminal, thus providing secrecy. This chapter offers a comparison between constructions based on classical LDPC codes and spatially coupled LDPC codes, and shows that the latter guarantees low information leakage rate.

List of Publications

[74] **Pierrot, A. J.**, Bloch, M. R., "Joint Channel Intrinsic Randomness and Channel Resolvability". In: *Proceedings of the 2013 IEEE Information Theory Workshop (ITW)*. Sept. 2013, pp. 1–5

[76] **Pierrot, A. J.**, Bloch, M. R., "Strongly Secure Communications Over the Two-Way Wiretap Channel". In: *IEEE Transactions on Information Forensics and Security* 6.3 (Sept. 2011), pp. 595–605

[77] **Pierrot, A. J.**, Bloch, M. R., "LDPC-Based Coded Cooperative Jamming Codes". In: *Proceedings of the IEEE Information Theory Workshop*. Lausanne, Switzerland, Sept. 2012, pp. 462–466

[75] **Pierrot, A. J.**, Chou, R. A., Bloch, M. R., "Experimental Aspects of Secret Key Generation in Indoor Wireless Environments". In: *Proceedings of the 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. June 2013, pp. 669–673

CHAPTER 2

INFORMATION-THEORETIC SECURITY[1]

In this chapter, several metrics are introduced to clarify notation and recall several well-known properties, which appear with further details in [16, 23]. These information-theoretic properties are the base to introduce several notions including:

- the *channel capacity*, which represents the maximum number of bits one can reliably transmit;
- the problem of *source coding with side information*, which considers the problem of compression for a correlated source of randomness;
- the *channel intrinsic randomness*, which defines the maximum uniform randomness that can be extracted from a channel independently of its input;
- the *channel resolvability*, which corresponds to the process of transforming a uniform random number into another one with a different distribution.

All these notions are the essential ingredients to analyze the security of multi-user schemes and are useful to provide asymptotical limits, error exponents, finite-length results, and insight into the design of practical systems. In particular, channel intrinsic randomness and source coding with side information are both essential results for secret-key generation, while channel resolvability is used to show the fundamental limits of strongly secure schemes.

2.1 Tools of Information Theory

2.1.1 Entropy and Mutual Information

The *entropy*, which was introduced by Shannon, is a statistical metric of the information, or, more precisely, of the uncertainty of the outcome of a random variable.

[1]Parts of the material in Section 2.3 have appeared in [74]: **Pierrot, A. J.**, Bloch, M. R., "Joint Channel Intrinsic Randomness and Channel Resolvability". In: *Proceedings of the 2013 IEEE Information Theory Workshop (ITW)*. Sept. 2013, pp. 1–5. ©IEEE 2013.

Definition 1 Let $X \in \mathcal{X}$ be a discrete random variable with distribution p_X. The *Shannon entropy* (or entropy for short) of X is defined as

$$\mathbb{H}(X) \triangleq -\sum_{x \in \mathcal{X}} p_X(x) \log p_X(x),$$

with the convention that $0 \log 0 \triangleq 0$. Unless otherwise specified, all logarithms are taken to the base two and the unit for entropy is called a bit. ◊

If $\mathcal{X} = \{0, 1\}$, then X is a binary random variable and its entropy solely depends on the parameter $p = \mathbb{P}(X = 0)$. The *binary entropy function* is defined as

$$\mathbb{H}_{\mathrm{b}}(p) \triangleq -p \log p - (1-p) \log(1-p). \tag{2.1}$$

For instance for $X \sim \mathcal{B}(p)$, if p is equal to $1/2$, the uncertainty on X is maximal and $\mathbb{H}_{\mathrm{b}}(p) = 1$.

Lemma 2.1 For any discrete random variable $X \in \mathcal{X}$

$$0 \leqslant \mathbb{H}(X) \leqslant \log |\mathcal{X}|.$$

The equality $\mathbb{H}(X) = 0$ holds if and only if X is a constant while the equality $\mathbb{H}(X) = \log |\mathcal{X}|$ holds if and only if X is uniform on $\mathcal{X}$. ◊

This first simple lemma is of primary importance since the notion of uniformity is one of the pivotal aspects of information-theoretic security. Uniformity corresponds to the maximum possible uncertainty of a random variable, which is one of the desirable properties for a secret-key.

The notion of entropy can be further extended to joint and conditional distributions.

Definition 2 Let $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$ be two discrete random variables with joint distribution p_{XY}. The *joint entropy* of X and Y is defined as

$$\mathbb{H}(XY) \triangleq -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_{XY}(x, y) \log p_{XY}(x, y). \qquad \diamond$$

Definition 3 Let $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$ be two discrete random variables with joint distribution p_{XY}. The *conditional entropy* of X given Y is defined as

$$\mathbb{H}(Y|X) \triangleq -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p_{XY}(x,y)\log p_{Y|X}(y|x)\,. \qquad \diamond$$

Using Bayes' rule, one can obtain the following relationship called the *chain rule of entropy*

$$\mathbb{H}(XY) = \mathbb{H}(X) + \mathbb{H}(Y|X)\,.$$

This also generalizes to the entropy of a random vector $X^n = (X_1, \ldots, X_n)$ as

$$\mathbb{H}(X^n) = \sum_{i=1}^{n} \mathbb{H}\left(X_i|X^{i-1}\right),$$

with the convention that $\mathbb{H}\left(X_1|X^0\right) \triangleq \mathbb{H}(X_1)$.

Lemma 2.2 ("**Conditioning does not increase entropy**") Let $X \in \mathcal{X}$ be a discrete random variable, $Y \in \mathcal{Y}$ either discrete or continuous, with joint distribution p_{XY}. Then,

$$\mathbb{H}(X|Y) \leqslant \mathbb{H}(X)\,. \qquad \diamond$$

In other words, this lemma asserts that the knowledge of Y cannot increase the uncertainty about X. From an information-theoretic point of view, it illustrates that it is important to take into account the side information an eavesdropper can obtain since it may reduce its uncertainty on the content of the legitimate communication.

The introduction of joint and conditional entropies illustrates how two random variables can be related. For instance, the *mutual information* between two variables X and Y represents the difference between the uncertainty on X and the uncertainty on X when Y is known.

Definition 4 Let $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$ be two discrete random variables with joint distribution p_{XY}. The *mutual information* between X and Y is defined as

$$\mathbb{I}(X;Y) \triangleq \mathbb{H}(X) - \mathbb{H}(X|Y)\,.$$

Let $X \in \mathcal{X}$, $Y \in \mathcal{Y}$ and $Z \in \mathcal{Z}$ be discrete random variables with joint distribution p_{XYZ}. The conditional mutual information between X and Y given Z is

$$\mathbb{I}(X;Y|Z) \triangleq \mathbb{H}(X|Z) - \mathbb{H}(X|YZ)\,. \qquad \diamond$$

The mutual information between X and Y can be also viewed as the expectation of the following random variable:

$$I(X;Y) \triangleq i_{X;Y}(X;Y) \text{ with } i_{X;Y} \triangleq \log \frac{p_{XY}}{p_X p_Y}.$$

Using the chain rule of entropy, one obtains the *chain rule of mutual information*

$$\mathbb{I}(X^n;Y) = \sum_{i=1}^{n} \mathbb{I}\left(X_i;Y|X^{i-1}\right),$$

with the convention that $\mathbb{I}\left(X_1;Y|X^0\right) \triangleq \mathbb{I}(X_1;Y)$.

Lemma 2.3 Let $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$ be two discrete random variables with joint distribution p_{XY}. Then,

$$0 \leqslant \mathbb{I}(X;Y) \leqslant \min\left(\mathbb{H}(X), \mathbb{H}(Y)\right).$$

The equality $\mathbb{I}(X;Y) = 0$ holds if and only if X and Y are independent. The equality $\mathbb{I}(X;Y) = \mathbb{H}(X)$ (resp. $\mathbb{I}(X;Y) = \mathbb{H}(Y)$) holds if and only if X is a function of Y (resp. Y is a function of X). ◊

Along with the notion of uniformity, the notion of independence is also important for information-theoretic security. Ensuring that the eavesdropper in a security scheme observes something completely independent of the legitimate communication guarantees the maximum desirable level of security. Guaranteeing independence avoids justifying security by means of computational complexity as it is done in classic cryptography.

Lemma 2.4 Let $X \in \mathcal{X}$, $Y \in \mathcal{Y}$ and $Z \in \mathcal{Z}$ be three discrete random variables with joint distribution p_{XYZ}. Then,

$$0 \leqslant \mathbb{I}(X;Y|Z) \leqslant \min\left(\mathbb{H}(X|Z), \mathbb{H}(Y|Z)\right).$$

The equality $\mathbb{I}(X;Y|Z) = 0$ holds if and only if X and Y are conditionally independent given Z. In this case, $X \to Y \to Z$ is called a Markov chain. The equality $\mathbb{I}(X;Y|Z) = \mathbb{H}(X|Z)$ (resp. $\mathbb{I}(X;Y|Z) = \mathbb{H}(Y|Z)$) holds if and only if X is a function of Y and Z (resp. Y is a function of X and Z). ◊

Lemma 2.5 (Data Processing Inequality) Let $X \in \mathcal{X}$, $Y \in \mathcal{Y}$ and $Z \in \mathcal{Z}$ be three discrete random variables such that $X \to Y \to Z$ forms a Markov chain. Then,

$$\mathbb{I}(X;Y) \geqslant \mathbb{I}(X;Z). \qquad \diamond$$

This inequality is equivalent to $\mathbb{H}(X|Y) \leqslant \mathbb{H}(X|Z)$, which means that, on average, processing Y can only increase the uncertainty about X. This property illustrates, for instance, that it is important not to suppose than the eavesdropper processes its observation to fully assess the security of a secrecy scheme.

Lemma 2.6 (Fano's Inequality) Let $X \in \mathcal{X}$ be a discrete random variable and let $\hat{X}$ be any estimate of X that takes values in the same alphabet $\mathcal{X}$. Let $\mathbf{P}_e \triangleq \mathbb{P}\big(X \neq \hat{X}\big)$ be the probability of error obtained when estimating X with $\hat{X}$. Then,

$$\mathbb{H}\big(X|\hat{X}\big) \leqslant \mathbb{H}_b(\mathbf{P}_e) + \mathbf{P}_e \log(|\mathcal{X}| - 1),$$

where $\mathbb{H}_b(\mathbf{P}_e)$ is the binary entropy function defined earlier. $\diamond$

Fano's inequality is another element of several proofs since it relates the information-theoretic quantity $\mathbb{H}\big(X|\hat{X}\big)$ to an operational quantity, the probability of error $\mathbf{P}_e$.

Since encoding and decoding operations correspond to applying a function to random variables, it is important to characterize the effects of such processing on the information metrics mentioned above. If there is no general rule for all the functions, some classes of functions exhibit interesting behaviors from an information-theoretic perspective. It is, for instance, the case for convex functions.

Definition 5 A function $f : \mathcal{I} \to \mathbb{R}$ defined on a set $\mathcal{I}$ is *convex* on $\mathcal{I}$ if for all $(x_1, x_2) \in \mathcal{I}^2$ and for all $\lambda \in [0, 1]$,

$$f(\lambda x_1 + (1-\lambda)x_2) \leqslant \lambda f(x_1) + (1-\lambda) f(x_2).$$

If the equation above holds with strict inequality, f is strictly convex on $\mathcal{I}$. A function $f : \mathcal{I} \to \mathbb{R}$ defined on a set $\mathcal{I}$ is (strictly) concave on $\mathcal{I}$ if the function $-f$ is (strictly) convex on $\mathcal{I}$. $\diamond$

Applying a convex function to a random variable yields the following properties.

Lemma 2.7 (Jensen's inequality) Let $X \in \mathcal{X}$ be a random variable and let $f : \mathcal{X} \to \mathbb{R}$ be a convex function. Then,

$$\mathbb{E}(f(X)) \geqslant f(\mathbb{E}(X)).$$

If f is strictly convex, then equality holds if and only if X is a constant. ◊

By exploiting the convexity of $-\log$ and Jensen's inequality, several properties ensue.

Lemma 2.8 Let $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$ be two discrete random variables with joint distribution p_{XY}. Then

- $\mathbb{H}(X)$ is a concave function of p_X;
- $\mathbb{I}(X;Y)$ is a concave function of p_X for $p_{Y|X}$ fixed;
- $\mathbb{I}(X;Y)$ is a convex function of $p_{Y|X}$ for p_X fixed. ◊

2.1.2 Rényi Entropy

There exist several variations of the notion of entropy in the literature. These variations are useful tools in information theory even if their operational meaning is not as intuitive as the Shannon entropy or the mutual information. The first variation is called the Rényi entropy [84] and is meant to be a generalization of the entropy that would preserve fundamental properties, such as the additivity of independent events.

Definition 6 Let $X \in \mathcal{X}$ be a discrete random variable with distribution p_X. Let $\epsilon \geqslant 0$. The *Rényi entropy* of order $\alpha \in \mathbb{R}_+ \backslash \{1\}$ of X is defined as

$$\mathbb{H}_\alpha(X) \triangleq \frac{1}{1-\alpha} \log \sum_{x \in \mathcal{X}} p_X(x)^\alpha. \tag{2.2}$$

◊

One can see the Rényi entropy as a counterpart of the p-norms in Euclidean spaces. The limit of $\mathbb{H}_\alpha(X)$, when α goes to one, is equal to the Shannon entropy $\mathbb{H}(X)$. The second order Rényi entropy is also called the collision entropy (also denoted $\mathbb{H}_c(X)$) and appears in

the problem of privacy amplification. The limit of $\mathbb{H}_\alpha(X)$ as α goes to infinity is called the min-entropy

$$\mathbb{H}_\infty(X) \triangleq \lim_{\alpha\to\infty} \frac{1}{1-\alpha} \log \sum_{x\in\mathcal{X}} p_X(x)^\alpha = -\log \max_{x\in\mathcal{X}} P_X(x). \tag{2.3}$$

In 2004, Renner and Wolf [83] further extended the notion of Rényi entropy for operational purposes in secret-key generation.

Definition 7 Let $X \in \mathcal{X}$ be a discrete random variable with distribution p_X. Let $\epsilon \geqslant 0$. The *ϵ-smooth Rényi entropy* of order α, with $\alpha \in \mathbb{R}_+^* \setminus \{1\}$, of X is defined as

$$\mathbb{H}_\alpha^\epsilon(X) \triangleq \frac{1}{1-\alpha} \inf_{q_X \in \mathcal{B}^\epsilon(X)} \log \sum_{x\in\mathcal{X}} q_X(x)^\alpha, \tag{2.4}$$

where $\mathcal{B}^\epsilon(X) \triangleq \{q_X, \mathbb{V}(p_X, q_X) \leqslant \epsilon\}$ is the set of distributions q_X that are ϵ close to p_X in terms of variational distance. ◊

The quantity $\mathbb{H}_\alpha^\epsilon(X)$ converges as α goes to infinity to the ϵ-smooth min-entropy. All these metrics have conditional formulations [6, 33, 51], in particular for two random variables X and Y, one definition of the conditional ϵ-smooth min-entropy of X given Y is

$$\mathbb{H}_\infty^\epsilon(X|Y) = \max_{q_{XY}\in\mathcal{B}^\epsilon(XY)} \min_{y\in\mathcal{Y}} \min_{x\in\mathcal{X}} \log \frac{p_Y(y)}{q_{XY}(XY)}. \tag{2.5}$$

2.1.3 Other Metrics

The previous section has illustrated the spirit of information theory, which is how one can relate the notion of information to quantities as abstract as probability distributions. In this section, several other metrics are introduced, and even if their meaning is not as intuitive as the entropy or the mutual information, they are primal intermediaries to conduct proofs.

Definition 8 Let X and X' be two discrete random variables defined on the same alphabet $\mathcal{X}$. The *(total) variational distance* between X and X' is

$$\mathbb{V}(p_X, p_{X'}) \triangleq \max_{\mathcal{A}\subseteq\mathcal{X}} \left(\mathbb{P}_X(\mathcal{A}) - \mathbb{P}_{X'}(\mathcal{A})\right) \equiv \frac{1}{2} \sum_{x\in\mathcal{X}} \left|p_X(x) - p_{X'}(x)\right|. \qquad \diamond$$

This distance is simply half of the L_1 distance between two distributions and all the associated properties (in particular the triangular inequality) hold.

Proposition 2.9 (**Triangular inequality**) Let X, X', and X'' be three discrete random variables defined on the same alphabet $\mathcal{X}$. Then,

$$\mathbb{V}(p_{X,X''}) \leqslant \mathbb{V}(p_X, p_{X'}) + \mathbb{V}(p_{X'}, p_{X''}). \qquad \diamond$$

There is also a counterpart of the data processing inequality for the variational distance.

Proposition 2.10 (**Data Processing Inequality**) Let X and X' be two random variables defined on the same alphabet $\mathcal{X}$. Consider a function $f : \mathcal{X} \to \mathcal{Y}$ mapping X to Y and X' to Y'. Then,

$$\mathbb{V}(p_Y, p_{Y'}) \leqslant \mathbb{V}(p_X, p_{X'}). \qquad \diamond$$

Fortunately, the variational distance can be related to the entropy through the following inequality [26].

Proposition 2.11 (**Csiszár & Körner**) Let X and X' be two discrete random variables defined on the same alphabet $\mathcal{X}$. Then,

$$\left|\mathbb{H}(X) - \mathbb{H}(X')\right| \leqslant \mathbb{V}(p_X, p_{X'}) \log \frac{|\mathcal{X}|}{\mathbb{V}(p_X, p_{X'})}. \qquad \diamond$$

This proposition can also relate the mutual information to the variational distance [50].

Corollary 2.12 Let X and Y be two discrete random variables defined on their respective alphabets $\mathcal{X}$ and $\mathcal{Y}$. If $\text{card}|\mathcal{X}| \geqslant 4$, then

$$\mathbb{I}(X, Y) \leqslant \mathbb{V}(p_{XY}, p_X p_Y) \log \frac{|\mathcal{X}|}{\mathbb{V}(p_{XY}, p_X p_Y)}. \qquad \diamond$$

Another important metric that relates two probability distributions, without being an actual distance, is the Kullback-Leiber (KL) divergence.

Definition 9 Let X and X' be two discrete random variables defined on the same alphabet $\mathcal{X}$. The *Kullback-Leiber (KL) divergence* between X and X' is

$$\mathbb{D}(p_X \| p_{X'}) \triangleq \sum_{x \in \mathcal{X}} p_X(x) \log \frac{p_X(x)}{p_{X'}(x)},$$

if $\forall x \in \mathcal{X},\ p_{X'}(x) = 0 \Rightarrow p_X(x) = 0$, with the convention $0 \log 0 = 0$. $\diamond$

The mutual information can be expressed with the Kullback-Leiber divergence between specific distributions.

Lemma 2.13 Let $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$ be two discrete random variables with joint distribution p_{XY} and respective marginal distributions p_X and p_Y. Then

$$\mathbb{I}(X;Y) \equiv \mathbb{D}(p_{XY} \| p_X p_Y). \qquad \diamond$$

The KL divergence also relates to the variational distance through Pinsker's inequality [26, 78].

Proposition 2.14 (Pinsker) Let X and X' be two discrete random variables defined on the same alphabet $\mathcal{X}$. Then,

$$\mathbb{V}(p_X, p_{X'}) \leqslant \sqrt{2\ln(2)\mathbb{D}(p_X \| p_{X'})}. \qquad \diamond$$

Similar to the entropy, the KL divergence can also be extended to a Rényi divergence [84].

Definition 10 Let X and X' be two discrete random variables defined on the same alphabet $\mathcal{X}$. The *Rényi divergence* of order $\alpha \in \mathbb{R}_+^* \setminus \{1\}$, between X and X' is

$$\mathbb{D}_\alpha(p_X \| p_{X'}) \triangleq \frac{1}{\alpha - 1} \log \sum_{x \in \mathcal{X}} p_X(x)^\alpha q_X(x)^{1-\alpha},$$

with the conventions $0/0 = 0$ and $x/0 = \infty$ for $x \neq 0$. $\diamond$

Note that, when α goes to one, the Rényi divergence is strictly equivalent to the Kullback-Lever divergence. The reader is invited to refer to [30] for a complete list of properties for the Rényi divergence.

2.1.4 Typical Sequences

The notion of typical sequence is at the center of information theory proofs. Intuitively, a typical sequence is a sequence whose statistics are representative of the behavior of its associated random variable. These sequences have interesting properties presented in this subsection. Two notions of typicality exist: the strong typicality is the most intuitive, while the weak typicality is mostly an information theory concept.

2.1.4.1 Strongly Typical Sequences

Let $x^n \in \mathcal{X}^n$ be a sequence whose n elements are in a finite alphabet $\mathcal{X}$.

Notation $N(a; x^n)$ denotes the number of occurrences of a symbol $a \in \mathcal{X}$ in the sequence x^n, and $\{N(a; x^n)/n : a \in \mathcal{X}\}$ represents the *empirical distribution* of x^n.

Definition 11 (Strong typical set) Let p_X be a distribution on a finite alphabet $\mathcal{X}$ and let $\epsilon > 0$. A sequence $x^n \in \mathcal{X}^n$ is (strongly) ϵ-typical with respect to p_X if

$$\forall a \in \mathcal{X} \quad \left|\frac{1}{n}N(a; x^n) - p_X(a)\right| \leqslant \epsilon\, p_X(a)\,.$$

The set of all ϵ-typical sequences with respect to p_X is called the strong typical set and is denoted by $\mathcal{T}_\epsilon^n(X)$. ◊

This definition is quite intuitive since a typical sequence has an empirical distribution "close" to p_X. Typical sequences are particularly useful in information theory because of a result known as the *asymptotic equipartition property* (AEP for short).

Theorem 2.15 (AEP) Let p_X be a distribution on a finite alphabet $\mathcal{X}$ and let $0 < \epsilon < \min_{x\in\mathcal{X}} p_X(x)$. Let X^n be a sequence of independent and identically distributed (i.i.d.) random variables with distribution p_X. Then,

$$1 - \delta_\epsilon(n) \leqslant \mathbb{P}\big(X^n \in \mathcal{T}_\epsilon^n(X)\big) \leqslant 1$$

$$(1 - \delta_\epsilon(n))2^{n(\mathbb{H}(X)-\delta(\epsilon))} \leqslant \left|\mathcal{T}_\epsilon^n(X)\right| \leqslant 2^{n(\mathbb{H}(X)+\delta(\epsilon))}$$

$$\forall x^n \in \mathcal{T}_\epsilon^n(X) \quad 2^{-n(\mathbb{H}(X)+\delta(\epsilon)} \leqslant p_{X^n}(x^n) \leqslant 2^{-n(\mathbb{H}(X)-\delta(\epsilon))}.$$ ◊

The three inequalities given in this theorem can be translated as:

1. for n sufficiently large, the probability that a sequence is strongly typical is close to one;
2. the number of strongly typical sequences is close to $2^{-n\mathbb{H}(X)}$, making a direct connection with the entropy of a random variable;
3. strongly typical sequences are almost uniformly distributed.

Remark There exist explicit expressions for $\delta_\epsilon(n)$ and $\delta(\epsilon)$ [55], but the rough characterization used in Theorem 2.15 is sufficient for the subsequent analysis. The exact dependence on n

or ϵ is not needed because of some operational properties such as $\delta(\epsilon) \pm \delta(\epsilon) = \delta(\epsilon)$ and $\delta_\epsilon(n) \pm \delta_\epsilon(n) = \delta_\epsilon(n)$.

The notion of typicality generalizes to multiple random variables. Assume that $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$ is a pair of sequences with elements in finite alphabets $\mathcal{X}$ and $\mathcal{Y}$. The number of occurrences of a pair $(a, b) \in \mathcal{X} \times \mathcal{Y}$ in the pair of sequences (x^n, y^n) is denoted by $N(a, b; x^n, y^n)$.

Definition 12 (Jointly typical set) Let p_{XY} be a joint distribution on the finite alphabet $\mathcal{X} \times \mathcal{Y}$ and let $\epsilon > 0$. Sequences $x^n \in \mathcal{X}^n$ and $y^n \in \mathcal{Y}^n$ are ϵ-jointly typical with respect to p_{XY} if

$$\forall (a, b) \in \mathcal{X} \times \mathcal{Y} \quad \left| \frac{1}{n} N(a, b; x^n, y^n) - p_{XY}(a, b) \right| \leqslant \epsilon \, p_{XY}(a, b) \, .$$

The set of all ϵ-jointly typical sequences with respect to p_{XY} is called the jointly typical set and is denoted by $\mathcal{T}_\epsilon^n(XY)$. ◊

Since $\mathcal{T}_\epsilon^n(XY) \subseteq \mathcal{T}_\epsilon^n(X) \times \mathcal{T}_\epsilon^n(Y)$, if two sequences x^n and y^n are jointly typical, then they are also individually typical. The following corollary of Theorem 2.15 is a direct consequence of this property.

Corollary 2.16 (Joint AEP) Let p_{XY} be a joint distribution on the finite alphabets $\mathcal{X} \times \mathcal{Y}$ and let $0 < \epsilon < \min_{(x,y) \in \mathcal{X} \times \mathcal{Y}} p_{XY}(x, y)$. Let (X^n, Y^n) be a sequence of i.i.d. random variables with joint distribution p_{XY}. Then,

$$1 - \delta_\epsilon(n) \leqslant \mathbb{P}\big((X^n, Y^n) \in \mathcal{T}_\epsilon^n(XY)\big) \leqslant 1$$

$$(1 - \delta_\epsilon(n)) 2^{n(\mathbb{H}(XY) - \delta(\epsilon))} \leqslant \left| \mathcal{T}_\epsilon^n(XY) \right| \leqslant 2^{n(\mathbb{H}(XY) + \delta(\epsilon))}$$

$$\forall (x^n, y^n) \in \mathcal{T}_\epsilon^n(XY) \quad 2^{-n(\mathbb{H}(XY) + \delta(\epsilon))} \leqslant p_{X^n Y^n}(x^n, y^n) \leqslant 2^{-n(\mathbb{H}(XY) - \delta(\epsilon))}. \qquad \text{◊}$$

A conditional version of the AEP exists for conditional typical sets.

Definition 13 Let p_{XY} be a joint distribution on the finite alphabets $\mathcal{X} \times \mathcal{Y}$ and let $\epsilon > 0$. Let $x^n \in \mathcal{T}_\epsilon^n(X)$. The set

$$\mathcal{T}_\epsilon^n(XY|x^n) \triangleq \left\{ y^n \in \mathcal{Y}^n : (x^n, y^n) \in \mathcal{T}_\epsilon^n(XY) \right\}$$

called the *conditional typical set* with respect to x^n. ◊

Theorem 2.17 (Conditional AEP) Let p_{XY} be a joint distribution on the finite alphabets $\mathcal{X} \times \mathcal{Y}$ and suppose $0 < \epsilon' < \epsilon \leqslant \min_{(x,y)\in\mathcal{X}\times\mathcal{Y}} p_{XY}(x,y)$. Let $x^n \in \mathcal{T}^n_{\epsilon'}(X)$ and let $\tilde{Y}^n$ be a sequence of random variables such that

$$\forall y^n \in \mathcal{Y}^n \quad p_{\tilde{Y}^n}(y^n) = \prod_{i=1}^{n} p_{Y|X}(y_i|x_i)\,.$$

Then,

$$1 - \delta_{\epsilon\epsilon'}(n) \leqslant \mathbb{P}\big(\tilde{Y}^n \in \mathcal{T}^n_\epsilon(XY|x^n)\big) \leqslant 1$$

$$(1 - \delta_{\epsilon\epsilon'}(n))2^{n(\mathbb{H}(Y|X)-\delta(\epsilon))} \leqslant \left|\mathcal{T}^n_\epsilon(XY|x^n)\right| \leqslant 2^{n(\mathbb{H}(Y|X)+\delta(\epsilon))}$$

$$\forall y^n \in \mathcal{T}^n_\epsilon(XY|x^n) \quad 2^{-n(\mathbb{H}(Y|X)+\delta(\epsilon))} \leqslant p_{Y^n|X^n}(y^n|x^n) \leqslant 2^{-n(\mathbb{H}(Y|X)-\delta(\epsilon))}.$$ ◊

2.1.4.2 Weakly Typical Sequences

Even if the notion of strong typicality is intuitive, it does not apply to continuous random variables. A weaker definition exists to cope with continuous random variables [23] by defining a typical sequence as a sequence whose empirical entropy is close to the true entropy of the corresponding random variable. The discrete formulation also has practical operational advantages, which will be exploited in the subsequent chapters.

Definition 14 (Weakly typical set) Let p_X be a distribution on a finite alphabet $\mathcal{X}$ and let $\epsilon > 0$. A sequence $x^n \in \mathcal{X}^n$ is (weakly) ϵ-typical with respect to p_X if

$$\left|-\frac{1}{n}\log p_{X^n}(x^n) - \mathbb{H}(X)\right| \leqslant \epsilon.$$

The set of a all weakly ϵ-typical sequences with respect to p_X is called the weakly typical set and is denoted $\mathcal{A}^n_\epsilon(X)$. ◊

For weak typicality, the AEP is directly obtained from the weak law of large numbers.

Theorem 2.18 (AEP) Let p_X be a distribution on a finite alphabet $\mathcal{X}$ and let $\epsilon > 0$. Let X^n be a sequence of independent and identically distributed (i.i.d.) random variables with distribution p_X. Then,

- for n sufficiently large, $\mathbb{P}\big(X^n \in \mathscr{A}_\epsilon^n(X)\big) > 1-\epsilon$;
- if $x^n \in \mathscr{A}_\epsilon^n(X)$, then $2^{-n(\mathbb{H}(X)+\epsilon)} \leqslant p_{X^n}(x^n) \leqslant 2^{-n(\mathbb{H}(X)-\epsilon)}$;.
- for n sufficiently large, $(1-\epsilon)2^{n(\mathbb{H}(X)-\epsilon)} \leqslant \left|\mathscr{A}_\epsilon^n(X)\right| \leqslant 2^{n(\mathbb{H}(X)+\epsilon)}$. ◊

The notion of joint typicality follows in a similar way.

Definition 15 (Jointly weak typical set) Let p_{XY} be a joint distribution on the finite alphabets $\mathscr{X} \times \mathscr{Y}$ and let $\epsilon > 0$. Sequences $x^n \in \mathscr{X}^n$ and $y^n \in \mathscr{Y}^n$ are *jointly (weakly) ϵ-typical* with respect to p_{XY} if

$$\left|-\frac{1}{n}\log p_{X^nY^n}(x^n,y^n) - \mathbb{H}(XY)\right| \leqslant \epsilon$$
$$\left|-\frac{1}{n}\log p_{X^n}(x^n) - \mathbb{H}(X)\right| \leqslant \epsilon$$
$$\left|-\frac{1}{n}\log p_{Y^n}(y^n) - \mathbb{H}(Y)\right| \leqslant \epsilon$$

The set of a all jointly weakly ϵ -typical sequences with respect to p_{XY} is called the jointly weakly typical set and is denoted $\mathscr{A}_\epsilon^n(XY)$. ◊

Theorem 2.19 (joint AEP) Let p_{XY} be a joint distribution on the finite alphabets $\mathscr{X} \times \mathscr{Y}$ and let $\epsilon > 0$. Let (X^n, Y^n) be a sequence of i.i.d. random variables with joint distribution p_{XY}. Then,

- for n sufficiently large, $\mathbb{P}\big((X^n,Y^n) \in \mathscr{A}_\epsilon^n(XY)\big) > 1-\epsilon$;
- if $(x^n,y^n) \in \mathscr{A}_\epsilon^n(XY)$, then $2^{-n(\mathbb{H}(XY)+\epsilon)} \leqslant p_{X^nY^n}(x^n,y^n) \leqslant 2^{-n(\mathbb{H}(XY)-\epsilon)}$;.
- for n sufficiently large, $(1-\epsilon)2^{n(\mathbb{H}(XY)-\epsilon)} \leqslant \left|\mathscr{A}_\epsilon^n(XY)\right| \leqslant 2^{n(\mathbb{H}(XY)+\epsilon)}$. ◊

Even if there is no exact counterpart to the conditional AEP given in Corollary 2.17 for weakly typical sequences, the following result still holds.

Theorem 2.20 Let p_{XY} be a joint distribution on the finite alphabets $\mathscr{X} \times \mathscr{Y}$ and let $\epsilon > 0$. Let $\tilde{Y}^n$ be a sequence of i.i.d. random variables with distribution p_Y and let $\tilde{X}^n$ be an independent sequence of i.i.d. random variables with distribution p_X then,

$$\mathbb{P}\big((\tilde{X}^n, Y^n) \in \mathscr{A}_\epsilon^n(XY)\big) \leqslant 2^{-n(\mathbb{I}(X;Y)-\delta(\epsilon))}.$$ ◊

2.1.5 Markov Chains and Functional Dependence Graphs

For some information theoretic models and coding schemes, it may be rather difficult to describe the different relationships between random variables. A functional dependence graph is a convenient tool to analyze these possibly complicated relationships by graphically identifying the different Markov chains.

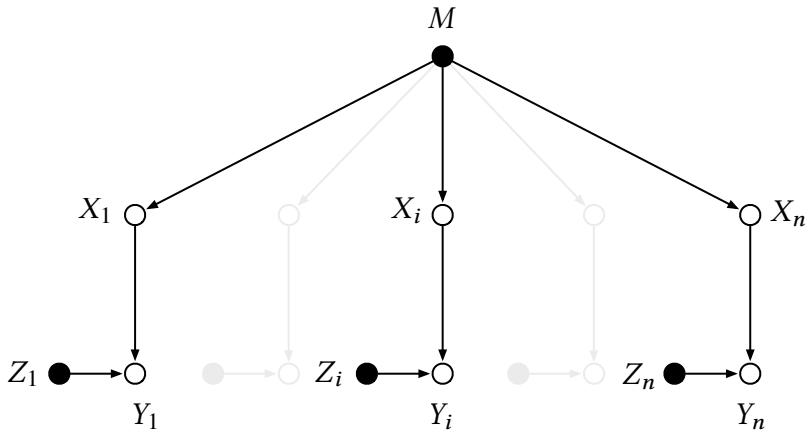

Figure 2.1 – *Functional dependence graph of Example 2.1.5. Independent random variables are indicated by* •*, while functions of these random variables are indicated by* ∘*.*

Definition 16 (Functional dependence graph) For m independent random variables and n functions of these variables, a functional dependence graph is a directed graph with $m+n$ vertices. Each edge between two nodes represents the existence of a mapping between two random variables. ◊

Example Let $M \in \mathcal{M}$ and $Z^n \in \mathbb{R}^n$ be two independent random variables, and $\{f_i\}_n$ be a set of functions from $\mathcal{M}$ to $\mathbb{R}^n$. For $i \in [\![1, n]\!]$, define the random variables $X_i = f_i(M)$ and $Y_i = X_i + Z_i$. The functional dependence graph of the random variables M, X_i^n, Y_i^n, and Z_i^n is illustrated in Figure 2.1.

Definition 17 (d-separation) Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be disjoint subsets of vertices in a functional dependence graph $\mathcal{G}$. The subset $\mathcal{Z}$ is said to *d-separate* $\mathcal{X}$ from $\mathcal{Y}$ if there exists no path between a vertex of $\mathcal{X}$ and a vertex of $\mathcal{Y}$ after the following operations have been performed:

- construct the subgraph $\mathcal{G}'$ consisting of all vertices in $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, as well as the edges and vertices encountered when moving backward starting from any of the vertices in $\mathcal{X}$, $\mathcal{Y}$, or $\mathcal{Z}$;
- in the subgraph $\mathcal{G}'$, delete all edges coming out of $\mathcal{Z}$;
- remove all arrows in $\mathcal{G}'$ to obtain an undirected graph. ◊

The usefulness of d-separation is justified by the following theorem.

Theorem 2.21 Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be disjoint subsets of the vertices in a functional dependence graph. If $\mathcal{Z}$ d-separates $\mathcal{X}$ from $\mathcal{Y}$, and if one collects the random variables in $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, in the random vectors X, Y, and Z, respectively, then $X \to Z \to Y$ forms a Markov chain. ◊

Theorem 2.21 is particularly useful in the converse proofs of channel coding theorems.

Example From the functional dependence graph of Figure 2.1, for any $i \neq j$ $X_i \to X_j \to Y_j$.

2.1.6 Information-Theoretic Bounds

Some additional inequalities are useful to bound the probabilities of rare events.

Lemma 2.22 (Markov's inequality) Let X be a non-negative real-valued random variable. Then,

$$\forall a > 0 \quad \mathbb{P}(X \geqslant a) \leqslant \frac{\mathbb{E}_X(X)}{a}. \qquad \diamond$$

This Lemma induces a result that is useful in the code selection step of achievability proofs relying on a random code generation.

Lemma 2.23 (Selection lemma) Let $X_n \in \mathcal{X}_n$ be a random variable and let $\mathcal{F}$ be a finite set of functions $f : \mathcal{X}_n \to \mathbb{R}^+$ such that $|\mathcal{F}|$ does not depend on n and

$$\forall f \in \mathcal{F} \quad \mathbb{E}_{X_n}(f(X_n)) \leqslant \delta(n).$$

Then, there exists a specific realization x_n of X_n such that

$$\forall f \in \mathcal{F} \quad f(x_n) \leqslant \delta(n). \qquad \diamond$$

Another useful consequences of Markov's is the Chernoff bound.

Lemma 2.24 (Chernoff bound) Let X be a real-valued random variable. For all $a > 0$,

$$\forall s > 0 \quad \mathbb{P}(X \geqslant a) \leqslant \mathbb{E}_X\big(e^{sX}\big)e,^{-sa}$$
$$\forall s < 0 \quad \mathbb{P}(X \leqslant a) \leqslant \mathbb{E}_X\big(e^{sX}\big)e.^{-sa} \qquad \diamond$$

2.2 Coding Primitives

The previous section provides the necessary tools to analyze many communication systems from an information-theoretic perspective. This section introduces four fundamental primitives: channel capacity, source coding with side information, channel intrinsic randomness, and channel resolvability. The analysis of multi-user communications relies on several of these primitives, which introduces the pivotal proof mechanisms used throughout this dissertation. The example of point-to-point communications illustrates the fundamental aspects of these coding primitives.

2.2.1 Channel Capacity

Introducing security constraints in the communication model does not mean ignoring the problem of reliability in communications. The goal is indeed twofold: providing reliable communications for the legitimate users, and ensuring security with respect to an external eavesdropper. The problem of reliable point-to-point communications has been introduced by Shannon in his seminal work [88] and was studied extensively for other channels (esp. the multiple-access channel).

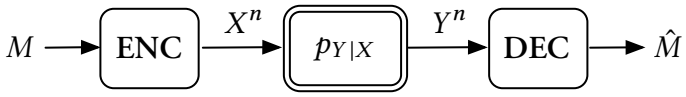

Figure 2.2 – *Point-to-point communications.*

The channel capacity for the problem of point-to-point communications depicted in Figure 2.2 is defined as the maximum bit rate a user can reliably transmit on a noisy channel.

Theorem 2.25 (Channel Coding Theorem) The capacity of a discrete memoryless channel (DMC) $\left(\mathcal{X}, p_{Y|X}, \mathcal{Y}\right)$ is

$$C = \max_{p_X} \mathbb{I}(X;Y).$$

In other words, if $R < C$ then R is an achievable transmission rate and an achievable transmission rate must satisfy $R \leqslant C$. ◊

Theorem 2.25 proves that the channel capacity is equal to the maximum over all possible input distributions of the mutual information between the channel input and the channel output.

The proof of this theorem relies on a *random coding argument*, which consists in drawing the symbols of codewords independently at random. Showing that there exists a code such that the probability of error goes to zero as n (the code blocklength) goes to infinity yields an upper bound on the transmission rate. This random code is only a technical tool and not a practical code that can be used as an actual coding scheme. The mechanisms presented in the following achievability proof are essential to tackle more complicated models.

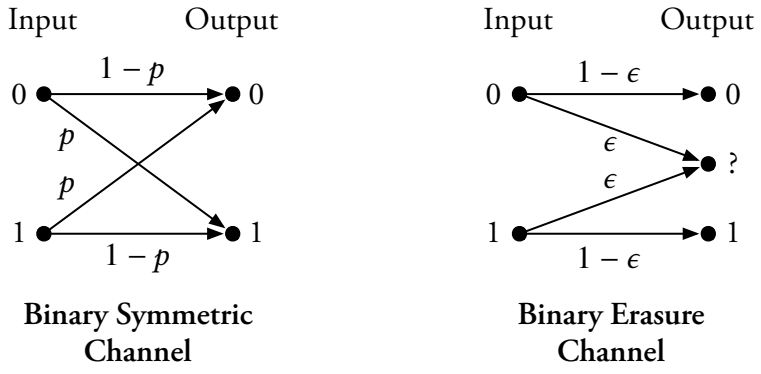

Figure 2.3 – *Binary symmetric channel and binary erasure channel.*

Example Figure 2.3 depicts two binary channels that are particularly useful in information theory. The binary symmetric channel BSC(p) flips bit values with a crossover probability p and has a capacity of $1 - \mathbb{H}_\mathrm{b}(p)$ bits. The binary erasure channel BEC(ϵ) erases bits a with a probability ϵ and has a capacity of $1 - \epsilon$ bits.

The additive white Gaussian noise (AWGN) channel also has a prominent role in information and communication theory to describe a practical communication scheme. The AWGN channel captures, in particular, the impact of thermal noise and interferences in wireless communications. The channel output at each instant $i \geqslant 1$ is $Y_i = X_i + N_i$, where X_i denotes the transmitted symbol and $\{N_i\}_{i\geqslant 1}$ are i.i.d. random variables with distribution $\mathcal{N}(0, \sigma^2)$. Without further restriction the capacity of the AWGN channel is infinite; however, by adding an average power constraint in the form of

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\big(X_i^2\big) \leqslant P,$$

, the capacity becomes finite.

Theorem 2.26 The capacity of a Gaussian channel is given by

$$C = \frac{1}{2}\log\left(1 + \frac{P}{\sigma^2}\right),$$

where P denotes the power constraint and σ^2 is the variance of the noise. ◊

2.2.2 Source Coding with Side Information

Along with channel coding, the problem of source compression is at the core of information theory. Information-theoretic tools are used to derive fundamental limits regarding the bit rate at which a source can be compressed without any loss. For instance, for a source $(\mathcal{X}, p_X)$, the minimum number of bits that must be stored or transmitted is $n\mathbb{H}(X)$, where n is the length of the source sequence.

The fundamental paper by Slepian and Wolf [89] considers the separate encoding of correlated sources as illustrated in Figure 2.4.

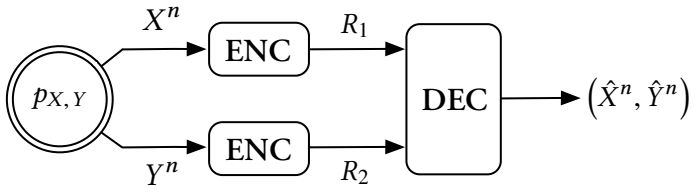

Figure 2.4 – *Separate encoding of correlated sources* (Slepian-Wolf problem).

Consider a DMS $(\mathcal{X}\mathcal{Y}, p_{XY})$ with two outputs X and Y with joint distribution p_{XY}. As depicted in Figure 2.4, the outputs are processed by two different encoders that compress X^n into a message of rate R_1 and compresses Y^n into a message of rate R_2. Both these messages are processed jointly by a single decoder, whose goal is to estimate X^n and Y^n. Since this model considers two different encoders one can expect that the best thing to do is to encode X at a rate $R_1 > \mathbb{H}(X)$ and Y at a rate $R_2 > \mathbb{H}(Y)$. This procedure guarantees that the probability of error vanishes as n goes to infinity but exploits neither the correlation of the source nor the common decoder. It turns out that it suffices to ensure that the sum rate $R_1 + R_2$ is greater than $\mathbb{H}(XY)$, which is in general smaller than $\mathbb{H}(X) + \mathbb{H}(Y)$. To achieve such a surprising result the encoders and the decoder must be properly designed [16, 23, 89].

Definition 18 A $(2^{kR_1}, 2^{kR_2}, k)$ *distributed source code* $\mathscr{C}_k$ for the discrete memoryless source $(\mathcal{X}\mathcal{Y}, p_{XY})$ consists of

- two message sets $\mathcal{M}_1 = [\![1, 2^{R_1}]\!]$ and $\mathcal{M}_2 = [\![1, 2^{nR_2}]\!]$;
- an encoding function $f_1 : \mathcal{X}^k \to \mathcal{M}_1,\ x^k \mapsto m_1$;
- an encoding function $f_2 : \mathcal{Y}^k \to \mathcal{M}_2,\ y^k \mapsto m_2$;
- a decoding function $g : \mathcal{M}_1 \times \mathcal{M}_2 \to (\mathcal{X}^k \times \mathcal{Y}^k) \cup \{?\},\ (m_1, m_2) \mapsto (\hat{x}^k, \hat{y}^k)$, where ? represents the error symbol. ◊

The performance of a code $\mathscr{C}_k$ is measured in terms of the average probability of error

$$\mathbf{P}_{\mathrm{e}}(\mathscr{C}_k) \triangleq \mathbb{P}\Big((\hat{X}^k, \hat{Y}^k) \neq (X^k, Y^k) \Big| \mathscr{C}_k\Big).$$

Definition 19 A rate pair (R_1, R_2) is *achievable* if there exists a sequence of $(2^{kR_1}, 2^{kR_2}, k)$ codes $\{\mathscr{C}_k\}_{k \geqslant 1}$ such that

$$\lim_{k \to \infty} \mathbf{P}_{\mathrm{e}}(\mathscr{C}_k) = 0.$$

The achievable rate region is defined as

$$\mathscr{R}^{sw} \triangleq \mathrm{cl}(\{(R_1, R_2) : (R_1, R_2) \text{ is achievable}\}). \qquad \text{◊}$$

The achievable rate region with separate encoding was first characterized by Slepian and Wolf. The region is often called the Slepian-Wolf region and codes for the distributed source coding problem are often referred to as Slepian-Wolf codes.

Theorem 2.27 (Slepian-Wolf Theorem) The achievable rate region with separate encoding for a source $(\mathcal{U}\mathcal{V}, p_{UV})$ is

$$\mathscr{R}^{\text{SW}} \triangleq \left\{ (R_1, R_2) : \begin{array}{r} R_1 \geqslant \mathbb{H}(X|Y), \\ R_2 \geqslant \mathbb{H}(Y|X), \\ R_1 + R_2 \geqslant \mathbb{H}(XY) \end{array} \right\} \qquad \diamond$$

Figure 2.5 illustrates the typical shape of the Slepian-Wolf region $\mathscr{R}^{\text{SW}}$.

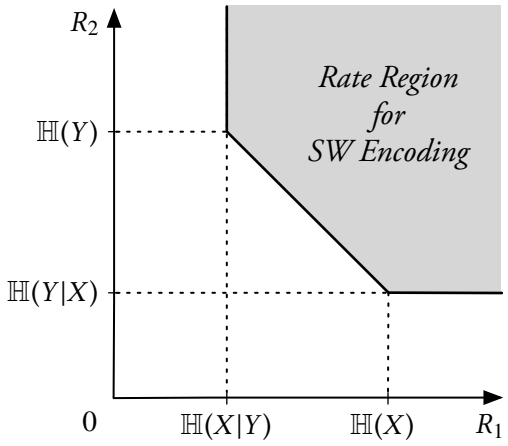

Figure 2.5 – *The Slepian-Wolf region for a discrete memoryless source* $(\mathcal{X}\mathcal{Y}, p_{XY})$.

Proof (Heuristic) The underlying principle of the proof is called random binning. It consists in labeling the typical sequences in such a way that the decoder can use the labels to select the right sequence with high probability. Exploiting the properties of the joint typicality is a solution to achieve a better compression rate. Let $\epsilon > 0$ and $k \in \mathbb{N}^*$. Let $R_1 > 0$ and $R_2 > 0$ be rates to be specified later. The $(2^{kR_1}, 2^{kR_2}, k)$ code $\mathscr{C}_k$ is constructed as follows.

- *Binning:* For each sequence $x^k \in \mathcal{T}_\epsilon^k(X)$, draw an index uniformly at random in the set $[\![1, 2^{kR_1}]\!]$. For each sequence $y^k \in \mathcal{T}_\epsilon^k(Y)$, draw an index uniformly at random in the set $[\![1, 2^{kR_2}]\!]$. The index assignments define the encoding functions

$$f_1 : \mathcal{X}^k \to [\![1, 2^{kR_1}]\!] \quad \text{and} \quad f_2 : \mathcal{Y}^k \to [\![1, 2^{kR_2}]\!],$$

  which are revealed to all parties.

- *Encoder 1:* given the observation x^k, if $x^k \in \mathcal{T}_\epsilon^k(X)$, output $m_1 = f_1(x^k)$; otherwise output $m_1 = 1$.

- *Encoder 2:* given the observation y^k, if $y^k \in \mathcal{T}_\epsilon^k(Y)$, output $m_2 = f_2(y^k)$; otherwise output $m_2 = 1$.

- *Decoder:* given messages m_1 and m_2, output $\hat{x}^k$ and $\hat{y}^k$ if they are the unique sequences such that $(\hat{x}^k, \hat{y}^k) \in \mathcal{T}_\epsilon^k(XY)$ and $f_1(\hat{x}^k) = m_1$, $f_2(\hat{y}^k) = m_2$; otherwise, output ?.

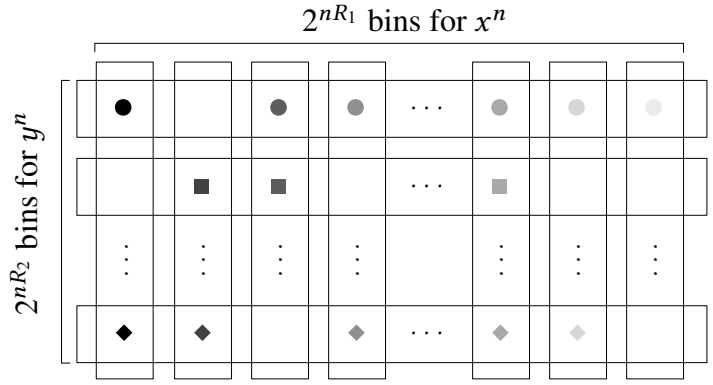

Figure 2.6 – *Binning procedure for the Slepian-Wolf coding. Each pictogram represents one of the* $2^{n\mathbb{H}(XY)}$ *jointly typical pairs* (x^n, y^n).

Figure 2.6 represents how the binning procedure work. The bins are designed in such a way that each of the $2^{n\mathbb{H}(XY)}$ jointly typical pairs (x^n, y^n) receives a unique pair of indices. The reader is invited to refer to [16, 23] for the detailed technical proof. ∎

One of the special cases of the Slepian-Wolf problem consists in supposing that one the components of the DMS $(\mathcal{X}\mathcal{Y}, p_{XY})$, say Y, is directly available at the decoder as side information and only X should be compressed. This problem is known as source coding with side information.

Corollary 2.28 (Source coding with side information) Consider a DMS $(\mathcal{X}\mathcal{Y}, p_{XY})$ and assume that $(\mathcal{X}, p_X)$ should be compressed knowing that $(\mathcal{Y}, p_Y)$ is available as side information at the decoder. Then,

$$\inf\{R : R \text{ is an achievable compression rate}\} = \mathbb{H}(X|Y)\,. \qquad \diamond$$

This result confirms that a source with low entropy can be compressed more than a source with higher entropy. Corollary 2.28 plays a pivotal role in physical layer security and, in particular, for secret-key generation. To distill a secret-key from the physical medium, two users must observe a correlated source of randomness. This correlated source of randomness does not provide identical observations, thus requiring a reconciliation procedure. To agree on a common key, the two users exchange messages that allow them to agree on a common sequence, which is then used to distill a key. Source coding with side information generalizes the problem of source coding when additional side information is available either to both users or only one.

2.2.3 Channel Intrinsic Randomness

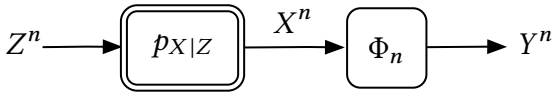

Figure 2.7 – *Channel Intrinsic Randomness (CIR) basic scheme.*

The Channel Intrinsic Randomness (CIR for short) is defined in [13] as the "maximum bit rates that can be extracted from a channel output *independently* of an input with known statistics." This problem is of primary importance to analyze secret-key generation schemes to provide strong security.

Formally, consider a discrete memoryless channel $(\mathcal{Z}, \mathcal{X}, p_{X|Z})$, and suppose that the distribution controlling the channel is p_Z. Channel intrinsic randomness consists in designing a mapping ϕ_n such that, for any $\epsilon > 0$

$$\mathbb{V}\big(p_{Z^n}\phi_n(X^n), p_{Z^n}q_{Y^n}\big) \leqslant \epsilon,$$

where q_{Y^n} represents a desired target distribution. For an arbitrary distribution q_{Y^n} is called *channel number generation*.

The following proposition yields a condition on the process $\{Y^n\}_{n\geqslant 1}$ to ensure the existence of such a mapping ϕ_n.

Proposition 2.29 (Achievability of CIR) If $\mathbb{H}(X|Z) > \mathbb{H}(Y)$, then

$$\forall 0 < \epsilon < 0.1,\ \exists \phi_n : \mathcal{X}^n \to \mathcal{Y}^n,\ \mathbb{V}\big(p_{Z^n}\phi_n(X^n), p_{Z^n}q_{Y^n}\big) \leqslant \epsilon, \tag{2.6}$$

and

$$\forall 0 < \epsilon < 0.1,\ \exists \phi'_n : \mathcal{X}^n \to \mathcal{Y}^n,\ \mathbb{D}\left(p_{Z^n}\phi_n(X^n)\|p_{Z^n}q_{Y^n}\right) \leqslant \epsilon. \tag{2.7}$$

◊

In [13], the author presents a result that holds for sources with memory, but all the analysis done in the subsequent chapters will only focus on channels without memory.

Proof This results can be proven as a corollary of Corollary 2.31 and Pinsker's inequality. ∎

2.2.4 Channel Resolvability

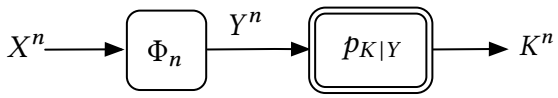

Figure 2.8 – *Channel Resolvability basic scheme.*

The paradigm behind the wiretap coding consists in ensuring that the observation of the eavesdropper is independent of the secret messages. A sufficient way to guarantee security consists in ensuring that the eavesdropper's statistics are not affected by the messages exchanged by the legitimate parties. In that case, the eavesdropper cannot resolve what message was transmitted since it doesn't observe significant variations on its end. This approach called *channel resolvability* ensures strongly secure communications, while capacity based approaches only provide

weak secrecy. Channel resolvability [39] aims at simulating a target distribution at the output of a channel by controlling its input with a uniform random number.

Formally, consider a discrete memoryless channel $(\mathcal{Y}, \mathcal{K}, p_{K|Y})$, and suppose that the distribution controlling the channel is p_K. Channel intrinsic randomness consists in designing a mapping ϕ_n such that, for any $\epsilon > 0$

$$\mathbb{V}\big(p_{Z^n \phi_n(X^n)}, p_{Z^n} p_{K^n}\big) \leqslant \epsilon. \tag{2.8}$$

The same result holds for the KL divergence.

2.3 Joint Analysis of Channel Intrinsic Randomness and Resolvability

This section introduces a scheme that includes and extends the problems of channel resolvability and channel intrinsic randomness. The model consists of a discrete memoryless source that is sent through a first discrete memoryless channel, which output is then processed to feed the input of a second channel. The objective is to simulate a random process with fixed distribution at the second channel output independently from the first channel input. A joint scheme is introduced to find joint exponents and asymptotic limits that can be specialized to the well-known schemes previously mentioned. This scheme also provides, for instance, a direct extension of channel intrinsic randomness with a non-uniform target distribution.

The joint approach does not improve asymptotic limits since separation holds in some cases, but the joint exponents are larger than the tandem exponents that would be obtained by using an intermediate uniform random number. Even if the model is more general, the subsequent results are connected to some of the separation approaches investigated in [96, 97] and to channel resolvability with non-uniform input [14, 42].

2.3.1 Definitions and Assumptions

Consider the setting illustrated in Figure 2.9 consisting of:

- a discrete memoryless source (DMS) $(\mathcal{Z}, q_Z)$ that outputs i.i.d. sequences $Z^n \in \mathcal{Z}^n$;
- a discrete memoryless channel (DMC) $(\mathcal{Z}, W_1, \mathcal{X})$ with transition probability W_1, which outputs i.i.d sequences $X^n \in \mathcal{X}^n$ when Z^n is present at the input;

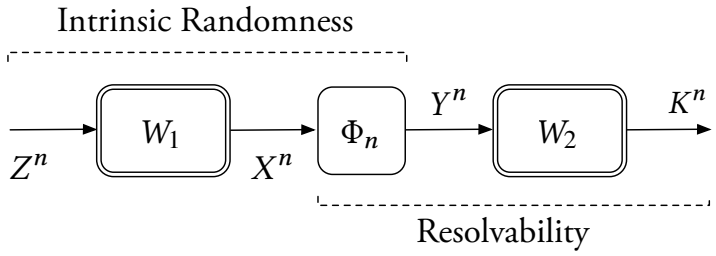

Figure 2.9 – *Joint channel intrinsic randomness and resolvability.*

- an encoding function $\phi_n : \mathcal{X}^n \to \mathcal{Y}^n$, $X^n \mapsto Y^n$;
- a second DMC $(\mathcal{Y}, W_2, \mathcal{K})$ with transition probability W_2, which outputs $K^n \in \mathcal{X}^n$ when Y^n is present at the input.

The problem consists in simulating a target i.i.d distribution q_{K^n} at the output of the channel W_2 independently of Z^n. The target probability q_{K^n} is defined via a fixed i.i.d distribution q_{Y^n} as

$$q_{K^n}(k^n) = \sum_{y^n \in \mathcal{Y}^n} W_2(k^n|y^n) q_{Y^n}(y^n).$$

In what follows, $\tilde{Y}$ and $\tilde{K}$ denote the target processes with distribution q_Y and q_K, respectively.

Mathematically, the constraints are formalized as follows.

- The simulated distribution p_{K^n} and the target distribution q_{K^n} should be asymptotically close in terms of KL divergence:

$$\lim_{n\to\infty} \mathbb{D}(p_{K^n} \| q_{K^n}) = 0.$$

- The simulated process K^n and the input process Z^n should be asymptotically statistically independent:

$$\lim_{n\to\infty} \mathbb{D}(p_{K^n Z^n} \| p_{K^n} q_{Z^n}) = 0.$$

These two conditions can be merged by requiring that

$$\lim_{n\to\infty} \mathbb{D}(p_{K^n Z^n} \| q_{K^n} q_{Z^n}) = 0.$$

For particular choices of DMSs and DMCs in the scheme mentioned above:

- if $\mathcal{Y} = \mathcal{K}$, $W_2 = id_{\mathcal{K}}$, and $q_{K^n} \sim \mathcal{U}[\![1, 2^{nR}]\!]$, it corresponds to *channel intrinsic randomness*;

- if $W_1 = q_X$, $q_{X^n} \sim \mathcal{U}[\![1, 2^{nR}]\!]$, it corresponds to *channel resolvability*.

2.3.2 Achievability and Exponents

2.3.2.1 Joint Exponent Derivation

The joint exponent derivation derives from a proof technique introduced by Hayashi [40, 41] (see also [48]). First, the joint coding scheme is constructed at random; the encoding function ϕ_n is randomly defined by mapping every sequence $x^n \in \mathcal{X}^n$ to a sequence $y^n \in \mathcal{Y}^n$ drawn according to q_{Y^n}. The corresponding random variable is denoted Φ_n.

For $0 < \alpha < 1$,

$$\begin{aligned}\mathbb{D}(p_{K^nZ^n}\|q_{K^n}q_{Z^n}) &\overset{(a)}{\leqslant} \sum_{z^n\in\mathcal{Z}^n} q_{Z^n}(z^n)\mathbb{D}(p_{K^n|Z^n=z^n}\|q_{K^n}) \\ &\overset{(b)}{\leqslant} \sum_{z^n\in\mathcal{Z}^n} q_{Z^n}(z^n)\mathbb{D}_{1+\alpha}(p_{K^n|Z^n=z^n}\|q_{K^n}),\end{aligned} \tag{2.9}$$

Inequality (a) follows from the law of total probability, and (b) holds because the Rényi divergence is increasing with respect to α (see for instance [29]).

By taking the expectation of inequality (2.9), over all encoding function Φ_n and with Jensen's inequality,

$$\begin{aligned}&\mathbb{E}_{\Phi_n}(\mathbb{D}(p_{K^nZ^n}\|q_{K^n}q_{Z^n})) \\ &\qquad\leqslant \frac{1}{\alpha}\sum_{z^n\in\mathcal{Z}^n} q_{Z^n}(z^n) \times \log\left(\sum_{k^n\in\mathcal{K}^n} q_{K^n}(k^n)^{-\alpha}\mathbb{E}_{\Phi_n}\left(p_{K^n|Z^n}(k^n|z^n)^{1+\alpha}\right)\right).\end{aligned} \tag{2.10}$$

By the law of total probability and observing that $Z^n \to X^n \to Y^n \to K^n$ forms a Markov chain and for a given realization ϕ_n of Φ_n,

$$p_{K^n|Z^n}(k^n|z^n) = \sum_{y^n\in\mathcal{Y}^n} W_2(k^n|y^n) \sum_{x^n\in\mathcal{X}^n} W_1(x^n|z^n)\mathbb{1}\left\{\phi_n(x^n) = y^n\right\}. \tag{2.11}$$

Therefore,

$$\mathbb{E}_{\Phi_n}\left(p_{K^n|Z^n}(k^n|z^n)^{1+\alpha}\right) = \sum_{x^n\in\mathcal{X}^n}\sum_{y^n\in\mathcal{Y}^n}\sum_{\phi_n(x^n)\in\mathcal{Y}^n} q_{Y^n}(\phi_n(x^n))W_1(x^n|z^n)W_2(k^n|y^n)\mathbb{1}\left\{\phi_n(x^n)=y^n\right\} \times \mathbb{E}^{x^n}_{\Phi_n}\left(\sum_{\tilde{x}^n,\tilde{y}^n} W_1(\tilde{x}^n|z^n)W_2(k^n|\tilde{y}^n)\mathbb{1}\left\{\phi_n(\tilde{x}^n)=\tilde{y}^n\right\}\right)^{\alpha}, \quad (2.12)$$

where $\mathbb{E}^{x^n}_{\Phi_n}$ is the expectation over all possible mapping Φ_n where the value $\phi_n(x^n)$ is fixed. This expectation can be upper bounded as follows

$$\begin{aligned}
&\mathbb{E}^{x^n}_{\Phi_n}\left(\sum_{\tilde{x}^n,\tilde{y}^n} W_1(\tilde{x}^n|z^n)W_2(k^n|\tilde{y}^n)\mathbb{1}\left\{\phi_n(\tilde{x}^n)=\tilde{y}^n\right\}\right)^{\alpha}\\
&\quad\overset{(a)}{\leqslant} \mathbb{E}^{x^n}_{\Phi_n}\left(W_1(x^n|z^n)W_2(k^n|y^n)\mathbb{1}\left\{\phi_n(x^n)=y^n\right\}\right)^{\alpha}\\
&\qquad + \mathbb{E}^{x^n}_{\Phi_n}\left(\sum_{\tilde{x}^n\neq x^n}\sum_{\tilde{y}^n\neq y^n} W_1(\tilde{x}^n|z^n)W_2(k^n|\tilde{y}^n)\mathbb{1}\left\{\phi_n(\tilde{x}^n)=\tilde{y}^n\right\}\right)^{\alpha}\\
&\quad\overset{(b)}{\leqslant} \left(W_1(x^n|z^n)W_2(k^n|y^n)\right)^{\alpha}\\
&\qquad + \left(\sum_{\tilde{x}^n\neq x^n}\sum_{\tilde{y}^n\neq y^n} \mathbb{E}^{x^n}_{\Phi_n}\left(W_1(\tilde{x}^n|z^n)W_2(k^n|\tilde{y}^n)\mathbb{1}\left\{\phi_n(\tilde{x}^n)=\tilde{y}^n\right\}\right)\right)^{\alpha}\\
&\quad\overset{(c)}{\leqslant} \left(W_1(x^n|z^n)W_2(k^n|y^n)\right)^{\alpha}\\
&\qquad + \left(\underbrace{\sum_{\tilde{x}^n\neq x^n} W_1(\tilde{x}^n|z^n)}_{=1}\underbrace{\sum_{\tilde{y}^n\neq y^n} W_2(k^n|\tilde{y}^n)q_{Y^n}(\tilde{y}^n)}_{q_{K^n}(k^n)}\right)^{\alpha}\\
&\quad\leqslant \left(W_1(x^n|z^n)W_2(k^n|y^n)\right)^{\alpha} + q_{K^n}(k^n)^{\alpha}. \qquad (2.13)
\end{aligned}$$

Inequality (a) is obtained by splitting the sum and because $x \mapsto x^{\alpha}$ is concave for $0 < \alpha < 1$; inequality (b) because of Jensen's inequality and $\mathbb{1}\left\{\phi_n(x^n)=y^n\right\} \leqslant 1$; and inequality (c) because $\mathbb{E}^{x^n}_{\Phi_n}(\mathbb{1}\left\{\phi_n(\tilde{x}^n)=\tilde{y}^n\right\}) = q_{Y^n}(\tilde{y}^n)$ for $\tilde{x}^n \neq x^n$.

Plugging (2.13) into (2.12), yields after some calculations,

$$\mathbb{E}_{\Phi_n}\left(p_{K^n|Z^n}(k^n|z^n)^{1+\alpha}\right) \leqslant q_{K^n}(k^n)^{1+\alpha}\left(1+\sum_{x^n\in\mathscr{X}^n} W_1(x^n|z^n)^{1+\alpha}\sum_{y^n\in\mathscr{Y}^n} q_{Y^n}(y^n)\left(\frac{W_2(k^n|y^n)}{q_{K^n}(k^n)}\right)^{1+\alpha}\right). \quad (2.14)$$

$$\begin{aligned}
&\mathbb{E}_{\Phi_n}(\mathbb{D}(p_{K^nZ^n}\|q_{K^n}q_{Z^n}))\\
&\leqslant \frac{1}{\alpha}\sum_{z^n\in\mathscr{Z}^n} q_{Z^n}(z^n)\\
&\quad\times\log\left(\sum_{k^n\in\mathscr{K}^n} q_{K^n}(k^n)\left(1+\sum_{x^n\in\mathscr{X}^n} W_1(x^n|z^n)^{1+\alpha}\sum_{y^n\in\mathscr{Y}^n} q_{Y^n}(y^n)\left(\frac{W_2(k^n|y^n)}{q_{K^n}(k^n)}\right)^{1+\alpha}\right)\right)\\
&\leqslant \frac{1}{\alpha}\log\left(1+\sum_{z^n\in\mathscr{Z}^n}\sum_{k^n\in\mathscr{K}^n} q_{Z^n}(z^n)q_{K^n}(k^n)\sum_{x^n\in\mathscr{X}^n} W_1(x^n|z^n)^{1+\alpha}\sum_{y^n\in\mathscr{Y}^n} q_{Y^n}(y^n)\left(\frac{W_2(k^n|y^n)}{q_{K^n}(k^n)}\right)^{1+\alpha}\right)\\
&\leqslant \frac{1}{\alpha}\sum_{k^n\in\mathscr{K}^n}\sum_{y^n\in\mathscr{Y}^n} W_2(k^n|y^n)^{1+\alpha}q_{Y^n}(y^n)q_{K^n}(k^n)^{-\alpha}\sum_{z^n\in\mathscr{Z}^n}\sum_{x^n\in\mathscr{X}^n} W_1(x^n|z^n)^{1+\alpha}q_{Z^n}(z^n) \quad (2.15)
\end{aligned}$$

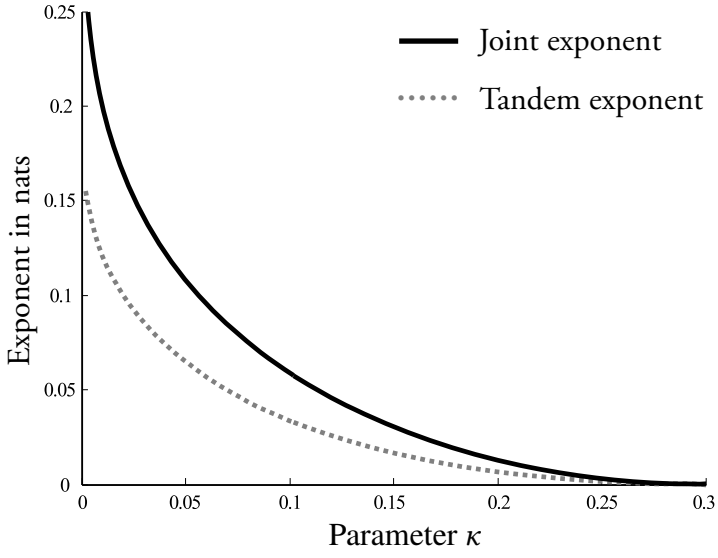

Figure 2.10 – *Comparison of joint and tandem exponents for different values of* κ *when* $\omega = 0.3$*, for channel intrinsic randomness and source resolvability.*

Finally, equation (2.15) brings the following proposition

Proposition 2.30 There exists a mapping $\phi_n : \mathcal{X}^n \to \mathcal{Y}^n$, such that

$$\mathbb{D}(p_{K^nZ^n} \| q_{K^n} q_{Z^n}) \leqslant e^{-nE_j(q_Z, q_K, W_1, W_2)},$$

with

$$E_j(q_Z, q_K, W_1, W_2) = \max_{\alpha \in [0,1]} E_c(\alpha, q_Z, W_1) - E_r(\alpha, q_Y, W_2), \tag{2.16}$$

and

$$\begin{aligned} E_c(\alpha, q_Z, W_1) &= -\log\left(\sum_{z \in \mathcal{Z}} q_Z(z) \sum_{x \in \mathcal{X}} W_1(x|z)^{1+\alpha}\right) \\ E_r(\alpha, q_Y, W_2) &= \log \sum_{k \in \mathcal{K}} \sum_{y \in \mathcal{Y}} W_2(k|y)^{1+\alpha} q_Y(y) q_K(k)^{-\alpha}. \end{aligned} \tag{2.17}$$

◊

Remark Using the same upper bound as in [41],

$$\begin{aligned} E_r(\alpha, q_Y, W_2) &\leqslant \log\left(\sum_{k^n \in \mathcal{K}^n} \left(\sum_{y^n \in \mathcal{Y}^n} q_{Y^n}(y^n) W_2(k^n|y^n)^{\frac{1}{1-\alpha}}\right)^{1-\alpha}\right) \\ &\triangleq E_r'(\alpha, q_Y, W_2). \end{aligned} \tag{2.18}$$

Corollary 2.31 If $\mathbb{H}(X|Z) \geqslant \mathbb{I}(\tilde{Y}; \tilde{K})$, then there exists a mapping $\phi_n : \mathcal{X} \to \mathcal{Y}$, such that

$$\lim_{n \to \infty} \mathbb{D}(p_{K^nZ^n} \| q_{K^n} q_{Z^n}) = 0.$$

◊

Proof The key behind the asymptotic results consists in observing that $\mathbb{E}_{\Phi_n}(\mathbb{D}(p_{K^nZ^n} \| q_{K^n} q_{Z^n}))$ goes to 0 as n goes to infinity if $E_j(q_Z, q_K, W_1, W_2) > 0$.

First notice that $E_j(0, q_Z, q_K, W_1, W_2) = 0$, meaning that $E_j(q_Z, q_K, W_1, W_2) \geqslant 0$. It is therefore sufficient to show that the derivative of $s \mapsto E_j(0, q_Z, q_K, W_1, W_2)$ in 0 is positive to ensure $E_j(q_Z, q_K, W_1, W_2) > 0$ since $E_j(\alpha, q_Z, q_K, W_1, W_2)$ is non-negative, non-decreasing, convex in α. Note that

$$\left.\frac{\partial E_c(\alpha, q_Z, W_1)}{\partial \alpha}\right|_{\alpha=0} = \mathbb{H}(X|Z). \tag{2.19}$$

In addition, $-E_r'(\alpha, q_Y, W_2)$ is non-negative, non-decreasing, convex in α, and

$$-\left.\frac{\partial E_r'(\alpha, q_Y, W_2)}{\partial \alpha}\right|_{\alpha=0} = -\mathbb{I}(\tilde{Y}; \tilde{K}). \tag{2.20}$$

Therefore, the joint exponent is positive if

$$\mathbb{H}(X|Z) - \mathbb{I}(\tilde{Y};\tilde{K}) > 0.$$

■

2.3.2.2 Comparison with Tandem Exponents

Figure 2.10 illustrates the gain of joint channel intrinsic randomness and source resolvability. The problem of simulating a random process with a target distribution q_{K^n} from the output of the channel W_1 independently from its input corresponds to having $\mathcal{Y}^n = \mathcal{K}^n$, $W_2 = id_{\mathcal{K}^n}$. For $q_{Z^n} = q_Z^n$ with $q_Z \sim \mathcal{B}(1/2)$, $W_1 = \mathrm{BSC}(\omega)$, and $q_{K^n} = q_K^n$ with $q_K \sim \mathcal{B}(\kappa)$ the joint exponent becomes

$$E_j^{\mathrm{cir}}(\kappa,\omega) = \max_{\alpha\in[0,1]} E_j^{\mathrm{cir}}(\alpha,\kappa,\omega), \tag{2.21}$$

where

$$E_j^{\mathrm{cir}}(\alpha,\kappa,\omega) = -\log\left(\omega^{1+\alpha} + (1-\omega)^{1+\alpha}\right) - \log\left(\kappa^{1-\alpha} + (1-\kappa)^{1-\alpha}\right). \tag{2.22}$$

The separate approach consists in achieving the desired result in two distinct and independent steps:

1. The DMC W_1 is first used to extract an intermediate uniform random variable that takes values in $[\![1, 2^{nR}]\!]$ (channel intrinsic randomness);

2. The uniform random variable is then used to generate the desired target process q_{K^n} (source resolvability).

The exponents for these steps are derived as special cases of the joint exponent.

- For channel intrinsic randomness, the exponent to optimize is

$$E_1(\alpha, q_Z, W_1, R) = -\log\left(\sum_{z\in\mathcal{Z}} q_Z(z) \sum_{x\in\mathcal{X}} W_1(x|z)^{1+\alpha}\right) - \alpha R, \tag{2.23}$$

  which consists in taking $E_r(\alpha, 2^{-nR}, id_{\mathcal{K}}) = \alpha R$ in the joint exponent.

- Similarly, for source resolvability, the exponent to optimize is

$$E_2(\alpha, q_K, R) = \alpha R - \log \sum_{k^n\in\mathcal{K}^n} q_{K^n}(k^n)^{1-\alpha}. \tag{2.24}$$

Although there is no claim of optimality behind this result, note that these exponents are close to the best exponents obtained by Hayashi [41] for variational distance. A slight improvement of the exponents can be obtained in some cases, as shown by Watanabe [102].

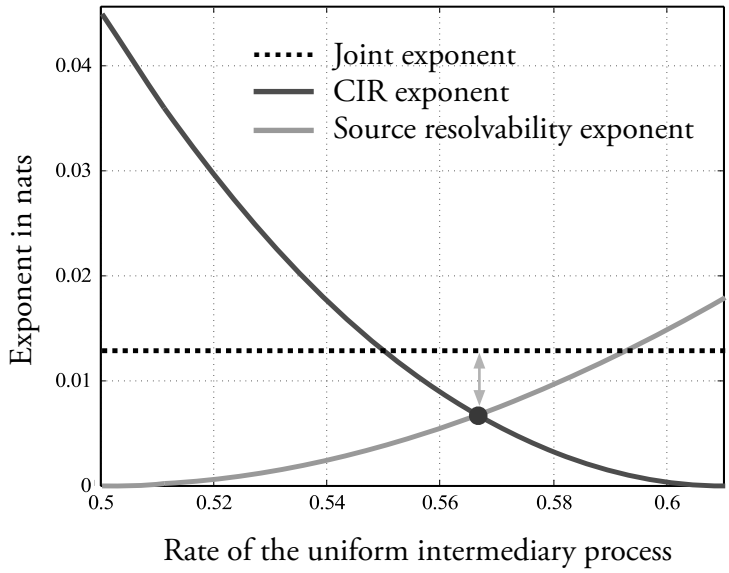

Figure 2.11 – *Determining the tandem exponent as the intersection of CIR and source resolvability exponents and comparison with the joint exponent for* $\kappa = 0.2$ *and* $\omega = 0.3$.

As illustrated in Figure 2.11, the tandem exponent corresponds to the intersection of the functions

$$R \mapsto \max_{\alpha \in [0,1]} E_1(\alpha, q_Z, W_1, R),$$

$$\text{and } R \mapsto \max_{\alpha \in [0,1]} E_2(\alpha, q_K, R).$$

In general, this intersection is below the optimal value of the joint exponent.

2.3.3 Converse

The following proposition presents a converse result for the joint scheme.

Proposition 2.32 For any DMC $(\mathscr{X}, W_1, \mathscr{Y})$ and any DMSs $(\mathscr{Z}, p_Z)$ and $(\mathscr{Y}, q_Y)$, joint channel intrinsic randomness and source resolvability requires

$$\mathbb{H}(X|Z) \geqslant \mathbb{H}(\tilde{Y}). \qquad \diamond$$

Proof Assume that joint channel intrinsic randomness and source resolvability is possible. For any $\epsilon > 0$, there exists a mapping ϕ_n such that $\mathbb{D}(p_{Y^nZ^n}\|q_{K^n}q_{Z^n}) \leqslant \epsilon$. Using [26, Lemma 2.7] and Pinsker's inequality gives the counterparts of Fano's equality in traditional source and channel coding.

- Since $\mathbb{D}(p_{Y^n}\|q_{Y^n}) \leqslant \delta(\epsilon)$, then

$$\frac{1}{n}\left|\mathbb{H}(Y^n) - \mathbb{H}(\tilde{Y}^n)\right| \leqslant \delta(\epsilon),$$

  where $\delta(\epsilon)$ denotes an arbitrary function of ϵ going to 0 with ϵ.

- Since $\mathbb{D}(p_{Y^nZ^n}\|p_{Y^n}q_{Z^n}) \leqslant \delta(\epsilon)$,

$$\frac{1}{n}\mathbb{I}(Y^n;Z^n) \leqslant \delta(\epsilon).$$

Therefore,

$$\begin{aligned}
\mathbb{H}(\tilde{Y}) \overset{\text{(a)}}{=} \frac{1}{n}\mathbb{H}(\tilde{Y}^n) &\overset{\text{(b)}}{\leqslant} \frac{1}{n}\mathbb{H}(Y^n) + \delta(\epsilon) \\
&= \frac{1}{n}\mathbb{H}(Y^n|Z^n) + \frac{1}{n}\mathbb{I}(Y^n;Z^n) + \delta(\epsilon) \\
&\overset{\text{(c)}}{\leqslant} \frac{1}{n}\mathbb{H}(Y^n|Z^n) + \delta(\epsilon) \\
&\overset{\text{(d)}}{\leqslant} \frac{1}{n}\mathbb{H}(X^n|Z^n) + \delta(\epsilon) \\
&\overset{\text{(e)}}{\leqslant} \mathbb{H}(X|Z) + \delta(\epsilon). \qquad (2.25)
\end{aligned}$$

Steps (a) and (e) follow since the source is memoryless, (b) arises from the continuity of the entropy rate, (c) from of the near independence of Y^n and Z^n, and (d) from the data processing inequality since Y^n is a function of X^n. ∎

Proposition 2.33 For any DMC $(\mathscr{Z}, W_1, \mathscr{X})$, any DMS $(\mathscr{Z}, p_Z)$ and $(\mathscr{Y}, q_Y)$, and any additive noise channel $(\mathscr{Y}, W_2, \mathscr{K})$, joint channel intrinsic randomness and channel resolvability requires

$$\mathbb{H}(X|Z) \geqslant \mathbb{I}(\tilde{Y};\tilde{K}). \qquad \diamond$$

Proof If the DMC $(\mathcal{Y}, W_2, \mathcal{K})$ is an additive channel, then $\tilde{K} = \tilde{Y} + E$, where E is some additive noise independent of the input $\tilde{Y}$. In this case,

$$\begin{aligned}
\mathbb{I}(\tilde{Y};\tilde{K}) \overset{(*)}{=} \frac{1}{n}\mathbb{I}(\tilde{Y}^n;\tilde{K}^n) &= \frac{1}{n}\mathbb{H}(\tilde{K}^n) - \frac{1}{n}\mathbb{H}(\tilde{K}^n|\tilde{Y}^n) \\
&\leqslant \frac{1}{n}\mathbb{H}(K^n) - \frac{1}{n}\mathbb{H}(E^n) + \delta(\epsilon) \\
&= \frac{1}{n}\mathbb{H}(K^n) - \frac{1}{n}\mathbb{H}(K^n|Y^n) + \delta(\epsilon) \\
&= \frac{1}{n}\mathbb{I}(K^n;Y^n) + \delta(\epsilon) \\
&\leqslant \frac{1}{n}\mathbb{H}(Y^n) + \delta(\epsilon). \qquad (2.26)
\end{aligned}$$

Step $(*)$ comes from the memoryless nature of the source and channel. Following the same steps as in the proof of Lemma 2.32 gives $\frac{1}{n}\mathbb{H}(Y^n) \leqslant \mathbb{H}(X|Z) + \delta(\epsilon)$. ∎

2.3.4 Discussion

If the channels are not memoryless, the analysis with the KL divergence does not carry over. Nevertheless, to obtain asymptotical results in the general case, one may use variational distance by replacing the criterion (2.3.1) by

$$\lim_{n\to\infty} \mathbb{V}(p_{K^nZ^n}, q_{K^n}q_{Z^n}) = 0. \qquad (2.27)$$

Using [38], the following sufficient condition obtained with a separate approach ensures the existence of an encoding mapping ϕ_n:

$$\text{p-liminf}_{n\to\infty} \frac{1}{n}H(X^n|Z^n) > \text{p-limsup}_{n\to\infty} \frac{1}{n}I(\tilde{Y}^n;\tilde{K}^n), \qquad (2.28)$$

where $H(X^n|Z^n)$, and $I(\tilde{Y}^n;\tilde{K}^n)$ are defined as

$$H(X^n|Z^n) \triangleq \log \frac{1}{p_{X^n|Z^n}(X^n|Z^n)} \quad \text{and} \quad I(\tilde{Y}^n;\tilde{K}^n) \triangleq \log \frac{W_2(\tilde{K}^n|\tilde{Y}^n)}{q_K(\tilde{K}^n)},$$

and

$$\begin{aligned}
\text{p-limsup}_{n\to\infty} \Xi_n &\triangleq \inf\left\{\xi \middle| \lim_{n\to\infty} \mathbb{P}(\Xi_n > \xi) = 0\right\} \\
\text{p-limsup}_{n\to\infty} \Xi_n &\triangleq \sup\left\{\xi \middle| \lim_{n\to\infty} \mathbb{P}(\Xi_n < \xi) = 0\right\}
\end{aligned}$$

for an arbitrary sequence of random variable $\{\Xi_n\}_{n=1}^{\infty}$ [39]. Note that, for a memoryless process,

$$\text{p-liminf}_{n\to\infty} \frac{1}{n} H(X^n|Z^n) = \mathbb{H}(X^n|Z^n), \quad \text{and} \quad \text{p-limsup}_{n\to\infty} \frac{1}{n} I(\tilde{Y}^n; \tilde{K}^n) = \mathbb{I}(\tilde{Y}^n; \tilde{K}^n)). \tag{2.29}$$

A matching converse when $W_2 = id_{\mathcal{X}}$ can be found in [13].

Appendix 7.1 briefly provides some details of the construction of practical codes with polar codes [7].

CHAPTER 3

THE TWO-WAY WIRETAP CHANNEL[1]

The wiretap channel is limited to some particular aspects of a multi-user scheme, but does not take into account cooperation, jamming, and feedback as means to increase secure communication rates. The two-way wiretap channel, in which users communicate over a noisy bidirectional channel while an eavesdropper observes interfering signals, combines all the effects present with multiple users because users have the possibility of cooperating while simultaneously jamming the eavesdropper. This model was first investigated by Tekin and Yener [92, 93] who showed that jamming with noise or controlled interference between codewords could provide secrecy gains. However, this strategy, called *cooperative jamming*, does not exploit feedback. It was later shown by He Yener [46] and Bloch [12] that strategies based on the feedback can perform strictly better than cooperative jamming alone. Recently, El Gamal *et al.* [28] proposed an achievable region for the two-way wiretap channel combining cooperative jamming and a secret-key exchange mechanism to transfer secure rate between users.

From a practical perspective, the two-way wiretap channel captures some of the limitations of real systems because all communications are intrinsically rate-limited. The two-way wiretap channel also generalizes many models; for instance, the works of Amariucai and Wei [3], Gündüz *et al.* [37], and Lai *et al.* [59] are special cases that focus on secure communication for one user only. Similarly, the model of Ardestanizadeh *et al.* [5] is a two-way wiretap channel in which one of the links is confidential and unheard by the eavesdropper. Many works on secret-key agreement with rate-limited public communication can be analyzed within this framework [24, 103] as well.

This chapter extends existing results in several directions: it is possible to design powerful coding schemes by partially decoupling the feedback and the interference and by relying on the

[1]Parts of the material in this chapter have appeared in [76]: **Pierrot, A. J.**, Bloch, M. R., "Strongly Secure Communications Over the Two-Way Wiretap Channel". In: *IEEE Transactions on Information Forensics and Security* 6.3 (Sept. 2011), pp. 595–605. ©IEEE 2011.

strategies presented in Section 1.3: cooperative jamming, secret-key exchange and secret-key generation.

Strong secrecy results, which require the eavesdropper to obtain a negligible amount of information instead of a negligible rate of information, exploit the concept of channel resolvability [11, 38–40, 90] to analyze cooperative jamming. Channel resolvability provides a conceptually convenient interpretation of cooperative jamming, which allows analyzing what happens when transmitting beyond the capacity of the eavesdropper's channel.

The outline of this chapter is as follows. Section 3.1 introduces the definitions pertaining to the two-way wiretap channel and a wiretap code. Section 3.2 presents a region of strongly secure rates achievable with cooperative jamming based on channel resolvability. This first step yields a result similar to what Tekin and Yener [92, Theorem 2] obtained for weak secrecy. In Section 3.3, the region is improved by introducing the secret-key exchange mechanism proposed in [28, 46]. The region is further extended by performing secret-key generation from a source induced by the noise used in cooperative jamming. Finally, Section 3.4 illustrates the achievable region in the Gaussian case.

3.1 Problem Statement

The problem of secure communication over a two-way wiretap channel is illustrated in Figure 3.1, in which:

- a legitimate user called Alice (or transmitter 1) sends message M_1 and estimates M_2;
- another legitimate user called Bob (or transmitter 2) sends message M_2 and estimates M_1;
- an eavesdropper called Eve observes Z^n.

The channel is supposed to be full-duplex, which means Alice and Bob communicate *simultaneously* over the channel. This assumption is relevant for some communication systems; however, it may be hard to realize in practice and many experimental communication systems operate with half-duplex, potentially yielding lower rates.

Definition 20 A *two-way wiretap channel*, denoted by

$$\left(\mathscr{X}_1, \mathscr{X}_2, \mathscr{Y}_1, \mathscr{Y}_2, \mathscr{Z}, \left\{p_{Y_1^n Y_2^n Z^n | X_1^n X_2^n}\right\}_{n \geqslant 1}\right),$$

consists of two arbitrary input alphabets $\mathscr{X}_1$ and $\mathscr{X}_2$, three arbitrary output alphabets $\mathscr{Y}_1$, $\mathscr{Y}_2$ and $\mathscr{Z}$, and a sequence of transition probabilities $\left\{p_{Y_1^n Y_2^n Z^n | X_1^n X_2^n}\right\}_{n \geqslant 1}$ such that:

$$\forall n \in \mathbb{N}^*, \forall \left(x_1^n, x_2^n\right) \in \mathscr{X}_1^n \times \mathscr{X}_2^n,$$
$$\sum_{y_1^n \in \mathscr{Y}_1^n} \sum_{y_2^n \in \mathscr{Y}_2^n} \sum_{z^n \in \mathscr{Z}^n} p_{Y_1^n Y_2^n Z^n | X_1^n X_2^n}\left(y_1^n, y_2^n, z^n | x_1^n, x_2^n\right) = 1. \quad (3.1)$$
$\diamond$

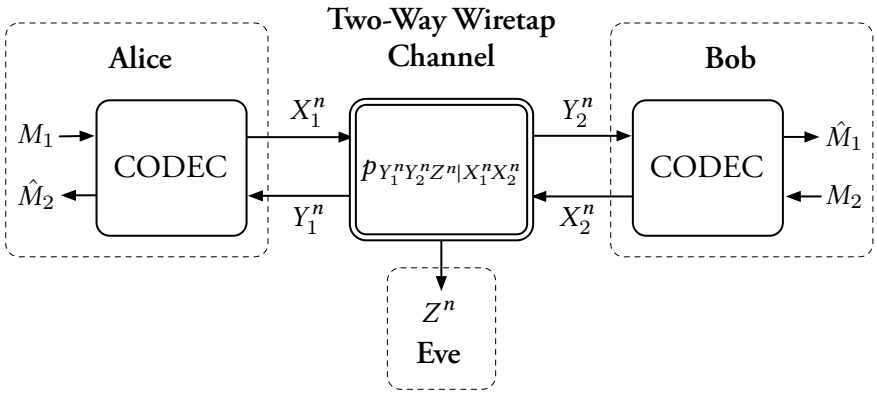

Figure 3.1 – *Communication over a two-way wiretap channel.*

The subsequent analysis is limited to a memoryless wiretap channel, but this approach generalizes in part to arbitrary channels using information spectrum methods [38].

Definition 21 A *memoryless* two-way wiretap channel, denoted by

$$\left(\mathscr{X}_1, \mathscr{X}_2, \mathscr{Y}_1, \mathscr{Y}_2, \mathscr{Z}, p_{Y_1 Y_2 Z | X_1 X_2}\right),$$

is a two-way wiretap channel for which:

$$\forall \left(x_1^n, x_2^n, y_1^n, y_2^n, z^n\right) \in \mathscr{X}_1^n \times \mathscr{X}_2^n \times \mathscr{Y}_1^n \times \mathscr{Y}_2^n \times \mathscr{Z}^n,$$
$$p_{Y_1^n Y_2^n Z^n | X_1^n X_2^n}\left(y_1^n, y_2^n, z^n | x_1^n, x_2^n\right) = \prod_{i=1}^{n} p_{Y_1 Y_2 Z | X_1 X_2}\left(y_1^{(i)}, y_2^{(i)}, z^{(i)} | x_1^{(i)}, x_2^{(i)}\right). \quad \diamond$$

A code for the two-way wiretap channel is formally defined as follows.

Definition 22 A $\left(2^{nR_1}, 2^{nR_2}, n\right)$ *two-way wiretap channel code* $\mathscr{C}_n$ consists of:

1. Two message alphabets: $\mathscr{M}_1 = [\![1, 2^{nR_1}]\!]$ and $\mathscr{M}_2 = [\![1, 2^{nR_2}]\!]$.

2. Two local sources of randomness $\left(\mathscr{R}_1, p_{R_1}\right)$ and $\left(\mathscr{R}_2, p_{R_2}\right)$ independent of the channel and messages.

3. Two sets of encoding functions that map a message and past channel observations to a channel input symbol:

   - n encoding functions for transmitter 1:

   $$\forall i \in [\![1, n]\!],\ f_1^{(i)} : \mathscr{M}_1 \times \mathscr{Y}_1^{i-1} \times \mathscr{R}_1 \to \mathscr{X}_1;$$

   - n encoding functions for transmitter 2:

   $$\forall i \in [\![1, n]\!],\ f_2^{(i)} : \mathscr{M}_2 \times \mathscr{Y}_2^{i-1} \times \mathscr{R}_2 \to \mathscr{X}_2.$$

4. Two decoding functions that map channel observations to a message or an error symbol "?":

   - $g_1 : \mathscr{Y}_1^n \times \mathscr{M}_1 \times \mathscr{R}_1 \to \mathscr{M}_2 \cup \{?\}$;
   - $g_2 : \mathscr{Y}_2^n \times \mathscr{M}_2 \times \mathscr{R}_2 \to \mathscr{M}_1 \cup \{?\}$.

The performance of a code $\mathscr{C}_n$ is assessed in terms of the following quantities:

- the probability of error:

$$\mathbf{P}_{\mathrm{e}}(\mathscr{C}_n) \triangleq \mathbb{P}\big((M_1, M_2) \neq (\hat{M}_1, \hat{M}_2) | \mathscr{C}_n\big);$$

- the information leakage to the eavesdropper:

$$\mathbf{L}(\mathscr{C}_n) \triangleq \mathbb{I}\left(Z^n; M_1 M_2 | \mathscr{C}_n\right).$$

$\diamond$

Definition 23 A rate pair (R_1, R_2) is *achievable* for a two-way wiretap channel if there exists a sequence of codes $\{\mathscr{C}_n\}_{n\geqslant 1}$ meeting the reliability and strong secrecy constraints:

- $\lim_{n\to\infty} \mathbf{P}_{\mathrm{e}}(\mathscr{C}_n) = 0$ (reliability);
- $\lim_{n\to\infty} \mathbf{L}(\mathscr{C}_n) = 0$ (strong secrecy). ◊

Definition 24 The *strong secrecy capacity* region $\bar{\mathscr{R}}^{2\mathrm{W}}$ is defined as:

$$\bar{\mathscr{R}}^{2\mathrm{W}} \triangleq \mathrm{cl}\left(\{(R_1, R_2) : (R_1, R_2) \text{ is achievable}\}\right),$$

whereas $\mathscr{R}^{2\mathrm{W}}$ denotes the *weak* secrecy capacity region. ◊

It is rather difficult to obtain a closed-form expression for the entire region of achievable rate pairs (R_1, R_2). In principle, the coding scheme in Definition 22 could simultaneously exploit the *interference* of transmitted signals at the eavesdropper's terminal and *feedback*. To obtain some insight, it is simpler to partially *decouple* these two effects.

- First, the interference penalizes the eavesdropper and increases secure communication rates. The interference can be of two types: interference between codewords or jamming with noise.
- Next, the feedback allows to increase the secrecy rate by means of *key exchange* and *key generation*. With secret-key exchange, one user sacrifices part of its secure communication rate to *exchange* a secret-key, whereas, with secret-key generation, both users exploit channel randomness to *distill* keys. Those keys are then used to encrypt messages with a one-time pad.

3.2 Resolvability-Based Cooperative Jamming

3.2.1 Cooperative Jamming

A natural attempt to increase secure communication rates consists in jamming Eve with noise, in order to decrease her signal-to-noise ratio. This strategy, called *cooperative jamming* [62], forces one user to stop transmitting information to jam the eavesdropper. To overcome this

limitation, Alice and Bob can use codewords whose interference also has a detrimental effect on Eve without sacrificing as much information rate. This scheme is called *coded cooperative jamming* and was introduced by Tekin and Yener [92, 93]. It is possible to combine both strategies and have Alice and Bob perform coded cooperative jamming while simultaneously jamming the eavesdropper with noise. Simultaneous cooperative jamming can be implemented by prefixing an artificial discrete memoryless channel (DMC) before the two-way wiretap channel (TWWTC) and sending codewords through the concatenated channels. This technique is therefore called *prefixing* in [28].

Note that cooperative jamming does not exploit feedback, which corresponds to using only two encoding functions $f_1^{(1)}$ and $f_2^{(1)}$ in Definition 22. Using cooperative jamming is then equivalent to studying the simplified channel model illustrated in Figure 3.2, in which the eavesdropper observes the output of a multiple-access channel.

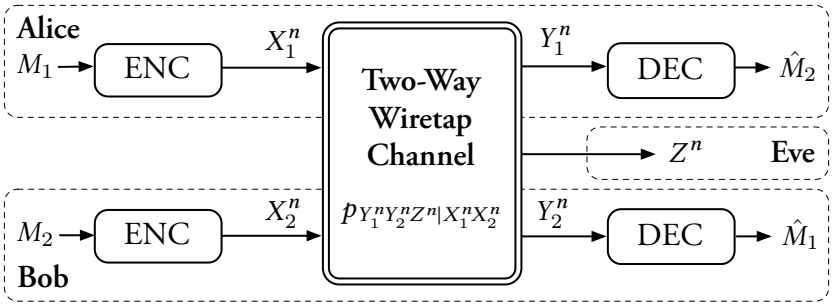

Figure 3.2 – *Communication over a two-way wiretap channel without feedback.*

3.2.2 Achievable Region

A first achievable region can be derived using the notion of channel resolvability, where uniformly distributed *auxiliary* messages $M_1' \in \mathscr{M}_1' \triangleq [\![1, 2^{nR_1'}]\!]$ and $M_2' \in \mathscr{M}_2' \triangleq [\![1, 2^{nR_2'}]\!]$ respectively play the role of the sources of randomness $(\mathscr{R}_1, p_{R_1})$ and $(\mathscr{R}_2, p_{R_2})$. Proposition 3.1 provides the region for rates (R_1, R_2, R_1', R_2').

Proposition 3.1

$$\mathscr{R} = \underset{R_1,R_2}{\text{Proj}} \bigcup_{p\in\mathscr{P}} \left\{ \begin{pmatrix} R_1 \\ R_2 \\ R_1' \\ R_2' \end{pmatrix} \in \mathbb{R}_+^4 \;\middle|\; \begin{array}{c} R_1 + R_1' \leqslant \mathbb{I}(Y_2; C_1|X_2) \\ R_2 + R_2' \leqslant \mathbb{I}(Y_1; C_2|X_1) \\ R_1' + R_2' \geqslant \mathbb{I}(C_1C_2; Z) \\ R_1' \geqslant \mathbb{I}(C_1; Z) \\ R_2' \geqslant \mathbb{I}(C_2; Z) \end{array} \right\} \subset \bar{\mathscr{R}}^{\text{2W}}, \tag{3.2}$$

where Proj_{R_1,R_2} is the projection on the plane of rates (R_1, R_2) and

$$\mathscr{P} = \{p_{X_1X_2C_1C_2Y_1Y_2Z} \text{ factorizing as: } p_{Y_1Y_2Z|X_1X_2}\, p_{X_1|C_1}\, p_{C_1}\, p_{X_2|C_2}\, p_{C_2}\}. \tag{3.3}$$

◊

Remark A similar result has been independently established by Yassaee and Aref [108] using a related technique based on approximation of output statistics. However, their proof only holds for discrete memoryless channels because it involves strongly typical sequences. The following proof relies on Steinberg's results [90], which hold for Gaussian memoryless channels.

Proof Two types of transmitted messages are considered to introduce randomness. First, the main messages must be transmitted between Alice and Bob reliably and securely with respect to Eve. Second, the auxiliary messages are used to perform coded cooperative jamming and introduce randomness to mislead the eavesdropper. Although auxiliary messages do not carry information by themselves, Alice and Bob must decode them *reliably*. The scheme also includes prefixing DMCs to perform simultaneous cooperative jamming.

The proof uses a random coding argument with fixed distributions p_{C_1}, p_{C_2}, $p_{X_1|C_1}$, and $p_{X_2|C_2}$ and a fixed $\epsilon > 0$.

Code creation The code consists of randomly generated codewords with encoding and decoding functions.

- **Code generation:** Generate $\lceil 2^{nR_1} \rceil \lceil 2^{nR_1'} \rceil$ i.i.d. sequences $c_1^n(\mu_1)$ with $\mu_1 = (i, j) \in \mathscr{M}_1 \times \mathscr{M}_1'$ according to p_{C_1}, and $\lceil 2^{nR_2} \rceil \lceil 2^{nR_2'} \rceil$ i.i.d. sequences $c_2^n(\mu_2)$ with $\mu_2 = (i, j) \in \mathscr{M}_2 \times \mathscr{M}_2'$ according to p_{C_2}. Here, i represents the index of the main message and j the index of the

auxiliary message. The random variable C_n represents the generated code, and $\mathscr{C}_n$ one of its realizations.

- **Encoding:** If Alice wants to send $\mu_1 \in \mathscr{M}_1 \times \mathscr{M}_1'$, she computes $c_1^n(\mu_1)$ and transmits x_1^n obtained by sending c_1^n through a DMC with transition probability $p_{X_1|C_1}$.

  Similarly, if Bob wants to send $\mu_2 \in \mathscr{M}_2 \times \mathscr{M}_2'$, he computes $c_2^n(\mu_2)$ and transmits x_2^n obtained by sending c_2^n through a DMC with transition probability $p_{X_2|C_2}$.

  The DMCs with transition probability $p_{X_1|C_1}$ and $p_{X_2|C_2}$ are prefix channels used for simultaneous cooperative jamming to confuse the eavesdropper.

- **Decoding:** the decoder is a typical set decoder.

  For simplicity, let $\mathscr{A}_{1,\epsilon}^n \triangleq \mathscr{A}_\epsilon^n(X_1, C_2, Y_1)$ and $\mathscr{A}_{2,\epsilon}^n \triangleq \mathscr{A}_\epsilon^n(X_2, C_1, Y_2)$.

  If y_1^n is received by Alice, she selects $\hat{\mu}_2$ such that $\left(x_1^n, c_2^n(\hat{\mu}_2), y_1^n\right) \in \mathscr{A}_{1,\epsilon}^n$. If such a tuple exists and is unique, output $\hat{\mu}_2$, otherwise declare an error ($\hat{\mu}_2 = ?$).

  Similarly, if y_2^n is received by Bob, he selects $\hat{\mu}_1$ such that $\left(x_2^n, c_1^n(\hat{\mu}_1), y_2^n\right) \in \mathscr{A}_{2,\epsilon}^n$. If such a tuple exists and is unique, output $\hat{\mu}_1$, otherwise declare an error ($\hat{\mu}_1 = ?$).

Probability of error analysis The probability of error is defined as

$$\mathrm{P_e}(\mathscr{C}_n) \triangleq \mathbb{P}\left((M_1, M_2, M_1', M_2') \neq (\hat{M}_1, \hat{M}_2, \hat{M}_1', \hat{M}_2') \middle| \mathscr{C}_n\right). \tag{3.4}$$

The following lemma provides conditions on rates such that this probability of error goes to zero as n goes to infinity.

Lemma 3.2 (Probability of error) For $\epsilon > 0$,

$$\begin{cases} R_1 + R_1' < \mathbb{I}(Y_2; C_1|X_2) \\ R_2 + R_2' < \mathbb{I}(Y_1; C_2|X_1) \end{cases} \Rightarrow \lim_{n\to\infty} \mathbb{E}(\mathrm{P_e}(C_n)) \leqslant \delta(\epsilon). \tag{3.5}$$

◊

Proof See Appendix 7.3.1 on page 128. ∎

Remark Lemma 3.2 provides conditions for reliable communication regardless of the eavesdropper, which is intuitive because reliability only depends on what is happening between the two legitimate users. However, this differs from the proof in [92], in which additional reliability constraints that depend on the eavesdropper are introduced to compute the leakage.

Leakage analysis The leakage is defined as

$$\mathbf{L}(\mathscr{C}_n) \triangleq \mathbb{I}(Z^n; M_1 M_2 | \mathscr{C}_n). \tag{3.6}$$

Lemma 3.3 (Leakage) For $\epsilon > 0$,

$$\begin{cases} R'_1 + R'_2 > \mathbb{I}(C_1 C_2; Z) \\ R'_1 > \mathbb{I}(C_1; Z) \\ R'_2 > \mathbb{I}(C_2; Z) \end{cases} \Rightarrow \lim_{n\to\infty} \mathbb{E}(\mathrm{L}(C_n)) \leqslant \delta(\epsilon). \tag{3.7}$$

♢

Proof See Appendix 7.3.2 on page 130. ∎

Remark Interpreting this result requires to precisely understand the role of auxiliary messages M'_1 and M'_2. These messages replace the sources of randomness and Lemma 3.3 offers lower bounds on the rates of these auxiliary messages. This result is also intuitive: more randomness must be introduced in the encoding process to prevent Eve from recovering the messages. Lemma 3.3 demonstrates that there exist minimum values of the rates that allow zero asymptotic leakage.

It is important to remember that, because of the constraints imposed by Lemma 3.2, increasing auxiliary message rates reduces the amount of information one can transmit through the channel.

Code selection Lemma 2.23 ("Selection Lemma") proves the existence of a specific sequence of codes $\{\mathscr{C}_n\}_{n\geqslant 1}$ such that

$$\lim_{n\to\infty} \mathbf{P}_{\mathrm{e}}(\mathscr{C}_n) \leqslant \delta(\epsilon) \text{ and } \lim_{n\to\infty} \mathbf{L}(\mathscr{C}_n) \leqslant \delta(\epsilon).$$

Remark Each code $\mathscr{C}_n$ consists of a pair of codes $(\mathscr{C}_1, \mathscr{C}_2)$. Although the codes $\mathscr{C}_1$ and $\mathscr{C}_2$ are generated according to independent distributions, note that both codes are *jointly* selected; therefore, the codes are used independently by Alice and Bob but optimized jointly to guarantee secrecy.

Conclusion For any $\epsilon > 0$, there exists a code such that the probability of error and the leakage are smaller than $\delta(\epsilon)$. Therefore, it is possible to create a sequence of codes $\{\mathscr{C}_n(\epsilon_n)\}_{n\geqslant 1}$ with $\epsilon_n \underset{n\to\infty}{\longrightarrow} 0$.

Combining the rate constraints in (3.5) and (3.7) yields the result provided in Proposition 3.1. ∎

The secure achievability region is obtained by elimination of the auxiliary rates R_1' and R_2'.

Corollary 3.4 (Strongly secure achievable region)

$$\mathscr{R} = \bigcup_{p\in\mathscr{P}} \left\{ \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} \in \mathbb{R}_+^2 \;\middle|\; \begin{aligned} R_1 &\leqslant \mathbb{I}(Y_2; C_1|X_2) - \mathbb{I}(C_1; Z) \\ R_2 &\leqslant \mathbb{I}(Y_1; C_2|X_1) - \mathbb{I}(C_2; Z) \\ R_1 + R_2 &\leqslant \mathbb{I}(Y_2; C_1|X_2) + \mathbb{I}(Y_1; C_2|X_1) \\ &\quad - \mathbb{I}(C_2C_2; Z) \end{aligned} \right\} \subset \bar{\mathscr{R}}^{2\mathrm{W}}, \tag{3.8}$$

where

$$\mathscr{P} = \{p_{X_1X_2C_1C_2Y_1Y_2Z} \text{ factorizing as: } p_{Y_1Y_2Z|X_1X_2}\, p_{X_1|C_1}\, p_{C_1}\, p_{X_2|C_2}\, p_{C_2}\}. \tag{3.9}$$

◊

Proof (Corollary 3.4) With (3.2),

$$\begin{aligned} R_1 + R_1' + R_2 + R_2' &\leqslant \mathbb{I}(Y_2; C_1|X_2) + \mathbb{I}(Y_1; C_2|X_1) \\ R_1 + R_2 &\leqslant \mathbb{I}(Y_2; C_1|X_2) + \mathbb{I}(Y_1; C_2|X_1) - R_1' - R_2' \\ R_1 + R_2 &\leqslant \mathbb{I}(Y_2; C_1|X_2) + \mathbb{I}(Y_1; C_2|X_1) - \mathbb{I}(C_2C_2; Z). \end{aligned}$$

Using (3.7),

$$\begin{aligned} R_1 + R_1' &\leqslant \mathbb{I}(Y_2; C_1|X_2) \\ R_1 &\leqslant \mathbb{I}(Y_2; C_1|X_2) - R_1' \\ R_1 &\leqslant \mathbb{I}(Y_2; C_1|X_2) - \mathbb{I}(C_1; Z). \end{aligned}$$

Similarly, $R_2 \leqslant \mathbb{I}(Y_1; C_2|X_1) - \mathbb{I}(C_2; Z)$. ∎

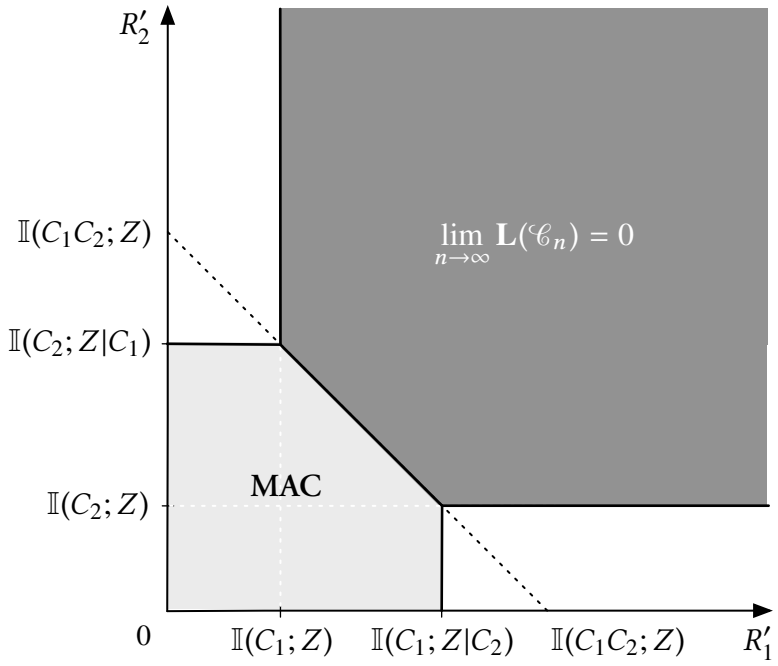

Figure 3.3 – *Constraints on R'_1 and R'_2 in* (3.2).

The region described in (3.8) is identical to the one obtained by Tekin and Yener in [92, 93] for weak secrecy. However, a closer look at the proof shows that their result is obtained by projecting the region $\mathscr{R}'$ defined as:

$$\mathscr{R}' = \bigcup_{p\in\mathscr{P}} \left\{ \begin{pmatrix} R_1 \\ R_2 \\ R'_1 \\ R'_2 \end{pmatrix} \in \mathbb{R}_+^4 \;\middle|\; \begin{array}{r} R_1 + R'_1 \leqslant \mathbb{I}(Y_2; C_1|X_2) \\ R_2 + R'_2 \leqslant \mathbb{I}(Y_1; C_2|X_1) \\ R'_1 + R'_2 = \mathbb{I}(C_1C_2; Z) \\ R'_1 \leqslant \mathbb{I}(C_1; Z|C_2) \\ R'_2 \leqslant \mathbb{I}(C_2; Z|C_1) \end{array} \right\}. \tag{3.10}$$

This region differs from 3.2 only for the constraints on auxiliary message rates (R'_1, R'_2). The difference is illustrated in Figure 3.3, where the dark area corresponds to constraints (3.7) and the light one to the constraints on (R'_1, R'_2) in (3.10). Note that the latter corresponds to the achievable region of a Multiple-Access Channel (MAC). This is not surprising because the proof of (3.10) relies explicitly on the analysis of the probability of error for the eavesdropper, who

obtains its signal through a virtual multiple-access channel as depicted in Figure 3.2. In contrast, the present approach analyzes the secrecy constraint directly, which leads to lower bounds on the auxiliary message rate required to confuse the eavesdropper. The projection of both regions on the plane of rates (R_1, R_2) is the same because it corresponds to having the auxiliary message rates on the diagonal edge, which is common to both. In terms of code structure, the approach of Tekin and Yener consists in augmenting the number of auxiliary messages until the leakage to the eavesdropper become negligible, which only happens on the slope of the MAC region. The constraints directly yield a region with negligible leakage and the approach consists in finding the minimum number of auxiliary messages needed to confuse the eavesdropper, thus augmenting the number of secret messages to the maximum possible value.

3.3 Secret-Key Exchange and Secret-Key Generation

The results in the previous section exploit the benefits of coded cooperative jamming and simultaneous cooperative jamming but do not consider the possibility of feedback. In particular, two mechanisms leverage feedback.

- First, the techniques presented in [46] and [28] allow transferring secret rate from one user to the other.
- Next, the randomness introduced by cooperative jamming is used to induce a source to distill secret-keys. Results for secret-key agreement with rate-limited public communication [24, 103] prove useful since information can be only exchanged through a rate-limited channel.

3.3.1 Key Exchange

Key exchange takes place on top of the cooperative jamming scheme by splitting the main and auxiliary messages into multiple parts. The sub-messages facilitate the exchange of a secret-key using the secret channel and encrypt part of the public message. Because of the "secret rate transfer," the rates must be redefined accordingly.

Consider a code for cooperative jamming with secret message rates (R_1, R_2) and auxiliary messages rates (R'_1, R'_2). For $i \in \{1, 2\}$, the main message M_i is split into two parts.

- A *key*, which is used for encryption by the other user $K_i \in \mathcal{K}_i = [\![1, 2^{nR_i^{\mathrm{k}}}]\!]$; user i needs to sacrifice a part of its secret message to transmit this key.
- A *secret* message $M_i^{\mathrm{s}} \in \mathcal{M}_i^{\mathrm{s}} = [\![1, 2^{nR_i^{\mathrm{s}}}]\!]$; this part corresponds to the part of the secret message user i does not sacrifice.

The auxiliary message M'_i is also split into two parts.

- An *encrypted* version $M_i^{\mathrm{e}} \in \mathcal{M}_i^{\mathrm{e}} = [\![1, 2^{nR_i^{\mathrm{e}}}]\!]$ of a message $M_i^{\notin}$. Encryption is done with a secret-key provided by the other user and a one-time pad to ensure perfect secrecy [35].
- An *open* message $M_i^{\mathrm{o}} \in \mathcal{M}_i^{\mathrm{o}} = [\![1, 2^{nR_i^{\mathrm{o}}}]\!]$; this corresponds to the part of the message that remains public and is still perfectly decipherable by the receiver.

By convention, if no secret-key is available $R_i^{\mathrm{e}} = 0$, otherwise, by construction, $R_1^{\mathrm{e}} \leqslant R_2^{\mathrm{k}}$, $R_2^{\mathrm{e}} \leqslant R_1^{\mathrm{k}}$. Since $\mathcal{M}_i = \mathcal{K}_i \times \mathcal{M}_i^{\mathrm{s}}$ and $\mathcal{M}'_i = \mathcal{M}_i^{\mathrm{e}} \times \mathcal{M}_i^{\mathrm{o}}$, the various rates relate as

$$R_1 = R_1^{\mathrm{s}} + R_1^{\mathrm{k}}, \quad R_2 = R_2^{\mathrm{s}} + R_2^{\mathrm{k}}, \quad R'_1 = R_1^{\mathrm{o}} + R_1^{\mathrm{e}}, \text{ and } R'_2 = R_2^{\mathrm{o}} + R_2^{\mathrm{e}}. \tag{3.11}$$

Although (R_1, R_2, R'_1, R'_2) still represents the rates provided by cooperative jamming, they are no longer the rates of interest after the secret-key exchange. In fact, part of the auxiliary message is encrypted while part of the secret message is sacrificed to exchange a key. Thus, the rates to consider are the following:

- a pair of secret rates: $\tilde{R}_1 = R_1^{\mathrm{s}} + R_1^{\mathrm{e}}$ and $\tilde{R}_2 = R_2^{\mathrm{s}} + R_2^{\mathrm{e}}$;
- a pair of public rates: $\tilde{R}'_1 = R_1^{\mathrm{o}}$ and $\tilde{R}'_2 = R_2^{\mathrm{o}}$.

Remark Because the secret-key sent by one user cannot be used simultaneously by the other, the secret-key exchange scheme must operate in several rounds. The secret-key comes from the previous one, except in the first round where no secret-key is available. A code of length n is

used B times, giving a new code of length $n' = Bn$; the first message does not use secret-key exchange, but the next $B-1$ do. If the communication rate for the first message is R^*, for the other ($B-1$ is R, and the overall rate is

$$\bar{R} = \frac{nR^* + n(B-1)R}{nB} \underset{B\to\infty}{=} R.$$

Thus, the first round incurs a negligible rate penalty as B goes to infinity.

For weak secrecy, the authors of [28] prove the following proposition:

Proposition 3.5 (El Gamal *et al.*)

$$\mathscr{R}^{\mathrm{F}} = \bigcup_{p\in\mathscr{P}} \left\{ \begin{pmatrix} R_1 \\ R_2 \\ R_1' \\ R_2' \end{pmatrix} \in \mathbb{R}_+^4 \;\middle|\; \begin{array}{r} R_1 \leqslant \mathbb{I}(Y_2; C_1|X_2) \\ R_2 \leqslant \mathbb{I}(Y_1; C_2|X_1) \\ R_1 + R_2 \leqslant \mathbb{I}(Y_2; C_1|X_2) + \mathbb{I}(Y_1; C_2|X_1) \\ - \mathbb{I}(C_2C_2; Z) \end{array} \right\} \subseteq \mathscr{R}^{\mathrm{2W}}, \tag{3.12}$$

where:

$$\mathscr{P} = \{p_{X_1X_2C_1C_2Y_1Y_2Z} \text{ factorizing as: } p_{Y_1Y_2ZX_1X_2}\, p_{X_1|C_1}\, p_{C_1}\, p_{X_2|C_2}\, p_{C_2}\}. \tag{3.13}$$

◇

Based on Section 3.2, this result also holds for strong secrecy: $\mathscr{R}^{\mathrm{F}} \subset \bar{\mathscr{R}}^{\mathrm{2W}}$.

Remark Comparing (3.8) and (3.12) proves that secret-key exchange improves the individual bounds on R_1 and R_2, but not the bound on the sum-rate. Individual bounds on R_1 and R_2 correspond to the capacity of the channel between the two users and cannot be improved; therefore, any improvement in the region should modify the sum-rate constraint.

3.3.2 Key Generation from Induced Source

Key exchange requires sacrificing part of the secret rate of one user. In addition, the channel randomness introduced for simultaneous cooperative jamming can be used to *extract* secret-keys. Although users must exchange additional messages to agree on a common secret-key, secret-key generation only comes at the expense of *public* message rate.

The next section focuses on the Gaussian two-way wiretap channel and explicitly demonstrates how cooperative jamming induces a discrete memoryless source that can be used to distill

a secret-key. The existence of this DMS is not obvious because the noise introduced by cooperative jamming is already exploited to harm the eavesdropper. It is a priori unclear if it could be used simultaneously as a source of common randomness for secret-key agreement. First, one must clarify what knowledge of the eavesdropper needs to be considered to identify the DMS. Let K be the secret-key to distill and let M_1, M_2, M_1' and M_2' be the secret and auxiliary messages used in the cooperative jamming code. The secret-key and secret messages must be independent and both must be hidden from the eavesdropper, that is, for $\epsilon > 0$, $\mathbb{I}(M_1M_2K; Z^n) \leqslant \epsilon$. Notice that

$$\begin{aligned}\mathbb{I}(M_1M_2K; Z^n) &= \mathbb{I}(M_1M_2; Z^n) + \mathbb{I}(K; Z^n|M_1M_2) \\ &= \mathbb{I}(M_1M_2; Z^n) + \mathbb{I}(K; Z^nM_1M_2). \end{aligned} \tag{3.14}$$

The term $\mathbb{I}(M_1M_2; Z^n)$ can be smaller than $\epsilon/2$ by construction of the cooperative jamming code, which requires $\mathbb{I}(K; Z^nM_1M_2) \leqslant \epsilon/2$; this means that the secret-key must be hidden from an eavesdropper having access not only to the observation Z^n but also to the secret messages M_1 and M_2. Furthermore, note that

$$\mathbb{I}(M_1M_2K; Z^n) = \mathbb{I}(M_1M_2; Z^n) + \mathbb{I}\big(K; Z^nM_1M_2M_1'M_2'\big) - \mathbb{I}\big(K; M_1'M_2'|Z^nM_1M_2\big). \tag{3.15}$$

If the cooperative jamming code is chosen to provide the highest secrecy rate, the public message rates must be chosen to lie on the boundary of the region in (3.8), for which [92, 93] shows that $\mathbb{H}\big(M_1'M_2'|Z^nM_1M_2\big) \leqslant \epsilon/2$. In this case,

$$\mathbb{I}(M_1M_2K; Z^n) \geqslant \mathbb{I}(M_1M_2; Z^n) + \mathbb{I}\big(K; Z^nM_1M_2M_1'M_2'\big) - \frac{\epsilon}{2}. \tag{3.16}$$

In other words, the secret-key K must be kept secret from an eavesdropper observing not only the channel output Z^n, but also knowing the secret and auxiliary messages transmitted by the legitimate users.

3.3.3 Achievable Region with Secret-Key Exchange and Secret-Key Generation

Adding secret-key exchange and secret-key generation primitives to the coding scheme creates new dependencies between the random variables of the different blocks. Splitting the messages

also introduces additional rate constraints and increases the dimension of the achievable rate region.

3.3.3.1 Fundamental Considerations

The results of Lemma 3.3 relies on resolvability results for the multiple-access channel by Steinberg [90]. However, this proof does not take into account several subtleties that appear when secret-key exchange and secret-key generation are performed over the two-way wiretap channel:

- the auxiliary messages M_1' and M_2' are no longer uniformly distributed;
- the eavesdropper does not have prior knowledge of the public messages sent;
- there exist dependencies between blocks since the secret-key is used across different blocks.

Non-uniform auxiliary messages The proof of Lemma 3.3 relies on the assumption that the auxiliary messages are uniformly distributed. Part of the auxiliary message is encrypted with a secret-key that is almost uniform, while the other part is used for the public message communication. Since both these sub-messages are not exactly uniform, Lemma 3.3 cannot be used directly and needs to be extended for the case of non-uniform messages.

Rate-limited secret-key generation is similar to the Wyner-Ziv problem [27], in which two users distill identical sequences from a correlated source by exchanging messages over a rate-limited public channel. In a similar fashion, the codebook generation operates as follows:

- the observation $\tilde{X}^n$ is sent trough a virtual channel, whose transition probability $p_{U|X}$ is controlled by Alice;[2]
- sequences U^n are labeled with three indices $F \in [\![1, 2^{nR_f}]\!]$, $K \in [\![1, 2^{nR_k}]\!]$, $C \in [\![1, 2^{nR_c}]\!]$ independently and uniformly distributed.

Using results from Chapter 2, several results follows:

[2]The case where Bob is generating the secret-key is similar.

1. channel intrinsic randomness ensures that indices are uniform and independent of $\tilde{Z}^n$, i.e. $\mathbb{V}\left(p_{FKC\tilde{Z}^n}, u_F u_K u_C p_{\tilde{Z}^n}\right) \leqslant \epsilon$, if

$$R_f + R_k + R_c < \mathbb{H}(U|\tilde{Z}); \tag{3.17}$$

2. Slepian-Wolf coding guarantees that it is possible to reconstruct U by observing $\tilde{Y}^n$ knowing indices F and K, i.e. there is a decoding function g such that $\mathbb{P}\left(U \neq g(\tilde{Y}^n, C)\right) \leqslant \epsilon$, if

$$R_f + R_c > \mathbb{H}(U|\tilde{Y}); \tag{3.18}$$

3. channel intrinsic randomness with respect to $\tilde{X}^n$ establishes that index C is independent of $\tilde{X}^n$, i.e $\mathbb{V}\left(p_{C\tilde{X}^n}, u_C p_{\tilde{X}^n}\right) \leqslant \epsilon$, if

$$R_c < \mathbb{H}(U|\tilde{X}). \tag{3.19}$$

These three constraints imply $R_f > \mathbb{I}\left(U;\tilde{X}\right) = \mathbb{I}\left(U;\tilde{Y}\right)$. The philosophy behind this approach is to design a code for the virtual source $(U, \tilde{X}, \tilde{Y}, \tilde{Z})$ based on the random binning of U with three indices. The index F represents the public message that will be exchanged over the rate-limited public channel; the index K represents the secret-key generated by both parties. The role of index C is less intuitive since it is used to select a code among all possible codes provided by this procedure. The first and third constraints guarantees secrecy, the second guarantee reliability.

The next step consists in showing that designing a code for the virtual source $(U, \tilde{X}, \tilde{Y}, \tilde{Z})$ is equivalent to designing a code for secret-key generation. The source joint probability is denoted $p_{U\tilde{X}\tilde{Y}\tilde{Z}}$, while the original joint probability representing the problem of secret-key generation is denoted $\tilde{p}_{U\tilde{X}\tilde{Y}\tilde{Z}}$. Including the auxiliary indices the total joint probability for the virtual source is

$$\begin{aligned} &p_{U\tilde{X}^n\tilde{Y}^n\tilde{Z}^nFKC}(u, x^n, y^n, z^n, f, k, \mathscr{C}) \\ &\qquad = p_{FKC|U}(f, k, \mathscr{C}|u) p_{U|\tilde{X}^n}(u|x^n) p_{\tilde{X}^n\tilde{Y}^n\tilde{Z}^n}(x^n, y^n, z^n), \end{aligned} \tag{3.20}$$

where Φ, Ψ and Υ represents the random encoding functions respectively giving indices F, K, and C for a given U. Note that

$$p_{FKC|U}(f,k,\mathscr{c}|u) = \mathbb{1}\{\Phi(u)=f, \Psi(u)=k, \Upsilon(u)=\mathscr{c}\}.$$

To show that the code previously designed works also for the problem of secret-key generation with rate-limited public communications, it suffices to verify that $p_{U\tilde{X}\tilde{Y}\tilde{Z}}$ is close to $\tilde{p}_{U\tilde{X}\tilde{Y}\tilde{Z}}$.

$$\begin{aligned}
\mathbb{V}\left(p_{U\tilde{X}^n\tilde{Y}^n\tilde{Z}^n}, \tilde{p}_{U\tilde{X}^n\tilde{Y}^n\tilde{Z}^n}\right) &= \mathbb{V}\left(p_{U\tilde{X}\tilde{Y}\tilde{Z}}, \tilde{p}_{U\tilde{X}^n\tilde{Y}^n\tilde{Z}^n}\right) \\
&= \mathbb{V}\left(p_{U\tilde{X}^nFKC}, u_C p_{FK|U}\, p_{U|C\tilde{X}^n} p_{\tilde{X}^n}\right) \\
&= \mathbb{V}\left(p_{FK|U}\, p_{U|C\tilde{X}^n} p_{C\tilde{X}^n}, u_C p_{FK|U}\, p_{U|C\tilde{X}^n} p_{\tilde{X}^n}\right) \\
&= \mathbb{V}\left(p_{C\tilde{X}^n}, u_C p_{\tilde{X}^n}\right) \leqslant \epsilon. \qquad (3.21)
\end{aligned}$$

The index C needs to be eliminated so that the same code can be used for any source observation with an argument similar to the selection lemma.

$$\begin{aligned}
\mathbb{P}\left(U \neq g(\tilde{Y}^n, C)\right) \leqslant \epsilon &= \sum_{\mathscr{c}=1}^{2^{nR_c}} \mathbb{P}\left(U \neq g(\tilde{Y}^n, \mathscr{c}) \middle| C = \mathscr{c}\right) u_C(\mathscr{c}) \\
&= \mathbb{E}_I\left(\mathbb{P}\left(U \neq g(\tilde{Y}^n, C) \middle| C\right)\right) \leqslant \epsilon. \qquad (3.22)
\end{aligned}$$

Markov's inequality implies $\mathbb{P}_I\left(\mathbb{P}\left(U \neq g(\tilde{Y}^n, C)\middle|C\right) > 2\epsilon\right) \leqslant 1/2$, thus there exists a particular codebook $\mathscr{C}_0$ such that $\mathbb{P}\left(U \neq g(\tilde{Y}^n, \mathscr{C}_0)\right) \leqslant 2\epsilon$

This coding scheme also guarantees that the message F and the secret-key K are almost uniform in terms of variational distance. The following lemma extends lemma 3.3 when auxiliary messages M_1' and M_2' are not exactly uniform.

Lemma 3.6 If the following three conditions

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{H}_2(M_1', M_2') > \mathbb{I}(X_1X_2;Z),\ \lim_{n\to\infty}\frac{1}{n}\mathbb{H}_2(M_1') > \mathbb{I}(X_1;Z),\ \text{and}\ \lim_{n\to\infty}\frac{1}{n}H_2(M_2') > \mathbb{I}(X_2;Z) \qquad (3.23)$$

are satisfied, then

$$\exists \beta > 0,\ \mathbb{E}_{C_n}(\mathbb{V}(p_{M_1M_2Z^n}, p_{M_1}p_{M_2}p_{Z^n})) \leqslant 2^{-\beta n}. \qquad (3.24)$$

$\diamond$

Proof See Appendix 7.3.3 on page 133. ∎

The proof of this lemma also relies on resolvability arguments but does not directly provide constraints on the communication rate. Instead, this lemma yields three constraints on the second order Rényi entropy of the auxiliary messages. For the encrypted message M_i^{e} (for $i \in \{1, 2\}$), the encryption is performed with a one-time pad by using a secret-key K_i. The secret-key K_i corresponds to the index K presented above and obtained from the previous round of communication is almost uniform, that is $\mathbb{V}\left(p_{K_i}, u_{K_i}\right) \leqslant \epsilon$, where u_{K_i} is the uniform distribution with same support as K_i. The encryption operation is $M_i^{\mathrm{e}} \triangleq M_i^{\not{\mathrm{e}}} \oplus K_i$ where $M_i^{\not{\mathrm{e}}}$ is the part of the auxiliary messages that needs to be encrypted. Additionaly, $q_{M_i^{\mathrm{e}}}$ represents the distribution of the encrypted message when $M_i^{\not{\mathrm{e}}}$ is uniformly distributed. Since the encryption is performed with a one-time pad $q_{M_i^{\mathrm{e}}} \equiv u_{K_i}$. The triangular inequality yields

$$\mathbb{V}\left(p_{M_i^{\mathrm{e}}}, u_{K_i}\right) \leqslant \mathbb{V}\left(p_{M_i^{\mathrm{e}}}, q_{M_i^{\mathrm{e}}}\right) + \underbrace{\mathbb{V}\left(q_{M_i^{\mathrm{e}}}, u_{K_i}\right)}_{\equiv 0}. \tag{3.25}$$

Using the data processing inequality for the variational distance,

$$\mathbb{V}\left(p_{M_i^{\mathrm{e}}}, q_{M_i^{\mathrm{e}}}\right) \leqslant \mathbb{V}\left(p_{K_i}, u_{K_i}\right) \leqslant \epsilon. \tag{3.26}$$

Therefore $\mathbb{V}\left(p_{M_i^{\mathrm{e}}}, u_{K_i}\right) \leqslant \epsilon$.

The remaining part of the auxiliary message corresponds to the open message M_i^{o} that is used as a public channel for the secret-key generation. This message corresponds to the index F in the scheme presented before, which is such that $\mathbb{V}\left(p_{M_i^{\mathrm{o}}}, u_F\right) \leqslant \epsilon$.

As for the total auxiliary message $p_{M_i'} = p_{M_i^{\mathrm{o}}} p_{M_i^{\mathrm{e}}}$, and since M_i^{o} and M_i^{e} are independent, then

$$\begin{aligned}
\mathbb{V}\left(p_{M_i'}, u_{M_i'}\right) &= \mathbb{V}\left(p_{M_i^{\mathrm{o}}} p_{M_i^{\mathrm{e}}}, u_{M_i^{\mathrm{o}}} u_{K_i}\right) \\
&\leqslant \mathbb{V}\left(p_{M_i^{\mathrm{o}}} p_{M_i^{\mathrm{e}}}, p_{M_i^{\mathrm{o}}} u_{K_i}\right) - \mathbb{V}\left(p_{M_i^{\mathrm{o}}} u_{K_i}, u_{M_i^{\mathrm{o}}} u_{K_i}\right) \\
&= \mathbb{V}\left(p_{M_i^{\mathrm{e}}}, u_{K_i}\right) - \mathbb{V}\left(p_{M_i^{\mathrm{o}}}, u_{M_i^{\mathrm{o}}}\right) \\
&\leqslant 2\epsilon
\end{aligned} \tag{3.27}$$

Prior knowledge of the eavesdropper The constraint (3.17) incidentally guarantees that the public messages exchanged for the reconciliation are independent of the eavesdropper's statistics.

Dependencies between random variable Secret-key exchange and secret-key generation require passing messages through different blocks. This introduces additional dependencies among the random variables that must be taken into account in the analysis. For the sake of simplicity, suppose both Alice and Bob simultaneously distill and use a secret-key.

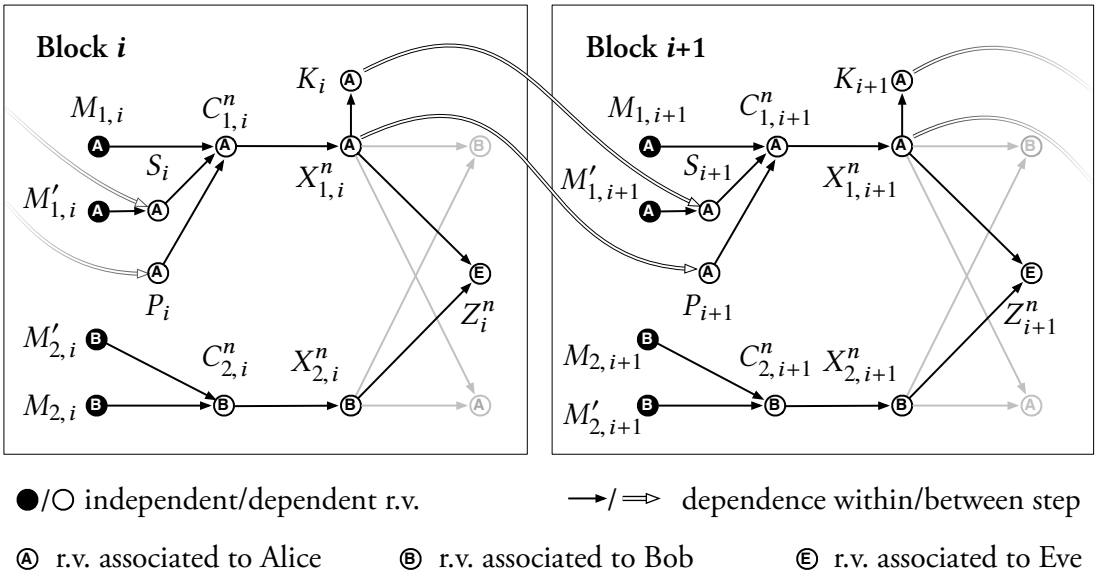

Figure 3.4 – *Dependence graph when secret-key generation is used over the two-way wiretap channel (Alice generates/uses the secret-key).*

Figure 3.4 illustrates the initial steps of secret-key generation, when Alice both generates and uses the secret-key, and represents the dependencies between the different blocks. For the secrecy analysis the following term must be upper bounded

$$\mathbb{I}\left(M_{1,1:B}M'_{1,2:B}M_{2,1:B};Z^n_{1:B}\right),$$

where B represents the total number of blocks. The following lemma bounds the increase in information brought by the observation of Z_{i+1}.

Lemma 3.7 Let $i \in [\![1, B-1]\!]$. Define $\mathbf{L}_i \triangleq \mathbb{I}\left(M_{1,1:B}M'_{1,2:B}M_{2,1:B}; Z_{1:i}^n\right)$. We have

$$\mathbf{L}_{i+1} - \mathbf{L}_i \leqslant \delta_n(\epsilon). \tag{3.28}$$

◇

Proof See Appendix 7.3.4 on page 137. This proof is given when Alice both generates and uses the secret-key, but the result still holds for the three other configurations. ∎

We then have

$$\begin{aligned}
\mathbf{L}_1 &= \mathbb{I}\left(M_{1,1:B}M'_{1,2:B}M_{2,1:B}; Z_1^n\right) \\
&= \mathbb{I}\left(M_{1,1}M_{2,1}; Z_1^n\right) + \mathbb{I}\left(M_{1,2:B}M'_{1,2:B}M_{2,2:B}; Z_1^n \middle| M_{1,1}M_{2,1}\right) \\
&\leqslant \delta_n(\epsilon) + \mathbb{I}\left(M_{1,2:B}M'_{1,2:B}M_{2,2:B}; Z_1^n \middle| M_{1,1}M_{2,1}\right) \\
&\leqslant \delta_n(\epsilon) + \mathbb{I}\left(M_{1,2:B}M'_{1,2:B}M_{2,2:B}; Z_1^n M_{1,1}M_{2,1}\right) \\
&\overset{(*)}{=} \delta_n(\epsilon),
\end{aligned}$$

where $(*)$ follows from independence of $M_{1,2:B}M'_{1,2:B}M_{2,2:B}$ and the random variables of Block 1.

Hence, the strong secrecy over multiple blocks follows from Lemma 3.7 by remarking that

$$\begin{aligned}
\mathbb{I}\left(M_{1,1:B}M'_{1,2:B}M_{2,1:B}; Z_{1:B}^n\right) &= \mathbf{L}_1 + \sum_{i=1}^{B-1}(\mathbf{L}_{i+1} - \mathbf{L}_i) \\
&\leqslant \delta_n(\epsilon) + (B-1)(\delta_n(\epsilon)) \\
&= B\delta_n(\epsilon).
\end{aligned}$$

3.3.3.2 Practical Considerations

Let one assume the existence of a DMS $(\tilde{X}, \tilde{Y}, \tilde{Z})$ independent of all other observations induced by cooperative jamming. The statistics of the DMS depend both on channel statistics and the code used for communications but does not need to be characterized completely.

First, note that substituting the rates defined in (5.17) in Proposition 3.1 provides a set of constraints that achievable rates must satisfy:

$$R_1^{\mathrm{s}} + R_1^{\mathrm{k}} + R_1^{\mathrm{o}} + R_1^{\mathrm{e}} \leqslant \mathbb{I}(Y_2; C_1|X_2) \tag{3.29}$$

$$R_2^{\mathrm{s}} + R_2^{\mathrm{k}} + R_2^{\mathrm{o}} + R_2^{\mathrm{e}} \leqslant \mathbb{I}(Y_1; C_2|X_1) \tag{3.30}$$

$$R_1^{\mathrm{o}} + R_1^{\mathrm{e}} + R_2^{\mathrm{o}} + R_2^{\mathrm{e}} \geqslant \mathbb{I}(C_1C_2; Z) \tag{3.31}$$

$$R_1^{\mathrm{o}} + R_1^{\mathrm{e}} \geqslant \mathbb{I}(C_1; Z) \tag{3.32}$$

$$R_2^{\mathrm{o}} + R_2^{\mathrm{e}} \geqslant \mathbb{I}(C_2; Z). \tag{3.33}$$

Introducing the secret-key generation by splitting the codewords in multiple parts requires to consider slightly different rates. For instance, the rates of interest for Alice become the following.

- $\tilde{R}_1^{\mathrm{s}} = R_1^{\mathrm{s}}$ since secret-key generation does not change anything for the secret message rate.
- $\tilde{R}_1^{\mathrm{k}} = R_1^{\mathrm{k}}$ since secret-key generation does not change anything for the secret-key rate used for secret-key exchange.
- $\tilde{R}_1^{\mathrm{o}} = R_1^{\mathrm{o}} - \bar{R}_1^{\mathrm{p}} - \bar{R}_1^{\mathrm{e}}$. A part $\bar{R}_1^{\mathrm{p}}$ of the open message rate is used for secret-key generation while a part $\bar{R}_1^{\mathrm{e}}$ is used to transmit an encrypted message with that same key. Notice that the constraint $R_1^{\mathrm{o}} \geqslant \bar{R}_1^{\mathrm{p}} + \bar{R}_1^{\mathrm{e}}$ must be satisfied.
- $\tilde{R}_1^{\mathrm{e}} = R_1^{\mathrm{e}} + \bar{R}_1^{\mathrm{e}}$, thus increasing the encrypted message rate thanks to the secret-key generation mechanism.

Alice's total secure communication rate is then $\tilde{R}_1^{\mathrm{s}} + \tilde{R}_1^{\mathrm{e}}$. Bob modifies his rates in a similar way.

Even if the DMS was characterized exactly, the secret-key capacity of a source with rate-limited public communication is not known. Therefore, one may consider a suboptimal secret-key agreement strategy in which a single user sacrifices a fraction $\bar{R}^{\mathrm{p}}$ of its open message rate for communication and a single user uses the secret-key generated from the source for encryption.

Table 3.1 – *Additional constraints for secret-key generation.*

	Public Communication	
Encryption	**Alice**	**Bob**
Alice: $\bar{R}_1^{\mathrm{e}} = \bar{R}_1^{\mathrm{k}},\ \bar{R}_2^{\mathrm{k}} = 0$	$\bar{R}_1^{\mathrm{p}} + \bar{R}_1^{\mathrm{k}} \leqslant R_1^{\mathrm{o}}$	$\bar{R}_1^{\mathrm{p}} \leqslant R_2^{\mathrm{o}},\ \bar{R}_1^{\mathrm{k}} \leqslant R_1^{\mathrm{o}}$
Bob: $\bar{R}_2^{\mathrm{e}} = \bar{R}_2^{\mathrm{k}},\ \bar{R}_1^{\mathrm{k}} = 0$	$\bar{R}_2^{\mathrm{p}} \leqslant R_1^{\mathrm{o}},\ \bar{R}_2^{\mathrm{k}} \leqslant R_2^{\mathrm{o}}$	$\bar{R}_2^{\mathrm{p}} + \bar{R}_2^{\mathrm{k}} \leqslant R_2^{\mathrm{o}}$

In this case, the optimal secret-key generation rate $\bar{R}^{\mathrm{k}}$ that can be distilled with rate-limited public communication is known [24, Theorem 2.6]:

$$\bar{R}^{\mathrm{k}} < \mathbb{I}\big(V;\tilde{X}|U\big) - \mathbb{I}\big(V;\tilde{Z}|U\big),$$

where the random variables U and V are such that $U \to V \to \tilde{Y} \to \tilde{X}\tilde{Z}$ forms a Markov chain and

$$\mathbb{I}\big(V;\tilde{Y}\big) - \mathbb{I}\big(V;\tilde{X}\big) \leqslant \bar{R}^{\mathrm{p}}.$$

Additional constraints on $\bar{R}^{\mathrm{p}}$ and $\bar{R}^{\mathrm{k}}$ exists, depending on which user performs public discussion and which user encrypts a message. Specifically, there exist four possible configurations presented in Table 3.1.

3.4 Gaussian Two-Way Wiretap Channel

This section focuses on the evaluation of some achievable regions based on the strategies developed in previous sections for the Gaussian two-way wiretap channel. The relationships between inputs and outputs are

$$\begin{cases} Y_1^n = \sqrt{g_1}X_1^n + X_2^n + N_{21}^n, \\ Y_2^n = X_1^n + \sqrt{g_2}X_2^n + N_{12}^n, \\ Z^n = \sqrt{h_1}X_1^n + \sqrt{h_2}X_2^n + N_e^n. \end{cases} \tag{3.34}$$

In addition, the inputs are subject to the power constraints:

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\Big(\big(X_1^{(i)}\big)^2\Big) \leqslant \rho_1 \text{ and } \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\Big(\big(X_2^{(i)}\big)^2\Big) \leqslant \rho_2.$$

The additive noises N_{21}^n, N_{12}^n, and N_e^n are supposed i.i.d. zero-mean unit-variance Gaussian vectors. Cooperative jamming is performed by prefixing an additive white Gaussian noise (AWGN) channel before each input of WTC_2. This is equivalent to considering a modified channel with inputs C_1 and C_2, with:

$$X_1^n = C_1^n + N_{11}^n \quad \text{and} \quad X_2^n = C_2^n + N_{22}^n, \tag{3.35}$$

where $N_{11}^n \sim \mathcal{N}(0, \rho_1^n \mathbf{I}_n)$ and $N_{22}^n \sim \mathcal{N}(0, \rho_2^n \mathbf{I}_n)$ correspond to the noise introduced by users to realize cooperative jamming. For $\rho_1^n \leqslant \rho_1$ and $\rho_2^n \leqslant \rho_2$, the modified power constraints are

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left(\left(C_k^{(i)}\right)^2\right) \leqslant \rho_k - \rho_k^n = \rho_k^c,\ k = 1, 2.$$

With minor modifications of the proof presented in Section 3.2 to account for the power constraints, an achievable region for the Gaussian two-way wiretap channel can be obtained by substituting the random variables $C_1 \sim \mathcal{N}(0, \rho_1^c)$ and $C_2 \sim \mathcal{N}(0, \rho_2^c)$ in the bounds obtained earlier.

3.4.1 Randomness Source Extraction

Key extraction requires an explicit characterization of the DMS induced by cooperative jamming. Define:

$$\begin{cases} \tilde{X}_1 = N_{11},\ \tilde{Y}_1 = \tilde{X}_2 + N_{21}, \\ \tilde{X}_2 = N_{22},\ \tilde{Y}_2 = \tilde{X}_1 + N_{12}, \\ \tilde{Z} = \sqrt{h_1}\,\tilde{X}_1 + \sqrt{h_2}\,\tilde{X}_2 + N_e. \end{cases} \tag{3.36}$$

Proposition 3.8 The triple $(\tilde{X}, \tilde{Y}, \tilde{Z})$ with $\tilde{X} = (\tilde{X}_1, \tilde{Y}_1)$ and $\tilde{Y} = (\tilde{X}_2,\ \tilde{Y}_2)$ is an independent DMS that can be used for secret-key generation. Alice, Bob and Eve observe the components $\tilde{X}$, $\tilde{Y}$, and $\tilde{Z}$, respectively. ◊

Proof Since Alice generates the noise N_{11}, she obtains $\tilde{X}_1$ directly. She also observes the channel output $Y_1^n = C_2^n + N_{22}^n + N_{21}^n + \sqrt{g_1}\,X_1^n$; since she knows X_1^n and can decode C_2^n with high probability, she can obtain $\tilde{Y}_1$. Similarly, Bob can obtain $\tilde{X}_2$ and $\tilde{Y}_2$.

Eve observes the channel output $Z^n = \sqrt{h_1}\,X_1^n + \sqrt{h_2}\,X_2^n + N_e^n$ and, as discussed in Section 3.3,

she obtains C_1^n and C_2^n as side-information. Therefore, Eve can compute $\tilde{Z}$, but one needs to show that $\tilde{Z}^n$ is a sufficient statistic for $\tilde{X}^n$ and $\tilde{Y}^n$ given (Z^n, C_1^n, C_2^n). This is the case since

$$\mathbb{H}\big(\tilde{X}^n, \tilde{Y}^n | Z^n C_1^n C_2^n\big) \overset{\text{(a)}}{=} \mathbb{H}\big(\tilde{X}^n, \tilde{Y}^n | \tilde{Z}^n C_1^n C_2^n\big) \overset{\text{(b)}}{=} \mathbb{H}\big(\tilde{X}^n, \tilde{Y}^n | \tilde{Z}^n\big), \tag{3.37}$$

because (a): $ZC_1C_2 \mapsto \tilde{Z}C_1C_2$ is bijective and (b): $(\tilde{X}, \tilde{Y}, \tilde{Z})$ is independent of (C_1, C_2). ∎

The DMS $(\tilde{X}, \tilde{Y}, \tilde{Z})$ is a vector source, for which there is no known closed-form expression for the secret-key capacity with rate-limited public-rate communication. However, if Alice and Bob ignore one of their observations, the DMS reduces to a degraded scalar source. By [103, Corollary 2], there exists a closed-form function f

$$f : R_p \mapsto \frac{1}{2} \log \frac{\mathrm{Var}(Y|XZ)e^{-2R_p} + \mathrm{Var}(Y|Z)(1 - e^{-2R_p})}{\mathrm{Var}(Y|XZ)}, \tag{3.38}$$

which relates the public communication rate to the secret-key generation rates.

- If Alice ignores $\tilde{Y}_1$ and Bob ignores $\tilde{X}_2$, $\tilde{Y}_2 \to \tilde{X}_1 \to \tilde{Z}$: if Bob sends public messages at rate $\bar{R}_p$, Alice and Bob can distill a secret-key at rate $\bar{R}_k$, such that $\bar{R}_k \leqslant f\left(\bar{R}_p, p_{\tilde{Y}_2\tilde{X}_1\tilde{Z}}\right)$.
- If Bob ignores $\tilde{Y}_2$ and Alice ignores $\tilde{X}_1$, $\tilde{Y}_1 \to \tilde{X}_2 \to \tilde{Z}$: if Alice sends public messages at rate $\bar{R}_p$, Alice and Bob can distill a secret-key at rate $\bar{R}_k$, such that $\bar{R}_k \leqslant f\left(\bar{R}_p, p_{\tilde{Y}_1\tilde{X}_2\tilde{Z}}\right)$.

With the explicit characterization of the DMS induced by cooperative jamming and the function f, it is possible to compute new achievable rates for the strategy described in Section 3.3. Although there is no known closed-form expression for the resulting region $\bar{\mathscr{R}}^{\mathrm{K}}$, it is possible to prove that the region strictly includes the region $\bar{\mathscr{R}}^{\mathrm{F}}$ in Proposition 3.5 strengthened for strong secrecy.

Proposition 3.9 Consider $\bar{\mathscr{R}}^{\mathrm{F}}$, $\bar{\mathscr{R}}^{\mathrm{K}}$, and $\bar{\mathscr{R}}^{\mathrm{2W}}$ defined as before. In general: $\bar{\mathscr{R}}^{\mathrm{F}} \subseteq \bar{\mathscr{R}}^{\mathrm{K}} \subseteq \bar{\mathscr{R}}^{\mathrm{2W}}$. There exist channels such that $\bar{\mathscr{R}}^{\mathrm{F}} \subset \bar{\mathscr{R}}^{\mathrm{K}}$.

Proof The first statement follows directly from the definition of the secret-key generation strategy in Section 3.3. For the second statement, consider the example of a channel for which:

$$0 < \rho_1 < \frac{h_2 - 1}{h_1} \text{ and } 0 < \rho_2 < \frac{h_1 - 1}{h_2}.$$

In this case $\mathbb{I}(C_1; Y_2|X_2) < \mathbb{I}(C_1; Z)$ and $\mathbb{I}(C_2; Y_1|X_1) < \mathbb{I}(C_2; Z)$, thus the region $\bar{\mathscr{R}}^{\mathrm{F}}$ is empty. However, since the power constraints are non-zero, a fraction of the power can be used to induce a DMS while still maintaining a positive auxiliary message rate. By [103, Theorem 3], it is possible to obtain a non-zero secret-key rate and, therefore, a non-zero secure communication rate. ∎

Remark Although the proof yields $\bar{\mathscr{R}}^{\mathrm{F}} \subset \bar{\mathscr{R}}^{\mathrm{K}}$ by exhibiting a dummy example, the numerical results in the next section clearly confirm improvements in various cases.

3.4.2 Results

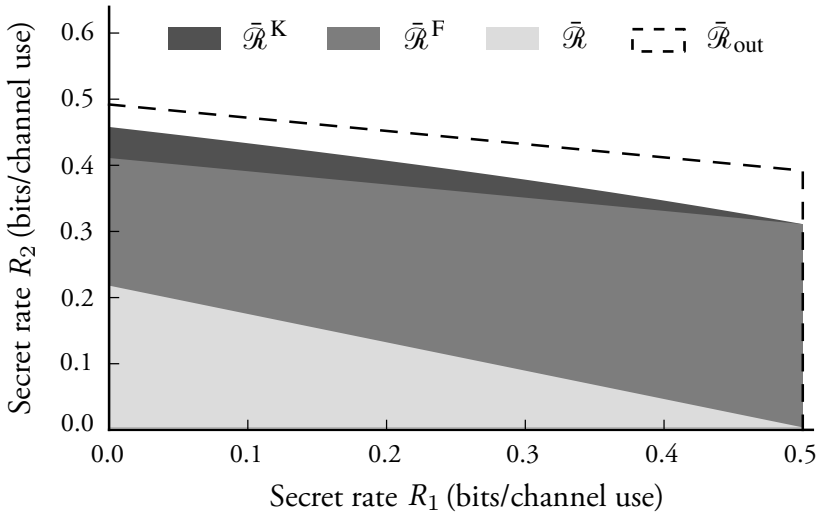

Figure 3.5 – *Region evaluation for* $g_1 = g_2 = 1$, $\rho_1 = 1$, $\rho_2 = 100$, $h_1 = 1$, $h_2 = 0.1$.

The first situation considers that the secret-key generation provides little gain because cooperative jamming leaves little room for improvement. Figure 3.5 illustrates the achievable region with simultaneous cooperative jamming and secret-key exchange $\bar{\mathscr{R}}^{\mathrm{F}}$, and the achievable region that also includes secret-key generation $\bar{\mathscr{R}}^{\mathrm{K}}$ obtained for $\rho_1 = 1$, $\rho_2 = 100$, $h_1 = 1$, $h_2 = 0.1$, and $g_1 = g_2 = 1$. The upper left corner of the region corresponds to a situation in which Alice uses all her power to jam ($\rho_1^n = \rho_1$) while Bob uses all his power to transmit ($\rho_2^c = \rho_2$). A secret-key can be extracted from the source induced by Alice; however, since $h_1 = 10$, Eve's

source observation is highly correlated to Alice's, which severely limits secret-key rates. The upper right corner of the region corresponds to a situation in which both Alice and Bob use their entire power to transmit (i.e., $\rho_1^c = \rho_1$ and $\rho_2^c = \rho_2$) and, therefore, no source is induced. For comparison, the figure also illustrates the region $\bar{\mathcal{R}}$ obtained using secret-key generation alone (no cooperative jamming). Although this strategy is clearly suboptimal, note that there exist practical coding schemes for secret-key generation over Gaussian channels; therefore, the region $\bar{\mathcal{R}}$ provides an estimate of rates achievable with current codes. These results can be compared to the best-known outer region $\bar{\mathcal{R}}_{\text{out}}$ for the two-way wiretap channel computed by He and Yener [43].

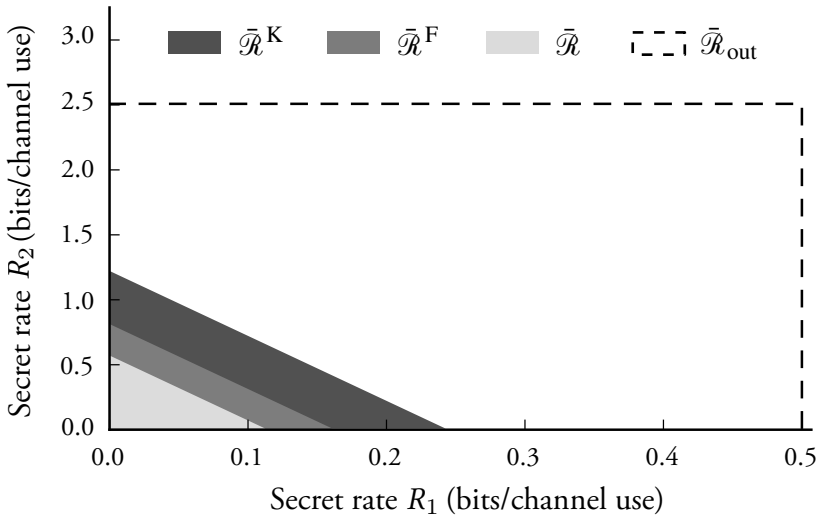

Figure 3.6 – *Region evaluation for* $g_1 = g_2 = 1$, $\rho_1 = \rho_2 = 1$, $h_1 = h_2 = 1.5$.

Figure 3.6 illustrates the regions $\bar{\mathcal{R}}^{\mathrm{F}}$ and $\bar{\mathcal{R}}^{\mathrm{K}}$ obtained for $\rho_1 = \rho_2 = 1$, $h_1 = h_2 = 1.5$, and $g_1 = g_2 = 1$. This corresponds to a situation in which Eve obtains a better observation than either Alice or Bob and in which Alice and Bob have little power available. In such a case, secret-key generation provides a significant improvement.

Finally, Figure 3.7 illustrates the region $\bar{\mathcal{R}}^{\mathrm{F}}$ and $\bar{\mathcal{R}}^{\mathrm{K}}$ obtained for $\rho_1 = \rho_2 = 0.9$, $h_1 = h_2 = 10$, and $g_1 = g_2 = 1$. In this case $\bar{\mathcal{R}}^{\mathrm{F}} = \varnothing$, but secret-key generation is possible.

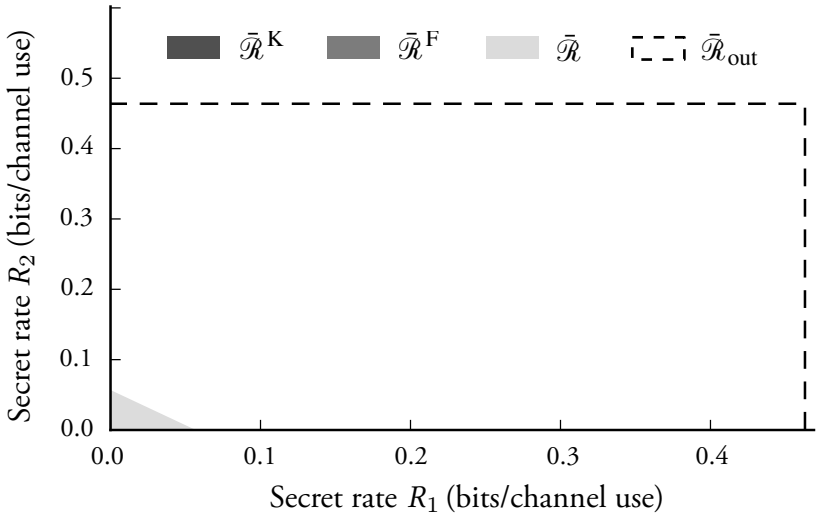

Figure 3.7 – *Region evaluation for* $\rho_1 = \rho_2 = 0.9$, $h_1 = h_2 = 10$, $g_1 = g_2 = 1$.

3.5 Conclusion and Discussion

This chapter provides an achievable region of strongly secure rates based on strategies that partially decouple the use of interference and feedback for secrecy. The coded cooperative strategy exploits the interference, while secret-key exchange and secret-key generation exploit the feedback. This combined approach shows significant improvements in the Gaussian case. More recently, several results have considered new models of multi-user communications. For instance in [46], the authors consider both full-duplex communications and half-duplex communications with an untrusted relay. For the first model, not exploiting the feedback induces an unbounded loss in secrecy rate in some cases. For the latter model, the penalty is negligible if the relay power is unlimited. The authors derive an achievable region with a simple cooperative jamming scheme that does not exploit feedback while showing near-optimal performances.

CHAPTER 4

EXPERIMENTAL ASPECTS OF SECRET-KEY GENERATION[1]

As seen in the previous chapters, *physical-layer security* promises new ways of providing secrecy through the utilization of the intrinsic randomness present in any communication medium, such as noise and interferences. Although the theory behind physical-layer security has been extensively studied, designing and implementing a physical-layer security system in a wireless environment remains a challenge. Without additional experimental validation, there is a risk that the models considered in theoretical works may be fairly disconnected from real systems since they rely on assumptions that cannot be met in a real wireless setting.

Table 4.1 – *Literature comparison.*

Reference	Experiments	Statistics	Security Analysis
[100, 101, 111]	No	Postulated or N/A	Non information-theoretic
[19, 63, 98]	Simulations		
[52, 72, 79, 80]	Yes		
[61]		$\mathbb{I}(X;Z) = 0$ because of decorrelation	
[49, 99, 105, 109]		Estimated from experimental measurements	Asymptotic information-theoretic
Proposed	Yes	Estimated from experimental measurements	Finite Length

As summarized in Table 4.1, several works have already experimentally investigated the generation of secret-keys from wireless channels. In fact, the gains of wireless channels provide a natural source of randomness, for which reciprocity guarantees that legitimate users obtain strongly correlated channel observations, while diversity ensures that the observations of

[1]Parts of the material in this chapter have appeared in [75]: **Pierrot, A. J.**, Chou, R. A., Bloch, M. R., "Experimental Aspects of Secret Key Generation in Indoor Wireless Environments". In: *Proceedings of the 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. June 2013, pp. 669–673. ©IEEE 2013.

a third-party eavesdropper disclose little information about the legitimate users' measurements. However, while these works are often motivated by an information-theoretic formulation, the information-theoretic model is not fully developed. In particular, common assumptions about the eavesdropper's statistics, such as the decorrelation of channel gains at distances larger than half the wavelength, as well as the use of asymptotic values of secret-key generation, lead to over-simplifications of the protocols and over-estimations of the achievable information-theoretic secret-key rates. Consequently, while canonical theoretical models of wireless channels have proved incredibly useful to design reliable communication systems, their use for the design of secret-key generation systems requires more care. Similarly, checking that generated keys pass statistical tests [6]–[9], which have been primarily designed for mathematical cryptography and only verify some desirable statistical property of the key, does not guarantee information-theoretic secrecy.

The objective of this chapter is to investigate the practical effect of eavesdropper's statistics by implementing a secret-key generation system from wireless channel gain with software-defined radios, and by carrying out a careful information-theoretic analysis. The weakness of previously reported system lies in the modeling of the source of randomness, but not in the operation of the subsequent protocol; therefore, this chapter does not attempt to develop a complete secret-key generation system. The results obtained in this chapter yields the following conclusions:

1. assuming that the eavesdropper does not get any information because of decorrelation with distance is not exact in a real wireless system;

2. the existence of a correlated eavesdropper's observation makes the evaluation of the finite-length secret-key rates much more intricate.

Section 4.1 recalls the basic principles of secret-key generation from channel variations, with the underlying mathematical formalism and assumptions. Section 4.2 describes the experimental setup used to induce a source of randomness by leveraging the variations of the channel

gains. Section 4.3 assesses the robustness of the scheme regarding the diversity assumption. Section 4.4 presents an achievable secret-key rate with a finite number of samples, which is then evaluated with the experimental measurements. Finally, Section 4.5 provides some conclusions and discussions regarding the possible limitations for the design of practical physical-layer security systems with an example of potential application.

4.1 Key Generation from Channel Variations

This section formalizes a generic process that extracts a source of randomness from the channel and then describes how the secret-key is generated from such a source.

4.1.1 Secret-Key Generation Strategy

4.1.1.1 Randomness in Wireless Channels

In a wireless channel with impulse response $h(t)$, the signal $i(t)$ emitted by a terminal and the corresponding signal $o(t)$ received at another are related as

$$o(t) = (h * i)(t) + w(t).$$

The thermal noise $w(t)$ appears at the receivers because of the thermal agitation in the electronic circuits. This noise is highly random, unpredictable, and has a short coherence time. It is not suitable for secret-key generation since there is no correlation among the thermal noises at different terminals.

The impulse response $h(t)$ of a wireless channel between two terminals results from the reflections and attenuations underwent by the transmitted signal along different paths. This chapter focuses on narrowband channels with approximately 1 MHz bandwidth, for which the received signal is essentially a delayed version of the original one attenuated by a random complex gain $G(t)\exp(\jmath\phi(t))$. This complex gain accounts for the aggregated effect of attenuation and phase change of each path; note that reciprocity guarantees that the gain $G_{AB}(t)$ between two points A and B is the same as the gain $G_{BA}(t)$ between the points B and A. The coherence time during which the gain $G(t)$ and the phase $\phi(t)$ remain constant scales approximately as $T_c \approx \lambda/v$, where λ is the wavelength and v is the characteristic speed of the environment. For

instance, in the experimental setup described later, the objects around the receiver move at about one meter per second, so that the coherence time is on the order of milliseconds for wireless communications at 2.5 GHz. It is easier to exploit the randomness of the channel gain $G(t)$ since precise measurements of the phase $\phi(t)$ require a precise synchronization of the terminals.

4.1.1.2 Randomness Extraction

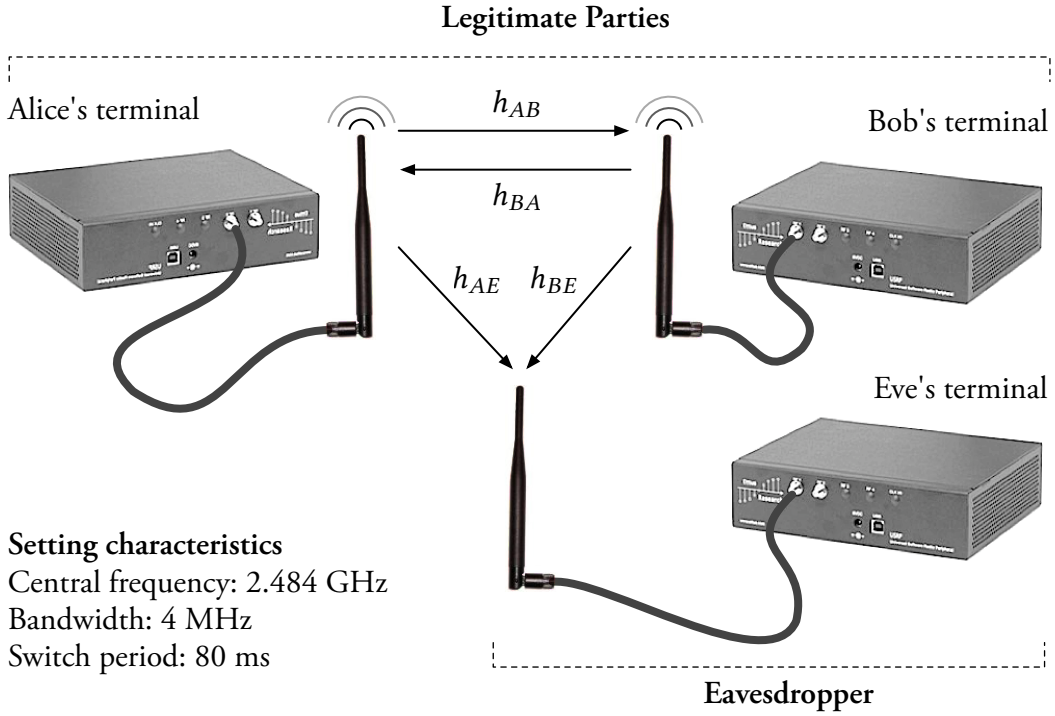

Figure 4.1 – *Experimental testbed with software-defined radios.*

Following common practice, and as illustrated in Figure 4.1, the channel gains between the two legitimate terminals are measured in Time-Division-Duplex mode as follows. The first terminal, Alice, sends a complex probe signal $b(t)$ with unit energy, whose duration β is much smaller than the coherence time T_c, so that the channel gain G remains almost constant over the pulse duration. The second terminal, Bob, measures from the channel a delayed and faded version $\tilde{b}(t)$ of $b(t)$. Using his knowledge of the probe signal, Bobs matches $\vec{b}(t) = \tilde{b}(t + T_d)$ and $b(t)$ as

$$T_d = \text{argmax}_\tau(\vec{b} * b)(\tau). \tag{4.1}$$

Then, Bob estimates the channel gain from Alice to Bob as

$$G_{AB} = \sqrt{\int_{\beta} \left|\vec{b}(t)\right|^2 \mathrm{d}t}, \tag{4.2}$$

by measuring the energy of the delayed received probe signal. Simultaneously, an eavesdropper, Eve, obtains a channel gain G_{AE}. Alice estimates the channel gain G_{BA} in a similar fashion from Bob's probe signal, which also provides Eve with a channel gain G_{BE} in the process.

4.1.2 Mathematical Formalism

Once n measurements are performed, Alice, Bob, and Eve, effectively observe the components of a noisy source $(\mathcal{X}^n \mathcal{Y}^n \mathcal{Z}^n, p_{X^n Y^n Z^n})$, in which X^n consists of n channel gains G^n_{BA}, Y^n consists of the channel gain G^n_{AB}, and Z^n consists of both sequences G^n_{AE} and G^n_{BE}. A secret-key generation strategy $\mathcal{S}_n$ for the source $(\mathcal{X}^n \mathcal{Y}^n \mathcal{Z}^n, p_{X^n Y^n Z^n})$ with unlimited public communications consists of the following operations.

- *Reconciliation:* Alice transmits a public message F over the public authenticated channel, which allows Bob to construct an estimate $\hat{X}^n$ of X^n from Y^n and F.

- *Privacy amplification:* Alice chooses a function G uniformly at random in a family of universal$_2$ hash functions, which is disclosed to all parties. Alice then computes $G(X^n) \in \mathcal{K}$ while Bob computes $G(\hat{X}^n)$. Setting $\mathcal{K} \triangleq [\![1, 2^{nR}]\!]$, R is called the *secret-key rate*.

In principle, Alice and Bob could interactively exchange messages, but this strategy is restricted to unidirectional operation. The secret-key generation strategy $\mathcal{S}_n$ must ensure the following:

1. *reliability*, measured with the probability of disagreement

$$\mathbf{P}_{\mathrm{d}}(\mathcal{S}_n) \triangleq \mathbb{P}(K \neq \hat{K} | \mathcal{S}_n);$$

2. *(strong) secrecy*, measured by the *leakage*

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

3. *(strong) uniformity*, measured by

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

Computing the aforementioned metrics requires the knowledge of source statistics, including that of the eavesdropper.

A secret-key rate $R \triangleq \frac{1}{n}\log|\mathcal{K}|$ is achievable if the three above metrics tend to zero as n goes to infinity, and the supremum of achievable secret-key rates is called the *secret-key capacity* C_s. Most recent works have focused on the calculation of C_s, which is an asymptotic limit obtained for infinitely many realizations of the source; in contrast, the analysis conducted in Section 4.4 focuses on a finite length behavior that only requires $\mathbf{L}(\mathcal{S}_n)$ and $\mathbf{U}(\mathcal{S}_n)$ to be small, but non-zero.

4.1.3 Assumptions behind the secret-key generation model

The secrecy guaranteed by a secret-key generation strategy rely on three common assumptions.

Availability of an authenticated public channel of unlimited capacity This assumption is not unreasonable if one aims at generating low secret-key rates for which the amount of public communication is negligible compared to the channel capacity. If one explicitly introduces a rate limitation, reconciliation with vector quantization can be used [21, 24] without fundamentally affecting the operation of the secret-key agreement strategy.

Existence of enough randomness Mobility in the environment is required to ensure that wireless channel gains have enough entropy. Mobility results from the movements of objects around the terminals or the terminals themselves; in indoor wireless environments, this channel gains experience variability as soon as people move around the communication terminals.

Knowledge of eavesdropper's statistics In an information-theoretic secret-key generation model, one requires the knowledge of the statistical dependencies between Eve's observations and the legitimate users' to assess the secrecy of the keys. Unfortunately, there exists no indirect way to estimate these statistical dependencies of the eavesdropper without performing measurements at Eve's terminal position. In addition, as pointed out in [52], the statistics

should not be influenced by the eavesdropper to prevent the induction of artificial, deterministic, and predictable variations of the channel parameters. In principle, if Eve only has partial control of the environment, any uncontrolled movement would suffice to induce variations she cannot predict. This assumption might be even more easily satisfied in wideband and MIMO systems since it would require more advanced capabilities from the attacker.

The knowledge of the eavesdropper's statistics is the most crucial assumption for the proper operation of a secret-key generation system. This could be avoided by operating in a quantum setting, e.g. [104], but such systems are only efficiently implemented in optics. In the classical wireless setting, the assumption is often circumvented by assuming that there exists enough *diversity* in the environment, so that one can either assume that $\mathbb{I}(G_{AB}; G_{AE}G_{BE}) = 0$ meaning the eavesdropper's observations are completely independent of the legitimate users', or, at least, that $\mathbb{I}(G_{AB}; G_{AE}G_{BE})$ is upper bounded. However, it is crucial to precisely assess the conditions under which the diversity assumption may hold, so as to define situations in which secret keys can be safely generated. Moreover, it is considerably easier to analyze secrecy assuming that $\mathbb{I}(G_{AB}; G_{AE}G_{BE}) = 0$ and that the eavesdropper only observes public communication. Privacy amplification and reconciliation are simply linked using the result of Cachin and Maurer [18], and counting the number of bits disclosed during privacy amplification is sufficient to establish the final key length. In contrast, when $\mathbb{I}(G_{AB}; G_{AE}G_{BE}) \neq 0$, the final secret key length depends on the eavesdropper's statistics and one must factor in the effect of statistical deviations from the mean when using a finite number of samples n.

4.2 Experimental Source Induction

This section describes the experimental setup and the procedure to characterize the statistics of the wireless channel gains. The gain measurements are reported for a source that is induced using a communication chain representative of a typical wireless system.

4.2.1 Setup Description

The system features three USRPs (see Appendix 7.2 for more details) with XCVR2450 daughterboards that operate in the 2.5 GHz and 5 GHz bands, typically used for Wi-Fi communications. However, the bandwidth is not wide enough (8 MHz maximum) to allow for actual Wi-Fi communications according to the IEEE standards, which would require over 20 MHz of bandwidth. Consequently, the reported secret-key rate is likely to be much smaller than what could be obtained on top of an IEEE802.11 transmission. RF signals are transmitted using standard Wi-Fi antennas with a transmission power below 100 mW.

The experiments are conducted in two ordinary office rooms representative of an indoor environment: one is the former Arcom wireless communication laboratory at Georgia Tech Lorraine, and the other is a conference room. The choice of which room used for the experiments was only motivated by convenience.

All radios are connected to a single computer that processes the transmission and reception data stream. There is no external hardware synchronization between the terminals, and the same configuration is used for all terminals, both in hardware and software. The first samples of the data stream are used to overcome hardware discrepancies by scaling all measurements to obtain the same received energy. After calibration, the scaling is kept constant through each experiment since no significant drift was observed during acquisition.

4.2.2 Communication Chain

All experiments were conducted using the three-user setup represented in Figure 4.1. The modulation frequency is 2.484 GHz, which corresponds the 14^{th} WLAN channel. This channel is not used in for Wi-Fi communications and does not interfere with other Wi-Fi channels. Since XCVR2450 daughterboards are limited to half-duplex operation, the channel gains cannot be measured simultaneously. This limitation occurs in a majority of communication systems since full-duplex communication requires complex hardware and software implementation [10]. To circumvent the problem, the radios are continuously switching between the Rx and Tx modes. Because of further hardware restrictions, the minimum half-duplex cycle time is limited to

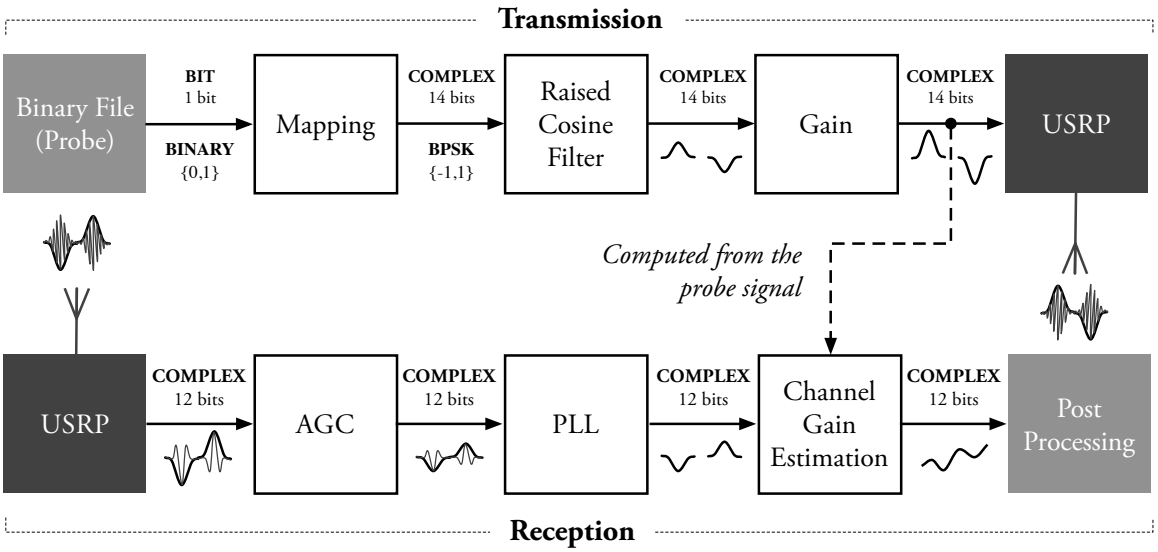

Figure 4.2 – *Communication chain for channel gain estimation.*

80 ms. Consequently, motion in the environment is limited to 1 $m.s^{-1}$ so that the channel gain would not vary much between an Rx/Tx switch, thus maintaining channel reciprocity. If the hardware allowed faster half-duplex cycle to capture faster fades, higher secret-key generation rates would be achieved in a high mobility environment, but the security analysis would remain essentially the same.

4.2.3 Characterization of Induced Source Statistics.

The estimation of the channel gains is performed using a probe message sent through the communication chain described in Figure 4.2. The gain present in the transmission chain allows power control and is kept constant throughout the entire duration of the experiment. During the reception, the USRP performs demodulation and analog-to-digital conversion. An automatic gain controller (AGC) scales the received signal to match the optimal range of the subsequent processing block. Note that it is tuned to be slow enough not to remove the gain variations over the timescale of interest. Because the system operates at a high carrier frequency, a phase-locked loop (PLL) is used to suppress any residual modulation resulting from minor differences between modulation and demodulation frequencies. The demodulated signal is then used to compute the transmission gain. Note that all parties know the probe signal and that the

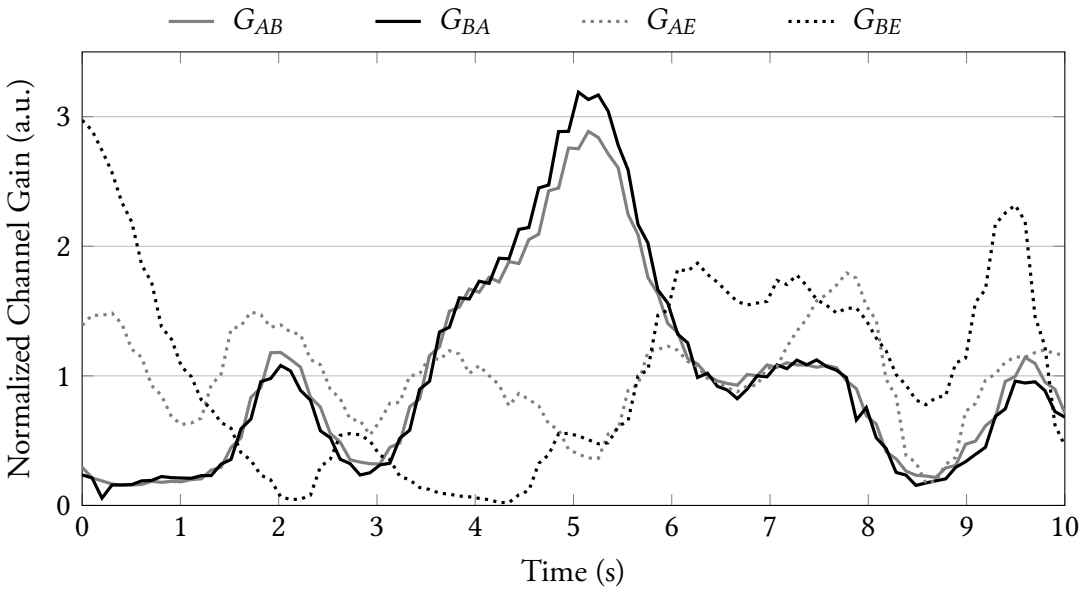

Figure 4.3 – *Channel gain measurements.*

transmission chain behavior is entirely deterministic so that all users also know the shaped signal and can compute the channel gain. The probe signal, which is a fixed randomly-generated sequence, is also used to synchronize the different radios in software.

A total of 500 gain measurement experiments were acquired, each lasting approximately ten seconds. Alice and Bob's terminals were separated by 1.5m and Eve's terminal was approximately 1m away from both Alice and Bob. Figure 4.3 illustrates the evolution of the following four channel gains: the channel gain G_{AB} from Alice to Bob, its reverse G_{BA} from Bob to Alice, the channel gain G_{AE} from Alice to Eve, and the channel gain G_{BE} from Bob to Eve. It appears that G_{AB} closely follows G_{BA}, confirming the existence of channel reciprocity. Eve's channel gains G_{AE} and G_{BE} are seemingly not related to the channel gains G_{AB} and G_{BA}, potentially confirming the existence of channel diversity. According to Jake's model, diversity should hold as soon as Eve is farther from Alice and Bob than the coherence distance, which is $\ell_c = \lambda/2 \approx$ 6 cm, at 2.484 GHz. The next section analyzes diversity more precisely.

It is desirable to operate on a *memoryless* source of randomness to make the statistical characterization tractable. Decimating the raw measurements in Fig. 4.3 is a way to remove the

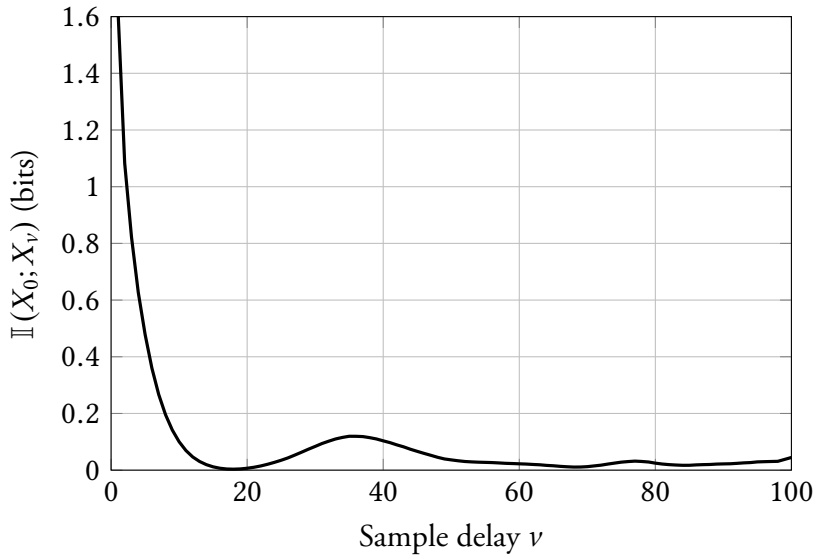

Figure 4.4 – *Evolution of the self-correlation of channel gains.*

time correlation by keeping a single sample per coherence interval; since the coherence time is on the order of magnitude of λ/v, T_c should be on the order of one second. It is possible to obtain a more precise characterization of the value of T_c with an estimation of the mutual information $\mathbb{I}(X_0; X_\nu)$ between a sample X_0 and the ν-th next sample X_ν, obtained by viewing each of the experimental time series as the realization of the same ergodic random process. The lower $\mathbb{I}(X_0; X_\nu)$ is, the less dependent the samples are. To use more samples for the estimation, the gains are supposed wide sense stationary, which was empirically confirmed by verifying that the quantity $\mathbb{I}(X_0; X_\nu)$ remained the same for different choices of X_0. Unless mentioned otherwise, all information metrics are estimated with the technique presented in [73]. As illustrated in Figure 4.4, the mutual information $\mathbb{I}(X_0; X_\nu)$ decays rapidly and vanishes after a dozen samples, corresponding to approximately one second, as expected. Operating on the down-sampled measurements instead of the original measurements would result in a lower achievable secret-key rate, which might seem an unnecessary simplification since it is possible to characterize achievable secret-key rates for sources with memory [17, 19]. However, without an accurate parametric model, the estimation of the statistics of a source with memory turns out to be a much more difficult problem.

The final step is then to estimate the joint statistics p_{XYZ} of the memoryless source, which poses two challenges. First, one would in principle need to analyze the estimation error and include it in the subsequent calculation of achievable key rates; the results presented in this chapter do not take this into account and assume that the estimation is accurate enough to be used as the true joint statistics. Second, the measurements only provide *quantized* measurements X_Q, Y_Q and Z_Q of the true channel gains X, Y, and Z, respectively. The quantization of X and Y is not critical since it reduces achievable secret-key rates by only affecting the reconciliation step. However, the quantization of Z results in an underestimation of the eavesdropper's knowledge, therefore one should assume that the eavesdropper can keep Z continuous. Since it is not possible to acquire a continuous-valued Z with the software-defined radios, the first step consists in constructing a histogram corresponding to $p_{X_Q Y_Q Z_{Q'}}$ from the measured data, where $Z_{Q'}$ is a quantized version of Z. Then the histogram is interpolated with respect to $Z_{Q'}$ to obtain $p_{X_Q Y_Q Z}$. The raw data was acquired with a 14 bits resolution, which was further quantized to obtain a 4-bit resolution for X_Q and Y_Q, and a 6.5-bit resolution for $Z_{Q'}$.

4.3 Statistics of the Channel Gain Observations

4.3.1 Empirical Gain Distribution

The histogram of the values of the source is represented Figure 4.5. In a wireless environment with many diffractive and reflective objects, one can show that (see [94]) channel gains follow a Rician distribution of the form

$$f_R(r) = \frac{2(K+1)r}{\Omega} \exp\left(-K - \frac{(K+1)r^2}{\Omega}\right) I_0\left(2\sqrt{\frac{K(K+1)}{\Omega}}r\right),$$

with $(K, \Omega) \in \mathbb{R}_+^2$.

The parameter K is the ratio between the line of sight energy transmission and the reflected or diffracted paths energy. In this case, estimating the parameters K and Ω from the empirical moments [2] yields the distribution in Figure 4.5.

The distribution fitting yields the following parameters $K = 2{,}48$ and $\Omega = 1$. This value for K confirms a strong line of sight between the two terminals. Regarding Ω, its small value can be justified because of the small number of diffractive objects.

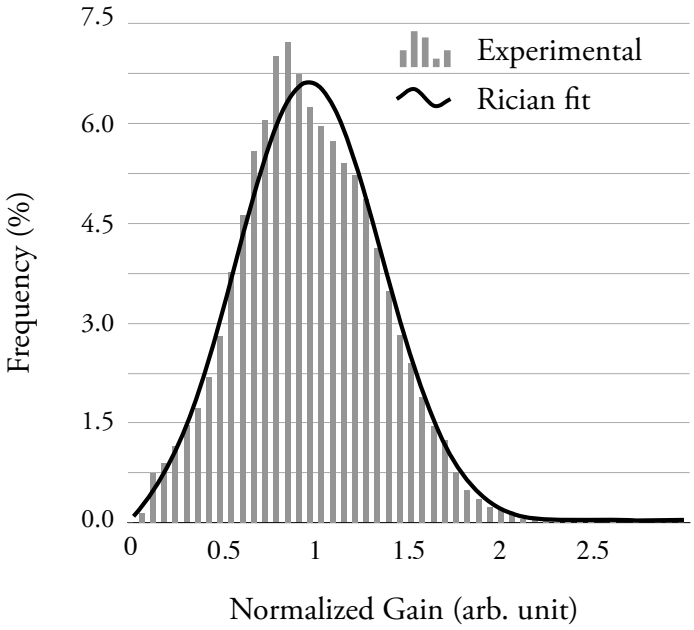

Figure 4.5 – *Gain distribution*

4.3.2 Robustness of the Diversity Assumption

Diversity is desirable to facilitate the generation of keys at a high rate so that the eavesdropper's observations are not correlated to legitimate users'. The ideal case would be a perfect statistical independence between these observations. However, even intuitively, it seems unlikely to happen; for instance, if Eve is close to either Alice or Bob, she may be able to obtain a strongly correlated observation of the channel gain. In an ideal case, the correlation would decay with distance, but this also depends on the configuration of the objects populating the environment.

A series of measurements[2] was conducted in the building conference room to verify whether diversity holds in a narrowband wireless setting. The room is about 40 square meters and is furnished as illustrated in Figure 4.6. Experiments were conducted off-hours to avoid any unwanted motion outside of the room. Two radios were placed in the middle of the room on the conference table two meters apart. A third radio terminal represents the eavesdropper, which was moved in the room across 60 possible positions. The channel gains G_{AB} and G_{AE} obtained by Bob and Eve is then used to evaluate information leakage. These experiments only

[2]Acknowledgment: Alan Dong for his invaluable help for those experiments.

involved one-way communications (Alice-to-Bob and Alice-to-Eve), thus avoiding the problem of half-duplex operation and allowing to gather data at a faster pace. Each experiment lasted one minute, yielding about 50,000 channels gain values at a 1 kHz sampling rate. The first and last hundred of samples were removed to eliminate artifacts when starting and stopping the radios.

The results of the measurement campaign are presented in Figures 4.6 and 4.7. Eve is placed across the positions indicated by the black "+" marks, which correspond to a coarse square grid of one meter and additional positions to cover interesting locations and the room borders. The brightness represents the *normalized secrecy-rate* (for $G_{AB} = G_{BA}$) between the gains obtained by Eve and those obtained by Bob, which is computed as

$$\begin{aligned} \text{Normalized secrecy-rate} &\triangleq \frac{\mathbb{I}(G_{AB}; G_{BA}) - \mathbb{I}(G_{AB}; G_{AE})}{\mathbb{I}(G_{AB}; G_{BA})} \\ &= 1 - \frac{\mathbb{I}(G_{AB}; G_{AE})}{\mathbb{H}(G_{AB})}, \end{aligned} \tag{4.3}$$

where G_{AB} and G_{AE} are the channel gains measured by Bob and Eve, respectively. This normalization compensates the entropy variations of the wireless channel gains across different experiments. This quantity is close to one (white) when the gains are independent, and equal to zero (black) when there is a one-to one mapping between G_{AB} and G_{AE}.

The first series of measurement, reported in Figure 4.6, is conducted without movement and serves as a benchmark. In this situation, there is no fluctuation of the channel gains, except those induced by the noise at the receivers' terminals. Therefore, the quantity $\mathbb{I}(G_{AB}; G_{AE})$ is small since the receiver noise is independent from one radio to another. When Eve and Alice use the same antenna, the darker spot comes from the electronic coupling between the terminals.

In a second series of measurements, represented in Figure 4.7, the operator is walking in the upper left corner of the room. High correlations appear when Bob and Eve's antennas are huddled together, and a fast decay of this correlation with distance, with leaked information reaching almost zero after a few centimeters. However, the leaked information increases again further away, even reaching values as high as 10% in the upper left corner. Since this corner

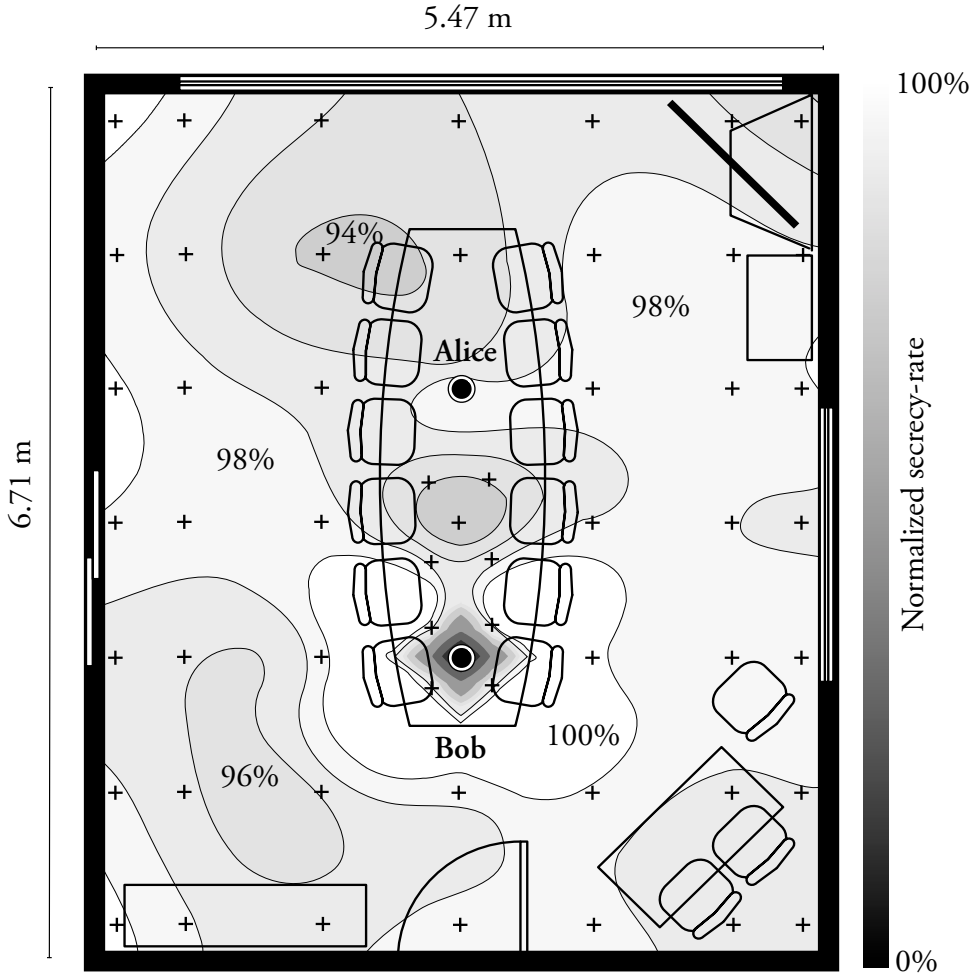

Figure 4.6 – *Normalized secrecy rate without motion.*

was actually the area where the operator was moving during the experiment, it suggests that measurements close to the motion source provide a better insight into the legitimate channel fluctuations. Therefore, defining a simple zone of guaranteed secrecy for key generation is not straightforward. From a security standpoint this proves that one cannot ignore the information leaked to the eavesdropper when channel variations come from the motion in the environment. It confirms that a secret-key generation *must* include a stage of privacy amplification to deal with unforeseeable levels of leaked information.

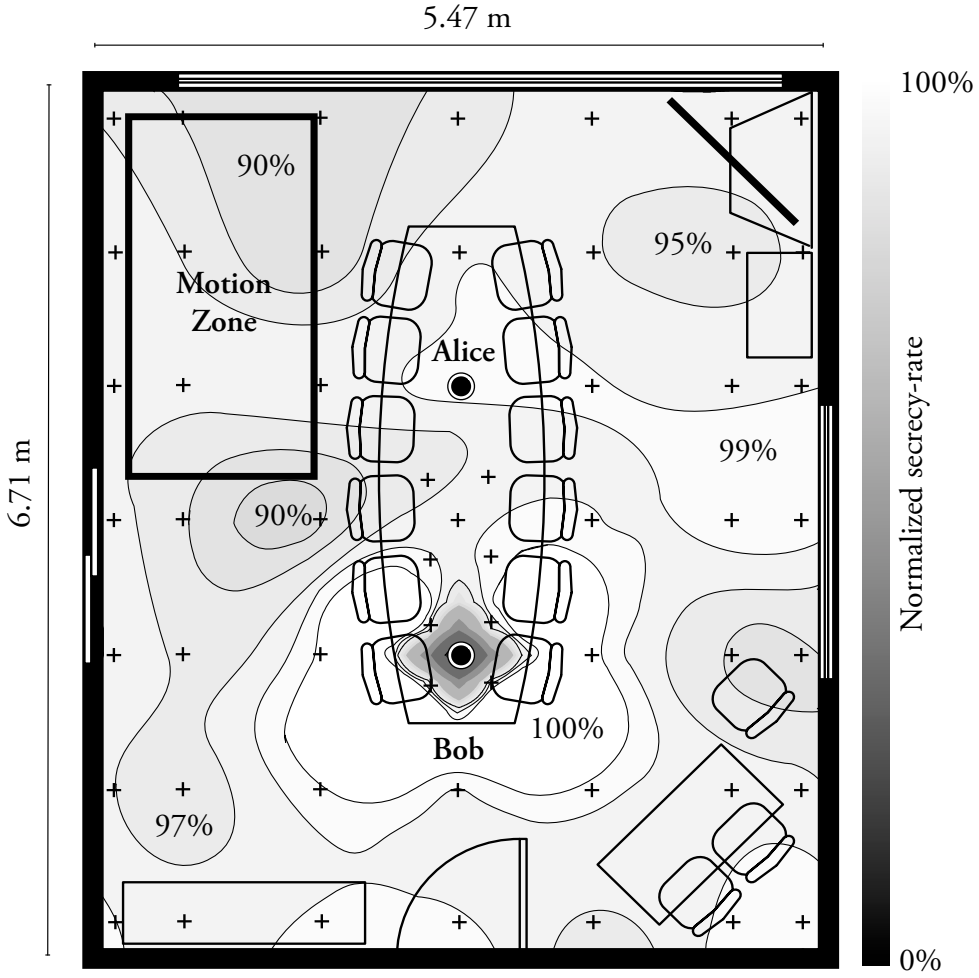

Figure 4.7 – *Normalized secrecy rate with motion.*

4.3.3 Environment Influence

The environment has a significant influence on the secret-key generation rate since the gain fluctuations used as a source of randomness are induced by external motion. The number of secret-key bits generated per second mainly depends on two factors: the entropy of the channel gain value provided by the source and the number of samples provided by the source per second. The latter depends on the coherence time of the channel since the decimation of the gain measurements relies on this value.

The experimental testbed used to analyze the influence of the environment is slightly different from previous experiments and consist of two radios placed on a cart to estimate the channel gain in one direction. The experiments were performed in the following situations:

1. the cart is moved in a hallway at a slow pace (about 2 km/h);
2. the cart is moved in a hallway at a normal pace (about 4 km/h);
3. the cart is moved in a hallway at a fast pace (about 6 km/h);
4. the cart is parked in the hallway with nobody moving around;
5. the cart is parked in the hallway between classes;
6. the cart is spun inside the building lobby;
7. the cart is parked outside;
8. the cart is moved outside at a fast pace (about 6 km/h);
9. one radio is shaken, while the other stays still.

The gain measurements are further filtered to remove the measurement noise and to solely keep the gain fluctuations, using a running average filter with a span of 20 ms, which is much lower than the typical time of the gain fluctuations for motions slower than 10 km/h.

Table 4.2 – *Gain measurement entropy and secret-key rate upper bound in various situations.*

Situation	1	2	3	4	5	6	7	8	9
Gain entropy (b.p.g.v.)	6.86	7.49	8.33	2.00	2.81	9.78	6.15	9.95	8.61
Coherence Time (ms)	510	328	181	18,904	9,474	322	4,230	57	148
Max secret-key rate (bit/sec)	13.4	22.8	46.0	0.11	0.30	30.4	1.45	175	58.1

Table 4.2 confirms that the gain entropy only depends on the presence of motion. The gain entropy is limited to 12 bits, which is the resolution of the analog-to-digital converters. The coherence time appears to be more sensitive to motion, which is what one would expect

from the direct relationship between the coherence time and the velocity of moving objects in the environment. This table also provides an upper bound on the secret-key bitrate by taking the ratio of the gain entropy over the coherence time, even though this may be far from what can be obtained considering the entire secret-key generation scheme with reconciliation and privacy amplification. As expected, the secret-key rates are higher when there is more mobility in the environment; bitrates are also higher in an outdoor environment, probably because the experiments conducted in the hallway involved a limited number of scatterers.

4.4 Secret-Key Generation in the Finite Blocklength Regime

Once the statistics $p_{X_Q Y_Q Z}$ of the source are characterized, one can easily compute asymptotic achievable secret-key rates $\mathbb{I}(X_Q;Y_Q)-\mathbb{I}(X_Q;Z)$. However, these rates may be far too optimistic when operating on a finite number of samples, and it is crucial to avoid overestimating the number of secret bits that one can effectively extract with reconciliation and privacy amplification. The analysis in 4.4.1 is based on the detailed study of privacy amplification with continous eavesdropper's observation, which differs from the finite-length analysis in [95, 102] restricted to discrete observations. The numerical results in 4.4.2 are also obtained for the memoryless source $p_{X_Q Y_Q Z}$ characterized experimentally in Section 4.2, and not from computer simulations. The only approximation in this analysis is that the source statistics estimated in Section 4.2.3 correspond to the true statistics; the entire analysis in Section 4.4.1 is exact.

4.4.1 Finite-Length Analysis for a Continuous Observation Z

As stated in 4.2.3, the following finite-length analysis assumes that the eavesdropper's observation Z is continuous and that the estimation of the statistics of the source made in Section 4.2.3 is accurate.

The finite key-length analysis relies on a sequential strategy [16, 67], in which the reconciliation step is performed with error correction codes and the privacy amplification step is performed with hash functions. However, since Z is continuous, it is not possible to directly

use previous approaches [9, 67], which are only valid for discrete random variables, or that can only be extended to continuous random variables in an asymptotic regime [20, 21].[3]

The reconciliation protocol is performed on the quantized versions X_Q and Y_Q of X and Y, and $\mathbf{P}_{\mathrm{e}}^{\mathrm{rec}}$ corresponds to the probability of error of the reconciliation step, while ℓ_{rec} represents the number of information bits leaked during the process.

A lower bound on $\mathbb{H}(K|GZ^nF)$ is needed to determine the final secret-key length obtained after privacy amplification with a hash function G chosen at random. This quantity represents the uncertainty obtained by the eavesdropper with its own observation Z^n, the knowledge of G, and the public message F. The strategy $\mathscr{S}_n$ is also known by the eavesdropper, but is omitted in the subsequent analysis to simplify the notation.

The goal is to bound the equivocation $\mathbb{H}(K|Z^nFG)$. This section explains how to bound this quantity as

$$k - \delta_\epsilon(n) \leqslant \mathbb{H}(K|Z^nFG) \leqslant k, \tag{4.4}$$

where k represents the size of the output of the hash function G used to distill the key. The finite length analysis consists in estimating k for a given n. Using the results in [102, Corollary 2] for privacy amplification, it follows that

$$\mathbb{E}\left(\mathbb{V}(p_{KZ^nFG}, u_{K^n} p_{Z^nFG})\right) \leqslant 2\epsilon + \frac{1}{2}\sqrt{2^{nR - \mathbb{H}_\infty^\epsilon\left(X_Q^n|Z^nFG\right)}}. \tag{4.5}$$

This expression involves the ϵ-smooth min-entropy of X_Q^n given the information available at the eavesdropper's terminal. This quantity cannot be directly evaluated, but relates to other information theoretic metrics that can be estimated.

Lemma 4.1 Let S and U be two random variables and let $r > 0$. Then, with probability at least $1 - 2^{-r}$,

$$\mathbb{H}_\infty^\epsilon(S) - \mathbb{H}_\infty^\epsilon(S|U = \bar{u}) \leqslant \log|\mathcal{U}| + r. \tag{4.6}$$

$\diamond$

[3] Note that, in principle, it is possible to quantize Z since by [78][32][23, Section 8.5][8, Lemma 2][21, Lemma 1.2], for any $\delta > 0$, if a quantized version $Z_{Q'}$ of Z is fine enough,

$$|\mathbb{I}(K;AZ) - \mathbb{I}(K;FZ_{Q'})| < \delta.$$

Proof See the proof in Appendix 7.3.5 on page 138. ∎

Lemma 4.1 proves that, with high probability, the decrease caused by conditioning the ϵ-smooth min-entropy on $\mathcal{U}$ is bounded by $\log|\mathcal{U}| + r$. To leverage this lemma, consider the following random variable

$$\Upsilon \triangleq \mathbb{1}\{\mathbb{H}_\infty^\epsilon(X_Q^n|Z^nG) - \mathbb{H}_\infty^\epsilon(X^n|Z^nG, F = f_{rec}) \leqslant \log|\mathcal{F}| + \sqrt{n}\}, \tag{4.7}$$

which is such that $\mathbb{P}(\Upsilon = 1) \geqslant 1 - 2^{\sqrt{n}}$.

From [47, Theorem 1], the ϵ-smooth entropy $\mathbb{H}_\infty^\epsilon\left(X_Q^n|Z^n\right)$ is lower bounded by the conditional entropy $\mathbb{H}(X_Q|Z)$ as

$$\forall 0 < \epsilon < 1,\ \mathbb{H}_\infty^\epsilon\left(X_Q^n|Z^n\right) \geqslant n\mathbb{H}\left(X_Q|Z\right) - \sqrt{2n}\log(|\mathcal{X}|+3)\sqrt{\log\epsilon^{-1}}. \tag{4.8}$$

Since Lemma 4.1 guarantees that

$$\mathbb{H}_\infty^\epsilon(X_Q^n|Z^nG, F = f_{rec}, \Upsilon = 1) \geqslant \mathbb{H}_\infty^\epsilon(X_Q^n|Z^nG, \Upsilon = 1) - \underbrace{\log|\mathcal{F}|}_{\ell_{\text{rec}}} - \sqrt{n}, \tag{4.9}$$

combining (4.8) and 4.9 yields

$$\mathbb{H}_\infty^\epsilon(X_Q^n|Z^nG, F = f_{rec}, \Upsilon = 1) \geqslant n\mathbb{H}\left(X_Q|Z\right) - \sqrt{2n}\log(|\mathcal{X}|+3)\sqrt{\log\epsilon^{-1}} - \ell_{\text{rec}} - \sqrt{n}. \tag{4.10}$$

Combining this equation with (4.5), provides an upper bound on $\mathbb{E}\left(\mathbb{V}(p_{KZ^nFG}, u_{K^n}p_{Z^nFG})\right)$ that can be evaluated. The last step consists in relating this variational distance to the leakage $\mathbf{L} \triangleq \mathbb{I}(K|Z^nFG)$, thanks to Csiszàr inequality:

$$\begin{aligned}
\mathbb{V}\left(p_{KZ^nFG,\Upsilon=1}, u_{K^n}p_{Z^nFG,\Upsilon=1}\right)\log\frac{2^k}{\mathbb{V}\left(p_{KZ^nFG,\Upsilon=1}, u_{K^n}p_{Z^nFG,\Upsilon=1}\right)} & \\
\geqslant |k + \mathbb{H}(Z^nFG|\Upsilon = 1) - \mathbb{H}(KZ^nFG|\Upsilon = 1)| & \\
= |k - \mathbb{H}(K|Z^nFG\Upsilon = 1)| & \\
\geqslant |\mathbb{H}(K) - \mathbb{H}(K|Z^nFG\Upsilon = 1)| & \\
\triangleq \mathbb{I}(K|Z^nFG) \triangleq \mathbf{L} &
\end{aligned} \tag{4.11}$$

4.4.2 Numerical Evaluation

The experiments provide four different channel measurements. Since both Alice and Bob transmit probe signals, they can respectively obtain the channel gains G_{BA} and G_{AB}. Incidentally, Eve can also estimate the channel gains G_{AE} and G_{BE}. The source of randomness is therefore (X, Y, Z), where $X = G_{BA}$, $Y = G_{AB}$, and $Z = (G_{AE}, G_{BE})$.

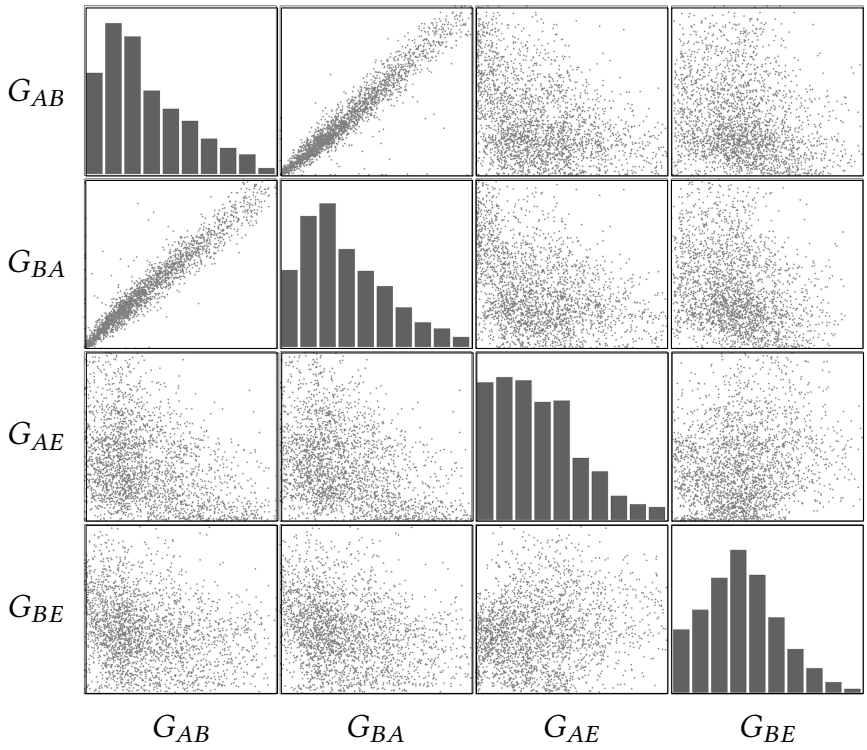

Figure 4.8 – *Correlations between the different normalized channel gains.*

Figure 4.8 illustrates how the channel gain measurements conducted in Section 4.2.3 are correlated with each other. This figure shows that the channel gains G_{BA} and G_{AB} are strongly correlated and confirms the hypothesis of reciprocity. The correlation is, however, not perfect, thus justifying the reconciliation phase in the secret-key generation protocol. On the other hand, there is no substantial correlation either between both Eve's observations G_{AE} and G_{BE}, or between Eve's observations and the legitimate users' observations. This second observation confirms the hypothesis of diversity when Eve is not located near one of the antennas of the legitimate parties. Looking solely at the distributions of the channel gains yields similar conclusions

since G_{BA} and G_{AB} have similar distributions (reciprocity), while G_{AE} and G_{BE} have different distributions due to the different path followed by the electromagnetic waves (diversity).

If the reconciliation protocol has an efficiency $\beta \in [0, 1]$ (see [16]), then $l_{\text{rec}} = n(\mathbb{H}(X) - \beta I(X;Y))$ bits are leaked to the eavesdropper. The best case scenario, obtained for $\beta = 1$, would yield $\mathbb{H}(X|Y)$ bits leaked during the reconciliation step. Note that the value of β may change depending on the blocklength n.

Asymptotically, it is known that the corresponding achievable secret-key rate is $R_{\text{low}} \triangleq \mathbb{I}(X;Y) - \mathbb{I}(X;Z)$, which is a lower bound of the secret-key capacity C_s [66].

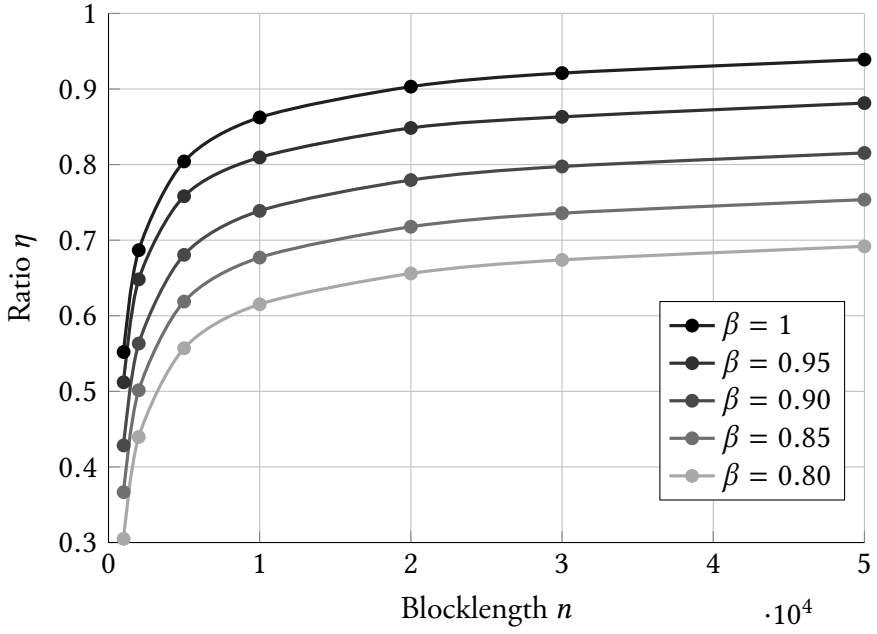

Figure 4.9 – *Ratio η, for* $\mathbf{U} < 10^{-3}$ *and* $\mathbf{L} < 10^{-3}$.

The gap between the finite and the asymptotic regime is quantified using the ratio

$$\eta \triangleq \frac{k/n}{R_{\text{low}}},$$

where $R_{\text{low}} \approx 1.46$ bits is a known achievable rate, which is upper bounded by the secrecy capacity C_s, which is such that $C_s \leqslant \mathbb{I}(X;Y|Z) \approx 1.76$ bits.

4.5 Conclusion and Discussion

The main limitation of a secret-key generation scheme relying on channel variations is the difficulty to clearly establish the conditions to provide secrecy without precise knowledge of the eavesdropper's statistics. The experiments show that there is no simple relationship between the eavesdropper's proximity and the correlation of its observations. Even worse, it appears that if the eavesdropper induces motion, it might be able to obtain an accurate estimation of the legitimate users' own observations. This is consistent with the observations made in [52] when the adversary controls the environment to induce a predictable channel behavior.

As already seen in [65, 109], there is also a strong relationship between the environment mobility and the achievable secret-key rate. The secret-key generation scheme should account for these variations to avoid extracting more randomness than possible. For instance, depending on the degree of mobility of the environment, the coherence time will change and the length of the secret-key should change accordingly.

Estimating the source statistics not only allows one to assess the security of the system, but also to define the source induction functions and to select the hash function output size. However, the estimation of the source statistics is a problem in itself, especially because the eavesdropper's observations are not available to legitimate users.

Finally, the achievable secret-key generation rate also depends on the blocklength. Thus a lightweight scheme that operates with short codes will have much lower performances than a scheme operating with infinitely long codes.

To increase secret-key generation rates, one possibility is to consider other channel parameters to obtain a source with more randomness. In theory, the channel phase is uniformly distributed and may exhibit more entropy than the gain magnitude. However, designing a secret-key generation scheme using the channel phase would be technically challenging and would require complex synchronization between terminals.

To maintain a minimum level of secrecy regardless of the environment properties, it is crucial to address the lack of eavesdropper channel state information (CSI) for the legitimate

users. For instance, one can obtain a significant advantage by using multiple antennas [20, 43, 99]. Again, this improvement would make the system more complex, and such choice would depend on the final desired level of secrecy.

CHAPTER 5

PRACTICAL CODED COOPERATIVE JAMMING[1]

Historically, coding theory has focused on designing codes that are good for reliability, but not directly meant to do binning for the wiretap codes. However, the growing interest for physical-layer security has spurred the development of new coding techniques for secrecy based on powerful error-correcting codes, such as low-density parity-check (LDPC) codes and polar codes.

The underlying ideas behind coding for secrecy, which are motivated by the analysis of the wiretap channel by Wyner [107], are twofold. Randomness should be introduced to confuse the eavesdropper by sending a random sequence along with the actual message and the codebook should exhibit a binning structure. One approach [4] consists in employing two-edge LDPC codes with a coset encoding technique in which the binning is controlled by the two-edge structure. Another approach [54, 106] is to use punctured LDPC codes to create a nested linear code structure. Both constructions exhibit performances close to the weak secrecy capacity when employed with spatially coupled LDPC codes [4, 53, 57, 82]. A last approach consists in exploiting polar codes to achieve the secrecy capacity as described in [22, 64] and references therein. All these works open up new directions for the design of practical codes for secrecy.

This section presents a code construction based on punctured LDPC codes for the multiple-access channel to perform coded cooperative jamming. This scheme guarantees weak secrecy and its performances are analyzed for both classical and spatially coupled LDPC codes.

The outline of this chapter is as follows. Section 5.1 introduces and motivates the two-way wiretap model considered in this chapter. Section 5.2 introduces spatially-coupled LDPC codes for the multiple-access channel. Section 5.3 describes how to use SC-LDPC codes to provide

[1]Parts of the material in this chapter have appeared in [77]: **Pierrot, A. J.**, Bloch, M. R., "LDPC-Based Coded Cooperative Jamming Codes". In: *Proceedings of the IEEE Information Theory Workshop*. Lausanne, Switzerland, Sept. 2012, pp. 462–466. ©IEEE 2012.

secrecy and assesses the performances of the proposed coding scheme. Section 5.4 concludes this chapter and discusses some of the model assumptions.

5.1 Coded Cooperative Jamming for the Two-Way Wiretap Channel

5.1.1 General Model

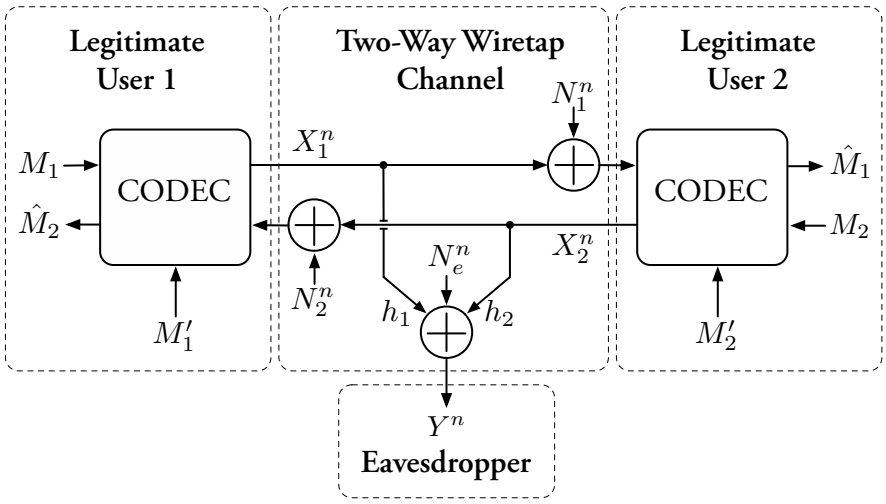

Figure 5.1 – *Communications over the Gaussian two-way wiretap channel.*

Consider a special case of the model presented in the previous section called *the binary-input memoryless Gaussian two-way wiretap channel*, which is presented in Figure 5.1. N_1 and N_2 represent the noise at legitimate users' terminals. From the eavesdropper's standpoint,

$$Z = \sqrt{h_1}X_1 + \sqrt{h_2}X_2 + N_e, \tag{5.1}$$

where

- $X_1, X_2 \in \{-1, 1\}$ are BPSK modulated symbols;
- $+$ is the usual real number addition;
- N_e is a centered Gaussian noise with variance σ^2 and independent of channel inputs; and gains are unitary $h_1 = h_2 = 1$.

The hypothesis of equal-gained interferences may appear restrictive since this situation is unlikely to naturally arise in a real wireless setting as shown in Section 5.4. However, one can

imagine a situation in which the eavesdropper is an honest-but-curious third user that also communicates, but should not access some of the information exchanged over the network. Since this user communicates over the network, it is possible to estimate its channel parameters and to induce pure interferences at its terminal by appropriately scaling power.

This model neglects the interferences between codewords at the legitimate receivers' terminals since the interferences that occur between exchanged codewords may be canceled by the knowledge of the transmitted sequences.

5.1.2 Erasure MAC

To make code design more tractable, the model is further simplified thanks to the following lemma.

Lemma 5.1 Any memoryless q-input channel $\left(\mathscr{X}, p_{Y|X}, \mathscr{Y}\right)$ (the output alphabet can be either continuous or finite) is stochastically degraded with respect to the *q-ary erasure channel* with erasure probability ϵ^* defined as

$$p_{Z|X} = \begin{cases} \epsilon & \text{if } Z = ? \\ 1-\epsilon & \text{if } Z = X \\ 0 & \text{otherwise} \end{cases}$$

$$\epsilon^* \triangleq \int_{\mathscr{Y}} \min_{u \in [\![0,q-1]\!]} p_{Y|X}(y|u) \mathrm{d}y.$$

◊

Proof Let $\left(\mathscr{X}, p_{Z|X}, \mathscr{Z}\right)$ be a q-ary erasure channel with erasure probability $\epsilon^* \notin \{0,1\}$:

$$p_{Z|X} = \epsilon^* \mathbb{1}\{Z = ?\} + (1-\epsilon^*)\mathbb{1}\{Z = X\},$$

with $\mathscr{X} = [\![0, q-1]\!]$ and $\mathscr{Z} = \mathscr{X} \cup \{?\}$. Now, define the channel $\left(\mathscr{Z}, p_{Y|Z}, \mathscr{Y}\right)$ as follows

$$p_{Y|Z}(y|?) = \frac{1}{\epsilon^*} \min_{u \in [\![0,q-1]\!]} p_{Y|X}(y|u), \tag{5.2}$$

$$p_{Y|Z}(y|z) = \frac{1}{1-\epsilon^*} \left(p_{Y|X}(y|z) - \min_{u \in [\![0,q-1]\!]} p_{Y|X}(y|u) \right), \quad \text{if } z \in [\![0, q-1]\!]. \tag{5.3}$$

For any $(x, y) \in \mathscr{X} \times \mathscr{Y}$,

$$\begin{aligned}\sum_{z\in\mathscr{X}} p_{Y|Z}(y|z)\, p_{Z|X}(z|x) &= p_{Y|Z}(y|?)\,\epsilon^* + \sum_{z\in[\![0,q-1]\!]} p_{Y|Z}(y|z)\,(1-\epsilon^*)\mathbb{1}(z=x)\\ &= \min_{u\in[\![0,q-1]\!]} p_{Y|X}(y|u) + p_{Y|X}(y|x) - \min_{u\in[\![0,q-1]\!]} p_{Y|X}(y|u)\\ &= p_{Y|X}(y|x)\,. \end{aligned} \tag{5.4}$$

■

In particular, this shows that the channel $(\{-2, 0, 2\}, p_{Y|X_1+X_2}, \mathbb{R})$ is stochastically degraded with respect to a ternary erasure channel. This observation allows to further simplify the model by considering the *erasure multiple-access channel* defined as follows

$$Z = \begin{cases} X_1 + X_2 & \text{with probability } 1-\epsilon \\ ? & \text{with probability } \epsilon \end{cases}, \tag{5.5}$$

where $\epsilon = 2Q\,(2/\sigma)$. This channel is not strictly equivalent to the initial model, but the data processing inequality ensures that the information received by the eavesdropper from the actual channel is smaller than the information it would obtain from the output of the erasure multiple-access channel.

The achievable communication rates (R_1, R_2) for such an erasure channel must satisfy

$$\begin{aligned} R_1 &\leqslant \mathbb{I}(X_1; Y|X_2) = 1-\epsilon,\\ R_2 &\leqslant \mathbb{I}(X_2; Y|X_1) = 1-\epsilon,\\ R_1 + R_2 &\leqslant \mathbb{I}(X_1, X_2; Y) \triangleq C_{\text{sum}} = 3(1-\epsilon)/2, \end{aligned} \tag{5.6}$$

which is obtained with uniform input symbols. Equivalently, the maximum erasure rate correctable by codes of rates R_1 and R_2 is

$$\epsilon_{\text{Shannon}} = \min\left(1-R_1, 1-R_2, 1-2/3(R_1+R_2)\right). \tag{5.7}$$

5.1.3 Leakage Analysis

Codes for the MAC are used in the following way to obtain secrecy for the two-way wiretap channel.

- X_1^n (resp. X_2^n) corresponds to the codeword sent by legitimate user 1 (resp. user 2) over the channel and obtained from secure message $M_1 \in [\![1, 2^{nR_1}]\!]$ (resp. $M_2 \in [\![1, 2^{nR_2}]\!]$) randomly encoded with an auxiliary message $M_1' \in [\![1, 2^{nR_1'}]\!]$ (resp. $M_2' \in [\![1, 2^{nR_2'}]\!]$).
- Z^n corresponds to the observation of an eavesdropper at the output of the erasure multiple-access channel. This eavesdropper must not get any information about the secure messages M_1 and M_2.

The secrecy metric is chosen to be the *leakage rate* (weak secrecy criterion); for a given code $\mathscr{C}_n$, ensuring secrecy requires

$$\lim_{n\to\infty} \underline{\mathbf{L}}(\mathscr{C}_n) = 0, \text{ with } \underline{\mathbf{L}}(\mathscr{C}_n) \triangleq \frac{1}{n}\mathbb{I}(Z^n; M_1, M_2|\mathscr{C}_n). \tag{5.8}$$

It can be shown that (for simplicity, the conditioning on $\mathscr{C}_n$ is dropped)

$$\begin{aligned}
n\underline{\mathbf{L}}(\mathscr{C}_n) &= \mathbb{I}(M_1M_2; Z^n) \\
&= \mathbb{I}\big(M_1M_2M_1'M_2'; Z^n\big) - \mathbb{I}\big(M_1'M_2'; Z^n|M_1M_2\big) \\
&= \mathbb{I}\big(X_1^nX_2^n; Z^n\big) + \underbrace{\mathbb{I}\big(M_1M_2; Z^n|X_1^nX_2^n\big)}_{=0\ (M_1M_2\to X_1^nX_2^n\to Z^n)} - \mathbb{I}\big(X_1^nX_2^n; Z^n|M_1M_2\big) \\
&= \mathbb{I}\big(X_1^nX_2^n; Z^n\big) - \mathbb{H}\big(M_1'M_2'|M_1M_2\big) + \mathbb{H}\big(M_1'M_2'|M_1M_2Z^n\big) \\
&= \mathbb{I}\big(X_1^nX_2^n; Z^n\big) - \mathbb{H}\big(M_1'M_2'\big) + \mathbb{H}\big(M_1'M_2'|M_1M_2Z^n\big) \\
&\leqslant nC_{\text{sum}} - n(R_1' + R_2') + \mathbb{H}\big(M_1'M_2'|M_1M_2Z^n\big).
\end{aligned} \tag{5.9}$$

Hence, to provide secrecy, the code rates R_1' and R_2' must compensate C_{sum} and minimize $\mathbb{H}\big(M_1'M_2'|M_1M_2Z^n\big)$. This last condition corresponds to considering a virtual receiver capable of decoding (M_1', M_2') from (M_1, M_2, Z^n). For code that allows this with an error probability P_e', Fano's inequality ensures

$$\mathbb{H}\big(M_1'M_2'|M_1M_2Z^n\big) \leqslant \mathbb{H}_{\text{b}}\big(P_e'\big) + nP_e'(R_1' + R_2').$$

5.2 LDPC Codes for the MAC

The previous chapters have shown that coded cooperative jamming is one of the mechanisms for providing secrecy in a multi-user scheme. However, the analysis only provides some insight

into the general structure of the code without providing any efficiently implementable codes. This section introduces *spatially coupled low-density parity-check codes* (SC-LDPC) for the MAC as presented in [56].

5.2.1 Spatially-Coupled LDPC Codes

Spatially coupled LDPC codes, SC-LDPC for short, are based on the standard LDPC codes that aim at increasing the belief propagation (BP) threshold by coupling individual LDPC codes. This construction, which has been introduced by Felström and Zigangirov [53] and further analyzed by Kudekar *et al.* [57], is capacity-achieving for the binary erasure channel. The capacity-based approach for secrecy relies on the use of good codes for reliability. This subsection presents the basic construction of SC-LDPC codes to clarify the mathematical definitions and properties of such ensembles. These codes also exhibit better floors than classical irregular LDPC codes and better finite-length performances than polar codes.

5.2.1.1 The (l,r, L) Ensemble

The construction of the (l,r, L) ensemble [57] consists in repeating and coupling several LDPC protographs, which are based on small regular (l,r) LDPC codes.

Definition 25 A protograph $\mathcal{P} = (\mathcal{V}, \mathcal{C}, \mathcal{E})$ is a bipartite graph that consists of a set of variable nodes $\mathcal{V}$, a set of check nodes $\mathcal{C}$, and a set of edges $\mathcal{E}$. Each edge $e \in \mathcal{E}$ connects a variable node $v_e \in \mathcal{V}$ to a check node $c_e \in \mathcal{C}$. A variable node and a check node are connected by multiple edges. ◊

For the construction of the (l,r, L) ensemble [57] where $\mathtt{r} = k\mathtt{l}$ and l is odd, the protograph $\mathcal{P}_{SC}$ consists of one check node and k variable nodes, where each variable node is connected to the unique check node by r edges. For example, Figure 5.2 depicts the protograph of a standard (3,6)-regular ensemble. The protograph is then repeated $2L + 1$ times and then for each check node:

- keep only one edge and disconnect the $\mathtt{l} - 1$ remaining ones;

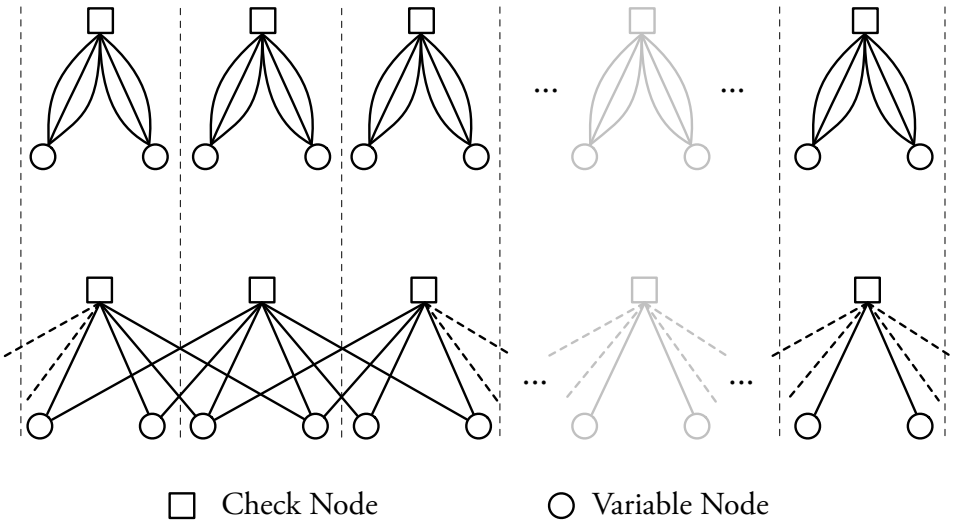

Figure 5.2 – *Spatially-coupled LDPC ensemble construction.*

- connect a pair of edges to the first adjacent check nodes, a pair to the second adjacent variable nodes, etc.;
- if there is no available adjacent check node, remove the edge.

This construction based on protograph can be generalized for cases where $\mathtt{r}$ is not a multiple of $\mathtt{l}$, but $\mathtt{l}$ is still odd.

1. Take $2L+1$ sets of M variable nodes $v_{i,j}$ with $i \in [\![-L, L]\!]$ and $j \in [\![1, M]\!]$. The total number of variable nodes is therefore $M(2L+1)$.
2. Take $2(L+\hat{\mathtt{l}})+1$ sets of $M\mathtt{l}/\mathtt{r}$ check nodes $c_{i,j}$ with $i \in [\![-L-\hat{\mathtt{l}}, L+\hat{\mathtt{l}}]\!]$, $\hat{\mathtt{l}} = (\mathtt{l}-1)/2$, and $j \in [\![1, M\mathtt{l}/\mathtt{r}]\!]$.
3. At each position $i \in [\![-L-\hat{\mathtt{l}}, L+\hat{\mathtt{l}}]\!]$, there are exactly $M\mathtt{l}$ check node sockets. Randomly connect the check nodes to the variable nodes $v_{i,\cdot}$ with $i \in [\![i-\hat{\mathtt{l}}, i+\hat{\mathtt{l}}]\!]$.

5.2.1.2 The $(\mathtt{l}, \mathtt{r}, L, w)$ *Ensemble*

The $(\mathtt{l}, \mathtt{r}, L)$ ensemble is rather complicated to analyze and can be simplified by introducing another level of randomization in the construction. The construction consist of

1. $2L+1$ sets of M variable nodes $v_{i,j}$ with $i \in [\![-L, L]\!]$ and $j \in [\![1, M]\!]$;

2. an infinity of candidate check nodes $c_{i,j}$ with $i \in \mathbb{N}$, $\hat{1} = (1-1)/2$, and $j \in [\![1, M1/r]\!]$. Each check node still has $M1$ sockets.

The construction considers an additional parameter w used to further randomize how variable and check nodes are connected. Each of the 1 connections of a variable node at position i is chosen from the range $[\![i, i+w-1]\!]$. Each of the r connections of a check node at position i is randomly chosen from the range $[\![i-w+1, i]\!]$.

Design Rate [57] The design rate of the $(1, r, L, w)$ ensemble for $w \leqslant 2L$ is

$$R_d(1, r, L, w) = \left(1 - \frac{1}{r}\right) - \frac{1}{r(2L+1)}\left(w + 1 - 2\sum_{i=0}^{w}\left(\frac{i}{w}\right)^r\right) \tag{5.10}$$

Performances The performances of the $(1, r, L, w)$ ensemble are extensively presented in [57]. In particular, the belief propagation (BP) threshold for this ensemble reaches the maximum-a-posteriori (MAP) threshold for the binary erasure channel.

5.2.2 Spatially-Coupled LDPC Codes for the MAC

This subsection introduces the construction of Spatially-Coupled LDPC Codes for the MAC as presented in [56]. As illustrated in Figure 5.3, this construction consists in connecting the variable nodes of two spatially-coupled LDPC codes pairwise through *functional nodes*.

Consider the $(1_1, r_1, 1_2, r_2, L, w)$[2] ensemble for the MAC, where

- $(1_1, r_1)$ and $(1_2, r_2)$ represent the degree of variable nodes and check nodes respectively, for the first and second codes;
- M variable nodes are placed in positions $[\![-L; L]\!]$, where M corresponds to a free parameter that controls the length of the code;
- $M \cdot 1_i/r_i$ (for $i \in \{1, 2\}$) check nodes are placed in positions $[\![-L; L+w-1]\!]$;
- finally, the parameter w controls how connections are chosen since a variable node in position j can only be connected to a check node in position $[\![j, j+w-1]\!]$.

[2] In the present chapter, $(1, r, L, w)$ denotes the ensemble $(1, r, 1, r, L, w)$.

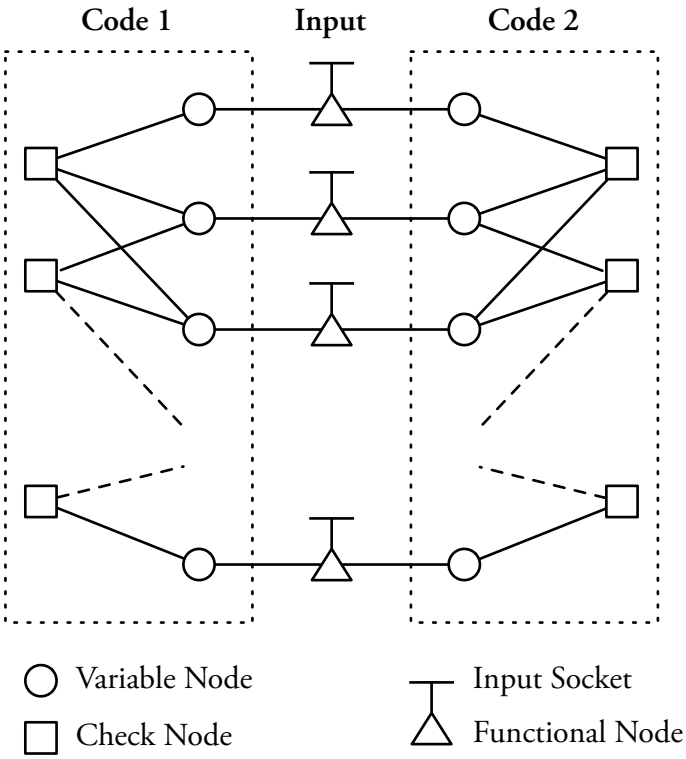

Figure 5.3 – *LDPC construction for the MAC.*

Decoder During the decoding process, variable nodes from associated factor graphs exchange information through the functional nodes:

- if the received symbol is -2 or 2, there is no uncertainty about the symbol values;
- if the received symbol is 0, the symbol pairs are either $(-1, 1)$ or $(1, -1)$, which can be used to find a bit value whenever its counterpart is discovered;
- if the received symbol is ?, there is no relationship between paired symbol values;
- a functional node receives from variable nodes extrinsic information;
- a functional node receives the channel a priori as an input;

- a functional node sends to variable nodes intrinsic information in the following sense

$$\begin{aligned}\lambda_{int,i}^{(1)} &= \log \frac{\mathbb{P}(Z_i|X_i^{(1)}=0)}{\mathbb{P}(Z_i|X_i^{(1)}=1)} \\ &= \log \frac{\sum_x \mathbb{P}(Z_i|X_i^{(1)}=0, X_i^{(2)}=x) e^{x\lambda_{ext,i}^{(2)}}}{\sum_x \mathbb{P}(Z_i|X_i^{(1)}=1, X_i^{(2)}=x) e^{x\lambda_{ext,i}^{(2)}}} \\ \lambda_{int,i}^{(2)} &= \log \frac{\sum_x \mathbb{P}(Z_i|X_i^{(2)}=0, X_i^{(1)}=x) e^{x\lambda_{ext,i}^{(1)}}}{\sum_x \mathbb{P}(Z_i|X_i^{(2)}=1, X_i^{(1)}=x) e^{x\lambda_{ext,i}^{(1)}}}.\end{aligned}$$

Design rate For $i \in \{1,2\}$, the design rates of such a multi-user code are provided by [56, Lemma 3]

$$R_{d,i} = 1 - \frac{\mathtt{l}_i}{\mathtt{r}_i} - \frac{\mathtt{l}_i}{\mathtt{r}_i}\frac{1}{2L+1}\left(w + 1 - 2\sum_{j=0}^{w}\left(\frac{j}{w}\right)^{\mathtt{r}_i}\right). \tag{5.11}$$

Note that this rate goes to the usual rate $1 - \mathtt{l}_i/\mathtt{r}_i$ as L goes to infinity for a fixed w.

Density Evolution equations for the BEC If $x_i^{(t)}$ (resp. $y_i^{(t)}$) corresponds to the probability that a variable-to-check (resp. a check-to-variable) message carries an erasure for user $i \in \{1,2\}$, the density evolutions (DE) equations are

$$y_j^{(1)} = 1 - \left(1 - \frac{1}{w}\sum_{k=0}^{w-1} y_{j-k}^{(1)}\right)^{\mathtt{r}_1-1} \tag{5.12}$$

$$x_j^{(1)} = \left(\epsilon + \frac{1-\epsilon}{2}\left(\frac{1}{w}\sum_{k=0}^{w-1} y_{j+k}^{(2)}\right)^{\mathtt{l}_2}\right)\left(\frac{1}{w}\sum_{k=0}^{w-1} y_{j+k}^{(1)}\right)^{\mathtt{l}_1-1} \tag{5.13}$$

$$y_j^{(2)} = 1 - \left(1 - \frac{1}{w}\sum_{k=0}^{w-1} y_{j-k}^{(2)}\right)^{\mathtt{r}_2-1} \tag{5.14}$$

$$x_j^{(2)} = \left(\epsilon + \frac{1-\epsilon}{2}\left(\frac{1}{w}\sum_{k=0}^{w-1} y_{j+k}^{(1)}\right)^{\mathtt{l}_1}\right)\left(\frac{1}{w}\sum_{k=0}^{w-1} y_{j+k}^{(2)}\right)^{\mathtt{l}_2-1}, \tag{5.15}$$

where

- $x_j^{(i)}$ denotes the average erasure probability sent from the variable node side to the check node side for user $i \in \{1,2\}$, for a node at position $j \in [\![-L, L]\!]$;
- $y_j^{(i)}$ denotes the average erasure probability flowing from the check node side to the variable node side for user i, for a node at position $j \in [\![-L, L+w-1]\!]$.

5.3 Punctured LDPC Codes for the MAC

5.3.1 Code Construction

The code construction is an extension of the notion of code puncturing [54, 106] for the wiretap channel to the two-way channel defined earlier. The puncturing scheme operates as follows.

1. Choose LDPC codes $(\mathscr{C}_1, \mathscr{C}_2)$ for the MAC [56] with parity-check matrices H_1 and H_2 of the same size.

2. Split the matrix $H_1 \in \mathfrak{M}_{m-l,m}(\mathbb{F}_2)$ as $H_1 = [A_1\ B_1]$, with $B_1 \in \mathrm{GL}_{m-l}(\mathbb{F}_2)$. Similarly, split the matrix $H_2 \in \mathfrak{M}_{m-l,m}(\mathbb{F}_2)$ as $H_2 = [A_2\ B_2]$, with $B_2 \in \mathrm{GL}_{m-l}(\mathbb{F}_2)$.

3. Form the temporary codeword $\tilde{X}_1^m = [m_1, m_1', s_1] \in \mathscr{C}_1$ with $m_1 \in \mathbb{F}_2^k$, $m_1' \in \mathbb{F}_2^{l-k}$ $s_1 = [m_1, m_1']A_1^{\mathsf{T}}(B_1^{-1})^{\mathsf{T}} \in \mathbb{F}_2^{m-l}$. The second user does the same with $\tilde{X}_2^m = [m_2, m_2', s_2] \in \mathscr{C}_2$ with $m_2 \in \mathbb{F}_2^k$, $m_2' \in \mathbb{F}_2^{l-k}$ $s_2 = [m_2, m_2']A_2^{\mathsf{T}}(B_2^{-1})^{\mathsf{T}} \in \mathbb{F}_2^{m-l}$.

4. Puncture this code by only keeping $X_1^n = [m_1', s_1]$ and $X_2^n = [m_2', s_2]$, with $n = m - k$.

Puncturing induces a nested linear code structure, which is an algebraic counterpart of the binning suggested by information-theoretic proofs. The rate of the mothercode is denoted by $R_d \triangleq l/m$ and the secret rate for user $i \in \{1, 2\}$ by $R_i \triangleq k/n$. If the fraction of punctured bit is p,

$$p = \frac{k}{m} \quad \Rightarrow \quad R_i \triangleq \frac{k}{n} = \frac{k}{m-k} = \frac{p}{1-p}. \tag{5.16}$$

The quantity

$$R_i' \triangleq \frac{l-k}{m-k} = \frac{l-pm}{m-pm} = \frac{R_d - p}{1-p} \tag{5.17}$$

represents the auxiliary message rate for user $i \in \{1, 2\}$.

Encoding Use the punctured LDPC codes for data transmission as follows. User 1 sets m_1 to be the secret message, m_1' the auxiliary message and s_1 the check sum defined above, and sends $[m_1', s_1]$ over the channel. User 2 performs similar encoding with independent sequences m_2 and m_2'.

Decoding for the legitimate receivers If user $i \in \{1, 2\}$ receives r_i, its decoder performs belief propagation with $[?^{m-n}, r_i^n]$ as the channel *a priori* since it has no access to the punctured bits, to estimate $\tilde{X}_i^m$.

Decoding for the virtual receiver To minimize the last term of (5.9), a virtual user should be able to retrieve messages M_1' and M_2' provided (M_1, M_2, Z^n). The decoder estimates $\tilde{X}_1^m$ and $\tilde{X}_2^m$ using its channel observation and knowing the punctured bits.

Puncturing increases the decoding capabilities of the virtual receiver, thus providing secrecy. However, when the number of punctured symbol increases, the communication rate decreases.

5.3.2 Leakage Analysis

Consider the $(\mathtt{l}_1, \mathtt{r}_1, \mathtt{l}_2, \mathtt{r}_2, L, w)$ ensemble for the MAC presented in the previous section. From Charlie's point of view, it consists in setting the a priori inputs of coupling nodes to the actual value of m. The density evolution equations remain the same, but the initialization is different. The a priori informations become

- Bob: $[?\ldots?\,,\ d_i'\,,\ e_i']$, therefore the density evolution initialization is

$$x^{(0)} = \frac{|c|}{n} + \frac{n - |c|}{n}\epsilon.$$

- Charlie: $[c,\ d_i''\,,\ e_i'']$, therefore the density evolution initialization is

$$x^{(0)} = \frac{n - |c|}{n}\epsilon.$$

Recall that since the virtual user must be able to recover (M_1', M_2') from (M_1, M_2, Y^n), the puncturing consists in providing the virtual user with the values of the messages at the punctured position. If p denotes the puncturing rate, note that this is not equivalent to setting the erasure probability to $p\epsilon$ in the density evolution equations derived in [56]. Indeed, whenever a symbol is punctured, the virtual receiver obtains the values of x_1 and x_2 at this position, not $x_1 + x_2$; therefore, not only the erasures are removed in positions of punctured bits, but the interferences when the symbol pairs are $(-1, 1)$ and $(1, -1)$ are also resolved.

Variable-to-check equation A variable node returns an erasure if all of the following happens:

1. the variable node is not punctured, which occurs with probability $1-p$;
2. the $\mathtt{l}_1 - 1$ adjacent edges carry erroneous messages;
3. the received value is erased (with probability ϵ), or the received value corresponds to a collision (with probability $(1-\epsilon)/2$) and the opposite variable node is unknown.

Therefore, the average probability of erasure is

$$x_j^{(1)} = (1-p)\left(\epsilon + \frac{1-\epsilon}{2}\left(\frac{1}{w}\sum_{k=0}^{w-1} y_{j+k}^{(2)}\right)^{\mathtt{l}_2}\right)\left(\frac{1}{w}\sum_{k=0}^{w-1} y_{j+k}^{(1)}\right)^{\mathtt{l}_1-1}. \tag{5.18}$$

Check-to-variable equation A check node sends an erasure to a variable node if at least one of the $\mathtt{r}_1 - 1$ other adjacent edges is an erasure:

$$y_j^{(1)} = 1 - \left(1 - \frac{1}{w}\sum_{k=0}^{w-1} x_{j-k}^{(1)}\right)^{\mathtt{r}_1-1}. \tag{5.19}$$

Similar equations are obtained for user 2.

The following proposition allows to avoid the analysis of density evolution for every possible pair of codewords.

Proposition 5.2 ([85]) The average behavior of the previous density evolution can be obtained for X_1^n corresponding to the all-zero codeword and X_2^n corresponding to a type one-half codeword (equal number of -1 and 1). In that case $X^n = X_1^n + X_2^n = X_2^n$ corresponds also to a type one-half codeword.

Proof The proof can be found in [85] for the Z-channel and holds for the multiple-access channel. ∎

Assume that $X_1^n + X_2^n$ is a type one-half codeword. Define the set

$$\mathscr{H}_\delta = \left\{X^n, \frac{1}{2} - \frac{\delta}{\sqrt{n}} \leqslant \frac{1}{n}\mathrm{wt}(X^n) \leqslant \frac{1}{2} + \frac{\delta}{\sqrt{n}}\right\}.$$

The probability that a codeword does not belong to that set is small and because of the symmetries in the decoding scheme, the behavior of the decoder is the same for all type one-half codewords. The analysis can be restricted to the codewords $X_1^n = \mathbf{1}$ and $X_2^n = \mathbf{h}$, where $\mathbf{h}$ is of type one-half.

Notice that a type one-half codeword $X^n = X_1^n + X_2^n$ captures the behavior of the code ensemble since, for the erasure multiple-access channel, it represents a sequence with half interferences and half known values, which is a typical interference pattern.

5.3.3 Reliability Analysis

Suppose the channel from user 1 to user 2 is a BEC(ϵ_1) and the channel from user 2 to user 1 is a BEC(ϵ_2). Puncturing the codewords is equivalent to increasing the erasure probability respectively to $p + \epsilon_1$ and $p + \epsilon_2$. If the codes $\mathscr{C}_1$ and $\mathscr{C}_2$ respectively have threshold ϵ_1^* and ϵ_2^*, one can ensure reliable communications if

$$\epsilon_1 \leqslant \epsilon_1^* - p \;\text{ and }\; \epsilon_2 \leqslant \epsilon_2^* - p. \tag{5.20}$$

The code thresholds ϵ_1^* and ϵ_2^* can be found using the density evolution equations for the $(\mathtt{l}, \mathtt{r}, L, w)$ SC-LDPC code ensemble

$$\forall j \in \{1,2\},\;\; y_j = 1 - \left(1 - \frac{1}{w}\sum_{k=0}^{w-1} y_{j-k}\right)^{\mathtt{r}-1} \text{and}\;\; x_j = \epsilon\left(\frac{1}{w}\sum_{k=0}^{w-1} y_{j+k}\right)^{\mathtt{l}-1}. \tag{5.21}$$

For instance, $\epsilon^*(3,6,200,3) = 0.4807$, $\epsilon^*(4,8,200,4) = 0.4893$ and $\epsilon^*(5,10,200,5) = 0.4908$.

5.3.4 Numerical Results

The density evolution equations [85] with puncturing allow to determine the asymptotic behavior of the SC-LDPC code ensemble for different code parameters and puncturing rate p as the blocklength goes to infinity. The erasure probability thresholds in Table 5.1 are computed in various cases with the density evolution equations.

The secrecy performance of the scheme is measured by the leakage rate. Whenever the threshold of the code ensemble is not exceeded, it is possible to decode auxiliary messages M_1'

Table 5.1 – *Advantages of spatially coupled LDPC codes.*

Code parameters			(3,6,3,6)-MAC-LDPC			(3,6,3,6,200,3)		
p	R_s	$\epsilon_{\text{Shannon}}$	R_p	ϵ^*	δ	R_p	ϵ^*	δ
0	0	0.3333	0.5	0.1226	0.2107	0.4977	0.3322	0.0011
0.1	0.1111	0.4074	0.4444	0.2136	0.1938	0.4419	0.4052	0.0022
0.2	0.2500	0.5000	0.3750	0.3210	0.1790	0.3722	0.4958	0.0042
0.3	0.4286	0.6190	0.2857	0.4497	0.1693	0.2825	0.6109	0.0081
0.4	0.6667	0.7778	0.1667	0.6079	0.1699	0.1629	0.7621	0.0157
Code parameters			**(4,8,4,8,200,4)**			**(5,10,5,10,200,5)**		
p	R_s	$\epsilon_{\text{Shannon}}$	R_p	ϵ^*	δ	R_p	ϵ^*	δ
0	0	0.3333	0.4965	0.3332	0.0001	0.4953	0.3333	0.0000
0.1	0.1111	0.4074	0.4406	0.4072	0.0002	0.4392	0.4074	0.0000
0.2	0.2500	0.5000	0.3706	0.4994	0.0006	0.3691	0.4999	0.0001
0.3	0.4286	0.6190	0.2807	0.6178	0.0012	0.2790	0.6188	0.0002
0.4	0.6667	0.7778	0.1609	0.7750	0.0028	0.1588	0.7772	0.0006

Note that the value R_p for $p = 0$ corresponds to the mothercode rate. The quantity δ corresponds to the gap between ϵ^* and $\epsilon_{\text{Shannon}}$.

and M_2' provided M_1, M_2, and Z^n, with a probability arbitrarily close to one. For a given code $\mathscr{C}$ in the ensemble with threshold ϵ^*, if $\epsilon \geqslant \epsilon^*$, then $\mathbf{L}(\mathscr{C}) \leqslant C_{\text{sum}}(\epsilon^*) - 2R_p$, where R_p is the auxiliary message rate given in (5.17).

In Figure 5.4, the leakage rate is plotted with respect to the secret and auxiliary message rates. It shows that, the higher the number of punctured bits, the higher the secrecy rate is, since the punctured part of the codewords carries secret messages. However, the leakage rate increases because the gap between the erasure threshold and the Shannon limit increases with the puncturing rate. Figure 5.4 also shows that coupled codes have better performances than classical LDPC codes and that their performance improve as the degree increases. Leakage values smaller than 10^{-3} bits can be reached, meaning that, on average, less than one per thousand bits is not secured. Finally, even if classical LDPC codes of degrees (4,8) and (5,10) do not work for the multiple-access channel (see [56]), they can be used for secrecy. However, they exhibit

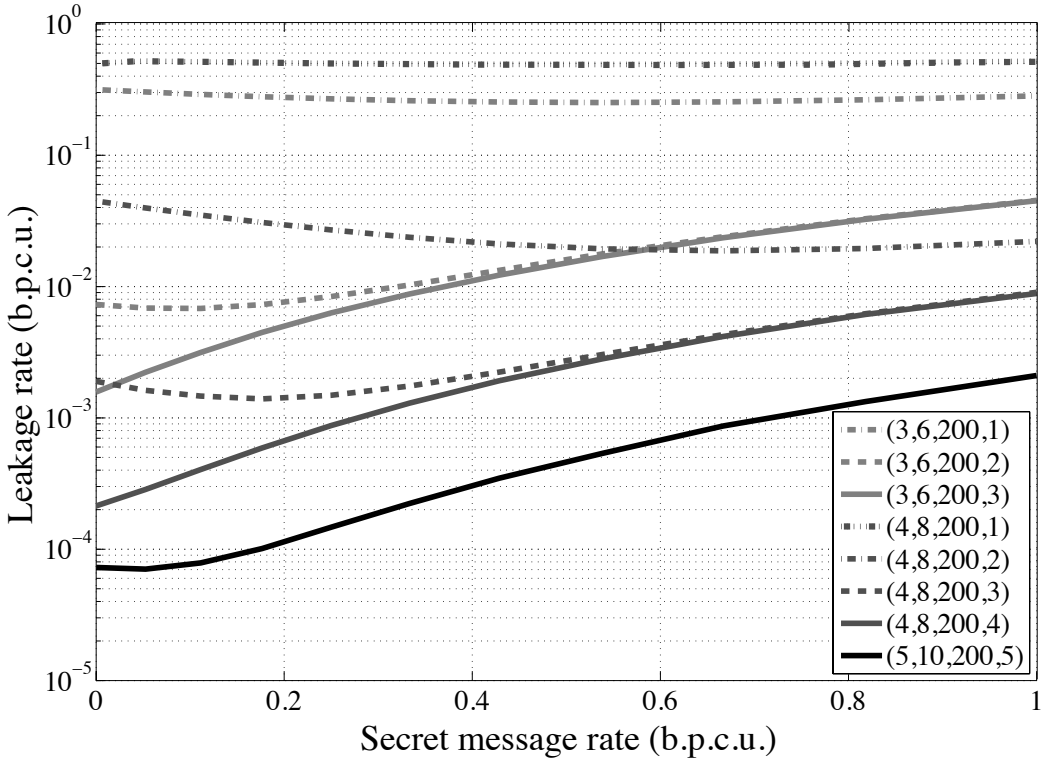

Figure 5.4 – *Leakage v.s. communication rate for various codes and puncturing rates. Codes with* $w = 1$ *correspond to classical LDPC codes.*

poor performance. For instance, the leakage rate is about 0.5 bits for a classical (4,8)-LDPC code.

Finite length effects Density evolution shows that spatially coupled LDPC codes provide good results for secrecy with low leakage rate. However, for actual codes, the effects of finite length decrease the performance for two reasons:

- the difference between C_{sum} and $R'_1 + R'_2$ may not be negligible since one must remain below $C_{\text{sum}}(\epsilon^*)$ to have P'_e small;
- if the probability of error is non-zero, the term $\mathbb{H}\left(M'_1 M'_2 | M_1 M_2 Y^n\right)$ is not negligible and the upper bound on the leakage worsens.

Consequently, the predicted performances of actual codes can be far from asymptotic results. For instance, spatially coupled codes with $L = 25$, $M = 216$, $\mathtt{l} = 3$, $\mathtt{r} = 6$, $w = 3$, and

$n = 11{,}016$ bits provide leakage rate in the order of one half, which is far from the expected asymptotic performances for such codes. Even if these codes appears to be better than a classical LDPC code, they are not sufficient to provide low leakage rate.

These poor performances are likely caused by the small size of the code and the looseness of the upper bound. Studying longer codes and comparing their performances with other constructions is a subject of future investigation.

Remark The previous scheme only provides weak secrecy and only guarantees a low leakage *rate*. Constructions based on polar codes for the multiple-access channel [86], could provide strong secrecy.

5.4 Conclusion and Discussion

This chapter presents a constrcution based on spatially-coupled LDPC codes for coded cooperative jamming over the two-way wiretap channel. The construction uses random puncturing to design a nested code structure to create codebooks, whose codewords detrimentally interfere at the eavesdropper's terminal. Spatially-coupled LDPC codes show significant advantages over classical LDPC codes in terms of leakage and secret rate. However, such a scheme presents some limitations in a real setting.

- The design involves a random an uniform puncturing, which is not practical.
- The model relies on the assumption that the interferences are perfect because the channel gains between Alice and Eve and between Bob and Eve are both unitary. Guaranteeing perfect interference also requires a precise synchronization between the legitimate users, which is challenging for a real communication system.

Uniform Puncturing In [54], the authors investigate the impact of replacing the random puncturing by an optimized puncturing scheme obtained with Differential Evolution [91]. They show gains of up to 0.4 dB over random puncturing. However, it is unclear how these conclusions extend to spatially-coupled LDPC codes over the two-way wiretap channel. If it

appears that a fixed puncturing pattern decreases the performance of the system, it is still possible to compensate the loss with a higher puncturing rate in practice. Increasing the puncturing rate directly impacts the legitimate users since they would need to increase their transmit power to compensate the rate loss.

Gain Imbalance The previous analysis relies on having equal gains to create perfect interference. Such an assumption is rather unrealistic in a real wireless environment. An experimental setting similar to the one presented in the previous section is a good way to assess the validity of the model introduced in Section 5.1. The following experiment is conducted using WARP programmable radios, which performs better than the USPRs presented in the previous chapter. The experimental setting consists of the following.

1. Two WARP radios (Alice and Bob) separated by 1 m that communicate on the 11^{th} WLAN channel (2.462 GHz).

2. A third radio (Eve) that receives on the 11^{th} WLAN channel and that can be moved to different locations.

3. The first radio transmits a sinusoidal wave at 5 MHz, while the other transmits a sinusoidal wave at 5.2 MHz.

4. The third radio uses the waves to compute the received gains G_{AE} and G_{BE}.

5. All the terminals are connected to the same computer and synchronized. The radios all use the same type of isotropic antenna and the transmit gains for Alice and Bob are the same.

Figure 5.5 shows how the gains G_{AE} and G_{BE} are actually imbalanced at the third radio terminal. The left side of the map represents the theoretical behavior one could expect from the simplified Friis equation [34],

$$\frac{P_r}{P_t} = G_t G_r \left(\frac{\lambda}{4\pi d}\right)^2. \tag{5.22}$$

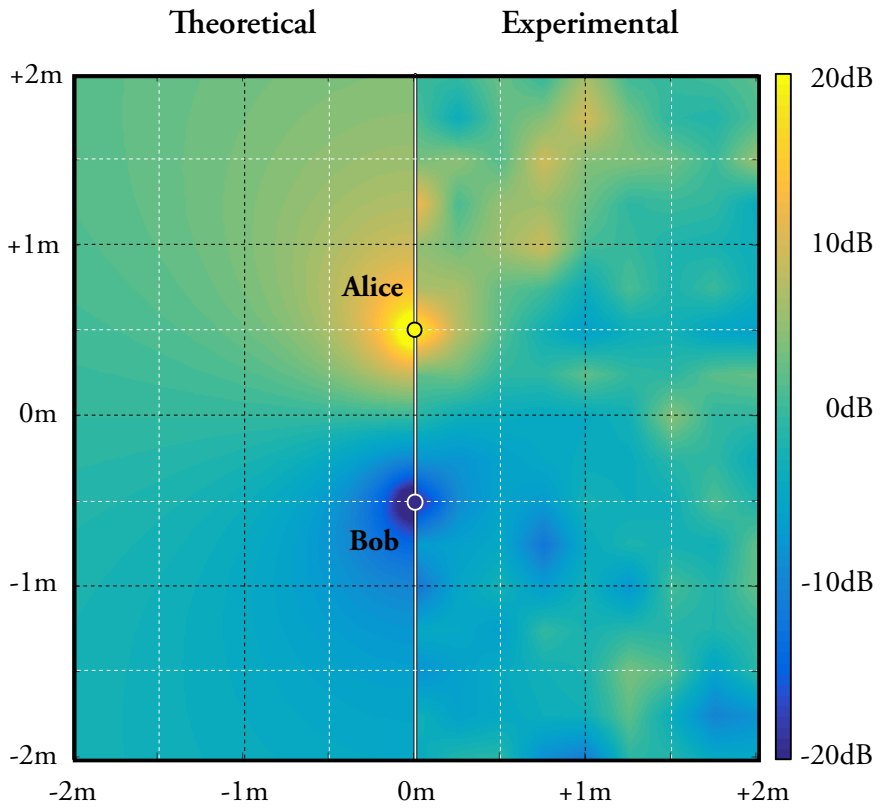

Figure 5.5 – *Gain imbalance between G_{AE} and G_{BE}.*

This equation represents the ratio of the received power P_r to the transmitted power P_t for isotropic receiving and transmitting antennas with respective gains G_r and G_t separated by a distance d; additionally, λ corresponds the wavelength of the carrier frequency. If both transmit antennas have the same gain, the gain ratio G_{AE} to G_{BE} in decibel is simply

$$\left.\frac{G_{AE}}{G_{BE}}\right|_{\text{dB}} = 20\log_{10}\frac{d_{BE}}{d_{AE}}, \tag{5.23}$$

where d_{AE} and d_{BE} respectively represent the distances between Alice and Eve and between Alice and Bob. The left-hand side of the map in Figure 5.5 shows this theoretical gain imbalance, while the right-hand side corresponds the actual gain imbalance measured with the WARPS.

Figure 5.5 shows that, in theory, gains should be balanced if Eve is far enough from Alice and Bob; however, experimental measurements show that this is not the case in a real wireless setting. There are several reasons that can explain this difference.

- Equation (5.22) is a very simple approximation of the path loss that does not take into account some properties of the antennas, such as their imperfect isotropy, their reflection coefficient of the antennas, and their polarization vectors.

- The environment is not free space, and the presence of reflecting and absorbing objects also affects the propagation of the wireless signal.

Managing this gain imbalance at the eavesdropper's terminal is challenging, but one can consider the following workarounds.

- The simple solution would consist in compensating the gain imbalance by adapting the transmit power of Alice and Bob's terminal. However, this solution requires a good knowledge of Eve's statistics.

- The best way to theoretically guarantee the performances of the coded cooperative jamming code is to refine the model presented in Section 5.1 to account for the gain imbalance. This solution is rather complicated since there is no simple degraded channel model in that case.

- Finally, one could design codes based on the simple theoretical model presented in Section 5.1 and verify how they perform in a real setting.

Terminal Synchronization The synchronization between Alice and Bob is another critical element to guarantee perfect interference at eve's terminal. If the codewords are not fully synchronized, the eavesdropper can gain additional information regarding the transmitted codewords. The radios should transmit codewords at the same time with each bit perfectly aligned, but it is also important to have synchronized local oscillators so that the signals are modulated at exactly the same frequency. Software-defined radios can achieve both goals, either by using an external clock or by using a GPS disciplined oscillator (GPSDO).

CHAPTER 6
CONCLUSION

6.1 Contributions

This dissertation has developed a comprehensive study of the coding mechanisms for multi-user physical layer security, with a particular emphasis on understanding how to best exploit interference and feedback for secrecy. The originality of the study lies in the combination of information theory, coding theory, and experiments conducted with software-defined radio. The results obtain have confirmed the promising possibilities of multi-user physical layer security but have also highlighted some of the practical challenges faced when deploying physical-layer security systems in a real wireless setting.

Chapter 2 has introduced the basic information-theoretic primitives and technical tools required to develop physical layer security schemes. In particular, this chapter has emphasized the crucial roles of channel resolvability and channel intrinsic randomness as primitives for the design of strongly secure schemes. One original result developed in this chapter is a joint exponent of channel resolvability and channel intrinsic randomness, which illustrates how to use some of the more intricate proof techniques required in subsequent chapters.

From the primitives presented in Chapter 2, Chapter 3 has developed an information-theoretic analysis of the coding mechanisms for the two-way wiretap channel. In particular, the analysis combines resolvability results for the multiple-access channel with secret-key generation and secret-key exchange. The resulting communication scheme has improved the state of the art by not only providing strong secrecy but also increasing the set of known achievable rates. The performance of this coding scheme has been illustrated for the Gaussian two-way wiretap channel, which is a model of practical interest.

Motivated by the crucial role of secret-key generation for the coding scheme in the coding scheme of Chapter 3, Chapter 4 has examined the practical limitations of a secret-key generation system in a real wireless setting. In particular, the study has focused on commonly made

assumptions regarding the reciprocity and diversity of the wireless channel. The rate penalty incurred by the use of a final number of samples as well as imperfect diversity and reciprocity have been precisely quantified using experimental channel gain measurements acquired with software-defined radios. While the results do not compromise the validity of the approach, they do suggest that many previously reported results may be overly optimistic.

Finally, Chapter 5 has proposed a code construction for the Gaussian two-way wiretap channel based on spatially-coupled LDPC codes. The resulting coding scheme exhibits low leakage rate but relies on perfect interferences at the eavesdropper's terminal, thus requiring good synchronization between the legitimate parties. This chapter has also investigated some of the limitations of the model used to design codes by comparing it again experimental measurements.

6.2 Perspectives

The work presented in this dissertation could be extended in several directions.

Strongly secure practical coded cooperative jamming codes The codes developed in Chapter 5 only offer weak secrecy. Polar codes could be used to provide strong secrecy [22], but the construction of low-complexity codes with good performance at reasonable block length might prove challenging.

Implementation of practical coded cooperative jamming schemes As discussed in Chapter 5, it may be challenging to design a practical coded cooperative jamming scheme since codeword interference is not a simple signal addition. As in Chapter 4, it would be valuable to implement such a system on software-defined radios to assess its performances in a real wireless environment. Since coded cooperative jamming requires an almost perfect synchronization of the terminals, it is important to assess the impact of imperfect synchronization on the overall secrecy rate. One could also expect that the position of the eavesdropper affects its abilities to gather information about the legitimate communications, although recent works [44, 45] suggest that this could be circumvented.

Hybrid solutions Chapter 4 describes how to generate a secret key between different terminals, but it does not investigate how the secret-key could be later used to provide secrecy. One could imagine a situation in which the secret-key serves in classical cryptographic primitives. For instance, one could used secret-key generation to refresh the WPA key of a Wi-Fi network periodically. This hybrid solution using both physical-layer security and cryptography would offer an additional layer of protection at no cost. The difficulty lies in the analysis of the overall solution since physical-layer security and cryptography do not rely on the same tools. Recent results [71] offer promising research directions.

CHAPTER 7

APPENDIX

7.1 Coding with Polar Codes

It is difficult to provide a simple joint coding scheme based on polar codes for the problem considered in Section 2.3. However, this appendix briefly illustrates how polar codes achieve the maximal achievable rates for some sources and channels with a separate approach. What follows assumes that $n \triangleq 2^m$ for some $m \in \mathbb{N}$.

For channel intrinsic randomness, consider the situation where Z^n is i.i.d. and $\mathcal{B}(\zeta)$, and X^n is such that $X^n = Z^n \oplus W^n$ with W^n i.i.d. $\mathcal{B}(\omega)$ and independent of Z^n. Polar codes (see [7] for notation) are used to encode X^n, the output of the encoder is denoted $U^n = X^n G^{\otimes m}$, where $G^{\otimes m}$ is the Arıkan transform. For $\delta_n \triangleq 2^{-n^\beta}$ with $\beta \in (0; 1/2)$, define the bit sets

$$\mathcal{G}_n^{\text{cir}} \triangleq \left\{ i \in [\![1, n]\!], \mathbb{H}\left(U_i | U^{i-1} Z^n\right) \geqslant 1 - \delta_n \right\}, \tag{7.1}$$

$$\mathcal{B}_n^{\text{cir}} \triangleq \complement_{[\![1,n]\!]} \mathcal{G}_n^{\text{cir}}. \tag{7.2}$$

One can then show [1, 22] that

$$\lim_{n\to\infty} \frac{\log |\mathcal{G}_n^{\text{cir}}|}{n} = \mathbb{H}(X|Z) \quad \text{and} \quad \lim_{n\to\infty} \mathbb{D}\left(p_{U[\mathcal{G}_n^{\text{cir}}]Z^n} \middle\| q_{U^n} q_{Z^n} \right) = 0, \tag{7.3}$$

where $p_{U[\mathcal{G}_n^{\text{cir}}]}$ is the distribution of the bits of U^n in position $\mathcal{G}_n^{\text{cir}}$ and $q_{U^n} \triangleq \mathcal{U}[\![1, 2^{nR}]\!]$.

For channel resolvability, if $\tilde{Y}^n$ is i.i.d. according to $\mathcal{B}(1/2)$ and if the channel W_2 is symmetric, the corresponding channel output $\tilde{K}^n$ is simulated as follows. Sequence Y^n is created from an n bit sequence V^n as $Y^n = V^n G^{\otimes n}$. The bits of V^n belong to either one of the sets

$$\mathcal{G}_n^{\text{res}} \triangleq \left\{ i \in [\![1, n]\!], C\left(W_n^{(i)}\right) \geqslant \delta_n \right\} \text{ and } \mathcal{B}_n^{\text{res}} \triangleq \complement_{[\![1,n]\!]} \mathcal{G}_n^{\text{res}}, \tag{7.4}$$

where $C(W_n^{(i)})$ is the capacity of the i -th bit-channel for bit V_i. If uniform bits are transmitted in positions $\mathcal{G}_n^{\text{res}}$, references [15, 64] prove that these codes are channel resolvability codes of rate $\mathbb{I}\left(\tilde{Y}; \tilde{K}\right)$, i.e.

$$\lim_{n\to\infty} \frac{\log |\mathcal{G}_n^{\text{res}}|}{n} = \mathbb{I}\left(\tilde{Y}; \tilde{K}\right) \text{ and } \lim_{n\to\infty} \mathbb{D}\left(p_{K^n} \| q_{K^n}\right) = 0. \tag{7.5}$$

Consequently, one can combine channel intrinsic randomness and channel resolvability by transmitting the nearly uniform bits of U^n obtained in positions $\mathcal{G}_n^{\text{cir}}$ on the position $\mathcal{G}_n^{\text{res}}$ of V^n, which is possible if $\mathbb{H}(X|Z) > \mathbb{I}(\tilde{Y};\tilde{K})$.

The solution in [64] guarantees both weak secrecy and reliability, but does not ensure reliability under a strong secrecy criterion. A solution to this problem consists in designing multi-block polar coding schemes with a secret-key exchange mechanism used to encrypt part of the non-secure bits [22, 87]. Such a construction achieves the secrecy capacity of symmetric degraded wiretap channels while also ensuring reliability.

7.2 Universal Software Radio Peripherals (USRPs)

USRPs (Universal Software Radio Peripheral) [31] are software-defined radios that can be used to communicate in various frequency ranges (DC-5GHz). USRPs are commonly used with the GNURadio [36] software suite that offers a wide range of possibilities for experimental communication systems.

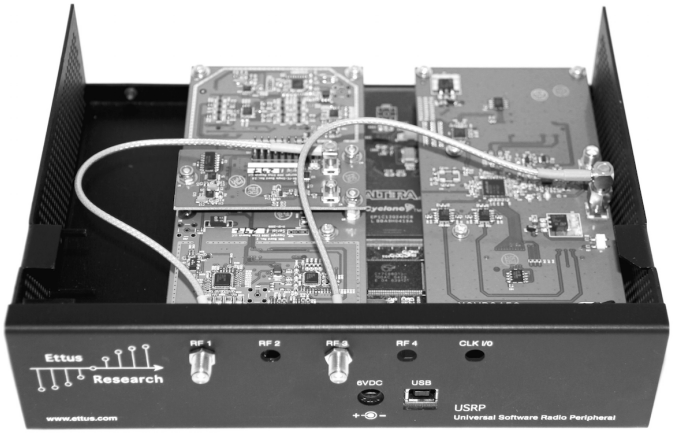

Figure 7.1 – *USRP1 with two daughterboards plugged in.*

At a high level, the architecture of a USRP presented in Figures 7.1 and 7.2 consists of a motherboard and between one and four daughterboards. The motherboard handles the processing of baseband signals while the daughterboards are designed to achieve frequency translation of the baseband signals.

Table 7.1 – *Comparisons of the different USPR models.*

Features	USRP1	N200	N210	E100	E110	E310	B200	B210	X300	X310
FPGA	Altera Cyclone	Spartan 3A1800	Spartan 3A3400	OMAP3 720 MHz	OMAP3 720 MHz	Zynq 7020	Spartan 675	Spartan 6150	Kintex7 325T	Kintex7 410T
Min freq.	DC	DC	DC	DC	DC	70 MHz	70 MHz	70 MHz	DC	DC
Max freq.	6 GHz	6 GHz	6 GHz	6 GHz	6 GHz	6 GHz	6 GHz	6 GHz	6 GHz	6 GHz
Bandwidth	16 MHz	50 MHz	50 MHz	8 MHz	8 MHz	56 MHz	56 MHz	56 MHz	120 MHz	120 MHz
Connection	USB 2.0	GbE	GbE	GbE	GbE	GbE/USB	USB 3.0	USB 3.0	GbE	GbE
MIMO	$\checkmark_{opt}$	$\checkmark_{opt}$	$\checkmark_{opt}$	×	×	✓	×	✓	✓	✓
DBoards	✓	✓	✓	✓	✓	×	×	×	✓	✓
Standalone	×	×	×	✓	✓	✓	×	×	×	×
GPS Sync	$\checkmark_{opt}$	$\checkmark_{opt}$	$\checkmark_{opt}$	$\checkmark_{opt}$	$\checkmark_{opt}$	✓	$\checkmark_{opt}$	$\checkmark_{opt}$	$\checkmark_{opt}$	$\checkmark_{opt}$
Price	$707	$1,515	$1,717	$1,310	$1,515	$2,700	$675	$1,100	$3,900	$4,800

There is a wide family daughterboards, covering different frequency bands from DC to 5 GHz. While the experiments described in Chapter 4 are performed with the first version of the USRP, several new versions have came out with improved characteristics. The current offer by Ettus is summarized in Table 7.1.

According to the frequency range used, different antennas may be used. Most experiments were performed in the 2.5 GHz and 5 GHz bands. The Wi-Fi-compliant antennas are perfectly adapted to transmit these signals. Although the center frequencies are 2.5 GHz and 5 GHz, the emission bandwidth is limited by the transfer speeds of the USB port to 8 MHz. This limitation makes USRPs incompatible with the IEEE802.11 protocol.

Figure 7.3 provides a representation of the internal operation of a USRP:

FPGA The FPGA (Field-Programmable Gate Array) performs many preprocessing operations to allow data to be transmitted through the USB port. The FPGA and the USB controller to which it is connected, are fully programmable via USB2.0. The FPGA is in charge of the

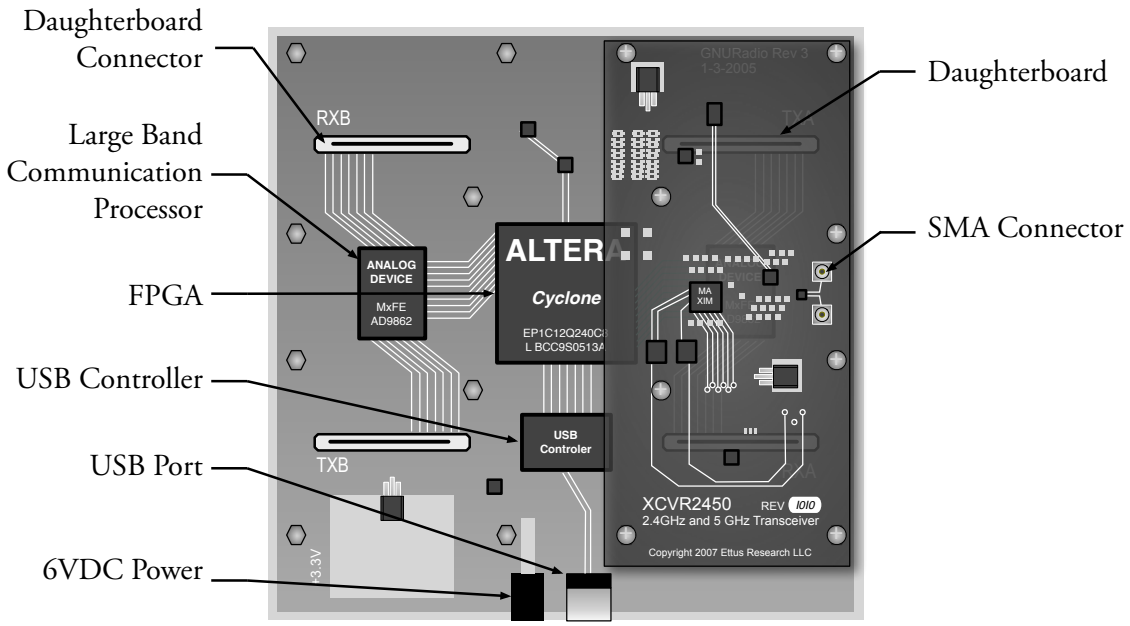

Figure 7.2 – *Overview of a software-defined radio (USRP1).*

distribution of signals between different daughterboards. It also allows various operations such as interpolation and decimation.

Analog/Digital converters The motherboard has four A/D converters of 12 bits. The sampling rate of the converter is 64 million samples per second. Although ADCs have a cutoff frequency of 150 MHz, if the bandwidth is greater than 32 MHz aliasing may occur. It is possible to use other sampling frequencies whenever they are in multiples of 128 MHz.

The board also contains four digital/analog converters of 14 bits. The sampling rate is 128 MHz. While the Shannon frequency is 64 MHz, it is best to stay below 50 MHz to make filtering easier. This offers a total of four inputs and four outputs with real numbers or two inputs and two outputs with complex numbers.

Daughterboards Daughterboards are pluggable circuits that, depending on the model, have different features. The motherboard can accommodate two Rx boards and two Tx boards or two dual mode transmitter boards. These boards are in charge of frequency translation. Each board has two SMA connectors for connecting the card to the antennas. Several daughterboards are available:

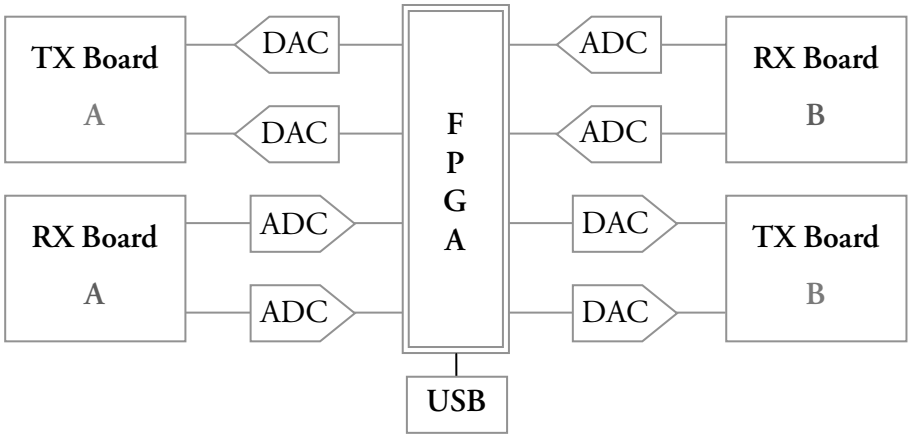

Figure 7.3 – *General diagram of a software-defined radio.*

- Basic RX / Basic TX: these boards do not perform any operation apart from connecting the converters outputs to SMA connectors.
- LFRX / LFTX: it is quite similar to the two previous boards, except that they include differential amplifiers to work with continuous signals.
- TVRX: Equipped with a MicroTuner, it can receive UHF and VHF signals and is therefore useful for receiving FM signals or TV.
- DBSRX: it can translate the baseband signals to frequencies between around 800 MHz and 2.4GHz. It also contains a channel filter adjustable between 1 MHz and 60 MHz.
- XCVR2450: it can work in the frequency range of IEEE 802.11n. It works in both transmission and reception. This is the board that is used for the experiments described in Chapter 4.

The ease of use software-defined radios largely stems from the fact that most of the processing is performed on baseband signals. The general idea is to observe that a bounded spectrum signal can be seen as a signal in baseband (centered around zero frequency) modulated by a signal of desired frequency. The modulation frequency corresponds to a translation of the signal spectrum.

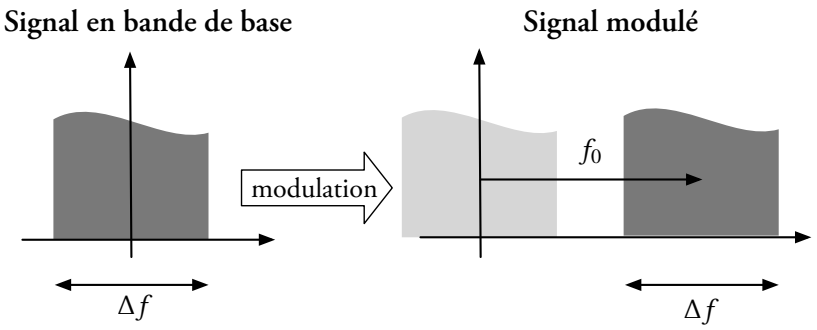

Figure 7.4 – *Modulation.*

Signal emission Several steps are necessary to emit a signal and are summarized in Figure 7.5 below.

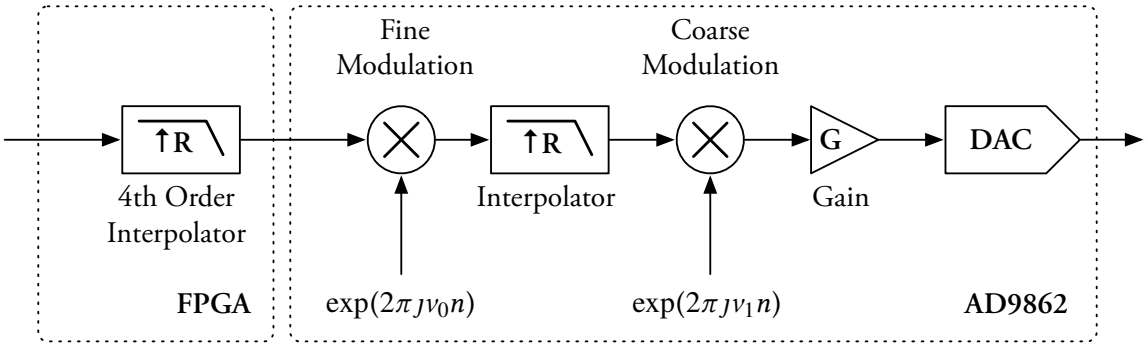

Figure 7.5 – *Signal modulation and transmission.*

The main steps are:

- **Interpolation**: to avoid clogging the USB port, the interpolator transmits only some of the points needed to represent signals. The interpolation is then necessary to achieve the number of samples the DAC needs. However, it is important to note that the interpolation rate affects the bandwidth of the transmitted signal.

- **Coarse modulation**: a first modulation allows coarsely approaching the center frequency of the desired modulation.

- **Fine modulation**: a second modulation allows precisely adjusting the center frequency of translation.

- **Gain**: A variable gain allows adapting the amplitude of signals to be transmitted at the level necessary to emit the wave with the antennas.

- **Digital/Analog conversion**: Finally, the digital signal is converted into an analog signal to be transmitted by the antennas in the form of an electromagnetic wave.

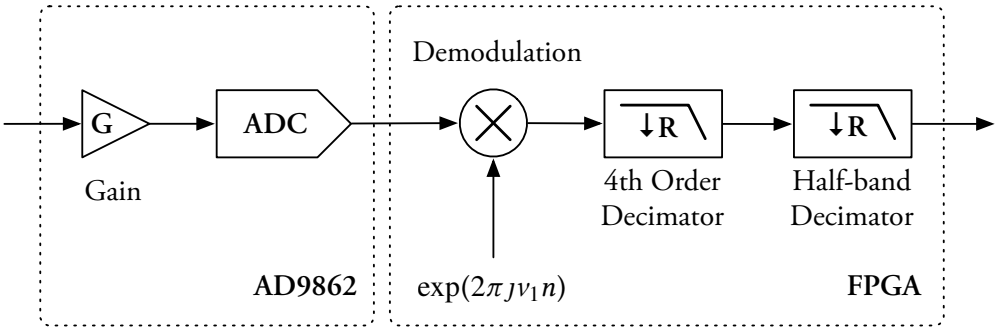

Figure 7.6 – *Signal reception and demodulation.*

Signal Reception The reception is performed with a mechanism similar to that of the emission. Again, it is important to be attentive to the decimation rate used. The higher this rate, the narrower the bandwidth of the received signal is. The demodulation is done only once. If the modulation and demodulation frequencies are slightly different (which is the case with asynchronous transmission), a modulation frequency (equal to the difference between the deviation of the frequency modulation and demodulation) will appear. It is necessary to find mechanisms to overcome this issue.

7.3 Proofs of Lemmas

7.3.1 Proof of Lemma 3.2

Define: $M_1 \triangleq \lceil 2^{R_1} \rceil$, $M_2 \triangleq \lceil 2^{R_2} \rceil$, $M_1' \triangleq \lceil 2^{R_1'} \rceil$ and $M_2' \triangleq \lceil 2^{R_2'} \rceil$, and let C be the random variable representing a randomly generated code. Let $\epsilon > 0$.

Without loss of generality, the analysis can be restricted to the transmission of a single message since

$$\begin{aligned}
\mathbb{E}(\mathbf{P}_{\mathrm{e}}(C)) &= \mathbb{E}\left(\sum_{\mu_1 \in \mathscr{M}_1 \times \mathscr{M}_1'} \sum_{\mu_2 \in \mathscr{M}_2 \times \mathscr{M}_2'} \mathbf{P}_{\mathrm{e}}(C|\mu_1, \mu_2 \text{ sent}) \mathbb{P}(\mu_1, \mu_2 \text{ sent})\right) \\
&= \sum_{\mu_1 \in \mathscr{M}_1 \times \mathscr{M}_1'} \sum_{\mu_2 \in \mathscr{M}_2 \times \mathscr{M}_2'} \mathbb{P}(\mu_1, \mu_2 \text{ sent}) \underbrace{\mathbb{E}_C(\mathbf{P}_{\mathrm{e}}(C|\mu_1, \mu_2 \text{ sent}))}_{\mathbb{E}_C(\mathbf{P}_{\mathrm{e}}(C|\mathbf{1},\mathbf{1} \text{ sent}))} \\
&\overset{\text{(a)}}{=} \mathbb{E}(\mathbf{P}_{\mathrm{e}}(C|\mathbf{1}, \mathbf{1} \text{ sent})) \\
&\overset{\text{(b)}}{\leqslant} \mathbb{E}(\mathbb{P}(\mu_1 \neq \hat{\mu}_1|\mathbf{1}, \mathbf{1} \text{ sent}, C)) + \mathbb{E}(\mathbb{P}(\mu_2 \neq \hat{\mu}_2|\mathbf{1}, \mathbf{1} \text{ sent}, C)).
\end{aligned} \tag{7.6}$$

Equality (a) follows from the symmetry of the random code construction, while inequality (b) comes from the union bound.

Each term can be expressed using the following events:

- $\mathscr{E}(i,j) = \left\{\left(X_1^n, C_2^n(i,j), Y_1^n\right) \in \mathscr{A}_{1,\epsilon}^n\right\}$
- $\mathscr{F}(i,j) = \left\{\left(X_2^n, C_1^n(i,j), Y_2^n\right) \in \mathscr{A}_{2,\epsilon}^n\right\}$

The probability of error is therefore

$$\mathbb{E}(\mathbb{P}(\mu_2 \neq \hat{\mu}_2|\mathbf{1}, \mathbf{1} \text{ sent}, C)) = \mathbb{P}\left(\mathscr{E}^{\mathrm{c}}(1,1) \cup \bigcup_{(i,j)\neq\mathbf{1}} \mathscr{E}(i,j)\right) \leqslant \underbrace{\mathbb{P}(\mathscr{E}^{\mathrm{c}}(1,1))}_{\leqslant \delta(\epsilon)} + \sum_{(i,j)\neq\mathbf{1}} \mathbb{P}(\mathscr{E}(i,j)) \tag{7.7}$$

$$\begin{aligned}
\sum_{(i,j)\neq\mathbf{1}} \mathbb{P}(\mathscr{E}(i,j)) &\leqslant \delta(\epsilon) + \sum_{(i,j)\neq\mathbf{1}} \sum_{(x_1^n, c_2^n, y_1^n) \in \mathscr{A}_{1,\epsilon}^n} p_{X_1^n Y_1^n}(x_1^n, c_1^n) p_{C_2^n}(c_2^n) \\
&\leqslant \delta(\epsilon) + \underbrace{M_2 M_2'}\, 2^{n(\mathbb{H}(X_1 C_2 Y_1)+\epsilon)} 2^{-n(\mathbb{H}(C_2)-\epsilon)} 2^{-n(\mathbb{H}(X_1 Y_1)-\epsilon)} \\
&\leqslant 2^{n(R_2+R_2'+\delta(n))} \\
&= \delta(\epsilon) + 2^{n(R_2+R_2'-\mathbb{I}(C_2;Y_1|X_1)+\delta(n)+3\epsilon)}.
\end{aligned} \tag{7.8}$$

Remark The second inequality follows from the union bound and the penultimate from the AEP.

If the rate constraints in (3.5) are satisfied, there exists a $\gamma > 0$ such that $R_2 + R_2' < \mathbb{I}(Y_1; C_2|X_1) - \gamma$, and choosing ϵ such that $\gamma + 3\epsilon > 0$, then

$$\lim_{n\to\infty} \mathbb{E}_C(\mathbb{P}(\mu_2 \neq \hat{\mu}_2|\mathbf{1}, \mathbf{1} \text{ sent}, C)) \leqslant \delta(\epsilon).$$

By symmetry and a similar reasoning on $\mathscr{F}$,

$$\lim_{n\to\infty} \mathbb{E}_C(\mathbb{P}(\mu_1 \neq \hat{\mu}_1|\mathbf{1}, \mathbf{1} \text{ sent}, C)) \leqslant \delta(\epsilon).$$

7.3.2 Proof of Lemma 3.3

Let$\epsilon > 0$.

An upper bound for the leakage using divergence

$$\begin{aligned}
\mathbf{L}(\mathscr{C}) &\triangleq \mathbb{I}(Z^n; M_1M_2|\mathscr{C}) \qquad (7.9)\\
&= \mathbb{D}\left(p_{M_1M_2Z^n|\mathscr{C}} \| p_{M_1M_2|\mathscr{C}} p_{Z^n|\mathscr{C}}\right)\\
&= \sum_{m_1=1}^{M_1}\sum_{m_2=1}^{M_2}\sum_{z^n\in\mathcal{Z}^n} p_{M_1M_2Z^n|\mathscr{C}}(m_1, m_2, z^n) \log_2 \frac{p_{M_1M_2Z^n|\mathscr{C}}(m_1, m_2, z^n)}{p_{M_1M_2|\mathscr{C}}(m_1, m_2) p_{Z^n|\mathscr{C}}(z^n)}\\
&= \sum_{m_1=1}^{M_1}\sum_{m_2=1}^{M_2}\sum_{z^n\in\mathcal{Z}^n} p_{Z^n|M_1M_2,\mathscr{C}}(z^n|m_1, m_2) p_{M_1M_2}(m_1, m_2) \log_2 \frac{p_{Z^n|M_1M_2,\mathscr{C}}(z^n|m_1, m_2)}{p_{Z^n|\mathscr{C}}(z^n)}\\
&= \frac{1}{M_1M_2}\sum_{m_1=1}^{M_1}\sum_{m_2=1}^{M_2} \mathbb{D}\left(p_{Z^n|m_1,m_2,\mathscr{C}} \| p_{Z^n|\mathscr{C}}\right)\\
&\leqslant \frac{1}{M_1M_2}\sum_{m_1=1}^{M_1}\sum_{m_2=1}^{M_2} \mathbb{D}\left(p_{Z^n|m_1,m_2,\mathscr{C}} \| p_{Z^n|\mathscr{C}}\right) + \mathbb{D}\left(p_{Z^n|\mathscr{C}} \| p_{Z^n}\right) \text{ because: } \mathbb{D}(\cdot\|\cdot) \succcurlyeq 0\\
&= \frac{1}{M_1M_2}\sum_{m_1=1}^{M_1}\sum_{m_2=1}^{M_2} \mathbb{D}\left(p_{Z^n|m_1,m_2,\mathscr{C}} \| p_{Z^n}\right). \qquad (7.10)
\end{aligned}$$

Taking the expectation and using the symmetry of the random code construction, yields

$$\begin{aligned}
\mathbb{E}_C(\mathbf{L}(C)) = \mathbb{I}(Z^n; M_1M_2|C) &\leqslant \frac{1}{M_1M_2}\sum_{m_1=1}^{M_1}\sum_{m_2=1}^{M_2} \mathbb{E}_C\big(\mathbb{D}\left(p_{Z^n|m_1,m_2,C} \| p_{Z^n}\right)\big)\\
&= \mathbb{E}_C\big(\mathbb{D}\left(p_{Z^n|1,1,C} \| p_{Z^n}\right)\big). \qquad (7.11)
\end{aligned}$$

An upper bound for the divergence:

$$\begin{aligned}
\mathbb{D}(p_{Z^n|1,1,\mathscr{C}}\|p_{Z^n}) &= \sum_{z^n\in\mathscr{Z}^n} p_{Z^n|1,1,\mathscr{C}}(z^n)\log_2\frac{p_{Z^n|1,1,\mathscr{C}}(z^n)}{p_{Z^n}(z^n)} \\
&= \sum_{z^n\in\mathscr{Z}^n} p_{Z^n|1,1,\mathscr{C}}(z^n)\log_2\frac{p_{Z^n|1,1,\mathscr{C}}(z^n)}{p_{Z^n}(z^n)}\mathbb{1}\left\{\frac{p_{Z^n|1,1,\mathscr{C}}(z^n)}{p_{Z^n}(z^n)}\leqslant\epsilon\right\} \\
&\quad+\sum_{z^n\in\mathscr{Z}^n} p_{Z^n|1,1,\mathscr{C}}(z^n)\log_2\frac{p_{Z^n|1,1,\mathscr{C}}(z^n)}{p_{Z^n}(z^n)}\mathbb{1}\left\{\frac{p_{Z^n|1,1,\mathscr{C}}(z^n)}{p_{Z^n}(z^n)}>\epsilon\right\} \\
&\leqslant \epsilon+\sum_{z^n\in\mathscr{Z}^n} p_{Z^n|1,1,\mathscr{C}}(z^n)\log_2\frac{p_{Z^n|1,1,\mathscr{C}}(z^n)}{p_{Z^n}(z^n)}\mathbb{1}\left\{\frac{p_{Z^n|1,1,\mathscr{C}}(z^n)}{p_{Z^n}(z^n)}>\epsilon\right\}. \qquad (7.12)
\end{aligned}$$

By the law of total probability:

$$\begin{aligned}
p_{Z^n}(z^n) &= \sum_{c_1^n\in\mathscr{C}_1^n}\sum_{c_2^n\in\mathscr{C}_2^n} p_{Z^n|C_1^nC_2^n}(z^n|c_1^n,c_2^n)\,p_{C_1^n}(c_1^n)\,p_{C_2^n}(c_2^n) \\
&\geqslant \sum_{c_1^n\in\mathscr{B}_1(1)}\sum_{c_2^n\in\mathscr{B}_2(1)} p_{Z^n|C_1^nC_2^n}(z^n|c_1^n,c_2^n)\,p_{C_1^n}(c_1^n)\,p_{C_2^n}(c_2^n) \\
&\geqslant \underbrace{M_1'M_2'p_{\min_1}^n p_{\min_2}^n}_{p_{\min}^n}\,p_{Z^n|1,1,\mathscr{C}}(z^n),
\end{aligned}$$

where $\mathscr{B}_i(m)$, for $i\in\{1,2\}$, represents the set of codewords associated with message m for user i. To obtain the last inequality, the p.d.f. of C_1 and C_2 must have compact support, which can always be obtained by cropping (to get a compact support) and scaling (to keep an unit area) the p.d.f. If $\check{C}_1$ and $\check{C}_2$ are derived from C_1 and C_2 with discrete or Gaussian p.d.f., it is possible to get $\mathbb{V}\left(p_{C_1},p_{\check{C}_1}\right)$ and $\mathbb{V}\left(p_{C_2},p_{\check{C}_2}\right)$ as small as desired. Since $\mathbb{I}(X;Y)$ viewed as a function of p_X, with $p_{Y|X}$ fixed, is continuous, the mutual informations involved in (3.7) are hardly modified by using $\check{C}_1$ and $\check{C}_2$ instead of C_1 and C_2.

By taking the expectation of (7.12) and with [90],

$$\mathbb{E}_C(\mathrm{L}(C))=\mathbb{I}(Z^n;M_1M_2|C)\leqslant\epsilon+n\log_2\left(\frac{1}{p_{\min}}\right)J_\epsilon, \qquad (7.13)$$

where J_ϵ is defined as:

$$J_\epsilon \triangleq \sum_{c_1^{(2)} \in \mathscr{C}_1^n} \cdots \sum_{c_1^{(M_1')} \in \mathscr{C}_1^n} \sum_{c_2^{(2)} \in \mathscr{C}_2^n} \cdots \sum_{c_2^{(M_2')} \in \mathscr{C}_2^n} p_{C_1^n}\left(c_1^{(2)}\right) \cdots p_{C_1^n}\left(c_1^{(M_1')}\right) p_{C_2^n}\left(c_2^{(2)}\right) \cdots p_{C_2^n}\left(c_2^{(M_2')}\right)$$
$$\times \sum_{c_1^{(1)} \in \mathscr{C}_1^n} \sum_{c_2^{(1)} \in \mathscr{C}_2^n} \sum_{z^n \in \mathscr{Z}^n} p_{C_1^n C_2^n Z^n}\left(c_1^{(1)}, c_2^{(1)}, z^n\right) \mathbb{1}\Bigg\{ \frac{1}{M_1' M_2'} \exp_2\left(i_{C_1^n C_2^n; Z^n}\left(c_1^{(1)}, c_2^{(1)}; z^n\right)\right)$$
$$+ \frac{1}{M_1' M_2'} \sum_{(i,j) \neq (1,1)} \exp_2\left(i_{C_1^n C_2^n; Z^n}\left(c_1^{(i)}, c_2^{(j)}; z^n\right)\right) > 1 + 4\epsilon \Bigg\}.$$

An upper bound for J_ϵ: Recall the following assumption from (3.7):

$$\begin{cases} R_1' + R_2' > \mathbb{I}(C_1 C_2; Z) \\ R_1' > \mathbb{I}(C_1; Z) \\ R_2' > \mathbb{I}(C_2; Z) \end{cases}$$

The quantity J_ϵ can be upper bounded as (see [90]):

$$J_\epsilon \leqslant J^{(1)} + J^{(2)} + J^{(3)} + J^{(4)}, \tag{7.14}$$

where $J^{(1)},\ J^{(2)},\ J^{(3)}$, and $J^{(4)}$ are defined as:

$$J^{(1)} = \mathbb{P}\Big(\frac{1}{M_1' M_2'} \exp_2\left(i_{C_1^n C_2^n; Z^n}\left(C_1^{(1)} C_2^{(1)}; Z^n\right)\right) > \epsilon \Big)$$
$$J^{(2)} = \mathbb{P}\Big(\frac{1}{M_1' M_2'} \sum_{i=1}^n \sum_{j=1}^n \exp_2\left(i_{C_1^n C_2^n; Z^n}\left(C_1^{(i)} C_2^{(j)}; Z^n\right)\right) > 1 + \epsilon \Big)$$
$$J^{(3)} = \mathbb{P}\Big(\frac{1}{M_1' M_2'} \sum_{i=1}^n \exp_2\left(i_{C_1^n C_2^n; Z^n}\left(C_1^{(i)} C_2^{(1)}; Z^n\right)\right) > \epsilon \Big)$$
$$J^{(4)} = \mathbb{P}\Big(\frac{1}{M_1' M_2'} \sum_{j=1}^n \exp_2\left(i_{C_1^n C_2^n; Z^n}\left(C_1^{(1)} C_2^{(j)}; Z^n\right)\right) > \epsilon \Big).$$

Now, each term are studied individually.

- First, $J^{(1)}$ can be upper bounded as follows:

$$J^{(1)} \leqslant \mathbb{P}(\exp_2(I\big(C_1^n C_2^n; Z^n\big)) > \exp_2(-n\epsilon + nR_1' + nR_2'))$$
$$= \mathbb{P}\Big(\frac{1}{n} I\big(C_1^n C_2^n; Z^n\big) > -\epsilon + R_1' + R_2' \Big)$$
$$= \mathbb{P}\Bigg(\frac{1}{n} \sum_{i=1}^n I\big(C_{1,i} C_{2,i}; Z_i\big) > R_1' + R_2' - \epsilon \Bigg)$$

since $R'_1 + R'_2 > \mathbb{I}(C_1C_2; Z)$, there exists $\gamma > 0$ such that $R'_1 + R'_2 > \mathbb{I}(C_1C_2; Z) + \gamma$. By taking $\epsilon < \gamma$ to guarantee $R'_1 + R'_2 - \epsilon > \mathbb{I}(C_1C_2; Z) + \underbrace{\gamma - \epsilon}_{>0}$ and to apply the Chernoff bound, then

$$\exists \alpha_{\gamma-\epsilon} > 0,\ J^{(1)} \leqslant e^{-\alpha_{\gamma-\epsilon} n}.$$

- With a similar reasoning, the terms $J^{(2)}$, $J^{(3)}$, and $J^{(4)}$ can also be upper bounded by quantities going exponentially to zero under the given conditions.

Conclusion: J_ϵ goes exponentially to zero as n goes to infinity. Therefore,

$$\lim_{n\to\infty} \mathbb{E}(\mathbf{L}(C)) \leqslant \delta(\epsilon).$$

7.3.3 Proof of Lemma 3.6

Define the joint distribution $q_{X_1^n X_2^n Z^n}$ as

$$\forall (x_1^n, x_2^n, z^n) \in \mathcal{X}_1^n \times \mathcal{X}_2^n \times \mathcal{Z}^n,\ q_{X_1^n X_2^n Z^n}(x_1^n, x_2^n, z^n) = p_{Z^n|X_1^n X_2^n}(z^n|x_1^n x_2^n) p_{X_1^n}(x_1^n) p_{X_2^n}(x_2^n).$$

First, the variational distance between distributions $p_{M_1M_2Z^n}$ and $p_{M_1M_2}p_{Z^n}$ can be upper bounded as

$$\begin{aligned}
\mathbb{V}(p_{M_1M_2Z^n}, p_{M_1M_2}p_{Z^n}) &= \mathbb{E}_{M_1M_2}(\mathbb{V}(p_{Z^n|M_1M_2}, p_{Z^n})) \\
&\leqslant \mathbb{E}_{M_1M_2}((\mathbb{V}(p_{Z^n|M_1M_2}, q_{Z^n}) + (\mathbb{V}(q_{Z^n}, p_{Z^n})) \\
&\leqslant 2\mathbb{E}_{M_1M_2}((\mathbb{V}(p_{Z^n|M_1M_2}, q_{Z^n})), \qquad (7.15)
\end{aligned}$$

By averaging over every code and by symmetry of the random code construction,

$$\mathbb{E}_{C_n}(\mathbb{V}(p_{M_1M_2Z^n}, p_{M_1M_2}p_{Z^n})) \leqslant 2\mathbb{E}_{C_n}(\mathbb{V}(p_{Z^n|M_1=1,M_2=1}, q_{Z^n})),$$

where

$$p_{Z^n|M_1=1,M_2=1}(z^n) = \sum_{m'_1=1}^{M'_1} \sum_{m'_2=1}^{M'_2} W_{Z^n|X_1^n X_2^n}(z^n|x_1^n(1,m'_1), x_2^n(1,m'_2)) p_{M'_1}(m'_1) p_{M'_2}(m'_2).$$

For any $z^n \in \mathcal{Z}^n$,

$$\begin{aligned}\mathbb{E}_{C_n}\left(p_{Z^n|M_1=1,M_2=1}\right) &= \mathbb{E}_{C_n}\left(\sum_{m_1'=1}^{M_1'}\sum_{m_2'=1}^{M_2'} W_{Z^n|X_1^nX_2^n}(z^n|x_1^n(1,m_1'),x_2^n(1,m_2'))p_{M_1'}(m_1')p_{M_2'}(m_2')\right)\\ &= \sum_{m_1'=1}^{M_1'}\sum_{m_2'=1}^{M_2'} \mathbb{E}_{C_n}\left(W_{Z^n|X_1^nX_2^n}(z^n|x_1^n(1,m_1'),x_2^n(1,m_2'))\right)p_{M_1'}(m_1')p_{M_2'}(m_2')\\ &= q_{Z^n}(z^n). \end{aligned} \tag{7.16}$$

Now define

$$\begin{aligned}p^{(1)}(z^n) \triangleq \sum_{m_1'=1}^{M_1'}\sum_{m_2'=1}^{M_2'} & W_{Z^n|X_1^nX_2^n}(z^n|x_1^n(1,m_1'),x_2^n(1,m_2'))p_{M_1'}(m_1')p_{M_2'}(m_2')\\ &\times\mathbb{1}\{(x_1^n(1,m_1'),x_2^n(1,m_2'),z^n)\in\mathcal{T}_\epsilon^n(X_1X_2Z)\}\\ p^{(2)}(z^n) \triangleq \sum_{m_1'=1}^{M_1'}\sum_{m_2'=1}^{M_2'} & W_{Z^n|X_1^nX_2^n}(z^n|x_1^n(1,m_1'),x_2^n(1,m_2'))p_{M_1'}(m_1')p_{M_2'}(m_2')\\ &\times\mathbb{1}\{(x_1^n(1,m_1'),x_2^n(1,m_2'),z^n)\notin\mathcal{T}_\epsilon^n(X_1X_2Z)\},\end{aligned}$$

which are used to upper bound $\mathbb{V}(p_{Z^n|M_1=1,M_2=1}, q_{Z^n})$ as

$$\begin{aligned}\mathbb{V}(p_{Z^n|M_1=1,M_2=1}, q_{Z^n}) \leqslant & \sum_{z^n\notin\mathcal{T}_\epsilon^n(Z)} |p_{Z^n|M_1=1,M_2=1}(z^n) - q_{Z^n}(z^n)|\\ &+ \sum_{z^n\in\mathcal{T}_\epsilon^n(Z)} \left|p^{(1)}(z^n) - \mathbb{E}(p^{(1)}(z^n))\right| + \sum_{z^n\in\mathcal{T}_\epsilon^n(Z)} \left|p^{(2)}(z^n) - \mathbb{E}(p^{(2)}(z^n))\right|\end{aligned}$$

The first and last terms involving non-typical and non-jointly typical sequences vanish as n goes to infinity. By using Jensen's inequality, the concavity of $x \mapsto \sqrt{x}$ guarantees

$$\mathbb{E}\left|p^{(1)}(z^n) - \mathbb{E}(p^{(1)}(z^n))\right| \leqslant \sqrt{\mathrm{Var}(p^{(1)}(z^n))}$$

The variance of $\mathrm{Var}(p^{(1)}(z^n))$ is

$$\begin{aligned}\mathrm{Var}(p^{(1)}(z^n)) \leqslant \left(\sum_{m_1'=1}^{M_1'}\sum_{m_2'=1}^{M_2'} p_{M_1'}(m_1')^2 p_{M_2'}(m_2')^2\right)\mathbb{E}\big(W^2_{Z^n|X_1^nX_2^n}(z^n|X_1^n(1,1),X_2^n(1,1))\\ \times\mathbf{1}\{(X_1^n(1,1),X_2^n(1,1),z^n)\in\mathcal{T}_\epsilon^n(X_1X_2Z)\}\big)\\ +\left(\sum_{m_1'=1}^{M_1'} p_{M_1'}(m_1')^2\right)\mathbb{E}\big(W_{Z^n|X_1^nX_2^n}(z^n|X_1^n(1,1),X_2^n(1,1))\\ \times W_{Z^n|X_1^nX_2^n}(z^n|X_1^n(1,1),X_2^n(1,2))\mathbf{1}\{(X_1^n(1,1),X_2^n(1,2),z^n)\in\mathcal{T}_\epsilon^n(X_1X_2Z)\}\big)\\ +\left(\sum_{m_2'=1}^{M_2'} p_{M_2'}(m_2')^2\right)\mathbb{E}\big(W_{Z^n|X_1^nX_2^n}(z^n|x_1^n(1,1),x_2^n(1,1))\\ \times W_{Z^n|X_1^nX_2^n}(z^n|X_1^n(1,2),X_2^n(1,1))\mathbf{1}\{(X_1^n(1,2),X_2^n(1,1),z^n)\in\mathcal{T}_\epsilon^n(X_1X_2Z)\}\big)\end{aligned}$$

Using the AEP on the first term yields

$$\begin{aligned}\mathbb{E}\big(W^2_{Z^n|X_1^nX_2^n}(z^n|X_1^n(1,1),X_2^n(1,1))\mathbf{1}\{(X_1^n(1,1),X_2^n(1,1),z^n)\in\mathcal{T}_\epsilon^n(X_1X_2Z)\}\big)\\ \leqslant \exp_2(-n(\mathbb{H}(Z|X_1X_2)+\mathbb{H}(Z)-\delta_\epsilon)).\end{aligned}\tag{7.17}$$

Similarly, the second term is upper bounded by

$$\begin{aligned}\mathbb{E}\big(W_{Z^n|X_1^nX_2^n}(z^n|X_1^n(1,1),X_2^n(1,1))W_{Z^n|X_1^nX_2^n}(z^n|X_1^n(1,1),X_2^n(1,2))\\ \times\mathbf{1}\{(X_1^n(1,1),X_2^n(1,2),z^n)\in\mathcal{T}_\epsilon^n(X_1X_2Z)\}\big)\\ \leqslant \exp_2(-n(\mathbb{H}(Z|X_1)+\mathbb{H}(Z)-\delta_\epsilon)),\end{aligned}\tag{7.18}$$

and the third term by

$$\begin{aligned}\mathbb{E}\big(W_{Z^n|X_1^nX_2^n}(z^n|X_1^n(1,1),X_2^n(1,1))W_{Z^n|X_1^nX_2^n}(z^n|X_1^n(1,2),X_2^n(1,1))\\ \times\mathbf{1}\{(X_1^n(1,2),X_2^n(1,1),z^n)\in\mathcal{T}_\epsilon^n(X_1X_2Z)\}\big)\\ \leqslant \exp_2(-n(\mathbb{H}(Z|X_2)+\mathbb{H}(Z)-\delta_\epsilon)).\end{aligned}\tag{7.19}$$

Combining the previous upper bounds,

$$\begin{aligned}
\mathrm{Var}(p^{(1)}(z^n)) \leqslant & \left(\sum_{m_1'=1}^{M_1'}\sum_{m_2'=1}^{M_2'} p_{M_1'}(m_1')^2 p_{M_2'}(m_2')^2\right)\exp_2(-n(\mathbb{H}(Z|X_1X_2)+\mathbb{H}(Z)-\delta_\epsilon)) \\
& + \left(\sum_{m_1'=1}^{M_1'} p_{M_1'}(m_1')^2\right)\exp_2(-n(\mathbb{H}(Z|X_1)+\mathbb{H}(Z)-\delta_\epsilon)) \\
& + \left(\sum_{m_2'=1}^{M_2'} p_{M_2'}(m_2')^2\right)\exp_2(-n(\mathbb{H}(Z|X_2)+\mathbb{H}(Z)-\delta_\epsilon)) \\
= & \exp_2(-\mathbb{H}_2(M_1'M_2') - n(\mathbb{H}(Z|X_1X_2)+\mathbb{H}(Z)-\delta_\epsilon)) \\
& + \exp_2(-\mathbb{H}_2(M_1') - n(\mathbb{H}(Z|X_1)+\mathbb{H}(Z)-\delta_\epsilon)) \\
& + \exp_2(-\mathbb{H}_2(M_2') - n(\mathbb{H}(Z|X_2)+\mathbb{H}(Z)-\delta_\epsilon)),
\end{aligned}$$

where $\mathbb{H}_2$ is the second order Rényi entropy. Finally,

$$\begin{aligned}
& \sum_{z^n\in\mathcal{T}_\epsilon^n(Z)} \left|p^{(1)}(z^n) - \mathbb{E}\left(p^{(1)}(z^n)\right)\right| \\
& \quad \leqslant \exp_2(\mathbb{H}(Z))\exp_2\left(-\frac{n}{2}\left(\frac{1}{n}\mathbb{H}_2(M_1'M_2') + \mathbb{H}(Z|X_1X_2) + \mathbb{H}(Z) - \delta_\epsilon\right)\right) \\
& \qquad + \exp_2(\mathbb{H}(Z))\exp_2\left(-\frac{n}{2}\left(\frac{1}{n}\mathbb{H}_2(M_1') + \mathbb{H}(Z|X_1) + \mathbb{H}(Z) - \delta_\epsilon\right)\right) \\
& \qquad + \exp_2(\mathbb{H}(Z))\exp_2\left(-\frac{n}{2}\left(\frac{1}{n}\mathbb{H}_2(M_2') + \mathbb{H}(Z|X_2) + \mathbb{H}(Z) - \delta_\epsilon\right)\right), \\
& \quad = \exp_2\left(-\frac{n}{2}\left(\frac{\mathbb{H}_2(M_1'M_2')}{n} + \mathbb{I}(X_1X_2;Z) - \delta_\epsilon\right)\right) \\
& \qquad + \exp_2\left(-\frac{n}{2}\left(\frac{\mathbb{H}_2(M_1')}{n} + \mathbb{I}(X_1;Z) - \delta_\epsilon\right)\right) \\
& \qquad + \exp_2\left(-\frac{n}{2}\left(\frac{\mathbb{H}_2(M_2')}{n} + \mathbb{I}(X_2;Z) - \delta_\epsilon\right)\right), \qquad (7.20)
\end{aligned}$$

Therefore, if

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{H}_2(M_1',M_2') > \mathbb{I}(X_1X_2;Z),\ \lim_{n\to\infty}\frac{1}{n}\mathbb{H}_2(M_1') > \mathbb{I}(X_1;Z),\ \text{and}\ \lim_{n\to\infty}\frac{1}{n}\mathbb{H}_2(M_2') > \mathbb{I}(X_2;Z), \qquad (7.21)$$

the sum vanishes as n goes to infinity.

7.3.4 Proof of Lemma 3.7

Let $i \in [\![1, B-1]\!]$. We have

$$
\begin{aligned}
\mathbf{L}_{i+1} - \mathbf{L}_i &= \mathbb{I}\left(M_{1,1:B} M'_{1,2:B} M_{2,1:B}; Z^n_{1:i+1}\right) - \mathbb{I}\left(M_{1,1:B} M'_{1,2:B} M_{2,1:B}; Z^n_{1:i}\right) \\
&= \mathbb{I}\left(M_{1,1:B} M'_{1,2:B} M_{2,1:B}; Z^n_{i+1} \middle| Z^n_{1:i}\right) \\
&= \mathbb{I}\left(M_{1,1:i+2:B} M'_{1,1:i+2:B} M_{2,1:i+2:B}; Z^n_{i+1} \middle| Z^n_{1:i} M_{1,1:i+1} M'_{1,1:i+1} M_{2,1:i+1}\right) \\
&\qquad + \mathbb{I}\left(M_{1,1:i+1} M'_{1,2:i+1} M_{2,1:i+1}; Z^n_{i+1} \middle| Z^n_{1:i}\right) \\
&\overset{(a)}{\leqslant} \mathbb{I}\left(M_{1,1:i+2:B} M'_{1,1:i+2:B} M_{2,1:i+2:B}; Z^n_{1:i+1} M_{1,1:i+1} M'_{1,1:i+1} M_{2,1:i+1}\right) \qquad (7.22) \\
&\qquad + \mathbb{I}\left(M_{1,1:i+1} M'_{1,2:i+1} M_{2,1:i+1} Z^n_{1:i}; Z^n_{i+1}\right) \\
&\overset{(b)}{=} \mathbb{I}\left(M_{1,1:i+1} M'_{1,2:i+1} M_{2,1:i+1} Z^n_{1:i}; Z^n_{i+1}\right) \qquad (7.23) \\
&= \mathbb{I}\left(M_{1,i+1} M'_{1,i+1} M_{2,i+1}; Z^n_{i+1}\right) + \mathbb{I}\left(M_{1,1:i} M'_{1,2:i} M_{2,1:i} Z^n_{1:i}; Z^n_{i+1} \middle| M_{1,i+1} M'_{1,i+1} M_{2,i+1}\right) \\
&\overset{(c)}{\leqslant} \delta_n(\epsilon) + \mathbb{I}\left(M_{1,1:i} M'_{1,2:i} M_{2,1:i} Z^n_{1:i}; Z^n_{i+1} \middle| M_{1,i+1} M'_{1,i+1} M_{2,i+1}\right) \qquad (7.24) \\
&\leqslant \delta_n(\epsilon) + \mathbb{I}\left(M_{1,1:i} M'_{1,2:i} M_{2,1:i} Z^n_{1:i}; Z^n_{i+1} M_{1,i+1} M'_{1,i+1} M_{2,i+1}\right) \\
&= \delta_n(\epsilon) + \mathbb{I}\left(M_{1,1:i} M'_{1,2:i} M_{2,1:i} Z^n_{1:i}; M_{1,i+1} M'_{1,i+1} M_{2,i+1}\right) \\
&\qquad + \mathbb{I}\left(M_{1,1:i} M'_{1,2:i} M_{2,1:i} Z^n_{1:i}; Z^n_{i+1} \middle| M_{1,i+1} M'_{1,i+1} M_{2,i+1}\right) \\
&= \delta_n(\epsilon) + \mathbb{I}\left(M_{1,1:i} M'_{1,2:i} M_{2,1:i} Z^n_{1:i}; Z^n_{i+1} \middle| M_{1,i+1} M'_{1,i+1} M_{2,i+1}\right) \\
&\overset{(d)}{\leqslant} \delta_n(\epsilon) + \mathbb{I}\left(M_{1,1:i} M'_{1,2:i} M_{2,1:i} Z^n_{1:i}; Z^n_{i+1} P_{i+1} K_i \middle| M_{1,i+1} M'_{1,i+1} M_{2,i+1}\right) \qquad (7.25) \\
&= \delta_n(\epsilon) + \mathbb{I}\left(M_{1,1:i} M'_{1,2:i} M_{2,1:i} Z^n_{1:i}; P_{i+1} K_i \middle| M_{1,i+1} M'_{1,i+1} M_{2,i+1}\right) \\
&\qquad + \mathbb{I}\left(M_{1,1:i} M'_{1,2:i} M_{2,1:i} Z^n_{1:i}; Z^n_{i+1} \middle| P_{i+1} K_i M_{1,i+1} M'_{1,i+1} M_{2,i+1}\right) \\
&\overset{(e)}{\leqslant} \delta_n(\epsilon) + \mathbb{I}\left(C^n_{1,1:i} C^n_{2,1:i} Z^n_{1:i}; K_i P_{i+1}\right) \qquad (7.26) \\
&= \delta_n(\epsilon) + \mathbb{I}\left(C^n_{1,i} C^n_{2,i} Z^n_i; K_i P_{i+1}\right) + \mathbb{I}\left(C^n_{1,1:i-1} C^n_{2,1:i-1} Z^n_{1:i-1}; K_i P_{i+1} \middle| C^n_{1,i} C^n_{2,i} Z^n_i\right) \\
&\overset{(f)}{\leqslant} \delta_n(\epsilon), \qquad (7.27)
\end{aligned}
$$

where (a) holds by the chain rule and positivity of mutual information, (b) holds by independence of $M_{1,1:i+2:B} M'_{1,2:i+2:B} M_{2,1:i+2:B}$ with all the random variables of the previous blocks, (c) holds thanks to the result on secret-key generation, in (d) the random variables P_{i+1} and K_i are

introduced, (e) holds because of the data processing inequality and because the conditioning on P_{i+1} and K_i breaks the dependencies between the random variables of block $(i+1)$ and the random variables of the previous blocks, (f) holds because (P_{i+1}, K_i) and $(C^n_{1,1:i-1}, C^n_{2,1:i-1}, Z^n_{1:i-1})$ are independent and also because P_{i+1} and K_i break the dependencies between consecutive blocks.

7.3.5 Proof of Lemma 4.1

To show that $\mathbb{P}_U(\mathbb{H}^\epsilon_\infty(S) - \mathbb{H}^\epsilon_\infty(S|U=\bar{u}) > \log|\mathcal{U}| + r)$, first note that

$$\begin{aligned}
\mathbb{H}^\epsilon_\infty(S|U=\bar{u}) &= \max_{q_{S\bar{U}} \in \mathcal{B}^\epsilon(Su)} \min_{s\in\mathcal{S}} \log \frac{p_U(\bar{u})}{q_{S\bar{U}}(s,\bar{u})} \\
&= -\min_{q_{S\bar{U}} \in \mathcal{B}^\epsilon(Su)} \max_{s\in\mathcal{S}} \log\left(q_{S\bar{U}}(s,\bar{u})\right) + \log\left(p_U(\bar{u})\right) \\
&\geqslant -\min_{q_S \in \mathcal{B}^\epsilon(S)} \max_{s\in\mathcal{S}} \log\left(q_S(s)\right) + \log\left(p_U(\bar{u})\right) \\
&= \mathbb{H}^\epsilon_\infty(S) + \log\left(p_U(\bar{u})\right) \qquad (7.28)
\end{aligned}$$

where the inequality follows because for all $(s,u) \in \mathcal{S}\times\mathcal{U}$, $q_{S\bar{U}}(s,\bar{u}) \leqslant p_S(s)$. Hence,

$$\begin{aligned}
\mathbb{P}_U(\mathbb{H}^\epsilon_\infty(S) - \mathbb{H}^\epsilon_\infty(S|U=\bar{u}) > \log|\mathcal{U}| + r) &\leqslant \mathbb{P}(\log p_U(U) < -\log|\mathcal{U}| - r) \\
&= \mathbb{P}\left(p_U(U) < \frac{2^{-r}}{|\mathcal{U}|}\right) \\
&= \sum_{u\in\mathcal{U}} p_U(u)\, \mathbb{1}\left\{p_U(u) < \frac{2^{-r}}{|\mathcal{U}|}\right\} \\
&\leqslant 2^{-r}. \qquad (7.29)
\end{aligned}$$

REFERENCES

[1] Abbe, E., "Low complexity constructions of secret keys using polar coding". In: *Proceedings of the IEEE Information Theory Workshop (ITW)*. Vol. 36. 2012 (cit. on p. 122).

[2] Abdi, A., "On the estimation of the K parameter for the Rice fading distribution". In: *IEEE Communications Letters* 5.3 (2001), pp. 92–94 (cit. on p. 86).

[3] Amariucai, G. T., Wei, S., "Secrecy Rates of Binary Wiretapper Channels Using Feedback Schemes". In: *Proceedings of the 42nd Annual Conference on Information Sciences and Systems CISS 2008*. Mar. 2008, pp. 624–629 (cit. on p. 47).

[4] Andersson, M., "Equivocation of Eve Using Two Edge Type LDPC Codes for the Binary Erasure Wiretap Channel". In: *Proceedings of the 44th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE. 2010, pp. 2045–2049 (cit. on p. 99).

[5] Ardestanizadeh, E., "Wiretap Channel With Secure Rate-Limited Feedback". In: *IEEE Transactions on Information Theory* 55.12 (Dec. 2009), pp. 5353–5361 (cit. on p. 47).

[6] Arimoto, S., "Information Measures and Capacity of Order α for Discrete Memoryless Channels". In: *Topics in Information Theory*. (Colloquia Mathematica Societatis János Bolyai) 16 (1977), pp. 41–52 (cit. on p. 19).

[7] Arıkan, E., "Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels". In: *IEEE Transactions on Information Theory* 55.7 (2009), pp. 3051–3073 (cit. on pp. 46, 122).

[8] Barros, J., Bloch, M., "Strong Secrecy for Wireless Channels (Invited Talk)". In: *Information Theoretic Security*. Springer, 2008, pp. 40–53 (cit. on p. 93).

[9] Bennett, C. H., "Generalized Privacy Amplification". In: *IEEE Transactions on Information Theory* 41.6 (Nov. 1995), pp. 1915–1923 (cit. on pp. 10, 93).

[10] Bharadia, D., McMilin, E., Katti, S., "Full Duplex Radios". In: *Proceedings of the ACM SIGCOMM 2013 Conference*. Aug. 2013 (cit. on p. 82).

[11] Bloch, M., Laneman, J. N., "On the Secrecy Capacity of Arbitrary Wiretap Channels". In: *Proceedings of the 46th Allerton Conference on Communication, Control, and Computing*. Monticello, IL, Sept. 2008, pp. 818–825 (cit. on p. 48).

[12] Bloch, M., "Channel Scrambling for Secrecy". In: *Proceedings of the IEEE International Symposium on Information Theory*. Seoul, Korea, July 2009, pp. 2452–2456 (cit. on p. 47).

[13] Bloch, M., "Channel Intrinsic Randomness". In: *Proceedings of the IEEE International Symposium on Information Theory*. Austin, TX, June 2010, pp. 2607–2611 (cit. on pp. 34, 35, 46).

[14] Bloch, M. R., Kliewer, J., "On Secure Communication with Constrained Randomization". In: *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*. Cambridge, MA, July 2012, pp. 1172–1176 (cit. on p. 36).

[15] Bloch, M. R., Luzzi, L., Kliewer, J., "Strong Coordination with Polar Codes". In: *Proceedings of the 50th Annual Allerton Conference on Communication, Control, and Computing*. 2012, pp. 565–571 (cit. on p. 122).

[16] Bloch, M., Barros, J., *Physical-layer Security: from Information Theory to Security Engineering*. Cambridge University Press, 2011 (cit. on pp. xvi, 13, 31, 33, 92, 96).

[17] Bloch, M., Laneman, J., "Strong Secrecy From Channel Resolvability". In: *IEEE Transactions on Information Theory* 59.12 (Dec. 2013), pp. 8077–8098 (cit. on pp. 5, 7, 85).

[18] Cachin, C., Maurer, U. M., "Linking Information Reconciliation and Privacy Amplification". In: *Journal of Cryptology* 10.2 (Mar. 1997), pp. 97–110 (cit. on p. 81).

[19] Chen, C., Jensen, M., "Secret Key Establishment Using Temporally and Spatially Correlated Wireless Channel Coefficients". In: *IEEE Transactions on Mobile Computing* 10.2 (Feb. 2011), pp. 205–215 (cit. on pp. 75, 85).

[20] Chou, R., Bloch, M., "Secret-Key Generation with Arbitrarily Varying Eavesdropper's Channel". In: *Proceedings of the IEEE GlobalSIP 2013*. 2013 (cit. on pp. 93, 98).

[21] Chou, R., Bloch, M., "Separation of Reliability and Secrecy in Rate-Limited Secret-Key Generation". In: *IEEE Transactions on Information Theory* 60.8 (Aug. 2014), pp. 4941–4957 (cit. on pp. 80, 93).

[22] Chou, R., Bloch, M., Abbe, E., "Polar Coding for Secret-key Generation". In: *Proceedings of the 2013 IEEE Information Theory Workshop (ITW)*. Sept. 2013, pp. 1–5 (cit. on pp. 99, 120, 122, 123).

[23] Cover, T. M., Thomas, J. A., *Elements of Information Theory*. 2nd. Wiley-Interscience, 2006 (cit. on pp. xvi, 13, 24, 31, 33, 93).

[24] Csiszàr, I., Narayan, P., "Common Randomness and Secret Key Generation with a Helper". In: *IEEE Transactions on Information Theory* 46.2 (Mar. 2000), pp. 344–366 (cit. on pp. 47, 58, 69, 80).

[25] Csiszàr, I., Narayan, P., "Secrecy Capacities for Multiple Terminals". In: *IEEE Transactions on Information Theory* 50.12 (Dec. 2004), pp. 3047–3061 (cit. on p. 8).

[26] Csiszàr, I., Körner, J., *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, 2011 (cit. on pp. 20, 21, 44).

[27] El Gamal, A., Kim, Y.-H., *Network Information Theory*. Cambridge University Press, 2011 (cit. on p. 62).

[28] El Gamal, A., "New Achievable Secrecy Rate Regions for the Two-Way Wiretap Channel". In: *Proceedings of the IEEE Information Theory Workshop*. Cairo, Egypt, Jan. 2010, pp. 1–5 (cit. on pp. 47, 48, 52, 58, 60).

[29] Erven, T., Harremoës, P., "Rényi Divergence and Majorization". In: *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*. Austin, TX, June 2010, pp. 1335–1339 (cit. on p. 38).

[30] Erven, T., Harremoës, P., "Rényi Divergence and Kullback-Leibler Divergence". In: *Information Theory, IEEE Transactions on* 60.7 (July 2014), pp. 3797–3820 (cit. on p. 21).

[31] Ettus, M., *Universal Software Radio Peripheral (USRP)*. 2008 (cit. on pp. xvi, 123).

[32] Fano, R. M., *Transmission of Information: A Statistical Theory of Communication*. New-York: John Wiley & Sons, Inc., 1961 (cit. on p. 93).

[33] Fehr, S., Berens, S., "On the Conditional Rényi Entropy". In: *IEEE Transactions on Information Theory* 60.11 (Nov. 2014), pp. 6801–6810 (cit. on p. 19).

[34] Friis, H. T., "A note on a Simple Transmission Formula". In: *Proceedings of the IRE* 34.5 (1946), pp. 254–256 (cit. on p. 116).

[35] G. D. Forney, J., "On the Role of MMSE Estimation in Approaching the Information-Theoretic Limits of Linear Gaussian Channels: Shannon Meets Wiener". In: *Proceedings of the 41st Annual Allerton Conference on Communication, Control, and Computing*. Monticello, IL, Oct. 2003, pp. 430–439 (cit. on p. 59).

[36] *GNU Radio Website*. `http://www.gnuradio.org`. Accessed: Apr 23, 2015 (cit. on p. 123).

[37] Gündüz, D., Brown, D. R., Poor, H. V., "Secret Communication with Feedback". In: *Proceedings of the 2008 International Symposium on Information Theory and its Applications, ISITA2008*. Dec. 2008, pp. 1–6 (cit. on p. 47).

[38] Han, T. S., *Information-Spectrum Methods in Information Theory*. Springer, 2002 (cit. on pp. 45, 48, 49).

[39] Han, T., Verdú, S., "Approximation Theory of Output Statistics". In: *IEEE Transactions on Information Theory* 39.3 (May 1993), pp. 752–772 (cit. on pp. 36, 46, 48).

[40] Hayashi, M., "General Nonasymptotic and Asymptotic Formulas in Channel Resolvability and Identification Capacity and their Application to the Wiretap Channels". In: *IEEE Transactions on Information Theory* 52.4 (Apr. 2006), pp. 1562–1575 (cit. on pp. 38, 48).

[41] Hayashi, M., "Exponential Decreasing Rate of Leaked Information in Universal Random Privacy Amplification". In: *IEEE Transactions on Information Theory* 57.6 (2011), pp. 3989–4001 (cit. on pp. 38, 41, 43).

[42] Hayashi, M., Matsumoto, R., "Secure Multiplex Coding with Dependent and Non-uniform Multiple Messages". In: *Proceedings of the 50th Annual Allerton Conference on Communication, Control, and Computing*. 2012, pp. 954–959 (cit. on p. 36).

[43] He, X., Yener, A., "Providing Secrecy When the Eavesdropper Channel is Arbitrarily Varying: A Case for Multiple Antennas". In: *Proceedings of the 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. 2010, pp. 1228–1235 (cit. on pp. 73, 98).

[44] He, X., Yener, A., "The Role of Channel States in Secret Key Generation". In: *Proceedings of the 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*. Sept. 2010, pp. 2681–2686 (cit. on p. 120).

[45] He, X., Yener, A., "Secrecy When the Eavesdropper Controls its Channel States". In: *Proceedings of the 2011 IEEE International Symposium on Information Theory (ISIT),* July 2011, pp. 618–622 (cit. on p. 120).

[46] He, X., Yener, A., "The Role of Feedback in Two-Way Secure Communications". In: *IEEE Transactions on Information Theory* 59.12 (Dec. 2013), pp. 8115–8130 (cit. on pp. 47, 48, 58, 74).

[47] Holenstein, T., Renner, R., "On the Randomness of Independent Experiments". In: *IEEE Transactions on Information Theory* 57.4 (Apr. 2011), pp. 1865–1871 (cit. on p. 94).

[48] Hou, J., Kramer, G., "Informational Divergence Approximations to Product Distributions". In: *Proceedings of the 13th Canadian Workshop on Information Theory (CWIT)*. June 2013, pp. 76–81 (cit. on p. 38).

[49] Imai, H., Kobara, K., Morozov, K., "On the Possibility of Key Agreement Using Variable Directional Antenna". In: *Proceedings of the 1st Joint Workshop on Information Security*. 2006, pp. 153–157 (cit. on pp. 10, 75).

[50] Imre, "Almost Independence and Secrecy Capacity". In: *Problems of Information Transmission* 32.1 (Jan. 1996), pp. 40–47 (cit. on p. 20).

[51] Iwamoto, M., Shikata, J., "Information Theoretic Security for Encryption Based on Conditional Rényi Entropies". In: *Information Theoretic Security*. Vol. 8317. Lecture Notes in Computer Science. Springer International Publishing, 2014, pp. 103–121 (cit. on p. 19).

[52] Jana, S., "On the Effectiveness of Secret Key Extraction from Wireless Signal Strength in Real Environments". In: *Proceedings of the 15th International Conference on Mobile Computing and Networking*. 2009, pp. 321–332 (cit. on pp. 75, 80, 97).

[53] Jimenez Felstrom, A., Zigangirov, K., "Time-varying Periodic Convolutional Codes with Low-density Parity-check Matrix". In: *IEEE Transactions on Information Theory* 45.6 (1999), pp. 2181–2191 (cit. on pp. 99, 104).

[54] Klinc, D., "LDPC Codes for the Gaussian Wiretap Channel". In: *IEEE Transactions on Information Forensics and Security* 6.3 (2011), pp. 532–540 (cit. on pp. 99, 109, 115).

[55] Kramer, G., *Topics in Multi-User Information Theory*. Vol. 4. Foundations and Trends in Communications and Information Theory 4-5. NOW Publishers, 2008, pp. 265–444 (cit. on p. 22).

[56] Kudekar, S., Kasai, K., "Spatially Coupled Codes over the Multiple Access Channel". In: *Proceedings of the IEEE International Symposium on Information Theory*. 2011, pp. 2816–2820 (cit. on pp. 104, 106, 108–110, 113).

[57] Kudekar, S., Richardson, T., Urbanke, R., "Threshold Saturation via Spatial Coupling: Why Convolutional LDPC Ensembles Perform So Well over the BEC". In: *IEEE Transactions on Information Theory* 57.2 (Feb. 2011), pp. 803–834 (cit. on pp. xvi, 99, 104, 106).

[58] Lai, L., El Gamal, H., "The Relay-Eavesdropper Channel: Cooperation for Secrecy". In: *IEEE Transactions on Information Theory* 54.9 (Sept. 2008), pp. 4005–4019 (cit. on p. 8).

[59] Lai, L., El Gamal, H., Poor, H. V., "The Wiretap Channel With Feedback: Encryption Over the Channel". In: *IEEE Transactions on Information Theory* 54.11 (Nov. 2008), pp. 5059–5067 (cit. on p. 47).

[60] Leung-Yan-Cheong, S., "On a Special Class of Wiretap Channels". In: *IEEE Transactions on Information Theory* 23.5 (1977), pp. 625–627 (cit. on p. 5).

[61] Li, Z., "Securing Wireless Systems via Lower Layer Enforcements". In: *Proceedings of the 2006 ACM Workshop on Wireless Security*. Los Angeles, California, USA, Sept. 2006, pp. 33–42 (cit. on p. 75).

[62] R. Liu and W. Trappe, eds. *Securing Wireless Communications at the Physical Layer*. Springer, 2010 (cit. on pp. 8, 51).

[63] Madiseh, M., "Secret Key Extraction in Ultra Wideband Channels for Unsynchronized Radios". In: *Proceedings of the 6th Annual Communication Networks and Services Research Conference*. 2008, pp. 88–95 (cit. on pp. 10, 75).

[64] Mahdavifar, H., Vardy, A., "Achieving the Secrecy Capacity of Wiretap Channels Using Polar Codes". In: *IEEE Transactions on Information Theory* 57.10 (2011), pp. 6428–6443 (cit. on pp. 7, 99, 122, 123).

[65] Mathur, S., "Radio-telepathy: Extracting a Secret Key From An Unauthenticated Wireless Channel". In: *Proceedings of the 14th ACM International Conference on Mobile Computing and Networking*. ACM. 2008, pp. 128–139 (cit. on p. 97).

[66] Maurer, U., "Secret Key Agreement by Public Discussion from Common Information". In: *IEEE Transactions on Information Theory* 39 (1993), pp. 733–742 (cit. on p. 96).

[67] Maurer, U., Wolf, S., "From Weak to Strong Information-theoretic Key Agreement". In: *Proceedings of the 2000 IEEE International Symposium on Information Theory*. 2000, pp. 18– (cit. on pp. 10, 92, 93).

[68] Menezes, A. J., Van Oorschot, P. C., Vanstone, S. A., *Handbook of Applied Cryptography*. CRC press, 1996 (cit. on p. xvi).

[69] Moon, T. K., *Error Correction Coding : Mathematical Methods And Algorithms*. John Wiley & Sons, 2005 (cit. on p. xvi).

[70] Murphy, P., Hunter, C., Welsh, E., *Wireless Open-Access Research Platform (WARP)*. 2008 (cit. on p. xvi).

[71] Oggier, F., Mihaljevic, M., "An Information-Theoretic Security Evaluation of a Class of Randomized Encryption Schemes". In: *IEEE Transactions on Information Forensics and Security* 9.2 (Feb. 2014), pp. 158–168 (cit. on p. 121).

[72] Patwari, N., "High-Rate Uncorrelated Bit Extraction for Shared Secret Key Generation from Channel Measurements". In: *IEEE Transactions on Mobile Computing* 9.1 (2010), pp. 17–30 (cit. on pp. 10, 75).

[73] Peng, H., Long, F., Ding, C., "Feature Selection Based on Mutual Information Criteria of Max-dependency, Max-relevance, and Min-redundancy". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27.8 (2005), pp. 1226–1238 (cit. on p. 85).

[74] **Pierrot, A. J.**, Bloch, M. R., "Joint Channel Intrinsic Randomness and Channel Resolvability". In: *Proceedings of the 2013 IEEE Information Theory Workshop (ITW)*. Sept. 2013, pp. 1–5 (cit. on pp. 10, 12, 13).

[75] **Pierrot, A. J.**, Chou, R. A., Bloch, M. R., "Experimental Aspects of Secret Key Generation in Indoor Wireless Environments". In: *Proceedings of the 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. June 2013, pp. 669–673 (cit. on pp. 11, 12, 75).

[76] **Pierrot, A. J.**, Bloch, M. R., "Strongly Secure Communications Over the Two-Way Wiretap Channel". In: *IEEE Transactions on Information Forensics and Security* 6.3 (Sept. 2011), pp. 595–605 (cit. on pp. 10, 12, 47).

[77] **Pierrot, A. J.**, Bloch, M. R., "LDPC-Based Coded Cooperative Jamming Codes". In: *Proceedings of the IEEE Information Theory Workshop*. Lausanne, Switzerland, Sept. 2012, pp. 462–466 (cit. on pp. 11, 12, 99).

[78] Pinsker, M. S., *Information and Information Stability of Random Variables and Processes*. Holden Day, 1964 (cit. on pp. 21, 93).

[79] Premnath, S. N., "Secret Key Extraction from Wireless Signal Strength in Real Environments". In: *IEEE Transactions on Mobile Computing* 12.5 (May 2013), pp. 917–930 (cit. on p. 75).

[80] Premnath, S. N., "Efficient High-Rate Secret Key Extraction in Wireless Sensor Networks Using Collaboration". In: *ACM Transactions on Sensor Networks* 11.1 (July 2014), 2:1–2:32 (cit. on p. 75).

[81] Proakis, J. G., *Digital Communications*. 4th. McGraw-Hill, 2001 (cit. on p. xvi).

[82] Rathi, V., "Rate-equivocation Optimal Spatially Coupled LDPC Codes for the BEC Wiretap Channel". In: *Proceedings of the Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*. Aug. 2011, pp. 2393–2397 (cit. on p. 99).

[83] Renner, R., Wolf, S., "Smooth Rényi Entropy and Applications". In: *Proceedings of the 2004 International Symposium on Information Theory (ISIT)*. 2004 (cit. on p. 19).

[84] Rényi, A., "On Measures of Entropy and Information". In: *Proceedings of the 4^{th} Berkeley Symposium on Mathematical Statistics and Probability*. Vol. 1. Berkeley, CA, 1961, pp. 547–561 (cit. on pp. 18, 21).

[85] Richardson, T., Urbanke, R., *Modern Coding Theory*. Cambridge University Press, 2008 (cit. on pp. 111, 112).

[86] Şaşoğlu, E., Telatar, E., Yeh, E., "Polar Codes for the Two-User Multiple-Access Channel". In: *IEEE Transactions on Information Theory* 59.10 (Oct. 2013), pp. 6583–6592 (cit. on p. 115).

[87] Şaşoğlu, E., Vardy, A., "A New Polar Coding Scheme for Strong Security on Wiretap Channels". In: *Proceedings of the 2013 IEEE International Symposium on Information Theory (ISIT)*. July 2013, pp. 1117–1121 (cit. on p. 123).

[88] Shannon, C. E., "A Mathematical Theory of Communication". In: *The Bell System Technical Journal* 27 (July 1948), pp. 379–423, 623–656 (cit. on p. 28).

[89] Slepian, D., Wolf, J. K., "Noiseless Coding of Correlated Information Sources". In: *IEEE Transactions on Information Theory* 19.4 (July 1973), pp. 471–480 (cit. on pp. 30, 31).

[90] Steinberg, Y., "Resolvability Theory for the Multiple-Access Channel". In: *IEEE Transactions on Information Theory* 44.2 (Mar. 1998), pp. 472–487 (cit. on pp. 48, 53, 62, 131, 132).

[91] Storn, R., Price, K., "Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces". In: *Journal of Global Optimization* 11.4 (1997), pp. 341–359 (cit. on p. 115).

[92] Tekin, E., Yener, A., "The General Gaussian Multiple-Access and Two-Way Wiretap Channels: Achievable Rates and Cooperative Jamming". In: *IEEE Transactions on Information Theory* 54.6 (June 2008), pp. 2735–2751 (cit. on pp. 8, 47, 48, 52, 55, 57, 61).

[93] Tekin, E., Yener, A., "Correction to: "The Gaussian Multiple Access Wire-Tap Channel" and "The General Gaussian Multiple Access and Two-Way Wire-Tap Channels: Achievable Rates and Cooperative Jamming"". In: *IEEE Transactions on Information Theory* 56.9 (Sept. 2010), pp. 4762–4763 (cit. on pp. 8, 47, 52, 57, 61).

[94] Tse, D., Viswanath, P., *Fundamentals of Wireless Communication*. Cambridge University Press, 2005 (cit. on p. 86).

[95] Tyagi, H., Watanabe, S., "A bound for Multiparty Secret Key Agreement and Implications for a Problem of sSecure Computing". In: *Advances in Cryptology–EUROCRYPT 2014*. Springer, 2014, pp. 369–386 (cit. on p. 92).

[96] Vembu, S., Verdú, S., "Generating Random Bits from An Arbitrary Source: Fundamental Limits". In: *IEEE Transactions on Information Theory* 41.5 (Sept. 1995), pp. 1322–1332 (cit. on p. 36).

[97] Visweswariah, K., Kulkarni, S. R., Verdu, S., "Separation of Random Number Generation and Resolvability". In: *IEEE Transactions on Information Theory* 46.6 (Sept. 2000), pp. 2237–2241 (cit. on p. 36).

[98] Wallace, J., "Secure Physical Layer Key Generation Schemes: Performance and Information Theoretic Limits". In: *Proceedings of the IEEE International Conference on Communications (ICC)*. 2009, pp. 1–5 (cit. on pp. 10, 75).

[99] Wallace, J., Sharma, R., "Automatic Secret Keys From Reciprocal MIMO Wireless Channels: Measurement and Analysis". In: *IEEE Transactions on Information Forensics and Security* 5.3 (2010), pp. 381–392 (cit. on pp. 75, 98).

[100] Wang, Q., Xu, K., Ren, K., "Cooperative Secret Key Generation from Phase Estimation in Narrowband Fading Channels". In: *IEEE Journal on Selected Areas in Communications* 30.9 (2012), pp. 1666–1674 (cit. on p. 75).

[101] Wang, Q., "Fast and Scalable Secret Key Generation Exploiting Channel Phase Randomness in Wireless Networks". In: *Proceedings of the 2011 IEEE INFOCOM*. 2011, pp. 1422–1430 (cit. on p. 75).

[102] Watanabe, S., Hayashi, M., "Non-asymptotic Analysis of Privacy Amplification via Rényi Entropy and Inf-spectral Entropy". In: *Proceedings of the 2013 IEEE International Symposium on Information Theory (ISIT)*. Istanbul, Turkey, July 2013, pp. 2715–2719 (cit. on pp. 43, 92, 93).

[103] Watanabe, S., Oohama, Y., "Secret Key Agreement from Vector Gaussian Sources by Rate-Limited Public Communication". In: *Proceedings of the IEEE International Symposium on Information Theory*. Austin, TX, June 2010, pp. 2597–2601 (cit. on pp. 47, 58, 71, 72).

[104] Weedbrook, C., "Gaussian Quantum Information". In: *Reviews of Modern Physics* 84.2 (2012), pp. 621–669 (cit. on p. 81).

[105] Wilson, R., Tse, D., Scholtz, R. A., "Channel Identification: Secret Sharing Using Reciprocity in Ultrawideband Channels". In: *IEEE Transactions on Information Forensics and Security* 2.3 (Sept. 2007), pp. 364–375 (cit. on p. 75).

[106] Wong, C. W., Wong, T. F., Shea, J. M., "LDPC Code Design for the BPSK-Constrained Gaussian Wiretap channel". In: *Proceedings of the 2011 Globecom Workshops*. Dec. 2011, pp. 898–902 (cit. on pp. 99, 109).

[107] Wyner, A. D., "The Wire-Tap Channel". In: *The Bell System Technical Journal* 54.8 (Oct. 1975), pp. 1355–1367 (cit. on pp. 4, 99).

[108] Yassaee, M. H., Aref, M. R., "Multiple Access Wiretap Channels with Strong Secrecy". In: *Proceedings of the IEEE Information Theory Workshop*. Dublin, Ireland, Sept. 2010 (cit. on p. 53).

[109] Ye, C., "Information-Theoretically Secret Key Generation for Fading Wireless Channels". In: *IEEE Transactions on Information Forensics and Security* 5.2 (2010), pp. 240–254 (cit. on pp. 10, 75, 97).

[110] Zafer, M., Agrawal, D., Srivatsa, M., "A Note on Information-Theoretic Secret Key Exchange over Wireless Channels". In: *Proceedings of the 47th Annual Allerton Conference on Communication, Control, and Computing*. 2009, pp. 754–761 (cit. on p. 10).

[111] Zhang, J., Kasera, S., Patwari, N., "Mobility Assisted Secret Key Generation Using Wireless Link Signatures". In: *Proceedings of the International Conference on Computer Communications*. 2010, pp. 1–5 (cit. on pp. 10, 75).

VITA

Alexandre J. Pierrot received the Diplôme d'Ingénieur from Supélec, Gif-sur-Yvette, France, in 2011 and the M.Sc. degree in Electrical and Computer Engineering from the Georgia Institute of Technology, Atlanta, GA, in 2011, where he is currently pursuing the Ph.D. degree. He is working with Prof. Matthieu R. Bloch in the communication architecture research group (Arcom) at Georgia Tech. His research interests include physical layer security, digital communications, and digital signal processing.